probabilistic models, and to develop the optimization frameworks from raw data design, to conceive their combination with expected to design the deep architectures able to extract cues difficult and has to be done with care. The joint training of probabilistic and deep neural models is present PhD work, this kind of models will be combined with and associated solutions based on variational inference. In the which we have explored generative/Bayesian probabilistic models in multi-person tracking growing exponentially with time. A clear example of this is found on the temporal dynamics, leads to a combinatorial problem to multiple persons, together with the impact this assignment has. The fact that multiple features are extracted and can be assigned efficiently represent a scene, and develop and evaluate the observation-to-person assignment problem and account for a time-varying number of people. Therefore, we have to conceptualize parametric models that are able to faithfully and efficiently represent a scene, and develop and evaluate the associated parameter estimation algorithms. Importantly, for the sake of interpretability, the representation should be structured into a set of individual cues per person plus a set of collective cues. We will inspire on state-of-the-art techniques for visual person (body, face) detection and for person description (appearance, pose, orientation), on the one side, and on speech processing (speech enhancement, speech and speaker automatic recognition), on the other side. Part of the informative features will be extracted using learnable parametric methods, e.g. deep neural networks (DNNs). We will need to investigate how to fine-tune these architectures to satisfy the goals of the project, and to adapt to the data distribution of multi-person conversational scenarios. Once these features are conceived and learned, we will be able to perform joint inference of individual and collective cues, and define and address the combinatorial assignment problem.

The fact that multiple features are extracted and can be assigned to multiple persons, together with the impact this assignment has on the temporal dynamics, leads to a combinatorial problem growing exponentially with time. A clear example of this is found in multi-person tracking [3, 4] and in sound separation [5] for which we have explored generative/Bayesian probabilistic models and associated solutions based on variational inference. In the present PhD work, this kind of models will be combined with DNNs. The joint training of probabilistic and deep neural models is difficult and has to be done with care [6]. The PhD student will be expected to design the deep architectures able to extract cues from raw data design, to conceive their combination with probabilistic models, and to develop the optimization frameworks.
and algorithms able to soundly optimize for the overall set of parameters. In this regard the experience of the team in studying generative/Bayesian probabilistic models and deep neural architectures [7] will serve as a basis to build upon.

References


Principales activités

The PhD work will take place at Inria Grenoble, in Montbonnot-Saint-Martin, in the Perception Team, headed by Radu Horaud. It will be supervised by Laurent Girin (Professor Grenoble-INP, http://www.gipsa-lab.grenoble-inp.fr/~laurent.girin/cv.html) & Xavier Alameda-Pineda (Inria Research Scientist, http://xavirema.eu/).

Compétences

Research Master’s degree, or equivalent, in a discipline connected to signal and information processing, computer vision and machine learning. A particular interest/experience in speech/audio processing, visual recognition, and/or multimodal fusion is a plus.

Strong motivation for the research work. Ability to work both independently and to collaborate within a small team. Oral and written English communication skills are highly appreciated.

Computer skills: MATLAB, Python, Deep Learning Toolkits (e.g. Keras, Pytorch).

Avantages

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking (after 6 months of employment) and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

Rémunération

Salary (before taxes) : 1982€ gross/month for 1st and 2nd year. 2085€ gross/month for 3rd year.