Development of a Wearable Vision Substitution Prototype for Blind and Visually Impaired that Assists in Everyday Conversations

Anna Kushnir

info@anna-Kushnir.de

Nicholas H. Müller

nicholas.mueller@fhws.de

Socio-Informatics and Societal Aspects of Digitalization

Würzburg, Germany

Presenter introduction

Anna Kushnir

- Master’s degree in Information Systems at the University of Applied Sciences Würzburg-Schweinfurt (FHWS).
- Research Assistant at FHWS in the field of Socio-Informatics.
- Current Work:
  Further development of the presented vision substitution prototype

Contents

Introduction

Introduction

- Everyday face-to-face communication situations use a variety of communication channels.
- In addition to verbalized information, several non-verbal cues are communicated, which have to be interpreted by the communication partners.
- These include for instance:
  - Facial expressions
  - Intonations
  - Gestures

Introduction

Sighted people

- Perceive the face-to-face communication through the senses of seeing and hearing.
- Are usually able to interpret the facial expressions of their communication partner, which allows them to recognize emotions and the emotional valence of their communication partner.

Blind and visually impaired people

- Perceive the face-to-face communication through the sense of hearing.
- Can recognize emotions and the emotional valence only through intonation, as the interlocutor's facial expressions cannot be seen.
- Are not able to recognize the emotions if their interlocutor is listening and therefore only non-verbal cues are transmitted.

➢ can **only partially** perceive and analyze the communication situation.

Introduction

- A lot of research has been done to meet a wide range of needs of people with visual disabilities.
- Unfortunately not enough attention has been given to the development of Sensory Substitution Devices (SSDs) that satisfy the need for access to non-verbal communication.

Introduction

Shortcoming of the most existing SSDs:

- The most SSDs provide only aural feedback, do not support live analysis or do not provide functionalities that support face-to-face communications (Orcam, n. s.; Microsoft, n. s. and McDaniel et al. 2008).

- **Problem of aural feedback:**
  - People with a visual impairment rely on their hearing to perceive their environment.
  - Aural feedback of a SSD may interfere with the hearing of one's own speech or the one of the communication partner.

➢ What is needed is a system that communicates non-verbal cues in **real-time** whose output is based on a **different sense than the hearing** on which verbal communication is perceived.

Objective

Objective

- **Development of a SSD prototype for blind and visually impaired that assists in face-to-face communication situations.**
  - Recognition of the seven basic emotions according to Ekman by utilizing the Facial Action Coding System (FACS).
  - Translating the recognized basic emotions into tactile stimuli, that are conveyed in real time to the blind or visually impaired user.

Fundamentals

The basis of this research work are the basic emotions according to Paul Ekman.

- For the so-called basic emotions **universal facial expressions** were empirically proven. People can read and interpret these universal facial expressions regardless of their culture (Ekman, 1999; Ekman et al., 2010)
- A basic emotion describes a emotion family with different forms and intensity levels. The basic emotions include the emotion groups of **anger**, **happiness**, **sadness**, **disgust**, **contempt**, **fear** and **surprise** (Ekman et al., 2010; Ekman & Cordaro, 2011).
- Unlike moods, emotions **last only a few seconds to minutes** (Ekman et al., 2010).

The basis of this research work are the basic emotions according to Paul Ekman.

- Other non-verbal cues that can be observed during a conversation can also indicate emotions (Ekman et al., 2010; Schmidt-Atzert, 2014).

- What is special about facial expressions, however, is that they **appear** in a conversation situation **even when a person is silent**, for example while listening (Ekman et al., 2010).

Fundamentals

The basis for the analysis of facial expressions and the emotion recognition in this research work is the Facial Action Coding System (FACS).

- FACS is a anatomically based coding system for all visually perceptible facial muscle movements and a common standard recognizing basic emotions and their intensity (Ekman et al., 2002).
- This is an extensive collection of photographs and texts on individual facial muscles, their interaction and the resulting facial expressions (Ekman et al., 2002).
- In order to determine the emotions of a facial expression using FACS, an expression has to be deconstructed in Action Units, on which the FACS is based (Ekman, 1982; Ekman et al., 2002).
- Action Units are assigned to almost every visible movement of facial muscles and can combine single or multiple muscle movements (Ekman et al., 2002).

SSD prototype construction and design

SSD prototype construction and design

In the conference paper the work in progress of developing the SSD prototype was described.

As a first step, a SSD prototype was developed that **conveys the emotional valence** of the user's interlocutor.

This prototype's architecture was formed by four main components:

1) Camera Unit
2) Notebook / Noldus FaceReader Software
3) Wearable microcontroller
4) Wearable haptic device

SSD prototype construction and design

1) Camera Unit

- The camera unit is the **sensor** of the prototype. It's main function is to **record the interlocutor's face** during the communication.
- By recording the facial expressions, the camera unit **creates an analysis basis** for the prototype.
- The **Logitech Brio 4K** webcam was used for the camera unit of the prototype.
- Although the entire prototype could not be used as a mobile device because of the webcam, this construction was still sufficient for initial studies.

SSD prototype construction and design

2) Notebook / Noldus FaceReader Software

- The notebook and the FaceReader Software form the **coupling system** of the SSD Prototype.
- The notebook is used to run the FaceReader Software (Noldus, n. s.).
- The software **analyzes and categorizes** the facial expressions from the live video stream of the sensor and is, therefore, the main part of the coupling system.
- It utilizes the FACS and is able to **recognize basic emotions** and their intensity according to Ekman:

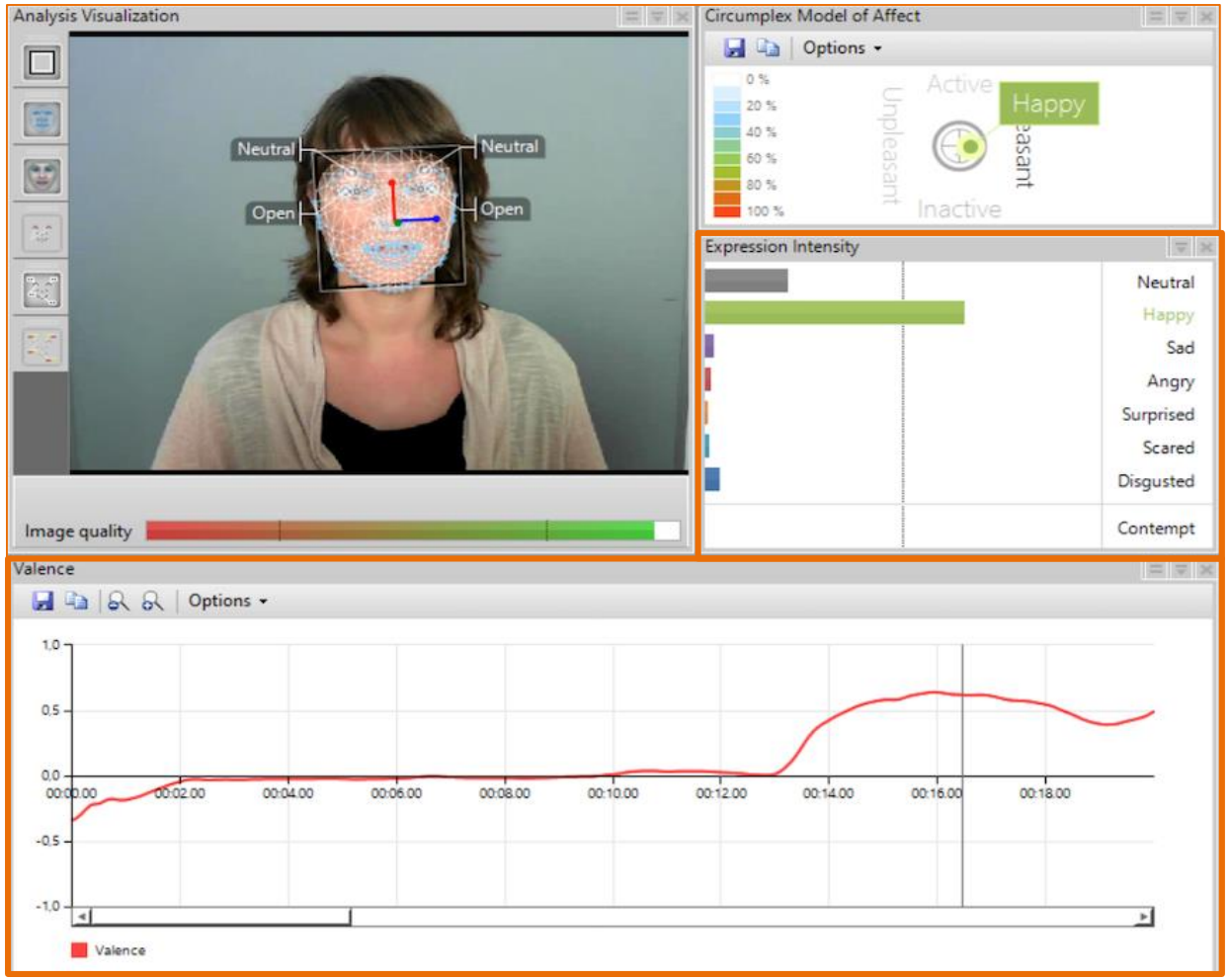
  - **Happiness**
  - **Anger**
  - **Sadness**
  - **Disgust**
  - **Contempt**
  - **Fear**
  - **Surprise**

2) Notebook / Noldus FaceReader Software

- The **emotional valence** is automatically calculated by the Software during analysis.
- It results from the difference between the intensity of the positive emotion and the intensity of the most pronounced negative emotion.
- **Happiness** is the only positive emotion, while **sadness**, **anger**, **fear** and **disgust** are considered to be negative emotions in the calculation.
- A special case is the emotion **surprise**, which can be either positive or negative. Therefore it is not used for the calculation of the emotional valence.

SSD prototype construction and design

2) Notebook / Noldus FaceReader Software



Basic Emotions

Emotional Valence

SSD prototype construction and design

3) Wearable Microcontroller

- After the analysis in the FaceReader Software whether the valence of the interlocutor is positive or negative, the results are sent to this unit.

- The function of this component is to **control the haptic device** depending on the sent emotional valence. Both together form the **stimulator** of the SSD prototype.

- **Component design:**
  - The main unit of the microcontroller is an Arduino Nano V3.3 board.
  - The connection to the notebook is made possible using the Bluetooth module HC-05-6.
  - For the power supply of the microcontroller a power bank is used.

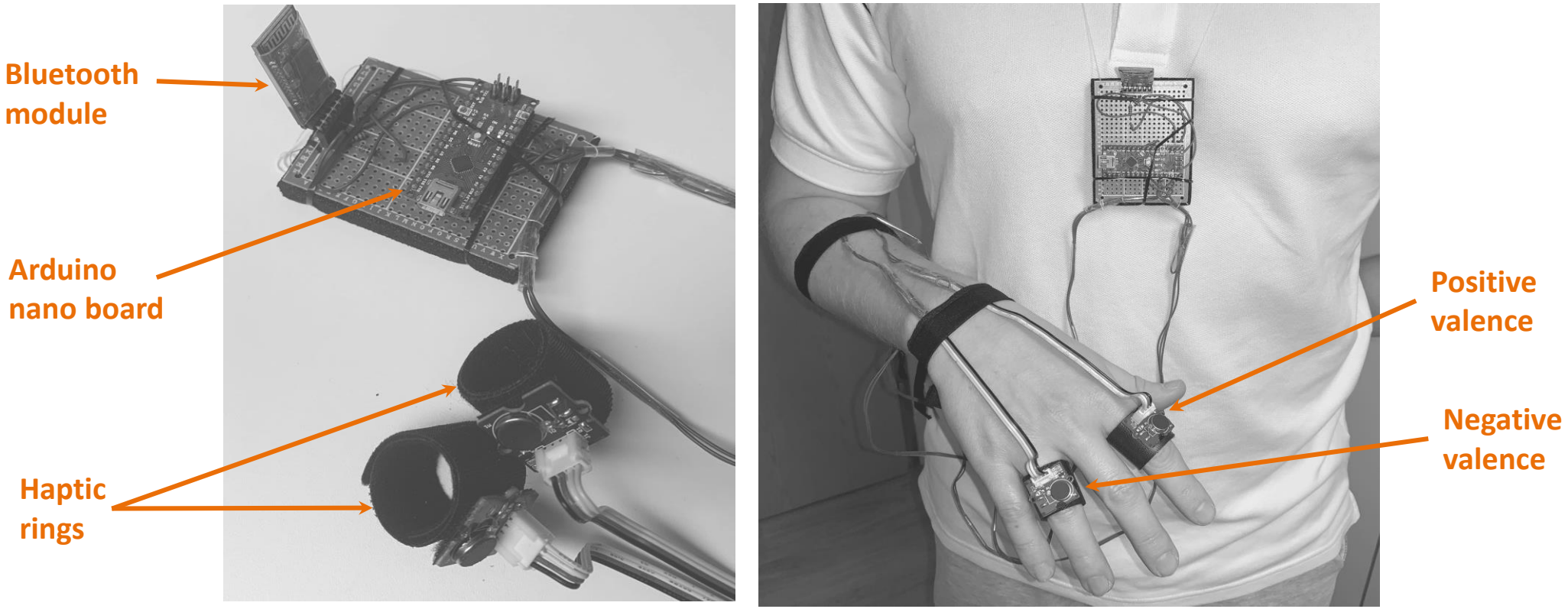
SSD prototype construction and design

4) Wearable Haptic Device

- The function of the haptic device is to **communicate the emotional valence** to the user via vibration signals.
- The wearable haptic device consists of two vibration rings, one of which is used for the positive and one for the negative valence.
- **Component design:**
  - Groove vibration motors were used for the haptic device, which are attached to resizable rings and can be worn carried on the non-dominant hand.
  - The vibration motors are connected by cables to the controlling unit, which is worn around the neck.
- In this way, a **mute and unobtrusive** signal transmission during a conversation can be ensured.

SSD prototype construction and design

Wearable microcontroller and haptic device



Further research

Further research

The described prototype has been further developed in the past few months.

- As announced in the outlook of the paper, the prototype has been extended by the seven basic emotions.

- However, before the prototype was extended, a study was conducted in which different tactile temperature stimuli were compared and **a suitable feedback method** for conveying emotions was selected.

Further research

In a quantitative study, tactile stimuli vibration, heat and cold were compared with one another (Kushnir & Müller, 2020).

- **Objective:**
  - To investigate which stimuli are best suited for conveying emotions during a conversation.
  - The focus was on tactile stimuli, as these do not restrict the communication interfaces speaking and hearing.

- **Result:**
  - Selection of the **vibration feedback** for the prototype's feedback as it was recognized more often and faster than the temperature feedback, which is important for conveying emotions that often last only a few seconds (Ekman et al. 2010).
  - In addition the vibration feedback was rated the least distracting by the majority of the study participants.

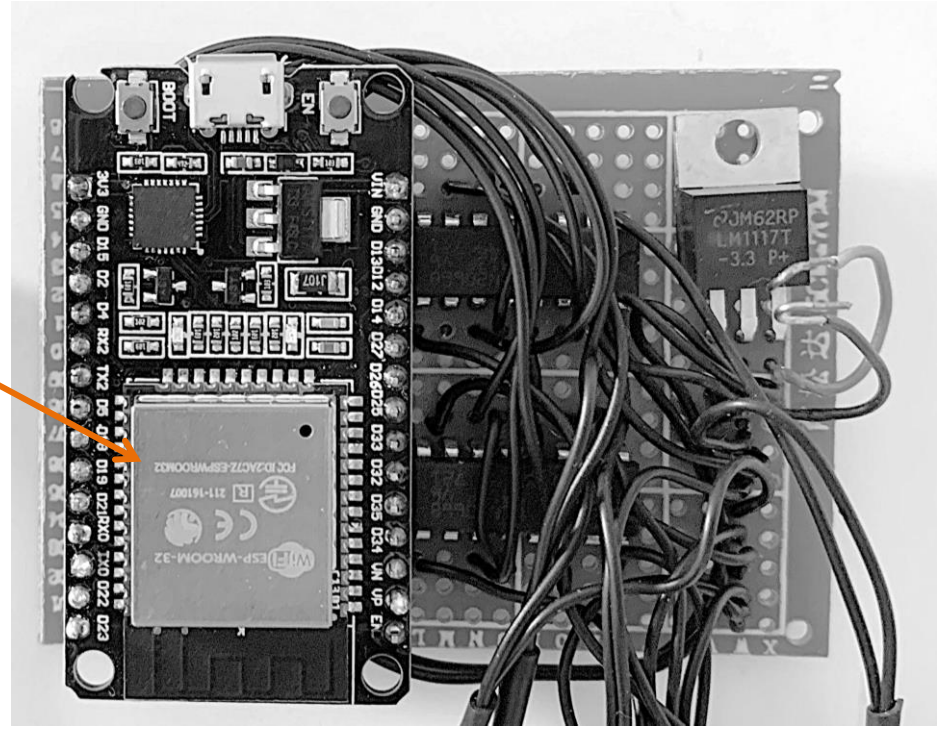
Further developments

Further developments

- **Adjustments of the Camera Unit**
  - The prototype described used a webcam to record facial expressions, whereby the device was not mobile.
  - For this reason, the webcam has been replaced by a **HD Mini Camera** (Spyschool, n. s.), which is attached to the temple of eyeglasses.
  - In this way, the conversation partner's face is **always in view** of the camera during a face-to-face communication.
  - In addition the HD Mini Camara uses a Smartphone for power supply, which also sends the live stream to the analysis software using the CameraFi App.

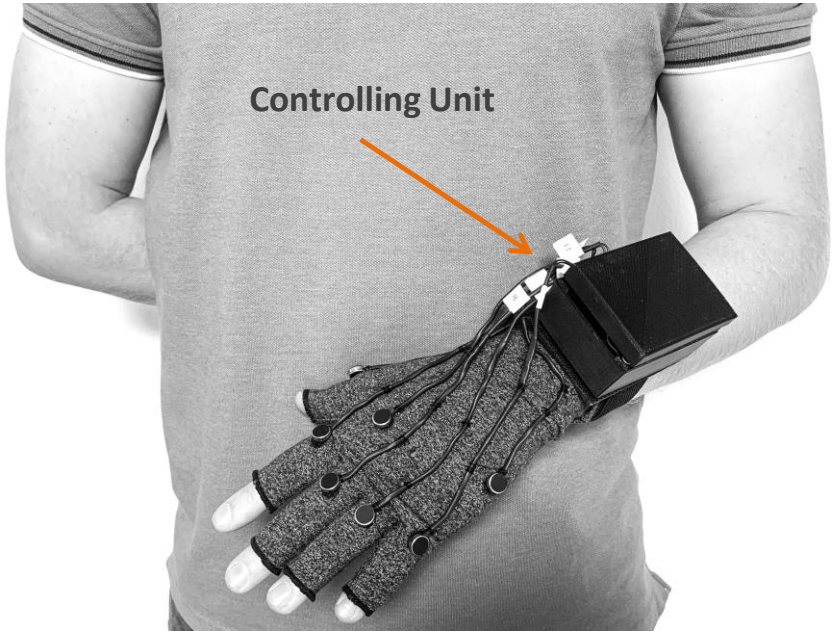
Further developments

- **Adjustments of the Wearable Microcontroller (1)**
  - Instead of the Arduino Nano board the successor prototype uses the **ESP32 WROOM-32U.**
  - In contrast to most Arduino controllers, the ESP32 has a Bluetooth module and also offers WiFi connection to the chip.



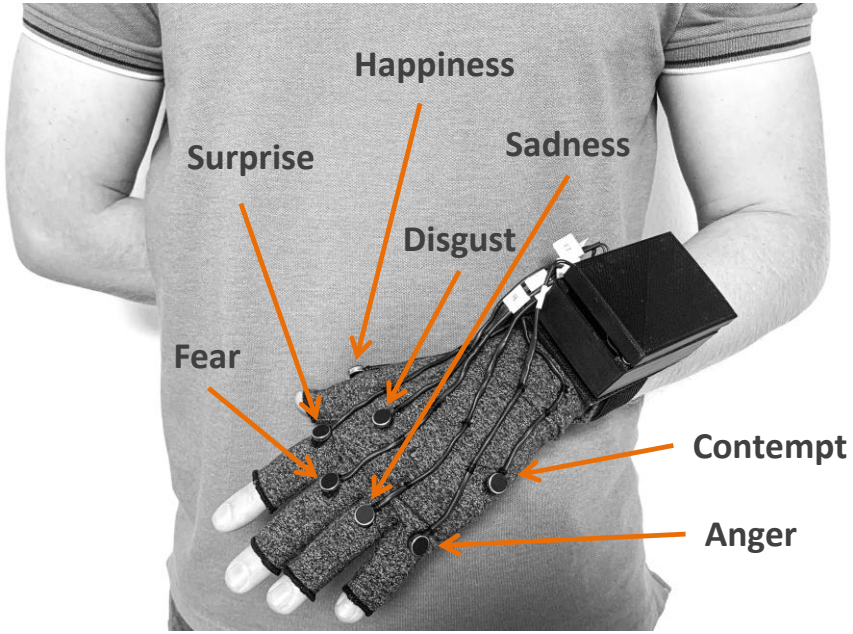
Further developments

- **Adjustments of the Wearable Microcontroller (2)**
  - Further the power supply by the power bank has been replaced by the light and flat **BL-5C Nokia battery**.
  - Thanks to the adjustments, it was possible to make the controlling unit light and compact.
  - The controlling unit including the battery is therefore no longer worn on the neck, but on the wrist.
  - The length of the cables that connect the controlling chip with the haptic interface is thus reduced to a minimum and ensures more freedom movement.



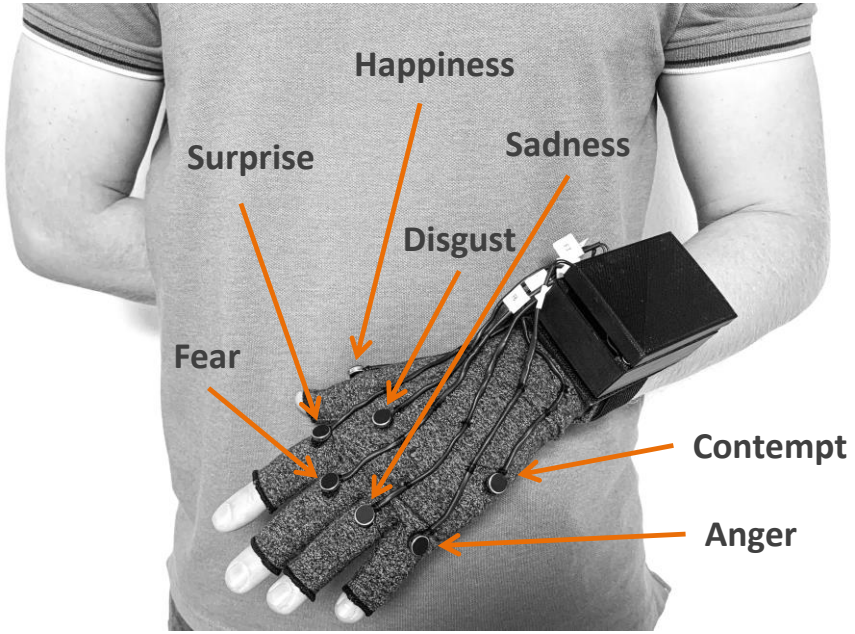
Further developments

- **Adjustments of the Haptic Interface (1)**
  - The feedback of the haptic device is no longer limited to the emotional valence, but reports the **seven basic emotions** according to Paul Ekman.
  - In this way the user is has access to a more detailed emotional state of the interlocutor.
  - Further the vibration rings have been replaced by a **haptic glove** and **smaller vibration elements**.



Further developments

- **Adjustments of the Haptic Interface (2)**
  - The haptic glove is based on a set of vibration motors attached to an elastic glove.
  - Each of the seven vibration motors, which are attached to the fingers the back of the hand, signal a basic emotion.
  - The eighth vibration motor is attached to the palm and vibrates when the software cannot recognize a face.



Conclusion

Conclusion

- A tactile vision substitution prototype that conveys the interlocutor‘s emotions during a communication situation has been designed.
- The emotion recognition is based on FACS and Ekman‘s basic emotions. This task is taken over by the Noldus FaceReader Software.
- The emotion identified by FACS is passed on as a vibration signal via the developed haptic interface.

Outlook

Outlook

- In order to provide the user with even more nuances of non-verbal communication, the prototype is to be expanded to include feedback on the intensity of emotions. This would make it possible to differentiate within an emotion family.

- Furthermore it is important to note that there is still room for improvement in terms of mobility.
  - For better handiness, the size of the Controlling Unit should be reduced.
  - The FACS analysis could be shifted to the Controlling Unit or the smartphone. In this way, the notebook would no longer be needed.

References

- Ekman, P. (1982). *Emotion in the Human Face* (2nd. edition). Cambridge University Press.
- Ekman, P. (1999). Basic emotions. In *Handbook of cognition and emotion*, pp. 45–60. John Wiley & Sons Ltd. https://doi.org/10.1002/0470013494.ch3
- Ekman, P., & Cordaro, D. (2011). What is Meant by Calling Emotions Basic: *Emotion Review*. https://doi.org/10.1177/1754073911410740
- Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial Action Coding System. The Manual on CD ROM*. Network Information Research Corporation.
- Ekman, P., Kuhlmann-Krieg, S., & Reiss, M. (2010). *Gefühle lesen [In English: Emotions Revealed]*, (2nd. edition). Spektrum Akademischer Verlag. http://recherche.bibkatalog.de/InfoGuideClient.fwssis/start.do?Login=wofwsw&Query=540="978-3-8274-2568-3„
- Kushnir, A., & Müller, N. H. (2020). Haptic Feedback in Everyday Conversation Situations. In HCI International 2020 – Posters, pp. 239–244. Springer International Publishing. https://doi.org/10.1007/978-3-030-50726-8_31
- McDaniel, T., Krishna, S., Balasubramanian, V., Colbry, D., & Panchanathan, S. (2008). Using a haptic belt to convey non-verbal communication cues during social interactions to individuals who are blind. *2008 IEEE International Workshop on Haptic Audio visual Environments and Games*, pp. 13–18. https://doi.org/10.1109/HAVE.2008.4685291
- Microsoft. (n. s.). *Seeing AI* [Online]. Available from: https://www.microsoft.com/de-de/ai/seeing-ai [Accessed: 01.11.2020].
- Noldus. (n. s.). *White paper on FaceReader methodology*. [Online]. Available from: https://info.noldus.com/free-white-paper-on-facereader-methodology [Accessed: 01.11.2020].
- Orcam. (n. s.). *OrCam MyEye*. [Online]. Available from: https://www.orcam.com/de/myeye2/ [Accessed: 01.11.2020].
- Schmidt-Atzert, L. (2014). *Emotionspsychologie* (2nd. revised and enlarged edition). Verlag W. Kohlhammer.
- Spyschool. (n. s.). *HD Mini Camera*. [Online]. Available from: https://www.spyschool.de/ [Accessed: 01.11.2020].