

Détection de ruptures en ligne par estimateur SVM de support de densité

Frédéric DESOBRY, Manuel DAVY

Institut de Recherche en Communication et Cybernétique de Nantes (IRCCyN)
1 rue de la Noë, BP 92101, 44321 Nantes Cedex 01, France
{Frederic.Desobry, Manuel.Davy}@irccyn.ec-nantes.fr

Résumé – Nous proposons une méthode d'apprentissage automatique visant à détecter des ruptures dans une suite de vecteurs. Notre algorithme appartient à la famille des méthodes dites à noyau et repose sur la transposition du problème étudié dans un espace hilbertien muni d'un noyau reproduisant. Un indice de stationarité est construit dans cet espace, où l'on compare deux cercles (qui résultent de l'estimation de deux supports de densité) à l'aide d'un critère de type Fisher. Une rupture correspond à une distance importante entre les centres des deux cercles (relativement à leur rayon). Nous montrons que cet indice peut en fait être calculé dans l'espace de départ; des simulations illustrent l'efficacité de la méthode sur des données réelles.

1 Introduction

La détection de ruptures est un problème de traitement du signal classique, et de nombreuses approches y ont été consacrées. Certaines reposent sur la connaissance d'un modèle statistique du signal; les techniques basées sur le rapport de vraisemblance généralisé (RVG) [1] et la théorie de la détection bayésienne font montre d'excellentes performances. Cependant, il existe certaines applications pour lesquelles il est difficile d'élaborer un modèle statistique précis du signal: des approches sans modèle doivent alors être envisagées. Nous proposons dans cet article un algorithme non-paramétrique de détection de ruptures en ligne.

Cet algorithme est en deux étapes. En premier lieu, des descripteurs \vec{x}_t (ou vecteurs) localisés en temps sont extraits en ligne du signal étudié. Il peut par exemple s'agir de coefficients cepstraux calculés sur une fenêtre glissante, d'une transformée de Fourier à court terme, etc. Puis nous définissons un indice de stationarité $I(t)$ calculé dans l'espace des descripteurs (appelé par la suite espace des données, et noté \mathcal{X}), et défini de façon géométrique dans un espace transformé \mathcal{F} induit par un noyau défini positif $k(\cdot, \cdot)$ ¹.

Cet article est organisé comme suit. Dans la section suivante, nous présentons succinctement l'estimation de support de densité par SVM. Nous décrivons notre algorithme de détection de ruptures en section 3; en particulier, nous expliquons de quelle façon l'indice de stationarité est défini dans l'espace transformé, puis calculé dans l'espace des données.

2 L'estimation SVM de support de densité

Soit un ensemble d'apprentissage $\mathbf{x} = (\vec{x}_1, \dots, \vec{x}_m)$ composé de m points de l'espace des données \mathcal{X} . Nous définissons un algorithme d'apprentissage \mathcal{A} par :

$$\mathcal{A} : \begin{array}{l} \cup_{m=1}^{\infty} \mathcal{X}^m \longrightarrow \mathcal{H} \\ \{\vec{x}_1, \dots, \vec{x}_m\} \longmapsto \mathcal{A}(\mathbf{x}) \end{array}$$

où \mathcal{H} est l'espace des fonctions indicatrices $\mathbf{I}_{\mathcal{R}}$ telles que $\mathbf{I}_{\mathcal{R}}(\vec{x}) = 1$ si $\vec{x} \in \mathcal{R}$ et 0 si $\vec{x} \notin \mathcal{R}$, pour tout sous-ensemble \mathcal{R} de \mathcal{X} . Puis nous définissons une application ϕ de \mathcal{X} dans un espace transformé \mathcal{F} :

$$\phi : \begin{array}{l} \mathcal{X} \longrightarrow \mathcal{F} \\ \vec{x} \longmapsto \mathbf{x} = \phi(\vec{x}) \end{array}$$

L'espace \mathcal{F} est muni d'un produit scalaire $\langle \cdot, \cdot \rangle$, et restreint \mathcal{H} à l'ensemble des fonctions sur des régions de décision de la forme $\mathcal{R} = \{\mathbf{x} : f_{\mathbf{x}}(\vec{x}) > 0\}$ où la fonction de décision $f_{\mathbf{x}}$ est définie de la façon suivante :

$$f_{\mathbf{x}} : \begin{array}{l} \mathcal{X} \longrightarrow \mathbb{R} \\ \vec{x} \longmapsto f_{\mathbf{x}}(\vec{x}) = \langle \mathbf{x}, \mathbf{w} \rangle - \rho \end{array} \quad (1)$$

Dans l'éq. (1), \mathbf{w} est combinaison linéaire des images des vecteurs d'apprentissage,

$$\mathbf{w} = \sum_{i=1}^m \alpha_i \mathbf{x}_i, \quad \text{avec } \alpha_i \geq 0 \quad \forall i = 1, \dots, m. \quad (2)$$

Les paramètres \mathbf{w} et ρ caractérisent $\mathcal{A}(\mathbf{x})$ pour un ensemble d'apprentissage donné \mathbf{x} , et se déterminent en résolvant

$$\max_{\mathbf{w}, \xi, \rho} -\frac{1}{2} \|\mathbf{w}\|^2 - \frac{1}{\nu m} \sum_{i=1}^m \xi_i + \rho \quad (3)$$

sous les contraintes : $\langle \mathbf{w}, \mathbf{x}_i \rangle \geq \rho - \xi_i, \quad \xi_i \geq 0.$

où $0 \leq \nu \leq 1$. Sous forme duale, nous obtenons $\alpha_i, i = 1, \dots, m$ en minimisant (par rapport aux α_i et ρ)

$$W = +\frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j \langle \mathbf{x}_i, \mathbf{x}_j \rangle \quad \text{sous les contraintes}$$

$$0 \leq \alpha_i \leq \frac{1}{\nu m} \quad \text{pour } i = 1, \dots, m \quad \text{et} \quad \sum_{i=1}^m \alpha_i = 1.$$

¹Au sujet de la relation entre noyaux définis positifs et espace transformé, voir, par exemple [2].

Tous les vecteurs d'apprentissage \vec{x}_i pour lesquels $f_{\mathbf{x}}(\vec{x}_i) \leq 0$ sont appelés *vecteurs-support* (SV) ; seuls les poids α_i correspondant à ces points ne sont pas nuls ; les SV suffisent donc à déterminer $f_{\mathbf{x}}(\cdot)$. Les SV se divisent en deux ensembles : les *SV de marge*, qui annulent la fonction de décision ($f_{\mathbf{x}}(\vec{x}_i) = 0$), et les *SV hors-marge* qui sont tels que $f_{\mathbf{x}}(\vec{x}_i) < 0$. On peut montrer [3] que ν majore le taux de SV hors-marge et minore le taux de SV ; de plus, ces deux taux sont asymptotiquement égaux à ν avec la probabilité 1 (en faisant des hypothèses faibles sur la distribution à l'origine des données).

Enfin, notons qu'il n'est jamais nécessaire de calculer $\phi(\vec{x})$: en effet, l'éq. (2) implique que le calcul de $\langle \mathbf{w}, \mathbf{x}_j \rangle$ ne fait intervenir les images des données d'apprentissage qu'à travers leur produit scalaire $\langle \mathbf{x}_i, \mathbf{x}_j \rangle = k(\vec{x}_i, \vec{x}_j)$. Un noyau k est un produit scalaire dans quelque espace transformé s'il répond aux conditions de Mercer [4]. Ces conditions sont satisfaites par un grand nombre de noyaux, en particulier le noyau gaussien,

$$k(\vec{x}_i, \vec{x}_j) = \exp\left(-\frac{1}{2\sigma^2} \|\vec{x}_i - \vec{x}_j\|^2\right). \quad (4)$$

En conclusion, la région \mathcal{R} contient les SV de marge et les non-SV, et définit une estimation du support de la densité à l'origine des points de l'ensemble d'apprentissage.

3 Algorithme de détection de ruptures en ligne

Nous introduisons dans cette section un algorithme de détection en ligne de ruptures dans la distribution des vecteurs \vec{x}_t . La suite des \vec{x}_t peut être directement la sortie d'un système ; mais dans les applications usuelles de traitement du signal, elle provient plus typiquement d'une étape de prétraitement, comme l'extraction de descripteurs.

3.1 Description de l'algorithme

Considérons à l'instant t deux sous-ensembles de \mathbf{x} : $\mathbf{x}_1 = (\vec{x}_{t-m_1}, \dots, \vec{x}_{t-1})$ et $\mathbf{x}_2 = (\vec{x}_t, \dots, \vec{x}_{t+m_2-1})$. L'estimation SVM du support de la densité à l'origine de \mathbf{x}_1 et \mathbf{x}_2 est effectuée via le calcul des paramètres \mathbf{w}_1, ρ_1 et \mathbf{w}_2, ρ_2 , et définit les régions de décision \mathcal{R}_1 et \mathcal{R}_2 . L'idée sous-jacente est qu'une rupture à l'instant T dans la distribution des vecteurs \vec{x}_t sera reflétée par des différences de localisation entre les régions \mathcal{R}_1 et \mathcal{R}_2 : en pratique, nous construisons un indice $I(t)$ qui exprime la dissimilarité entre \mathbf{x}_1 et \mathbf{x}_2 via une mesure de la dissimilarité entre \mathcal{R}_1 et \mathcal{R}_2 . (Le calcul de $I(t)$ pour t fixé est décrit dans la sous-section suivante). Une fois $I(t)$ calculé, on incrémente t et l'on met à jour les sous-ensembles \mathbf{x}_1 et \mathbf{x}_2 . Les paramètres SVM \mathbf{w}_1, ρ_1 et \mathbf{w}_2, ρ_2 sont également mis à jour, par la technique séquentielle présentée dans [5]. Cette technique permet d'éviter de recalculer complètement \mathbf{w}, ρ pour chaque nouveau t .

Des ruptures sont finalement détectées chaque fois que l'indice $I(t)$ présente un pic supérieur au seuil η (cette étape est commune à la plupart des méthodes de détection de ruptures). Dans la sous-section suivante, nous montrons comment des considérations géométriques dans l'espace transformé permettent de définir un nouvel indice de stationnarité.

3.2 Géométrie de l'espace transformé

Soit k un noyau défini positif et normalisé : $k(\vec{x}, \vec{x}) = 1, \forall \vec{x} \in \mathcal{X}$. Les images des vecteurs d'apprentissage $\mathbf{x}_t = \phi(\vec{x}_t)$ sont situées sur l'hypersphère \mathcal{S} de rayon 1 et centrée sur l'origine $\mathbf{0}$ de \mathcal{F} . Considérons un ensemble d'apprentissage, par exemple \mathbf{x}_1 ; le problème d'optimisation défini par l'éq. (3) admet l'interprétation géométrique suivante. Les paramètres \mathbf{w}_1, ρ_1 définissent un hyperplan \mathcal{W}_1 de \mathcal{F} , orthogonal à \mathbf{w}_1 et distant de $\rho_1/\|\mathbf{w}_1\|$ de $\mathbf{0}$ (voir fig. 1). \mathcal{W}_1 sépare les images des vecteurs d'apprentissage de l'origine $\mathbf{0}$. Dans \mathcal{F} , l'image de \mathcal{R}_1 est la portion de \mathcal{S} délimitée par \mathcal{W}_1 , et située dans le demi-espace ne contenant pas $\mathbf{0}$. Sa frontière est un hypercercle \mathcal{C}_1 de \mathcal{F} . L'hypersphère \mathcal{S} étant de rayon 1, \mathcal{C}_1 est entièrement défini par \mathbf{w}_1, ρ_1 . De même, pour \mathbf{x}_2 , la résolution de l'éq. (3) conduit à un hyperplan \mathcal{W}_2 et un hypercercle \mathcal{C}_2 défini par \mathbf{w}_2, ρ_2 .

Les régions images de \mathcal{R}_1 et \mathcal{R}_2 dans \mathcal{F} sont de forme géométrique simple (leurs frontières sont des hypercercles \mathcal{C}_1 et \mathcal{C}_2). Elles peuvent donc être comparées via leurs centre et rayon respectifs, qui s'expriment en fonction de \mathbf{w}_1, ρ_1 et \mathbf{w}_2, ρ_2 .

Une façon simple de comparer \mathcal{C}_1 et \mathcal{C}_2 (i.e., \mathcal{R}_1 et \mathcal{R}_2), et donc de construire $I(t)$, consiste à considérer un rapport des distances inter-régions et intra-régions dans \mathcal{F} , tel :

$$I(t) = \frac{\text{distance entre les centres des cercles } \mathcal{C}_1 \text{ et } \mathcal{C}_2}{\text{rayon de } \mathcal{C}_1 + \text{rayon de } \mathcal{C}_2}$$

Si $\mathbf{w}_1 = \mathbf{w}_2$, les deux ensembles d'apprentissage sont situés au même endroit, et $I(t) = 0$. Il s'agit bien de ce que nous souhaitons obtenir dans une telle situation. Nous laissons de côté la situation théorique pathologique où \mathbf{w}_1 et \mathbf{w}_2 sont simplement colinéaires, les centres des cercles étant confondus pour des rayons évoluant différemment avec le temps. Supposons désormais que \mathbf{w}_1 et \mathbf{w}_2 ne soient pas colinéaires. Il existe alors au moins un triplet $(\mathbf{p}_1, \mathbf{p}_2, \mathcal{C})$ tel que :

$$\begin{cases} \mathbf{p}_1 \in \mathcal{S} \cap \mathcal{W}_1, & \mathbf{p}_2 \in \mathcal{S} \cap \mathcal{W}_2 \\ \mathcal{C} \text{ est un cercle de rayon 1 et de centre } \mathbf{0} \\ \mathbf{c}_1, \mathbf{c}_2, \mathbf{p}_1, \text{ et } \mathbf{p}_2 \text{ sont situés sur } \mathcal{C}. \end{cases} \quad (5)$$

où \mathbf{c}_1 (respectivement \mathbf{c}_2) est le point situé à l'intersection de la droite passant par $\mathbf{0}$ et orientée selon \mathbf{w}_1 (respectivement \mathbf{w}_2) et \mathcal{S} (voir fig. 1).

\mathbf{c}_1 et \mathbf{c}_2 sont les centres (géométriques) des deux régions occupées par les images par ϕ des vecteurs d'apprentissage, et \mathbf{p}_1 et \mathbf{p}_2 possèdent des propriétés identiques aux SV de marge puisqu'ils sont sur les hyperplans séparateurs. Pourtant il est possible que ces quatre points ne possèdent pas d'antécédent par ϕ dans l'espace des données. $I(t)$ peut cependant s'exprimer en termes de produits scalaires dans l'espace transformé, donc à l'aide de noyaux dans l'espace des données. $I(t)$ peut en effet s'écrire

$$I(t) = \frac{|\widehat{\mathbf{c}_1 \mathbf{c}_2}|}{|\widehat{\mathbf{c}_1 \mathbf{p}_1}| + |\widehat{\mathbf{c}_2 \mathbf{p}_2}|} \quad (6)$$

où $\widehat{\cdot}$ est l'arc de cercle. Comme \mathcal{S} est de rayon 1, le numérateur est de valeur égale à l'angle entre \mathbf{c}_1 et \mathbf{c}_2 , et

$$\widehat{\mathbf{c}_1 \mathbf{c}_2} = \widehat{\mathbf{c}_1 \mathbf{0} \mathbf{c}_2} = \cos^{-1}\left(\frac{\langle \mathbf{w}_1, \mathbf{w}_2 \rangle}{\|\mathbf{w}_1\| \cdot \|\mathbf{w}_2\|}\right) \quad (7)$$

En utilisant l'éq. (2), nous avons

$$\frac{\langle \mathbf{w}_1, \mathbf{w}_2 \rangle}{\|\mathbf{w}_1\| \cdot \|\mathbf{w}_2\|} = \frac{\sum_{i,j} \alpha_{1i} \alpha_{2j} k(\vec{x}_{1i}, \vec{x}_{2j})}{\left(\sum_{i,j} \alpha_{1i} \alpha_{1j} k(\vec{x}_{1i}, \vec{x}_{1j}) \cdot \sum_{i,j} \alpha_{2i} \alpha_{2j} k(\vec{x}_{2i}, \vec{x}_{2j})\right)^{0.5}}$$

où α_{1i} (resp. α_{2i}) correspond à \mathbf{w}_1 et \vec{x}_{1i} (resp. \mathbf{w}_2 et \vec{x}_{2i}). Des calculs similaires mènent à l'expression de l'arc de cercle entre \mathbf{c}_1 et \mathbf{p}_1 :

$$\mathbf{c}_1 \widehat{\mathbf{p}}_1 = \cos^{-1} \left(\frac{\rho_1}{\left(\sum_{i,j} \alpha_{1i} \alpha_{1j} k(\vec{x}_{1i}, \vec{x}_{1j}) \right)^{0.5}} \right)$$

Nous obtenons un résultat semblable pour $\mathbf{c}_2 \widehat{\mathbf{p}}_2$, ce qui montre que I est une fonction de t pouvant être calculée dans \mathcal{X} en utilisant uniquement $k(\cdot, \cdot)$.

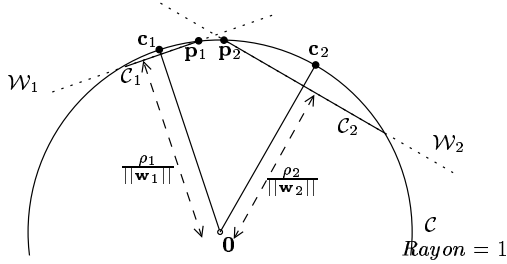


FIG. 1 – Dans l'espace transformé \mathcal{F} , les images des vecteurs d'apprentissage sont situées sur l'hypersphère \mathcal{S} de centre \mathbf{O} et de rayon 1. L'estimateur de densité sur \mathbf{x}_1 (resp. \mathbf{x}_2) fournit un hypercercle \mathcal{C}_1 (resp. \mathcal{C}_2) dans l'hyperplan \mathcal{W}_1 (resp. \mathcal{W}_2).

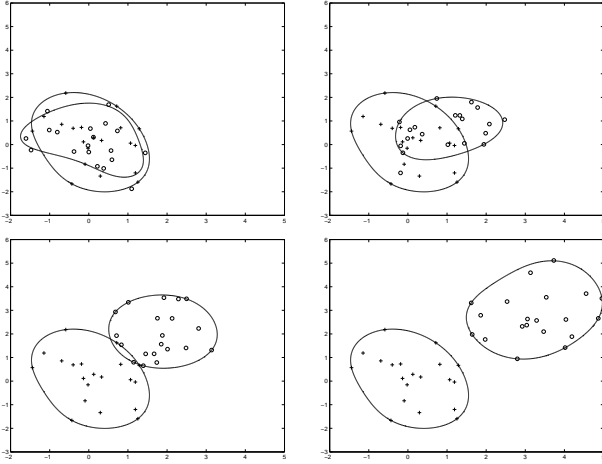


FIG. 2 – Dans l'espace des données, évolution des séparatrices fournies par les deux estimateurs de densité lorsque la moyenne de l'un des deux ensembles change (dans l'ordre : de gauche à droite puis de haut en bas). Deux surfaces avec un faible recouvrement correspondent à un changement.

3.3 Discussion

On déduit de l'éq. (6) que $I(t)$ est un rapport de type Fisher. Des ruptures seront donc détectées si elles sont liées à des changements de moyenne (terme $\mathbf{c}_1 \widehat{\mathbf{c}}_2$), relativement à l'échelle des données (obtenue grâce à $\mathbf{c}_1 \widehat{\mathbf{p}}_1$ et $\mathbf{c}_2 \widehat{\mathbf{p}}_2$). Notons que $I(t)$ est robuste aux outliers, puisque l'estimateur SVM de support de densité considère systématiquement que certains vecteurs d'apprentissage sont des outliers (le nombre de SV hors-marge est contrôlé par le paramètre ν). Bien que fondé sur les mêmes principes que le coefficient de Rayleigh utilisé dans le discriminant de Fisher par noyau (KFD) [2], notre coefficient $I(t)$ est différent en ce sens que dans le cadre de la KFD, le coefficient de Rayleigh est maximisé par rapport à la direction de projection \mathbf{w} alors que dans notre méthode, il résulte du pro-

blème d'optimisation SVM.

On peut également considérer une approche biclasse, où \mathbf{x}_1 et \mathbf{x}_2 sont vus comme des ensembles d'apprentissage correspondant à deux classes différentes. En entraînant un algorithme SVM de classification biclasse sur ces deux ensembles, il est possible d'interpréter la marge entre les deux classes comme un indice de stationnarité. Cette approche présente cependant plusieurs défauts : tout d'abord, lorsqu'aucune rupture n'a lieu, \mathbf{x}_1 et \mathbf{x}_2 sont localisés au même endroit ; il est donc artificiel de construire une séparatrice dans ces conditions. Deuxièmement, des vecteurs considérés comme des outliers dans l'approche classification ne le seront pas forcément en utilisant deux estimateurs monoclasses : dans le premier cas en effet, la façon dont les outliers sont sélectionnés dans un ensemble (par exemple, \mathbf{x}_1) est gouvernée par l'autre ensemble (par exemple, \mathbf{x}_2), et pas par leur comportement intrinsèque vis-à-vis du processus \vec{x}_t . Troisièmement, le coût algorithmique est plus lourd avec la classification SV biclasse puisque l'apprentissage avec $m_1 + m_2$ vecteurs (dans $\mathbf{x}_1 \cup \mathbf{x}_2$) est plus coûteux que l'apprentissage séparé de m_1 , puis m_2 vecteurs.

Finalement, l'approche proposée en [6] est en relation avec un problème légèrement différent, mais en rapport toutefois : la détection précoce d'anormalités (lorsqu'aucun retard à la détection n'est toléré). Dans [6], un candidat \vec{x}_t est testé par un unique détecteur de nouveauté entraîné sur l'ensemble d'apprentissage $\{\vec{x}_{t-m}, \dots, \vec{x}_{t-1}\}$. Un indice de stationnarité est construit avec la sortie du détecteur :

$$I_o(t) = \langle \mathbf{w}_o(t), \phi(\vec{x}_t) \rangle - \rho_o(t) \quad (8)$$

Des simulations sur des signaux de musique où les \vec{x}_t sont des transformées de Fourier à court terme ont fourni de bons résultats, même si la détection précoce d'anormalité n'était pas adéquate dans le contexte de la segmentation de musique, où un retard à la détection peut être accepté.

4 Simulations

Dans cette section nous mettons en œuvre la méthode SVM de détection de ruptures et comparons ses performances à celles de deux autres techniques de détection de ruptures : le Rapport de Vraisemblance Généralisé (RVG, voir par exemple [1]), et une technique reposant sur des distances entre imagerie temps-fréquence [7]. Contrairement au RVG, cette dernière technique est une approche sans modèle. Nous testons ces trois méthodes sur des signaux artificiels et sur des signaux réels.

Les données artificielles sont composées de 2000 réalisations d'un bruit blanc gaussien (2048 points) filtrées par un AR d'ordre 4 de module 0.99 et dont les fréquences sont tirées uniformément entre 0.05 et 0.45. Les 1000 premières réalisations voient les paramètres de l'AR brusquement modifiés lorsque l'on atteint la moitié du signal ; pour les 1000 suivantes, ils ne présentent aucun changement.

L'algorithme SVM réalise son apprentissage sur deux ensembles de descripteurs, tous deux extraits de la pseudo Wigner Ville lissée (fenêtre temporelle : 25 points, fenêtre fréquentielle : 67 points) du signal de départ, qui consiste en une réalisation du bruit filtré. Chaque vecteur d'apprentissage est une imagerie temps-fréquence de 12 colonnes, et chaque ensemble d'apprentissage en comporte 20. Les paramètres SVM proprement dits sont : $\sigma = 1.5$ pour le noyau gaussien, et $\nu = 0.2$.

La même représentation temps-fréquence est utilisée pour la méthode à base de distances entre imassettes TF, avec des imassettes de largeur 150, et le choix de la distance de Kolmogorov [7]. Nous fournissons au RVG la forme correcte du modèle : AR, avec plusieurs ordres (3, 4 et 6).

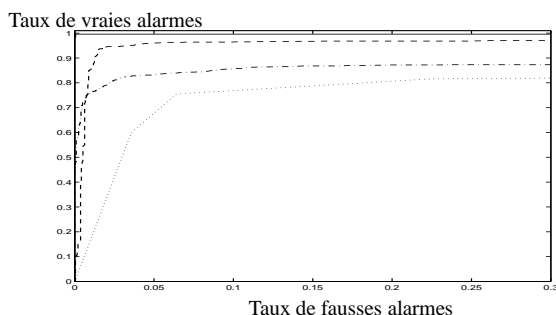


FIG. 3 – Courbes COR pour la méthode à base de distances entre imassettes temps-fréquence (pointillés), RVG avec un ordre inférieur (mixte), méthode SVM (tiret), et RVG avec un ordre correct ou supérieur (trait plein). Ces deux dernières méthodes offrent de bons résultats.

Nous traçons en figure 3 les courbes COR correspondant aux différentes méthodes. Le RVG muni d'un AR d'ordre 4 (ordre correct) ou 6 (ordre supérieur) présente de très bonnes performances. Le détecteur SVM se comporte correctement avec par exemple 95% de vrais positifs pour 2% de faux positifs, alors qu'il ne possède aucune information a priori sur les données analysées. La méthode à base de distances TF et le RVG avec un modèle d'ordre inférieur donnent de mauvais résultats. Nous pouvons observer des résultats semblables sur les histogrammes des instants de détection des vrais positifs, représentés en figure 4. Notons que bien que se servant des mêmes descripteurs, la méthode SVM fournit de bien meilleurs résultats que la méthode à base de distances TF.

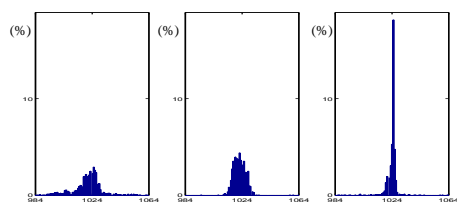


FIG. 4 – Histogrammes des instants estimés de rupture : % d'instants détectés dans le voisinage admissible de l'instant théorique de rupture (1024) pour la méthode à base de distances entre imassettes (à gauche), la méthode SVM (au centre), et le RVG avec le bon ordre du modèle (à droite).

Nous avons également mis en œuvre la méthode SVM sur des signaux de musique, où elle s'est montrée efficace. La figure 5 expose le traitement du même signal de musique qu'en [6]. Tous les changements de dynamiques sont correctement détectés (ce qui est confirmé par l'écoute du morceau), ce qui n'était pas le cas avec la méthode à base de distances TF, voir [6]. De plus notre indice est lisse, avec des pics marqués, et un bon contraste entre ces pics et le reste du signal.

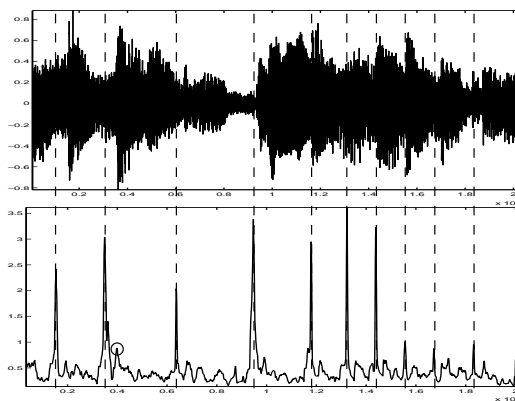


FIG. 5 – Signal de musique (haut) et l'indice de détection de rupture SVM associé (bas). Pour un seuil η valant 0.85, tous les changements (pointillés) sont correctement détectés, avec un seul faux positif (cercle).

5 Conclusion

Dans ce papier, nous avons introduit une technique à noyau originale pour réaliser la détection en-ligne de rupture dans des signaux. Des simulations ont montré son bon comportement comparativement à d'autres algorithmes sur des données synthétiques, et de meilleures performances sur des signaux de musique, à la fois en termes de précision de la détection, et de contraste de l'indice.

Références

- [1] M. Basseville and I. Nikiforov, *Detection of Abrupt Changes - Theory and Application*, Prentice-Hall, April 1993.
- [2] A. Smola and B. Schölkopf, *Learning with Kernels*, MIT press, 2002.
- [3] B. Schölkopf, J.C. Platt, J. Shawe-Taylor, A.J. Smola, and R.C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural Computation*, vol. 13, no. 7, pp. 1443–1471, 2001.
- [4] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, N.Y., 1995.
- [5] A. Gretton and F. Desobry, "On-line one-class ν -support vector machines. an application to signal segmentation," in *IEEE ICASSP*, Hong-Kong, China, April 2003.
- [6] M. Davy and S. Godsill, "Detection of abrupt spectral changes using support vector machines. an application to audio signal segmentation," in *IEEE ICASSP-02*, Orlando, USA, May 2002.
- [7] H. Laurent and C. Doncarli, "Stationarity index for abrupt changes detection in the time-frequency plane," *IEEE Signal Processing Letters*, vol. 5, no. 2, pp. 43 – 45, February 1998.