

**CFA '18 LE HAVRE ■ 23-27 avril 2018**  
**14<sup>ème</sup> Congrès Français d'Acoustique**



**Constitution d'une base de données physiquement valide pour les approches de localisation de sources par Deep Learning sur antennes microphoniques intelligentes**

H. Pujol, É. Bavu et A. Garcia

Lab. de Mécanique des Struct. et des Syst. Couplés (CNAM Paris), 2 rue Conté (Case courrier 2D6R10),  
75003 Paris, France  
hadrien.pujol@lecnam.net

Afin de s'affranchir des modèles de sources, d'environnement, de capteurs et d'antennes, nous développons actuellement une approche de localisation de sources par Deep Learning, basée sur l'utilisation de réseaux de neurones profonds.

L'explosion de l'efficacité de ces méthodes dans les domaines de la reconnaissance d'image et de parole repose en partie sur l'architecture du réseau de neurones et des techniques de rétro-propagation utilisées, mais également sur l'utilisation de bases de données réalistes de très grandes dimensions.

Dans notre cas, la qualité de l'apprentissage pour la localisation de sources sonores repose également sur la qualité et le réalisme de la base de données d'apprentissage "labellée", c'est à dire sur l'obtention de données mesurées pour un grand nombre de sources, dont on connaît parfaitement la position (le "label"), dans des environnements potentiellement bruités et réverbérants. Pour cela, elle doit également être acquise sur une antenne réelle de capteurs, avec ses défauts, le plus souvent négligés par les approches de type "modèles".

Nous présenterons la manière dont nous constituons notre base de données grâce à une synthèse de champ en temps réel par Ambisonie d'ordres élevés sur la sphère de spatialisation "SpherBedev" du Cnam. Le champ est reconstruit physiquement avec une grande précision dans le "sweet spot" contenant l'antenne microphonique. De plus, nous utilisons un grand nombre de type de signaux émis, grâce à des bases de données ouvertes, classiquement utilisées pour de la reconnaissance de sons urbains ou de signaux vocaux. Nous ajoutons également à cette synthèse physique des profils de bruit spatialisés, enregistrés grâce au microphone ambisonique MemsBedev. L'influence de l'environnement de mesure peut également être encodé en composantes ambisoniques, ouvrant la voie à l'obtention d'excellents résultats de localisation de sources par Deep Learning dans un grand nombre de cas expérimentaux.