

# Implementação Eficiente do Método da Correlação Cruzada Normalizada para Detecção de *Double-Talk*

Rafael Dutra Demetri e Rui Seara

**Resumo**—Este artigo discute uma implementação do método da correlação cruzada normalizada (NCC) para detecção de *double-talk* em sistemas de cancelamento de eco. A abordagem proposta depende apenas do conhecimento dos sinais do *far-end* e do *near-end*, não levando em conta outras variáveis do sistema. A implementação considerada utiliza uma versão de baixa complexidade computacional do algoritmo RLS para efetuar uma parte importante dos cálculos, eliminando assim a necessidade de algumas usuais aproximações. Uma redução adicional de complexidade é obtida se a detecção de *double-talk* for realizada em uma frequência de amostragem menor do que a do sistema. Resultados de simulação mostram muito bom desempenho da abordagem proposta quando comparada com outras técnicas da literatura.

**Palavras-Chave**—Cancelamento de eco, *double-talk*, método da correlação cruzada normalizada.

**Abstract**—This paper discusses an implementation of the normalized cross-correlation method (NCC) for double-talk detection in echo cancellation systems. The proposed approach relies only on the far-end and near-end signals. A low complexity version of the RLS algorithm is used to perform most of the required calculations, thereby avoiding some commonly used approximations. Further computational complexity savings can be obtained by performing double-talk detection in a reduced sampling rate. Simulation results point out very good performance for the proposed approach as compared with other techniques.

**Keywords**—Echo cancellation, double-talk, normalized cross-correlation method.

## I. INTRODUÇÃO

Sistemas de cancelamento de eco são utilizados para remover ecos indesejáveis existentes em sistemas de telefonia. Os tipos de eco mais comuns nesses sistemas são o eco acústico, que ocorre quando existe um acoplamento acústico entre o microfone e o alto-falante de um sistema viva-voz, e o eco de linha, que ocorre na telefonia analógica [e em sistemas de voz sobre IP (VoIP)] devido a imperfeições nos acoplamentos dos circuitos de quatro para dois fios [1]. Nas situações em que o atraso entre o sinal de voz e o correspondente sinal de eco é grande (maior do que 20 ms), a existência de eco pode causar desconforto aos usuários e prejudicar a fluência de uma conversação telefônica. Nesses casos, pode-se melhorar o conforto dos usuários e a qualidade da conversação através da utilização de um cancelador de eco.

Rafael Dutra Demetri e Rui Seara, LINSE – Laboratório de Circuitos e Processamento de Sinais, Departamento de Engenharia Elétrica, Universidade Federal de Santa Catarina, Florianópolis, SC, E-mails: {rdemetri, seara}@linse.ufsc.br.

Este trabalho foi parcialmente financiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e pela empresa Dígito Tecnologia Ltda.

O diagrama de blocos de um sistema de cancelamento de eco típico é mostrado na Fig. 1. A notação utilizada é a seguinte:  $x(n)$  é o sinal de voz do usuário distante (*far-end*),  $h(n)$  representa a resposta ao impulso do caminho do eco,  $\hat{h}(n)$  denota a resposta ao impulso do filtro adaptativo (que é modelada geralmente por um filtro FIR),  $y(n)$  é a soma do sinal de eco  $d(n)$  com o sinal de voz do usuário próximo (*near-end*)  $v(n)$ ,  $\hat{d}(n)$  representa a réplica do sinal de eco e  $e(n)$ , o erro residual.

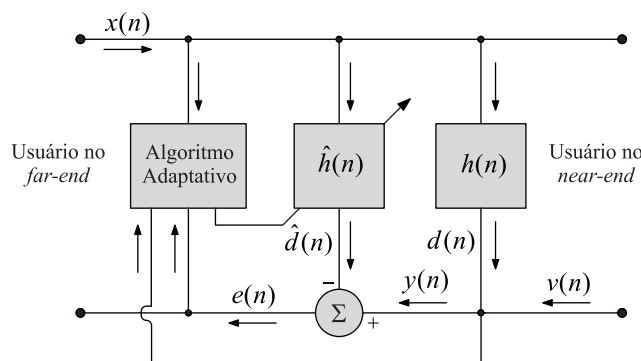


Fig. 1. Diagrama de blocos de um sistema de cancelamento de eco típico.

Os coeficientes do filtro  $\hat{h}(n)$  são adaptados buscando minimizar alguma função custo [que, em geral, depende do sinal de erro  $e(n)$ ]. Um dos algoritmos de filtragem adaptativa mais utilizados é o LMS (*least-mean-square*), devido principalmente à sua simplicidade, robustez e baixa complexidade computacional [2]. No LMS, a função custo considerada é o valor médio do erro quadrático. Na prática, utiliza-se geralmente variações do LMS, tais como o *normalized LMS* (NLMS) e o *proportional NLMS* (PNLMS) [3].

Uma outra família de algoritmos adaptativos é a baseada no método dos mínimos quadrados, no qual a função custo é obtida como uma soma ponderada de amostras passadas do erro quadrático. Aplicando um método recursivo para determinar os coeficientes que minimizam a função custo, chega-se ao algoritmo *recursive least-squares* (RLS) [2]. Comparado ao LMS, o RLS tem velocidade de convergência tipicamente uma ordem de grandeza maior e sua convergência não depende das características espectrais de  $x(n)$  [2]. Uma desvantagem do RLS em relação ao LMS é a sua elevada complexidade computacional. O RLS, em sua forma padrão, tem complexidade  $\mathcal{O}(N^2)$  [4] (onde  $N$  é a ordem do filtro adaptativo), fazendo com que implementações de filtros adaptativos de ordem elevada demandem grande poder computacional.

Independentemente da técnica utilizada para adaptação dos coeficientes, o desempenho de um cancelador de eco é afetado

se a detecção dos instantes de atividade de fala nos dois lados da linha (*double-talk*) não é adequadamente efetuada. Durante o *double-talk*, o algoritmo adaptativo deve interromper o processo de adaptação, evitando dessa forma uma degradação na estimativa do caminho do eco [1]. Devido à sua importância para sistemas de cancelamento de eco, diversos métodos para a detecção de *double-talk* vêm sendo propostos na literatura [5]-[9].

O método da correlação cruzada normalizada (NCC) [10] para detecção de *double-talk* tem como principal característica a facilidade de uso em um grande número de situações sem necessidade de maiores ajustes. Em sua formulação original, os cálculos envolvem apenas os sinais do *far-end* e *near-end*. A implementação direta desse método, no entanto, é geralmente evitada, pois envolve a inversão de uma matriz  $N \times N$ . Logo, na prática, busca-se utilizar versões simplificadas desse método [10], [11], que dependem também dos coeficientes do filtro adaptativo  $\hat{h}(n)$ .

Neste trabalho, é apresentado um método para a implementação direta do NCC que elimina a necessidade de inversão da matriz  $N \times N$ . A chave para esse método advém da constatação de que a parte de maior carga computacional do NCC pode ser efetuada através do algoritmo RLS. Dessa forma, pode-se utilizar uma das versões de baixa complexidade desse algoritmo, disponíveis na literatura [12]-[15], para implementar o NCC, evitando a dependência de  $\hat{h}(n)$ .

Este trabalho é organizado como segue. A Seção II apresenta as técnicas usuais para detecção de *double-talk*. A Seção III discute em maior detalhe o método da correlação cruzada normalizada. A Seção IV descreve a implementação do método da correlação cruzada normalizada baseada no algoritmo RLS. Na Seção V, resultados de simulação obtidos são mostrados e discutidos. Finalmente, a Seção VI apresenta as conclusões e comentários deste trabalho.

## II. ALGORITMOS PARA DETECÇÃO DE DOUBLE-TALK

Existem diferentes métodos para detecção de *double-talk* propostos na literatura. Em geral, esses métodos operam como descrito a seguir:

- i) Uma variável de decisão  $\xi(n)$  é calculada, utilizando os sinais disponíveis  $x(n)$ ,  $y(n)$ ,  $\epsilon(n)$  e  $\hat{d}(n)$ , e os coeficientes da resposta ao impulso  $\hat{h}(n)$ .
- ii) A variável  $\xi(n)$  é comparada a um limiar pré-definido  $T$ , e *double-talk* é assumido quando

$$\xi(n) < T. \quad (1)$$

Algumas abordagens para detecção de *double-talk* podem falhar sob certas condições na prática. Em particular, para algumas técnicas, não se pode dispor de um limiar  $T$  que seja independente das características do sistema (como, por exemplo, potência dos sinais envolvidos ou atenuação da resposta ao impulso). Dessa forma, o valor de  $T$  deve ser ajustado para cada caso, o que pode ser indesejável em certas aplicações.

Uma outra condição que pode levar à falha é a dependência do algoritmo de detecção de *double-talk* do valor dos coeficientes do filtro adaptativo  $\hat{h}(n)$ . Quando essa dependência existe, alguns fatores devem ser levados em conta, a saber:

- Uma falha na detecção de *double-talk* pode fazer o algoritmo adaptativo divergir, o que por consequência pode impactar a detecção de *double-talk* nos instantes seguintes.
- Uma mudança abrupta na resposta ao impulso do sistema pode ser interpretada, de forma errada, como sendo *double-talk*, impedindo a adaptação do filtro à nova condição.
- Na inicialização do cancelador de eco [ou seja, quando  $\hat{h}(n) = 0$ ], deve existir uma estratégia que permita a adaptação do filtro antes do início da detecção de *double-talk*.

Algumas técnicas conhecidas para detecção de *double-talk* são brevemente discutidas na sequência.

### A. Algoritmo Geigel

O algoritmo *Geigel* [5] é um algoritmo bastante simples para detecção de *double-talk*. Nesse algoritmo, a variável de detecção é

$$\xi(n)_G = \frac{\|\mathbf{x}(n)\|_\infty}{|y(n)|} \quad (2)$$

onde  $\mathbf{x}(n) = [x(n) \ x(n-1) \ \dots \ x(n-N+1)]^T$  é o vetor de amostras do *far-end* e  $\|\cdot\|_\infty$  caracteriza a norma infinita de um vetor. Esse esquema de detecção é baseado na comparação de nível entre os sinais  $y(n)$  e  $x(n)$ . São consideradas as últimas  $N$  amostras de  $x(n)$  visando compensar o atraso do caminho do eco. O limiar  $T$  é escolhido de forma a compensar um valor específico de atenuação do caminho do eco. Uma desvantagem do algoritmo *Geigel* é que, em muitas situações, não é possível conhecer *a priori* esse valor de atenuação.

### B. Método da Correlação Cruzada

Um outro método de detecção de *double-talk* utiliza a correlação entre o vetor de amostras do *far-end*  $\mathbf{x}(n)$  e o sinal do *near-end*  $y(n)$  para detecção de *double-talk* [8]. A medida de correlação considerada é o vetor de coeficientes de correlação cruzada, dado por

$$\mathbf{c}_{xy} = \frac{E[\mathbf{x}(n)y(n)]}{\sqrt{E[x(n)^2]E[y(n)^2]}} \quad (3)$$

$$= \frac{\mathbf{r}_{xy}}{\sigma_x \sigma_y} \quad (4)$$

onde  $E(\cdot)$  denota o valor esperado,  $\mathbf{r}_{xy} = E[\mathbf{x}(n)y(n)]$  é o vetor de correlação cruzada entre  $\mathbf{x}(n)$  e  $y(n)$  e  $\sigma_x$  e  $\sigma_y$  são os desvios-padrão de  $x(n)$  e  $y(n)$ , respectivamente. O coeficiente de correlação cruzada aproxima-se de 1 quando a correlação é máxima, e de 0 quando é mínima. Dessa forma, a variável de detecção é dada por

$$\xi_{xy} = \|\mathbf{c}_{xy}\|_\infty. \quad (5)$$

A principal desvantagem do método da correlação cruzada é que o valor ótimo do limiar  $T$  pode ter grande variação dependendo das condições de operação [10].

### III. MÉTODO DA CORRELAÇÃO CRUZADA NORMALIZADA

O método da correlação cruzada normalizada para detecção de *double-talk* é baseado no método da correlação entre  $\mathbf{x}(n)$  e  $y(n)$ . A sua principal vantagem é obter um valor constante para  $\xi(n)$  quando  $v(n) = 0$ , independente de quaisquer características do sistema. A seguir, é apresentada a derivação desse algoritmo e as abordagens usuais para a sua implementação.

Da Fig. 1, observa-se que o sinal  $y(n)$  recebido pelo cancelador de eco é

$$y(n) = d(n) + v(n). \quad (6)$$

Dessa forma, a variância do sinal  $y(n)$  é dada por

$$\sigma_y^2 = \sigma_d^2 + \sigma_v^2. \quad (7)$$

Já a variância do sinal  $d(n)$  pode ser escrita como

$$\sigma_d^2 = \mathbf{h}^T \mathbf{R}_{\mathbf{xx}} \mathbf{h} \quad (8)$$

onde  $\mathbf{h} = [h(0) \ h(1) \ \dots \ h(N-1)]^T$  é a resposta ao impulso do sistema e  $\mathbf{R}_{\mathbf{xx}} = E[\mathbf{x}(n)\mathbf{x}^T(n)]$  é a matriz de autocorrelação do sinal de entrada. Dado que  $d(n) = \mathbf{h}^T \mathbf{x}(n)$ , tem-se

$$\mathbf{h} = \mathbf{R}_{\mathbf{xx}}^{-1} \mathbf{r}_{\mathbf{x}d} \quad (9)$$

onde  $\mathbf{r}_{\mathbf{x}d} = E[\mathbf{x}(n)d(n)]$  é o vetor de correlação cruzada entre  $\mathbf{x}(n)$  e  $d(n)$ . Assim, pode-se reescrever (8) como

$$\sigma_d^2 = \mathbf{r}_{\mathbf{x}d}^T \mathbf{R}_{\mathbf{xx}}^{-1} \mathbf{r}_{\mathbf{x}d}. \quad (10)$$

Então, assumindo que o sinal de entrada  $x(n)$  e o sinal de voz do *near-end*  $v(n)$  são não-correlacionados, obtém-se

$$\begin{aligned} \mathbf{r}_{\mathbf{x}y} &= E[\mathbf{x}(n)y(n)] = E\{\mathbf{x}(n)[d(n) + v(n)]\} \\ &= E[\mathbf{x}(n)d(n)] \\ &= \mathbf{r}_{\mathbf{x}d}. \end{aligned} \quad (11)$$

Assim,  $\sigma_d^2$  pode ser expresso como

$$\sigma_d^2 = \mathbf{r}_{\mathbf{x}y}^T \mathbf{R}_{\mathbf{xx}}^{-1} \mathbf{r}_{\mathbf{x}y}. \quad (12)$$

A variável de detecção do NCC é obtida calculando-se a razão entre  $\sigma_d^2$  e  $\sigma_y^2$ . Portanto,

$$\xi_{\text{NCC}} = \frac{\sigma_d^2}{\sigma_y^2} = \frac{\mathbf{r}_{\mathbf{x}y}^T \mathbf{R}_{\mathbf{xx}}^{-1} \mathbf{r}_{\mathbf{x}y}}{\sigma_y^2}. \quad (13)$$

O cálculo de  $\xi_{\text{NCC}}$ , de acordo com (13), depende apenas dos sinais  $x(n)$  e  $y(n)$ , que são os sinais disponíveis para o cancelador de eco.

Agora, reescrevendo (13) como

$$\xi_{\text{NCC}} = \frac{\sigma_d^2}{\sigma_y^2} = \frac{\sigma_d^2}{\sigma_d^2 + \sigma_v^2} \quad (14)$$

observa-se que, quando existe *double-talk* [ $v(n) \neq 0$ ], o denominador de (14) torna-se maior do que seu numerador. Dessa forma, a detecção de *double-talk* pode ser realizada considerando  $T = 1$  em (1). Contudo, costuma-se na prática utilizar um valor  $T < 1$ , visando compensar possíveis erros de estimação das grandezas estatísticas e a presença de ruído.

As grandezas estatísticas requeridas para o cálculo de  $\xi_{\text{NCC}}$  podem ser estimadas recursivamente como segue:

$$\mathbf{r}_{\mathbf{x}y}(n) = \mathbf{x}(n)y(n) + \lambda \mathbf{r}_{\mathbf{x}y}(n-1) \quad (15)$$

$$\sigma_y^2(n) = y^2(n) + \lambda \sigma_y^2(n-1) \quad (16)$$

$$\mathbf{R}_{\mathbf{xx}}(n) = \mathbf{x}(n)\mathbf{x}^T(n) + \lambda \mathbf{R}_{\mathbf{xx}}(n-1) \quad (17)$$

onde  $0 < \lambda \leq 1$  é o fator de esquecimento. Tal procedimento é adequado do ponto de vista prático, pois pode ser implementado sem excessiva carga computacional. Dessa forma, a variável de detecção pode ser obtida como

$$\xi_{\text{NCC}}(n) = \frac{\mathbf{r}_{\mathbf{x}y}(n)^T \mathbf{R}_{\mathbf{xx}}^{-1}(n) \mathbf{r}_{\mathbf{x}y}(n)}{\sigma_y^2(n)}. \quad (18)$$

Visando reduzir a complexidade computacional da implementação, pode-se evitar o cálculo de  $\mathbf{R}_{\mathbf{xx}}^{-1}(n)$  assumindo que o filtro adaptativo tenha convergido. Nesse caso, tem-se

$$\mathbf{h} = \mathbf{R}_{\mathbf{xx}}^{-1} \mathbf{r}_{\mathbf{x}y} \approx \hat{\mathbf{h}}(n) \quad (19)$$

onde  $\hat{\mathbf{h}}(n) = [\hat{h}_0(n) \ \hat{h}_1(n) \ \dots \ \hat{h}_{N-1}(n)]^T$  é o vetor de coeficientes do filtro adaptativo. Assim, é possível aproximar (18) por

$$\xi_{\text{NCC}}(n) \approx \frac{\mathbf{r}_{\mathbf{x}y}^T(n) \hat{\mathbf{h}}(n)}{\sigma_y^2(n)}. \quad (20)$$

Assumindo ainda a convergência do filtro adaptativo, uma outra estratégia para simplificar a implementação do NCC é utilizar a réplica do eco  $\hat{d}(n)$  no cálculo da variância do sinal  $d(n)$ . Assim,

$$\sigma_d^2(n) \approx \hat{\sigma}_d^2(n) = \hat{d}^2(n) + \lambda \sigma_d^2(n-1). \quad (21)$$

Então, substituindo (21) em (18), obtém-se a aproximação

$$\xi_{\text{NCC}}(n) \approx \frac{\sigma_d^2(n)}{\sigma_y^2(n)}. \quad (22)$$

O cálculo aproximado de  $\xi_{\text{NCC}}(n)$  através de (20) ou (22) é sedutor do ponto de vista de complexidade computacional. No entanto, em ambos os casos, a qualidade da aproximação (e, portanto, o desempenho do detector de *double-talk*) depende de a convergência dos coeficientes do filtro adaptativo ter sido alcançada.

### IV. IMPLEMENTAÇÃO DO NCC BASEADA NO RLS

É possível eliminar a dependência entre detecção de *double-talk* e convergência do filtro adaptativo [existente em (20) e (22)] implementando diretamente (18). Um possível caminho para tal implementação é obter uma estimativa de  $\mathbf{R}_{\mathbf{xx}}(n)$  a partir de (17) e determinar a sua inversa para todo  $n$ . No entanto, a inversão de uma matriz  $N \times N$  acarreta quase sempre uma carga computacional alta. A mesma matriz  $\mathbf{R}_{\mathbf{xx}}^{-1}(n)$  pode também ser obtida de forma recursiva, utilizando para tal o ganho de *Kalman*  $\mathbf{k}(n)$  [2]. Portanto,

$$\mathbf{R}_{\mathbf{xx}}^{-1}(n) = \frac{1}{\lambda} [\mathbf{I} - \mathbf{k}(n)\mathbf{x}^T(n)] \mathbf{R}_{\mathbf{xx}}^{-1}(n-1). \quad (23)$$

Ainda assim, a implementação de (23) e (13) implica uma complexidade  $\mathcal{O}(N^2)$  [sem contar o cálculo de  $\mathbf{k}(n)$ ], que pode ser proibitivamente elevada para filtros de ordem elevada.

Uma alternativa mais interessante para determinar  $\xi_{\text{NCC}}(n)$  é levar em conta que o vetor de coeficientes  $\hat{\mathbf{h}}_{\text{RLS}}(n)$ , obtido pelo algoritmo RLS, pode ser escrito como [2]

$$\hat{\mathbf{h}}_{\text{RLS}}(n) = \mathbf{R}_{\mathbf{xx}}^{-1}(n) \mathbf{r}_{\mathbf{x}y}(n) \quad (24)$$

com  $\mathbf{r}_{xy}(n)$  e  $\mathbf{R}_{xx}(n)$  dados, respectivamente, por (15) e (17). Dessa forma, é possível utilizar o algoritmo RLS como um procedimento eficiente para resolver (24). A variável  $\xi_{\text{NCC}}(n)$  pode então ser obtida por

$$\xi_{\text{NCC}}(n) = \frac{\mathbf{r}_{xy}^T(n) \hat{\mathbf{h}}_{\text{RLS}}(n)}{\sigma_y^2(n)}. \quad (25)$$

Aqui, os coeficientes determinados pelo algoritmo RLS não devem ser interpretados como aqueles de um filtro adaptativo em uma aplicação convencional de cancelamento de eco. Tais coeficientes são adaptados inclusive durante os instantes de *double-talk* e, dessa forma, não há garantia de que  $\hat{\mathbf{h}}_{\text{RLS}}(n)$  estime adequadamente o caminho do eco.

O algoritmo RLS, em sua forma padrão, calcula explicitamente  $\mathbf{R}_{xx}^{-1}(n)$  de acordo com (23), tendo complexidade  $\mathcal{O}(N^2)$ . No entanto, existem diversas estratégias de implementação do RLS, com diferentes características, que apresentam menor complexidade. Algumas abordagens específicas tiram proveito da estrutura transversal do filtro adaptativo, resultando em uma complexidade  $\mathcal{O}(N)$ . Dentre essas, a de menor complexidade é o *fast transversal filter* (FTF) [12], que tem complexidade  $7N$ . Entretanto, o FTF é reconhecido ser numericamente instável quando implementado com precisão finita devido à acumulação de erros de arredondamento [13]. Uma modificação do FTF, denominada *stabilized FTF* (SFTF) [13], visa melhorar suas características numéricas, com um pequeno aumento de complexidade ( $8N$ ). Tanto o algoritmo FTF quanto o SFTF têm em sua estrutura interna preditores lineares *backward* e *forward*, de ordem  $N$ . No caso do SFTF, a atualização dos preditores demanda uma complexidade  $6N$ , enquanto a atualização de  $\hat{\mathbf{h}}_{\text{RLS}}(n)$ ,  $2N$ .

Em determinados casos, a maior parte da informação dos sinais envolvidos pode ser recuperada utilizando preditores de ordem  $N_p \leq N$ . Baseando-se nessa idéia, foi proposto o algoritmo *fast Newton transversal filter* (FNTF) [14]. No FNTF, o vetor  $\hat{\mathbf{h}}_{\text{RLS}}(n)$  de ordem  $N$  é obtido através da extrapolação de preditores de ordem  $N_p$ . Essa extrapolação tem complexidade computacional  $2N$ , que é o limite inferior da complexidade do FNTF. Para  $N_p = N$ , o FNTF obtém a mesma solução do algoritmo RLS; já para  $N_p = 0$ , as características de convergência são similares às do LMS.

No decorrer deste trabalho, é considerado o uso do FNTF para o cálculo de  $\hat{\mathbf{h}}_{\text{RLS}}(n)$ . A atualização dos preditores é realizada de maneira idêntica à do SFTF (com complexidade  $6N_p$ ), resultando assim em uma complexidade computacional total de  $6N_p + 2N$ . Dessa forma, o algoritmo proposto para detecção de *double-talk* pode ser implementado seguindo os passos:

- i) Cálculo de  $\hat{\mathbf{h}}_{\text{RLS}}(n)$  utilizando o FNTF.
- ii) Atualização de  $\mathbf{r}_{xy}(n)$  e  $\sigma_y^2(n)$  usando (15) e (16).
- iii) Cálculo de  $\xi_{\text{NCC}}(n)$  através de (25).

Na Tabela I, é mostrada a complexidade computacional de cada uma das etapas da implementação do NCC proposta. Observa-se que a ordem  $N_p$  dos preditores tem grande impacto na complexidade computacional total do método.

Além do mais, é possível reduzir a complexidade do NCC se forem consideradas algumas suposições sobre os

TABELA I  
COMPLEXIDADE COMPUTACIONAL DA IMPLEMENTAÇÃO PROPOSTA

Etapa	Complexidade
Cálculo de $\hat{\mathbf{h}}_{\text{RLS}}(n)$	$6N_p + 2N$
Atualização recursiva de $\mathbf{r}_{xy}(n)$ e $\sigma_y^2(n)$	$2N$
Cálculo de $\xi_{\text{NCC}}(n)$	$N$
TOTAL	$6N_p + 5N$

sinais envolvidos. Em particular, considerando que sinais de voz têm a maior parte de sua energia concentrada em baixas frequências, pode-se (em determinados casos) realizar a detecção de *double-talk* usando uma taxa de amostragem menor do que a taxa do sistema, sem grande prejuízo de desempenho. Assim, a dimensão  $N$  dos vetores utilizados na detecção de *double-talk* reduz-se para

$$N_r = \frac{N}{L} \quad (26)$$

onde  $L$  caracteriza o fator de redução da taxa de amostragem. Considerando ainda que o número total de operações é também reduzido por um fator  $L$  (devido à redução na taxa de amostragem), a complexidade total do método torna-se

$$C_r = \frac{6N_p + 5N_r}{L} = \frac{6N_p}{L} + \frac{5N}{L^2} \quad (27)$$

## V. RESULTADOS DE SIMULAÇÃO

Nesta seção, são apresentados resultados experimentais visando avaliar o desempenho das técnicas apresentadas para detecção de *double-talk*. De maneira geral, o desempenho de um detector de *double-talk* pode ser caracterizado pelas seguintes grandezas:

- *Probabilidade de falso alarme*  $P_f$ . Probabilidade de falsa detecção de *double-talk*, ou seja, de *double-talk* ser declarado quando não há atividade do usuário do *near-end* [ $v(n) = 0$ ].
- *Probabilidade de detecção*  $P_d$ . Probabilidade da correta detecção de *double-talk*, ou seja, de *double-talk* ser declarado quando existe atividade do usuário do *near-end* [ $v(n) \neq 0$ ].
- *Probabilidade de perda*  $P_m = 1 - P_d$ . Probabilidade de *double-talk* não ser detectado quando existe atividade do usuário do *near-end* [ $v(n) \neq 0$ ].

Tipicamente, existe uma relação de compromisso entre a baixa probabilidade de falso alarme  $P_f$  e a baixa probabilidade de perda  $P_m$ , que é controlada pelo valor de  $T$  em (1). Em sistemas de cancelamento de eco, a penalidade associada às perdas é elevada quando comparada às dos falsos alarmes, já que essas situações podem acarretar divergência do filtro adaptativo. Já os falsos alarmes causam, no máximo, uma redução na velocidade de convergência do filtro.

A maneira escolhida aqui para avaliar o desempenho das técnicas de detecção de *double-talk* é através do valor de  $P_m$  obtido para um dado valor de  $P_f$ . Nos experimentos realizados, a probabilidade de falso alarme  $P_f$  é medida como

$$P_f = \frac{N_{x,DT}}{N_x} \quad (28)$$

onde  $N_x$  é o número de amostras em que há somente atividade do usuário do *far-end* [ $x(n) \neq 0$  e  $v(n) = 0$ ] e  $N_{x,DT}$  é o número de amostras desse conjunto nas quais é declarado *double-talk*. Já a probabilidade de perda  $P_m$  é medida como

$$P_m = 1 - \frac{N_{x,v,DT}}{N_{x,v}} \quad (29)$$

onde  $N_{x,v}$  é o número de amostras em que há atividade dos usuários do *far-end* e do *near-end* [ $x(n) \neq 0$  e  $v(n) \neq 0$ ] e  $N_{x,v,DT}$  é o número de amostras desse conjunto nas quais é declarado *double-talk*. Nos experimentos realizados, o valor de  $T$  é ajustado de modo a obter  $P_f = 0,1$ .

Quando se compara diferentes técnicas para detecção de *double-talk*, é interessante medir  $P_m$  para diferentes valores da razão *near-end* para *far-end* (NFR), a qual caracteriza uma medida da energia do sinal do *near-end* em relação à do *far-end*. Assim,

$$\text{NFR} = 10 \log_{10} \frac{\sigma_v^2}{\sigma_y^2} \quad (30)$$

Os arquivos de voz correspondentes ao *far-end* e *near-end* foram extraídos de uma base de dados de sinais de voz, amostrados em 8 kHz. A duração do sinal do *far-end* é de aproximadamente 16 s e a do sinal de *near-end*, de aproximadamente 11 s. Para obter os diferentes valores de NFR, é aplicado um ganho correspondente ao sinal do *near-end*. O vetor dos coeficientes da resposta ao impulso  $h$  utilizado nos experimentos corresponde às primeiras  $N_h = 256$  amostras de uma resposta ao impulso  $h(n)$  medida (através de um processo de identificação de sistemas) em uma situação real de eco de linha.

O desempenho do método da correlação cruzada normalizada depende do fator de esquecimento  $\lambda$  utilizado. Na Fig. 2, é mostrada a curva de probabilidade de perda  $P_m$  obtida para diversos valores de  $\lambda$ , com  $\hat{h}_{\text{RLS}}(n)$  calculado através do algoritmo RLS convencional, considerando  $N = 256$ ,  $P_f = 0,1$  e  $\text{NFR} = -5$  dB. Observa-se que os melhores resultados são obtidos para valores de  $\lambda$  inferiores a 0,996.

No entanto, o algoritmo SFTF (e, conseqüentemente, o FNTF) pode apresentar instabilidade numérica caso o valor de  $\lambda$  seja muito pequeno. Em [13], é mostrado que um comportamento numérico satisfatório é obtido se

$$1 \geq \lambda \geq \lambda_{\text{MIN}} \quad (31)$$

com

$$\lambda_{\text{MIN}} = 1 - \frac{0,4}{N} \quad (32)$$

Para  $N = 256$ , tem-se  $\lambda_{\text{MIN}} \approx 0,9985$ . No entanto, observa-se pela Fig. 2 que a faixa de valores para  $\lambda$  definida em (31) não corresponde à de melhor desempenho. A abordagem aqui escolhida para obter estabilidade numérica sem comprometer o desempenho é utilizar, para o cálculo de  $\hat{h}_{\text{RLS}}(n)$ , um fator de esquecimento, designado como  $\lambda_{\text{RLS}}$ , diferente daquele denotado por  $\lambda$ , usado em (15) e (16). Os valores considerados nos experimentos são  $\lambda_{\text{RLS}} = 0,9999$  e  $\lambda = 0,995$ .

Um outro parâmetro que influencia o desempenho do método da correlação cruzada normalizada baseado no FNTF são as ordens  $N_p$  dos preditores. Na Fig. 3, é mostrada a curva

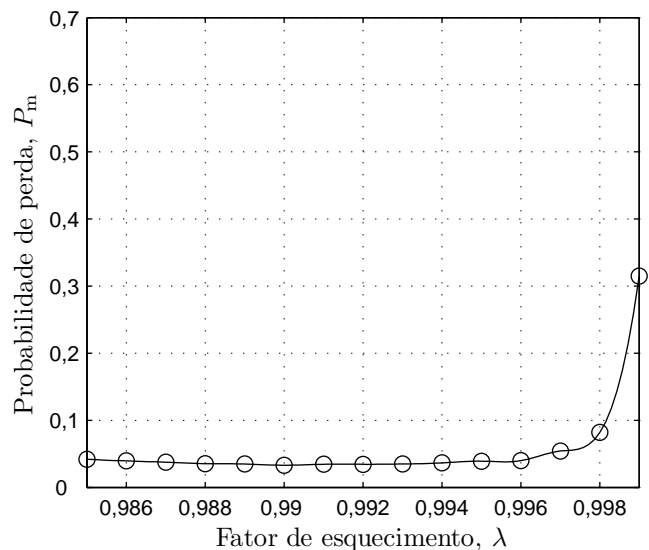


Fig. 2. Influência do fator de esquecimento  $\lambda$ , para NFR de  $-5$  dB e probabilidade de falso alarme  $P_f$  de 0,1.

da probabilidade de perda  $P_m$  em função de  $N_p$ . Observa-se que a redução da ordem dos preditores não prejudica substancialmente o desempenho do método. Mais ainda, o melhor caso corresponde a  $N_p = 0$ . Esse valor de  $N_p$  é utilizado nos experimentos que seguem.

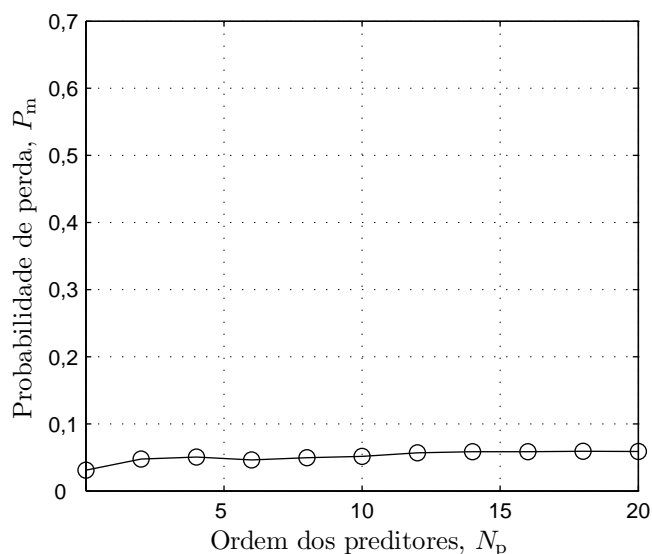


Fig. 3. Influência da ordem dos preditores  $N_p$ , considerando NFR de  $-5$  dB e  $P_f$  de 0,1.

Na Fig. 4, é mostrada uma comparação entre os desempenhos dos métodos da correlação cruzada normalizada [utilizando tanto o SFTF quanto o FNTF para o cálculo de  $\hat{h}_{\text{RLS}}(n)$ ], da correlação cruzada e do algoritmo de Geigel. São considerados valores de NFR entre  $-20$  dB e  $10$  dB, que correspondem a valores usualmente considerados em aplicações práticas. Já o limiar  $T$  é ajustado para  $P_f = 0,1$ . No método da correlação cruzada, as grandezas estatísticas  $r_{xy}$ ,  $\sigma_x$  e  $\sigma_y$  são estimadas de maneira idêntica às do NCC.

Observa-se que o método da correlação cruzada normalizada apresenta o melhor desempenho, para todos os valores de NFR considerados. Constata-se também que a utilização do FNTF (com  $N_p = 0$ ) ao invés do SFTF para o cálculo de  $\hat{\mathbf{h}}_{\text{RLS}}(n)$  não afeta os resultados do algoritmo. O desempenho do método da correlação cruzada convencional aproxima-se do obtido pelo NCC apenas para valores elevados de NFR (aproximadamente, maior do que  $-8$  dB) e o do algoritmo de Geigel é, de maneira geral, o pior dos três.

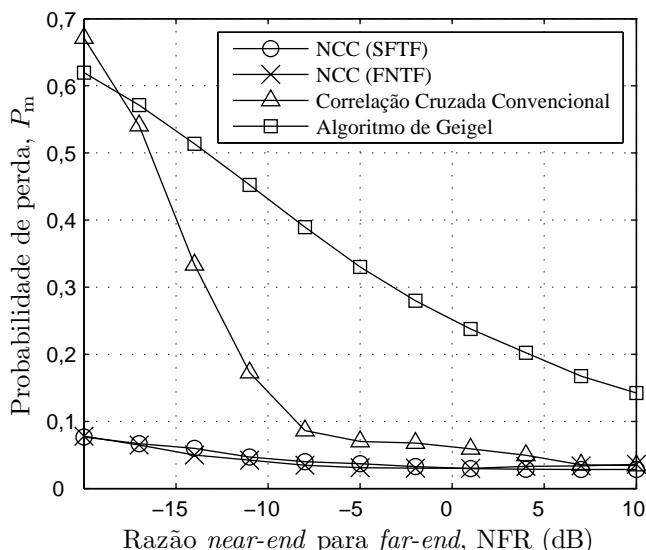


Fig. 4. Comparação dos valores de  $P_m$  obtidos por diferentes técnicas, considerando uma probabilidade de falso alarme  $P_f$  de 0,1.

Na Fig. 5, é mostrada uma comparação do desempenho do método da correlação cruzada normalizada baseado no FNTF com diferentes fatores de redução da taxa de amostragem do sistema. Observa-se que até a frequência de amostragem  $f_s = 2$  kHz (que corresponde ao fator de redução  $L = 4$ ), não há qualquer significativa perda de desempenho.

## VI. CONCLUSÕES

Neste trabalho, é apresentada uma implementação do método da correlação cruzada normalizada para detecção de *double-talk*. A principal característica desse método é que o limiar de detecção não depende das características dos sinais ou do sistema. Na implementação proposta, não existe também dependência dos coeficientes do filtro de cancelamento de eco. A parcela mais significativa dos cálculos necessários pode ser realizada através do algoritmo RLS. Assim, utilizando o algoritmo FNTF (uma versão de baixa complexidade do RLS), com preditores de ordem zero, o NCC pode ser implementado com complexidade de  $5N$ . Pode-se adicionalmente reduzir a complexidade computacional se a detecção de *double-talk* for realizada usando uma taxa de amostragem inferior à do sistema. Resultados de simulação mostraram que a implementação proposta apresenta um desempenho superior quando comparado ao de outros métodos da literatura. Além do mais, observa-se que a detecção de *double-talk* pode ser realizada usando uma taxa de amostragem tão baixa quanto

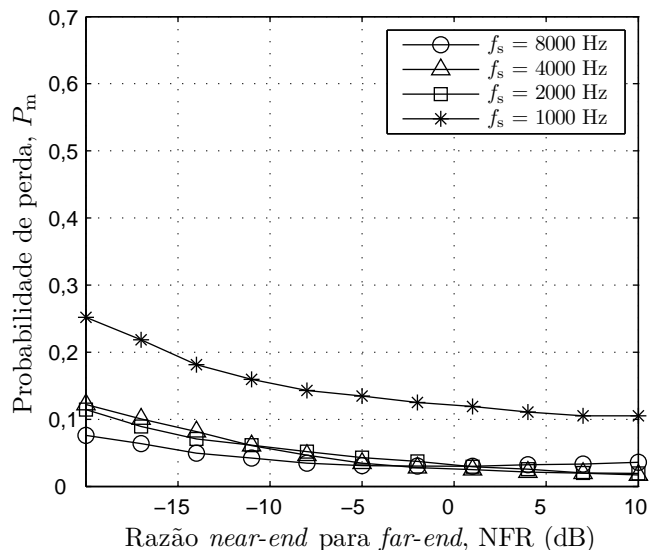


Fig. 5. Probabilidade de perda  $P_m$  para diferentes frequências de amostragem.

$f_s = 2$  kHz, sem significativo prejuízo ao desempenho. Nessa situação, a complexidade da abordagem proposta pode ser tão baixa quanto  $0,3N$ .

## REFERÊNCIAS

- [1] L. S. Gay and J. Benesty, *Acoustic Signal Processing for Telecommunication*, 1st ed. Norwell, USA: Kluwer Academic Publishers, 2000.
- [2] S. Haykin, *Adaptive filter theory*, 3rd ed. New Jersey, USA: Prentice Hall, 2000.
- [3] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 5, pp. 508–518, Sept. 2000.
- [4] B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications*, 1st ed. New York, USA: John Wiley & Sons, 2000.
- [5] D. L. Duttweiler, "A twelve-channel digital echo canceler," *IEEE Trans. Commun.*, vol. 26, no. 5, pp. 647–653, May 1978.
- [6] H. Ye and B.-X. Wu, "A new double-talk detection algorithm based on the orthogonality theorem," *IEEE Trans. Commun.*, vol. 39, no. 11, pp. 1542–1545, Nov. 1991.
- [7] T. Gansler, M. Hansson, C.-J. Ivarsson, and G. Salomonsson, "A double-talk detector based on coherence," *IEEE Trans. Commun.*, vol. 44, no. 11, pp. 1421–1427, Nov. 1996.
- [8] J. H. Cho, D. R. Morgan, and J. Benesty, "An objective technique for evaluating doubletalk detectors in acoustic echo cancelers," *IEEE Trans. Speech Audio Process.*, vol. 7, pp. 718–724, Nov. 1999.
- [9] K. Ochiai, T. Araseki, and T. Ogihara, "Echo canceler with two echo path models," *IEEE Trans. Commun.*, vol. 25, no. 6, pp. 589–595, Jun 1977.
- [10] J. Benesty, D. R. Morgan, and J. H. Cho, "A new class of doubletalk detectors based on cross-correlation," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 2, pp. 168–172, Mar. 2000.
- [11] M. Waseem, M. Omer, A. Bhatti, and S. Shah, "A new doubletalk detector with fuzzy decision thresholding," in *Proc. IEEE Students Conf.*, Lahore, Pakistan, Aug. 2002, vol.1, pp. 117–121.
- [12] J. Cioffi and T. Kailath, "Fast recursive least squares transversal filters for adaptive filtering," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 2, pp. 304–337, Apr. 1984.
- [13] D. Slock and T. Kailath, "Numerically stable fast transversal filters for recursive least squares adaptive filtering," *IEEE Trans. Signal Process.*, vol. 39, no. 1, pp. 92–114, Jan. 1991.
- [14] G. Moustakides and S. Theodoridis, "Fast newton transversal filters—a new class of adaptive estimation algorithms," *IEEE Trans. Signal Process.*, vol. 39, no. 10, pp. 2184–2193, Oct. 1991.
- [15] K. Maoche and D. Slock, "Fast subsampled-updating stabilized fast transversal filter (FSU SFTF) RLS algorithm for adaptive filtering," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2248–2257, Aug. 2000.