

UNE NOUVELLE STRUCTURE BLANCHISSANTE POUR
ANNULATION D'ECHO ACOUSTIQUE

Mamadou MBOUP¹, Madeleine BONNET^{1,2}

1. Ecole Supérieure d'Electricité

Plateau du Moulon, 91190 Gif sur Yvette, France

et Groupement de Recherche CNRS Traitement du Signal et Image

2. Université Paris V, UFR MSI, 12 rue Cujas, 75005, Paris

RÉSUMÉ

L'annulation d'écho acoustique relève de l'identification adaptative d'un système avec entrée et sortie connues. Pour pallier la faible vitesse de convergence de l'algorithme LMS, due à la forte corrélation de la parole en entrée du système, une nouvelle structure d'annulation d'écho est présentée.

1. INTRODUCTION

Dans une salle d'audioconférence, le signal x issu du haut-parleur (fig. 1) est filtré acoustiquement par la salle ; le signal résultant y est capté par le micro, il représente un écho gênant pour l'usager distant. L'annulation d'écho, qui permet d'éviter cet inconvénient consiste à identifier le canal acoustique puis à soustraire de l'écho son estimation \hat{y} .

La figure 1 donne la structure classique de l'annulation d'écho où $H(n)$ représente le filtre adaptatif servant à identifier le chemin d'écho inconnu. La sortie de ce filtre, soustraite de l'écho, fournit l'écho résiduel utilisé pour l'adaptation du filtre :

$$e(n) = y(n) - \hat{y}(n) \quad (1)$$

Le canal acoustique est caractérisé par une durée significative de 30 à 250 ms [1] ; son identification par un filtre transverse nécessite plusieurs centaines voire plusieurs milliers de coefficients à la fréquence de 8kHz. Le canal acoustique est d'autre part sujet à de fortes ou rapides variations dues aux mouvements des personnes ou au déplacement d'objets dans la salle.

L'algorithme des moindres carrés moyens LMS est, de par sa faible complexité et sa stabilité, intéressant pour un traitement en temps réel.

Il est bien connu [2] que la vitesse de convergence de l'algorithme LMS est inversement proportionnelle à la longueur du filtre adaptatif et au degré de corrélation du signal d'entrée x . Or, ce signal est ici la parole qui est fortement non stationnaire en niveau et en spectre. Il est donc nécessaire de concevoir des algorithmes qui s'adaptent plus rapidement que le LMS.

L'idée de blanchir le signal d'entrée pour accélérer la vitesse de convergence a déjà été mise en application dans des travaux récents. Dans [3], le filtre P qui blanchit l'entrée est aussi placé sur le chemin d'écho (fig. 2) ; il est alors nécessaire de disposer du filtre inverse pour obtenir l'écho résiduel. Cet inconvénient est

ABSTRACT

Acoustic echo cancellation deals with adaptive system identification with known input-output. In order to fasten the convergence speed of the LMS algorithm, which is rather slow with a correlated input signal - as speech -, this paper presents a new cancellation structure, including a prewhitening process.

levé dans [4] en dédoublant le filtre adaptatif qui identifie le canal d'écho (fig.3). Il en découle cependant une charge de calcul supplémentaire qui rend plus coûteuse l'implémentation en temps réel.

Nous présentons ici deux algorithmes, de faible complexité, utilisant un préblanchiment, où les opérations de filtrage et d'adaptation sont séparées.

2. STRUCTURE CLASSIQUE BLANCHISSANTE

Le temps de convergence de l'algorithme LMS est proportionnel à la dispersion de la matrice de covariance du signal d'entrée. Le but recherché est de réduire ce temps sans modifier la relation entrée-sortie du canal d'écho.

La figure 4 présente une structure d'annulation d'écho strictement équivalente à la structure classique de la fig.1. En effet, le signal de parole x se trouve directement à l'entrée de l'annuleur, contrairement aux figures 2 et 3, où l'annuleur reçoit un signal blanchi. Sur ces figures, la sortie x_f du filtre adaptatif blanchissant P représente l'erreur de prédiction de l'entrée x et s'écrit :

$$x_f(n) = x(n) - \sum_{i=1} p_i x(n-i) \quad (3)$$

L'adaptation de P , rendue nécessaire par le caractère non stationnaire de la parole, est obtenue ici par un algorithme classique de type gradient stochastique.

Dans la structure proposée, le filtre P n'intervient que dans l'adaptation de l'annuleur. Ce même filtre est alors appliqué à l'écho résiduel e pour délivrer l'erreur d'adaptation e_f du nouvel algorithme

$$H(n) = H(n-1) + \mu e_f(n) X_f(n) \quad (4)$$

où $X_f(n)$ représente le vecteur des échantillons passés de l'entrée blanchie et μ est un pas d'adaptation judicieusement choisi.



Remarquons que cet algorithme est précisément l'algorithme LMS appliqué à la structure classique de la figure 1 qui aurait pour entrée le signal x_f . En effet, e_f s'écrit

$$e_f(n) = e(n) - \sum_{i=1}^p p_i e(n-i) \quad (5)$$

ce qui donne

$$e_f(n) = [y(n) - \hat{y}(n)] - \sum_{i=1}^p p_i [y(n-i) - \hat{y}(n-i)] \quad (6)$$

et, en utilisant (3) :

$$e_f(n) = \sum_j h_j^* x_f(n-j) - \sum_j h_j(n-1) x_f(n-j) \quad (7)$$

où h_j^* et $h_j(n)$ sont respectivement les coefficients de la réponse impulsionnelle du chemin d'écho inconnu et de l'annuleur à l'instant n .

Le signal x_f étant moins corrélé que le signal d'origine x , il est naturel que l'algorithme converge plus rapidement, ce qui est vérifié par les simulations détaillées au paragraphe 4.

La minimisation de la moyenne quadratique de l'écho résiduel est obtenue via celle de la moyenne quadratique de sa version filtrée e_f . Le filtrage (de l'écho résiduel), qui peut être vu comme un lissage, risque cependant d'atténuer l'information instantanée sur les variations rapides du chemin d'écho, contenue dans e . C'est pourquoi nous proposons maintenant un algorithme qui supprime ce lissage.

3. UN NOUVEL ALGORITHME BLANCHISSANT

Sur la figure 5, la suppression du filtrage de e préserve l'information instantanée. L'algorithme d'adaptation devient donc :

$$H(n) = H(n-1) + \mu e(n) X_f(n). \quad (8)$$

Pour analyser cet algorithme, nous définissons le vecteur déviation

$$V(n) \triangleq H(n-1) - H^*(n) \quad (9)$$

où H^* est la réponse du chemin inconnu variant avec le temps. Ce vecteur peut être décomposé selon

$$V(n) = V_t(n) + V_b(n) + V_r(n), \quad (10)$$

où la composante transitoire $V_t(n)$, la composante due au bruit de sortie $V_b(n)$ et la composante retard $V_r(n)$ s'écrivent [2] :

$$V_t(n) = [\mathbf{I} - \mu X_f(n) X_f^T(n)] V_t(n-1); \quad V_t(0) = V(0), \quad (11)$$

$$V_b(n) = [\mathbf{I} - \mu X_f(n) X_f^T(n)] V_b(n-1) + \mu e_o(n) X_f(n) \quad (12)$$

où e_o est un bruit additif en sortie, que l'on suppose indépendant et centré ;

$$V_r(n) = [\mathbf{I} - \mu X_f(n) X_f^T(n)] V_r(n-1) - D(n) \quad (13)$$

avec

$$D(n) = H^*(n) - H^*(n-1). \quad (14)$$

Dans ce qui suit, nous supposons que

- i) les vecteurs $X(n)$ et $V(n-1)$ sont indépendants
- ii) le signal d'entrée est stationnaire
- iii) la prédiction de l'entrée est optimale.

3.1 absence de biais

Nous supposons ici que le canal est invariant, c'est-à-dire que $D(n) = \mathbf{0}$. Montrons que l'algorithme (8) est sans biais. L'erreur s'écrit :

$$e(n) = e_o(n) + \{H^* - H(n-1)\}^T X(n) \quad (15)$$

D'après l'indépendance des vecteurs $X(n)$ et $H(n-1)$, il vient

$$\begin{aligned} E\{e(n) X_f(n)\} &= E\{e_o X_f(n)\} \\ &+ E\{X_f(n) X_f^T(n)\} E\{H^* - H(n-1)\} \end{aligned} \quad (16)$$

Le vecteur $X(n)$ se décompose en le vecteur $X_f(n)$ et en une composante de prédiction qui lui est orthogonale donc (16) s'écrit

$$E\{e(n) X_f(n)\} = \mathbf{0} + E\{X_f(n) X_f^T(n)\} E\{H^* - H(n-1)\} \quad (17)$$

$$\Rightarrow E\{e(n) X_f(n)\} = \mathbf{0} \Leftrightarrow E\{H(n-1)\} = H^* \quad (18)$$

ce qui prouve que l'algorithme est non biaisé.

3.2 Analyse au premier ordre

3.2.1 convergence de la moyenne

Supposons maintenant que le canal est toujours invariant dans le temps et que le bruit est nul ; on a donc $V(n) = V_t(n)$. D'après les hypothèses, on peut écrire :

$$E\{V(n)\} = [\mathbf{I} - \mu \mathbf{R}_{X_f}(n)] E\{V(n-1)\}, \quad (19)$$

où $\mathbf{R}_{X_f}(n)$ est la matrice de covariance diagonale des échantillons x_f .

On en déduit que $E\{V(n)\}$ tend vers $\mathbf{0}$ ssi $|1 - \mu \lambda_i\{\mathbf{R}_{X_f}\}| < 1$,

ou d'une façon équivalente ssi $0 < \mu < \frac{1}{\lambda_{\max}\{\mathbf{R}_{X_f}\}}$.

3.2.2 temps de convergence

Comme pour l'algorithme LMS, le temps de convergence est défini par

$$\tau = \frac{\rho}{2}, \quad (20)$$

où $\rho = \lambda_{\max} / \lambda_{\min}$ représente la dispersion de la matrice de covariance du signal d'entrée de l'algorithme. En vertu de l'hypothèse iii), toutes les valeurs propres de \mathbf{R}_{X_f} sont égales donc $\rho = 1$. On obtient bien une amélioration de la vitesse de convergence par rapport à l'algorithme LMS pour lequel $\rho > 1$.

3.3 Analyse au second ordre

3.3.1 comportement transitoire

Le comportement de la moyenne quadratique de la composante transitoire définie en (11) détermine la convergence de l'algorithme. On pose :

$$\mathbf{V}(n) \triangleq E\{V_t(n) V_t^T(n)\}; \quad \mathbf{\Omega}(n) \triangleq \mathbf{I} - \mu X_f(n) X_f^T(n);$$

$$\varepsilon(n) \triangleq E\{|V_t(n)|^2\}; \quad \bar{\mathbf{\Omega}}(n) \triangleq E\{\mathbf{\Omega}^T(n) \mathbf{\Omega}(n)\}. \quad (21)$$

Il vient alors, en vertu de i) :

$$\varepsilon(n) = \text{Tr}\{\mathbf{V}(n)\} = \text{Tr}\{\bar{\mathbf{Q}}(n)\mathbf{V}(n-1)\}, \quad (22)$$

où $\text{Tr}\{\mathbf{M}\}$ représente la trace de la matrice \mathbf{M} .

Si la constante d'adaptation μ est suffisamment faible, la matrice $\bar{\mathbf{Q}}(n)$ est définie positive. On obtient alors :

$$\lambda_{\min}\{\bar{\mathbf{Q}}(n)\} \varepsilon(n-1) \leq \varepsilon(n) \leq \lambda_{\max}\{\bar{\mathbf{Q}}(n)\} \varepsilon(n-1). \quad (23)$$

Désignons par \mathbf{R}_X et $\sigma_{x_f}^2$ respectivement la matrice de corrélation du signal d'entrée x et la variance de son erreur de prédiction. Il découle des hypothèses ii) et iii) que les valeurs propres extrêmes de la matrice $\bar{\mathbf{Q}}$ sont données par :

$$\alpha_m \triangleq \lambda_{\min}\{\bar{\mathbf{Q}}\} = 1 - \mu [2 - \mu L \lambda_{\min}\{\mathbf{R}_X\}] \sigma_{x_f}^2, \quad (24)$$

$$\alpha_M \triangleq \lambda_{\max}\{\bar{\mathbf{Q}}\} = 1 - \mu [2 - \mu L \lambda_{\max}\{\mathbf{R}_X\}] \sigma_{x_f}^2 \quad (25)$$

L désignant la longueur du filtre H .

La déviation quadratique moyenne transitoire $\varepsilon(n)$ décroît donc exponentiellement vers zéro ssi $\alpha_M < 1$, c'est à dire ssi :

$$0 < \mu < \frac{2}{L \lambda_{\max}\{\mathbf{R}_X\}}. \quad (26)$$

Cette vitesse de décroissance est maximale pour un choix de μ tel que :

$$\mu_{opt} = \frac{1}{L \lambda_{\max}\{\mathbf{R}_X\}}. \quad (27)$$

D'autre part pour $\mu \ll \mu_{opt}$ alors on a $\alpha_m = \alpha_M \triangleq \alpha$ d'où

$$\varepsilon(n) = K\alpha^n; \quad K = \varepsilon(0). \quad (28)$$

3.3.2 Comportement asymptotique

Nous considérons dans cette partie le cas général où le bruit additif n'est pas nul et où le canal est variable dans le temps. Après convergence, l'écart entre les filtres optimal et adaptatif, $H^*(n) - H(n-1)$, est alors dû au bruit et au retard de l'adaptation.

On définit l'erreur quadratique moyenne résiduelle (EQMR) par :

$$m(n) = E\{|V(n-1)X^T(n)|^2\} \quad (29)$$

Comme pour le vecteur déviation, l'EQMR peut être découpée selon [5] :

$$\begin{aligned} m(n) &= E\{|V_b(n-1)X^T(n)|^2\} + E\{|V_r(n-1)X^T(n)|^2\} \\ &= m_b(n) + m_r(n). \end{aligned} \quad (30)$$

Si l'on suppose que le vecteur $D(n)$, représentant les variations du canal, est aléatoire, indépendant et centré, alors on montre que l'on a asymptotiquement :

$$m_b = \frac{\mu L \sigma_{x_f}^2 E_o}{2 - \mu L \sigma_{x_f}^2} \quad (31)$$

$$m_r = \frac{1}{\mu} \frac{L \sigma_D^2}{2 - \mu L \sigma_{x_f}^2}; \quad (32)$$

où $E_o \triangleq E\{|e_o(n)|^2\}$ et $N \sigma_D^2 \triangleq E\{|D(n)|^2\}$.

L'erreur quadratique moyenne totale est alors :

$$\begin{aligned} \varepsilon &\triangleq E\{|e(n)|^2\} = E\{|e_o(n) - V(n-1)X^T(n)|^2\} \\ &= E_o + m_b + m_r \end{aligned} \quad (33)$$

$$\Rightarrow \varepsilon = E_o + \frac{\mu L \sigma_{x_f}^2 E_o}{2 - \mu L \sigma_{x_f}^2} + \frac{1}{\mu} \frac{L \sigma_D^2}{2 - \mu L \sigma_{x_f}^2}. \quad (34)$$

Ce résultat est celui d'un schéma classique (fig.1), piloté par l'algorithme LMS, avec en entrée le signal x_f . Le comportement asymptotique de notre algorithme est alors strictement équivalent à celui de l'algorithme LMS.

4. Simulations

Les simulations ont été effectuées sur des réponses tronquées de salles ; le signal en entrée est un bruit USASI (ayant un spectre similaire au spectre moyen de la parole). Les prédicteurs utilisés sont transverses d'ordre 2.

La figure 6 compare la convergence des algorithmes relatifs aux figures 1 (LMS), 4 (S1) et 5 (S2) avec un canal d'écho invariant dans le temps. Conformément à la théorie, le nouvel algorithme présente un comportement asymptotique identique à celui du LMS et, il converge beaucoup plus rapidement (12500 itérations) que l'anneur classique (28000) itérations.

La figure 7 montre la poursuite d'un chemin d'écho variant dans le temps. Là encore le nouvel algorithme obtient de meilleures performances que le LMS.

Conclusion

La structure blanchissante proposée ici, ne modifie pas les relations entrée-sortie des filtres, contrairement aux structures précédemment connues. Une amélioration importante des propriétés transitoires de l'algorithme LMS a été obtenue, tout en assurant un comportement asymptotique identique. Enfin, la complexité de notre algorithme est comparable à celle de l'algorithme LMS (ordre 2 pour le prédicteur vs 512 pour l'anneur).

Bibliographie

- [1] A. GILLOIRE: "Identification de structures transverses. Application aux réponses acoustiques", Traitement du Signal, vol.6, n° 5, 1990, pp.347-351.
- [2] O. MACCHI, M. BELLANGER: "Filtrage adaptatif transverse", Traitement du Signal, vol 5, n° 3, 1988, pp 115-132.
- [3] E. HANSLER and U. SCHULTEISS: "On line identification of Loudspeaker-Enclosure-Microphone-System for compensating acoustical echoes", Proc. IFAC *Identification and system Parameter Estimation*, 1985 New-York, pp. 1841-1846.
- [4] E. HANSLER: "Adaptive echo compensation applied to the hands-free telephone problem", Proc. *ISCAS*, 1990, pp. 279-282.
- [5] B. WIDROW, S. D. STEARNS: "Adaptive signal processing", Prentice Hall, Englewood Cliffs, New Jersey, 1985.

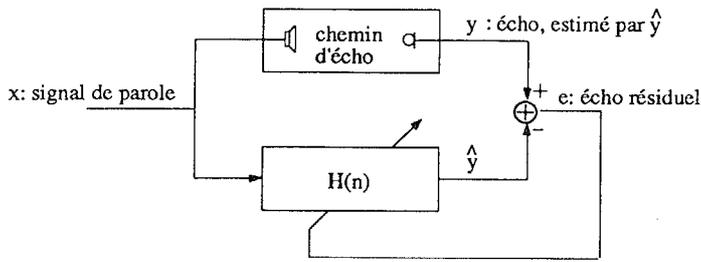


Figure 1: Structure classique d'annulation adaptative d'écho

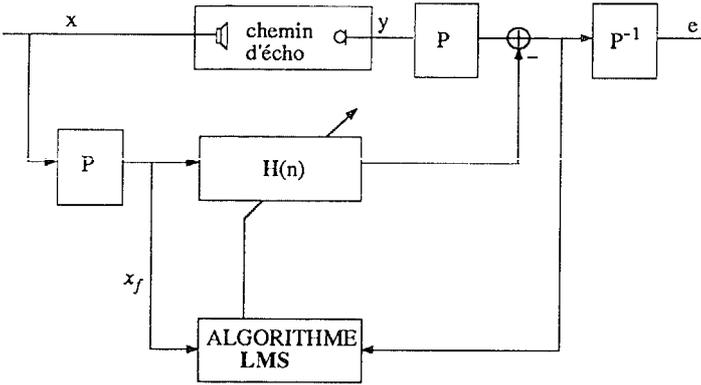


Figure 2: Préblanchiment sur le chemin d'écho, nécessitant le filtre inverse [3]

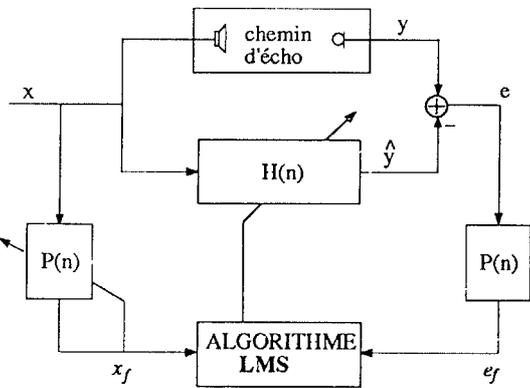


Figure 4: Structure classique avec préblanchiment (S1)

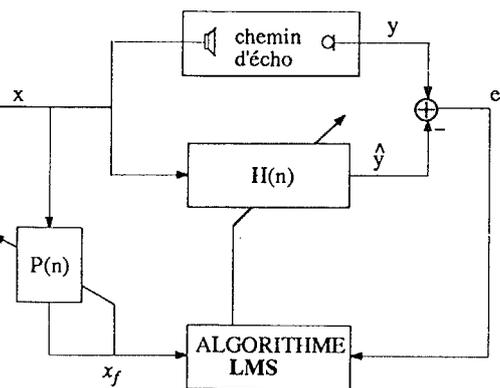


Figure 5: Nouvelle structure blanchissante (S2)

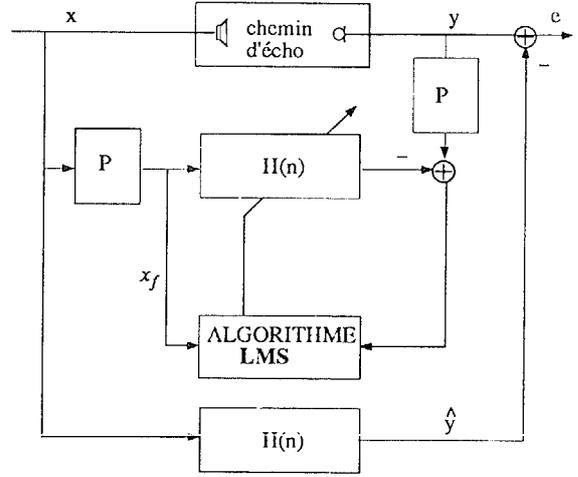


Figure 3: Préblanchiment nécessitant la recopie de l'annuleur [4]

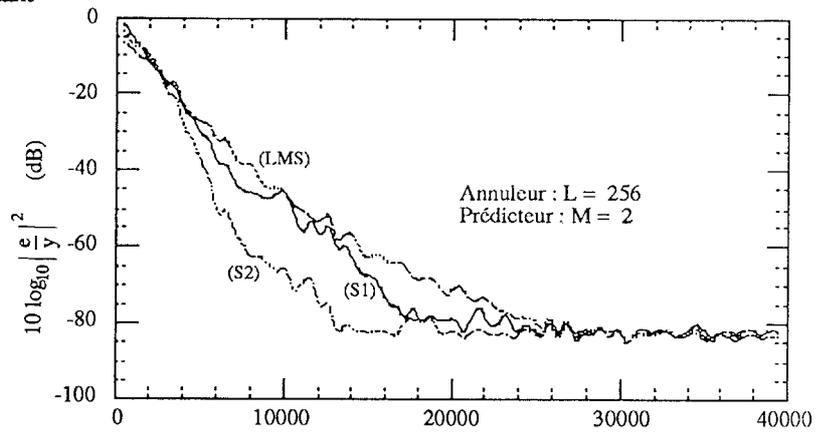


Figure 6: Vitesse de convergence

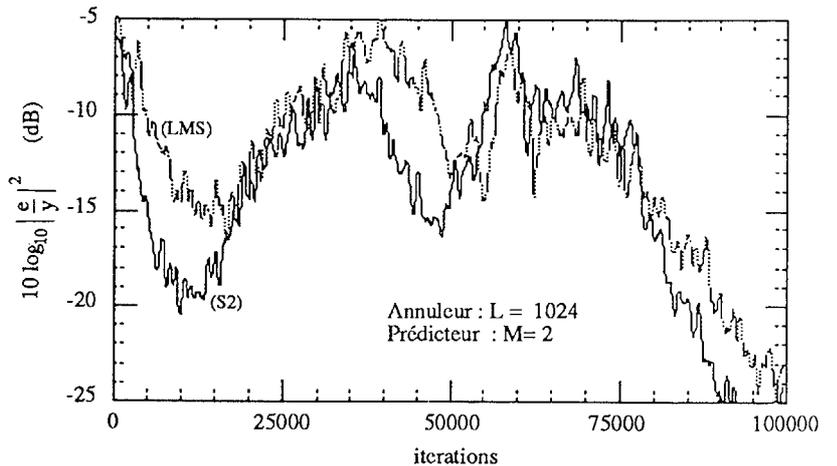


Figure 7: Poursuite des variations du chemin d'écho