

DEPARTMENT OF PHILOSOPHY, LINGUISTICS AND THEORY OF SCIENCE

Look on my thesis, ye mighty:

Gaze Interaction and Social Robotics

Vidya Somashekarappa

April 2024

Thesis submitted for the Degree of Doctor of Philosophy in Computational Linguistics, to be publicly defended, by due permission of the Dean of the Faculty of Arts at the University of Gothenburg, on April 25, 2024, at 13:15 in J330, Humanisten, Reströmögatan 6, Gothenburg.

*Opponent: Maîtreöse de conférences/senior lecturer Dominique Knutsen,
Univerité de Lille*



UNIVERSITY OF GOTHENBURG

Abstract

Title: Look on my thesis, ye mighty: Gaze Interaction and Social Robotics
Author: Vidya Somashekarappa
Supervisors: Asad Sayeed and Christine Howes
Language: English
ISBN: 978-91-8069-659-3 (print)
ISBN: 978-91-8069-660-9 (pdf)
Online: <http://hdl.handle.net/2077/80112>
Keywords: gaze behaviour, multimodal dialogue, machine learning, computer vision, human-robot Interaction, social robotics

Gaze, a significant non-verbal social signal, conveys attentional cues and provides insight into others' intentions and future actions. The thesis examines the intricate aspects of gaze in human-human dyadic interaction, aiming to extract insights applicable to enhance multimodal human-agent dialogue. By annotating various types of gaze behavior alongside speech, the thesis explores the meaning of temporal patterns in gaze cues and their correlations. On the basis of leveraging a multimodal corpus of dyadic taste-testing interactions, the thesis further investigates the relationship between laughter, pragmatic functions, and accompanying gaze patterns. The findings reveal that laughter serves different pragmatic functions in association with distinct gaze patterns, underscoring the importance of laughter and gaze in multimodal meaning construction and coordination, relevant for designing human-like conversational agents. The thesis also proposes a novel approach to estimate gaze using a neural network architecture, considering dynamic patterns of real-world gaze behavior in natural interaction. The framework aims to facilitate responsive and intuitive interaction by enabling robots/avatars to communicate with humans using natural multimodal dialogue. This framework performs unified gaze detection, gaze-object prediction, and object-landmark heatmap generation. Evaluation on annotated datasets demonstrates superior performance compared to previous methods, with promising implications for implementing a contextualized gaze-tracking behavior in robotic interaction. Finally, the thesis investigates the impact of different gaze patterns from a robot on Human-Robot Interaction (HRI). The results suggest that manipulating robot gaze based on human-human interaction patterns positively influences user perceptions, enhancing anthropomorphism and engagement.