

Annulation d'échos acoustiques par filtrage non linéaire

J.-P. COSTA, Laurent BARCAROLI, Axel LAGRANGE

IUP GMI d'Avignon
Agroparc BP1228 – 84911 Avignon Cedex 9, France
jpc@iup.univ-avignon.fr

Résumé – De nos jours, l'utilisation d'un téléphone portable en mode main libre fait intervenir un filtre linéaire adaptatif, dans le but de compenser les échos acoustiques. Dans ce mode de fonctionnement, le signal émis par le haut parleur est de fort niveau; il en résulte que des distorsions non linéaires interviennent dans la propagation du signal, entre le haut parleur et le microphone. Différents types de filtres sont proposés pour compenser ces échos, incluant les filtres non linéaires en cascade ainsi que le filtre bilinéaire. Le filtre bilinéaire est un sous modèle du NARMAX (Nonlinear Autoregressive Moving Average with eXogenous inputs). Par la suite, nous présenterons une évaluation des performances basée sur une mesure standard ERLE (Echo Return Loss Enhancement), entre les différents filtres utilisés.

Abstract – The miniaturization of GSM handsets creates nonlinear acoustical echoes between microphones and loudspeakers when the signal level is high (hands-free communication). Several methods including nonlinear cascade filters and a bilinear filter are proposed to compensate these echoes. A bilinear filter is a restricted NARMAX (Nonlinear Autoregressive Moving Average with eXogenous inputs) filter. We will present an evaluation based on the standard ERLE (echo return loss enhancement) measure, between a simple linear adaptive FIR filter and various nonlinear filters. These experiments are carried out first on a simulated communication system, then on experimental signals.

1 Introduction

De nos jours, l'utilisation d'un téléphone portable en mode main libre fait intervenir un filtre linéaire adaptatif, dans le but de compenser les échos acoustiques. Dans ce mode de fonctionnement, le signal émis par le haut parleur est de fort niveau; il en résulte que des distorsions non linéaires interviennent dans la propagation du signal, entre le haut-parleur et le microphone. Ces perturbations non linéaires peuvent être de type mécanique (vibrations), mais également de type électronique (saturation du microphone). Donc, une relation linéaire ne permet pas de modéliser entièrement le comportement du système [1], [2],[3]. Traditionnellement l'AEC (Acoustical Echo Cancellation) employé ne peut compenser ce type d'échos [4].

Notre objectif est de proposer un filtre permettant d'annuler les différents échos, au niveau du microphone, issus du haut parleur (cf figure 1). Cette étude a été proposée par la société Stepmind qui développe un circuit intégré GSM/GPRS/EDGE réalisant toutes les fonctions bande de base pour un terminal mobile GSM. Dans ce cadre, la plupart des fonctions traitement du signal sont développées pour être intégrées sur un DSP virgule fixe 16 bits. Le filtre proposé devra donc respecter les contraintes de taille mémoire et de temps réel de l'application.

De nombreuses méthodes ont été étudiées dans le cadre d'annulation d'échos non linéaires. On peut citer :

- les filtres de Volterra [5]. Ces filtres sont une généralisation du filtre RIF en non linéaire, cependant il nécessite une grande complexité de calcul (section 3).
- Les réseaux de neurones [6]; cette structure en cascade offre de nouvelles perspectives mais nécessite l'apport d'une nouvelle référence par l'intermédiaire d'un nouveau microphone [4].
- Structure NARMAX; par définition c'est le modèle paramétrique le plus général. Cependant, du fait de sa com-

plexité de structure il nécessite une phase de pré-identification [7].

- Plus récemment des structures de filtres en cascade ont été proposées [2],[4]. Dans l'article [2] un modèle non polynomial basé sur une structure de type Wiener-Hammerstein est proposé, avec une fonction de saturation sans mémoire; alors que dans l'article [4] un modèle polynomial de type hammerstein est présenté.

Dans le paragraphe 2 nous présenterons les différents modèles utilisés dans cette étude. Des commentaires seront proposés dans le paragraphe 3. Enfin, les résultats en simulation et sur signaux expérimentaux seront regroupés dans la table 2 du paragraphe 5.

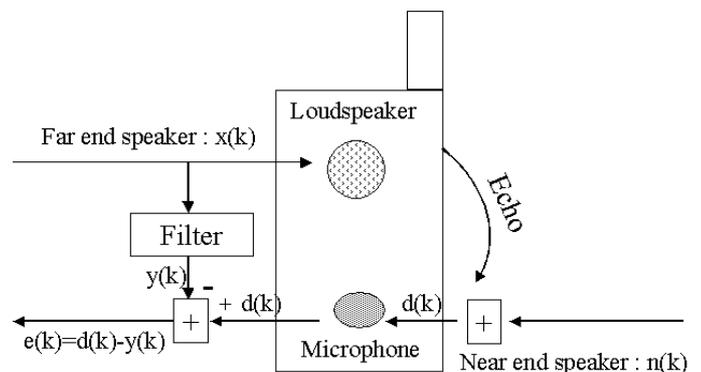


FIG. 1: Diagramme AEC.

2 Les filtres

Plusieurs méthodes ont été proposées pour répondre aux spécificités de cette application, comme il a été décrit dans le paragraphe précédent. Cependant, les filtres de Wiener (polynomial), Wiener-Hammerstein (polynomial) et filtre Bilinéaire (sous modèle du NARMAX) n'ont jusqu'à présent et à notre connaissance, jamais fait l'objet d'une étude dans ce cadre de travail.

Soit y_k et x_k respectivement l'entrée et la sortie du filtre au temps k .

2.1 Le filtre RIF

La relation d'entrée-sortie est définie comme suit :

$$y_k = \sum_{i=0}^{m_x-1} h_i x_{k-i}, \quad (1)$$

où m_x correspond à la mémoire du filtre. Ce filtre est communément utilisé dans les fonctions d'annulation d'échos des DSP du commerce.

2.2 Filtres non linéaires

2.2.1 Modèle de Volterra

Le filtre de Volterra de mémoire m_x , de degré de nonlinéarité D est défini par :

$$y_k = \sum_{i=1}^D \sum_{j_1, \dots, j_i=0}^{m_x-1} \theta_{j_1, \dots, j_i}^i x_{k-j_1} \cdots x_{k-j_i} \quad (2)$$

Le filtre de Volterra est très attractif car c'est la généralisation directe d'une description linéaire des systèmes. De plus, il caractérise le comportement d'un grand nombre de systèmes physiques. [8].

2.2.2 Modèle Bilinéaire

Le modèle Bilinéaire est un modèle paramétrique qui contient des termes croisés (entre l'entrée et la sortie du filtre) [9]. Ce modèle est une sous-classe du modèle NARMAX (Nonlinear Autoregressive Moving Average with exogenous inputs) [7] :

$$y_k = \sum_{i=0}^{m_{x1}-1} \alpha_i x_{k-i} + \sum_{j=1}^{m_{y1}} \beta_j y_{k-j} + \sum_{i=0}^{m_{x2}-1} \sum_{j=1}^{m_{y2}} \gamma_{ij} x_{k-i} y_{k-j} \quad (3)$$

Ce rebouclage de la sortie en entrée permet d'obtenir une représentation concise de systèmes non-linéaires.

2.3 Modèles non linéaires en cascade

Le principal avantage de ces structures est de limiter fortement le nombre de paramètres à estimer pour une complexité égale, cf paragraphe 3.

2.3.1 Modèle d'Hammerstein

Il correspond à un filtre polynomial sans mémoire suivi d'un filtre linéaire RIF.

$$\begin{aligned} u_k &= \sum_{i=1}^D a_i x^i \\ y_k &= \sum_{i=0}^{m_x-1} h_i u_{k-i} \end{aligned} \quad (4)$$

2.3.2 Modèle de Wiener

C'est un filtre RIF suivi d'un filtre polynomial sans mémoire.

$$u_k = \sum_{i=0}^{m_x-1} h_i x_{k-i}, \quad y_k = \sum_{i=1}^D a_i u_k^i \quad (5)$$

2.3.3 Modèle de Wiener-Hammerstein

Il est composé d'un filtre RIF, d'un filtre polynomial sans mémoire suivi d'un filtre RIF.

$$\begin{aligned} u_k^1 &= \sum_{i=0}^{m_{x1}-1} h_i x_{k-i}, \quad u_k^2 = \sum_{i=1}^D a_i (u_k^1)^i \\ y_k &= \sum_{i=0}^{m_{x2}-1} h_i u_{k-i}^2 \end{aligned} \quad (6)$$

3 Commentaires et limitations

3.1 Volterra vs structures en cascade

Soit y_k^V , y_k^W et y_k^H respectivement la séquence de sortie du modèle de Volterra, Wiener et Hammerstein. Fixons $D = 2$ et $m_x = 2$:

$$y_k^V = \theta_0^1 x_k + \theta_1^1 x_{k-1} + \theta_0^2 x_k^2 + \theta_1^2 x_{k-1}^2 + \theta_{01}^2 x_k x_{k-1} \quad (7)$$

$$y_k^W = \underline{a_1 h_0} x_k + a_1 h_1 x_{k-1} + a_2 h_0 x_k^2 + a_2 h_1 x_{k-1}^2 + \dots + a_2 h_0 h_1 x_k x_{k-1} \quad (8)$$

$$y_k^H = \underline{a_1 h_0} x_k + a_1 h_1 x_{k-1} + a_2 h_0 x_k^2 + a_2 h_1 x_{k-1}^2 \quad (9)$$

A partir de cet exemple on peut noter plusieurs points importants :

- Les modèles de Volterra et Wiener contiennent plus de termes que le modèle d'Hammerstein $x_k x_{k-1}$ (tous les termes croisés sont présents).
- Le nombre de paramètres du modèle de Volterra est beaucoup plus important (cf table 1) que pour les modèles en cascade.
- Peu de paramètres sont nécessaires pour les modèles de Wiener et d'Hammerstein (à complexité non linéaire égale). En revanche, la forte contrainte associée est que les paramètres des termes sont dépendants entre eux. ($\underline{a_1 h_0}$ pour le terme en x_k). Ceci peut conduire à des problèmes d'estimation. Il en est de même pour le modèle Wiener-Hammerstein.
- Les modèles non linéaires en cascade peuvent être considérés comme des sous-classes du modèle de Volterra.

3.2 Nombre de paramètres à estimer

Le nombre de paramètres à estimer est une caractéristique importante dans la complexité calculatoire ainsi que pour la convergence des algorithmes mis en oeuvre.

Prenons les exemples suivants :

- EXEMPLE 1 : $D = 3$, $m_x = 10$, $m_{x1} = m_{x2} = 10$ and $m_{y1} = 0$ and $m_{y2} = 1$.
- EXEMPLE 2 : $D = 3$, $m_x = 20$, $m_{x1} = m_{x2} = 20$ and $m_{y1} = 0$ and $m_{y2} = 1$.

Les résultats sont présentés dans la table 1.

Modèle	Nb de paramètres	EX1	EX2
FIR	m_x	10	20
Volterra	$(D + m_x)! / (D! m_x!)$	285	1770
Hammerstein	$m_x + D$	13	23
Wiener	$m_x + D$	13	23
Wiener-Hammerstein	$m_{x1} + m_{x2} + D$	23	43
Bilinear	$m_{x1} + m_{y1} + m_{x2} m_{y2}$	20	40

TAB. 1: Nombre de paramètres.

On montre aisément que le nombre de paramètres du modèle de Volterra croît rapidement avec D et/ou m_x . Une conséquence importante, est qu'un nombre important d'échantillons est nécessaire afin d'estimer ces paramètres avec une précision raisonnable, ce qui est contraire avec l'aspect temps réel de notre application. Pour toutes ces raisons, le modèle de Volterra ne sera pas considéré par la suite; de plus, on limitera le degré du polynôme à $D = 3$ [6].

4 Algorithmes Adaptatifs

Les paramètres (exceptés les a_i des structures en cascade) sont mis à jour avec l'algorithme NLMS (gradient stochastique), en minimisant l'erreur quadratique moyenne. Les paramètres des filtres polynômiaux sont quant à eux mis à jour avec l'algorithme RLS, qui permet une convergence plus rapide, ce qui constitue un élément important dans notre application d'AEC [4]. De plus, des erreurs *a priori* et *a posteriori* sont utilisées pour le calcul de la mise à jour des paramètres. Ceci conduit à augmenter la qualité de l'annulation d'échos [10] (cf description de l'algorithme en [4]).

5 Expérimentations

5.1 Résultats de simulations

Traditionnellement le canal de communication non linéaire peut être modélisé par une saturation (fonction sans mémoire) suivie d'un propagation linéaire. Cette modélisation ne tient donc pas compte des propagations mécaniques (vibrations).

La fonction sigmoïde [11], définie par la suite, permet de modéliser des saturations (figure 4 pour $\alpha = 1, 2$ et 5).

$$\varphi(u) = \left(\frac{2}{1 + \exp(-\alpha u)} - 1 \right) \beta. \quad (10)$$

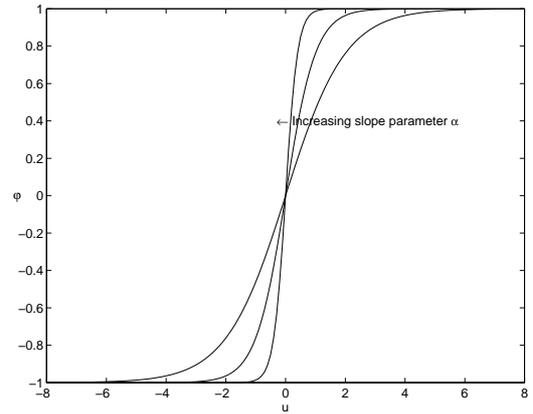


FIG. 2: Fonction Sigmoïde.

On choisit comme système linéaire $H(z) = \frac{1}{1 - \tau_1 z^{-1}}$, avec $\tau_1 = 0.2$. La séquence d'entrée $\{x_k\}$ est une séquence i.i.d. $\mathcal{N}(0, 1)$.

En "single talk", les résultats seront présentés en utilisant l'Echo Return Loss Enhancement (ERLE), défini par :

$$ERLE_{dB} = 10 \log \left[\frac{\mathbf{d}^t \mathbf{d}}{\mathbf{e}^t \mathbf{e}} \right] \quad (\text{single talk}),$$

et pour le "double talk", nous utilisons l'inverse de l'erreur quadratique moyenne normalisée :

$$Er_{dB} = 10 \log \left[\frac{\mathbf{d}^t \mathbf{d}}{(\mathbf{e} - \mathbf{n})^t (\mathbf{e} - \mathbf{n})} \right] \quad (\text{double talk}),$$

où \mathbf{d} , \mathbf{e} , \mathbf{n} sont respectivement les séquences observée, transmise et de parole en local (cf figure 1). Le degré du polynôme D ainsi que les tailles mémoires m_x , m_{x1} des filtres sont fixés à : $D = 3$, $m_x = 10$, $m_{x1} = 10$. Pour le modèle bilinéaire, les résultats indiqués ont été obtenus pour : $m_{y1} = 3$, $m_{x2} = 3$, $m_{y2} = 4$; et pour le modèle Wiener-Hammerstein pour : $m_{x2} = 10$. Les résultats indiqués dans la table 2 ont été obtenus en faisant la moyenne sur 100 réalisations.

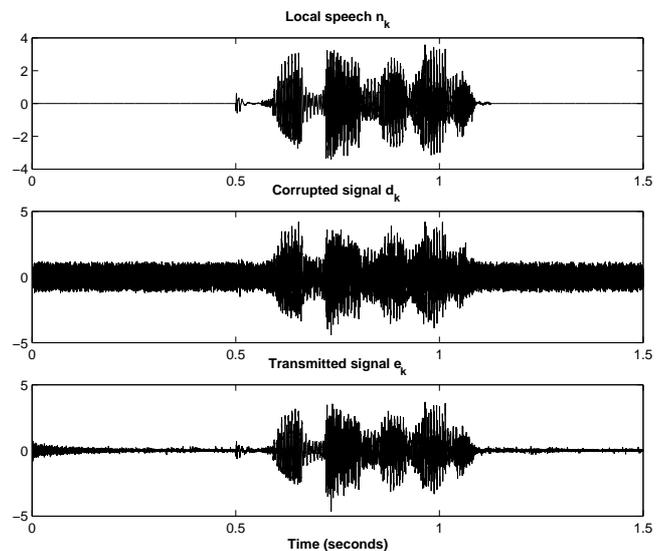


FIG. 3: En simulation.

5.2 Signaux réels

Dans ce paragraphe nous comparons les résultats d'annulations d'échos obtenus par les filtres non linéaires et comparé avec le filtre RIF adaptatif traditionnellement implémenté sur des signaux enregistrés en conditions réelles (sur GSM mode main libre) sans "double talk" ($n_k = 0$). On fixe le degré du polynôme à $D = 3$ et la mémoire à $m_x = 256$. Les résultats ont été obtenus pour $m_{x1} = 100, m_{y1} = 70, m_{x2} = 2, m_{y2} = 4$ pour le modèle bilinéaire et $m_{x1} = 100, m_{x2} = 150$ pour le Wiener-Hammerstein.

MODÈLE	SIMULATION		EXPÉRIMENTAL
	ERLE	E_r	ERLE
FIR	8.79	14.44	14.32
Hammerstein	17.28	16.61	17.49
Wiener	16.61	11.38	14.85
Wiener-Hammerstein	18.43	13.76	12.54
Bilinear	9.74	13.52	15.23

TAB. 2: Résultats pour l'annulation d'échos non linéaire.

5.3 Commentaires

- Les meilleurs résultats de simulation (en terme d'ERLE) ont été obtenus par le modèle polynômial Wiener-Hammerstein, ainsi que par le modèle d'Hammerstein pour le "double talk".
- Le filtre bilinéaire, du fait structure complexe a des résultats préliminaires intéressants et nécessite une étude plus fine avant d'être rejeté définitivement.
- Il faut noter que les signaux récoltés au cours de cette expériences n'ont pas été mis en situation de saturation du haut parleur, ce qui explique la faible différence de performances entre le filtre RIF et les filtres non linéaires.

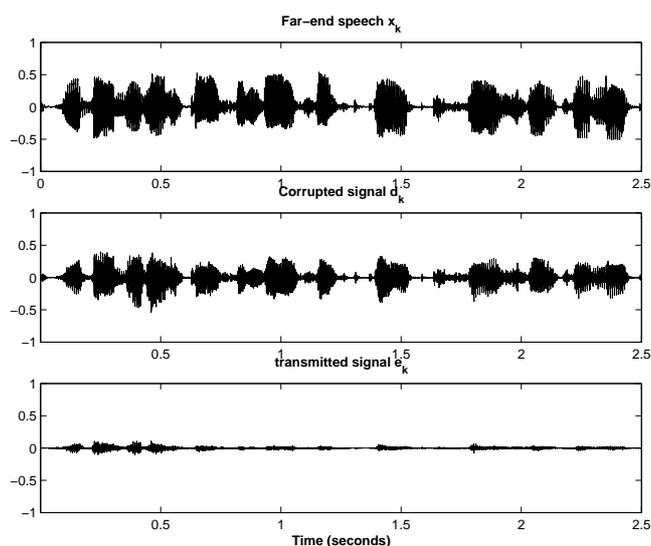


FIG. 4: Experimental results.

6 Synthèse

De nouvelles structures de filtre ont été proposé dans le cadre d'une application d'annulation d'échos non linéaire : des filtres

en cascades et le filtre bilinéaire. Ces structures ont été comparées avec le filtre généralement implémenté le RIF adaptatif (NLMS). Pour l'ensemble des tests effectués le modèle Hammerstein semble avoir le meilleur comportement. Il faut nuancer ces résultats, car ils ont été obtenus que sur une série de signaux provenant d'une seule et même expérience. Il reste donc à valider cette étude sur de nouveaux signaux. De plus, nous n'avons pas traité le cas où le système était perturbé par un bruit additif. Cette hypothèse n'est pas réaliste dans le cadre de notre application GSM en mode main libre. Il reste donc à valider ces résultats en présence de bruit ambiant (voiture, blabla noise,...).

Nous tenons à remercier Messieurs F. Kuech et A. Stenger de l'université Erlangen-Nuremberg de nous avoir fourni leur signaux expérimentaux.

Références

- [1] F. Heinle, R. Rabenstein, and A. Stenger, "A measurement method for the linear and nonlinear properties of electro-acoustic transmission systems.," *Signal Processing*, vol. 64, no. 1, pp. 49–60, 1998.
- [2] B. S. Noll et D. L. Jones, "Nonlinear echo cancellation for hands-free speakerphones," *NSIP*, September 1997.
- [3] J-P. Costa, L. Pronzato, and E. Thierry, "Nonlinear prediction by kriging, with application to noise cancellation," *Signal Processing*, vol. 80, pp. 553–566, 2000.
- [4] A. Stenger and W. Kellermann, "RLS-Adapted Polynomial for Nonlinear Acoustic Echo Cancelling," *Signal Processing*, vol. 80, pp. 1747–1760, Sep. 2000.
- [5] A. Stenger, W. Kellermann, and R. Rabenstein, "Nonlinear acoustic echo cancellation with 2nd order adaptive volterra filters," *Icassp*, 1999.
- [6] A. N. Birkett and R. A. Goubran, "Acoustic echo cancellation using nlms-neural network structures," *Icassp*, vol. 3, 1995.
- [7] J-P. Costa, T. Pitarque, and E. Thierry, "Using Orthogonal Least Square For Adaptive Nonlinear Filtering of GSM Signals," *Icassp*, vol. 3, pp. 2397–2400, 1997.
- [8] R. D. Nowak and B. D. Van Veen, "Random and Pseudorandom Inputs for Volterra Filter Identification," *IEEE trans. on Signal Processing*, vol. 42, no. 8, august 1994.
- [9] F. Fnaiech and L. Ljung, "Recursive identification of bilinear systems," *Int. J. Control*, vol. 45, no. 2, pp. 453–470, 1987.
- [10] J-P. Costa, A. Lagrange, and A. Arliaud, "Acoustic echo cancellation using nonlinear cascade filters," *Icassp, Hong-Kong*, 2003.
- [11] S. Haykin, "Adaptive Filter Theory", Third Edition, Prentice Hall, Inc, A Simon & Schuster Company, New Jersey, 1996.