

TRAITEMENT DES REPONSES IMPULSIONNELLES LONGUES EN ANNULATION D'ECHO ACOUSTIQUE

I.Kammoun⁽¹⁾⁽²⁾, *M.Jaidane*⁽¹⁾

(1) L.S.Télécoms, ENIT, Campus Universitaire, le Belvédère, Tunis
meriem.jaidane@enit.rnu.tn, Fax:(216)1872729

(2) Département TSI, ENST, 46 rue Barrault-75634 Paris Cedex 13
ikammoun@exite.com, Fax:(216)4210609

Résumé -Dans les systèmes de prise de son mains libres, lorsque la réponse impulsionnelle de couplage haut-parleur/microphone est longue, l'identification adaptative de cette réponse dans l'annulation d'écho acoustique nécessite une charge de calcul très importante. Dans le but de réduire la complexité de l'algorithme de gradient stochastique normalisé (NLMS), un algorithme d'adaptation sélective est proposé dans [1]. L'étude par simulation de cet algorithme a montré son intérêt dans le contexte radio-mobile et son insuffisance dans le contexte de visioconférence du fait de la longueur de la réponse impulsionnelle. Pour pallier à cette insuffisance, nous proposons dans cette étude un nouveau couple structure décimée/algorithme sélectif pour l'annuleur d'écho dans le cas de la visioconférence.

1 Introduction

Dans une salle de visioconférence, le canal acoustique caractérisant le chemin d'écho entre le haut-parleur et le microphone est caractérisé par une réponse impulsionnelle pouvant atteindre 250 ms [2]. Son identification par un filtre transverse peut alors nécessiter plusieurs milliers de coefficients, principalement lorsque l'on effectue une transmission de la parole en bande élargie.

L'estimation se fait habituellement en temps réel à l'aide d'algorithmes de type gradient stochastique (LMS) se caractérisant par une simplicité de mise en oeuvre et un bon conditionnement numérique. De tels algorithmes déjà peu complexes permettent de traiter des réponses impulsionnelles de quelques centaines de coefficients (contexte radio-mobile) mais deviennent insuffisants pour de longues réponses impulsionnelles. Plusieurs variantes de l'algorithme LMS (normalisé ou non) ont été proposées pour réduire encore leur complexité. L'algorithme développé dans [3] n'adapte que le coefficient associé à l'élément de données qui a la valeur absolue maximale dans la mémoire du filtre à chaque itération. La stabilité d'un tel algorithme est

Abstract - In the hands free communications, when the impulse response of the coupling loud-speaker/microphone is long, the adaptive identification of this response in acoustic echo cancellation requires very important load calculation. In order to reduce the complexity of the classical Normalized Least Mean Square (NLMS) adaptive algorithm, a selective coefficient adaptive algorithm was proposed in [1]. The simulation study of this algorithm proved his importance in the mobile radio context and his lack of performances in the videoconference one due to the impulse response length. In order to improve these results, we propose in this paper a new couple decimated structure/selectif algorithm for the acoustic echo cancellation in the videoconference case.

malheureusement très sensible à la distribution de probabilité de l'entrée. Les deux algorithmes LMS d'adaptation partielle proposés dans [4] n'adaptent qu'un ensemble prédéterminé de M coefficients à chaque itération. Pour l'un d'eux, les M coefficients d'indice i_1, \dots, i_M à l'itération k sont choisis tel que $(k - i_j + 1) \bmod \frac{L}{M} = 0$ (où L est la longueur du filtre adaptatif). Le choix de ces coefficients n'est pas basé sur une étude théorique. Dans un autre travail [5], deux algorithmes NLMS et APA (Algorithme de Projection Affine), d'adaptation partielle sélective par blocs ont été développés selon le principe de perturbation minimale. De tels algorithmes ont prouvé leurs performances dans un contexte d'annulation d'écho. Dans [1], un algorithme d'adaptation partielle sélective (Mmax NLMS), consiste à identifier les coefficients ayant une estimation de l'amplitude de leurs gradients la plus importante à chaque itération. Une étude exacte des performances de cet algorithme a été réalisée dans [7]. Une étude par simulation est montrée dans la section 2. Une structure permettant de réduire la complexité des algorithmes adaptatifs est présentée dans la section 3. L'algorithme proposé, basé sur cette structure et l'algorithme Mmax NLMS est présenté dans la section 4.

2 Insuffisance du Mmax NLMS pour les réponses impulsionnelles longues

L'idée de base de l'identification à l'aide du Mmax NLMS [1] est que la sensibilité de l'erreur relative à un coefficient du filtre adaptatif se reflète dans le degré d'importance des composantes du vecteur gradient.

En considérant l'algorithme NLMS, Le gradient estimé à l'itération k dans la direction du $i^{\text{ème}}$ coefficient est $(-2e_k x_{k-i+1})$, où e_k est l'erreur a priori définie par $e_k = y_k - X_k^T \cdot H_k$. Cette erreur représente le signal de parole locale auquel se rajoute l'écho résiduel : meilleure sera l'estimation de l'écho et plus faible sera l'écho résiduel. La sortie du microphone y_k est l'observation supposée scalaire, $X_k = (x_k, x_{k-1}, \dots, x_{k-L+1})^T$ est le vecteur de l'observation de l'entrée (la sortie du haut-parleur) de longueur L et L est la longueur du filtre modélisant la réponse impulsionnelle. Toutes les composantes du gradient comportent la quantité $-2 e_k$. L'idée du Mmax NLMS est donc de choisir les coefficients associés aux M valeurs les plus grandes de $|x_{k-i+1}|$, pour l'adaptation. L'algorithme proposé dans [1] est le suivant :

$$H_{k+1}(i) = \begin{cases} H_k(i) + \frac{\mu}{X_k^T X_k} \cdot e_k \cdot x_{k-i+1} & \text{si } i \text{ correspond} \\ & \text{à l'un des } M \text{ premiers maximums} \\ & \text{de } |x_{k-i+1}|, 1 \leq i \leq L \\ H_k(i) & \text{sinon} \end{cases} \quad (1)$$

$\mu > 0$ est le pas d'adaptation, on adapte uniquement M parmi L coefficients ; $H_k(i_1), \dots, H_k(i_M)$. On montre [1] que si $\sum_{j=1}^M x_{k-i_j+1}^2$ approche $X_k^T X_k$, la vitesse de convergence de l'algorithme d'adaptation partielle approche celle de l'algorithme NLMS complètement adaptatif. Remarquons qu'avec $M = L$, on retrouve l'algorithme NLMS. Dans [1], on montre également sous certaines conditions sur l'entrée que l'algorithme converge même dans le cas extrême $M = 1$. L'analyse de convergence a été faite en utilisant l'hypothèse d'indépendance des vecteurs d'observation. Une étude exacte des performances de cet algorithme a été réalisée dans [7]. Dans [1], l'algorithme Mmax NLMS a été testé uniquement avec des filtres transverses de longueur 200.

Ici, pour montrer l'insuffisance de cet algorithme dans le cas des réponses impulsionnelles longues, l'algorithme est testé dans le contexte radio mobile et le contexte de visioconférence. Pour les deux simulations suivantes comme pour toutes les simulations présentées dans ce papier, le signal d'entrée fortement corrélé est obtenu en faisant passer un signal blanc gaussien de moyenne nulle à travers le filtre $H(z) = \frac{1}{1-1,58 \cdot z^{-1} + 0,81 \cdot z^{-2}}$. Un bruit

blanc est ajouté à ce signal ($\text{SNR} \simeq 51\text{dB}$). Le critère de performances considéré est l'énergie moyenne de l'erreur résiduelle, calculée sur des blocs de N échantillons, soit $EMRE = 10 \log_{10} \langle e_k^2 \rangle$, où $\langle \rangle$ désigne une moyenne temporelle effectuée sur N échantillons.

• Simulation dans un contexte radio-mobile

L'analyse de la réponse impulsionnelle d'une voiture montre que les 250 premiers coefficients ont des valeurs significatives ($L=250$). La figure 1 illustre l'évolution temporelle de l'énergie moyenne résiduelle en dB pour différentes valeurs de M ($M=1$, $M=25$ et $M=L$). On remarque que l'algorithme converge dans le cas $M=1$ et qu'il donne des résultats très proches du NLMS complètement adaptatif (CA) pour $M=25$.

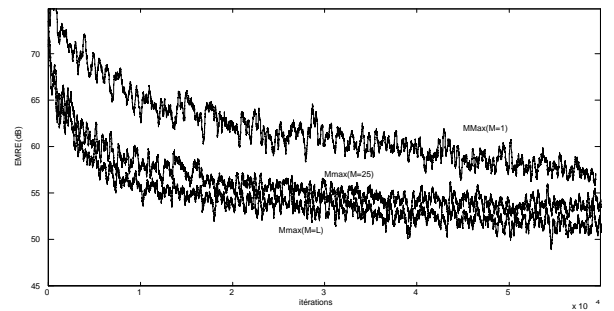


Figure 1 – EMRE(dB) pour un signal corrélé en contexte radio-mobile

• Simulation dans un contexte de visioconférence

Considérons la réponse impulsionnelle d'une salle de visioconférence où les 2000 premiers coefficients ont des valeurs significatives (fréquence d'échantillonnage à 8 kHz). D'après la figure 2, on remarque que le résultat obtenu avec $M = 1$ est très insuffisant. $M = 30$ permet bien sûr d'avoir un résultat meilleur mais qui est d'une part complexe du fait de la procédure de recherche des 30 premiers maximums à chaque itération et d'autre part encore loin du résultat obtenu avec le NLMS complètement adaptatif comme c'était le cas dans le contexte radio mobile.

Afin d'améliorer le résultat obtenu avec M faible dans le cas de la visioconférence, nous proposons d'appliquer l'algorithme Mmax NLMS à une structure dite décimée/interpolée inspirée de [6]. Cela permet de tirer profit, d'une part, de la réduction de complexité par l'algorithme Mmax NLMS, et d'autre part, des avantages de la structure décimée/interpolée.

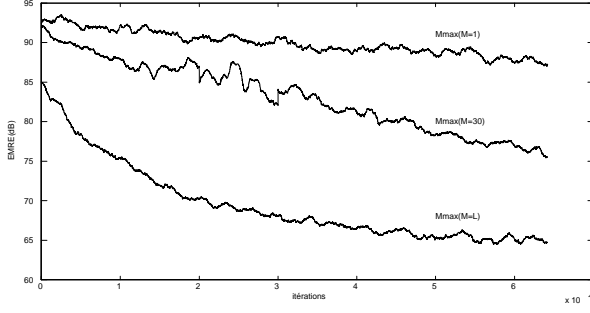


Figure 2 – EMRE(dB) pour un signal corrélé en visioconférence

3 Principe de la structure décimée/interpolée

L'idée de base est d'exploiter la redondance des coefficients de la réponse impulsionnelle du chemin d'écho dans le but d'éliminer des échantillons de la réponse impulsionnelle qui seront recréés en utilisant un schéma d'interpolation. L'avantage de la structure décimée/interpolée (déc/interp) [6] est qu'habituellement en annulation d'écho, on tronque le modèle de la réponse impulsionnelle longue, en identifiant les premiers coefficients du modèle. Généralement on identifie uniquement 300 parmi les 4000 coefficients alors qu'avec une structure déc/interp une meilleure prise en compte de l'ensemble de la réponse est possible. On a choisi dans ce papier une décimation de 1 sur 2 coefficients et une compensation du coefficient décimé par une combinaison linéaire entre le coefficient précédent et le coefficient suivant adaptés. Cette approche est aisément généralisable et selon la structure de la réponse impulsionnelle traitée, une décimation de 1 sur P est possible, associée à différents schémas d'interpolation.

Le filtre modèle est alors régi par l'équation suivante:

$$\hat{y}_k = \sum_{i=0}^{N-1} [H_k(2i).x_{k-2i} + H_k(2i+1).x_{k-2i-1}] \quad (2)$$

Les $H_k(2i)$ peuvent être adaptés avec un algorithme d'adaptation donné (par exemple le NLMS). Et les $H_k(2i+1)$ sont obtenus par interpolation.

Soit $H_k(2i+1) = \alpha.H_k(2i) + (1-\alpha).H_k(2i+2)$ où α est un coefficient réel vérifiant $0 \leq \alpha \leq 1$. La figure 3, où on considère le NLMS comme algorithme d'adaptation pour les $H_k(2i)$, montre ainsi comment les performances de la structure déc/interp avec $2N=1500$ dépassent celles de la structure tronquée à 750 coefficients, et que la courbe obtenue avec la structure déc/interp est presque confondue avec celle obtenue avec la structure tronquée à 1500 coefficients.

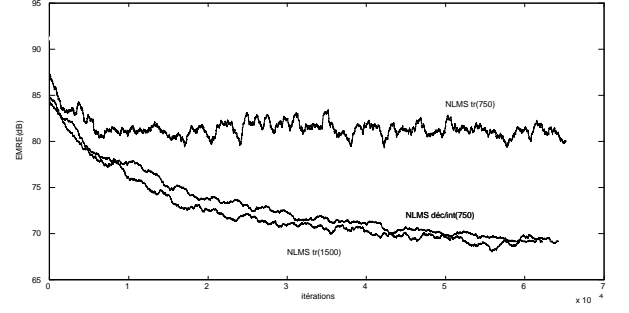


Figure 3 – Intérêt de la structure décimée/interpolée pour les RI longues

4 Application du Mmax NLMS à une structure décimée/interpolée

Le principe de l'algorithme proposé est d'appliquer l'algorithme Mmax NLMS à la structure décimée interpolée définie dans la section 3. Il faut ainsi chercher les M coefficients qui assurent le maximum du vecteur gradient. Avec la structure (2), le gradient instantané dans la direction du coefficient $2i$ (les seuls à être adaptés) à l'itération k est :

$$\begin{cases} -2.e_k.[x_k + \alpha.x_{k-1}] & \text{pour } i = 0 \\ -2.e_k.[x_{k-2i} + \alpha.x_{k-2i-1} + (1-\alpha).x_{k-2i+1}] & \text{pour } 1 \leq i \leq N-1 \\ -2.e_k.[x_{k-2N+1} + (1-\alpha).x_{k-2N+2}] & \text{pour } i = 2N-1 \end{cases}$$

On propose donc d'adapter à chaque itération k les coefficients associés aux M valeurs les plus grandes de $|x_k + \alpha.x_{k-1}|$, $|x_{k-2i} + \alpha.x_{k-2i-1} + (1-\alpha).x_{k-2i+1}|$ ($1 \leq i \leq N-1$) et $|x_{k-2N+1} + (1-\alpha).x_{k-2N+2}|$. Le filtre adaptatif est gouverné par:

$$H_{k+1}(2i) = H_k(2i) + \frac{\mu}{X_k^T X_k} . e_k . (x_{k-2i} + \alpha.x_{k-2i-1} + (1-\alpha).x_{k-2i+1})$$

si i correspond à l'un des M premiers maximums de $|x_{k-2i} + \alpha.x_{k-2i-1} + (1-\alpha).x_{k-2i+1}|$ avec $i = 0, 1, \dots, N'-1$

où $N' = N/2$.

On montre que la vitesse de convergence de l'algorithme proposé approche celle de l'algorithme NLMS appliqué à la structure décimée/interpolée. On définit pour cela comme dans [1] l'erreur a priori $e(k) = y_k - X_k^T . H_k$ et l'erreur a posteriori $e_p(k) = y_k - X_k^T . H_k$. On a

$$e_p(k) = e(k) - \sum_{j=0}^{2N-1} x_{k-j} . (H_{k+1}(j) - H_k(j)).$$

On montre facilement dans ce cas :

$$e_p(k) = e(k) . [1 - \frac{\mu}{X_k^T X_k} . \sum_{j=0}^{M-1} (x_{k-2j} + \alpha.x_{k-2j-1})]$$

$$+(1 - \alpha).x_{k-2i_j+1})^2] \quad (3)$$

où i_0, i_1, \dots, i_{M-1} correspondent aux M coefficients qui seront adaptés. Dans le cas d'application de l'algorithme NLMS d'adaptation complète à la structure déc/interp, on trouve la même expression avec $M=N$.

Ainsi lorsque $\sum_{j=0}^{M-1} (x_{k-2i_j} + \alpha.x_{k-2i_j-1} + (1 - \alpha).x_{k-2i_j+1})^2$ approche $\sum_{j=0}^{N-1} (x_{k-2i_j} + \alpha.x_{k-2i_j-1} + (1 - \alpha).x_{k-2i_j+1})^2$, la vitesse de convergence de l'algorithme d'adaptation partielle appliqué à la structure décimée/interpolée approche celle de l'algorithme d'adaptation complète appliqué à cette même structure.

Considérons la réponse impulsionnelle d'une salle de visioconférence (tronquée à 1500 coefficients) et appliquons l'algorithme proposé à une structure décimée à 750 coefficients ($2N'=1500$). L'interpolation est faite avec $\alpha = 1/2$. On compare les performances de l'algorithme Mmax dans une telle situation avec l'algorithme Mmax appliqué à une simple structure à 1500 coefficients ($M=30$ et $M=60$). On remarque d'après la figure 4 que les deux courbes obtenues avec l'algorithme proposé ($M=30$) et l'algorithme Mmax NLMS ($M=60$) sont presque confondues. Le résultat obtenu avec l'algorithme proposé est par ailleurs bien meilleur que celui obtenu avec l'algorithme Mmax NLMS pour $M=30$ et une structure classique.

La structure décimée interpolée adaptée par l'algorithme Mmax NLMS bénéficie ainsi d'une double réduction de complexité:

- celle liée à la décimation car l'interpolation reste moins coûteuse qu'une adaptation, quelque soit l'algorithme utilisé.
- celle liée à l'algorithme Mmax car il est moins complexe de rechercher les M coefficients qui assurent le maximum du gradient sur $N/2$ coefficients au lieu de N.

5 Conclusion

Les annuleurs d'écho adaptatifs classiques présents dans les systèmes de prise de son mains-libres présentent des limitations lorsqu'ils doivent traiter des réponses impulsionnelles longues comme c'est le cas dans les systèmes de visioconférence en bande élargie. L'algorithme proposé dans ce papier est basé sur l'application d'une adaptation sélective des coefficients à une structure transverse décimée/interpolée (décimation à l'ordre 2). Il devient possible d'obtenir des performances similaires à celle d'une structure classique d'adaptation avec une complexité plus faible. Les résultats obtenus avec une décimation

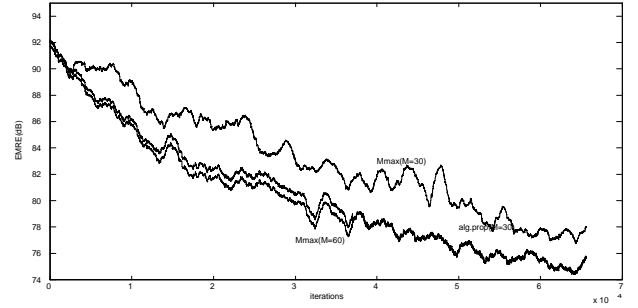


Figure 4 – EMRE(dB) pour le Mmax et le Mmax déc/interp

faible (1 sur 2) sont encourageants. Une amélioration plus conséquente devrait être obtenue avec une décimation plus forte associée à une interpolation adéquate.

REFERENCES

- [1] T. Aboulnasr and K. Mayyas, "Complexity Reduction of the NLMS Algorithm via Selective Coefficient Update", IEEE Trans. on Signal Processing, vol. 47, n°5, May 1999.
- [2] A. Gilloire, "Identification de Structures Transverses. Applications aux Réponses Acoustiques", Traitement du signal, vol.6, n°5, 1990, pp.347 – 351.
- [3] S.C. Douglas, "Analysis and Implementation of the Max-NLMS Adaptive Filter", in proc. 29th Asilomar Conf. Signals Syst. Comput., Pacific Grove, CA, Oct. 1995, vol. 1, pp. 659-663.
- [4] S.C. Douglas, "Adaptive Filters employing Partial Updates", IEEE Trans. Circuits syst. II, vol.44, pp. 209-216, march. 1997.
- [5] K. Doğançy and O. Tanrikulu, "Selective-Partial-Update NLMS and Affine Projection Algorithms for Acoustic Echo Cancellation", ICASSP 2000, Turkey.
- [6] Y. Neuvo, C.Y. Dong and S.K. Mitra, "Interpolated Finite Impulsive Response Digital Filters". IEEE Trans. on Acoust., Speech and Signal Processing, Vol. ASSP-32, pp. 563-570, june 1984.
- [7] I. Kammoun, M. Jaidane, "Exact Performances Analysis of a Selective Coefficient Adaptive Algorithm in Acoustic Echo Cancellation", ICASSP 2001, Salt Lake City.