What are they looking at? Automatic Simultaneous Dyadic Gaze Detection from Videos

Vidya Somashekarappa
University of Gothenburg, Gothenburg, Vastra Gotaland County, Sweden

Aram Karimi
university of gothenburg, Gothenburg, Västra Götaland County, Sweden

Christine Howes
University of Gothenburg, Gothenburg, Sweden

Asad Sayeed
University of Gothenburg, Göteborg, Sweden

Abstract

Gaze is cardinal for interaction since it allows us to visually interact with the environment and understand the attention and intent of others. Gaze also plays a critical role in HRI tasks such as object recognition and manipulation, as a robot can use gaze to direct its attention to specific objects or areas of interest in its environment. In this study, we automate gaze estimation for various types of gaze behaviours (such as turn taking, joint attention, gaze following, gaze aversion and mutual attention) in a natural dyadic interaction using videos without wearable cameras or eye trackers that are implementable on a robot. There is no single dataset that covers all of the different gaze and scene combinations that is address in this paper. We propose a model that utilizes the manual annotation of gaze targets in a natural dialogue setting and generate simultaneous gaze prediction for both parties in the video, along with attention heatmaps that provide exclusive information of the target object-of-interest in the scene, by also providing out-of-scene gaze predictions. Our model performs better than the baseline methods that currently exist and the data that was generated is available for different categories of gaze

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.*