

Avaliação de desvios vocais por meio de escalogramas Wavelet

A. V. N. de França*, L. O. Pinheiro Filho*, S. E. N. Correia*, S. L. N. C. Costa*

*Instituto Federal da Paraíba, João Pessoa, Brasil

e-mail: anderson.franca@academico.ifpb.edu.br

Abstract - *Voice is the most common communication tool, result of the interaction of several physiological systems. The vocal disorders refer to processes that affect vocal production, usually involving vocal folds and laryngeal mucosa, and may be caused by organic or functional factors, causing changes in voice quality, height and intensity. This work aims to determine graphic patterns to identify vocal deviations in a non-invasive way, using the scalogram obtained from the continuous wavelet transform. Morlet wavelet function is employed in two hundred and fifty-six levels of resolution. Through the obtained scalograms, for the healthy voice signals and those containing the deviations roughness, breathiness and strain, it was possible to perceive variations in the regularity of the patterns, as well as in the intensity and frequency bands between the signals that have different deviations. A standardization of these aspects between the signals of similar vocal classification is shown by scalograms. These graphic patterns can be used in the evaluation of the vocal quality for screening and phonotherapeutic treatment.*

Palavras-chave: *wavelet, escalograma, desvios vocais.*

Introdução

Em nossa sociedade tem-se uma constante evolução do pensamento em prol do desenvolvimento de serviços para benefícios da própria comunidade [1]. Com isso, é possível observar o crescimento de várias pesquisas acerca da qualidade da voz, bem como da criação de ferramentas de auxílio para diagnósticos ou tratamentos de patologias que afetam a voz, dentre outras.

Não há uma definição plenamente aceita como cânone para classificação da qualidade de uma voz, sendo obtida a partir da opinião dos ouvintes para cada falante, podendo variar em função dos parâmetros utilizados para a definição [2].

Neste sentido, para a avaliação médica da qualidade vocal são utilizados dois importantes recursos [3]: a análise perceptivo-auditiva, que consiste num método totalmente subjetivo, onde um especialista treinado para tal função ouve o paciente e define, a partir de sua experiência, numa escala

analógico-visual o grau de intensidade com que o desvio vocal se apresenta na voz sob análise; e a análise acústica, que tem sido utilizado mais recentemente e refere-se à uma observação e extração de parâmetros do sinal de áudio gerado pelo paciente utilizando-se *softwares* específicos, proporcionando ao profissional de saúde um método objetivo e complementar para avaliação das vozes.

Geralmente, a análise acústica é realizada com base nas características do sinal no domínio do tempo e da frequência, pelo fato de ser onde se tem as alterações mais visíveis dos sinais de vozes desviadas quando comparados com os sinais de vozes consideradas saudáveis, pelos métodos subjetivos. Um exemplo dessas alterações encontra-se na adição de ruído e perturbações em baixas frequências quando a voz é rugosa, e em altas frequências, quando soprosa [4]. Vale salientar que uma voz pode ter características de vários desvios, ainda que em graus de intensidade diferentes.

Vários métodos têm sido empregados na análise tempo-frequência de sinais. Entre eles se destaca a Transformada Wavelet por fornecer a análise dos sinais em distintas faixas de frequência. Por meio das Wavelets é possível observar tanto as características globais, componentes de baixas frequências, quanto locais, componentes de alta frequência, dos sinais analisados.

O escalograma Wavelet é uma representação gráfica em duas dimensões que fornece a densidade de energia do sinal ao longo das escalas de frequência, permitindo a detecção das frequências mais representativas de um sinal, o que pode ser útil, na análise acústica, para apontar uma desordem vocal.

Nayak et al. [5] propuseram o uso de escalogramas Wavelet para análise de sinais de voz saudáveis e afetados por paralisia e hiperfunção. Foi empregada a wavelet Chapéu Mexicano e os padrões visuais obtidos foram capazes de fazer a diferenciação entre as classes de sinais analisadas.

Em Correia et al [6] os escalogramas obtidos através da família Wavelet Chapéu Mexicano, considerando 128 escalas, foi usado para análise de sinais de voz saudáveis e apresentando nódulos, paralisia e edema de Reinke nas pregas vocais. Os padrões obtidos foram atrativos para identificação

das patologias laringeas analisadas.

Neste trabalho, o uso de escalogramas Wavelet como ferramenta na avaliação da qualidade vocal é proposto para a identificação da presença de desvios vocais, assim como o grau de acentuação em que se apresentam. Foi escolhida a Wavelet de Morlet por possuir maior semelhança com os sinais de voz analisados, gerando padrões gráficos mais representativos. Três tipos de desvios vocais são considerados: rugosidade, soproidade e tensão. A ideia central é determinar um padrão visual para cada tipo de desvio analisado, que possa servir de auxílio em triagem e tratamentos fonoterápicos.

Na seção seguinte é apresentada a metodologia do trabalho, onde será exposto todo o procedimento realizado, seguida dos resultados obtidos e a respectiva análise. Por fim, a seção Conclusões apresenta os comentários finais e a perspectiva de trabalhos futuros.

Materiais e métodos

A Figura 1 ilustra as etapas componentes do processo de avaliação dos desvios vocais, onde após escolhida a base de dados, calculou-se os coeficientes da Transformada Wavelet para cada sinal sob análise, com os quais foram obtidos os escalogramas para a identificação visual das características próprias a cada tipo de sinal de voz.

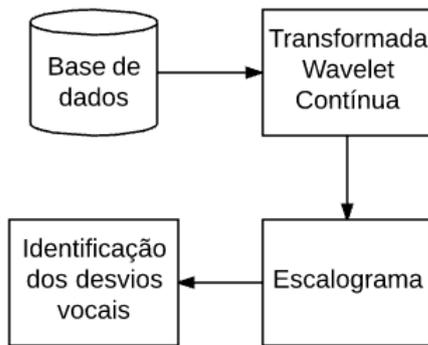


Figura 1 – Fluxograma da avaliação de desvios vocais utilizando escalogramas

Base de Dados

Os sinais de vozes empregados correspondem a emissões da vogal do português brasileiro /ε/ ("é"), sustentada por 1 segundo, obtidos por um sintetizador (VoiceSim), 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. Esses sinais foram obtidos e armazenados a uma taxa de amostragem de 44100 Hz, 16 bits por amostra. Foram selecionados inicialmente 90 sinais de voz, divididos em quatro conjuntos: 10 saudáveis, 29 soprosas, 30 rugosas e 21 tensas. Os sinais de vozes desviadas foram escolhidos de modo a conter intensidades de leve a intenso.

A escolha dessa base de dados se deve ao fato de que os sinais de voz sintetizados possibilitam um controle de cada desvio existente e do seu grau de intensidade, auxiliando na identificação de características comuns para cada tipo de desvio, o que não ocorre nos sinais de vozes reais.

Transformada Wavelet Contínua

A Transformada Wavelet Contínua é uma ferramenta utilizada na exploração de características de sinais não estacionários para extrair informações de variações em certas bandas de frequências ao longo do tempo. Trata-se de uma integral linear cujos coeficientes, para uma determinada função $f(t)$, são obtidos por meio da Equação 1 [7].

$$W_x^\psi(a, b; \psi) = a^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(t) \psi_{a,b}^*(t) dt \quad (1)$$

As variáveis a e b representam os parâmetros de escala e de translação, respectivamente, e variam continuamente [7]. O termo $\psi_{a,b}^*(t)$ consiste nas wavelets filhas geradas pela dilatação e translação da wavelet mãe $\psi(t)$ definida na Equação 2 [7]:

$$\psi_{a,b}(t) = a^{-\frac{1}{2}} \psi\left(\frac{t-b}{a}\right) \quad (2)$$

A wavelet de Morlet é definida na Equação 3 [8]:

$$\psi = C \cdot e^{-\frac{t^2}{2}} \cdot \cos(5t) \quad (3)$$

em que C é uma constante usada para normalização. É possível visualizar o gráfico da função de Morlet na Figura 2.

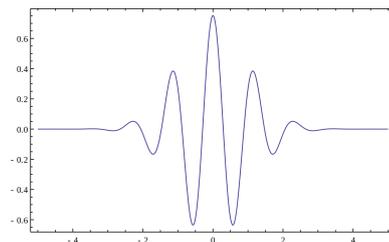


Figura 2 – Forma de onda da Função de Morlet

Escalogramas

Para cada sinal de voz é realizada a sua decomposição por meio da Wavelet de Morlet em 256 níveis de resolução e, em seguida, obtido o seu escalograma. Vários testes foram realizados, considerando diferentes durações do sinal no tempo, com o intuito de determinar a quantidade de amostras do sinal de voz que fornece uma melhor representação visual no escalograma, sendo adotado um valor padrão de aproximadamente 45,35 ms, com um total de 2000 amostras.

O escalograma Wavelet é o resultado do quadrado do módulo dos coeficientes gerados pela Transformada Wavelet Contínua, de acordo com a Equação 4 [9].

$$SG_x(a, b; \psi) = |W_x(a, b; \psi)|^2 \quad (4)$$

Cada pixel que compõe o escalograma é calculado isoladamente baseando-se nas informações de escala, tempo e da wavelet filha, logo possui variação de intensidade para cada mínima variação de tempo ou escala. O padrão de cores resultante compreende informações sobre o tamanho e a localização de determinados eventos no domínio do tempo.

Resultados

Resultados preliminares apontam diferenças nos escalogramas dos sinais analisados, sendo maior a evidência para os sinais de vozes tensos. Na Figura 3 pode-se observar os escalogramas Wavelet de dois sinais de vozes distintos da base de dados, ambos classificados como saudáveis.

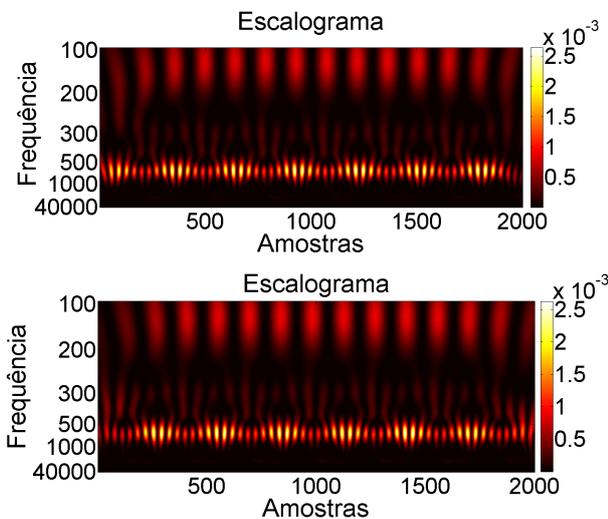


Figura 3 – Escalogramas de dois sinais de vozes saudáveis

As Figuras de 4 - 6 ilustram os escalogramas das vozes desviadas. Dois exemplos de cada tipo de desvio são considerados. Ao se comparar com os escalogramas Wavelets dos sinais vocais disfônicos da base de dados, é possível visualizar a diferença entre os escalogramas de sinais desviados e os saudáveis, além de um padrão, ainda que com algumas pequenas variações, para a distribuição de energia no sinal nas diferentes faixas de frequência de cada tipo de desvio vocal utilizado para os testes.

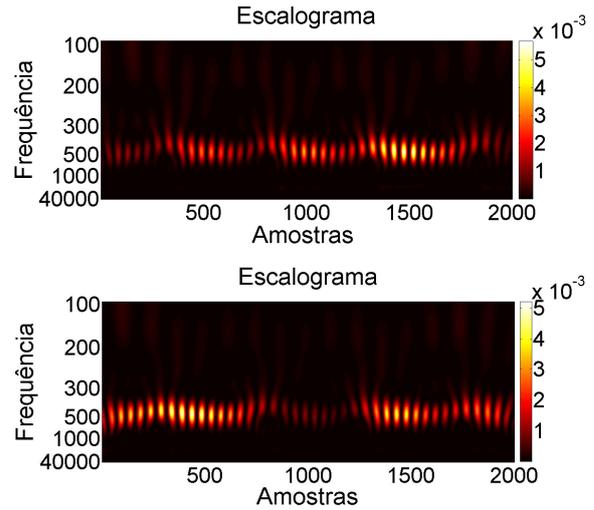


Figura 4 – Escalogramas de dois sinais de vozes rugosas

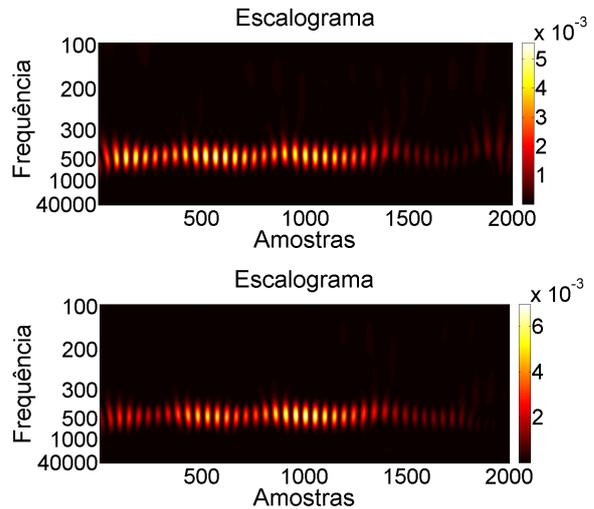


Figura 5 – Escalogramas de dois sinais de vozes soprosas

Discussão

Ao observar os escalogramas obtidos, pode ser visualizada a periodicidade do sinal através da regu-

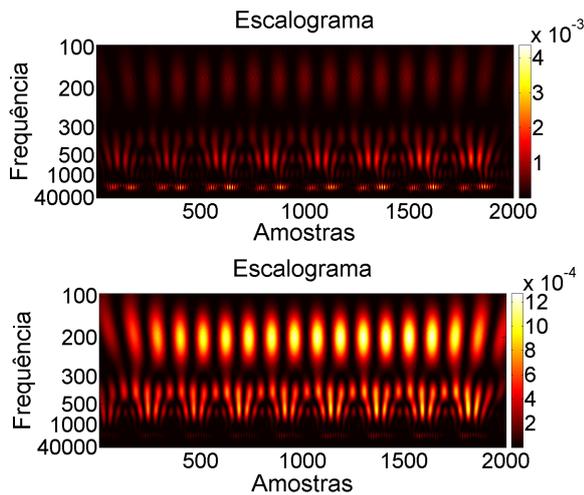


Figura 6 – Escalogramas de dois sinais de vozes tensas

laridade dos padrões exibidos nas Figuras 3 - 6. A emissão da vogal sustentada em vozes saudáveis é periódica, enquanto para sinais disfônicos a periodicidade é alterada. Nos sinais de vozes com tensão nota-se a periodicidade do sinal, porém com a energia distribuída em várias faixas de frequência, além de interrupções mais bruscas dada a forma rápida como a densidade de energia aumenta e diminui ao longo do tempo. Observa-se também que entre os sinais rugosos e soprosos há uma diminuição da energia nas baixas frequências, quando comparados aos sinais de vozes saudáveis, sendo maior nos sinais soprosos. Para os sinais tensos, há um aumento de energia em frequências mais altas, em torno de 1000Hz.

Além disso é possível identificar a similaridade nos padrões dos escalogramas com sinais de mesma classe.

Conclusões

Neste trabalho foi proposto o uso de escalogramas gerados com os coeficientes da Transformada Wavelet Contínua como parâmetro para discriminação visual da qualidade vocal e constatou-se que o escalograma foi capaz de fornecer os padrões que viabilizam a identificação de desvios vocais. Este método pode ser empregado para auxiliar os profissionais de saúde na triagem ou no tratamento fonoterápico de forma complementar às avaliações subjetivas realizadas pelos mesmos.

Em trabalhos futuros pretende-se utilizar a Transformada Wavelet Discreta para a extração de medidas, a exemplo da energia e entropia, a partir dos coeficientes Wavelet para utilizá-las na

detecção destes desvios vocais.

Agradecimentos

Os autores do trabalho agradecem ao IFPB pelo financiamento da pesquisa.

Referências

- [1] F. Wuyts, M. De Bodt, G. MOLENBERGHS, M. Remacle, L. Heylen, B. Millet, C. Van Lierde, J. Raes, and P. Van de Heyning, "The dysphonia severity index: An objective measure of vocal quality based on a multiparameter approach," 2000.
- [2] D. Childers and C. Lee, "Vocal quality factors: analysis, synthesis, and perception.," *The Journal of the Acoustical Society of America*, vol. 90, no. 5, pp. 2394–2410, 1991.
- [3] P. A. Pontes, V. P. Vieira, M. I. R. Gonçalves, and A. A. Pontes, "Características das vozes roucas, ásperas e normais: análise acústica espectral comparativa," *Revista Brasileira de Otorrinolaringologia*, 2002.
- [4] S. N. Awan and N. Roy, "Acoustic prediction of voice type in women with functional dysphonia," *Journal of Voice*, vol. 19, no. 2, pp. 268–282, 2005.
- [5] B. P. S. A. R. . A. U. V. Nayak, J., "Classification and analysis of speech abnormalities.," 2005.
- [6] S. E. N. Correia, S. L. N. C. Costa, and C. B. Aguiar, "Pré-diagnósticos de patologias laríngeas baseado em escalograma wavelet," *IX Congresso de Pesquisa e Inovação da Rede Norte e Nordeste de Educação Tecnológica*, 2014.
- [7] Z. Peng, F. Chu, and Y. He, "Vibration signal analysis and feature extraction based on reasigned wavelet scalogram," *Journal of Sound and Vibration*, vol. 253, no. 5, pp. 1087–1100, 2002.
- [8] I. Daubechies, *Ten lectures on wavelets*. SIAM, 1992.
- [9] Z. Peng and F. Chu, "Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography," *Mechanical systems and signal processing*, vol. 18, no. 2, pp. 199–221, 2004.