

2019-01579 - PhD Position F/M Multimodal interaction data generation using deep adversarial learning

Type de contrat : CDD de la fonction publique
Niveau de diplôme exigé : Bac + 5 ou équivalent
Fonction : Doctorant

A propos du centre ou de la direction fonctionnelle

Grenoble Rhône-Alpes Research Center groups together a few less than 800 people in 35 research teams and 9 research support departments.

Staff is localized on 5 campuses in Grenoble and Lyon, in close collaboration with labs, research and higher education institutions in Grenoble and Lyon, but also with the economic players in these areas.

Present in the fields of software, high-performance computing, Internet of things, image and data, but also simulation in oceanography and biology, it participates at the best level of international scientific achievements and collaborations in both Europe and the rest of the world.

Contexte et atouts du poste

The PhD student will be co-supervised by Dr. Dominique Vaufreydaz (Pervasive team, supervisor, <https://research.vaufreydaz.org/>) and by Dr. Xavier Alameda-Pineda (Perception team, co-supervisor, <https://xavirema.eu/>). This thesis will benefit from the background of the two teams in multimodal perception and interaction.

The funding of the thesis is already acquired through the IRS MIDGen Project (University Grenoble Alpes). The Ph.D. student will be hired by University and work at Inria.

Mission confiée

The funded project MIDGen (Multimodal Interaction Data Generation) addresses researches within the so-called "Ambient Assisted Living" (AAL) research field, more precisely on assistance to elderly or frail people. This topic is of interest as a societal challenge in the near future. The current researches on assistive systems for elderly range from smartphone helping applications to companion robots at home. The main challenge in building such companion robots is to provide social competence in perceiving, reasoning and expressing the social and emotional aspects of interactions with human, also known as the "social presence". To fulfill the prerequisites of this social presence, multimodal perception algorithms must accurately perceive social signals emitted by humans. Many techniques have emerged in the state-of-the-art in recent years, taking advantage of the progresses of Deep Learning. Some are remarkably efficient on tasks such as detecting people in images, detecting or recognizing speech. The MIDGen project focuses on the need of current Deep Learning perception algorithms for a huge amount of training data, data that are not available in sufficient quantity for our research on multimodal interactions with elderly people.

In the context of the project, the collection and annotation of large-scale data sets poses ethical and privacy concerns, and requires enormous resources. We plan to overcome this issue by learning data generators (e.g. generative adversarial networks and variational auto-encoders) from already available and less sensitive corpora (e.g. political debates). These generators will be able to synthesize large-scale data sets automatically annotated by the generation process. Limited progresses depicted in the literature showed that it is possible to generate facial cues or full body poses of a single individual, but little is known of how to generate audio-visual data describing the interaction between two or more people (speech turns, conversational gestures, people's gaze, etc.).

Principales activités

The MIDGen proposal is structured into workpackages:

- controllable and domain-adaptive mono-modal data generators
- their multi-modal counterpart
- evaluation

The main challenge of the first two workpackages is to develop generators that are controllable and domain-adaptive. We decided to first address the controllability and adaptation problems of several mono-modal data generators (*workpackage 1*) to later on investigate multimodality (*workpackage 2*). The evaluation (*workpackage 3*) will be carried out throughout the project based on our expertise on multimodal scene analysis and human robot interaction. An intrinsic assessment of the quality of the generated multimodal data will be conducted to understand if the generated data is realistic enough to detect low-level features (people's position, speaking status, ...). Using Amiqua4Home (<https://amiqua4home.inria.fr/>) and/or Domus experimental platforms (<http://multicom.imag.fr/multicom.imag.fr/spipc5a9.html?article114>), the last evaluation steps will validate our generative approach on real live interactions with robots in ecological situations.

The PhD candidate will start with the generation of monomodal interaction data (Workpackage 1) during the whole first year. It will be followed by Workpackage 2, on multimodal data until the 30th month of the thesis. At the same time, evaluations will be conducted to refine and validate the results (Workpackage 3). The dissemination of

Informations générales

- **Thème/Domaine :** Robotique et environnements intelligents
- **Ville :** Montbonnot
- **Centre Inria :** CRI Grenoble - Rhône-Alpes
- **Date de prise de fonction souhaitée :** 2019-10-01
- **Durée de contrat :** 3 ans
- **Date limite pour postuler :** 2019-06-04

Contacts

- **Equipe Inria :** PERVASIVE
- **Directeur de thèse :**
Vaufreydaz Dominique /
dominique.vaufreydaz@inria.fr

A propos d'Inria

Inria, l'institut national de recherche dédié aux sciences du numérique, promeut l'excellence scientifique et le transfert pour avoir le plus grand impact. Il emploie 2400 personnes. Ses 200 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3000 scientifiques pour relever les défis des sciences informatiques et mathématiques, souvent à l'interface d'autres disciplines. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 160 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

Consignes pour postuler

Sécurité défense :

Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement :

Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.

Attention : Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

the work publish results in top conference (ICMI, ACM MM, ECCV, CVPR, IROS, ICRA).

Compétences

The candidate must hold a Master in Computer Sciences or in Applied Mathematics, with a background in signal processing, machine learning. Knowledge in Deep Learning is appreciated. Good programming skills are also required (C++, python).

Avantages

- Partial reimbursement of public transport costs
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

Rémunération

1768,55€/ month before taxes