Master Internship: from images to social interactions understanding

Sorbonne Université – ISIR and LIP6

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.

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.

MatchNMingle dataset: http://matchmakers.ewi.tudelft.nl/matchnmingle/pmwiki/index.php?n=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..

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