

Discriminação entre sinais de vozes saudáveis e patológicos por meio da análise da imagem do espaço de fase reconstruído

Pablo H. U. de Pinho, Maria F. K. B. Couras, Silvana L. N. C. Costa e Suzete E. N. Correia

Resumo— Neste trabalho, são empregadas técnicas de processamento de sinais baseadas na análise dinâmica não linear para investigar a alteração vocal provocada pela presença de patologia laringeas, a partir das modificações no espaço de fase dos sinais analisados. São utilizados dois métodos de extração de características para obter medidas das imagens do espaço de fase que mostram o comportamento da dinâmica vocal do sistema: método da contagem de caixas ponderadas e método da similaridade. Com o auxílio de um classificador MLP (*Multilayer Perceptron*) é realizada a classificação com as medidas extraídas individualmente e de forma combinada. Obteve-se acurácia de 99% na classificação de vozes patológicas e saudáveis. Os métodos empregados se mostram promissores e podem ser viáveis na implementação de um sistema para detecção de alterações vocais provocadas pela presença de patologias laringeas.

Palavras-Chave— *Reconstrução do Espaço de Fase, Redes Neurais, Patologias Laringeas, Processamento de Sinais de voz.*

Abstract— In this work, signal processing techniques based on nonlinear dynamic analysis are used to investigate the vocal changes caused by the presence of laryngeal pathology, based on the phase space modifications of the analyzed signals. Two methods of features extraction are used to obtain measurements of the phase space images that show the behavior of the vocal dynamics of the system: weighted box counting method and similarity method. With the aid of an MLP (*Multilayer Perceptron*) classifier, classification with the measurements extracted individually and in combination is carried out. Accuracy of 99% was obtained in the discrimination between healthy and pathological voices. The methods employed are promising and may be feasible in the implementation of a system for detecting vocal changes caused by the presence of laryngeal pathologies.

Keywords— *Reconstruction of Phase Space, Neural Networks, Laryngeal Diseases, Speech Signal Processing.*

I. INTRODUÇÃO

Diversos métodos têm sido propostos para avaliar a presença de distúrbios vocais, tais como o método de compressão PPM (*Prediction by Partial Matching*), associado às características tradicionais dos sinais analisados para detectar a presença de patologias atingindo acurácia de 92% [1]. Podem ser citadas também a análise cepstral aliada a alguns parâmetros acústicos tradicionais (frequência fundamental, energia, potência, formantes, entre outras) obtendo taxa de acurácia de 81% na classificação de paralisia nas pregas vocais [2]. No entanto, a definição do melhor método e das características ou parâmetros mais adequados para detectar a presença de patologias laringeas ainda é objeto de estudo por especialistas. Pesquisas indicam que a análise não linear pode quantificar padrões que não foram bem

avaliados utilizando as medidas de perturbação, concluindo que este tipo de análise pode quantificar padrões aperiódicos apresentados em patologias severas [3, 4].

Técnicas da análise da dinâmica não-linear têm sido empregadas para analisar um sinal caótico. A maioria delas são baseadas na análise do espaço de fase reconstruído, combinado com métodos estatísticos, para determinar algumas medidas quantitativas que possam caracterizar o sinal, tais como dimensão de correlação, expoentes de *Lyapunov*, função de informação mútua e medidas de entropia, entre outras [4-7].

Couras et. al [8] e Pinho et. al [9] empregaram técnicas de processamento de sinais de voz e imagem baseadas na análise da dinâmica não-linear para classificação e caracterização de patologias laringeas. Estas técnicas empregadas basearam-se nos trabalhos de Roopaei et al. [10] e Amann et al. [11], em que foram extraídas características das imagens do espaço de fase reconstruído a partir dos métodos da Contagem de Caixas e da Diferença. Foram obtidas taxas de acurácia na classificação entre sinais de vozes afetadas por edema e sinais saudáveis superiores a 96%, mostrando que o uso de características extraídas do espaço de fase para diagnosticar a presença de patologias laringeas é bastante interessante e promissor [8, 9].

Além dos métodos da Contagem de Caixas e da Diferença, Roopaei et al. [10] propuseram dois outros métodos, denominados Método da Contagem de Caixas Ponderadas e Método da Similaridade, os quais baseiam-se também nas imagens dos espaços de fase reconstruído de sinais biomédicos (eletrocardiogramas). Este trabalho tem como principal objetivo empregar o Método da Contagem de Caixas Ponderadas e o Método da Similaridade para caracterizar e discriminar sinais de vozes saudáveis e sinais de vozes afetados por patologias na laringe (edema de Reinke e paralisia nas pregas vocais). Para classificação desses sinais foram empregadas as Redes neurais artificiais MLP (*Multilayer Perceptron*).

O artigo está organizado da seguinte forma: Na Seção II é descrito o método da reconstrução do espaço de fase. Na Seção III é apresentada a descrição da base de dados e da metodologia empregada. Na Seção IV são apresentados os resultados obtidos e discussão, seguida da Seção V, em que são apresentadas as conclusões.

II. RECONSTRUÇÃO DO ESPAÇO DE FASE

A técnica da reconstrução do espaço de fase é baseada no Teorema da Imersão de Takens [12]. Este teorema demonstra que através do método das coordenadas defasadas é possível reconstruir propriedades do espaço de fase a partir de uma

série temporal em que vetores m -dimensionais são reconstruídos [11, 13]:

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

$x(t)$: série temporal, m : dimensão de imersão, τ : tempo de atraso e $i = 1, 2, \dots, n$ é a quantidade de estados do sistema. O comportamento do traçado obtido na reconstrução do espaço de fase, usando o tempo de atraso ótimo (τ ótimo), representa a dinâmica vocal. Quanto mais regular o espaço de fase reconstruído, mais periodicidade o sinal possui [13].

O cálculo do passo de reconstrução ótimo é baseado na Teoria da Informação, onde utiliza o método da informação mútua média, proposto por Fraser e Swinney (1986) [14, 15]. Ele é determinado pelo ponto em que ocorre o primeiro valor mínimo na curva de informação mútua (FMMI), que representa a autocorrelação da série, medindo a dependência do sinal consigo mesmo [16]. Esta curva oferece a melhor estimativa do tempo de atraso ou passo de reconstrução ótimo para o sinal em análise, realçando a dinâmica aleatória e não determinística do sinal, representado por τ ótimo.

III. MATERIAIS E MÉTODOS

Nesta seção é feita a descrição da base de dados e da metodologia empregada no trabalho.

A. Base de Dados

Foram utilizados sinais disponibilizados da base de dados *Disordered Voice Database*, Modelo 4337, gravada pelo *Massachusetts Eye and Ear Infirmary* (MEEI) [17]. Esta base contém pronúncias sustentadas da vogal /a/. Foram selecionados 149 sinais, dos quais 53 sinais são de vozes saudáveis e 96 sinais são de vozes patológicas (44 sinais são de vozes afetadas por edema e 52 sinais são afetados por paralisia nas pregas vocais).

Os sinais foram amostrados na frequência de 25 kHz para sinais patológicos e 50 kHz para sinais saudáveis, em que os últimos foram sub-amostrados para 25 kHz com o intuito de preservar o tamanho do segmento em amostras para obtenção do espaço de fase. Os sinais utilizados possuem duração média de um segundo para os sinais de vozes patológicas e de três segundos para os sinais de vozes saudáveis, sendo que não há influência das diferentes durações dos sinais de voz, pois no processamento, os sinais são delimitados como será visto na seção C. Estes sinais são previamente processados, não contendo intervalos de silêncio antes ou depois da locução.

B. Metodologia

Inicialmente, é realizada a aquisição do sinal. Nesta pesquisa, os sinais de voz foram selecionados da base de dados descrita na Seção A. Posteriormente, é realizada a extração do tempo de atraso ótimo (τ ótimo) para cada sinal de voz a partir da curva de informação mútua. O valor do τ ótimo obtido foi empregado para realizar a reconstrução do espaço de fase. Estes espaços de fase foram armazenados em imagens quadradas (512 x 512 *pixels*). As características foram extraídas através do método da similaridade e do método da contagem de caixas ponderadas [10]. Esses valores foram armazenados como padrões de referência e foram utilizados pelo classificador MLP para realizar a classificação dos sinais

Pinho, P. H. U., Couras, M. F. K. B., Costa, S. C. e Correia, S.E.N, Programa de Pós-graduação em Engenharia Elétrica, Instituto Federal da Paraíba, João Pessoa-PB, Brasil, E-mails: pablohenriqueifpb@gmail.com, kallynna.mary@gmail.com, silvana@ifpb.edu.br e suzete@ifpb.edu.br. Este trabalho foi parcialmente financiado pela COPEX e PRPIPG - IFPB.

em saudável ou patológico, ou ainda, discriminando a patologia (edema ou paralisia nas pregas vocais).

C. Reconstrução do Espaço de Fase

A reconstrução do espaço de fase foi realizada por meio da *toolbox* CRP (*Cross Recurrence Plots*) implementada em ambiente MATLAB[®] v.7.9. [18]. Para realizar a reconstrução do espaço de fase é necessário estimar o tempo de atraso ótimo para cada sinal.

Tempo de Atraso (τ): O tempo de atraso para a reconstrução do espaço de fase (τ ótimo) foi encontrado através da curva de informação mútua [16], obtida a partir do programa *Visual Recurrence Analysis* (VRA). Como o valor do τ ótimo varia para cada sinal analisado, foram calculados os valores médio (τ_o med), mínimo (τ_o min) e máximo (τ_o max) para cada classe (normal, edema e paralisia), com os quais foram construídos os espaços de fase para cada sinal. Os valores dos tempos de atraso obtidos para cada classe estão apresentados na Figura 1. Observa-se que os sinais de vozes saudáveis apresentam valores mais baixos para o tempo de atraso ótimo e mais próximos da mediana, enquanto que os sinais patológicos apresentam valores mais altos e com maior dispersão.

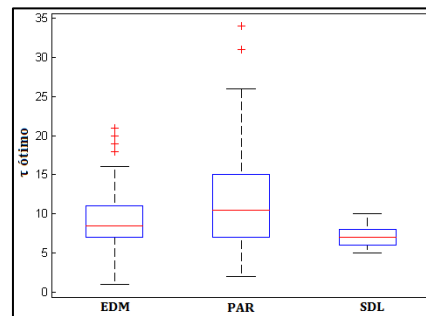


Fig. 1. – Boxplots para os valores do τ ótimo para as classes edema (EDM), paralisia (PRL) e saudável (SDL).

O valor médio para tempo de atraso ótimo (τ_o med) obtido para os sinais afetados por edema é 9, para os sinais afetados por paralisia é 12 e para os sinais saudáveis é 7.

A reconstrução do espaço de fase foi realizada para 10 ciclos do sinal, como proposto por Dajer [4], sendo que cada ciclo corresponde a um período de *pitch* [5]. Foram reconstruídos os espaços de fase utilizando os quatro valores obtidos para o tempo de atraso (τ ótimo, τ_o min, τ_o med e τ_o max), obtendo assim, para cada sinal, quatro espaços de fase, a partir dos quais serão obtidas as características pelos métodos empregados. Na Figura 2 são apresentados os espaços de fase reconstruídos para sinal de voz afetado por edema nas pregas vocais, considerando os quatro atrasos de tempo correspondentes.

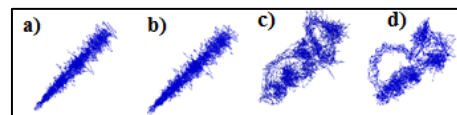


Fig. 2. Espaços de fase reconstruídos para um sinal de voz afetado por edema nas pregas vocais (a) τ ótimo; b) τ_o min; c) τ_o med e d) τ_o max).

D. Método da Contagem de Caixas Ponderadas

No método da contagem de caixas ponderadas, a imagem do espaço de fase é dividida em caixas, que correspondem a um grupo de *pixels*. Essa divisão é realizada utilizando grades de diversos tamanhos (40x40, 20x20, 10x10 e 5x5). Neste método, é observado se parte da trajetória do espaço de fase está dentro ou não das caixas. Se estiver dentro da caixa,

atribui-se um peso para cada caixa visitada de acordo com a importância dada aos *pixels* contidos na caixa, como é observado na Expressão (2) [10]:

$$\text{Peso do pixel visitado} = \begin{cases} 2, & \text{Se o pixel é visitado e se possui} \\ & \text{pelo menos um pixel vizinho visitado} \\ 1, & \text{Se o pixel analisado é visitado} \\ 0, & \text{Se o pixel analisado não é visitado} \end{cases} \quad (2)$$

Em Roopaei et al. [10] é afirmado que esse método amplia o valor dos *pixels* contados para um sinal caótico ou atenua o valor de *pixels* contados para um sinal rítmico de acordo com a Expressão (2). Após a atribuição dos pesos da Expressão (2) é contabilizada a quantidade de caixas visitadas de acordo com a atribuição dos pesos e posteriormente é realizado o cálculo do valor de d para cada imagem como na Equação (3).

$$d_{met}(\tau) = \frac{n^{\circ} \text{ de caixas visitadas de acordo com o peso}}{n^{\circ} \text{ total de caixas}}, \quad (3)$$

em que $d_{met}(\cdot)$: índice de complexidade do sistema e o índice met corresponde ao método aplicado, sendo $d_{ccp}(\cdot)$, correspondente à contagem das caixas pelo método da contagem de caixas ponderadas e $d_{sim}(\cdot)$ pelo método da similaridade. Com os valores de $d_{met}(\cdot)$, obtidos a partir das imagens dos espaços de fase, é determinada a diferença absoluta entre estes valores, considerando os diferentes atrasos, cujos valores dependem de cada classe. Na Figura 3 pode ser observado o método aplicado para a grade 10. Foram utilizados os espaços de fase reconstruídos para o τ ótimo mínimo ($\tau_o \text{ min}$) e o τ ótimo médio ($\tau_o \text{ med}$).

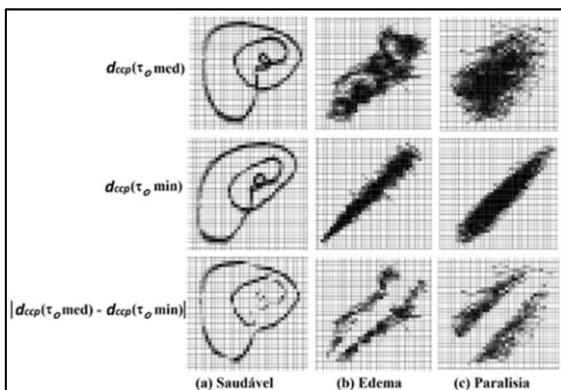


Fig. 3. Imagens obtidas a partir do método da contagem de caixas ponderadas com a grade 10 aplicado em um sinal de voz (a) saudável; (b) afetado por edema; e (c) afetado por paralisia nas pregas vocais.

E. Método da Similaridade

Este método procura estimar o valor da autossimilaridade das imagens dos espaços de fase, medindo o número de *pixels* visitados e a similaridade do comportamento das trajetórias do espaço de fase reconstruído [10]. O método utiliza duas imagens do espaço de fase reconstruído com base em dois tempos de atrasos diferentes. As imagens são divididas em caixas como no método da Seção D. Os valores dos *pixels* são projetados em uma matriz $[A_{ij}]$ com o mesmo tamanho da imagem (i se refere à linha e j se refere à coluna em que se encontram). Posteriormente, são atribuídos pesos para cada *pixel*, onde o *pixel* visitado pelo espaço de fase tem o valor de um e o *pixel* não visitado tem o valor de zero.

Para as duas matrizes encontradas, considerando tempos de atraso distintos, os *pixels* homólogos são multiplicados criando, assim, uma nova matriz. Esta multiplicação, pode ser

interpretada como a aplicação do operador lógico AND. A nova matriz gera uma imagem que representa a intercessão entre as imagens com os diferentes tempos de atraso utilizados. Posteriormente, é calculado o valor do índice $d_{met}(\cdot)$, utilizando a Equação (3) para a nova imagem encontrada. Na Figura 4 está ilustrado o método da similaridade, utilizando a grade 10, com os espaços de fase reconstruídos com os valores de $\tau_o \text{ min}$ e $\tau_o \text{ med}$.

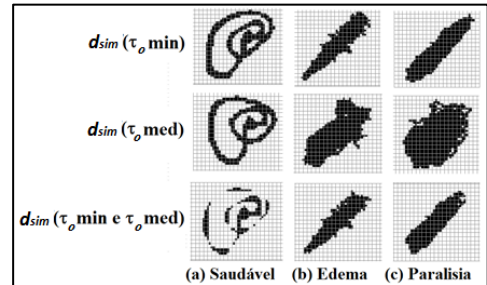


Fig. 4. Imagens do espaço de fase, com a grade 10 aplicado em um sinal de voz (a) saudável; (b) afetado por edema; e (c) afetado por paralisia nas pregas vocais.

F. Classificação dos sinais de Voz

Os sinais foram classificados através de um classificador MLP. Este classificador possui uma rede *feedforward* com algoritmo de aprendizado supervisionado do Gradiente Conjugado Escalonado (SCG) [19, 20]. A quantidade de neurônios na camada de entrada depende da quantidade de características utilizadas para a classificação, na camada oculta foram empregados 09 neurônios e na camada de saída 02 neurônios. Foram considerados quatro casos de classificação: (i) Patologia Vs. Saudável, em que na classe Patologia estão incluídos os sinais afetados por edema e por paralisia nas pregas vocais; (ii) Edema Vs. Saudável; (iii) Paralisia Vs. Saudável; e (iv) Edema Vs. Paralisia. Os sinais utilizados foram separados, aleatoriamente, em dois conjuntos distintos, em que 2/3 foram utilizados para treinamento e 1/3 foram utilizados para teste. A fim de realizar a validação cruzada e dar maior confiabilidade a rede foi executada 10 vezes.

Para os testes, foram realizadas diversas combinações das características de entrada para cada método, a fim de encontrar a combinação que pudesse maximizar os resultados da acurácia. Os métodos não foram combinados, apenas as características dentro dos métodos. Os padrões de treinamento e teste utilizados são as características extraídas a partir dos métodos empregados. Para o Método da Contagem de Caixas Ponderadas foram empregadas sete características: (i) $d_{ccp}(\tau_o \text{ min})$; (ii) $d_{ccp}(\tau_o \text{ med})$; (iii) $d_{ccp}(\tau_o \text{ max})$; (iv) $d_{ccp}(\tau \text{ ótimo})$; (v) Diferença absoluta entre $\tau_o \text{ med}$ e $\tau_o \text{ min}$ $|d_{ccp}(\tau_o \text{ med}) - d_{ccp}(\tau_o \text{ min})|$; (vi) Diferença absoluta entre $\tau_o \text{ max}$ e $\tau_o \text{ med}$ $|d_{ccp}(\tau_o \text{ max}) - d_{ccp}(\tau_o \text{ med})|$ e (vii) Diferença absoluta entre $\tau_o \text{ max}$ e $\tau_o \text{ min}$ $|d_{ccp}(\tau_o \text{ max}) - d_{ccp}(\tau_o \text{ min})|$. Para o Método da Similaridade foram usadas três características: (i) $d_{sim}(\tau_o \text{ min e } \tau_o \text{ med})$; (ii) $d_{sim}(\tau_o \text{ med e } \tau_o \text{ max})$ e (iii) $d_{sim}(\tau_o \text{ min e } \tau_o \text{ max})$.

IV. RESULTADOS

Na Tabela I são apresentados os resultados para a classificação com as medidas individuais para o Método da Similaridade (MSIM). Observa-se que as classificações obtidas para as três classes foram excelentes de acordo com o critério de Hosmer e Lemeshow [21], onde a melhor acurácia média encontrada foi na distinção entre vozes afetadas por edema nas pregas vocais e vozes saudáveis, utilizando como medida $d_{sim}(\tau_o \text{ min e } \tau_o \text{ med})$ e grade 40. De maneira geral, a

grade que obteve melhor desempenho foi a grade 40. Na Tabela II, é apresentada a classificação com as medidas combinadas. As taxas de acurácia média encontradas são superiores 99%, com taxa de especificidade e sensibilidade superiores a 98%. As características $d_{sim}(\tau_o \text{ min e } \tau_o \text{ med})$ e $d_{sim}(\tau_o \text{ med e } \tau_o \text{ max})$ estão presentes em todas as combinações, associadas a outras medidas para a discriminação dos sinais.

Na Tabela III são apresentados os resultados para a classificação com as medidas individuais para o Método da Contagem de Caixas Ponderadas (MCCP) e na Tabela IV, é apresentada a classificação com as medidas combinadas. Na Tabela III observa-se que a classificação obtida para a distinção entre sinais de vozes afetados por edema e sinais saudáveis apresentou a melhor acurácia, sensibilidade e especificidade com taxas superiores a 95% entre as três classificações, obtendo este mesmo desempenho para as grades 5 e 10. A classificação que obteve menor desempenho foi na distinção entre vozes afetadas por paralisia e vozes

saudáveis, com a grade 40, apresentando acurácia de aproximadamente 87%. Nas três classificações, a característica que melhor classificou individualmente foi $|d_{ccp}(\tau_o \text{ med}) - d_{ccp}(\tau_o \text{ min})|$.

Na Tabela IV, as classificações foram eficientes, obtendo taxas de acurácia, sensibilidade e especificidade superiores a 95%, sendo que a melhor grade para as três classificações foi a grade 40. As características $d_{ccp}(\tau \text{ ótimo})$, $d_{ccp}(\tau_o \text{ min})$ estiveram presentes em todas as combinações, associadas a outras medidas para a discriminação entre as classes. Na distinção entre patologias e vozes saudáveis e na distinção entre sinais de vozes afetados por paralisia nas pregas vocais e sinais saudáveis foram utilizadas três características.

Na Tabela V são apresentados os melhores resultados com as características individuais para o método da similaridade e método da contagem de caixas ponderadas referente à discriminação entre sinais de vozes afetados por edema e sinais afetados por paralisia.

TABELA I. RESULTADO PARA A CLASSIFICAÇÃO COM AS CARACTERÍSTICAS INDIVIDUAIS - MÉTODO DA SIMILARIDADE

Classificação	Acurácia (%)	Sensibilidade (%)	Especificidade (%)	Medida	Grade
Patologia vs. Normal	93,29 ± 0,15	95,74 ± 0,46	88,99 ± 0,70	$d_{sim}(\tau_o \text{ min e } \tau_o \text{ med})$	40
Edema vs. Normal	96,91 ± 4,81	100,0 ± 5,26	94,64 ± 0,54	$d_{sim}(\tau_o \text{ min e } \tau_o \text{ med})$	40
Paralisia vs. Normal	92,27 ± 0,44	96,22 ± 1,27	89,66 ± 0,14	$d_{sim}(\tau_o \text{ min e } \tau_o \text{ max})$	40

TABELA II. RESULTADO PARA A CLASSIFICAÇÃO COM AS CARACTERÍSTICAS COMBINADAS - MÉTODO DA SIMILARIDADE

Classificação	Acurácia (%)	Sensibilidade (%)	Especificidade (%)	Medida	Grade
Patologia vs. Normal	99,33 ± 0,22	98,15 ± 0,46	99,48 ± 0,26	$d_{sim}(\tau_o \text{ min e } \tau_o \text{ med}), d_{sim}(\tau_o \text{ med e } \tau_o \text{ max}), d_{sim}(\tau_o \text{ min e } \tau_o \text{ max})$	20
Edema vs. Normal	99,48 ± 0,22	100,0 ± 0,33	100,0 ± 0,38	$d_{sim}(\tau_o \text{ min e } \tau_o \text{ med}), d_{sim}(\tau_o \text{ med e } \tau_o \text{ max})$	20
Paralisia vs. Normal	99,05 ± 0,19	98,15 ± 0,32	100,0 ± 0,39	$d_{sim}(\tau_o \text{ min e } \tau_o \text{ med}), d_{sim}(\tau_o \text{ min e } \tau_o \text{ max}), d_{sim}(\tau_o \text{ med e } \tau_o \text{ max})$	20

TABELA III. RESULTADO PARA A CLASSIFICAÇÃO COM AS CARACTERÍSTICAS INDIVIDUAIS - MÉTODO DA CONTAGEM DE CAIXAS PONDERADAS

Classificação	Acurácia (%)	Sensibilidade (%)	Especificidade (%)	Medida	Grade
Patologia vs. Normal	91,28 ± 0,14	96,63 ± 0,31	83,33 ± 0,07	$ d_{ccp}(\tau_o \text{ med}) - d_{ccp}(\tau_o \text{ min}) $	40
Edema vs. Normal	97,94 ± 0,52	95,65 ± 0,05	100,0 ± 0,94	$ d_{ccp}(\tau_o \text{ med}) - d_{ccp}(\tau_o \text{ min}) $	10
	97,94 ± 2,36	95,65 ± 0,43	100,0 ± 3,18	$ d_{ccp}(\tau_o \text{ med}) - d_{ccp}(\tau_o \text{ min}) $	5
Paralisia vs. Normal	86,67 ± 1,15	95,12 ± 2,15	80,95 ± 0,59	$ d_{ccp}(\tau_o \text{ med}) - d_{ccp}(\tau_o \text{ min}) $	40

TABELA IV. RESULTADO PARA A CLASSIFICAÇÃO COM AS CARACTERÍSTICAS COMBINADAS - MÉTODO DA CONTAGEM DE CAIXAS PONDERADAS

Classificação	Acurácia (%)	Sensibilidade (%)	Especificidade (%)	Medida	Grade
Patologia vs. Normal	95,97 ± 0,26	97,91 ± 0,34	93,40 ± 0,66	$d_{ccp}(\tau \text{ ótimo}), d_{ccp}(\tau_o \text{ min}), d_{ccp}(\tau_o \text{ med})$	40
Edema vs. Normal	98,97 ± 0,34	97,78 ± 0,34	100,0 ± 0,49	$d_{ccp}(\tau \text{ ótimo}), d_{ccp}(\tau_o \text{ min}), d_{ccp}(\tau_o \text{ med}), d_{ccp}(\tau_o \text{ med}) - d_{ccp}(\tau_o \text{ min}) , d_{ccp}(\tau_o \text{ max}) - d_{ccp}(\tau_o \text{ min}) $	40
Paralisia vs. Normal	97,62 ± 0,24	96,33 ± 0,40	100,0 ± 0,31	$d_{ccp}(\tau \text{ ótimo}), d_{ccp}(\tau_o \text{ min}), d_{ccp}(\tau_o \text{ max}) - d_{ccp}(\tau_o \text{ min}) $	40

Os métodos implementados apresentaram taxas de classificação superiores a 79% na discriminação entre patologias. Sendo que o método da contagem de caixas ponderadas possui a vantagem de ter a grade 40 no melhor resultado, que proporciona menor custo computacional que a grade 5.

Na Tabela VI são apresentados os melhores resultados com as medidas combinadas para o método da similaridade e

método da contagem de caixas ponderadas referente à discriminação entre sinais de vozes afetados por edema e sinais afetados por paralisia.

Com a combinação das medidas (Tabela VI), a acurácia aumentou para cerca de 85%, o que demonstra o aumento do potencial discriminativo das medidas de forma combinada, na discriminação entre patologias. Ambos os métodos obtiveram resultados semelhantes, onde o diferencial foi a quantidade de medidas utilizadas por cada método, tendo destaque o método

da contagem de caixas ponderadas, com o uso de apenas três medidas na etapa de classificação.

TABELA V. RESULTADO PARA A CLASSIFICAÇÃO ENTRE PATOLOGIAS – MSIM E M CCP (CARACTERÍSTICAS INDIVIDUAIS).

Classe	Acurácia(%)	Método	Medida/ Grade
Edema Vs. Paralisia	81,77 ± 0,61	MCCP	$d_{ccp}(\tau_o \text{ min})$
	79,17 ± 0,39	MSIM	$d_{sim}(\tau_o \text{ min e } \tau_o \text{ med})$
			40 5

TABELA VI RESULTADO PARA A CLASSIFICAÇÃO ENTRE PATOLOGIAS – MSIM E M CCP (CARACTERÍSTICAS COMBINADAS).

Classe	Acurácia(%)	Método	Medida/ Grade
Edema Vs. Paralisia	85,42 ± 0,84	MCCP	$d_{ccp}(\tau_o \text{ min}), d_{ccp}(\tau_o \text{ med}), d_{ccp}(\tau_o \text{ max}) - d_{ccp}(\tau_o \text{ min}) $
	85,94 ± 1,03	MSIM	$d_{sim}(\tau_o \text{ min e } \tau_o \text{ med}), d_{sim}(\tau_o \text{ min e } \tau_o \text{ max}), d_{sim}(\tau_o \text{ med e } \tau_o \text{ max})$
			20 20

V. CONCLUSÕES

Os métodos da similaridade e da contagem de caixas ponderadas obtiveram bons resultados, mas de maneira individual, o maior destaque foi dado para o método da contagem de caixas ponderadas. Esperava-se que as medidas extraídas com o τ ótimo obtivessem maior êxito na classificação, pois esta seria a característica que melhor demonstraria a dinâmica do sistema, o que não ocorreu, pelo menos de maneira individual. As características de maneira individual que melhor classificaram os sinais foram as medidas que utilizaram $\tau_o \text{ min}$ e $\tau_o \text{ med}$. Entretanto, na classificação com as medidas combinadas, as características extraídas com o τ ótimo, estão presentes nos melhores resultados de classificação.

Quando comparados os resultados obtidos com os trabalhos de Couras et al. e Pinho et al. [8, 9] observa-se que os métodos da similaridade e da contagem de caixas ponderadas obtiveram resultados semelhantes em alguns casos de classificação, mas em outros casos (como na discriminação entre vozes afetadas por paralisia e vozes saudáveis e entre vozes afetadas por edema e vozes afetadas por paralisia nas pregas vocais) os resultados foram melhores, obtendo taxas de acurácia superiores a 95%.

Há variação nos resultados, de acordo com a grade empregada. Entre as quatro grades, de maneira geral, as grades 10 e 40 obtiveram maior frequência nos resultados com as características individuais, sendo possível escolher uma dessas grades para aplicar os métodos. A grade 40 possui maior destaque porque ela requer menor custo computacional comparada as outras grades. Em alguns casos, a classificação com as características individuais foram bem semelhantes à classificação com as características combinadas, como no caso da discriminação entre sinais de voz afetados por edema e vozes saudáveis. De uma maneira geral, pode-se observar que as medidas extraídas a partir de imagens do espaço de fase reconstruído apresentam potencial discriminativo excelente de acordo com [21], para identificar características inerentes de cada classe e, assim, discriminar as classes, estudadas nesta pesquisa, com precisão.

AGRADECIMENTOS

Os autores agradecem a PRPIPG e a COPEX do Instituto Federal de Educação Ciência e Tecnologia da Paraíba, Campus João Pessoa, pelo financiamento parcial da pesquisa e ao PPgEE pelo apoio científico.

REFERÊNCIAS

- [1] H. P. Barbosa, J. M. Fachine and J. E. Rangel, “Classificação de Patologias da Fala a partir do PPM”, *Congresso Brasileiro de Inteligência Computacional*, Mar. 2016. [Online].
- [2] N. F. R. Alves, “Diagnóstico Inteligente de Patologias da Laringe”, Dissertação de Mestrado, Escola Superior de Tecnologia e Gestão Instituto Politécnico de Bragança, POR, 2016. [Online].
- [3] J. Y. Lee, S. H. Choi, J. J. Jiang, M. S. Hahn, H-S. Choi, “Perturbation and Nonlinear Dynamic Analysis of Sustained Vowels in Saudavel and Pathological Voices”, *Journal of the Korean Society of Speech Science*, vol.2, n° 1, pp. 113-120, Fev. 2010.
- [4] M. E. Dajer, *Análise de Sinais de Voz com Padrões Visuais de Dinâmica Vocal*, Tese de doutorado, Escola de Engenharia de São Carlos, Universidade de São Paulo, SP, 2010. [Online].
- [5] M. Behlau, *Voz, o Livro do Especialista*. Vol. 1, Ed. Revinter, Rio Janeiro-RJ, 2008.
- [6] V. J. D. Vieira, *Avaliação de Distúrbios da Voz por meio de Análise de Quantificação de Recorrência*, Dissertação de mestrado, Instituto Federal da Paraíba, PB, p.218, 2014.
- [7] R. Behroozmand, F. Almasganj, M. H. Moradi, “Pathological Assessment of Vocal Fold Nodules and Polyp Using Acoustic Perturbation and Phase Space Features”, *IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, vol.2, Mai. 2006.
- [8] M. F. K. B. Couras, P. H. U. de Pinho, E. C. S. Dantas, S. L. N. C. Costa, S. E.N. Correia, “Detecção De Patologias Laringeas usando o Método de Contagem de Caixas no Espaço de Fase de Sinais de Fala”, *XXV CBEb*, Nov. 2016. [Online].
- [9] P. H. U. de Pinho, M. F. K. B. Couras, E. C. S. Dantas, S. L. N. C. Costa, S. E.N. Correia, “Classificação de patologias laringeas por meio de características do espaço de fase reconstruído”, *XXXIV SBrT*, Set. 2016. [Online].
- [10] M. Roopaei, R. Boostani, R. R. Sarvestani, M. A. Taghavi, Z. Azimifar, “Chaotic based Reconstructed Phase Space Features for Detecting Ventricular Fibrillation”, *Biomedical Signal Processing and Control*, vol. 5, pp. 318-327, Out. 2010. [Online].
- [11] A. Amann, R. Tratnig, K. Unterkofler, “Detecting Ventricular Fibrillation by Time-Delay Methods”, *IEEE Transactions on Biomedical Engineering*, vol. 54, pp. 174-177, Jan. 2007. [Online].
- [12] F. Takens, “Detecting Strange Attractores in Turbulence”, *Dynamical Systems and Turbulence*, Warwick 1980. pp. 366-381, 1981.
- [13] D. G. Galdino, *Análise Acústica não Linear dos Padrões Visuais de Dinâmica Vocal (PVDV) de Homens Adultos*, Dissertação de mestrado, Escola de Engenharia de São Carlos, Instituto de Química de São Carlos da Universidade de São Paulo, SP, 2012.
- [14] T. M. Cover, J. A. Thomas, “Elements of Information Theory”, (*Wiley Series in Telecommunications and Signal Processing*), 2ª Edição, Hardcover, 23 Sep 2007.
- [15] M. A. Savi, *Dinâmica Não Linear e Caos*. Editora E-papers, 2006.
- [16] A. M. Fraser, H. L. Swinney, “Independent Coordinates for Strange Attractors from Mutual Information”, *Physical Review A*, vol. 33, no. 2, p. 1134, 1986.
- [17] Kay Elemetrics, *Disordered Voice Database*, Model 4337. 3th ed. United States, 1994.
- [18] N. Marwan, *Cross Recurrence Plot Toolbox for Matlab, Version 5.17 (r28.16)*, 2013. Disponível em <http://tocsy.pik-potsdam.de/CRPtoolbox/index.html>. Acesso em fevereiro de 2013.
- [19] S. Haykin, *Redes Neurais, Princípios e Práticas*, Porto Alegre, RS. BOOKMAN, 2001.
- [20] B. Simsek, Y. Tansel, E. Hüseyin, “Hybridizing a Fuzzy Multi-response Taguchi Optimization Algorithm with Artificial Neural Networks to solve Standard Ready-Mixed Concrete Optimization Problems” *International Journal of Computational Intelligence Systems*, Vol. 9, No. 3; pp.525-543, 2016.
- [21] D. W. Jr. Hosmer, S. Lemeshow, *Applied logistic regression*. 2ª ed. New York: John Wiley & Sons; 2000.