

Audio Signal Enhancement for the Socially Intelligent Robot Haru

Internship Offer

Dates: Min. 3 months between September 1st 2019 and April 1st 2020

Internship advisors: Antoine Deleforge (Inria Researcher, Team MULTISPEECH)

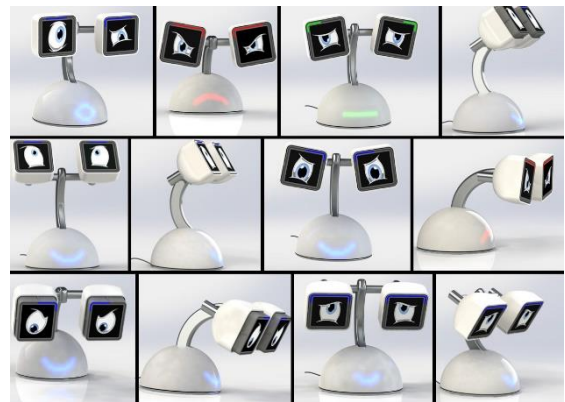
Contact: antoine.deleforge@inria.fr

Location: Center Inria of Nancy – Grand Est. 615 Rue du Jardin-Botanique, 54600 Villers-lès-Nancy, France.

Requirements: Bachelor degree in signal processing/machine learning or higher. Good programming skills in C/C++ and Python. Previous experience in audio and robotics is a plus.

Topic description:

Haru is an autonomous robotic platform consisting of a base and two screens mimicking a face with two eyes. The robot is equipped with actuators (for the whole body, the neck, and the eyes) as well as cameras and microphones for visual and auditory sensing. In addition, it may rely on microphones distributed in the environment. An international consortium of academic and industrial collaborators is working on making the robot “socially intelligent”, i.e., make it interact with users naturally by detecting who is speaking when, to whom, with which emotion and by reacting accordingly, e.g., by displaying images on its screens.



Among this consortium, Inria is responsible for conducting research in collaboration with the Honda Research Institute (HRI) in Japan on two audio signal-processing tasks: speech enhancement and auditory speaker localization. These two tasks are related to each other, since better speaker localization can translate into better speech enhancement. The goal of the internship is to extend and adapt sound source localization [1] and speech enhancement [2] methods developed at Inria so that they can be later integrated into the Haru interaction framework by HRI. Specific challenges include adaptations to online rather than batch processing, to multiple speaker conditions, to distributed microphone arrays and robustness to reverberant environments.

Internship Conditions:

The internship will take place at Inria Nancy under the supervision of Antoine Deleforge.

Bibliography:

[1] Lebarbenchon, Romain, et al. "Evaluation of an open-source implementation of the SRP-PHAT algorithm within the 2018 LOCATA challenge." arXiv preprint arXiv:1812.05901 (2018).

[2] Nugraha, Aditya Arie, Antoine Liutkus, and Emmanuel Vincent. "Multichannel audio source separation with deep neural networks." IEEE/ACM Transactions on Audio, Speech, and Language Processing 24.9 (2016): 1652-1664.