CHRISTOPH AMMA, ANDREAS FISCHER, THORSTEN STEIN, HERMANN SCHWAMEDER, TANJA SCHULTZ

Emotionserkennung auf der Basis von Gangmustern

Einleitung

Zwischenmenschliche Kommunikation ist ein vielschichtiger Prozess, bei dem Informationen auf unterschiedlichen Ebenen und durch verschiedene Modalitäten ausgetauscht werden. Es werden sowohl faktische als auch emotionale Inhalte übermittelt. Dies geschieht sowohl verbal, als auch nonverbal. Menschen erfassen den emotionalen Zustand ihres Partners und passen ihr Verhalten der emotionalen Situation entsprechend an. Verglichen hierzu, ist die Kommunikation zwischen Menschen und technischen Systemen auf aktive Eingaben durch den Menschen und statische Reaktionen des Systems reduziert. Um Menschen eine möglichst natürliche und intuitive Kommunikation mit technischen Systemen zu ermöglichen, sollten diese auch nonverbale Kommunikationsinhalte erkennen und interpretieren können (Robinson & Kaliouby, 2009). Menschen sind beispielsweise in der Lage, durch die Beobachtung von Gang und Körperhaltung anderer Menschen, Rückschlüsse auf deren emotionalen Zustand zu ziehen und ihr Verhalten entsprechend anzupassen. Unsere Arbeit stellt einen ersten Schritt in Richtung der Nachbildung dieser Fähigkeit in technischen Systemen dar. Hierfür muss untersucht werden, welche Bewegungsmerkmale Rückschlüsse auf den emotionalen Zustand eines Menschen zulassen und wie diese Zustände auf Basis der Merkmale mit automatischen Verfahren erkannt werden können (vgl. z.B. Janssen et al., 2008). In der vorliegenden Arbeit untersuchen wir die Auswirkungen emotionaler Zustände auf das Gangmuster der Probanden und die Möglichkeit diese automatisch zu klassifizieren.

Methode

Biomechanische Messwertaufnahme

Die Datenerhebung der Gangmuster erfolgte unter Laborbedingungen bei zehn Probanden, von denen fünf weiblich und fünf männlich waren. Die Gehbewegung fand auf einem Laufband der Firma HP Cosmos mit einer Länge von 300cm und einer Breite von 125cm statt. Bei den Probanden konnte auf Grund des Studienfachs (Sportwissenschaft/-pädagogik) davon ausgegangen werden, dass sie entweder bereits Erfahrung mit Gehen auf dem Laufband vorweisen konnten, oder aber schnell an die Laufbandumgebung adaptieren konnten. Die Wahrscheinlichkeit, dass die Probanden eine unnatürliche Gehbewegung während der Messwert-

aufnahme zeigten wurde dadurch minimiert, dass vor der Aufzeichnung eine Adaptationsphase von 192 Sekunden stattfand, die direkt in die Messung überleitete. Die Probanden gingen für den gesamten Zeitraum der Eingewöhnungsphase und der Messwertaufnahme bei einer selbst gewählten Gehgeschwindigkeit. Während der Emotionsinduktion über einen Beamer und einen Kopfhörer wurden die Probanden kinematisch, dynamisch und elektromyographisch vermessen.

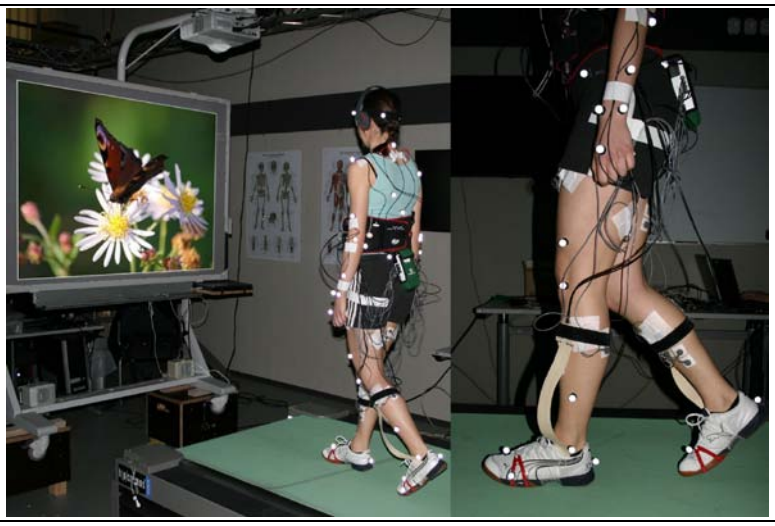

Abb. 1: Versuchsaufbau

Kinematische Messwertaufnahme: Für die Erfassung der kinematischen Parameter wurde ein Vicon MX System mit 14 Kameras verwendet, dass mit einer Frequenz von 200Hz Daten über den gesamten Zeitraum erfasste. Reflektive Marker mit einem Durchmesser von 14mm wurden an den Probanden gemäß dem Vicon PluginGait Markerset angebracht. Zur Datensynchronisation wurde ein Triggersignal mit dem Vicon A/D-Wandler eingelesen.
Dynamische Messwertaufnahme: Die dynamische Messwertaufnahme erfolgte in den Schuhen der Probanden mit Einlegesohlen der Firma Novel (Pedar). Die Druckverteilung am Fuß wurde mit einer Frequenz von 100Hz erfasst und zeitsynchron zur Kinematik aufgezeichnet. Für jeden Probanden standen Sportschuhe vom gleichen Modell und Pedarsohlen in der entsprechenden Größe zur Verfügung.
Elektromyografische Messwertaufnahme: Die Aufzeichnung der Muskelaktivität erfolgte mit einem telemetrischen EMG-System der Firma Noraxon mit einer Frequenz von 1500Hz hardwareseitig und 1000Hz softwareseitig. Das System wurde durch ein analoges Signal mit Vicon und Pedar synchronisiert. Für die Analyse der Musteränderungen beim Gehen wurden ausschließlich Muskelgruppen der unteren Extremität abgeleitet. Elektroden wurden auf die MM. Gastroknemius lateralis, MM. Soleus, MM. Bizeps femoris, MM. Tibialis und MM. Rektus femoris beider Seiten angebracht. Zusätzlich wurden über das EMG-System die Herzfrequenz (HF), sowie die elektrodermale Aktivität (EDA) an der Innenseite der nicht-dominanten Hand aufgzeichnet.

Emotionsinduktion

Für die Induktion emotionaler Zustände werden in der Fachliteratur verschiedene Techniken vorgeschlagen (Sokolowski, 2008). Die Befunde deuten daraufhin, dass die Induktionseffekte insgesamt für negative Stimmungen stärker ausfallen als für positive Stimmungen und die Darbietung von Filmen und Geschichten am wirksamsten ist (Westerman et al., 1996), wobei anzumerken ist, dass es keine Reize gibt, die bei allen Menschen zu jedem Zeitpunkt eine bestimmte Emotion anregen können (Sokolowski, 2008). Auf der Basis der vorliegenden Befunde zur Emotionsinduktion, wurden den Probanden visuelle Reize in Form einer Bilderdiashow in Kombination mit akustischen Reizen präsentiert. Die Bilder wurden aus dem International Affective Picture Set (Lang et al., 2008) und die Geräusche aus dem International Affective Digital Sounds System (Bradley et al., 1999) gewählt. Die Bilder und Geräusche wurden in einer Untersuchung in den zwei emotionalen Dimensionen „Erregung“ und „Valenz“ bewertet und können somit als validierte Referenzdaten verwendet werden. Es wurden insgesamt drei jeweils extreme Ausprägungen in beiden Dimensionen gewählt und entsprechende Blöcke von Bildern (6 Bilder pro Block) und Geräuschen (4 Geräusche pro Block) zusammengestellt. Davon hatten zwei Blöcke maximale Valenz, werden also von Probanden in der Regel als positiv empfunden, einmal mit minimaler und einmal mit maximaler Erregung. Der dritte Block umfasste Daten mit minimaler Valenz und maximaler Erregung (Abb. 2).

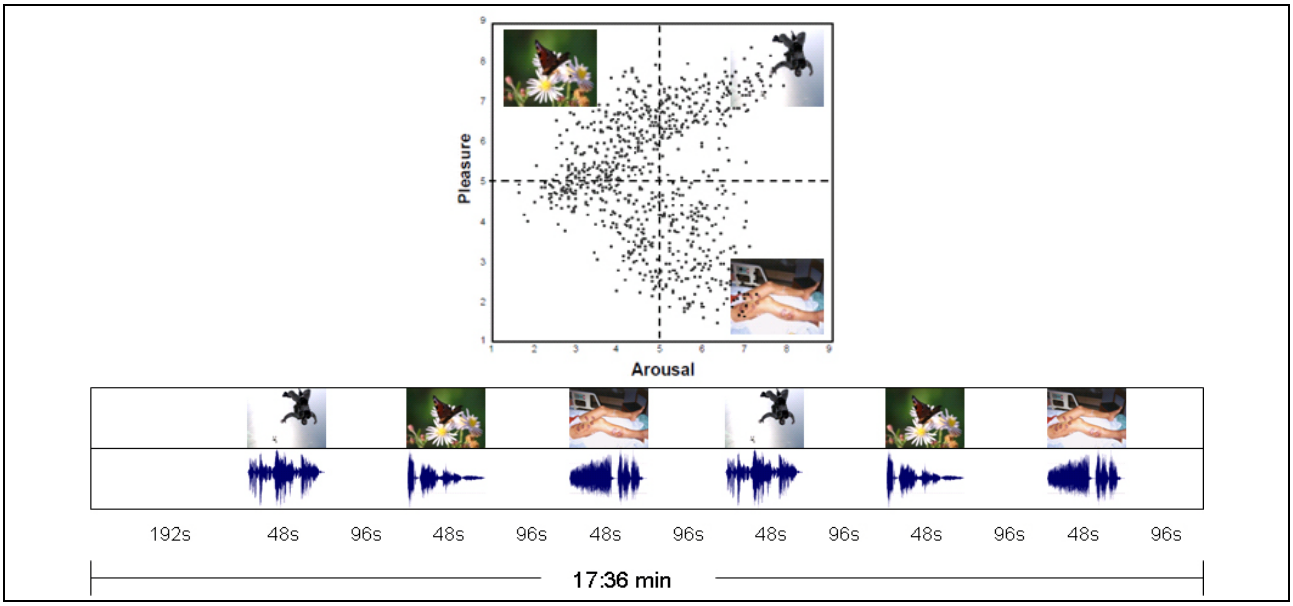

Abb. 2: Einteilung der Bilder und Töne nach den Dimensionen Erregung (Arousal) und Valenz (Pleasure) und drei extreme Beispiele (oben). Im unteren Teil der Abbildung ist der Ablauf der Intervention zur Emotionsinduktion dargestellt. Zur Gewöhnung an die Untersuchungssituation war der erste neutrale Block von doppelter Länge. Die Bilder ähneln den Originalen (Quelle: wikimedia.org).

Von jedem Block wurden zwei unterschiedliche Versionen erstellt, wobei jeder der insgesamt 6 Blöcke 48 Sekunden lang war. Die Blöcke wurden den Probanden auf eine Tafel in ihrem Sichtfeld projiziert und die Geräusche per Kopfhörer vorgespielt (Abb. 1). Zwischen den Blöcken wurde jeweils für 96 Sekunden ein weißer Bild-

schirm ohne Ton als neutraler Reiz präsentiert (Abb. 2). Nach dem Versuch füllten alle Probanden einen Fragebogen aus, indem sie selbst beurteilten, wie sie die Bilderfolgen der einzelnen Blöcke empfunden haben. Hierfür wurde die etablierte SAM Skala verwendet, die auch bei der oben erwähnten Validierungsstudie eingesetzt wurde.

Bewegungsklassifikation

Zur Modellierung und Klassifikation der Kinematik Daten wurden Hidden Markov Modelle (HMM) eingesetzt (Rabiner, 1989). HMMs sind ein stochastisches Verfahren zur Modellierung sequentieller Daten. Die Daten wurden hierzu automatisch in einzelne Gangzyklen segmentiert. Die Gangzyklen, die im Bereich ±1s um die Blockgrenzen lagen wurden aus den Daten entfernt um die Zyklen eindeutig den Blöcken zuordnen zu können. Als Merkmale wurden die Gelenkwinkel der unteren Extremitäten über die Zeit verwendet. Für jeden der drei induzierten emotionalen Zustände sowie den neutralen Zustand wurde ein HMM trainiert. Um Effekte, die sich möglicherweise aus der Änderung des Gangmusters über die gesamte Versuchszeit ergeben könnten auszuschließen, wurde auf den Daten des jeweils ersten Blocks einer Klasse trainiert und auf den Daten des zweiten Blocks getestet. Die Menge der Testdaten war somit auch disjunkt zur Menge der Trainingsdaten. Als Topologie für die HMMs wurden links-rechts Modelle mit 10 Zuständen und reflexiven Kanten in jedem Zustand gewählt. Die Emissionswahrscheinlichkeiten wurden mit Gaußschen Mischverteilungen modelliert. Die Gangzyklen aus der Testmenge wurden einzeln klassifiziert. Als Gütemaß wurde die Erkennungsrate (ER) verwendet. Diese gibt das Verhältnis korrekt klassifizierter Gangzyklen zur Gesamtzahl der Zyklen an.

Ergebnisse

Die Selbstbeurteilung der 10 Probanden auf der Basis der neunstufigen SAM-Skala (1=minimal, 9=maximal) deutet daraufhin, dass die gewählte Methode der Emotionsinduktion die gewünschten Effekte tendenziell erzielt hat (Tab. 1).

Tab. 1: Deskriptive Statistik (Mittelwert und Standardabweichung) der Selbstbeurteilung (N=20) der Probanden auf der Basis der neunstufigen SAM-Skala (1=minimal, 9=maximal). Dabei sind die beabsichtigten Wirkungen: Block 1 (B1) max. Valenz und min. Erregung; Block 2 (B2) max. Valenz und max. Erregung; Block 3 (B3) min. Valenz und max. Erregung.

	B1 Valenz	**B1 Erregung**	**B2 Valenz**	**B2 Erregung**	**B3 Valenz**	**B3 Erregung**
Mittelwert	7,1	2,9	7,9	5,6	1,8	8,1
Standardabw.	0,8	1,2	1,1	2,3	0,9	0,8

Tabelle 2 zeigt die Ergebnisse der Klassifikation der Einzelgangzyklen mittels HMMs auf Basis der Kinematikdaten. Die Klassifikation wurde auf den Daten von 9 Probanden durchgeführt, da die Daten einer Probandin nicht nutzbar waren. Während des Experimentes gab es keine menschlich beobachtbaren auffälligen Ände-

rungen im Gangmuster der Probanden über die Zeit. Die ersten Ergebnisse der automatischen Klassifikation deuten daraufhin, dass eine Klassifikation verschiedener emotionaler Zustände mit HMMs auf der Basis von Winkelverläufen der unteren Extremitäten beim Gehen auf dem Laufband nicht möglich ist.

Tab. 1: Erkennungsraten (ER) auf den Kinematikdaten (Winkelverläufe der unteren Extremitäten) der Einzelgangzyklen für vier Klassen.

Daten	Mittelwert ER	Standardabw. ER
Alle (5m/4w)	33%	10%
Weiblich (4w)	41%	10%
Männlich (5m)	28%	7%

Diskussion

Die bisherigen Ergebnisse auf Basis intrinsischer kinematischer Parameter der unteren Extremitäten zeigen noch keine befriedigende Erkennungsleistung. Derzeit arbeiten wir daran, komplexere Merkmale zu extrahieren, wie beispielsweise statistische Größen der einzelnen Phasen eines Gangzyklus. Zusätzlich werden wir eine Klassifikation auf der Basis mehrer Gangzyklen, beispielsweise per Mehrheitsentscheid auswerten, da eine Erkennung von Einzelgangzyklen für kein relevantes Anwendungsszenario nötig ist. Die Auswertung der EMG, Pedar, EDA und HF Daten steht noch aus. Darüber hinaus könnte eine Fusion der einzelnen Signale im Falle komplementärer Informationen ebenfalls die Erkennungsleistung verbessern. Grundsätzlich muss auch hinterfragt werden, ob der Versuchsaufbau mit einem Laufband zu restriktiv war, d.h. die durch das Laufband vorgegebene konstante Geschwindigkeit die Bewegungsvariabilität zu stark einschränkt.

Literatur

Bradley, M. M. & Lang, P. J. (1999). *International affective digitized sounds (IADS): Stimuli, instruction manual and affective ratings* (Tech. Rep. No. B-2). University of Florida.

Janssen, D. et al. (2008). Recognition of Emotions in Gait Patterns by Means of Artificial Neural Nets. *Journal of Nonverbal Behaviour, 32*, 79-92.

Lang, P.J., Bradley, M.M. & Cuthbert, B.N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. Technical Report A-8. University of Florida

Rabiner, L. (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE, 77*, 257–286.

Robinson, P. & Kaliouby, Rana el (2009). Computation of emotions in man and machines. *Phil. Trans. R. Soc. B, 364*, 3441-3447.

Sokolowski, K. (2008). Emotion. In Müsseler, J. (Hrsg.), *Allgemeine Psychologie* (2. Aufl.) (S. 294-333). Berlin: Springer.

Westermann, R., Spies, K., Stahl, G. & Hesse, F.W. (1996). Relative effectiveness and validity of mood induction procedures: a meta-analysis. *European Journal of Social Psychology, 26*, 557-580.