

# Cancelamento de retorno local em aparelhos telefônicos para deficientes auditivos

Reynaldo P. Condori e Phillip M. S. Burt

**Resumo**—Em um aparelho telefônico para deficientes auditivos, para poder compensar a deficiência do usuário é necessário impedir a instabilidade de malha fechada que produziria microfonia, o que é feito neste trabalho pelo cancelamento adaptativo do retorno local da híbrida. Determinamos inicialmente um objetivo de cancelamento diretamente dependente da amplificação máxima a ser introduzida. Verificamos então que o uso convencional dos algoritmos LMS,  $\epsilon$ -NLMS e RLS não alcança este objetivo. Iniciamos então a adaptação antes da conversação, acrescentando um filtro *notch* para eliminar o tom de discar. Com isto, o algoritmo LMS é suficiente para alcançar o objetivo mencionado.

**Palavras-Chave**—Cancelamento de retorno local, microfonia, filtragem adaptativa.

**Abstract**—In a telephone set for the hearing impaired, in order to compensate for the user's impairment, it is necessary to avoid closed-loop instability that would produce howling, which is done in this work by adaptive cancellation of the local return in the hybrid. Initially we establish a goal for local return cancellation that depends directly on the maximum amplification to be introduced. We verify then that the conventional use of the LMS,  $\epsilon$ -NLMS and RLS algorithms doesn't attain this goal. We start then the adaptation before the conversation, adding a notch filter to suppress the dialing tone. With this procedure, the LMS algorithm is sufficient for attaining the mentioned goal.

**Keywords**—Local return cancellation, howling, adaptive filtering.

## I. INTRODUÇÃO

Para melhorar a qualidade de vida de pessoas com deficiência auditiva 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 a cápsula e o microfone do aparelho (eco acústico), pode se tornar instável devido à introdução da amplificação, dando origem a microfonia (*howling*). 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. Aqui, portanto, seguimos esta segunda abordagem.

O problema básico como colocado acima constitui um problema padrão de filtragem adaptativa. Há, no entanto, certas particularidades importantes na aplicação em questão, como discutido a seguir. De início, consideremos o dimensionamento da amplificação introduzida para uma efetiva compensação da deficiência auditiva do usuário. Trata-se de um problema consideravelmente complexo e que está fora do escopo deste trabalho. É suficiente considerar aqui, no entanto, que pode ser desejável em algum momento introduzir até 55 dB de ganho em alguma frequência [1], sendo que a viabilidade disto 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. Uma forma de fazer isto é proposta na Seção III.

Além disto, 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 da cápsula conforme o lado ativo da conversação e fazendo a adaptação somente quando o usuário local 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, o que é analisado na Seção IV.

Finalmente, como visto na Seção V, o atendimento do objetivo acima leva ao problema de filtragem adaptativa na presença de interferência senoidal, para o qual é considerada a adição de um filtro *notch*.

## II. DEFINIÇÃO DO PROBLEMA

A Figura 1 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 a cápsula 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  $\mathbf{x}(n) = [x(n) x(n-1) \cdots x(n-N+1)]^T$  e

$\mathbf{h} = [h_0 \ h_1 \ \dots \ h_{N-1}]^T$  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.

A função de transferência de malha aberta do esquema da Figura 1 é  $G(z)W(z)H(z)$  e, pelo critério de Nyquist de estabilidade, poderá ocorrer microfonia em uma frequência  $\omega_0$  tal que

$$|G(e^{j\omega_0})W(e^{j\omega_0})H(e^{j\omega_0})| \geq 1 \quad (1)$$

$$\angle G(e^{j\omega_0})W(e^{j\omega_0})H(e^{j\omega_0}) = 0, \quad (2)$$

onde  $z_0 = e^{j\omega_0}$ . O problema em questão, portanto, consiste em modificar de alguma forma a resposta impulsiva de retorno local  $h(n)$  para que as condições acima não sejam verificadas.

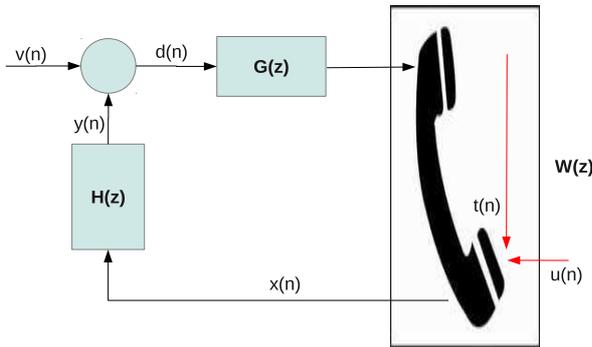


Fig. 1. Esquema básico de um aparelho telefônico para deficientes auditivos.

### III. 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 (1), uma condição suficiente para isto é

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

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  $|W(e^{j\omega})H(e^{j\omega})|$  assumisse valores próximos ao limite superior unitário sem nunca porém ultrapassá-lo. Portanto, no projeto do aparelho para uso convencional, é de esperar que, para evitar a 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  $\omega$ ,  $|W(e^{j\omega})H(e^{j\omega})| \leq |W(e^{j\omega})| \|H(e^{j\omega})\|_\infty$ , onde  $\|H(e^{j\omega})\|_\infty \triangleq \max_\omega |H(e^{j\omega})|$ . Para a não ocorrência de microfonia, uma condição suficiente é então

$$\|W(e^{j\omega})\| \|H(e^{j\omega})\|_\infty < 1, \text{ para todo } \omega,$$

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

$$\|W(e^{j\omega})\|_\infty < \frac{1}{\|H(e^{j\omega})\|_\infty}. \quad (3)$$

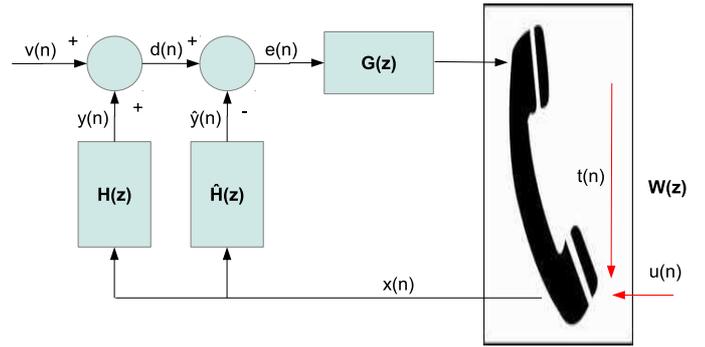


Fig. 2. Esquema básico de um aparelho telefônico para deficientes auditivos com cancelador de retorno local.

Consideremos agora o esquema básico de um aparelho telefônico para deficientes auditivos com um cancelador de retorno local, apresentado na Figura 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

$$\|G(e^{j\omega})\|_\infty \|W(e^{j\omega})\|_\infty \|H(e^{j\omega}) - \hat{H}(e^{j\omega})\|_\infty < 1, \quad (4)$$

pois

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

para todo  $\omega$ . Por sua vez, com (3), 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) é satisfeita. Reorganizando esta expressão, chegamos ao objetivo a ser estipulado ao bloco de cancelamento adaptativo de retorno local:

$$K \triangleq \frac{\|H(e^{j\omega}) - \hat{H}(e^{j\omega})\|_\infty}{\|H(e^{j\omega})\|_\infty} < \frac{1}{\|G(e^{j\omega})\|_\infty}, \quad (5)$$

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

### IV. DESEMPENHO COM SINAL DE FALA

Verifiquemos agora o desempenho, em termos do objetivo (5), dos algoritmos adaptativos largamente conhecidos LMS,  $\epsilon$ -NLMS e RLS [2], [3]. Para tanto, consideremos, como [4],

$$\mathbf{h}^T = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.4, 0.3, 0.2, 0.1\} \quad (6)$$

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, normalizados para potência unitária. Com (5) e com o que foi discutido na Introdução, 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

$$\max_{n \geq 16000} K(n) < -55 \text{ dB}. \quad (7)$$

Consideramos que a adaptação é feita quando apenas o usuário local está ativo e, portanto,  $v(n)$  é devido ao ruído ambiente captado no monofone do usuário distante e também ao ruído somado ao longo da transmissão. Além disso, consideramos que, estando em modo *half-duplex*, o sinal para a cápsula 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 3 temos os resultados para os algoritmos LMS e  $\epsilon$ -NLMS e na Figura 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  $\epsilon$ -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.

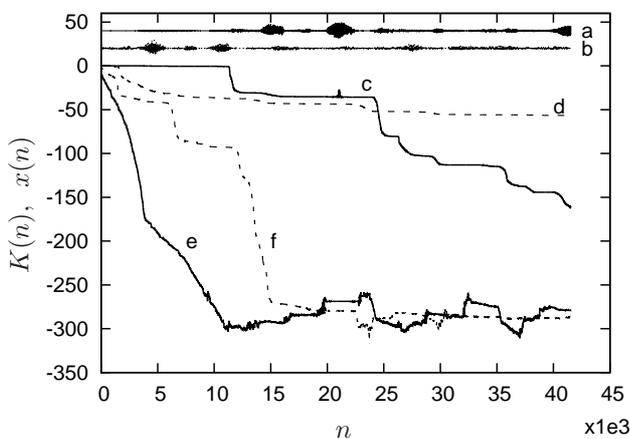


Fig. 3. Adaptação com sinais de fala, sem ruído de medida. Traços a e b: sinais  $x(n)$  A e B, respectivamente (fora de escala); c e d:  $K(n)$  para LMS com  $\mu = 0,007$  (próximo ao máximo  $\mu$ ), para sinais A e B, respectivamente; e e f:  $K(n)$  para  $\epsilon$ -NLMS com  $\mu = 0,1$  e  $\epsilon = 0$ , para sinais A e B, respectivamente.

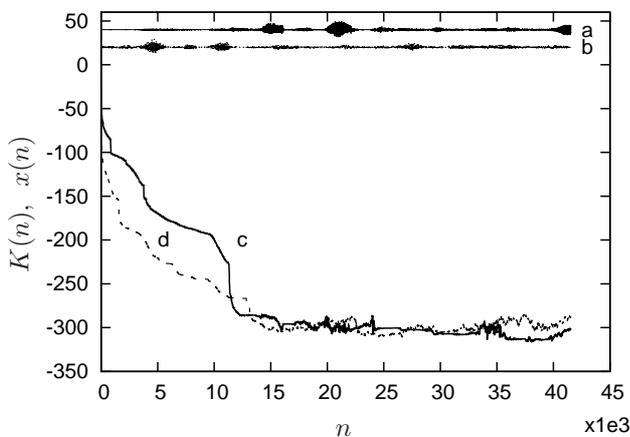


Fig. 4. Adaptação com sinais de fala, sem ruído de medida. Traços a e b: sinais  $x(n)$  A e B, respectivamente (fora de escala); c e d:  $K(n)$  para RLS com  $\lambda = 0,999$  e  $\mathbf{P}(0) = 10^6 \mathbf{I}$ , para sinais A e B, respectivamente.

Consideremos agora, mais realisticamente, ruído de medida não-nulo, com  $v(n)$  branco e gaussiano, com uma relação

sinal/ruído (SNR)  $E[y^2(n)]/E[v^2(n)] = 30$  dB. Os resultados médios para 10 realizações do ruído estão na Figura 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 (7).

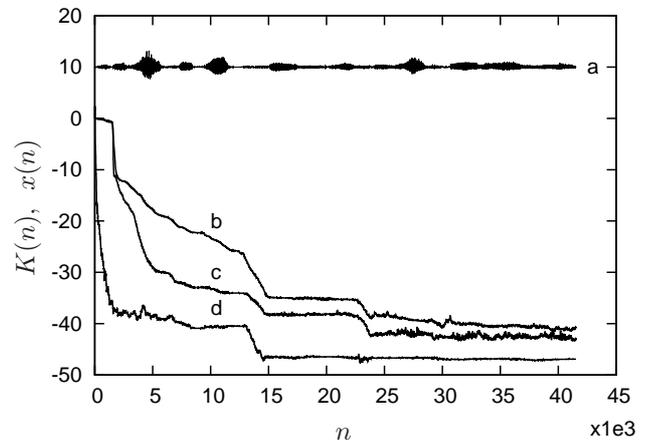


Fig. 5. Adaptação com sinais de fala, ruído de medida com SNR = 30 dB. Traço a: sinal  $x(n)$  B (fora de escala); b, c e d:  $K(n)$  para, respectivamente, LMS com  $\mu = 0,005$ ,  $\epsilon$ -NLMS com  $\mu = 0,005$  e  $\epsilon = 0,09$ , RLS com  $\lambda = 1$  e  $\mathbf{P}(0) = 10^3 \mathbf{I}$ .

## V. 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) 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,  $n = 48000$ . 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

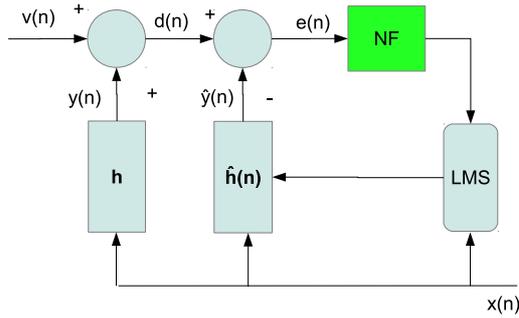
$$v(n) = v_I(n) + v_N(n),$$

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* com função de transferência

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

onde  $\omega_0 = 2\pi \times 425/16000$  e, para um efeito bastante reduzido fora da frequência 425 Hz,  $\rho = 0,99$ . Como representado na Figura 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 2 está desativada). Na prática, a frequência

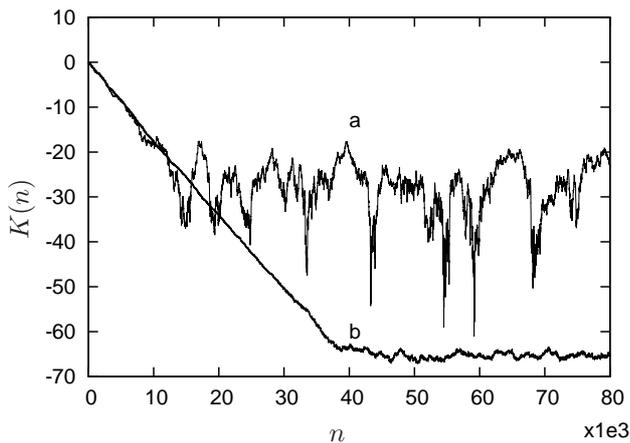
Fig. 6. Adição de filtro *notch* para eliminar tom de discar.

do tom de discar pode apresentar variações de até algumas de dezenas de Hz em relação à frequência nominal e um filtro *notch* adaptativo [5] 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.

Apesar de  $v_N(n)$  não conter mais o ruído ambiente captado da outra extremidade da conexão, consideramos ainda a SNR  $E[y^2(n)]/E[v_N^2(n)] = 30$  dB. Consideramos também, realisticamente, uma relação sinal/interferência (SIR)  $E[y^2(n)]/E[v_I^2(n)] = 0$  dB. Na Figura 7 temos o resultado deste procedimento para o algoritmo LMS. Podemos ver que o objetivo

$$\max_{n \geq 48000} K(n) < -55 \text{ dB.} \quad (8)$$

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

Fig. 7. Adaptação com sinal de entrada branco e gaussiano, ruído de medida com SNR = 30 dB e interferência senoidal com SNR = 0 dB. Traços a e b: LMS,  $\mu = 0,0002$ , sem e com o filtro *notch*, respectivamente.

## VI. CONCLUSÃO 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,  $\epsilon$ -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 é considerar a continuidade da adaptação durante a conversação, de modo a rastrear estas variações.

## REFERÊNCIAS

- [1] E. Zwicker and H. Fastl, *Psycho-acoustics*, Springer-Verlag, New York, 1999.
- [2] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Upper Saddle River, 2004.
- [3] A. H. Sayed, *Fundamentals of Adaptive Filtering*, Wiley, New Jersey, 2003.
- [4] Z. Banjac et al., "Local echo canceller with optimal input design for true full-duplex speech scrambling system," *IEEE Transactions on Signal Processing*, vol. 50, no. 8, pp. 1877–1882, Aug. 2002.
- [5] P. A. Regalia, *Adaptive IIR filtering in signal processing and control*, Marcel Dekker, New York, 1995.