

Titre de la thèse	Apprentissage sur représentations de Wigner-Ville de signaux bioacoustiques fortement modulés en phase et fréquence : étude de la résolution de la vitesse et de l'effet Doppler des biosonars.
-------------------	---

Ecole Doctorale	ED548
Laboratoire	Laboratoire d'Informatique et Systèmes (LIS)
Discipline	Informatique et Sciences de la Mer
Directeur(s) de Thèse & Encadrant(s)	directeur : Hervé Glotin, co-directrice : HDR Dr Adeline Paiement co-encadrement : Dr Sébastien Paris

Description du sujet de recherche

Contexte, originalité et pertinence par rapport à l'état de l'art : Le lien montré par Ville dans les années 1950 entre la fonction d'ambiguïté et la représentation temps-fréquence donne l'importance de cette grandeur dans l'étude des signaux fortement modulés en phase et en fréquence $BT \gg 1$ (étendu spectrale B, durée T) (Max et al 1986). De tels signaux sont très fréquents dans les systèmes bioacoustiques. Par exemple l'émission acoustique du cachalot étudiée par Jiang et al 2021, a permis de calculer à l'aide de la fonction d'ambiguïté des paramètres du biosonar du cachalot, tels que la résolution de sa portée, sa résolution de la vitesse et sa tolérance à l'effet Doppler. En somme, le cachalot modifie constamment les formes d'onde des signaux transmis pendant les phases d'approche et de capture afin de s'adapter aux différentes conditions écologiques.

Cette thèse remettra en contexte l'étude des biosonar par apprentissage optimisée de la représentation temps-fréquence à partir de la distribution de Wigner-Ville (Balestrierio, Glotin et al 2022). Cette représentation permettra une meilleure discrimination des formes d'ondes.

Ce développement de méthode sera appliqué sur des données réelles, enregistrements sonores bruts pour en déduire des comportements, des signatures acoustiques et des motifs de fréquence bioacoustiques dans différents écosystèmes.

Cette approche permettrait d'ajouter un contexte du paysage sonore des enregistrements de terrain et ainsi d'interpréter le signal en contexte et suivant l'information recherchée par l'animal. Cela permettrait aussi de réaliser une possible amélioration des algorithmes de localisation et d'interprétation du comportement de la mégafaune.

Objectifs et Méthodes : Le cœur de l'approche est dans la continuité de nos travaux en apprentissage de représentation temps-fréquence de signaux adaptés voir intelligents, comme la parole, les chants d'oiseaux, ou les biosonars. Ces travaux très originaux et qui revisitent la transformée de Wigner-Ville des années 1940, à la mode du Deep learning de 2020, sont publiés dans Balestrierio Glotin Baraniuk 2022 et se décomposent suivant cet algorithme:

1. Sélection par biotopes, espèces et milieux d'un jeu de données acoustiques large bande à bon rapport signal à bruit, avec ou sans effet Doppler.
2. Calculer leur décomposition Wigner Ville (Wigner, 1932) suivant nos algorithmes disponibles. Contrairement aux représentations TFR communément utilisées, la transformée de VW est bilinéaire. Cette transformée combine des filtres sinusoïdaux complexes comme la transformée de Fourier et l'auto-corrélation du signal et est défini par :

$$WV_x(t, f) = \int_{-\infty}^{\infty} x\left(t - \frac{\tau}{2}\right) x^*\left(t + \frac{\tau}{2}\right) e^{-if\tau} d\tau.$$

Cette étape sera faite en parallèle et large scale, notamment on pourra faire la WVD par fenêtres sur HPC du LIS Dyni.

3. La dernière étape consiste à apprendre les filtres gaussiens adaptés au signal original, par réseau convolutionnel CNN dont la première couche de chaque CNN réalise une convolution 2D par des densités de gaussiennes avec la transformée K par noyaux gaussiens 2D de la transformée de WV:

$$K_x(t, f) = \int_{\mathbb{R} \times [0, 2\pi)} WV_x(\tau, \omega) \Phi_f(t - \tau, \omega) d\tau d\omega,$$

les noyaux gaussiens étant :

$$\Phi_f(\tau, \omega) = \mathcal{N}\left(\begin{pmatrix} \tau \\ \omega \end{pmatrix}; \begin{pmatrix} 0 \\ \mu_f(f) \end{pmatrix}, \begin{pmatrix} \sigma_t(f)^2 & \rho(f) \\ \rho(f) & \sigma_f(f)^2 \end{pmatrix}\right) \mathbb{1}_{\{\omega \in [0, 2\pi)\}},$$

et les paramètres appris par cette première couche apprennent la moyenne et la variance de chacun des K filtres gaussiens 2D du modèle.

Un exemple de représentation de WV appris par IA par cette méthode est Fig. 1.

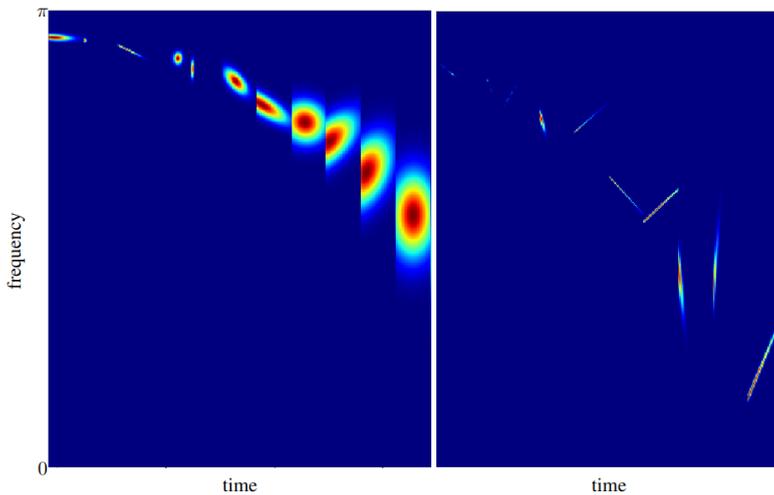


Fig 1: Représentation WV appris : A gauche sur parole (humaine), à droite sur oiseaux (d'après Balestriero, Glotin et al 2022)

Le corpus sera choisi et bien représentatif de la variété bioacoustique de notre monde, tant en sous-marin qu'en terrestre, notamment par le fait que DYNi collecte depuis 15 ans 600 To de signaux variés de part le monde.

Retombées attendues : Il sera notamment question de l'usage des antennes pour la séparation, le suivi et trace de pistes de grands sondeurs et classer par espèces et individus les clics ou les vocalises de cétacés. Un des impacts sera la jonction de la mesure en multi capteurs (antenne stéréo ou plus) en milieux sous-marins et terrestres pour séparer les sources, classer les espèces.

L'application de ce modèle à notre corpus de cachalots de l'île Maurice que nous constituons depuis 10 ans en coll. avec F. Sarano Longitude 181 sera très intéressante puisque contenant le comportement et déplacement des individus cachalots usant de leur biosonar.

En terrestre il sera question de préciser la classification d'oiseaux par TF optimisée.

Une surveillance plus précise des populations par le suivi de leurs adaptations acoustiques suite au changement climatique et la pression anthropophonique, apportera une meilleure compréhension des comportements. C'est une clef pour le suivi

de la biodiversité telle que requise par nos collaborateurs PELAGOS, PNPC, Office Français de la Biodiversité ou le Ministère de la transition écologique.

Mots clés : Wigner-Ville, Effet Doppler, TDOA, Triangulation, Détecteur, Classifieur, YOLO, Ondelettes, Répertoire vocale

Références :

Balestriero, R., Glotin, H. & Baranuik, R. (2022). Interpretable and Learnable Super-Resolution Time-Frequency Representation. Proceedings of the 2nd Mathematical and Scientific Machine Learning Conference, in Proceedings of Machine Learning Research 145:118-152 Available from <https://proceedings.mlr.press/v145/balestriero22a.html>.

Becker S, Marcel Ackermann, Sebastian Lapuschkin, Klaus-Robert Muller, and Wojciech Samek. Interpreting and explaining deep neural networks for classification of audio signals. arXiv preprint arXiv:1807.03418, 2018.

Best P. (2022). Automated Detection and Classification of Cetacean Acoustic Signals. Machine Learning. Phd Thesis, Université de toulon (dir Glotin). <https://hal.archives-ouvertes.fr/tel-03826638/document>

Deloustal N., Chavin S., Glotin H. (2023), Rapport de recherche V1 Suivi et identification temps-fréquence bioacoustique par transfert deep learning sur YOLO : gestion des chœurs. http://sabiod.lis-lab.fr/pub/YOLO_DYNI.pdf

Deloustal N., Glotin H., (2022), Rapport de recherche, Veille bioacoustique de l'avifaune en Guadeloupe. http://sabiod.lis-lab.fr/pub/LIS_OFB_GUADELOUPE_RAPPORT-1.pdf

Ferrari M. (2020) Study of a Biosonar Based on the Modeling of a Complete Chain of Emission-Propagation-Reception with Validation on Sperm Whales, Phd Thesis, Université Picardie Jules Verne, (dir Glotin & Asch) <https://hal.archives-ouvertes.fr/tel-03078625/document>

Glotin, Mischenko, Giraudet, 2015, brevet international, JOINT CONSTRAINTS IMPOSED ON MULTIBAND TIME TRANSITIVE AND DOPPLER-EFFECT DIFFERENCES, FOR SEPARATING, CHARACTERIZING, AND LOCATING SOUND SOURCES VIA PASSIVE ACOUSTICS (54) Titre : CONTRAINTES CONJOINTES DE TRANSITIVITE DE DIFFERENCES TEMPORELLES ET EFFET DOPPLER MULTIBANDES POUR LA SEPARATION, CARACTERISATION, ET LOCALISATION DE SOURCES SONORES PARACOUSTIQUE PASSIVE, <https://patentimages.storage.googleapis.com/50/2d/de/a0db00f3ca53fb/WO2015177172A1.pdf>

Jiang, J., Li, S., Wang, X., Li, C., Sun, Z., Liu, W., ... & Fu, X. (2021). Study of the relationship between sound signals and behaviors of a sperm whale during the hunting process. *Applied Acoustics*, 174, 107745.

Lostanlen V, Justin Salamon, Andrew Farnsworth, Steve Kelling, and Juan Pablo Bello. Birdvox-full-night: A dataset and benchmark for avian flight call detection. In 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 266–270. IEEE, 2018

Poupard M. (2020) Contributions en Méthodes Bioacoustiques Multiéchelles: Spécifiques, populationnelles, individuelles et comportementale, Phd Thesis, Université de Toulon (dir Glotin Soriano Lengagne) http://sabiod.univ-tln.fr/pub/poupard/cv/m_poupard_phd_08012021.pdf

Max et al., Éléments en traitement du signal, 4ième édition 1986

Encadrement et conditions matérielles pour le doctorant

Serveur HPC LIS DYNI - 500 To de données Bioacoustiques de l'équipe en milieux marin et terrestre, les données les plus propres seront considérées. Accès à la ferme GPU du LIS pour les apprentissages.

Compétences attendues et personnes à contacter

Bonnes connaissances en bioacoustique marine et comportement de la faune marine, Master Sciences de la Mer ou Master Informatique

Personne(s) à contacter :

glotin@univ-tln.fr

adeline.paiement@univ-tln.fr