

# Thèse de Doctorat

Grégoire LAFAY

**Simulation de scènes sonores environnementales :  
application à l'analyse sensorielle et à l'analyse automatique.**  
**Simulation of environmental acoustic scenes: application to sensory and computational analyses.**

## Résumé

La présente thèse traite de l'analyse de scènes extraites d'environnements sonores, résultat auditif du mélange de sources émettrices distinctes et concomitantes. Ouvrant le champ des sources et des recherches possibles au-delà des domaines plus spécifiques que sont la parole ou la musique, l'environnement sonore est un objet complexe. Son analyse, le processus par lequel le sujet lui donne sens, porte à la fois sur les données perçues et sur le contexte de perception de ces données.

Tant dans le domaine de la perception que de l'apprentissage machine, toute expérience suppose un contrôle fin de l'expérimentateur sur les stimuli proposés. Néanmoins, la nature de l'environnement sonore nécessite de se placer dans un cadre écologique, c'est à dire de recourir à des données réelles, enregistrées, plutôt qu'à des stimuli de synthèse.

Conscient de cette problématique, nous proposons un modèle permettant de simuler, à partir d'enregistrements de sons isolés, des scènes sonores dont nous maîtrisons les propriétés structurelles -- intensité, densité et diversité des sources. Appuyé sur les connaissances disponibles sur le système auditif humain, le modèle envisage la scène sonore comme un objet composite, une somme de sons sources.

Nous investissons à l'aide de cet outil deux champs d'application. Le premier concerne la perception, et la notion d'agrément perçu dans des environnements urbains. L'usage de données simulées nous permet d'apprécier finement l'impact de chaque source sonore sur celui-ci. Le deuxième concerne la détection automatique d'événements sonores et propose une méthodologie d'évaluation des algorithmes mettant à l'épreuve leurs capacités de généralisation.

## Mots-clés

Simulation de données, détection d'événements sonores, analyse de scènes auditives, évaluation perceptive, paysage sonore urbain, qualité des paysages sonores, cognition ancrée, psychologie cognitive.

## Abstract

This thesis deals with environmental scene analysis, the auditory result of mixing separate but concurrent emitting sources. The sound environment is a complex object, which opens the field of possible research beyond the specific areas that are speech or music. For a person to make sense of its sonic environment, the involved process relies on both the perceived data and its context.

For each experiment, one must be, as much as possible, in control of the evaluated stimuli, whether the field of investigation is perception or machine learning. Nevertheless, the sound environment needs to be studied in an ecological framework, using real recordings of sounds as stimuli rather than synthetic pure tones.

We therefore propose a model of sound scenes allowing us to simulate complex sound environments from isolated sound recordings. The high level structural properties of the simulated scenes -- such as the type of sources, their sound levels or the event density -- are set by the experimenter. Based on knowledge of the human auditory system, the model abstracts the sound environment as a composite object, a sum of sound sources.

The usefulness of the proposed model is assessed on two areas of investigation. The first is related to the soundscape perception issue, where the model is used to propose an innovative experimental protocol to study pleasantness perception of urban soundscape. The second tackles the major issue of evaluation in machine listening, for which we consider simulated data in order to powerfully assess the generalization capacities of automatic sound event detection systems.

## Key Words

Data simulation, sound event detection, auditive scenes analysis, perception assessment, cognition, urban soundscape, soundscape quality, grounded cognition, cognitive psychology.