

Classification pour la détection de défauts de surface sur rail

Quentin MAYOLLE^{1,2}, Philippe VANHEEGHE^{1,2}, Pierre-Marie ROGIER¹

¹IRT Railenium, Valenciennes, France.

²Univ. Lille, CNRS, Centrale Lille, UMR 9189 - CRISTAL, Lille France.

quentin.mayolle@railenium.eu, philippe.vanheeghe@centralelille.fr
pierre-marie.rogier@railenium.eu

Résumé – Dans cet article, nous proposerons des méthodes de classification automatique sur des signaux issus de mesures de contrôle non destructif pour le rail effectuées en laboratoire. Nous testons d’abord les performances de ces méthodes sur des signaux issus d’EMAT (Electro-Magnetic Acoustic Transducers), outil permettant l’émission d’ultrasons par des transducteurs magnétiques, et dont l’intérêt pour la détection de défauts de surface a été démontrée récemment. Nous proposerons dans un deuxième temps le calcul de nouvelles caractéristiques à utiliser pour la classification, afin de s’affranchir de la problématique liée aux variations d’amplitude entre les mesures, due au dispositif de mesure.

Abstract – In this article, we suggest automatic classification methods for signals coming from laboratory Non Destructive Testing on railways. In this article, we have first tested the performances of these methods on signals coming from EMAT (Electro-Magnetic Acoustic Transducers), a tool for generating ultrasonic waves with magnetic transducers, which has been shown to be an effective way to detect surface defects. We further suggest the calculation of new characteristics to be used for classification, to bypass the problems linked to changes of amplitude between measures, due to the acquisition device.

1 Introduction

Les rails de chemin de fer subissent d’importantes contraintes dues aux passages répétés des trains. Ces sollicitations peuvent entraîner la formation de fissures. Pour prévenir les ruptures de rail, il est important de pouvoir les détecter le plus tôt possible. Les défauts de surfaces apparaissent dans la partie supérieure du rail, où a lieu le contact roue/rail. Le contrôle non destructif (NDT : Non Destructive Testing) permet d’obtenir une information sur l’état de la surface du rail sans l’endommager. De nombreuses techniques existent, on peut citer, parmi les plus connues : l’inspection visuelle, l’utilisation de capteurs à ultrasons, et les courants de Foucauld [1]. La classification automatique a été employée sur des signaux obtenus par ces techniques avec succès, notamment avec des SVM (Support Vector Machines) [2][3]. Les Electro-Magnetic Acoustic Transducers (EMAT) ont récemment démontré leur efficacité pour la détection de fissures [8]. Cette méthode présente l’avantage de générer une onde ultrasonore en surface du rail sans nécessiter de contact avec celui-ci [10]. Ceci permet d’embarquer ces capteurs pour réaliser des mesures sur un engin d’auscultation mobile, et d’obtenir un signal, malgré une variation du lift-off (l’espace entre le capteur et la surface du rail), contrairement aux techniques à ultrasons classiques.

Dans cet article, nous montrerons que les outils de classification cités précédemment donnent aussi de bons résultats sur les signaux EMAT. Dans la perspective d’un engin d’auscultation mobile, le lift-off est susceptible de changer d’une mesure à l’autre. L’augmentation du lift-off diminuant expo-

mentiellement l’amplitude du champ magnétique généré [8], il est donc nécessaire de pouvoir fournir aux classifieurs des observations qui ne sont pas influencées par la géométrie du capteur ou l’environnement. Oukhellou [4] a mis en évidence l’importance de caractéristiques invariantes aux homothéties et translations temporelles. De plus, elles doivent pouvoir être facilement interprétables et ne pas dépendre de la fréquence d’échantillonnage. Nous proposons le calcul de caractéristiques permettant de répondre aux 4 critères précédents, basées sur l’analyse vibratoire des signaux [5]. Nous montrerons que leur utilisation améliorera les scores de classification lorsque les conditions expérimentales changent.

Cette classification de signaux issus d’EMAT a été testée sur des mesures obtenues sur un banc d’essai. Des défauts ont été usinés sur un rail, sur lequel les mesures ont été réalisées. Dans un premier temps, nous présenterons le principe de la mesure par EMAT, ainsi que le banc d’essai. Ensuite, nous décrivons la méthodologie employée : l’extraction de caractéristiques des signaux, et les méthodes de classification des acquisitions. Une dernière partie exposera les résultats obtenus sur ces signaux.

2 Experimentation

2.1 Principe de la mesure par EMAT

La technologie des EMAT permet d’obtenir des informations sur les couches extérieures du champignon du rail. Un transducteur magnétique va produire une force de Lorentz en

surface du rail, qui va générer une onde de Rayleigh de longueur d'onde $\lambda = 5 \text{ mm}$, et qui va se déplacer à une vitesse de l'ordre de 2980 m/s . La propagation de l'onde peut être perturbée par une fissure. En fonction de la profondeur de celle-ci ou de son inclinaison par rapport à la surface, on peut observer des phénomènes de réflexion ou d'atténuation de l'onde [8].

2.2 Présentation du banc d'essai

La partie mesure du banc d'essai est constituée de deux transducteurs, les EMAT, composés d'une bobine et d'un aimant permanent. Ils sont montés sur une plateforme reliée à une courroie motorisée permettant de déplacer l'ensemble le long du rail. Le principe du dispositif est représenté en figure 1. Deux types de mesure sont alors possibles. Lors des mesures dites en « statique », le dispositif ne bouge pas lors de l'acquisition. Il est seulement déplacé cm par cm entre chaque acquisition. Lors des mesures dites en « dynamique », le dispositif est translaté le long du rail à une vitesse de 2 m/s , et des acquisitions sont réalisées périodiquement durant la translation.

Le rail utilisé pour l'expérimentation possède une longueur de 1 m . Quatre défauts ont été usinés, de profondeurs respectives $4, 8, 7$ et 4.1 mm , et espacés de 10 cm .

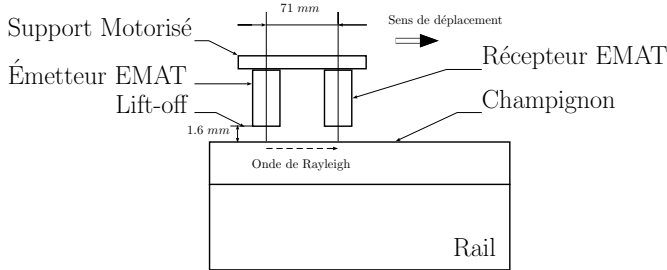


FIGURE 1 – Schéma de principe du banc de mesure

On réalise une acquisition de la manière suivante : on démarre l'enregistrement du signal et on génère simultanément une onde ultrasonore en surface du rail. La durée de l'enregistrement, égale à $200 \mu s$, permet de capter des réflexions. La fréquence d'échantillonnage f_e est fixée à 25 MHz , pour obtenir ainsi 5000 points. En statique, une acquisition est réalisée tous les centimètres. En dynamique, le dispositif est en mouvement, et la fréquence des acquisitions est réglée de manière à en obtenir également une tous les centimètres.

3 Méthodologie

3.1 Extraction de caractéristiques

On se propose d'utiliser deux ensembles de caractéristiques pour la classification, calculées pour chaque acquisition.

Dans un premier temps, nous avons calculé des caractéristiques courantes en NDT [2] : énergie, maximum de l'enveloppe, instant de ce maximum, mode fréquentiel, densité spectrale

correspondant à ce mode, et, suivant Oukhellou [4], écarts entre le minimum et le maximum pour les parties réelles et imaginaires de la transformée de Fourier, ainsi que la différence entre les fréquences correspondantes. Ces caractéristiques formeront l'ensemble de caractéristiques noté « Car 1 ».

Dans un deuxième temps, nous avons calculé des caractéristiques statistiques indépendantes de l'amplitude du signal, souvent utilisées en analyse vibratoire [5]. En normalisant la densité spectrale discrète, on obtient une densité de probabilité discrète. Des statistiques peuvent être ainsi calculées. Elles ont l'avantage d'être invariantes aux changements d'échelles (homothéties), aux translations temporelles, et aux changements de fréquence d'échantillonnage.

Si l'on note s un signal de N points, ayant pour fréquence d'échantillonnage f_e , alors on notera $S(f_k)$ la transformée de Fourier discrète associée à la fréquence f_k , avec $f_k = \frac{k f_e}{N}$, $k = 0..N$. En utilisant la densité spectrale de puissance, l'application $f_n \mapsto \frac{|S(f_n)|^2}{\sum_{k=0}^{N/2-1} |S(f_k)|^2}$, $k = 0..N$ définit une densité de probabilité sur $\{f_k, k = 0..(N/2 - 1)\}$.

On définit ainsi :

— le moment ordinaire d'ordre 0 (énergie spectrale) :

$$m_0 = \sum_{k \geq 0} S(f_k) \quad (1)$$

— le moment ordinaire d'ordre i :

$$m_i = \left(\sum_{k \geq 0} f_k^i S(f_k) \right) / m_0 \quad (2)$$

— le moment centré d'ordre i :

$$\mu_i = \left(\sum_{k \geq 0} (f_k - m_1)^i S(f_k) \right) / m_0 \quad (3)$$

— la fréquence empirique d'extrema : $\nu_p = \sqrt{m_4/m_2}$; la fréquence empirique de zéros : $\nu_0 = \sqrt{m_2}$; le coefficient alpha d'ordre 2 : $\alpha_2 = m_2/\sqrt{m_4}$.

On obtient les caractéristiques suivantes : le mode de la distribution, la densité associée à ce mode, le centre de gravité m_1 , l'écart type $\sqrt{\mu_2}$, le coefficient d'asymétrie $\mu_3/\mu_2^{3/2}$ et le kurtosis $\gamma_2 = \mu_4/\mu_2^2 - 3$ de la distribution, ainsi que ν_p , ν_0 et α_2 . L'ensemble de caractéristiques sera noté « Car 2 ».

En prenant la valeur absolue de la transformée de Fourier pour une fréquence donnée, le terme lié au changement de phase disparaît. L'utilisation du spectre de puissance permet ainsi de résoudre le problème de l'invariance à la translation temporelle. De plus, la normalisation résout celui du changement par homothétie. Les moments synthétisent une information sur l'ensemble du spectre, qui n'est plus directement dépendante de la fréquence d'échantillonnage. L'analyse vibratoire a de plus montré le sens physique de ces variables [5]. L'ensemble de caractéristiques répond ainsi aux 4 critères fixés précédemment.

3.2 Classification d'acquisitions EMAT

Nous avons séparé l'ensemble de mesures statiques en deux : un ensemble d'apprentissage S1, comprenant deux défauts, et un de test T1, avec deux autres défauts. Les acquisitions en dynamique formeront le deuxième ensemble de test T2. Les

algorithmes de classification sont entraînés sur S1, et on testera les capacités des algorithmes à généraliser, d'abord sur d'autres défauts en statique (sur T1), puis sur les acquisitions T2.

Trois méthodes de classification ont été choisies. En premier lieu, l'algorithme de classification par processus gaussien de Rasmussen et Williams [7]. Son principal intérêt est de ne pas nécessiter d'optimisation des hyperparamètres par cross-validation. En effet, on peut directement les sélectionner en maximisant la log vraisemblance marginale. Cette méthode de classification est non-paramétrique, et nécessite de choisir une fonction de covariance k : la version anisotropique du noyau Squared Exponential a été choisie, permettant d'attribuer un paramètre $l_i \in \mathbb{R}_+^*$ à chaque dimension $i = 1..d$, et ainsi de régler la contribution de chacune à la décision :

$$k(x, x') = \sigma \exp \left(-\frac{1}{2} \sum_{i=1}^d \frac{|x_i - x'_i|^2}{l_i^2} \right), \sigma \in \mathbb{R}_+^*$$

En deuxième lieu, l'algorithme de classification probabiliste par régression logistique, simple d'implémentation [9]. Enfin, en dernier lieu les SVM linéaires, majoritairement utilisées, en NDT, pour la classification des signaux [2][6]. Les hyperparamètres de régularisation des deux derniers algorithmes ont été choisis par cross-validation (3 plis) sur S1.

Les trois algorithmes utilisés présentent la particularité de pouvoir fournir un score pour chaque observation de T1 et T2. Les classifications par processus gaussien (GP) et par régression logistique permettent d'obtenir une probabilité d'appartenance à la classe « défaut », entre 0 et 1. Les SVM donnent également un score : la distance des observations à l'hyperplan séparateur.

4 Résultats

Suite aux expérimentations sur le banc, on dispose de deux ensembles de données. Le premier correspond aux 92 acquisitions statiques de 5000 points, S1 et T1. Le deuxième correspond à celles dynamiques T2, et en comporte 95, avec le même nombre de points. Tous les ensembles peuvent être scindés en 4 parties, chacune comportant une zone de défaut. Chaque défaut produit plusieurs acquisitions labellisées « défaut ».

4.1 Signaux expérimentaux

La figure 2 représente différents signaux enregistrés via le récepteur EMAT, en présence ou non de défaut, en statique et en dynamique. Par exemple, on observe dans le signal a) l'onde qui arrive directement au récepteur, entre 0 et 20 μs , ainsi que deux échos aux alentours de 80 μs , liés à des réflexions aux bords ou aux fissures. On observe, dans le cas des acquisitions dynamiques, une dégradation du signal, supposée être la conséquence des variations du lift-off.

4.2 Résultats de la classification

La métrique utilisée pour l'évaluation des performances est l'AUC (Area Under the Curve) de la courbe ROC (Receiver

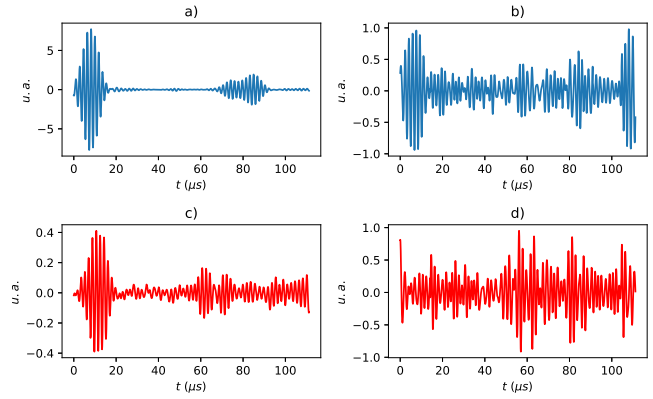


FIGURE 2 – Exemple de signaux acquis par le récepteur EMAT a) En Statique Sans Défaut b) En Dynamique Sans Défaut c) En Statique Avec Défaut d) En Dynamique Avec Défaut

Operating Characteristic). La courbe ROC mesure la performance de classifieurs binaires fournissant un score de classification pour chaque observation. En faisant varier le seuil permettant d'attribuer une classe en fonction de ce score, on modifie les taux de faux positifs et de vrais négatifs. L'aire sous cette courbe est l'AUC. Un score AUC de 0.5 correspond à un classifieur non-informatif. Un score AUC de 1 correspond à un classifieur parfait.

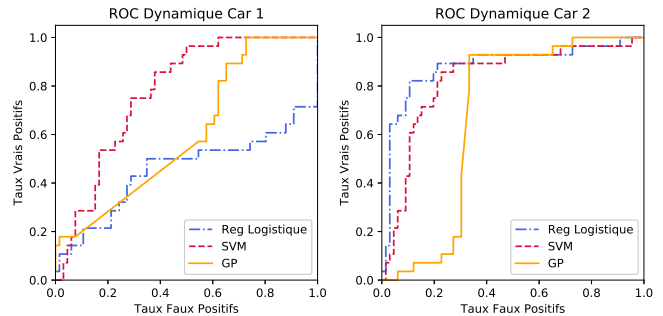


FIGURE 3 – Courbes ROC pour les ensembles de variables

La figure 3 illustre les performances des méthodes de classification sur les acquisitions dynamiques par leur courbes ROC. On remarque que la régression logistique réalise une classification nettement moins bonne que les deux autres méthodes. En particulier, pour un taux de faux positif élevé, la courbe associée est bien en dessous des autres. Une analyse plus poussée des caractéristiques a été nécessaire pour expliquer ces mauvaises performances. Des variables corrélées sur l'ensemble d'apprentissage S1 ne le sont plus sur l'ensemble de test. La régression logistique étant sensible aux corrélations entre variables, cela dégrade ses performances. Sur l'ensemble dynamique, on remarque pour Car 1 que certaines corrélations entre des variables changent de signe. Avec une régularisation insuffisante, cela pourrait expliquer le mauvais score obtenu. Le ta-

bleau 1 présente les scores AUC obtenus avec les trois méthodes de classification, testées sur T1 et T2, avec les deux ensembles de caractéristiques. On notera « Car i -stat » le score obtenu sur T1 avec les caractéristiques i , et « Car i - dyn » le score obtenu sur T2 avec les caractéristiques i .

TABLE 1 – Résultats AUC

	Rég Log	SVM	GP
Car 1 - stat	0.99	0.96	0.98
Car 2 - stat	0.57	0.69	0.79
Car 1 - dyn	0.45	0.77	0.60
Car 2 - dyn	0.88	0.83	0.68

On constate une diminution des performances en utilisant uniquement l'ensemble de caractéristiques Car 1, lors de la classification des signaux d'acquisitions dynamiques. Les performances sont meilleures en utilisant l'ensemble Car 2. En effet, l'énergie du signal est une caractéristique importante, traduisant l'effet direct d'une atténuation du signal par une fissure. La perte de cette information dégrade les performances des algorithmes de classification. Néanmoins, l'identification des défauts reste toujours possible. Selon les conditions de la prise de mesure, il est nécessaire de changer le seuil qui délimite l'appartenance aux classe « défaut » ou « non-défaut » à partir du score donné. En dynamique, des acquisitions bruitées et moins précises sont attendues. Le seuil doit donc être diminué. De plus, en acquisition dynamique, le lift-off n'est pas maintenu fixe tout au long du rail. L'amplitude maximale du signal est plus susceptible de changer d'une acquisition à une autre. L'utilisation des caractéristiques Car 2 permet donc de fournir plus de robustesse à la classification, notamment lors de l'utilisation des méthodes SVM et des GP. Les bonnes performances des SVM montrent qu'un modèle linéaire serait suffisant pour cette tâche de classification.

5 Conclusion

Dans cet article, nous avons montré que la classification de signaux EMAT dans un contexte NDT des rails est possible. La classification en utilisant des caractéristiques simples calculées à partir du signal temporel et du spectre donne des résultats positifs. Néanmoins, la recherche d'une méthodologie plus robuste aux changements d'amplitude d'une acquisition à l'autre, nous a conduit à utiliser d'autres caractéristiques calculées à partir du spectre. Leur utilisation a dégradé les performances des algorithmes de classification dans le cas statique, où les conditions d'acquisition sont constantes durant l'expérience. Dans le cas dynamique, le lift-off est variable, et l'ensemble Car 2 permet d'obtenir cette fois de meilleurs résultats que l'ensemble Car 1. Les capacités de généralisation ont ainsi augmenté grâce à ces nouvelles variables. Une piste d'amélioration serait de pouvoir fusionner les résultats des deux méthodes, en insérant une information sur le lift-off, par un autre capteur.

Ces résultats ouvrent la voie à mesures à plus hautes vitesses, où d'autres types de perturbations sont attendus.

Ce travail fait l'objet d'un co-financement de la région Hauts-de-France - Eurotunnel - Railenium dans le cadre d'un projet entre Eurotunnel et Railenium.

Références

- [1] M. Ph Papaelias, C. Roberts, and C. L. Davis. *A review on non-destructive evaluation of rails : State-of-the-art and future development*. Proceedings of the Institution of Mechanical Engineers, Part F : Journal of Rail and Rapid Transit, vol. 222, no. 4, pp. 367–384, 2008.
- [2] A. Sophian, G. Tian, and M. Fan. *Pulsed eddy current non-destructive testing and evaluation : A review*. Chinese Journal of Mechanical Engineering, vol. 30, no. 3, pp. 500–514, 2017.
- [3] T. Szugs, A. Krüger, G. Jansen, B. Beltman, S. Gao, H. Mümmel, R. Ahlbrink. *Combination of ultrasonic and eddy current testing with imaging for characterization of rolling contact fatigue*. 19th World Conference on Non-Destructive Testing, 2016.
- [4] L. Oukhellou. *Paramétrisation et classification de signaux en contrôle non destructif. Application à la reconnaissance des défauts de rails par courants de foucault*. Université Paris Sud - Paris XI, 1997.
- [5] M. Mršnik, J. Slavič, and M. Boltežar. *Frequency-domain methods for a vibration-fatigue-life estimation - Application to real data*. International Journal of Fatigue, vol. 47, pp. 8–17, 2013.
- [6] M. Bentoumi, G. Millérioux, G. Bloch, L. Oukhellou, P. Aknin. *Classification de Défauts de Rail par SVM*. Congrès International IEEE Signaux, Circuits et Systèmes SCS'04, Monastir, Tunisie, p.18-21, mars 2004
- [7] Carl Edward Rasmussen and Christopher K. I Williams. *Gaussian processes for machine learning*. MIT Press, chap. 3 & chap. 5, 2008
- [8] B. Ndao. *Génération et détection sans contact des ondes de Rayleigh par méthodes ultrasons-laser et EMAT en mode statique et dynamique : application à la détection de défauts surfaciques dans le champignon du rail*. Thèse de doctorat. Électronique, Acoustique et télécommunications. Valenciennes, 2016
- [9] Christopher M. Bishop. *Pattern recognition and machine learning*. Springer, chap. 4, 2009
- [10] J.L. Rose, C.M. Lee, T.R. Hay, Y. Cho, I.K. Park. *Rail Inspection with Guided Waves*. Asia-Pacific Conference on NDT, Auckland, New Zealand, 5th - 10th Nov, 2006