

Détection statistique temps-fréquence dans un contexte acoustique à faible RSB

Florian DADOUCHI¹, Avelino VILLA¹, Cédric GERVAISE^{1,2}, Cornel IOANA¹, Julien HUILLERY³, Jérôme I. MARS¹

¹Université Grenoble-Alpes, Gipsa-Lab, CNRS, 38000 Grenoble, France

²Chaire CHORUS, Fondation Grenoble INP, 46 Avenue Viallet, 38031 Grenoble, France

³École Centrale de Lyon, Laboratoire Ampère, CNRS, 69134 Ecully, France

¹[prenom.nom]@gipsa-lab.grenoble-inp.fr, ²cedric.gervaise@chorusacoustics.com, ³julien.huillery@ec-lyon.fr

Résumé – Ce travail aborde la détection statistique à partir du spectrogramme de signaux non-stationnaires multi-composantes dans des environnements où le rapport signal sur bruit (RSB) est faible. Un double test d'hypothèses binaires est formulé. Le premier test correspond à un détecteur d'énergie appliqué à l'échelle du pixel temps-fréquence (TF). Pour atteindre une bonne probabilité de détection à faible RSB, ce test doit cependant accepter de nombreuses fausses alarmes. Afin de réduire ce nombre de fausses alarmes et ainsi améliorer les performances globales de détection, un second test est proposé. Ce test considère la densité de détection sur des régions TF. Sous l'hypothèse "bruit seul", nous montrons que cette densité de détection suit une loi binomiale corrélée du fait de la redondance du spectrogramme. Afin d'estimer cette loi et résoudre le second test, une méthode numérique utilisant des spectrogrammes assistants est introduite. Pour un point de fonctionnement du détecteur (Pd,Pfa), l'ajout du second test permet de réduire le RSB de 6 dB, au prix d'une perte de résolution TF. La méthode est illustrée avec succès sur des signaux réels d'acoustique passive.

Abstract – This work deals with the statistical detection of multicomponent non-stationary signals at low signal-to-noise ratio using the spectrogram. A double binary hypothesis test (BHT) framework is proposed. The first BHT is a classical time-frequency (TF) energy detector, the positives of which are composed of true detections as well as false alarms. So as to discriminate between those two kinds of positives, a second BHT is proposed based on the density of detection in the TF regions of the spectrogram. A high-density of detections is synonymous with the presence of signal, while a low density of detections is more likely due to false alarms. The Neyman-Pearson detection criterion is used for both of the BHTs. According to the redundancy of the spectrogram, the density of detection under the null hypothesis is distributed as a correlated binomial variable. We introduce the assistant spectrograms to numerically estimate this distribution and solve the second BHT. The method is illustrated with real-world signals from a passive acoustic context.

1 Introduction

La détection de signaux non-stationnaires dans des contextes de RSB faible est un problème d'intérêt dans de nombreux domaines. En acoustique sous-marine [1] ou en bioacoustique [2], la nature des signaux émis (modulations de fréquence non linéaires, impulsions large bande et "bursts") exige le développement d'outils de traitement du signal spécifiques. De nombreuses difficultés s'additionnent : 1. Non-stationnarité des signaux ; 2. Signaux multi-composantes/ Nombre de signaux inconnu/ Croisement des signaux dans le plan TF ; 3. Grande variabilité des RSB ; 4. Bruit de fond coloré et/ou non stationnaire ; 5. Grandes bases de données ; 6. Forte directivité des sources/Fading. On peut noter que pour les mêmes performances de détection, un gain de 6 dB en terme de RSB permet d'augmenter la portée de détection d'un facteur 2 et le volume surveillé d'un facteur 8. Un enjeu majeur est ensuite l'estimation des modulations de fréquence non linéaires qui forment une large classe de signaux émis par les animaux ou lors des

activités anthropiques. Dans ce but, plusieurs approches développées dans la littérature visent à détecter les pics d'une représentation TF, généralement le spectrogramme puisqu'il offre un cadre bien maîtrisé [3, 4, 5, 6]. Dans un contexte réel, un élément clé reste l'absence d'hypothèse a-priori sur le signal d'intérêt (puissance, forme, fréquence et nombre de composantes).

Ce papier considère une méthode statistique robuste de détection TF qui comprend deux tests d'hypothèses binaires (BHT). Le premier test, nommé BHT1, opère une détection à l'échelle du pixel TF en acceptant un grand nombre de fausses alarmes pour assurer une bonne probabilité de détection. Le second test, nommé BHT2, travaille à l'échelle d'une région TF, et a pour vocation de rejeter les fausses alarmes du premier test en utilisant un critère de densité de détection. Comme décrit dans [7, 8], BHT1 se base sur la modélisation des coefficients du spectrogramme liés au bruit avec des variables aléatoires du χ^2 [4] et l'utilisation du critère de Neyman-Pearson pour une probabilité de fausse alarme (pfa) donnée. Ce premier test opère en chaque pixel TF et renvoie un spectrogramme binaire. L'ori-

ginalité de la méthode présentée ici repose sur l'ajout d'un second test d'hypothèses, noté BHT2. Celui-ci ajoute une information régionale au processus de détection en considérant la densité des pics détectés dans une région TF du spectrogramme binaire. Une région avec une haute densité de détection sera considérée comme "signal" et une région avec une faible densité de détection sera considérée comme "bruit seul". Cette méthode de détection est appelée FADA (False Alarm Density Analysis). La seconde contribution de cette étude est de prendre en considération la redondance du spectrogramme : nous montrons qu'il est plus approprié de décrire la distribution de probabilité de la densité des pixels TF sélectionnés par BHT1 avec une loi binomiale corrélée. Devant la difficulté analytique de cette loi, nous proposons une méthode d'estimation numérique utilisant des spectrogrammes assistants. Cette méthode vise à reproduire la structure de dépendance entre les coefficients du spectrogramme.

2 Détection TF locale

Soit $x[n] = s[n] + r[n]$, une observation contenant un signal d'intérêt $s[n]$ corrompu par un bruit additif gaussien, non-stationnaire et coloré $r[n]$. Le spectrogramme du signal $x[n]$ est défini selon

$$S_x[n, k] = \left| \sum_{p=0}^{M-1} x[nD + p]w[p]e^{-i2\pi k \frac{p}{M}} \right|^2, \quad (1)$$

avec $w[n]$ une fenêtre d'analyse de longueur M et D l'incrément temporel entre deux analyses successives.

Le premier test BHT1 consiste à faire le choix entre l'absence ou la présence du signal en chaque point $[n, k]$ du spectrogramme, soit :

$$\text{BHT1} : \begin{cases} H_0 : [n, k] \text{ est un point "bruit seul"} \\ H_1 : [n, k] \text{ est un point "signal"} \end{cases} \quad (2)$$

Résoudre ce test est équivalent à fixer un seuil $S_{th}[n, k]$ variable en fonction du point d'étude avec lequel $S_x[n, k]$ est comparé. Utilisant un critère de Neyman-Pearson, ce seuil est donné selon [4]

$$S_{th}[n, k] = -\gamma_r[n, k] \ln(P_{fa1}), \quad (3)$$

où $\gamma_r[n, k]$ représente la densité spectrale de puissance temporelle du bruit [9] et P_{fa1} la pfa choisie pour BHT1. L'application de ce seuil sur le spectrogramme du signal conduit au spectrogramme binaire $\phi[n, k]$ obtenu selon

$$\phi[n, k] = \begin{cases} 1 \text{ si } S_x[n, k] > S_{th}[n, k] \\ 0 \text{ sinon} \end{cases} \quad (4)$$

Comme illustré sur la figure 1, le spectrogramme binaire ainsi obtenu contient des détections vraies et des fausses alarmes. Le second test d'hypothèses introduit dans la section suivante utilise une information régionale pour différencier ces deux types de détections.

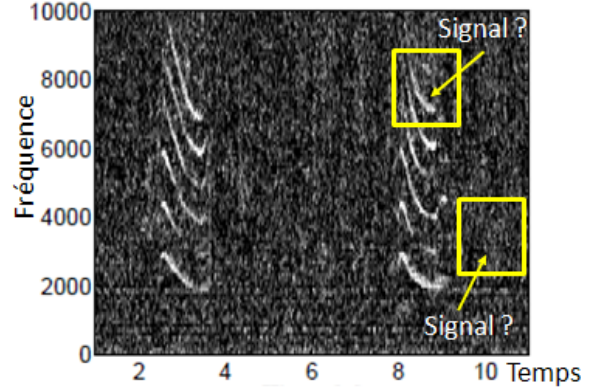


FIGURE 1 – Spectrogramme binaire $\phi[n, k]$ contenant de vraies détections et des fausses alarmes ($P_{fa1} = 0.05$).

3 Détection TF régionale

Dans cette section le deuxième test d'hypothèses BHT2 est formulé et les outils nécessaires à sa résolution sont présentés. BHT2 utilise comme entrée le spectrogramme binaire $\phi[n, k]$ obtenu par BHT1 et fournit en sortie un deuxième spectrogramme binaire $\psi[n, k]$ contenant moins de fausses alarmes et plus de détections vraies.

3.1 Test d'hypothèses binaires BHT2

Soit R une région TF arbitraire de taille A . Le test BHT2 est relatif à la question suivante : la région R contient-elle une partie du support TF du signal d'intérêt ou n'abrite-t-elle que du bruit ?

$$\text{BHT2} : \begin{cases} H_0^R : R \text{ est une région "bruit seul"} \\ H_1^R : R \text{ est une région "signal"} \end{cases} \quad (5)$$

Sur la carte de décision $\phi[n, k]$ issu de BHT1 (4), les fausses alarmes se répartissent uniformément dans le plan TF. Au contraire, les détections vraies dues aux signaux cohérents apparaissent regroupées. La décision prise pour BHT2 est alors basée sur la densité des détections observées sur $\phi[n, k]$. Une plus grande quantité de détections est attendue dans les régions TF contenant des signaux par rapport à celles ne contenant que du bruit. La densité de détection $\rho[n, k]$ dans une région R autour du point $[n, k]$ est définie selon :

$$\rho[n, k] = \frac{1}{A} \sum_{[m, l] \in R} \phi[m, l]. \quad (6)$$

De façon identique à BHT1, BHT2 est résolu avec un critère de Neyman-Pearson en comparant la densité ρ avec un seuil ρ_{th} qui dépendra de la pfa choisie (notée P_{fa2} pour ce second test). La carte de décision $\psi[n, k]$ issue de BHT2 est obtenue selon

$$\psi[n, k] = \begin{cases} 1 \text{ si } \rho[n, k] > \rho_{th}[n, k] \\ 0 \text{ sinon} \end{cases} \quad (7)$$

Dans les deux sous-sections suivantes, on s'intéresse à la détermination du seuil ρ_{th} à partir de la distribution de probabilité de ρ sous l'hypothèse H_0^R .

3.2 Analyse de la densité des fausses alarmes

Le premier test BHT1 utilise un critère de Neyman-Pearson avec une pfa constante. Sous H_0 , les variables aléatoires $\phi[n, k]$ obtenues selon (4) sont donc identiquement distribuées selon une loi de Bernoulli de paramètre P_{fa1} , soit

$$\begin{cases} p_{H_0}(\phi = 1) = \mathbb{P}\{S_x > S_{th}|H_0\} = P_{fa1} \\ p_{H_0}(\phi = 0) = 1 - P_{fa1} \end{cases} \quad (8)$$

Notons que le caractère non-blanc et non-stationnaire du bruit $r[n]$ est pris en compte dans le calcul du seuil de détection $S_{th}[n, k]$ donné en (3). En conséquence la distribution de $\phi[n, k]$ est indépendante des propriétés TF du bruit (sous réserve de la connaissance ou d'une bonne estimation de $\gamma_r[n, k]$).

Si l'on suppose que, dans le cas "bruit seul", les coefficients du spectrogramme sont indépendants, alors les variables aléatoires $\phi[n, k]$ sont également indépendantes. Sous ces conditions la densité ρ définie en (6) suit, sous l'hypothèse H_0^R , une loi binomiale $\mathcal{B}(P_{fa1}, A)$ normalisée par un facteur $1/A$. Sa loi de probabilité s'écrit, $\forall q \in \{0, 1, \dots, A\}/A$

$$p_{H_0^R}(\rho = q) = \binom{A}{Aq} P_{fa1}^{Aq} (1 - P_{fa1})^{A-Aq}. \quad (9)$$

Cependant, en raison de la redondance intrinsèque du spectrogramme cette hypothèse d'indépendance n'est pas valable. Par construction (forme de la fenêtre d'analyse $w[n]$ et recouvrement, voir (1)), les coefficients du spectrogramme sont généralement temporellement et fréquentiellement corrélés. Afin de prendre en considération cette redondance, il devient nécessaire de modéliser la densité de fausses alarmes dans une région R "bruit seul" avec une loi binomiale corrélée. L'expression analytique de cette loi de probabilité fait intervenir les moments d'ordres supérieurs des coefficients du spectrogramme [10]. Ces moments étant en pratique difficiles à obtenir, une méthode numérique pour estimer cette loi à partir de spectrogrammes assistants est présentée dans la section suivante. À partir de la loi obtenue, le choix de P_{fa2} permet de déterminer le seuil ρ_{th} utilisé dans (7).

3.3 Estimation par les spectrogrammes assistants

L'idée est de recréer la structure de corrélation des coefficients du spectrogramme en appliquant à un bruit blanc gaussien la même transformation que celle utilisée sur le signal étudié. Le processus est le suivant :

- (1) Génération d'un bruit blanc gaussien centré de variance unité.
- (2) Calcul de son spectrogramme avec les mêmes paramètres que ceux utilisés pour l'analyse du signal.
- (3) Application de BHT1 (on obtient ici uniquement des fausses alarmes).
- (4) Calcul de la carte des densités de détection $\rho[n, k]$ sur les régions R .
- (5) Estimation par histogramme de la loi de probabilité de ρ sous l'hypothèse H_0^R .

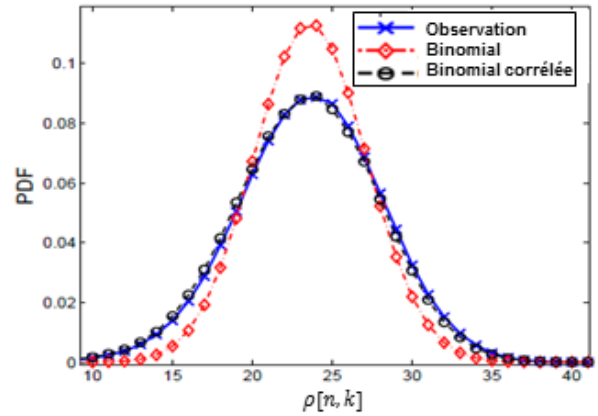


FIGURE 2 – Estimation de la distribution de $\rho[n, k]$ sous l'hypothèse H_0^R avec les spectrogrammes assistants à partir de données réelles (ERATO-09, SHOM, 45 secondes d'enregistrement) et comparaison avec le modèle binomial décorréolé donné en (9).

Il est fondamental de noter qu'il n'est pas nécessaire d'utiliser en entrée de cette méthode un signal identique au bruit $r[n]$ contenu dans le signal $x[n]$ étudié. En effet, comme indiqué dans la section 3.2, la distribution des variables $\phi[n, k]$ est indépendante des statistiques temps-fréquence du bruit et seule la structure de corrélation induite par l'outil d'analyse (le spectrogramme) influence la loi de probabilité de la densité de détection ρ .

La figure 2 illustre la pertinence de cette méthode d'estimation sur un signal réel de bruit sous-marin. Le spectrogramme est seuillé avec $P_{fa1} = 0.5$, les détections sont comptées sur des régions R de taille $A = 7 \times 7$ et la loi de probabilité est estimée à partir de l'histogramme des valeurs observées. Comparée à la loi binomiale décorrélée donnée en (9), la distribution calculée via les spectrogrammes assistants estime très correctement celle des observations.

4 Illustration sur données réelles

La capacité de l'algorithme à faire la distinction entre les vraies détections et les fausses alarmes est illustrée figure 3 sur un signal acoustique de dauphins (ERATO-09, SHOM). L'exemple présenté, qui n'a pas vocation de preuve, est caractéristique des résultats obtenus avec la méthode FADA sur de grandes bases de données et illustre bien les points forts et les points faibles de cette méthode. La même pfa (10^{-3}) est appliquée pour les deux tests BHT1 et BHT2. Dans le cas BHT1 seul (figure 3-b), la vocalise entourée en rouge (RSB local de 4 dB) n'est pas détectée alors que les autres vocalises le sont (RSB local > 10 dB). Dans le cas de l'utilisation de FADA (BHT1 et BHT2 en série, figure 3-c), la vocalise est bien détectée. Globalement, la méthode FADA procure un gain de RSB de l'ordre de 6 dB. Par ailleurs, il doit être noté que le support TF s'élargit, amenant une perte de résolution en raison de l'approche

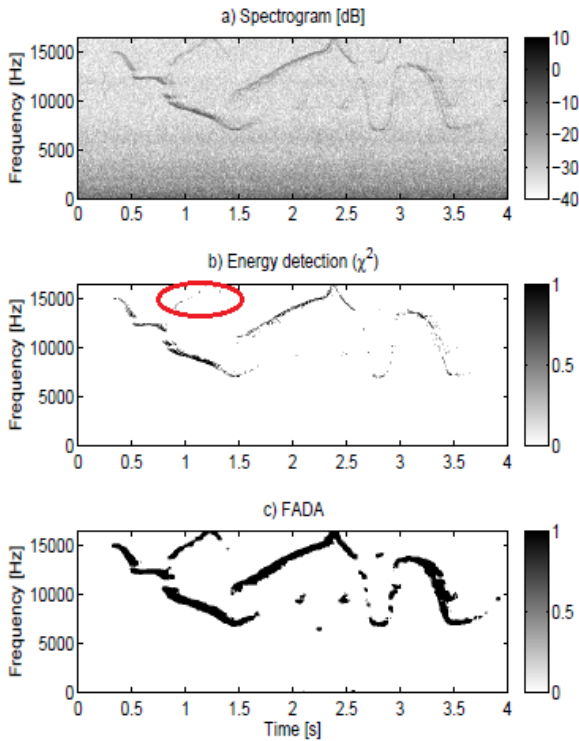


FIGURE 3 – Détection de signature acoustique de dauphins communs (*Delphis delphis*) sur un enregistrement réel d’acoustique passive réalisé dans le golfe de Gascogne : a) Spectrogramme $S_x[n, k]$ du signal, b) Résultat $\phi[n, k]$ du test BHT1 avec $P_{fa1} = 10^{-3}$, c) Résultat $\psi[n, k]$ du test BHT2 avec $P_{fa2} = 10^{-3}$.

régionale (ici R est de taille $A = 7 \times 7$ pixels TF). Toutefois, cette perte de résolution ne modifie pas considérablement la forme des modulations et autorise les traitements ultérieurs de reconnaissance, classification et dénombrement.

5 Conclusion

La contribution principale de ce travail est la formulation d’un double test d’hypothèses binaires pour la détection temps-fréquence de signaux multi-composantes et non-stationnaires dans un environnement à faible RSB. Le premier test est basé sur l’énergie du signal, tandis que le deuxième est fondé sur la densité de détections dans une région temps-fréquence. L’ensemble des deux tests binaires, appelé FADA (Fausse Alarme Density Analysis), repose sur l’absence d’hypothèses a-priori sur le signal d’intérêt. Le premier test d’hypothèses BHT1 engendre un spectrogramme binaire contenant de vraies détections et des fausses alarmes. Pour réduire le nombre de fausses alarmes on utilise un deuxième test BHT2 : si la densité des détections est supérieure au seuil obtenu selon le critère de Neyman-Pearson, la région contient du signal, sinon elle contient seulement du bruit. En tenant compte de la redondance intrin-

sèque du spectrogramme, il est montré que la densité de fausses alarmes suit une distribution binomiale corrélée. On propose une méthode numérique permettant d’estimer cette loi de probabilité via des spectrogrammes assistants. Nous avons illustré les points forts (amélioration de la probabilité de détection) et les limitations (perte de résolution TF) de la méthode sur des données réelles.

Remerciements

Ce travail a été réalisé grâce à une bourse DGA, l’appui de la région Rhône-Alpes (bourse explora-doc) et dans l’ANR-12-ASTR-0021-03. Nous remercions le SHOM (Dr. Y. Stephan) pour les données ERATO 2009.

Références

- [1] M.C. Caldwell, D.K. Caldwell and P. L Tyack, Review of the signature-whistle hypothesis for the Atlantic bottlenose dolphin. *The bottlenose dolphin*, p.199-225, S. Leatherwood and R.S.R. Reeves, 2012.
- [2] D.R. Griffin, Audible and ultrasonic sounds of bats. *Cellular and Molecular Life Sciences*, 7(12) :448-453, 1951.
- [3] P. Flandrin, *Time-frequency/time-scale analysis*, Academic Press, 1998.
- [4] J. Huillery, F. Millioz, and N. Martin, On the description of spectrogram probabilities with a chi-squared law. *IEEE Transactions on Signal Processing*, 56(6) :2249-2258, 2008.
- [5] D. Gillespie, Detection and classification of right whale calls using an edge detector operating on a smoothed spectrogram. *Canadian Acoustics*, 32(2) :39-47, 2004.
- [6] I. R Urazghildiiev and C. W Clark, Acoustic detection of north Atlantic right whale contact calls using spectrogram-based statistics. *Journal of the Acoustical Society of America*, 122(2) :769-776, 2007.
- [7] F. Dadouchi, C. Gervaise, C. Ioana, J. Huillery, J.I. Mars, Segmentation temps-fréquence statistique : Analyse de la répartition spatiale des détections sur spectrogramme N fois seuillé. *XXIV Colloque du GRETSI*, 2013.
- [8] F. Dadouchi, C. Gervaise, C. Ioana, J. Huillery, J.I. Mars, Automated segmentation of linear time-frequency representation of marine mammals sounds. *Journal of the Acoustical Society of America*, 134(3) :2546-25, 2013.
- [9] J. Huillery, F. Millioz, and N. Martin, Gaussian Noise Time-Varying Power Spectrum Estimation with Minimal Statistics. *IEEE Transactions on Signal Processing*, 62(22) :5892-5906, 2014.
- [10] RR Bahadur, A representation of the joint distribution of responses to n dichotomous items. *studies in item analysis and prediction*, 158-168, 1961.