

Comparação de métodos para supressão de ruídos em sinal de voz utilizando Transformada Wavelet

Gustavo dos Santos Cardoso
DECOM - UNICAMP
Universidade Estadual de Campinas
São Paulo, Campinas
Email: gustavo_16a@hotmail.com

Gustavo Peglow Kuhn
DEE - IFsul
Instituto Federal Sul-rio-grandense
Rio Grande do Sul, Pelotas
Email: gustavowolgep@gmail.com

Samuel dos Santos Cardoso
IEB - UFSC
Universidade Federal de Santa Catarina
Santa Catarina, Florianópolis
Email: samuel.ifsul.ee@gmail.com

Abstract—This paper discusses two methods for noise reduction in speech signals through Discrete Wavelet Transform using the db6 Wavelet. The methods described were Hard-Thresholding and Soft-Thresholding. These methods use statistical information from the transformed signal to obtain a threshold for denoising. In order to compare the methods, tests have been performed using speech signals contaminated with white noise content. Different signal noise ratio were used. The methods were compared using the output signal noise ratio and Itakura-Saito distance. Results show better performance for the Hard-Thresholding method.

Index Terms—Speech signal, Wavelet Transform, Hard-Thresholding, Soft-Thresholding.

I. INTRODUÇÃO

Em muitas aplicações que envolvem a transmissão de sinais de voz, a presença de ruídos aditivos pode reduzir significativamente a qualidade e a inteligibilidade dos sinais [1], [2]. Neste contexto, vários pesquisadores tem despendido esforços em pesquisas voltadas a melhoria de técnicas de processamento de sinais de áudio no intuito de reduzir ou eliminar o ruído aditivo [3], [4], [5], [6]. Existem muitos métodos para remoção e supressão de ruídos, desde métodos simples a métodos complexos, com menor ou maior custo computacional [2]. Entre as técnicas utilizadas em processamento de sinais de voz destacam-se as baseadas no uso de artifícios como a Transformada Wavelet (TW), filtros adaptativos [7], redes neurais [8], subtração espectral [9], entre outras. Dentre esses métodos, os que utilizam a TW destacam-se por serem relativamente simples, rápidos e eficientes [5].

A metodologia baseada na TW consiste em decompor janelas do sinal de entrada em coeficientes no domínio frequência e calcular um limiar para corte dos ruídos baseado no desvio padrão. Esse limiar é então utilizado para separar quais coeficientes representam o sinal e quais representam o ruído. Após separados o sinal é reconstruído por meio da Transformada Wavelet Inversa (TWI) sem as componentes do ruído [1], [10], [11].

Tendo em vista a simplicidade e eficiência da TW, neste artigo será realizada a análise entre duas metodologias de supressão de ruídos baseadas na Transformada Wavelet Discreta (TWD) conhecidas como *Soft-thresholding* e *Hard-thresholding*.

Nas seções seguintes são detalhados os processos de desenvolvimento e teste da proposta. Na seção II é retratada a metodologia do trabalho. Na seção III são descritos os experimentos e resultados alcançados. Por fim, a seção IV apresenta a conclusão.

II. METODOLOGIA

A. O Princípio da remoção de ruído pela Transformada Wavelet

Em casos práticos, os sinais dos ruídos geralmente aparecem associados à sinais de alta frequência e os sinais úteis aparecem como sinais de baixa frequência [3]. Como os ruídos estão associados as altas frequências, quando o sinal de voz com ruído aditivo é decomposto pela Transformada Wavelet Discreta o ruído fica associado aos componentes de alta frequência dos filtros Wavelet e, por consequência, o sinal de interesse fica associado aos coeficientes de baixa frequência [3].

No contexto descrito, o processo de remoção do ruído de um sinal dimensional funciona da seguinte forma: primeiro o sinal é decomposto pela TWD com o objetivo de selecionar limiares adequados para remoção do ruído nos coeficientes Wavelets do sinal transformado; e posteriormente o sinal é reconstruído por meio da Transformada Wavelet Discreta Inversa (TWDI) reduzindo as componentes com ruído [11]. Embora esse processo consiga eliminar boa parte dos ruídos que estão associados as altas frequências, ruídos de baixa frequência permanecem intrínsecos ao sinal filtrado [1], [10].

Nesse cenário, um dos principais fatores que afetam a qualidade na remoção do ruído são a supressão do ruído próximo ao limiar entre as baixas e altas frequências e a seleção da função do limiar para a associação dos coeficientes a estas frequências [1], [5], [10].

B. Transformada Wavelet e Transformada Wavelet Inversa

Seja $s(t)$ um sinal de tempo contínuo e $\psi_{a,b}(t)$ a função Wavelet, defini-se a TW ($Ws(a,b)$) do sinal $s(t)$ como [1]:

$$Ws(a,b) = \int_{-\infty}^{+\infty} s(t) \cdot \psi_{a,b}(t) dt \quad (1)$$

Em (1) a função Wavelet $\psi_{a,b}(t)$ é definida por (2) que está associada a uma função de escala representada por (3):

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

$$\phi_{a,b}(t) = \frac{1}{\sqrt{a}} \cdot \phi\left(\frac{t-b}{a}\right) \quad (3)$$

onde b representa o fator de deslocamento e a o fator de escala da Wavelet.

Alguns autores como em [1], [10] sugerem filtros Wavelet com largura de banda de uma oitava, o que implica na discretização de $a = 2^m$ e $b = 2^m \cdot k$, onde k e m são números inteiros. Assim, as funções Wavelet $\psi_{a,b}(t)$ e de escala $\phi_{a,b}(t)$ podem ser reescritas por (4) e (5), respectivamente [12].

$$\psi_{a,b}(t) = 2^{\frac{m}{2}} \cdot \psi\left(\frac{t - 2^m \cdot k}{2^m}\right) \quad (4)$$

$$\phi_{a,b}(t) = 2^{\frac{m}{2}} \cdot \phi\left(\frac{t - 2^m \cdot k}{2^m}\right) \quad (5)$$

Desta maneira é possível obter uma decomposição diádica do sinal, na qual a função de escala é usada essencialmente para definir os coeficientes das aproximações do sinal filtrado e a função Wavelet é utilizada para definir os coeficientes dos detalhes do mesmo sinal [1], [10]. Neste contexto, a decomposição diádica permite uma análise multirresolução com o uso de bancos de filtros associados as funções Wavelet e de escala para a decomposição do sinal em vários níveis, definindo deste jeito a TWD [1], [13], [14], [15].

O processo de decomposição Wavelet consiste na passagem de um sinal discreto $S[n]$ por um filtro passa-baixas L_D e por um filtro passa-altas H_D , onde os coeficientes destes filtros estão associados respectivamente as funções $\psi_{a,b}(t)$ e $\phi_{a,b}(t)$ [16]. Após a filtragem do sinal por cada um dos filtros é feita a subamostragem do sinal resultante com um fator de 2. Após este processo são obtidos os coeficientes das aproximações c_j e os coeficientes dos detalhes d_j , ambos relativos ao sinal. A descrição do processo de decomposição Wavelet está ilustrado na Figura 1.

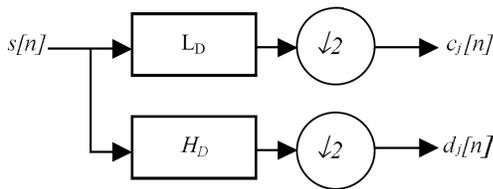


Figura 1. Decomposição Wavelet de 1 nível

O processo de decomposição pode ser repetido diversas vezes até que o número de amostras nos coeficientes c_j e d_j seja de apenas uma amostra, compondo a chamada árvore de decomposição Wavelet diádica, apresentada na Figura 2 [1], [3].

Após sua decomposição o sinal S pode ser reconstruído através de um processo inverso chamado de Transformada Wavelet Discreta Inversa. Esse processo consiste em realizar uma sobre-amostragem com zeros nos coeficientes c_j e d_j , após

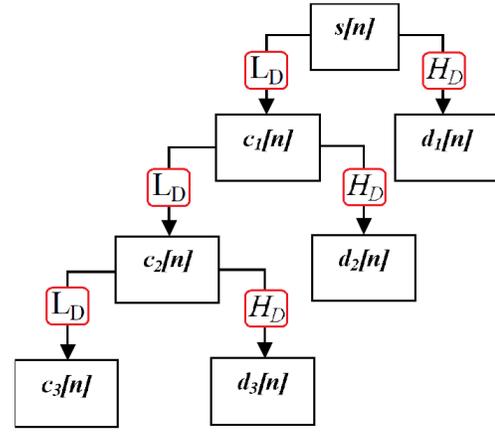


Figura 2. Decomposição Wavelet em múltiplos níveis

isso esses coeficientes passam pelos filtros de reconstrução L_R e H_R para que então sejam obtidos os coeficientes de aproximação de um nível acima. Este processo deve ser feito até se obter o sinal reconstruído e os filtros L_R e H_R devem ser projetados de maneira que os mesmos removam os efeitos de distorção que os filtros de decomposição L_D e H_D possam causar [1], [13], [14], [15], [17]. O processo da TWDI é ilustrado na Figura 3.

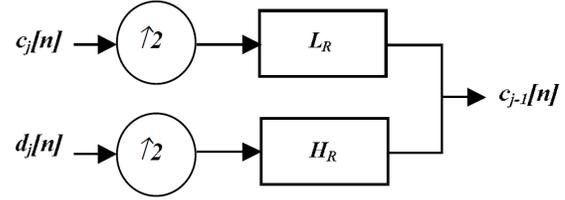


Figura 3. Reconstrução Wavelet em 1 nível

C. Métodos de Redução de Ruído a Partir de Limiares

1) *Limiar universal* λ : um dos trabalhos mais divulgados sobre aplicações em supressão de ruídos é descrito por [18] e [19]. Em particular, [18] descreve algumas regras para a seleção do limiar. Entre estas regras está a de seleção do limiar universal λ , o qual é utilizado nos métodos *Hard-Thresholding* e *Soft-Thresholding*. O limiar universal λ é definido por (6):

$$\lambda = \sigma \cdot \sqrt{2 \cdot \log_2(N)} \quad (6)$$

sendo N o número de amostras (comprimento) do sinal ruidoso e σ o desvio padrão do ruído, o qual foi estimado por (7):

$$\sigma = \left(\frac{1}{N-1} \sum_{i=1}^N (s_i - \bar{s})^2 \right)^{\frac{1}{2}} \quad (7)$$

sendo N o número de amostras utilizadas para calcular o σ , s_i a amostra do sinal e \bar{s} a média do sinal.

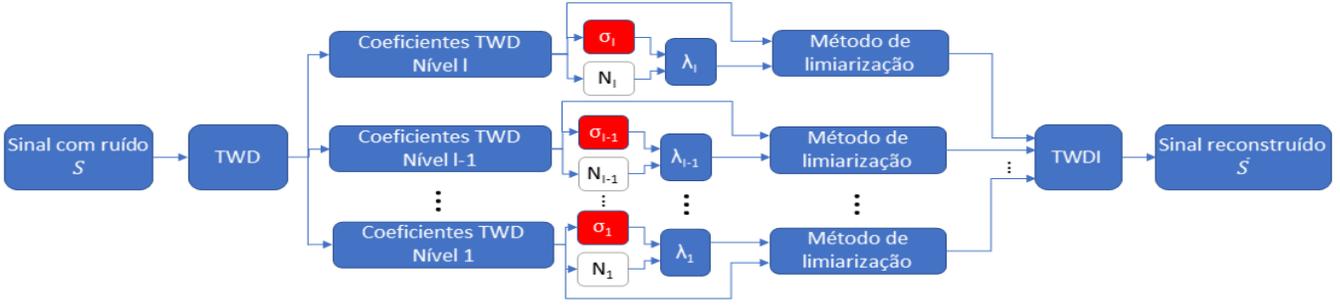


Figura 4: Diagrama de blocos do processo de remoção de ruído

2) *Hard-Thresholding*: o método *Hard-Thresholding* consiste em zerar os coeficientes que possuem valores em módulo menores que o valor do limiar e manter os coeficientes com valores em módulo maiores ou iguais ao do limiar [18], [19]. Esse método não altera as propriedades locais do sinal, mas devido à descontinuidade ao zerar os coeficientes do sinal, o mesmo causa uma certa deturpação no sinal reconstruído. A descrição matemática deste método é dada por (8).

$$\bar{s}_i = \begin{cases} s_i, & |s_i| \geq \lambda \\ 0, & |s_i| < \lambda \end{cases} \quad (8)$$

onde \bar{s}_i é o coeficiente obtido, s_i é a amostra do sinal analisada e λ é o limiar de comparação.

3) *Soft-Thresholding*: o método *Soft-Thresholding* busca diminuir as descontinuidades que podem ser geradas quando se utiliza o *Hard-Thresholding*. Este método consiste também em zerar os coeficientes que possuem valores em módulo menores que o valor do limiar [5], [19]. No entanto, diferentemente do método anterior, os coeficientes que possuem valores maiores que o do limiar terão o módulo de seus valores reduzidos com o valor de λ , o que implica na modificação do sinal reconstruído inclusive nos trechos em que ocorre a fala. O cálculo dos coeficientes é dado por (9):

$$\bar{s}_i = \begin{cases} \text{sgn}(s_i)(|s_i| - \lambda), & |s_i| \geq \lambda, \\ 0, & |s_i| < \lambda. \end{cases} \quad (9)$$

onde \bar{s}_i é o coeficiente obtido, s_i é a amostra do sinal analisada, λ é o limiar de comparação e sgn é a função que retorna se o sinal da amostra é positivo ou negativo.

D. Descrição dos Algoritmos utilizados

Nesta subseção são descritos os passos realizados para execução dos métodos. Primeiro o sinal de voz é carregado junto com a informação da frequência de amostragem dele. A frequência de amostragem é necessária para saber quantas amostras (N_l) compõem um intervalo de 300 ms, pois neste intervalo, considera-se que há somente o ruído aditivo para estimar o desvio padrão do mesmo. Após, decompõem-se o sinal em cinco níveis com a TWD usando a wavelet Daubeachies de sexta ordem - db6 [20]. Então, para determinar

o desvio padrão, em cada nível, utiliza-se janelas de 23 ms com sobreposição de 50% até os 300 ms de sinal para estimar a média do desvio padrão. Desta forma a média do desvio padrão em um determinado nível (σ_l) é gerado, sendo l o nível da decomposição da TWD. Após, utiliza-se a fórmula do limiar universal, para obter-se um limiar para cada nível da TWD denominado λ_l . Depois, os métodos de redução de ruído por limiares são aplicados em cada nível de decomposição da TWD com seus respectivos limiares e médias de desvios padrões. Por fim, utilizando a TWDI, realiza-se a reconstrução do sinal com o ruído reduzido. A Figura 4 mostra o diagrama de blocos do processo realizado.

III. EXPERIMENTOS E RESULTADOS

A. Descrição do Experimento

Para os experimentos foram utilizados sinais limpos gravados a uma frequência de amostragem de 44,1 kHz, aos quais foram adicionados ruído branco gaussiano aditivo. Para testar e comparar os métodos, diversos valores de relação sinal ruído de entrada para o ruído branco adicionado foram utilizados, sendo estes de: 1 dB, 2 dB, 3 dB, 4 dB, 5 dB, 6 dB, 8 dB, 10 dB, 12 dB, 15 dB, 18 dB, 22 dB, 28 dB, 30 dB, 35 dB e 45 dB. Foram realizados ao total 32 testes, 16 para o método *Hard-Thresholding* e outros 16 para o método *Soft-Thresholding*.

A implementação dos algoritmos e as simulações foram todas realizadas no software Matlab (*MathWorks Inc.*) versão R2013a, o qual possui uma grande coleção de funções para a análise Wavelet [3], [4], [6], [17].

Para avaliar e comparar os resultados dos algoritmos descritos na subseção II.c foram utilizadas as medidas de relação sinal ruído após a filtragem (SNR_o) e a distância de Itakura-Saito (d_{IS}). A SNR_o fornece uma medida da qualidade do sinal em relação ao ruído, entretanto durante o processo de filtragem podem ocorrer certas distorções do sinal original. Neste caso em que há distorções, a distância de Itakura-Saito d_{IS} fornece melhores resultados como medida de distorção do sinal e pode ser calculada conforme (10) [21]:

$$d_{IS} = \sum_{i,\omega} \frac{P_{i,\omega}}{\hat{P}_{i,\omega}} - \ln \left(\frac{P_{i,\omega}}{\hat{P}_{i,\omega}} \right) - 1 \quad (10)$$

onde $P_{i,\omega}$ é a potência espectral do sinal real (original sem ruído) em um determinado instante i na frequência ω e $\hat{P}_{i,\omega}$ é a potência espectral estimada após a filtragem.

A relação sinal ruído SNR_o após o processo de filtragem do sinal é expressa por (11) [4]:

$$SNR_o = 10 \cdot \log \left(\frac{P_{sinal}}{P_{ruído}} \right) \quad (11)$$

onde P_{sinal} é a potência do sinal e $P_{ruído}$ é a potência do ruído remanescente.

Além das métricas numéricas foram feitas algumas análises qualitativas sobre os gráficos dos sinais no domínio do tempo (sinal original e reconstruído) e de seus respectivos espectrogramas.

B. Resultados

Para melhor entender o uso dos dois métodos de limiarização (*Hard-Thresholding* e *Soft-Thresholding*) foram realizadas análises qualitativas sobre as imagens do sinal de voz contaminado com ruído branco e relação sinal ruído de entrada SNR_i de 1 dB e 15 dB, pois estes sinais ressaltam de maneira clara as características apontadas pelos métodos.

Na primeira análise, o sinal de voz foi contaminado com ruído branco, obtendo uma relação sinal ruído de entrada (SNR_i) de 1 dB. Os resultados obtidos no domínio do tempo para o método *Hard-Thresholding* estão ilustrados na Figura 5, na qual é feita uma comparação realizada entre o sinal contaminado, o sinal filtrado e o sinal original.

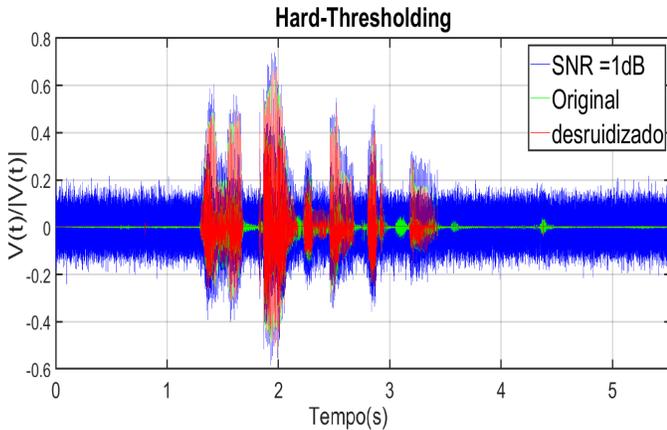


Figura 5. Sinais de fala no domínio tempo sendo o azul o sinal com $SNR_i(\text{dB})=1$, em verde o original e vermelho o reconstruído

Analisando a Figura 5 percebe-se no sinal em vermelho (filtrado) uma considerável redução de ruído, principalmente em instantes de silêncio. Entretanto, além da redução do ruído ocorrem distorções na fala próximo aos períodos de silêncio. Essas constatações qualitativas corroboram com a relação sinal ruído de saída (SNR_o) de 11,1 dB e uma distorção espectral medida pela distância de Itakura-Saito (d_{IS}) de 0,1839.

Além das análises realizadas no domínio do tempo, outro parâmetro bastante utilizado para avaliar os resultados oriundos do processamento é o espectrograma. As Figuras 6(a), 6(b)

e 6(c) correspondem aos espectrogramas do sinal de voz sem ruído, sinal contaminado e sinal estimado (filtrado).

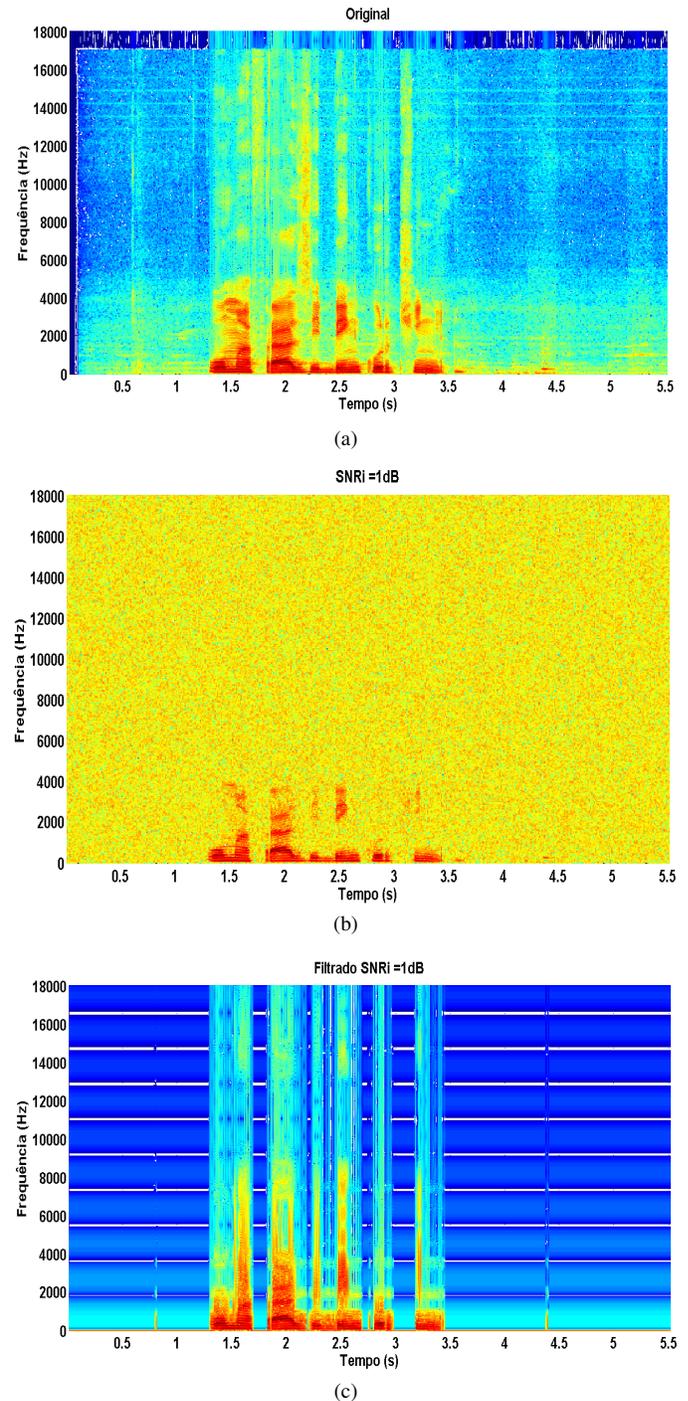


Figura 6. (a) Espectrograma do sinal original; (b) espectrograma do sinal com ruído aditivo e $SNR_i(\text{dB}) = 1$; (c) espectrograma do sinal ($SNR_i=1\text{dB}$) após a filtragem com o método limiar *Hard-Thresholding*

Conforme a análise dos três espectrogramas, percebe-se na Figura 6(c) que entre as frequências de 2000 a 4000 Hz a intensidade do sinal é ressaltada pelo algoritmo baseado na limiarização *Hard-Thresholding*. Embora o algoritmo tenha

acentuado as faixas mais importantes da palavra elétrica presentes no espectrograma do sinal reconstruído (filtrado) ele também distorceu o sinal, como é possível verificar entre os instantes de 3 a 3,5 segundos.

Além da análise do método de limiarização *Hard-Thresholding* ter sido realizado no sinal de voz com SNR_i igual a 1 dB, também foram feitas análises do método com o sinal apresentando SNR_i igual a 15 dB. Neste caso, os resultados obtidos foram razoavelmente melhores. Conforme a Figura 7 percebe-se que o algoritmo baseado no método *Hard-Thresholding* distorceu menos o sinal reconstruído em relação ao original. Isso pode ser visto entre as faixas de tempo de 3 a 5 segundos.

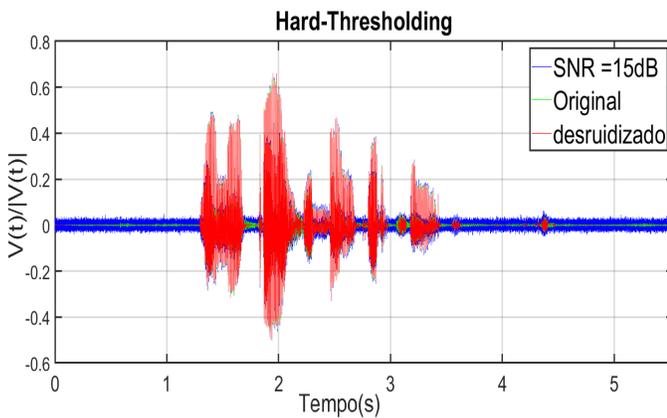


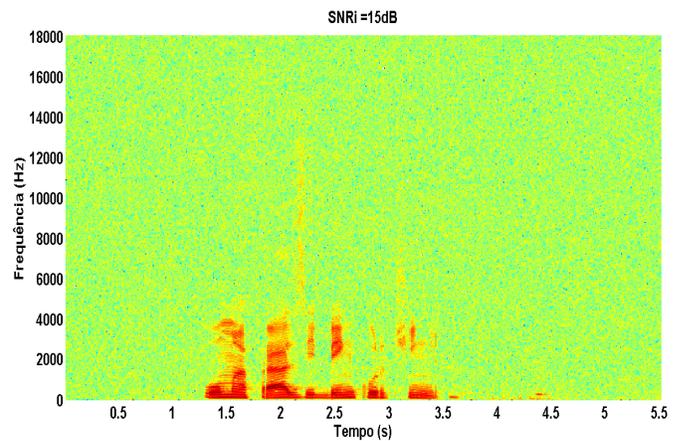
Figura 7. Sinais de fala no domínio tempo sendo o azul o sinal com $SNR_i(\text{dB})=15$, em verde o original e vermelho o reconstruído

Na análise do espectrograma do sinal reconstruído para este caso percebe-se (ver Figura 8(b)) que o sinal foi menos distorcido. Isso pode ser visto nos trechos temporais entre 3 e 3,5 segundos onde as frequências na faixa de 2000 a 6000 Hz também foram ressaltadas muito similarmente ao espectro do sinal original. Os parâmetro SNR_o e d_{IS} ratificam as análises qualitativas com os respectivos valores, 21,3045 dB e 0,0748.

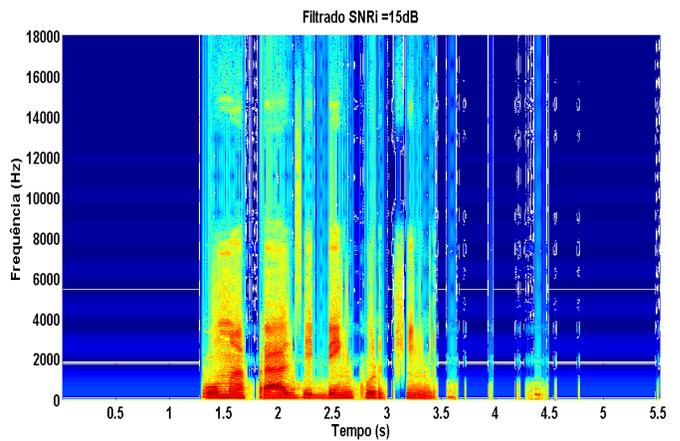
Os valores citados mostram que houve um acréscimo na relação sinal ruído de aproximadamente 4 dB após a filtragem e que ocorreu menor distorção do sinal, visto que quanto mais próxima de zero for a distância de Itakura-Saito menor será a distorção. Neste contexto, o método *Hard-Thresholding* se mostrou melhor para a filtragem em sinais de voz com um nível de ruído mais brando.

Além das análises feitas sobre o método de limiarização *Hard-Thresholding* essas mesmas análises foram realizadas sobre o método *Soft-Thresholding*.

Na primeira análise para método *Soft-Thresholding*, o sinal foi contaminado com ruído branco, obtendo uma relação sinal ruído de entrada (SNR_i) de 1 dB como no método anterior. Neste caso, os resultados obtidos no domínio do tempo estão ilustrados na Figura 9, na qual é possível observar algumas especificidades deste método. É possível observar na figura que houve uma significativa redução da amplitude do sinal filtrado incluindo os trechos em que ocorre a fala (palavra elétrica). Esta redução ocorre devido a supressão dos



(a)



(b)

Figura 8. (a) Espectrograma do sinal com ruído aditivo e $SNR_i(\text{dB}) = 15$; (b) espectrograma do sinal ($SNR_i=15\text{dB}$) reconstruído

coeficientes que possuem valores em módulo maiores que o limiar λ neste método.

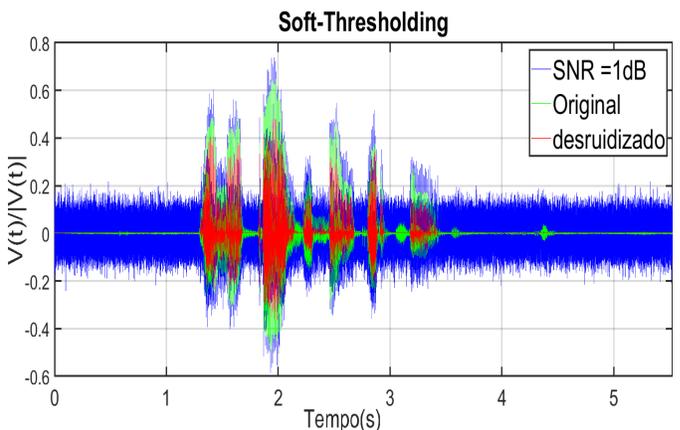


Figura 9. Sinais de fala no domínio tempo sendo o azul o sinal com $SNR_i(\text{dB})=1$, em verde o original e vermelho o filtrado com o método limiar *Soft-Thresholding*

Nesse contexto, uma das especificidades deste método é que

sua aplicação resulta em maior distorção do sinal, visto que ocorrem alterações inclusive onde há a palavra elétrica.

Na análise do espectrograma do sinal reconstruído para este caso é possível observar (ver Figura 10) estas distorções na palavra elétrica entre os trechos temporais de 2 a 2,5 segundos. Neste trecho ocorre uma significativa redução da potência do sinal entre as frequências de 2000 a 4000 Hz. Para este caso os valores de SNR_o e d_{IS} foram respectivamente, 7,3712 dB e 0,2477.

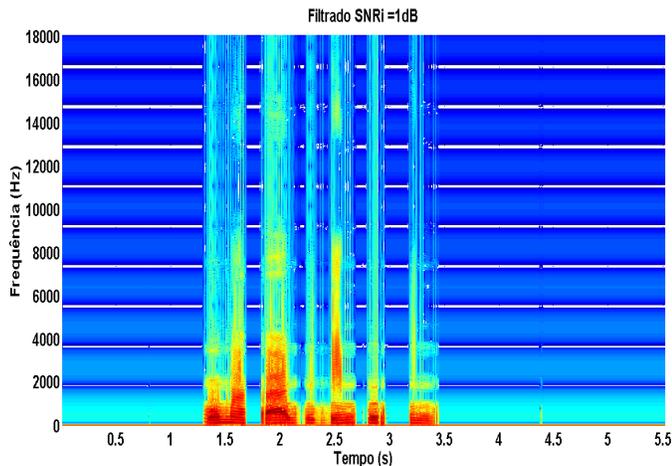


Figura 10. Espectrograma do sinal ($SNR_i=1dB$) após a filtragem com o método limiar Soft-Thresholding

Na segunda análise para o método *Soft-Thresholding*, o sinal foi contaminado com ruído branco, obtendo uma relação sinal ruído de entrada (SNR_i) de 15 dB. Mesmo neste caso ocorreu a distorção do sinal onde há a palavra elétrica como já era previsto. A redução da amplitude do sinal provida por este método pode ocasionar problemas de supressão demasiada da fala em trechos onde a intensidade da palavra elétrica é muito branda como, por exemplo, no trecho próximo a 3 segundos ilustrado na Figura 11. Excluindo esse problema o sinal foi reconstruído com bastante similaridade ao original.

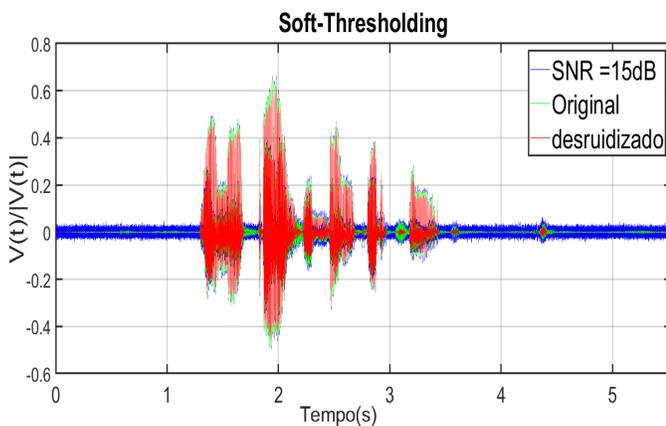


Figura 11. Sinais de fala no domínio tempo sendo o azul o sinal com $SNR_i(dB)=15$, em verde o original e vermelho o filtrado com o método limiar Soft-Thresholding

No espectrograma do sinal reconstruído para este caso é possível observar (ver Figura 12) a redução da intensidade da palavra elétrica justamente no trecho próximo aos 3 segundos. Neste trecho ocorre uma significativa redução da potência do sinal entre as frequências de 2000 a 4000 Hz. Para este caso os valores de SNR_o e d_{IS} foram respectivamente, 16,7 dB e 0,1738.

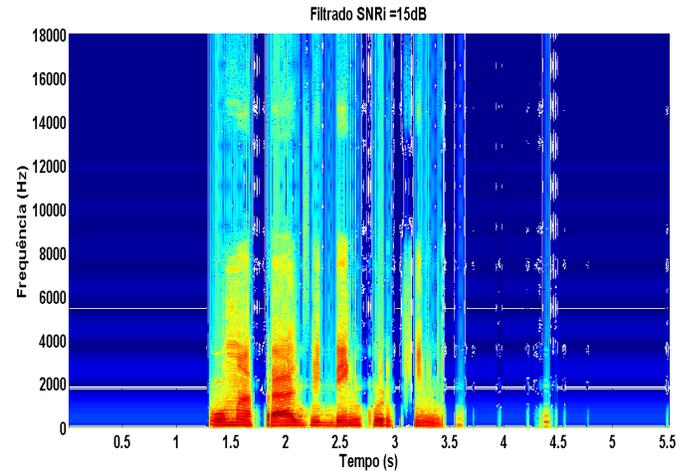


Figura 12. Espectrograma do sinal ($SNR_i=15dB$) após a filtragem com o método limiar Soft-Thresholding

Após as análises qualitativas dos sinais de voz no domínio do tempo e no espectrograma foram realizadas análises sobre as métricas descritas na subseção anterior para várias relações de SNR_i também descritas anteriormente. Os resultados estão descritos nas Tabelas I e II, e ilustrados pelos gráficos das Figuras 13(a) e 13(b), onde as métricas SNR_o e d_{IS} foram ilustradas em função da SNR_i .

Tabela I

Resultados do método *Hard-Thresholding* para diferentes relações de SNR_i

Método	$SNR_i(dB)$	d_{IS}	$SNR_o(dB)$
<i>Hard-Thresholding</i>	1	0,1839	11,0984
	2	0,1775	11,9748
	3	0,1768	12,9563
	4	0,1751	13,4596
	5	0,1762	14,3962
	6	0,1733	14,9325
	8	0,1434	16,2670
	10	0,1691	17,7188
	12	0,1047	19,1444
	15	0,06548	21,3045
	18	0,0396	23,1232
	22	0,0213	25,6839
	28	0,0076	29,5643
	30	0,0065	30,9783
	35	0,0037	34,2066
45	0,0003	39,5480	

Tabela II
Resultados do método *Soft-Thresholding* para diferentes relações de SNR_i

Método	SNR_i (dB)	d_{IS}	SNR_o (dB)
<i>Soft-Thresholding</i>	1	0,2477	7,3712
	2	0,2345	7,9786
	3	0,2286	8,6373
	4	0,2219	9,0972
	5	0,2120	9,9042
	6	0,2115	10,7501
	8	0,2019	11,8563
	10	0,1948	13,2419
	12	0,1851	14,5349
	15	0,1738	16,6979
	18	0,1470	18,6278
	22	0,1084	21,2626
	28	0,0572	25,0408
	30	0,0434	26,5046
	35	0,0250	29,3680
45	0,0052	34,0335	

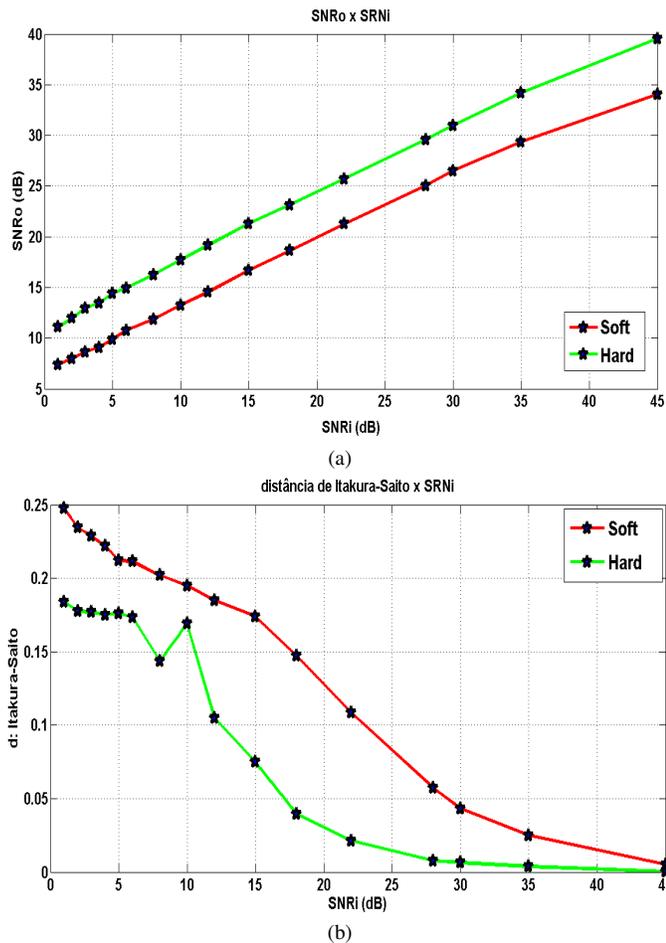


Figura 13. (a) Relação sinal ruído após filtragem SNR_o x relação sinal ruído do sinal com ruído aditivo SNR_i ; (b) distância de Itakura-Saito x relação sinal ruído do sinal com ruído aditivo SNR_i

De modo geral, os resultados mostram que o método de limiarização *Hard-Thresholding* gera menos distorção no sinal que o método *Soft-Thresholding* para diferentes relações de SNR_i . Isto ocorre visto que no método *Soft-Thresholding* os coeficientes que são iguais ou estão acima do limiar λ descrito em (9) também são modificados, ou seja, mesmo onde há a palavra elétrica o sinal é filtrado. Tal fato faz com que a amplitude do sinal seja suprimida como visto nas Figuras 9 e 11. Além disso, o método *Hard-Thresholding* mostrou-se superior em relação a sua capacidade de filtrar ruído aditivo produzindo SNR_o maiores que o outro método. Os resultados obtidos concordam com resultados adquiridos em [6].

IV. CONCLUSÃO

Este artigo mostrou o estudo sobre dois métodos para a redução de ruído em sinais de voz baseados na transformada Wavelet Discreta: *Hard-Thresholding* e *Soft-Thresholding*. Para a comprovação da eficiência destas técnicas utilizou-se as métricas de relação sinal ruído após a filtragem SNR_o e distância de Itakura-Saito d_{IS} . Após os testes realizados, concluiu-se através das métricas que a metodologia *Hard-Thresholding* apresentou melhor desempenho tanto em relação a redução do ruído quanto em relação a distorção. Entretanto é importante salientar que a escolha da Wavelet utilizada para o desenvolvimento do filtro pode influenciar o desempenho de ambos os métodos. Neste contexto, em trabalhos futuros, distintas funções Wavelet como a db5, a symlet e outras devem ser utilizadas na concepção dos filtros para que se realize testes mais conclusivos sobre o desempenho geral destes métodos.

REFERÊNCIAS

- [1] S. Mallat, "A wavelet tour of signal processing," Elsevier/Academic Press, 2009.
- [2] S. K. Prasad, S. S. Natrajan, and S. Kalaivani, "Efficiency analysis of noise reduction algorithms: Analysis of the best algorithm of noise reduction from a set of algorithms," in *2017 International Conference on Inventive Computing and Informatics (ICICI)*. IEEE, 2017, pp. 1137–1140.
- [3] M. Saric, L. Bilicic, and H. Dujmic, "White noise reduction of audio signal using wavelets transform with modified universal threshold," *University of Split, R. Boskovic b. b HR*, vol. 21000, 2005.
- [4] R. Aggarwal, J. K. Singh, V. K. Gupta, S. Rathore, M. Tiwari, and A. Khare, "Noise reduction of speech signal using wavelet transform with modified universal threshold," *International Journal of Computer Applications*, vol. 20, no. 5, pp. 14–19, 2011.
- [5] M. A. Q. Duarte *et al.*, "Redução de ruído em sinais de voz no domínio wavelet," 2005.
- [6] L. Jing-yi, L. Hong, Y. Dong, and Z. Yan-sheng, "A new wavelet threshold function and denoising application," *Mathematical Problems in Engineering*, vol. 2016, 2016.
- [7] S. V. Vaseghi, "Advanced digital signal processing and noise reduction," John Wiley & Sons, 2008.
- [8] M. Ishwarya, "An improved online tamil character recognition using neural networks," *International Journal of Advanced Science and Technology*, vol. 42, pp. 1–10, 2012.
- [9] M. Berouti, R. Schwartz, and J. Makhoul, "Enhancement of speech corrupted by acoustic noise," in *ICASSP'79. IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4. IEEE, 1979, pp. 208–211.
- [10] G. Strang and T. Nguyen, "Wavelets and filter banks," SIAM, 1996.
- [11] R. Merry and M. Steinbuch, "Wavelet theory and applications," *literature study, Eindhoven university of technology, Department of mechanical engineering, Control systems technology group*, 2005.

- [12] E. Hernández and G. Weiss, "A first course on wavelets," CRC press, 1996.
- [13] C. Taswell, "The what, how, and why of wavelet shrinkage denoising," *Computing in science & engineering*, vol. 2, no. 3, pp. 12–19, 2000.
- [14] S. Tsai, "Wavelet transform and denoising," Master's Thesis, URN: etd-12062002-152858, Chapter 4, 2002.
- [15] C. Gargour, M. Gabrea, V. Ramachandran, and J.-M. Lina, "A short introduction to wavelets and their applications," *IEEE circuits and systems magazine*, vol. 9, no. 2, pp. 57–68, 2009.
- [16] A. Bruce, H.-Y. Gao, and A. Bruce, "Applied wavelet analysis with S-plus," Springer New York, 1996, vol. 168.
- [17] S. G. Mihov, R. M. Ivanov, and A. N. Popov, "Denoising speech signals by wavelet transform," *Annual Journal Of Electronics*, no. 6, pp. 2–5, 2009.
- [18] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *biometrika*, vol. 81, no. 3, pp. 425–455, 1994.
- [19] D. L. Donoho, "De-noising by soft-thresholding," *IEEE transactions on information theory*, vol. 41, no. 3, pp. 613–627, 1995.
- [20] I. Codello and W. Kuniszyk-Józkowiak, "Wavelet analysis of speech signal," *Annales Universitatis Mariae Curie-Sklodowska, section AI-Informatica*, vol. 6, no. 1, 2007.
- [21] M. Togami, "Multi-channel itakura saito distance minimization with deep neural network," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 536–540.