

AMELIORATION DE L'ESTIMATION DES PARAMETRES
D'UN MODELE NON-STATIONNAIRE. APPLICATION A
DES SIGNAUX DE PAROLE BRUITES

ZERUBIA J., BARLAUD M., MENEZ J., ALENGRIN G

LASSY - UA 814 du CNRS - Université de Nice
41 Bd Napoléon III - 06041 Nice cedex -
(GRECO 69 "SARTA" du CNRS)

RESUME :

L'estimation des paramètres d'un modèle de prédiction linéaire (AR ou ARMA) d'un signal vocal bruité est un problème délicat. On observe généralement l'apparition d'un biais sur les paramètres du prédicteur.

Dans cet article, nous présentons une méthode non-stationnaire évolutive basée sur les moindres carrés. La correction du biais sur les coefficients de la partie AR du modèle ARMA est faite en estimant la puissance du bruit perturbateur. Cela conduit à résoudre un système d'équations non-linéaires. Nous proposons un algorithme qui calcule les paramètres biaisés, puis la puissance du bruit ainsi que les paramètres non-biaisés de façon récursive. Enfin, nous présentons les résultats expérimentaux obtenus avec un signal de parole réel pour un rapport S/B de 10 dB.

1. INTRODUCTION

L'estimation des paramètres d'un modèle de prédiction linéaire (AR ou ARMA) d'un signal vocal bruité est généralement biaisée [6]. De nombreuses méthodes de correction de biais ont été développées en supposant le signal de parole stationnaire par intervalle [8]. Plus récemment, des méthodes non-stationnaires adaptées au caractère lentement variable du signal de parole ont été proposées. La méthode que nous présentons dans cet article est évolutive [3],[4],[5],[7] : les coefficients AR du modèle ARMA sont des combinaisons linéaires d'une base de fonctions convenablement choisie. La méthode repose sur les moindres carrés. La correction du biais sur les coefficients AR est effectuée grâce à l'estimation de la puissance du bruit perturbateur [1]. On montre que l'estimation simultanée de cette puissance et des coefficients non-biaisés revient à résoudre un système d'équations non-linéaires. Nous proposons un algorithme de résolution récursif : nous calculons les paramètres biaisés ; puis nous estimons la puissance du bruit et les paramètres non-biaisés.

Enfin, nous présentons les résultats expérimentaux obtenus avec un signal de parole réel pour un rapport S/B égal à 10 dB

2. POSITION DU PROBLEME

Considérons un modèle ARMA évolutif dont la partie MA est à coefficients constants. Le signal non-bruité peut s'écrire :

SUMMARY :

For vocal signals, parameter-estimation in the presence of noise of an AR or ARMA model is a difficult problem. Usual techniques give biased parameters. In this paper, we describe a non-stationary method based on a modified least-squares algorithm. The bias on the AR time-varying parameters of the ARMA model is corrected by computing the variance of the disturbing noise. After some calculus, we get a system of non-linear equations. A recursive algorithm to solve the system is proposed : the first step computes the biased parameters, then the second one gives the noise variance and the unbiased parameters. At last, we show experimental results using a real speech signal with an SNR equal to 10 dB.

$$s_n = - \sum_{i=1}^p a_i(n) s_{n-i} + \sum_{j=0}^m d_j u_{n-j} \quad (1)$$

les paramètres $a_i(n)$ s'expriment à partir d'une base de fonctions $f_k(n)$:

$$a_i(n) = \sum_{k=0}^q a_{ik} f_k(n) \quad i \in [1,p] \quad (2)$$

En supposant que le bruit b_n est additif, stationnaire, blanc, gaussien et centré - $N(0, \sigma^2)$ - l'observation y_n s'écrit alors :

$$y_n = s_n + b_n \quad (3)$$

en utilisant (1) et (3), il vient :

$$y_n + \sum_{i=1}^p a_i(n) y_{n-i} = b_n + \sum_{i=1}^p a_i(n) b_{n-i} + \sum_{j=0}^m d_j u_{n-j} \quad (4)$$

Nous savons que sur l'intervalle $[1, N]$, l'estimateur des moindres carrés α du vecteur paramètre θ du modèle peut être calculé par :



$$\hat{\alpha}_N = + \left[\begin{matrix} \Sigma & Z & Z^T \\ n=1 & n-1 & n-1 \end{matrix} \right]^{-1} \left[\begin{matrix} \Sigma & Z \\ n=1 & n-1 \end{matrix} \right] y_n \quad (5)$$

avec $\theta^T = [-a_{10} \dots -a_{1q} \dots -a_{p0} \dots -a_{pq} \quad 0 \dots 0]$

$$Z^T = [y_{n-1} \quad f(n) \dots y_{n-1} \quad f(n), \quad u_n \dots u_{n-m}]$$

Cet estimateur est biaisé.

3. ALGORITHME DE CORRECTION DU BIAIS

L'algorithme que nous proposons est récursif, il est composé de deux parties :

- la première calcule les coefficients ARMA biaisés par la méthode classique des moindres carrés.

- la seconde calcule les coefficients non-biaisés ainsi que la puissance σ^2 du bruit perturbateur.

* étape 1

Pour éviter une inversion de matrice, nous utilisons les formules suivantes :

$$\hat{\alpha}_N = \hat{\alpha}_{N-1} + P_{N-1} Z_{N-1}^{-1} (y_N - Z_{N-1}^T \hat{\alpha}_{N-1}) \quad (6)$$

$$P_N = P_{N-1} - P_{N-1} Z_{N-1}^{-1} Z_{N-1}^T P_{N-1} \quad (7)$$

$$Q_N = 1 + Z_{N-1}^T P_{N-1} Z_{N-1} \quad (8)$$

avec $P_N = \left[\begin{matrix} \Sigma & Z \\ n=1 & n-1 \end{matrix} \right]^{-1}$

* étape 2

Pour corriger le biais de l'estimateur $\hat{\alpha}$, nous devons estimer la variance du bruit. Du fait de l'ergodicité du bruit, nous avons la relation suivante [8] :

$$\sigma_N^2 = \frac{R}{N + \hat{\theta}^T D \hat{\alpha}} \quad (9)$$

avec $R = \sum_{n=1}^N \epsilon(n-1)^2$

$$\epsilon(n-1) = y_n - \hat{y}_n \quad \text{et} \quad \hat{y}_n = Z_{n-1}^T \hat{\alpha}_{n-1}$$

$$D_N = \begin{bmatrix} F & & 0 \\ N & & \\ 0 & & 0 \end{bmatrix}$$

et

$$F_N = \sum_{n=1}^N \begin{bmatrix} F & 0 \\ n & \\ 0 & F \end{bmatrix}$$

où $F_n = \begin{bmatrix} f(n) \\ .0 \\ . \\ f(n) \end{bmatrix}$ [f(n)...f(n)]

Dans l'algorithme, le calcul de R est fait récursivement en utilisant la formule ci-dessous :

$$R_N = R_{N-1} + \epsilon^2(N-1) / Q_N \quad (10)$$

Enfin, le calcul de l'estimateur non-biaisé $\hat{\theta}_N$ est obtenu par [8] :

$$\hat{\theta}_N = \hat{\alpha}_N + \frac{\sigma_N^2}{N} P_N D_N \hat{\theta}_{N-1} \quad (11)$$

4. RESULTATS EXPERIMENTAUX

Nous avons tout d'abord étudié la méthode sur un signal synthétique fourni par l'ENST dans le cadre du GRECO "SARTA". Puis nous avons traité des signaux de parole réels. Nous présentons ici les résultats obtenus sur une fenêtre de 512 points. Nous avons choisi comme base de fonctions des polynômes de Legendre normalisés [2], les ordres p, q et m étant respectivement fixés à 8, 1 et 4. La figure 1 montre le signal non-bruité. Nous pouvons comparer les Densités Spectrales de Puissance (représentation temps-fréquence en pseudo 3D) du signal original (fig. 2), du signal bruité à 10 dB (fig. 3) et du signal traité par la méthode proposée (fig. 4).

5. CONCLUSION

Les premiers résultats obtenus sur un signal de parole montrent que la méthode proposée permet d'obtenir une meilleure information formantique même à un rapport signal sur bruit assez faible. Cependant quelques problèmes subsistent. On note, par exemple, l'apparition de pics parasites pour des rapports S/B inférieurs ou égaux à 15 dB. Ceci est dû à une trop grande correction du biais des paramètres du modèle. Ces effets doivent pouvoir cependant être corrigés.

BIBLIOGRAPHIE

- [1] Alengrin, Barlaud, Menez : "Unbiased parameter-estimation of non-stationary signals in noise" -IEEE ASSP - Oct. 86.
- [2] Charbonnier, Barlaud, Alengrin, Menez : "results on an AR modeling of non-stationary signals" -Signal Proc.- March 87.
- [3] Grenier : "Time-dependent ARMA modeling of non-stationary signals" - IEEE ASSP - Aug. 83.
- [4] Grenier : "Modélisation de signaux non-stationnaires" Thèse d'Etat. Univ. Paris-Sud - Orsay - 84.
- [5] Hall, Oppenheim, Willsky - "Time-varying parametric modeling of speech" - IEEE Decision Contr. Conf. - New-Orleans - 77.
- [6] Kay : "The effects of noise on the autoregressive spectral estimator". IEEE - ASSP - Oct. 79.
- [7] Omnes-Chevalier : "Analyse et restitution du signal de parole par modèles évolutifs" - Thèse de Docteur-Ingénieur - ENST-Paris-86.
- [8] Zerubia : "Modélisation d'un signal à partir d'observations bruitées, application à la réduction du bruit pour des signaux de parole". Thèse de Docteur-Ingénieur - Univ. de Nice - 86.

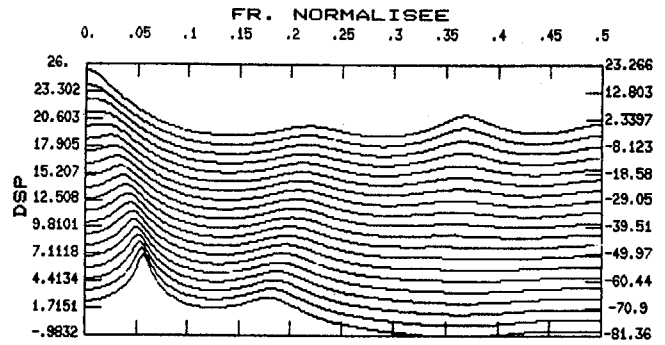


Fig. 3 : DSP du signal bruité (10 dB)

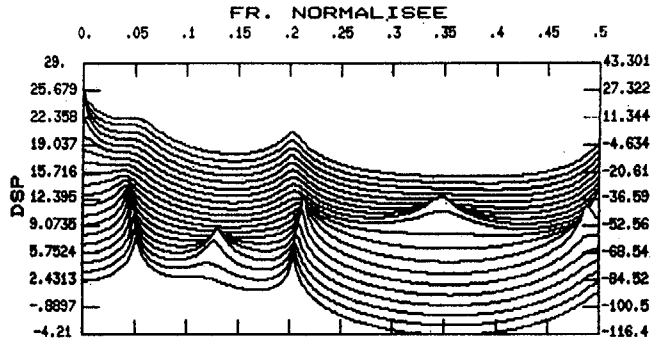


Fig. 4 : DSP du signal traité (10 dB)

SIGNAL NON BRUITE

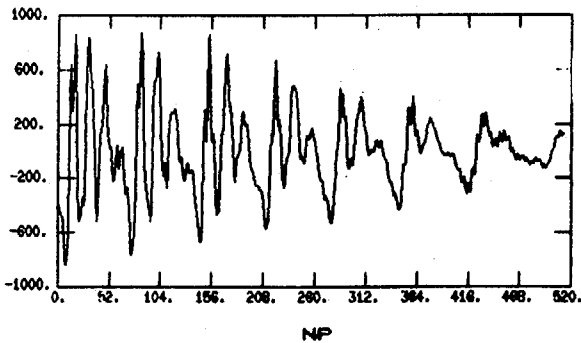


Fig. 1 : Signal de parole non bruité

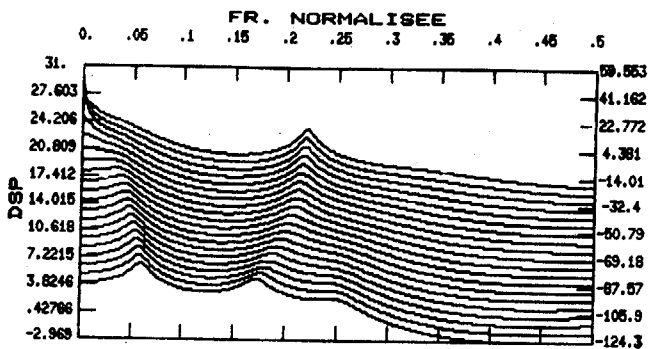


Fig. 2 : DSP du signal non bruité

