

Fiche de stage

Sujet du stage : Des images à la compréhension des interactions sociales / From images to social interactions understanding

Encadrant·e : Isabelle Bloch / Catherine Pelachaud

Date de début du stage : Février-Avril 2022

Durée du stage : 6 mois

Niveau d'études souhaité : Master 2

Laboratoire d'accueil : ISIR (*Institut des Systèmes Intelligents et de Robotique*), Campus Pierre et Marie Curie, 4 place Jussieu, 75005 Paris.

Personne à contacter

Prénom Nom : Catherine Pelachaud et Isabelle Bloch

Email : catherine.pelachaud@upmc.fr; isabelle.bloch@sorbonne-universite.fr

Envoyer votre candidature par mail, avec [*sujet du stage*] en objet, votre CV, vos diplômes, les noms de vos référents et une lettre de motivation.

Date limite de dépôt de la candidature : janvier 2022

Description du stage (en français)

Sujet : Des images à la compréhension des interactions sociales

Résumé :

Les études sur l'interaction humain-humain ont introduit le concept de F-formation (Kendon, 1990) qui définit trois zones : privée, sociale et publique. Les participants qui ont une interaction sociale se placent dans une certaine formation spatiale. Ils peuvent se faire face, être côte à côte... Leur position et leur comportement, tels que l'orientation du corps, le comportement du regard, peuvent indiquer une grande quantité d'informations ; ils peuvent révéler des informations sur leur niveau d'engagement, leur centre d'intérêt, mais aussi la qualité de leur relation, leur degré d'intimité, pour n'en citer que quelques-uns. La position et le comportement des participants évoluent continuellement pour s'adapter aux comportements des autres et pour obéir à certaines normes socioculturelles.

Dernièrement, des modèles computationnels ont été conçus pour détecter si les personnes forment un groupe et sa formation en fonction de la proxémique et des comportements (Cabrera-Quiros et al, 2018). Une analyse plus poussée peut être poursuivie pour caractériser la dynamique de l'interaction sociale entre les participants. De tels modèles peuvent ensuite être utilisés pour piloter les comportements des robots lorsqu'ils interagissent avec des humains.

Objectifs du stage :

L'objectif de ce stage est d'analyser les interactions de groupe et leur évolution dans le temps. Nous nous appuyons sur des données existantes (images et vidéos) d'interaction de groupe qui ont été annotées à différents niveaux (activité, parole, rire, comportement non-verbal). Nous utiliserons d'abord la base de données MatchNMI (Raman&Hung, 19).

Sous la co-tutelle de :

Plusieurs étapes sont prévues :

1. Réaliser une étude bibliographique sur la détection des F-formations, en se concentrant en particulier sur les vidéos et l'évolution temporelle des formations et l'évolution temporelle des formations.
2. Effectuer des tests en utilisant la méthode développée par V. Fortier l'année dernière.
3. Étendre le modèle aux vidéos.
4. Effectuer des analyses d'actions sociales telles que la prédiction de qui sera le prochain orateur ou la relation sociale entre les interlocuteurs.
5. En fonction des réalisations du projet, évaluer les résultats en utilisant des agents virtuels.

Jeu de données MatchNMingle :

<http://matchmakers.ewi.tudelft.nl/matchnmingle/pmwiki/index.php?%20%20n=Main.TheDataset>

References:

- A. Kendon: Conducting interaction: Patterns of behavior in focused encounters. Cambridge University Press, 1990.
- L. Cabrera-Quiros, A. Demetriou, E. Gedik, L. v. d. Meij and H. Hung. The MatchNMingle dataset: a novel multi-sensor resource for the analysis of social interactions and group dynamics in-the-wild during free-standing conversations and speed dates. Transactions on Affective Computing, 2018
- C. Raman, H. Hung: Towards automatic estimation of conversation floors within F-formations. 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW), 2019.
- V. Fortier : Détecter, analyser et comprendre les interactions sociales à partir d'images. Rapport de stage de M2, Sorbonne Université, 2021.
- S. Thompson, A. Gupta, A.W. Gupta, A. Chen, and M. Vázquez. Conversational Group Detection with Graph Neural Networks. In International Conference on Multimodal Interaction, pp. 248-252, 2021.

Profil recherché : étudiant en master d'ingénierie ou d'informatique, étudiant en ingénierie.

Compétences requises : analyse d'images, langage de programmation

Candidatures : envoyez votre CV, vos diplômes, les noms de vos référents et une lettre de motivation (en format pdf) à catherine.pelachaud@upmc.fr, isabelle.bloch@sorbonne-universite.fr

Description du stage (en anglais)

Subject: From images to social interactions understanding

Abstract:

Studies in human-human interaction have introduced the concept of F-formation (Kendon, 1990) that defines three zones: private, social and public. Participants during social interaction place themselves in certain spatial formation. They can face each other, be side by side... Their position and behavior such as body orientation, gaze behavior can indicate a great quantity of information; they can reveal information about their level of engagement, their focus of interest but also the quality of their relationship, their degree of intimacy, to name a few. Participants' position and behavior evolve continuously to accommodate others' behaviors and to obey to some socio-cultural norms.

Lately computational models have been designed to detect if people form a group and its formation based on proxemics and behaviors (Cabrera-Quiros et al, 2018). Further analysis can be pursued to characterize the dynamics of the social interaction between participants. Such models can then be used to drive the behaviors of robots when interacting with humans.

Sous la co-tutelle de :

Internship Objectives:

The aim of this internship is to analyze group interaction and their evolution over time. We will rely on existing data (images and videos) of group interaction that have been annotated at different levels (activity, speaking, laughing, non-verbal behavior). We will first make use of the database MatchNMingle (Raman&Hung,19). Several steps are foreseen:

1. Perform a literature survey on F-formation detection, focusing in particular on videos and time evolution of the formations.
 2. Perform tests using the method previously developed by V. Fortier last year.
 3. Extend the model to videos.
 4. Perform analysis of social actions such as predicting who will be the next speaker or the social relationship between interactants.
 5. Depending on the project achievements, evaluate the results using virtual agents.
- <http://matchmakers.ewi.tudelft.nl/matchnmingle/pmwiki/index.php?%20%20n=Main.TheDataset>

References:

- A. Kendon: Conducting interaction: Patterns of behavior in focused encounters. Cambridge University Press, 1990.
- L. Cabrera-Quiros, A. Demetriou, E. Gedik, L. v. d. Meij and H. Hung. The MatchNMingle dataset: a novel multi-sensor resource for the analysis of social interactions and group dynamics in-the-wild during free-standing conversations and speed dates. Transactions on Affective Computing, 2018
- C. Raman, H. Hung: Towards automatic estimation of conversation floors within F-formations. 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW), 2019.
- V. Fortier : Détecter, analyser et comprendre les interactions sociales à partir d'images. Rapport de stage de M2, Sorbonne Université, 2021.
- S. Thompson, A. Gupta, A.W. Gupta, A. Chen, and M. Vázquez. Conversational Group Detection with Graph Neural Networks. In International Conference on Multimodal Interaction, pp. 248-252, 2021.

Required Profile: Master student in engineering or computer science, engineering student

Required skills: Image analysis, programming language

Applications: send CV, grades, names of referent and motivation letter (in pdf format) to catherine.pelachaud@upmc.fr, isabelle.bloch@sorbonne-universite.fr