

Coeficientes de Hilbert-Huang-Hurst aplicados à Classificação de Distúrbios Vocais

Vinicus J. D. Vieira, Emilly Duarte, Silvana C. Costa e Suzete E. N. Correia

Resumo—Neste trabalho, o vetor de coeficientes Hilbert-Huang-Hurst, conhecido como HHHC (*Hilbert-Huang-Hurst Coefficients*), é aplicado como um conjunto de características acústicas para classificação de vozes saudáveis e vozes afetadas por patologias laríngeas. Essa abordagem leva em consideração os aspectos não lineares e não estacionários relacionados à produção vocal, empregando a decomposição empírica de modos (EMD, *Empirical Mode Decomposition*) e realizando a extração do expoente de Hurst. Uma variação da EMD, conhecida como *Ensemble* EMD (EEMD), também é empregada na extração do vetor HHHC. Nos experimentos são consideradas as patologias paralisia, edema e nódulos nas pregas vocais. Como classificadores, são utilizados a Análise Discriminante Linear (*Linear Discriminant Analysis* – LDA) e a Análise Discriminante Quadrática (*Linear Discriminant Analysis* – QDA). O desempenho da classificação atinge uma acurácia média máxima de 90,55% com um vetor HHHC de dimensão 4. Os resultados indicam que o vetor HHHC é uma medida promissora na classificação de distúrbios da voz.

Palavras-Chave—Processamento de sinais de voz, Classificação de patologias laríngeas, decomposição empírica de modos, expoente de Hurst.

Abstract—In this work, the Hilbert-Huang-Hurst coefficients vector, known as HHHC, is applied as a set of acoustic features for the classification of healthy voices and voices affected by laryngeal pathologies. This approach takes into account the nonlinear and non-stationary aspects related to vocal production, employing empirical mode decomposition (EMD) and extracting the Hurst exponent. A variation of EMD, known as Ensemble EMD (EEMD), is also employed in the HHHC extraction. The experiments consider pathologies such as paralysis, edema, and nodules on the vocal folds. As classifiers, it is employed Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA). The classification performance achieves a maximum average accuracy of 90.55% with a 4-dimensional HHHC vector. The results indicate that the HHHC vector is a promising measure in classifying voice disorders.

Keywords—Speech signal processing, laryngeal pathology classification, empirical mode decomposition, Hurst exponent.

I. INTRODUÇÃO

A voz humana desempenha um papel fundamental na comunicação e expressão, sendo uma ferramenta vital tanto no contexto social quanto profissional. No entanto, a presença de distúrbios vocais pode impactar significativamente a qualidade de vida e a eficácia da comunicação. Os distúrbios da voz podem manifestar-se de diversas formas, incluindo rouquidão, aspereza, falta de volume e até mesmo a completa perda da voz [1], [2]. A análise acústica de distúrbios vocais desempenha um papel crucial nesse contexto, permitindo uma

avaliação objetiva das características acústicas da voz [3]. Essa abordagem não apenas auxilia no diagnóstico diferencial, mas também no monitoramento da progressão dos distúrbios e na avaliação dos resultados de intervenções terapêuticas.

Na literatura, diversos trabalhos buscam por conjuntos de características robustas para a discriminação de distúrbios vocais [4–6]. Alguns estudos propõem métricas acústicas fundamentadas em um modelo linear de produção vocal [3], [7] e outros sugerem medidas baseadas em um modelo não linear [4], [6], [8]. Além disso, a presença de patologias laríngeas provocam variações acústicas não estacionárias no sinal de voz [9]. O grau de não estacionariedade no processo de produção vocal tem influência na extração de medidas acústicas lineares e não lineares para classificação de vozes saudáveis e patológicas [10].

O vetor HHHC foi proposto em [11] como um método de análise de variações acústicas não lineares e não estacionárias em sinais de voz afetados por estados emocionais. Nesta abordagem, a decomposição EMD é empregada para obter representações de um sinal em diferentes regiões de frequência. Além disso, o expoente de Hurst é estimado em cada uma dessas regiões, sendo cada valor de Hurst um componente do vetor HHHC. Na sua proposta original, foi observado que os melhores resultados de classificação foram obtidos com a utilização de uma variação da EMD, conhecida como *ensemble* EMD (EEMD).

Neste trabalho, é realizada uma investigação da aplicabilidade do vetor HHHC no contexto de distúrbios da voz introduzidos por patologias laríngeas. A principal contribuição deste estudo é empregar dois tipos de decomposição adaptativa (EMD e EEMD) a fim de verificar se a extração do vetor HHHC é robusta para classificação de distúrbios vocais.

O restante do texto está organizado da seguinte forma: Na Seção II são apresentadas brevemente as definições relacionadas ao vetor HHHC. Na Seção III é apresentada a metodologia empregada neste trabalho. Na Seção IV são apresentados os resultados obtidos nos experimentos realizados e, na Seção V, é apresentada a conclusão.

II. COEFICIENTES DE HILBERT-HUANG-HURST

O vetor HHHC, ou vetor de Coeficientes de Hilbert-Huang-Hurst, é um atributo não linear que caracteriza os efeitos das variações acústicas na fonte de excitação (sistema de produção vocal) [12]. A decomposição baseada na EMD é utilizada na tarefa de detecção de distúrbios no sinal de voz. Ou seja, a partir da decomposição, as variações acústicas introduzidas pelos distúrbios vocais são enfatizadas. Na proposta original do HHHC, foi empregada uma alternativa à EMD, conhecida como EEMD.

A partir da EMD ou da EEMD, o expoente de Hurst, que é relacionado com a fonte de excitação [13], é estimado do

Vinicus J. D. Vieira, Depto. de Processamento de Voz, SiDi, Alpha-ville, Campinas, SP - Brasil; Emilly Duarte, Silvana C. Costa e Suzete E. N. Correia, Programa de Pós-Graduação em Engenharia Elétrica, Instituto Federal de Educação, Ciência e Tecnologia da Paraíba (PPGEE/IFPB). E-mails: {vinicius.dv@sidi.org.br, emilly.josy@academico.ifpb.edu.br, silvana@ifpb.edu.br, suzete@ifpb.edu.br}.

processo de decomposição resultante para formar a matriz de atributos. Essas abordagens e a metodologia de extração do vetor HHHC são apresentados a seguir.

A. EMD

A abordagem EMD consiste em analisar um sinal $x(t)$ que apresenta dois extremos consecutivos (máximo ou mínimo) nos pontos t_- e t_+ . Isso resulta na definição de uma componente de alta frequência do sinal, também conhecida como detalhe $d(t)$, e uma componente de tendência local, ou resíduo $r(t)$, tal que:

$$x(t) = d(t) + r(t), \quad t_- \leq t \leq t_+. \quad (1)$$

O conjunto de detalhes locais, obtidos a partir de todos os extremos consecutivos de $x(t)$, compõe a primeira Função de Modo Intrínseco (IMF - *Intrinsic Mode Function*) da decomposição. Essa separação entre componentes de alta e baixa frequência é repetida de forma iterativa sobre o sinal residual $r(t)$, resultando em um conjunto de IMFs e um resíduo de baixa frequência.

Para realizar a EMD em um sinal $x(t)$, os seguintes passos são seguidos [14], [15]:

- 1) Identificação de todos os extremos de $x(t)$, ou seja, os pontos de máximo $x_{max}(t)$ e mínimo $x_{min}(t)$ locais;
- 2) Obtenção das envoltórias $e_{max}(t)$ e $e_{min}(t)$, interpolando os pontos de máximo e mínimo, respectivamente. Para isso, é utilizada a interpolação polinomial de terceiro grau por meio do método de *spline*¹;
- 3) Cálculo do resíduo como a média entre as envoltórias: $r(t) = (e_{min}(t) + e_{max}(t)) / 2$;
- 4) Extração das componentes de detalhes: $d(t) = x(t) - r(t)$;
- 5) Iteração sobre o sinal residual $r(t)$.

Para que o passo (5) seja eficaz, a componente de detalhe $d(t)$ obtida no passo (4) deve obedecer a duas propriedades: a primeira está relacionada à diferença entre a quantidade de extremos e a quantidade de cruzamentos em zero, a qual deve ser nula ou igual a um; a outra propriedade diz respeito ao valor médio definido pelas envoltórias dos máximos e mínimos, o qual deve ser nulo. Se essas propriedades não forem atendidas, os passos (1-4) são repetidos, com $d(t)$ substituindo $x(t)$. Esse processo, conhecido como *sifting*, é repetido até garantir que a nova função $d(t)$ seja considerada uma IMF [14].

O algoritmo do método EMD garante que qualquer sinal $x(t)$ possa ser decomposto em um número finito (K) de iterações e pode ser expresso como:

$$x(t) = \sum_{k=1}^K \text{IMF}_k(t) + r(t), \quad (2)$$

em que $\text{IMF}_k(t)$, $1 \leq k \leq K$, são as funções de detalhes $d(t)$ obtidas no passo (4) de cada iteração e $r(t)$ é o sinal residual resultante da última iteração.

¹Método matemático que utiliza funções cúbicas segmentadas para aproximar uma curva suave que passa por um conjunto de pontos dados.

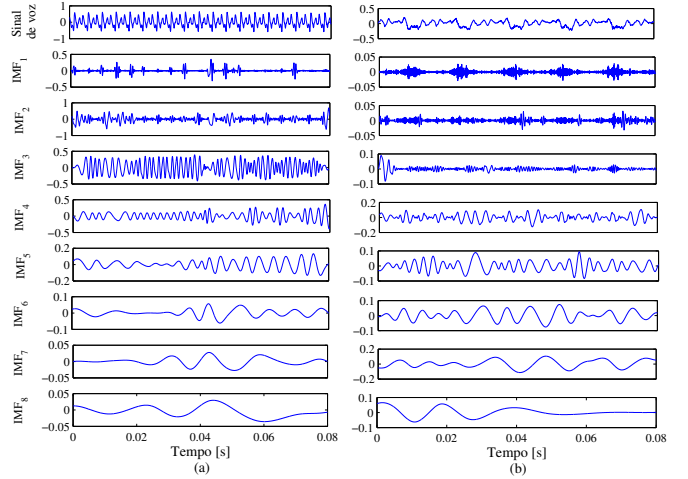


Fig. 1. Primeiras oito IMFs obtidas com a EMD a partir de segmentos de 80 ms de sinais de voz: (a) saudável; (b) patológico (paralisia).

Exemplos da decomposição empírica de modos são apresentados na Figura 1, com $K = 8$. São observados trechos de 80 ms de um sinal de voz saudável e um sinal de voz afetado por paralisia nas pregas vocais. A redução no número de extremos de um modo para o próximo implica que, localmente, as primeiras IMFs possuem oscilações mais rápidas (altas frequências) que as IMFs de maior índice. Apesar de se tratar de uma mesma tarefa de emissão sonora (vogal sustentada /a/), é possível notar diferenças entre os dois sinais no que diz respeito ao padrão de oscilação e à amplitude dessas oscilações nas diferentes IMFs. Portanto, torna-se razoável supor que haja um potencial discriminativo entre sinais de vozes saudáveis e patológicas quando se observa características intrínsecas nas IMFs.

B. EEMD

A EMD pode apresentar um fenômeno conhecido como *mode mixing* (ou mistura entre modos), que acontece quando uma IMF pode conter sinais de diferentes escalas ou quando sinais de escalas similares estão contidos em diferentes IMFs [16].

O método EEMD [16] foi proposto para evitar o *mode mixing*. Este método utiliza ruído Gaussiano branco na decomposição EMD [15], [17]. As IMFs na EEMD são definidas como sendo a média das IMFs correspondentes obtidas na EMD aplicado a uma quantidade I de sinais corrompidos $x^i(t)$, pela adição de seqüências de ruído Gaussiano branco, $w^i(t)$, ao sinal original, $x(t)$. Assim,

$$x^i(t) = x(t) + w^i(t), \quad i = 1, \dots, I. \quad (3)$$

Os K modos $\text{IMF}_k^i(t)$ do sinal corrompido são obtidos da aplicação do EMD em $x^i(t)$. O k -ésimo modo de $x(t)$ é dado por:

$$\overline{\text{IMF}}_k = \frac{1}{I} \sum_{i=1}^I \text{IMF}_k^i(t). \quad (4)$$

Em termos gerais, a ideia básica do método EEMD é adicionar o ruído branco de forma que ele seja distribuído

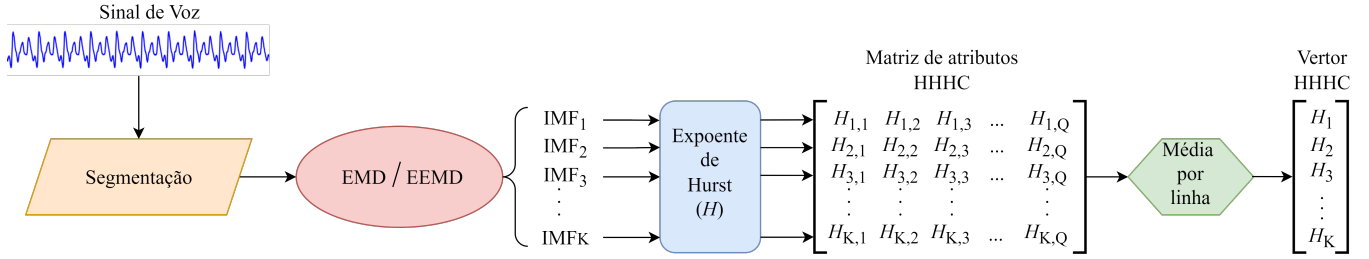


Fig. 2. Metodologia de extração do vetor HHHC, considerando como entrada um sinal de voz.

uniformemente em todo o espaço tempo-frequência. Com a estrutura similar a um banco de filtros² resultante da EMD, o ruído branco adicionado ao sinal posiciona as componentes de frequência de forma mais apropriada possível em uma escala mais adequada do que a EMD.

C. Expoente de Hurst

O expoente de Hurst ($0 \leq H \leq 1$), ou coeficiente de Hurst, expressa a dependência temporal ou grau de escala de um processo estocástico [18]. Essa medida também pode ser definida pela taxa de decaimento da função coeficiente de autocorrelação (*Autocorrelation Coefficient Function – ACF*) $\rho(k)$ ($-1 < \rho(k) < 1$) com $k \rightarrow \infty$. Seja um sinal de voz representado por um processo estocástico $x(t)$, com a função coeficiente de autocorrelação normalizada ($\rho(k)$) definida por

$$\rho(k) = \frac{E[(x(t) - \mu_x)(x(t+k) - \mu_x)]}{E[(x(t) - \mu_x)^2]}, \quad (5)$$

em que μ_x é a média de $x(t)$ e k é o atraso. O comportamento assintótico de $\rho(k)$ é dado por

$$\rho(k) \sim H(2H - 1)k^{2(H-2)}, \quad k \rightarrow \infty. \quad (6)$$

D. O vetor HHHC

A extração do vetor HHHC é realizada em duas etapas: decomposição pelo método EMD ou EEMD, e uma estimativa multicanal do expoente de Hurst. Um exemplo da estimativa do vetor HHHC é mostrado na Figura 2. A decomposição baseada em EMD é aplicada no sinal de entrada, que pode ser um trecho sonoro do sinal de voz dividido em Q segmentos. Então, os valores de H são obtidos quadro a quadro de cada uma das K IMFs. Neste caso, da IMF₁ a IMF_K, tem-se H_1 a H_K para cada um dos seguimentos, formando a matriz de atributos HHHC. Assim, o vetor HHHC $[H_1, H_2, \dots, H_K]$, do sinal de voz de entrada, é construído como a média da matriz de atributos HHHC.

III. METODOLOGIA

Para o desenvolvimento deste trabalho, foi utilizada a base de voz desenvolvida pelo *Kay Massachusetts Eye and Ear Infirmary (MEEI) Voice and Speech Lab* [19], conhecida como *Disordered Voice Database, Model 4337*. Dessa base, que consiste de sinais de vogal sustentada (*/a/*), são analisados 53

casos de vozes saudáveis e 114 vozes afetadas por patologias laringeas (52 sinais de vozes afetadas por paralisia nas pregas vocais, 44 sinais de vozes afetadas por edema de Reinke e 18 sinais de vozes afetadas por nódulos vocais). A taxa de amostragem considerada na composição da base foi 50 kHz para os sinais saudáveis e 25 kHz para os sinais patológicos. Por convenção, neste trabalho adota-se 25 kHz também para os sinais saudáveis (por meio de *downsampling*) para evitar qualquer influência deste contexto na classificação.

Para o cenário experimental deste trabalho, é utilizada a linguagem de programação Python. A segmentação dos sinais de voz ocorre com um tamanho de quadro de 80 ms, com sobreposição de 50%. Para a obtenção das IMFs com as abordagens EMD e EEMD (com $I = 100$), é utilizada a biblioteca PyEMD³. A estimativa do expoente de Hurst é realizada com o método R/S [20], disponível na biblioteca *hurst*⁴. Dois classificadores são empregados para analisar o potencial discriminativo do vetor HHHC: o LDA (*Linear Discriminant Analysis*) e o QDA (*Quadratic Discriminant Analysis*). Esses classificadores são implementados por meio da biblioteca *scikit-learn*⁵.

O método *k-fold* de validação cruzada, com $k = 10$, é utilizado para dar mais confiabilidade aos resultados, pois tem sido comumente utilizado em classificação de distúrbios vocais [6], [7], [21]. Para retirar qualquer influência do tamanho amostral nos resultados, no processo de classificação foram selecionados aleatoriamente, entre todas as patologias, 53 sinais, para atingir a mesma quantidade de sinais saudáveis.

As medidas de acurácia, sensibilidade e especificidade são utilizadas para analisar o desempenho dos classificadores. Essas medidas estão relacionadas à capacidade de um classificador em diagnosticar uma doença (sensibilidade), diagnosticar um estado saudável (especificidade), bem como medir seu desempenho global (acurácia) [6], [9].

IV. RESULTADOS

Na Figura 3 são apresentados os valores médios do expoente de Hurst extraído de cada uma das oito IMFs dos sinais de vozes saudáveis e patológicos. Ou seja, são apresentadas as médias de cada um dos componentes do vetor HHHC. Note-se que, com EMD, há uma separação mais evidente entre saudável e patológico nos índices de número 1, 2, 3 e 6. Ainda, é possível observar que os valores de Hurst para patologias

³<https://github.com/laszukdawid/PyEMD> [último acesso em 22/03/2024].

⁴<https://github.com/Mottl/hurst> [último acesso em 22/03/2024].

⁵<https://scikit-learn.org/stable/index.html> [último acesso em 22/03/2024].

²A decomposição EMD pode ser considerada um banco de filtros diádico para o caso de ruído Gaussiano branco [16].

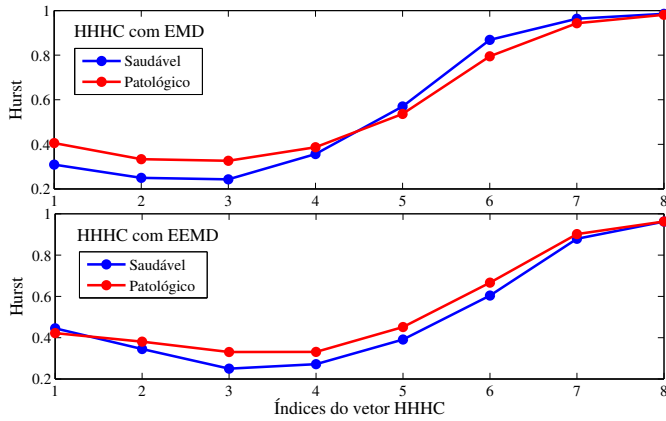


Fig. 3. Média dos valores do Expoente de Hurst, para cada uma das oito IMFs (índices do vetor HHHHC) obtidas com EMD e EEMD, respectivamente, considerando as classes de sinais saudáveis e patológicos.

são maiores em relação aos sinais saudáveis considerando as quatro primeiras IMFs (índices do vetor HHHHC). Este comportamento se inverte a partir do quinto índice. Por outro lado, considerando a decomposição EEMD, nota-se que para quase todos os índices do vetor HHHHC os sinais patológicos possuem maiores valores médios de Hurst. Neste caso, pode ser observada uma separação maior para os índices de número 3, 4, 5 e 6. De maneira geral, observa-se que a EMD permite uma ênfase das diferenças entre saudável e patológico nas frequências mais altas do sinal (primeiras IMFs). A EEMD, por sua vez, ao aplicar ruído branco no sinal, parece enfatizar as diferenças entre saudável e patológico em frequências intermediárias. A seguir, são apresentados os resultados da classificação com LDA e QDA, como forma de realizar uma análise objetiva da discriminação entre saudável e patológico utilizando o vetor HHHHC.

A. Resultados com o classificador LDA

Os resultados da classificação com LDA considerando o vetor HHHHC extraído com EMD e EEMD, respectivamente, são apresentados na Tabela I. Na classificação individual (Ind.) e considerando todas as possíveis combinações (Comb.) entre índices do vetor HHHHC, o classificador LDA obteve acurácia média acima de 80%. Como destaque no contexto da EMD, os maiores valores de acurácia, sensibilidade e especificidade foram obtidos com um conjunto de 5 medidas, ou seja, valores de expoente de Hurst obtidos de cinco IMFs. Em termos de acurácia ($\mu = 88,64\%$), o desempenho foi superior à classificação individual em aproximadamente 5 pontos percentuais (p.p.). Por outro lado, no contexto da EEMD, a acurácia e a sensibilidade atingiram seus valores máximos com a combinação de 6 medidas, sendo a acurácia aproximadamente 8 p.p. mais elevada em relação à classificação individual. A especificidade, por sua vez, atingiu seu valor médio máximo com três combinações (4 a 4, 5 a 5 e 6 a 6).

B. Resultados com o classificador QDA

Na Tabela II são apresentados os resultados da classificação utilizando QDA para os casos da extração do HHHHC com

TABELA I

DESEMPENHO DO CLASSIFICADOR LDA COM HHHHC BASEADO EM EMD E EM EEMD.

Decomposição com EMD				
Comb.	Ac (%)	Sen (%)	Esp (%)	Medidas
Ind.	82,18±3,53	90,32±3,48	74,89±6,40	H_2
2a2	83,27±3,57	87,40±4,32	77,13±4,75	H_1, H_3
3a3	86,00±2,33	89,07±3,77	79,99±4,90	H_1, H_2, H_3
4a4	87,73±2,68	90,14±3,50	82,90±4,99	H_1, H_2, H_3, H_4
5a5	88,64±2,91	92,14±3,43	83,14±4,85	H_1, H_3, H_4, H_6, H_7
6a6	87,82±2,62	91,57±3,58	81,48±5,10	$H_1, H_2, H_3, H_4, H_7, H_8$
7a7	86,82±2,71	90,14±3,50	81,48±5,10	$H_1, H_2, H_3, H_4, H_6, H_7, H_8$
8	86,00±2,95	87,57±3,53	81,48±5,10	Todas
Decomposição com EEMD				
Comb.	Ac (%)	Sen (%)	Esp (%)	Medidas
Ind.	80,91±3,88	82,04±4,86	78,48±5,83	H_3
2a2	84,73±4,02	83,98±4,78	84,90±6,45	H_3, H_4
3a3	84,82±3,99	83,80±4,87	84,67±6,87	H_3, H_7, H_8
4a4	87,55±3,88	85,62±5,09	88,00±6,29	H_3, H_5, H_7, H_8
5a5	87,73±4,17	84,81±5,75	88,00±6,29	H_3, H_4, H_5, H_7, H_8
6a6	88,64±3,93	86,81±5,18	88,00±6,29	$H_1, H_3, H_4, H_5, H_7, H_8$
7a7	86,73±4,23	84,81±5,75	86,57±6,17	$H_1, H_2, H_3, H_4, H_5, H_6, H_7, H_8$
8	85,73±4,01	83,56±5,54	86,57±6,17	Todas

TABELA II

DESEMPENHO DO CLASSIFICADOR QDA COM HHHHC BASEADO EM EMD E EM EEMD.

Decomposição com EMD				
Comb.	Ac (%)	Sen (%)	Esp (%)	Medidas
Ind.	81,27±3,91	91,57±3,58	69,89±7,79	H_2
2a2	87,00±2,26	89,57±3,61	83,65±4,12	H_2, H_3
3a3	88,73±2,21	88,48±6,26	88,08±5,06	H_1, H_2, H_4
4a4	90,55±2,71	89,90±6,35	87,67±6,08	H_1, H_2, H_3, H_4
5a5	89,64±2,09	91,81±3,53	86,42±4,91	H_1, H_2, H_3, H_4, H_5
6a6	90,55±3,00	86,48±7,09	93,00±2,73	$H_1, H_2, H_3, H_4, H_7, H_8$
7a7	89,82±2,71	87,48±3,90	93,75±2,45	$H_1, H_3, H_4, H_5, H_6, H_7, H_8$
8	88,73±3,49	86,29±5,06	93,75±2,45	Todas
Decomposição com EEMD				
Comb.	Ac (%)	Sen (%)	Esp (%)	Medidas
Ind.	80,91±3,88	82,04±4,86	78,48±5,83	H_3
2a2	88,45±3,89	90,38±4,92	83,33±7,07	H_3, H_5
3a3	88,45±3,89	88,38±5,65	85,00±6,65	H_3, H_5, H_6
4a4	89,45±2,82	93,99±3,33	81,33±7,24	H_2, H_3, H_5, H_7
5a5	88,55±3,97	87,29±4,70	90,00±5,37	H_3, H_5, H_6, H_7, H_8
6a6	89,45±2,88	88,56±4,43	88,00±6,72	$H_1, H_2, H_3, H_4, H_5, H_8$
7a7	87,45±3,61	89,64±3,67	83,00±7,46	$H_2, H_3, H_4, H_5, H_6, H_7, H_8$
8	85,55±3,70	86,06±5,08	83,00±7,46	Todas

EMD e EEMD, respectivamente. Assim como pode ser observado na classificação com LDA, a classificação com QDA proporciona valores de acurácia acima de 80% em todos os cenários possíveis. Considerando a decomposição EMD, o maior valor médio de acurácia ($\mu = 90,55\%$) foi atingido com uma combinação de quatro medidas do expoente de Hurst, sendo aproximadamente 9 p.p. acima da classificação com uma única medida. O maior valor de sensibilidade foi obtido na combinação 5 a 5 e o maior valor de especificidade foi obtido nas combinações de 7 e de 8 medidas de Hurst. No contexto da classificação com EEMD, o maior valor de acurácia ($\mu = 89,45\%$) também aparece em uma combinação de quatro medidas e com 9 p.p. acima da acurácia obtida com a classificação individual. O destaque de sensibilidade também ocorre na combinação de 4 medidas e o maior valor de especificidade ocorre com a combinação de 5 medidas.

C. Discussão

A seleção de métricas acústicas confiáveis para representar cada tipo de distúrbio vocal é uma tarefa desafiadora, pois depende de vários fatores, como o tipo e a gravidade da lesão, a severidade dos seus efeitos e a presença de ruído no sinal

de voz analisado [7], [21]. Em [22] a EMD foi empregada para estimação de frequência fundamental e classificação de patologias vocais com sinais sintéticos e reais, atingindo acurácia acima de 90% em ambos os contextos.

No que diz respeito ao expoente de Hurst, esta medida já foi empregada individualmente para classificação de desordens vocais [23], atingindo acurácia de aproximadamente 87%. Neste trabalho, o vetor HHHC proporciona ao classificador um desempenho acima de 90% unindo a EMD com a estimação do expoente de Hurst. Vale ressaltar que a decomposição EMD tem um custo computacional menor que a EEMD, uma vez que esta última precisa acrescentar todas as realizações de ruído branco (neste trabalho, 100 realizações) no processamento.

Na proposta original do HHHC [11], uma fusão desta medida com o índice de não estacionariedade (INS – *index of non-stationarity*) representou uma melhoria acima de 3 p.p. na classificação de estados emocionais. Como o INS também proporcionou desempenho acima de 90% ao classificador [9], em trabalhos futuros pretende-se fazer tal fusão (HHHC+INS) a fim de melhorar a acurácia. Todavia, os resultados encontrados neste trabalho já apresentam acurácia aceitável para auxílio clínico (acurácia $\geq 70\%$) [7].

V. CONCLUSÕES

Este trabalho apresentou a análise da classificação de vozes saudáveis e vozes patológicas empregando o vetor HHHC. Para tanto, foi investigado o potencial discriminativo desta medida acústica utilizando, na sua extração, as decomposições EMD e EEMD a fim de observar qual das duas abordagens proporciona um melhor desempenho de classificação. Para tanto, um classificador linear (LDA) e outro não linear (QDA) foram utilizados nos experimentos. Os resultados obtidos mostraram que ambos os classificadores atingem um desempenho acima de 80% de acurácia em todos os cenários testados. O maior valor de acurácia foi de 90,55%, com QDA, empregando um vetor HHHC com 4 componentes e sendo extraído com EMD. Estes resultados indicam que o vetor HHHC tem potencial para ser empregado na classificação de patologias laringeas por meio da voz.

AGRADECIMENTOS

Este trabalho não seria possível sem o apoio do Instituto SiDi e do Programa de Pós-Graduação em Engenharia Elétrica do Instituto Federal de Educação, Ciência e Tecnologia da Paraíba.

REFERÊNCIAS

- [1] L. W. Lopes, H. F. d. Silva, D. d. S. Evangelista, J. D. d. Silva, L. B. Simões, P. O. Costa e Silva, M. F. B. d. Lima-Silva, and A. A. F. d. Almeida, “Relação entre os sintomas vocais, intensidade do desvio vocal e diagnóstico laríngeo em pacientes com distúrbios da voz,” in *CoDAS*, vol. 28, pp. 439–445, SciELO Brasil, 2016.
- [2] M. L. V. Masson, L. P. Ferreira, and M. Maeno, “Distúrbio de voz relacionado ao trabalho: um olhar sobre o passado, o presente e o futuro,” *Revista Brasileira de Saúde Ocupacional*, vol. 49, p. edcin9, 2024.
- [3] S. L. do Nascimento Cunha Costa, *Análise Acústica, Baseada no Modelo Linear de Produção da Fala, para Discriminação de Vozes Patológicas*. PhD thesis, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Campina Grande, 2008.
- [4] V. J. Vieira, S. C. Costa, S. L. Correia, L. W. Lopes, W. C. d. A. Costa, and F. M. de Assis, “Exploiting nonlinearity of the speech production system for voice disorder assessment by recurrence quantification analysis,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 28, no. 8, p. 085709, 2018.
- [5] A. Akbari and M. K. Arjmandi, “An efficient voice pathology classification scheme based on applying multi-layer linear discriminant analysis to wavelet packet-based features,” *Biomedical Signal Processing and Control*, vol. 10, pp. 209–223, 2014.
- [6] W. C. de Almeida Costa, *Análise Dinâmica Não Linear de Sinais de Voz para Detecção de Patologias Laríngeas*. PhD thesis, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Campina Grande, 2012.
- [7] L. W. Lopes, L. B. Simões, J. D. da Silva, D. da Silva Evangelista, A. C. d. N. e Ugulino, P. O. C. Silva, and V. J. D. Vieira, “Accuracy of acoustic analysis measurements in the evaluation of patients with different laryngeal diagnoses,” *Journal of Voice*, vol. 31, no. 3, pp. 382–e15, 2017.
- [8] J. J. Jiang, Y. Zhang, and C. McGilligan, “Chaos in voice, from modeling to wavelet measurement,” *Journal of Voice*, vol. 20, no. 1, pp. 2–17, 2006.
- [9] V. J. D. Vieira, S. C. Costa, and S. E. N. Correia, “Índice de não estacionariedade aplicado à classificação de desordens vocais,” in *Anais do XL Simpósio Brasileiro de Telecomunicações e Processamento de Sinais*, SBtT 2022, 2022.
- [10] V. J. D. Vieira, S. C. Costa, and S. E. N. Correia, “Non-stationarity-based adaptive segmentation applied to voice disorder discrimination,” *IEEE Access*, vol. 11, pp. 54750–54759, 2023.
- [11] V. J. D. Vieira, *Análise de Variações Acústicas Não Estacionárias e seu Efeito na Detecção de Múltiplas Emoções e Condições de Estresse*. PhD thesis, Universidade Federal de Campina Grande, 2018.
- [12] V. Vieira, R. Coelho, and F. M. de Assis, “Hilbert–huang–hurst-based non-linear acoustic feature vector for emotion classification with stochastic models and learning systems,” *IET Signal Processing*, vol. 14, no. 8, pp. 522–532, 2020.
- [13] L. Zão, D. Cavalcante, and R. Coelho, “Time-frequency feature and AMS-GMM mask for acoustic emotion classification,” *IEEE Signal Processing Letters*, vol. 21, pp. 620–624, May 2014.
- [14] N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, N. Yen, C. Tung, and H. Liu, “The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis,” *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, pp. 903–995, March 1998.
- [15] P. Flandrin, G. Rilling, and P. Gonçalvès, “Empirical mode decomposition as a filter bank,” *IEEE Signal Processing Letters*, vol. 11, pp. 112–114, February 2004.
- [16] Z. Wu and N. E. Huang, “Ensemble Empirical Mode Decomposition: a Noise-Assisted Data Analysis Method,” *Advances in Adaptive Data Analysis*, vol. 1, no. 01, pp. 1–41, 2009.
- [17] Z. Wu and N. E. Huang, “A Study of the Characteristics of White Noise using the Empirical Mode Decomposition Method,” in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 460, pp. 1597–1611, The Royal Society, 2004.
- [18] H. E. Hurst, “Long-Term Storage Capacity of Reservoirs,” *Trans. Amer. Soc. Civil Eng.*, vol. 116, pp. 770–808, 1951.
- [19] K. Elemetrics, “Kay elemetrics corp. disordered voice database.” Model 4337, 03 Ed., 1994.
- [20] B. B. Mandelbrot and J. R. Wallis, “Robustness of the rescaled range r/s in the measurement of noncyclic long run statistical dependence,” *Water resources research*, vol. 5, no. 5, pp. 967–988, 1969.
- [21] V. J. D. Vieira, “Avaliação de distúrbios da voz por meio de análise de quantificação de recorrência,” Master’s thesis, Programa de Pós-Graduação em Engenharia Elétrica, Instituto Federal de Educação, Ciência e Tecnologia da Paraíba, 2014.
- [22] G. Schlotthauer, M. E. Torres, and H. L. Rufiner, “Pathological voice analysis and classification based on empirical mode decomposition,” *Development of multimodal interfaces: active listening and synchrony*, vol. 5967, pp. 364–381, 2010.
- [23] W. C. d. A. Costa, S. L. d. N. C. Costa, F. M. d. Assis, and B. G. Aguiar Neto, “Classificação de sinais de vozes saudáveis e patológicas por meio da combinação entre medidas da análise dinâmica não linear e codificação preditiva linear,” *Revista Brasileira de Engenharia Biomédica*, vol. 29, pp. 3–14, 2013.