



Proposition de stage de Césure Etudiant Ingénieur informatique Science de Données et Traitement du Signal

Sujet

Faisabilité de la reconnaissance automatique de signaux produits par une communauté de grillons au sein d'enregistrements sonores d'ambiance forestière de Nouvelle-Calédonie

Contexte

Dans le contexte de la dégradation rapide de la biodiversité sous pressions anthropiques, il est nécessaire de développer des proxys intégrateurs de la qualité des habitats naturels. Au cours des 10 dernières années, le développement de méthodes d'évaluation de la biodiversité par l'acoustique passive a connu un grand progrès avec l'utilisation d'enregistreurs automatiques de plus en plus performants pour enregistrer les ambiances sonores des milieux naturels (Aide et al 2013). Ces méthodes permettent d'envisager une évaluation/caractérisation de la biodiversité terrestre et la mise en œuvre de suivis environnementaux (qualification de l'intégrité des habitats, détection de perturbations, détection d'espèces emblématiques, ...) pour étudier des dynamiques d'écosystèmes et de communautés animales. On parle aujourd'hui d'éco-acoustique pour ce champ de recherche (Sueur et al 2015). Les méthodes (méta-acoustique) utilisées permettent, soit une approche globale des productions sonores, soit des approches plus fines et détaillées avec identification de productions sonores spécifiques de groupes biologiques d'intérêt. Ce dernier point nécessite des temps d'écoute prolongés pour des experts capables de reconnaître les productions sonores produites par les différents groupes taxonomiques d'intérêt. Ce temps et cette expertise deviennent alors des facteurs limitant pour une utilisation en routine de ces méthodes.

En Nouvelle-Calédonie, les travaux pionniers d'éco-acoustique menés ces dernières années par l'IRD IMBE & le MNHN ISYEB ont permis de confirmer le potentiel de l'acoustique environnemental passive pour l'évaluation et le suivi de la biodiversité à long terme. En outre, ces travaux ont permis de mettre en évidence la place prépondérante des grillons dans les ambiances sonores des milieux naturels de Nouvelle-Calédonie (Gasc et al 2013, 2018), notamment par leur contribution exceptionnelle à la biophonie (ensemble des signaux produits par les animaux). Il a été également montré que les communautés de grillons ont un fort potentiel pour être utilisées comme un indicateur fiable de la qualité des habitats de Nouvelle-Calédonie. Ces précédents travaux ont généré un jeu unique de données acoustiques collectées en octobre-Novembre 2013, avec 3 modalités d'habitats de complexité sonore croissante et représentatives des habitats du sud de la Nouvelle-Calédonie dans un contexte de succession écologique (maquis, paraforestier et forêt). Au total 12 sites ont pu être enregistrés en continu (4 pour chacune des modalités), selon un pas de temps d'enregistrement de 2 min toutes les 28 minutes pendant au moins 2 semaines. Ce jeu de données a fait l'objet en partie d'une écoute experte sur un pas de temps de 6 jours (Thèse de Jeremy Anso). Parallèlement une base de données a pu également être constituée avec des chants représentatifs des 19 espèces de grillons identifiés dans les enregistrements.

Objectifs

Pour aller plus loin dans la faisabilité d'un outil acoustique, nous souhaitons explorer les possibilités d'automatisation de la reconnaissance des espèces de grillons. Des méthodes ont été développées en bioacoustique, par exemple, pour distinguer les sons de bateaux de ceux des baleines ou pour l'identification d'espèces d'oiseaux (Fagerlund et al. 2007, Zaugg et al. 2010, Malfanteet al. 2017). Ces méthodes sont le plus souvent basées sur le calcul et l'extraction des descripteurs discriminants. L'explosion de l'acquisition de ces séries temporelles (Time-Series) et la naissance de science de données, ces dernières

années, ont permis le développement de nouvelles méthodes pour décrire et caractériser ces séries dans le but de les identifier par des méthodes de fouille de données ou d'apprentissage automatique (Zheng et al. 2014, Hu et al. 2013, Paparrizos et al. 2015). L'objectif du stage est d'exploiter ces nouvelles méthodes pour confirmer ou infirmer des hypothèses des experts et découvrir d'autres connaissances cachées.

En particulier, il s'agira d'évaluer les résultats de ces techniques dans différentes conditions de complexité acoustique : uniquement le taxon cible vs un environnement acoustiques composés des taxons cibles et d'autres taxons chanteurs (oiseaux, sauterelles, criquets, cigales). Il s'agira également de se focaliser sur quelques espèces de grillons d'intérêt particulier (un chant bien distinctif et un potentiel bio-indicateur des modalités environnementales). Aussi, le stagiaire aura pour mission les objectifs suivants :

- 1- Automatiser la reconnaissance de l'occurrence de chants de grillons dans les séquences enregistrées à partir de méthodes classiques de fouille de données et le développement d'algorithmes *ad hoc* (classements de signaux)
- 2- Détecter des patrons simples correspondant à des espèces représentatives des différentes modalités d'habitat dans un contexte de plus ou moins grande complexité sonore (du maquis vers la forêt dense) ;
- 3- Mettre en œuvre éventuellement un processus automatique pour filtrer les « bruits » issus de la géophonie (pluie, vent ...)

Références

- Aide T.M., Corrada-Bravo C., Campos-Cerqueira M., Milan C., Vega G., Alvarez R. (2013) Real-time bioacoustics monitoring and automated species identification. *PeerJ* 1:e103.
- Anso J. (2016) Maintien à long terme des communautés d'insectes forestiers dans un contexte de changement global : Interactions entre communautés forestières de grillons et espèces envahissantes de fourmis en Nouvelle-Calédonie. Thèse de doctorat de l'université de la Nouvelle-Calédonie, spécialité : Biologie des populations et écologie, Ecole Doctorale du Pacifique, Nouméa, 431 pp.
- Fagerlund S. (2007) Bird Species Recognition Using Support Vector Machines. *EURASIP Journal on Advances in Signal Processing*, 2007(1).
- Gasc A., Sueur J., Pavoine S., Pellens R., Grandcolas P. (2013) Biodiversity sampling using a global acoustic approach: contrasting sites with microendemics in New Caledonia? *PLoS ONE*, 8: e65311.
- Gasc A., Anso J., Sueur J., Jourdan H., Desutter-Grandcolas L. (2018) Cricket calling communities as an indicator of the invasive ant *Wasmannia auropunctata* in an insular biodiversity hotspot. *Biological Invasions* 20: 1099-1111.
- Hu B., Chen Y., Keogh E. (2013) Time series classification under more realistic assumptions. In *Proceedings of the 2013 SIAM International Conference on Data Mining* (pp. 578-586). Society for Industrial and Applied Mathematics.
- Malfante M., Mura M.D., Boullay B., Métaxian J-P., Mars J (2017) Apprentissage statistique: classification automatique de signaux volcano-sismiques. XXVIème colloque GRETSI, Juan les Pins, France. 2017, Proceeding Actes de la conférence GretsI 2017
- Paparrizos J., Gravano L. (2015) k-shape: Efficient and accurate clustering of time series. In *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data* (pp. 1855-1870). ACM.
- Sueur J., Farina A. (2015) Ecoacoustics: the Ecological Investigation and Interpretation of Environmental Sound. *Biosemiotics* 8(3) : 493–502
- Zaugg S., Van Der Schaar M., Houégnigan L., Gervaise C., André M. (2010) Real-time acoustic classification of sperm whale clicks and shipping impulses from deep-sea observatories. *Applied Acoustics*, 71(11) :1011–1019, 2010
- Zheng Y., Liu Q., Chen E., Ge Y., Zhao J.L. (2014) Time series classification using multi-channels deep convolutional neural networks. In *International Conference on Web-Age Information Management* (pp. 298-310). Springer, Cham.

Encadrement : Nazha Selmaoui- Folcher (MCF HDR UNC, ISEA), Amandine Gasc (CR IRD, IMBE) , Hervé Jourdan (IR IRD, IMBE)

Lieux : IMBE, UMR CNRS 7263 / IRD 237 / AMU /UAPV, Centre IRD de Nouméa, BPA5, 98848
Nouméa cedex, Nouvelle-Calédonie ; **ISEA (équipe Data Science)**, Université de la Nouvelle Calédonie,
98800 Nouméa

Durée : 5 à 6 mois (idéalement 16 Août 2018 au 15 fév. 2019)

Conditions : Indemnités forfaitaires de stage (570 euros et logement pris en charge sur site campus IRD)
Billet d'avion Paris/Nouméa/Paris pris en charge également

Compétences recherchées : Dynamisme, autonomie, expérience en fouille de séries temporelles (Time Series), science de données et/ou en traitement du signal, intérêt pour l'évaluation de l'environnement et l'écologie, connaissances en acoustique souhaitée mais non nécessaire.

Pour des raisons administratives (délai d'obtention de titre de séjour spécifique à la Nouvelle-Calédonie), seuls les candidats de nationalité française pourront être sélectionnés pour ce stage.

Cv et lettre de motivation à envoyer **avant le 01/08/2018** à nazha.selmaoui@univ-nc.nc,
herve.jourdan@ird.fr et amandine.gasc@ird.fr