

UTILISATION CONJOINTE DE MODELES INTERNE ET EXTERNE POUR LA CLASSIFICATION ET LA LOCALISATION DE DEFAUTS

Application au diagnostic de la transmission voie-machine du TGV.

Latifa Oukhellow¹, Patrice Aknin², François Vilette³

¹ CERTES–Université Paris XII, 61 avenue du Général de Gaulle, 94100 Créteil

² INRETS-LTN, 2 avenue du Général Malleret Joinville, 94114 Arcueil Cedex

³ SNCF, IPM MR 3, 17 rue d'Amsterdam, 75008 Paris

Mots-Clés : diagnostic, circuit de voie, détection de rupture, algorithme GLR, algorithme CUSUM

Un système automatique de diagnostic des circuits de voie de la transmission voie-machine du TGV a été mis au point pour détecter le dysfonctionnement de certains composants placés à la voie.

L'ESSENTIEL

- La ligne ferroviaire grande vitesse est segmentée en zones contiguës dites « zones isolées », auxquelles sont associées un système de signalisation appelé circuit de voie (CdV). Ces CdV transmettent au train des informations telles que la vitesse limite d'avancement.
- Un système de diagnostic des défauts du CdV a été développé en se basant uniquement sur l'analyse du courant dit de court-circuit (Icc) mesuré à l'aide d'un véhicule d'inspection.
- Le système de diagnostic combine deux types de modélisation : une modélisation interne fondée sur la connaissance du fonctionnement physique du système et une modélisation externe du courant Icc construite à partir de l'expertise acquise.
- Le modèle interne du CdV a été mis au point en tenant compte de tous les éléments ponctuels placés à la voie et des éléments répartis (rail, ballast...). Ce modèle permet de calculer l'évolution théorique du courant Icc le long de la zone isolée, avec la possibilité de simuler l'apparition de différents types de défauts.
- Le modèle externe permet de représenter le signal Icc par un nombre de paramètres réduit à l'aide de paramétrisations de type autorégressive, d'ondelettes ou d'heuristiques.
- Le diagnostic du système fournit deux informations : la classe de fonctionnement (sans défaut/avec défaut) et la localisation du défaut si le système est défectueux. La classe de fonctionnement est obtenue à l'aide d'un réseau de neurones et la localisation du défaut fait appel à des algorithmes de détection de rupture dans les signaux.

SYNOPSIS

- The high speed railway track is divided into contiguous segments called "isolated zones". A track circuit is associated to each isolated zone and it transmits informations from track to train such as the speed limit.

- The information transmission system diagnosis is based only on the analysis of the short-circuited current Icc measured by a specific inspection vehicle.
- The diagnosis system presented in this paper uses both an internal modeling based on its physical behaviour and an external one that requires a priori knowledges on the signal to be analysed.
- An internal modeling of the track circuit takes into account all limited elements of the track or distributed ones (rail, ballast...). It is used to estimate the Icc current evolution with the possibility of different defect type simulation.
- An external modeling is performed in order to obtain a data representation space with reduced dimension and to make the classification task easier. Both parametric modeling such as autoregressive, wavelet and heuristic parametrization of the signal are tested.
- The system achieves a classification task (defect case/ no defect case) by means of a neural network. In the defect case a localization unit using detection algorithms determines the defect position.

1. INTRODUCTION

La bonne compréhension du comportement d'un système physique nécessite de développer des modèles. Pour certaines applications, un haut niveau d'analyse conduit à des modèles internes complexes où un grand nombre de paramètres doit être ajusté pour reproduire le phénomène physique observé. Cette approche est souvent incompatible avec des contraintes de diagnostic temps réel et une approche externe basée sur l'expertise acquise par les opérateurs sur les observations du système est souvent préférée.

Cet article présente un système de diagnostic de la transmission voie-machine des TGV, basée sur une

utilisation conjointe de modèles interne et externe (cf. figure 1).

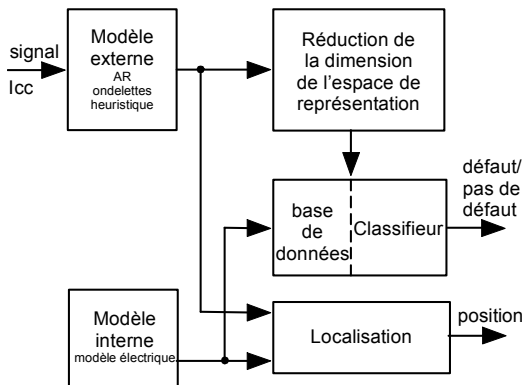


Figure 1 : Système de diagnostic de la transmission voie-machine du TGV

La modélisation externe permet d'obtenir un espace de représentation des données de faible dimension facilitant la tâche de classification. Si une observation est affectée à la classe « défaut », une procédure de localisation utilisant une modélisation interne et/ou externe détermine ensuite la position du défaut.

2. PRINCIPE DE LA TRANSMISSION VOIE-MACHINE

Sur le réseau ferré français, les lignes à grande vitesse sont toutes équipées d'un système de transmission voie-machine qui utilisent les rails pour transmettre, sur une fréquence porteuse donnée, un certain nombre d'informations telle que la vitesse maximale à ne pas dépasser sur une section de voie donnée [1]. Un CdV comprend (cf. figure 2) :

- un émetteur branché à l'une des extrémités de la section de voie et qui délivre un courant alternatif.
- une ligne de transmission constituée des deux fils de rail dont on adapte l'impédance à l'aide de condensateurs d'accord disposés régulièrement.
- un récepteur branché à l'autre extrémité de la section de voie.

Les courants délivrés par le CdV "remontent" vers le récepteur et subissent des atténuations inévitables tout au long de la section de voie ainsi délimitée. Pour limiter ce phénomène, un accord électrique est nécessaire pour compenser la nature selfique de la ligne ferroviaire. Cet accord s'effectue à l'aide de condensateurs régulièrement espacés et connectés entre les fils de rail. Le nombre de condensateurs d'accord dépend de la valeur de la fréquence porteuse et de la longueur de la section de voie (ou zone isolée).

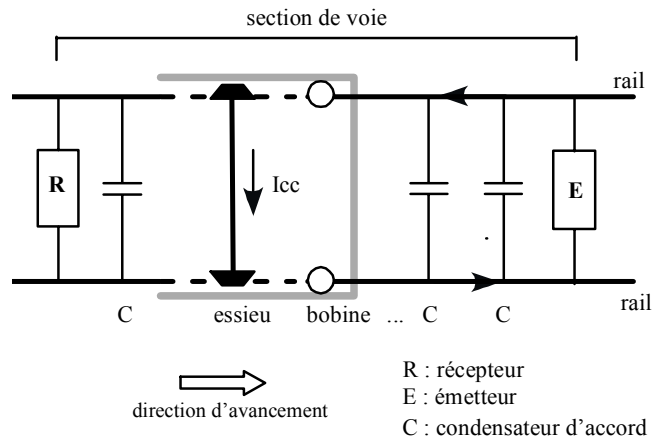


Figure 2 : Schéma électrique d'un CdV

Le diagnostic de fonctionnement du système complet a pour principal objectif de vérifier si le signal reçu sur le récepteur n'est pas trop atténué, ce qui nuirait à son bon fonctionnement et à la qualité des informations transmises au train. Pour ce faire, un véhicule spécifique d'inspection est utilisé. Il est équipé de bobines de mesure qui visent les rails et délivrent une tension proportionnelle au courant I_{cc} qui les parcourt alors que le premier essieu du train court-circuite électriquement les deux rails. La figure 4 donne un exemple d'évolution du courant I_{cc} . Nous détaillons dans cet article la fonction « détection d'absence de condensateur d'accord » du système de diagnostic opérant sur le courant I_{cc} .

Dans un premier temps, un modèle électrique du CdV est détaillé. Ce modèle permet de simuler l'évolution du courant I_{cc} lorsque le train se déplace de l'émetteur vers le récepteur.

3. MODELISATION DU CIRCUIT DE VOIE

3.1. Modèle interne

Dans cette approche, la voie ferrée est considérée comme une ligne à constantes réparties [1] d'impédance caractéristique Z_c et de constante de propagation θ :

$$Z_c = \sqrt{\frac{R + j2\pi fL}{G + j2\pi fC}}$$

et $\theta = \sqrt{(R + j2\pi fL)(G + j2\pi fC)}$

où f est la fréquence, R la résistance longitudinale, L la capacité longitudinale, G la conductance transversale, C la capacité transversale. La matrice de

transfert d'une voie de longueur L considérée comme un quadripôle est donnée par :

$$Q = \begin{bmatrix} \text{ch}(\theta L) & -Z_c \text{sh}(\theta L) \\ -\frac{1}{Z_c} \text{sh}(\theta L) & \text{ch}(\theta L) \end{bmatrix}$$

Le CdV est donc composé d'une succession de quadripôles linéaires passifs et symétriques, alternance de quadripôles équivalents à des portions de voie et de quadripôles représentant les différents éléments ponctuels placés à la voie tels que les condensateurs de compensation, les joints d'émission et de réception... (cf. figure 3). A l'aide de la théorie des lignes et d'équations matricielles, le modèle électrique permet de calculer le courant en chaque point du CdV, pour chaque position du train, ce dernier réalisant un court-circuit mobile sur la zone modélisée.

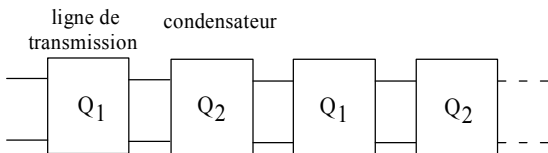


Figure 3 : Schéma équivalent d'un CdV

La figure 4 présente l'exemple d'un courant I_{cc} réel et simulé pour un CdV de longueur 1500 m. Une centaine de paramètres physiques doivent être ajustés pour le calcul du modèle. En dépit de cette complexité, le modèle interne fournit de bons résultats, très voisins des observations (cf. figure 4). On peut noter la présence d'une oscillation principale ($\lambda \approx 400$ m) sur le courant I_{cc} due à une désadaptation de la voie. L'espacement régulier des condensateurs d'accord se traduit par l'apparition d'arches sur le courant (forme de chaînette inversée) de période égale au pas de compensation.

L'intérêt du modèle interne est double. D'une part, il permet de calculer l'évolution théorique du courant I_{cc} le long du CdV, avec différents défauts susceptibles de survenir sur le système et donc d'analyser finement leur manifestation. D'autre part, sachant que l'on dispose de peu de données réelles avec défaut, ce modèle a permis la constitution d'une base représentative de signaux I_{cc} , simulés avec ou sans défauts (2200 signaux générés pour des paramètres d'entrée variables comme la longueur du CdV, l'existence et la position des condensateurs absents, la fréquence porteuse et le rapport signal à bruit).

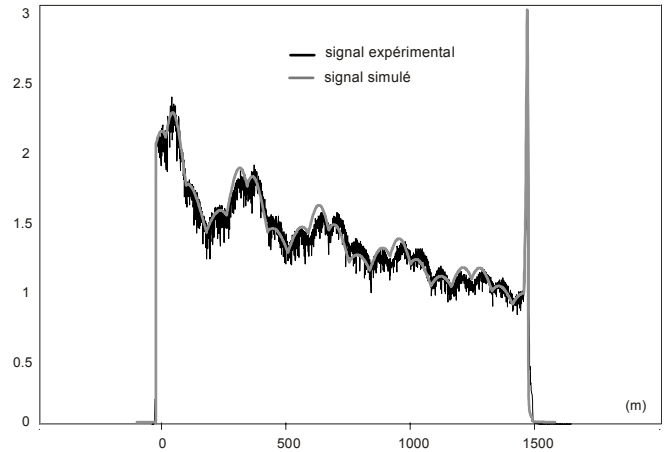


Figure 4 : Exemple d'un signal I_{cc} expérimental et simulé

3.2. Modèles externes

Deux types de modélisation externes ont été élaborées et comparées en terme de performances de classification :

- une modélisation autorégressive combinée avec une décomposition en ondelettes discrètes. Dans cette approche, le signal est considéré comme un processus autorégressif d'ordre p , généré en filtrant un bruit blanc de variance unitaire. La transformée en ondelettes discrètes du signal est également appliquée pour le décomposer, par projections successives, en approximations (composantes basse fréquence) et détails (composantes haute fréquence). A l'issue de cette paramétrisation, le signal I_{cc} est représenté par un jeu de 20 paramètres convenablement sélectionnés à l'aide du critère d'Akaike et de la procédure d'orthogonalisation de Gram-Schmidt [2].

- Une approche heuristique basée sur la connaissance experte acquise lors de l'étude a permis d'établir un modèle à 6 paramètres pour approximer le courant I_{cc} (cf. figure 5). Le modèle est donné par l'expression :

$$f(x) = \frac{a_1 \cdot e^{a_2 x}}{1 + a_3 \cdot \cos(a_4 x + a_5) \cdot e^{a_6 x}}$$

Ce modèle relativement simple ne permet pas d'appréhender la totalité du contenu informatif du signal I_{cc} (les chaînettes inversées ne sont pas modélisées). Mais nous verrons qu'il est pertinent -car parcimonieux- pour traiter le problème de diagnostic posé. Par exemple, la décroissance régulière le long du CdV de l'onde à 400 m a été identifiée comme un signe de « bonne santé » du système et celle-ci est

parfaitement chiffrée par le seul coefficient a_3 .

L'estimation des coefficients du modèle à partir du signal réel, s'effectue par minimisation d'une fonction de coût basée sur l'erreur quadratique moyenne à l'aide de la méthode du simplexe [3]. Seuls 4 coefficients sont utilisés pour représenter le signal, le facteur d'échelle a_1 et la fréquence spatiale a_5 n'étant pas pertinents pour discriminer les différentes classes de défaut. L'étude du comportement du modèle externe a permis d'extraire également un ensemble complémentaire de 4 descripteurs tels que l'erreur de modélisation, les variations de constantes de temps locales, le taux de décroissance des enveloppes inférieures et supérieures.... Au total, le signal est ainsi représenté par un ensemble réduit de 8 paramètres.

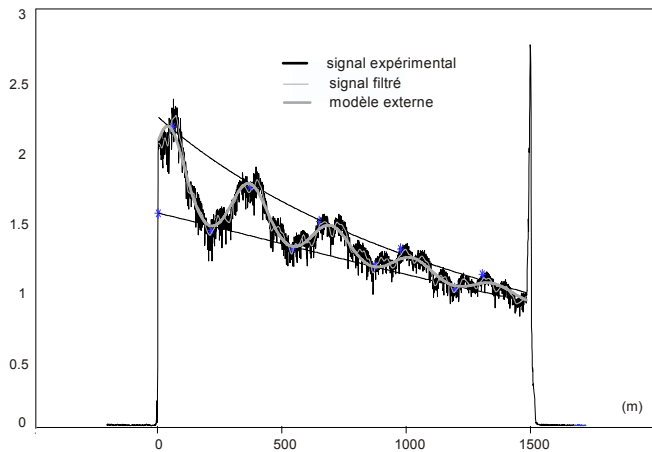


Figure 5 : Approximation du signal I_{cc} à l'aide du modèle externe

4. RESULTATS DE CLASSIFICATION

Nous avons choisi de réaliser la tâche de classification à l'aide d'un réseau de neurones de type perceptron multicouches à une seule couche cachée (7 neurones à activation sigmoïdale) [4]. La classification est ici effectuée en termes de Défaut/Non défaut et l'apprentissage a été focalisé sur la classe des défauts « condensateur d'accord absent ». Le tableau 1 donne les performances obtenues sur la base des 2200 signaux paramétrés par les deux modèles externes décrits au paragraphe 3.2. Les résultats ont été estimés par la méthode du Leave One Out. Dans cette méthode de validation, on effectue l'apprentissage sur la base entière moins un exemple et on teste le diagnostic sur l'exemple écarté ; la performance est donc une grandeur statistique moyennée sur le nombre d'exemples [4].

	Bonnes détection	Fausse alarmes
AR + ondelettes	60.8 %	29.4 %
Modèle heuristique	97.6 %	0.4 %

Tableau 1 : Performances de classification

Le modèle heuristique est plus parcimonieux que la paramétrisation par modèle AR et coefficients d'ondelettes (seulement 8 coefficients représentent le signal contre 20 pour le second modèle). De plus, la prise en compte d'une quantité importante d'informations a priori dans ce modèle explique son meilleur comportement en terme de discrimination entre classes.

5. LOCALISATION DES DEFAUTS

Lorsqu'un CdV est classé comme défectueux, la position du défaut doit être estimée. Trois approches seront détaillées pour la localisation du défaut, toutes basées sur la recherche des modifications locales des caractéristiques du signal I_{cc} .

5.1. Localisation du défaut avec modèle interne

La localisation du défaut s'effectue ici à l'aide d'une procédure qui calcule le modèle interne du CdV, après le retrait d'un des condensateurs d'accord. Autant de modèles que de condensateurs sont calculés. Le modèle retenu est celui qui présente la plus faible erreur quadratique avec le signal réel (cf. figure 6). Cette approche suppose que la quasi-totalité des paramètres internes au modèle ont été préalablement réglés correctement.

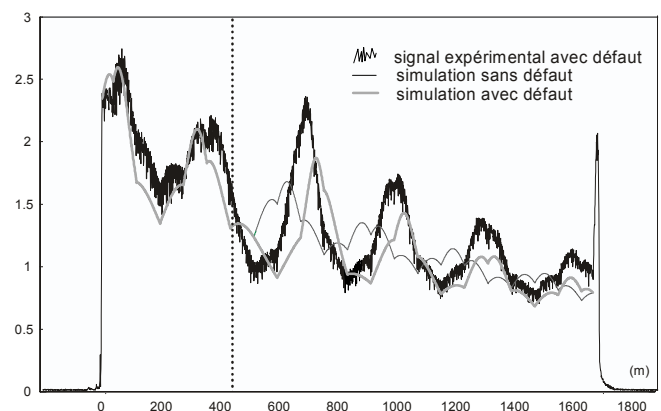


Figure 6 : Localisation du défaut avec le modèle interne

5.2. Localisation du défaut avec modèle externe

Dans ce cas, la localisation du défaut est basée sur la détection de rupture dans le modèle externe décrit dans le paragraphe 3.2. Deux algorithmes de détection ont été utilisés pour estimer des distances entre modèles : le rapport du maximum de vraisemblance généralisé (GLR Generalized Likelihood Ratio) et l’algorithme des sommes cumulées (CUSUM) [5]. Le but est de comparer, au sens d’une métrique donnée, les paramètres de différents modèles estimés sur des fenêtres d’observations différentes.

- Rapport de vraisemblance généralisé (GLR)

La localisation du défaut est basée sur un test de détection utilisant une mesure du rapport de vraisemblance généralisé. Deux modèles M_1 , M_2 sont identifiés sur deux fenêtres à mémoire croissante et décroissante, obtenues en additionnant ou en omettant une cellule du CdV. Un modèle global est également estimé sur la totalité du signal (cf. figure 7). Cette implantation de l’algorithme est légèrement différente du cas usuel où l’on fait intervenir une fenêtre de référence à mémoire croissante et une fenêtre glissante de longueur fixe [6]. Elle est adaptée à l’étude non causale d’un signal de longueur fixe (la longueur de la zone isolée). La distance entre les trois modèles est donnée par la relation :

$$d_i = L \log(\sigma_0^2) - L_1 \log(\sigma_1^2) - L_2 \log(\sigma_2^2)$$

où L , L_1 , L_2 sont les longueurs des trois fenêtres, σ_0^2 , σ_1^2 , σ_2^2 les erreurs résiduelles associées aux trois modèles et i l’indice de la cellule de compensation.

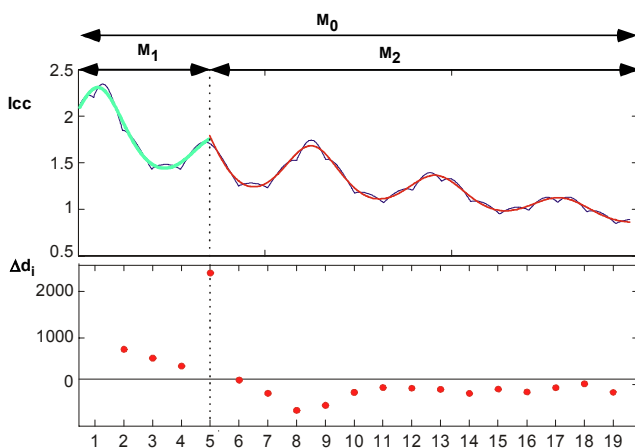


Figure 7 : Localisation du défaut par l’algorithme GLR

La détection est activée si la variation de la distance d_i dépasse un seuil judicieusement choisi. Le comportement de cet algorithme sur un signal I_{cc} est illustré figure 7. On peut noter que la variation de la distance d_i atteint effectivement un maximum à la position du défaut. Cette méthode de localisation donne de bons résultats, quelque soit la position du défaut.

- Algorithme des sommes cumulées (CUSUM)

Cet algorithme estime deux modèles : un modèle global à long terme et un modèle local à court terme, qui sont différemment affectés par un changement dans le signal. Les paramètres de ces modèles sont calculés sur une fenêtre à mémoire croissante et une fenêtre glissante de longueur fixe incluse dans la fenêtre croissante (cf. figure 8).

La distance entre les deux modèles M_1 et M_2 estimée par l’algorithme des sommes cumulées combine les erreurs de modélisation et d’innovations [5] :

$$d_i = \frac{1}{2} \log\left(\frac{\sigma_1^2}{\sigma_2^2}\right) + \frac{\varepsilon_1^2}{2\sigma_1^2} - \frac{\varepsilon_2^2}{2\sigma_2^2}$$

où σ_1 , σ_2 sont les erreurs résiduelles et ε_1 , ε_2 les innovations. La figure 8 montre le comportement de cet algorithme sur le même exemple que celui de la figure 7.

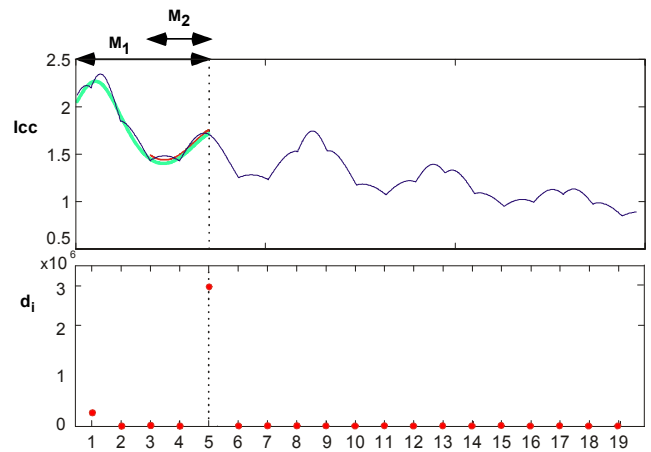


Figure 8 : Localisation du défaut par l’algorithme des sommes cumulées

Le choix de la longueur de la fenêtre M_2 résulte d’un compromis entre la précision de l’estimation des paramètres du modèle local et l’énergie relative des composantes basse-fréquence induite par le défaut. Une fenêtre glissante de 2 cellules fournit d’assez bons résultats pour notre base de données.

5.3. Comparaison des trois méthodes de localisation

L'utilisation du modèle interne est particulièrement intéressante dans la mesure où la même démarche peut servir à localiser d'autres types de défaut. En revanche, le principal inconvénient de cette approche est le nombre élevé de paramètres du modèle électrique qui rend son ajustement difficile. Si les paramètres sont ajustés de façon précise, la localisation fonctionne très bien.

Le rapport du maximum de vraisemblance permet une estimation robuste de la position du défaut sans connaissance a priori. L'absence de condensateur est bien identifiée par cette approche. Le cas de défaut de gravité moindre (condensateur avec pertes) peut présenter des difficultés car le contraste entre modèles n'est pas toujours suffisant.

L'algorithme CUSUM fournit un meilleur contraste en termes de distance mais il induit une zone morte en début de CdV (2 à 3 cellules), où aucun défaut ne peut être repéré.

6. CONCLUSION

Cet article montre l'intérêt d'utiliser à la fois des modélisations interne et externe pour réaliser un diagnostic précis d'un système complexe. Le modèle interne est un outil de simulation et d'analyse complet du comportement électrique des CdV. Il peut servir également à la localisation des défauts mais s'avère relativement lourd à manipuler. Son principal intérêt a été d'augmenter la connaissance experte du système grâce à sa capacité de générer un grand nombre de dysfonctionnements difficilement observables sur les signaux réels, sur une période d'observation nécessairement finie.

Cette connaissance experte acquise sur le système a permis de développer un modèle externe du signal Icc de complexité réduite et fournissant de bonnes performances en classification comparée au modèle interne. Des résultats encourageants pour la localisation des défauts ont également été obtenus à l'aide de l'algorithme du rapport de vraisemblance généralisé utilisant le modèle externe.

7. REFERENCES

- [1] F. Fessant, P. Aknin, F. Vilette, M. Antoni. *Modélisation électrique du circuit de voie, élément du système de transmission voie-machine des TGV*. Revue 3EI n°27. pp.46-52. 2001
- [2] L. Oukhellou, P. Aknin, H. Stoppiglia, G. Dreyfus. *A new decision criterion for feature selection. Application to the Classification of Non Destructive Testing Signatures*. IX European Signal Processing Conference. Vol 1 pp.411-414. Greece. 1998.
- [3] J.A. Nelder, R. Mead, *A simplex method for function minimization.*, Computer journal. Vol 7 pp.308-313. 1965
- [4] C.M. Bishop *Neural network for pattern recognition*. Clarendon press. 1995.
- [5] M. Basseville, I.V. Nikiforov, *Detection of abrupt changes : theory and application*, Prentice Hall, 1997.
- [6] C. Villemur, F. Castanié, "*Détection de défauts et ségmentation de signaux de contrôle non destructif*", GRETSI, pp.73-76. 1987

LES AUTEURS

Latifa OUKHELLOU est Maître de Conférence à l'Université Paris XII-Créteil, Docteur de l'Université Paris Sud - Orsay. Ses recherches actuelles portent sur le traitement du signal et l'analyse de données appliquées à des problématiques de diagnostic.

Patrice AKNIN est Directeur de Recherche à l'INRETS. Ancien élève de l'ENS Cachan, Docteur de l'Université Paris Sud - Orsay, il anime aujourd'hui l'équipe "Diagnostic des systèmes de transports guidés" au Laboratoire des Technologies Nouvelles de l'INRETS. email : aknin@inrets.fr

François VILETTE est Expert en Ingénierie de Maintenance des systèmes de signalisation, d'énergie et de traction électrique au Département Maintenance Développement Infrastructure de la SNCF.