

# Aplicação do espectro marginal de Hilbert na análise de desvios vocais

Thales Bezerra, João Souza, Maísa Santos, Silvana Costa e Suzete Correia

**Resumo**—Neste trabalho, os gráficos de amplitude-frequência fornecidos pelo espectro marginal de Hilbert são usados para caracterizar tanto sinais de vozes saudáveis, como os que apresentam os desvios vocais: rugosidade, sopro e tensão. Os resultados demonstram que os padrões obtidos parecem ser apropriados para diferenciar os sinais analisados.

**Palavras-Chave**— *Análise acústica, Decomposição de Modo Empírico, desvios vocais.*

**Abstract**— In this work, the amplitude-frequency graphs provided by Hilbert's marginal spectrum are used to characterize both healthy voice signals as those with vocal deviations: roughness, breathiness and strain. The results demonstrate that the obtained patterns seem to be appropriate to differentiate the analyzed signals.

**Keywords**— *Acoustic analysis, Empirical Mode Decomposition, vocal deviations.*

## I. INTRODUÇÃO

A voz disfônica provoca um vasto número de impressões na sua percepção, tais como rouquidão (rugosidade), sensação de escape de ar ao falar (sopro) e/ou esforço na emissão de som (tensão), o que pode gerar consequências desfavoráveis ao falante.

Desvios vocais podem ser caracterizados pela composição dos seus componentes harmônicos e ruídos, na distribuição do espectro acústico. Na prática clínica, fonoaudiólogos, empregam o espectrograma, gráfico bidimensional, obtido a partir da *Short Time Fourier Transform* (STFT), como ferramenta para a avaliação da qualidade vocal. A rugosidade está correlacionada à presença de ruído e perdas de componentes harmônicos em alta frequência no sinal de voz, devido às irregularidades das vibrações das pregas vocais. Por outro lado, as vozes soprosas são caracterizadas pela presença de ruído nas altas frequências [1].

A Decomposição de Modo Empírico é um método empregado para a decomposição de sinais não estacionários, como sinais de voz, que se baseia na interpolação para traçar sucessivamente Funções de Modo Intrínseco (*Intrinsic Mode Functions* - IMFs) [2]. Quando a transformação de Hilbert é aplicada às IMFs, a frequência instantânea e a amplitude do sinal podem ser determinadas. Este processo é chamado de transformada de Hilbert-Huang (HHT). A maior vantagem, é que esse método é totalmente adaptável, sem a necessidade de selecionar uma função base a priori, para a decomposição do sinal.

A Transformada de Hilbert-Huang (HHT) proporciona a

---

Thales Bezerra, João Souza, Maísa Santos, Silvana Costa e Suzete Correia, Instituto Federal de Educação, Ciência e Tecnologia da Paraíba, João Pessoa – PB, Brasil, E-mails thalesmartinsbezerra9@gmail.com, jvfelixxy@gmail.com, maisalfsantos@gmail.com, silvana@ifpb.edu.br, suzete@ifpb.edu.br. Este trabalho foi parcialmente financiado pelo IFPB e CAPES.

representação, em duas dimensões, das amplitudes e frequências de modo instantâneo, para sinais com múltiplas componentes de frequência. O espectro marginal de Hilbert é obtido a partir da HHT e reflete com maior precisão a variação da amplitude do sinal em cada ponto de frequência, quando comparado com a representação clássica obtida através da STFT, permitindo uma melhor caracterização de sinais que têm uma natureza não estacionária [3].

Abreu *et al.* [4] empregaram medidas temporais e espectrais, extraídas das IMFs, para detectar os desvios de rugosidade e sopro de sinais de voz. Uma acurácia de 95,00% foi obtida na detecção entre sinais de vozes saudáveis e soprosos. Kaleem *et al.* [5] compararam o espectro marginal de Hilbert de um sinal de voz saudável com o de um sinal de voz patológico e concluíram que a diferença entre os espectros pode distinguir os sinais. Características temporais e espectrais foram consideradas para as IMFs de 1-10, obtendo uma classificação de 95,7% de acurácia.

Neste trabalho, pretende-se empregar o espectro marginal de Hilbert na avaliação da qualidade vocal, através da caracterização de sinais de voz saudáveis e afetados pelos desvios de rugosidade, sopro e tensão.

## II. ESPECTRO MARGINAL DE HILBERT

Na Decomposição de Modo Empírico, o sinal original  $x[n]$  pode ser representado como a soma de  $K$  IMFs  $c_i$  e o resíduo  $r_K$ , [2],

$$x[n] = \sum_{i=1}^K c_i[n] + r_K. \quad (1)$$

O espectro tempo-frequência de Hilbert  $H[w,n]$  de um sinal é obtido omitindo-se o componente residual e considerando a frequência instantânea  $w_i$  e amplitude  $a_i$ ,

$$H[w,n] = \text{Re} \sum_{i=1}^K a_i[n] \exp(j \sum w_i[n] dn), \quad (2)$$

em que  $\text{Re}$  representa a parte real de  $H[w,n]$ .

O espectro marginal de Hilbert  $h(w)$  representa a contribuição de energia, de cada valor de frequência, em todo o comprimento do sinal, sendo definido como [5]:

$$h(w) = \sum_{n=1}^N H[w,n], \quad (3)$$

em que  $N$  é o número total de amostras.

## III. MATERIAIS E MÉTODOS

### A. Base de dados

Foram usadas emissões da vogal sustentada /E/, amostradas a uma taxa de 44.100 amostras/s, com e quantizadas com 16 bits/amostra. Os sinais fazem parte de um banco de vozes

sintetizadas, produzido no Departamento de Ciência da Computação, da Universidade de Brasília, em colaboração com os Laboratórios de Imagem, Processamento de Sinal e Acústica da Universidade Livre de Bruxelas [6].

### B. Metodologia

Para cada sinal de voz, foi realizada a Decomposição de Modo Empírico para seis IMFs. O espectro de Hilbert foi obtido e convertido para escala de cinza. Em seguida, as ocorrências em cada frequência do espectro de Hilbert foram somadas, gerando o espectro marginal. Foram selecionados da base, 05 sinais de vozes saudáveis e 15 sinais de vozes com desvios vocais, sendo 05 para cada tipo de desvio considerado, rugosidade, sopro e tensão.

## IV. RESULTADOS

As Figuras 1-4 representam, respectivamente, os espectros marginais de Hilbert para os sinais de vozes saudáveis, soprosas, rugosas e tensas. Gráficos com características distintas, para os diversos tipos de sinais analisados podem ser observados. Diferentemente do que ocorre nos sinais de vozes saudáveis (Figura 1), pode ser percebida a presença de ruídos em frequências acima de 3500Hz, nos sinais de vozes desviadas. A presença de sub-harmônicos no espectro, indica irregularidade à fonação e presença de outras fontes sonoras à emissão vocal [7].

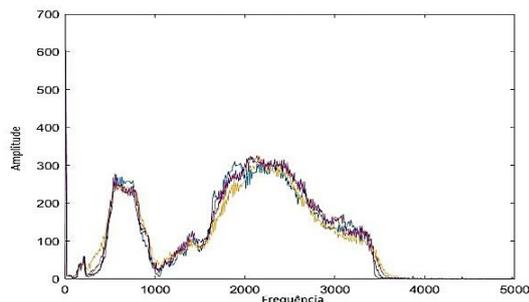


Fig. 1. Espectro marginal de Hilbert de vozes saudáveis.

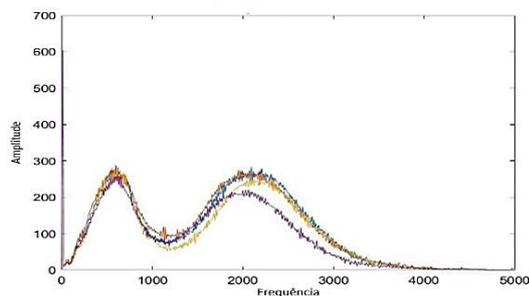


Fig. 2. Espectro marginal de Hilbert de vozes soprosas.

A sopro pode ser evidenciada, na Figura 2, pelo aumento da amplitude na faixa de frequências em torno de 1000 Hz. A rugosidade (Figura 3) é evidenciada pela presença de ruído em todo o espectro e pela diminuição da amplitude nas frequências acima de 2000 Hz. O espectro marginal de Hilbert para as vozes tensas (Figura 4), é o que mais se distingue dos demais, possuindo uma menor amplitude nas baixas frequências e uma amplitude maior nas frequências próximas de 4000 Hz.

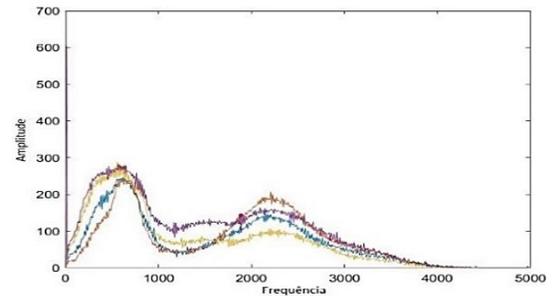


Fig. 3. Espectro marginal de Hilbert de vozes rugosas.

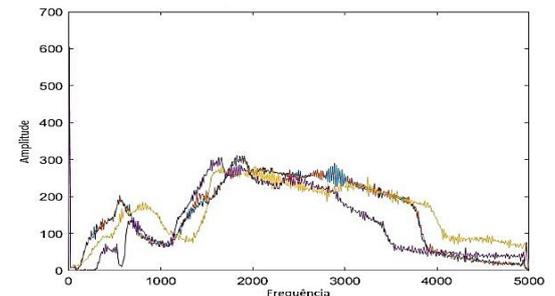


Fig. 4. Espectro marginal de Hilbert de vozes tensas.

## V. CONSIDERAÇÕES FINAIS

O estudo a partir de vozes sintetizadas pôde proporcionar maior controle sobre os desvios vocais presentes no sinal de voz, de modo a se obter com os espectros marginais, padrões gráficos representativos para as classes analisadas. Observou-se, nos resultados obtidos, que o espectro marginal de Hilbert pode ser empregado para detectar a presença de desvios vocais.

## AGRADECIMENTOS

Ao IFPB e a CAPES pelo apoio financeiro. Ao professor Leonardo Lopes pela base de dados.

## REFERÊNCIAS

- [1] M. Behlau, *A Voz do Especialista - Volume 1*. Editora Revinter, Rio de Janeiro, 2001.
- [2] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung e H. 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 A: mathematical, physical and engineering sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
- [3] C. Li, G. Yu, B. Fu, H. Hu, X. Zhu e Q. Zhu. "Fault separation and detection for compound bearing-gear fault condition based on decomposition of marginal Hilbert spectrum", *IEEE Access*, vol. 7, pp. 110518–110530, 2019.
- [4] S. R. de Abreu, S. E. N. Correia e S.L.do N. C. Costa. "Classificação de desvios vocais utilizando o método de decomposição em modos empíricos", *XXXVI Simpósio Brasileiro de Telecomunicações e Processamento de Sinais - SBrT2018*, 2018.
- [5] M. Kaleem, B. Ghoraani, A. Guergachi e S. Krishnan. "Pathological speech signal analysis and classification using empirical mode decomposition", *Med Biol Eng Comput*, vol 51, pp.811–821, 2013.
- [6] J. C. Lucero, J. Schoentgen, M. Behlau. "Physics-based synthesis of disordered voices". In: *Proceedings of Interspeech 14th Annual Conference of the International Speech Communication Association*, p. 25-9, 2013.
- [7] I. Titze. "Workshop on acoustic voice analysis: Summary statement," National Center for Voice and Speech, Denver, CO, 1995.