# Proposal for a Master Thesis

Topic: **STFT Bin Selection for Sparsity-based Algorithms**

Description: A broad range of speech signal processing algorithms, e.g., direction of arrival estimation, source localization, blind source separation, or source counting, is based on the assumption of sparsity of the mixture of sources in the STFT domain [1]. Under this assumption, only one source is active at any time/frequency tile in a speech signal mixture, such that the corresponding short-time Fourier spectra of the single sources do not overlap. Most of these algorithms apply clustering algorithms to separate the source signals w.r.t. a specific feature, e.g., the direction of arrival.

However, many of the observed STFT tiles do not contain useful information due to reverberation, noise or speech inactivity. The performance of these algorithms can be improved by selecting only the best STFT tiles for the subsequent clustering. At the same time, the computational load is reduced due to the smaller data set which has to be clustered.

The aim of this thesis is the implementation and evaluation of measures for the selection of STFT tiles for algorithms based on the assumption of sparsity of speech signal mixtures in the STFT domain. First the signal energy of an STFT tile, the speech presence probability [2] and the coherent-to-diffuse power ratio [3] should be investigated. The selection of the STFT tiles can be applied for algorithms for the estimation of directions of arrival [4] or the localization and tracking of multiple speakers [5]. In addition, a literature survey of similar selection procedures is a mandatory part of the thesis.

http://people.csail.mit.edu/mrub/VisualMic/
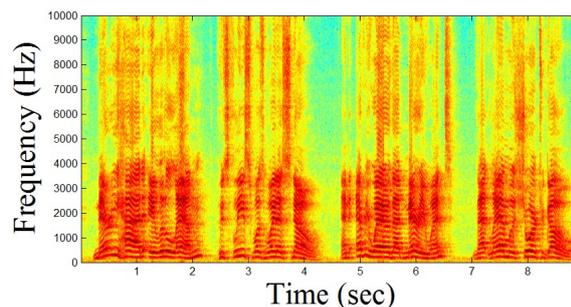
[1] S. Rickard. "Sparse Sources Are Separated Sources." EUSPICO, Florence, Italy, 2006.

[2] T. Gerkmann and R. C. Hendriks. "Noise Power Estimation Based on the Probability of Speech Presence." WASPAA, 145–48. New Paltz, USA, 2011.

[3] A. Schwarz and W. Kellermann. "Coherent-to-Diffuse Power Ratio Estimation for Dereverberation." TASLP. 23, no. 6 (Jun. 2015): 1006–18.

[4] S. Araki, H. Sawada, R. Mukai and S. Makino. "DOA Estimation for Multiple Sparse Sources with Arbitrarily Arranged Multiple Sensors." J. of Signal Process. Syst. 63, no. 3 (Jun. 2011): 265–75.

[5] O. Schwartz and S. Gannot. "Speaker Tracking Using Recursive EM Algorithms." TASLP 22, no. 2 (Feb. 2014): 392–402.

Supervisor: M.Sc. Andreas Brendel, room 01.178 (Wetterkreuz 15), Andreas.Brendel@FAU.de

Professor: Prof. Dr.-Ing. Walter Kellermann

Available: Immediately