

Neue Methoden der Interaktion zwischen Mensch und Maschine

Arvid KAPPAS

*Department of Psychology and Methods,
Jacobs University Bremen, Campus Ring 1, D-28759 Bremen*

Kurzfassung: Die Herausforderung, in künstlichen Systemen soziale Intelligenz zu implementieren, wird von neuen Forschungsfeldern wie Social Robotics oder Affective Computing angenommen. Neue Entwicklungen zeigen, wie die Interaktion von Mensch und Maschine durch die Einbindung von verbalem und nonverbalem Verhalten natürlicher und möglicherweise auch effektiver sein kann. Dieser Beitrag umreißt kurz wichtige Aspekte von Kommunikation und Interaktion im Allgemeinen und diskutiert generelle Ideen für Produktionsumgebungen, die auch auf den Flugzeugbau zutreffen. Es wird aufgerufen, sich an interdisziplinären Entwicklungsarbeiten zu beteiligen, um die Möglichkeiten für den Flugzeugbau einzuschätzen und dann entsprechend zu implementieren.

Schlüsselwörter: Kommunikation, Mensch-Maschinen Interaktion, Nonverbale Kommunikation, Affective Computing, Social Robotics, Emotionen

1. Menschliche Kommunikation und Kommunikation zwischen Menschen und Maschinen

Im Laufe der letzten fünfzig Jahre hat sich die Interaktion zwischen Mensch und Maschine, im Sinne des Menschen als Bedieners und der Maschine als des Objektes, welches bedient wird, deutlich in die Richtung von *Kommunikation und Interaktion* zwischen Mensch und Maschine verschoben. Zum einen gibt es Zustände von Maschinen, die offensichtlich sind. Zum Beispiel kann es sicht- und/oder hörbar sein, dass eine Maschine an- oder ausgeschaltet ist. Eine Schranke kann zu oder auf sein; ein Verschluss kann ver- oder entriegelt sein; die Position eines Objektes in einem Raum kann offensichtlich sein. Zum anderen gibt es Zustände von Maschinen oder Prozessen, die nicht direkt ersichtlich sind und die daher durch eine Vielzahl von Anzeigen dargestellt werden, die von abstrakten Symbolen, z.B. durch unterschiedlich farbige Leuchten oder simple Töne, über analoge Displays, die bestimmte Kennwerte darstellen, bis hin zu textlicher Kommunikation auf Bildschirmen reichen. Die Form solcher Kommunikation ist zum einen eine Funktion technischer Rahmenbedingungen bzw. Beschränkungen, als auch der Versuch Fehler in der Kommunikation zu verhindern. So sollte z.B. die Redundanz bestimmter Zustandsmitteilungen erhöht werden, wenn Umweltbedingungen schwierig sind, z.B. laute oder wechselhaft beleuchtete Arbeitsumgebungen. Umgekehrt kommuniziert der Mensch mit der Maschine mit Knöpfen, Schaltern, Hebeln, Rädern, etc. wenn es nur wenige Freiheitsgrade der Kontrolle oder Kommunikation mit Maschinen gibt, oder mit Keyboards unterschiedlicher Komplexität im Falle vieler Freiheitsgrade. Oft aber nicht immer liegt hier die Initiative beim Menschen; es ist aber auch möglich, dass eine Maschine eine Bestätigung anfordert z.B. dass eine bestimmte Information wahrgenommen wurde.

Die Interaktion zwischen Menschen sieht deutlich anders aus. Sprache ist sicher ein Medium, das hier sehr im Vordergrund steht, aber darüber hinaus ist verbale Kommunikation in den meisten Fällen auch eingebettet in eine reiche Umgebung *nonverbaler Kommunikation*. Kommunikationswissenschaftler, Psychologen, Biologen und andere versuchen seit langem die Komplexität menschlicher Kommunikation und Interaktion zu beschreiben und zu verstehen. Im einfachsten Fall lässt sich ein solcher Prozess durch das klassische Kommunikationsmodell von Shannon (1948) darstellen (siehe Abb. 1). Hier gibt es einen Sender und einen Empfänger; im Zentrum steht die Nachricht, die übermittelt wird.

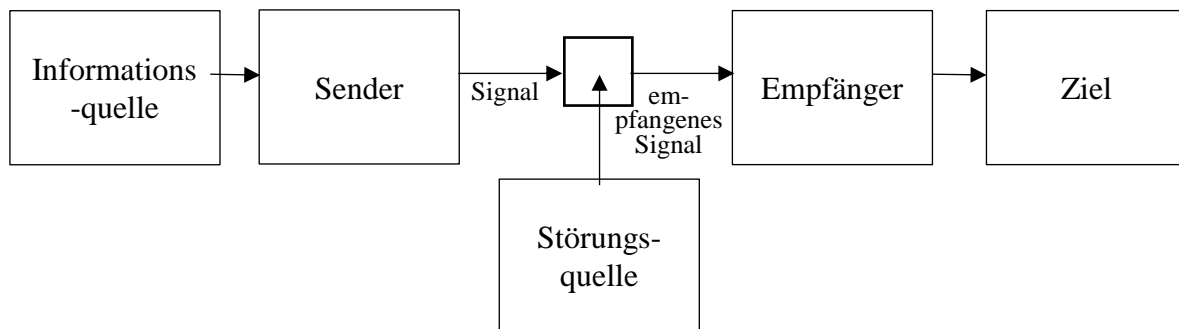


Abbildung 1: Kommunikationsmodell nach Shannon (1948).

In der tatsächlichen Kommunikationssituation zwischen Menschen werden allerdings in einem Voll duplex-Modus Nachrichten über viele Kanäle zwischen den Teilnehmern gleichzeitig gesendet. Die Signale sind teils redundant, teils unterschiedlich, manchmal auch widersprüchlich. Nicht alle gesendeten Informationen werden auch wahrgenommen und umgekehrt, glauben Menschen Signale empfangen zu haben, die nie gesendet wurden, sondern die Konsequenz von Erwartungen oder Fehlwahrnehmungen sind. Da beide (oder mehrere) Teilnehmer zur gleichen Zeit senden und empfangen, sind sie Teil eines dynamischen Systems, das durch Imitation oder andere Echtzeitprozesse den oder die Teilnehmer in Echtzeit beeinflusst und reguliert. Da es Rückkopplungsschleifen auch innerhalb von Personen gibt (z.B. durch den Einfluss von Ausdruck oder Haltung auf interne Zustände), geht es bei der menschlichen Kommunikation um ineinander verschachtelte Rückkopplungsschleifen (siehe Abb. 2). Zieht zum Beispiel ein Zuhörer die Augenbrauen zusammen, so kann der Sender Inhalt, aber auch die Form der Kommunikation dynamisch verändern – z.B. weiter ausholen, Wiederholungen einbauen, langsamer sprechen, lauter sprechen, Rückfragen integrieren.

In den letzten Jahren hat ein Interesse an Computersystemen mit sozialer Intelligenz zugenommen. Diese Systeme können Ausdrucksverhalten wahrnehmen und verfügen oft über die Möglichkeit auch Ausdrücke in verschiedenen Kanälen darzustellen. Zwei neue Forschungsfelder sind hier besonders relevant: *Social Robotics* und *Affective Computing* (Picard, 1997). Während der Schwerpunkt bei ersterem auf der sozialen Interaktion liegt, geht es im zweiten um emotionale Ausdrücke und andere Prozesse, da Emotionen eine besondere Rolle für nonverbales Verhalten spielen (siehe Calvo, D’Mello, Gratch, & Kappas, 2015). Diese Felder sind mittlerweile so weit gereift, dass sie eigene wissenschaftliche Organisationen und Verbände besitzen. Es gibt dedizierte wissenschaftliche Zeitschriften und Konferenzen. Allerdings spielen diese Entwicklungen bisher keine

große Rolle in Produktionsanlagen oder anderen industriellen Prozessen, wo Roboter in verschiedenen Formen nicht nur sichtbar, sondern normal sind. Dies ist ein geeigneter Zeitpunkt zu evaluieren, ob es hier neue Möglichkeiten gibt, die die Effektivität/Produktivität und Sicherheit verbessern können und gleichzeitig das Interface zwischen Menschen und künstlichen Systemen generell verbessern. Im Folgenden möchte ich zunächst ein Fallbeispiel aus einem anderen Kontext aus meiner eigenen Arbeitsgruppe – das EMOTE Projekt – beschreiben, um zu zeigen wie *Affective Computing* Interaktionen zwischen Mensch und Maschine unterstützen kann.

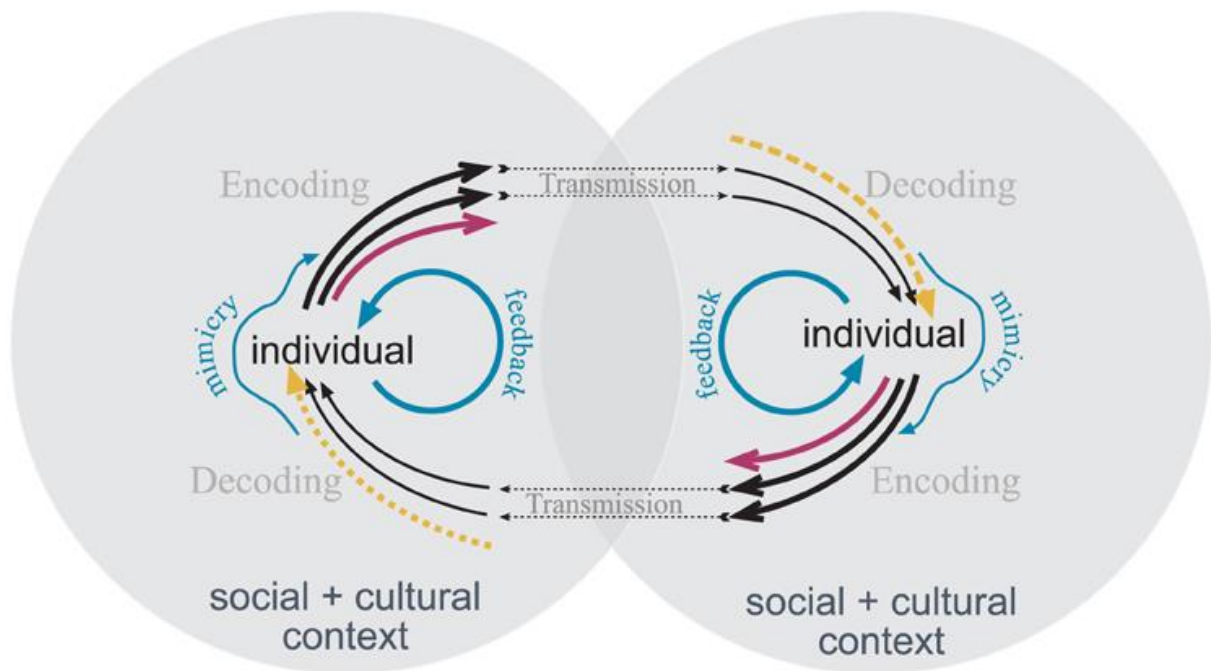


Abbildung 2: Das Superlinsenmodell der Kommunikation von Kappas (2013). Das Modell beinhaltet unterschiedliche Kanäle, inklusive Kanäle, die nicht empfangen werden (können), sowie synchrone Interaktionen und interne Feedbackschleifen.

2. Das EMOTE-Projekt

EMOTE war ein von der EU unterstütztes Projekt, welches im Rahmen des siebten Rahmenprogramms gefördert wurde und von 2012 bis 2016 lief. Es ging hierbei um die Verwendung von Informations- und Kommunikationstechnologie in einem schulischen Kontext. Das Projekt involvierte Partner aus Portugal, England, Schottland, Schweden und Deutschland. Es ging um die Entwicklung empathischer Roboter zur Verwendung als Tutor für Kinder in der Schule. Zwei Anwendungen wurden entwickelt, in denen ein NAO Roboter (Aldebaran/Softbank) mit einem oder zwei Kindern Konzepte aus der Erdkunde übt.

Es gibt seit den achtziger Jahren ein großes Interesse an computerunterstütztem Lehren und Lernen, aber hier werden Schlussfolgerungen über den Wissensstand der Schüler vor allem über deren Fehler gezogen. Wenn ich z.B. eintippe, dass $8 \times 7 = 55$ ist, dann würde man vermuten, dass ich die Multiplikation nicht beherrsche oder mich vertippt habe. Was allerdings *Affective Computing* ermöglicht, ist einzu-

schätzen, dass es ein Problem gibt, *bevor* der Fehler gemacht wurde. Dies kann erreicht werden, indem Ausdrucksverhalten und einige physiologische Parameter zu einer Einschätzung des emotionalen Zustandes herangezogen werden und in Bezug gebracht werden mit anderem Verhalten, das möglicherweise mit Leistung und Verständnis in Bezug steht. Die Verwendung physiologischer Veränderungen hat eine gewisse Tradition mittlerweile bei der Einschätzung von Überlastung/Stress. Die Kombination von Gesichtsbewegungen, z.B. dem Schließen der Augen in beim Fahren von Autos wird bereits in Produktionsmodellen verschiedener Hersteller zum Einschätzen von Müdigkeit verwendet. Der Unterschied zu einer Anwendung wie EMOTE ist, dass hier die emotionalen Ausdrücke verwendet werden, um über den Umweg der Emotion Rückschlüsse auf das Verständnis zu nehmen. Nur um dies zu unterstreichen: es geht hier nicht nur um unterschiedliche Grade von Belastung, oder Überlastung, sondern auch ob etwas zu leicht ist, oder Spaß macht weil das Niveau richtig ist.



Abbildung 3: Zwei Kinder in Interaktion mit einem NAO-Roboter und einem Touch Table im Kontext des EMOTE-Projektes (Bild: EMOTE).

Wichtig für das Verständnis des Ansatzes ist, dass der Roboter in EMOTE nur das Front-end eines Systems ist. Außer dem Roboter beinhaltet dieses einen Touchtable mit dem Spiel (in Abb. 3. ging es um gemeinsames Lernen in einem Spiel zur nachhaltigen Entwicklung), mehrere Kameras, Infrarotsensor (Kinect), sowie einen portablen Sensor für Hautleitfähigkeit und Bewegung der Kinder. Aus dem Status des Spiels, dem Verhalten der Kinder inklusive deren nonverbaalem Verhalten und einfacher physiologischer Veränderungen kann so ein internes Modell der Lerner berechnet werden, das Wissen, Verhalten und emotionalen Zustand in Verbindung setzt.

EMOTE wurde in drei Ländern getestet. Verschiedene Varianten, die auf unterschiedliche Roboter, virtuelle versus physische Roboter oder Langzeiteffekte abzielten, zeigen, dass die von uns propagierten Systeme im Prinzip möglich sind. De facto sind die preiswerten Plastikroboter, die wir verwendet haben, noch nicht robust genug, um einer Klasse von Kindern zu widerstehen, aber es ging auch um

die Plausibilität solcher Systeme. Das Team an der Jacobs University befasste sich hauptsächlich mit den emotionalen Konzepten aus psychologischer Sicht, dem Vergleich der portablen Sensoren mit kontrollierten Laborumgebungen sowie der Untersuchung neuer Interaktionsmodalitäten z.B. der Entwicklung von akustischen Signalen für die Roboter. Mehr Information über EMOTE kann man auf der Projektseite finden (<http://emote-project.eu>).

3. Entwicklungsmöglichkeiten für neue Interaktionsformen zwischen Mensch und Maschine

Aus der kurzen Diskussion am Anfang dieses Beitrags geht hervor, dass die Interaktion und Kommunikation zwischen Mensch und Maschine einfach und zuverlässig sein soll. Fehlkommunikation kann zu katastrophalen Konsequenzen führen und muss dringend vermieden werden. Daher werden zum Beispiel Interaktionen mit Maschinen in natürlicher Sprache bisher kaum für kritische Prozesse verwendet, da auch hier Fehler passieren, wenn der Kontext nicht sehr eingeeengt ist. Eine Frage, die nur mit *Ja* oder *Nein* beantwortet werden kann, ist mittlerweile sehr zuverlässig mit einer Benutzeroberfläche in natürlicher Sprache zu beantworten. Eine Frage wie „was soll ich als nächstes tun?“ hingegen bietet viel Raum für Missverständnisse.

Es soll aber betont werden, dass die effektive Kommunikation zwischen Menschen zeigt, dass menschliche Systeme sehr schnell, sehr komplexe Signale vermitteln können. Hierbei sind nicht alle Prozesse willentlich. Wenige Menschen kontrollieren strategisch die Bewegungen ihrer Augenbrauen oder der Blinzelfrequenz – bzw. sind überhaupt in der Lage dies zu tun. Daher bietet die Möglichkeit auch solche unwillkürlichen Bestandteile von Kommunikation in die Interaktion zwischen Mensch und Maschine einzubinden neue Wege, die mehr Natürlichkeit einfließen lässt und damit auch potentiell dem Menschen gerechter wird. Sowohl Wahrnehmung von Ausdrucksverhalten, als auch der Ausdruck selbst sind sehr schnell und somit prinzipiell besser für kritische Interaktionen geeignet als eine Tastatur oder ein komplexes Bedienfeld. Umweltfaktoren wie laute Umgebungen sind kein prinzipieller Hinderungsgrund zur Verwendung von Sprache – visuelle Verfahren oder die direkte Messung der Aktivität von Gesichtsmuskeln können diese Herausforderungen erfolgreich umgehen (Schultz et al., 2013). Schlechte oder wechselnde Beleuchtungsverhältnisse können durch neue Kameras in unterschiedlichen Formaten und der Verwendung von Infrarotaufnahmen neutralisiert werden. Schließlich bietet die Verwendung physiologischer Sensoren, die einfach in Kleidung oder z.B. in Armbänder integriert sind, weiterreichende Möglichkeiten, die Menschen typischerweise nicht verwenden können (siehe auch Schultz et al., 2013).

Dies ist also ein guter Zeitpunkt auf der einen Seite zu erfassen, an welchen Stellen im Produktionsprozess, z.B. im Flugzeugbau, solche Informationen sinnvoll verwendet werden könnten und auf der anderen Seite zu testen, wie zuverlässig solche an menschliche Kommunikationsformen orientierte Verfahren in realen Arbeitskontexten verwendet werden könnten. Dies bedarf einer interdisziplinären Zusammenarbeit von Ingenieuren und Psychologen auf der Entwicklungsseite und Praktiker für die notwendigen Tests in der „realen Welt“. Sicher ist, dass dies kein *Science Fiction* ist und wir in der näheren Zukunft in den verschiedensten Kontexten zuhause und bei der Arbeit Robotern und anderen Systemen mit sozialer Intelligenz begegnen werden.

4. Literatur

- Calvo, R. A., D'Mello, S. K., Gratch, J. & Kappas, A. (Eds.) (2015). The Oxford Handbook of Affective Computing, New York: Oxford University Press
- Kappas A (2013) Social regulation of emotion: messy layers. *Frontiers in Psychology*, 4, 51. doi: 10.3389/fpsyg.2013.00051.
- Picard, R. (1997). *Affective Computing*. MIT Press. Cambridge.
- Schultz, T., Amma, C., Wand, M., Heger, D., & Putze, F. (2013). Biosignale-basierte Mensch-Maschine Schnittstellen. *Automatisierungstechnik*, 61, 760 – 769.
- Shannon, C.E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27, 379-423, 623-656.

Danksagung: Das EMOTE-Projekt wurde von der Europäischen Kommission European Commission gefördert (EU FP7 ICT-317923 EMOTE, EMbOdied-perceptive Tutors for Empathy-based learning).



Gesellschaft für
Arbeitswissenschaft e.V.

Fokus Mensch im Flugzeugbau

Herbstkonferenz der
Gesellschaft für Arbeitswissenschaft

Airbus Operations GmbH und
Zentrum für Angewandte Luftfahrt-
forschung (ZAL) in Hamburg

29. und 30. September 2016

GfA Press

**Dokumentation der Herbstkonferenz der Gesellschaft für Arbeitswissenschaft e.V.
vom 29. und 30. September 2016, Hamburg**

Herausgegeben von der Gesellschaft für Arbeitswissenschaft e.V.
Dortmund: GfA-Press, 2016
ISBN 978-3-936804-21-8

NE: Gesellschaft für Arbeitswissenschaft: Konferenzband

Als Manuskript zusammengestellt. Dieser Konferenzband ist nur in der Geschäftsstelle (s. u.)
erhältlich.

Alle Rechte vorbehalten.

© **GfA-Press, Dortmund**
Schriftleitung: Matthias Jäger

im Auftrag der Gesellschaft für Arbeitswissenschaft e.V.

Ohne ausdrückliche Genehmigung der Gesellschaft für Arbeitswissenschaft e.V. ist es nicht gestattet,
den Konferenzband oder Teile daraus in irgendeiner Form (durch Fotokopie, Mikrofilm oder ein
anderes Verfahren) zu vervielfältigen.

USB-Print: Marlen Manke

Screendesign und Umsetzung

© 2016 fröse multimedia, Frank Fröse

office@internetkundenservice.de · www.internetkundenservice.de