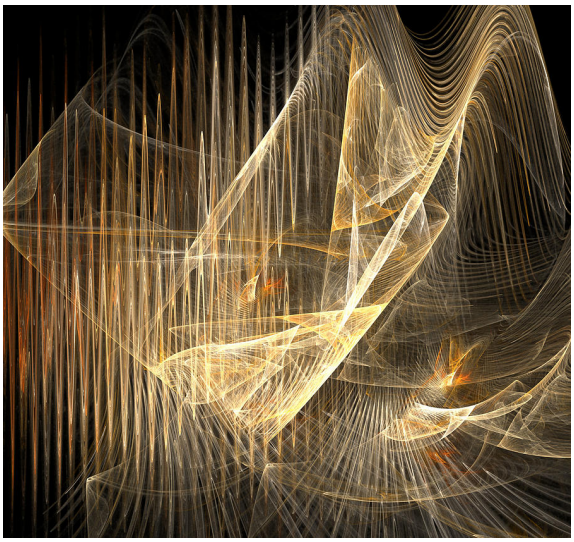
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Distributed Computing

Prof. R. Wattenhofer

Audio Source Separation with Deep Learning

Source separation is an established research field in signal processing. Most people have probably heard of a specific kind of source separation problem, called the "Cocktail Party Problem", where one hears a mixture of many voices and wants to focus in on one particular conversion.



In general, the input signal consists of one or more channels (mono vs. stereo), and one would like to separate it into its constituent sources, e.g., voice and instruments. In this thesis we want to continue previous work we have done on mono source separation. In particular, we want to improve our existing deep learning models, come up with and evaluate novel data augmentation and post processing methods, and investigate the differences between mono and stereo source separation. The latter topic is interesting because it is not entirely clear (neither from a biological nor signal processing standpoint), why exactly stereo source separation is easier. One aspect is of course the spatial information contained in a stereo signal, but it is likely not the whole story.

If this sounds interesting to you, do not hesitate to contact us so we can have a chat.

Requirements: Knowledge in Deep Learning, or solid background in Machine Learning. Implementation experience is an advantage. You should be able to read and understand the first 12 chapters of the "Deep Learning Book" by Goodfellow et al. (available for free online from MIT press). If you are interested in the topic but new to deep learning we expect you to complete an introductory deep learning course before applying for the thesis, such as Andrew Ng's coursera course (use the free trial!)[1] or this Udacity course[2].

Interested? Please contact us for more details!

Contacts

- Gino Brunner: brunnegi@ethz.ch, ETZ G63
- Oliver Richter: richtero@ethz.ch, ETZ G63

[1] https://www.coursera.org/specializations/deep-learning
[2] https://classroom.udacity.com/courses/ud730