

**Uso de equações de diferenças na obtenção de filtros para redução  
de ruído em sinais de voz no domínio wavelet**

**Caio Cesar Enside de Abreu**

**Orientador:** Prof. Dr. Francisco Villarreal Alvarado

**Coorientador:** Prof. Dr. Marco Aparecido Queiroz Duarte

Dissertação apresentada ao PPGEE,  
Faculdade de Engenharia - UNESP –  
Campus de Ilha Solteira, para obtenção  
do título de Mestre em Engenharia  
Elétrica. Área de Conhecimento:  
Automação.

Ilha Solteira – SP

Fevereiro/2013

FICHA CATALOGRÁFICA

Desenvolvido pelo Serviço Técnico de Biblioteca e Documentação

A162u Abreu, Caio Cesar Enside de .  
    Uso de equações de diferenças na obtenção de filtros para redução de ruído em sinais de voz no domínio wavelet / Caio Cesar Enside de Abreu. -- Ilha Solteira: [s.n.], 2013  
    96 f. : il.

Dissertação (mestrado) - Universidade Estadual Paulista. Faculdade de Engenharia de Ilha Solteira. Área de conhecimento: Automação, 2013

Orientador: Francisco Villarreal Alvarado  
Co-orientador: Marco Aparecido Queiroz Duarte  
Inclui bibliografia

1. Wavelets. 2. Diferenças finitas. 3. Redução de ruído não-limiar.

**CERTIFICADO DE APROVAÇÃO**


**TÍTULO:** Uso de equações de diferenças na obtenção de filtros para redução de ruído em sinais de voz no domínio wavelet

**AUTOR:** CAIO CESAR ENSIDE DE ABREU


**ORIENTADOR:** Prof. Dr. FRANCISCO VILLARREAL ALVARADO

**CO-ORIENTADOR:** Prof. Dr. MARCO APARECIDO QUEIROZ DUARTE

Aprovado como parte das exigências para obtenção do Título de Mestre em Engenharia Elétrica ,  
Área: AUTOMAÇÃO, pela Comissão Examinadora:

  
Prof. Dr. FRANCISCO VILLARREAL ALVARADO  
Departamento de Matemática / Faculdade de Engenharia de Ilha Solteira

  
Prof. Dr. JOZUE VIEIRA FILHO  
Departamento de Engenharia Elétrica / Faculdade de Engenharia de Ilha Solteira

  
Prof. Dr. LEANDRO DE CAMPOS TEIXEIRA GOMES  
Diretoria de Tecnologia de Serviços / Fundação Centro de Pesquisa e Desenvolvimento em  
Telecomunicações - CPqD

Data da realização: 20 de fevereiro de 2013.

Dedico esta dissertação aos meus pais,  
Rosângela e Haroldo.

## AGRADECIMENTOS

---

- Primeiramente agradeço a Deus, por abrir portas e indicar o caminho certo a seguir, e também por conceder a saúde e o discernimento necessários para alcançar meus objetivos pessoais.
- Ao meu orientador Professor Dr. Francisco Villarreal Alvarado, pela confiança depositada ao abrir as portas do mestrado para mim, sempre acreditando em minha capacidade, e é claro, pelas correções e sugestões, nunca se esquecendo do rigor matemático.
- Ao Professor Dr. Marco Aparecido Queiroz Duarte, meu coorientador, pelas ideias, sugestões e correções, as quais foram fundamentais para o desenvolvimento desta pesquisa.
- Ao amigo que se apresentou durante o período das disciplinas, Professor Dr. Jozué Vieira Filho, pelos conselhos, sugestões, e a atenção concedida sempre atendendo às minhas visitas em sua sala, mesmo quando o tempo lhe era curto.
- À minha mãe Rosângela Enside de Abreu, que sempre me apoiou e nunca mediu esforços para que eu pudesse alcançar meus objetivos, sempre lembrando de meu nome e da minha pesquisa em suas preces.
- À minha namorada Grazielle Toledo, pelo conforto, estímulo e compreensão nos momentos difíceis.
- A minha companheira de apartamento e irmã Tatiane Aparecida Enside de Abreu. Obrigado por me aturar!
- Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro.
- A todos os meus familiares e amigos.

## RESUMO

---

O método proposto neste trabalho tem por objetivo superar algumas deficiências que os métodos que utilizam limiar apresentam quando se almeja a redução de ruído em sinais de voz. Dentre elas, destaca-se a dificuldade no processamento de sinais contaminados com ruído colorido. Com base neste fato, desenvolveu-se um método eficiente para o processamento de vários tipos de ruído colorido, que é o tipo de ruído presente em situações reais. A metodologia de redução de ruído proposta consiste na estimação do sinal de saída a partir do original, no domínio wavelet, sem o uso de limiar. O sinal é estimado por equações de diferenças finitas. Para a estimação do ruído é usado um operador, que é aplicado em cada uma das equações de diferenças. Uma combinação polinomial é realizada de maneira a condensar todas as equações de diferenças em uma única função de transferência. Esta por sua vez sofrerá um ajuste sigmoidal visando uma melhor adequação entre as amplitudes dos sinais originais e processados. Porém, antes disto, propõe-se uma pré-filtragem realizada por um operador de pré-processamento. O filtro é obtido no último passo, quando é feito o ajuste sigmoidal. A principal mudança, em relação às metodologias anteriores, está na construção de um filtro que atue no sinal levando em consideração informações contidas em cada coeficiente ou em uma pequena vizinhança do mesmo. Com isso, não se faz necessário o uso de um único valor (valor do limiar) como referência para processar todos os coeficientes de uma faixa de frequência do sinal. Comparando o método proposto com o método desenvolvido por Soares et al. (2011), constatou-se a obtenção de melhores resultados.

**Palavras-chave:** Wavelets. Diferenças finitas. Redução de ruído não-limiar.

## ABSTRACT

---

The method proposed in this paper aims to overcome some deficiencies of the methods which use the threshold feature when the objective is to reduce noise in speech signals. Among them, it was highlight the difficult in processing signals contaminated with colored noise. Based on this fact, it was developed a method that is efficient for the processing of various types of colored noise, which is the predominant noise in real situations. The proposed noise reduction method consists in estimating the output signal from the original, in wavelet domain, without using a threshold value. The signal is estimated by finite difference equations. For the noise estimation an operator, which is applied to each difference equation. A polynomial combination is performed in order to condense all difference equations in a single transfer function, which, in turn, will suffer a sigmoidal adjustment, seeking a better adequacy between the amplitudes of the original and processed signals. However, before this, a pre-filtering performed by an operator of pre-processing is proposed. The filter is obtained in the last step, which is the sigmoidal adjustment. The main change is in the construction of a filter that act on the signal taking into account the information contained in each coefficient or in a small neighborhood of it. Thus, it is not necessary to use a single value (threshold value) as reference to process all the coefficients of a frequency band of the signal. The proposed method was compared the one developed by Soares et al. (2011), and better results were obtained.

**Keywords:** Wavelets. Finite differences. Nonthreshold noise reduction.

## LISTA DE ILUSTRAÇÕES

---

Figura 1 – Esquema para a recuperação exata de um sinal de voz contaminado por ruído aditivo.....	21
Figura 2 – Gráfico de uma wavelet.....	26
Figura 3 – Caixas de Heisenberg para a WFT (a) e para a CWT (b).....	30
Figura 4 – Árvore de decomposição wavelet (banco de filtros de análise).....	33
Figura 5 – Árvore de reconstrução wavelet (banco de filtros de síntese).....	33
Figura 6 - Trecho de um sinal de voz e sua transformada wavelet.....	36
Figura 7 - Diagrama do Método Proposto por Soares (2009).....	41
Figura 8 - Diagrama do método proposto.....	43
Figura 9 - Sinal limpo $s(n)$ .....	47
Figura 10 - Sinal contaminado com ruído colorido $y(n)$ .....	47
Figura 11 - Trecho de voz do sinal $Y[n]$ , no domínio wavelet, obtido a partir da contaminação do sinal limpo $s[n]$ por ruído colorido.....	48
Figura 12 - Curvas das diferenças centradas de primeira a quarta ordem, respectivamente, quando aplicadas ao segmento visualizado na Figura 11.....	48
Figura 13 - Curvas das diferenças avançadas de primeira a quarta ordem, respectivamente, quando aplicadas ao segmento visualizado na Figura 11.....	49
Figura 14 - Sinal processado $\hat{y}(n)$ a partir do sinal original limpo $s(n)$ , usando a função $E_{cp}$ e a transformada wavelet inversa.....	52

Figura 15 - Sinal processado $\hat{y}(n)$ a partir do sinal $y(n)$ contaminado por ruído colorido, usando a função $E_{cp}$ e a transformada wavelet inversa.....	52
Figura 16 - Esquema de frequências de um banco de filtros.....	53
Figura 17 - Gráfico do operador $Dif$ .....	54
Figura 18 - Sinal processado $\hat{y}(n)$ a partir do sinal original limpo $x(n)$ , usando o operador $Dif$ e a transformada wavelet inversa (a); Zoom na região demarcada (b).....	55
Figura 19 - Curva gerada pela combinação das duas funções sigmóides na equação (34).....	56
Figura 20 - Sinal de voz limpo $s(n)$ e sinal processado $\hat{y}(n)$ com ruído colorido, usando o método proposto e a transformada wavelet inversa.....	57
Figura 21 - Sinal processado $\hat{y}(n)$ com ruído colorido, usando o método proposto e a transformada wavelet inversa.....	58
Figura 22 - Resultados obtidos por meio de avaliações objetivas.....	62
Figura 23 - Forma de onda do Sinal F limpo (a); contaminado com ruído de carro (b); contaminado com ruído em restaurante (c); contaminado com ruído de metrô (d).....	63
Figura 24 - Forma de onda dos sinais processados, a partir do Sinal F limpo, contaminado com: ruído de carro (a), ruído em restaurante (b), ruído de metrô (c).....	65
Figura 25 - Forma de onda do Sinal B limpo (a); contaminado com ruído de carro (b); contaminado com ruído em restaurante (c); contaminado com ruído de metrô (d).....	66
Figura 26 - Forma de onda dos sinais processados, a partir do Sinal B limpo, contaminado com: ruído de carro (a), ruído em restaurante (b), ruído de metrô (c) .....	68
Figura 27 - Comparações dos valores de SNR obtidos pelo método proposto e pela metodologia de Soares et al. (2011).....	72
Figura 28 - Comparações das notas PESQ obtidas pelo método proposto e pelo método de Soares et al. (2011).....	73
Figura 29 - Comparações das correlações obtidas entre os sinais originais e processados pelo método proposto, e pelo método proposto em Soares et al. (2011).....	74
Figura 30 - Sinal de voz limpo $x(n)$ .....	85
Figura 31 - Sinal de voz contaminado $y(n)$ (10 dB).....	85
Figura 32 - Sinal de voz processado $\hat{y}(n)$ , usando o método proposto a partir do sinal de voz limpo $s(n)$ contaminado com ruído de carro (10 dB).....	86
Figura 33 - Sinal de voz limpo $s(n)$ (azul) e sinal processado $\hat{y}(n)$ (vermelho), usando o método proposto a partir do sinal de voz limpo contaminado com ruído de carro (10 dB).....	86

Figura 34 - Sinal de voz contaminado $y(n)$ (5 dB).....	87
Figura 35 - Sinal de voz processado $\hat{y}(n)$ , usando o método proposto a partir do sinal de voz limpo $s(n)$ contaminado com ruído de carro (5 dB).....	88
Figura 36 - Sinal de voz limpo $x(n)$ (azul) e sinal processado $\hat{y}(n)$ (vermelho) a partir do sinal de voz limpo contaminado com ruído de carro (5 dB).....	88
Figura 37 - Gráfico da equação (43) para diferentes valores de $c$ .....	91
Figura 38 - Sinal de voz limpo $x(n)$ .....	93
Figura 39 - Sinal de voz ruidoso $y(n)$ , obtido por contaminação do sinal original $x(n)$ com ruído em restaurante (10 dB).....	94
Figura 40 - Sinal de voz limpo $s(n)$ (azul) e sinal processado $\hat{y}(n)$ (vermelho) usando o método proposto a partir do sinal de voz limpo contaminado com ruído em restaurante.....	94
Figura 41 - Sinal de voz limpo $s(n)$ (azul) e sinal processado $\hat{y}(n)$ (vermelho) usando o método proposto, sem o operador <i>Dif</i> , a partir do sinal de voz limpo contaminado com ruído em restaurante.....	95

## **LISTA DE TABELAS**

---

Tabela 1 - MOS e a qualidade do discurso correspondente.....	22
Tabela 2 - DMOS e o correspondente nível de degradação.....	22
Tabela 3 - Análise da variação dos valores de SNR dos sinais processados, quando comparados ao valor da média das SNR dos sinais limpos.....	62
Tabela 4 - Resultados obtidos para todos os sinais utilizados nas simulações, levando em consideração os valores de SNR para os sinais limpo e processado.....	70
Tabela 5 - PESQ dos sinais ruidosos.....	70
Tabela 6 - PESQ dos sinais processados.....	71
Tabela 7 - Avaliações objetivas para o sinal $s(n)$ em particular.....	89
Tabela 8 - Avaliações objetivas para o sinal processado $y(n)$ , em particular.....	95

## LISTA DE SÍMBOLOS

---

$\mathbb{Z}$	Conjunto dos números inteiros;
$\mathbb{R}$	Conjunto dos números reais;
$\mathbb{C}$	Conjunto dos números complexos;
$\psi(t)$	Função wavelet;
$\hat{\Psi}(\omega)$	Transformada de Fourier da função wavelet;
$\varphi(t)$	Função escala;
$\langle f, g \rangle$	Produto interno entre $f$ e $g$ ;
$C^k$	Classe das funções com derivadas contínuas até a ordem $k$ ;
$a$	Parâmetro de escala da função wavelet;
$b$	Parâmetro de translação da função wavelet;
$\text{supp}(f)$	Suporte da função $f$ ;
$L^2(\mathbb{R})$	Espaço das funções mensuráveis de Lebesgue de quadrado integrável;
$\sigma_e^2$	Variância do erro;
$\sigma_x^2$	Variância do sinal original;
dB	Decibel;
log	Logaritmo;
$ x $	Valor absoluto de $x$ ;
$\ x\ $	Norma de $x$ ;
$W_f(a, b)$	Transformada wavelet contínua de $f$ ;
$\lambda$	Limiar usado para filtragem do sinal;
$f(t)$	Função no domínio do tempo;

$\hat{f}$	Transformada de Fourier da função $f$ ;
$\bar{f}$	Conjugado complexo da função $f$ ;
$H(z)$	Filtro passa-baixas associado a uma wavelet;
$G(z)$	Filtro Passa-altas associado a uma wavelet;
WFT	Transformada de Fourier janelada;
$(\downarrow 2)$	Operador de dizimação de ordem 2 ou <i>downsampling</i> ;
$(\uparrow 2)$	Operador de inserção de zeros de ordem 2 ou <i>upsampling</i> ;
$c_i[n]$	Coefficientes de aproximação de nível $i$ da DWT;
$d_i[n]$	Coefficientes de detalhes de nível $i$ da DWT;
$L$	Número máximo de faixas de frequências na decomposição pela DWT;
$T$	Matriz $T$ ;
$\sigma$	Parâmetro de inclinação da sigmoide;
$s[n]$	Sinal de voz limpo no domínio do tempo;
$w[n]$	Ruído;
$y[n]$	Sinal de voz contaminado com ruído colorido;
$S[n]$	Sinal de voz no domínio wavelet;
$W[n]$	Ruído no domínio wavelet;
$Y[n]$	Sinal contaminado no domínio wavelet;
$\hat{Y}[n]$	Sinal processado no domínio wavelet;
$\hat{y}[n]$	Sinal processado no domínio do tempo;
$c_i^j$	Coefficientes das diferenças finitas centradas de ordem $j$ na posição $i$ ;
$a_i^j$	Coefficientes das diferenças finitas avançadas de ordem $j$ na posição $i$ ;
$C^j$	Vetor de diferenças finitas centradas de ordem $j$ ;
$A^j$	Vetor de diferenças finitas avançadas de ordem $j$ ;
$d_i$	Coefficientes de um vetor qualquer de diferenças finitas;
$D$	Vetor qualquer de diferenças finitas;
$g(D)$	Função que aplica o operador SNRP no vetor $D$ ;
$\alpha$	Parâmetro da função $g$ , $\alpha \in ]0,1[$ ;
$E_{cp}$	Função que realiza a combinação polinomial;
$e_i$	Coefficientes da função $E_{cp}$ ;
$H[n]$	Filtro construído a partir do método proposto;
$h_i$	Coefficientes do filtro $H[n]$ ;

$\tau$  Desvio padrão da janela em processamento;

$\beta$  Valor baseado na potência do ruído.

## LISTA DE ABREVIATURAS

---

AN – Amplitude Numérica;

dBN – Função wavelet de Daubechies de ordem N;

TW – Transformada wavelet;

TWI – Transformada wavelet inversa;

CWT – Transformada wavelet contínua;

DWT – Transformada wavelet discreta;

IDWT – Transformada wavelet discreta inversa;

SNR – Relação Sinal/Ruído;

*TRH – Threshold;*

MATLAB – *Matrix Laboratory;*

PESQ - *Perceptual Evaluation of Speech Quality;*

cp – Combinação polinomial;

ITU-T – *International Telecommunication Union - Telecommunication Standardization Sector;*

MOS – *Mean Opinion Score;*

DMOS - *Degradation Mean Opinion Score;*

SNRP – Relação Sinal/Ruído *a Priori;*

OMS – Operador Média Simples;

SNRPR – Relação Sinal/Ruído *a Posteriori.*

*Dif* – Operador de pré-filtragem.

## SUMÁRIO

---

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>15</b>
1.1	Considerações preliminares e motivação.....	15
1.2	Organização do texto.....	17
1.3	Contribuições do trabalho.....	18
<b>2</b>	<b>CONSIDERAÇÕES SOBRE REDUÇÃO DE RUÍDO EM SINAIS DE VOZ.....</b>	<b>19</b>
2.1	Métodos de redução de ruído.....	20
2.2	Medidas subjetivas de qualidade.....	21
2.3	Medidas objetivas de qualidade.....	22
<b>3</b>	<b>ANÁLISE WAVELET.....</b>	<b>24</b>
3.1	Funções Wavelets.....	25
3.2	Transformada Wavelet Contínua.....	28
3.3	Transformada Wavelet Contínua Inversa.....	30
3.4	Transformada Wavelet Discreta.....	31
3.5	Transformada Wavelet Discreta Inversa.....	31
3.6	Interpretação da DWT como um banco de filtros.....	32
<b>4</b>	<b>MÉTODOS DE REDUÇÃO DE RUÍDO EM SINAIS DE VOZ NO DOMÍNIO WAVELET.....</b>	<b>35</b>
4.1	Redução de ruído em sinais de voz por métodos que utilizam limiares.....	37

4.2	Redução de ruído em sinais de voz por métodos não-limiarees.....	39
4.2.1	<i>Análise de um método não-limiar</i> .....	39
<b>5</b>	<b>METODOLOGIA PROPOSTA</b> .....	<b>42</b>
5.1	Objetivos do método proposto.....	43
5.2	O uso das equações de diferenças para estimar o sinal de saída.....	45
5.3	Estimação do ruído.....	49
5.4	A combinação polinomial para estimar o sinal de saída.....	50
5.5	O operador <i>Dif</i> .....	53
5.6	Obtenção do filtro de redução de ruído.....	56
<b>6</b>	<b>ANÁLISE E DISCUSSÃO DOS RESULTADOS</b> .....	<b>60</b>
6.1	Testes computacionais.....	61
6.2	Comparações a partir do método proposto por Soares et al. (2011).....	72
<b>7</b>	<b>CONSIDERAÇÕES FINAIS</b> .....	<b>75</b>
	<b>REFERÊNCIAS</b> .....	<b>78</b>
	<b>APÊNDICE A - DEMONSTRAÇÃO DE TEOREMA</b> .....	<b>83</b>
	<b>APÊNDICE B - ANÁLISE VISUAL DO PROCESSAMENTO REALIZADO PELO MÉTODO PROPOSTO</b> .....	<b>84</b>
	<b>APÊNDICE C - SOBRE A EXPRESSÃO UTILIZADA PARA O CÁLCULO DO OPERADOR <i>Dif</i></b> .....	<b>90</b>
	<b>APÊNDICE D - ANÁLISE DA INFLUÊNCIA DO OPERADOR <i>Dif</i> NA QUALIDADE DO SINAL PROCESSADO</b> .....	<b>93</b>

---

# 1 INTRODUÇÃO

## 1.1 Considerações Preliminares e Motivação

Quando se deseja estudar cientificamente um problema do universo físico, isto é, deixar de lado a visão intuitiva do senso comum e analisar os fatos com um maior teor de criticidade e abrangência, deve-se usar uma linguagem adequada que facilite e racionalize o pensamento. A modelagem matemática é a ferramenta adequada. Segundo D'Ambrosio (2003), a modelagem matemática é um processo valioso para encarar situações reais, culminando com a solução efetiva de um problema real e não com a simples resolução de um problema artificial. No universo do processamento digital de sinais, um fenômeno físico é modelado por meio de uma função, que é chamada de Sinal (GOMES; VELHO; GOLDSTEIN, 1997).

Existem, porém, variados tipos de sinais do universo físico que não podem ser modelados. Isto se deve ao fato destes sinais possuírem um alto teor de complexidade. O que reforça ainda mais a necessidade de compreendê-los. Sua análise é feita a partir de outras ferramentas matemáticas, capazes de evidenciar suas principais propriedades. Dentre elas, destacam-se as transformadas integrais. As transformadas mais utilizadas no âmbito do processamento de sinais são a transformada de Fourier e a transformada wavelet (SANCHEZ, 2008; GOMES; VELHO; GOLDSTEIN, 1997).

Um sinal de voz até pode ser modelado, porém, devido à grande variação de tons e timbres, esta modelagem se torna muito complexa. Processar sinais de voz é muito importante, tendo em vista a vasta gama de aplicações como codificação, redução de ruído, reconhecimento automático de fala, dentre outras (SOARES, 2009).

A redução de ruído em sinais de voz vem sendo amplamente explorada. Deve-se isso ao fato de que, muitas vezes, o sinal transmitido pelos meios de telecomunicações, por exemplo, pode conter algum tipo de ruído (DELLER; PROAKIS; HANSEN, 1993). Considera-se ruído qualquer elemento, ou sinal, que tem a capacidade de reduzir a inteligibilidade de uma informação de som (NOCETI FILHO, 2004). Um sinal de voz pode ser contaminado por ruído, também, ao ser gravado, captado por um microfone ou devido ao meio de reprodução do mesmo. Em todos os casos, o ruído causa um desconforto ao ouvinte, provocando a perda da inteligibilidade do sinal.

Quando o assunto é redução de ruído em sinais de voz, o objetivo é melhorar a qualidade do sinal. Existem vários métodos para a redução de ruído. Alguns usam a transformada de Fourier (VIEIRA FILHO, 1996) e outros, a transformada wavelet (DUARTE, 2005; DUARTE, 2010; SOARES et al., 2008; SOARES, et al., 2011). Dentre os que utilizam a transformada wavelet, alguns são baseados em redução por Limiar, ou seja, eliminam ou atenuam os coeficientes do sinal transformado cujos valores absolutos estão abaixo de um valor estipulado a priori (SOARES et al., 2007), e outros considerados Não-Limiar (SOARES, 2009). Os métodos que usam limiar são os mais utilizados na literatura, porém, possuem grandes limitações que aos poucos estão sendo superadas por métodos que não fazem uso do limiar (SOARES et al., 2008; SOARES et al., 2011).

A metodologia de redução de ruído proposta neste trabalho consiste na estimação do sinal de saída a partir de um sinal ruidoso, no domínio wavelet, sem o uso do limiar. O sinal é estimado pelas equações de diferenças usadas na resolução numérica de derivadas e equações diferenciais, descritas no capítulo 5. Para a estimação do ruído é usado o operador SNRP proposto por Soares et al. (2011), que é aplicado em cada uma das equações de diferenças. Uma combinação polinomial é realizada de maneira a condensar todas as equações de diferenças em uma única função de transferência; esta, por sua vez, sofrerá um ajuste sigmoidal visando uma melhor adequação entre as amplitudes dos sinais originais e processados. Porém, antes disto, propõe-se uma pré-filtragem, realizada por um operador denominado *Dif*, cuja definição e justificativa encontram-se nos Apêndices C e D. O filtro é obtido no último passo, que consiste num ajuste sigmoidal.

A principal motivação para o desenvolvimento deste trabalho é a possibilidade de superar algumas deficiências que os métodos que utilizam limiares apresentam. A principal mudança está na construção de um filtro que atua no sinal, levando em consideração informações contidas em cada coeficiente em processamento, ou em uma vizinhança do mesmo. O tamanho desta vizinhança varia de acordo com a ordem da equação de diferenças usada na construção do filtro. Com isso, não se faz necessário o uso de um único valor (valor do limiar) como referência para processar todos os coeficientes de uma faixa de frequência do sinal.

## 1.2 Organização do Texto.

Inicialmente, apresenta-se a principal ideia por trás dos métodos de redução de ruído em sinais de voz, assim como as principais medidas de qualidade utilizadas pela literatura especializada para a avaliação do desempenho dos mesmos; este assunto é abordado no capítulo 2. No capítulo 3, apresenta-se um estudo mais detalhado sobre a teoria wavelet, dando ênfase aos principais conceitos matemáticos que serão considerados neste trabalho. O capítulo 4 trata de uma análise dos métodos mais recentes de redução de ruído em sinais de voz baseados em wavelet, destacando as vantagens e desvantagens desses métodos. No capítulo 5 é apresentada e discutida a metodologia proposta para redução de ruído em sinais de voz, no domínio wavelet, sem o uso de limiar. A apresentação dos resultados obtidos e a avaliação do método proposto serão tratadas no capítulo 6. Por fim, no capítulo 7 são apresentadas as conclusões e as considerações finais do presente trabalho.

Ao final deste trabalho, apresentam-se quatro apêndices de suma relevância para a compreensão de como o método proposto atua em um sinal de voz. Estes apêndices tratam dos seguintes assuntos:

**Apêndice A** – Demonstração do teorema onde se afirma que combinações lineares de uma base do espaço  $L^2(\mathbb{R})$  podem ser utilizadas para se obter aproximações de qualquer função do espaço gerado por tal base.

**Apêndice B** – Análise visual do processamento realizado pelo método proposto. A análise é realizada a partir da forma de onda de um sinal em particular, escolhido arbitrariamente.

**Apêndice C** – Sobre a expressão utilizada para o cálculo do operador *Dif*.

**Apêndice D** – Faz-se uma análise da influência do operador *Dif* no sinal processado. Tal análise é realizada considerando as formas de onda dos sinais original e processado, quando se utiliza o operador *Dif* e quando não se utiliza o mesmo. Além da análise visual, mostram-

se também as diferenças comprovadas pelas análises (avaliações) objetivas referentes ao uso deste operador.

### **1.3 Contribuições do Trabalho.**

As principais contribuições do Trabalho são:

- A apresentação de uma metodologia de redução de ruído no domínio wavelet que dispensa o uso de limiares.
- A construção de um filtro que atua no sinal levando em consideração informações contidas em cada coeficiente do sinal ou em uma vizinhança do mesmo.
- Introdução de um processo de pré-filtragem, que permite uma forte redução de ruído no sinal ruidoso minimizando perdas nos coeficientes referentes à voz.

---

## 2 CONSIDERAÇÕES SOBRE REDUÇÃO DE RUÍDO EM SINAIS DE VOZ

Em muitos sistemas de comunicação, a presença de ruído de fundo pode causar perda da qualidade ou inteligibilidade do sinal de voz (DELLER; PROAKIS; HANSEN, 1993). Ruído é qualquer sinal que tenha a capacidade de reduzir a inteligibilidade de uma informação de som (NOCETI FILHO, 2004).

Em algumas situações práticas, como, por exemplo, em um sistema de comunicação móvel, observa-se o sinal de voz juntamente com o ruído aleatório proveniente do canal de comunicação e do ambiente do locutor. Como normalmente trata-se o sinal no lado do receptor, não é possível um conhecimento detalhado do perfil do ruído, mas apenas do sinal ruidoso (ANTUNES, 2006). Neste trabalho, assume-se que os sinais ruidosos em questão possam ser descritos pelo modelo aditivo:

$$y[n] = s[n] + w[n], \tag{1}$$

em que  $s[n]$  é o sinal de voz,  $w[n]$  é o ruído e  $n$  representa o índice de tempo discreto.

Quando o assunto é redução de ruído em sinais de voz, o objetivo é melhorar a qualidade do sinal. Existem vários métodos para a redução de ruído. Alguns usando a transformada de Fourier (VIEIRA FILHO, 1996) e outros no domínio Wavelet. Dentre os que utilizam a transformada wavelet, a maioria é baseada em métodos que utilizam limiares

para a filtragem do sinal (DUARTE, 2005), porém métodos não-limiar têm surgido com resultados satisfatórios (SOARES, 2009).

## 2.1 Métodos de redução de ruído

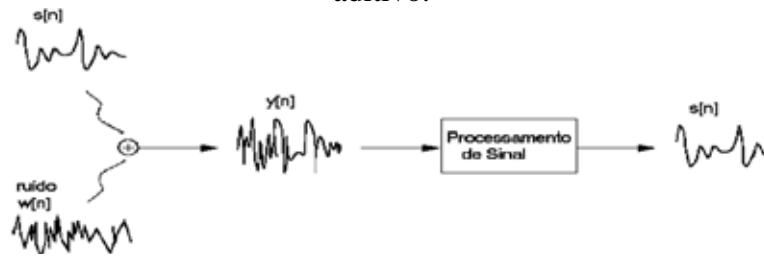
A redução de ruído em sinais de qualquer dimensão é uma área relevante da engenharia. Em situações práticas, o ambiente no qual se faz a aquisição do sinal, dificilmente está livre da presença de ruído. Neste sentido, existem vários métodos para eliminar ou atenuar ruídos (LOUIS; MAAB; RIEDER, 1998).

Em sua essência, os métodos de processamento de sinais de voz podem ser divididos em dois blocos: os que fazem uso de um limiar e os que não usam limiares. Os métodos que usam limiar eliminam ou atenuam os coeficientes do sinal transformado, cujos valores absolutos estejam abaixo de um valor estipulado a priori (SOARES et al., 2007). Alguns métodos de redução de ruído em sinais de voz baseados em limiar apresentam bons resultados (por exemplo, Seok and Bae (1997); Sheikhzadeh and Abutalebi (2001) e Lallouani, Gabrea; Gargour (2004)). Porém, estes métodos podem confundir ruído com voz, pelo fato de que alguns coeficientes de voz possuem valores absolutos abaixo do limiar usado (DUARTE, 2005). Outro problema que ocorre é que, quando o ruído é adicionado, muitos coeficientes redundantes assumem valores acima do limiar estipulado. Deste modo, na aplicação do limiar, muitos coeficientes que deveriam ser eliminados ou atenuados acabam sendo preservados, causando pouca eficiência na redução de ruído (SOARES et al., 2008). Além dos problemas já citados acerca dos métodos que utilizam limiares, outro importuno seria o fato de que os mesmos possuem funções de transferência descontínuas.

Para Soares et al. (2008), uma forma de sanar os problemas apresentados pelos métodos que utilizam limiares, seria o uso de um método que pudesse atuar no sinal de acordo com o próprio conteúdo ruidoso de seus coeficientes. Fica implícito que os autores se referem à criação de um método não-limiar.

A contaminação do sinal de voz  $s[n]$  para estes métodos pode ser feita com o ruído gaussiano branco ou com algum tipo de ruído colorido (ruído produzido pelos pneus de um carro viajando por uma rodovia, por exemplo). A melhoria dos sinais de voz, aqui considerados, consiste em procurar recuperar o sinal original a partir do sinal ruidoso  $y[n]$ . A Figura 1 representa a situação ideal, na qual se recupera o sinal original de maneira exata.

Figura 1: Esquema para a recuperação exata de um sinal de voz contaminado por ruído aditivo.



Fonte: Elaborado pelo próprio autor.

O nível de influência do ruído em um sinal pode ser medido de várias formas. Dentre as principais, destaca-se a SNR (*Signal to Noise Ratio*). Conhecida também como relação sinal/ruído. Expressa em decibéis (dB), a SNR também pode ser utilizada para avaliar o grau de melhoria que um algoritmo de redução de ruído causa no sinal. Em outras palavras, esta técnica também pode ser utilizada como uma medida de qualidade (NOCETI FILHO, 2004).

Nas próximas subseções, a SNR será discutida mais detalhadamente. Neste sentido, também serão apresentadas algumas medidas de qualidade de voz frequentemente abordadas na literatura especializada.

## 2.2 Medidas subjetivas de qualidade

Medidas de qualidade que utilizam seres humanos para a classificação dos sinais processados, são chamadas medidas subjetivas de qualidade de voz (YANG, 1999). Estas são baseadas em comparações dos dados dos sinais de voz, original e processado, por um ou mais ouvintes, que fazem uma classificação subjetiva ao longo de uma escala predeterminada (DELLER; PROAXIS; HANSEM, 1993). O procedimento é simples, porém requer uma grande quantidade de tempo e custo.

As medidas subjetivas são baseadas na ideia de que a maioria das respostas dos ouvintes é semelhante para um mesmo sinal. Neste sentido, apresentam-se a um grupo de ouvintes os sinais de voz, original e processado, para que baseado nas respostas, possa-se fazer uma análise estatística dos resultados finais (YANG, 1999).

Existem na literatura vários tipos de medidas subjetivas. Dentre elas, duas se destacam: a *Mean Opinion Score* (MOS) e a *Degradation Mean Opinion Score* (DMOS). Ambas são utilizadas com frequência para estimar o desempenho dos sistemas de telecomunicações (YANG, 1999).

- MOS – Ouvintes são utilizados para classificar a qualidade global de uma frase ou enunciado que está sendo testado, sem ter como referência o sinal original (YANG, 1999). A classificação é feita em cinco categorias, conforme a Tabela 1.
- DMOS – Os ouvintes são indagados sobre a taxa de irritação ou nível de degradação de uma frase ou enunciado, comparando o sinal de teste com o original (YANG, 1999). A classificação é feita em cinco categorias, conforme a Tabela 2.

Tabela 1: MOS e a qualidade do discurso correspondente.

Classificação	Qualidade do discurso
5	Excelente
4	Bom
3	Regular
2	Insatisfatório
1	Ruim

Fonte: Yang (1999).

Tabela 2: DMOS e o correspondente nível de degradação.

Classificação	Nível de degradação
5	Inaudível
4	Audível mas não incômodo
3	Levemente incômodo
2	Incômodo
1	Muito incômodo

Fonte: Yang (1999).

### 2.3 Medidas objetivas de qualidade

Atualmente, os métodos mais precisos para a avaliação da qualidade de voz são os métodos subjetivos. Porém, por razões de custo e tempo necessários para a utilização destes métodos, muitos esforços estão sendo empregados no desenvolvimento de medidas objetivas (HU; LOIZOU, 2008).

Para Deller; Proakis; Ransen (1993), as medidas objetivas de qualidade são baseadas em comparações matemáticas entre os sinais de voz original e processado. A avaliação por meio de medidas objetivas de qualidade proporciona meios repetitivos, precisos e quantitativos de se comparar o desempenho de algoritmos de melhoria em sinais de voz (SOARES, 2009).

Para Quackenbush et al. (1988) e Hu e Loizou (2008), as medidas objetivas mais conhecidas e amplamente utilizadas são a *Signal-to-noise ratio* (SNR), a *segmental SNR* (SNRseg), que é uma variação da SNR original e a *perceptual evaluation of speech quality* (PESQ).

Dentre as três medidas mencionadas, a PESQ apresenta uma dificuldade maior em sua realização. Proposta pela ITU (*International Telecommunications Union*), a avaliação realizada pela PESQ é baseada em características psicoacústicas do ouvido humano. Ela representa um modelo cognitivo e os resultados são expressos por meio de uma escala análoga à da avaliação MOS, ou seja, sua pontuação varia entre 1 e 5. Resultados considerados bons são aqueles cuja nota PESQ é maior ou igual a três, caso contrário, o sinal será considerado de baixa qualidade (BEERENDS ET AL., 2002).

Em contrapartida, a SNR é de fácil implementação. Ela representa um erro médio sobre o tempo e a frequência do sinal processado (SOARES, 2009). Como estas medidas fazem uma comparação matemática dos sinais de voz original e processado, os mesmos precisam estar sincronizados. Caso contrário, o resultado obtido terá pouco a ver com as distorções introduzidas. A SNR, especificamente para um sinal de voz, é calculada da seguinte forma (DELLER; PROAKIS; HANSEN, 1993):

$$\text{SNR} = 10 \log_{10} \left( \frac{\sum_{i=0}^{N-1} s_i^2}{\sum_{i=0}^{N-1} w_i^2} \right), \quad (2)$$

sendo  $s_i$  e  $w_i$  segmentos de voz e ruído, respectivamente. O número de amostras do segmento escolhido é representado por  $N$ .

---

### 3 ANÁLISE WAVELET

A primeira referência ao termo wavelet data de 1910, em uma tese de Alfred Haar (HAAR, 1910). Esta nova ideia de função foi difundida somente algumas décadas depois, onde passou a ser conhecida como a primeira função wavelet (SOARES, 2009). A teoria conhecida hoje acerca de wavelets foi apresentada em meados da década de 1980 por Jean Morlet, Yves Meyer e Alex Grossman, que integravam a equipe do Centro de Física Teórica de Marseille, na França (DUARTE, 2005). Em 1989, Stephane Mallat publicou um trabalho no qual utilizou bases de wavelets ortonormais para decompor uma imagem. Nesse trabalho, Mallat interligou o conceito de wavelets com a teoria de processamento digital de sinais (MALLAT, 1989). Este feito culminou com uma grande difusão da teoria wavelet.

As funções wavelets podem distinguir as características locais de um sinal em diferentes escalas, e ainda, cobrir toda a região no qual o sinal é estudado por meio de translações. Pelo fato de a grande maioria das wavelets possuírem suporte compacto, elas são muito úteis na análise de sinais não-estacionários. Neste sentido, a análise wavelet leva certa vantagem com relação à análise de Fourier (LIMA, 2004). As funções wavelets constituem uma ferramenta matemática para decompor funções, permitindo que as mesmas sejam descritas de uma forma que apresente os mais finos detalhes, ou de uma forma grosseira que permita uma visão “global”. É interessante ressaltar que aproximação ou representação de funções, conhecidas ou não, por meio de funções especiais pode ser vista como um tema central da análise matemática (BLATTER, 1998).

As aplicações das funções wavelets se concentram principalmente na área de processamento de sinais. A análise de sinais é parte essencial das atividades científicas e tecnológicas contemporâneas. Algumas das aplicações de wavelets se dão em telecomunicações, medicina, televisão, compressão de imagens, análise de transitórios em linhas de potência, caracterização de sinais acústicos, física quântica, descontaminação de sinais, neurofisiologia, análise de sinais biomédicos (eletrocardiogramas, mamografias digitais, eletroencefalogramas, sequências de DNA), previsão de comportamento de mercado financeiro, entre outras (DUARTE, 2005).

### 3.1 Funções Wavelets

Como a análise de Fourier, a análise wavelet lida com expansões de funções em termos de um conjunto base de funções. Em contrapartida à análise de Fourier, a análise wavelet expande funções, não em termos de polinômios trigonométricos, mas em termos de wavelets, que são geradas por dilatação e translação de uma função fixa dita wavelet mãe (LEE AND YAMAMOTO, 1994).

É denotado por  $L^2(\mathbb{R})$  o espaço das funções, mensuráveis no sentido de Lebesgue, de quadrado integrável, comumente denominado de espaço das funções de energia finita, isto é, se  $\psi(t) \in L^2(\mathbb{R})$  então (BLATTER, 1998):

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty.$$

Fisicamente, a condição acima afirma que a energia total do sinal é finita. Lembrando que o espaço  $L^2[a, b]$  é definido como segue:

$$L^2[a, b] = \left\{ f: [a, b] \rightarrow \mathbb{C}; \int_a^b |f(t)|^2 dt < \infty \right\},$$

onde se pode ter  $a = -\infty$  e/ou  $b = \infty$ . O produto interno em  $L^2[a, b]$  é definido como:

$$\langle f, g \rangle_{L^2} = \int_a^b f(t) \overline{g(t)} dt \quad f \text{ e } g \in L^2[a, b].$$

É interessante salientar que, caso o intervalo  $[a, b]$  seja finito, têm-se a integral de Riemann (BOGGESS AND NARCOWICH, 2001).

Para facilitar a compreensão, considere  $a = 0$  e  $b = 1$ . O conjunto de funções  $\{1, t, t^2, t^3 \dots\}$  é linearmente independente e pertence a  $L^2[0,1]$ . A função  $f(t) = 1/t$  é um exemplo de função que não pertence a  $L^2[0,1]$ . Note que

$$\int_0^1 \left(\frac{1}{t}\right)^2 dt = \infty.$$

**Definição 1:** Uma função  $\psi(t) \in L^2(\mathbb{R})$  é denominada wavelet se sua transformada de Fourier  $\hat{\psi}(\omega)$  satisfaz a condição (chamada condição de Admissibilidade):

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty. \quad (3)$$

Dessa condição, tem-se que  $\lim_{\omega \rightarrow 0} \hat{\psi}(\omega) = 0$  (DAUBECHIES, 1992). Deste modo, se  $\hat{\psi}(\omega)$  é contínua então  $\hat{\psi}(0) = 0$ . Como

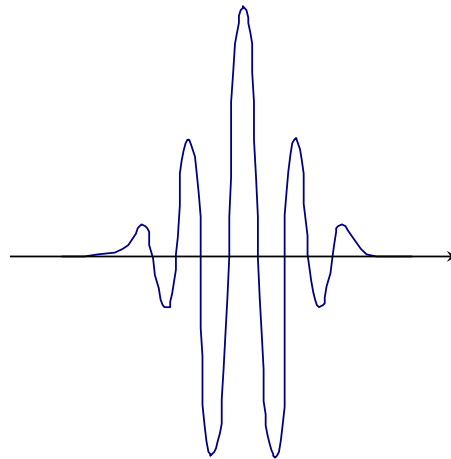
$$\hat{\psi}(\omega) = \int_{-\infty}^{\infty} \psi(t) e^{-j\omega t} dt,$$

segue que

$$\int_{-\infty}^{\infty} \psi(t) dt = 0. \quad (4)$$

Analisando a equação (4), conclui-se que  $\psi(t)$  deve oscilar de modo a cancelar as áreas positivas e negativas, tornando assim a integral igual a zero. Na Figura 2 é mostrado um exemplo de wavelet que ilustra a característica geométrica de  $\psi(t)$ .

Figura 2: Gráfico de uma Wavelet.



Fonte: Duarte (2005).

Dada uma função  $\psi(t) \in L^2(\mathbb{R})$ , as funções wavelets associadas, via relação de translação e escala, são definidas por:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \quad (5)$$

onde a função  $\psi(t)$  é dilatada por um fator  $a$  e transladada por  $b$ . O fator  $|a|^{-1/2}$  é de natureza técnica, e é usado para garantir que  $\|\psi_{a,b}\| = 1$  (BLATTER, 1998). A função  $\psi(t)$  é dita “wavelet mãe”, pois é a partir dela que é gerada uma família de wavelets, denominadas “wavelets filhas”, por meio da equação (5).

As funções wavelets possuem suporte compacto, ou decaem exponencialmente a zero quando  $t \rightarrow \pm\infty$  (LEE AND YAMAMOTO, 1994). O que implica que as wavelets são funções que possuem uma boa localização no tempo. Contribuem para uma análise local de funções, o que difere das funções base da análise de Fourier, que são não-nulas em todo o intervalo de definição, e portanto contribuem globalmente. As funções utilizadas para expandir funções na análise de Fourier são as funções trigonométricas, seno e cosseno (FIGUEIREDO, 1977).

Antes de prosseguir, fazem-se necessárias algumas definições.

**Definição 2:** Uma coleção de funções  $\{\varphi_n \in L^2(\mathbb{R}); n \in \mathbb{Z}\}$  é um conjunto ortogonal se  $\langle \varphi_n, \varphi_m \rangle_{L^2} = 0$  para  $n \neq m$ .

Uma coleção de funções  $\{\varphi_n \in L^2(\mathbb{R}); n \in \mathbb{Z}\}$  é um conjunto ortonormal se  $\{\varphi_n\}$  for ortogonal e  $\|\varphi_n\| = 1$  para cada  $n \in \mathbb{Z}$ .

**Definição 3.** Um espaço de Hilbert é um espaço vetorial  $V$  com produto interno  $\langle \cdot \rangle$  e com norma definida por  $\|x\| = \sqrt{\langle x, x \rangle}$  ( $x \in V$ ).

**Teorema.** Suponha que  $V_0$  seja um subespaço de um espaço de funções de Hilbert  $V$  munido com produto interno  $\langle \cdot \rangle$ . Suponha que  $\{\varphi_1, \dots, \varphi_n\}$  seja uma base ortonormal de  $V_0$ . Se  $f \in V_0$ , então:

$$f = \sum_{j=1}^n \langle f, \varphi_j \rangle \varphi_j.$$

Demonstração: ver Apêndice A.

Esta aproximação para  $f$  pode ser relaxada considerando um erro  $\varepsilon$  na aproximação como segue.

**Definição 4.** Considere um espaço de funções de Hilbert  $V$ . Uma coleção de funções  $\{\varphi_n; n \in \mathbb{Z}\} \subset V$  é um conjunto ortonormal completo se esse conjunto for ortonormal e satisfazer a seguinte condição

$$\left\| \sum_{j=-n}^n \langle f, \varphi_j \rangle \varphi_j \right\| < \varepsilon,$$

para todo  $f$  pertencente a  $V$  e todo  $\varepsilon > 0$ .

A Definição 4 afirma que combinações lineares, de uma dada base  $\{\varphi_n\}$  do espaço  $L^2(\mathbb{R})$ , podem ser utilizadas para se obter aproximações de qualquer função do espaço gerado por essa base. Os conjuntos ortonormais completos também são chamados de bases ortonormais do espaço (GOMES; VELHO; GOLDSTEIN, 1997).

### 3.2 Transformada Wavelet Contínua

A transformada wavelet contínua ou CWT (do inglês *Continuous Wavelet Transform*) fornece uma análise tempo-frequência similar à transformada de Fourier janelada (WFT, do inglês *Windowed Fourier Transform*), porém, com uma diferença fundamental, explicada a seguir.

Considere  $f(t), \psi(t) \in L^2(\mathbb{R})$ . A transformada wavelet contínua de  $f(t)$ , utilizando as funções  $\psi_{a,b}(t)$ , é definida como segue:

$$W_f(a, b) = \int_{-\infty}^{\infty} f(t) |a|^{-p} \psi\left(\frac{t-b}{a}\right) dt, \quad b \in \mathbb{R}; a \in \mathbb{R}^*; p \geq 0. \quad (6)$$

Uma similaridade entre a transformada de Fourier Janelada e a transformada wavelet está no fato de que ambas tomam o produto interno de  $f(t)$  com uma família de funções indexadas por duas variáveis,  $g_{\omega,b}(t) = g(t-b)e^{j\omega t}$  para a WFT e  $\psi_{a,b}(t) = |a|^{-p}\psi((t-b)/a)$  para a CWT. Essas funções geralmente são chamadas de funções moduladoras.

Uma diferença relevante está no fato de que a WFT introduz uma escala fixa e analisa o sinal independentemente dos valores de  $\omega$ . Já a CWT possui a propriedade de adaptar a escala aos valores de  $\omega$  (DAUBECHIES, 1992).

Na equação (6), o parâmetro  $a$  é o fator de escala. Faz-se uma analogia entre  $a$  e a escala utilizada em mapas. Se a escala aumenta,  $\psi_{a,b}$  aumenta no tempo, capturando comportamentos de tempo longo. Se a escala diminui,  $\psi_{a,b}$  sofre uma contração no tempo, capturando comportamentos de tempo curto, transitórios. Escala grande, significa visão global, enquanto que escala pequena significa visualização dos detalhes. Em outras palavras,  $|a|$  grande implica escala larga, ou baixa frequência de  $\psi_{a,b}$ ,  $|a|$  pequeno implica escala fina, ou alta frequência de  $\psi_{a,b}$  (DAUBECHIES, 1992).

A transformada wavelet sanou algumas deficiências da transformada de Fourier janelada. A principal delas é o caso em que os detalhes do sinal são menores que a largura da janela. Quando isto ocorre, têm-se um problema semelhante ao que ocorre com a

transformada de Fourier: esses detalhes não serão localizados, apesar de serem detectados. Caso as características sejam maiores que a largura da janela, tem-se o problema inverso (GOMES; VELHO; GOLDSTEIN, 1997).

O parâmetro  $b$  da equação (6) corresponde ao deslocamento da origem sobre o eixo das abscissas. Analogamente à WFT, este parâmetro é responsável por deslizar a função moduladora  $\psi_{a,b}$  sobre o eixo horizontal, a fim de cobrir todo o sinal.

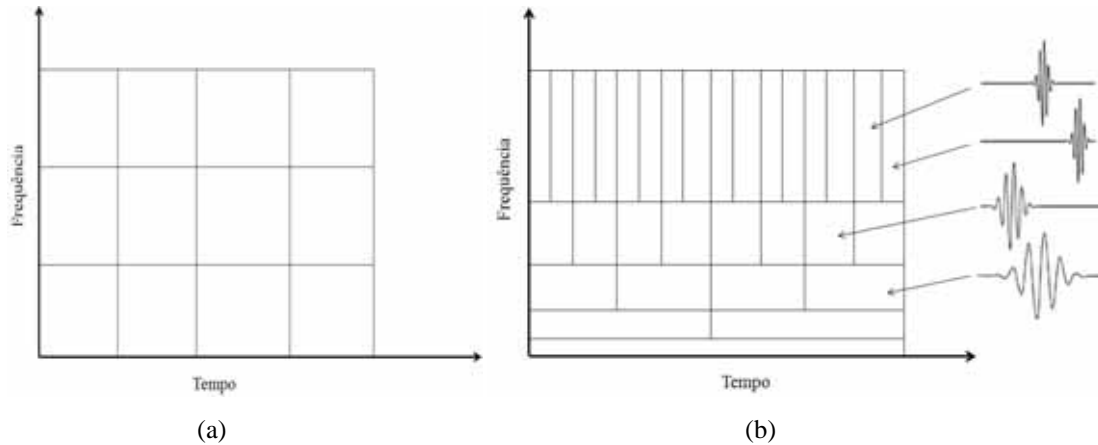
A transformada wavelet contínua, ou simplesmente transformada wavelet (WT), admite várias interpretações. Uma delas é que, por definição,  $W_f(a, b) = \langle f, \psi_{a,b} \rangle$ . Outra interpretação muito útil à teoria de processamento de sinais é que a WT corresponde a uma operação de filtragem de um sinal  $f(t)$  por um filtro cujos coeficientes são gerados pela função wavelet que está sendo utilizada na análise (STRANG AND NGUYEN, 1996).

Um fato interessante é que o escalamento possibilita a compressão ( $a < 1$ ) ou dilatação ( $a > 1$ ) da wavelet mãe  $\psi(t)$ . A wavelet mãe quando escalonada e deslocada no tempo, que é a operação de translação, origina as wavelets filhas  $\psi_{a,b}(t)$  (GOMES; VELHO; GOLDSTEIN, 1997).

Quando a transformada wavelet é aplicada a um sinal  $f(t)$ , obtém-se os coeficientes wavelets, que são resultados da operação  $\langle f, \psi_{a,b} \rangle$ . Estes coeficientes representam a semelhança entre a função  $f(t)$  e as wavelets filhas (VETTERLI; KOVAČEVIĆ; GOYAL, 2011).

Pensando em uma análise tempo-frequência, uma diferença entre a transformada de Fourier Janelada e a transformada wavelet está na forma das funções de análise, ou comumente chamadas funções moduladoras,  $g_{\omega,b}$  e  $\psi_{a,b}$ . As funções  $g_{\omega,b}$  consistem de envelopes deslocados no tempo de  $b$  unidades e “preenchidos” por oscilações de altas frequências, todos possuindo a mesma largura. Em contrapartida, as wavelets  $\psi_{a,b}$  têm a capacidade de adaptar a largura às frequências constituintes. Quando é constituída por baixas frequências, a wavelet é bem larga, já para altas frequências a wavelet será estreita (ARAÚJO, 2007). Neste sentido, pensando em analisar fenômenos de curta duração, como singularidades em funções, a transformada wavelet é superior à transformada de Fourier Janelada. A representação destes envelopes no plano tempo-frequência é chamada Caixa de Heisenberg (ARAÚJO, 2007). Nas Figuras 3 (a) e 3 (b), respectivamente, encontram-se as Caixas de Heisenberg para a WFT e a CWT.

Figura 3: Caixas de Heisenberg para a WFT (a) e para a CWT (b).



Fonte: Adaptado de Araújo (2007).

### 3.3 Transformada Wavelet Contínua Inversa

Se o sinal  $f(t)$  for decomposto utilizando uma wavelet  $\psi(t)$  que satisfaça a condição de admissibilidade:

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty,$$

onde  $\hat{\psi}(\omega)$  é a transformada de Fourier de  $\psi(t)$ , então é possível reconstruir o sinal através da transformada wavelet contínua inversa definida pela equação:

$$f(t) = \frac{1}{c_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_f(a, b) \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \frac{da db}{|a|^2}$$

e, por conveniência,

$$f(t) = \frac{1}{c_\psi |a|^2 \sqrt{|a|}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_f(a, b) \psi\left(\frac{t-b}{a}\right) da db. \quad (7)$$

Nota-se que as equações (6) e (7) usam o mesmo núcleo,  $\psi_{a,b} = |a|^{-\frac{1}{2}} \psi\left(\frac{t-b}{a}\right)$ .

Segundo Daubechies (1992), a equação (7) pode ser encarada de dois modos:

- Um modo de reconstrução de  $f(t)$ , desde que sua transformada wavelet inversa seja conhecida;
- Um modo de representação de  $f(t)$ , como uma superposição das wavelets filhas.

### 3.4 Transformada Wavelet Discreta

Da discretização da CWT origina-se a transformada wavelet discreta (DWT, do inglês *Discrete Wavelet Transform*). Para isso, faz-se necessário a discretização dos parâmetros de escala  $a$  e de translação  $b$  (HERNANDES; WEISS, 1996):

$$a = a_0^m,$$

$$b = na_0^m b_0,$$

onde  $m, n \in \mathbb{Z}$ ,  $a_0 > 1$  e  $b_0 \neq 0$ .

Substituindo  $a$  e  $b$  na equação (6), tem-se a transformada wavelet discreta:

$$W_f(m, n) = \frac{1}{\sqrt{a_0^m}} \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t - na_0^m b_0}{a_0^m}\right) dt. \quad (8)$$

Da equação (8) e das condições de discretização, fazem-se necessárias algumas observações (SOARES, 2009):

- A transformada wavelet discreta é definida apenas para valores de escalas positivos  $a_0 > 1$ ;
- O passo de translação é proporcional à escala  $b = na_0^m b_0$ ;
- A transformada wavelet discreta produz um conjunto finito de coeficientes wavelet  $W_f(m, n)$ ;
- O processamento é realizado sobre o tempo contínuo.

### 3.5 Transformada Wavelet Discreta Inversa

Segundo Daubechies (1992), uma função  $f(t)$  pode ser reconstruída através de seus coeficientes wavelets discretos se a wavelet mãe gerar um frame.

**Definição 5:** Uma coleção de funções  $\{\varphi_n; n \in \mathbb{Z}\}$  é considerada um *frame* se existem constantes  $A$  e  $B$ , satisfazendo  $0 < A \leq B < \infty$ , tal que para toda  $f(t) \in L^2(\mathbb{R})$ :

$$A\|f\|^2 \leq \sum_{n=-\infty}^{\infty} |\langle f, \varphi_n \rangle|^2 \leq B\|f\|^2.$$

As constantes  $A$  e  $B$  são denominadas cotas ou limites do frame.

Os limites  $A$  e  $B$  de um frame governam a precisão da reconstrução do sinal na transformada wavelet discreta inversa. Quanto mais próximos forem os valores de  $A$  e  $B$ , mais precisa será a reconstrução (BLATTER, 1998). A função  $f(t)$  é reconstruída pela equação

$$f(t) = \frac{2}{A+B} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} W_f(m, n) \psi_n(t).$$

Se  $A = B$  o frame é chamado de *tight*. Um frame *tight* é exato se  $A = B = 1$ , implicando que todas as wavelets formam uma base ortonormal para  $L^2(\mathbb{R})$  e, então, a reconstrução de qualquer função será exata (DAUBECHIES, 1992).

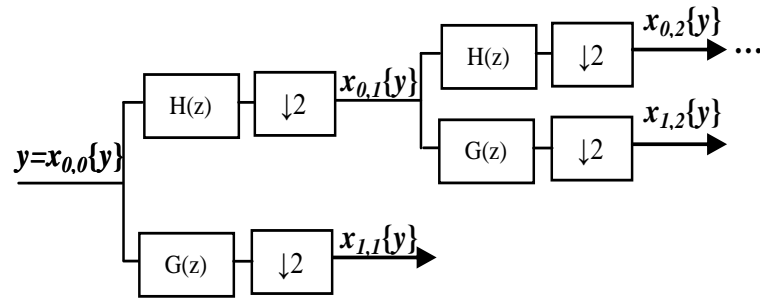
### 3.6 Interpretação da DWT como um banco de filtros

Em processamento digital de sinais, um problema relevante consiste em realizar uma boa representação do sinal que se deseja analisar, utilizando um número pequeno de sinais básicos. Estes sinais podem ser senoides ou wavelets. No ambiente das funções wavelets, os filtros digitais desempenham um papel importante. A transformada wavelet discreta pode ser interpretada como um banco de filtros (STRANG AND NGUYEN, 1996).

Para entender a função que um filtro pode desempenhar em situações práticas, considera-se o problema de compressão de sinais. Sinais presentes na natureza possuem um conteúdo altamente oscilatório. Para efeito de compressão, pode-se trabalhar com os componentes de baixas e altas frequências através de filtros passa-baixas e passa-altas. Um filtro passa-baixas não permite que as oscilações de mais alta frequência do sinal passem por ele, obtendo assim os componentes de baixas frequências. O filtro passa-altas permite que tais oscilações passem por ele, obtendo desta forma os componentes de alta frequência do sinal. A partir desta decomposição, os filtros de síntese podem reconstruir o sinal original de maneira exata. O processo de compressão é feito entre a fase de análise (decomposição) e a fase de síntese (reconstrução). Frequências que são mal representadas podem ser intencionalmente esquecidas, desconsideradas. A eliminação de tais coeficientes não prejudicará o processo de reconstrução (STRANG AND NGUYEN, 1996).

Um algoritmo eficiente para calcular a transformada wavelet discreta (DWT) foi proposto por Mallat (MALLAT, 1989; MALLAT, 1999). Este método, chamado de Transformada Wavelet Rápida, usa um banco de filtros digitais numa estrutura de árvore (DAUBECHIES, 1992; STRANG AND NGUYEN, 1996), conforme ilustrado na Figura 4. Esta estrutura é chamada de árvore de decomposição wavelet.

Figura 4. Árvore de decomposição wavelet (banco de filtros de análise).



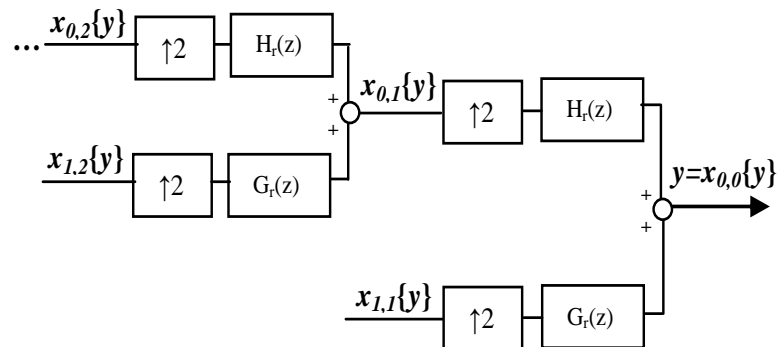
Fonte: Adaptado de Mallat (1999).

Na Figura 4,  $y$  representa o sinal a ser decomposto e os coeficientes  $x_{0,j}\{y\}$  e  $x_{1,j}\{y\}$  estão relacionados, respectivamente, com as aproximações e os detalhes do sinal no nível de resolução  $j$  ( $j = 1, 2, \dots$ ).  $H$  e  $G$  são filtros passa-baixas e passa-altas, respectivamente, convenientemente escolhidos de forma a garantir que o banco de filtros seja invertível.  $(\downarrow 2)$  denota uma operação de subamostragem (*downsampling*), que consiste em remover uma a cada duas amostras do sinal. Cada filtro, seguido do operador de subamostragem é chamado ou de canal passa-baixas ou de canal passa-altas.

Como as wavelets têm uma natureza passa-faixas (RIOUL AND VETTERLI, 1991), os detalhes estão associados a faixas de frequências do sinal analisado, cada uma situada aproximadamente entre  $2^{-(j+1)}\omega_s$  e  $2^{-j}\omega_s$ , sendo que  $\omega_s$  é a taxa de amostragem.

Algumas restrições sobre os filtros  $H$  e  $G$  garantem que a árvore de decomposição wavelet seja invertível, isto é, que o sinal  $y$  possa ser reconstruído a partir de seus coeficientes de aproximação e detalhes (STRANG AND NGUYEN, 1996). A estrutura inversa da DWT, apresentada na Figura 5, é chamada de árvore de reconstrução wavelet.

Figura 5. Árvore de reconstrução wavelet (banco de filtros de síntese).



Fonte: Adaptado de Mallat (1999).

Na Figura 5,  $H_r$  e  $G_r$  são filtros de reconstrução associados aos filtros  $H$  e  $G$  apresentados na Figura 4.  $(\uparrow 2)$  denota a operação de sobreamostragem (*upsampling*), que consiste em inserir um zero entre cada duas amostras do sinal. Cada um dos filtros usados no processo de decomposição e reconstrução tem  $2k$  pesos. Se o banco de filtros da DWT é ortogonal e se  $h_i$  e  $g_i$  são os respectivos pesos dos filtros  $H$  e  $G$ , na Figura 4, então os pesos  $h_i^r$  e  $g_i^r$  dos respectivos filtros de reconstrução  $H_r$  e  $G_r$  são obtidos a partir da equação (9) (STRANG AND NGUYEN, 1996).

$$\begin{cases} h_i^r = h_{(2k-1)-i}, & i = 0, \dots, 2k - 1 \\ g_i^r = g_{(2k-1)-i} & i = 0, \dots, 2k - 1 \end{cases} \quad (9)$$

---

## 4 MÉTODOS DE REDUÇÃO DE RUÍDO EM SINAIS DE VOZ NO DOMÍNIO WAVELET

O procedimento básico de eliminação de ruído em sinais de voz, no domínio wavelet, pode ser resumidamente descrito como se segue:

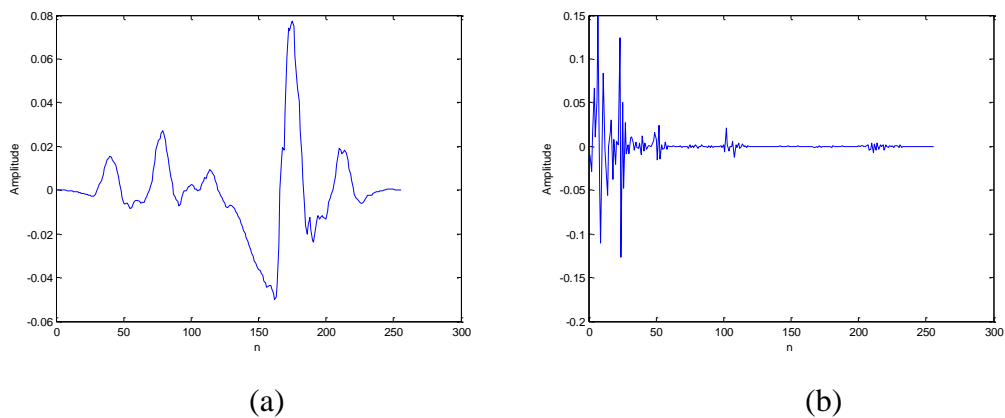
- Considere um quadro do sinal ruidoso (técnica de janelamento);
- Calcule a sua transformada wavelet;
- Aplique uma técnica de filtragem de ruído no sinal transformado;
- Calcule a transformada wavelet inversa destes coeficientes;
- Use a parte central do quadro para compor a saída;
- Repita os procedimentos anteriores para todos os blocos.

A técnica de janelamento referida é comumente empregada com 50% de sobreposição. Geralmente, a função limiar é aplicada apenas aos coeficientes de detalhes. Considera-se que os coeficientes de aproximação representam a estrutura básica do sinal original (SOARES, 2009).

A transformada wavelet vem sendo amplamente utilizada em métodos de processamento de sinais de voz. Isto se deve principalmente ao fato já mencionado na seção 3.2, de que com a TW é possível ajustar a escala para se obter máxima resolução no processo de decomposição, minimizando possíveis perdas. Sendo assim, a transformada wavelet se torna uma ferramenta eficiente no processo de modelagem de sinais não-estacionários, como é

o caso dos sinais de voz (STRANG AND NGUYEN, 1996). Outra característica explorada pelos métodos que usam wavelets é que, no domínio wavelet, a energia do sinal está concentrada principalmente em um número pequeno de coeficientes (MORETTIN, 1999). Se comparados com os demais, estes coeficientes assumem valores relativamente maiores. Neste sentido, pode-se eliminar ou atenuar os coeficientes com valores absolutos próximos de zero para combater o ruído enquanto que se preserva a informação relevante do sinal original. A Figura 6 ilustra este fato. Nela consta um trecho de um sinal de voz limpo e sua transformada wavelet. O segmento é um trecho de voz de um sinal em português (a), e o mesmo decomposto usando a função wavelet de Daubechies de ordem 10 (b). Utilizou-se uma janela de Hanning com 256 pontos para a aquisição do sinal.

Figura 6: Trecho de um sinal de voz (a) e sua transformada Wavelet (b).



Fonte: Elaborado pelo próprio autor.

A seleção de uma função wavelet para o tratamento de um sinal de voz pode ser baseada em alguns critérios. Para Duarte (2005), uma boa escolha seria aquela que minimiza o erro na reconstrução do sinal e maximiza a relação sinal/ruído (SNR). Um segundo critério a se considerar é o fato de que grande parte da energia de um sinal de voz está concentrada nas baixas frequências (DELLER; PROAKIS; HANSEN, 1993). Sendo assim, a melhor escolha para uma base wavelet estaria fundamentada na propriedade de conservação de energia nos coeficientes de baixa frequência.

As seções subsequentes deste capítulo têm por objetivo passar a ideia geral por trás dos métodos que utilizam o limiar e dos métodos que não o utilizam.

#### 4.1 Redução de ruído em sinais de voz por métodos que utilizam limiares

Um fator importante em processamento de sinais de voz é reduzir potencialmente o ruído, ao mesmo tempo em que se mantém a qualidade ou a inteligibilidade do sinal. Para isso, faz-se o uso das transformadas integrais. Existem métodos que utilizam a transformada de Fourier (VIEIRA FILHO, 1996) e outros que utilizam a transformada wavelet discreta (DWT) (DONOHO, 1995; DAE-SUNG et al., 2002; DUARTE, 2005; SOARES, 2009).

Para Nievergelt (1999) e Morettin (1999), a vantagem da transformada wavelet está no fato de que muitos coeficientes são irrelevantes na reconstrução do sinal. Assim, uma atuação mais drástica nestes coeficientes, não afetaria uma reconstrução perfeita do sinal original. É interessante lembrar, que este fato propicia aplicações como compressões de sinais no domínio wavelet (DUARTE et al., 2003).

As primeiras funções de limiar que surgiram foram as *Hard thresholding* e a *Soft thresholding* (ANTUNES, 2006). Apesar dessas funções ainda serem fortemente utilizadas na literatura, nasceram outras funções com a finalidade de se obter melhores resultados. Cita-se como exemplo a função *Sigmoidal thresholding* (DUARTE, 2005).

Sendo  $Y$  a representação do sinal  $y(n)$  no domínio wavelet, as equações (12), (13) e (15) contêm as funções de transferência para os métodos citados anteriormente.

*Hard thresholding.* Este método elimina os coeficientes que estão abaixo do limiar. Apresentando grande eficiência no caso de compressão de sinais (SOARES et al., 2008):

$$y_{thr} = \begin{cases} Y, & \text{se } |Y| > \lambda \\ 0, & \text{se } |Y| \leq \lambda \end{cases} \quad (12)$$

*Soft thresholding.* Neste caso, assume-se que os componentes ruidosos são distribuídos igualmente em todos os coeficientes wavelets. Deste modo, todos os coeficientes acima do limiar são reduzidos (SOARES et al., 2008).

$$y_{thr} = \begin{cases} \text{sgn}(Y)(|Y| - \lambda) & \text{se } |Y| > \lambda \\ 0 & \text{se } |Y| \leq \lambda \end{cases} \quad (13)$$

*Sigmoidal thresholding.* Este método atenua os coeficientes que estão abaixo do limiar usando a função sigmoide apresentada na equação (14). Os coeficientes são atenuados por um fator dependente de seus próprios valores.

$$\text{Sigmoide}(t) = \frac{1 - e^{-\gamma t}}{1 + e^{-\gamma t}}, \quad (14)$$

$$y_{thr} = \begin{cases} Y, & \text{se } |Y| > \lambda \\ |Y| \text{Sigmoide}(Y), & \text{se } |Y| \leq \lambda \end{cases} \quad (15)$$

Existem vários métodos para se calcular o limiar, dentre eles se destaca o método *VisuShrink* proposto por Donoho e Johnstone (1994). Este limiar é comumente denominado limiar universal. Justifica-se pelo fato de ser válido para sinais gerais (BALAKRISHNAN; BORGES; PARCHMENT, 2010):

$$\lambda = \theta \sqrt{2 \log(N)}$$

onde  $N$  é o comprimento de  $Y$  e  $\theta = \text{mediana}(|Y|)/0,6745$ , que representa a estimativa do ruído ou ainda uma aproximação do desvio padrão do ruído.

Tanto o limiar *hard* como o *soft* introduzem descontinuidades no tempo e na frequência do sinal em processamento, o que pode ser observado como falhas no espectrograma do mesmo. Na prática, este fato é notado por um sussurro, um desconforto auditivo no momento de transição de um trecho considerado ruidoso para um trecho de voz (DUARTE, 2005).

Alguns métodos que fazem o tratamento de um sinal levando em conta o limiar se mostraram eficientes, principalmente quando combinados com alguma outra metodologia (SOARES, 2009). Porém, estes apresentam problemas quando alguns coeficientes ruidosos possuem valores maiores ou iguais aos coeficientes de voz. Isto ocorre com frequência quando o ruído presente no sinal for o ruído colorido, que é o tipo de ruído geralmente presente em situações reais. Neste caso, o algoritmo pode confundir voz com ruído, eliminando coeficientes importantes e inserindo distorções na fala. O problema inverso acontece quando, após a contaminação, muitos coeficientes redundantes assumem valores acima do limiar. O que acarreta um excesso de ruído no sinal processado. Por fim, outro problema está relacionado à utilização de um único valor de limiar para todas as faixas de frequências do sinal (DUARTE, 2005).

Uma forma de se evitar a degradação do sinal consiste na identificação dos trechos de voz e silêncio. Alguns métodos propostos usam medidas baseadas na distribuição de energia, em cada janela do sinal analisado, para a identificação destes trechos (DUARTE, VIEIRA FILHO; VILLARREAL, 2009; SHEIKHZADEH AND ABUTALEBI, 2001; SEOK AND BAE, 1997). Após a identificação, pode-se utilizar um segundo limiar para atuar em trechos considerados de voz.

Duarte (2005) afirma que um bom algoritmo de redução de ruído baseado no limiar deve identificar trechos de silêncio/voz, assim como trechos *voiced/unvoiced*. Trechos considerados *unvoiced* são aqueles em que o som das letras podem se confundir com ruído. Dependendo do contexto, pode ser o caso das letras *s, f* e *x*. De um modo mais geral, quando

as cordas vocais vibram e produzem som, este é denominado *voiced*. Caso as cordas vocais não vibrem durante a produção do som, temos um segmento *unvoiced* (DELLER; PROAKIS; HANSEN, 1993; ATAL AND RABINER, 1976). Desta forma, o método de limiar deve ser aplicado de maneira diferente também para estas regiões.

## 4.2 Redução de ruído em sinais de voz por métodos não-limiare

Como citado na seção anterior, métodos para redução de ruído em sinais de voz baseados em limiars, podem ser muito delicados. Eles podem facilmente introduzir distorções, ocasionando desconfortos ao ouvinte. Uma forma de sanar os problemas apresentados por estes métodos seria o uso de uma técnica que pudesse atuar no sinal, de acordo com seu próprio conteúdo ruidoso (SOARES et al., 2008). Neste sentido, o limiar já não se faz necessário.

A principal ideia por trás dos métodos não-limiare é o uso de um único operador. Este tem por finalidade atuar de maneira equivalente em todo o sinal, de modo a reduzir o ruído sem a inserção de distorções. Para isso, é desejável que este operador seja contínuo, e ainda, que seu gráfico seja uma curva suave. Levando em conta este fato, alguns autores fazem o uso de um ou mais operadores que atuam de maneira análoga em todo o sinal. Aproveitam-se as características particulares de cada um, combinando-os de maneira a manter as principais características do sinal de voz (SOARES, 2009; QIANG AND WAN, 2003).

Os métodos considerados não-limiare vêm se mostrando eficientes. Isto se deve principalmente ao fato de suas funções de transferências não apresentarem discontinuidades. Porém, como era de se esperar, apresentam alguns problemas. Analogamente aos métodos que usam limiars como, por exemplo, dificuldades no tratamento de regiões *unvoiced*.

### 4.2.1 Análise de um método não-limiar

Como visto nas seções anteriores, apesar de eficientes, os métodos de redução de ruído baseados em wavelets podem apresentar alguns inconvenientes. Assim, pode-se dizer que a busca por métodos baseados em wavelets que não utilizem limiars é um dos objetivos dos pesquisadores da área de processamento de sinais.

Soares et al. (2011) apresentam uma metodologia para redução de ruído, sem o uso de limiar, que proporciona resultados satisfatórios. Este método consiste na aplicação de três operadores parciais e independentes, com o objetivo de tornar o sinal ruidoso inteligível. O

filtro para a redução de ruído é constituído pela combinação polinomial de três operadores mais um ajuste sigmoidal. Este último é aplicado com o intuito de manter as amplitudes dos sinais, original e processado, no mesmo intervalo. A seguir é apresentada uma descrição detalhada deste método.

Seja  $Y[n] = \{y_0, \dots, y_{N-1}\}$  um sinal ruidoso no domínio wavelet. Soares et al. (2011) propõem os seguintes operadores:

- Operador Média Simples (OMS)

$$p_i = \frac{|y_i| + |y_{i+1}|}{2}, \text{ para } i = 0, \dots, N - 2, \text{ e } p_{N-1} = p_{N-2}. \quad (16)$$

- Operador relação sinal/ruído a priori (SNRP)

$$z_i = \frac{|y_i|}{\alpha + |y_i|}, \text{ para } i = 0, \dots, N - 2 \text{ e } z_{N-1} = z_{N-2}. \quad (17)$$

- Operador relação sinal/ruído a posteriori (SNRPR)

$$r_i = \frac{|y_i|}{\alpha + |y_{w_i}|}, \text{ para } i = 0, \dots, N - 2 \text{ e } r_{N-1} = r_{N-2}. \quad (18)$$

Nas equações (17) e (18),  $\alpha$  é um valor compreendido no intervalo [0,1]. O vetor  $y_{w_i}$  representa a média do ruído presente nos correspondentes componentes do último trecho de silêncio do sinal em processamento. O operador SNRPR é aplicado apenas nas regiões de voz do sinal.

A combinação polinomial dos operadores dados nas equações (16), (17) e (18) é feita segundo a equação (19).

$$f_i = p_i^3 + z_i^2 + r_i \quad (19)$$

O objetivo da equação (19) é minimizar as distorções produzidas pelos operadores OMS e SNRP.

Num primeiro momento, a redução de ruído se daria pela multiplicação dos coeficientes  $f_i$ , pelos coeficientes  $y_i$  do sinal ruidoso. Porém, esta operação aumenta consideravelmente a amplitude do sinal processado, excedendo a amplitude do sinal original. Neste sentido, é feito um ajuste sigmoidal nos coeficientes  $f_i$ :

$$h_i = \left| \frac{1 - e^{-\sigma f_i}}{1 + e^{-\sigma f_i}} \cdot \frac{1 - e^{\sigma f_i}}{1 + e^{\sigma f_i}} \right|. \quad (20)$$

O parâmetro  $\sigma$  controla a inclinação das duas sigmoides envolvidas na equação (20). Para fortes níveis de contaminação, é necessário utilizar um valor pequeno para  $\sigma$ , a fim de efetuar

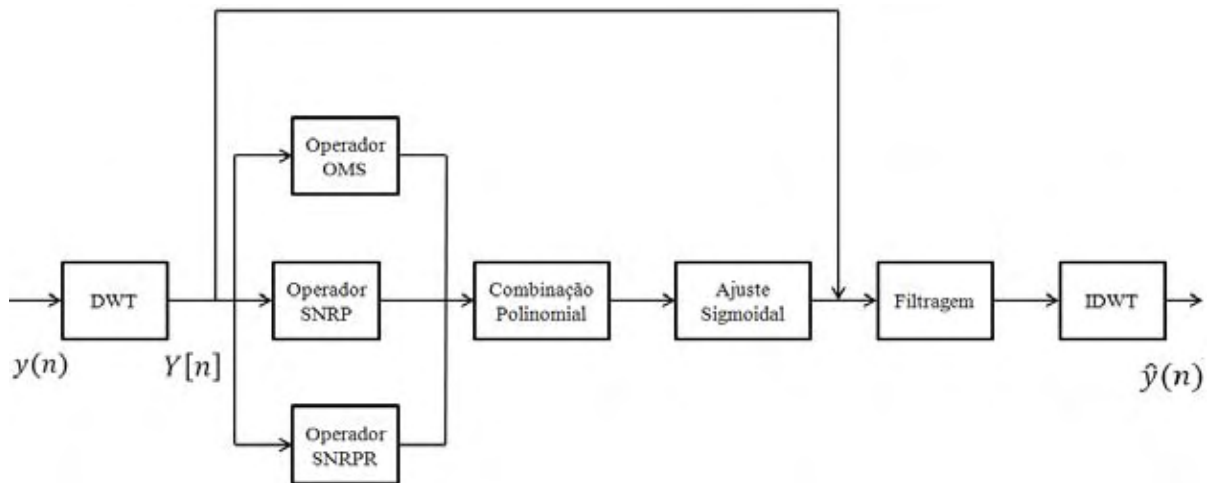
uma maior redução de ruído. A equação (20) gera os coeficientes do filtro utilizado para a redução do ruído, e, normaliza a amplitude do sinal processado no intervalo  $[-1,1]$ .

O processo de filtragem se faz através do produto entre os coeficientes  $h_i$  do filtro  $H$  e os correspondentes coeficientes  $y_i$  do sinal ruidoso  $Y[n]$ , no domínio wavelet. Obtêm-se assim os coeficientes do sinal processado  $\hat{Y}[n]$ , no domínio wavelet, conforme a equação (21).

$$\hat{Y}_i = y_i \cdot h_i \quad (21)$$

O sinal processado resultante  $\hat{y}[n]$  é obtido através da IDWT aplicada nos coeficientes de  $\hat{Y}[n]$ . O diagrama de blocos do método proposto por Soares et al. (2011) é mostrado na Figura 7.

Figura 7: Diagrama do Método Proposto por Soares et al. (2011).



Fonte: Adaptado de Soares (2009).

Os testes realizados usando este método mostraram que o mesmo é mais eficiente quando o ruído em questão é o ruído branco. Vale ressaltar que, bons resultados também foram alcançados quando o ruído presente no sinal era o colorido. A grande vantagem deste método é o uso de uma função de transferência que não utiliza o limiar para reduzir o ruído. A função, sendo contínua, não introduz distorções nos trechos de voz do sinal, causando também uma redução uniforme do nível de ruído em todo o sinal (SOARES, 2009). O uso da função sigmoide faz com que o filtro construído para atuar na redução de ruído, não atinja uma amplitude maior do que 1. Este fato impede que a amplitude do sinal processado ultrapasse a amplitude do sinal original.

---

## 5 METODOLOGIA PROPOSTA

Em processamento de sinais de voz, a busca por métodos de redução de ruído que não utilizem limiares vem se tornando mais frequente. Deve-se isso à necessidade de sanar alguns inconvenientes gerados por métodos que utilizam limiar (ver seção 4.1).

O atrativo em métodos não-limiares é a possibilidade de um tratamento equivalente para todas as faixas de frequência do sinal analisado, evitando desconfortos auditivos ao ouvinte. Soares et al. (2011) propôs um método que realiza uma redução de ruído uniforme em todo o sinal. Deve-se isto à obtenção de um filtro que atua no sinal levando em consideração o nível de ruído presente em cada coeficiente, e não em cada janela, como é o caso da maioria dos métodos que utilizam limiar.

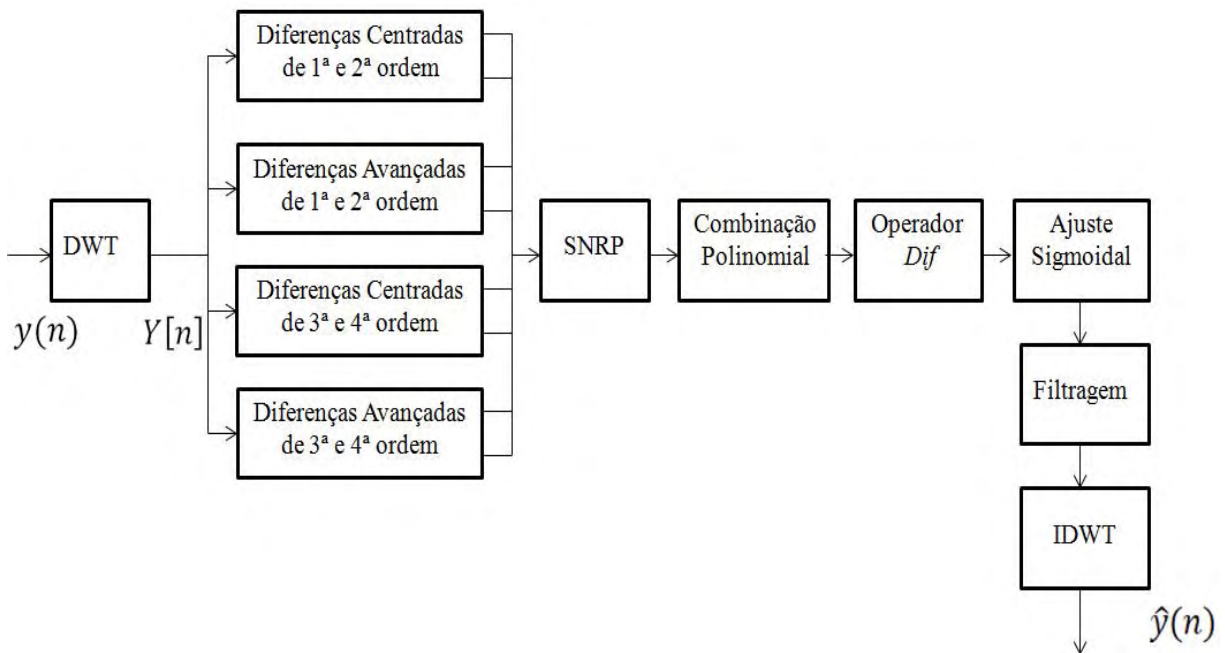
Com base nas observações feitas anteriormente, busca-se um método de redução de ruído em sinais de voz, no domínio wavelet, que possa sanar algumas das dificuldades citadas. O objetivo é a construção de um filtro que atue no sinal levando em consideração informações contidas em cada coeficiente ou em uma vizinhança do mesmo. Com isso, não se faz necessário o uso de um único valor (valor do limiar) como referência para processar todos os coeficientes de uma faixa de frequência do sinal.

## 5.1 Objetivos do método proposto

A metodologia de redução de ruído proposta neste trabalho consiste na estimação do sinal de saída a partir de um sinal ruidoso, no domínio wavelet, sem o uso de limiar. O sinal é estimado por equações de diferenças finitas (SPERANDIO; MENDES; SILVA, 2003; DUARTE, 2010). Para a estimação do ruído é usado o operador SNRP proposto por Soares et al. (2011), que é aplicado em cada uma das equações de diferenças. Uma combinação polinomial é realizada de maneira a condensar todas as equações de diferenças em uma única função de transferência. Esta, por sua vez, sofrerá um ajuste sigmoidal, visando uma melhor adequação entre as amplitudes dos sinais original e processado. Porém, antes disto, propõe-se uma pré-filtragem, realizada por um operador denominado *Dif*. O filtro é obtido no último passo, que consiste em um ajuste sigmoidal.

A Figura 8 apresenta o esquema do método proposto neste trabalho.

Figura 8: Diagrama do método proposto.



Fonte: Elaborado pelo próprio autor.

De acordo com o diagrama da Figura 8, dado um sinal ruidoso  $y(n)$  no domínio do tempo, o método proposto é aplicado da seguinte forma:

- 1- Aplica-se a transformada wavelet discreta a  $y(n)$ , com  $y(n) = s(n) + w(n)$ , sendo  $s(n)$  o sinal de voz limpo e  $w(n)$  o ruído adicionado, para obter o sinal transformado  $Y[n]$ .
- 2- Calculam-se as equações de diferenças a partir do sinal transformado  $Y[n]$ . As restrições para o uso de cada equação e a justificativa para usá-las serão assuntos das seções subsequentes.
- 3- A partir das equações de diferenças, obtêm-se as curvas determinadas pelo operador SNRP.
- 4- Após a obtenção das curvas relacionadas às diferenças e ao operador SNRP, é realizada uma combinação polinomial das mesmas, aproveitando assim as informações inerentes a cada equação.
- 5- A curva obtida pela combinação polinomial recebe um tratamento específico nas altas frequências, realizado pelo operador *Dif*. As características deste operador, e a razão em utilizá-lo, serão explicadas nas próximas seções deste capítulo.
- 6- Um ajuste sigmoidal é realizado, assim como em Soares et al. (2011), para que os coeficientes da curva obtida no passo anterior permaneçam no intervalo  $[-1,1]$  e, desta forma, a amplitude do sinal original seja mantida.
- 7- Feito o ajuste sigmoidal, o filtro utilizado no processo de filtragem está concluído. A filtragem do ruído é feita através do produto dos coeficientes do filtro pelos respectivos coeficientes do sinal  $Y(n)$ , no domínio wavelet.
- 8- Por fim, após a filtragem, o sinal processado  $\hat{y}(n)$  é reconstruído usando a transformada wavelet discreta inversa.

Na sequência, serão apresentados detalhadamente todos os processos descritos acima. Para isto será considerado o sinal de voz limpo  $s(n)$ , no formato *wav*, que modela a seguinte frase:

$$s(n): \begin{cases} \text{"Hedge apples may stain} \\ \text{your hands green."} \end{cases}$$

A partir do sinal de voz limpo  $s(n)$ , com a contaminação com ruído colorido  $w(n)$ , obtêm-se o sinal ruidoso  $y(n) = s(n) + w(n)$  (vide Figura 10). Ainda para a descrição dos processos propostos, serão considerados também:

- Os sinais  $S[n]$  e  $W[n]$ , obtidos usando a transformada wavelet discreta de Daubechies (db10) a partir dos sinais  $s(n)$  e  $w(n)$ , respectivamente;

- O sinal ruidoso  $Y[n]$ , obtido após a contaminação da seguinte forma:  

$$Y[n] = S[n] + W[n].$$

Determinou-se o uso da função wavelet de Daubechies de ordem 10 (db10), levando em consideração os estudos realizados por Duarte (2005). O autor verificou, para diversas funções wavelets, alguns aspectos fundamentais para o processamento de um sinal de voz. Tais aspectos foram: nível de redução de ruído, distorção e tempo de processamento. Para medir o nível de distorção, foi usada a distância cepstral média. A função wavelet que se sobressaiu durante os ensaios foi a Daubechies de ordem 10.

## 5.2 O uso das equações de diferenças para estimar o sinal de saída

A estimação do sinal de saída é feita através das equações de diferenças usadas na diferenciação numérica e resolução numérica de equações diferenciais. Além do uso das diferenças centradas e avançadas de primeira e segunda ordem, proposto por Duarte (2010), propõe-se o uso das diferenças de terceira e quarta ordem.

Considere um sinal de voz no domínio wavelet  $Y[n] = \{y_0, \dots, y_{n-1}\}$ . Os coeficientes obtidos pelas diferenças finitas centradas e avançadas são representados respectivamente por  $c_i^j$  e  $a_i^j$ . O índice  $j$  representa a ordem das equações e o índice  $i$  representa uma posição arbitrária, no tempo discreto  $n$ , do sinal em processamento. Seguem abaixo as equações referentes a todas as diferenças finitas (SPERANDIO; MENDES; SILVA, 2003).

- Diferenças centradas de primeira a quarta ordem:

$$c_i^1 = \frac{1}{2}(y_{i+1} - y_{i-1}), \quad (22)$$

$$c_i^2 = y_{i+1} - 2y_i + y_{i-1}, \quad (23)$$

$$c_i^3 = \frac{1}{2}(y_{i+2} - 2y_{i+1} + 2y_{i-1} - y_{i-2}), \quad (24)$$

$$c_i^4 = y_{i+2} - 4y_{i+1} + 6y_i - 4y_{i-1} + y_{i-2}. \quad (25)$$

- Diferenças avançadas de primeira a quarta ordem:

$$a_i^1 = y_{i+1} - y_i, \quad (26)$$

$$a_i^2 = y_i - 2y_{i+1} + y_{i+2}, \quad (27)$$

$$a_i^3 = -y_i + 3y_{i+1} - 3y_{i+2} + y_{i+3}, \quad (28)$$

$$a_i^4 = y_i - 4y_{i+1} + 6y_{i+2} - 4y_{i+3} + y_{i+4}, \quad (29)$$

Desta forma, obtêm-se, respectivamente, os vetores de mesmo comprimento de  $Y$ :

$$C^1[n] = \{c_0^1, c_1^1, \dots, c_{n-1}^1\},$$

$$C^2[n] = \{c_0^2, c_1^2, \dots, c_{n-1}^2\},$$

$$C^3[n] = \{c_0^3, c_1^3, \dots, c_{n-1}^3\},$$

$$C^4[n] = \{c_0^4, c_1^4, \dots, c_{n-1}^4\},$$

$$A^1[n] = \{a_0^1, a_1^1, \dots, a_{n-1}^1\},$$

$$A^2[n] = \{a_0^2, a_1^2, \dots, a_{n-1}^2\},$$

$$A^3[n] = \{a_0^3, a_1^3, \dots, a_{n-1}^3\},$$

$$A^4[n] = \{a_0^4, a_1^4, \dots, a_{n-1}^4\},$$

todos no domínio wavelet.

Uma interpretação pertinente relativa às diferenças finitas seria que, quando se analisa um determinado coeficiente de um sinal de voz, levam-se em consideração informações contidas em uma “vizinhança” do mesmo. O tamanho dessa vizinhança varia de acordo com a ordem da equação de diferenças usada. Neste sentido, quanto maior o grau das equações de diferenças, maior a vizinhança onde se extrai as informações utilizadas na construção do respectivo coeficiente de filtro.

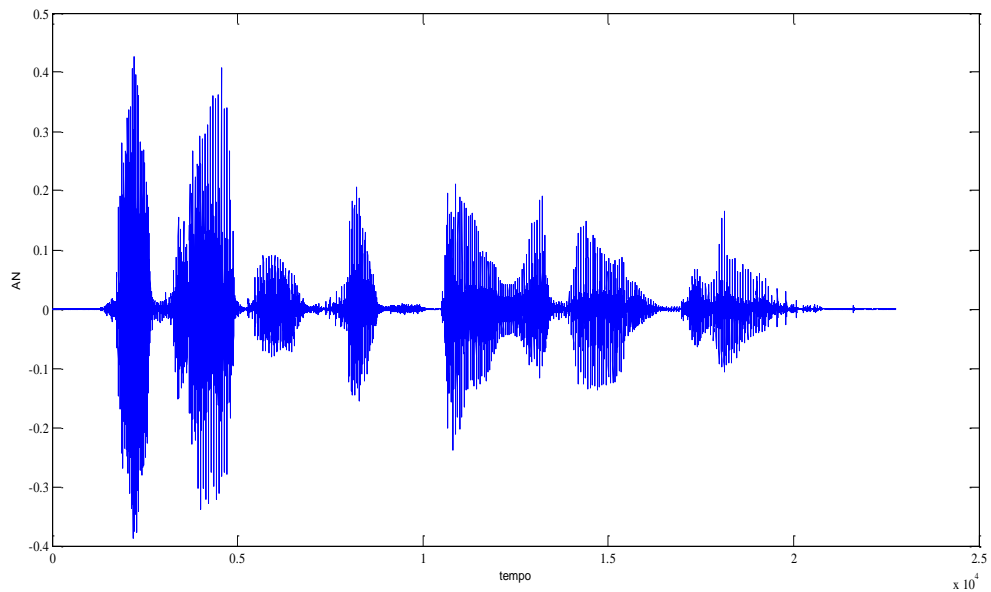
Por conta das características oscilatórias de um sinal de voz, as equações de diferenças não podem ser aplicadas diretamente para a redução de ruído. É evidente que, caso um coeficiente seja negativo, ao efetuar uma diferença entre outro coeficiente negativo, o mesmo pode mudar de sinal. Em um todo, este fato leva a inserção de fortes distorções no sinal processado, tornando-o ininteligível. Além disso, estas equações não estimam o ruído presente no sinal, apenas evidenciam o seu comportamento.

Durante o processamento, em trechos de silêncio não se faz necessária a utilização das diferenças de terceira e quarta ordem. Isto se justifica, pelo fato de que as equações de ordens mais elevadas têm o objetivo de minimizar perdas de coeficientes significativos na reconstrução do sinal. Neste sentido, em trechos considerados de silêncio aplicam-se apenas as diferenças centradas e avançadas de primeira e segunda ordem. Em contrapartida, quando o trecho em processamento é um trecho de voz, necessita-se de maior quantidade de informações para construção do filtro. Desta forma, para estas regiões calculam-se as diferenças centradas e avançadas de primeira, segunda, terceira e quarta ordem. A intenção é evitar perdas de regiões consideradas delicadas durante o processamento de um sinal de voz, como é o caso de regiões *unvoiced*.

Nas Figuras 9, 10, 11, 12 e 13 constam, respectivamente, o sinal de voz  $s(n)$  usado para as ilustrações, o sinal ruidoso  $y(n)$  contaminado com ruído colorido (ruído de carro), um

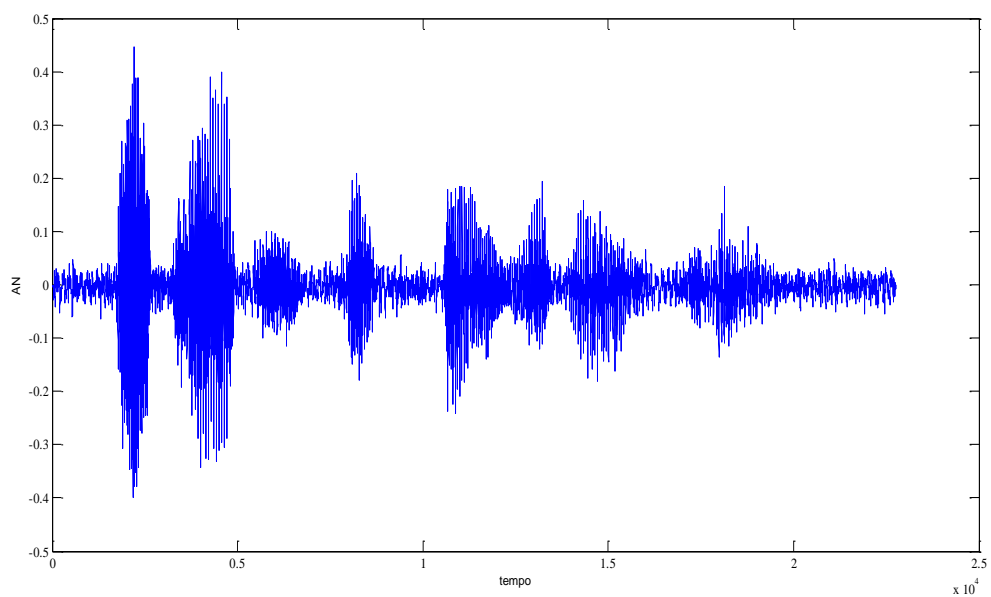
trecho vozeado no domínio wavelet, as curvas produzidas pelas diferenças centradas e as curvas produzidas pelas diferenças avançadas, quando aplicadas a este segmento.

Figura 9: Sinal limpo  $s(n)$ .



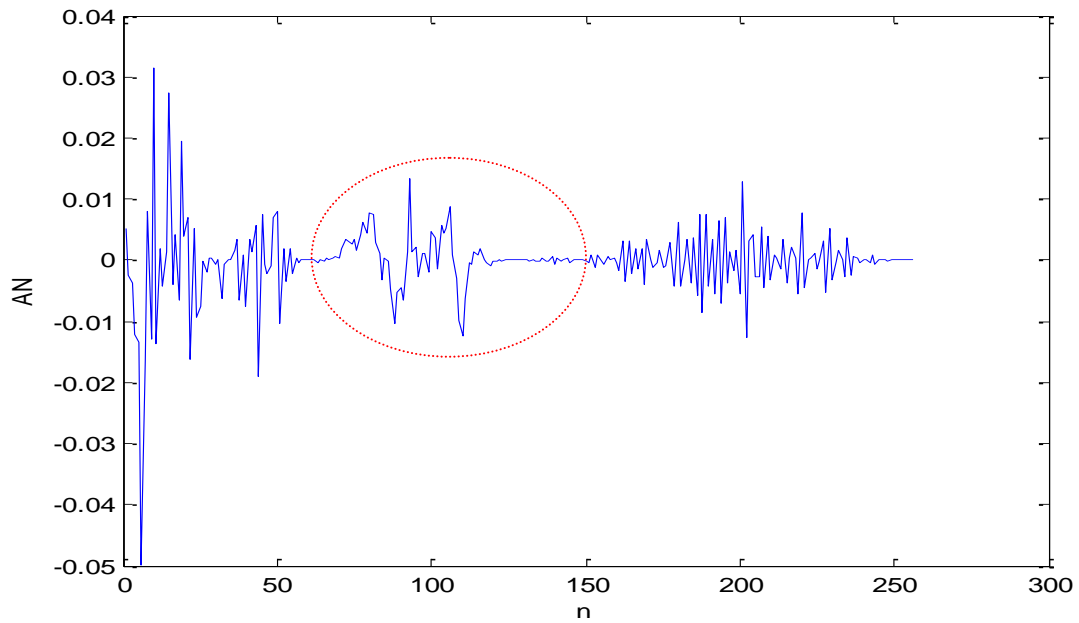
Fonte: Elaborado pelo próprio autor.

Figura 10: Sinal contaminado com ruído colorido  $y(n)$ .



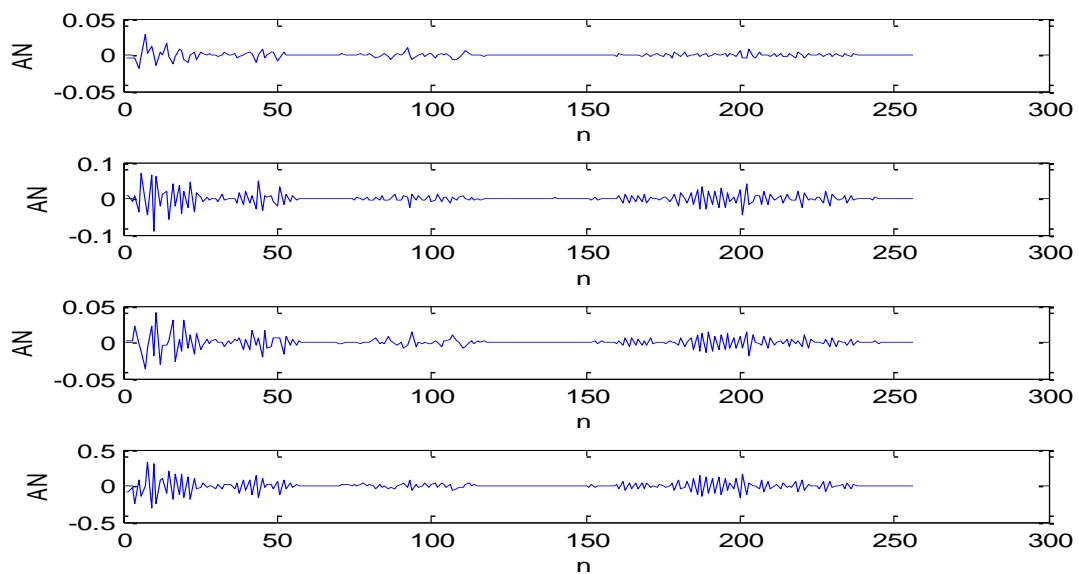
Fonte: Elaborado pelo próprio autor.

Figura 11: Trecho de voz do sinal  $Y[n]$ , no domínio wavelet, obtido a partir da contaminação do sinal limpo  $s[n]$  por ruído colorido.



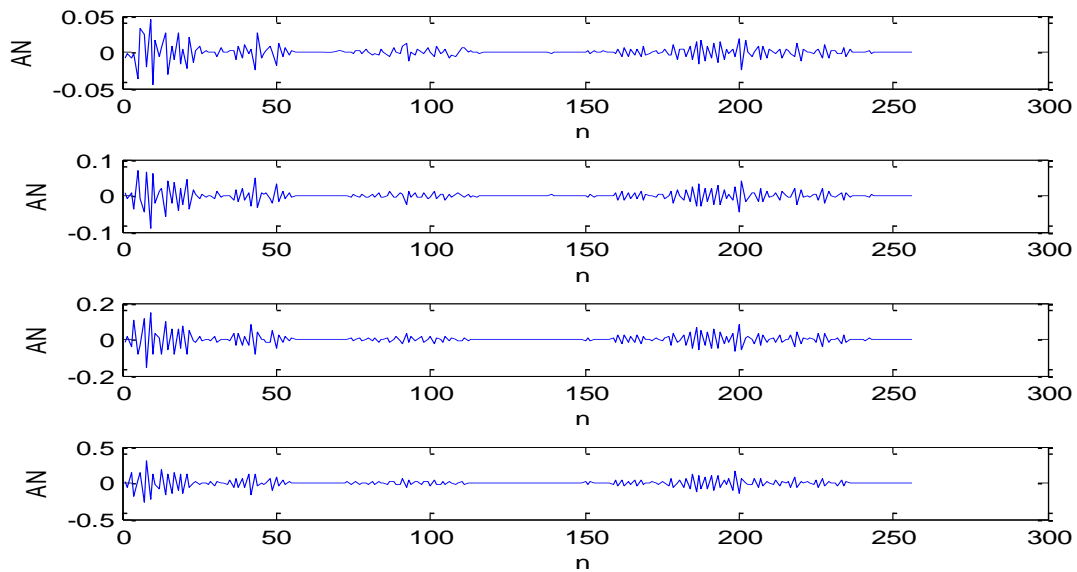
Fonte: Produzido pelo próprio autor.

Figura 12: Curvas das diferenças centradas de primeira a quarta ordem, respectivamente, quando aplicadas ao segmento visualizado na Figura 11.



Fonte: Elaborado pelo próprio autor.

Figura 13: Curvas das diferenças avançadas de primeira a quarta ordem, respectivamente, quando aplicadas ao segmento visualizado na Figura 11.



Fonte: Elaborado pelo próprio autor.

Fazendo uma análise do sinal no domínio wavelet (Figura 11) e das diferenças (Figuras 12 e 13), nota-se que as diferenças seguem as principais características do sinal ao qual são aplicadas: mantêm-se as características oscilatórias do sinal no domínio wavelet, ao mesmo tempo em que reduzem ou eliminam coeficientes irrelevantes. Nota-se também que na “região” destacada na Figura 11, onde vários coeficientes são nulos ou muito pequenos, ocorre uma forte atenuação, por parte das diferenças, com relação aos demais coeficientes. Este fato é melhor observado a partir das diferenças avançadas (ver Figura 13). Isto se deve principalmente ao fato de que, ao processar um determinado coeficiente, levam-se em consideração informações presentes nos coeficientes que estão próximos ao mesmo.

Outro fato importante é que, na primeira quarta parte do sinal no domínio wavelet, onde está concentrada a maioria dos coeficientes relevantes para o processo de reconstrução, as diferenças produzem um ganho na amplitude destes coeficientes. O que é muito útil para a construção de um filtro que tenha como função preservar estes coeficientes.

### 5.3 Estimação do ruído

Para estimar o ruído, utiliza-se o operador relação sinal/ruído *a priori* (SNRP) proposto por Soares et al. (2011). Este é aplicado a cada uma das diferenças finitas, calculadas para o sinal em processamento. Para um vetor qualquer de diferenças finitas  $D[n]$

no domínio wavelet, com coeficientes  $d_i$ ,  $i = 0, \dots, n - 1$ , define-se a função  $g(D)$  como segue:

$$g(d_i) = \frac{|d_i|}{\alpha + |d_i|}, \quad \text{para } i = 0, \dots, n - 2 \text{ e } g(d_{n-1}) = g(d_{n-2}), \quad (30)$$

onde  $\alpha \in ]0, 1[$ .

O operador  $g(D)$  faz a estimativa do ruído presente em cada coeficiente do vetor de diferenças  $D[n]$ , característico dos métodos baseados na SNR *a priori* (COHEN, 2004). Sendo assim, o uso deste operador para o tratamento de sinais de voz não causa distorções nos trechos de voz (SOARES, 2009).

Por fim, o sinal estimado é dado pela combinação polinomial da função dada na equação (30), quando aplicada às diferenças apresentadas na seção 5.2. Os detalhes desta operação serão tratados na próxima seção.

#### 5.4 Combinação polinomial para se estimar o sinal de saída

O processamento via combinação polinomial combina as curvas produzidas pela função  $g(D)$ , quando aplicada a cada uma das diferenças. O objetivo é reduzir o ruído e evitar distorções no sinal. O processamento usado em trechos de silêncio difere do usado em trechos de voz. Sendo assim, faz-se necessário o uso de um detector silêncio/voz.

Considere  $D$  como sendo o conjunto formado pelos vetores das diferenças centradas  $C^j = \{c_i^j \mid 0 \leq i \leq N - 1\}$  e avançadas  $A^j = \{a_i^j \mid 0 \leq i \leq N - 1\}$ , de ordem  $j$  ( $1 \leq j \leq 4$ ) e comprimento  $N - 1$ . Considere também  $E[n] = \{e_0, \dots, e_{N-1}\}$ , o conjunto imagem da função em questão.

A combinação polinomial proposta no trabalho consiste na introdução da função  $E_{cp} : D \rightarrow E[n]$  definida por  $E_{cp}(c_i^j) = e_i$  e  $E_{cp}(a_i^j) = e_i$ , para  $i = 0, \dots, N - 1$  e  $1 \leq j \leq 4$ , cujos elementos  $e_i$  são obtidos de acordo com a equação (31). Tal função é responsável pela estimação do sinal de saída, que posteriormente, recebe o ajuste sigmoidal abordado na seção 5.6. Para entender como  $N$  e  $n$  se relacionam, considere um sinal ruidoso de comprimento  $N = 2^n$  a ser processado:

$$y[n] = \left\{ y_{2k}, y_{2k+1} : 0 \leq k \leq \frac{2^n - 2}{2} \right\}.$$

Este sinal pode ser representado como segue:

$$y[n] = \{y_0, y_1, \dots, y_{2k}, y_{2k+1}, \dots, y_{2^{n-2}}, y_{2^{n-1}}\} = \{y_0, y_1, \dots, y_{2k}, y_{2k+1}, \dots, y_{N-2}, y_{N-1}\},$$

considerando que  $y \in \mathbb{R}^N$ .

Na prática, toma-se uma janela de comprimento  $N = 2^n$  do sinal  $y[n]$  a ser processado, verifica-se se a janela está em uma região de silêncio ou de voz. Para fins didáticos, cria-se uma variável binária denominada *decisão*. Se a janela estiver numa região de silêncio, atribui-se à variável o valor 0 (zero). Caso a região seja de voz, atribui-se à mesma o valor 1 (um). Sendo assim, os coeficientes  $e_i$  do conjunto  $E[n]$  são obtidos da seguinte forma:

$$e_i = \begin{cases} g(c_i^1) + g(a_i^1) + g(c_i^2) + g(a_i^2) & \text{se } \textit{decisão} = 0 \\ g(c_i^1) + g(a_i^1) + g(c_i^2) + g(a_i^2) + g((c_i^3)^2) + g((a_i^3)^3) + g((c_i^4)^3) + g((a_i^4)^3) & \text{se } \textit{decisão} = 1. \end{cases} \quad (31)$$

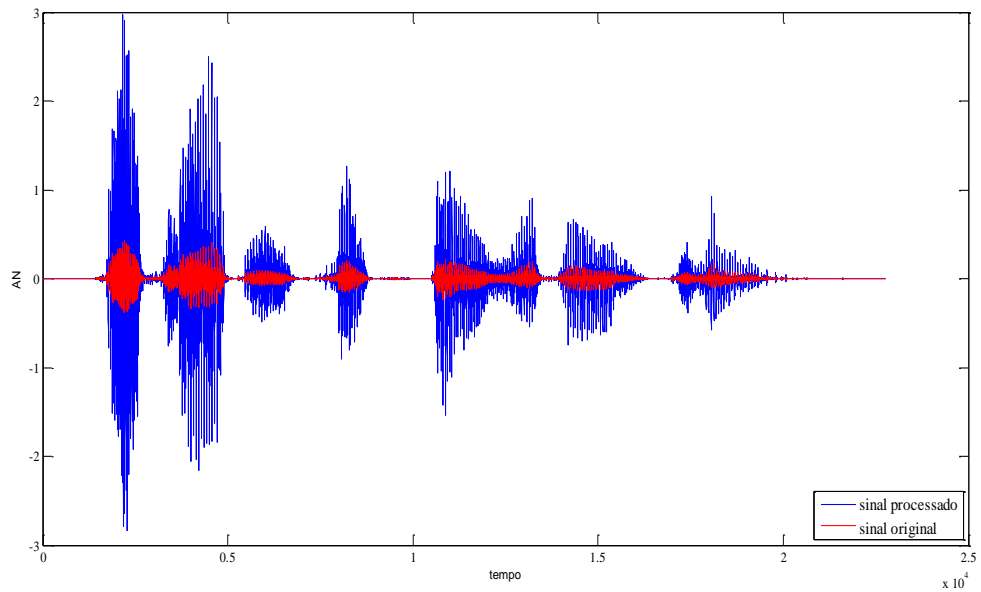
As potências aplicadas às diferenças de terceira e quarta ordem foram determinadas de modo empírico, com o objetivo de reduzir o ruído presente nos trechos de voz. Com isso, acrescentam-se às diferenças de primeira e segunda ordem apenas coeficientes realmente importantes ao processo de reconstrução. O objetivo é melhorar a qualidade da voz no sinal processado.

É evidente que, com esta operação, cada coeficiente recebe uma atenuação que depende apenas de seu próprio valor. Os valores destes coeficientes, em geral, são menores que um. Assim, quanto menor o valor do coeficiente, maior será a atenuação. Em contrapartida, feita a combinação conforme a equação (31), a amplitude da função  $E_{cp}$  será maior que um, extrapolando a amplitude do sinal original após o processamento.

A Figura 14 apresenta, em vermelho, uma representação do sinal de voz limpo  $s(n)$  e, em azul, uma representação do mesmo processado pela função  $E_{cp}$ . O processamento se faz pela multiplicação dos coeficientes da função  $E_{cp}$ , pelos respectivos coeficientes do sinal  $s[n]$ , no domínio wavelet. Já a Figura 15 trás uma representação do sinal de voz processado  $\hat{y}(n)$ , usando a função  $E_{cp}$  e a transformada wavelet inversa, a partir do sinal de voz limpo  $s(n)$ , contaminado com ruído colorido.

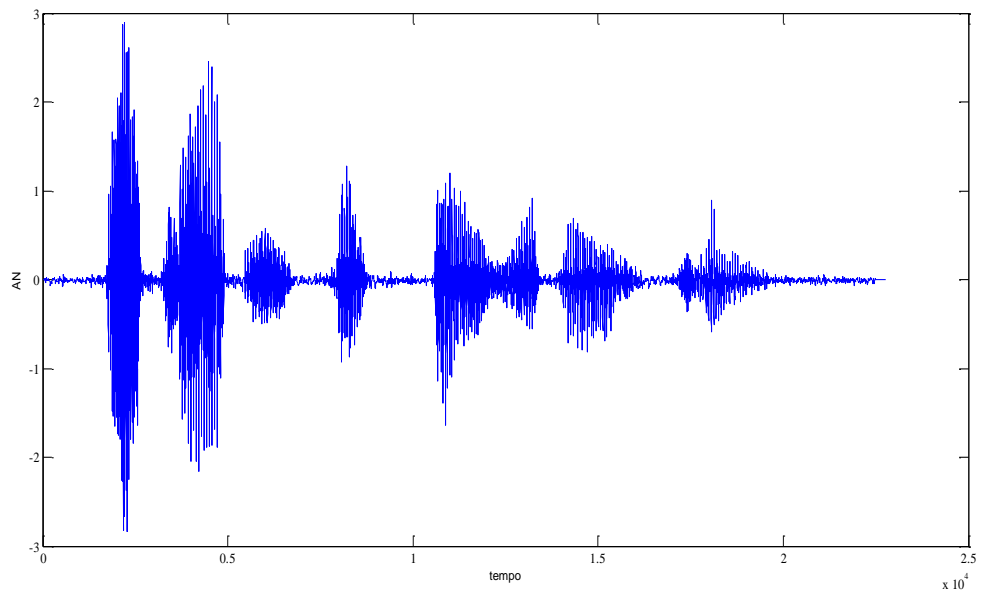
Analisando a Figura 15, verifica-se que há redução de ruído, porém, ressaltam-se os coeficientes referentes à voz. Este fato é essencial para não haver perda nas amplitudes intermediárias ou baixas do sinal no domínio do tempo, durante o processo de filtragem. Este fato será assunto das seções subsequentes.

Figura 14: Sinal processado  $\hat{y}(n)$  a partir do sinal original limpo  $s(n)$ , usando a função  $E_{cp}$  e a transformada wavelet inversa.



Fonte: Elaborado pelo próprio autor.

Figura 15: Sinal processado  $\hat{y}(n)$  a partir do sinal  $y(n)$  contaminado por ruído colorido, usando a função  $E_{cp}$  e a transformada wavelet inversa.



Fonte: Elaborado pelo próprio autor.

Uma desvantagem a se considerar neste processamento é o tratamento diferenciado para silêncio e voz. Este tipo de processamento pode gerar desconforto ao ouvinte, pois causa “estampidos” na audição do sinal (SOARES, 2009).

O uso das potências nas diferenças de terceira e quarta ordem se justifica pelo fato de que, com isso, a maioria absoluta dos coeficientes ruidosos é eliminada. Desta forma, acrescentam-se ao processamento somente coeficientes relevantes, minimizando os efeitos indesejáveis.

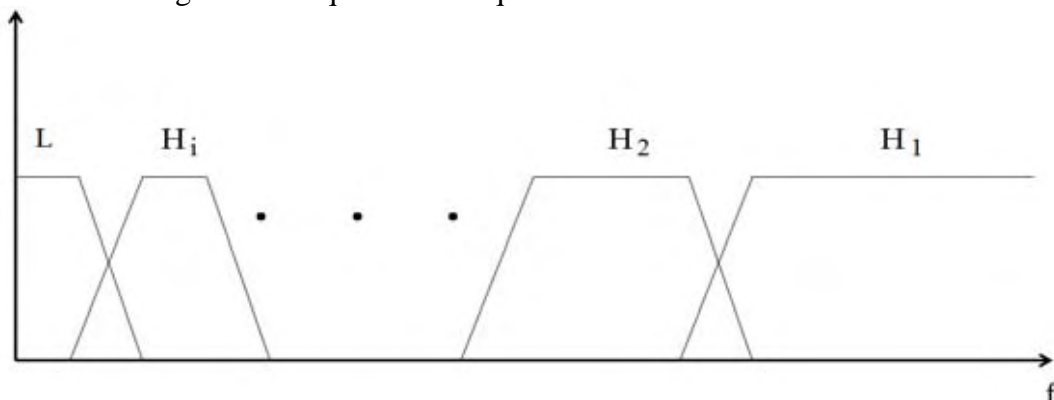
## 5.5 O operador *Dif*

Alguns métodos de redução de ruído que utilizam limiar, a fim de alcançar melhores resultados, diferenciam o tratamento para regiões consideradas *unvoiced* (DUARTE, 2005). Isto se deve ao fato de que estas regiões se confundem ao ruído, isto é, as chances das mesmas ficarem abaixo do limiar calculado é muito grande.

Apesar do método proposto não utilizar nenhum tipo de limiar, destinou-se uma atenção também aos trechos considerados *unvoiced*. Neste sentido, propõe-se um operador que tem como função minimizar estas perdas, denominado *Dif*.

Para entender o funcionamento deste operador, considera-se o esquema de um banco de filtros (ver Figura 16), onde o sinal no domínio wavelet pode ser dividido em diferentes faixas de frequência. Os sons *unvoiced* estão localizados em quase todo o espectro, reconhecidos com o aumento de energia em altas frequências quando comparados às regiões *voiced* (DUARTE, 2005). O operador proposto *Dif* realiza um tratamento específico nas altas frequências, tendo como objetivo eliminar ruído ao mesmo tempo em que evidencia as regiões *voiced* e *unvoiced*. Este processo pode ser encarado como uma pré-filtragem.

Figura 16: Esquema de frequências de um banco de filtros.



Fonte: Elaborado pelo próprio autor.

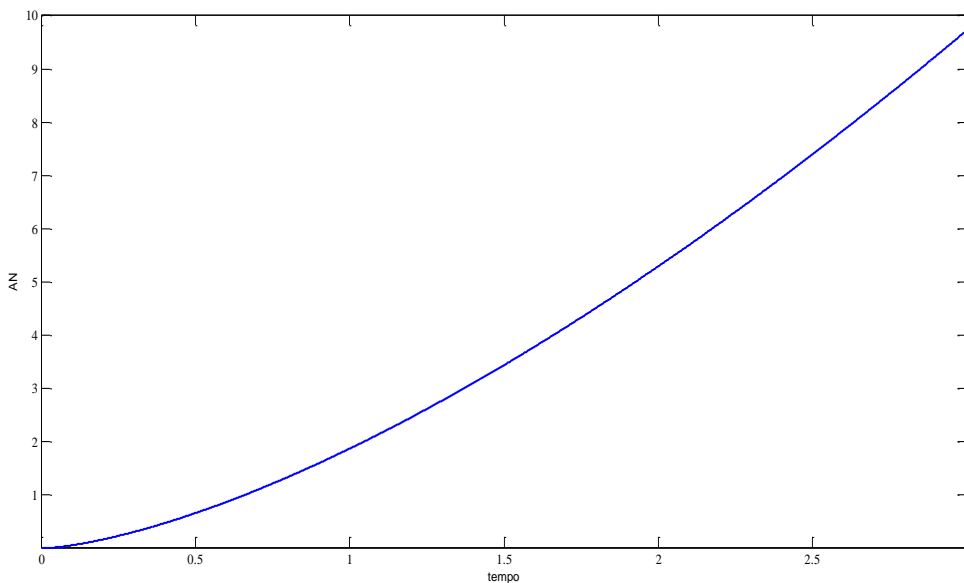
Em particular, divide-se o sinal de voz no domínio wavelet em três faixas diferentes e equivalentes:  $L$ ,  $H_2$  e  $H_1$ . O operador  $Dif$  é aplicado sobre os coeficientes  $e_i$  obtidos do processo de combinação polinomial conforme a equação (32). No Apêndice D, é apresentada uma explicação mais detalhada sobre a razão da expressão do operador  $Dif$ .

$$Dif(e_i) = \begin{cases} \sqrt{\frac{7}{2}} \cdot (e_i)^3 & \text{se } e_i \text{ está contido em } H_2 \text{ ou } H_1 \\ e_i & \text{, caso contrário.} \end{cases} \quad (32)$$

Em outras palavras, os coeficientes da primeira terça parte do espectro não serão alterados, pois eles já foram trabalhados pelos processos anteriores. Este operador propicia uma maior redução de ruído, minimizando as perdas de coeficientes relevantes. Na Figura 17 é mostrado o gráfico da função definida pela equação (32). Note que os coeficientes com amplitude abaixo de 0,5 receberão uma leve atenuação, enquanto que os coeficientes com amplitude superior receberão um ganho em amplitude.

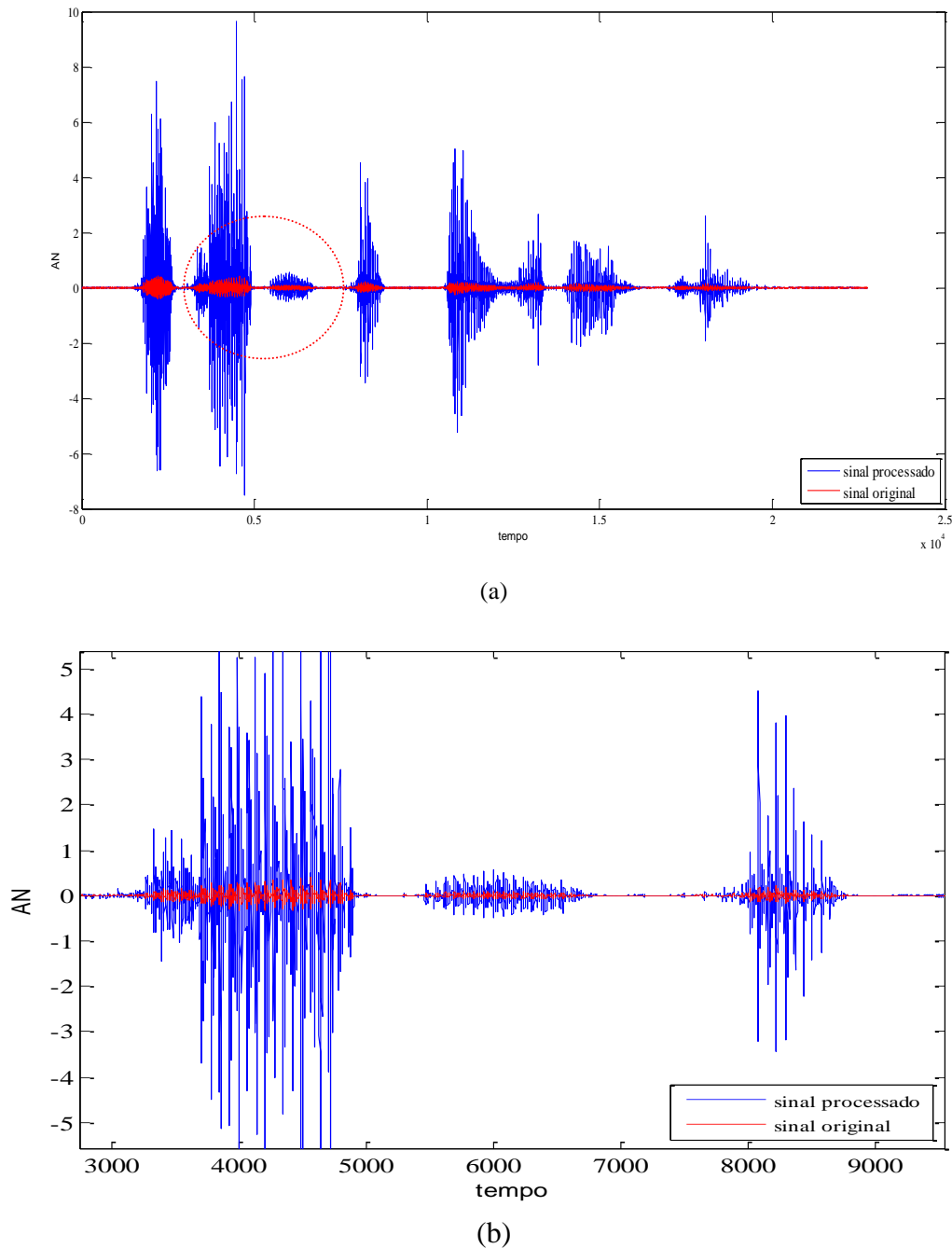
É evidente que este processamento causará um aumento demasiado na amplitude do sinal processado. No entanto, justifica-se seu uso pelo fato de evidenciar regiões *unvoiced* e *voiced*, ver Figura 18 (a) e 18 (b). Outro fato relevante, é que este operador atua de maneira equivalente em todas as janelas, evitando ‘estampidos’, incômodos ao ouvinte.

Figura 17: Gráfico do operador  $Dif$ .



Fonte: Elaborado pelo próprio autor

Figura 18: Sinal processado  $\hat{y}(n)$  a partir do sinal original limpo  $x(n)$ , usando o operador *Dif* e a transformada wavelet inversa (a); Zoom na região demarcada (b).



Fonte: Elaborado pelo próprio autor.

Analisando a Figura 18 (b), o primeiro trecho de voz é um trecho *unvoiced*. Note que esta região ficou bem evidenciada, quando comparada ao sinal original. Este fato é muito importante para garantir uma forte redução de ruído, aliada ao mínimo de distorção possível, durante o processo de filtragem. O Apêndice C destaca a contribuição do operador *Dif* no processamento proposto, nele consta uma análise da influência deste operador.

## 5.6 Obtenção do filtro de redução de ruído

A fim de preservar a amplitude original do sinal, o filtro  $H[n] = \{h_0, h_1, \dots, h_{n-1}\}$ , utilizado no processo de filtragem, é adquirido via ajuste sigmoidal dos coeficientes do sinal estimado. Os coeficientes de  $H$  são obtidos conforme a equação (33) (SOARES 2009; SOARES et al., 2011):

$$h_i = \left| \frac{1-e^{-\sigma e_i}}{1+e^{-\sigma e_i}} \cdot \frac{1-e^{\sigma e_i}}{1+e^{\sigma e_i}} \right|. \quad (33)$$

O parâmetro  $\sigma$  controla a inclinação das duas sigmoides e é calculado conforme a seguinte equação (SOARES et al., 2011):

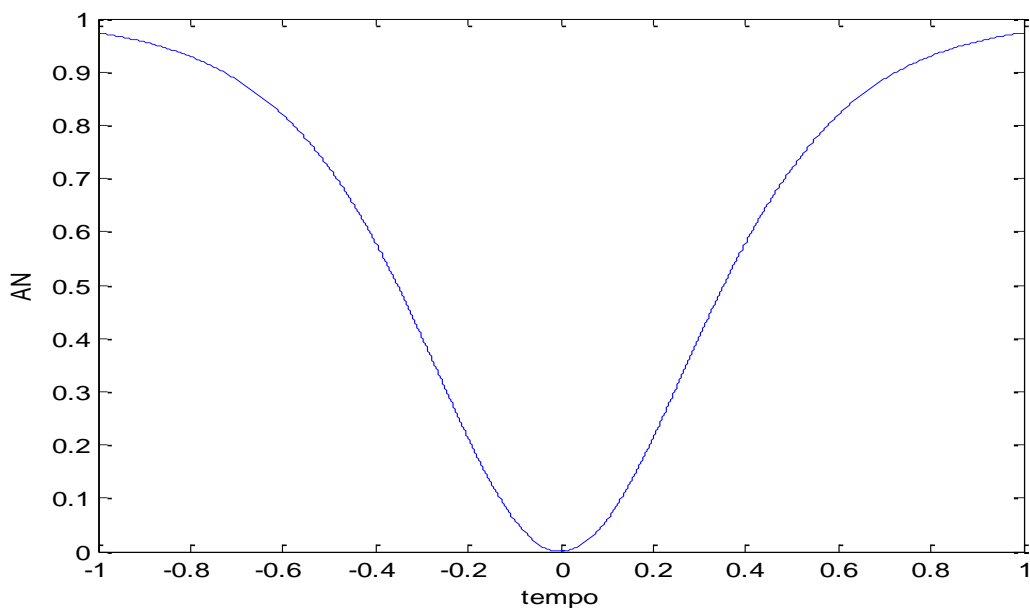
$$\sigma = \beta \cdot \tau. \quad (34)$$

sendo  $\tau$  o desvio padrão da janela em processamento e  $\beta$  um valor baseado na potência do ruído.

O objetivo deste ajuste é combinar os processamentos propostos pelas equações de diferenças apresentadas na seção 5.2, mantendo a amplitude do sinal original.

A equação (33) gera os coeficientes do filtro utilizado para a redução do ruído, e, normaliza a amplitude do sinal processado no intervalo  $[-1,1]$ . Na Figura 19, é mostrado o gráfico gerado pela combinação das duas funções sigmoides envolvidas na equação (33).

Figura 19: Curva gerada pela combinação das duas funções sigmoides na equação.



Fonte: Elaborada pelo próprio autor.

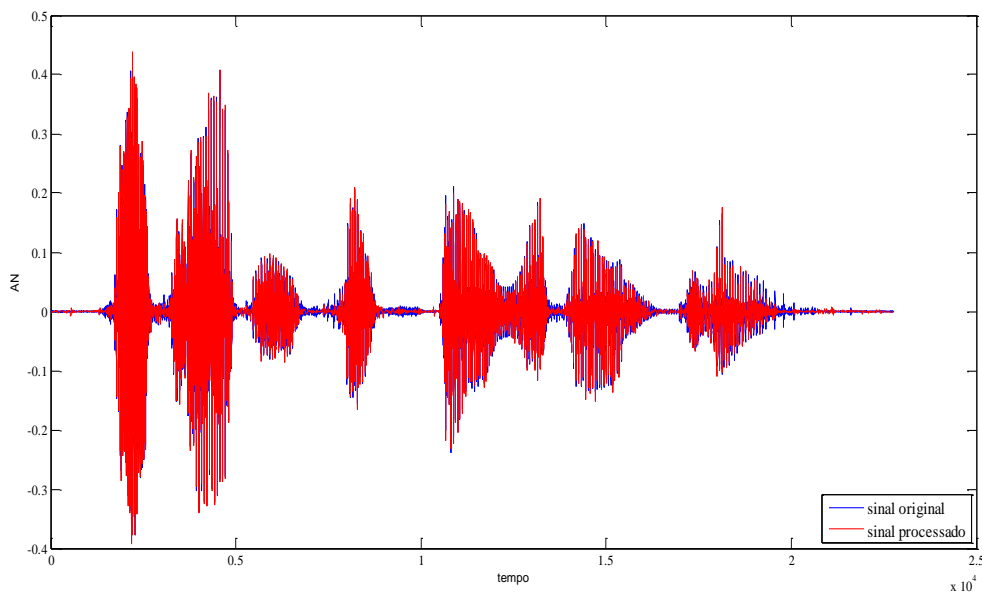
Para gerar a curva da Figura 19, admitiu-se  $\sigma = 5$ . Vale ressaltar que, quando o ajuste sigmoidal é aplicado ao sinal de voz, a curva de atenuação não é tão suave como a apresentada na Figura 19. Deve-se isto ao fato de que um sinal de voz não possui características lineares como o vetor tempo utilizado para gerar esta curva.

O processo de filtragem se faz através do produto entre os coeficientes  $h_i$  do filtro  $H$ , construído a partir do método proposto, pelos correspondentes coeficientes  $y_i$  do sinal ruidoso  $Y[n]$ , no domínio wavelet. Obtém-se desse modo, o sinal processado  $\hat{Y}[n] = \{\hat{y}_0, \hat{y}_1, \dots, \hat{y}_{n-1}\}$  de comprimento  $n$ . Cada coeficiente  $\hat{y}_i$  é obtido conforme a equação (35):

$$\hat{y}_i = y_i \cdot h_i. \quad (35)$$

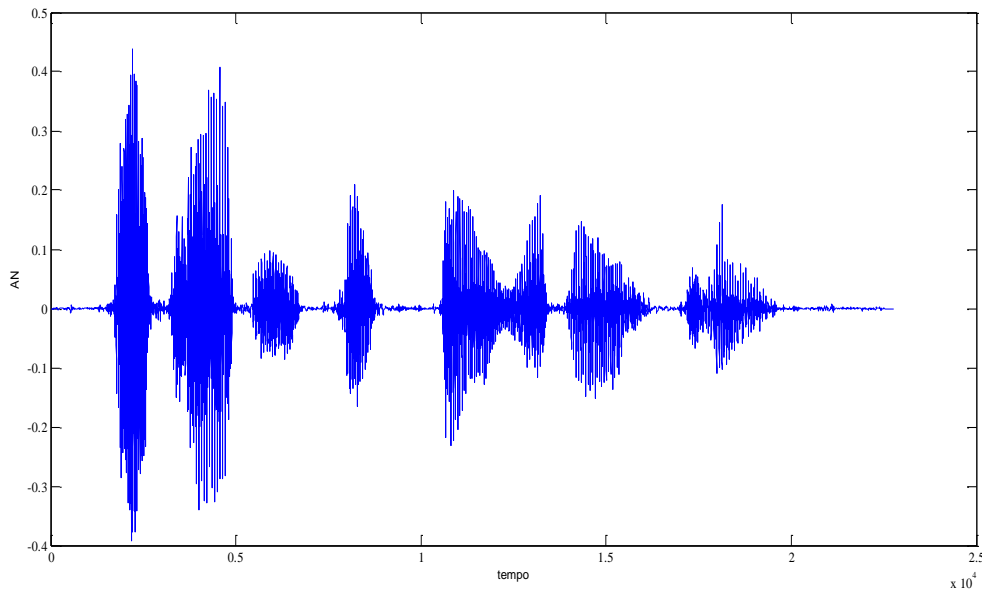
Na Figura 20 apresenta-se, em azul, uma representação do sinal de voz limpo  $s(n)$  e, em vermelho, uma representação do sinal processado  $\hat{y}(n)$  com ruído colorido. Observando esta figura, verifica-se que o método proposto é eficiente na redução de ruído. Além disso, efetua uma redução uniforme ao longo de todo o sinal.

Figura (20): Sinal de voz limpo  $s(n)$  e sinal processado  $\hat{y}(n)$  com ruído colorido, usando o método proposto e a transformada wavelet inversa.



Fonte: Elaborado pelo próprio autor.

Figura 21: Sinal processado  $\hat{y}(n)$  com ruído colorido, usando o método proposto e a transformada wavelet inversa.



Fonte: Elaborado pelo próprio autor.

Para concluir a análise, a Figura 21 apresenta somente o sinal processado  $\hat{y}(n)$  com ruído colorido.

Outro fato importante é que a amplitude do sinal é conservada. Embora haja diferenças entre os dois sinais, essas diferenças são mínimas e não representam fortes distorções no sinal processado.

A combinação polinomial proposta neste trabalho foi obtida após exaustivos testes, obtendo discussões e resultados satisfatórios. O que diferencia este método, em relação aos demais métodos de redução de ruído no domínio wavelet, é: a construção de um filtro que atua no sinal levando em consideração informações contidas em cada coeficiente e numa vizinhança do mesmo; a aplicação de um operador que não atua sobre todo o sinal no domínio wavelet, no caso o operador *Dif*; e o fato de não haver nenhum tipo de limiar durante a realização da filtragem.

Pode-se afirmar que o método aqui apresentado contribui para a melhoria do método proposto por Soares et al. (2011) nos seguintes aspectos:

- Estende-se o conceito de observar o ruído presente em cada coeficiente do sinal ruidoso, durante a construção dos respectivos coeficientes de filtro, para uma vizinhança de cada coeficiente. O tamanho de tal vizinhança é determinado pelo grau das equações de diferenças finitas.

- O operador dedicado à estimação do sinal não é aplicado diretamente ao sinal ruidoso no domínio wavelet, e sim ao sinal obtido pelas diferenças.
- Propõe-se um operador que tem como função minimizar as perdas de amplitudes do sinal durante o processamento.

Sobre a proposta de Duarte (2010), o método proposto se mostra vantajoso por aumentar a ordem das diferenças finitas usadas na obtenção do filtro, melhorando a estimação do ruído, por considerar regiões de silêncio e voz separadas e, também, pelo tratamento *voiced/unvoiced* dos sinais processados. Além disso, Duarte (2010) não faz combinação polinomial com as diferenças finitas e sim uma combinação linear, através de uma soma ponderada.

No capítulo seguinte, realiza-se a apresentação e discussão dos resultados obtidos pelo método proposto, aplicando-o a vários sinais de voz contaminados por diferentes tipos de ruído.

---

## 6 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Neste capítulo, são apresentados os resultados das implementações do método proposto e as avaliações dos sinais processados. Para constatar a eficiência do método, verificaram-se os níveis de redução de ruído e de distorção usando duas medidas objetivas de qualidade, a SNR (*Signal to Noise Ratio*) e a PESQ (*perceptual evaluation of speech quality*).

A SNR pode ser calculada de acordo com a equação (36). Toma-se a razão entre um segmento de voz e um segmento de silêncio de cada sinal da seguinte forma (DELLER; PROAKIS; HANSEN, 1993):

$$\text{SNR} = 10 \log_{10} \left( \frac{\sum_{i=0}^{N-1} x_i^2}{\sum_{i=0}^{N-1} w_i^2} \right), \quad (36)$$

sendo  $N$  o número de amostras do segmento escolhido. Neste trabalho, com base nos sinais utilizados, quando possível adotou-se  $N = 5000$ , caso contrário o comprimento do segmento utilizado foi de 1000 amostras. Os valores de SNR obtidos são resultantes da média das SNR's calculadas para três segmentos distintos.

Após a obtenção do sinal processado  $\hat{y}(n)$  e de sua respectiva SNR, cabe destacar as seguintes observações (SOARES, 2009):

- Caso, após o processamento, a SNR do sinal processado seja muito baixa em relação à SNR do sinal limpo (original), significa que houve pouca redução de ruído.

- Se, após o processamento, a SNR do sinal processado for muito maior que a SNR do sinal limpo, isto implicará uma distorção nos trechos de voz do mesmo. Deve-se isto a uma possível redução exagerada do ruído e a confusão de coeficientes de voz com ruído.

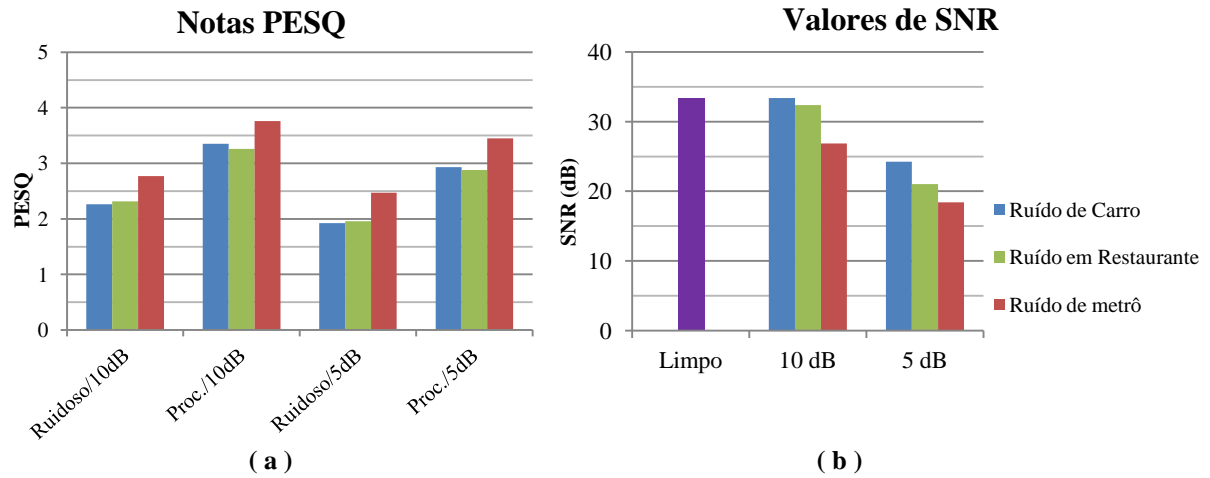
Proposta pela ITU, (2001) (*International Telecommunications Union*), a avaliação realizada pela PESQ é baseada em características psicoacústicas do ouvido humano. Representa um modelo cognitivo e os resultados são expressos por meio da avaliação MOS (*Mean Opinion Score*). A pontuação PESQ varia entre 1 e 5. Resultados considerados bons são aqueles cuja nota PESQ é maior ou igual a 3, caso contrário, o sinal será considerado de baixa qualidade (BEERENDS et al., 2002).

## 6.1 Testes Computacionais

Durante os testes foram utilizados treze sinais de voz, com voz masculina e feminina, nos idiomas inglês e português, todos sugeridos pela ITU-T (*Test Signals for Telecommunication Systems*). As simulações foram realizadas no ambiente MATLAB e a função wavelet utilizada na aplicação da DWT foi a wavelet de Daubechies de ordem 10 (db10). Foram utilizados arquivos no formato *wav*, com amplitude normalizada no intervalo (-1,1). A técnica de janelamento foi empregada utilizando a janela de *Hanning* com 50% de sobreposição entre os segmentos. O comprimento de cada janela é de 256 amostras, permitindo a obtenção de até 8 níveis de resolução na decomposição do sinal.

A contaminação dos sinais originais foi feita com três tipos de ruído colorido, presentes em situações reais: ruído de carro, em restaurante e de metrô. Os sinais foram analisados com dois níveis de SNR, 5dB e 10dB. A Figura 22 (a) apresenta as médias das notas PESQ obtidas para os sinais ruidosos e processados, assim como a Figura 22 (b) exibe as médias de SNR dos sinais limpos e processados. Na Tabela 3, é feita uma análise das variações das SNR dos sinais processados, quando comparados aos valores dos sinais limpos. Um exemplo mais detalhado, visualizando as formas de ondas dos sinais limpos e processados, para um sinal em particular, é abordado no Apêndice B.

Figura 22: Resultados obtidos por meio de avaliações objetivas.



Fonte: Elaborado pelo próprio autor.

Tabela 3: Análise da variação dos valores de SNR dos sinais processados, quando comparados ao valor da média das SNR dos sinais limpos.

	10dB (%)	5dB (%)
Ruído de Carro	99,444	71,647
Ruído em restaurante	96,517	61,402
Ruído de Metrô	79,337	53,688

Fonte: Elaborado pelo próprio autor.

Analisando os valores das SNR apresentados na Figura 22 (b), e, levando em consideração as observações feitas anteriormente, verifica-se que o método proposto reduz significativamente o ruído presente no sinal sem inserir fortes distorções. Os valores de SNR dos sinais processados estiveram sempre próximos aos valores dos sinais originais. Observa-se que os valores apresentados na Tabela 3, para os sinais com 10 dB de contaminação, ficam próximos de 100%. O que representaria uma situação ideal, pois se assim fosse, as SNR dos sinais processados seriam iguais as SNR dos sinais limpos, indicando um ótimo desempenho do método na redução do ruído.

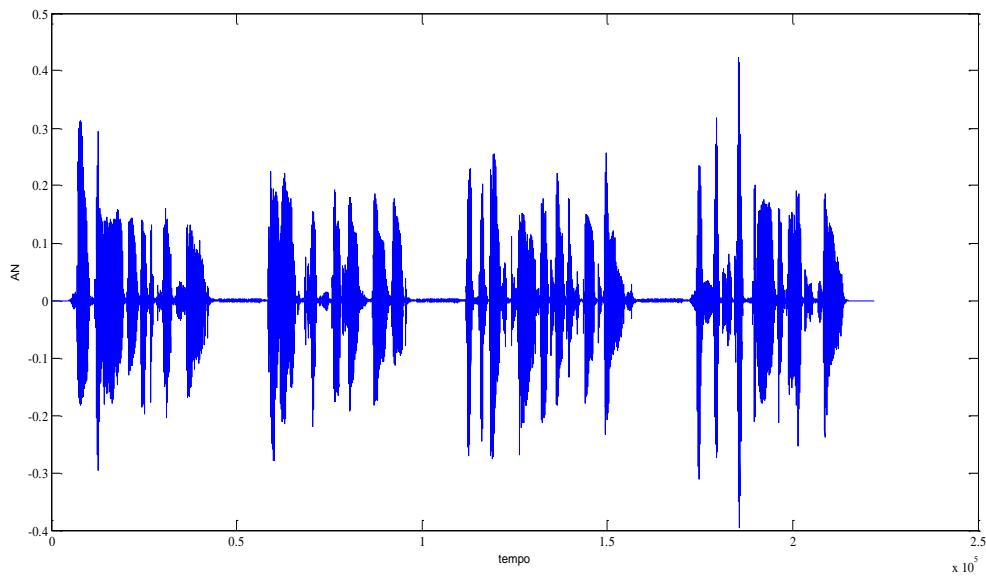
Como a SNR é uma medida de qualidade objetiva que serve para avaliar o nível do ruído presente em um sinal, ela não mede a inteligibilidade dos sinais processados. Para isso, devem-se levar em consideração as notas PESQ dos mesmos. Sendo assim, analisando as notas PESQ apresentadas na Figura 22 (a), constata-se que não houve distorções nos sinais processados. Além disso, estes sinais podem ser considerados de boa qualidade do ponto de vista auditivo, pois em média, obtiveram notas muito próximas ou superiores a 3.

Para ilustrar o processamento proposto, segue abaixo as formas de ondas de dois sinais em particular, um em voz feminina e outro em voz masculina. A ordem é a seguinte: sinal limpo, contaminados (para os três tipos de ruído) e processados.

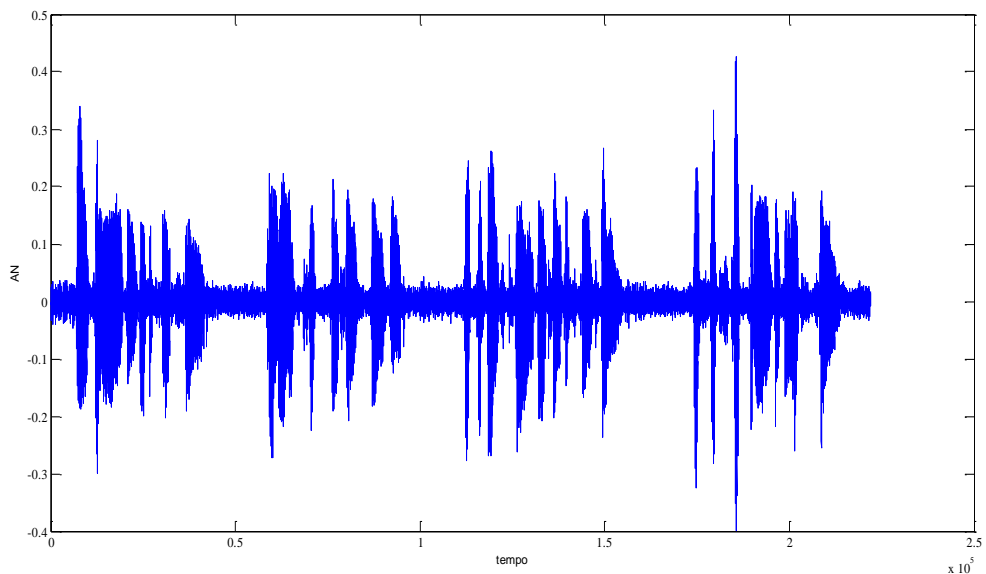
O primeiro sinal apresentado, na voz feminina, modela a seguinte frase:

Sinal F:  $\left\{ \begin{array}{l} \textit{Heave the line over the port side. A lathe cuts and trims any wood.} \\ \textit{It's a dense crowd in two distinct ways. His hip struck the knee} \\ \textit{of the next player.} \end{array} \right.$

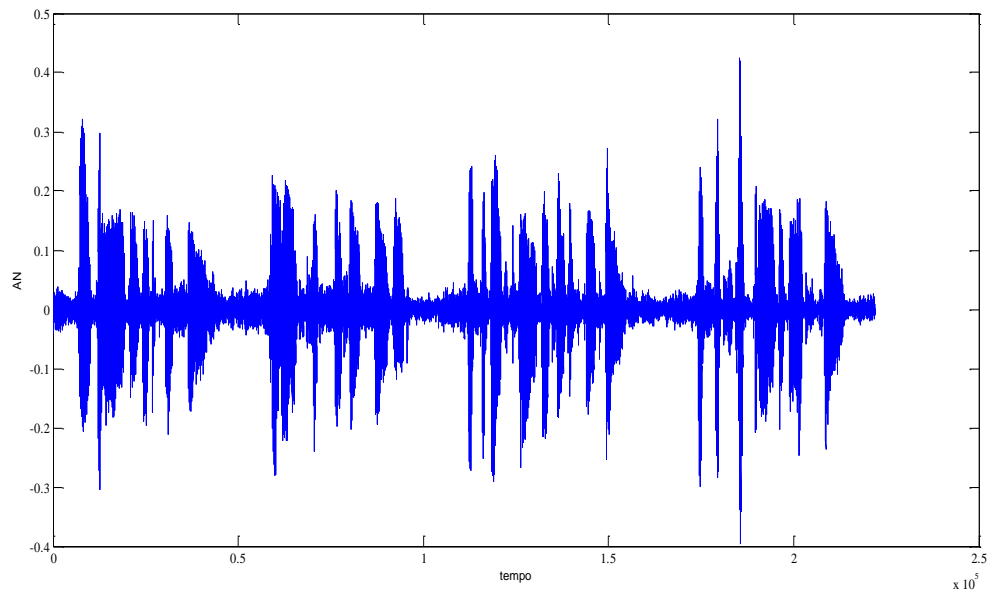
Figura 23: Forma de onda do Sinal F limpo (a); contaminado com ruído de carro (b); contaminado com ruído em restaurante (c); contaminado com ruído de metrô (d).



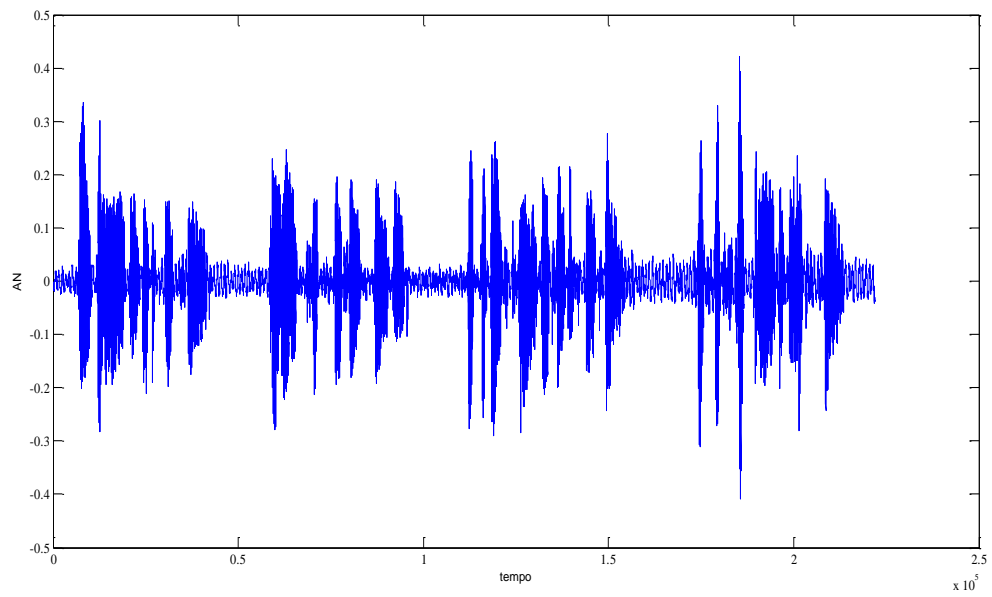
(a)



(b)



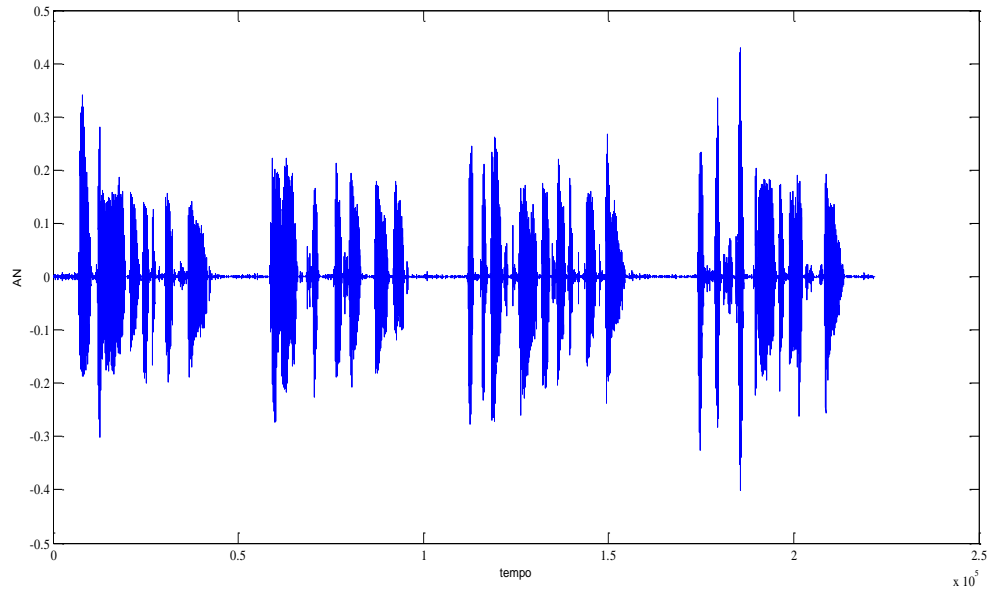
(c)



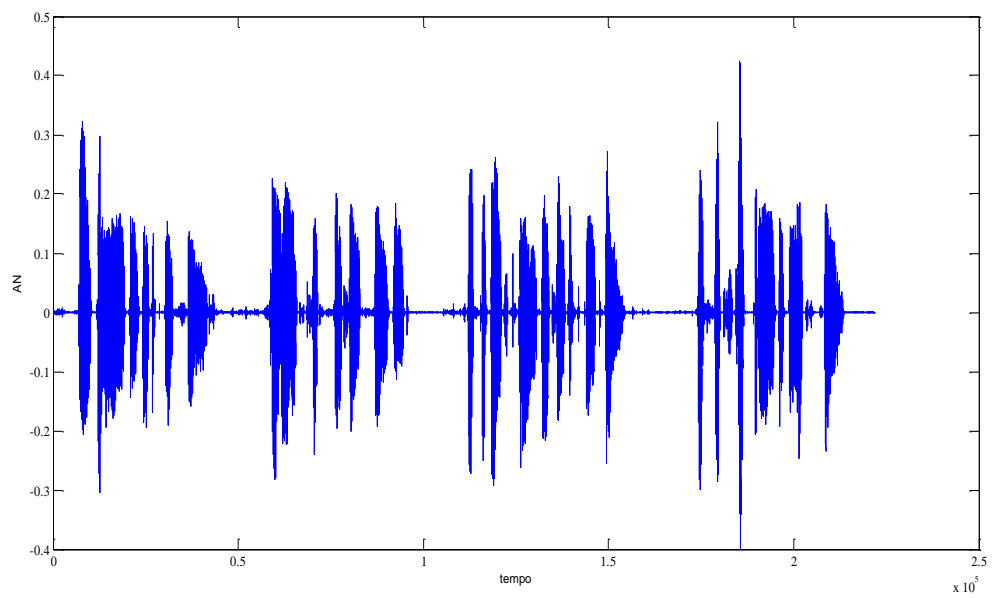
(d)

Fonte: Elaborado pelo próprio autor.

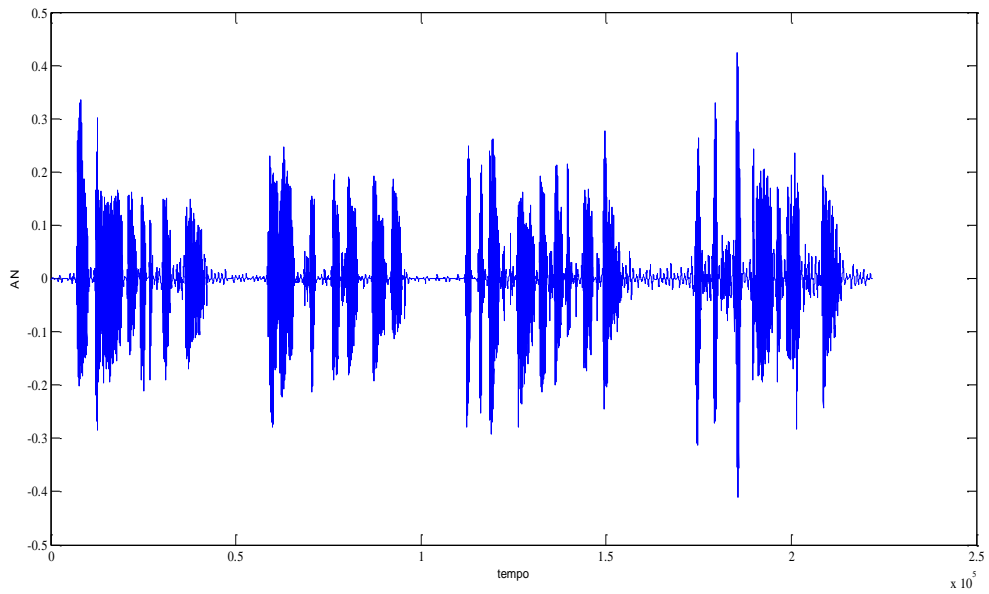
Figura 24: Forma de onda dos sinais processados, a partir do Sinal F limpo, contaminado com: ruído de carro (a), ruído em restaurante (b), ruído de metrô (c).



(a)



(b)



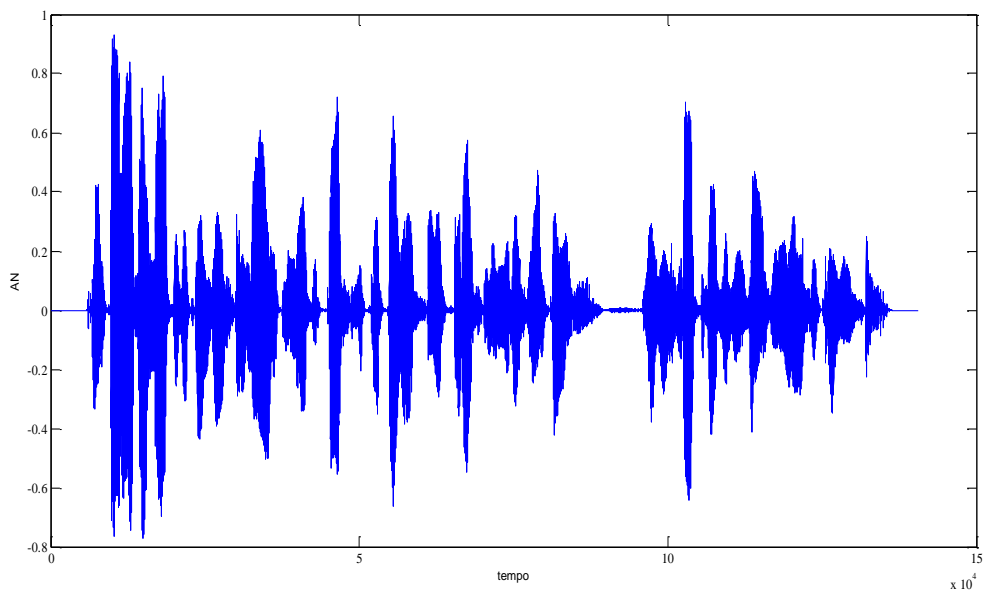
(c)

Fonte: Elaborado pelo próprio autor.

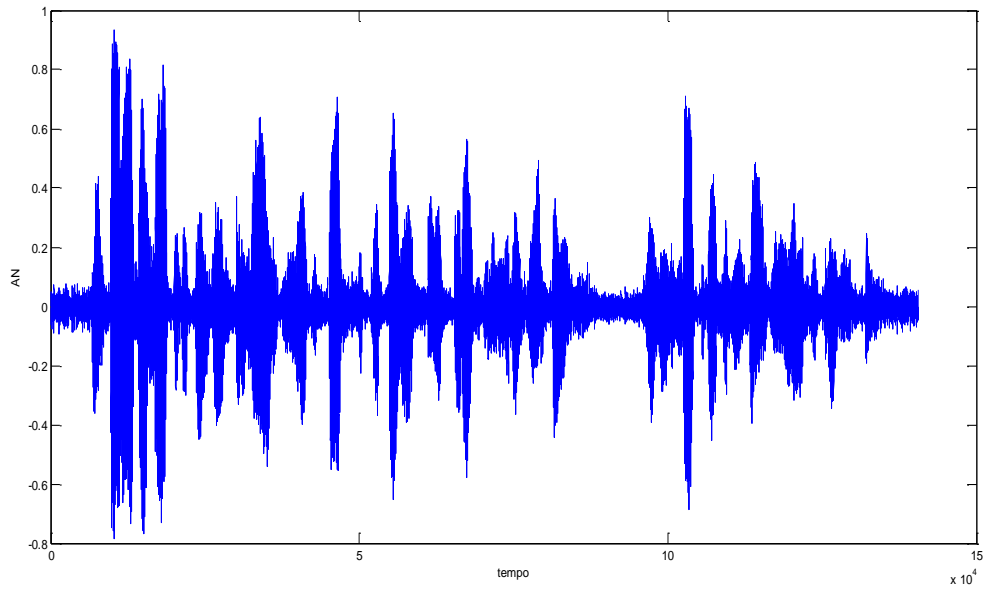
O segundo sinal, na voz masculina, modela a seguinte frase:

Sinal B: *Comparado ao método convencional, o novo teste de propagação oferece inúmeras vantagens, diz o chefe da divisão de meios de transmissão".*

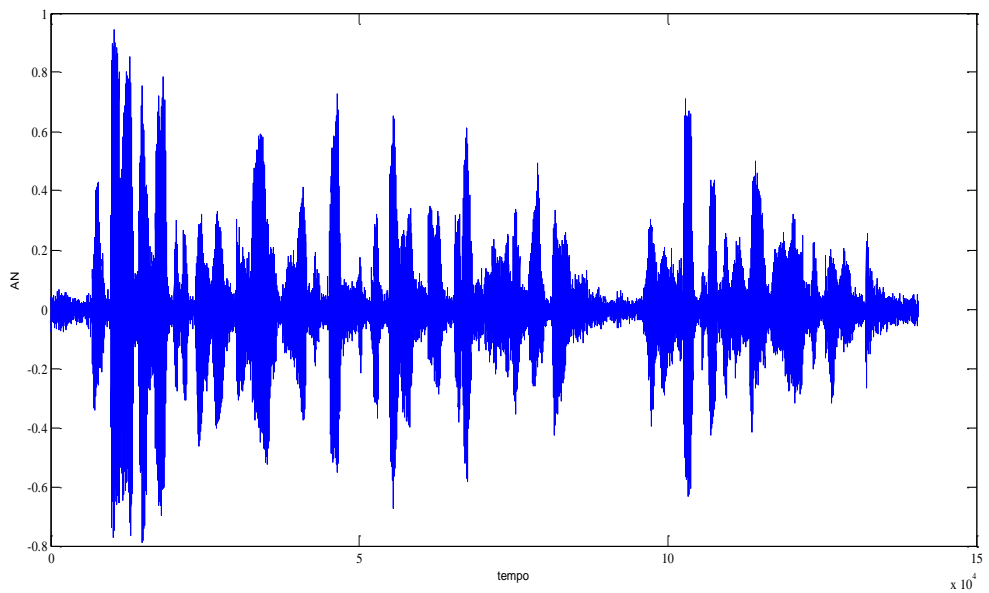
Figura 25: Forma de onda do Sinal B limpo (a); contaminado com ruído de carro (b); contaminado com ruído em restaurante (c); contaminado com ruído de metrô (d).



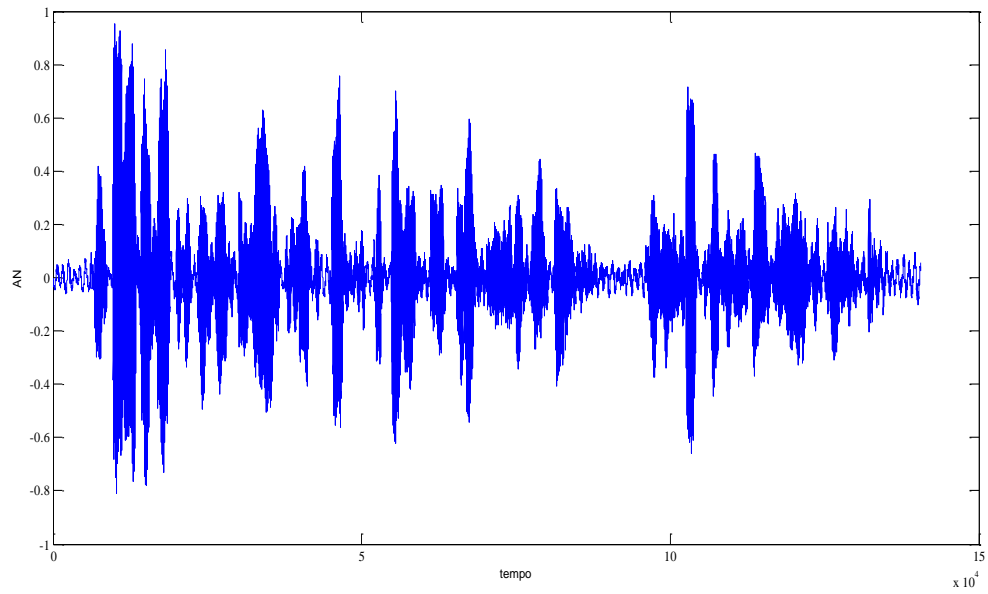
(a)



(b)



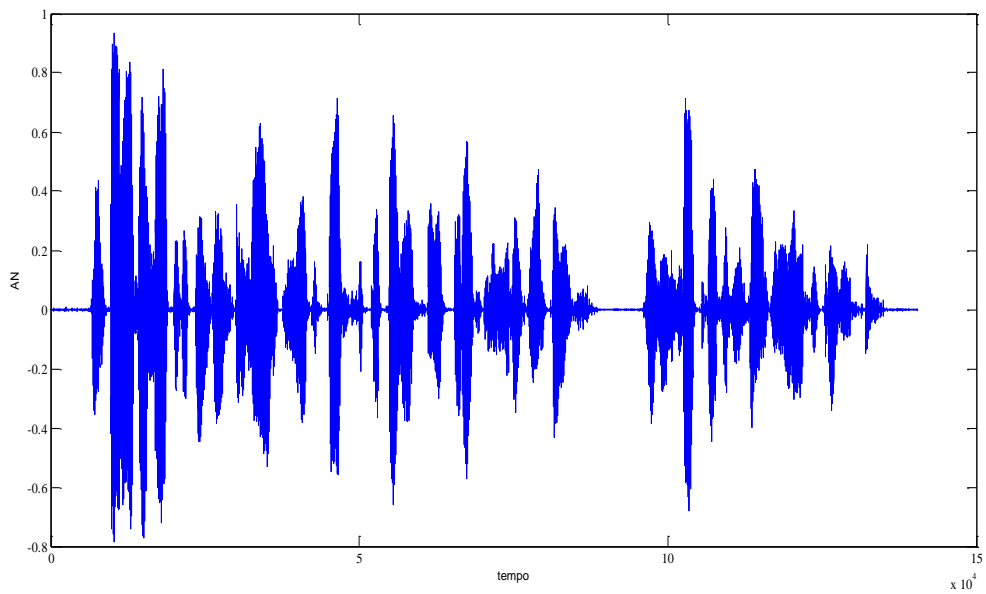
(c)



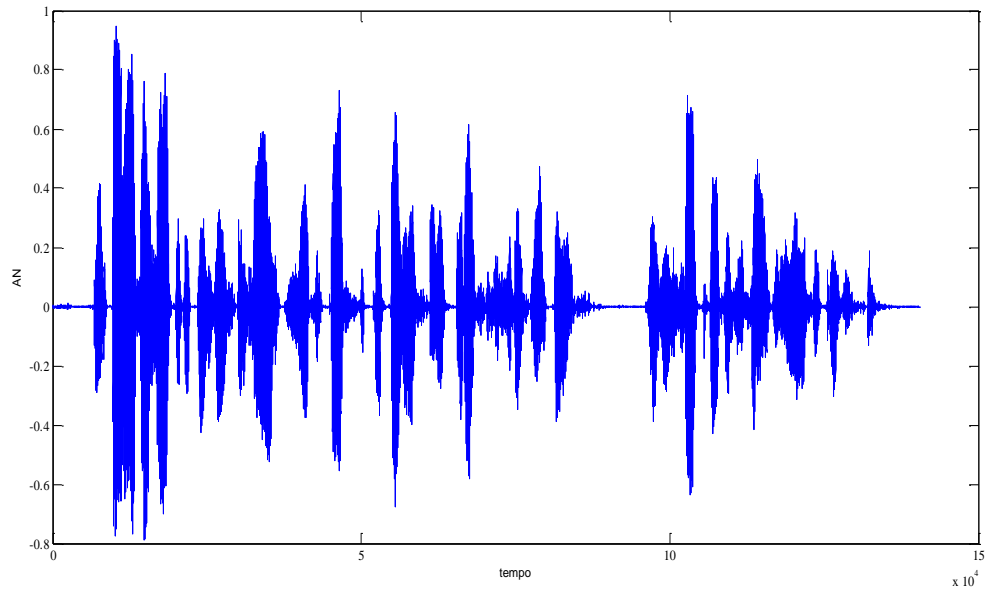
(d)

Fonte: Elaborado pelo próprio autor.

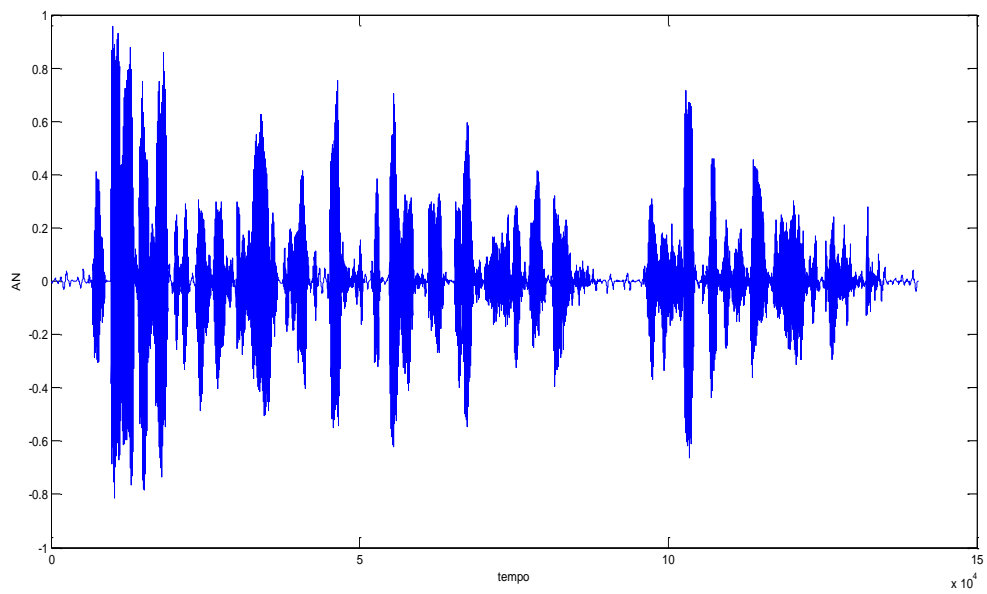
Figura 26: Forma de onda dos sinais processados, a partir do Sinal B limpo, contaminado com: ruído de carro (a), ruído em restaurante (b), ruído de metrô (c).



(a)



(b)



(c)

Fonte: Elaborado pelo próprio autor.

Para uma análise mais rigorosa, a Tabela 4 trás os resultados obtidos, para cada sinal utilizado nas simulações, tendo como referência a SNR. Constan na mesma, o gênero, idioma, nível de contaminação e tipo de ruído utilizado na contaminação dos sinais originais, assim como a SNR dos sinais limpos.

Tabela 4: Resultados obtidos para todos os sinais utilizados nas simulações, levando em consideração os valores de SNR para os sinais limpos e processados.

	Gênero/ idioma	10dB			5dB			Limpo
		C	R	M	C	R	M	
Sinal A	Masc./inglês	32,920	29,244	29,005	24,473	19,234	23,370	32,694
Sinal B	Masc./português	37,074	33,790	27,078	30,361	20,997	15,082	34,220
Sinal C	Fem./inglês	21,567	23,135	20,922	11,360	16,394	10,789	26,537
Sinal D	Masc./inglês	26,286	28,880	21,209	11,691	20,4106	12,166	28,479
Sinal E	Masc./português	31,164	34,340	27,013	20,274	23,024	16,314	38,470
Sinal F	Fem./inglês	35,714	29,711	28,823	29,985	25,090	22,448	32,034
Sinal G	Masc./inglês	39,504	39,230	30,792	31,104	25,089	22,487	38,909
Sinal H	Masc./inglês	33,939	29,649	23,525	29,007	16,072	15,440	35,570
Sinal I	Masc./inglês	39,630	38,240	22,196	30,950	23,410	15,071	39,443
Sinal J	Masc./inglês	34,882	31,634	20,322	23,452	17,103	13,176	38,053
Sinal K	Masc./inglês	39,905	29,016	23,398	27,634	18,758	15,517	38,480
Sinal L	Fem./português	28,592	27,872	16,303	20,205	17,244	10,031	28,099
Sinal M	Fem./português	30,383	28,723	27,920	24,612	24,569	23,178	30,904

Fonte: Elaborado pelo próprio autor.

Legenda: C (ruído de carro), R (ruído em restaurante), M (ruído de metrô), Masc.(Masculino) e Fem. (Feminino).

As Tabelas 5 e 6 apresentam as notas obtidas, para cada sinal separadamente, a partir da avaliação realizada pela PESQ. Na Tabela 5, constam as notas para os sinais ruidosos e na Tabela 6, constam as notas obtidas para os sinais processados.

Tabela 5: PESQ dos sinais ruidosos.

	Gênero/ idioma	10dB			5dB		
		C	R	M	C	R	M
Sinal A	Masc./inglês	2,072	2,299	2,608	1,831	1,988	2,443

Sinal B	Masc./português	2,401	2,343	2,731	2,056	1,884	2,511
Sinal C	Fem./inglês	2,483	2,473	3,13	2,113	2,092	2,965
Sinal D	Masc./inglês	2,622	2,810	2,906	2,20	2,379	2,636
Sinal E	Masc./português	2,112	2,254	2,663	1,852	1,891	2,389
Sinal F	Fem./inglês	2,162	2,360	2,711	1,807	1,991	2,4177
Sinal G	Masc./inglês	2,327	2,342	2,773	2,056	2,075	2,533
Sinal H	Masc./inglês	1,986	2,114	2,407	1,737	1,755	2,148
Sinal I	Masc./inglês	2,463	2,349	2,824	2,106	2,043	2,526
Sinal J	Masc./inglês	2,255	2,349	2,810	1,899	2,002	2,472
Sinal K	Masc./inglês	2,270	2,225	2,748	1,878	1,936	2,4276
Sinal L	Fem./português	2,142	2,073	2,623	1,722	1,710	2,298
Sinal M	Fem./português	2,156	2,080	2,833	1,743	1,735	2,316

Fonte: Elaborado pelo próprio autor.

Legenda: C (ruído de carro), R (ruído em restaurante), M (ruído de metrô), Masc.(Masculino) e Fem. (Feminino).

Tabela 6: PESQ dos sinais processados.

	Gênero/ idioma	10dB			5dB		
		C	R	M	C	R	M
Sinal A	Masc./inglês	3,432	3,311	3,732	2,978	2,960	3,489
Sinal B	Masc./português	3,568	3,428	4,032	3,132	2,884	3,665
Sinal C	Fem./inglês	3,413	3,562	4,464	2,995	3,251	4,100
Sinal D	Masc./inglês	3,503	3,562	3,921	3,124	3,251	3,512
Sinal E	Masc./português	3,334	3,15	3,464	3,023	2,800	3,382
Sinal F	Fem./inglês	3,261	3,210	3,647	2,798	2,871	3,384
Sinal G	Masc./inglês	3,465	3,211	3,625	3,06	2,903	3,409
Sinal H	Masc./inglês	3,093	3,101	3,394	2,846	2,686	3,237
Sinal I	Masc./inglês	3,544	3,337	3,864	3,131	2,905	3,504
Sinal J	Masc./inglês	3,218	3,223	3,744	2,745	2,810	3,425
Sinal K	Masc./inglês	3,255	3,091	3,628	2,774	2,778	3,343
Sinal L	Fem./português	3,262	3,11	3,649	2,760	2,643	3,137
Sinal M	Fem./português	3,244	3,06	3,685	2,73	2,656	3,282

Fonte: Elaborado pelo próprio autor.

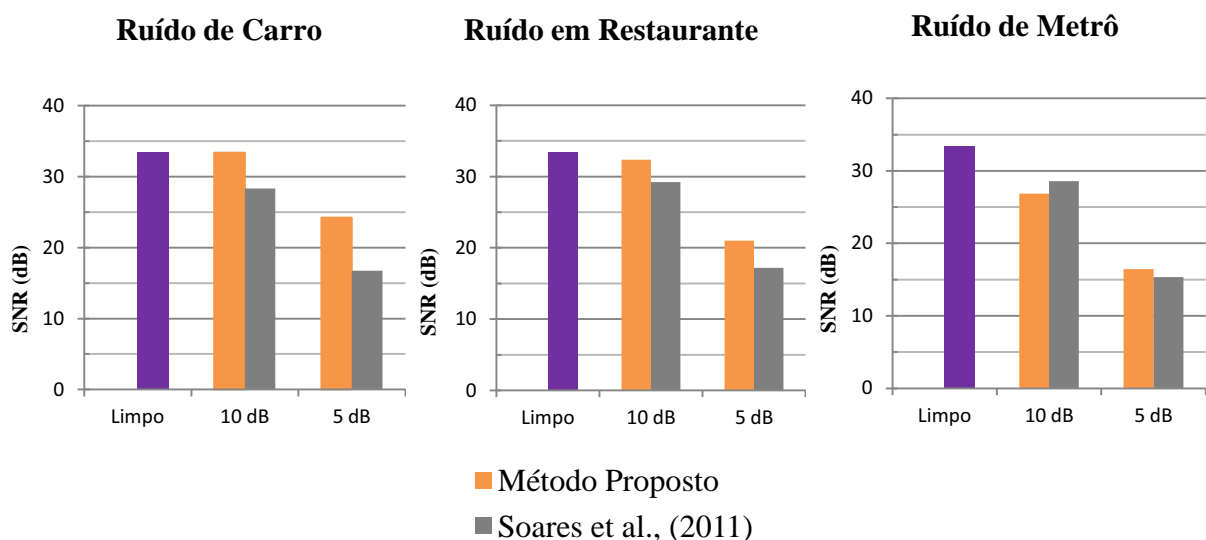
Legenda: C (ruído de carro), R (ruído em restaurante), M (ruído de metrô), Masc.(Masculino) e Fem. (Feminino).

## 6.2 Comparações a partir do método proposto por Soares et al. (2011)

A fim de validar as contribuições do presente trabalho com relação à melhoria do método proposto por Soares et al. (2011), cujos pontos principais já foram mencionados na seção 5.6, apresenta-se nesta seção uma comparação formal entre ambos métodos.

Durante as simulações, foram utilizados exatamente os mesmos sinais e os mesmos níveis de contaminação utilizados para a obtenção dos resultados expostos na seção 6.1. Na Figura 27 são mostradas as comparações dos valores de SNR, em média, para os sinais processados utilizando os dois métodos.

Figura 27. Comparações dos valores de SNR obtidos pelo método proposto e pela metodologia de Soares et al. (2011)

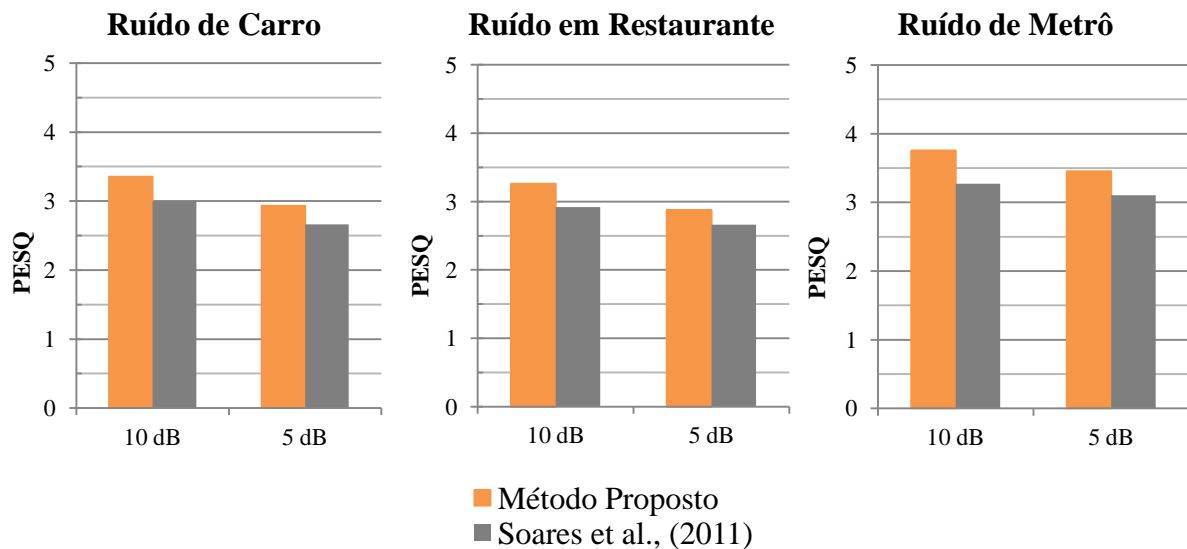


Fonte: Elaborado pelo próprio autor.

Nota-se que o Método Proposto neste trabalho foi mais eficiente que o método proposto por Soares et al. (2011), no quesito redução de ruído, quando o ruído em questão é o ruído de carro ou o ruído de restaurante. No caso do ruído de metrô, o método proposto em Soares et al. (2011) levou ligeira vantagem, porém apenas para uma contaminação de 10 dB.

Como um bom método de redução de ruído em sinais de voz deve aliar uma forte redução de ruído à preservação da qualidade psicoacústica do sinal, as comparações a partir das notas PESQ para os sinais processados ganham mais força. Sendo assim, na Figura 28 apresentam-se as avaliações realizadas usando a PESQ.

Figura 28. Comparações das notas PESQ obtidas pelo método proposto e pelo método de Soares et al. (2011)



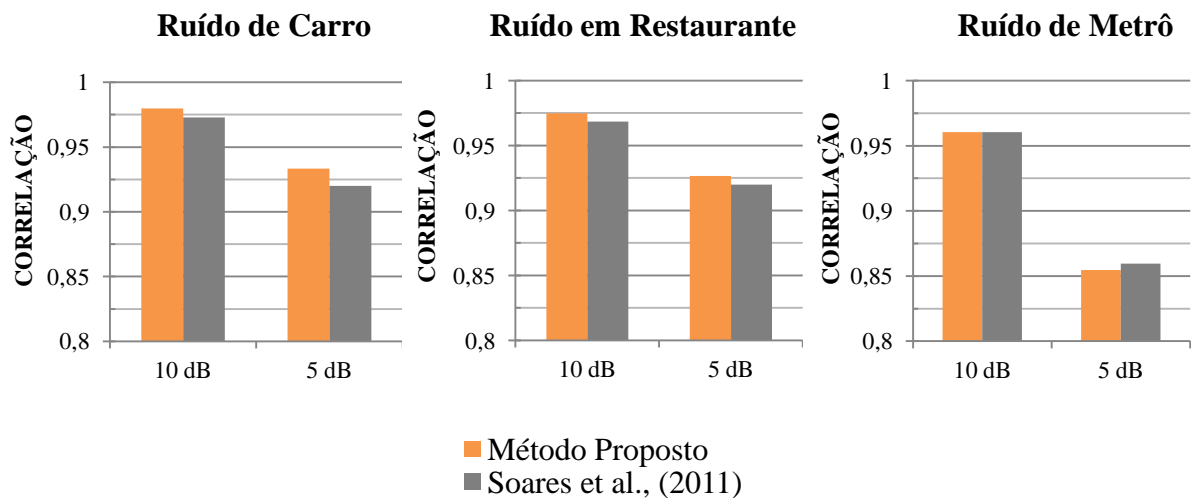
Fonte: Elaborado pelo próprio autor.

Com base na Figura 28, conclui-se que o Método Proposto neste trabalho supera o método proposto em Soares et al. (2011), com relação a qualidade psicoacústica dos sinais processados, para os três tipos de ruído utilizados. Deve-se isto principalmente ao uso das equações de diferenças finitas e à proposta de um operador dedicado a minimizar perdas de coeficientes *unvoiced*.

Outra forma de comparar ambos os métodos seria através do Coeficiente de Correlação de Pearson. O coeficiente de correlação de Pearson varia no intervalo  $[-1,1]$ , mensurando a relação mútua entre dois termos (FIELD, 2005). Coeficiente de correlação igual a  $-1$  ou  $1$  implica em uma correlação perfeita entre o sinal original e o sinal processado. Caso um processamento receba uma nota de correlação igual a  $0$ , significa que o sinal processado não está correlacionado com o original. Em outras palavras, o sinal processado pouco tem a ver com o sinal original. Uma interpretação para a ocorrência de uma correlação negativa entre duas variáveis seria que, caso o valor de uma aumente, a outra sempre diminui. Fica como exemplo o par de variáveis  $y_1 = 2x$  e  $y_2 = -2x$ , que possuem um coeficiente de correlação igual a  $-1$ .

Na Figura 29 constam as médias das correlações obtidas entre os sinais originais e processados, para o Método proposto e para o método proposto por Soares et al. (2011). Note que neste caso, as notas de correlações variam no intervalo  $[0,1]$ .

Figura 29. Comparações das correlações obtidas entre os sinais originais e processados pelo método proposto e pelo método proposto em Soares et al. (2011).



Fonte: Elaborado pelo próprio autor.

Com base na Figura 29, nota-se que ambos os métodos obtiveram forte correlação entre os sinais originais e processados. Para Santos (2007), coeficiente de correlação compreendido entre  $[0,8, 1[$ , indica uma forte correlação entre as variáveis observadas.

Para efeito de comparações, verifica-se que o Método Proposto se sobressaiu com relação ao método proposto por Soares et al. (2011), quando o ruído em questão é o ruído de carro ou em restaurante. Já para o ruído de metrô, Soares et al. (2011) levou ligeira vantagem, principalmente para uma contaminação de 5 dB. Deve-se isto ao fato do mesmo ter obtido em alguns casos, também para o ruído de metrô, melhores notas de SNR.

Levando em consideração as comparações realizadas nesta seção, e atribuindo um peso maior aos resultados utilizando as notas PESQ (que refletem de uma maneira mais precisa a qualidade do processamento), conclui-se que, de um modo geral, os resultados obtidos pelo Método Proposto neste trabalho superam os resultados obtidos por Soares et al. (2011).

---

## 7 CONSIDERAÇÕES FINAIS

No presente trabalho foram estudadas técnicas de redução de ruído em sinais de voz no domínio wavelet. O objetivo foi propor uma metodologia que seja eficiente para o tratamento de sinais contaminados por ruído colorido.

Para isso, foram estudados os fundamentos da teoria wavelet e da redução de ruído em sinais de voz e os principais métodos de redução de ruído em sinais de voz no domínio wavelet.

A transformada wavelet discreta é uma ferramenta poderosa em métodos de processamento de sinais de voz, tanto para compressão, quanto para melhoria da qualidade do sinal. Quando o objetivo é melhorar a qualidade do sinal de voz, o problema consiste em determinar um método de filtragem que seja eficiente na eliminação do ruído de fundo, e ainda, que minimize as perdas em trechos de voz. Os métodos que usam limiar são os mais comuns quando o processamento ocorre no domínio wavelet. Estes se baseiam na premissa de que muitos coeficientes são redundantes e, quando eliminados, não comprometem a reconstrução do sinal. O maior problema desta técnica consiste em determinar a melhor forma de se calcular e aplicar o limiar.

A proposta deste trabalho foi realizar estudos e pesquisas, de modo a culminar com a determinação de um método não-limiar para redução de ruído em sinais de voz no domínio wavelet. O sinal de saída foi estimado, a partir do original, pelas equações de diferenças finitas. Acrescentam-se informações contidas em uma vizinhança para a construção de cada

coeficiente do filtro. Para a estimação do ruído foi usado o operador SNRP, que é aplicado em cada uma das equações de diferenças. Uma combinação polinomial foi realizada de maneira a condensar todas as equações de diferenças em uma única função de transferência. Além disso, foi realizada uma pré-filtragem, preparando o sinal estimado pelas diferenças para o processo de ajuste de amplitudes. Desta forma, conseguiu-se um filtro que atue no sinal levando em consideração informações contidas em cada coeficiente ou em uma vizinhança do mesmo.

A eficiência do método foi avaliada a partir de testes computacionais. Utilizaram-se treze sinais de voz e três tipos de ruído presente em situações cotidianas. Para verificar a qualidade dos sinais processados, foram avaliados os níveis de redução de ruído e distorção usando como referência os valores de SNR. O grau de inteligibilidade do sinal processado também foi avaliado. Para isso, tomou-se como referência as notas PESQ atribuídas aos mesmos.

Os valores de SNR dos sinais processados estiveram sempre muito próximos dos valores dos sinais limpos, indicando boa redução de ruído sem inserir distorções nos trechos de voz dos mesmos. As notas PESQ estiveram sempre superiores a 3, para os sinais com nível de contaminação de 10 dB. Para um nível de contaminação de 5 dB, as notas se mantiveram muito próximas a 3, em alguns casos até superando esta marca. Este fato indica uma boa qualidade dos sinais processados. A redução de amplitude que ocorre durante o processamento é praticamente imperceptível à audição humana, sendo identificada principalmente para os sinais com nível de contaminação de 5 dB (ver apêndice B).

A principal vantagem do método não-limiar proposto neste trabalho, em relação aos demais métodos de redução de ruído no domínio wavelet, é a construção de um filtro que atue no sinal levando em consideração informações contidas em cada coeficiente em processamento, ou, em uma vizinhança do mesmo. Realiza-se um tratamento equivalente para todo o sinal no domínio wavelet, evitando desconfortos auditivos ao ouvinte. Com isso, não se faz necessário o uso de um único valor (valor do limiar) como referência para processar todos os coeficientes de uma faixa de frequência do sinal.

Os resultados apresentados neste trabalho são aceitáveis quando o ruído em questão é o ruído colorido. As comparações realizadas a partir do método proposto por Soares et al. (2011) ratificaram a eficiência e as contribuições deste trabalho. Um ponto ainda em discussão é o uso do operador *Dif*, este requer um pouco mais de atenção para que, durante o processamento, não evidencie o ruído, e sim somente coeficientes relevantes, minimizando assim a redução de amplitude nos sinais processados. Testes e modificações continuam sendo

realizados na busca de melhorias nos sinais processados e na busca de bons resultados, também, para outros tipos de ruído.

---

## REFERÊNCIAS

ANTUNES JÚNIOR, I. **Redução de ruído em sinais de voz usando curvas especializadas de modificação dos coeficientes da transformada em co-seno.** 2006. 112 f. Tese (Doutorado)-Escola Politécnica da Universidade de São Paulo, Universidade de São Paulo, São Paulo, 2006.

ARAÚJO, L. C. **Banco de Filtros Wavelets com Fator de Escala Maior que Dois.** 2007. 172 f. Dissertação (Mestrado) - Departamento de Engenharia Elétrica, Universidade Federal de Minas Gerais, Belo Horizonte, 2007.

ATAL, B. S.; RABINER, L. R. A Pattern Recognition Approach to Voiced – Unvoiced – Silence Classification with Applications to Speech Recognition, **IEEE Transaction on Acoustics, Speech and Signal Processing**, New York, v. 24, n. 3, p. 201-212, 1976.

BALAKRISHNAN, V.; BORGES, N.; PARCHMENT, L. **Wavelet Denoising and Speech Enhancement**, Baltimore: Department of Electrical and Computer Engineering, The Johns Hopkins University, 2010.

BEERENDS, J. G.; HEKSTRA, A. P.; RIX, A. W.; HOLLIER, M. P. Perceptual evaluation of speech quality (PESQ) the new ITU standard for end-to-end speech quality assessment part ii: psychoacoustic model. **AES Journal**, New York, v. 50, n.10, p. 765-778, 2002.

BLATTER, C. **Wavelets a Primer.** Natick: A.K. Peters, 1998. 202 p.

BOGGESS, A.; NARCOWICH, F.J. **A first course in Wavelets with Fourier Analysis.** Upper Saddle River: Prentice Hall, 2001. 283 p.

COHEN, I. Speech enhancement using a noncausal a priori SNR estimator. **IEEE Signal Processing Letters**, New York, v.11, n.9, p. 725-728, 2004.

D'AMBRÓSIO U. **Educação Matemática: da teoria à prática**. 10 ed. Campinas: Papirus, 2003. 121p.

DAE-SUNG, K.; JUNG-GO, C.; GEUN-TAEK, R; MUN-SEOB, B.; HYEON-DEOK, B. Noise reduction using coefficients smoothing wavelet transforms. In: IEEE PROCEEDINGS OF INTERNATIONAL CONFERENCE ON SIGNAL PROCESSING, APPLICATION AN TECHNOLOGY, ICSPAT , 2002, San José. **Proceedings of the...** San José: IEEE, 2002. p. 029-033.

DAUBECHIES, I. **Ten lectures on wavelets**. Philadelphia: SIAM Books, 1992. 357p.

DELLER, J. L.; PROAKIS, J. G.; HANSEN, J. H. L. **Discrete-time processing of speech signals**. New York: Macmillan, 1993. 908p.

DONOHO, D. L.; JOHNSTONE, I. M. Ideal spatial adaptation via wavelet shrinkage. **Biometrika**, London, v. 81, n.3, p. 425-455, 1994.

DONOHO, D. L. De-noising by soft-thresholding. **IEEE Transactions on Information Theory**, New York, v.41, n. 3, p. 613-627, 1995.

DUARTE, M. A. Q.; OLIVEIRA, L. C. O. ; VILLARREAL, F.; DÍAZ, L. A. Compressão de sinais elétricos usando a transformada de wavelet. **TEMA: Tendências em Matemática Aplicada e Computacional**, São Carlos, v.4, n.1, p. 21-30, 2003.

DUARTE, M. A. Q. **Redução de ruído em sinais de voz no domínio Wavelet**. 2005. 105 f. Tese (Doutorado em Automação) - Faculdade de Engenharia, Universidade Estadual Paulista, Ilha Solteira, 2005.

DUARTE, M. A. Q.; VIEIRA FILHO, J.; VILLARREAL, F., Um Método Simples e Eficiente para Detecção de Atividade de Voz Usando a Transformada Wavelet, In: CONGRESSO BRASILEIRO DE MATEMÁTICA APLICADA E COMPUTACIONAL, CNMAC, 9, 2009, Cuiabá-MT, **Anais...** Cuiabá: CNMAC 2009, p. 1022 – 1028, 2009.

DUARTE, M. A. Q. **Processamento de sinais de voz no domínio wavelet**. Cassilândia: UEMS, 2010. (Relatório de Pesquisa)

FIELD, A. **Discovering Statistics Using SPSS**. London: Sage, 2005. 135 p.

FIGUEIREDO, D. G. **Análise de Fourier e equações diferenciais parciais**. Rio de Janeiro: IMPA. 1977. 274 p.

GOMES, J., VELHO, L., GOLDSTEIN, S., **Wavelets: teoria, software e aplicações**, Rio de Janeiro, IMPA, 1997. 217 p.

HAAR, A. Zur. Theorie der orthogonalen functionen-systeme. **Mathematische Annalen**, Berlin, v.69, p. 331-371, 1910.

HERNANDES, E; WEISS, G; **A First Course on Wavelets**. Boca Raton: CRC press, 1996. 489 p.

HU Y.; LOIZOU, P.C.; Evaluation of Objective Quality Measures for Speech. **IEEE Transactions on Audio, Speech, and Language Processing**, Dallas v. 16 n. 1, p. 229-238, 2008.

INTERNATIONAL TELECOMMUNICATIONS UNION. **Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs**. Geneva: ITU, 2001. (ITU-T P.862).

LALLOUANI, A.; GABREA, M.; GARGOUR, C. S. Wavelet based speech enhancement using two different threshold-based denoising algorithms. In: ELECTRICAL AND COMPUTER ENGINEERING, CANADIAN CONFERENCE, 2004, Canada. **Conference...** Canadá: IEEE, 2004. v.1, p. 315-318, 2004. Disponível em: <http://ieeexplore.ieee.org/xpl/tocresult.jsp?isnumber=29618&isYear=2004>. Acesso em: 29 out. 2012.

LEE, D. T. L.; YAMAMOTO A. Wavelet analysis: theory and applications. **Hewlett-Packard Journal**, Palo Alto, v. 45, p. 44-54, 1994. Disponível em: <<http://www.hpl.hp.com/hpjournal/94dec/dec94a6a.pdf>>. Acesso em: 12 dez. 2012.

LIMA, P. C. Wavelets: uma introdução. **Matemática Universitária**, Belo Horizonte, n. 33, p. 13-44, 2004. Disponível em: <[http://matematicauniversitaria.ime.usp.br/Conteudo/n33/n33\\_Artigo02.pdf](http://matematicauniversitaria.ime.usp.br/Conteudo/n33/n33_Artigo02.pdf)>. Acesso em: 20 dez. 2012.

LOUIS, A. K.; MAAB, P.; RIEDER, A. **Wavelets theory and applications**. Chichester: John Wiley & Sons, 1998. 324 p.

MALLAT, S. A theory for multiresolution representation signal decomposition: the Wavelet representation. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, New York, v.11, n.7, p. 674-693, 1989.

MALLAT, S. **A Wavelet tour of signal processing**. 2. ed. New York. Academic Press, 1999. 637 p.

MELLO, C. A. **Análise Wavelet**. 2008. Universidade Federal do Pernambuco. Centro de Informática. Disponível em: <[http://www.cin.ufpe.br/~cabm/pds/PDS\\_Aula08%20Wavelet.pdf](http://www.cin.ufpe.br/~cabm/pds/PDS_Aula08%20Wavelet.pdf)>. Acesso em 22/10/2011.

MORETTIN, P. A. **Ondas e ondaletas: da análise de fourier à análise de ondaletas**. São Paulo: Edusp, 1999. 276 p.

NIEVERGELT, Y. **Wavelets made easy**. Boston: Birkhäuser, 1999. 297 p.

NOCETI FILHO, S. **Fundamentos sobre ruídos**. Parte II – definição, caracterização e tipos de ruídos. Universidade Federal de Santa Catarina. Disponível em: <<http://www.etelj.com.br/etelj/artigos/151280402b0bc1accfaea913d6301caf.pdf>>. Acesso em 23 jan. 2012.

OPPENHEIM, A. V.; SCHAFER, R. W.; BUCK, J. R. **Discrete-time signal processing**. New Jersey: Prentice Hall, 1998.

QIANG, F.; WAN, E. A. **A novel speech enhancement system based on wavelet denoising**. School of Science and Engineering at OHSU, Center of Spoken 146 Language Understanding. 2003. Disponível em: <<http://cslu.cse.ogi.edu/publications/ps/fu03.pdf>>. Acesso em: 04 fev. 2012.

QUACKENBUSH S. R.; BARNWELL III T. P. and CLEMENTS M. A. **Objective Measures of Speech Quality**, Prentice Hall, Englewood Cliffs, 1988.

RIOUL, O.; VETTERLI, M. Wavelets and signal processing, **Signal Processing Magazine IEEE** New York, v. 8, n. 4, p. 14-38, 1991.

SANCHEZ, F.L. **Análise cepstral baseada em diferentes famílias de transformada Wavelet**. 2008. 96 f. Dissertação (Mestrado) – Escola de Engenharia de São Carlos/Faculdade de Medicina de Ribeirão Preto/Instituto de Química de São Carlos da Universidade de São Paulo, 2008.

SANTOS, C. **Estatística descritiva: manual de auto-aprendizagem**. Lisboa: Silabo, 2007. 312 p.

SHEIKHZADEH, H.; ABUTALEBI, H. R. An improved wavelet-based speech enhancement system. In: EUROPEAN CONFERENCE ON SPEECH COMMUNICATION AND TECHNOLOGY, EUROSPEECH, 7, 2001, Aalborg. **Proceedings of the...** Aalborg: ISCA, p. 1855-1858, 2001.

SEOK, J. W.; BAE, K. S. Speech enhancement with reduction of noise components in the wavelet domain. In: IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, ICASSP'97, 1997, Munich. **Proceedings of the...** Munich: IEEE, 1997. v. 2, p. 1323-1326.

SOARES, W. C.; DUARTE, M. A. Q.; VILLARREAL, F.; VIEIRA FILHO. Análise de métodos de redução de ruído em sinais de voz no domínio wavelet. In: BRAZILIAN CONFERENCE ON DYNAMICS, CONTROL AND THEIR APPLICATIONS, DINCON'2007, 6, 2007, São José do Rio Preto. **Proceedings of the...** São José do Rio preto: SBMAC, 2007 p.918 – 923.

SOARES,W.C. **Um método não-limiar para redução de ruído em sinais de voz no domínio Wavelet**. 2009. 205 f. Tese (Doutorado em Automação)- Faculdade de Engenharia, Universidade Estadual Paulista, Ilha Solteira, 2009.

SOARES, W. C.; DUARTE, M. A. Q.; VILLARREAL, F.; VIEIRA FILHO, J. Análise de métodos de redução de ruído por limiar no domínio wavelet. **TEMA: Tendências em Matemática Aplicada e Computacional**, São Carlos, v. 9, n. 3, p. 471-480, 2008.

SOARES, W. C.; VILLARREAL, F.; DUARTE, M. A. Q.; VIEIRA FILHO. Wavelets in a Problem of Signal Processing, **Novi Sad Journal of Mathematics**, Novi Sad, v. 41, n. 1, p. 11-20, 2011.

SPERANDIO, D.; MENDES, J. T.; SILVA, L. H. M. **Cálculo Numérico**: características matemáticas e computacionais dos métodos numéricos. São Paulo: Prentice Hall, 2003. 354 p.

STRANG, G.; NGUYEN, T. **Wavelets and Filter Banks**. Wellesley-Cambridge Press, 1996. 490 p.

VETTERLI, M.; KOVAČEVIĆ J.; GOYAL, V. K. **Fourier and wavelet signal processing**. (October 30, 2011). Disponível em: <http://fourierandwavelets.org/>. Acesso em: 22 jan. 2012.

VIEIRA FILHO, J. **Redução de ruído em sinais de voz nos sistemas rádio móveis veiculares**. 1996. 113 f. Tese (Doutorado) - Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, 1996.

YANG, W. **Enhanced modified bark spectral distortion (EMBSD)**: An objective speech quality measure based on audible distortion and cognition model. 1999. 177 f. Tese (Doctor) – Graduate Board in Philosophy, Temple University, Philadelphia, 1999.

---

## APÊNDICE A - DEMONSTRAÇÃO DE TEOREMA

**Teorema.** Suponha que  $V_0$  seja um subespaço de um espaço de funções de Hilbert  $V$  munido com produto interno  $\langle \cdot \rangle$ . Suponha que  $\{\varphi_1, \dots, \varphi_n\}$  é uma base ortonormal de  $V_0$ . Se  $f \in V_0$ , então:

$$f = \sum_{j=1}^n \langle f, \varphi_j \rangle \varphi_j.$$

Demonstração:

Seja  $f \in V_0$  uma função qualquer. Essa função é uma combinação linear, dos  $\varphi_j$ , da forma:

$$f = \sum_{j=1}^n \alpha_j \varphi_j, \text{ com } \alpha_j \in \mathbb{C}. \quad (37)$$

Para calcular cada constante  $\alpha_k$ , tomamos o produto interno em ambos os lados na equação (37) com  $\varphi_k$ :

$$\langle f, \varphi_k \rangle = \sum_{j=1}^n \langle \alpha_j \varphi_j, \varphi_k \rangle. \quad (38)$$

O único termo não nulo na parte direita da equação (38) ocorre quando  $j = k$ . Isto se deve ao fato de que os  $\varphi_j$  são ortonormais. Desta forma:

$$\langle f, \varphi_k \rangle = \langle \alpha_k \varphi_k, \varphi_k \rangle = \alpha_k \langle \varphi_k, \varphi_k \rangle = \alpha_k.$$

---

## APÊNDICE B - ANÁLISE VISUAL DO PROCESSAMENTO REALIZADO PELO MÉTODO PROPOSTO

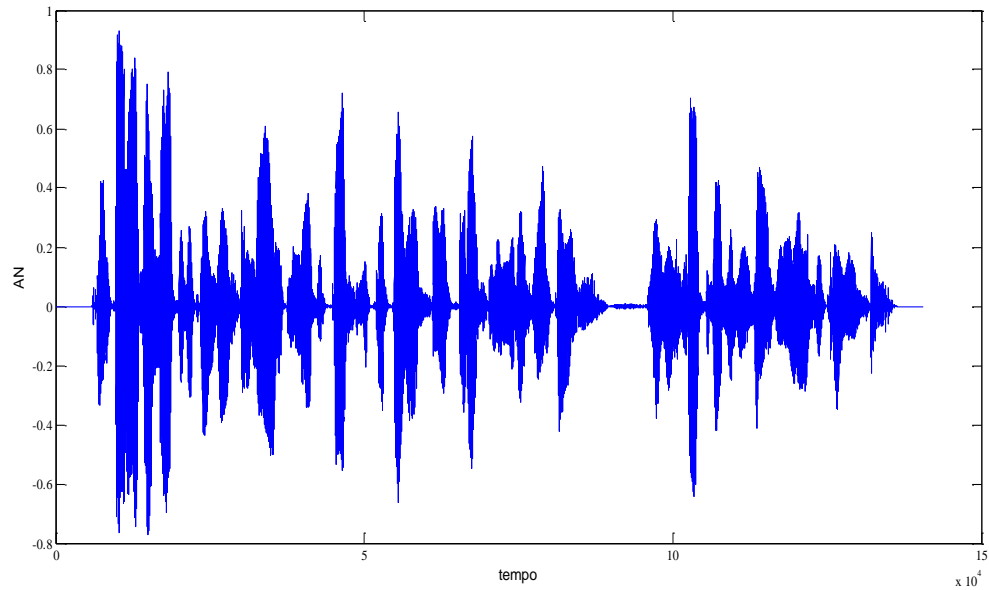
Este apêndice trata da análise da forma de onda de um sinal em particular. A análise será feita a partir de um sinal, escolhido arbitrariamente, considerando as formas de onda dos sinais ruidosos e processados. O ruído adicionado é o ruído de carro com dois níveis de contaminação, 10 dB e 5 dB. A frase utilizada é a seguinte:

$s(n)$ : *Comparado ao método convencional, o novo teste de propagação oferece inúmeras vantagens, diz o chefe da divisão de meios de transmissão".*

O sinal limpo  $s(n)$  foi amostrado 16 kHz e quantizado com 16 bits por amostra. A técnica de janelamento foi empregada utilizando a janela de Hanning com 50% de sobreposição entre os segmentos. Cada janela é composta por 256 amostras do sinal  $y(n) = s(n) + w(n)$ , onde  $w(n)$  é o ruído produzido pelos pneus de um carro a 90 km/h em uma rodovia. A função utilizada na realização da DWT foi a wavelet de Daubechies de ordem 10 (db10).

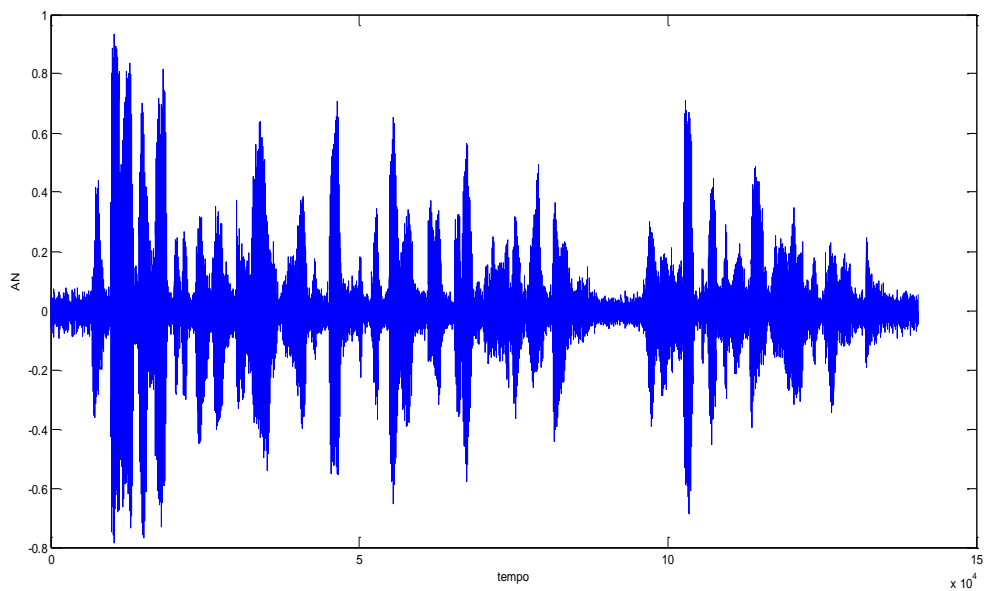
As Figuras 30 e 31, trazem as formas de onda do sinal  $s(n)$  limpo e do sinal ruidoso  $y(n)$ , obtido do sinal original após contaminação com nível de 10 dB.

Figura 30. Sinal de voz limpo  $x(n)$ .



Fonte: Elaborado pelo próprio autor.

Figura 31. Sinal de voz contaminado  $y(n)$  (10 dB).

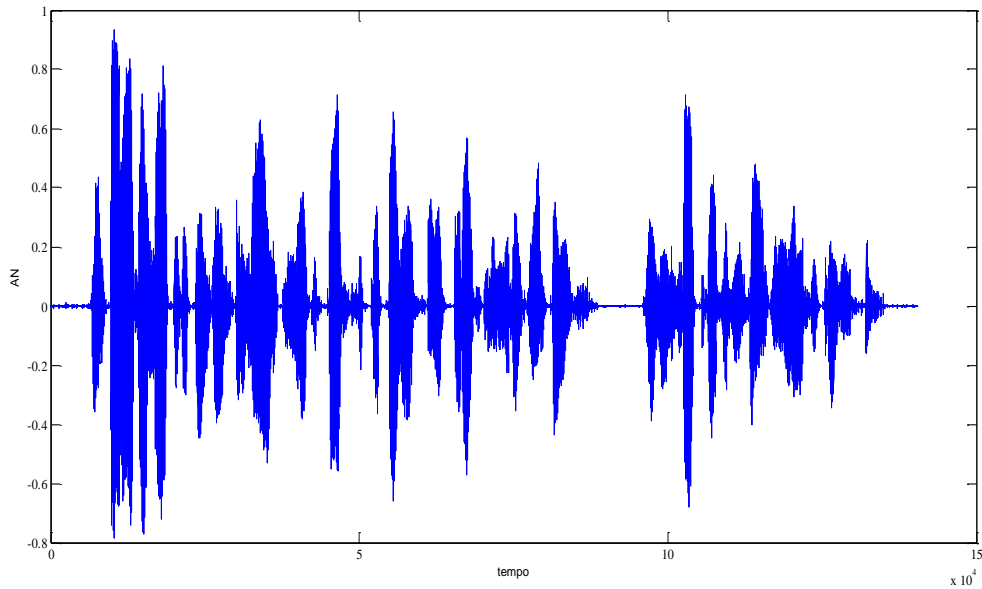


Fonte: Elaborado pelo próprio autor.

Na Figura 32, consta o sinal de voz processado,  $\hat{y}(n)$ , usando o método proposto, quando o sinal de entrada é o sinal contido na Figura 31. Nota-se que o método é eficiente na redução de ruído. Para melhor comparar, a Figura 33 traz o sinal original  $s(n)$ , em azul, e o

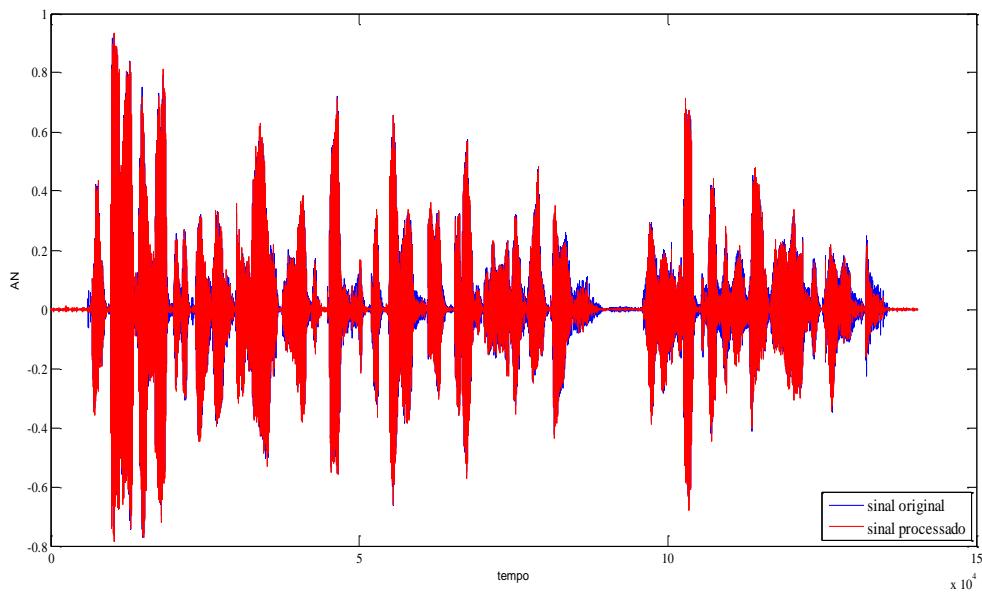
sinal processado,  $\hat{y}(n)$ , em vermelho. Observa-se que apesar da forte redução de ruído, a perda de amplitude dos coeficientes de voz não é relevante.

Figura 32. Sinal de voz processado  $\hat{y}(n)$ , usando o método proposto, a partir do sinal de voz limpo  $s(n)$  contaminado com ruído de carro (10 dB).



Fonte: Elaborado pelo próprio autor.

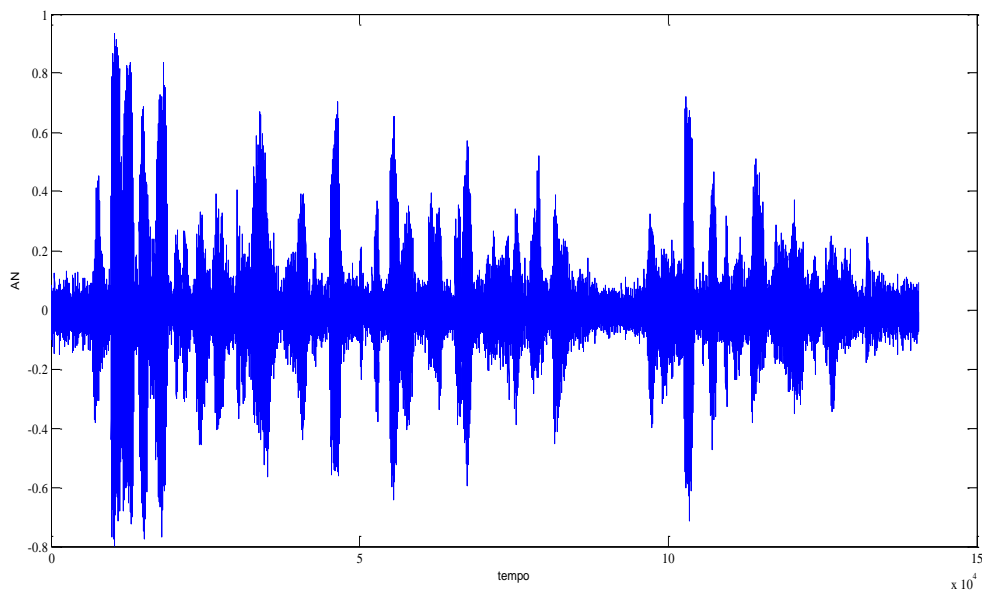
Figura 33. Sinal de voz limpo  $s(n)$  (azul) e sinal processado,  $\hat{y}(n)$  (vermelho), usando o método proposto, a partir do sinal de voz limpo contaminado com ruído de carro (10 dB).



Fonte: Elaborado pelo próprio autor.

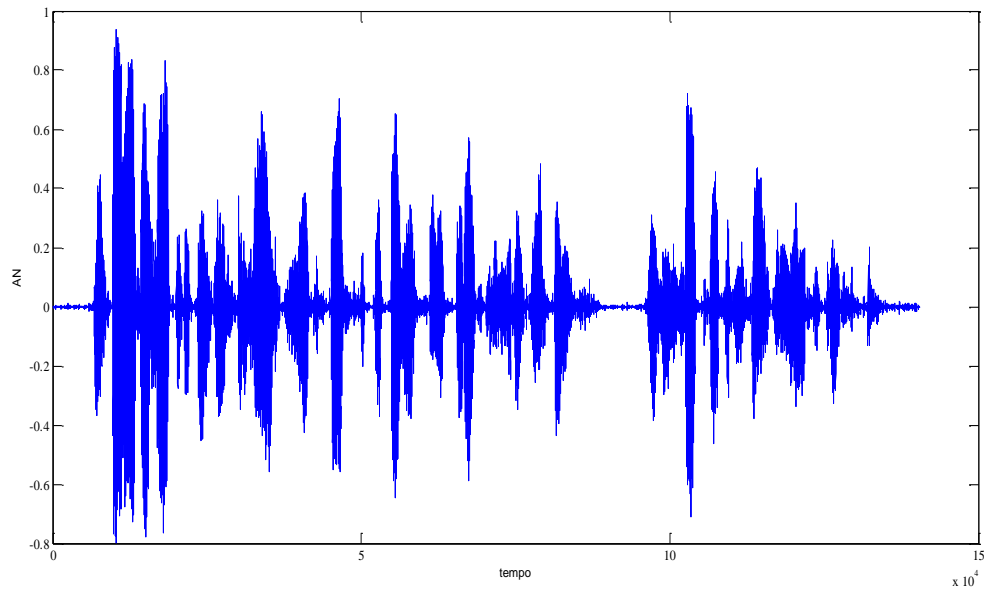
Para analisar o método, agora para o processamento do sinal  $y(n)$  com um nível de contaminação de 5 dB, têm-se as Figuras 34, 35 e 36. As Figuras 34 e 35 trazem, respectivamente, a forma de onda do sinal ruidoso  $y(n)$  e a forma de onda do sinal processado pelo método proposto,  $\hat{y}(n)$ . A fim de analisar a redução de amplitude dos coeficientes de voz, têm-se a Figura 36. Em azul, o sinal limpo  $s(n)$ , e em vermelho o sinal processado  $\hat{y}(n)$ . Nota-se que, embora as perdas sejam maiores que no caso do processamento para 10 dB, as mesmas também não são significativas e não comprometem a inteligibilidade do sinal. Também para este caso, a redução de ruído efetuada pelo método foi significativa.

Figura 34. Sinal de voz contaminado  $y(n)$  (5 dB).



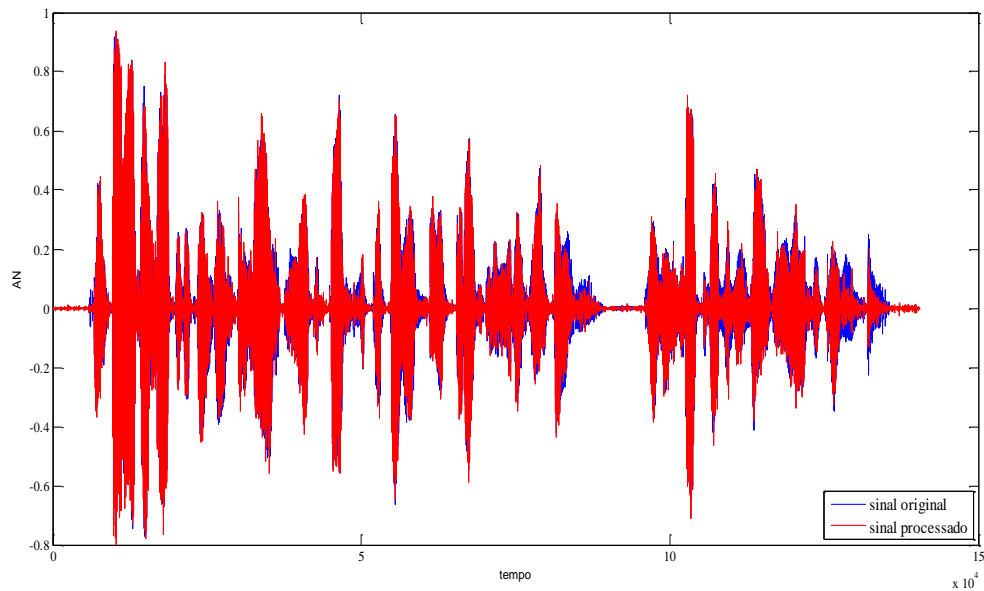
Fonte: Elaborado pelo próprio autor.

Figura 35. Sinal de voz processado,  $\hat{y}(n)$ , usando o método proposto a partir do sinal de voz limpo  $s(n)$  contaminado com ruído de carro (5 dB).



Fonte: Elaborado pelo próprio autor.

Figura 36. Sinal de voz limpo  $x(n)$  (azul) e sinal processado  $\hat{y}(n)$  (vermelho), a partir do sinal de voz limpo contaminado com ruído de carro (5 dB).



Fonte: Elaborado pelo próprio autor.

Para se ter uma ideia das notas fornecidas pelas avaliações objetivas para este sinal, a tabela abaixo contém as notas PESQ e os valores de SNR para este caso em particular. A SNR

do sinal limpo é de 34,22 dB e as notas PESQ dos sinais ruidosos para uma contaminação de 10dB e 5dB valem, respectivamente, 2,401 e 2,056.

Tabela 7. Avaliações objetivas para o sinal  $s(n)$  em particular.

	Contaminação de 10 dB	Contaminação de 5dB
PESQ	3,568	3,132
SNR (dB)	37,074	30,361

Fonte: Elaborado pelo próprio autor.

Tomando como referência a Tabela 7 e as formas de onda dos sinais processado apresentadas neste apêndice, constata-se que o método reduz significativamente o ruído presente no sinal, sem a inserção de fortes distorções.

---

## APÊNDICE C - SOBRE A EXPRESSÃO UTILIZADA PARA A DEFINIÇÃO DO OPERADOR *Dif*.

Como definido no Capítulo 5 (equação (32)), o operador *Dif* é aplicado no sinal em processamento da seguinte forma:

$$Dif(e_i) = \begin{cases} \sqrt{\frac{7}{2}} \cdot (e_i)^3 & \text{se } e_i \text{ está contido em } H_2 \text{ ou } H_1. \\ e_i & \text{caso contrário.} \end{cases} \quad (39)$$

O objetivo deste apêndice é explicar a razão em se utilizar a expressão que aparece na equação (39), em particular, “como” surgiu a função definida pela equação (40).

$$f(t) = \sqrt{\frac{7}{2}} t^3. \quad (40)$$

Durante as simulações, notou-se que apesar da eficiência em reduzir ruído, o método proposto encontrava dificuldades no tratamento de regiões *unvoiced*. Observou-se isto, através da comparação entre as formas de onda dos sinais original e processado. O que acontecia era a redução de amplitudes destes segmentos nos sinais processados. Na busca de minimizar estas perdas, surgiu a necessidade de utilizar um operador de pré-processamento. Este teria a função de produzir um ganho em amplitude nestes segmentos, antes do processo de filtragem. Desta forma, a redução de amplitudes seria menor.

Buscou-se, então, um operador que pudesse evidenciar tais segmentos. Porém, tal operador deveria apresentar um gráfico que fosse uma curva nem muito suave, nem muito

abrupta, pois ele teria que ressaltar os coeficientes *unvoiced* e, ao mesmo tempo, não evidenciar os coeficientes ruidosos.

Com esse objetivo considerou-se o seguinte problema de valor inicial (PVI)

$$\frac{dy}{dt} = \frac{21t^2}{4y}, \quad y(0) = 0. \quad (41)$$

A solução geral, da equação diferencial considerada, é a família de curvas dadas por:

$$2y^2 - 7t^3 = c; \quad y \neq 0, \quad (42)$$

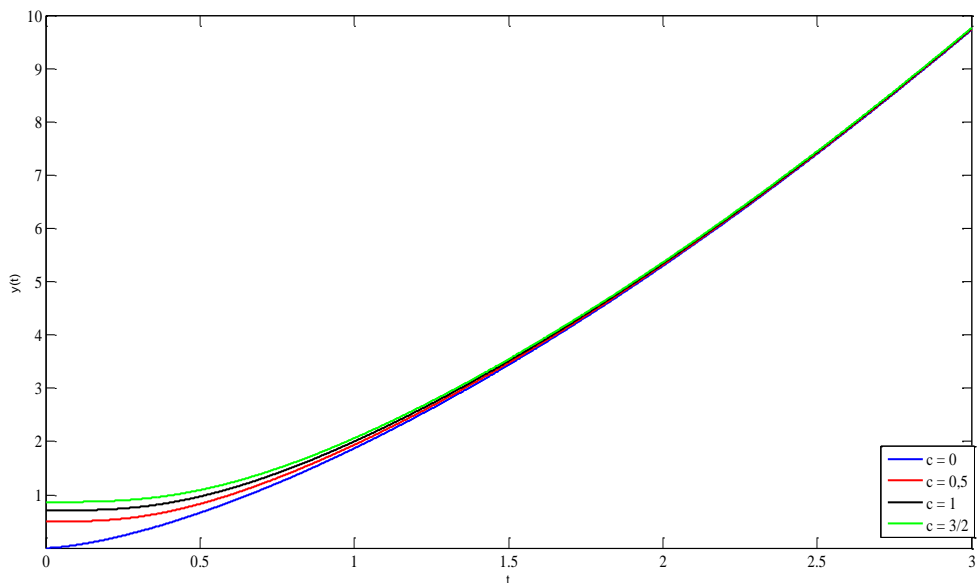
sendo  $c$  uma constante qualquer.

Fixado  $c$ , a equação (42), que está na forma implícita, pode ser escrita de maneira explícita, conforme a equação (43).

$$y = \pm \sqrt{\frac{7t^3 + c}{2}}, \quad \text{com } t \geq \left(-\frac{c}{7}\right)^{1/3}. \quad (43)$$

Para as aplicações serão consideradas funções  $y(t) \geq 0$ . Observe que, para cada valor de  $c$ , tem-se uma curva diferente. A Figura 37 mostra algumas destas curvas.

Figura 37. Gráfico da equação (43) para diferentes valores de  $c$ .



Fonte: Elaborado pelo próprio autor.

Analisando as curvas mostradas na Figura 37, nota-se que a curva mais adequada para a aplicação em questão é aquela obtida quando  $c = 0$ , pois as outras não contêm a origem do plano  $(t, y)$ .

Por outro lado, usando a condição inicial  $y(0) = 0$ , do PVI considerado, tem-se claramente que  $c = 0$ . Assim, a partir da equação (43), a solução do PVI considerado é dada por

$$y = \pm \sqrt{\frac{7}{2}} t^3, \quad \text{para } t \geq 0. \quad (44)$$

A equação (44) contém a função utilizada para a definição do operador *Dif*, definição essa dada pela equação (39).

Observando as curvas da Figura 37, nota-se que a equação (44) pode ser usada como um operador que age suavemente no sinal ao qual for aplicado. Porém, atua com mais rigor nos coeficientes menos significativos (mais próximos de zero) e tende a manter os coeficientes mais importantes para a reconstrução do sinal.

---

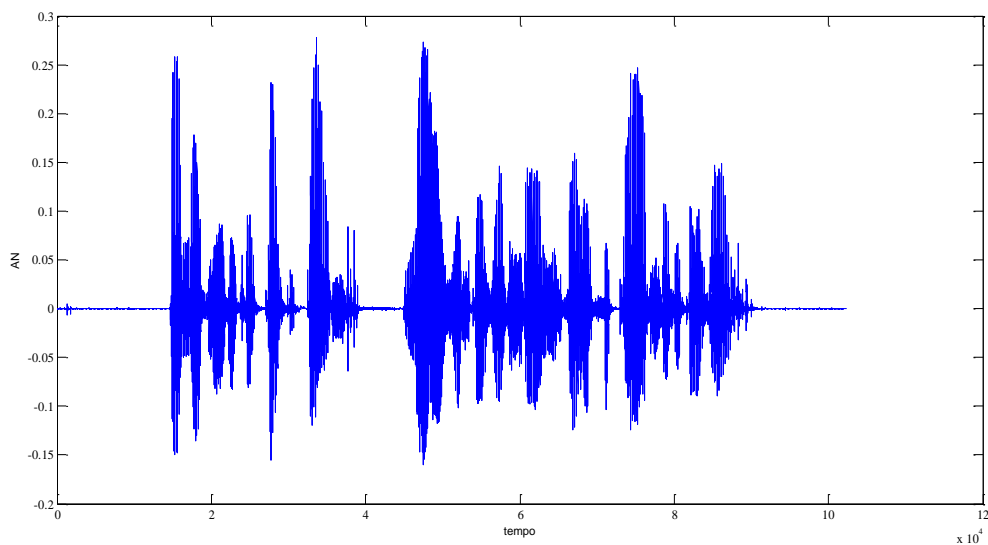
## APÊNDICE D - ANÁLISE DA INFLUÊNCIA DO OPERADOR *Dif* NA QUALIDADE DO SINAL PROCESSADO

Neste apêndice será realizada uma análise da influência do operador *Dif* no sinal processado. Para isso será considerado um sinal de voz escolhido arbitrariamente. Para fins didáticos, o nível de contaminação será de 10 dB. O sinal ruidoso  $y(n)$  é obtido por contaminação do sinal limpo  $s(n)$  com ruído em restaurante. O sinal modela a seguinte frase:

$$s(n): \begin{cases} \text{"A medida que o tempo passa,} \\ \text{"mais nos convencemos da eficácia do convênio."} \end{cases}$$

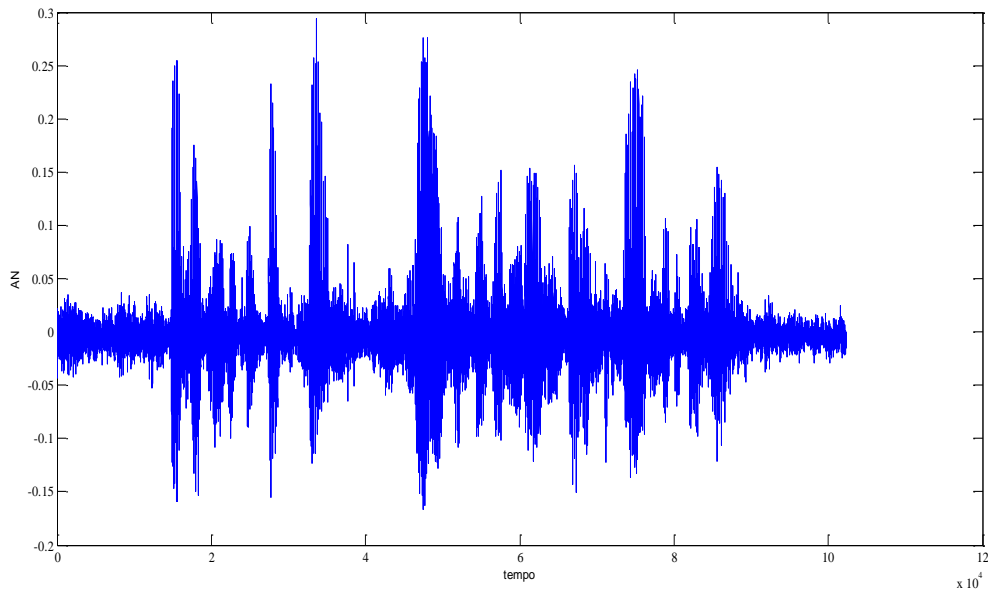
Nas Figuras 38 e 39, constam os sinais de voz limpo e ruidoso, respectivamente.

Figura 38. Sinal de voz limpo  $x(n)$ .



Fonte: Elaborado pelo próprio autor.

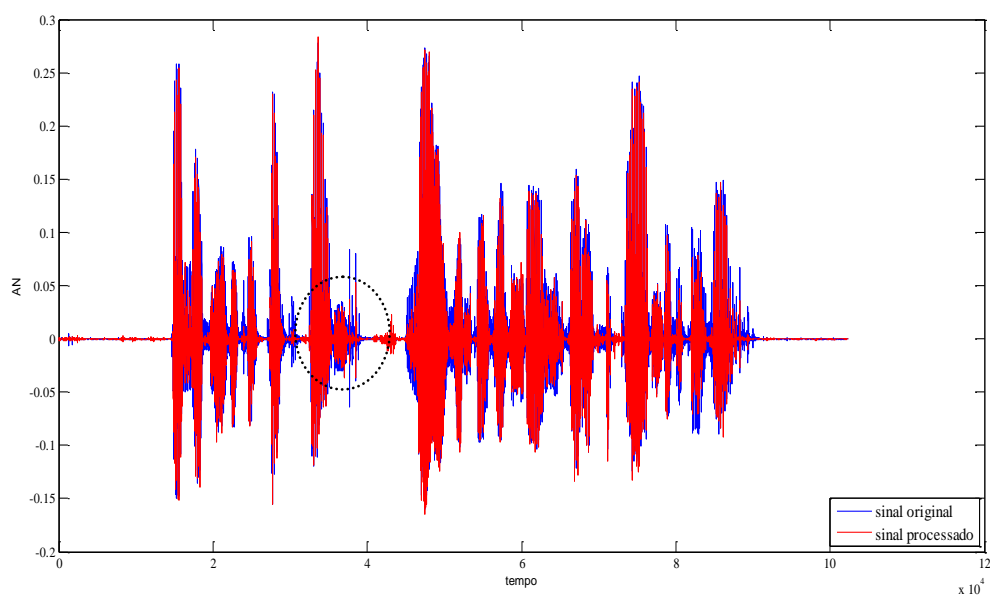
Figura 39. Sinal de voz ruidoso  $y(n)$ , obtido por contaminação do sinal original  $x(n)$  com ruído em restaurante (10 dB).



Fonte: Elaborado pelo próprio autor.

A Figura 40 traz, em azul, a forma de onda do sinal limpo  $s(n)$  e, em vermelho, a forma de onda do sinal processado  $\hat{y}(n)$  usando o método proposto.

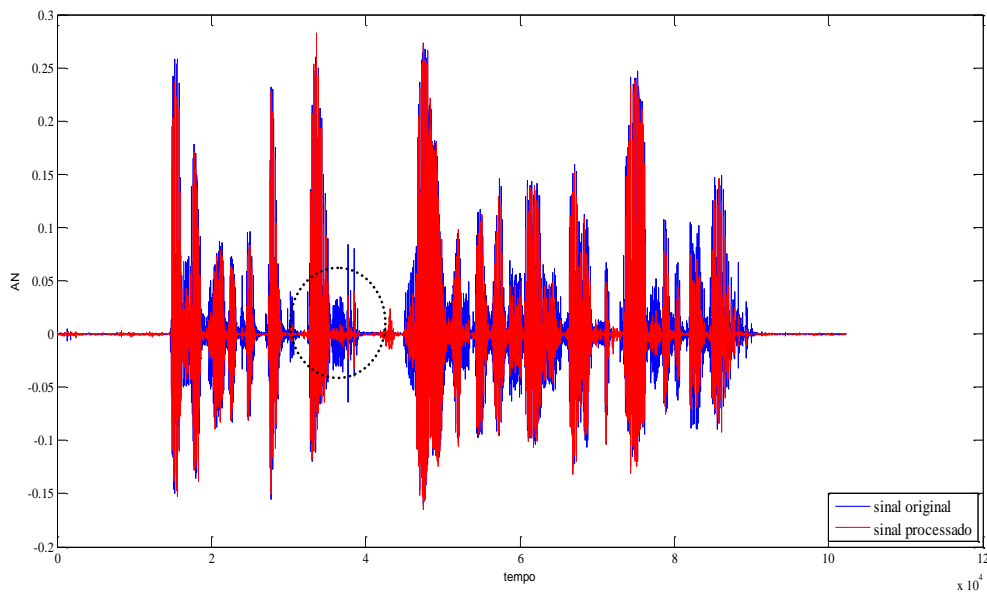
Figura 40. Sinal de voz limpo  $s(n)$  (azul) e sinal processado  $\hat{y}(n)$  (vermelho) usando o método proposto a partir do sinal de voz limpo contaminado com ruído em restaurante.



Fonte: Elaborado pelo próprio autor.

Para se obter uma análise visual da influência do operador *Dif*, na conservação da amplitude do sinal processado, considere também a Figura 41. Nela consta, em azul, a forma de onda do sinal  $s(n)$  e, em vermelho, a forma de onda do sinal processado  $\hat{y}(n)$  pelo método proposto, porém, sem usar o operador *Dif*.

Figura 41. Sinal de voz limpo  $s(n)$  (azul) e sinal processado  $\hat{y}(n)$  (vermelho) usando o método proposto, sem o operador *Dif*, a partir do sinal de voz limpo contaminado com ruído em restaurante.



Fonte: Elaborado pelo próprio autor.

Analisando as Figuras 40 e 41, observa-se que o operador *Dif* influencia diretamente para uma maior conservação da amplitude do sinal processado. Este fato é evidenciado ao longo de todo o sinal, porém, mais nítido na região circunscrita em ambas as figuras. Como era de se esperar, minimizando as perdas de amplitude dos coeficientes referentes à voz, através do operador *Dif*, consegue-se uma maior qualidade do sinal processado. A Tabela 8 traz as avaliações objetivas, ressaltando a influência do operador *Dif*.

Tabela 8. Avaliações objetivas para o sinal processado  $y(n)$ , em particular.

	Sem o operador <i>Dif</i>	Com o operador <i>Dif</i>
PESQ	2,912	3,150
SNR (dB)	35,937	36,006

Fonte: Elaborado pelo próprio autor.

Analisando a Tabela 8, verifica-se que o operador proposto *Dif* influencia diretamente na qualidade do sinal processado. Esta influência é mínima, porém positiva. Este fato é ratificado com o aumento da nota PESQ quando o mesmo é utilizado. Já para a redução de ruído, ele é indiferente. Note que as variações da SNR são insignificantes.