Detecting Bids for Eye Contact Using a Wearable Camera

Zhefan Ye, Yin Li, Yun Liu, Chanel Bridges, Agata Rozga, James M. Rehg

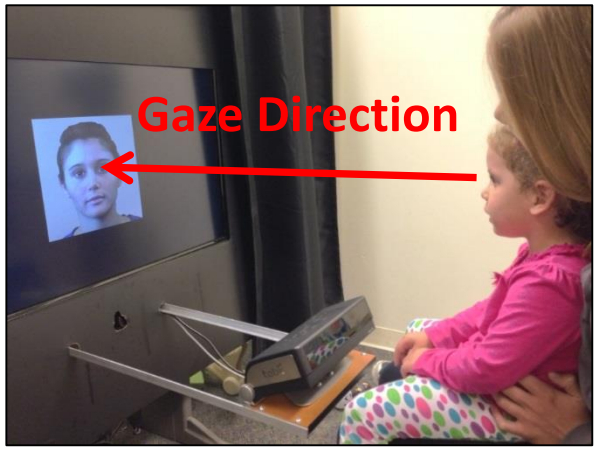
Center for Behavioral Imaging, School of Interactive Computing,
Georgia Institute of Technology

Georgia Institute of Technology

Limitations of Gaze Measurement Technologies

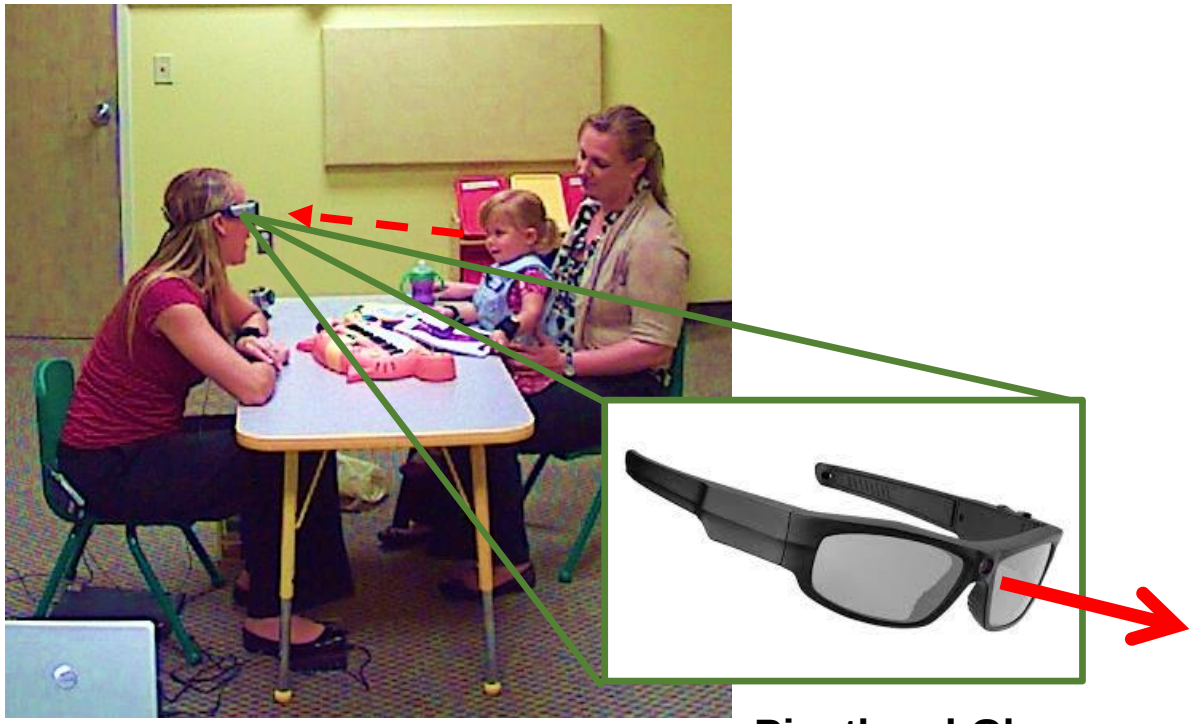
- ***Table-mounted eye tracker***
  - Need to map measured Gaze Direction to Point of Regard (PoR)
  - Good for monitors, bad for people

PoR



- ***Wearable eye tracker***
  - Challenges with calibration
  - Poor tolerance by children

First-Person View of Eye Contact



Pivothead Glasses

First-Person View of the Child

Multimodal Dyadic Behavior Dataset (MMDB)

- ***160 sessions***
  - Adult-child
  - Semi-constructed
  - 3-5 minutes
  - 15-30 months

- ***Multimodal***
  - Video
  - Audio
  - Physiological

cbi.gatech.edu/mmdb/

Contributions

- *Demonstrate feasibility of eye contact detection based on first-person video*
- *Collect the first dataset of first-person videos from adult-child social interactions*
- *Evaluate accuracy of first-person approach at both frame and event levels*

Agenda

Frame-level Detection

Event-level Detection

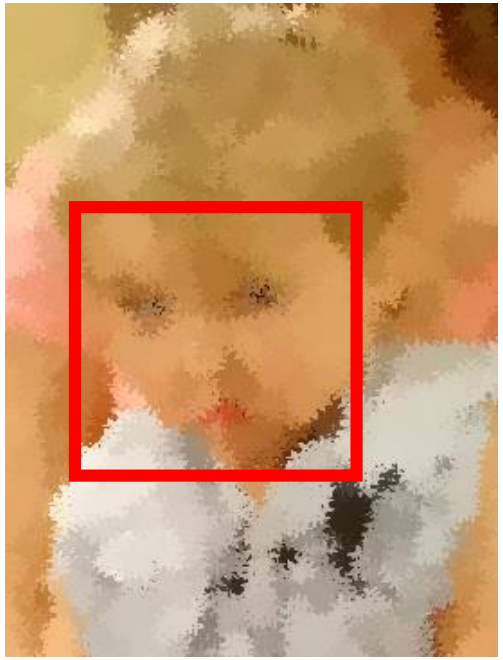
Experiments and Results

Conclusion

Detect and Localize Face

- ***Face Detection***
  - OMRON OKAO Vision

- ***Facial Landmarks & Head Pose***
  - IntraFace (De la Torre et al., CMU)



OMRON OKAO Vision

IntraFace

Ambiguity in Gaze Estimation

Pose-dependent **E**gocentric **E**ye **C**ontact Approach (**PEEC**)

Head Pose and Eye Regions

Head Pose

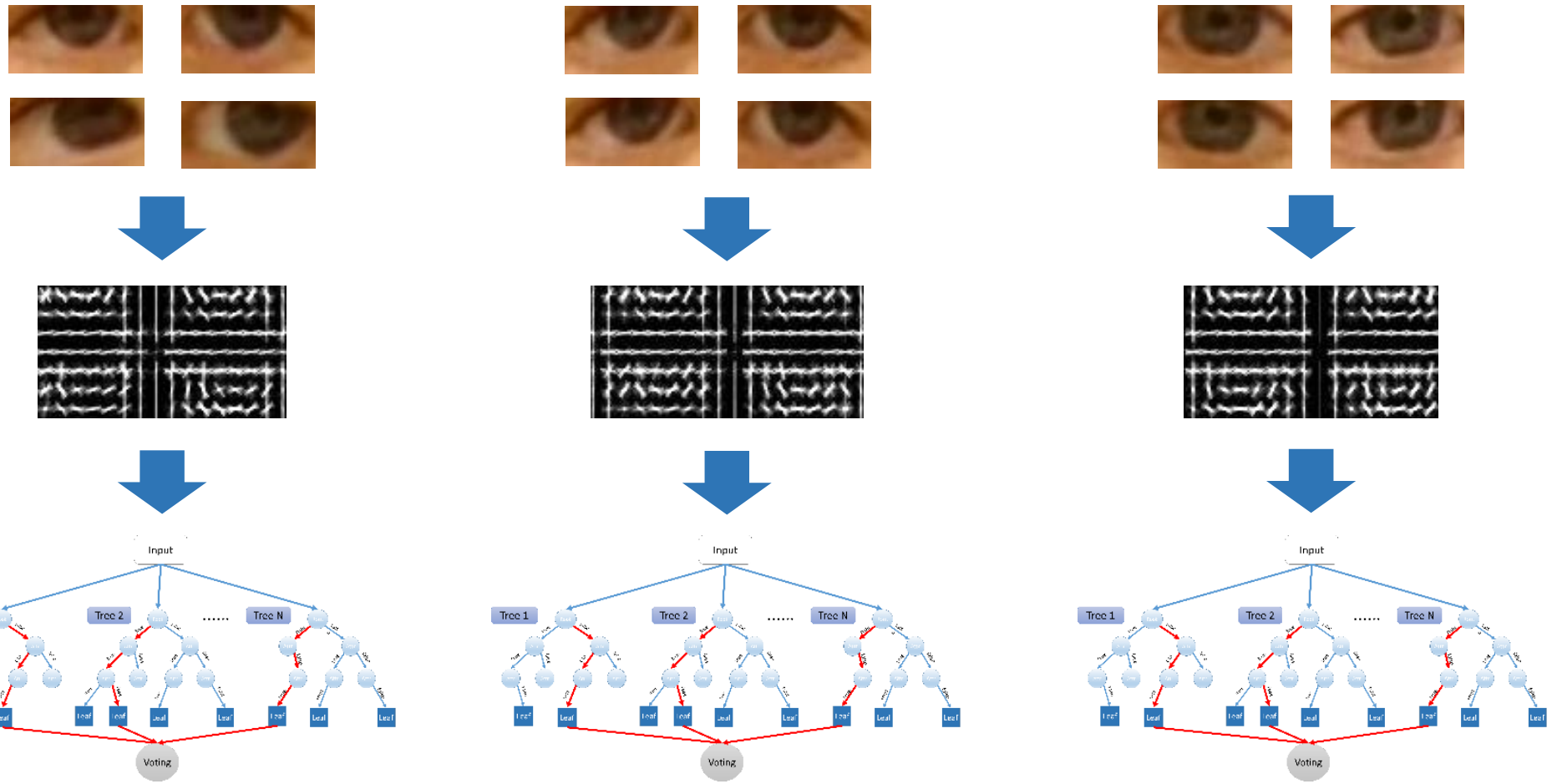
Eye Regions

Pose-Dependent Classifier

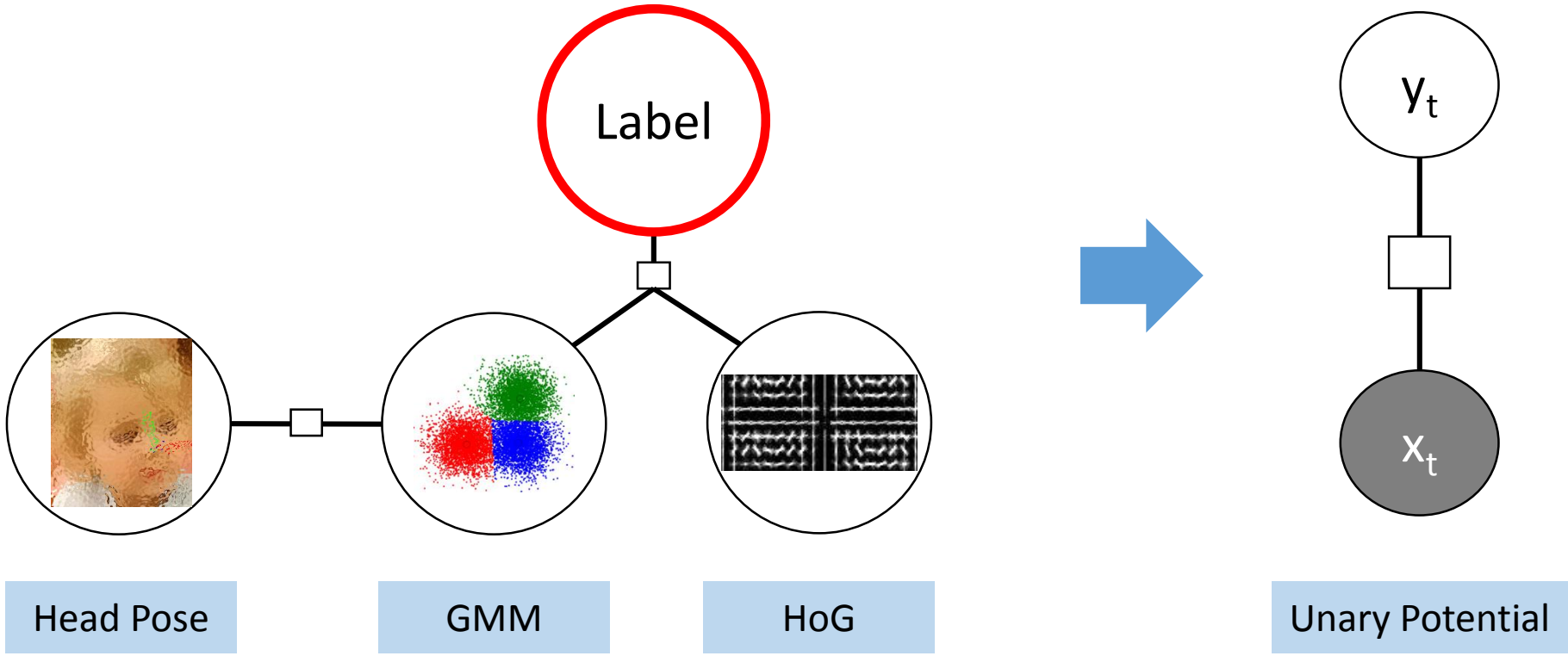
Eye Regions

HoG

Random Forest



Frame-level Model



Agenda

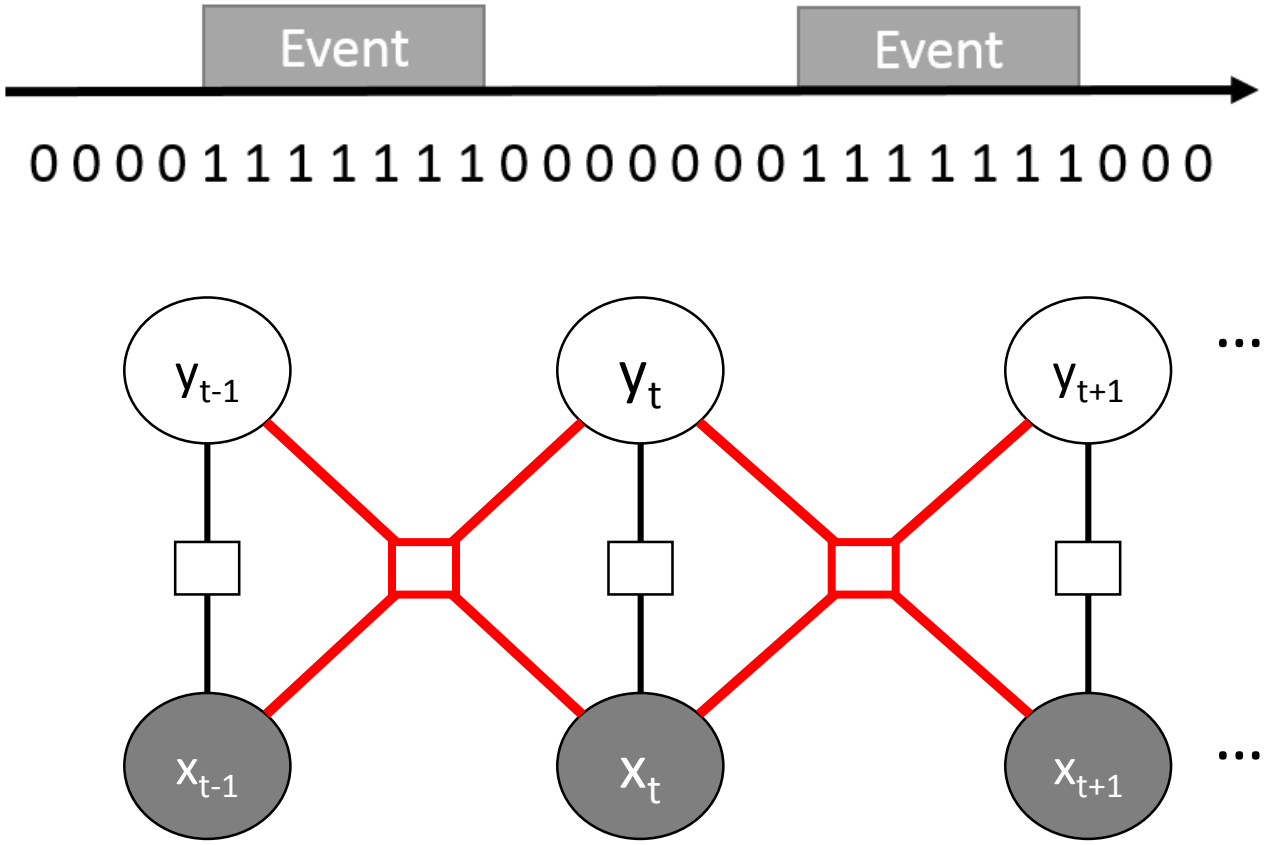
Frame-level Detection

Event-level Detection

Experiments and Results

Conclusion

Linear-Chain CRF



Agenda

Frame-level Detection

Event-level Detection

Experiments and Results

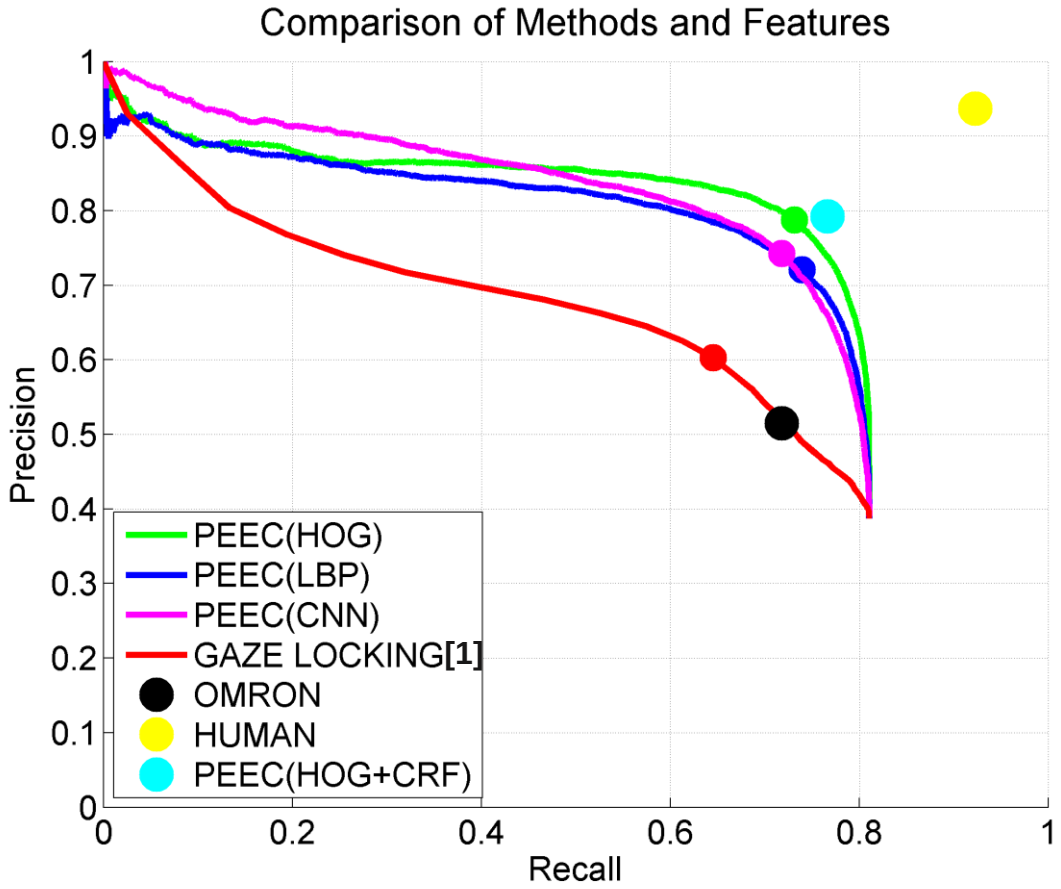
Conclusion

Ground Truth and Experiments

- *12 annotated sessions*
- *5 annotators*
  - Majority vote
- *Leave-one-subject-out cross-validation*

Video Result

Frame-level Result



Error Case 1

Face Occlusion

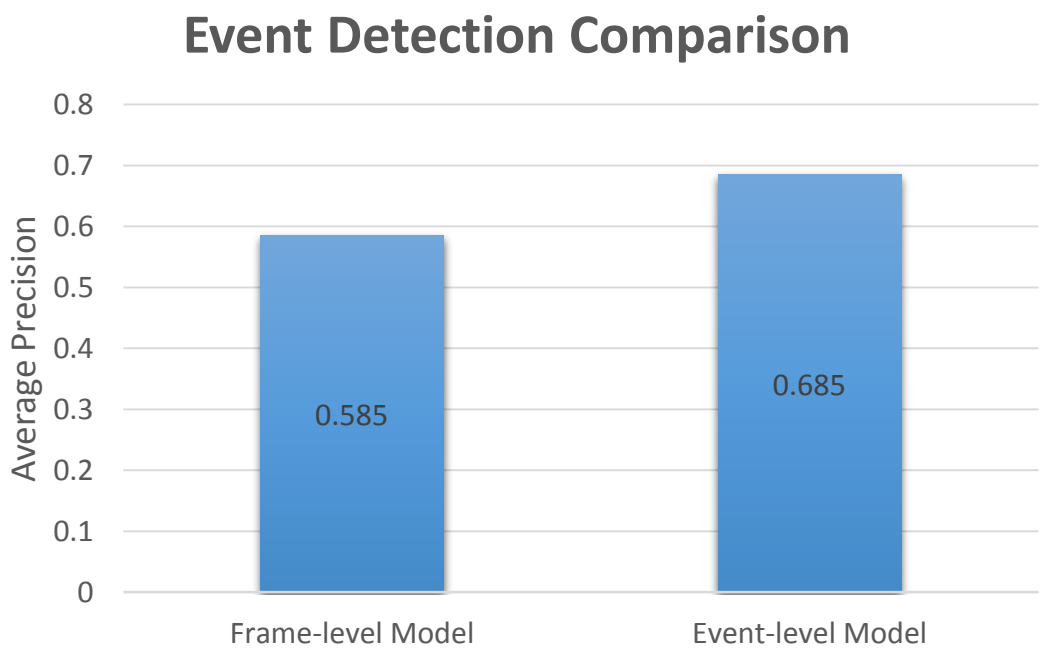
Error Case 2

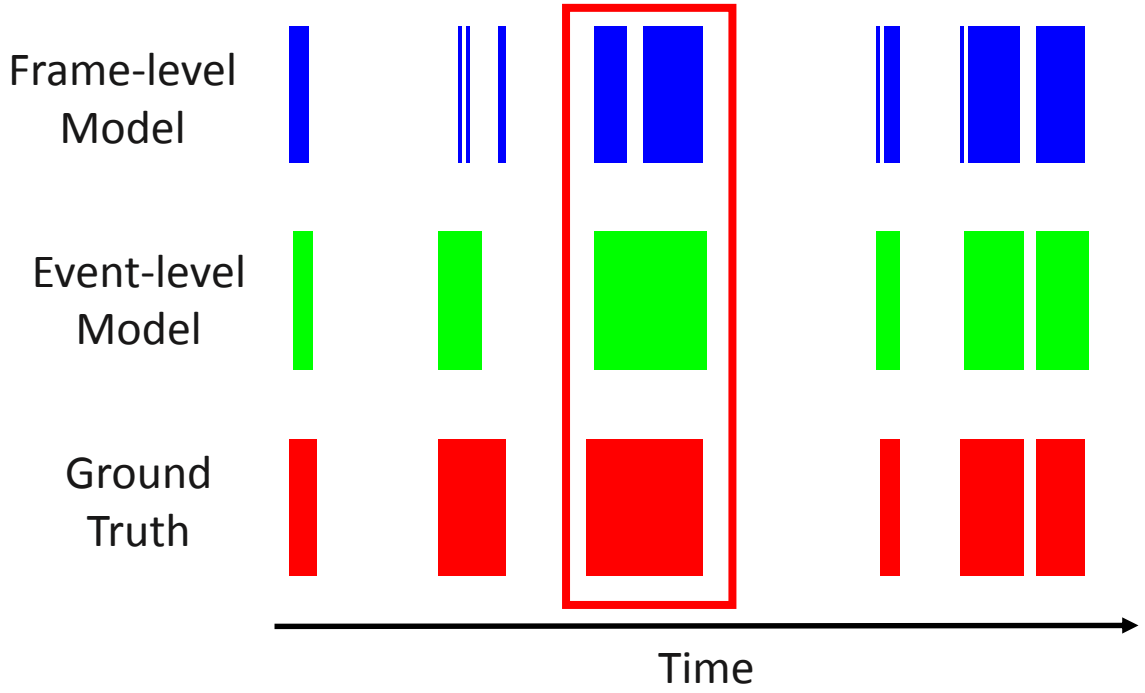
Face Rotation

[1] B. Smith, Q. Yin, S. Feiner, and S. Nayar. Gaze Locking: Passive Eye Contact Detection for Human Object Interaction. In *ACM Symposium on UIST*, pages 271–280, Oct 2013.

Event-level Result

- *Average Precision*





Conclusion

- *Non-invasive detection of eye contact during face-to-face social interactions can be achieved through first-person vision*
- *Developed a prototype system for automatically-measuring the frequency and duration of eye contact events*
- *Temporal smoothing can improve the accuracy of event level detection*
- *Dataset of first-person social interactions available to the research community via MMDB*

Q&A

Thank you for your attention