

10ème Congrès Français d'Acoustique

Lyon, 12-16 Avril 2010

La reconnaissance automatique de sources sonores dans l'environnement

Boris Defreville¹, Stéphane Bloquet², Guillaume Filippi², Christine Aujard²

¹ ORELIA SAS, 1 rue du Port du Valvins F-77210 AVON, boris.defreville@orelia.fr

² 01dB-Metravib 200 Chemin des Ormeaux F-69578 LIMONEST, stephane.bloquet@areva.com guillaume.filippi@areva.com
christine.aujard@areva.com

L'environnement sonore urbain est composé d'un ensemble complexe et varié de sources sonores, en évolution permanente. Les transports routiers et ferroviaires, les aéronefs, les industries et les commerces, les chantiers et les activités locales sont à l'origine des ambiances citadines. De plus, les pratiques innovantes et efficaces de surveillance dans l'environnement conduisent à des mesures longues durées où l'opérateur intervient ponctuellement à distance. L'identification des diverses composantes sonores reste néanmoins fondamentale. Comment distinguer et reconnaître avec précision les sources de bruit dans une ambiance acoustique complexe? Le logiciel OSSR permet de reconnaître de manière fiable les événements sonores et de discriminer automatiquement la nature des sources de bruit, parfois même dans des conditions difficiles pour l'oreille humaine... Un exemple opérationnel concernant la détection de bruit d'inverseurs de poussée (ou bruits de reverse) à proximité d'un aéroport est évoqué ci-après, mettant en avant l'intérêt majeur de détecter automatiquement les sources acoustiques dans un contexte de gestion environnemental durable.

1 Introduction

Les ambiances sonores urbaines évoluent en permanence au fil du temps. Rythmées par la prédominance des passages de véhicules routiers et/ou ferroviaires, des mouvements d'aéronefs, des activités industrielles et/ou humaines, des phénomènes mécaniques et/ou naturels... les ambiances s'avèrent successivement bruyantes, gênantes, mais également parfois agréables [1].

Décideurs, acteurs de la ville et professionnels de l'acoustique s'interrogent fréquemment sur les composantes de notre environnement sonore :

- Quelle est l'impact d'un projet de déviation de l'accès à un centre ville ?
- L'influence des trains est-elle forte ou faible, dans l'ambiance sonore locale ?
- Quelle est la contribution énergétique liée aux avions dans le bruit global ?
- Quels sont les niveaux de bruit générés par l'industrie? par l'infrastructure voisine ?
- Est-ce que les émergences sonores mesurées lors de la surveillance correspondent à des passages de véhicules routiers, de trains, d'avions, ... ?

La directive européenne 2002/49/CE relative à l'évaluation et la gestion du bruit dans l'environnement vise à cartographier les situations acoustiques et à programmer des plans d'actions limitant les effets du bruit. De plus, elle impose une information régulière des populations locales sur ces sujets. Gestionnaires de collectivités territoriales, d'infrastructures terrestres, d'activités aéroportuaires et industrielles sont conjointement impliqués dans ces démarches menées à l'échelle européenne. Pour optimiser leur planification et

leur communication auprès des riverains, tous souhaitent aujourd'hui accéder à la connaissance précise des niveaux sonores et de l'origine des bruits propagés dans l'environnement.

Ainsi, de nombreux sites se dotent de moyens de surveillance acoustique de type Oper@ afin de suivre l'évolution des ambiances sonores en continu et pendant de longues durées. Afin de mieux détecter et cibler les événements sonores présents lors des campagnes de mesures, les utilisateurs peuvent avoir recours à plusieurs solutions. La première est de corréler les données de mesures avec d'autres données extérieures. Par exemple, un opérateur peut être présent lors de la mesure afin de coder un certain nombre d'événements. Il peut également effectuer des écoutes à posteriori des enregistrements sonores. Les aéroports qui disposent de données radars peuvent les utiliser pour trier un certain nombre d'événements qui viennent polluer leurs mesures. Certains encore utilisent des données de comptage ou même des photographies leur permettant de mieux identifier les types d'événements sonores. La seconde méthode consiste à trier les événements à partir de la mesure du niveau sonore moyen L_{Aeq} en fixant un seuil et une durée minimale à un événement.

Enfin, grâce au logiciel de reconnaissance de sources sonores OSSR, les utilisateurs de stations de mesures longues durées ont désormais à leur disposition un nouvel outil leur permettant de mieux exploiter leurs mesures (voir [2, 3] pour des exemples d'utilisation).



Figure 1 : système de mesure Oper@, pour la surveillance acoustique longue durée. L'utilisation couplée du logiciel OSSR permet de reconnaître les sources à l'origine des bruits mesurés.

Par la suite, nous décrivons de manière simplifiée les concepts technologiques du logiciel OSSR, nous exposons un exemple d'utilisation pour la détection des bruits d'inverseurs de poussée et nous abordons les possibilités d'initier la détection de diverses sources sonores, en particulier d'origine industrielles.

2 Principes et méthodes de reconnaissance automatique à l'aide du logiciel OSSR

La fonction première de l'oreille est de détecter et reconnaître les sources de bruits qui nous entourent [4]. De manière schématisée, le cerveau analyse les signaux en provenance de l'oreille, les interprète et les associe instantanément aux images correspondantes des sources sonores grâce à l'utilisation d'un certain nombre d'indices acoustiques pertinents. Ces indices acoustiques ont été obtenus au préalable lors d'une phase d'apprentissage ou d'entraînement. La réalisation d'une reconnaissance de sources par un ordinateur « en aveugle » nécessite, de la même manière que l'oreille humaine, de définir un ensemble de descripteurs audio pertinents. Ces descripteurs sont équivalents aux indices acoustiques et une phase d'apprentissage est également nécessaire à leur extraction.

Aussi, la reconnaissance de sources sonores est basée sur la reconnaissance de formes. C'est un ensemble de techniques et de méthodes visant à identifier des motifs à partir de données brutes afin de prendre une décision dépendant de la catégorie attribuée à ce motif. Il s'agit d'une branche de l'intelligence artificielle qui fait largement appel aux techniques d'apprentissage automatique et aux statistiques.

La première étape dans un système de reconnaissance de forme est d'extraire les indices acoustiques caractéristiques d'une classe de bruit.

Sur la figure 2 qui représente le sonagramme d'une séquence constituée d'événements sonores (de gauche à droite : un train, une moto, un avion, une voiture, des voix, une voiture et un train), nous pouvons repérer visuellement certaines de ces caractéristiques. Principalement, les niveaux d'énergie ne sont pas distribués de la même manière selon les événements.

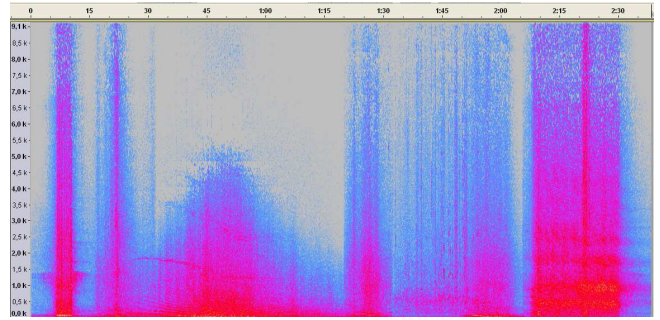


Figure 2 : sonagramme d'une séquence sonore de 2min et 45s représentation des véhicules, des trains, un avion et des voix humaines.

Si nous nous focalisons sur le passage de l'avion et en effectuons une représentation spectrale instantanée (instant 35 de la figure 2), nous pouvons faire apparaître trois caractéristiques des bruits d'avions à réaction sur la figure 3 : des harmoniques en dessous de 1000 Hz, une raie spectrale aux alentours de 2000 Hz ainsi qu'une chute brutale de l'énergie à partir de 4000 Hz.

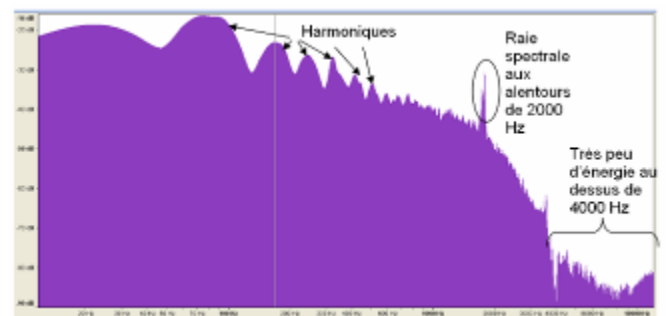


Figure 3 : représentation du spectre à la seconde 35 de l'avion de la figure 2. Nous observons 3 caractéristiques propres aux bruits des avions à réaction.

Ces trois caractéristiques varient selon la mesure (distance au sonomètre), la configuration de vol (décollage, atterrissage, approche) ainsi que le type et le niveau du bruit de fond. Toutes ces variations et leurs limites doivent être prises en compte si l'on veut pouvoir reconnaître un grand nombre d'avions dans un grand nombre de situations. Cela est rendu possible grâce à une phase d'apprentissage.

La figure 4 montre la répartition d'échantillons sonores de la classe avions (points bleus) ainsi que des échantillons sonores d'autres classes (points rouges) dans l'espace représenté par deux caractéristiques propres aux avions. La phase d'apprentissage consiste à établir des frontières entre ces deux classes. Cette frontière est alors mémorisée dans un modèle et tout nouvel échantillon sonore à classer est alors projeté dans cet espace. La zone dans laquelle il se trouve permet de conclure s'il s'agit ou non d'un avion. Notons toutefois que dans la réalité, l'espace formé est à n dimensions, n correspondant au nombre de caractéristiques ou indices acoustiques. La ligne verte de la figure 4 correspond à une frontière trop basique, la courbe en violet caractérise un effet de sur-apprentissage. La courbe en noire est la courbe adéquate.

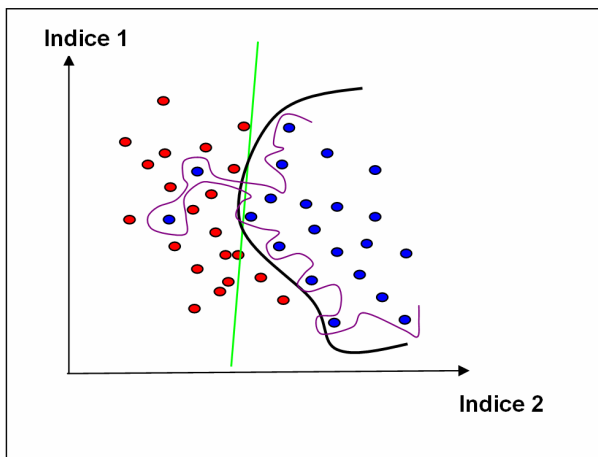


Figure 4 : illustration de différentes stratégies pour la phase d'apprentissage.

Enfin, il faut signaler que la phase d'apprentissage est effectuée à partir d'un nombre d'échantillons sonores relativement restreint (une centaine d'avions) mais d'une riche diversité qui permet de couvrir un large spectre de types d'avions et de conditions de mesures.

Les systèmes de reconnaissance traditionnels reposent généralement sur l'utilisation d'un même corpus de descripteurs audio, quelle que soit la source à reconnaître. Si de telles méthodes présentent de bons résultats dans le domaine de la reconnaissance vocale où les dictionnaires de mots ainsi que leur agencement sont bien connus, elles ne sont pas appropriées pour la reconnaissance des bruits de l'environnement qui sont particulièrement complexes, stationnaires et souvent mélangés.

La méthode de reconnaissance présentée ici permet de trouver des indices discriminants et orthogonaux pour chacune des classes de bruit à détecter. Cette approche a prouvé tant sur le plan théorique que pratique qu'elle apportait des résultats plus pertinents que l'approche traditionnelle [5, 6, 7].

3 Utilisation et performances du logiciel OSSR

Complémentaire des systèmes de mesure, le logiciel OSSR (Orelia Sound Source Recognition) réalise la reconnaissance automatique des sources de bruit contenues dans un signal audio. Proposant trois modules de détection (bruit routier, bruit ferroviaire, bruit aérien), le logiciel OSSR reconnaît les sources de bruit surveillées. Les apparitions de sources spécifiques peuvent être détectées sans contrainte de niveau acoustique et lors de multi-exposition sonore en détectant la source de bruit dominante. Le traitement de nombreux fichiers de données peut être réalisé automatiquement. Il consiste en une analyse successive des enregistrements. L'ensemble des fonctionnalités est accessible via une interface utilisateur représentée sur la figure 5.

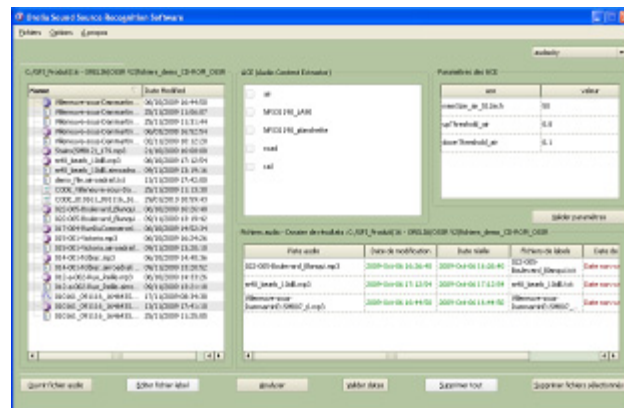


Figure 5 : interface d'utilisation du logiciel OSSR.

Sélectif, OSSR extrait les apparitions des sources sonores liées au transport (avions, trains, véhicules légers, poids lourds, 2 roues,...), en excluant les contributions des autres sources de bruit. Il permet le codage automatique des événements détectant la présence de sources spécifiques mesurées.

Le logiciel OSSR traite les fichiers de formats .wav et .mp3 : le débit minimum des fichiers audio que le système traite est égal à 40kbps, la fréquence d'échantillonnage minimale est de 24kHz. La résolution est de 16 à 32 bits. L'export est réalisable vers dBTrait, logiciel de traitement de données acoustiques. L'éditeur de fichiers sonores Audacity™ permet également de représenter les informations codées. Il est enfin possible pour l'utilisateur de créer un format d'export grâce à l'utilisation du xslt [8], langage de transformation XML. Cela permet d'obtenir facilement une compatibilité entre les données de codages obtenues à l'aide d'OSSR et les logiciels de traitement et de représentation des mesures utilisés par les acousticiens. Le logiciel OSSR est compatible avec les versions Pro des systèmes d'exploitation Windows XP™, Windows Vista™, Windows 7™ et Windows Server 2003™.

Les performances du logiciel OSSR sont sensiblement comparables à celles de l'oreille humaine. En appliquant la technique de la validation croisée (10-folds cross-validation) [9], la performance de chacun des trois modules de détection disponibles est supérieure à 95%. Les tests sont effectués à partir d'échantillons très proches du point de vue du timbre de la classe à détecter. Le traitement des informations est beaucoup plus rapide que la durée de la mesure. Par exemple, la vitesse d'analyse est supérieure à 20 fois la durée de l'enregistrement sur un PC équipé d'un processeur Intel Core Duo E6750@2.66GHz.

Le système est également évolutif et il est apte à détecter tous types de sources sonores, suite à une phase d'apprentissage. L'apprentissage de la détection d'une nouvelle source se fait à partir d'un certain nombre de fichiers de signaux enregistrés par les stations de mesure. Un module de détection générique est alors développé. Il permet la reconnaissance de la source pour le traitement des fichiers. Ce module est appelé Audio Content Extractor (ACE). Un exemple de création d'un nouveau module de détection pour une étude spécifique est évoqué ci-après.

4 Cas opérationnel de détection des inverseurs de poussée (ou bruits de reverse)

L'utilisation des inverseurs de poussée est pratique courante lors de l'atterrissage des avions sur les aéroports car ils permettent de ne pas trop solliciter les freins des roues et de raccourcir la distance d'atterrissage. Étant très bruyants, les inverseurs de poussée sont interdits en période nocturne dans certains aéroports [10, 11], car lorsque le moteur s'ouvre et augmente en puissance, l'isolation sonore est brisée.

Une expérimentation de détection acoustique d'inverseurs de flux (poussée pour les jets et les pas d'hélices des turboréacteurs) a été menée sur la plateforme de Nice - Côte d'Azur en 2009 [12]. Le bruit a été surveillé en continu par une station de mesure Oper@. Les niveaux de bruit globaux et les données spectrales 1/3 octave ont été mesurés toutes les secondes. De plus, un fichier audio a été enregistré sur dépassement de seuil ($L_{A_{Max}}$ supérieur à 75 dB(A) avec 20 secondes avant et après le niveau max). Une zone d'analyse des trajectoires a sélectionné les trajectoires d'arrivée. Une corrélation basée sur le temps d'apparition des trajectoires dans la zone et le temps d'apparition des événements acoustiques a été opérée. Cette corrélation a permis de consolider les résultats de la reconnaissance acoustique et d'identifier le type d'avion à l'origine du bruit d'inverseurs de flux observé.



Figure 6 : emplacement du capteur sonore sur Nice Côte d'Azur pour la détection acoustique des inverseurs de flux en atterrissage.

Une analyse manuelle quasi systématique des enregistrements sonores du mois de mars 2009 a été réalisée dans les domaines temporel et fréquentiel. Elle a permis de constituer un échantillon test d'événements d'inverseurs de flux.

La DSNA a fourni une base de données d'apprentissage constituée de 66 fichiers audio de reverse, 26 fichiers audio de décollages jet ou hélicoptères et 37 fichiers audio de décollages. ORELIA a alors réalisé l'extraction d'un module de reconnaissance des bruits de reverse. La performance de ce dernier a été testée sur une base de données indépendante du corpus d'apprentissage.

Sur 195 bruits de reverse étudiés répartis sur une durée de 3 mois, 168 sont correctement détectés soit un taux correct de 86%. Pour l'ensemble de ces bruits de reverse identifiés, plus de 4608 avions ont également atteint ou dépassé le seuil de 75 dB(A). 5% de ces avions ont été identifiés comme reverse par le système. Cela est très peu compte tenu du fait que les bruits de reverse et les bruits d'avion au décollage proviennent de la même source sonore et ils ont donc un timbre très proche.

Chaque jour, il y a 150 arrivées dont 6 bruits de reverse. L'utilisation du logiciel OSSR permet de détecter 5 bruits de reverse. Il détecte également 8 avions qui ne sont pas des reverses. De la sorte, la technologie de reconnaissance sonore permet de réduire considérablement le nombre de fichiers audio à traiter de manière journalière. En moyenne, il permet de réduire de 87% le nombre de données à traiter par un opérateur.

Ainsi, après détermination d'un site de mesurage adapté sur l'emprise d'une plateforme aéroportuaire, il devient donc possible de mettre en place un système de suivi quasi automatisé des émissions acoustiques d'inverseurs de flux.

D'autres opérations sont à prévoir sur d'autres aéroports. Grâce notamment à l'implémentation de nouveaux descripteurs audio, notre objectif est de dépasser les 90% de réduction des données à traiter.

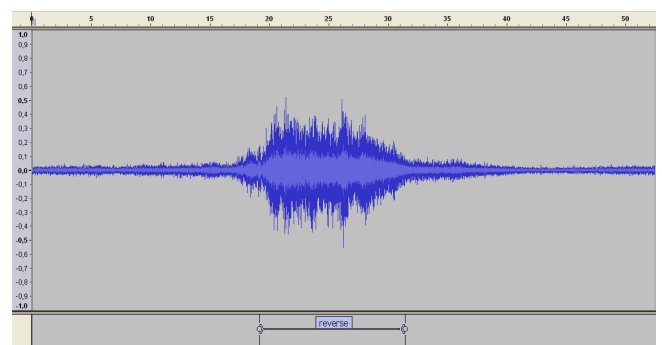


Figure 7 : exemple de visualisation sur Audacity® d'un codage d'un bruit de reverse effectué par le logiciel OSSR.

5. Ouverture vers la détection des sources industrielles

Les industries sont fréquemment implantées en périphérie des villes au sein de zones qui leur sont réservées. Malgré cela, il n'est pas rare que certaines habitations soient voisines de ces sites. Les bruits industriels constituent alors une réelle nuisance pour les riverains. Les établissements classés font l'objet d'une réglementation spécifique autorisant des émergences diurnes et nocturnes limitées, à respectivement 5dB(A) et 3dB(A) vis-à-vis de l'ambiance locale.

La surveillance acoustique de ces activités permet de contrôler la validité réglementaire des installations vis-à-vis du bruit émis dans l'environnement, de manière continue. Une reconnaissance automatique des sources sonores permet de justifier si les émergences caractérisées correspondent à des activités liées à l'industrie ou aux infrastructures voisines.

Ainsi, diverses solutions existent en termes de reconnaissance automatique de sources à l'aide du logiciel OSSR :

- La configuration sur site est telle qu'il s'avère judicieux de coder les véhicules circulant au voisinage du site. Dans ce cas, le bruit résiduel peut être attribué à l'industrie et les sources codées automatiquement à l'aide du module « bruit routier » voient leurs contributions énergétiques ôtées du bruit global.
- Les sources de bruit industriel sont particulièrement audibles dans le bruit global. Un apprentissage spécifique de reconnaissance de forme industrielle peut être initié. Un nouvel ACE (ou signature sonore) est constitué. Chaque événement correspondant pourra ensuite être détecté à l'aide du logiciel OSSR. L'industriel peut alors communiquer en toute transparence, sur l'impact sonore de son établissement sur l'environnement. Les sources « parasites » au site n'entrent pas dans les contributions énergétiques liées aux activités industrielles locales.

6 Conclusion

Innovation fondamentale pour l'environnement, le logiciel OSSR constitue en complément des systèmes de mesure longues durée tel Oper@. OSSR est un outil particulièrement adapté à la détection des sources lors de la surveillance acoustique des aéroports, de gestion sonore des villes, de surveillance acoustique des activités de loisirs...

En environnement urbain, l'utilisation du logiciel OSSR d'analyse de fichiers audio configuré selon l'emplacement des systèmes de mesure et les différentes sources sonores perçues révèle des performances intéressantes de l'ordre de 90 à 95% de réussite de détection.

Références

- [1] B. Defreville and C. Lavandier, "The contribution of sound source characteristics in the assessment of urban sound-scapes", *Acta Acustica united with Acustica* 92, 912–921 (2006).
- [2] B. Barbo, C. Rosin, J-M. Machet, B. Defreville. "A pattern recognition approach for aircraft noise detection", *Internoise* 2009.
- [3] Rosin C., Barbo B., Defreville B., "Monitoring du bruit des avions : une détection à partir du signal audio", *Congrès français d'acoustique* (2010).
- [4] J. Blauert and U. Jekosh, "Sound quality evaluation, a multi-layered problem", *Acta Acustica united with Acustica* 83, 747–753 (1997).
- [5] H. Almuallim and T. G. Dietterich. "Learning boolean concepts in the presence of many irrelevant features". *Artificial Intelligence*, 69(1-2):279–305, 1994.
- [6] D. Koller and M. Sahami. "Toward optimal feature selection". In *Proceedings of the Thirteenth International Conference on Machine Learning*, pages 284–292, 1996.
- [7] A. L. Blum and R. L. Rivest. "Training a 3-node neural networks is NP-complete". *Neural Networks*, 5:117 – 127, 1992.
- [8] World Wide Web Consortium. *eXtensible Stylesheet Language Transformations*. <http://www.w3.org/TR/xsl.1999>.
- [9] Ian H. Witten, Eibe Frank "Data Mining. Practical machine learning tools and techniques"(2nd edition). Morgan Kaufmann, June 2005.
- [10] Code de bonne conduite environnementale. Mars 2006. http://www.snpl.com/communication/communiques/lire/338/signature_a_marseille_d_un_code/
- [11] Préfecture de la Gironde. "Arrêté portant établissement du Plan de Prévention du Bruit dans l'Environnement de l'aérodrome de Bordeaux-Mérignac" – 17 mars 2009. http://www.dac-so.aviation-civile.gouv.fr/index.php?option=com_docman&task=doc_download&gid=103
- [12] D. Martin. "Étude sur la détection acoustique des inverseurs de flux" – DGAC Mission Environnement (2009).