MASTER'S THESIS:

Speaker Change Cues in Video Conferencing Tools

DESCRIPTION

Video conferences are generally perceived as being less pleasant than their real-life counterparts would be, where all participants are co-present. Why is that? In this project we want to explore the hypothesis that one contributing factor is the lack of smooth turn-taking (that is, the change from one speaker to the next).

Kousidis & Schlangen (2015) have shown that a robot that follows a conversation (by turning its head and moving its eyes, towards either of two participants of a conversation) is perceived favourably when it uses a predictive turn-taking model (anticipating speaker changes and potentially moving before the end of the speaker's utterance), compared to when it uses a reactive model (moving only after the end of the utterance). The "likeabilty" (and other factors) of the robot was judged by overhearers who watched a video of its performance.

This project will explore whether similar conversational cues can be provided within standard video conferencing setups (where the current speaker is often indicated by a highlighted frame around their video tile). In the first part of the project, such a setting will be realised, taking the same corpus of conversations as used in Kousidis & Schlangen (2015). Experimental materials will be produced and an experiment will be designed for letting them be judged by naive participants. (Given the current situation, most likely through crowdwork platforms.) The results will then be statistically analysed.

In a potential second part (or potentially an independent second project), a live setting will be created (using technology such as HTML5's WebRTC) where video chat streams can be manipulated. In a first study, it will be explored whether actual speakers can provide explicit turn-release cues (e.g., by pushing a key when they are done speaking), and whether relaying such signals to their interlocutor leads to increased interaction quality.

Requirements: Besides general programming skills, project 1 requires familiarity with video editing tools or programming kits (or the willingness to acquire it), and experience with conducting and analysing user experiments. Project 2 in addition requires programming skills for setting up the interaction interface.

The project(s) will be jointly supervised by Prof. Sebastian Möller (TU Berlin, Quality Labs) and Prof. David Schlangen (U Potsdam, colab.potsdam).

LITERATURE

Kousidis, Spyridon, and Schlangen, David, 2015; The Power of a Glance: Evaluating Embodiment and Turn-Tracking Strategies of an Active Robotic Overhearer; In Proceedings of AAAI Spring Symposium on Turn-taking and Coordination in Human-Machine Interaction 2015 (pdf)

CONTACT

Thilo Michael, thilo.michael@tu-berlin.de
(Deadline: 31.03.2021)