

Computational methods of analysing nonverbal cues



★ The NOVICOM project, led by **Daniel Gatica-Perez** and Marie Curie fellow **Oya Aran** is developing methods to analyse nonverbal group communication via computational models. This could lead to tools for accurate interpretation of social meaning, as an alternative to word-based conversation

When we speak to another person, or people, how we say what we say is as important as what we're saying. Unconsciously or otherwise, listeners pick up on a myriad of nonverbal information which helps to make their minds up about the person speaking, placing value judgements on their words.

Nonverbal communication is a series of wordless messages transmitted through aural cues (voice quality, speaking style, rhythm, intonation) and also through visual cues (gestures; body language and posture; facial expression and gaze). These cues are used by all of us every day to predict the mood and personality of others, as well as to make sense of social relations, in a very wide range of situations. It has been shown that, in many social situations, humans can interpret nonverbal cues with high accuracy.

The NOVICOM project, based at the Idiap Research Institute in Switzerland, is developing new computational methods to automatically detect and analyse nonverbal behaviour in face-to-face conversations. NOVICOM is looking at communicative gestures and postures, working towards the integration of

multi-modal cues for analysis of group interaction.

The research is being led by Daniel Gatica-Perez, senior researcher at Idiap, and Oya Aran, Marie Curie fellow at Idiap. Aran was first acquainted with the Marie Curie Intra-European Fellowships Programme, which is funding NOVICOM, through her doctoral advisor at Bogazici University, Turkey. Aran then contacted Gatica-Perez at Idiap. Together they decided on the project idea and wrote the NOVICOM project proposal.

"Social interaction is a fundamental aspect of human life and is also a key research area in psychology and cognitive science," Gatica-Perez said. "Social psychologists have been researching the dimensions of social interaction for decades and found out that a variety of social communicative cues strongly determine social behaviour and interaction outcomes. Many of these cues are consciously produced, in the form of spoken language. However, besides the spoken words, human interaction also involves nonverbal elements, which are extensively and often unconsciously used in human communication.

"Our aim in this project, based on several findings of social psychology, is to develop computational systems that can automatically analyse social behaviour by observing conversations via sensing devices such as cameras and microphones. We focus on group conversations and aim to predict aspects of the underlying social context, including both individual actions and interactions among the people in the group. Examples include turn-taking, interest, dominance, and roles."

Supporting role in analysis

A computational analysis of nonverbal communication has, the project argues, great potential value for social sciences. Tools that improve collective decision-making or keep people in the loop during teleconferences could use technology to enhance and improve communication. But can complete analysis of highly complex nonverbal human interactions ever be fully harnessed, even by sophisticated technological means?

"Although people are quite successful at interpreting social behaviour, we are not always correct," said Aran. "During social interaction, people make decisions based

on the social cues they receive, however these decisions may be subject to change as the interaction continues. Some behavioural cues are more vague and harder to analyse. The computational analysis of social interaction aims to create systems that are capable of understanding some elements of social context and reason accordingly, but clearly the scientific problems are very challenging.

“We expect that in the future these systems will make mistakes as humans do, quite likely more often, but it is important to realise that these systems do not aim to replace human assessment but to support the interaction and communication between people and machines, so people will always be making the final decisions.”

Useful applications

NOVICOM is treading on new ground in that a visual approach to capturing nonverbal communication hasn't really been explored with a high degree of accuracy. An aural approach has been the benchmark of quality so far, partly because audio capture systems were available earlier than their video equivalents. Video is also more sensitive to environmental conditions and requires adequate resolution and frame rates. NOVICOM is using publicly available data collected in previous projects in which its researchers have participated, which includes audio and video recordings of four-people conversations. It has also designed a portable audio-visual recording system, which includes two webcams and a microphone array and is able to record four people sitting around a table. This system allows for the capture of group interaction outside the laboratory with people who volunteer to participate in a number of group discussion scenarios.

The project is concentrating on a selected number of key research tasks in social interaction analysis, including the automatic estimation of dominance in groups, the level of interest or engagement of the members of the group during their interaction, and their agreement/disagreement status.

“In these situations, people unconsciously display visual cues, in the form of gestures and body postures, which partly reveal their social attributes,” said Gatica-Perez. “For each task, our specific objectives are twofold. First we attempt to automatically detect the visual nonverbal cues that are displayed during interaction.

Second, we investigate multimodal approaches that integrate visual and aural nonverbal cues to infer dominance, interest, or agreement.”

With people talking in the presence of microphones and video cameras, it could be assumed that those participating could be forgiven for not acting as naturally as they might otherwise, and therefore sending out the wrong signals. The project's sensing technology attempts to reduce these effects.

“Social interaction capture systems should not use sensors that affect the interaction,” said Aran. “The subjects should be able to act naturally without any distraction caused by the sensors that are used. For instance, to record audio, we use a microphone array device that avoids the need to attach microphones to people. Regarding cameras, while it is true that people might be distracted or feel self-conscious at the beginning of an interaction, when they are in a natural environment and focused on an engaging task, they rapidly tend to forget about the sensors and act naturally.”

So far, the project has concentrated on the modelling of dominance, investigating the effect of visual nonverbal clues to estimate the most dominant and least dominant person in a group conversation. Data was collected by observers, who reported on perceived levels of dominance. The team developed multimodal fusion techniques to utilise audio and visual nonverbal information jointly, eventually showing that the visual information is necessary in order to achieve better performance in dominance estimation.

Techniques that analyse interactions via visual cues, in addition to the aural ones, can be used as part of larger systems that, for example, support collective decision-making in meetings by providing key information about the participants' behaviour and the interaction itself.

“The outcome of the analysis could be used to promote an increase of attention or to reset the focus of a conversation whenever appropriate,” said Gatica-Perez. “Alternatively in a teleconferencing system, these techniques could help to keep the interest level of the remote users. Finally, and very importantly, the methods we are investigating could also be used to develop tools for self-assessment, training, and education. From these potential applications, we believe that the project can ultimately contribute to support our work and daily life.” ★

At a glance

Full Project Title

Automatic Analysis of Group Conversations via Visual Cues in Nonverbal Communication (NOVICOM)

Project Objectives

Automatic detection of visual nonverbal communication cues and multimodal integration of visual and aural nonverbal cues for small group interaction analysis.

Project Partners

Idiap Research Institute, Martigny, Switzerland

Contact Details

Project Coordinator,
Idiap Research Institute
Centre du Parc, Rue Marconi 19, PO Box
592, CH - 1920 Martigny, Switzerland
T: +41 277 217 758 (Oya Aran)
T: +41 277 217 733 (Daniel Gatica-Perez)
E: oya.aran@idiap.ch
E: gatica@idiap.ch
W: idiap.ch/~oaran

Daniel Gatica-Perez (right)
Oya Aran (left)



Project Coordinators

Daniel Gatica-Perez is a senior researcher at Idiap Research Institute, Switzerland, where he directs the social computing group, developing computational models to analyse human behaviour from sensor data. His research integrates methods from multimedia signal processing, machine learning, ubiquitous computing, and elements from social sciences to address questions related to the discovery, recognition, and prediction of behaviour in real life.

Oya Aran received the BS, MS and PhD degrees in Computer Engineering from Bogazici University, Istanbul, Turkey. She was awarded a Marie Curie Intra-European fellowship in 2009 with the NOVICOM project. She joined the Idiap Research Institute as a postdoctoral researcher and a Marie Curie fellow.

