

Séparation de sources par factorisation en matrices non-négatives et localisation de baleines à partir de sismomètres de fond de mer

Jean LECOULANT¹, Quentin SAINT-CHRISTOPHE^{1,2}, Abdel-Ouahab BOUDRAA¹,

¹IRENav (EA3634), École Navale/Arts-Métiers ParisTech, BCRM Brest, CC 600, 29240 Brest Cedex 9, France

²Lab-STICC (UMR CNRS 6285) ENSTA Bretagne, 2 Rue François Verny, 29200 Brest, France

Résumé – Dans ce travail, la factorisation en matrices non-négatives est proposée pour séparer les chants stéréotypés de deux sous-espèces de baleine bleue enregistrés par un seul sismomètre de fond de mer (OBS). Elle décompose le spectrogramme en une matrice de composantes fréquentielles et une matrice d'évolutions temporelles. Les composantes dont les pics d'évolution temporelle correspondent à l'intervalle inter-chant des baleines servent à reconstruire des spectrogrammes dominés par chaque chant. Pour la localisation, nous utilisons l'azigramme, une représentation de l'azimut de la source dans le plan temps-fréquence. On pondère l'azigramme par les spectrogrammes reconstruits et on construit une densité d'occurrences en azimut. On construit les densités d'occurrences en latitude-longitude et on localise les baleines grâce à plusieurs OBS.

Abstract – In this work, we use non-negative matrix factorization for source separation of two blue whale-subspecies stereotyped calls recorded by ocean bottom seismometers (OBS). Non-negative matrix factorization decomposes the spectrogram into a spectral-component matrix and a time-encoding matrix. We use components with peaks in the time-encoding matrix corresponding to the whale inter-call interval to reconstruct spectrograms dominated by each whale call. Our source localization uses azigram, a time-frequency plan representation of source azimuth. Recovered spectrograms are used to weigh azigram and to build azimuth occurrence densities. We propose a method to build latitude-longitude occurrence densities and to localize whales using several OBS.

1 Introduction

Les sismomètres de fond de mer (OBS : *ocean bottom seismometer*) permettent une couverture acoustique passive de l'océan mondial, en particulier par grands fonds, sur la bande ultra-basse fréquence (UBF) (≤ 100 Hz). Pour améliorer la compréhension de ce milieu complexe, il est indispensable de classifier et de localiser une grande diversité de sources acoustiques d'origines anthropique, biologique, géophysique et cryogénique.

Dans le domaine temps-fréquence (spectrogramme), des algorithmes de traitement d'image peuvent être utilisés pour la séparation de sources. La factorisation en matrices non-négatives (NMF : *non-negative matrix factorization*) est une technique efficace pour isoler dans une image des motifs de faible étendue, au contraire de l'analyse en composantes principales, qui génère des images de base couvrant la totalité de l'image de départ [7]. La NMF ne génère pas d'amplitudes négatives dans les spectrogrammes reconstruits, ce qui serait sans signification physique, et n'est pas dépendante de l'hypothèse de signal non-gaussien et de bruit gaussien requise par l'analyse en composantes indépendantes [4].

La NMF décompose le spectrogramme entre une matrice qui contient des composantes fréquentielles et une autre qui contient leurs évolutions temporelles. Lin *et al.* [8] ont utilisé le spectre de ces fonctions d'activation pour repérer les rythmes jour-nuit et séparer les chorus biologiques des autres sources acoustiques. Lecoulant *et al.* [5] ont étendu cette idée à la séparation de sources de chants stéréotypés de grandes baleines, caractérisés par une répétition à un intervalle régulier. Nous utilisons la NMF pour séparer les chants de deux sous-espèces de baleines bleues présentes dans l'Océan indien

austral. Nous mettons ensuite à profit l'information vectorielle acquise par les OBS pour obtenir l'azimut des sources séparées [12] et proposons un algorithme pour fusionner l'information azimutale obtenue sur plusieurs capteurs en une localisation en latitude-longitude.

Cet article présente les données utilisées en Sec. 2, décrit la méthode de séparation de sources en Sec. 3 et documente la méthode de localisation en Sec. 4. Les remarques de conclusion sont données en Sec. 5.

2 Données

Nous appliquons la méthode de séparation de sources et de localisation à des données librement accessibles issues du réseau d'OBS RHUM-RUM (Reunion Hotspot and Upper Mantle – Réunions Unterer Mantel) [1]. La profondeur moyenne de ce réseau (4176 m) le place dans le domaine des grands fonds tandis que la fréquence d'acquisition des capteurs (50 ou 100 Hz) correspond aux UBF. Chaque station du réseau comporte un hydrophone omnidirectionnel et trois accéléromètres qui donnent accès à la vitesse de déplacement dans les trois directions de l'espace. L'antenne SWIR (Southwest Indian Ridge) est la zone la plus dense du réseau RHUM-RUM, avec huit OBS dans un rectangle de 100 km sur 50 km centré sur le point de latitude 27,5° S et de longitude 65,5° E. Cette densité en capteurs facilite la fusion d'information nécessaire à la méthode de localisation multi-capteur que nous proposons.

3 Séparation de sources

Nous effectuons la séparation de sources mono-capteur en acoustique passive UBF grâce à la NMF. Cet algorithme d'optimisation décompose un spectrogramme de dimension $p \times q$ ($V_{p \times q}$) en une matrice contenant n composantes spectrales ($W_{p \times n}$) et une seconde matrice contenant leurs évolutions temporelles ($H_{n \times q}$). Lecoulant *et al.* [5] ont proposé d'utiliser des détecteurs basés sur une connaissance *a-priori* des motifs temps-fréquence associés aux différentes sources d'intérêt. Ces détecteurs permettent de grouper les composantes associées à chaque source pour obtenir un spectrogramme dominé par une unique source. Cette démarche permet une séparation de sources à un faible coût numérique et sans étape d'apprentissage, qui nécessiterait des données déjà bien séparées et annotées.

La résolution itérative de l'équation $V - W \times H = 0$ est initialisée par les matrices W_0 et H_0 , puis ces matrices sont mises à jour pour chercher à minimiser une fonction de coût L , de façon que $L(V - W_k \times H_k) \geq L(V - W_{k+1} \times H_{k+1})$. Nous initialisons W et H par une décomposition en valeurs singulières où les zéros sont remplacés par la moyenne de V . La fonction de coût est la norme de Frobenius (norme L2). Les matrices W et H sont itérées par mise à jour multiplicative : W_{k+1} et H_{k+1} reçoivent un produit de W_k , H_k , V et de leurs inverses et transposées [6].

Les détecteurs de chants de baleines sont basés sur la répétition des chants à un intervalle régulier : l'intervalle interchant (ICI : Inter-Call Interval). Ces détecteurs calculent le spectre des fonctions d'activation H afin de repérer une périodicité dans les composantes temporelles séparées par NMF. Si le spectre d'une fonction d'activation présente un pic entre $0, 9/ICI$ et $1, 1/ICI$, alors la composante est classifiée comme des chants émis par la baleine dont l'ICI correspond. L'intervalle autour de la fréquence $1/ICI$ rend la méthode robuste aux variations de $\pm 0,4$ s dans l'émission des chants de baleines [3]. L'ICI est de 66 s pour les appels en Z de la baleine bleue australe (BBA) [3] et de 103 s pour les chants de la baleine bleue pygmée de Madagascar (BBPM).

L'efficacité de la séparation de source peut être évaluée en utilisant le rapport signal-à-bruit proposé par Bouffaut *et al.* [2] pour les chants de la BBA et en l'appliquant aux données brutes et données séparées par NMF. À la station RR48 et pour la journée du 31 mai 2013, étudiée par Bouffaut *et al.* [2], on constate une amélioration moyenne du rapport signal-à-bruit de 33 dB.

La Figure 1 (a) présente un spectrogramme de la pression enregistrée pendant 36 minutes par la station RR48 du réseau RHUM-RUM le 15 mai 2013. Les sons impulsifs générés par les séismes de la dorsale sud-ouest indienne apparaissent sous forme de barres verticales. Ils durent peu de temps et ont une grande largeur de bande. Au moins deux espèces de grandes baleines sont présentes. Les chants d'une BBPM sont constitués de deux unités complexes qui couvrent la bande entre 13 et 34 Hz avec un ICI de 103 s [9]. Les chants de la BBA forment un Z dans le plan temps-fréquence : une unité haute à 26 Hz est suivie d'une descente continue et d'une unité basse à 18 Hz. Pour les deux sous-espèces, le spectrogramme montre une série entière de chants, séparée de la suivante et de la précédente par un intervalle inter-série qui correspond à la respiration de la baleine [3].

Les Figures 1 (b) et (c) présentent les spectrogrammes séparés grâce à la NMF avec 50 composantes et aux détecteurs d'appels en Z et de chants de BBPM. Bien que les détecteurs n'utilisent pas d'informations sur le spectre des chants de baleine, mais uniquement l'ICI, les deux motifs sont correctement séparés. De l'énergie provenant de chants de la BBPM est cependant utilisée à tort pour reconstituer les appels en Z, ce qui explique la présence d'énergie hors de la bande 18-26 Hz sur le spectrogramme Fig. (1b). Il reste à présent à localiser ces sources.

4 Localisation

Les OBS mesurent la pression p et les trois composantes du vecteur vitesse $\mathbf{v}$. Ces deux grandeurs permettent de construire l'intensité $\mathbf{I}$, un vecteur qui pointe à l'opposé de la direction d'arrivée donnée par le vecteur unité $\mathbf{u}$, et donc, dans la direction de propagation de l'onde acoustique. On travaille en complexe, dans le plan temps-fréquence :

$$\mathbf{I} = \frac{1}{2} \text{Re}(p^* \mathbf{v}) = -\frac{|p|^2}{\rho c} \mathbf{u}, \quad (1)$$

où Re indique la partie réelle, $*$ le complexe conjugué, ρ est la densité de l'eau et c , la vitesse du son. On peut facilement retrouver la direction d'arrivée :

$$\mathbf{u} = -\frac{\mathbf{I}}{\|\mathbf{I}\|}. \quad (2)$$

Les composantes horizontales du vecteur $\mathbf{u}$ permettent de construire l'azigramme, une matrice de même dimension que le spectrogramme et qui donne l'azimut de la source en fonction du temps et de la fréquence. Une estimation par noyau pondérée permet de construire une densité d'occurrence des azimuts sur une fenêtre en temps et en fréquence [12]. Faire glisser cette fenêtre selon l'axe des temps permet d'estimer la variation temporelle de l'azimut de la source.

Nous proposons d'intégrer la séparation de sources dans cette procédure de localisation à travers la pondération de l'estimation par noyau de la densité d'occurrence des azimuts. Pour la source g , on pondère l'estimation par le spectrogramme Q issu de la séparation de sources, en échelle logarithmique et normalisé entre 0 et 1.

La Figure 1 présente les azigrammes pondérés avec une dynamique de 50 dB en utilisant l'amplitude brute (d), l'amplitude des appels en Z séparés (e) et l'amplitude des chants de BBPM séparés (f). La BBA apparaît autour d'un azimut de 270° et la BBPM autour d'un azimut de 90° . Cette figure révèle aussi les limites de la méthode de séparation de sources, puisque le spectrogramme des appels en Z séparés apparaît en partie reconstitué à partir d'énergie provenant de l'azimut 90° , tandis que le spectrogramme des chants de BBPM séparés est en partie reconstitué à partir d'énergie provenant de l'azimut 270° .

Les densités d'occurrences en azimut présentées Figure 2 sont construites à partir des azigrammes Figure 1 en utilisant des bandes de fréquence différentes et des pondérations différentes. Pour localiser la BBA, on procède à l'estimation par noyau sur les fréquences comprises entre 18 et 26 Hz (Fig. 2a et b). Pour localiser la BBPM, on procède à l'estimation par noyau sur les fréquences comprises entre 33 et 35 Hz (Fig. 2c

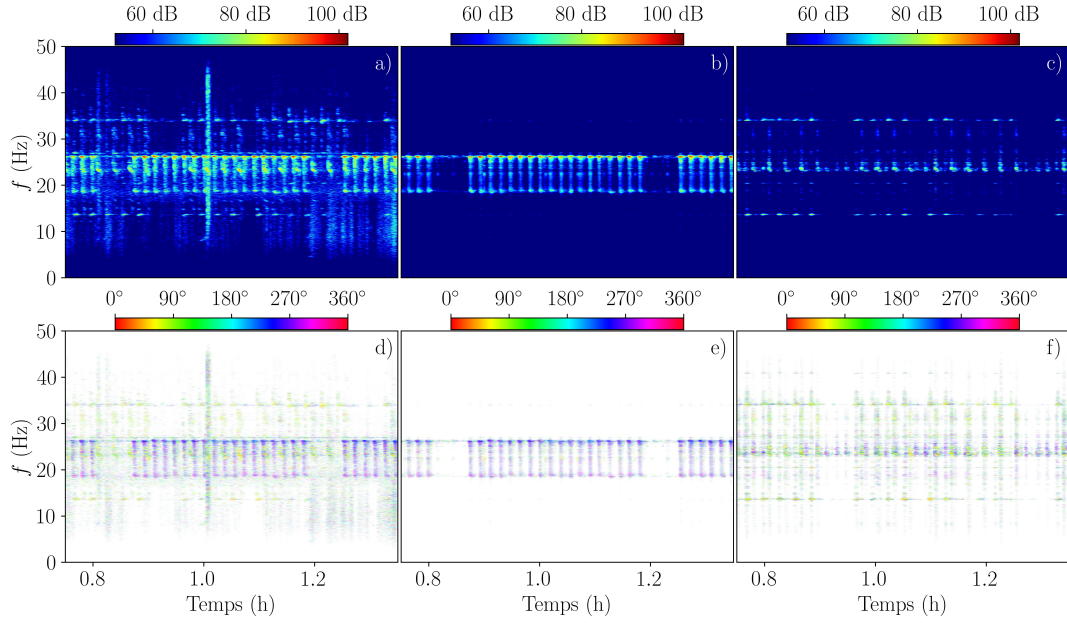


FIGURE 1 : Spectrogrammes (en dB) de la pression enregistrée par la station RR48 le 15 mai 2013 de 00h45 à 01h21, filtrée entre 5 et 46 Hz (longueur de la fenêtre : 4096 points, recouvrement : 98%) : (a) données brutes, (b) appels en Z séparés, (c) appels de la BBPM séparés. Azigrammes (en °) correspondant : (d) données brutes, (e) appels en Z séparés, (f) appels de la BBPM séparés.

et d), autour de l'unité la plus intense du signal – environ 34 Hz. Pour les deux cibles d'intérêt, la pondération par le résultat de la séparation de sources plutôt que par l'amplitude brute produit des densités d'occurrences plus piquées, qui conduiront à une moindre incertitude dans la localisation des baleines.

Pour passer d'une densité d'occurrences en azimuth à une densité d'occurrences en latitude-longitude, nous utilisons un algorithme proposé par Saint-Christophe [10]. On suppose que la source est présente dans un domaine de recherche qui est discrétisé. Pour chaque OBS, on doit calculer la densité d'occurrence que contient la case (i, j) du domaine de recherche à partir de la densité d'occurrence en azimuth P . La case (i, j) apparaît pour l'OBS localisé dans la case (i_0, j_0) avec un diamètre angulaire borné par :

$$\alpha = \left\lceil 2 \arctan \left(\frac{d}{2D} \right) \right\rceil, \quad (3)$$

où D est la distance de (i_0, j_0) à (i, j) et d la longueur de la diagonale de la case. Définir α ainsi revient à assimiler la surface de la case à celle du cercle dans lequel elle s'inscrit. La partie entière supérieure $\lceil \cdot \rceil$ permet de correspondre aux azimuths discrétisés que contient P . L'angle entre l'est et la droite qui passe par les centres des cases (i_0, j_0) et (i, j) est :

$$\gamma_0 = \arctan \left(\frac{j - j_0}{i - i_0} \right) + \begin{cases} 180 & \text{si } i < i_0 \\ 0 & \text{sinon} \end{cases}. \quad (4)$$

La densité d'occurrence dans la cellule (i, j) vaut alors :

$$\delta_k(i, j) = \sum_{\gamma=\gamma_0-\alpha/2}^{\gamma_0+\alpha/2} P(\gamma) S(\gamma, i, j), \quad (5)$$

où $S(\gamma, i, j)$ représente l'aire de la surface de la case (i, j) comprise entre les deux droites passant par le centre de la case (i_0, j_0) et faisant un angle avec l'est égal à, respectivement, $\gamma - 0, 5^\circ$ et $\gamma + 0, 5^\circ$.

La fusion de l'information provenant des différents OBS d'indice k s'effectue par un produit des densités d'occurrences en latitude-longitude :

$$\Delta(i, j) = \prod_k \delta_k(i, j) / \sum_{m, n} \prod_k \delta_k(m, n). \quad (6)$$

On peut finalement identifier le maximum de la densité d'occurrence $\Delta(i, j)$ à la position de la source acoustique, tandis que la largeur du mode autour de ce maximum renseigne sur l'incertitude en position.

La carte Figure 3 donne le résultat de cette méthode de localisation pour la BBA et la BBPM à 01h37. Pour la BBA, les faisceaux constitués des densités d'occurrences calculées pour chaque station concourent en une région limitée et permettant de localiser la baleine avec confiance. Pour la BBPM, les densités d'occurrences calculées pour chaque station dessinent un large cône d'incertitude qui ne permet pas de donner une position certaine pour la baleine. Ces deux incertitudes de localisation sont typiques d'une source située respectivement dans le réseau et hors du réseau d'OBS de l'antenne SWIR.

5 Conclusion

La NMF apparaît comme un outil efficace pour séparer deux types de chants de baleine et les isoler du bruit ambiant à partir d'une seule information : l'ICI. Pour ce faire, nous détectons les pics du spectre des fonctions d'activation correspondant à l'ICI recherché et reconstituons un spectrogramme séparé. Ce spectrogramme peut être utilisé avec l'azigramme pour évaluer une localisation en azimuth des baleines d'intérêt. En croisant les localisations en azimuth obtenues par plusieurs OBS, il est possible de localiser les baleines en latitude et longitude.

La méthode requiert encore une évaluation systématique en mesurant son apport en termes de taux de détection via filtrage adapté stochastique [2] ou par apprentissage profond

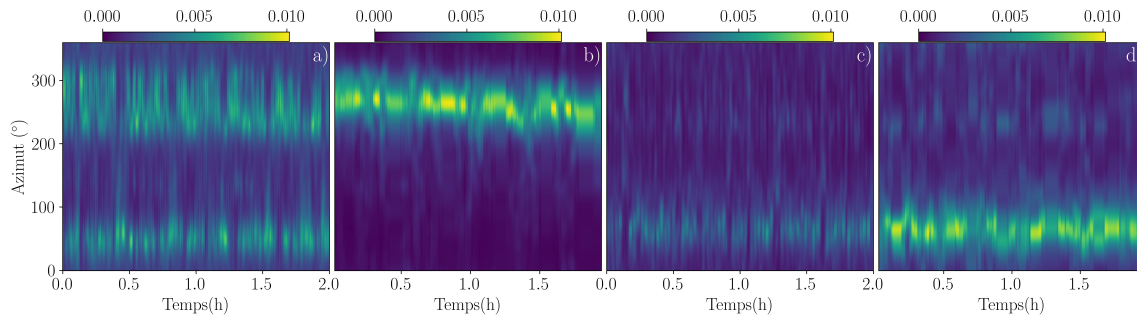


FIGURE 2 : Densités d’occurrences en azimut entre 18 et 26 Hz (a et b) pondérées par l’amplitude brute (a) et par l’amplitude des appels en Z séparés (b). Densités d’occurrences en azimut entre 33 et 35 Hz (c et d) pondérées par l’amplitude brute (c) et par l’amplitude des chants de BBPM séparés (d).

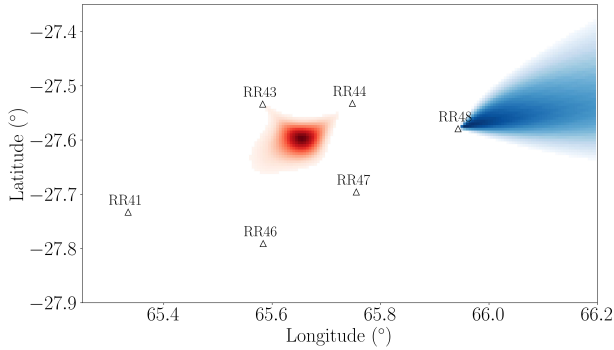


FIGURE 3 : Cartes des stations du réseau RHUM-RUM servant à la localisation (triangle) et densités d’occurrences de la BBA (en rouge) et de la BBPM (en bleu) à 01h37 le 15 mai 2013.

[11] et en relevant les temps d’arrivée aux hydrophones pour la localisation [3].

L’étape suivante consistera à tenir compte de l’azigramme lors de la procédure de NMF, afin d’éviter de constituer des classes à partir d’énergie provenant d’azimuts différents. La technique de séparation et de localisation présentée ici pourra également être appliquée à d’autres sources acoustiques, notamment des sources situées en surface, comme des navires.

Références

- [1] Guilhem BARRUOL et Karin SIGLOCH : Investigating La Réunion Hot Spot From Crust to Core. *Eos, Transactions American Geophysical Union*, 94(23):205–207, 2013.
- [2] Léa BOUFFAUT, Richard DRÉO, Valérie LABAT, Abdel-O. BOUDRAA et Guilhem BARRUOL : Passive stochastic matched filter for Antarctic blue whale call detection. *The Journal of the Acoustical Society of America*, 144(2):955–965, 2018.
- [3] Richard DRÉO, Léa BOUFFAUT, Emmanuelle LEROY, Guilhem BARRUOL et Flore SAMARAN : Baleen whale distribution and seasonal occurrence revealed by an ocean bottom seismometer network in the Western Indian Ocean. *Deep Sea Research Part II : Topical Studies in Oceanography*, 161:132–144, 2019.
- [4] A. HYVÄRINEN et E. OJA : Independent component analysis : algorithms and applications. *Neural Networks*, 13(4-5):411–430, 2000.
- [5] Jean LECOULANT, Abdel-Ouahab BOUDRAA et Samuel PINSON : Non-negative matrix factorization based single-channel source-separation of passive underwater acoustic signals in deep sea. *JASA Express Letters*, 5(2):026004, 2025.
- [6] Daniel LEE et H. Sebastian SEUNG : Algorithms for Non-negative Matrix Factorization. In T. LEEN, T. DIETTERICH et V. TRESP, éditeurs : *Advances in Neural Information Processing Systems*, volume 13. MIT Press, 2000.
- [7] Daniel D. LEE et H. Sebastian SEUNG : Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791, 1999.
- [8] Tzu-Hao LIN, Yu TSAO et Tomonari AKAMATSU : Comparison of passive acoustic soniferous fish monitoring with supervised and unsupervised approaches. *The Journal of the Acoustical Society of America*, 143(4):EL278–EL284, 2018.
- [9] Mark A. McDONALD, Sarah L. MESNICK et John A. HILDEBRAND : Biogeographic characterisation of blue whale song worldwide : using song to identify populations. *J. Cetacean Res. Manage.*, 8(1):55–65, 2023.
- [10] Quentin SAINT-CHRISTOPHE : *Fusion d’information multisource en monde ouvert – Vulnérabilité du processus et gestion des conflits*. Thèse de doctorat, ENSTA - IP Paris, 2025.
- [11] Elena SCHALL, İdil İlğaz KAYA, Elisabeth DEBUSCHERE, Paul DEVOS et Clea PARCERISAS : Deep learning in marine bioacoustics : a benchmark for baleen whale detection. *Remote Sensing in Ecology and Conservation*, 10(5):642–654, octobre 2024.
- [12] A TRABATTONI, G BARRUOL, R DREO, A O BOUDRAA et F R FONTAINE : Orienting and locating ocean-bottom seismometers from ship noise analysis. *Geophysical Journal International*, page ggz519, 2019.