Title: Modeling interaural cues with Complex Student distributions

Environment :
INRIA team Mistis (Statistics), Grenoble, and INRIA team Panama (audio signal processing)

Supervisors: Florence Forbes (florence.forbes@inria.fr) Web site: http://mistis.inrialpes.fr/
Antoine Deleforge (antoine.deleforge@inria.fr), Web site: https://team.inria.fr/panama/
In collaboration with INRIA team Perception.

Location: INRIA Grenoble

Topic:
Given a scene recorded with a camera and 2 microphones, can we determine where the sound comes from in the camera field of view? Or in other words, can we determine the "pixels that sound"? This is a fundamental question of audio-visual scene analysis, with numerous applications, notably in human-robot interaction. In this internship, the problem will be addressed using advanced statistical modeling and machine learning. The student will be asked to extend and implement a recent statistical model referred to as "Student Locally-linear Mapping"[4,5]. This model performs locally affine regression through a mixture of Student's t-distributions, and was showed to be particularly suited when observations are ratios of noisy complex-valued data [1,3]. This feature will be used to map the differences of level and phase recorded at microphones to the 2D position of the sound source in the image [2]. The student will use the audio visual dataset AVASM recorded at Inria Grenoble in order to train and test the model ().

Skills :
We look for candidates strongly motivated by challenging research topics. The applicant should have good background in statistics and computer science (good programming skills). The required knowledge includes ideally Bayesian statistics, EM algorithm, mixture models, graphical models. As regards to software developments, R, Matlab, or C languages are possible. The successful candidate should have good oral and writing communication skills in English.

References
[1] R. J. Baxley, B. T.Walkenhorst, and G. Acosta-Marum, “Complex gaussian ratio distribution with applications for error rate calculation in fading channels with imperfect csi,” in Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE. IEEE, 2010, pp. 1–5.
[2] E. Vincent, S. Arberet, and R. Gribonval, “Underdetermined instantaneous audio source separation via local gaussian modeling,” in Independent Component Analysis and Signal Separation. Springer, 2009, pp. 775–782.
[3] https://team.inria.fr/perception/the-avasm-dataset/
[4] E. Perthame and F. Forbes, SLLiM and hybrid SLLiM: Student mixtures for locally-linear mapping. Unpublished manuscript 2015.
[5] D. Wraith, F. Forbes, Location and scale mixtures of Gaussians with flexible tail behaviour: Properties, inference and application to multivariate clustering. Computational Statistics and Data Analysis, 90:61--73,2015

Key-words: Statistics, mixture models, high dimension, generalized Student distribution, sound source separation.