

# Benutzungsschnittstellen zur Vermittlung erkannter Emotionen

Yannick Bachteler<sup>1</sup>, Michael Balda<sup>2</sup>, Martin Christof Kindsmüller<sup>3</sup>

User-Centered Ubiquitous Computing, Fraunhofer FIT<sup>1</sup>

Metrilus GmbH<sup>2</sup>

Technische Hochschule Brandenburg, HCI-Group<sup>3</sup>

yannick.bachteler@fit.fraunhofer.de, michael.balda@metrilus.de,  
mck@th-brandenburg.de

## Zusammenfassung

In diesem Beitrag wird die Konzeption und prototypische Entwicklung von Benutzungsschnittstellen zur Ausgabe erkannter Emotion vorgestellt. Im Fokus stehen Menschen, die Schwierigkeiten mit der Emotionserkennung in Gesprächssituationen haben, insbesondere Menschen mit einer Autismus-Spektrum-Störung (ASS) oder starken Sehbeeinträchtigung. Mithilfe zielgruppengerechter und gebrauchstauglicher Benutzungsschnittstellen zur Darstellung der Emotionen, soll ihnen und ihren Gesprächspartner\*innen im Alltag geholfen werden. Die entwickelten Konzepte berücksichtigen sowohl die Microsoft HoloLens als auch die Dot Watch (Braille-Smartwatch).

## 1 Einleitung

Dem Thema Inklusion wird in den letzten Jahren zunehmend Beachtung geschenkt. Inklusion beschreibt dabei sowohl den Prozess als auch den Zustand der Zugehörigkeit und die uneingeschränkte Teilhabe aller Menschen in der Gesellschaft. Inklusion kann durch technologische Entwicklungen unterstützt und teilweise vielleicht sogar ermöglicht werden. Dieses Ziel wurde im Rahmen der „Neue Nähe“-Hackathons (Aktion Mensch, 2017) verfolgt, indem an technologischen „Ideen für mehr Nähe und weniger Barrieren zwischen Menschen mit und ohne Behinderung“ gearbeitet wurde.

Im Rahmen des Hackathons wurde ein Prototyp zur Emotionserkennung entwickelt, welcher Menschen mit einer ASS unterstützen soll. Der Prototyp ist in der Lage folgende Basisemotionen nach Ekman und Friesen (1976, Angst, Ärger, Trauer, Freude, Ekel und Überraschung, sowie neutral) zu erkennen und auf einem Monitor auszugeben. Die automatische Emotions-

erkennung basiert auf einer kombinierten Analyse von Videodaten (*facial expression*) und Audiodaten (*vocal affect expression*), wofür entsprechende Sensoren benötigt werden. Eine erkannte Emotion wird dabei als Emoticon visualisiert, das das Gesicht vollflächig überlagert.

## 2 Benutzer- und Anforderungsanalyse

Die Konzeption zielgruppengerechter Schnittstellen erfordert zunächst die Identifizierung potentieller Zielgruppen. Mittels Recherche und darauf aufbauender Interviews konnte die bereits bekannte Zielgruppe der Menschen mit einer ASS bestätigt und zusätzlich Menschen mit einer Sehbeeinträchtigung als mögliche Benutzer\*innen identifiziert werden. Beide Zielgruppen können Defizite in der Emotionserkennung haben. Diesen soll durch die angestrebte Softwarelösung entgegengewirkt werden. Dass Emotionen im Gesicht und Sprache nicht erkannt werden, bedeutet jedoch nicht, dass mit einer gegebenen Information über das Empfinden einer Person nicht umgegangen werden kann. Weiterhin konnten Kirst et al. (2015) zeigen, dass die faziale Emotionserkennung von Menschen mit einer ASS durch computergestütztes Training verbessert werden kann. Es wird angenommen, dass eine Software zur Emotionserkennung die Zielgruppe auch in Gesprächssituationen unterstützen kann.

Bei einem zu hohen Sehverlust kann der Gesichtsausdruck von Gesprächspartner\*innen nicht mehr erkannt werden. Die faziale Emotionserkennung ist daher auch bei Menschen mit einer entsprechenden Sehbeeinträchtigung nicht möglich. Dieser fehlende Reiz wird durch andere wahrgenommene Reize wie bspw. der Stimme oder anhand von Geräuschen weitestgehend ausgeglichen. Interviewpartner\*innen der Zielgruppe gaben jedoch an, dass eine zusätzliche, softwaregestützte Analyse ihnen zu mehr Sicherheit in Gesprächen verhelfen kann, weswegen sie als 2. Zielgruppe der Emotionserkennung identifiziert wurden.

Die Zielgruppen unterscheiden sich primär durch physiologische Charakteristiken. Diese waren in der Anforderungsanalyse zunächst bei der Betrachtung möglicher Geräte ausschlaggebend. Während für Menschen mit einer ASS der Fokus auf visuellen und auditiven Ausgabegeräten (z. B. Datenbrille, Smartwatch/-phone) lag, wurden für Menschen mit einer Sehbeeinträchtigung haptische, taktile und auditive Geräte (z. B. Braille-/Smartwatch, OrCam MyEye) betrachtet. Basierend auf der vorangegangenen Analyse wurden Anforderungen erhoben und in Form von User Stories festgehalten. Diese sind z. T. gerätespezifisch formuliert und lassen sich in zielgruppenspezifisch und -übergreifend einteilen. Desweiteren wurden Risiken identifiziert, die in unmittelbarer Verbindung mit der Nutzung einer Software zur Emotionserkennung stehen. Die Risiken sind zielgruppenunabhängig und betreffen insbesondere die Akzeptanz der Benutzer\*innen und des Umfelds.

### 3 Konzeption und Realisierung

Mit Blick auf die Konzeption der Emotionserkennung als Unterstützungssystem (und nicht als Lernsystem) erfolgte die Auswahl möglicher Geräte insbesondere hinsichtlich der besten Integration in Gesprächssituationen. Für Menschen mit einer ASS konnte eine AR-fähige Datenbrille als am besten geeignet gefolgert werden. Für sehbeeinträchtigte Menschen wird angenommen, dass die bestmögliche Lösung in einer Kombination von Datenbrille (OrCam MyEye) und Braille-Smartwatch (Dot Watch) liegt. Mangels Verfügbarkeit erfolgte die Konzeption jedoch nur für die Dot Watch, die in den Zielgruppeinterviews gegenüber der Datenbrille bevorzugt wurde.

Infolge der unterschiedlichen Bedürfnisse wurden zwei eigenständige Konzepte entwickelt. Das Konzept für Menschen mit einer ASS wurde für die AR-fähige Datenbrille Microsoft HoloLens erstellt. Der Austausch durch ein anderes Head Mounted Display kann zu Unterschieden in der Interaktion und der Größe des Sichtfeldes führen. Das grundlegende Konzept wird dadurch jedoch nicht beeinträchtigt. Die Konzeption erfolgte zunächst als LoFi-Prototyp. Hierbei wurden verschiedene Darstellungsvarianten betrachtet. Möglich war eine Ausgabe als Text oder als Emoticon, deren Farbe und/oder Position wiederum unterschiedlich sein konnte. Auch wurden die Anzeigedauer und der Übergang bei Emotionswechsel mit einbezogen. Ziel der Realisierung war ein Mock-Up zur Evaluation des Konzepts. Hierfür wurden zwei kurze Gesprächssituationen aufgezeichnet, welche mithilfe der Microsoft Emotionen-API ausgewertet wurden. Die erkannten Emotionen wurden durch die manuelle Nachbearbeitung sowohl als Text als auch als Emoticon eingeblendet. Ergebnis der Realisierung waren somit Videos, welche die Echtzeit-Emotionserkennung für die in Abbildung 1 dargestellten Varianten simuliert.



Abbildung 1: Die entsprechende Darstellung erfolgt jeweils relativ zum erkannten Gesicht.

Das Konzept für sehbeeinträchtigte Menschen wurde im Funktionsumfangs auf die Braille-Smartwatch Dot Watch ausgelegt und umfasst neben der Darstellung erkannter Emotionen mit Braille-Zeichen auf der Dot Watch auch eine Smartphone-App (iOS), welche mithilfe der betriebssysteminternen Bedienungshilfen gesteuert werden kann.

### 4 Evaluation und Ergebnisse

Die Evaluation des Konzepts der AR-fähigen Datenbrille sollte zeigen, ob die realisierten Darstellungen für die Ausgabe erkannter Emotionen im Gesichtsfeld geeignet sind. Um die Ausgabe im Gesichtsfeld zu simulieren wurden die Videos durch die Proband\*innen ( $n=25$ ) mit Hilfe einer Cardboard VR-Brille betrachtet. Die Bewertung erfolgte sowohl quantitativ (UEQ,

Laugwitz et al. 2006) als auch qualitativ durch einen Fragebogen. Bei keiner der Proband\*innen liegt eine ASS vor. Dennoch konnten erste Hinweise zur Gestaltung der Ausgabe (Darstellung, Anzeigedauer, Übergang bei Emotionswechsel) im Gesichtsfeld und der technischen Umsetzung festgehalten werden. Die Ergebnisse der qualitativen und quantitativen Umfragen zeigen, dass die Darstellung mit Emoticons bevorzugt wird. Die gesammelten Daten geben keine Hinweise darauf, dass die Ausgabe im Gesichtsfeld als störend oder ungeeignet empfunden wird.

Die Auswertung des UEQs ergab für beide Darstellungen durchweg positive ( $>0,8$ ) Bewertungen. Bei Betrachtung aller Datensätze liegt lediglich für die Attraktivität ein signifikanter Unterschied ( $p < .05$ ) vor. Im Vergleich der zuerst betrachteten Darstellung (13 Datensätze) ist der Unterschied nur für die Dimension der Durchschaubarkeit signifikant. Im abschließenden Fragebogen präferierten 72% der Proband\*innen Emoticons. Diejenigen mit einer Präferenz für Textausgabe, gaben an, dass sie Text zwar den Emoticons vorziehen, aber auch diesen nicht als geeignet empfinden. In der Nachbefragung konnten mögliche Verbesserungen und neue Visualisierungsideen gesammelt werden, welche Eingang in künftigen Untersuchungen finden werden. U.a. der Vorschlag mit mehr Farben zu arbeiten („Mund“ des Emoticons farblich unterschiedlich darstellen) oder auf ein reines Farbsystem zu setzen (z. B. Stimmungsbarometer).

Aufgrund fehlender Hardware konnte das Konzept der Braille-Smartwatch nicht ausreichend realisiert werden, um einen Benutzertest durchführen zu können. Dennoch konnte im Rahmen einer formativen Evaluation mithilfe von Interviews ( $n=3$ ) mit der Zielgruppe einzelne Aspekte des Konzepts diskutiert und deren Ergebnisse dokumentiert werden. Sowohl der unauffällige Einsatz als auch die anpassbare Zeichenlänge wurden hierbei positiv hervorgehoben.

Die Frage, ob die jeweilige Technologie gewinnbringend im Alltag eingesetzt werden kann, muss für beide Konzepte im Rahmen weiterführender Evaluationen beantwortet werden. Hierfür ist es notwendig (mehr) Proband\*innen beider Zielgruppen einzubeziehen. Die stete Weiterentwicklung der Technologien wird ebenso wichtig für den Erfolg der Konzepte sein, wie auch das Steigern der Akzeptanz von Benutzer\*innen und deren Gesprächspartner\*innen.

## Literaturverzeichnis

- Aktion Mensch. (2017). Für eine Neue Nähe: Hackathon in Berlin. Abgerufen am 03. 05 2018 von <https://www.aktion-mensch.de/neuenaeh/hackathon.html>
- Ekman, P., & Friesen, W. (1976). Pictures of Facial Affect. Consulting Psychologists Press.
- Kirst, S., Zoerner, D., Schütze, J., Lucke, U., & Dziobek, I. (2015). Zirkus Empathico: Eine mobile Applikation zum Training sozioemotionaler Kompetenzen. In Pongratz, H. & Keil, R. (Hrsg.): DeLFI 2015 - Die 13. e-Learning Fachtagung Informatik der Gesellschaft für Informatik e. V., S. 107-118.
- Laugwitz, B., Schrepp, M. & Held, T., (2006). Konstruktion eines Fragebogens zur Messung der User Experience von Softwareprodukten. In: Heinecke, A. M. & Paul, H. (Hrsg.): *Mensch und Computer 2006*. München: Oldenbourg Verlag. S. 125-134.