

# Avaliação de Diversas Bases Wavelets na Detecção de Distúrbios Vocais Infantis

Mikaelle O. Santos, Washington C. de A. Costa,  
Silvana C. Costa e Suzete E. N. Correia  
Programa de Pós-Graduação em Engenharia Elétrica  
Instituto Federal de Educação, Ciência e Tecnologia da  
Paraíba  
João Pessoa, Brasil

Leonardo W. Lopes  
Departamento de Fonoaudiologia  
Universidade Federal da Paraíba  
João Pessoa, Brasil

**Resumo**—A alta prevalência de desvios vocais na infância exige uma atenção especial na avaliação e diagnóstico das vozes infantis. Técnicas de processamento digital de sinais podem ser empregadas para auxiliar outros métodos de avaliação de distúrbios da voz, tais como análise otorrinolaringológica e análise perceptivo-auditiva. A transformada wavelet é uma ferramenta interessante para análise acústica por permitir a representação dos sinais em faixas de frequências diferentes. Neste trabalho, as medidas de energia normalizada e entropia dos coeficientes de detalhes da transformada wavelet discreta são empregadas na avaliação do grau de intensidade do desvio vocal em crianças. São avaliadas 45 wavelets da família de Daubechies em nove níveis de resolução. A classificação dos sinais é realizada empregando a Análise Discriminante Quadrática. A wavelet de Daubechies de ordem 12 apresentou melhor resultado em termos de acurácia e custo computacional.

**Palavras-Chave**—Análise Acústica; Reconhecimento de padrões; Transformada Wavelet; Análise Discriminante Quadrática.

## I. INTRODUÇÃO

Crianças com distúrbios de voz podem apresentar efeitos negativos no seu desenvolvimento social, educacional e físico [1]. A investigação e o diagnóstico precoce do desvio vocal infantil permite maior eficácia no tratamento. Entretanto, a avaliação de desordens vocais em crianças apresenta alguns desafios relacionados às dificuldades de cooperação das mesmas durante os exames tradicionais. Tais exames consistem em anamnese, seguida de exames físicos e visuais da laringe, realizados por otorrinolaringologistas, além daqueles, realizados por fonoaudiólogos que incluem a audição e avaliação da qualidade vocal do paciente.

Pesquisadores têm buscado, através do processamento digital de sinais, medidas discriminativas, que sejam capazes de imprimir uma boa avaliação da qualidade vocal, bem como o seu diagnóstico e monitoramento do tratamento. Essas técnicas são automáticas, de caráter não invasivo e de baixo custo, comparadas às técnicas comumente empregadas. A análise acústica permite ao avaliador determinar e quantificar a qualidade vocal do indivíduo através dos diferentes parâmetros acústicos que representam o sinal de voz, sendo um ótimo recurso para capturar desordens vocais.

A extração de características do sinal de voz, que representem bem o desvio vocal que se pretende investigar, é de fundamental importância para uma classificação mais acurada do tipo e do grau do desvio, para acompanhamento do processo de terapia fonoaudiológica. Uma classificação eficiente pode auxiliar o terapeuta a avaliar o quanto a terapia está sendo efetiva, de forma objetiva. Para tanto, é necessário que a técnica proposta tenha confiabilidade e apresente as informações das mudanças ocorridas no sinal antes e após a terapia vocal, necessárias para um diagnóstico mais preciso.

Distúrbios vocais infantis causam alterações na qualidade da voz, sendo a rouquidão e a soproalidade as mais frequentes [2], representadas como ruídos em faixas de frequências diferentes no espectro do sinal de voz. A rouquidão é proveniente da irregularidade de vibração das pregas vocais que geram ruídos nas baixas frequências. A soproalidade indica a presença de ruído de fundo, audível, que corresponde fisiologicamente à fenda glótica (abertura das pregas vocais) [3] e relaciona-se à ruídos presentes nas altas frequências [4].

As principais medidas acústicas utilizadas na detecção de desvios vocais são a frequência fundamental, o *Jitter* e o *Shimmer*. Existem outras características do sinal sonoro capazes de fornecer informações importantes, tais como os formantes, as medidas de ruído, a intensidade, e o tempo máximo de fonação [5]. A transformada wavelet é uma ferramenta matemática que permite a análise multirresolucional de sinais, ou seja, a análise de sinais em frequências distintas, de modo que, em cada frequência, aspectos diferentes sejam observados. Tal característica torna-a uma ferramenta útil e interessante para o processamento de sinais, aplicada na análise acústica.

Na literatura, encontram-se trabalhos que empregam as wavelets na detecção automática de desordens vocais, entretanto, a maioria é voltada para a discriminação de patologias laríngeas, empregando bases de dados de vozes adultas [6-8]. Como há diferenças no padrão vocal de adultos e crianças, é necessário o desenvolvimento de sistemas automáticos para detecção de desordens vocais voltados ao público infantil.

Outra característica da transformada wavelet é o fato de que existem diversas funções que podem ser usadas para a

representação dos sinais, as quais podem ser selecionadas de acordo com a aplicação. Em processamento digital de sinais de voz, as wavelets da família de Daubechies têm sido bastante empregadas [6]. Tais funções possuem filtros de comprimentos diferentes, são ortogonais e possuem suporte compacto [9].

Neste trabalho é apresentado um sistema que emprega a energia normalizada e entropia dos coeficientes de detalhes da transformada wavelet discreta para discriminar o grau de severidade do desvio vocal em crianças. O artigo está organizado da seguinte forma: na Seção II são apresentados os Materiais e Métodos, na Seção III os Resultados Obtidos, os quais são discutidos na Seção IV e, na Seção V, as Considerações Finais.

## II. MATERIAIS E MÉTODOS

### A. Base de Dados

Os sinais de voz empregados são provenientes da base de dados coletada pelo Laboratório Integrado de Estudos da Voz (LIEV), do Departamento de Fonoaudiologia da Universidade Federal da Paraíba [10], tendo sido aprovada a sua aquisição pelo comitê de ética em pesquisa da instituição de origem sob o protocolo número 775/10. A mesma contém 93 sinais de vozes infantis da vogal sustentada /ε/ de ambos os sexos, com idade variando entre 3 e 10 anos, sendo 48 meninas e 45 meninos, todos integrantes de uma escola vinculada a uma instituição de ensino pública federal. A coleta foi realizada em ambiente silencioso, com menos de 50 dB NPS, utilizando um computador pessoal e microfone *headset*, utilizando o *software* PRAAT, versão 5.1.44, com taxa de amostragem de 44.100 Hz.

A escolha da vogal /ε/ se dá devido às poucas variações que ocorrem na configuração do trato vocal durante a sua emissão, comparada às outras vogais orais do português brasileiro, apresentando harmônicos de maior intensidade e mais bem definidos [11]. Inicialmente, as vozes foram editadas, sendo eliminados os segundos iniciais e finais da emissão da vogal, devido a maior irregularidade nesses trechos, preservando-se o tempo mínimo de dois segundos para cada emissão. Foi realizada a normalização dos sinais, a fim de obter uma padronização na saída de áudio entre - 6 e 6 dB.

Para a análise perceptivo-auditiva da voz, foi utilizada a escala analógica visual (EAV), com uma métrica de 0 a 100 mm. A marcação mais próxima do zero representa uma menor alteração, e a mais próxima do cem, uma alteração maior, com uma possibilidade de gradação de 0 a 100 mm da alteração. Posteriormente foi realizada uma correspondência da escala numérica (EN) para a EAV, sendo o grau 1 (0-35,5 mm) relacionado às vozes saudáveis, grau 2 (35,6-50,5 mm) desvio leve, grau 3 (50,6-90,5 mm) desvio moderado e grau 4 (90,6-100 mm) desvio intenso [12].

Essa avaliação foi realizada por meio de consenso por três fonoaudiólogos especialistas em voz, com experiência em avaliação vocal perceptivo-auditiva. A sessão de avaliação perceptiva ocorreu em ambiente silencioso. Inicialmente, os juízes foram orientados de que as vozes deveriam ser consideradas saudáveis quando fossem socialmente aceitáveis para uma criança. No final da sessão de avaliação perceptiva, 10% das amostras foram repetidas aleatoriamente, para a

análise da confiabilidade através do Coeficiente Kappa de Cohen. O valor de kappa foi de 0,80, indicando uma boa concordância entre os avaliadores.

### B. Extração de Características

A transformada wavelet consiste na decomposição de um sinal  $x(t)$  em diferentes faixas de frequência através de translações e escalonamentos de uma função oscilatória, de média zero e comprimento finito, denominada wavelet mãe  $\psi(t)$ . Matematicamente, a transformada wavelet contínua é definida como [13]:

$$T(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t-b}{a}\right) dt, \quad (1)$$

em que  $a > 0$  é o fator de escala,  $b$  o fator de translação e  $a^{-1/2}$  o fator de normalização que mantém a mesma energia para todas as wavelets independente da escala utilizada. Os coeficientes wavelets fornecem informações da energia local do sinal em várias escalas. À medida que o fator de escala aumenta, há uma expansão da wavelet no tempo, as oscilações tornam-se lentas e é feita uma análise das baixas frequências do sinal. Por outro lado, se o fator de escala diminui, as oscilações aumentam e é feita uma análise das altas frequências do sinal.

Uma vez que a wavelet mãe é dilatada e transladada continuamente, uma quantidade redundante de coeficientes é obtida. Na versão discreta, as wavelets são geradas através de fatores  $a$  e  $b$  diádicos representados por  $a = 2^j$  e  $b = k \cdot 2^j$ , em que  $j$  e  $k$  são inteiros. Uma maneira de se implementar a transformada wavelet discreta é através do algoritmo multiresolução proposto por Mallat [14]. Para tanto, além da wavelet é empregada outra função denominada escalonamento  $\phi(t)$ . As funções  $\phi(t)$  e  $\psi(t)$  são representadas por :

$$\phi(t) = \sum_n h(n) \phi(2t - n), \quad (2)$$

$$\psi(t) = \sum_n g(n) \phi(2t - n), \quad (3)$$

em que  $h(n)$  e  $g(n)$  são as respostas ao impulso dos filtros passa-baixas e passa-altas, respectivamente. Da decomposição do sinal  $x(t)$ , dois conjuntos de coeficientes podem ser obtidos para cada nível de resolução  $j$ : os coeficientes de aproximação  $c_j(k)$  e os coeficientes de detalhe  $d_j(k)$ , os quais são definidos por:

$$c_j(k) = \sum_m h(m - 2k) c_{j-1}(m), \quad (4)$$

$$d_j(k) = \sum_m g(m - 2k) c_{j-1}(m). \quad (5)$$

Os coeficientes wavelets  $c_j(k)$  e  $d_j(k)$  são obtidos através da convolução do sinal com as respostas ao impulso dos filtros passa-baixas e passa-altas, respectivamente, seguido pela decimação por um fator 2. Da filtragem por  $h(n)$  obtém-se a aproximação do sinal, que é uma representação de baixa frequência do sinal original, enquanto da filtragem por  $g(n)$  é fornecida a diferença entre duas aproximações sucessivas do sinal original. Uma aproximação fornece a tendência geral do

sinal enquanto os detalhes os componentes de alta frequência do mesmo. Se o sinal original  $c_{j-1}(k)$  possui  $N$  amostras, então os sinais  $c_j(k)$  e  $d_j(k)$  possuem  $2^j N$  amostras. Dessa forma, a decomposição wavelet é formada por sequências que possuem o mesmo tamanho do sinal de entrada.

A Fig. 1 ilustra o algoritmo de decomposição da transformada wavelet discreta para os sinais de voz analisados, considerando nove níveis de resolução. Para cada nível é associada a faixa de frequências correspondente aos coeficientes de detalhes. Como os sinais de voz foram amostrados a uma taxa de 44.100 amostras/s, considera-se para análise frequências até 22.050 Hz. Pode ser observado que o sinal é decomposto em vários coeficientes de resolução mais baixa.

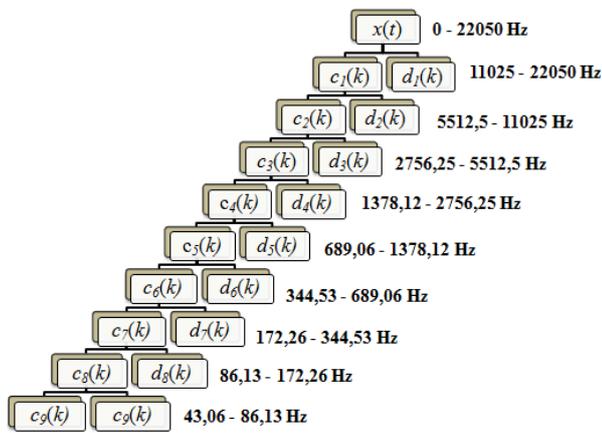


Fig. 1. Algoritmo de decomposição da transformada wavelet discreta para os sinais de voz analisados em nove níveis de resolução.

Algumas características podem ser extraídas a partir dos coeficientes obtidos pela decomposição wavelet de um determinado sinal. Nesta pesquisa, são utilizadas a energia normalizada  $E_j$  e a entropia de Shannon  $H_j$  dos coeficientes de detalhe, para cada um dos nove níveis de resolução considerados, de acordo com as Equações (6) e (7), respectivamente.

$$E_j = \frac{\sum_k |d_j(k)|^2}{\sum_k (|c_j(k)|^2 + |d_j(k)|^2)} \quad (6)$$

$$H_j = -\sum_k p_j(k) \log p_j(k), \quad (7)$$

em que as probabilidades de ocorrência dos coeficientes podem ser calculadas por:

$$p_j(k) = \frac{|d_j(k)|^2}{\sum_k |d_j(k)|^2}. \quad (8)$$

Existem diferenças espectrográficas entre sinais considerados saudáveis dos sinais que apresentam distúrbios vocais, de modo que a distribuição de energia ao longo da frequência deve ocorrer em faixas de frequência distintas entre eles, proporcionando uma separação entre as classes. De modo análogo, a entropia é usada por medir o grau de desordem de um sinal. Sendo assim, a mesma é influenciada pelas

irregularidades e aleatoriedade dos sistemas fisiológicos, a exemplo do sistema de produção vocal [3], podendo ser usada como medida na avaliação de distúrbios vocais.

São consideradas para avaliação 45 wavelets da família de Daubechies com ordens de 1 a 45. As wavelets de Daubechies são uma família, formada por várias funções que possuem filtros de comprimento diferentes [7]. Os nomes das wavelets são descritos por dbN, em que  $N$  é o comprimento do filtro ou ordem da wavelet.

### C. Classificação

As medidas de energia e entropia extraídas, em cada nível de resolução, são utilizadas individualmente e combinadas entre si para a classificação do grau da intensidade do desvio vocal ou da qualidade vocal predominante. Para a classificação, é utilizada a análise discriminante quadrática com validação cruzada.

O processo de validação cruzada é realizado dividindo os sinais analisados em 10 subconjuntos mutuamente exclusivos e de mesmo tamanho. Um subconjunto é utilizado para teste e os demais para treino. O procedimento é realizado 10 vezes alternando o subconjunto de teste. Ao final calcula-se a acurácia média. A escolha dos sinais deu-se de forma aleatória e sem repetição.

## III. RESULTADOS

Para mensurar a precisão dos classificadores empregados, três medidas foram consideradas: acurácia, sensibilidade (Sens) e especificidade (Esp). A acurácia mede a taxa de classificação correta global, refletindo a capacidade do classificador de identificar corretamente quando há e quando não há a presença do distúrbio. A sensibilidade representa a capacidade do classificador em identificar a presença do distúrbio quando ele de fato existe, enquanto a especificidade mede a capacidade do classificador em identificar corretamente a ausência do distúrbio quando, de fato, ele não existe.

Foram testadas as wavelets de Daubechies de ordem 1 a 45, a fim de investigar qual delas apresentava melhor desempenho, em termos de acurácia. Tais funções foram testadas quanto à capacidade de classificação em Voz Saudável x Voz Desviada (GG1 x (GG2 e GG3)), utilizando-se das nove medidas de energia. Os resultados obtidos são apresentados na Fig. 2.

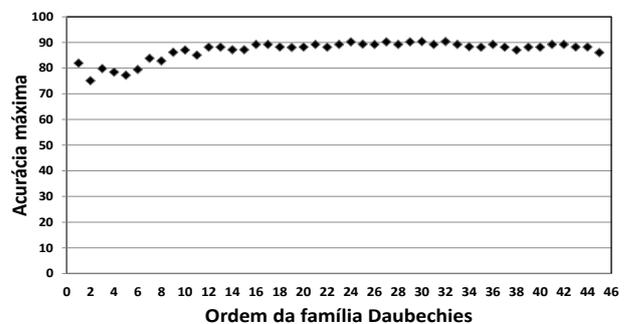


Fig. 2. Classificação GG1 x (GG2 e GG3) para as wavelets da família de Daubechies.

Inicialmente, à medida que se aumenta a ordem da wavelet, o desempenho do sistema também aumenta. Entretanto, a partir da ordem 12 o desempenho permanece praticamente inalterado. Quanto maior a ordem da wavelet, maior o número de coeficientes dos filtros passa-baixas e passa-altas empregado na decomposição, o que acarreta um maior custo computacional. Quanto menor a duração da função no tempo, mais características de alta frequência serão considerados.

Pelos motivos citados, optou-se por empregar a família db12 para a classificação dos sinais de vozes infantis. Foram considerados quatro casos distintos: GG1 x (GG2 e GG3), GG1 x GG2, GG1 x GG3 e GG2 x GG3. O primeiro caso de classificação pode ser utilizado para triagem vocal, onde o sinal de voz é considerado sem ou com desvio. Os demais casos de classificação podem ser utilizados para monitoramento da terapia fonoaudiológica, já que classificam os desvios vocais quanto a sua severidade.

As Tabelas I-IV apresentam os melhores resultados obtidos para todos os casos de classificação considerados. Inicialmente, a energia e a entropia dos coeficientes de detalhe das wavelets foram avaliadas individualmente para cada nível de resolução. Em seguida, as características foram agrupadas em um único vetor, contendo 18 medidas (9 de energia e 9 de entropia). Por fim, as características foram combinadas entre si, sendo selecionada a combinação que apresentou a melhor taxa de classificação.

#### IV. DISCUSSÃO

Neste trabalho, 45 wavelets da família de Daubechies foram avaliadas na classificação do grau de severidade de distúrbios vocais infantis. A família de ordem 12 destacou-se por manter a mesma taxa de acurácia com um menor custo computacional o que a torna interessante para aplicações em tempo real.

Os sinais desviados possuem ruídos nas altas e baixas frequências, caracterizados por sopro e rugosidade, respectivamente. Nas classificações entre as vozes saudáveis e vozes desviadas (Tabelas I-III), quando considerada as medidas individuais, a energia nas baixas e a entropia nas altas frequências foram determinantes na discriminação. Quando considerada a combinação das medidas, a energia no nível 1 de resolução esteve presente em todas as combinações, bem como a entropia no nível de resolução 5. Isto indica a sopro e rugosidade como um fator determinante para a separação das classes em termos de energia e a rugosidade em termos da entropia.

Na classificação entre sinais desviados (Tabela IV) observa-se que a energia nas mais baixas frequências é relevante, tendo em vista que ambas as classes possuem sopro e rugosidade. As entropias nas altas frequências tiveram maior poder de discriminação.

Na classificação GG1x(GG2 e GG3) os valores de especificidade (Esp), conforme mostrado na Tabela I, apresentam valores bem mais altos que a sensibilidade (Sens). Isto se deve ao fato do número de sinais sem desvio (GG1) ser bem menor que os sinais alterados (10 e 83, respectivamente). Essa disparidade tornou o sistema de classificação especialista em identificar a ausência do desvio.

TABELA I. CLASSIFICAÇÃO GG1 x (GG2 e GG3)

Medidas		Acurácia (%)	Sens (%)	Esp (%)
Individual	$E_5$	62,33 ± 5,28	40 ± 16,33	65 ± 5,93
	$H_1$	77,22 ± 6,72	0 ± 0	86,53 ± 7,52
Combinação	Todas	85,88 ± 2,31	10 ± 10	95,14 ± 1,98
	$E_1E_3E_4E_6E_7$ $H_5H_8$	95,77 ± 1,72	60 ± 16,33	100 ± 0

TABELA II. CLASSIFICAÇÃO GG1 x GG2

Medidas		Acurácia (%)	Sens (%)	Esp (%)
Individual	$E_6$	60 ± 5,21	50 ± 16,67	61,43 ± 6,08
	$H_1$	78,75 ± 5,29	30 ± 15,28	80 ± 13,34
Combinação	Todas	85 ± 2,50	10 ± 10	95,72 ± 2,18
	$E_1E_3E_4E_5$ $H_5H_8$	96,25 ± 1,90	70 ± 15,28	100 ± 0

TABELA III. CLASSIFICAÇÃO GG1 x GG3

Medidas		Acurácia (%)	Sens (%)	Esp (%)
Individual	$E_7$	76,67 ± 13,20	80 ± 13,34	75 ± 13,34
	$H_2$	75,00 ± 8,70	60 ± 16,34	90 ± 6,67
Combinação	Todas	65,00 ± 6,22	20 ± 13,34	100 ± 0
	$E_1E_6E_7H_5$	96,66 ± 3,34	90 ± 7,52	100 ± 0

TABELA IV. CLASSIFICAÇÃO GG2 x GG3

Medidas		Acurácia (%)	Sens (%)	Esp (%)
Individual	$E_6$	71,94 ± 4,32	72,86 ± 3,34	60 ± 16,34
	$H_2$	58,33 ± 5,99	54,29 ± 6,67	85 ± 10,68
Combinação	Todas	66,66 ± 7,86	80 ± 13,34	60 ± 14,54
	$E_6E_7$ $H_2H_3H_4H_5$	89,02 ± 2,33	94,29 ± 2,33	65 ± 15,73

No caso GG1xGG3, o nível de energia 7,  $E_7$ , apresentou um maior equilíbrio entre as medidas de desempenho, faixa na qual está a frequência fundamental (frequência de vibração das pregas vocais) das vozes infantis.

O mesmo ocorre em todas as classificações entre os sinais sem e com desvio (ver Tabelas II e III). A combinação das medidas, entretanto, proporcionou um pouco mais de equilíbrio ao sistema. Na classificação entre as vozes sem desvio e com desvio moderado (GG1xGG3), a discriminação entre os sinais ficou mais evidente com sensibilidade e especificidades acima de 90%, apresentando ainda a maior acurácia.

#### V. CONSIDERAÇÕES FINAIS

Os resultados obtidos indicam que a transformada wavelet de ordem 12 é adequada para avaliação do grau de intensidade dos desvios vocais em crianças. A energia normalizada e a entropia dos coeficientes wavelets conseguiram representar os distúrbios vocais com taxas de acurácia acima de 95 % foram obtidas na separação entre vozes saudáveis e desviadas. A especificidade e sensibilidade apresentaram melhores resultados com a combinação das medidas. No entanto, é necessário, para que o sistema seja especialista tanto na detecção de desvio, quanto na ausência do mesmo. Na discriminação entre o grau de severidade das vozes desviadas

uma acurácia de 89% foi alcançada, com sensibilidade de 94,29%.

Para todos os casos de classificação considerados, a combinação das medidas resultou num aumento significativo da acurácia e reduziu o vetor de características de 18 para, no máximo, 7 medidas. Desta forma, conclui-se que a análise por faixas de frequência através da transformada wavelet mostra-se como uma técnica promissora que pode ser considerada para ser empregada como uma ferramenta para análise acústica de desvios vocais.

#### AGRADECIMENTOS

Os autores agradecem ao IFPB, ao programa de pós-graduação, pelo suporte financeiro a esta pesquisa, bem como ao Laboratório Integrado de Estudos da Voz (LIEV) da UFPB, pela concessão da Base de Dados.

#### REFERENCES

- [1] V. J. D. Vieira, "Avaliação de distúrbios da voz por meio de análise de quantificação de recorrência", Dissertação de Mestrado, Instituto Federal de Educação Ciência e Tecnologia da Paraíba, Curso de Pós-Graduação em Engenharia Elétrica, João Pessoa, 2014.
- [2] R. Oliveira, L. Teixeira, A. Gama, A. Medeiros, "Análise perceptivo-auditiva, acústica e autopercepção vocal em crianças", *Jornal da Sociedade Brasileira de Fonoaudiologia*; 2011; vol. 23, no. 2, pp. 158-163.
- [3] A. Parraga, "Aplicação da transformada wavelet packet na análise e classificação de sinais de vozes patológicas," Master's thesis, Porto Alegre, Brasil, 2002.
- [4] J. Martens, H. Versnel, and P. H. Dejonckere, "The effect of visible speech in the perceptual rating of pathological voices," *Archives of Otolaryngology-Head & Neck Surgery*, vol. 133, no. 2, pp. 178-185, 2007.
- [5] G. Niedzielska, "Acoustic analysis in the diagnosis of voice disorders in children". *Int. J. Pediatr. Otorhinolaryngol.*, Mar 2001, 57(3), pp.189-93.
- [6] E. S. Fonseca, R. C. Guido, P. R. Scalassara, C. D. Maciel, and J. C. Pereira, "Wavelet timefrequency analysis and least squares support vector machines for the identification of voice disorders," *Computers in Biology and Medicine*, vol. 37, no. 4, pp. 571-578, 2007.
- [7] M. K. Arjmandi and M. Pooyan, "An optimum algorithm in pathological voice quality assessment using wavelet-packet-based features, linear discriminant analysis and support vector machine," *Biomedical Signal Processing and Control*, vol. 7, no. 1, pp. 3-19, 2012.
- [8] R. Behroozmand and F. Almasganj, "Optimal selection of wavelet-packet-based features using genetic algorithm in pathological assessment of patients' speech signal with unilateral vocal fold paralysis," *Computers in Biology and Medicine*, vol. 37, no. 4, pp. 474-485, 2007.
- [9] I. Daubechies et al., *Ten lectures on wavelets*, vol. 61. SIAM, 1992.
- [10] L. W. Lopes, I. L. B. Lima, L. N. A. Almeida, D. P. Cavalcante, and A. A. F. de Almeida, "Severity of voice disorders in children: correlations between perceptual and acoustic data", *Journal of Voice*, vol. 26, no. 6, pp. 819-e7, 2012.
- [11] M. I. R. Goncalves, P. A. L. Pontes, V. P. Vieira, A. A de L. Pontes, D. Curcio, N. G. De Blase. "Transfer function of Brazilian Portuguese oral vowels: a comparative acoustic analysis". *Braz. j. otorhinolaryngol*, vol.75, n.5, pp. 680-684, 2009.
- [12] R. Yamasaki, S. Leão, G. Madazio, M. Padovani, R. Azevedo, and M. Behlau, "Correspondência entre escala analógico-visual e a escala numérica na avaliação perceptivo-auditiva de vozes", *XVI Congresso Brasileiro de Fonoaudiologia*, pp. 24-27, 2008.
- [13] A. D. Poularikas, *The Transforms and Applications Handbook*, CRC Press, 2010.
- [14] S. Mallat, *A wavelet tour of signal processing*. Academic press, 1997.