

COMPARAISON DE STRUCTURES DE CORRELOFILTRE.
APPLICATION AU DEBRUITAGE DE LA PAROLE

Gérard FAUCON*

André GILLOIRE**

* Laboratoire Traitement du Signal - IRISA - UNIVERSITE DE RENNES I - Campus de Beaulieu - 35042 RENNES CEDEX

** Département C.M.C. - C.N.E.T. LANNION A - B.P. 40 - 22301 LANNION CEDEX

RESUME

Le problème posé est celui de l'estimation d'un signal de parole perturbé par un bruit additif lorsqu'une voie référence bruit seul est disponible. Une structure dérivée du corrélofiltre classique est ici proposée. La réponse impulsionnelle du filtre attaqué par le bruit de référence préalablement blanchi et devant générer une estimée du bruit perturbateur est ainsi obtenue : on estime la fonction d'intercorrélation entre la référence blanchie et le signal bruité dont on a soustrait certaines composantes du bruit perturbateur estimées à partir des points de la réponse impulsionnelle de rang inférieur. Les performances de ce corrélofiltre dit local sont analysées sur des signaux enregistrés dans une voiture et montrent son intérêt.

1 - POSITION DU PROBLEME

On s'intéresse ici au problème de l'estimation continue d'un signal de parole S perturbé par un bruit additif B lorsqu'on dispose sur une entrée principale de l'observation $V = S+B$ et sur une entrée annexe, d'une référence bruit seul B' corrélée avec le bruit perturbateur B . Les deux bruits sont indépendants du signal S . Ce problème peut être résolu par l'utilisation d'un corrélofiltre [1] lorsque le bruit référence est blanc ou a été préalablement blanchi. Ce système corrélofiltre a été largement étudié [2,3]. Une structure voisine de ce corrélofiltre est ici présentée. Elle en diffère par le calcul des coefficients : ceux-ci sont maintenant obtenus par une estimation de la fonction d'intercorrélation entre la référence blanchie et la voie d'observation de laquelle on a soustrait certaines composantes estimées du bruit perturbateur. Ce traitement est appliqué pour le débruitage de la parole sur des enregistrements effectués dans une voiture.

Les processus S, B, B' sont discrets et aléatoires. Les instants t_k de définition de l'ensemble discret (entrées, sortie, système) sont dus à un échantillonnage périodique du temps de période θ_e :

$$X_k = X(t_k) = X(k \theta_e) .$$

SUMMARY

We consider here the estimation of a noisy speech signal when a reference noise is available. The classical correlofilter is so modified : the impulse response of the filter giving the estimated disturbing noise is obtained from the whitened reference and from the noisy signal of which some estimated components of the disturbing noise are subtracted. This correlofilter named local is studied with real signals recorded in a car as inputs.

2 - CORRELOFILTRE GLOBAL ET CORRELOFILTRE LOCAL

Dans le problème de l'estimation continue d'un signal bruité avec référence bruit seul disponible, la méthode couramment utilisée pour l'estimation du signal S est l'obtention d'une estimée du bruit perturbateur par filtrage du bruit de référence, puis de soustraire cette estimée de l'observation $V = S+B$ pour donner le signal estimé \hat{S} . Dans le cas du débruitage de la parole, les statistiques sur le signal et sur les bruits ne sont pas connues et doivent être estimées à partir des données. De plus, les bruits considérés ne sont pas stationnaires et on souhaite effectuer un traitement en ligne. Pour la suite, on se place dans le cas causal avec des traitements mis à jour à chaque arrivée d'échantillons.

Le gain du filtre optimum (au sens des moindres carrés) devant générer B à partir de B' est $g(v) = \gamma_{BB'}(v) / \gamma_{B'}(v)$. Si un blanchiment du bruit de référence est effectué, on pourra envoyer le bruit blanchi noté N à l'entrée d'un filtre de gain $h(v) = \gamma_{BN}(v) / \gamma_N(v)$. Comme le signal est indépendant des deux bruits et si l'on choisit une densité spectrale $\gamma_N(v) = 1, \forall v$, la réponse impulsionnelle du filtre est alors donnée par la fonction d'intercorrélation $\Gamma_{VN}(\tau)$ entre la voie d'observation et la référence blanchie. On obtient alors le corrélofiltre (avec blanchiment préalable) que nous appellerons ici "corrélofiltre global". Ce type de traitement a largement



été étudié dans la littérature pour le cas stationnaire. Pour un processus N discret, $\gamma_N(v) = 1, \forall v$, signifie que les échantillons du bruit N sont décorrelés et de variance unité. Aussi, dans le cas de processus non stationnaires, nous estimerons les coefficients d'un filtre de dimension r ainsi :

$$h_{\ell,j} = \frac{A_{\ell,j}}{B_{\ell,j}}, 0 \leq \ell \leq r-1 \quad (1)$$

où $A_{\ell,j} = \beta A_{\ell,j-1} + (1-\beta) N_{j-\ell} V_j \quad (2)$

$$B_{\ell,j} = \beta B_{\ell,j-1} + (1-\beta) N_{j-\ell}^2 \quad (3)$$

Le signal estimé à l'instant j est alors donné par :

$$\hat{S}_j = V_j - \sum_{\ell=0}^{r-1} h_{\ell,j} N_{j-\ell} \quad (4)$$

β est un facteur d'oubli ($0 < \beta < 1$). Pour des processus stationnaires et en prenant pour l'estimation de $h_{\ell,j}$ une fenêtre rectangulaire à horizon croissant, on sait que l'erreur quadratique moyenne d'estimation dépend de la dimension finie du filtrage, du temps d'estimation fini et de la puissance du signal. Lorsque le temps d'estimation (lié à j) tend vers l'infini, l'erreur d'estimation est uniquement due à la dimension finie du filtrage. Pour des processus non-stationnaires, on utilise une fenêtre exponentielle et l'erreur d'estimation dépend du facteur d'oubli β .

Comment peut-on en gardant la même structure de base améliorer l'estimation des coefficients du filtre ? Le coefficient $h_{\ell,j}$ est donné, dans ce qui précède, par l'estimation du point correspondant de la fonction d'intercorrélation $\Gamma_{VN}(\ell, \theta_e)$. Cette estimation est biaisée par la présence du signal et par les composantes du bruit perturbateur non liées à $N_{j-\ell}$. Or, à chaque étage du corrélifiltre, on peut retrancher à la voie d'observation un bruit calculé à partir des étages de rang inférieur ou égal et utiliser cette différence pour calculer le coefficient de l'étage d'ordre supérieur. On obtient alors ce qui est appelé ici le corrélifiltre local (figure 1), par opposition au corrélifiltre global où l'on n'utilise pas d'"observations intermédiaires". Les équations qui régissent le corrélifiltre local de dimension r représenté figure 1 sont :

$$e_{\ell+1,j} = e_{\ell,j} - h_{\ell,j} N_{j-\ell}, 0 \leq \ell \leq r-1 \quad (5)$$

avec $e_{0,j} = V_j = S_j + B_j \quad (6)$

Le signal estimé est $\hat{S}_j = e_{r,j} \quad (7)$

Les coefficients estimés $h_{\ell,j}$ sont donnés par :

$$h_{\ell,j} = \frac{C_{\ell,j}}{B_{\ell,j}} \quad (8)$$

où $C_{\ell,j} = \beta C_{\ell,j-1} + (1-\beta) N_{j-\ell} e_{\ell,j} \quad (9)$

$B_{\ell,j}$ est donné par (3).

Notons que (7) pourra s'écrire comme (4)

$$\hat{S}_j = e_{r,j} = S_j + B_j - \sum_{\ell=0}^{r-1} h_{\ell,j} N_{j-\ell}$$

mais les valeurs des $h_{\ell,j}, 1 \leq \ell \leq r-1$, seront différentes dans les deux types de corrélifiltre, seule la valeur de $h_{0,j}$ sera la même.

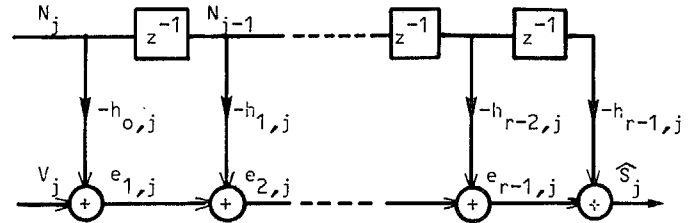


figure 1

3. STRUCTURE FILTRE EN TREILLIS + FILTRE TRANSVERSE

Les deux types de corrélifiltre décrits précédemment sont maintenant transposés dans une structure filtre en treillis + filtre transverse. Dans le problème de l'estimation d'un bruit à partir d'un autre bruit, la rapidité de convergence est augmentée en utilisant le principe d'orthogonalisation. Avant de procéder à l'estimation proprement dite, on effectue une orthogonalisation du bruit de référence. On applique ainsi le bruit de référence à un filtre en treillis pour obtenir des erreurs de prédiction rétrogrades orthogonales. Ces erreurs sont ensuite appliquées à un filtre transverse (figure 2). Cette structure filtre en treillis + filtre transverse est souvent utilisée dès lors que la contrainte sur la complexité (nombre d'opérations) le permet. Le calcul des coefficients du filtre transverse peut se faire de deux façons différentes, on obtient alors deux algorithmes transverses dont des études de la convergence ont déjà été présentées [4,5].

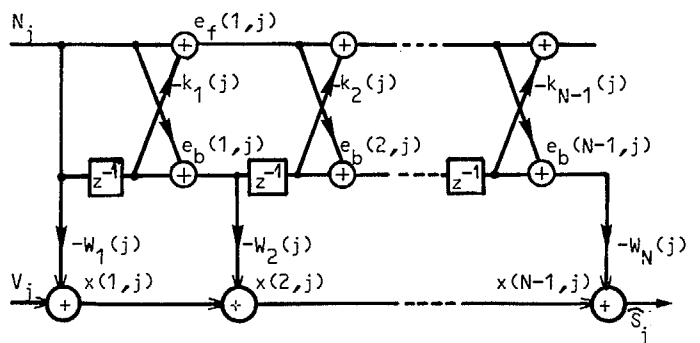


figure 2

Les résultats présentés au paragraphe suivant sur le corrélifiltre global et sur le corrélifiltre local laissent présager d'un meilleur résultat lorsque les coefficients du filtre transverse seront calculés de la même manière que dans le corrélifiltre local. Notre but ici est de déterminer si pour le problème de

débruitage de la parole qui nous intéresse, une différence de performances entre les deux algorithmes existe. Précisons d'abord l'algorithme retenu ici pour déterminer chaque coefficient de réflexion $k_n(j)$. La valeur de $k_n(j)$ sera calculée de façon récursive :

$$k_n(j) = A_{n-1}(j) / B_{n-1}(j) \quad (10)$$

$$A_{n-1}(j) = \beta A_{n-1}(j-1) + 2(1-\beta) e_f(n-1, j) \cdot e_b(n-1, j-1) \quad (11)$$

$$B_{n-1}(j) = \beta B_{n-1}(j-1) + (1-\beta) [e_b^2(n-1, j-1) + e_f^2(n-1, j)] \quad (12)$$

où $e_f(n, j)$ et $e_b(n, j)$ représentent respectivement les erreurs directes et les erreurs rétrogrades à l'instant j . L'équation régissant le filtre transverse à N coefficients est :

$$x(n, j) = x(n-1, j) - W_n(j) e_b(n-1, j) \quad 1 \leq n \leq N \quad (13)$$

$$\text{avec } x(0, j) = V_j = S_j + B_j \quad (14)$$

Le signal \hat{S}_j est donné par :

$$\hat{S}_j = x(N, j) = S_j + B_j - \sum_{n=1}^N W_n(j) e_b(n-1, j) \quad (15)$$

Les coefficients $W_n(j)$ pourront donc être déterminés de deux façons : soit à partir des intercorrélations estimées entre la voie principale V_j et chacune des erreurs rétrogrades (cas global) ; soit à partir des intercorrélations estimées entre la voie principale de laquelle on a soustrait certaines composantes de bruit et chacune des erreurs rétrogrades (cas local). Dans le cas global, on prend : $W_n(j) = C_{n-1}(j) / E_{n-1}(j)$ (16) où $C_{n-1}(j) = \beta C_{n-1}(j-1) + (1-\beta) V(j) e_b(n-1, j)$ (17) et $E_{n-1}(j) = \beta E_{n-1}(j-1) + (1-\beta) e_b^2(n-1, j)$ (18) Dans le cas local, on prend : $W_n(j) = D_{n-1}(j) / E_{n-1}(j)$ (19) où $D_{n-1}(j) = \beta D_{n-1}(j-1) + (1-\beta) x(n-1, j) e_b(n-1, j)$ (20) et $E_{n-1}(j)$ défini en (18).

4. COMPARAISON DES PERFORMANCES

4.1. Résultats sur le corrélofiltre global et sur le corrélofiltre local

Pour permettre une comparaison entre le corrélofiltre global et le corrélofiltre local, une évaluation des performances a été faite sur ordinateur pour chacun de ces deux types dans les mêmes conditions d'entrée. Dans un premier temps, le signal utile et les bruits sont simulés. Dans un deuxième temps, l'évaluation est faite sur des signaux réels enregistrés dans une voiture et échantillonnés à la fréquence de 8 kHz. Le signal et le bruit ont été enregistrés séparément de façon à superposer le signal sur des tranches de bruit choisies.

a) *Signaux simulés* : La figure 3 montre les résultats obtenus dans la situation suivante : un bruit blanc N_k (généralisé par le calculateur) est envoyé sur la voie de référence (aucun blanchiment n'est alors nécessaire), le bruit perturbateur est construit à partir du bruit blanc ainsi $B_k = N_k + 0.9 B_{k-1}$ et est ajouté au signal utile $S_k = 3 \sin(2\pi k / 5.1)$ sur la voie d'observation,

(S/B = -0.7 dB). La figure 4 illustre l'influence du blanchiment non parfait sur les performances des deux corrélofiltres. On utilise alors le bruit blanc précédent N_k légèrement filtré : $N'_k = N_k + 0.05 N'_{k-1} + 0.025 N'_{k-2}$ et N'_k est envoyé sur la voie de référence des corrélofiltres.

Les figures 3 et 4 présentent l'erreur quadratique moyenne d'estimation ϵ_k en fonction du temps. Cette erreur a été obtenue par moyennage sur 1000 simulations. Les bruits étant stationnaires, une fenêtre rectangulaire d'estimation des coefficients aurait pu être choisie. Nous avons en fait utilisé les algorithmes à fenêtre exponentielle présentés précédemment (avec β proche de 1) puisque ce sont eux qui seront utilisés sur des signaux réels. Pour les situations données, le corrélofiltre local s'avère plus performant que le corrélofiltre global. Un blanchiment non parfait est bien sûr préjudiciable aux deux corrélofiltres, l'erreur est augmentée, mais les performances obtenues restent très en faveur du corrélofiltre local.

b) *Performances sur signaux réels* : Le signal de parole est capté dans une voiture à l'arrêt par un microphone principal, le bruit de référence est capté par un second microphone dans la voiture en roulement. Ceci permet d'avoir une voie de référence dépourvue de signal utile. C'est un cas idéalisé car, dans la pratique, on recueille un signal de parole sur la référence. Le blanchiment du bruit référence B'_k peut être effectué soit par blocs soit de manière récursive. Il est réalisé ici par un filtre en treillis de dimension M dont les coefficients de réflexion sont calculés suivant les équations (10), (11), (12). L'erreur de prédiction directe à la sortie de la dernière cellule est utilisée comme référence pour le corrélofiltre. Pour les figures 5 et 6, on calcule les rapports (S/B) en entrée et en sortie sur des blocs de 256 échantillons. On note G le gain obtenu : $G(\text{dB}) = (S/B)_s - (S/B)_e$.

Pour la figure 5, le bruit perturbateur B_k est généré à partir du bruit de référence B'_k par le filtrage suivant : $B_k = B'_k + 0.9 B_{k-1}$. Pour les courbes de la figure 6, le bruit de référence et le bruit perturbateur ont été enregistrés à l'intérieur de la voiture par les deux microphones précédents. Une analyse de ces bruits (perturbateur et de référence) montre que leur énergie est prépondérante aux basses fréquences (en dessous 500 Hz) et que la cohérence entre les deux bruits est relativement bonne au dessous 500-600 Hz et faible au dessus. Ce résultat indique qu'il sera impossible de restituer parfaitement le bruit perturbateur à partir du bruit de référence. Cependant, les résultats obtenus avec les traitements étudiés ici montrent qu'on arrive à améliorer le rapport signal à bruit de façon appréciable. Le gain G obtenu est plus important avec le corrélofiltre local.



4.2. Résultats sur structure filtre treillis + filtre transverse

L'évaluation des performances pour cette structure est faite sur un signal de parole et des bruits enregistrés dans une voiture, dans les conditions utilisées pour la figure 6. Les rapports signal à bruit et le gain G définis précédemment sont calculés sur des blocs de 256 échantillons. La dimension du filtre en treillis est M , celle du filtre transverse est N . La figure 7 montre les résultats obtenus pour un jeu de paramètres donné.

4.3. Comparaison des résultats

Les différents résultats présentés montre l'amélioration apportée lorsque le calcul des coefficients du filtre d'estimation fait appel aux observations intermédiaires. Dans les courbes présentées, nous n'avons pas cherché à comparer les deux corrélofiltres (avec blanchiment préalable) avec les deux structures treillis. Or, l'utilisation du corrélofiltre nécessite un bruit de référence blanc ; en supposant que le bruit de référence est autorégressif, le blanchiment est ici réalisé par un filtre en treillis. Le tableau figure 8 compare les 4 systèmes ainsi désignés : C/G, corrélofiltre global, C/L, corrélofiltre local ; T/G, structure treillis avec filtre transverse global ; T/L structure treillis avec filtre transverse local. La valeur du gain G est obtenue par moyennage sur les 70 derniers blocs de traitements effectués sur 83 blocs (de 256 échantillons). Les meilleurs résultats sont obtenus par la structure treillis avec filtre transverse local qui ne fait aucune hypothèse sur le modèle du bruit de référence et qui utilise les estimations locales. Une analyse de l'erreur ou du gain G plus fine que celle donnée par le tableau figure 8 et faite sur chacun des blocs donne les mêmes résultats comparatifs.

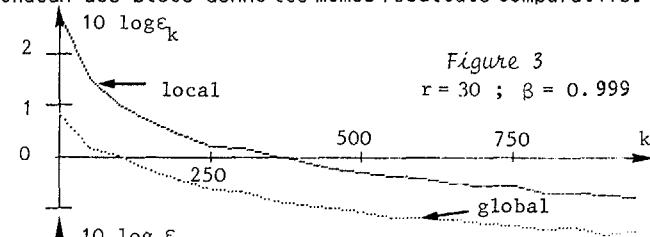


Figure 3

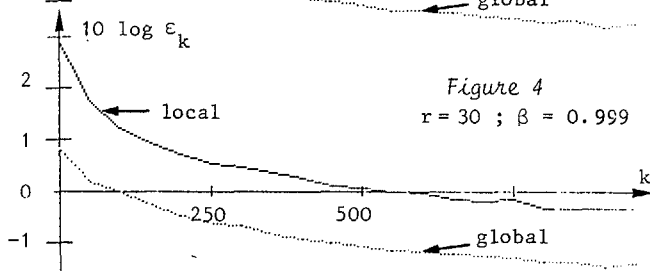
 $r = 30 ; \beta = 0.999$ 

Figure 4

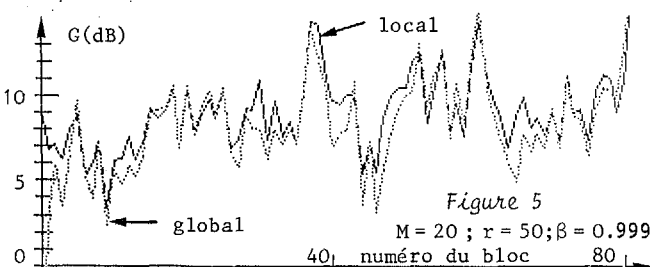
 $r = 30 ; \beta = 0.999$ 

Figure 5

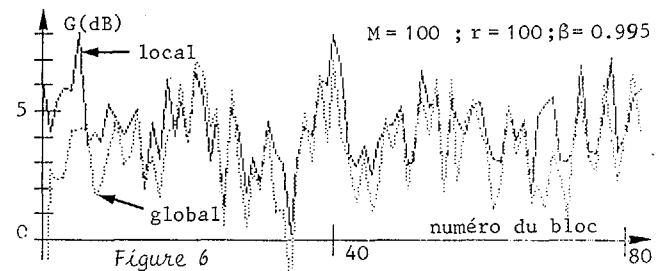
 $M = 20 ; r = 50 ; \beta = 0.999$ 

Figure 6

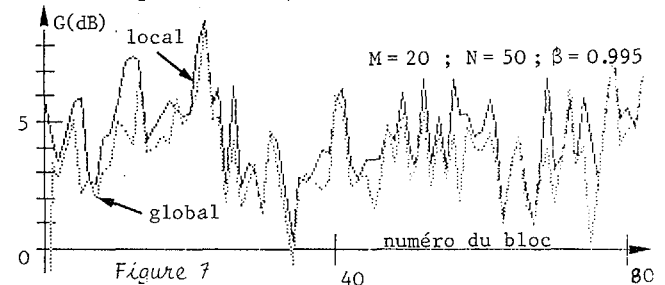


Figure 7

M	20	20	20	20	50	50	100
r=N	30	50	50	100	50	100	100
1-beta	0.005	0.001	0.005	0.005	0.005	0.005	0.005
C/G	1.40	2.50	3.13	2.99	2.57	3.70	3.57
C/L	1.50	2.58	3.43	5.48	2.77	4.97	4.27
T/G	3.74	3.11	3.73	2.97	4.18	3.99	4.38
T/L	3.90	3.37	4.46	5.76	4.50	5.74	5.80

Figure 8 (signaux réels)

5. CONCLUSION

Pour le problème de l'estimation d'un signal bruité lorsqu'on dispose d'un bruit de référence blanc, la modification du corrélofiltre classique en utilisant les estimées des intercorrélations entre la voie référence et la voie d'observation de laquelle on a soustrait certaines composantes estimées du bruit perturbateur est présentée et justifiée qualitativement. Le corrélofiltre obtenu, appelé corrélofiltre local, conduit à des résultats sensiblement meilleurs. Une étude faite sur des signaux enregistrés dans une voiture justifie l'utilisation de cette structure et sa transposition à la structure treillis+transverse pour laquelle des résultats sont également présentés.

BIBLIOGRAPHIE

- [1] H. MERMOZ : "Modularité" du traitement adaptatif d'antenne. Annales des Télécommunications, 29, n°1-2, 1974.
- [2] W. KOFMAN, A. SILVENT, J. LIENARD : Etude théorique et expérimentale du système corrélofiltre. Annales des Télécommunications, 37, n°3-4, 1982.
- [3] G. FAUCON : Influence of the variance of a white-noise used as reference in a correlofilter. Accepted in Signal Processing (à paraître, Juin 1987).
- [4] M. HONIG : Convergence models for lattice joint process estimators and least squares algorithms. IEEE Transactions on acoustics, speech and signal processing. Vol. ASSP-31, n°2, April 1983.
- [5] M. HONIG, D. MESSERSCHMITT : Convergence properties of an adaptive digital lattice filter. IEEE Transactions on acoustics, speech and signal processing. Vol. ASSP-29, n°3, June 1981.

Cette étude a été financée par le C.N.E.T. LANNION A