REYNALDO PAMPA CONDORI

**CANCELAMENTO DE RETORNO LOCAL EM
APARELHOS TELEFÔNICOS PARA DEFICIENTES
AUDITIVOS**

Dissertação apresentada à Escola Politécnica
da Universidade de São Paulo para obtenção
do título de Mestre em Engenharia Elétrica

Área de Concentração:
Sistemas Eletrônicos

Orientador :
Prof. Dr. Phillip Mark Seymour Burt

São Paulo

2012

Agradecimentos

Aos meus pais pelos seus esforços na procura de melhores oportunidades para mim e meus irmãos.

A minha irmã Elsa, pelo seu apoio nos momentos difíceis.

Ao meu orientador Prof. Dr. Phillip Mark Seymour Burt, pela sua paciência e apoio ao longo do desenvolvimento deste trabalho.

A Deus pelas suas bênçãos.

Sumário

Lista de Figuras	5
Lista de Tabelas	6
Lista de Siglas	7
Resumo	8
Abstract	10
1 Introdução	12
1.1 Objetivo do Trabalho	12
1.2 Organização do Trabalho	13
1.3 Publicação associada ao trabalho	14
2 Conceitos básicos	15
2.1 Breve descrição do funcionamento de um aparelho telefônico para deficientes auditivos	16
2.2 Tipos de realimentação nos aparelhos telefônicos para deficientes auditivos	18
2.2.1 Retorno local	18
2.2.2 Eco acústico	19
2.3 Análise da estabilidade em sistemas discretos realimentados	20
2.4 Sistema auditivo humano	21
2.5 Área de audição	22
2.6 Definição do problema	24
3 Aspectos Teóricos	26
3.1 Filtragem adaptativa	26
3.1.1 Filtro de Wiener	27
3.1.2 O algoritmo do gradiente descendente	28

3.1.3	O algoritmo Least Mean Square (LMS)	28
3.1.4	O algoritmo LMS normalizado	29
3.1.5	O algoritmo Recursive Least Square (RLS)	30
3.1.6	Complexidade computacional	32
3.2	Filtros <i>notch</i>	32
4	Cancelamento de retorno local em aparelhos telefônicos para deficientes auditivos	34
4.1	Amplificação para compensação de deficiência	35
4.2	Objetivo de cancelamento de retorno local	36
4.3	Desempenho com sinal de fala	38
4.4	Adaptação anterior à conversação	41
5	Conclusões e perspectivas	44
	Bibliografia	45

Lista de Figuras

2.1	Esquema simplificado das características do funcionamento do aparelho telefônico para deficientes auditivos.	16
2.2	Princípio de funcionamento da híbrida indutiva	19
2.3	Sistema realimentado.	20
2.4	Sistema auditivo humano. Figura reproduzida de [9]	22
2.5	Área auditiva típica de um ser humano. Figura adaptada de [11].	23
2.6	Esquema básico de um aparelho telefônico para deficientes auditivos	24
3.1	Representação em diagramas de blocos do problema da filtragem adaptativa. Figura adaptada de [13]	28
4.1	Efeito da perda auditiva sobre a area auditiva.	35
4.2	Esquema básico de um aparelho telefônico para deficientes auditivos com cancelador de retorno local	38
4.3	Adaptação com sinais de fala, sem ruído de medida. Traços a e b : sinais $x(n)$ A e B, respectivamente (fora de escala); c e d : $K(n)$ para LMS com $\mu = 0,007$ (próximo ao máximo μ), para sinais A e B, respectivamente; e e f : $K(n)$ para ϵ -NLMS com $\mu = 0,1$ e $\epsilon = 0$, para sinais A e B, respectivamente.	39
4.4	Adaptação com sinais de fala, sem ruído de medida. Traços a e b : sinais $x(n)$ A e B, respectivamente (fora de escala); c e d : $K(n)$ para RLS com $\lambda = 0,999$ e $\mathbf{P}(0) = 10^6 \mathbf{I}$, para sinais A e B, respectivamente.	40
4.5	Adaptação com sinais de fala, ruído de medida com SNR = 30 dB. Traço a : sinal $x(n)$ B (fora de escala); b,c e d : $K(n)$ para, respectivamente, LMS com $\mu = 0,005$, ϵ -NLMS com $\mu = 0,005$ e $\epsilon = 0,09$, RLS com $\lambda = 1$ e $\mathbf{P}(0) = 10^3 \mathbf{I}$	41

4.6	Adição de filtro <i>notch</i> para eliminar tom de discar.	42
4.7	Adaptação com sinal de entrada branco e gaussiano, ruído de medida com SNR = 30 dB e inteferência senoidal com SNR= 0 dB. Traços a e b : LMS, $\mu = 0,0002$, sem e com o filtro <i>notch</i> , respectivamente.	43

Lista de Tabelas

2.1	Descrição de alguns sinais na sinalização acústica. Informação extraída de [3].	18
2.2	Grau de perda auditiva. Informação extraída de [12].	24
3.1	Complexidade computacional por iteração dos algoritmos LMS, ϵ -NLMS e RLS. Informação extraída de [14].	32

Lista de Siglas

DR1	<i>dynamic range of hearing</i> (faixa dinâmica de audição)
DR2	<i>residual dynamic range</i> (faixa dinâmica residual)
DRP	<i>dynamic range</i> (faixa dinâmica)
DSP	<i>digital signal processor</i> (processador digital de sinal)
FIR	<i>finite impulse response</i> (resposta ao pulso unitário finita)
HL	<i>hearing level</i> (nível de audição)
HTL	<i>hearing threshold level</i> (nível de limiar de audição)
IIR	<i>infinite impulse response</i> (resposta ao pulso unitário infinita)
LMS	<i>least mean square</i> (algoritmo dos mínimos quadrados)
NLMS	<i>normalized least mean square</i> (algoritmo dos mínimos quadrados normalizado)
RLS	<i>recursive least squares</i> (algoritmo dos mínimos quadrados recursivos)
SIL	<i>sound intensity level</i> (nível de intensidade sonora)
SIR	<i>signal-to-interference ratio</i> (relação sinal/interferência)
SNR	<i>signal-to-noise ratio</i> (relação sinal/ruído)
SPL	<i>sound pressure level</i> (nível de pressão sonora)

Resumo

Em algumas ocasiões o uso de aparelhos é necessário para melhorar a audição de deficientes auditivos. Para melhorar a clareza das ligações telefônicas dessas pessoas, há interesse em aparelhos telefônicos que desempenham função semelhante aos aparelhos de audição (*hearing aids*), amplificando adequadamente o sinal recebido do outro lado da conexão. Um dos maiores problemas com tais aparelhos é que a malha fechada que inclui o retorno local na híbrida telefônica e o acoplamento acústico entre o alto-falante e o microfone do aparelho pode se tornar instável devido à introdução da amplificação, dando origem a microfonia. Uma abordagem típica deste problema nos aparelhos de audição é feita mediante o cancelamento adaptativo do eco acústico. No caso sendo considerado aqui, porém, também é possível eliminar a microfonia fazendo o cancelamento adaptativo do retorno local na híbrida, abordagem que é mais simples, devido à resposta impulsiva mais curta a ser compensada e sua menor variação temporal.

Portanto, esta dissertação tem como objetivo o desenvolvimento de um cancelador de retorno local usando filtragem adaptativa para fins de impedir a microfonia num aparelho telefônico para deficientes auditivos.

Para uma efetiva compensação de um determinado grau de deficiência auditiva, será visto que pode ser desejável em algum momento introduzir até uma certa quantidade de ganho em alguma frequência, sendo que a viabilidade disso depende da efetividade do cancelamento do retorno local, é definida uma medida de desempenho dos filtros adaptativos usados em termos da amplificação máxima que pode ser introduzida graças ao cancelamento do retorno local sem provocar microfonia. Considerando esta medida de desempenho, foi definido um objetivo de 55 dB de amplificação máxima a ser alcançado pelo cancelamento do retorno local. Verificamos então que o uso dos algoritmos LMS (*least mean square*), ϵ -NLMS (*normalized least mean square*) e RLS (*recursive least squares*) não alcança este objetivo

com sinal de fala como entrada. Visto isto, a adaptação antes da conversação é avaliada com sinal branco gaussiano como entrada, acrescentando-se um filtro *notch* para eliminar o tom de discar. Os resultados mostram que o algoritmo LMS é suficiente para alcançar o objetivo mencionado.

Abstract

On some occasions the use of apparatus are needed to improve the deafness of hearing impaired. To improve the clarity of the telephonic connections of these people, there is interest in telephone sets that exert similar function to the hearing aids, properly amplifying the received signal of the another side of the connection. One of the main problems with such apparatus is that the closed loop that includes local return in the telephone hybrid and acoustic coupling between the capsule and the microphone of the telephone set would become unstable due to the introduction of amplification, producing howling. A typical approach to this problem in hearing aids is made by the adaptive acoustic echo cancellation. In this case, however, it is also possible to eliminate howling by the adaptive local return cancellation in the hybrid, which is a simpler approach, due to shorter impulse response to be compensated and their lower temporal variation.

Therefore, the goal of this dissertation is to develop a local return canceller using adaptive filtering for the purpose of preventing howling in a telephone set for hearing impaired.

For an effective compensation of a certain degree of hearing impairment, it will be seen that it may be desirable at some point to introduce up to a certain amount of gain at some frequency. Whereas the viability of this depends on the effectiveness of the cancellation of the local return, a performance measure of used adaptive filters in terms of the maximum amplification which may be released due to the cancellation of the local return is defined without causing howling. Considering this performance measure was set a goal of 55 dB of maximum gain to be achieved by the local return cancelling. We verify then that the use of the LMS (least mean square), ϵ -NLMS (normalized least mean square) and RLS (recursive least squares) algorithms doesn't attain this goal with speech signal as input. Therefore, adaptation before the conversation is evaluated, with white gaussian signal as input, adding a notch filter to eliminate the dial tone. The results show that the LMS

algorithm is sufficient to achieve the mentioned goal.

Capítulo 1

Introdução

Em algumas ocasiões o uso de aparelhos é necessário para melhorar a audição de deficientes auditivos. Para melhorar a clareza das ligações telefônicas destas pessoas, há interesse em aparelhos telefônicos que desempenham função semelhante aos aparelhos de audição (*hearing aids*), amplificando adequadamente o sinal recebido do outro lado da conexão. Um dos maiores problemas com tais aparelhos é que a malha fechada que inclui o retorno local na híbrida telefônica (eco elétrico) e o acoplamento acústico entre o alto-falante e o microfone do aparelho (eco acústico), pode se tornar instável devido à introdução da amplificação, dando origem a microfonia (*howling*). Esse fenômeno pode ser percebido tipicamente como um apito que acontece quando se deixa um microfone perto demais de um alto-falante num sistema de amplificação. Um problema semelhante ocorre nos aparelhos de audição e uma das abordagens típicas para sua solução é o cancelamento de eco acústico por meio de filtragem adaptativa, abordagem que poderia ser seguida nos aparelhos telefônicos para deficientes auditivos. Neste caso, porém, também é possível eliminar a microfonia fazendo o cancelamento adaptativo do retorno local na híbrida, abordagem que é mais simples, devido à resposta impulsiva mais curta a ser compensada e sua menor variação temporal. Neste trabalho, portanto, eliminamos a microfonia mediante o cancelamento adaptativo do retorno local.

1.1 Objetivo do Trabalho

Esta dissertação tem como objetivo o desenvolvimento de um cancelador de retorno local usando filtragem adaptativa, para fins de impedir a microfonia num aparelho telefônico para

deficientes auditivos.

1.2 Organização do Trabalho

No Capítulo 2, inicialmente, se faz uma breve descrição do funcionamento de um aparelho telefônico para deficientes auditivos. Depois discutem-se as origens do retorno local e do eco acústico como fenômenos decorrentes do funcionamento do aparelho em questão. Devido à relação causa-consequência entre a instabilidade do aparelho e a microfonia, em seguida se define uma condição de estabilidade para um sistema com realimentação positiva (o caso do aparelho em questão). Então o sistema auditivo humano e as causas das deficiências auditivas são estudados. Na continuação, a área da audição humana é apresentada. Por fim, baseada na condição de estabilidade definida anteriormente, uma condição suficiente de estabilidade é determinada para o aparelho telefônico para deficientes auditivos.

No Capítulo 3 são estudados o filtro de Wiener e o algoritmo do gradiente descendente como preliminares para o estudo dos filtros adaptativos LMS, NLMS e o RLS. Em seguida, para fins de comparação as complexidades computacionais destes filtros são mostradas. O capítulo inclui o estudo do filtro digital *notch*, usado no Capítulo 4 para eliminar a interferência do tom de discar.

No Capítulo 4, inicialmente, a necessidade de expressar uma medida do desempenho dos filtros adaptativos em função do ganho máximo introduzido sem provocar microfonia é vista. Isto é feito na seção seguinte. Além disso, é definido um objetivo de 55 dB de amplificação máxima a ser alcançado pelo cancelamento do retorno local. Finalmente, o desempenho dos filtros adaptativos mencionados é avaliado. Primeiro, os algoritmos são avaliados com sinal de fala como entrada. Os resultados mostram que este método não alcança o objetivo determinado. Visto isto, a adaptação antes da conversação é avaliada com sinal branco gaussiano como entrada, acrescentando-se um filtro notch para eliminar o tom de discar. Os resultados mostram que o algoritmo LMS é suficiente para alcançar o objetivo mencionado. Portanto o análise do desempenho dos algoritmos restantes é desnecessário.

Finalmente, no Capítulo 5, são apresentadas as conclusões e perspectivas deste trabalho.

1.3 Publicação associada ao trabalho

Os resultados obtidos neste trabalho foram publicados nos anais do XXX Simpósio Brasileiro de Telecomunicações 2012, no artigo intitulado “Cancelamento de retorno local em aparelhos telefônicos para deficientes auditivos” em co-autoria com o Prof. Dr. Phillip Mark Seymour Burt.

Capítulo 2

Conceitos básicos

Microfonia é um problema comum em aparelhos de audição [1] e em sistemas de amplificação [2]. Neles a abordagem típica é feita sobre a realimentação acústica. Este tipo de problema também ocorre nos aparelhos telefônicos para deficientes auditivos. Mas cabe indicar que esses aparelhos apresentam, além do eco acústico, outro tipo de realimentação denominado retorno local, decorrente das características do funcionamento do aparelho telefônico. Neste trabalho é de importância entender estes dois tipos de realimentação (Seção 2.2), mas para isso, antes, se faz necessário conhecer o funcionamento dos aparelhos em questão (Seção 2.1). Na Seção 2.3 se faz um estudo das condições que um sistema de tempo discreto com realimentação positiva deve obedecer para ser estável. Isso para definir posteriormente na Seção 2.6 as condições que um aparelho telefônico para deficientes auditivos deve obedecer para o mesmo estado. Pelo fato de que a medida dos resultados desta dissertação está condicionada com a qualidade de vida do deficiente auditivo, se faz necessário entender o problema da perda de audição. Para isso, na Seção 2.4 se apresenta o sistema auditivo humano e as possíveis causas de deficiência auditiva. Na Seção 2.5 é definida a área de audição de uma pessoa normal e uma classificação típica de perda auditiva é apresentada. Por fim, define-se o problema que implica impedir a ocorrência da microfonia.

2.1 Breve descrição do funcionamento de um aparelho telefônico para deficientes auditivos

Tanto um aparelho telefônico convencional quanto um para deficientes auditivos apresentam visivelmente duas partes: o *handset* e o corpo do telefone como esquematizado na Figura 2.1. No *handset* podemos encontrar o alto-falante, do qual emergem os sons que são escutados pelo usuário, e o microfone, que converte a voz do usuário em sinais elétricos. O sinal elétrico do microfone é transportado por meio de dois fios até o corpo do telefone e da mesma forma sinais provenientes do corpo do telefone são transportados até o alto-falante. Também podemos notar que o corpo do telefone é conectado à central telefônica somente por meio de dois fios. Os sinais elétricos tanto do usuário distante quanto do usuário próximo são transportados nele. Isto é feito por razões econômicas. Portanto, algum dispositivo se faz necessário para fazer a conversão 4 fios / 2 fios, para permitir a comunicação *full duplex*. Este dispositivo é denominado híbrida [3]. Em outras palavras, a híbrida vai permitir que em só dois fios possam-se transportar tanto o sinal do usuário próximo como do usuário distante.

Figura 2.1: Esquema simplificado das características do funcionamento do aparelho telefônico para deficientes auditivos.

Na atualidade, a maioria dos aparelhos telefônicos para deficientes auditivos, para com-

pensar as perdas auditivas dos deficientes auditivos, fazem uso de técnicas de processamento digital de sinais. Então, para que os sinais elétricos analógicos provenientes do usuário distante (após passar pela híbrida), possam ser compensados à conveniência do deficiente auditivo antes devem ser convertidos em sinais de tempo discreto. E após disso, o sinal obtido do processamento deve ser convertido em sinal analógico para que possa ser escutado pelo usuário através do alto-falante. Estes conversores são denominados de conversores analógicos-digitais (A/D) e digitais-analógicos (D/A) como mostrado na Figura 2.1. Podemos notar também, na figura mencionada, a conversão analógica-digital dos sinais provenientes do microfone e a conversão digital-analógica com direção ao usuário distante. Como será visto, a conversão nesta direção, se faz necessária para transformar o sinal analógico decorrente do microfone em um sinal em tempo discreto, para seu posterior processamento, segundo a aplicação em questão. O processamento dos sinais é levado a cabo pelos conhecidos DSP's (*digital signal processors*).

Quando se estabelece uma ligação telefônica, se faz necessária uma comunicação entre o telefone do usuário e a central e também é possível que seja necessária uma comunicação entre centrais. O primeiro tipo de comunicação é denominada de sinalização do assinante e a segunda de sinalização intercentrais. Na Seção 4.4 será de interesse conhecer o que acontece após tirar o *handset* do corpo do telefone. Este período de tempo corresponde à sinalização do assinante. Na continuação, uma breve descrição do estabelecimento de uma ligação telefônica entre um usuário A que deseja-se comunicar com um usuário B será feita.

Uma vez que o usuário A tira o *handset* do gancho do corpo do telefone, o circuito deste conecta-se à central telefônica fluindo através da linha corrente direta, que é para a central um sinal de tentativa de ligação telefônica. A central após de algumas verificações envia o tom de discar que é um aviso ao usuário para discar o número do usuário B. Uma vez que a central do usuário B tenha sido conectada, ela envia um tom de toque (vide Tabela 2.1) para os usuários A e B. Se o usuário B tirar o *handset*, a conversação começa. Além do tom de discar e tom de toque podemos encontrar outros sinais, com diferentes funções, que são mostrados na Tabela 2.1.

Sinal	Frequência (Hz)	Descrição
Corrente de toque	$(25 \pm 2, 5)$	Sinal enviado ao usuário chamado, indicando que há uma ligação dirigida a ele.
Tom de discar	$(425 \pm 2, 5)$	Sinal enviado ao usuário chamador após a ocupação do circuito de linha associado, para que se inicie a marcação dos algarismos do telefone do usuário chamado.
Tom de ocupado	$(425 \pm 2, 5)$	Sinal enviado ao terminal chamador indicando que o terminal chamado está na condição de ocupado.
Tom de controle de chamada	$(425 \pm 2, 5)$	Sinal enviado pela central de destino ao terminal chamador para indicar que a cadeia de comutação foi completada e o terminal chamado está livre.

Tabela 2.1: Descrição de alguns sinais na sinalização acústica. Informação extraída de [3].

2.2 Tipos de realimentação nos aparelhos telefônicos para deficientes auditivos

Baseado no discutido na seção anterior, explicamos como ocorrem o retorno local e o eco acústico, tendo em consideração as características do telefone.

2.2.1 Retorno local

Para entender este tipo de realimentação se faz necessário estudar antes o princípio de funcionamento das híbridas. Um tipo de híbrida conhecido por híbrida indutiva será estudado. Na Figura 2.2 podem-se observar os elementos que a conformam. Um transformador balanceado tem o enrolamento secundário conectado ao alto-falante. O microfone é conectado ao centro do enrolamento primário. A impedância de balanceamento Z_b tem função de duplicar a impedância de linha Z_L . Portanto, quando Z_b iguala Z_L exatamente, a corrente elétrica gerada pelo microfone devido aos sons recebidos é dividida igualmente. Então correntes opostas fluirão pelo enrolamento primário, gerando voltagens de polaridades opostas no enrolamento secundário [4]. Portanto, nenhum sinal originado no microfone chegará até o

alto-falante, mas sim será transmitido através dos dois pares de fios até a central telefônica. Sinais chegados da central não serão cancelados e chegarão ao alto-falante. Mas cabe indicar que o balanceamento entre a impedância de linha e a impedância de balanceamento não é perfeita, permitindo que parte da energia proveniente do microfone chegue ao alto-falante, originando-se o denominado retorno local.

Figura 2.2: Princípio de funcionamento da híbrida indutiva

2.2.2 Eco acústico

Este tipo de realimentação ocorre quando há acoplamento acústico entre o alto-falante e o microfone do aparelho telefônico para deficientes auditivos. Este problema é comum em diversas aplicações, como em terminais viva-voz [5], sistemas de amplificação, sistemas de teleconferência [6] e aparelhos de audição, podendo ocasionar microfonia ou o desconforto de se ouvir sua própria fala atrasada de alguns instantes. O som produzido pelo alto-falante percorre muitos caminhos chegando até o usuário próximo, mas também uma certa quantidade dele alcança o microfone (vide Figura 2.1). O sinal que alcança o microfone é agora um sinal diferente do sinal que saiu do alto-falante. Ele foi modificado pelo caminho que existe entre o alto-falante e o microfone. A abordagem comum feita para solucionar este problema é o cancelamento do eco acústico. Este procedimento torna-se complicado pelo comprimento e variabilidade da resposta ao impulso do caminho do eco decorrentes das características do acoplamento acústico entre o microfone e o alto-falante.

2.3 Análise da estabilidade em sistemas discretos realimentados

Como se mencionou anteriormente, as origens da microfonia têm relação com a instabilidade de um sistema realimentado. Portanto, considerando sinais em tempo discreto, para evitar a microfonia é de importância definir quando um sistema discreto realimentado é estável.

A condição suficiente e necessária para a estabilidade de um sistema em tempo discreto, assumindo z o atraso unitário, é que todos os pólos da sua função de transferência no domínio z estejam fora da circunferência unitária [7]. Como será visto na Seção 2.6, nesta dissertação, nosso sistema de interesse será modelado como um sistema discreto com realimentação positiva. Podemos observar na Figura 2.3 um sistema simplificado com tais características. Calculando a sua função de transferência de malha fechada temos

$$\frac{A(z)}{1 - A(z)B(z)}. \quad (2.1)$$

Como mencionado, para que este sistema seja estável as raízes da sua equação característica definida como $F(z) = 1 - A(z)B(z)$ devem estar localizadas fora da circunferência unitária.

Figura 2.3: Sistema realimentado.

Segundo o teorema do módulo máximo [8], se a função de transferência em malha aberta de (2.1), $A(z)B(z)$, é analítica na região limitada pela circunferência unitária C e sobre ela, então o valor máximo de $|A(z)B(z)|$ ocorre sobre C , a menos que $A(z)B(z)$ seja uma constante. Então, assumindo que $A(z)B(z)$ é analítica dentro da circunferência unitária e na mesma circunferência, a expressão seguinte é verdadeira

$$|A(z)B(z)| \leq \|A(e^{j\omega})B(e^{j\omega})\|_{\infty}, \text{ para todo } |z| \leq 1. \quad (2.2)$$

onde $\|A(e^{j\omega})B(e^{j\omega})\|_{\infty} \triangleq \max_{\omega} |A(e^{j\omega})B(e^{j\omega})|$. Cabe notar, que

$$|A(e^{j\omega})B(e^{j\omega})| \leq \|A(e^{j\omega})B(e^{j\omega})\|_{\infty}, \text{ para todo } \omega. \quad (2.3)$$

Portanto, se $|A(e^{j\omega})B(e^{j\omega})| < 1$ é verdadeiro para todo ω , também

$$|A(z)B(z)| < 1, \text{ para } |z| \leq 1 \quad (2.4)$$

é verdadeiro. Da expressão (2.4) se deduz que $1 - A(z)B(z)$ não tem raízes dentro do círculo unitário, portanto (2.1) é estável, se

$$|A(e^{j\omega})B(e^{j\omega})| < 1, \text{ para todo } \omega. \quad (2.5)$$

2.4 Sistema auditivo humano

O sistema auditivo humano nos permite perceber os sons. Na continuação se faz uma breve descrição desse sistema. Na Figura 2.4 são mostrados as partes que o compõem. O ouvido é composto de três partes: o ouvido externo, o ouvido médio e o ouvido interno [9]. No ouvido externo o som é captado pela orelha e passa pelo canal auditivo até chegar ao tímpano. Então o som transmitido faz vibrar o tímpano [10]. A função do ouvido médio é transferir eficientemente a energia sonora do ar para o líquido contido dentro da cóclea. As vibrações são transmitidas até a cóclea no ouvido interno através de três pequenos ossos do ouvido médio conhecidos como martelo, estribo e bigorna. As vibrações transmitidas à cóclea ativam pequenas células ciliadas. Estas células transformam as vibrações em impulsos nervosos, os quais são captados pelo nervo acústico e enviados ao cérebro.

Quando o som não pode ser transmitido normalmente através do canal auditivo e/ou do ouvido médio para a cóclea, temos um caso de deficiência auditiva condutiva. As causas mais freqüentes neste caso são o acúmulo de cerume, a perfuração do tímpano, as infecções na orelha média (otites) e a fixação dos ossículos da audição (otoescleroses). Normalmente estes tipos de problemas são tratados por meio de tratamento médico com medicamentos ou por uma cirurgia.

Tipicamente, o dano nas células ciliadas, denominado deficiência auditiva sensorio-neural, é a maior causa de perda de audição. As duas causas mais comuns são a exposição a ruídos de alta intensidade e a surdez pela idade. Em muitos casos não existe cura para

Figura 2.4: Sistema auditivo humano. Figura reproduzida de [9]

esta deficiência e o uso de aparelhos de audição é recomendado.

2.5 Área de audição

Antes de abordar o tema desta seção, cabe indicar que a medida usual da amplitude do som é dada em termos de intensidade, ou seja, a energia do som transmitido por segundo através de uma unidade de área num meio do som [10]. Esta intensidade de som pode ser expressa em decibéis e em relação a uma referência como

$$10 \log_{10} \left(\frac{I_1}{I_0} \right) \quad (2.6)$$

onde I_1 é a intensidade do som e I_0 é a referência. Quando a intensidade sonora de referência usada é 10^{-12} W/m^2 e o meio do som é o ar, (2.6) é referido como níveis de intensidade sonora (SIL - *sound intensity levels*). De forma análoga os níveis de pressão sonora (SPL - *sound pressure level*) são definidos por

$$20 \log_{10} \left(\frac{p_1}{p_0} \right) \quad (2.7)$$

onde p_1 é a pressão sonora e p_0 é a referência igual a $20 \mu \text{ Pa}$.

Retomando o tema desta seção, a área de audição é um plano no qual sons audíveis podem ser ilustrados [11]. Na Figura 2.5 podemos apreciar a área auditiva típica de um ser humano. Podemos apreciar a faixa das frequências no eixo das abscissas (*frequency*)

numa escala logarítmica, que varia desde 20 Hz até 20 kHz aproximadamente. No eixo das ordenadas os níveis de sons são representados, no lado esquerdo, como níveis de pressão sonora e no lado direito como níveis de intensidade sonora.

A área de audição é limitada abaixo pela curva do limiar de audição (*threshold in quiet*) e acima pela curva do limiar da dor (*threshold of pain*). Dentro da área auditiva duas áreas são notórias, a área da música (*music*) e dentro dela a área da fala (*speech*). Também podemos apreciar dentro da área de audição o limite de risco de dano (*limit of damage risk*). A curva pontilhada, perto do limiar de audição, indica uma perda de audição.

Figura 2.5: Área auditiva típica de um ser humano. Figura adaptada de [11].

É de interesse saber que quando uma deficiência auditiva é suspeita uma forma de especificar os limiares de audição é especificá-los relativos a um limiar médio obtido de ouvintes jovens de boa saúde com audição “normal” para cada frequência. Neste caso as unidades são especificadas como dB HL (*hearing level*) em Europa e dB HTL (*hearing threshold level*) nos Estados Unidos. Por exemplo, para um limiar de audição de 40 dB HL a uma frequência de 4 kHz, entende-se que o limiar estimado desta pessoa fica 40 dB acima do limiar normal na frequência mencionada [10]. Na Tabela 2.2 é mostrada uma classificação do grau de perda auditiva, baseado na média dos limiares de audição para as frequências 500, 1000 e 2000 Hz mediante testes de *air conduction* [12].

Grau de perda auditiva	Faixa dB HL
Audição normal	< 15
Perda de audição leve	$16 - 25$
Perda de audição suave	$26 - 40$
Perda de audição moderada	$41 - 55$
Perda de audição moderada a severa	$56 - 70$
Perda de audição severa	$71 - 90$
Perda de audição profunda	≥ 91

Tabela 2.2: Grau de perda auditiva. Informação extraída de [12].

2.6 Definição do problema

A Figura 2.6 mostra o esquema básico de um aparelho telefônico para deficientes auditivos. O bloco $G(z)$ tem o objetivo de amplificar o sinal do usuário distante, de modo a compensar a deficiência auditiva do usuário local. Na Figura, os seguintes sinais são definidos: $t(n)$ é o eco acústico devido ao acoplamento $W(z)$ entre o alto-falante e o microfone do aparelho; $u(n)$ é o sinal local captado pelo microfone e $x(n) = t(n) + u(n)$ é o sinal a ser enviado ao usuário distante; $y(n)$ é o retorno local na híbrida, dado por $y(n) = \mathbf{h}^T \mathbf{x}(n)$, onde $\mathbf{x}(n) = [x(n) x(n-1) \cdots x(n-N+1)]$ e $\mathbf{h} = [h_0 h_1 \cdots h_{N-1}]$ contém as N amostras da resposta ao impulso do caminho de retorno local, considerado aqui como um sistema FIR; $d(n) = v(n) + y(n)$ é a entrada do bloco $G(z)$, onde $v(n)$ é o sinal recebido do usuário distante.

Figure 2.6: Esquema básico de um aparelho telefônico para deficientes auditivos

Definido a função de transferência de malha aberta da Figura 2.6 como $G(z)W(z)H(z)$, segundo o discutido na Seção 2.3 sobre a estabilidade de sistemas discretos com realimentação positiva, uma condição suficiente para que o sistema seja estável e que não ocorra microfonia para uma dada frequência ω é que seja verificado

$$|G(e^{j\omega})W(e^{j\omega})H(e^{j\omega})| < 1. \quad (2.8)$$

O problema em questão, portanto, consiste em modificar de alguma forma a resposta impulsiva de retorno local $h(n)$ para que (2.8) acima seja verificada.

Como se mencionou anteriormente, este problema será abordado nesta dissertação mediante o cancelamento do retorno local usando filtros adaptativos.

Capítulo 3

Aspectos Teóricos

Na continuação são discutidas as “ferramentas” usadas no Capítulo 4 para alcançar com sucesso os objetivos desta dissertação. Este capítulo é dividido em duas seções que correspondem ao estudo dos algoritmos dos filtros adaptativos e ao estudo dos filtros *notch*. Antes de abordar o estudo dos filtros adaptativos usados nesta dissertação, no início, são apresentados o filtro de Wiener e o algoritmo do gradiente descendente como preliminares para a abordagem do LMS estudado na Seção 3.1.3, o ϵ -NLMS na Seção 3.1.4 e o RLS na Seção 3.1.5 [13], [14]. Além disso, a Seção 3.1.6 apresenta para fins de comparação a complexidade computacional dos algoritmos dos filtros mencionados. Por fim, na Seção 3.2 são apresentados os filtros *notch*, usados para eliminar o tom de discar enviado pela central telefônica como será visto.

3.1 Filtragem adaptativa

O projeto de filtros digitais com coeficientes fixos requer especificações bem definidas. No entanto, há situações onde as especificações não estão disponíveis ou variam no tempo. A solução nestes casos é o uso de filtros digitais com coeficientes adaptativos conhecidos como filtros adaptativos [15]. Inicialmente, se faz uma revisão do filtro Wiener e o algoritmo do gradiente descendente. Embora eles não sejam utilizados nesta dissertação o seu estudo se faz necessário para ter uma melhor compreensão dos algoritmos aqui apresentados.

3.1.1 Filtro de Wiener

Consideremos os sinais $x(n)$ e $d(n)$ na Figura 3.1 estacionários no sentido amplo e relacionados estatisticamente. O filtro $\hat{H}(z)$ será definido como filtro de Wiener quando a sua saída $\hat{y}(n)$ for a melhor aproximação do sinal $d(n)$ no sentido de que o módulo quadrático médio do erro $e(n)$ é minimizado. Assumindo que $\hat{H}(z)$ é um filtro FIR (*finite impulse response*) com resposta impulsiva $\hat{\mathbf{h}} = [\hat{h}_0 \hat{h}_1 \cdots \hat{h}_{M-1}]^T$, definimos o erro como

$$e(n) = d(n) - \mathbf{x}^T(n) \hat{\mathbf{h}} \quad (3.1)$$

sendo $\mathbf{x}(n) = [x(n) x(n-1) \cdots x(n-M+1)]^T$. Assumindo $\hat{\mathbf{h}}$ fixo determinamos a média de $e^2(n)$

$$E[e^2(n)] = E[d^2(n)] - 2E[d(n) \mathbf{x}^T(n)] \hat{\mathbf{h}} + \hat{\mathbf{h}}^T E[\mathbf{x}(n) \mathbf{x}^T(n)] \hat{\mathbf{h}} \quad (3.2)$$

Definimos $\mathbf{R} = E[\mathbf{x}(n) \mathbf{x}^T(n)]$ e $\mathbf{p} = E[d(n) \mathbf{x}(n)]$ como a matriz de autocorrelação do vetor de entrada $\mathbf{x}(n)$ e o vetor de correlação cruzada entre os elementos do vetor $\mathbf{x}(n)$ e o sinal $d(n)$. Portanto, (3.2) pode ser reescrito como

$$E[e^2(n)] = E[d^2(n)] - 2\mathbf{p}^T \hat{\mathbf{h}} + \hat{\mathbf{h}}^T \mathbf{R} \hat{\mathbf{h}}. \quad (3.3)$$

Para determinar o vetor de coeficientes $\hat{\mathbf{h}}$ que minimiza esta função, calculamos o seu gradiente

$$\nabla_{\hat{\mathbf{h}}} E[e^2(n)] = 2\mathbf{R} \hat{\mathbf{h}} - 2\mathbf{p}. \quad (3.4)$$

Igualando este resultado a $\mathbf{0}$ e assumindo a não singularidade de $\mathbf{R}$, obtemos

$$\hat{\mathbf{h}}_0 = \mathbf{R}^{-1} \mathbf{p} \quad (3.5)$$

definido como a solução de Wiener. Embora com (3.5) se consiga a melhor estimativa de $d(n)$ no sentido quadrático médio, o uso prático dele em aplicações em tempo real é em geral pouco viável. Por exemplo, em aplicações em que o ambiente não seja estacionário. Além disso, o custo computacional necessário para calcular a solução de Wiener, poderá ser muito dispendioso, devido ao cálculo da inversa de $\mathbf{R}$.

Figura 3.1: Representação em diagramas de blocos do problema da filtragem adaptativa. Figura adaptada de [13]

3.1.2 O algoritmo do gradiente descendente

O algoritmo do gradiente descendente propõe determinar a solução de Wiener ajustando iterativamente $\hat{\mathbf{h}}(n)$ a partir de um valor inicial $\hat{\mathbf{h}}(0)$, diminuindo assim, a cada iteração, o valor da função custo até o seu valor mínimo. Os ajustes são obtidos segundo

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) + \frac{1}{2}\mu [-\nabla J(n)], \quad (3.6)$$

sendo $\hat{\mathbf{h}}(n+1)$ o novo vetor de coeficientes, $\nabla J(n)$ o gradiente da função custo $J(n)$ com respeito a $\hat{\mathbf{h}}(n)$, μ o passo de adaptação. É de se esperar que para $n \rightarrow \infty$, $\hat{\mathbf{h}}(n)$ convirja para a solução de Wiener. Similarmente a (3.4) determinamos que o gradiente da função custo é definido por

$$\nabla_{\hat{\mathbf{h}}(n)} J(n) = 2\mathbf{R}\hat{\mathbf{h}}(n) - 2\mathbf{p}. \quad (3.7)$$

Portanto, substituindo (3.7) em (3.6) obtemos

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) + \mu (\mathbf{p} - \mathbf{R}\hat{\mathbf{h}}(n)). \quad (3.8)$$

Cabe ressaltar, que a diferença do filtro de Wiener o algoritmo do gradiente descendente evita o cálculo da inversa de $\mathbf{R}$ na determinação da solução de Wiener.

3.1.3 O algoritmo Least Mean Square (LMS)

A medida exata do vetor gradiente em (3.6) implica o conhecimento a priori da matriz de autocorrelação $\mathbf{R}$ do vetor de entrada e do vetor de correlação cruzada $\mathbf{p}$ entre o vetor de entrada e o sinal desejado.

No algoritmo LMS, em (3.8), substituem-se $\mathbf{R}$ e $\mathbf{p}$ pelas suas estimativas instantâneas $\hat{\mathbf{R}}(n) = \mathbf{x}(n) \mathbf{x}^T(n)$ e $\hat{\mathbf{p}}(n) = d(n) \mathbf{x}(n)$. Obtendo-se

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) + \mu \mathbf{x}(n) \left[d(n) - \mathbf{x}^T(n) \hat{\mathbf{h}}(n) \right]. \quad (3.9)$$

Lembrando (3.1), (3.9) pode ser escrito como

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) + \mu \mathbf{x}(n) e(n), \quad (3.10)$$

sendo μ a constante que controla o tamanho do passo do algoritmo. Para garantir a convergência dos coeficientes na média o passo de adaptação deve ser escolhido de tal forma que obedeça

$$0 < \mu < \frac{2}{\lambda_{\max}}, \quad (3.11)$$

sendo $\lambda_{\max}$ o maior autovalor de $\mathbf{R}$.

3.1.4 O algoritmo LMS normalizado

Na equação do algoritmo LMS pode-se ver que $\mu e(n) \mathbf{x}(n)$ é diretamente proporcional ao vetor de entrada $\mathbf{x}(n)$. Portanto, um vetor $\mathbf{x}(n)$ de norma grande tende a causar uma mudança maior em $\hat{\mathbf{h}}$ do que um vetor $\mathbf{x}(n)$ de norma menor. Em aplicações envolvendo sinais de voz isso se torna um problema pelo fato dos sinais de voz apresentarem momentos de atividades separados por pausas, de modo que a norma varia consideravelmente. Para solucionar este problema o passo de adaptação é dividido pela norma quadrática do vetor $\mathbf{x}(n)$

$$\mu(n) = \frac{\tilde{\mu}}{\|\mathbf{x}(n)\|^2}. \quad (3.12)$$

Para evitar problemas com $\|\mathbf{x}(n)\|$ muito pequeno ou nulo é somado ao denominador uma constante positiva ϵ . Portanto o algoritmo ϵ -NLMS fica descrito por

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) + \frac{\tilde{\mu}}{\epsilon + \|\mathbf{x}(n)\|^2} e(n) \mathbf{x}(n). \quad (3.13)$$

A estabilidade do algoritmo será garantido se $\tilde{\mu}$ satisfaz

$$0 < \tilde{\mu} < 2. \quad (3.14)$$

3.1.5 O algoritmo Recursive Least Square (RLS)

Este algoritmo baseia-se no uso do método dos mínimos quadrados. Este método procura encontrar o $\hat{\mathbf{h}}$ que minimiza a função custo $\xi = \sum_{i=i_1}^{i_2} e^2(i)$, sendo $e(i) = d(i) - \mathbf{x}^T(i) \hat{\mathbf{h}}$, i_1 o limite inferior das observações tomadas dos sinais e i_2 o limite superior. O RLS permite encontrar o $\hat{\mathbf{h}}$ que minimiza a função custo no tempo n dado o $\hat{\mathbf{h}}$ que minimiza a função custo no tempo $n - 1$ [13]. Para o RLS a função custo pode ser definida como

$$\xi(n) = \sum_{i=1}^n \lambda^{n-i} e^2(i) \quad (3.15)$$

sendo λ o fator de esquecimento que determina a relevância dos dados antigos. Tal fator deve ser escolhido na faixa

$$0 \ll \lambda \leq 1. \quad (3.16)$$

O valor ótimo de $\hat{\mathbf{h}}(n)$ pelo qual (3.15) é minimizada, é definido por

$$\hat{\mathbf{h}}_0(n) = \mathbf{R}_{\mathbf{x}}(n)^{-1} \mathbf{r}_{d\mathbf{x}}(n), \quad (3.17)$$

sendo $\mathbf{R}_{\mathbf{x}}(n)$ a matriz de autocorrelação determinística de $\mathbf{x}(n)$ definida por $\mathbf{R}_{\mathbf{x}}(n) = \sum_{i=1}^n \lambda^{n-i} \mathbf{x}(i) \mathbf{x}(i)^T$ e $\mathbf{r}_{d\mathbf{x}}(n)$ o vetor de correlação cruzada determinístico entre d e $\mathbf{x}$ definido por $\mathbf{r}_{d\mathbf{x}}(n) = \sum_{i=1}^n \lambda^{n-i} \mathbf{x}(i) d(i)$. Tanto $\mathbf{R}_{\mathbf{x}}(n)$ quanto $\mathbf{r}_{d\mathbf{x}}(n)$ podem ser escritas da seguinte forma

$$\mathbf{R}_{\mathbf{x}}(n) = \lambda \mathbf{R}_{\mathbf{x}}(n-1) + \mathbf{x}(n) \mathbf{x}(n)^T \quad (3.18)$$

e

$$\mathbf{r}_{d\mathbf{x}}(n) = \lambda \mathbf{r}_{d\mathbf{x}}(n-1) + \mathbf{x}(n) d(n). \quad (3.19)$$

A complexidade computacional na determinação de (3.17) pode ser um problema devido à presença do cálculo da inversa da matriz de autocorrelação determinística. Uma forma recursiva é obtida aplicando a identidade de inversão de matrizes que diz que dadas as

matrizes $\mathbf{X}$ e $\mathbf{Y}$ ambas de ordem m por m e as matrizes $\mathbf{D}$ de ordem n por m e $\mathbf{Q}$ de ordem m por n , definidas positivas e relacionadas segundo a equação

$$\mathbf{X} = \mathbf{Y}^{-1} + \mathbf{D}\mathbf{Q}^{-1}\mathbf{D}^T, \quad (3.20)$$

podemos obter a inversa de $\mathbf{X}$ como

$$\mathbf{X}^{-1} = \mathbf{Y} - \mathbf{Y}\mathbf{D}(\mathbf{Q} + \mathbf{D}^T\mathbf{Y}\mathbf{D})^{-1}\mathbf{D}^T\mathbf{Y}. \quad (3.21)$$

Identificando elementos de (3.20) com os de (3.18) segundo $\mathbf{X} = \mathbf{R}_{\mathbf{x}}(n)$, $\mathbf{Y}^{-1} = \lambda\mathbf{R}_{\mathbf{x}}(n-1)$, $\mathbf{D} = \mathbf{x}(n)$ e $\mathbf{Q} = 1$ expressamos $\mathbf{R}_{\mathbf{x}}^{-1}(n)$ como

$$\mathbf{R}_{\mathbf{x}}^{-1}(n) = \lambda^{-1}\mathbf{R}_{\mathbf{x}}^{-1}(n-1) - \frac{\lambda^{-2}\mathbf{R}_{\mathbf{x}}^{-1}(n-1)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{R}_{\mathbf{x}}^{-1}(n-1)}{1 + \lambda^{-1}\mathbf{x}^T(n)\mathbf{R}_{\mathbf{x}}^{-1}(n-1)\mathbf{x}(n)}. \quad (3.22)$$

Além disso, definimos

$$\mathbf{P}(n) = \mathbf{R}_{\mathbf{x}}^{-1}(n) \quad (3.23)$$

como a matriz de correlação inversa e

$$\mathbf{k}(n) = \frac{\lambda^{-1}\mathbf{P}(n-1)\mathbf{x}(n)}{1 + \lambda^{-1}\mathbf{x}^T(n)\mathbf{P}(n-1)\mathbf{x}(n)} \quad (3.24)$$

como o vetor ganho. Então, podemos reescrever (3.22) como

$$\mathbf{P}(n) = \lambda^{-1}\mathbf{P}(n-1) - \lambda^{-1}\mathbf{k}(n)\mathbf{x}^T(n)\mathbf{P}(n-1). \quad (3.25)$$

Reorganizando (3.24) temos

$$\mathbf{k}(n) = \mathbf{P}(n)\mathbf{x}(n). \quad (3.26)$$

Utilizando (3.17) , (3.19) , (3.23) e (3.26) definimos $\hat{\mathbf{h}}(n)$ por

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mathbf{k}(n)\alpha(n), \quad (3.27)$$

sendo $\alpha(n)$ o erro de estimação a priori definido por

$$\alpha(n) = d(n) - \hat{\mathbf{h}}^T(n-1)\mathbf{x}(n). \quad (3.28)$$

O algoritmo RLS apresenta uma alta velocidade de convergência comparado com o LMS mas, em contrapartida, apresenta uma alta complexidade computacional. Cabe notar que o valor inicial de $\mathbf{P}$, deve ser escolhido de uma forma que assegure a não-singularidade da matriz de autocorrelação determinística.

Uma forma simples de fazer isso é modificar a expressão $\mathbf{R}_{\mathbf{x}}(n)$ segundo

$$\mathbf{R}_{\mathbf{x}}(n) = \sum_{i=1}^n \lambda^{n-i} \mathbf{x}(i) \mathbf{x}(i)^T + \delta \lambda^n \mathbf{I} \quad (3.29)$$

sendo $\mathbf{I}$ a matriz identidade e δ uma constante positiva pequena. Portanto colocando $n = 0$ na equação (3.29) temos

$$\mathbf{R}_{\mathbf{x}}(0) = \delta \mathbf{I}. \quad (3.30)$$

Correspondentemente, para $\mathbf{P}(0)$ temos

$$\mathbf{P}(0) = \delta^{-1} \mathbf{I}. \quad (3.31)$$

3.1.6 Complexidade computacional

Para fins de comparação, podemos observar na Tabela 3.1 a quantidade de operações aritméticas necessárias por iteração para os algoritmos LMS, ϵ -NLMS e RLS, considerando sinais reais.

Algoritmo	Multiplicações	Somas	Divisões
LMS	$2M + 1$	$2M$	0
ϵ -NLMS	$3M + 1$	$3M$	1
RLS	$M^2 + 5M + 1$	$M^2 + 3M$	1

Tabela 3.1: Complexidade computacional por iteração dos algoritmos LMS, ϵ -NLMS e RLS. Informação extraída de [14].

3.2 Filtros *notch*

Os filtros *notch* são utilizados para eliminar sinais de banda estreita ou senoidais de um sinal de banda larga [16]. Filtros *notch* são úteis em muitas aplicações, por exemplo, na remoção

do ruído de 60 Hz da rede elétrica de sinais de eletrocardiograma. Podem ser divididos em filtros *notch* FIR (*finite impulse response*) e IIR (*infinite impulse response*). Nesta dissertação, serão utilizados os filtros *notch* IIR de segunda ordem com os zeros restritos ao círculo unitário de ângulo igual à frequência *notch* e com os polos localizados no mesmo ângulo que os zeros, mas com a diferença que o seu raio ρ é restrito a $0 \ll \rho < 1$. A função de transferência do filtro com as características mencionadas é mostrada na equação seguinte

$$F(z) = \frac{z^2 - 2\cos\omega_0 z + 1}{z^2 - 2\rho\cos\omega_0 z + \rho^2} \quad (3.32)$$

onde ω_0 é definido por $\omega_0 = 2\pi\frac{f_0}{f_s}$, sendo f_s é a frequência de amostragem e f_0 a frequência que se deseja rejeitar. Esse filtro será de utilidade na Seção 4.4, para a eliminação do sinal de tom de discar.

Capítulo 4

Cancelamento de retorno local em aparelhos telefônicos para deficientes auditivos

As discussões dos capítulos anteriores foram feitas para construir uma base para o desenvolvimento deste capítulo. Como mencionado, a condição suficiente desenvolvida na Seção 2.6 para impedir a instabilidade e, em consequência, a microfonia no aparelho telefônico para deficientes auditivos, pode ser obtida mediante filtragem adaptativa de modo a cancelar o retorno local. Mas antes de abordar o cancelamento do retorno local uma medida do seu desempenho deve ser desenvolvida. Para isso, na Seção 4.1, consideramos o problema do dimensionamento da amplificação introduzida para uma efetiva compensação da deficiência auditiva do usuário. Nesta seção, será visto que é mais conveniente levantar o desempenho dos filtros adaptativos diretamente em termos da amplificação máxima que pode ser introduzida graças ao cancelamento do retorno local, e não em termos de medidas mais usuais como o erro quadrático médio em excesso. Uma forma de fazer isto é proposta na Seção 4.2. Considerando essa medida de desempenho, foi definido um objetivo de 55 dB de amplificação máxima a ser alcançada pelo cancelamento do retorno local. Finalmente, na Seção 4.3 o desempenho dos filtros adaptativos mencionados são avaliados. Primeiro, os algoritmos são avaliados com sinal de fala como entrada. Os resultados mostram que este método não alcança o objetivo determinado. Visto isto, a adaptação antes da conversação é avaliada na Seção 4.4, com sinal branco gaussiano como entrada, acrescentando-se um filtro notch para

eliminar o tom de discar. Os resultados mostram que o algoritmo LMS é suficiente para alcançar o objetivo mencionado.

4.1 Amplificação para compensação de deficiência

Abordar o dimensionamento da amplificação introduzida para uma efetiva compensação da deficiência auditiva do usuário pode ser um problema consideravelmente complexo e que está fora do escopo deste trabalho. Não obstante, será suficiente considerar aqui um exemplo simples. Na Figura 4.1, uma perda auditiva de 60 dB é considerada na área de audição apresentada na Figura 2.5. Portanto a faixa dinâmica de audição normal (DR1 - *dynamic range of hearing*) de aproximadamente 140 dB se reduz até uma faixa dinâmica residual (DR2 - *residual dynamic range*) de aproximadamente 80 dB. Portanto, boa parte da janela dos sons de fala telefônica (considerada na figura como a área que encerra o retângulo, inicialmente localizada abaixo da linha que define a perda auditiva), reduzindo-se sua faixa dinâmica (DRP - *dynamic range*) de 45 – 50 dB aproximadamente até 5 ou 10 dB. Então a parte da faixa dinâmica que foi perdida não será escutada pela pessoa com perda auditiva.

Figura 4.1: Efeito da perda auditiva sobre a área auditiva.

Para solucionar este problema, os aparelhos para deficientes auditivos devem deslocar a janela para cima, para que ela fique na faixa dinâmica residual.

Portanto dada uma perda auditiva em determinada frequência será necessário compensá-la com um certo ganho. As perdas auditivas podem variar ao longo da faixa de frequências. Mas cabe notar da Tabela 2.2 que para uma classificação de grau de perda auditiva será desejável em algum momento introduzir até uma certa quantidade de ganho em alguma frequência, sendo que a viabilidade disso dependerá da efetividade do cancelamento do retorno local. Neste contexto, é mais conveniente levantar o desempenho dos filtros adaptativos diretamente em termos da amplificação máxima que pode ser introduzida graças ao cancelamento do retorno local, e não em termos de medidas mais usuais como o erro quadrático médio em excesso, trabalho que será abordado na próxima seção.

4.2 Objetivo de cancelamento de retorno local

Partimos do pressuposto razoável de que o aparelho telefônico utilizado tenha sido projetado de modo a não apresentar microfonia com $G(z) \equiv 1$, condição que corresponde ao seu uso convencional, sem compensação auditiva. Com (2.8) uma condição suficiente para isto é

$$|W(e^{j\omega}) H(e^{j\omega})| < 1, \text{ para todo } \omega. \quad (4.1)$$

Na prática, porém, esta condição é pouco realista, no sentido de que o limite superior estipulado é bem maior do que o limite superior efetivamente verificado. De fato, dada a variabilidade do acoplamento acústico $W(e^{j\omega})$ com o posicionamento do monofone em relação ao usuário, é pouco provável que $|W(e^{j\omega}) H(e^{j\omega})|$ assumam valores próximos ao limite superior unitário sem nunca porém ultrapassá-lo. Portanto, no projeto do aparelho para uso convencional, é de se esperar que, para evitar microfonia, tenha sido empregada uma condição que resulta ser mais restritiva para $W(e^{j\omega})$.

Uma condição mais restritiva para $W(e^{j\omega})$, porém mais realista no sentido acima, pode ser obtida notando inicialmente que, para todo ω ,

$$|W(e^{j\omega}) H(e^{j\omega})| \leq |W(e^{j\omega})| \|H(e^{j\omega})\|_{\infty} \quad (4.2)$$

onde $\|H(e^{j\omega})\|_{\infty} \triangleq \max_{\omega} |H(e^{j\omega})|$ (vide Seção 2.3). Para a não ocorrência de microfonia, uma condição suficiente é então

$$|W(e^{j\omega})| \|H(e^{j\omega})\|_{\infty} < 1, \quad (4.3)$$

o que equivale a $\|W(e^{j\omega})\|_{\infty} \|H(e^{j\omega})\|_{\infty} < 1$, ou ainda,

$$\|W(e^{j\omega})\|_{\infty} < \frac{1}{\|H(e^{j\omega})\|_{\infty}}. \quad (4.4)$$

Consideremos agora o esquema básico de um aparelho telefônico para deficientes auditivos com um cancelador de retorno local, apresentado na Figura 4.2. Na figura, $\hat{H}(z)$ é uma estimativa de $H(z)$ a ser obtida adaptativamente. Uma condição suficiente para que não ocorra microfonia é

$$\|G(e^{j\omega})\|_{\infty} \|W(e^{j\omega})\|_{\infty} \|H(e^{j\omega}) - \hat{H}(e^{j\omega})\|_{\infty} < 1, \quad (4.5)$$

pois

$$\begin{aligned} \left| G(e^{j\omega}) W(e^{j\omega}) [H(e^{j\omega}) - \hat{H}(e^{j\omega})] \right| &\leq \\ &\leq \|G(e^{j\omega}) W(e^{j\omega}) [H(e^{j\omega}) - \hat{H}(e^{j\omega})]\|_{\infty} \leq \\ &\leq \|G(e^{j\omega})\|_{\infty} \|W(e^{j\omega})\|_{\infty} \|H(e^{j\omega}) - \hat{H}(e^{j\omega})\|_{\infty} < 1, \end{aligned}$$

para todo ω . Por sua vez, com (4.4), a condição

$$\|G(e^{j\omega})\|_{\infty} \frac{1}{\|H(e^{j\omega})\|_{\infty}} \|H(e^{j\omega}) - \hat{H}(e^{j\omega})\|_{\infty} < 1$$

garante que (4.5) seja satisfeita. Reorganizando esta expressão, chegamos ao objetivo a ser estipulado ao bloco de cancelamento adaptativo de retorno local:

$$K \triangleq \frac{\|H(e^{j\omega}) - \hat{H}(e^{j\omega})\|_{\infty}}{\|H(e^{j\omega})\|_{\infty}} < \frac{1}{\|G(e^{j\omega})\|_{\infty}}, \quad (4.6)$$

onde, com base no discutido na Seção 2.5 será adotado $\|G(e^{j\omega})\|_{\infty} = 55 \text{ dB}$.

Figura 4.2: Esquema básico de um aparelho telefônico para deficientes auditivos com cancelador de retorno local

4.3 Desempenho com sinal de fala

Verifiquemos agora o desempenho, em termos do objetivo (4.6), dos algoritmos adaptativos LMS, ϵ -NLMS e RLS. Para tanto, considerando o discutido na Seção 2.6, definamos, como [17],

$$\mathbf{h}^T = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.4, 0.3, 0.2, 0.1\} \quad (4.7)$$

e $x(n)$ dado pelo sinal em A) test/dr1/faks0/sa1.wav ou B) train/dr1/fcjf0/sa1.wav da base de dados TIMIT [18], normalizados para potência unitária. Mas é necessário considerar que, durante a adaptação, enquanto um certo nível de cancelamento de retorno local não for atingido, o aparelho deverá operar com funcionalidade limitada, em modo *half-duplex*, interrompendo adequadamente o sinal do microfone ou do alto-falante conforme o lado ativo da conversação e fazendo a adaptação somente quando o usuário próximo estiver ativo. Devido a isto, é necessário que um cancelamento de retorno local suficiente seja atingido em um tempo suficientemente pequeno para não prejudicar a conversação. Adotamos para este tempo um objetivo de 1 segundo após o início da conversação.

Portanto, com (4.6), o objetivo a ser alcançado pelo cancelamento de retorno local é obter (e manter depois) $K < -55$ dB em um tempo de 1 segundo, ou seja, em $n = 16 \times 10^3$ amostras, considerando uma taxa de amostragem de 16 kHz. Mais formalmente, o objetivo pode ser colocado como

$$\max_{n \geq 16000} K(n) < -55 \text{ dB}. \quad (4.8)$$

Também, considerando que a adaptação é feita quando apenas o usuário próximo está ativo, $v(n)$ é devido ao ruído ambiente captado no microfone do usuário distante e também ao ruído somado ao longo da transmissão. Além disso, consideramos que, estando em modo *half-duplex*, o sinal para o alto-falante é cortado, levando a $t(n) \equiv 0$ e $x(n) \equiv u(n)$.

Consideremos, inicialmente, o caso ideal de ruído de medida $v(n)$ nulo. Na Figura 4.3 temos os resultados para os algoritmos LMS e ϵ -NLMS e na Figura 4.4 temos os resultados para o algoritmo RLS. Os parâmetros de cada algoritmo foram ajustados priorizando a velocidade de convergência. Mesmo com este ajuste, podemos ver que em $n = 16 \times 10^3$ e para sinais de fala típicos de conversações, o LMS e o ϵ -NLMS ainda estão no transitório da adaptação, e este transitório depende fortemente do sinal particular $x(n)$. O algoritmo RLS está logo no final do transitório. Mesmo no caso ideal sem ruído, portanto, o uso do LMS e NLMS seria problemático.

Figura 4.3: Adaptação com sinais de fala, sem ruído de medida. Traços **a** e **b**: sinais $x(n)$ A e B, respectivamente (fora de escala); **c** e **d**: $K(n)$ para LMS com $\mu = 0,007$ (próximo ao máximo μ), para sinais A e B, respectivamente; **e** e **f**: $K(n)$ para ϵ -NLMS com $\mu = 0,1$ e $\epsilon = 0$, para sinais A e B, respectivamente.

Figura 4.4: Adaptação com sinais de fala, sem ruído de medida. Traços **a** e **b**: sinais $x(n)$ A e B, respectivamente (fora de escala); **c** e **d**: $K(n)$ para RLS com $\lambda = 0,999$ e $\mathbf{P}(0) = 10^6 \mathbf{I}$, para sinais A e B, respectivamente.

Consideremos agora, mais realisticamente, ruído de medida não-nulo, com $v(n)$ branco e gaussiano, com uma relação sinal/ruído (SNR - *signal-to-noise ratio*) $E[y^2(n)]/E[v^2(n)] = 30$ dB. Os resultados médios para 100 realizações do ruído estão na Figura 4.5, agora com os parâmetros dos algoritmos adaptativos ajustados para obter o menor valor possível para $\max_{n \geq 16000} K(n)$. Podemos ver que nenhum dos algoritmos considerados alcança o objetivo (4.8).

Figura 4.5: Adaptação com sinais de fala, ruído de medida com $\text{SNR} = 30$ dB. Traço **a**: sinal $x(n)$ B (fora de escala); **b, c** e **d** : $K(n)$ para, respectivamente, LMS com $\mu = 0,005$, ϵ -NLMS com $\mu = 0,005$ e $\epsilon = 0,09$, RLS com $\lambda = 1$ e $\mathbf{P}(0) = 10^3 \mathbf{I}$.

4.4 Adaptação anterior à conversação

Como visto na seção anterior, o procedimento convencional de iniciar a adaptação juntamente com a conversação telefônica não alcançou o desempenho necessário para a aplicação em questão aqui. Consideremos então que, quando não em uso, o aparelho telefônico, de forma automática e periódica, realiza o equivalente a tirar o fone do gancho, recebendo portanto o tom de discar de 425 Hz da central telefônica (vide Tabela 2.1), e em seguida faz a adaptação de $\hat{H}(z)$. Nesta situação, estabelecemos que um valor suficiente de cancelamento de retorno deva ser atingido em 3 segundos, ou seja, $n = 48000$. Este período de tempo foi escolhido de forma arbitrária considerando que o objetivo da adaptação deve ser atingido antes do tom de discar se transformar em tom de ocupado. Como o sinal $x(n)$ pode ser agora escolhido conforme desejado, adota-se, para maior velocidade de convergência, um sinal branco gaussiano de potência unitária. O ruído de medida é dado agora por

$$v(n) = v_I(n) + v_N(n), \quad (4.9)$$

onde $v_I(n)$ é o tom de discar e $v_N(n)$ é ruído somado ao longo da transmissão. Para eliminar a interferência do tom de discar no processo adaptativo, emprega-se um filtro notch segundo

(3.32) para $\omega_0 = \frac{2\pi 425}{16000}$ e, para um efeito bastante reduzido fora da frequência 425 Hz, $\rho = 0,99$. Como representado na Figura 4.6, durante a adaptação o sinal de erro $e(n)$ passa por este filtro notch antes de ser usado pelo algoritmo adaptativo. (Notar que a adaptação é realizada antes da conversação telefônica e, portanto, a parte que inclui o bloco $G(z)$ e o monofone na Figura 4.2 está desativada). Na prática, a frequência do tom de discar pode apresentar variações de até algumas de dezenas de Hz em relação à frequência nominal e um filtro notch adaptativo tem que ser usado. Sua adaptação partiria da frequência nominal de 425 Hz e seria bastante rápida, podendo ser realizada antes da injeção de $x(n)$ e adaptação de $\hat{H}(z)$. Não há problema, portanto, em considerar aqui um filtro *notch* fixo.

Figure 4.6: Adição de filtro *notch* para eliminar tom de discar.

Apesar de $v_N(n)$ não conter mais o ruído ambiente captado da outra extremidade da conexão, consideramos ainda a SNR

$$\frac{E[y^2(n)]}{E[v_N^2(n)]} = 30 \text{ dB}. \quad (4.10)$$

Consideramos também, realisticamente, uma relação sinal/interferência (SIR - *signal-to-interference ratio*)

$$\frac{E[y^2(n)]}{E[v_I^2(n)]} = 0 \text{ dB}. \quad (4.11)$$

Na Figura 4.7 temos o resultado deste procedimento para o algoritmo LMS. Podemos ver que o objetivo

$$\max_{n \geq 48000} K(n) < -55 \text{ dB} \quad (4.12)$$

é atingido com certa folga. Dado que o LMS já satisfaz este objetivo, não há motivação para considerar os algoritmos ϵ -NLMS e RLS, mais complexos computacionalmente.

Figura 4.7: Adaptação com sinal de entrada branco e gaussiano, ruído de medida com $\text{SNR} = 30$ dB e interferência senoidal com $\text{SNR} = 0$ dB. Traços **a** e **b**: LMS, $\mu = 0,0002$, sem e com o filtro *notch*, respectivamente.

Capítulo 5

Conclusões e perspectivas

Foi determinado um objetivo de desempenho para o cancelamento de retorno local em aparelhos telefônicos para deficientes auditivos. A utilização convencional dos algoritmos adaptativos LMS, ϵ -NLMS e RLS não consegue, porém, alcançar tal objetivo. Por outro lado, isto é obtido já pelo algoritmo LMS, com o procedimento proposto de injetar uma excitação e iniciar a adaptação antes do início da conversação, usando ainda um filtro notch sobre o sinal de erro da adaptação, para eliminar o tom de discar. Dado que a resposta impulsiva do retorno local pode variar, ainda que lentamente, durante a conversação, um passo seguinte neste trabalho a considerar é a continuidade da adaptação durante a conversação, de modo a rastrear estas variações.

Referências Bibliográficas

- [1] ENGBRETSON, A. M. Benefits of Digital Hearing Aids. *IEEE Engineering in Medicine and Biology*, v. 13, No. 2, p. 238-248, Apr. 1, 1994.
- [2] LOCTWASSANA, W.; PUNCHALARD, R.; LORSAWATSIRI, A.; KOSCCYAPORN; WARDKCIN, P. Adaptive howling suppressor in an audio amplifier system. *Communications, 2007. APCC 2007. Asia-Pacific Conference on*, p. 445-448, Oct. 18-20, 2007.
- [3] JESZENSKY, P. J. E. *Sistemas telefônicos*. Manole, Barueri, São Paulo, 2004.
- [4] NOLL, A. M. *Introduction to telephones and telephone systems*. Artech House telecommunications library, Boston-London, 1999.
- [5] CAREZIA, A. H. C. *Redução adaptativa de eco e de ruído para terminais viva voz*. Dissertação de Mestrado em Engenharia Elétrica - Escola Politécnica da Universidade de São Paulo, 2002.
- [6] BISPO, B. C. *Cancelamento de eco elétrico em redes telefônicas e acústico em sistemas de teleconferência*. Dissertação de Mestrado em Engenharia Elétrica - Universidade Federal do Rio de Janeiro, 2008.
- [7] SINHA, N. K. *Control systems*. 3ª edição, New Age International Publisher, New Delhi, 1998.
- [8] SPIEGEL, M. R. *Variáveis complexas*. Editora McGraw-Hill do Brasil, 1973 (Coleção Schaum).
- [9] "O ouvido humano".[Online]. <http://www.if.ufrj.br/teaching/fis2/ondas2/ouvido/ouvido.html>
- [10] MOORE B. C. J. *An introduction to the psychology of hearing*. 5ª edição, Academic Press, San Diego, 2003.

- [11] ZWICKER E.; FASTL H. *Psycho-acoustics*. 3^a edição, Springer-Verlag, New York, 1999.
- [12] DOBIE, R. A.; VAN HEMEL, S. B. *Hearing loss: determining eligibility for social security benefits*. The National Academies Press, Washington D.C., 2005.
- [13] HAYKIN, S. *Adaptive Filter Theory*. Prentice Hall, Upper Saddle River, 2004.
- [14] SAYED, A. H. *Fundamentals of adaptive filtering*. Wiley, New Jersey, 2003.
- [15] DINIZ, P. S. R. *Adaptive filtering algorithms and practical implementation*. 3^a edição, Springer Science+Business Media, LLC, New York, 2008.
- [16] GIL-CACHO, Pepe; VAN WATERSCHOOT, T.; MOONEN, M.; JENSEN, S. H. Regularized adaptive notch filters for acoustic howling suppression. *in Proc. 17th European Signal Process. Conf. (EUSIPCO '09)*, p. 2574-2578, Glasgow, Scotland, UK, Aug. 2009.
- [17] BANJAC, Z.; KOVA EVI, B.; MILOSAVLJEVI, M.; VEINOVI, M. Local echo canceller with optimal input design for true full-duplex speech scrambling system. *IEEE Transaction on Signal Processing*, p.1877-1882, V. 50, No.8, 2002.
- [18] The DARPA TIMIT Acoustic-Phonetic Continuous Speech Corpus (TIMIT) CD-ROM.