

Cancelamento de eco acústico não-linear com o uso de um modelo físico do alto-falante

Roberto de O. Possidente e Phillip M. S. Burt

Resumo— Apresenta-se um novo método de cancelamento de eco acústico não-linear, no qual introduz-se antes do filtro adaptativo um bloco não-linear fixo, derivado de um modelo físico do alto-falante. A complexidade computacional de diferentes alternativas de implementação de tal bloco não-linear é considerada e o melhor compromisso entre desempenho e complexidade computacional é levantado. Aponta-se também que o filtro adaptativo, ainda que puramente linear, compensa imperfeições do modelo não-linear do alto-falante, o que tem um papel importante no cancelamento de eco alcançado.

Palavras-Chave— Cancelamento de eco acústico não-linear, filtros adaptativos, processamento digital de sinais

Abstract— A new non-linear acoustic echo cancellation method is presented, in which a fixed non-linear block, derived from a physical model of the speaker, is introduced before the adaptive filter. The computational complexity of different implementation alternatives of such non-linear block is considered and the best trade-off between performance and computational complexity is assessed. It is also pointed out that the adaptive filter, although purely linear, compensates for imperfections of the non-linear model of the speaker, which has an important role in the achieved echo cancellation.

Keywords— Nonlinear acoustic echo cancellation, adaptive filters, digital signal processing

I. INTRODUÇÃO

Na abordagem usual do cancelamento do eco acústico em sistemas de viva-voz, o conjunto formado pelo amplificador, alto-falante, ambiente reverberante e microfone é modelado por um sistema linear [1].

Estendendo esta abordagem, trabalhos como [2], [3] e [4] analisam os canceladores de eco lineares (LAEC) levando em conta a não-linearidade do alto-falante: em [2] e [4] discutem-se as limitações dos LAEC nestas condições e em [3] avalia-se a robustez de alguns algoritmos para filtros adaptativos lineares.

Trabalhos como [5] e [6], por sua vez, consideram o uso de filtros adaptativos não-lineares. Em [5], um filtro de Volterra adaptativo de segunda ordem antecede um filtro adaptativo linear. Em [6], o caminho de eco acústico é integralmente modelado por um filtro de Volterra de segunda ordem.

Ainda que o bloco não-linear em cancelamento de eco acústico não-linear (NLAEC) seja adaptativo, as características estruturais adotadas para ele (como ordem e extensão de memória) são influenciadas pelo modelo não-linear adotado para o alto-falante. Frequentemente, porém, são adotados modelos particulares sem justificativa satisfatória. Visando sanar

Roberto de O. Possidente, e-mail: roberto.possidente@usp.br; Phillip M. S. Burt, Dept. de Eng. de Telecomunicações e Controle - EPUSP, e-mail: phillip@lcs.poli.usp.br.

esta deficiência, no presente trabalho propõe-se o uso de um modelo derivado de considerações físicas, como por exemplo [7], usado em [8] e [9] para calcular núcleos de Volterra do dispositivo. No sistema proposto aqui, tal modelo também é incorporado, de forma aproximada, na cadeia de geração da estimativa do eco acústico, levando-se em conta o melhor compromisso entre precisão e carga computacional de diferentes estruturas não-lineares possíveis. Aponta-se ainda a ação de compensação pelo filtro adaptativo linear de imprecisões do modelo não-linear.

II. SISTEMA PROPOSTO E ESTRUTURAS NÃO-LINEARES

O sistema de NLAEC proposto faz uso de um modelo físico do alto-falante, o qual supõe-se que o fabricante do equipamento de viva-voz possui. O sistema está na Figura 1, onde PNL é um bloco de pré-processamento (isto é, anterior ao filtro adaptativo) não-linear fixo derivado do modelo do alto-falante, W é um filtro adaptativo linear, H é a resposta impulsiva do ambiente, $u(n)$ é o sinal remoto a ser reproduzido no alto-falante e $r(n)$ é o sinal local capturado pelo microfone.

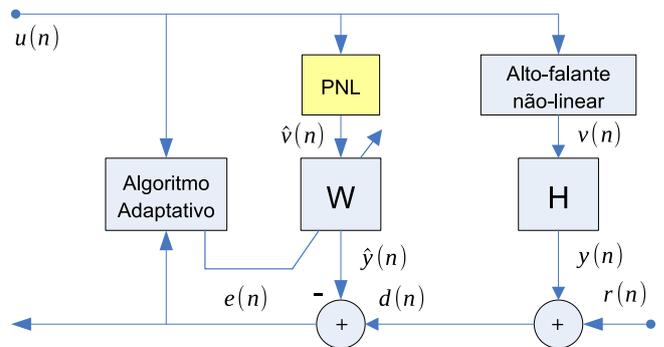


Fig. 1. NLAEC com pré-processamento não-linear

Consideramos a seguir diferentes formas de implementar o bloco PNL e a complexidade computacional de cada uma.

A. Filtro de Volterra convencional

Com os sinais definidos na Figura 1, a ação de um filtro de Volterra (VF) de ordem P e extensões de memória N_p , $p = 1, \dots, P$ é dada por

$$\hat{v}(n) = \sum_{p=1}^P \sum_{k_1=0}^{N_p-1} \dots \sum_{k_p=0}^{N_p-1} h_p(\mathbf{k}_p) u(n-k_1) \dots u(n-k_p) \quad (1)$$

onde $\mathbf{k}_p \triangleq (k_1, \dots, k_p)$ e $h_p(\mathbf{k}_p)$ é o núcleo de Volterra de ordem p utilizado. Com núcleos triangulares, isto é, $h_p(\mathbf{k}_p) =$

0 se $k_i > k_j$ para algum $i < j$, minimiza-se o número de operações aritméticas para o cômputo da saída [10]. O núcleo triangular de ordem p tem $\binom{N_p+p-1}{p}$ coeficientes e cada um requer p multiplicações, de modo que o número total de multiplicações por amostra é

$$N_{\text{mult}} = \sum_{p=1}^P \binom{N_p+p-1}{p} p \quad (2)$$

Deve-se notar que este número cresce rapidamente com a ordem p e as extensões de memória N_p .

B. Filtro de Volterra esparso interpolado

Admitindo-se certa redundância entre coeficientes vizinhos do núcleo $h_p(\mathbf{k}_p)$, pode-se substituí-lo por um núcleo esparso $h_p^{(s)}(\mathbf{k}_p)$. Adotando um fator de decimação L , tal núcleo é dado por $h_p^{(s)}(\mathbf{k}_p) = 0$ se $k_i \neq rL$ para algum i , com $r \in \mathbb{N}$, e $h_p^{(s)}(\mathbf{k}_p) = h_p(\mathbf{k}_p)$ caso contrário [11]. Os núcleos esparsos agem sobre a saída de um filtro interpolador de comprimento $2L - 1$ cuja entrada é $u(n)$. Com isso, além das $2L - 1$ multiplicações relativas a esse filtro, o número de multiplicações por amostra requeridas é dado por (2) com N_p trocado por

$$N_p^{(s)} = \left\lfloor \frac{N_p - 1}{L} + 1 \right\rfloor \quad (3)$$

C. Modelo PARAFAC-Volterra

Núcleos $h_p(k_1, \dots, k_p)$, $p < 1$, simétricos, isto é, invariantes a permutações dos índices k_i , podem ser escritos como uma decomposição PARAFAC

$$h_p(\mathbf{k}_p) = \sum_{r=1}^{R_p} \lambda_r \prod_{i=1}^p a_{p,r}(k_i), \quad (4)$$

para algum posto R_p [12]. Esta decomposição pode ser vista como a generalização da decomposição SVD de uma matriz simétrica, que corresponde ao caso $p = 2$. Absorvendo, para maior simplicidade, os fatores λ_r nos vetores $\mathbf{a}_{p,r}$ e introduzindo a expressão em (1), resulta para a parcela da saída referente a $h_p(\mathbf{k}_p)$

$$\hat{v}_p(n) = \sum_{r=1}^{R_p} \prod_{i=1}^p \hat{v}_{p,r}(n) = \sum_{r=1}^{R_p} [\hat{v}_{p,r}(n)]^p, \quad (5)$$

onde $\hat{v}_{p,r}(n) \triangleq \sum_{k=0}^{N_p-1} a_{p,r}(k)u(n-k)$. Pode-se ver que $\hat{v}_p(n)$ é obtido passando a entrada por R_p filtros de resposta impulsiva $a_{p,r}$, elevando a saída de cada um à potência p e somando os resultados.

Na prática, para um melhor compromisso entre complexidade computacional e precisão, pode-se usar um número de filtros $r_p < R_p$. Desta forma, o número de multiplicações por amostra requeridas é

$$N_{\text{mult}} = N_1 + \sum_{p=2}^P r_p(N_p + p - 1) \quad (6)$$

D. Derivação do bloco de pré-processamento

Partindo de um modelo em espaço de estados não-linear do alto-falante, os núcleos de um filtro de Volterra convencional definido por (1) podem ser calculados de forma eficiente, conforme descrito em [8] e [9]. Podem então ser usados diretamente no bloco PNL de pré-processamento.

De modo a reduzir a carga computacional do bloco PNL (às custas de alguma perda de precisão), estes núcleos podem ser substituídos por suas versões esparsas, conforme visto na Seção II-B. Outra opção para reduzir a complexidade computacional é o modelo PARAFAC-Volterra. Para tanto, a decomposição PARAFAC dos núcleos pode ser realizada, por exemplo, com o pacote *Tensorlab* para MATLAB [13].

III. COMPROMISSO ENTRE DESEMPENHO E COMPLEXIDADE COMPUTACIONAL

Nesta seção verifica-se o funcionamento do método de NLAEC proposto em um caso concreto. Para as três estruturas não-lineares discutidas na seção anterior, determina-se o bloco de pré-processamento a partir de um modelo de alto-falante automotivo para graves descrito em [7]. O compromisso entre desempenho e complexidade computacional de cada estrutura é então levantado.

A. Procedimento geral dos testes

Fazendo referência ao diagrama de blocos na Figura 1, o sinal de entrada $u(n)$ é o de um locutor masculino com duração de 10 s e frequência de amostragem 5 kHz, obtido da biblioteca TIMIT [14] e submetido a filtragem passa-baixas com corte em 500 Hz; sua amplitude é ajustada para corresponder a uma variância de 5 V^2 , situação em que o comportamento não-linear do alto-falante é apreciável; as amostras de saída do alto-falante $v(n)$ são calculadas por integração numérica do seu modelo físico, usando a função `ode15s` do MATLAB; a resposta impulsiva do ambiente H é obtida do banco de dados SMARD [15]; o filtro adaptativo W tem 256 coeficientes e é adaptado pelo algoritmo NLMS

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu}{\gamma + \hat{\mathbf{v}}(n)^T \hat{\mathbf{v}}(n)} e(n) \hat{\mathbf{v}}(n) \quad (7)$$

com $\gamma = 10^{-9}$ e $\mu = 1, 2$.

O desempenho do cancelamento é medido pelo ERLE (*Echo Return Loss Enhancement*):

$$\text{ERLE} = \frac{\sum_n y^2(n)}{\sum_n [\hat{y}(n) - y(n)]^2}, \quad (8)$$

onde as somatórias correspondem a blocos de 0,1 s de duração.

Para avaliação dos resultados, admite-se que o sinal de eco $y(n)$ é da mesma ordem de grandeza do sinal local $r(n)$, quando este está presente (o que é razoável em um sistema viva-voz). Neste caso, pode-se constatar que $\text{ERLE} < 20 \text{ dB}$ já é inconveniente para o usuário e $\text{ERLE} > 40 \text{ dB}$ é praticamente o mesmo que não ter eco.

B. Filtros de Volterra

Consideram-se inicialmente filtros de Volterra de ordens $P = 1$ (correspondendo a pré-processamento apenas linear) $P = 2$ e $P = 3$ e extensões de memória $N_p = 100$, $p = 1, 2, 3$. Para enfatizar os aspectos de desempenho em questão, consideramos inicialmente que o sinal local $r(n)$ é nulo. Os resultados de ERLE estão na Figura 2, onde, para melhor visualização, foi aplicada uma janela temporal de suavização de 4 amostras. O número de multiplicações de cada caso está na Tabela I.

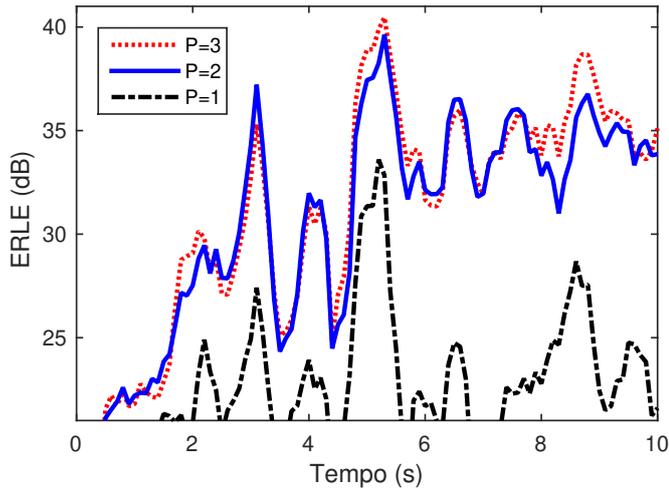


Fig. 2. ERLE para filtros de Volterra de ordem $P = 1, 2, 3$.

Modelo	N_{mult}
Filtro de Volterra, $P = 1$	100
Filtro de Volterra, $P = 2$	10200
Filtro de Volterra, $P = 3$	525300
Esparso interpolado	682
PARAFAC-Volterra	403

TABELA I

COMPLEXIDADE COMPUTACIONAL DO BLOCO PNL

Constata-se que o desempenho com ordem $P = 2$ ou $P = 3$ é bastante superior ao desempenho com $P = 1$, o que era esperado. Por outro lado, interessantemente, usar uma ordem $P = 3$, apesar de representar um aumento de 51 vezes na complexidade computacional, resulta apenas em um pequeno aumento de desempenho, o que será discutido na Seção IV. Desta forma, retemos apenas o filtro de Volterra de ordem $P = 2$ para a avaliação seguinte.

C. Modelos esparso interpolado e PARAFAC

Consideram-se agora os modelos esparso interpolado com $L = 4$ e PARAFAC-Volterra com $r_p = 3$, ambos de ordem $P = 2$ e $N_p = 100$. Os resultados estão na Figura 3, ainda para $r(n) = 0$. Verifica-se que o modelo PARAFAC-Volterra tem um desempenho praticamente igual ao do filtro de Volterra de ordem $P = 2$ e, ao mesmo tempo, conforme a Tabela I, tem uma complexidade computacional 25 vezes menor. O

modelo esparso interpolado também tem uma complexidade computacional menor mas seu desempenho não é tão bom.

Finalmente, os resultados com $r(n)$ dado por ruído branco Gaussiano 40 dB abaixo do sinal de eco estão na Figura 4. Como esperado, a presença de ruído reduz os valores alcançados de ERLE para os três modelos. O modelo PARAFAC-Volterra continua, porém, proporcionando o melhor compromisso entre desempenho e complexidade computacional.

No caso do alto-falante particular em questão conclui-se, portanto, que o modelo PARAFAC-Volterra descrito é a melhor alternativa de implementação do bloco PNL de pré-processamento não-linear do método de NLAEC proposto.

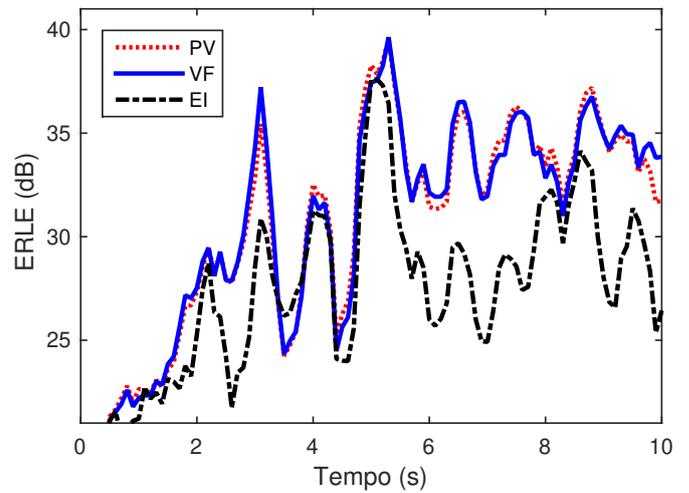


Fig. 3. ERLE para filtro de Volterra (VF) e modelos esparso interpolado (EI) e PARAFAC-Volterra (PV)

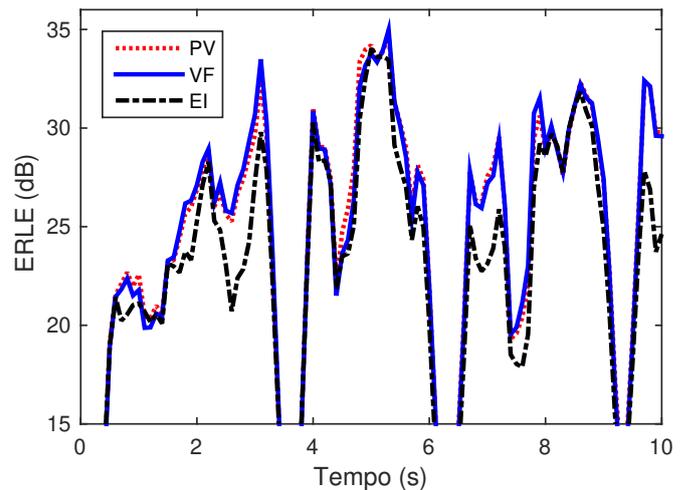


Fig. 4. ERLE para filtro de Volterra (VF) e modelos esparso interpolado (EI) e PARAFAC-Volterra (PV), com ruído branco gaussiano

IV. EFEITO DO FILTRO ADAPTATIVO

Fazendo novamente referência à Figura 1, a precisão da aproximação do alto-falante pelo bloco de pré-processamento

PNL pode ser medida pelo EAN (Erro de Aproximação Normalizado)

$$\text{EAN} = \frac{\sum_n v^2(n)}{\sum_n [\hat{v}(n) - v(n)]^2} \quad (9)$$

Na Figura 5 temos os valores de EAN para os filtros de Volterra. Nota-se que com ordem $P = 3$ consegue-se uma precisão consideravelmente melhor do que com $P = 2$. Em princípio, para uma resposta ambiente H não particularmente seletiva, esperaria-se que isso se refletisse em um ganho de desempenho de ERLE semelhante. Como visto na Seção III-B, porém, tal ganho não ocorre.

Pode-se concluir então que o filtro adaptativo, apesar de ser puramente linear, é capaz de compensar alguma imprecisão do modelamento não-linear do alto-falante. Este comportamento, cuja análise mais detalhada é deixada para um trabalho futuro, é bastante benéfico por duas razões: 1) como visto, permite que o bloco PNL tenha uma complexidade computacional bem menor; 2) além disso, irá atenuar o efeito das inevitáveis imprecisões entre o modelo físico usado para derivar o PNL e o alto-falante real em cada unidade do equipamento viva-voz.

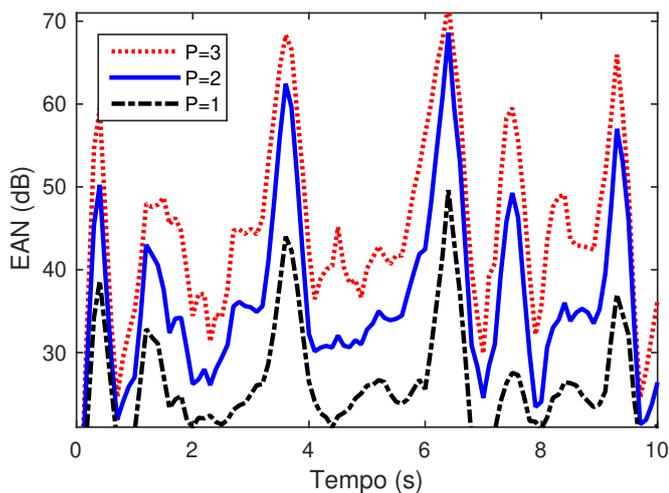


Fig. 5. EAN para filtros de Volterra de ordem $P = 1, 2, 3$.

V. CONCLUSÕES

Um novo método de cancelamento de eco acústico não-linear foi proposto e validado. O método consiste em usar um bloco de pré-processamento não-linear fixo derivado de um modelo físico do alto-falante de um sistema viva-voz. Tal modelo físico deve ser do conhecimento do fabricante do equipamento de viva-voz. Descreveu-se a complexidade computacional de diferentes alternativas de implementação de tal bloco não-linear. Para o caso de um alto-falante particular, foi levantado o compromisso entre desempenho e complexidade computacional destas alternativas, verificando-se que a melhor opção no caso é o uso de uma estrutura PARAFAC-Volterra de segunda ordem. Verificou-se também que o filtro adaptativo, ainda que puramente linear, compensa imperfeições do modelo não-linear do alto-falante, tendo um papel importante no desempenho alcançado pelo sistema de cancelamento de eco acústico.

REFERÊNCIAS

- [1] C. Breining, P. Dreiscitel, E. Hansler, A. Mader, B. Nitsch, H. Puder, T. Schertler, G. Schmidt, and J. Tilp, "Acoustic echo control. an application of very-high-order adaptive filters," *Signal Processing Magazine, IEEE*, vol. 16, no. 4, pp. 42–69, Jul 1999.
- [2] A.N. Birkett and R.A. Goubran, "Limitations of handsfree acoustic echo cancellers due to nonlinear loudspeaker distortion and enclosure vibration effects," in *Applications of Signal Processing to Audio and Acoustics, 1995., IEEE ASSP Workshop on*, Oct 1995, pp. 103–106.
- [3] M.I. Mossi, N.W.D. Evans, and C. Beaugeant, "An assessment of linear adaptive filter performance with nonlinear distortions," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, March 2010, pp. 313–316.
- [4] Riitta Niemistö and Tuomo Mäkelä, "On performance of linear adaptive filtering algorithms in acoustic echo control in presence of distorting loudspeakers," in *International Workshop on Acoustic Echo and Noise Control (IWAENC2003), Kyoto, Japan*, Japan, Kyoto, Sept. 2003.
- [5] A. Stenger, L. Trautmann, and R. Rabenstein, "Nonlinear acoustic echo cancellation with 2nd order adaptive volterra filters," in *Acoustics, Speech, and Signal Processing, 1999. Proceedings., 1999 IEEE International Conference on*, Mar 1999, vol. 2, pp. 877–880 vol.2.
- [6] A. Guerin, G. Faucon, and R. Le Bouquin-Jeannes, "Nonlinear acoustic echo cancellation based on volterra filters," *Speech and Audio Processing, IEEE Transactions on*, vol. 11, no. 6, pp. 672–683, Nov 2003.
- [7] D. Franken, K. Meerkotter, and J. Wassmuth, "Passive parametric modeling of dynamic loudspeakers," *Speech and Audio Processing, IEEE Transactions on*, vol. 9, no. 8, pp. 885–891, Nov 2001.
- [8] J. H. d. M. Goulart and P. M. S. Burt, "Efficient kernel computation for volterra filter structure evaluation," *IEEE Signal Processing Letters*, vol. 19, no. 3, pp. 135–138, March 2012.
- [9] P. M. S. Burt and J. H. de Moraes Goulart, "Efficient computation of bilinear approximations and volterra models of nonlinear systems," *IEEE Transactions on Signal Processing*, vol. 66, no. 3, pp. 804–816, Feb 2018.
- [10] V. J. Mathews and G. L. Sicuranza, *Polynomial signal processing*, Wiley, 2000.
- [11] E.L.O. Batista, O.J. Tobias, and R. Seara, "A sparse-interpolated scheme for implementing adaptive volterra filters," *Signal Processing, IEEE Transactions on*, vol. 58, no. 4, pp. 2022–2035, April 2010.
- [12] Gérard Favier, Alain Y. Kibangou, and Thomas Bouilloc, "Nonlinear system modeling and identification using Volterra-PARAFAC models," *International Journal of Adaptive Control and Signal Processing*, vol. 26, no. 1, pp. 30–53, Jan. 2012.
- [13] www.tensorlab.net.
- [14] J. S. Garofolo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, D. S. Pallett, and N. L. Dahlgren 1993.
- [15] J.K. Nielsen, J.R. Jensen, S.H. Jensen, and M.G. Christensen, "The single- and multichannel audio recordings database (smard)," in *International Workshop Acoustic Signal Processing*, Sep. 2014.