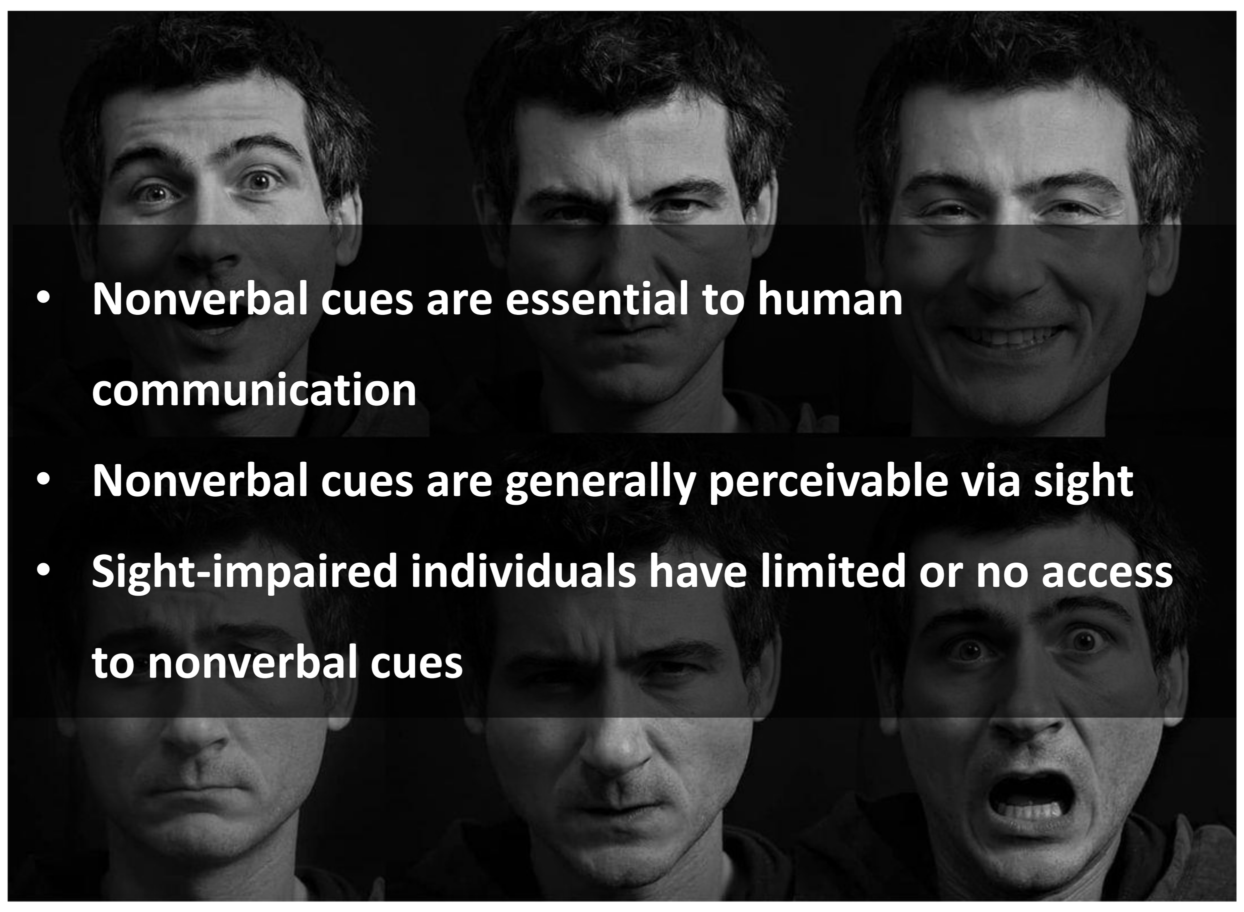
Augmented Reality and Affective Computing for Nonverbal Interaction Support of the Visually Impaired

Deniz Iren, Krist Shingjergji, Felix Bottger,
Corrie Urlings, Roland Klemke
deniz.iren@ou.nl

Open Universiteit

1. Introduction



- Nonverbal cues are essential to human communication
- Nonverbal cues are generally perceivable via sight
- Sight-impaired individuals have limited or no access to nonverbal cues

2. Background

Affective computing: aims at understanding and developing the technology for *detecting*, *interpreting*, *responding to* human affect

Affect

- moods and emotions
- observable through physiological signals (e.g., tone of voice, facial expressions, gestures)

Augmented Reality(AR)

- Most studies focus on sight
- AR also covers other sensory augmentation

Types of affective computing systems

- **SER**: Speech Emotion Recognition
  - Recurrent Neural Networks (RNN)
- **FER**: Face Expression Recognition
  - Convolutional Neural Networks (CNN)
- **GR**: Gesture Recognition
  - Markov Models / Finite State Machines

Important characteristics

- Performance, complexity, size

Related Work on Assistive AR

- Auditory and haptic substituting sight
- Navigation, obstacle avoidance, object detection [1]
- Enhancing sight; only for partially impaired

3. Proposed Solution

SOLUTION PROPOSITION

- Wearable technologies
- Augmented Reality
- Affective Computing

TECHNICAL CHALLENGES

- Processing images from a moving camera
- Battery / Compute Limitations

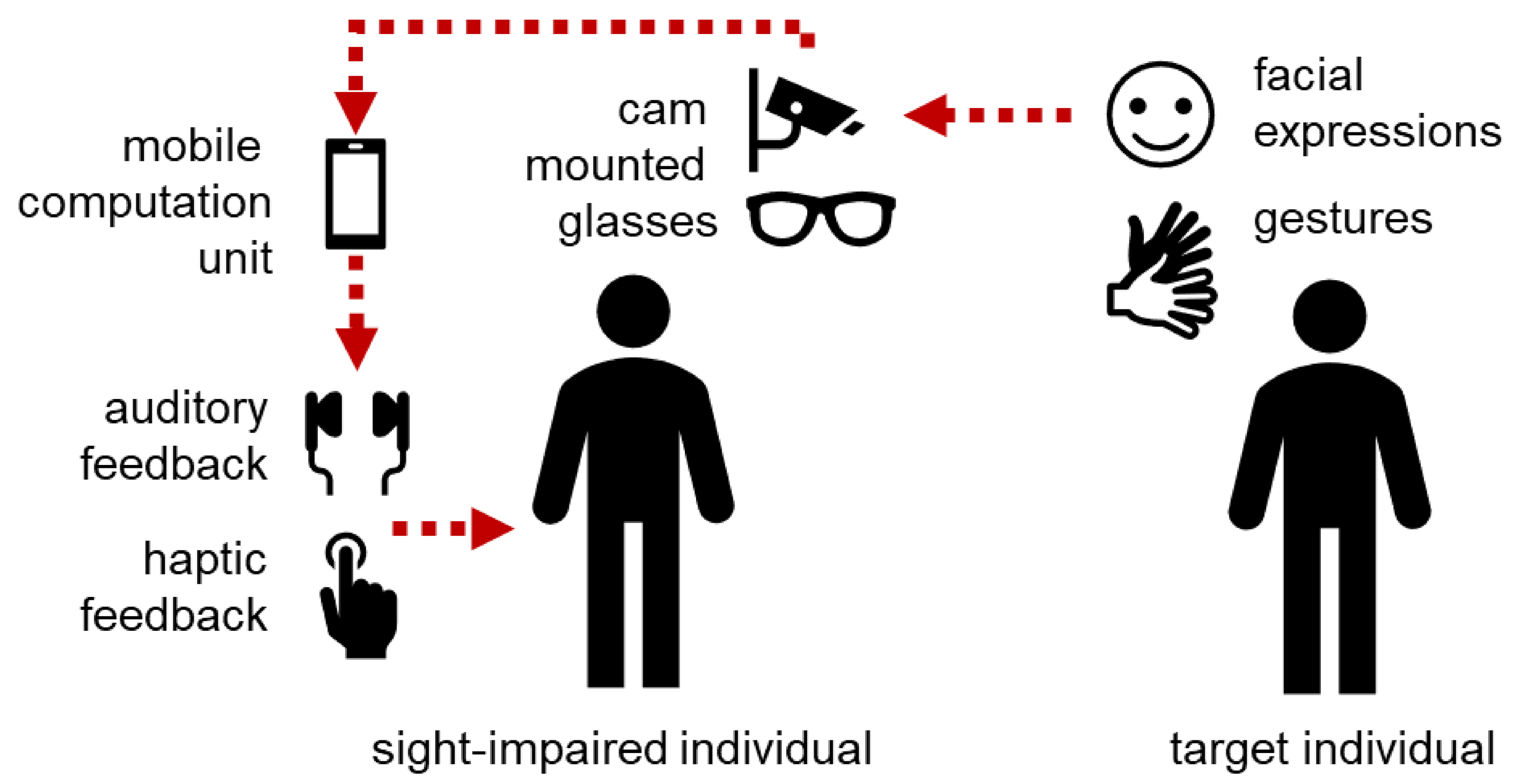
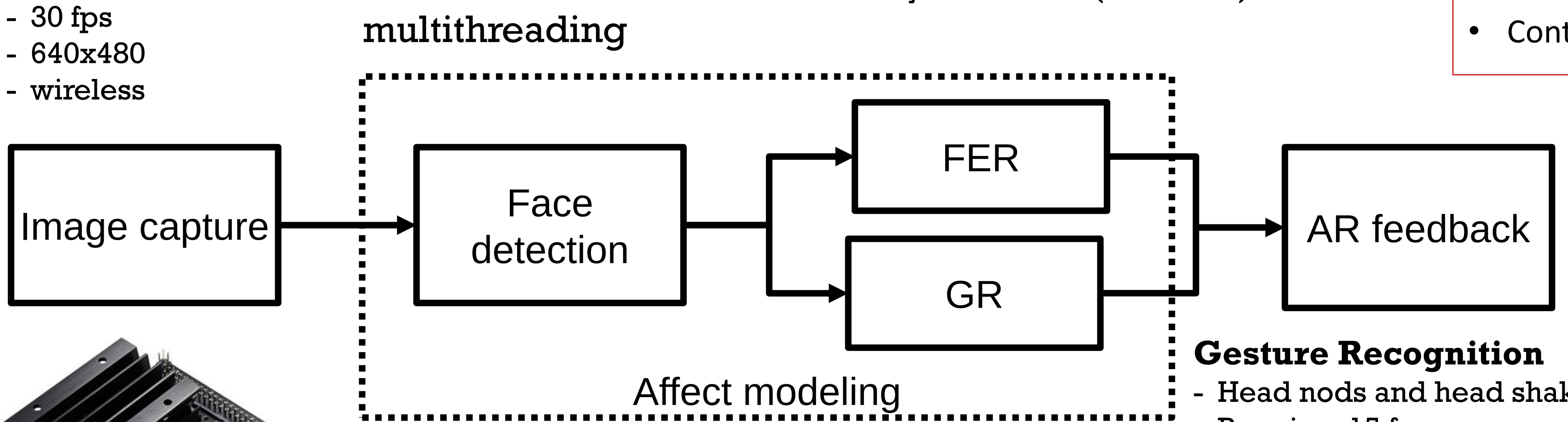


Figure 1. The conceptual schematics of the proposed solution

Aimcam Pro 2i
- 30 fps
- 640x480
- wireless

Facial Landmark Extraction
- Dlib
- (68x2) vectors

Facial Expression Recognition
- Based on AU detection
- Smile (AU6+AU12),
- Frown (AU4),
- Eyebrow raise (AU1+AU2)

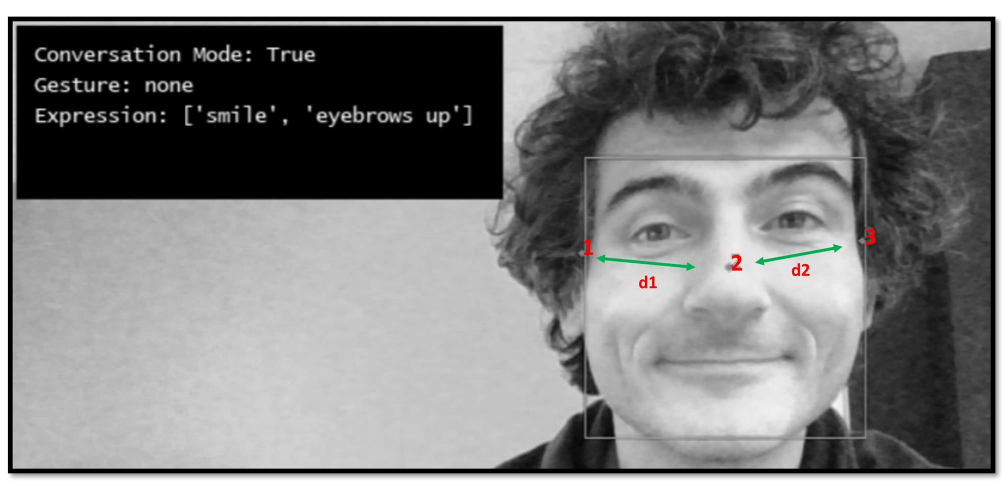


FER Model
- CNN
- 4 convolutional layers
- 32, 32, 64, 64 filters
- ReLU activation
- Trained on CK+ and DISFA
- Average testing accuracy F-1 = 77.12
- Model size: 13.5MB

Gesture Recognition
- Head nods and head shakes
- Requires 15 fps

NVIDIA Jetson Nano

Conversation Mode
- Activates either manually or when a face stays in the focus for a while
- To avoid unnecessary computation



AR: Haptic Feedback
- Custom built
- 24xTectonic vibration motors

Figure 2. The components of the prototype

4. Discussion

MOVING CAMERA

- Processing images from a moving camera
  - For FER; less problematic
  - For GR; requires special solution
  - Our design is robust against camera movements

RESOURCE LIMITATIONS

- Battery and compute Limits
- Avoid unnecessary computation
- Optimize performance/complexity of models

FEEDBACK TO THE USER

- What to convey?
- Avoiding information overload (e.g., aggregation, smoothing)
- How to convey?

WHY NOT EMOTIONS?

- Context dependent, cutural, and individual variances

Authors

Deniz IREN
Associate Professor
Department of Information Science,
Open Universiteit
deniz.iren@ou.nl

Krist Shingjergji
PhD Candidate
Department of Technology Enhanced Learning and Innovation,
Open Universiteit

Felix Böttger
Research Assistant

Corrie Urlings
Assistant Professor
Department of Technology Enhanced Learning and Innovation,
Open Universiteit

Roland Klemke
Professor
Department of Technology Enhanced Learning and Innovation,
Open Universiteit

Team HART
Special thanks to the students of Team HART

References

[1] M. Zahn and A. A. Khan. Obstacle avoidance for blind people using a 3d camera and a haptic feedback sleeve. arXiv preprint arXiv:2201.04453, 2022.

[2] S. Bromuri, A. P. Henkel, D. Iren, and V. Urovi. Using ai to predict service agent stress from emotion patterns in service interactions. Journal of Service Management, 2020.

[3] S. Latif, R. Rana, S. Khalifa, R. Jurdak, J. Qadir, and B. W. Schuller. Survey of deep representation learning for speech emotion recognition. IEEE Transactions on Affective Computing, 2021.

[4] K. Shingjergji, D. Iren, F. Bottger, C. Urlings, and R. Klemke. Inter- ¨ pretable explainability in facial emotion recognition and gamification for data collection. In 2022 10th International Conference on Affective Computing and Intelligent Interaction (ACII), pp. 1–8. IEEE, 2022.