



IRIT

Informatics Research
Institute of Toulouse

Within the framework of
ERC FACTORY



Monday 10 December 2018

10h30 – 12h00

INP-ENSEEIH, Salle des thèses

Simon LEGLAIVE

Inria Grenoble Rhône-Alpes

Audio Signal Modeling for Source Separation - From Hand-designed to Learned Probabilistic Priors

Abstract: Under-determined audio source separation is an ill-posed inverse problem whose goal is to recover audio source signals from the observation of one or several mixtures. On one hand, Bayesian approaches regularize the inverse problem by defining priors over the audio source signals. Informative priors will enforce audio-specific features to the solution, such as sparsity or low-rankness in the time-frequency domain. Designing such priors is therefore an important aspect of solving the audio source separation problem, and traditionally, priors are hand-designed based on some kind of expert knowledge. On the other hand, discriminative deep learning methods for source separation try to directly map the observed mixture to the separated sources. This is achieved by learning a deep neural network on a large dataset. Despite their impressive performance, such specially-trained systems cannot be adapted to slightly different configurations, such as a change in the number of microphones. Ideally, we would like to develop methods in between those two worlds, leveraging both the flexibility of Bayesian approaches and the efficiency of deep learning methods.

In the first part of this talk, we will present a Bayesian framework for under-determined audio source separation in multichannel reverberant mixtures. Source signals and mixing filters are modeled as Student's t latent random variables respectively in the time-frequency and in the time domains. We propose a variational inference algorithm in order to perform source separation.

In the second part of this talk, we will show how to learn directly from data a speaker-independent generative speech model based on deep neural networks. This model is used for a single-channel semi-supervised speech enhancement application. We propose a Monte Carlo expectation-maximization algorithm. Experiments show that the proposed method outperforms a semi-supervised non-negative matrix factorization baseline and a fully-supervised discriminative deep learning approach.

Seminar

05 61 55 65 10

info@irit.fr

www.irit.fr



Organized by
Department SI - Team SC