PhD position in machine learning in Toulouse
Representation learning for audio signal processing

Description of work

The short-time Fourier transform (STFT) is a classical & ubiquitous off-the-shelf front-end to many audio signal processing tasks such as denoising or source separation. In particular, the STFT can be interpreted as a one-layer convolutional neural net with pre-specified weights (Fourier transform) [1]. The aim of the thesis is to move from this rigid architecture to adaptive representations that involve learning parts of the network (such as the window [2] or the transform [3]) or considering deep variants [4, 5]. Target applications are in source separation [6] and phase retrieval [7].

Advisors

Thomas Oberlin (IRIT) thomas.oberlin@enseeiht.fr
Cédric Févotte (IRIT) cedric.fevotte@irit.fr

Place of work

The position is part of project **FACTORY** (*New paradigms for latent factor estimation*), funded by the **European Research Council** under a Consolidator Grant (2016-2021) and coordinated by Cédric Févotte (CNRS). FACTORY is hosted by the **Institut de Recherche en Informatique de Toulouse** (IRIT), a joint laboratory of CNRS and Toulouse universities & engineering schools. The physical location for the project is the ENSEEIHT campus, in a lively neighbourhood of the city center.

Candidate profile and application

Prospective applicants should have a MSc in **signal processing**, **machine learning**, **applied mathematics**, **statistics**, or a related discipline, good programming skills, and good communication skills in English, both written and oral.

Salary is ∼1400€ net per month and may be complemented with teaching or consulting activities (subject to availability). The position comes with health insurance & other social benefits. The targeted starting period is autumn 2018.

Applicants are requested to send a CV, academic transcripts and the contact details of one or two referees in a single PDF file. Applications and informal enquiries are to be emailed to the two advisors listed above. Applications will be collected until mid-September until and then on until a suitable candidate is found.

References

[1] M. Dörfler, R. Bammer and T. Grill. Inside the spectrogram: Convolutional Neural Networks in audio processing. *In Proc. International Conference on Sampling Theory and Applications (SampTA)*, 2017.

[2] N. Perraudin, N. Holighaus, P. L. Søndergaard and P. Balazs. Designing Gabor windows using convex optimization. *Applied Mathematics and Computation*, 2018.

[3] S. Ravishankar and Y. Bresler. Learning sparsifying transforms. *IEEE Trans. Signal Processing*, 2013.

[4] G. Hinton, S. Osindero and Y. W. Teh. A fast learning algorithm for deep belief nets. *Neural Computation*, 2006.

[5] J. Andén and S. Mallat. Deep scattering spectrum. *IEEE Transactions on Signal Processing*, 2014.

[6] S. Venkataramani, C. Subakan and P. Smaragdis. Neural network alternatives to convolutive audio models for source separation. *In Proc. IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, 2017.

[7] K. Oyamada, et al. Generative adversarial network-based approach to signal reconstruction from magnitude spectrograms. *arXiv:1804.02181*, 2018