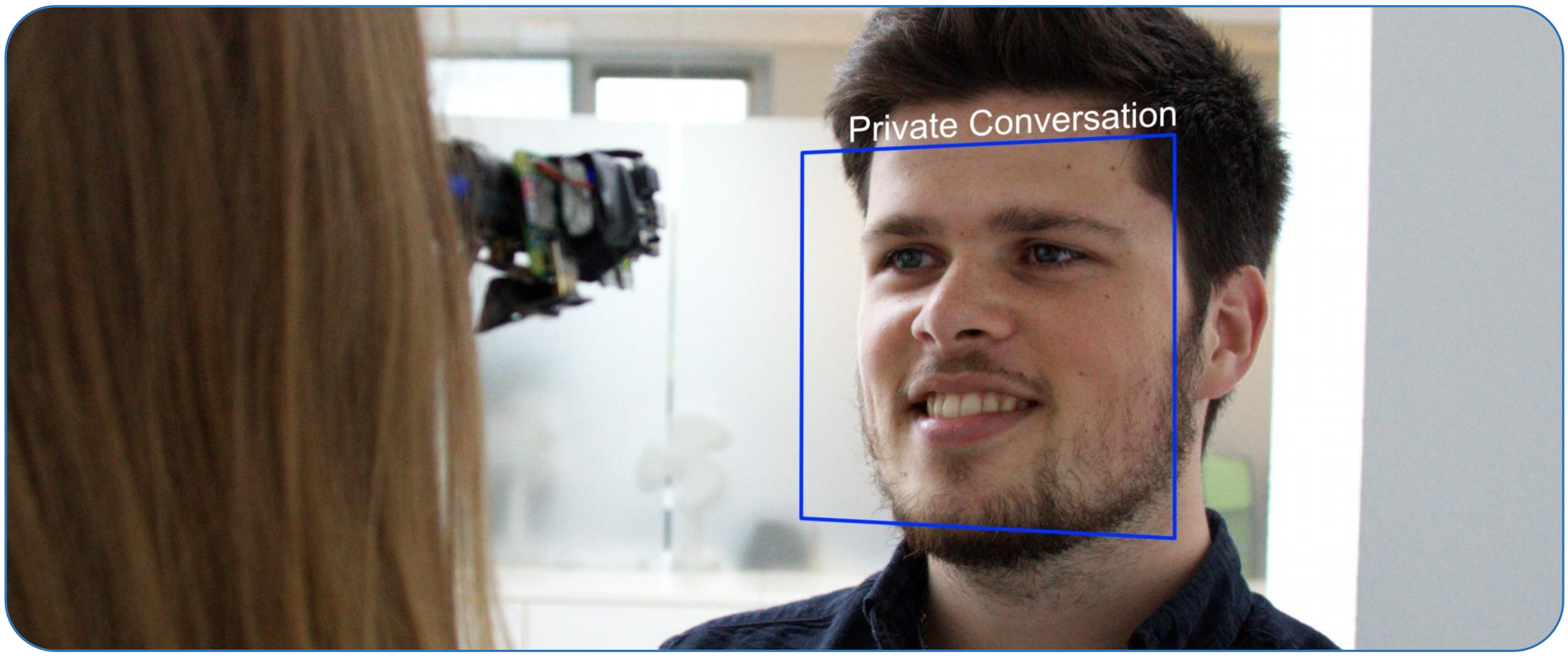
Towards Respectful Smart Glasses through Conversation Detection

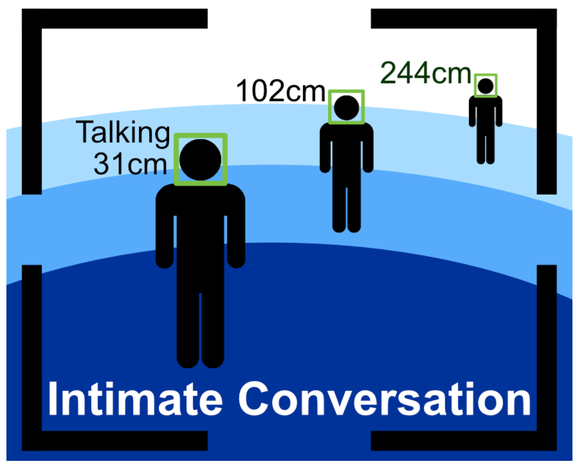
Franziska Meirose[1], Sven Schultze[1], Sebastian Kuehlewind[1],
Marion Koelle[1], Larbi Abdenebaoui[2], Susanne Boll[1]



Motivation

- Body-worn camera devices recently hit the consumer market
  - Threat to the users' as well as their bystanders' privacy
- Interpersonal conversations should be protected
- Prior work relied on audio data → privacy issues:
  - Content of conversation has to be captured
- Conversation detection relying only on image data should be considered

Scenario



- Based on human comfort zones
- Proximity of conversation partner is inferred from the camera's imagery
- Classification as intimate, personal, social or public
- Not restricted to the exemplary travel scenario

On their Interrail trip through Italy, Lara and Paul are using smart glasses to automatically capture short videos and still images of memorable moments. At the moment, they are on a guided tour in the city center of Siena.
When Lara is wearing the device while listening to the tour guide, the situation is classified as *public conversation* (distance > 300cm) and the experience may be captured as short clip. Ordering an ice cream or buying something from a market vendor would be classified as *social conversation* (between 300cm and 120cm distance), as a consequence, audio might be muted, and only still images would be captured. During lunch at a restaurant, a conversation with another member of her travel group or Paul would be classified as *private*, or *intimate*, depending on the proxemics. As a reaction, the system would stop the recording, and/or remind to Lara to take off the glasses to protect their privacy and to comply with social norms.

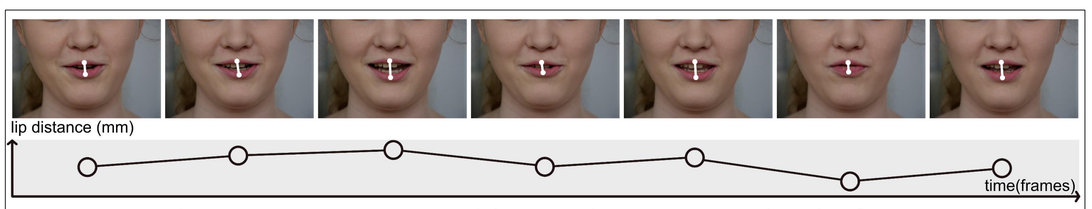
Prototype

Detection and Tracking

- Face Detection via Viola-Jones implementation of OpenCV every 30 Frames
- Median-Flow algorithm in between to track the faces

Analysis

- Calculate distance between camera and face
- Obtain facial landmarks to determine shape of mouth
- Calculate distance between the upper and the lower lip
  - Observe this value over ten frames
- Compute differences of the values of subsequent frames
  - Determine the average of the differences
  - Multiply with the standard deviation of all ten values
- Apply a threshold to decide whether the person is talking
- Combine analysis of all detected faces for conversation status
- Determine conversation zone with distance to the nearest face



Distances between upper and lower lip over multiple frames

Future Work

- Contextual privacy protection:
  Blur imagery, mute audio or cover the camera lens
- Subsequent long-term user study to provide insights on conversation detection and protection in-the-wild

Contact

University of Oldenburg[1]/OFFIS[2]
Department of Computing Science
Media Informatics and Multimedia Systems
Escherweg 2, 26122 Oldenburg
Phone: +49 (0) 441 – 97 22 213
Email: hci@uni-oldenburg.de