

Le lissage H_∞ : une alternative au filtrage de Kalman pour le rehaussement du signal de parole ?

D. Labarre¹, E. Grivel¹, N. Christov², M. Najim¹

¹ Equipe Signal & Image, UMR LAPS 5131, ENSEIRB BP 99, F-33402 Talence, France

Email : {david.labarre ; eric.grivel ; mohamed.najim}@laps.u-bordeaux1.fr

² Laboratoire d'Automatique, Génie Informatique et Signal, USTL, Cité Scientifique, 59655 Villeneuve d'Ascq, France

Email : nicolai.christov@univ-lille1.fr

Résumé – La mise en oeuvre d'un filtre de Kalman repose sur des hypothèses fortes sur le processus générateur et le bruit d'observation du système. En effet, ils doivent être tous les deux blancs et gaussiens et leurs variances connues a priori. De plus, les incertitudes liées à la modélisation du système ne sont pas prises en compte dans la pratique. Pour ces deux raisons, nous proposons d'étudier l'apport des techniques d'estimation H_∞ dans le contexte du rehaussement du signal de parole fondé sur un modèle AutoRégressif (AR). Ces approches, qui consistent à minimiser les effets des perturbations sur l'erreur d'estimation du signal dans le pire cas, sont a priori attrayantes car elles ne nécessitent aucune hypothèse statistique ni sur le processus générateur ni sur le bruit d'observation. En outre, comme la qualité perceptuelle du signal rehaussé dépend de l'étape d'estimation des paramètres AR, nous proposons ici une estimation conjointe du signal de parole et des paramètres AR. Plus particulièrement, un lissage H_∞ de type retard fixe fournit une estimation du signal de parole à partir de la dernière estimation des paramètres AR, alors que les paramètres AR sont mis à jour avec un filtrage H_∞ à partir de la dernière estimation lissée du signal de parole.

Abstract – In real cases, the use of a Kalman filter for signal processing may be questionable. Indeed, it does not take into account modelling uncertainties and requires a priori system information difficult to be estimated. In this paper, our purpose is to evaluate how relevant the H_∞ filters are in the framework of speech processing, where the driving process for instance is usually unknown. Indeed, these H_∞ techniques, which consist in minimizing the worst possible effects of the disturbances on the estimation error, seem appealing because no statistical information on the driving process and the additive observation noise of the system are required. However, the auditory quality of the enhanced speech still depends on the way the AR parameters are estimated because they set the spectral features of the signal to be estimated. Therefore, we propose a new method based on a cross-coupled algorithm involving two estimators running in parallel. A fixed-lag H_∞ smoothing provides an estimate of the signal, while a H_∞ filter updates the AR parameters. At time k , the signal is estimated from the latest updates of the AR parameters, while the AR parameters are estimated from the latest enhanced version of the signal.

1. Introduction

Dans de nombreuses applications du traitement du signal comme le traitement de la parole ou l'égalisation, des méthodes paramétriques fondées sur un modèle *a priori* du système sont souvent utilisées. Parmi les techniques d'estimation, le filtre de Kalman [1] présente l'avantage d'être récursif et optimal au sens des moindres carrés. Néanmoins, la mise en oeuvre de ce filtre repose sur des hypothèses fortes sur le processus générateur et le bruit d'observation du système. En effet, ils doivent être tous les deux blancs et gaussiens et leurs variances connues *a priori*. De plus, avec cette approche, les incertitudes liées à la modélisation du système ne sont pas prises en compte.

Pour pallier ces inconvénients, nous proposons d'étudier des approches alternatives fondées sur une estimation de type H_∞ . Dans ce cas, l'objectif est de minimiser les effets des perturbations sur l'erreur d'estimation du signal dans le pire cas. Cette approche est à première vue séduisante car aucune hypothèse statistique sur la représentation du système n'est nécessaire.

Bien que la théorie H_∞ ait été particulièrement étudiée dans le domaine de l'automatique, les applications en traitement du signal restent peu nombreuses. Ainsi, dans [8], Sayyarodsari *et al.* traitent le problème de l'annulation de bruit. Un filtre H_∞ est mis en oeuvre pour mettre à jour les paramètres du système, plutôt qu'un algorithme LMS

habituellement utilisé. Une adaptation plus rapide et une estimation plus lisse des paramètres sont alors obtenues. Les techniques d'optimisation H_∞ ont été également employées pour l'égalisation de canaux de communications numériques [3] et la conception de bancs de filtres [10]. Dans ces deux cas, les auteurs concluent que les approches H_2 et H_∞ donnent des résultats similaires en moyenne.

Ici, nous nous focalisons sur le problème du rehaussement d'un signal de parole s_k perturbé par un bruit additif n_k , à partir d'une unique séquence d'observations y_k définie comme suit :

$$y_k = s_k + n_k. \quad (1)$$

Le signal de parole s_k est modélisé par le processus AutoRégressif (AR) suivant :

$$s_k = -\sum_{i=1}^p a_i s_{k-i} + \alpha_k \quad (2)$$

où p désigne l'ordre du modèle, $\underline{\theta} = [a_1 \ \dots \ a_p]^T$ sont les paramètres AR et le processus générateur α_k est blanc, gaussien, de moyenne nulle et de variance σ_α^2 .

Dans [6], nous avons initié notre étude sur la pertinence du filtrage H_∞ en considérant deux protocoles de simulation.

Dans le premier, les paramètres du modèle sont estimés à partir du signal de parole non bruité, ce qui permet de caractériser le comportement du filtre H_∞ face aux incertitudes de modélisation. Le second consiste à estimer les paramètres AR à partir des observations bruitées afin d'analyser le comportement du filtre H_∞ vis-à-vis des erreurs d'estimation sur les paramètres AR. Les résultats obtenus nous ont permis de conclure que l'estimation des paramètres AR reste une étape clé dans la procédure de rehaussement car elle conditionne les propriétés spectrales du signal rehaussé.

Dès lors, l'approche proposée par Shen *et al.* dans [9], et déclinée dans le contexte des communications mobiles [2], est mise en défaut. En effet, pour estimer conjointement un signal et les paramètres du modèle AR associé à partir des observations bruitées, les auteurs utilisent deux filtres H_∞ opérant en série. Les paramètres AR sont tout d'abord estimés à partir des observations bruitées par filtrage H_∞ . Puis, ils sont utilisés dans le second filtrage pour retrouver les données. Bien que les auteurs justifient leur démarche par le manque d'information *a priori*, ils négligent l'étape d'estimation des paramètres AR.

Pour pallier ce problème, nous proposons ici d'utiliser deux estimateurs récursifs qui opèrent en parallèle et interagissent comme suit :

- Un lissage H_∞ de type retard fixe (*fixed-lag* en anglais) permet d'estimer le signal de parole en utilisant la dernière valeur de l'estimation des paramètres AR.
- Un filtrage H_∞ permet de mettre à jour l'estimation des paramètres AR en utilisant la dernière estimation lissée du signal de parole. Cf. Fig. 1.

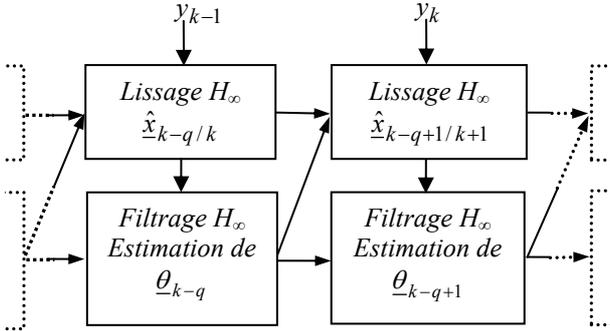


FIG. 1 : principe de la méthode proposée

L'article est organisé comme suit : la partie 2 traite du filtrage H_∞ , la partie 3 de l'estimation des paramètres AR et la partie 4 de l'ajustement des matrices de poids intervenant dans les équations du filtrage et du lissage H_∞ . La dernière partie présente des résultats de simulations sur l'estimation de paramètres AR et le rehaussement du signal de parole.

2. Filtrage et lissage H_∞

Outre les approches polynomiales [4], les solutions au problème du filtrage H_∞ actuellement les plus étudiées sont fondées sur une Représentation dans l'Espace d'Etat (REE) de la forme suivante :

$$\begin{cases} \underline{x}_k = \Phi_k \underline{x}_{k-1} + \Gamma \alpha_k \\ y_k = H \underline{x}_k + n_k \\ s_k = L \underline{x}_k \end{cases} \quad (3)$$

où \underline{x}_k désigne le vecteur d'état, Φ_k la matrice de transition, Γ le vecteur d'entrée, H le vecteur d'observation et L un vecteur permettant de se focaliser sur une combinaison linéaire particulière des variables d'état, ici s_k .

Ces approches ont l'avantage de conduire à une implantation plus simple du filtre comparée aux approches polynomiales.

Etant donné un niveau d'atténuation γ , l'estimation du signal \hat{s}_k doit satisfaire le critère suivant :

$$J_\infty = \sup_{\alpha_k, n_k} \frac{\sum_{k=0}^{N-1} \|s_k - \hat{s}_k\|_2^2}{\sum_{k=1}^N [\|\alpha_k\|_2^2 + \|n_k\|_2^2]} < \gamma^2 \quad (4)$$

où N est le nombre d'observations y_k disponibles.

Nous nous intéressons ici à une solution fondée sur une équation de Riccati [5]. Dans ce cas, l'estimation \hat{s}_k vérifiant la condition (4) existe si et seulement si :

$$(P_{k+1/k})^{-1} + H^T H - \gamma^{-2} L^T L > 0 \quad (5)$$

où la matrice $P_{k+1/k}$ satisfait la récursion de Riccati suivante :

$$P_{k+1/k} = \Phi_k P_{k/k} \Phi_k^T + \Gamma Q \Gamma^T \quad (6a)$$

$$P_{k/k} = P_{k/k-1} - P_{k/k-1} \begin{bmatrix} H^T & L^T \end{bmatrix} T^{-1} \begin{bmatrix} H \\ L \end{bmatrix} P_{k/k-1} \quad (6b)$$

où $T = \begin{bmatrix} R & 0 \\ 0 & -\gamma^2 \end{bmatrix} + \begin{bmatrix} H \\ L \end{bmatrix} P_{k/k-1} \begin{bmatrix} H^T & L^T \end{bmatrix}$ et Q et R sont des matrices de pondération permettant de régler le comportement du filtre H_∞ .

Les estimations *a posteriori* du vecteur d'état et du signal sont alors remises à jour comme suit :

$$\hat{s}_k = L \hat{\underline{x}}_{k/k} \quad (7a)$$

$$\hat{\underline{x}}_{k/k} = \hat{\underline{x}}_{k/k-1} + K_k \nu_k \quad (7b)$$

$$\hat{\underline{x}}_{k/k-1} = \Phi_k \hat{\underline{x}}_{k-1/k-1} \quad (7c)$$

où $\hat{\underline{x}}_{k/l}$ désigne l'estimation de \underline{x}_k étant données l observations, $K_k = P_{k/k-1} H (R + H P_{k/k-1} H^T)^{-1}$ le gain du filtre et $\nu_k = y_k - H \hat{\underline{x}}_{k/k-1}$ le processus d'innovation.

Un lissage H_∞ à retard fixe consiste alors à estimer le signal à l'instant $k-q$ étant données k observations, où q définit le retard. Dans la suite, on supposera que $q \geq p$. En construisant le vecteur d'état comme suit :

$$\underline{x}_k = [s_k \cdots s_{k-q}]^T, \quad (8)$$

les matrices de la REE (3) sont les suivantes :

$$H = \Gamma^T = [1 \ 0 \ \cdots \ 0],$$

$$L = [0 \ \cdots \ 0 \ 1]$$

et

$$\Phi_k = \begin{bmatrix} -a_1 & \cdots & -a_p & 0 & \cdots & 0 \\ 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 & 0 \end{bmatrix}.$$

Le filtrage H_∞ défini par les équations (6)-(7) permet d'obtenir les versions filtrées et lissées du signal de parole de la manière suivante :

$$\hat{s}_{k-q/k-q} = H \hat{x}_{k-q/k-q} \quad (9)$$

$$\hat{s}_{k-q/k} = L \hat{x}_{k/k} \quad (10)$$

Remarques :

1. Les filtres de Kalman et H_∞ ont la même structure. De plus, lorsque $\gamma \rightarrow \infty$, les deux filtres coïncident.
2. Le coût calculatoire du filtre H_∞ est sensiblement plus élevé que celui du filtre de Kalman.
3. Les approches fondées sur la théorie des jeux [11] peuvent être considérées comme un problème d'estimation H_∞ car elles traitent d'un problème de type minimax.

Comme les paramètres AR ne sont pas disponibles, nous proposons maintenant de les estimer avec un second filtre récursif appliqué à la version lissée $\hat{s}_{k-q/k}$ du signal.

3. Estimation des paramètres AR

Sur une trame d'analyse, le signal est supposé stationnaire et les paramètres AR sont constants :

$$\underline{\theta}_{k-q} = \underline{\theta}_{k-q-1}. \quad (11)$$

On exprime alors l'estimation lissée du signal $\hat{s}_{k-q/k}$ en fonction des paramètres AR $\underline{\theta}_{k-q}$. Etant donnée (7), si $\{k-i\}_{i=0,\dots,q-1}$ observations sont disponibles, on montre aisément que :

$$\hat{s}_{k-q/k} = \hat{s}_{k-q/k-q} + \sum_{i=1}^q K_{k-q+i}^{i+1} \nu_{k-q+i}. \quad (12)$$

où K_l^m désigne la $m^{\text{ième}}$ composante du vecteur gain du filtre à l'instant l . Or, d'après (7) et (9), la version filtrée du signal de parole à l'instant $k-q$ satisfait la relation suivante :

$$\begin{aligned} \hat{s}_{k-q/k-q} &= H \left(\Phi_{k-q} \hat{x}_{k-q-1/k-q-1} + K_{k-q} \nu_{k-q} \right) \\ &= - \left(\hat{x}_{k-q-1/k-q-1}^{1:p} \right)^T \underline{\theta}_{k-q} + K_{k-q}^1 \nu_{k-q} \end{aligned} \quad (13)$$

où $\hat{x}_{k/k}^{1:p}$ correspond aux p premières composantes de $\hat{x}_{k/k}$. Ainsi, en substituant (12) dans (13), on obtient :

$$\hat{s}_{k-q/k} = - \left(\hat{x}_{k-q-1/k-q-1}^{1:p} \right)^T \underline{\theta}_{k-q} + \nu_k^\theta, \quad (14)$$

avec $\nu_k^\theta = \sum_{i=0}^q K_{k-q+i}^{i+1} \nu_{k-q+i}$.

Les paramètres AR sont alors estimés avec un filtre H_∞ fondé sur la REE (11)-(14). Ce filtre met en jeu deux matrices de pondération R^θ et Q^θ et un niveau d'atténuation γ^θ .

Les niveaux d'atténuation des bruits γ et γ^θ peuvent être par exemple fixés à l'aide d'une procédure d'optimisation itérative. L'ajustement des matrices de pondération Q , R , Q^θ et R^θ est plus délicat et nous proposons d'étudier ce dernier point dans le paragraphe 4.

4. Réglage des matrices de poids

Les matrices de pondération Q , R , Q^θ et R^θ permettent d'ajuster les comportements des estimateurs H_∞ . Elles jouent le même rôle que les variances du processus générateur et du bruit d'observation respectivement dans le cadre H_2 , comme le mentionnent Shen *et al.* dans [9]. En outre, la matrice $P_{k+1/k}$ peut être interprétée comme une borne supérieure de la matrice de covariance de l'erreur dans le cadre H_2 [11].

Nous proposons donc d'adapter ces matrices en raisonnant par analogie au cadre H_2 . Nous supposons tout d'abord que les caractéristiques du bruit additif d'observation n_k varient lentement au cours du temps et l'ajustement de R est réalisé pendant les trames de silence.

La matrice Q est réglée selon les résultats présentés dans [7] de la manière suivante :

$$Q_k = \frac{k-1}{k} Q_{k-1} + \frac{1}{k} D M_k D^T. \quad (15)$$

avec $M_k = P_{k/k} - \Phi_{k-1} P_{k-1/k-1} (\Phi_{k-1})^T + K_k \nu_k^2 K_k^T$ et $D = (\Gamma^T \Gamma) \Gamma^T$.

Dans le filtre de Kalman [1], l'innovation ν_k est un bruit blanc de variance $H P_{k/k-1} H^T + R$. Ainsi, nous proposons d'ajuster la matrice R^θ de la manière suivante :

$$R_k^\theta = \sum_{i=0}^q \left\{ \left(K_{k-q+i}^{i+1} \right)^2 \left(H P_{k-q+i/k-q+i-1} H^T + R \right) \right\}. \quad (16)$$

Enfin, d'après l'équation (11), la valeur de Q^θ est fixée à 0.

5. Résultats expérimentaux

5.1 Estimation de paramètres AR

Dans un premier temps, nous étudions les performances de l'approche proposée pour l'estimation de paramètres AR, car cette étape conditionne le rehaussement du signal de parole.

Lorsque le processus AR est perturbé par un bruit additif, Shen *et al.* [9] appliquent directement un filtre H_∞ sur les observations bruitées pour estimer les paramètres AR. Nous réalisons donc une étude comparative entre le filtre de Kalman, le filtre H_∞ et la méthode proposée. Le protocole de l'expérience est le suivant :

512 échantillons d'un processus AR d'ordre 6 sont générés selon les pôles $p_{1,2} = 0.98e^{\pm j0.1\pi}$, $p_{3,4} = 0.97e^{\pm j0.3\pi}$ et $p_{5,6} = 0.8e^{\pm j0.84\pi}$. Ils sont ensuite perturbés par un bruit

blanc gaussien à moyenne nulle avec un Rapport Signal à Bruit (RSB) de 10 dB. 100 réalisations du bruit sont utilisées.

Les résultats du tableau 1 montrent que l'utilisation directe du filtre de Kalman et du filtre H_∞ sur des observations bruitées conduisent à des estimations on-line biaisées et très similaires. La méthode proposée permet de réduire significativement les erreurs d'estimation des paramètres AR.

TAB. 1 : Estimation de paramètres AR à partir d'observations bruitées avec différentes méthodes ; RSB = 10 dB.

Paramètres	Valeur désirée	Filtre de Kalman	Filtre H_∞ [9]	Méthode proposée
σ_α^2	1	10,11	10,23	3,12
a_1	-1,60	-0,99	-0,99	-1,45
a_2	0,45	0,03	0,04	0,35
a_3	0,87	0,31	0,33	0,68
a_4	-0,51	0,003	-0,01	-0,18
a_5	-0,56	-0,22	-0,23	-0,71
a_6	0,58	0,17	0,18	0,57

5.2 Rehaussement du signal de parole

Le signal /Waziwaza/, échantillonné à 16 kHz, est perturbé soit par un bruit blanc gaussien, soit par un bruit coloré à Moyenne Ajustée (MA) large bande. Nous avons mené une étude comparative entre la méthode proposée et celle présentée dans [9].

Même si le filtrage H_∞ permet de réduire l'influence des incertitudes du modèle sur l'estimation du signal de parole, l'estimation des paramètres AR reste une étape cruciale. Notre méthode a l'avantage d'une part de limiter les incertitudes d'estimation sur les paramètres comparées à l'approche présentée dans [9], et d'autre part de lisser le signal rehaussé, atténuant ainsi le bruit résiduel.

TAB. 2 : Gains du RSB : cas d'un bruit blanc gaussien

Bruit blanc gaussien	RSB d'entrée	5	10	15
	Méthode [9]		4.4	4.0
Méthode proposée		8.6	7.8	5.2

TAB. 3 : Gains du RSB : cas d'un bruit coloré

Bruit coloré de type MA	RSB d'entrée	5	10	15
	Méthode [9]		6.1	5.5
Méthode proposée		6.4	5.8	3.9

6. Conclusion

Dans cet article, notre objectif est d'évaluer la pertinence de l'estimation H_∞ pour le rehaussement du signal de parole fondé sur un modèle AR. Dans ce cadre, aucune hypothèse statistique n'est nécessaire pour caractériser le processus générateur et le bruit additif. Suite à l'étude menée dans [6], nous proposons une méthode de débruitage fondée sur deux estimateurs H_∞ en parallèle. Un lissage de type retard fixe utilise la dernière estimation des paramètres pour estimer le signal de parole, tandis qu'un filtrage H_∞ utilise la dernière version lissée du signal pour mettre à jour les paramètres AR. Cette approche a l'avantage de réduire les erreurs d'estimation sur les paramètres. Nous avons ainsi réalisé une étude comparative sur des données AR synthétiques qui montre la pertinence de l'approche proposée. Dans le contexte du rehaussement du signal de parole, l'utilisation d'un lissage permet en outre de réduire le bruit résiduel. Ainsi, les résultats expérimentaux d'une comparaison avec la seule méthode H_∞ référencée à notre connaissance confirment l'amélioration de la qualité auditive du signal de parole rehaussé obtenu avec notre approche.

Références

- [1] B. D. O. Anderson, J. B. Moore, *Optimal filtering*, Prentice Hall, N.Y., 1979.
- [2] J. Cai, X. Shen, J. W. Mark, *Robust Channel Estimation for OFDM Wireless Communication Systems – An H_∞ Approach*, IEEE Trans. on Wireless Communications, vol. 3, n°6, November 2004, pp.2060-2071.
- [3] A.T. Erdogan, B. Hassibi, T. Kailath, *On linear H_∞ equalization of communication channels*, IEEE Trans. on SP, vol. 48, n°11, November 2000, pp. 3227-3231.
- [4] M. J. Grimble, A. E. Sayed, *Solution of the H_∞ Optimal Linear Filtering Problem for Discrete-Time Systems*, IEEE Trans. on ASSP, vol. 38, n°7, July 1990, pp. 1092-1104.
- [5] B. Hassibi, A. H. Sayed, T. Kailath, *Indefinite-Quadratic Estimation and Control, A unified Approach to H_2 and H_∞ Theories*, SIAM edition, 1999.
- [6] D. Labarre, E. Grivel, M. Najim, *Relevance of H_∞ filtering for speech enhancement*, accepted at ICASSP'05.
- [7] D. Labarre, E. Grivel, M. Najim, E. Todini, *Two-Kalman Filters Based Instrumental Variable Techniques for Speech Enhancement*, Proc. IEEE MMSP, 2004, Siena, Italy.
- [8] B. Sayyarodsari, B. Hassibi, J. How, A. Carrier, *An H_∞ -optimal alternative to the FxLMS algorithm*, Proc. ACC, 1998, pp.1116-1121.
- [9] X. Shen, L. Deng, *A Dynamic System Approach to Speech Enhancement Using the H_∞ Filtering Algorithm*, IEEE Trans. on SAP, vol. 7, n°4, July 1999, pp. 391-399.
- [10] H. Vikalo, B. Hassibi, T. Kailath, *On H_∞ optimal signal reconstruction in noisy filter banks*, Proc. ICASSP, 1999, pp. 1493-1496.
- [11] I. Yaesh, U. Shaked, *Game Theory Approach to Optimal Linear State Estimation and Its Relation to the Minimum H_∞ -Norm Estimation*, IEEE Trans. on AC, vol. 37, n°6, June 1992, pp. 828-831.