Égaliseurs neuronaux rapides

Cyril Iskander, étudiant 2e cycle

Paul Fortier, directeur de recherche

H.T. Huynh, codirecteur de recherche

Abstract: Linear equalizers are known to perform poorly on highly dispersive or non-linear channels. Consequently, attempts have been made to incorporate some form of non-linearity in equalizers, enabling them to cope with the aforementioned situations. Recently, there has been rapid progress in the application of neural networks to equalization, however their slow speed of convergence and high computational complexity remain two burdens which need to be worked upon. This report presents some alternatives to the fast and efficient training of neural equalizers, using approaches derived from the theory of adaptive signal processing. Performances of these modern filters are shown to outperform conventional structures.

Résumé: La performance limitée des égaliseurs linéaires faces à des canaux fortement dispersifs ou non-linéaires a encouragé le développement de nouvelles structures non-linéaires. Des égaliseurs basés sur des réseaux de neurones ont été proposés, cependant la lenteur de la convergence et la complexité de calcul demeurent deux fardeaux importants pour une implémentation en temps réel. Le texte suivant résume des approches qui ont été prises pour permettre à un égaliseur neuronal de surpasser les performances des égaliseurs conventionnels.

Introduction: nécessité de l’égalisation

Le phénomène d’interférence intersymbole (ISI) dégrade la performance des systèmes de communication à travers des canaux dispersifs. On cherche à corriger ce phénomène par des techniques dites d’*égalisation* [1]. Il s’agit en principe d’estimer la réponse impulsionnelle du canal de façon adaptative. Récemment, les réseaux de neurones artificiels deviennent un outil de travail dans le domaine du traitement du signal. Des chercheurs commencent à l’exploiter dans le domaine de l’égalisation [2].

L’égalisation par réseaux neuronaux

Les réseaux de neurones qui ont été privilégiés à ce jour pour l’égalisation sont le *perceptron polynômial* (PP) [3], le *réseau à fonction de base radiale* (*Radial-Basis Function network,* ou RBF) [4] et le *perceptron multicouche* (PMC). Une étude sur ces trois types d'égaliseurs a donné lieu aux constatations suivantes:

- le PP possède une structure simple, mais le nombre de coefficients (et donc la complexité de l'algorithme) croît exponentiellement avec le nombre d'entrées. Ainsi, il est difficile de compenser pour des réponses impulsionnelles longues;
- le RBF ne possède que 2 couches, la deuxième étant linéaire. L'algorithme d'entraînement est donc simple. Cependant, le nombre de neurones nécessaires dans la 1ère couche croît exponentiellement avec la longueur de la réponse impulsionnelle du canal, qui de plus est *a priori* inconnue. Afin de déterminer un nombre de neurones suffisant, il est nécessaire d'incorporer aux algorithmes de base des procédures d'entraînement plus élaborées.
- le PMC possède la structure et les algorithmes d'entraînement les plus complexes. Cependant, avec 3 couches et les nombres de neurones par couche fixés à l'avance, le PMC est capable d'égaliser un large éventail de canaux.

On a décidé de se concentrer sur l'étude du PMC, plus robuste que le PP et le RBF.

L'égaliseur PMC

Parmi les caractéristiques requises des égaliseurs, les suivantes ont prépondérance: rapidité de convergence durant le régime transitoire, efficacité de poursuite durant le régime permanent, possibilité d'implémenter l'algorithme en temps réel, robustesse face à la précision finie des calculateurs, pour une implémentation physique. Des solutions possibles ont été développées pour chacun de ces aspects.

Rapidité de convergence: L'algorithme d'entraînement du PMC le plus répandu est la *back-propagation* (BP), qui est en fait une version généralisée du LMS appliqué à des structures non-linéaires. L'algorithme BP hérite donc de la lenteur de convergence du LMS, due à l'estimation stochastique du gradient. Cette lenteur est accentuée par le fait que la surface d'erreurs du PMC est non-convexe et hautement irrégulière, à cause des fonctions d'activation non-linéaires des neurones. Afin d'accélérer la convergence, on propose deux approches: l'ajustement des poids par des algorithmes utilisant le gradient déterministe (algorithmes de type Kalman) et le prétraitement des entrées du PMC par des algorithmes blanchisseurs. Les algorithmes de type Kalman que l'on a implémentés sont le EKF (*Extended Kalman Filter*) et le RLS (*Recursive Least Squares*); le RLS est moins performant que le EKF, mais sa complexité est moindre. Les algorithmes utilisés pour blanchir les entrées du PMC sont le GAL (*Gradient Adaptive Lattice*), le LSL (*Least-Squares Lattice*) et le RMGS (*Recursive Modified Gram-Schmidt*). Les deux premiers sont applicables seulement à la couche d'entrée du PMC, alors que le troisième peut être incorporé à chaque couche, mais est de complexité supérieure. Ces deux approches sont semblables du fait qu'elles visent à blanchir les entrées de l'égaliseur et donc le rendre insensible à un mauvais conditionnement de la matrice d'autocorrélation des entrées, ce qui est le cas dans un canal hautement dispersif.

Efficacité de poursuite: Les avis sont partagés sur le choix du LMS ou du RLS comme meilleur algorithme de poursuite dans un environnement non-stationnaire. Leurs performances varient selon le type de canal rencontré et les conditions de SNR (*Signal-to-Noise Ratio*). Les mêmes disparités sont présentes dans l'extension de ces algorithmes au PMC. Il existe cependant certains moyens d'améliorer la capacité de pistage du PMC: par exemple, en permettant les taux d'apprentissage de chaque couche (et même de chaque neurone) de s'adapter aux variations de la réponse impulsionnelle.

Complexité de l'algorithme: Les algorithmes de type Kalman proposés plus tôt ont une complexité qui les rendent difficilement implémentables en temps réel. Certaines versions rapides de ces algorithmes, tels le FRLS (*Fast RLS*) et le FTF (*Fast Transversal Filter*) peuvent être utilisés, mais uniquement au niveau de la couche d'entrée. Une diminution (moins importante) de la complexité est également possible en utilisant des formes factorisées de l'algorithme Kalman.

Stabilité numérique: Les algorithmes de type Kalman présentent un danger de divergence causée par les erreurs d'arrondis et de troncature dans une machine à précision finie. Des formes factorisées de l'algorithme Kalman, comme le SR-RLS (*Square-Root RLS*) sont plus robustes aux erreurs numériques. Quant aux algorithmes rapides FRLS et FTF, il est nécessaire de les remplacer par des versions stabilisées telles le CFRLS (*Covariance Fast RLS*) et le SFTF (*Stabilized Fast Transversal Filter*).

Résultats

Les algorithmes précédents sont testés sur un canal de Rummler. La figure 1 illustre des courbes de convergence obtenues avec un canal de Rummler à phase non-minimale, de réponse impulsionnelle: $h(k) = 0.3482\delta(k) + 0.3482\delta(k-1) + 0.8704\delta(k-2) + 0.3482\delta(k-3)$.

On utilise une modulation QAM et 16 entrées transversales pour chaque égaliseur, soit le LMS, le PMC entraîné par BP et le PMC entraîné par RLS. Les égaliseurs neuronaux sont plus lents à converger que les algorithmes traditionnels, mais l'erreur résiduelle à la convergence est nettement inférieure. Des expériences sont également en cours pour tester les performances des algorithmes sur un canal intérieur, avec variation temporelle de la réponse impulsionnelle et affaiblissement de Rice [5]. Les figures 2 et 3 représentent les courbes de convergence (non moyennées) à nouveau pour le LMS et le PMC-EKF.

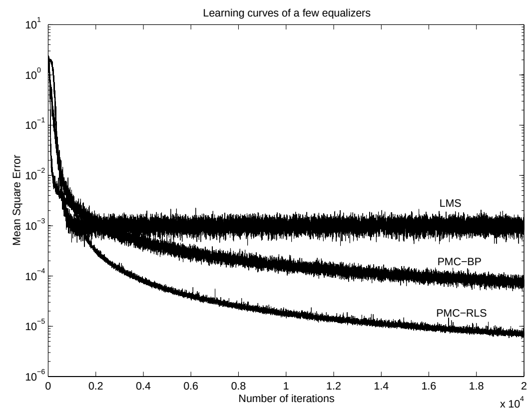

Figure 1 Convergence de 3 égaliseurs.

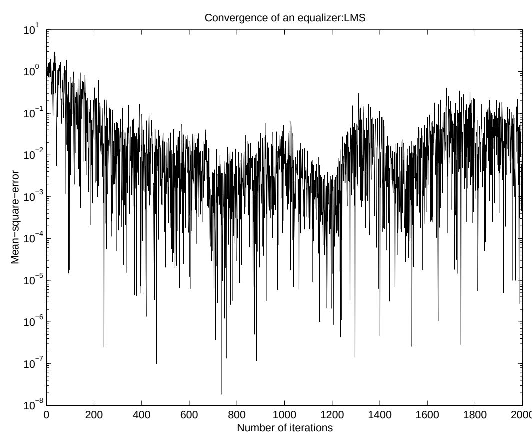

Figure 2 Convergence d'un égaliseur LMS : canal non-stationnaire.

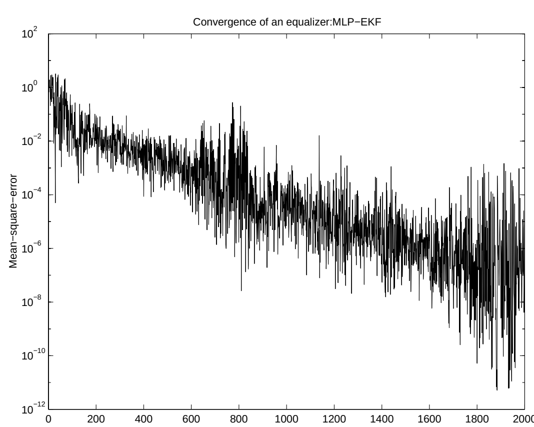

Figure 3 Convergence d'un égaliseur PMC-EKF: canal non-stationnaire.

Conclusion

Malgré leur complexité supérieure aux égaliseurs transversaux conventionnels, les égaliseurs neuronaux se distinguent par leur capacité à reconstituer de manière beaucoup plus précise la réponse impulsionnelle inverse du canal. Des travaux sont en cours pour utiliser des algorithmes *aveugles* (non-supervisés) pour entraîner des réseaux de neurones [6], [7].

Références

[1] Qureshi, S.U.H., “Adaptive equalization”, *Proceedings of the IEEE*, Vol. 73, no 9, pp.1349-1387, Sept. 1985.

[2] Chen, S., Gibson, G.J., Cowan, C.F.N., Grant, P.M., “Adaptive equalization of finite non-linear channels using multilayer perceptrons”, *Signal Processing*, Vol.20, pp.107-119, 1990.

[3] Chen, S., Gibson, G.J., Cowan, C.F.N., “Adaptive channel equalisation using a polynomial-perceptron structure”, *IEE Proceedings*, Vol. 137, Pt. I, no 5, oct. 1990.

[4] Chen, S., Gibson, G.J., Cowan, C.F.N., Grant, P.M., “Reconstruction of binary signals using an adaptive radial-basis-function equalizer”, *Signal Processing*, Vol. 22, pp. 77-93,1991.

[5] Rappaport, T.S., Seidel, S.Y., Takamizawa, K., “Statistical channel impulse response models for factory and open plan building radio communication system design”, *IEEE Transactions on Communications*, Vol. 39, no 5, pp. 794-807, mai 1991.

[6] You, C., Hong, D., “Nonlinear blind equalization schemes using complex-valued multilayer feedforward neural networks”, *IEEE Transactions on Neural Networks*, Vol. 9, no 6, pp. 1442-1455, nov. 1998.

[7] Mannerkoski, J., Taylor, D.P., “Blind equalization using least-squares lattice prediction”, *IEEE Transactions on Signal Processing*, Vol. 47, no 3, pp. 630-640, mars 1999.