

# Sobre a escolha dos filtros wavelet e da função janela para sistemas de melhoramento de voz baseados na DWT

Exayne Santos Mourão e Caio Cesar Enside de Abreu

**Resumo**— Este trabalho apresenta um estudo empírico sobre a escolha dos filtros wavelet e das funções janela utilizadas em sistemas de melhoramento de voz baseados na transformada wavelet discreta - DWT. Simulações foram realizadas combinando diversos filtros wavelet com diversos tipos de janela e realizando o processamento de várias sentenças. Os resultados mostraram que a função wavelet db 10, em conjunto com a janela de Hanning, produziram os melhores resultados.

**Palavras-Chave**— Filtros wavelet, janelamento, Melhoramento de voz.

**Abstract**— This paper presents an empirical study on the choice of the wavelet filters and window functions used in speech enhancement systems based on discrete wavelet transform - DWT. Simulations were performed by combining various wavelet filters with different window types and performing the speech processing. Results showed that the wavelet function db 10, together with the Hanning window, generated the best results.

**Keywords**— Wavelet filters, windowing, speech enhancement.

## I. INTRODUÇÃO

O meio mais usado para se comunicar é através da fala. No entanto, ao transmitir um sinal de voz através de um sistema de comunicação, este pode sofrer algum tipo de interferência. Um sinal de voz pode ser contaminado por ruído durante o processo de aquisição, por meio de um microfone, ou devido ao meio de reprodução. A presença do ruído afeta diretamente a qualidade e a inteligibilidade de um sinal de voz [2]. Neste sentido, uma importante área de pesquisa é aquela que desenvolve métodos de redução de ruído para sinais de voz. Dentre estes métodos, destacam-se aqueles que utilizam a DWT [2].

A implementação da DWT é baseada em um banco de filtros digitais, onde os filtros estão diretamente relacionados à função wavelet escolhida para o processamento [2]. Apesar da DWT ser amplamente utilizada para a redução do ruído presente em sinais voz [2], não foi encontrado na literatura uma análise mais detalhada sobre a escolha de tais filtros. Neste sentido, o objetivo deste trabalho é apresentar dados empíricos que auxiliem nesta escolha.

Além da escolha dos filtros wavelet, este trabalho aborda a escolha da janela mais adequada para a segmentação de

um sinal de fala. A aplicação da técnica de janelamento é essencial, pois, um sinal de voz pode possuir longa duração. Dessa forma, o custo computacional é reduzido consideravelmente [4].

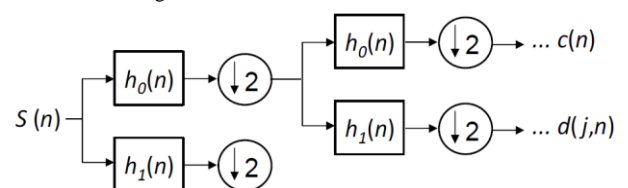
A fim de avaliar a qualidade do processamento realizado, adotando diferentes funções wavelet e diferentes tipos de janela, utilizaram-se neste trabalho algumas medidas de qualidade objetivas clássicas da área: PESQ (Perceptual Evaluation of Speech Quality); SNR (Signal to Noise Ratio) e MSE (Minimum Mean Square Error) [2].

## II. TRANSFORMADA WAVELET DISCRETA

A DWT é uma ferramenta eficiente no processamento de sinais não-estacionários [2]. Uma de suas principais características é a possibilidade de ajustar a escala e assim obter máxima resolução no processo de decomposição, minimizando possíveis perdas de informação [1]. Outra característica relevante da DWT é a concentração de energia do sinal transformado em um número pequeno de coeficientes. Esta característica fundamenta a aplicação de métodos de redução de ruído baseados em um limiar [2]. Os métodos baseados no limiar consideram com ruído os coeficientes do sinal cujo valor absoluto é menor que um determinado valor [1]. Dessa forma, é possível eliminar parte do ruído preservando as informações pertinentes do sinal original [2].

Como destacado anteriormente, a implementação da DWT é baseada em um banco de filtros digitais, como mostra a Figura 1.

Fig 1. Banco de filtros de análise da DWT.



Analisando a Figura 1,  $h_0$  e  $h_1$  representam filtros passa baixa e passa alta, respectivamente.  $c(n)$  e  $d(j,n)$  representam os coeficientes de aproximação e detalhes na escala  $j$ , respectivamente. De acordo com [3], os coeficientes de detalhes na escala  $j$  correspondem à frequência situada aproximadamente entre  $(2^{-j}f_s, 2^{-j-1}f_s)$ . Dessa forma, uma importante propriedade da DWT

desejável ao processamento de sinais de voz é a sucessiva separação de altas e baixas frequências [2].

### III. SIMULAÇÕES

O banco de dados utilizado neste estudo é denominado NOIZEUS [5]. O banco de dados NOIZEUS é composto por trinta sentenças foneticamente balanceadas na língua inglesa. As sentenças foram produzidas por três oradores do sexo masculino e três do sexo feminino, totalizando quinze sinais de fala na voz masculina e quinze na voz feminina. Todos os sinais possuem frequência de amostragem de 8kHz.

A fim de avaliar a influência da escolha da janela e dos filtros wavelet em sistemas de melhoramento de voz, dois experimentos foram realizados:

*Experimento 1.* Com o objetivo de verificar o erro (no sentido MSE) envolvido na operação de janelamento, quando aplicada em conjunto com o processo análise/síntese produzido pela DWT, fixou-se a escolha de uma função wavelet e variou-se o tipo de janela. Neste experimento, apenas sinais puros foram utilizados e foi fixado  $j=3$ . O janelamento foi realizado com 50% de sobreposição.

*Experimento 2.* A fim de avaliar a influência da escolha dos filtros wavelet no processo de redução do ruído em sinais de voz, fixou-se a janela mais apropriada (de acordo com o experimento 1) e variou-se a wavelet escolhida. Para cada escolha dos filtros, o limiar *soft* foi utilizado para a redução do ruído de fundo presente nos 30 sinais de teste.

Para a execução do experimento 1, a função wavelet escolhida foi a Daubechies de ordem 4 (db 4). Já para o experimento 2, utilizou-se sentenças contaminadas com o ruído Gaussiano branco. O nível de contaminação de todos os sinais foi fixado em 10dB.

As Tabelas I e II apresentam os resultados dos experimentos 1 e 2, respectivamente.

TABELA I. ERRO ENVOLVIDO NO PROCESSO DE JANELAMENTO

JANELA	MSE
Bartlett	$2,3030 \times 10^{-7}$
Hamming	$9,7399 \times 10^{-6}$
Hanning	$1,8639 \times 10^{-7}$
Kaiser	$1,2191 \times 10^{-6}$
Retangular	$1,8000 \times 10^{-3}$

Note, a partir da Tabela I, que a escolha do tipo de janela não deve exercer forte influência sobre o desempenho de um sistema de melhoramento de voz. Isso se justifica devido ao baixo erro envolvido no processo de janelamento. Dentre as janelas testadas, a janela de Hanning e a janela retangular forneceram o melhor e o pior resultado, respectivamente. Além disso, vale destacar que os resultados apresentados na Tabela I também incorporam um erro muito pequeno (garantido pela propriedade de perfeita reconstrução do banco de filtros [2]) relacionado à DWT.

Com base na Tabela I, a janela Hanning foi selecionada para a execução do experimento 2. Para a interpretação da Tabela II, vale ressaltar que quanto maior a nota PESQ, melhor será a qualidade do sinal processado. Avaliada em

decibéis, quanto maior for a SNR de um sinal de voz, menor será o nível do ruído de fundo.

TABELA II. DESEMPENHO DOS FILTROS WAVELET

WAVELET	PESQ	SNR (dB)
Haar	3,4328	13,4894
db 2	3,4524	13,7842
db 4	3,4681	13,8742
db 10	3,4868	14,9233
db 22	3,4204	13,6068
Coiflets 1	3,4499	13,8788
Coiflets 3	3,4570	14,2941
Symlets 2	3,4554	13,7842
Symlets 8	3,4472	14,0668
Biortogonal 1.1	3,4328	13,4874

Analisando a Tabela II, verifica-se que a escolha dos filtros wavelet influenciam diretamente no processo de redução do ruído de fundo em sinais de voz. A função wavelet que se destacou foi a db 10, proporcionando em média quase 5dB de melhoria na SNR. Além disso, conseguiu aliar forte atenuação do ruído com boa qualidade no processamento, verificado pela nota PESQ alcançada. Outro ponto destacado durante as simulações é que filtros com suporte muito longo ou muito curto, não são desejáveis para o melhoramento de sinais de voz. Veja as funções wavelet de Haar, db 22 e Biortogonal.

### IV. CONSIDERAÇÕES FINAIS

A partir das simulações realizadas neste trabalho, verificou-se que a escolha dos filtros wavelet influencia diretamente no resultado da operação de filtragem de sinais de voz. A função wavelet que gerou os melhores resultados foi a db 10. A função db 10 gera um conjunto de filtros com suporte desejável (20 coeficientes), não sendo muito curto e nem muito longo. Com relação ao tipo de janela utilizada nas simulações, verificou-se que a grande maioria, com exceção da janela retangular, apresenta bons resultados ao trabalharem em conjunto com o processo análise/síntese produzido pela DWT.

Como trabalho futuro pretende-se estudar os diversos tipos de limiarização wavelet.

### AGRADECIMENTOS

Os autores gostariam de agradecer a FAPEMAT (processo 419307/2015), pelo apoio ao desenvolvimento do projeto.

### REFERÊNCIAS

- [1] C. C. E. ABREU "Uso de equações de diferenças na obtenção de filtros para redução de ruído em sinais de voz no domínio wavelet", Dissertação (Mestrado em Engenharia Elétrica) – Universidade Estadual Paulista Júlio de Mesquita Filho, 2013.
- [2] M. A. Q. DUARTE, "Redução de ruído em sinais de voz no domínio wavelet", Tese (Doutorado em Engenharia Elétrica) — Universidade Estadual Paulista Júlio de Mesquita Filho, 2005.
- [3] S. A. MALLAT, *Wavelet tour of signal processing*, San Diego: Academic Press, 1998.
- [4] A. V. OPPENHEIM, *Discrete-time signal processing*. Englewood Cliffs, NJ: Prentice hall, 1989.
- [5] Y. HU, P. C. LOIZOU, "Subjective comparison and evaluation of speech enhancement algorithms", *Speech communication*, 49, (7) pp. 588-601, 2007.