

UNIVERSIDADE FEDERAL DE SANTA CATARINA

CENTRO TECNOLÓGICO

DEPARTAMENTO DE ENGENHARIA ELÉTRICA

**DESENVOLVIMENTO DE UM SIMULADOR DE
APARELHO AUDITIVO DIGITAL**

CNPQ/PIBIC

THIAGO HIDEKI AKINAGA

Orientador:

Márcio Holsbach Costa

Florianópolis

2010

RESUMO

Este trabalho apresenta um programa de simulação de um aparelho auditivo microprocessado contendo quatro tipos avançados de processamento de sinal de fala: redução de ruído de banda larga, redução de ruído de banda estreita, microfones direcionais adaptativos e compressão de amplitude. Para cada tipo de processamento foram incluídas diversas técnicas apresentadas na literatura da área, permitindo também a inclusão posterior de outras que venham a ser disponibilizadas. O sistema foi desenvolvido no aplicativo MATLAB[®] e apresenta uma interface com o usuário que permite de forma fácil e intuitiva tanto a configuração dos subsistemas quanto a análise, comparação e visualização dos resultados de processamento. O programa desenvolvido é uma ferramenta interessante para o estudo, compreensão e projeto de técnicas avançadas de processamento de sinais em aplicações de auxílio à audição.

Palavras-chave: *Software* de simulação; Aparelho auditivo digital; Processamento de sinais biomédicos.

AGRADECIMENTOS

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico pelo apoio financeiro.

À Grasieli de Oliveira Ramos pela revisão do texto e sugestões no *design* da interface.

À Ângela Crepaldi pelas discussões e contribuições para tornar a interface mais acessível.

A todos os colegas do Laboratório de Pesquisas em Processamento Digital de Sinais.

SUMÁRIO

RESUMO	i
AGRADECIMENTOS	ii
SUMÁRIO	iii
1 INTRODUÇÃO	1
1.1 Objetivos.....	1
1.2 Justificativa.....	1
2 FUNDAMENTOS TEÓRICOS	2
2.1 Anatomia e fisiologia do ouvido humano	2
2.2 Classificação da deficiência de audição	2
2.3 Técnicas de auxílio à audição	3
2.4 Aparelho auditivo digital	3
2.4.1 Microfones direcionais	4
2.4.2 Redução de ruído	4
2.4.3 Compressão de amplitude.....	5
2.4.4 Cancelamento de realimentação	5
3 MATERIAIS E MÉTODOS.....	7
3.1 Estrutura do programa	7
3.2 Microfones direcionais	8
3.3 Redução de ruído não correlacionado	10
3.4 Redução de ruído correlacionado	16
3.5 Compressão de amplitude.....	18
4 RESULTADOS	22
4.1 Ferramenta de simulação	22
5 DISCUSSÃO	27
6 PERSPECTIVAS FUTURAS	28
REFERÊNCIAS	29

1 INTRODUÇÃO

Há mais de 50 anos surgiam os primeiros aparelhos auditivos compostos por tubos de vácuo e posteriormente por transistores. Ao longo do tempo houve um expressivo avanço tecnológico no desenvolvimento de dispositivos semicondutores que permitiram o acesso, a um baixo custo, de microprocessadores com grande capacidade computacional. Tal processo evolutivo permitiu o uso de técnicas complexas de processamento digital de sinais em aparelhos auditivos. Os esforços por parte dos pesquisadores para desenvolver novos dispositivos e novas estratégias para melhorar a qualidade desses dispositivos aumentam a cada ano.

1.1 Objetivos

O objetivo deste trabalho é desenvolver um software que permita a simulação de técnicas avançadas de processamento de sinais voltadas para aparelhos auditivos digitais.

1.2 Justificativa

As ferramentas de simulação encontram-se presentes em grande parte das pesquisas científicas. Algumas vantagens da simulação computacional são: a economia de custos e a confiabilidade. Simuladores de aparelhos auditivos [1] permitem uma previsão do desempenho dos algoritmos de processamento antes que esses sejam implementados em um dispositivo real, permitindo a diminuição do tempo de desenvolvimento. Além disto, constituem-se como ferramentas interessantes para fins educacionais ou de popularização do conhecimento.

2 FUNDAMENTOS TEÓRICOS

Alguns conceitos de fundamental importância para a compreensão deste trabalho serão tratados neste capítulo. Para isto, espera-se que o leitor possua conhecimentos básicos sobre sinais e sistemas discretos, vetores, além de conceitos básicos sobre estatística.

Utilizar-se-á a seguinte notação para representação das equações matemáticas: matrizes e vetores serão denotados respectivamente por letras em negrito, maiúsculas e minúsculas, e variáveis escalares em letras minúsculas em itálico.

2.1 Anatomia e fisiologia do ouvido humano

Anatomicamente divide-se o ouvido humano em três partes: ouvido externo, ouvido médio e ouvido interno. O primeiro tem como principal função captar o som, e repassá-lo ao ouvido médio. Quando recebido, este som faz vibrar a membrana timpânica, onde a vibração é captada e amplificada por três ossículos chamados de bigorna, martelo, e estribo. Após isto, o movimento vibratório é repassado a outra membrana que divide o ouvido médio do interno. O movimento do estribo faz o fluido coclear se deslocar, esse fluido fica no interior da cóclea que possui o formato parecido com um caracol. A cóclea divide-se em três compartimentos separados por duas membranas (membrana Reissner e membrana basilar). Ali, células ciliadas possuem diferentes sensibilidades ao som em função da frequência, sendo a conversão para impulsos elétricos feita pelas fibras individuais (aproximadamente 30000 fibras) da porção auditiva do VIII nervo craniano.

2.2 Classificação da deficiência de audição

A perda total ou parcial da audição caracteriza a deficiência auditiva ou surdez, dependendo da origem. São diversos os motivos que levam os indivíduos a apresentar algum tipo de deficiência auditiva. As perdas podem ser classificadas basicamente em três padrões: condutiva, neurossensorial e mista. A perda por condução ocorre quando há existência de algum empecilho no caminho da transmissão do som do ouvido externo até o ouvido interno. Já na perda neurossensorial uma lesão nas células ciliadas sensoriais e nervosas impede que o som seja processado na cóclea e enviado para o cérebro. Por fim, a perda mista se caracteriza por uma deficiência em ambos os sistemas. Em relação ao nível da perda auditiva têm-se quatro classificações: leve, média, severa e profunda.

2.3 Técnicas de auxílio à audição

Para auxiliar os deficientes auditivos podem ser utilizadas alternativas como próteses auditivas comuns ou implantes cocleares. O primeiro tipo resolve grande parte dos problemas de audição, inclusive alguns dos mais graves. Entretanto, para que a solução seja efetiva, faz-se necessária a existência de pelo menos uma parcela dos nervos ciliares originais. Caso isto não ocorra, ou seja, o indivíduo não alcance uma discriminação maior de 40% em testes para reconhecimento de frases em apresentação aberta, deve-se verificar a possibilidade de se fazer o implante coclear [2].

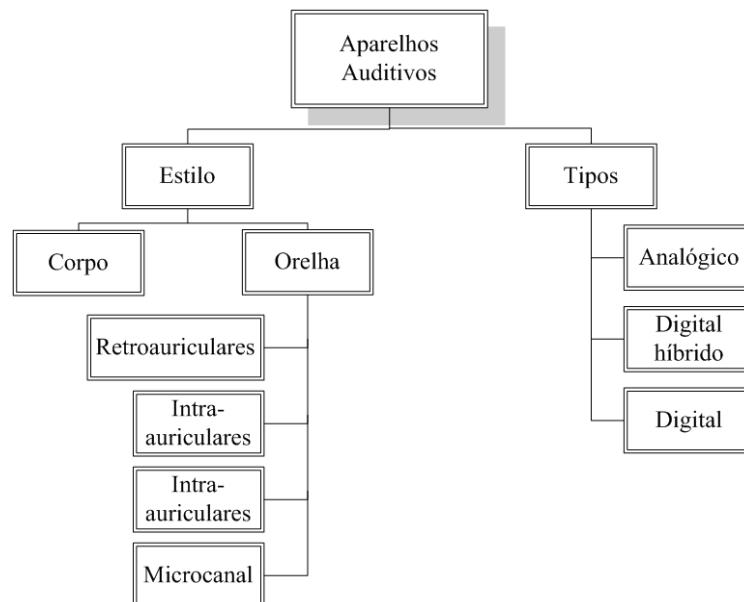


Figura 1 – Esquema resumido com os tipos e estilos de aparelhos auditivos.

2.4 Aparelho auditivo digital

Os aparelhos auditivos podem ser divididos em três tipos, analógico, digital híbrido e digital. O diagrama da Figura 1 mostra as principais categorias dos dispositivos. Em decorrência da rápida evolução tecnológica, possibilitando o aumento da capacidade computacional e a diminuição dos custos e tamanho dos sistemas microprocessados, os aparelhos auditivos digitais têm sido utilizados cada vez mais nos dias atuais. Entre as vantagens do aparelho auditivo digital destaca-se a possibilidade da utilização de técnicas complexas de processamento do sinal de fala, tais como: microfones direcionais adaptativos, técnicas de supressão ao ruído, compressão dinâmica adaptativa e cancelamento de realimentação. Contudo, a complexidade dos algoritmos de processamento acarreta um custo

mais elevado.

O diagrama em blocos de um aparelho auditivo digital é apresentado na Figura 2. Os diversos subsistemas que o compõe são descritos nas seções a seguir.

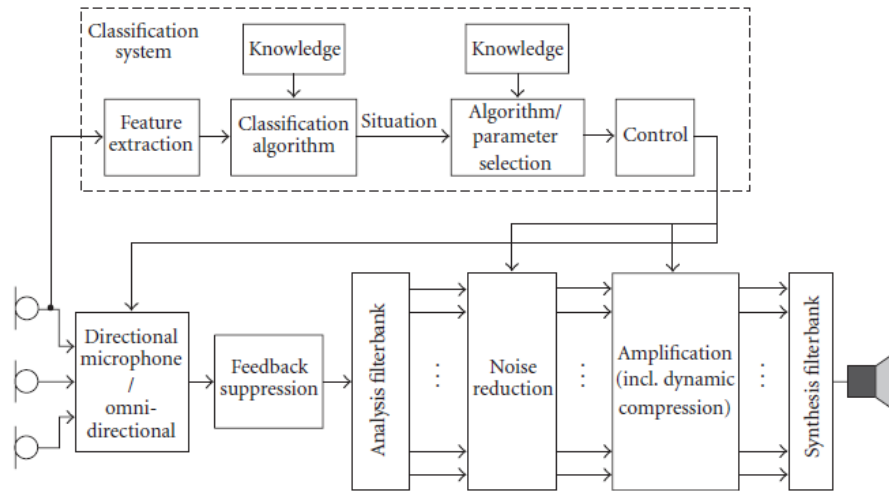


Figura 2 – Diagrama de blocos dos estágios de processamento de um aparelho auditivo digital. Extraído de [3].

2.4.1 Microfones direcionais

Os microfones direcionais são úteis nos aparelhos auditivos do tipo retroauriculares (BTE – *behind the ear*), onde ocorre uma perda da capacidade na determinação da direção da fonte sonora. Esta estratégia de processamento funciona bem em situações em que o sinal de interesse (geralmente a fala) e o ruído originam-se de fontes com posições espaciais diferentes. Há basicamente dois tipos de configuração dos aparelhos direcionais, o mais simples é chamado de direcionalidade fixa e o mais complexo denominado de direcionalidade adaptativa. Este último permite o rastreamento da direção de chegada da fala e/ou do ruído, permitindo uma maior atenuação das fontes sonoras não desejadas. Os aparelhos com microfones direcionais têm possibilitado nos últimos anos um aumento significativo da inteligibilidade da fala em ambientes ruidosos [3].

2.4.2 Redução de ruído

Nos modelos de aparelhos em que não é possível a implantação de mais de um microfone (para a implementação de direcionalidade na captação), ou ainda quando não existem diferenças significativas na posição espacial das fontes de ruído e de sinal de interesse, surge a necessidade da implementação de técnicas de redução de ruído. Os

sistemas de redução de ruído encontram-se presentes em grande parte dos dispositivos de auxílio à audição e podem ser divididos em sistemas de redução de ruído de banda larga (neste texto chamado de não-correlacionado) e sistemas de redução de ruído de banda estreita (aqui denominado de correlacionado) [3]. O primeiro tem o objetivo de reduzir ruídos sonoros caracterizados por componentes em frequência ao longo de todo o espectro. Em geral os métodos para supressão de ruído não-correlacionado utilizam informações estatísticas da fala e do ruído para melhorar a relação sinal ruído (SNR – *signal to noise ratio*). Por outro lado, quando o ruído de fundo possui componentes em faixas de frequência bem determinadas, nas quais as componentes em frequência da fala não são significativos, técnicas de filtragem seletiva em frequência, tais como o uso de filtros rejeita-faixa (*notch*), ou outros mais elaborados (como os abordados neste trabalho), são capazes de alcançar excelentes resultados com baixo custo computacional.

2.4.3 Compressão de amplitude

Uma pessoa com deficiência auditiva apresenta uma redução na faixa dinâmica das frequências audíveis quando comparada a uma pessoa com audição normal. Em caso de perda neurossensorial observa-se um fenômeno típico conhecido como *loudness recruitment* onde há um rápido crescimento na percepção da intensidade sonora com pequenos incrementos na intensidade do som. O sistema de compressão de amplitude tem como principal objetivo atuar no controle do ganho dado ao longo do processo de amplificação do sinal de fala para manter o sinal audível e confortável. Em geral o sistema reduz a faixa dinâmica do sinal de entrada para se adequar à excursão dinâmica apresentada pela audição do deficiente auditivo. Com isto consegue-se reduzir as perdas de informações durante a fala, música, sinais de alerta, entre outras informações.

2.4.4 Cancelamento de realimentação

A necessidade de grande amplificação do som para compensar as perdas dos indivíduos portadores de deficiência auditiva pode ocasionar uma realimentação acústica do sinal produzido pelo alto-falante no próprio microfone. Este fenômeno ocasiona um desconforto ao usuário do aparelho e limita o desempenho do dispositivo, uma vez que os valores de ganho do sistema de amplificação são limitados [4]. O caminho de realimentação depende do estilo do aparelho auditivo, do tamanho da abertura (passagem entre o microfone

e o alto-falante), da distância ao ouvido de objetos como chapéus, celulares ou mãos. Nos dois primeiros casos (estáticos) pode-se ajustar o algoritmo de cancelamento de realimentação de forma a limitar o ganho máximo do sistema em malha fechada para manter a estabilidade. Contudo no cotidiano do indivíduo é comum encontrar situações (atender telefone, por exemplo) onde se necessita de um comportamento dinâmico do sistema de cancelamento da realimentação [3].

3 MATERIAIS E MÉTODOS

Os materiais e métodos utilizados para o desenvolvimento do *software* de simulação de um aparelho auditivo digital são apresentados nesse capítulo. São apresentadas informações sobre a ferramenta de programação utilizada, a idéia geral da interface desenvolvida e o equacionamento básico de algumas das técnicas implementadas no *software*.

3.1 Estrutura do programa

Para construção da interface gráfica utilizou-se a ferramenta de desenvolvimento GUIDE (*Graphical User Interface Development Enviroment*) do *software* MATLAB®. Entre as vantagens desta ferramenta, tem-se: (1) a facilidade de programação devido à linguagem de alto nível; (2) a facilidade de integração com outras rotinas já disponibilizadas para o MATLAB; (3) a facilidade de migração entre sistemas operacionais [5].

Para implementação do simulador de um aparelho auditivo dividiu-se o sistema em três grandes blocos (Figura 3). O primeiro estágio consiste na entrada dos sinais de áudio a serem processados. O segundo estágio é composto pela etapa de processamento propriamente dita. Já o último estágio é responsável pela exibição dos resultados.

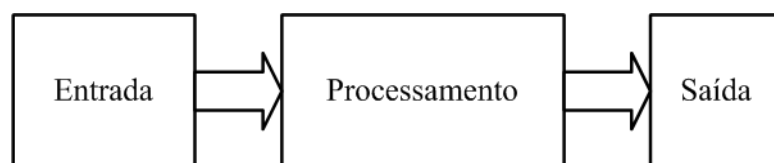


Figura 3 – Divisão em blocos do sistema de simulação.

No bloco de entrada do sistema tem-se a possibilidade de escolher o número de microfones do aparelho auditivo, podendo ser um ou dois. Para o caso de um único microfone é possível realizar a leitura de dois sinais de entrada, correspondendo a um sinal de fala e um sinal de ruído. Quando o subsistema de microfones direcionais é selecionado pelo usuário, permite-se a inclusão de um segundo microfone e um terceiro sinal, relacionado à fonte de ruído proveniente de uma direção diversa dos sinais anteriores.

No bloco de processamento encontram-se os subsistemas de microfone direcional adaptativo, redução de ruído (não-correlacionado e correlacionado) e compressão de amplitude. O sistema processa a informação bloco a bloco, ou seja, cada subsistema processa

toda a informação para passar ao estágio seguinte.

Como saída do sistema tem-se os sinais resultantes de cada estágio de processamento, além dos sinais de entrada e saída do aparelho auditivo.

3.2 Microfones direcionais

Uma das formas de implementar a direcionalidade utilizando dois microfones próximos em arranjo *end-fire* é fazer com que o atraso τ do sistema da Figura 4 seja modificado para que seu valor seja igual ao valor de atraso da propagação do ruído entre os microfones. No caso da modificação do atraso realizada de forma automática, tem-se a direcionalidade adaptativa.

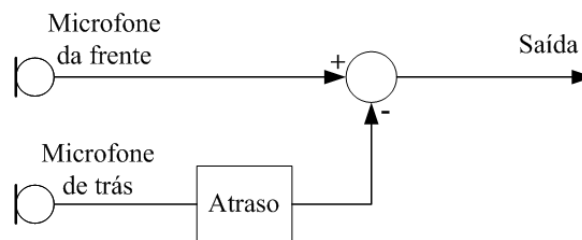


Figura 4 - Típico sistema de processamento de direção com dois microfones omnidirecionais em arranjo *end-fire*.

Seja d a distância entre os dois microfones e c a velocidade do som no ar, pode-se escrever o atraso como $\tau = k(d/c)$. Variando o ganho k obtêm-se padrões de recepção (polares) diferentes. A direção exatamente na frente do usuário do aparelho auditivo é representada como 0° , enquanto que 180° representa a direção diretamente atrás do usuário.

Para implementar esta unidade de atraso é necessário fazer uso de um filtro que realize um atraso fracionário. Para tal, utilizou-se o filtro de Thiran [6] que é um filtro passatudo com atraso de grupo maximamente plano. O filtro de Thiran também foi utilizado para mimetizar o sistema físico de forma a simular o atraso do som recebido pelos dois microfones em função da localização espacial das fontes.

O sistema proposto por Luo *et. al* em [7] é apresentado na Figura 5, onde os sinais recebidos no microfone frontal e no microfone traseiro são denominados $f(n)$ e $b(n)$, respectivamente. Definem-se ainda: d/c como a unidade de atraso entre os dois microfones; $e(n)$ saída do sistema; $W(n)$ ganho adaptativo; $z(n)$ saída da unidade de processamento do ganho adaptativo.

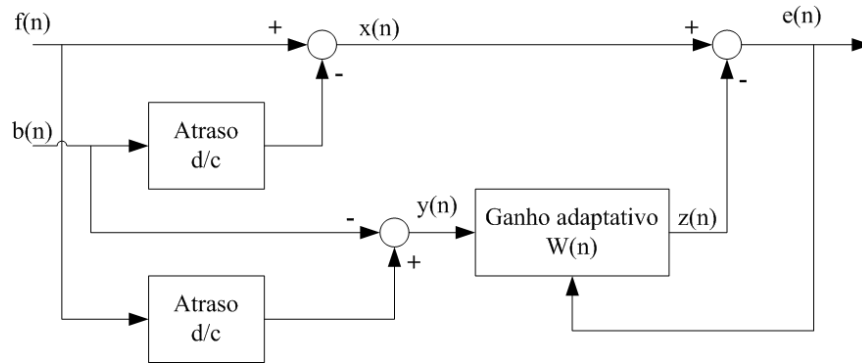


Figura 5 - Esquemático do sistema de formação de nulo adaptativo proposto em [7].

Da Figura 5 pode-se demonstrar que o padrão polar de $x(n)$ é um cardióide com nulo em 180° e o padrão polar de $y(n)$ é um cardióide com nulo em 0° . O padrão polar da saída do sistema inteiro é decorrente da combinação ponderada dos sinais $x(n)$ e $y(n)$, através do ganho $W(n)$. A relação do nulo da saída do sistema com o ganho $W(n)$ é dada pela Eq. (1), onde θ_{nulo} é o ângulo de nulo ao longo da linha entre os dois microfones.

$$W(n) = -\frac{1 + \cos(\theta_{nulo})}{1 - \cos(\theta_{nulo})} \quad (1)$$

Sendo τ o atraso na Figura 4, então o tempo de atraso relaciona-se com o ganho segundo a Eq. (2).

$$\tau = \frac{1 + W(n) d}{1 - W(n) c} \quad (2)$$

O ganho ótimo W_{opt} que minimiza a potência do ruído pode ser obtido como mostrado pela Eq. (3), onde R_{xx} é a potência de $x(n)$, R_{yy} a potência de $y(n)$ e R_{xy} é a correlação cruzada de $x(n)$ e $y(n)$.

$$W_{opt} = \frac{R_{xy}}{R_{yy}} \quad (3)$$

A partir desses resultados pode-se mostrar que o valor quadrático médio do sinal de saída $e(n)$ é uma função quadrática de $W(n)$ com um único mínimo W_{opt} . Com base no apresentado, pode-se demonstrar que o nulo do diagrama polar será direcionado para a direção de chegada do ruído quando o sinal de fala e o ruído estiverem presentes simultaneamente.

Para atualizar adaptativamente o ganho ótimo W_{opt} pode-se utilizar algoritmos adaptativos como o algoritmo LMS, o NLMS, ou ainda o RLS (*Recursive Least Squares*).

Entretanto, no presente sistema seguiu-se a proposta de [7] onde é utilizado um algoritmo adaptativo *frame-by-frame*¹. Assim, para calcular o ganho ótimo, primeiramente calcula-se a estimativa da potência de $y(n)$ e a correlação cruzada entre $x(n)$ e $y(n)$ no m -ésimo *frame*, onde M é a quantidade de amostras em cada *frame*. Utilizando fatores ajustáveis α tal que $0 \leq \alpha \leq 1$, tem-se um processamento *frame-by-frame* mais suave, podendo melhorar a estimativa de W_{opt} . As correlações são calculadas conforme as equações (4) e (5).

$$\hat{R}_{xy}(m) = \frac{\alpha}{M} \sum_{n=1}^M x(n)y(n) + (1 - \alpha)\hat{R}_{xy}(m - 1) \quad (4)$$

$$\hat{R}_{yy}(m) = \frac{\alpha}{M} \sum_{n=1}^M x(n)y(n) + (1 - \alpha)\hat{R}_{yy}(m - 1) \quad (5)$$

3.3 Redução de ruído não correlacionado

A escolha das técnicas para compor o sistema de redução de ruído não-correlacionado foi realizada levando-se em conta os critérios de importância científica e aplicabilidade prática. A técnica conhecida por logaritmo do mínimo erro quadrático médio (LogMMSE - *logarithm minimum mean square error*), desenvolvida por Ephraim e Malah [8] é uma técnica utilizada como padrão de comparação de algoritmos de redução de ruído em sinais de fala. Os mesmos autores apresentam em [9] outra técnica semelhante, conhecida por mínimo erro quadrático médio (MMSE - *minimum mean square error*), também bastante difundida no meio científico. Além destas duas técnicas mais conhecidas, optou-se por adicionar mais duas técnicas que podem ser implementadas em aparelhos auditivos devido ao seu baixo custo computacional. Estas técnicas são conhecidas como preditor adaptativo [10] (ALE - *adaptive line enhancement*) e filtro Wiener [14] (*Wiener filter*).

A estrutura do ALE apresentada na Figura 6 consiste basicamente em um preditor *backward* de ordem M , onde M representa a quantidade de atrasos que se necessita para descorrelacionar o ruído associado $z(n)$ (considerado aleatório, gaussiano e branco) do sinal de entrada $x(n)$ (considerado determinístico ou muito correlacionado). Supondo que o sinal $y(n)$ (saída do filtro adaptativo) e o sinal $x(n)$ (sinal contaminado) possuem um sinal de fala correlacionado, porém associado a um ruído descorrelacionado, o filtro adaptativo de

¹ Define-se *frame* como o trecho de amostras analisado, geralmente se utiliza alguns milissegundos, cerca de 20ms a 40ms, devido a não estacionariedade do sinal de fala.

resposta finita ao impulso (FIR – *finite impulse response*), mostrado na Figura 7, consegue estimar a parcela referente ao sinal de fala presente em $x(n)$, baseado em amostras passadas do sinal contaminado, minimizando o erro quadrático médio entre o sinal desejado e o sinal de saída.

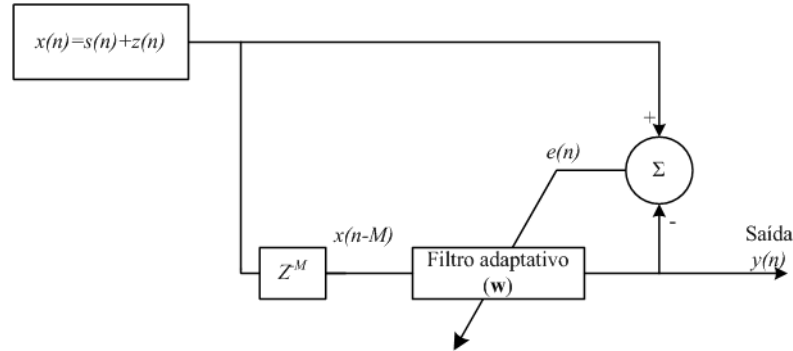


Figura 6 – Diagrama de blocos do *adaptive line enhancement*.

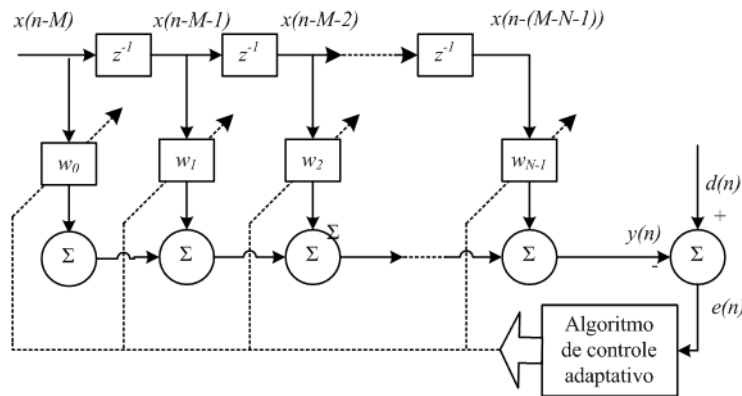


Figura 7 – Diagrama de blocos do filtro adaptativo.

Dado $x(n)=s(n)+z(n)$, $x(n)$ representa o sinal de voz contaminado, $s(n)$ representa o sinal de voz e $z(n)$ representa o sinal de ruído. A estimação da saída do filtro adaptativo é dada por:

$$y(n) = \sum_{m=0}^{N-1} w_m(n)x((n-M)-m) = \mathbf{w}^T \mathbf{x} \quad (6)$$

onde:

$$\mathbf{w} = [w_0(n) \quad w_1(n) \quad \cdots \quad w_{N-1}(n)]^T \text{ e}$$

$$\mathbf{x} = [x(n-M) \quad x(n-M-1) \quad \cdots \quad x(n-M-N+1)]^T$$

são respectivamente os coeficientes do filtro FIR e o sinal de entrada do combinador linear, cujo comprimento é determinado por N . O valor do erro instantâneo $e(n)$ é dado pela diferença entre o sinal desejado $d(n)$ e a saída do filtro $y(n)$.

Widrow e Hopf desenvolveram o algoritmo adaptativo LMS (*Least Mean Square*) que aproxima o erro quadrático médio ($E\{e^2(n)\}$) através do erro quadrático instantâneo ($e^2(n)$). Diante da complexidade de se obter diretamente o ponto que minimiza o erro quadrático médio, conhecido como a solução de Wiener, métodos de estimativa de gradiente são úteis para este fim. A equação de atualização dos coeficientes do algoritmo LMS é dada por:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e(n)\mathbf{x}(n) \quad (7)$$

Na Eq. (7) a constante μ representa o passo de aprendizagem ou adaptação. Nota-se pela Figura 6 e Figura 7 que o sinal $d(n)$ corresponde a $x(n)$. Com isto observa-se que devido ao atraso de decorrelação, a amostra predita baseada em amostras passadas, está atrasada em relação ao sinal $d(n)$. Dessa forma, o valor de M deve ser o menor possível para que não ocorra uma distorção significativa do sinal de interesse, e grande o suficiente para que o ruído seja decorrelacionado.

As técnicas LogMMSE e MMSE são estimadores não-lineares da magnitude do espectro do sinal de interesse (sinal de fala sem ruído) que utilizam modelos estatísticos e critérios de otimização da função densidade de probabilidade. Estes estimadores, quando aplicados na redução de ruído associado a um sinal de voz, utilizam os coeficientes da DFT (*discrete Fourier transform*) do sinal de voz contaminado com ruído. Os esforços se concentram em encontrar os coeficientes da DFT do sinal de voz sem ruído ambiente. Para compreensão dos estimadores LogMMSE e MMSE, se faz previamente uma breve abordagem dos estimadores *maximum likelihood*.

Seja $y(n)=x(n)+d(n)$ o sinal de voz contaminado, onde $x(n)$ representa o sinal de voz, e $d(n)$ representa o sinal de ruído associado. Aplicando a transformada de Fourier em ambos os lados, tem-se:

$$Y(\omega_k) = X(\omega_k) + D(\omega_k) \quad (8)$$

onde $\omega_k=2\pi k/N$, com $k=0,1,2,\dots,N-1$, sendo N o tamanho do *frame* analisado em amostras. Esta abordagem causa certa correlação entre os coeficientes da DFT. Entretanto na prática se usa janelamento e sobreposição (*overlap*). Apesar de esta abordagem violar a consideração de não-correlação entre os coeficientes, o erro é desprezível na prática [11]. Escrevendo a Eq. (8) na forma polar, tem-se:

$$Y(\omega_k) = Y_k e^{j\theta_y(k)} = X_k e^{j\theta_x(k)} + D_k e^{j\theta_d(k)} \quad (9)$$

Na abordagem do estimador *maximum likelihood* assumem-se X_k e θ_k

desconhecidos, porém determinísticos. A função densidade probabilidade de $D(\omega_k)$ é assumida ser de média zero, gaussiana complexa e de variância $\lambda_d/2$ da parte real e $\lambda_d/2$ da imaginária. Com isto a função densidade de probabilidade do sinal de voz contaminado terá variância λ_d e média $X_k e^{j\theta_s(k)}$ sendo gaussiana também.

Definido $\gamma_k = Y_k^2 / \lambda_d(k)$ como a SNR *a posteriori*, ou seja, baseada na observação do sinal corrompido, o espectro estimado do sinal de fala obtido por meio da otimização da função densidade de probabilidade em relação à X_k , pode ser escrito como:

$$\hat{X}(\omega_k) = \left[\frac{1}{2} + \frac{1}{2} \sqrt{\frac{\gamma_k - 1}{\gamma_k}} \right] Y(\omega_k) = G_{ML}(\gamma_k) Y(\omega_k) \quad (10)$$

Se assumirmos que o sinal de voz pode ser modelado como a composição de variáveis aleatórias independentes, gaussianas de média zero e variância λ_x desconhecida e determinística, pode-se obter uma estimativa para X_k , de maneira análoga ao caso anterior, dada pela Eq. (11),

$$\begin{aligned} \hat{X}_k &= \sqrt{Y_k^2 - D_k^2} = \sqrt{\frac{Y_k^2 - \lambda_d(k)}{Y_k^2}} Y(\omega_k) \\ &= \sqrt{\frac{\gamma_k - 1}{\gamma_k}} Y(\omega_k) \\ &= G_{PS}(\gamma_k) Y(\omega_k) \end{aligned} \quad (11)$$

onde esta estimativa é conhecida como *spectrum subtraction estimator*. Com algumas manipulações consegue-se expressar o estimador *Wiener filter* pela Eq. (12).

$$\begin{aligned} \hat{X}(\omega_k) &= \frac{\lambda_x(k)}{\lambda_x(k) + \lambda_d(k)} Y(\omega_k) \\ &= \frac{\gamma_k - 1}{\gamma_k} Y(\omega_k) \\ &= G_{PS}^2 Y(\omega_k) \end{aligned} \quad (12)$$

Na abordagem com estimadores bayesianos os parâmetros são considerados aleatórios. Com isto precisa-se estimar a variável aleatória utilizando critérios baseados no teorema de Bayes, que permite utilizar o conhecimento *a priori* dos parâmetros para agregar melhorias aos estimadores. Os dois estimadores apresentados são o MMSE e o LogMMSE.

Devido à importância da STSA (*short time spectral amplitude*) na qualidade e inteligibilidade da fala, muitos autores propuseram métodos para obter a estimativa da solução ótima a partir do espectro do sinal contaminado. Estes estimadores são baseados na

minimização do erro quadrático médio entre a magnitude real e a magnitude estimada (Eq. (13)), ou também conhecidos como estimadores MMSE.

$$e = E\{(\hat{X}_k - X_k)^2\} \quad (13)$$

Minimizando a esperança, utilizando o critério de Bayes, chega-se em:

$$\hat{X}_k = E[X_k | Y(\omega_0)Y(\omega_1)\dots Y(\omega_{N-1})] \quad (14)$$

Onde a Eq. (14) representa a média *a posteriori* da função densidade de probabilidade de X_k . O problema está em determinar a função de probabilidade dos coeficientes da DFT do ruído e da fala para o cálculo da média. Para contornar estes problemas Ephraim e Malah propuseram um modelo que assume:

1. Os coeficientes da transformada de Fourier têm distribuição de probabilidade gaussiana. A média dos coeficientes é zero e a variância dos coeficientes é variante no tempo;
2. Os coeficientes da transformada de Fourier são estatisticamente independentes e também não correlacionados.

Isto possibilita chegar à expressão que estima a magnitude do espectro de X_k dada por:

$$\begin{aligned} \hat{X}_k &= \sqrt{\lambda_k} \Gamma(1.5) \Phi(-0.5, 1; -\nu_k) \\ &= \frac{\sqrt{\pi}}{2} \frac{\sqrt{\nu_k}}{\gamma_k} \exp\left(-\frac{\nu_k}{2}\right) \left[(1 + \nu_k) I_0\left(\frac{\nu_k}{2}\right) + \nu_k I_1\left(\frac{\nu_k}{2}\right) \right] Y_k \\ &= G(\xi_k, \gamma_k) Y_k \end{aligned} \quad (15)$$

Onde $I_0(\cdot)$ e $I_1(\cdot)$ representam as funções de Bessel modificadas de ordem zero e um respectivamente. Os parâmetros da Eq. (15) são calculados por:

$$\nu_k = \frac{\xi_k}{1 + \xi_k} \gamma_k \quad (16)$$

$$\gamma_k = \frac{Y_k^2}{\lambda_d(k)} \quad (17)$$

$$\xi_k = \frac{\lambda_x(k)}{\lambda_d(k)} \quad (18)$$

onde a Eq. (17) representa a SNR *a posteriori* (relação sinal-ruído dos dados observados) e a Eq. (18) representa a SNR *a priori* (verdadeira relação sinal-ruído).

Estes estimadores LogMMSE são baseados na minimização do erro quadrático

médio entre o logaritmo da magnitude do sinal real e o logaritmo da magnitude do sinal estimado.

$$e = E \left\{ (\log X_k - \log \hat{X}_k)^2 \right\} \quad (19)$$

As referências sugerem que a minimização do erro quadrático médio da magnitude do logaritmo do espectro pode ser mais adequada para o processamento de fala [11]. A expressão que estima a magnitude do espectro é dada por:

$$\begin{aligned} \hat{X}_k &= \frac{\xi_k}{\xi_k + 1} \exp \left\{ \frac{1}{2} \int_{\nu_k}^{\infty} \frac{e^{-t}}{t} dt \right\} Y_k \\ &= G_{LSA}(\xi_k, \nu_k) Y_k \end{aligned} \quad (20)$$

Uma aproximação para se calcular a integral da Eq. (20) é dada por:

$$Ei(x) = \int_x^{\infty} \frac{e^{-x}}{x} dx \approx \frac{e^{-x}}{x} \sum_k \frac{k!}{x^k} \quad (21)$$

Na prática para calcular λ_d (variância do ruído) pode-se assumir o ruído estacionário e avaliar esta informação ao longo de um trecho do sinal no qual não seja detectada a existência de fala (para isso utiliza-se um detector de atividade vocal – VAD – *voice activity detector*). Para calcular ξ_k o problema é mais complexo, entretanto existem alguns métodos como, por exemplo, os métodos *maximum likelihood*, e *decision-directed*. O primeiro faz uma estimativa da relação sinal-ruído *a priori* dada por:

$$\hat{\xi}_k(m) = \max \left(\frac{1}{L} \sum_{j=0}^{L-1} \gamma_k(m-j) - 1, 0 \right) \quad (22)$$

onde m representa o *frame* em processamento. Na prática substitui-se o termo $1/L \sum_{j=0}^{L-1} \gamma_k(m-j)$ por uma média calculada recursivamente, dada por:

$$\bar{\gamma}_k(m) = a\bar{\gamma}_k(m-1) + (1-a) \frac{\gamma_k(m)}{\beta} \quad (23)$$

onde $0 \leq a < 1$ é uma constante passa baixas e $\beta \geq 1$ é uma fator de correção. Com isto a estimativa ($\hat{\xi}_k$) é calculada por:

$$\hat{\xi}_k(m) = \max(\bar{\gamma}_k(m) - 1, 0) \quad (24)$$

A segunda maneira de se estimar a relação sinal-ruído *a priori*, conhecida como *decision directed*, e pode ser expressa por:

$$\hat{\xi}_k(m) = a \frac{\hat{X}_k^2(m-1)}{\lambda_d(k, m-1)} + (1-a) \max(\gamma_k(m) - 1, 0) \quad (25)$$

Nota-se que o primeiro termo da Eq. (25) consiste na relação sinal-ruído *a priori* do *frame* anterior multiplicada por uma constante. Como condição inicial, utiliza-se:

$$\hat{\xi}_k(0) = a + (1-a) \max(\gamma_k(0) - 1, 0) \quad (26)$$

3.4 Redução de ruído correlacionado

O sistema de redução de ruído correlacionado processa o sinal decomposto em diversas bandas de frequência. Para decompor o sinal em diversos canais e processá-lo, utilizou-se uma estrutura semelhante à contida no microprocessador GA3280 da *Gennum Corporation* [12]. O diagrama em blocos do banco de filtros do GA3280 (TDFB - *time domain filter bank*) mostrado na Figura 8 é composto basicamente por quatro grandes módulos: um divisor de banda (*band splitter*); um detector de nível (*level detector*); um filtro passa-baixas (*smoothing filter*); e um limitador de amplitude (*output compression limiter*).

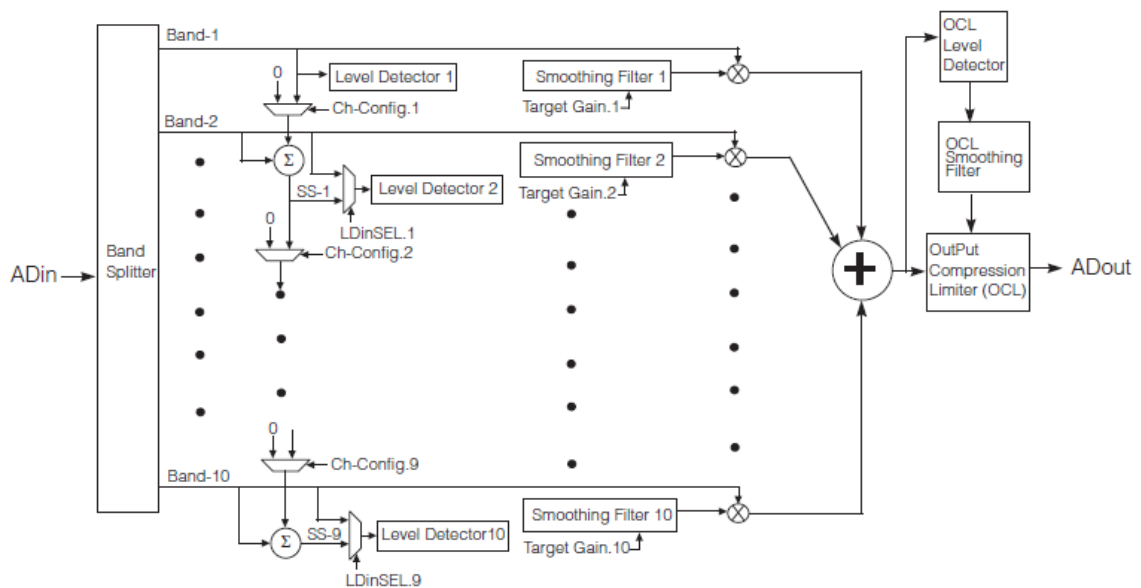


Figura 8 - Diagrama do *Time-Domain Filter Bank*. Extraído de [12].

O divisor de banda é formado por filtros passa-baixas e passa-altas. Os registradores definem o número de bandas em que se deseja separar o sinal de análise. No total são nove registradores que comandam o divisor de bandas, possibilitando no máximo 10 divisões na banda do sinal. Na Figura 9 mostra-se o funcionamento da interação entre os registradores e a divisão de bandas. O usuário define o número de bandas para decomposição (configurando

os registradores através da especificação da frequência de cruzamento) e o tipo de filtro e sua ordem. Até 6ª ordem os filtros são Butterworth, e da 6ª em diante pode-se escolher entre Butterworth e Elíptico. O manual sugere que se utilize o *Gennum Universal Integrated Development Enviroment (GUIDE™)* para calcular os coeficientes do divisor de bandas de acordo com a configuração escolhida.

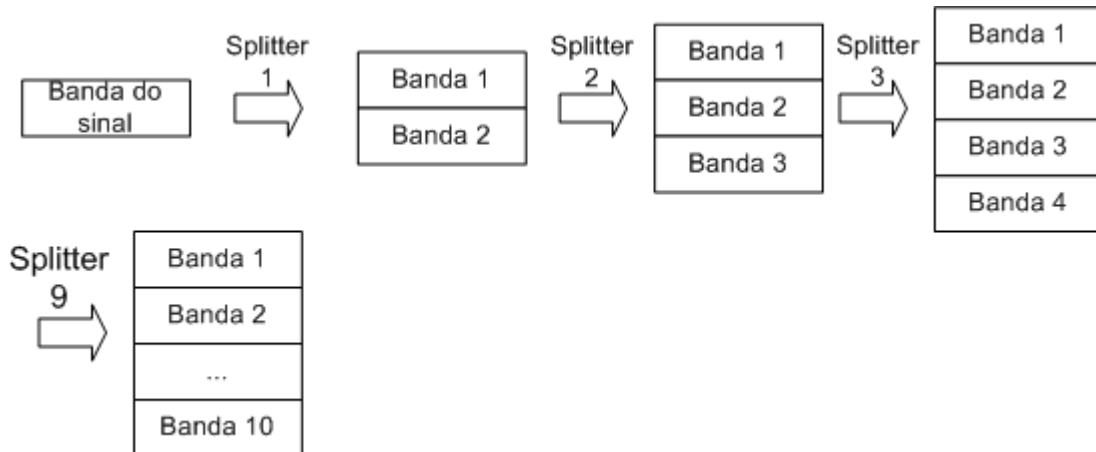


Figura 9 - Funcionamento da divisão de bandas.

Os detectores de níveis (*level detectors*) consistem em filtros recursivos de 1ª ordem, regidos pela função de transferência:

$$H(z) = \frac{2^{-(2+k)}}{1 - (1 - 2^{-(2+k)})z^{-1}} \quad (27)$$

onde o parâmetro k varia de 0 a 15 de acordo com as constantes de tempo de ataque (*attack*) e relaxamento (*release*) desejadas. Para implementação prática do detector de nível, implementou-se:

$$\begin{aligned} y(n) &= (1 - 2^{-(2+k)})y(n-1) + 2^{-(2+k)}x(n) \\ y(n) &= \beta y(n-1) + (1 - \beta)x(n) \end{aligned} \quad (28)$$

A Eq. (28) permite a implementação digital do detector de nível para aplicações em detectores de atividade vocal e compressores de amplitude. A constante k regula os tempos de *attack* e *release* do detector de nível. Pode-se demonstrar que a relação entre a constante k e os tempos em milissegundos é dada por:

$$t = \text{ceil} \left(\frac{-1}{\ln(1 - 2^{-(2+k)})} \right) \times T_s \times 1000 \quad (29)$$

Após a separação do sinal de entrada em bandas de frequência, utilizou-se um detector de atividade vocal em cada canal para supressão do ruído correlacionado naquela

frequência. O método utilizado baseia-se na proposta de Fang e Nilsson [13]. Inicialmente os detectores de níveis são utilizados para detecção do envelope do sinal de entrada. Em seguida utiliza-se o mesmo detector de nível com constante de tempo mais lenta para estimar o envelope do ruído presente. Após a estimação dos dois envelopes, faz-se uma comparação da ordem de grandeza das magnitudes dos envelopes da fala e do ruído, caso a magnitude da fala seja n vezes superior à magnitude do ruído, considera-se a existência de voz, caso contrário assume-se que o trecho em análise contém apenas ruído.

3.5 Compressão de amplitude

Para realizar a compressão de amplitude, o sistema utiliza o divisor de bandas explicado na seção anterior. São apresentadas duas técnicas, a primeira é chamada de compressão dinâmica (OCL – *output compression limiter*) e a outra conhecida por otimização adaptativa dinâmica (ADRO – *adaptive dynamic range optimization*).

Na compressão OCL monitora-se a potência média do sinal de entrada em cada canal disponibilizado pelo divisor de bandas, caso esta ultrapasse um limiar pré-estabelecido aplica-se um fator de atenuação ao sinal [14]. Como no processamento durante uma situação real não se tem acesso ao sinal por completo para o cálculo de sua potência média, utiliza-se a Eq. (30) para a estimação da potência média do sinal.

$$p(n) = \beta p(n-1) + (1-\beta)x^2(n) \quad (30)$$

onde $x(n)$ corresponde ao sinal de fala a ser processado, $p(n)$ equivale à potência média estimada de $x(n)$, e β a constante responsável pela constante de tempo. Se g corresponde ao ganho (menor que a unidade) aplicado ao sinal de entrada $x(n)$, então a saída $y(n)$ é dada por:

$$y(n) = gx(n) \quad (31)$$

Sendo assim, a potência do sinal de saída expressa em termos do ganho e da potência do sinal de entrada pode ser escrita como:

$$P_{out} = g^2 P_{in} \quad (32)$$

A Figura 10 mostra uma curva típica de compressão. Caso a potência do sinal de entrada do sistema do compressor seja menor que o limiar pré-estabelecido (p_{th}) a potência do sinal de saída será igual à potência do sinal de entrada. Se a potência do sinal de entrada estiver compreendida entre o limiar (p_{th}) e um valor máximo estabelecido (p_{max}), então a

saída será comprimida por uma razão s (conhecida por razão de compressão, ou *compression ratio*). Por fim, se a potência do sinal de entrada ultrapassar um valor máximo estabelecido (p_{max}) a saída será um valor constante igual à p_{sat} .

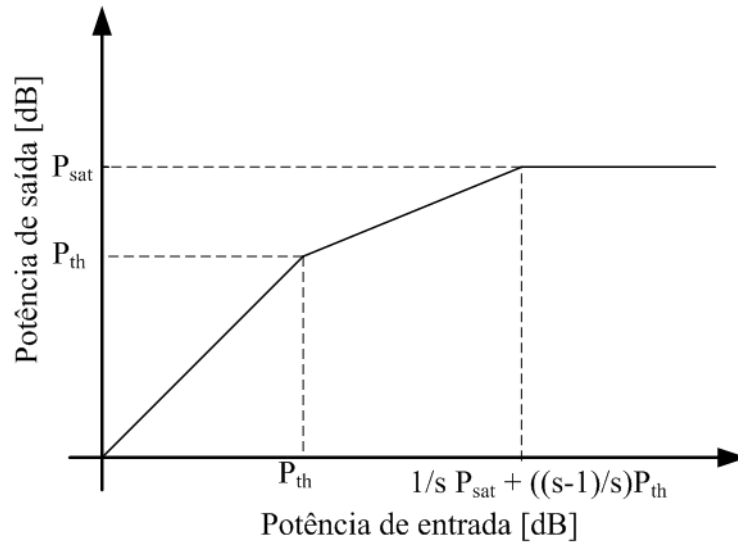


Figura 10 - Curva típica de uma compressão de amplitude. Adaptada de [14].

Com as definições feitas anteriormente, chega-se nas seguintes equações:

$$ganho = \begin{cases} 1 & , se p_{in} < p_{th} \\ \left(\frac{p_{in}}{p_{th}}\right)^{\frac{s-1}{s}} & , se p_{th} < p_{in} < \left(\frac{\frac{1}{1-s} p_{sat}^s}{p_{th}^s}\right) \\ \left(\frac{p_{sat}}{p_{in}}\right)^{\frac{1}{2}} & , se p_{in} > \left(\frac{\frac{1}{1-s} p_{sat}^s}{p_{th}^s}\right) \end{cases} \quad (33)$$

O gráfico da Figura 11 mostra a qualidade do som em função da constante de tempo e da razão de compressão s .

A definição de SPL (*sound pressure level*) se faz necessária para comparação com algumas curvas onde os valores fornecidos encontram-se em dB SPL. Por definição:

$$SPL = 10 \log_{10} \left(\frac{p_{rms}^2}{p_{ref}^2} \right) = 20 \log_{10} \left(\frac{p_{rms}}{p_{ref}} \right) \quad (34)$$

onde p_{ref} corresponde ao limiar da audição humana, $20\mu\text{Pa}$. Assim para definirmos uma constante p_{th} em função do seu valor em dB SPL, tem-se a seguinte relação:

$$p_{th} = p_{ref}^2 \times 10^{\frac{P_{th\text{SPL}}}{10}} \tag{35}$$



Figura 11 - Qualidade do sinal após a compressão em função das constantes de tempo e da razão de compressão. Adpatada de [16].

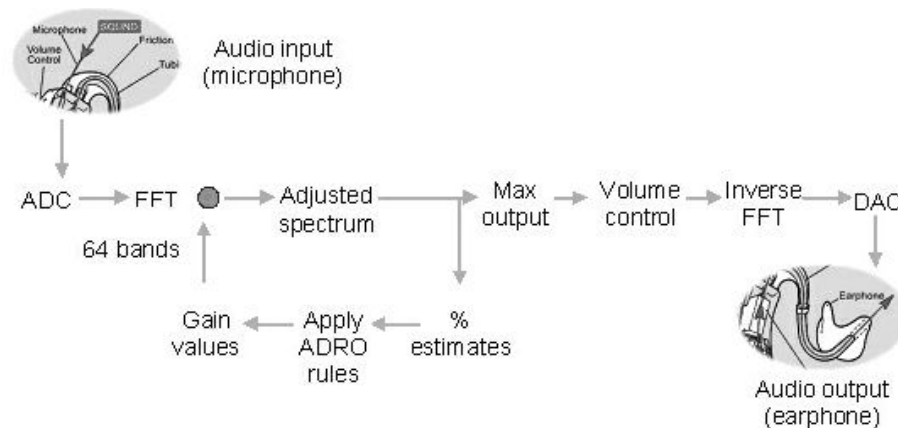


Figura 12 - Esquema básico de implementação da estratégia de amplificação ADRO. Extraída de [15].

O ADRO (*adaptive dynamic range optimization*) consiste em uma estratégia de amplificação que utiliza técnicas de processamento digital de sinais com o objetivo de proporcionar conforto, capacidade de audição, e inteligibilidade dos sons para usuários de aparelhos auditivos e/ou implantes cocleares. O esquema apresentado a seguir utiliza lógica difusa (*Fuzzy logic*) para aperfeiçoar a saída de cada banda de frequência do aparelho auditivo digital. A Figura 12 mostra o esquema de implementação do ADRO.

Na lista a seguir são definidas as quatro regras de compressão utilizadas para o esquema:

1. Conforto: Reduz lentamente o ganho (máximo 3dB/s equivale aproximadamente a 5s

de *attack* e *release*) no canal se o nível da saída exceder o limite de conforto mais que 10% do tempo;

2. **Audição:** Aumenta lentamente o ganho no canal se o nível da saída permanecer abaixo do limite de audição mais que 30% do tempo (esta regra só é testada caso a regra 1 esteja satisfeita);

3. **Proteção do ouvido:** Limita a saída em cada canal para não exceder um valor máximo;

4. **Supressão ao ruído:** Limita o ganho máximo em cada canal para evitar amplificar o ruído de fundo.

Para definir os limites de “conforto” e “audição” se faz necessário a implantação dos estimadores de frequência. Os estimadores de frequência funcionam da seguinte maneira. Há dois estimadores de frequência (um para o limiar de conforto e outro para o limiar de audição). Por meio de um processo iterativo mede-se a intensidade do som em intervalos regulares (20ms), compara-se o valor atual da intensidade com a estimativa atual, se for maior soma-se um valor a estimativa (passo de subida), caso contrário subtrai-se um valor a estimativa (passo de descida).

4 RESULTADOS

Este capítulo apresenta o sistema desenvolvido, mostrando suas funcionalidades, opções, e aplicações.

4.1 Ferramenta de simulação

Na Figura 13 apresenta-se a janela inicial do programa computacional desenvolvido. Observam-se os estágios descritos na Seção 3.1. Os botões “Fala” e “Ruído” formam o estágio de entrada. Os blocos referentes aos subsistemas de microfone direcional, redução de ruído não-correlacionado, redução de ruído correlacionado e compressão de amplitude constituem o bloco de processamento do sinal. O alto-falante representa a etapa de saída.

Os subsistemas não habilitados nesta versão do *software* encontram-se na cor cinza. As demais opções como os alto-falantes, o segundo microfone, e o botão “Simular”, somente ficarão habilitados quando o usuário assim o desejar. Por exemplo, para o alto-falante ficar disponível é imprescindível que se tenha um sinal de áudio. Para o botão “Simular” se tornar visível é preciso que toda a configuração para a simulação seja feita adequadamente.

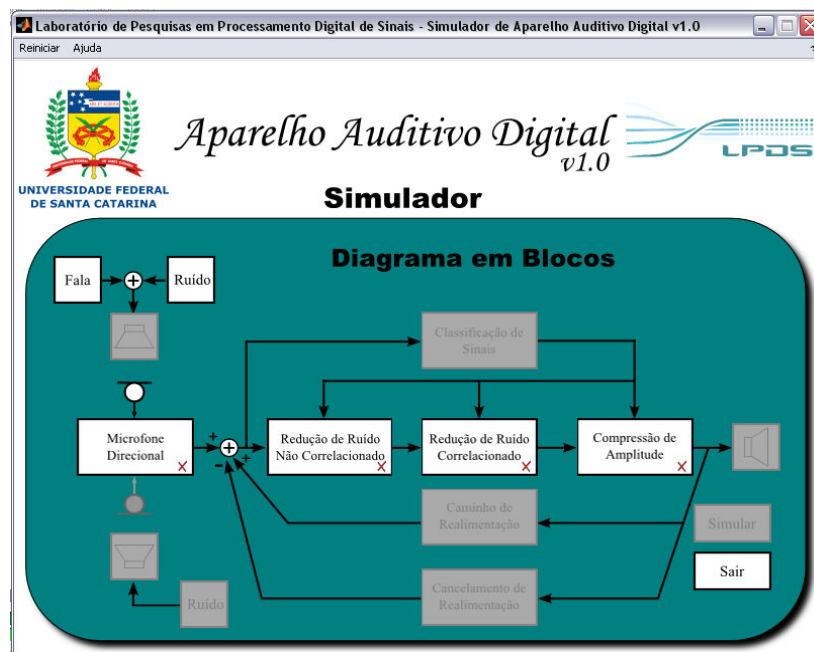


Figura 13 – Janela principal do *software* de simulação de um aparelho auditivo digital.

Para simular as diversas situações do cotidiano de um usuário de aparelho auditivo, foi implementada na etapa de entrada dois botões para carregar um sinal de fala e um sinal de ruído (formato *wave*), permitindo diversas combinações possíveis entre fala e ruído. Após a

leitura dos sinais, estes são somados de forma a simular a combinação dos sinais acústicos no ar (Figura 14a). Caso não seja de interesse utilizar uma das fontes sonoras pode-se desabilitar a fonte de fala ou ruído. A soma realizada possui um controle automático contra saturação.

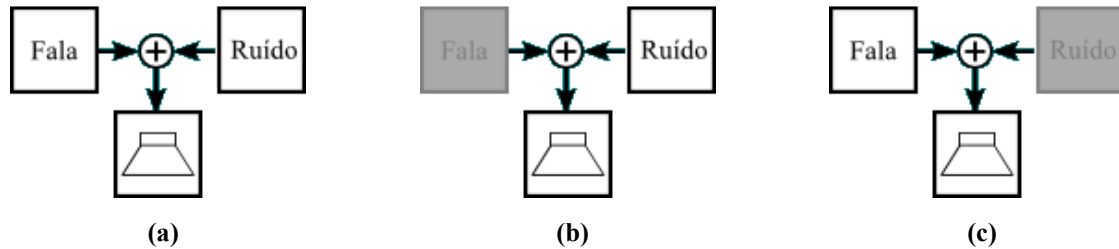


Figura 14 – Combinações para o sinal de entrada.

Após a leitura dos sinais de áudio pode-se ouvir, desenhar, e salvar o sinal enviado ao microfone do aparelho auditivo (fala associada ao ruído). Para possibilitar estas opções ao usuário, adicionou-se ao programa um menu que pode ser visto na Figura 15a. O controle durante a execução do sinal de áudio pode ser feito por meio da barra mostrada na Figura 15b.



Figura 15 – (a) Menu de acesso como opções para ouvir, desenhar, e salvar o sinal de áudio. (b) Barra de controle durante a execução de um sinal de áudio.

Na situação em que o subsistema de microfones direcionais não está habilitado, o sinal emitido pelo alto-falante na entrada é o mesmo que chega ao único microfone disponível. Caso contrário, após a simulação tem-se acesso aos sinais que efetivamente chegam a cada um dos microfones (Figura 16).

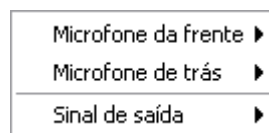


Figura 16 – Opções disponíveis para acessar o sinal recebido por cada microfone.

A partir da janela principal (Figura 13) pode-se acessar cada subsistema (microfones direcionais, redutores de ruído e compressor de amplitude) para a escolha da técnica desejada. Depois de selecionada uma determinada técnica pode-se configurar os respectivos parâmetros.

Na Figura 17 mostra-se as opções de configuração da técnica do nulo adaptativo implementada no subsistema de microfones direcionais. Existe a possibilidade de

configuração dos seguintes elementos: distância entre os microfones; fator de ponderação na estimação da correlação cruzada; se os sinais captados pelos microfones serão simulados (devido a um atraso artificial de recepção) ou se será realizada leitura de dois arquivos com sinais reais; localização das fontes sonoras; e equalização do sinal de saída.

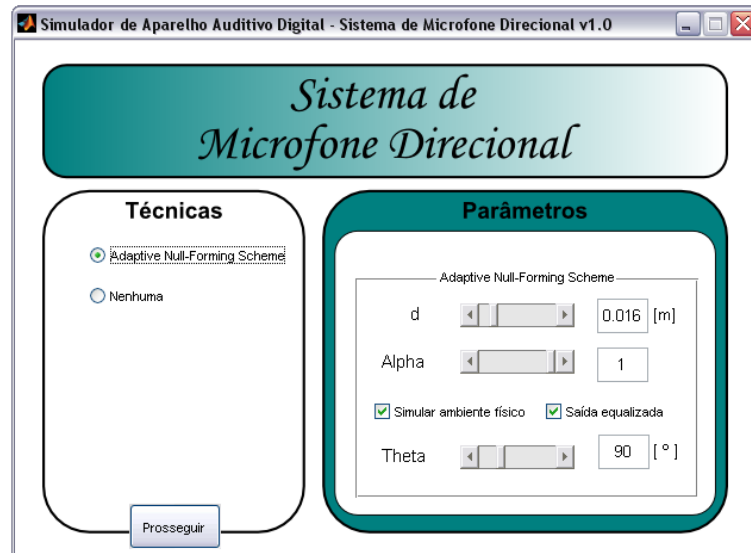


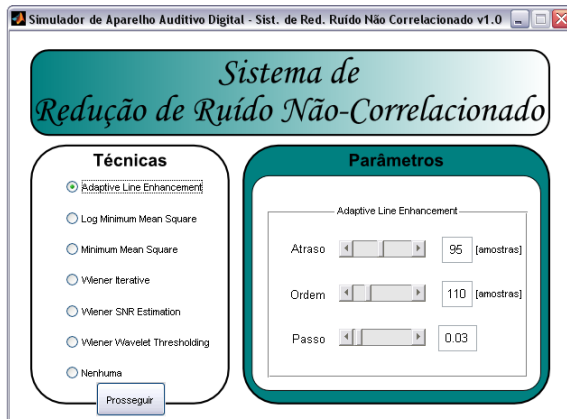
Figura 17 – Subsistema de microfones direcionais.

Para o subsistema de redução de ruído não-correlacionado as técnicas e seus respectivos parâmetros ajustáveis podem ser visualizados na Figura 18. Entre os diversos parâmetros que podem ser modificados, pode-se citar: o tamanho do *frame* de análise; o passo de adaptação; e as constantes de tempo dos filtros passa-baixas.

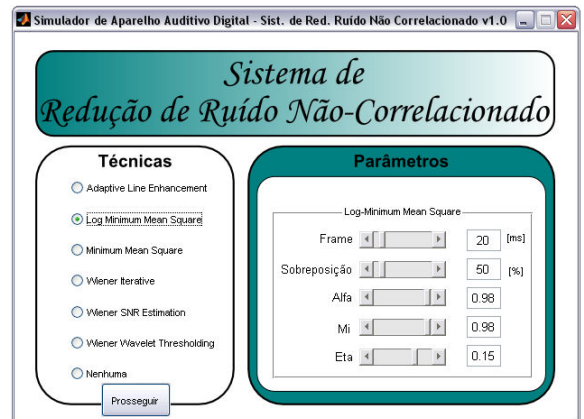
Os parâmetros da técnica utilizada no subsistema de redução de ruído correlacionado estão expostos na Figura 19. Observa-se que é possível ajustar as constantes de *attack* e *release* e também o limiar do detector de atividade vocal.

As opções disponibilizadas para as técnicas contidas no subsistema de compressão de amplitude são apresentadas na Figura 20.

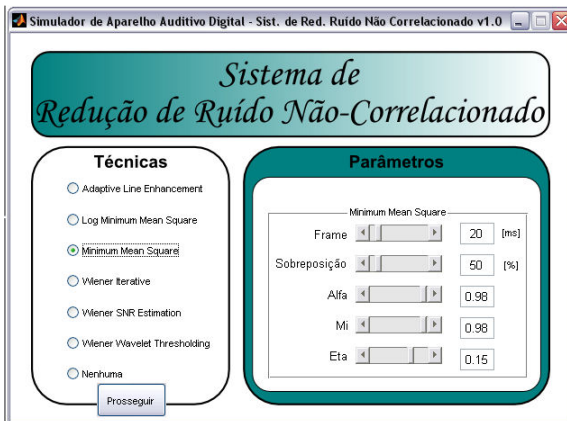
Para um melhor controle por parte do usuário sobre quais subsistemas estão sendo utilizados no processamento, optou-se pela implementação de um indicador (verdadeiro ou falso) em cada botão na janela principal. Na Figura 21a tem-se uma situação em que o usuário optou por habilitar apenas o subsistema de redução de ruído não-correlacionado. Observa-se o indicador na cor verde para o subsistema ativo (redução de ruído não-correlacionado) e na cor vermelha para os demais subsistemas.



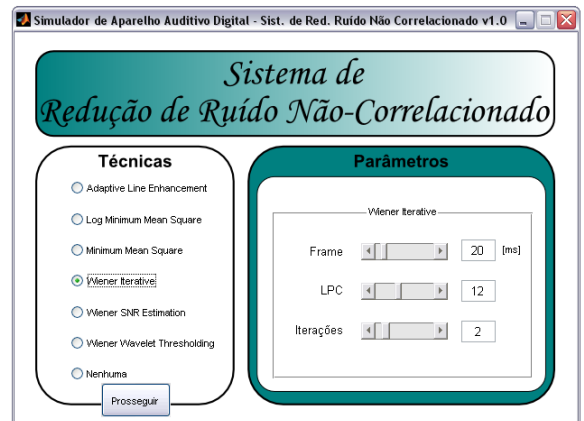
(a)



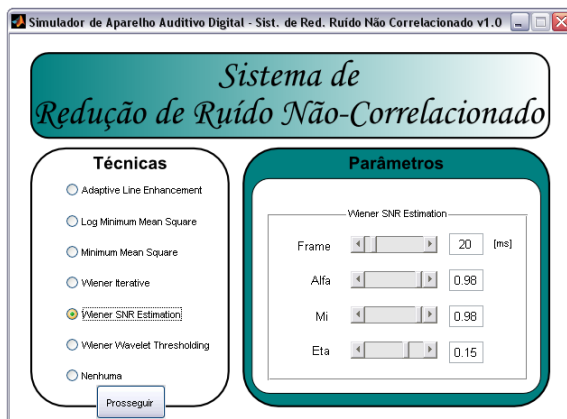
(b)



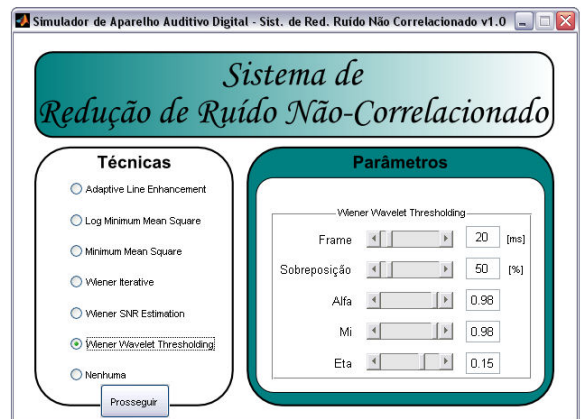
(c)



(d)



(e)



(f)

Figura 18 – Subsistema de redução de ruído não-correlacionado. (a) ALE; (b) LogMMSE; (c) MMSE; (d) Wiener *iterative*; (e) Wiener SNR *estimation*; (f) Wiener *wavelet thresholding*.

Após a seleção dos sinais de entrada e da configuração dos diversos subsistemas pode-se simular o processamento realizado pelo aparelho auditivo. Existem diversas combinações de simulação possíveis, permitindo a análise dos algoritmos em várias situações de trabalho. Após a finalização da simulação, pode-se acessar o sinal de saída

(correspondente ao som emitido pelo alto-falante dentro do ouvido do usuário). Também é possível visualizar os sinais nas saídas intermediárias de cada subsistema de processamento. As opções oferecidas para a avaliação dos sinais de áudio são reprodução, visualização e armazenamento dos sinais de áudio. Na Figura 21b mostra-se um exemplo do estado da janela após a simulação com todos os subsistemas habilitados.

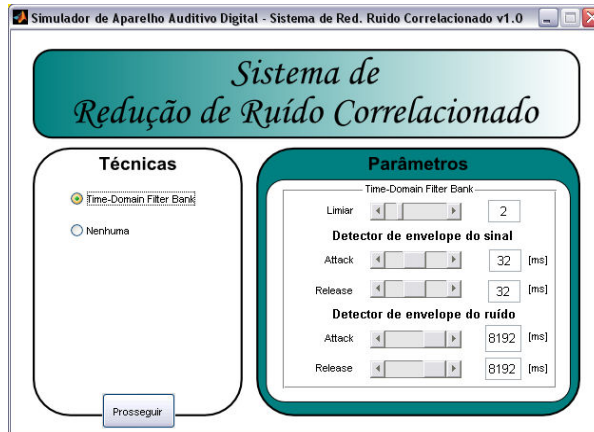
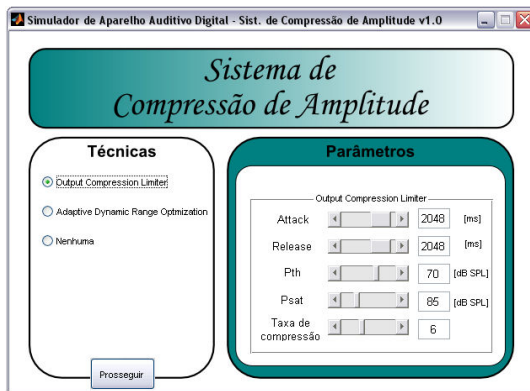
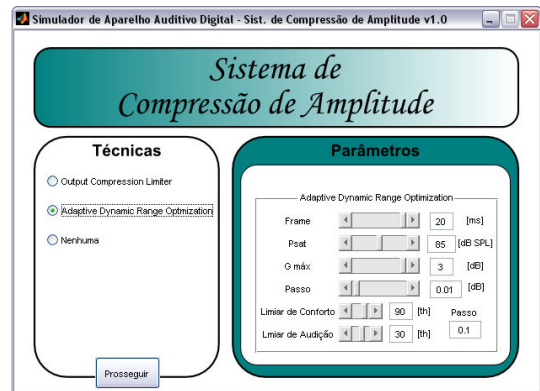


Figura 19 – Subsistema de redução de ruído correlacionado.

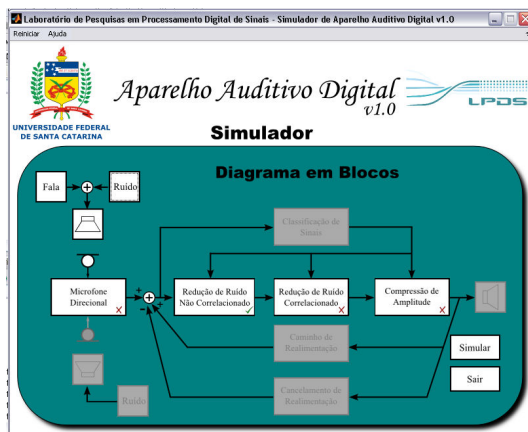


(a)

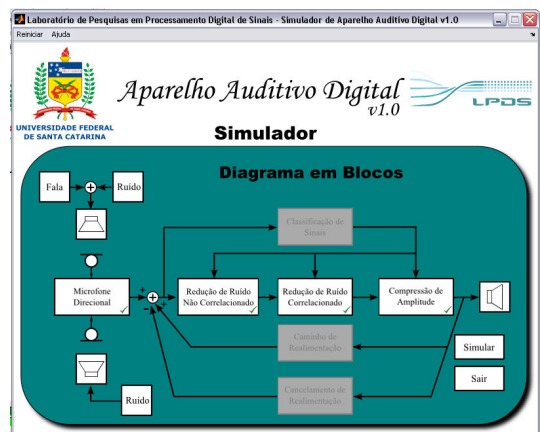


(b)

Figura 20 – Subsistema de compressão de amplitude. (a) Compressor dinâmico limitador. (b) Otimização dinâmica adaptativa.



(a)



(b)

Figura 21 – Exemplo do uso do software. (a) Indicador ativo/inativo. (b) Janela após simulação.

5 DISCUSSÃO

Para atingir os objetivos descritos na Seção 1.1 foi necessária a realização de uma ampla revisão da literatura científica e de patentes comerciais, descritas nas referências bibliográficas deste trabalho, bem como o estudo da arquitetura de um microprocessador dedicado para aplicações em aparelhos auditivos digitais.

Durante o desenvolvimento da interface priorizou-se a praticidade para o usuário. Para tornar isto possível, uma parte do tempo foi gasta no aprendizado de ferramentas para auxiliar no *design* das janelas do programa computacional e facilitar a interatividade com os botões, menus e gráficos.

O estágio de entrada, onde são carregados os sinais, apresenta algumas funcionalidades interessantes. Observa-se a possibilidade de várias combinações com diferentes tipos de voz e ruído. Nota-se ainda a possibilidade de desabilitar qualquer uma das fontes sonoras, por exemplo, para observar o efeito do processamento sem a presença de ruído. Com a função do microfone direcional tem-se a opção de utilizar sinais gravados no modo estéreo, bastando carregar os canais separadamente e informar a distância entre os microfones.

Na parte referente ao processamento, a configuração dos parâmetros utilizados na última simulação permanece salva na memória. Desta forma evita-se a necessidade de reconfiguração de todas as variáveis a cada simulação, caso se deseje alterar apenas um parâmetro. Isto se torna útil quando se está estudando a influência de algum parâmetro em específico. A indicação de ativo/inativo nos subsistemas permite verificar se o subsistema encontra-se com alguma técnica selecionada, sem a necessidade acessar janela referente ao subsistema em questão. O acesso aos sinais processados por cada subsistema permite avaliações objetivas e subjetivas.

Apesar do software desenvolvido não apresentar alguns subsistemas de processamento, conseguiu-se reproduzir algumas características representativas de um aparelho auditivo digital comercial. Além disto, o programa computacional tem flexibilidade para adicionar novas funcionalidades e também pode ser utilizado para motivação de estudantes no aprendizado de técnicas de processamento de fala.

6 PERSPECTIVAS FUTURAS

Como as operações realizadas no *software* são feitas em precisão infinita (precisão máxima do aplicativo MATLAB[®]), seria interessante incluir a possibilidade de realizar operações com aritmética de ponto fixo para simular o efeito da quantização da mesma forma que ocorre em um dispositivo real. Outra sugestão seria disponibilizar ao usuário a possibilidade da alteração do fator de atenuação das diversas bandas para a compensação da deficiência de audição.

Para agregar mais funcionalidades ao sistema desenvolvido, posteriormente devem ser implementados os subsistemas de cancelamento de realimentação, classificação de sinais e a realimentação de ruído. Como o processamento do sistema é feito em blocos, deve-se alterar a forma de processamento do sinal para processamento por amostra.

REFERÊNCIAS

- [1] VICEN-BUENO, R.; GIL-PITA, R.; UTRILLA-MANSO, M.; ÁLVAREZ-PÉREZ, L. A hearing aid simulator to test adaptive signal processing algorithms. In: **IEEE International Symposium on Intelligent Signal Processing**. Alcalá de Henares: 3-5 out 2007. p. 1-6.
- [2] FERREIRA, R.F. *et. al.* Resultados auditivos com implante coclear multicanal em pacientes submetidos a cirurgia no hospital das clinicas da faculdade de medicina da universidade de são Paulo. **Revista brasileira de otorrinolaringologia**, v. 70, n. 5, p. 632-637, 2004.
- [3] HAMACHER, V., *et. al.* Signal processing in high-end hearing aids: state of the art, challenges, and future trends. **EURASIP Journal on Applied Signal Processing**, v. 2005, n. 18, p. 2915-2929, 2005.
- [4] LUO, F.-L.; NEHORAI, A. Recent development in signal processing for digital hearing aids. **IEEE Signal Processing Magazine**, v. 26, p. 103-106, setembro 2006.
- [5] SMITH, S.T. **MATLAB: advanced GUI development**. Indianápolis: Dog ear publishing. 2006.
- [6] LAAKSO, T.I.; VÄLIMÄKI, V.; KARJALAINEN, M; LAINE, U.K. Splitting the Unit Delay – Tools for fractional delay filter design. **IEEE signal processing magazine**, 1053-5888/96. 1996.
- [7] LUO, F.L.; YANG, J.; PAVLOVIC, C.; NEHORAI, A. Adaptive null-forming scheme in digital hearing aids. **IEEE Transactions on Signal Processing**, 50(7):1583-1590. 2002.
- [8] EPHRAIM, Y.; MALAH, D. Speech enhancement using a minimum mean-square error log-spectral amplitude estimator. **IEEE Trans. Acoust., Speech, Signal Process.**, ASSP-23(2), 443-445(1985).
- [9] EPHRAIM, Y.; MALAH, D. Speech enhancement using a minimum mean- square error short-time spectral amplitude estimator”. **IEEE Trans. Acoust., Speech, Signal Process.**, ASSP-32, 1109-1121.(1984).
- [10] FARHANG-BOROUJENY, B. **Adaptive filters: theory and applications**. New York: John Willey & Sons, Inc.1998.
- [11] LOIZOU, P.C. **Speech enhancement: theory and practice**. Boca Raton: CRC Press. 2007.

- [12] Voyageur TD™ Getting Started, GENNUM, 2006.
- [13] FANG, X.; NILSSON, M.. **Noise Reduction Apparatus and Method**. United States Patent, 2004.
- [14] STETZLER, T., *et. al.* Low-power real-time programmable DSP development platform for digital hearing aids. **Texas instruments: Application report**, abril 2000.
- [15] BLAMEY, P.J. Adaptive dynamic range optimization (ADRO): a digital amplification strategy for hearing aids and cochlear implants. **Trends in Amplification**, v. 9, n. 2, p. 77-98, 2005.
- [16] ARMSTRONG, S. The dynamics of compression: some key elements explored. **The Hearing Journal**, vol. 46, no 11 , pp. 43-47, Novembro 1993.