



Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Emotionserkennung mittels Bio-Sensoren

Master Hauptseminar
Andreas Kamenz

06.12.2016



Inhaltsübersicht

- **Einleitung**
- **Grundlagen**
- **Lösungsansätze**
- **Bisherige Ergebnisse und weitere Ziele**
- **Zusammenfassung**



Problemstellung

- **Computer sollen menschliche Emotionen erkennen**
- **Ziel: Human Computer Interaction verbessern**
- **Beispiel: Companion Systeme sollen emotionalen Zustand erkennen und entsprechend reagieren**



[F1]



Anwendungsgebiete

- **Companion Systeme**
- **Spiele**
- **Smart Home**
- **HealthCare (Krankenhaus)**
- **Fitnessstudio**
- **Psychologie (Depressionserkennung)**



Emotionsmodelle (diskret)

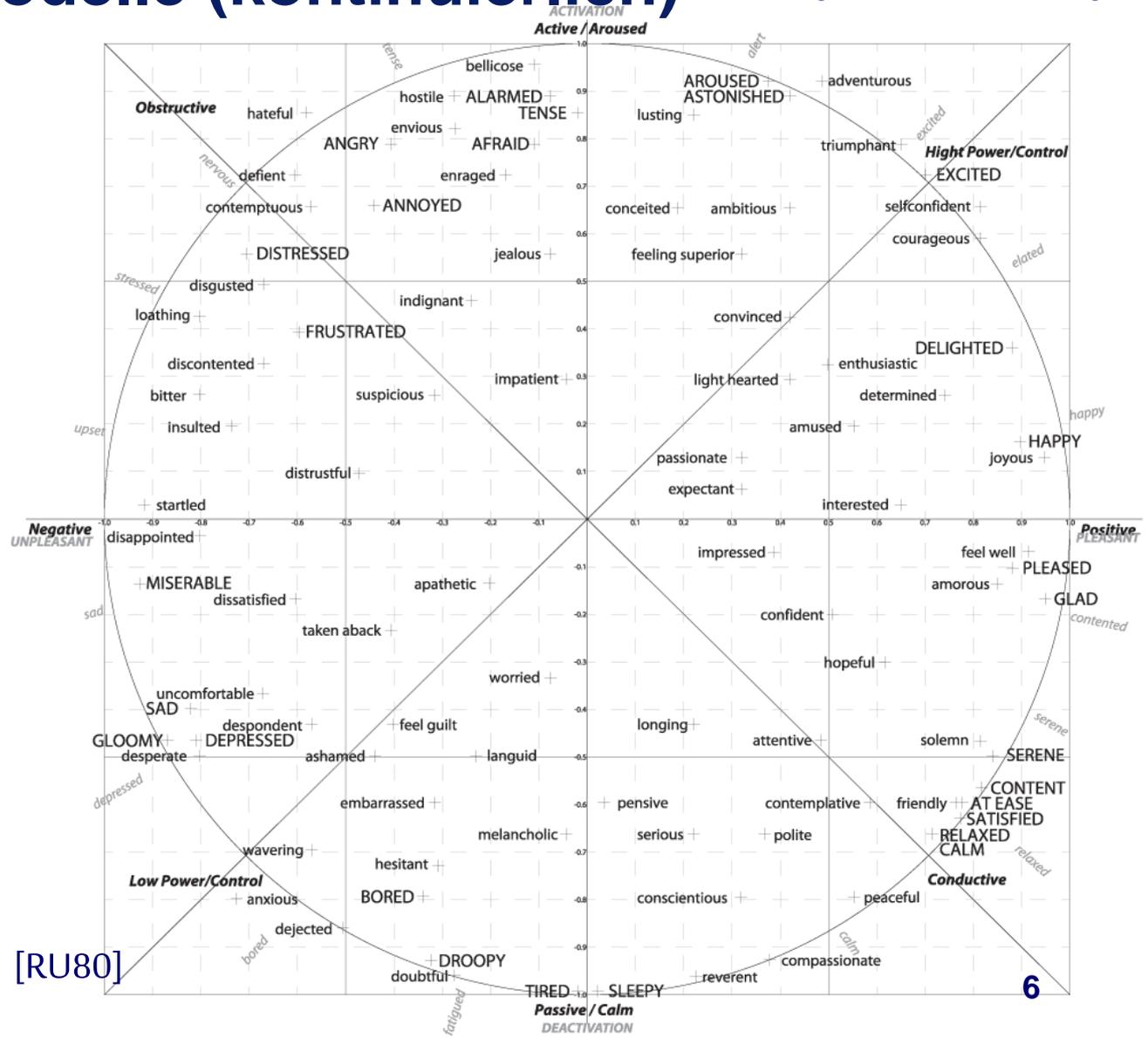
- **6 Grundemotionen nach Ekman [PE72]:**
Anger, Disgust, Fear, Joy, Sadness, Surprise
- **7 nach Ekman's Facial Action Coding System (FACS) [PE78]:**
Anger, Disgust, Fear, Happy, Sadness, Surprise, Contempt
- **24 Emotionen nach Baron-Cohen [BC04]:**

Afraid	Angry	Bored	Bothered
Disbelieving	Disgusted	Excited	Fond
Happy	Hurt	Interested	Kind
Liked	Romantic	Sad	Sneaky
Sorry	Sure	Surprised	Thinking
Touched	Unfriendly	Unsure	Wanting



Emotionsmodelle (kontinuierlich) Hochschule für Angewandte Wissenschaften Hamburg

- Arousal
- Valence



Emotionserkennung nr [RU80]



Teamprojekt / Schnittstellen

- **EmotionBike Projekt**
- **Ziel: Emotionserkennung**
- **Methoden:**
 - Gesichtserkennung
 - Eye-Tracking
 - Wärmebildkamera
 - Bio Sensoren



[F2]



EmotionBike [MU16]

Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences



Emotionserkennung mittels Bio-Sensoren

[F3]



Herausforderungen

- **Welche Indikatoren für Emotionen?**
- **Wie kombinieren?**
- **Wie sind sie zu deuten?**
- **Welches Dataset zur Klassifikation?**
- **Emotionen bei der Kommunikation kommen über verschiedene Kanäle [ME68]:**
 - **55% Mimik**
 - **38% Stimme**
 - **7% Sprache**



Lösungsansätze



Lösungsansatz 1

- **Videobasiertes Verfahren, z.B. [CH15]**
 1. Bildsequenzen, aus denen Features extrahiert werden
 2. Das Video wird als globaler Featurevektor repräsentiert
- **Problem:**
 - Anatomie von Gesichtern ist sehr unterschiedlich
 - Ausrichtung des Gesichtes ist wichtig
 - Wenn Erkennung in Echtzeit erfolgen soll,
 - dann schnelle Verfahren für Feature Extraktion (HOG, SIFT) & Klassifikation (SVM) nötig
- **Geschwindigkeit ist wichtiger als Erkennungsleistung**



Lösungsansatz 1 - Ergebnis

Hochschule für Angewandte Wissenschaften Hamburg
Hambura University of Applied Sciences

Dataset CK+

Method	Accuracy on CK+ (%)							
	An	Co	Di	Fe	Ha	Sa	Su	Ave.
CLM [32]	70.1	52.4	92.5	72.1	94.2	45.9	93.6	74.4
HOE [5]	76.4	65.4	83.6	73.3	92.1	88.6	92.8	82.3
LBP-TOP [4]	82.2	77.8	91.5	72.0	98.6	57.1	97.6	82.4
ITBN [6] (15-flod)	91.1	78.6	94.0	83.3	89.8	76.0	91.3	86.3
HOG 3D [7]	84.4	77.8	94.9	68.0	100	75.0	98.8	85.6
LSH-CORF [3]	71.3	–	90.8	79.0	92.6	90.5	96.6	86.8
3D LUT [21]	76.3	35.1	60.5	73.8	91.0	48.2	92.8	68.2
3DCNN-DAP [31]	91.1	66.7	96.6	80.0	98.6	85.7	96.4	87.9
STM [2]	–	–	–	–	–	–	–	91.1
McRiHOG	94.3	82.9	92.7	91.5	93.1	81.6	97.3	90.5

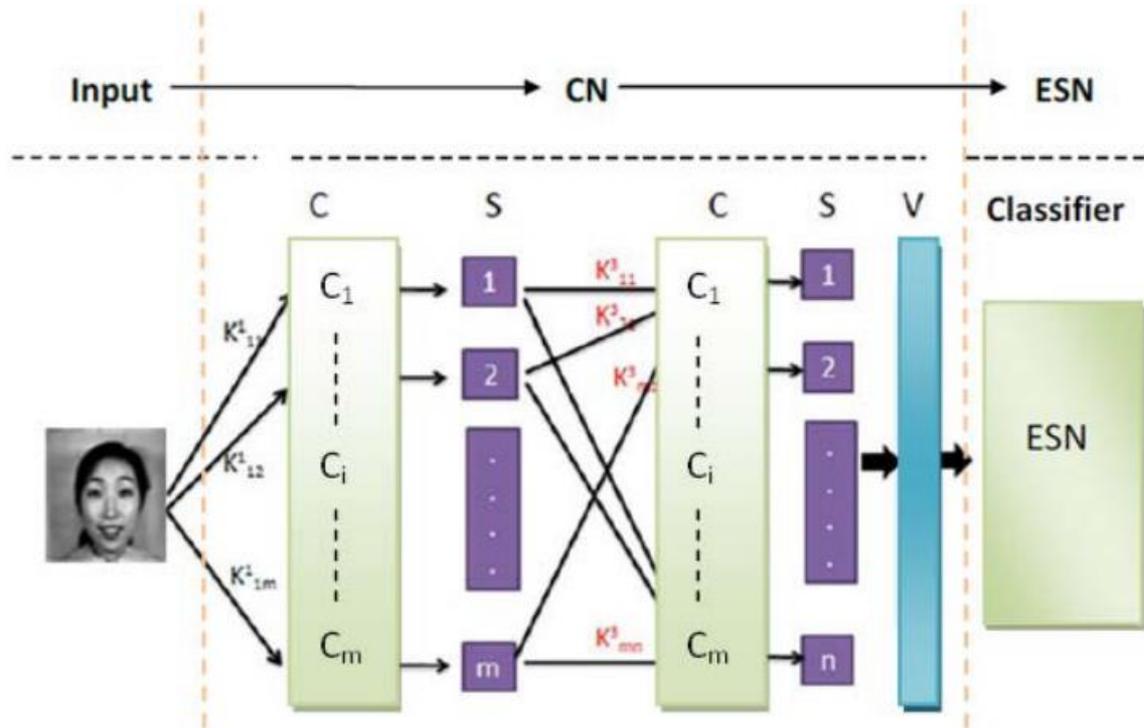
Dataset AFEW

Method	Accuracy on AFEW (%)						
	An	Di	Fe	Ha	Sa	Su	ave.
HOE [5]	11.2	16.5	9.0	33.5	15.3	28.3	19.0
LBP-TOP [4]	11.7	19.6	17.9	42.3	23.8	33.6	24.8
HOG 3D [7]	–	–	–	–	–	–	26.9
LSH-CORF [3]	23.1	12.8	38.6	9.7	21.1	10.9	19.4
3D LUT [21]	45.7	0	0	62.0	13.2	48.6	28.2
STM [2]	–	–	–	–	–	–	31.7
McRiHOG	68.2	0	48.1	83.3	32.0	91.6	53.6



Lösungsansatz 1

- Videobasiert, nach [WE15]
- Verbesserungsvorschlag:
 - Spezielle CNNs zur Feature Extraction
 - Random Architektur



[F5]



Lösungsansatz 2

- **Audiobasiertes Verfahren, z.B. [FE15]**
- **31 Low Level Descriptors und 42 statistische Funktionen**
- **Daraus resultieren ca. 2000 akustische Features**
- **SVM für die Klassifikation**

Database	Language	Family	Symbol	# Arousal		# Valence		# m	# f	kHz
				-	+	-	+			
Emo-DB [32]	German	Germanic	DE	248	246	352	142	5	5	20
DES [61]	Danish	Germanic	DK	104	156	156	104	2	2	20
Enterface [20]	English	Germanic	GB	215	857	427	645	34	8	16
SES [62]	Spanish	Romanic	ES	15	18	15	18	1	0	16
SRoL [18]	Romanian	Romanic	RO	154	154	154	154	11	8	22
Busim [45]	Turkish	Turkic	TR	242	242	242	242	3	8	16
Mandarin [13]	Mandarin	Sino-Tibetan	CN	60	180	120	120	3	3	22
Burmese [13]	Burmese	Sino-Tibetan	MM	69	177	108	138	3	3	22

[F6]



Lösungsansatz 2 - Ergebnis

Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

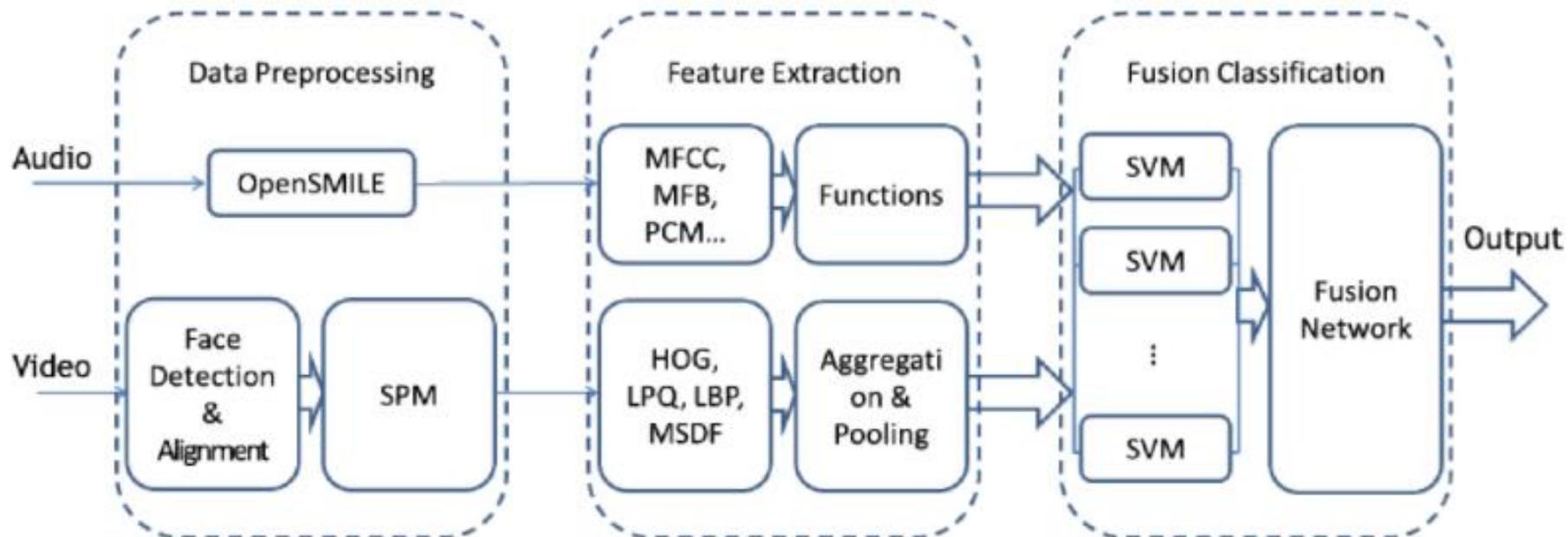
% UA test on:	AROUSAL (train on:)							
	DE	DK	GB	ES	RO	TR	CN	MM
DE	97.3*	50.3	71.3 [∩]	54.6 [∩]	50.0	69.0	60.9 [∩]	68.7 [∩]
DK	50.7 [∩]	95.0*	79.4	60.6	59.0	50.0	78.3	85.7
GB	56.4	62.6	87.7*	63.6	59.7	50.2	75.4	71.1
ES	52.6	65.3	54.2	100.0*	53.2	51.8	77.0	76.4
RO	63.1	68.0	78.9	60.7 [∩]	87.3*	52.8	65.4	54.0
TR	51.4	53.0	78.4	54.6 [∩]	56.8	88.4*	72.9	50.8
CN	72.0	65.0	76.8	72.7	68.1	63.8	99.5*	92.6
MM	58.1 [∩]	57.4 [∩]	57.9	54.6 [∩]	51.9	53.0	85.4	97.1*
	VALENCE							
DE	86.3*	58.5 [∩]	59.9	54.5	50.3	51.6	62.5	54.4
DK	50.8	68.4*	59.6 [∩]	51.5	52.5	58.6	55.5 [∩]	57.8 [∩]
GB	71.0	53.9 [∩]	79.4*	51.5	50.9	51.2	54.6 [∩]	52.4
ES	58.3 [∩]	54.2	61.3	100.0*	50.0	51.6	57.1 [∩]	64.3 [∩]
RO	61.3	52.0 [∩]	57.0 [∩]	54.5	56.4*	54.2 [∩]	55.0 [∩]	54.0
TR	67.2	57.3	50.4	51.6 [∩]	52.9	72.3*	50.5 [∩]	52.9 [∩]
CN	57.7 [∩]	54.6	54.2	54.5	50.7 [∩]	54.9	95.8*	83.7
MM	51.3 [∩]	51.6 [∩]	51.7	54.5	53.6 [∩]	50.7 [∩]	77.0	94.7*

UNWEIGHTED ACCURACY (UA) FOR CROSS-LANGUAGE POLARITY RECOGNITION; TRAIN ON ONE LANGUAGE, TEST ON ANOTHER LANGUAGE; MAIN DIAGONAL (*): INTRA-CORPUS CROSS-VALIDATION (NOT INCLUDED IN THE MEANS); [∩] INDICATES DATA-BASED MODEL-INVERSION CASES.



Lösungsansatz 3

- Kombination von Video & Audio, z.B. [SU15]
- Beste bekannte Erkennungsrate ~50%
- Dataset: AFEW

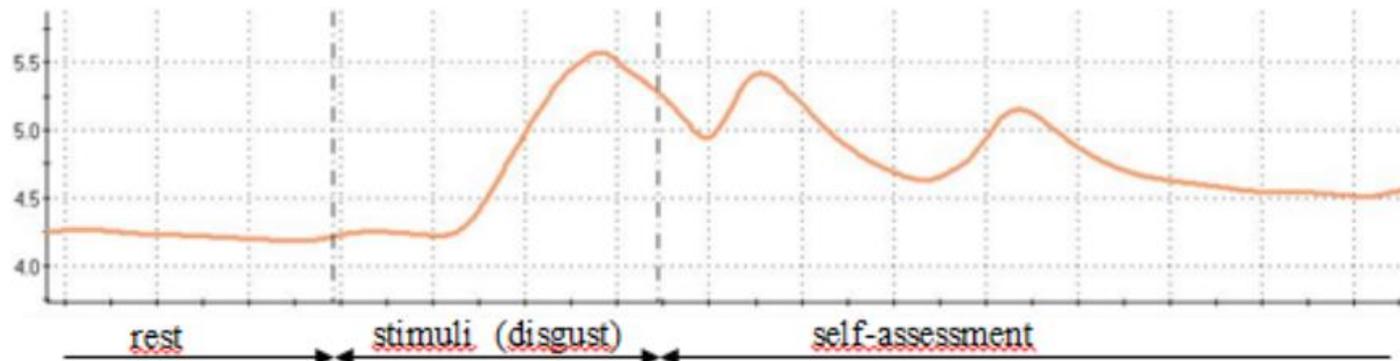


[F8]



Lösungsansatz 4

- **Physiologische Daten, z.B. [WI15]**
- **Emotionserkennung mit Sensoren für:**
 - Hautwiderstand (EDR)
 - Puls
 - Atmung
 - Körpertemperatur
- **Emotionen werden zeitversetzt erkannt, z.B. EDR:**



Emotionserkennung mittels Bio-Sensoren



Eigener Lösungsansatz

- **Bio-Sensoren zur Emotionserkennung**
 - Physiologische Daten erfassen
 - Features extrahieren und klassifizieren
- **Kombination von Video, physiologische und weitere Datenquellen (EmotionBike)**
- **Sehr aktuelles Thema**
- **Erstmalige Challenge zur Emotionserkennung: AV+EC [AV15], Dataset AFEW 6.0**



Eigener Lösungsansatz

Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

- **Messwerte:**
 - Hautwiderstand
 - Hauttemperatur
 - Puls
 - Atmung
 - EKG (Herz)
 - Beschleunigung



Emotionserkennung mittels Bio-Sensoren

[F10]



Bisherige Ergebnisse

- **Grundprojekt:**
- **Speziellen Befragungsmanager entwickelt**
- **System zur Datenspeicherung und Synchronisierung entwickelt**
- **Probemessungen mit folgenden Sensoren durchgeführt:**
 - **Polar Brustgurt H7**
 - **Fitbit Surge HR**
 - **Plux Biomedical Sensor Toolkit**





Bisherige Ergebnisse

- **Hauptprojekt:**
- **Integration des Agenten ins Gesamtsystem**
 - **Eventsystem zur Steuerung**
- **Features entwickelt**
 - **Peak Detection, Beschleunigung, Max/Min, Durchschnitt für EDA, Puls, Temperatur, Atmung**
- **Erste Versuche durchgeführt**



Beispiel

Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences



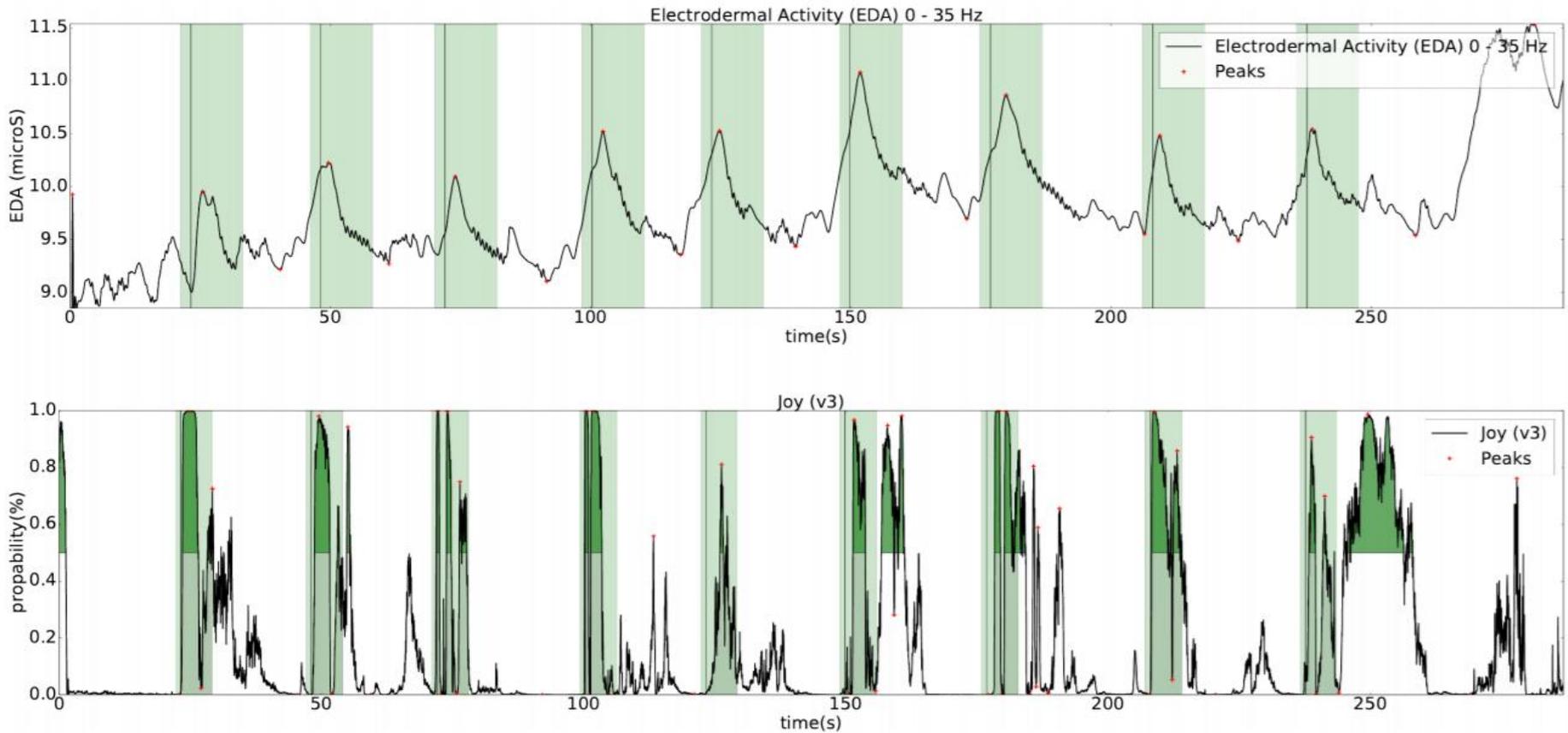
Emotionserkennung mittels Bio-Sensoren

[F12]



Beispiel

Hochschule für Angewandte Wissenschaften Hamburg
 Hamburg University of Applied Sciences



[F13]



Weitere Ziele

- **Masterarbeit:**
 - Emotionserkennung mit physiologischen Daten im Kontext Exergames
 - Sensordaten analysieren, auswerten und ggf. mit den anderen Lösungsansätzen vergleichen

- **Risiken:**
 - Klassifizierung anhand Features kann schwierig werden
 - Datenqualität



Zusammenfassung

- **Relativ neuer Forschungsbereich**
- **Viele aktuelle Veröffentlichungen**
- **Viel Verbesserungspotential bei Erkennungsraten**
- **Vorhandene Ergebnisse teilweise nicht vergleichbar (unterschiedliche Datasets)**



Vielen Dank für die Aufmerksamkeit!

06.12.2016

Referenzen 1/2

- [ME68] A. Mehrabian. (1968). Communication without words. *Psychological today*,2, 53-55.
- [PE72] P. Ekman, W. V. Friesen, and P. Ellsworth, Emotion in the human face. New York: Pergamon Press, 1972.
- [PE78] P. Ekman and W. V. Friesen, Facial action coding system: a technique for the measurement of facial movement. Palo Alto: Consulting Psychologists Press, 1978.
- [RU80] J. Russel, A Circumplex Model of Affect, *Journal of Personality and Social Psychology* 1980, Vol. 39, No. 6, pp. 1161-1178
- [BC04] S. Baron-Cohen, O. Golan, S. Wheelwright, and J. Hill, “Mind reading: the interactive guide to emotions.” DVD, 2004.
- [MU15] L. Müller et al., EmotionBike: A Study of Provoking Emotions in Cycling Exergames, In: Entertainment Computing - ICEC 2015, Volume 9353 of the series Lecture Notes in Computer Science pp. 155-168
- [CH15] Chen, Jinhui; Takiguchi, Tetsuya; Arika, Yasuo, Facial Expression Recognition with Multithreaded Cascade of Rotation-invariant HOG, In: International Conference on Affective Computing and Intelligent Interaction (ACII), pp. 155-168, (2015)
- [WE15] Wen, Guihua; Li, Huihui; Li, Danyang, An ensemble convolutional echo state networks for facial expression recognition, In: International Conference on Affective Computing and Intelligent Interaction (ACII), pp. 873–878, (2015)
- [FE15] Feraru, Silvia Monica; Schuller, Dagmar; Schuller, Björn, Cross-language acoustic emotion recognition: An overview and some tendencies, In: International Conference on Affective Computing and Intelligent Interaction (ACII), pp. 125–131, (2015)

Referenzen 2/2

- [SU15]** Chen, Jinhui; Takiguchi, Tetsuya; Arika, Yasuo, Facial Expression Recognition with Multithreaded Cascade of Rotation-invariant HOG, In: International Conference on Affective Computing and Intelligent Interaction (ACII), pp. 155-168, (2015)
- [WI15]** Szwoch, Wioleta, Emotion Recognition Using Physiological Signals, In: MIDI'15 Proceedings of the Multimedia, Interaction, Design and Innovation, Article No. 15, pp. 1-8, (2015)
- [AV15]** Ringeval, Fabien; Schuller, Björn; Valstar, Michel; Jaiswal, Shashank; Marchi, Erik; Lalanne, Denis; Cowie, Roddy; Pantic, Maja, AV+ EC 2015 – The First Affect Recognition Challenge Bridging Across Audio, Video, and Physiological Data, In: AVEC '15 Proceedings of the 5th International Workshop on Audio/Visual Emotion Challenge, (2015)
- [MU16]** L. Müller et al., IEEE SSCI 2016, Physiological Data Analysis for an Emotional Provoking Exergame

Abbildungen

- [F1] <http://i2.cdn.cnn.com/cnnnext/dam/assets/150622133225-pepper-robot-hugs-girl-medium-plus-169.jpg>
- [F2] [MU15]
- [F3] [MU16]
- [F4] [CH15]
- [F5] [WE15]
- [F6] [FE15]
- [F7] [FE15]
- [F8] [SU15]
- [F9] [WI15]
- [F10] © Victoria Bibaeva
- [F11] <http://plux.info/>
- [F12] [MU16]
- [F13] [MU16]