

Basics of video-based motion tracking for linguists with a use case example with MediaPipe

Olcay Türk (Universität Bielefeld)

A careful study of multimodal communication demands the consideration of various modes of signalling including but not limited to the gestures of face and hand. The enormity and the complexity of the annotation/analysis of these can be intimidating. We as researchers still largely rely on our qualitative insight and eyes to deal with such essential tasks despite the availability of reliable quantitative tools. MediaPipe is one such tool that can assist us in capturing gesture kinematics. In this talk, I will discuss the prerequisites and the basics of video-based motion capture and look at a few examples with MediaPipe, making use of existing public toolkits (e.g., The Envision Toolbox) [1].

[1] Cwiek, A., De Melo, G., Edelman, J., Owoyele, B., Pouw, W., Santuber, J. Trujillo, J.. Envision Toolbox: Multimodal (Signal) Processing and Analysis in Communication.
<https://github.com/WimPouw/EnvisionBootcamp2021> (**Alphabetical order)