

Thesis abstract: The widespread adoption of electronic health records within clinical settings has renewed interest in understanding physician-patient interactions. Previous work analyzing clinical interactions has mostly coupled patient surveys with manually annotated video interactions provided by human coders. Physician gaze is among the components of the non-verbal interaction which has been found to impact patient outcomes. The work described in this thesis illustrates an automated system for multi-video labeling of patient-physician interactions and shows that image features (in the form of body positioning coordinates and optical flow) can provide important visual aids for learning physician gaze with over 90% accuracy. While our approach focuses on physician gaze, it can be extended to capture other clinical human-human and human-technology interactions as well as connect these interactions to patient ratings of clinical interactions.