



universität
wien

MASTERARBEIT

Titel der Masterarbeit

Die Sehnsucht, das Innere des Menschen in seinem
Äußeren zu erkennen

Von der Physiognomik zur automatischen Gesichtsausdruckserkennung

Verfasser

Christoph Musik, Bakk.phil.

angestrebter akademischer Grad

Master of Arts (MA)

Wien, im September 2009

Studienkennzahl (lt. Studienblatt): 066/905

Studienrichtung (lt. Studienblatt): Masterstudium Soziologie

Betreuerin: Univ. Prof. Dr. Ulrike Felt

Danksagung

Ich möchte die Tradition der Danksagung gerne nutzen, um denjenigen menschlichen und nicht-menschlichen AkteurInnen zu danken, die mich im Laufe dieser Masterarbeit unterstützt und begleitet haben. Zunächst einmal bedanke ich mich außerordentlich bei meiner Betreuerin Univ. Prof. Dr. Ulrike Felt, die es mir ermöglicht hat, diese Arbeit in dieser Art und Weise schreiben zu können. Schon in meinem ersten Semester an der Universität habe ich sie als eine äußerst kompetente, unterstützende und motivierende Professorin erlebt. Dieser Eindruck hat sich im Laufe der Zeit verfestigt und ich kann zu diesem Zeitpunkt getrost sagen, dass ich mir keine bessere Betreuung vorstellen kann.

Daran anschließend möchte ich mich auch bei allen Mitarbeiterinnen und Mitarbeitern des Instituts für Wissenschaftsforschung und allen Teilnehmerinnen und Teilnehmern der von diesem Institut veranstalteten Internationalen Summer School in Raach am Hochgebirge danken. Dafür stellvertretend möchte ich Michael Strassnig und Astrid Mager sowie Susan Molyneux-Hodgson und Sheila Jasanoff nennen, die durch ihre hilfreichen Kommentare zu meiner Masterarbeit einen wertvollen Beitrag zum Gelingen dieser geleistet haben.

Danken möchte ich auch all jenen, die zu unterschiedlichen Zeitpunkten meine Masterarbeit kommentiert haben, mir Ratschläge, Tipps oder Verbesserungsvorschläge gegeben haben oder sie Korrektur gelesen haben. Hierbei möchte ich meine StudienkollegInnen Doris Harrasser, Kay Felder, Clemens Huber sowie Stephan Gruber und Daniel Messner erwähnen.

Einen entscheidenden Beitrag zu dieser Masterarbeit hat auch ein ganz bestimmter Akteur geleistet, mit dem ich wohl die meiste Zeit beim Schreiben verbracht habe. Vielen Dank für die treuen Dienste will ich meinem *TravelMate 290* aussprechen, den ich mit diesen Worten in seinen verdienten Ruhe(zu)stand schicken werde.

Zu guter Letzt möchte ich meinen FreundInnen und meiner Familie danken, die wohl alle selbst am besten wissen, was sie in welcher Weise beigetragen haben. Schwärmerisch und aus ganzem Herzen gilt mein besonderer Dank Helene A. Schlöglmann, die sich in vielfacher Weise verdient gemacht hat. Ich kann es kaum in Worte fassen, doch diese Mischung aus scharfer konstruktiver Kritik und reflektierenden Momenten zusammen mit emotionaler Nähe und leidenschaftlicher Aufopferung ist einfach einzigartig! Vielen Dank!

Inhaltsverzeichnis

DANKSAGUNG	3
INHALTSVERZEICHNIS	5
ABBILDUNGSVERZEICHNIS	7
EINLEITUNG	9
Erkenntnisinteresse und Fragen	10
Aufbau der Arbeit	12
ERSTER TEIL: HISTORISCHE ANALYSE	15
Lavater und die Physiognomik: Zwischen Wissenschaft und Magie	17
Omnipräsente Physiognomik	17
Johann Caspar Lavater (1741–1801) – Person und Werk	20
Die Physiognomik vor Lavater	22
Die physiognomisch-wissenschaftliche Praxis bei Lavater	23
Rezeption und Kritik an Lavaters Physiognomik	26
Von Lavater zu Darwin	31
Auf den Spuren der Physiognomik Lavaters: Franz Josphe Gall und Carl Gustav Carus	32
Der <i>mimic turn</i> : Lichtenberg, Engel, Bell und Piderit	34
Angekommen bei der Mimik: Charles Darwin (1809 – 1882)	36
Paul Ekman und Emotionen: The way to change the world is research	43
Paul Ekman (geb. 1934): Person und Werk	43
Die Frage der Universalität von Gesichtsausdrücken	45
Von Basisemotionen und Darbietungsregeln	48
Gesichtsausdrücke messen: Das Facial Action Coding System (FACS)	50
Rezeption und Kritik an Ekmans Forschung	51
ZWEITER TEIL: THEORIEN UND KONZEPTE	55
Welches Verständnis von Technologie haben wir?: Definitionen, Konzepte und Zugänge	57
Technologischer Determinismus	58
Social Shaping and Social Construction of Technology	59
Akteur – Netzwerk Theorie und die pragmatistische Technik- und Sozialtheorie	61
Wie wird Wissen produziert? Konzepte und Zugänge	65
Was ist technowissenschaftliches Wissen? Grenzziehungen, Objektivität, Expertise und Quantifizierung	66
Die besondere Rolle von technowissenschaftlichen Bildern	68
Implizites und nicht-explizites Wissen	70

DRITTER TEIL: EMPIRISCHE ANALYSE DER AUTOMATISCHEN GESICHTSAUSDRUCKSERKENNUNG	73
Methode, Feldzugang, Auswertung	79
Das ExpertInneninterview	80
Interviewleitfaden	83
Feldzugang	83
Auswertung.....	84
Kontextualisierung: Anwendungen zwischen Mensch-Maschine Interaktion und Surveillance 87	
Mensch – Maschine Interaktion.....	88
Surveillance	90
Sonstige Anwendungsfelder	92
Forschungsaktivitäten	95
Auf dem Weg zum Maschinensehen: Eine Geschichte von Standardisierungen und Reduktionen	99
Biologischer Körper: Gesicht	100
Universeller Anspruch nicht haltbar	101
Fehlende Einbindung von Kontextinformationen	103
Bearbeitung von Bildern (Normalisierung)	104
Die Grundwahrheit (ground truth) - Eine umstrittene Tatsache	107
Klassifikation der Kategorien	111
Repräsentation in Zahlen	112
Filtern und Glätten.....	114
CONCLUSIO.....	117
Warum die Sehnsucht, das Innere des Menschen in seinem Äußeren zu erkennen, eine Sehnsucht bleibt und trotzdem in Erfüllung geht	120
LITERATUR.....	123
Internetquellen	133
ZUSAMMENFASSUNG / ABSTRACT	135
Zusammenfassung auf Deutsch	135
Abstract English	136
ANHANG	137
Verzeichnis der geführten Interviews.....	137
Interviewleitfaden	137
LEBENS LAUF	141

Abbildungsverzeichnis

Abbildung 1:	J. C. Lavater, Stirnmessgerät.....	25
Abbildung 2:	Anonym, Vermessungen an der Silhouette Lavaters	26
Abbildung 3:	Darwin auf der englischen Zehnpfundnote.....	37
Abbildung 4:	Spott und Missachtung.....	39
Abbildung 5:	Enttäuschter und mürrischer Schimpanse	40
Abbildung 6:	Schrecken	41
Abbildung 7:	Ekel dargestellt von einem South Fore	47
Abbildung 8:	Echtzeit Gesichtsausdruckserkennung	87
Abbildung 9:	Bildschirmbild der Software FaceReader™	90
Abbildung 10:	Bearbeitetes Gesichtsbild mit Raster	106
Abbildung 11:	Fließende Geschlechtsidentität.....	116

Einleitung

In der Nacht des 2. Novembers 2007 wurde im italienischen Perugia die englische Austauschstudentin Meredith Kercher in ihrer eigenen Wohngemeinschaft brutal ermordet. Sie wurde vermutlich sexuell missbraucht, gewürgt, mit Messern traktiert und schließlich erstochen. Wegen Mordes angeklagt ist Merediths Mitbewohnerin, die zum Tatzeitpunkt 21-jährige US Amerikanerin Amanda Knox, welche in den Medien als „Engelsgesicht“ bezeichnet und so weltweit bekannt wurde. Wenn die deutsche *BILD Zeitung* fragt¹, ob die junge Frau mit dem Engelsgesicht zu einer solchen Bluttat im beschaulichen Perugia fähig ist, so ist das doch ein Hinweis darauf, dass durch diesen Fall etwas in Unruhe geraten ist. Wir glauben zu wissen, dass eine Person mit einem Engelsgesicht wie Amanda Knox nicht in der Lage sein kann, so eine schreckliche Tat zu vollziehen. Wieso wird hier aber überhaupt von einem Engelsgesicht gesprochen, was ist es, was die junge Frau zu einem Engel macht, obwohl sie vermutlich einen Mord begangen hat? Sagt ihr „Äußeres“ zunächst mehr aus als ihr „Inneres“, wird das „Innere“ etwa durch das „Äußere“ bestimmt? Im Äußeren des Menschen, insbesondere in seinem Gesicht, scheinen sich jedenfalls eine Vielzahl an Informationen über sein Inneres, etwa über seinen Charakter, sein Wesen oder auch seine Gefühle zu verbergen. Auch die Wissenschaften scheinen sich seit langer Zeit danach zu sehnen, die vielen Informationen und Codes des Gesichts lesbar zu machen, sie zu erkennen und zu dechiffrieren. So versuchte im 18.Jahrhundert insbesondere Johann Caspar Lavater die Physiognomik als Wissenschaft zu etablieren, indem er von den Umrissen eines menschlichen Gesichts auf seinen Charakter schließen wollte.

Ein relativ neues Forschungsgebiet in dieser Hinsicht sind die Arbeiten, die sich im Umfeld von automatischer Erkennung von Gesichtsausdrücken bewegen. Mit Hilfe der Mimik erkennt eine Software in Echtzeit, wie sich eine Person gerade fühlt; ein spezifischer Gesichtsausdruck wird einer der Grundemotionen zugeordnet. Anhand des Äußeren des Menschen, seiner sich ständig verändernden Mimik, soll auf sein Inneres, seine Gefühle geschlossen werden. So könnte die Fantasie George Orwell's aus seinem Roman *1984* bald schon Realität werden, wenn er schreibt: „Er hatte die ruhige optimistische Miene aufgesetzt, die zur Schau zu tragen ratsam war, wenn man dem Televisor das Gesicht zukehrte“ (Orwell 1949: 8).

1 http://www.bild.de/BILD/news/vermishtes/2009/01/15/das-engelsgesicht-und-der-sex-mord/der-anklaeger-steht-selber-wegen-amtsmissbrauch-vor_20gericht.html [verfügbar am 07.07.09]

Wenn es nach außen hin einfach und alltäglich scheint, einen Gesichtsausdruck richtig deuten zu können, so stecken hinter den Technologien, die einen Gesichtsausdruck automatisch erkennen sollen, hochkomplexe technische Prozesse, die einer *black box* gleichen, deren Inhalt für den Anwender oder die Anwenderin der Technologie nicht eingesehen und nachvollzogen werden kann. Dabei lassen sich beim Öffnen der *black box* eine Vielzahl an Transformationen und Übersetzungen erkennen, welche in Form von Standardisierungen, Normalisierungen, Reduktionen und Interpretationen in die nach außen hin funktionierende Technologie einfließen. Zieht man in Betracht, dass es zu einer zunehmenden Verflechtung von menschlichem Verhalten und technischen Anwendungen in *sozio-technischen Konstellationen* kommt, so muss man sich fragen, was es bedeutet, wenn plötzlich eine Maschine (z.B. ein Fahrkartenautomat oder ein Computer am Arbeitsplatz) erkennen kann, wie sich der vis-a-vis befindende Mensch fühlt und gegebenenfalls darauf reagiert. Die Beziehung von Mensch und Maschine wandelt sich grundlegend, vor allem weil sich die Interaktivität zwischen den beiden verändert. Maschinen können mehr und mehr autonom handeln und erreichen dadurch einen weit höheren Grad an Handlungsverantwortung (Rammert 2007: 79). Es ist also nicht länger dem Menschen vorbehalten, Urteile über das Gefühlsleben von Mitmenschen treffen zu können, indem man dem anderen in sein Gesicht blickt, zunehmend werden auch unsere Maschinen diese Fähigkeit besitzen. An diesem Punkt stellt sich die grundlegende Frage, wie Maschinen, Computer, Agenten, Roboter und weitere technische Anwendungen und Systeme überhaupt sehen und wahrnehmen lernen, wie lernen sie, ihre Umwelt zu erkennen und einzuschätzen. Die Lehrtätigkeit bzw. die Entwicklung dieser Fähigkeit des Sehens, Wahrnehmens und Erkennens ist ein hochkomplizierter sozio-technischer Prozess, der sorgfältig untersucht und reflektiert werden sollte.

Erkenntnisinteresse und Fragen

Mit dieser Masterarbeit möchte ich mich langsam der großen Frage annähern, wie unseren Maschinen und Robotern das Sehen und Erkennen von sowohl Objekten, als auch von Menschen, ihren Bewegungen, Interaktionen und Gefühlen beigebracht wird. Es genügt nicht, einen Computer mit einer Kamera auszustatten, vielmehr müssen Algorithmen geschaffen werden, die die Bilder, die der Computer über die Kamera erhält, mit Bedeutung anfüllen. Ich will mich dieser großen Frage und Aufgabe nähern, indem ich einen genaueren Blick auf gegenwärtig entstehende Technologien zur Gesichtsausdruckserkennung, also auf eine spezielle Anwendung im Bereich des Maschinensehens werfe. Ich verstehe insofern meine Masterarbeit als explorative Studie, die ich als eine Geschichte konzeptualisiert habe,

welche von der Sehnsucht erzählt, das Innere des Menschen in seinem Äußeren zu erkennen.

An dieser Stelle sollte angemerkt werden, dass die Technologien zur Gesichtsausdruckserkennung und ihre Entstehung nicht mehr als selbstverständlich angesehen werden können, sondern es muss hinterfragt werden:

Welches implizite Modell vom Sehen und Erkennen steckt in den Technologien?

Die theoretischen Annahmen hinter dieser Frage sind, dass eine Technologie von einer großen Zahl an sozialen Faktoren beeinflusst ist, aber noch viel mehr, dass sich eine Technologie in Ko-Produktion mit der Gesellschaft in einem *nahtlosen Netz* von unterschiedlichen AkteurInnen, und damit sind menschliche als auch nicht-menschliche Entitäten gemeint², entwickelt.

Wie sind also alltägliche, kulturell geschliffene Annahmen und Konzepte des Sehens mit den wissenschaftlichen und technischen Anwendungen des Maschinensehens verwoben?

Was passiert, wenn alltägliches, visuelles Wissen über Gesichtsausdrücke und Gefühle so transformiert und übersetzt wird, dass es für eine Maschine anwendbar wird?

Welchen Standardisierungen und Reduktionen unterliegt das Wissen, um in die Maschine integriert werden zu können und welche Auswirkungen hat dies wiederum auf das Wissen?

In meiner Arbeit geht es insbesondere darum, nachvollziehen zu können, welches Wissen in die Anwendungen und Technologien zur automatischen Gesichtsausdruckserkennung einfließt, wie dieses Wissen verarbeitet, transformiert und übersetzt wird, sodass es letztendlich in der technischen Anwendung funktioniert. Ich möchte sozusagen die *black box* der automatischen Gesichtsausdruckserkennung öffnen und beleuchten, welche Transformationen und Übersetzungen von unterschiedlichen AkteurInnen in die *black box* eingeflossen sind und weiterhin einfließen. Ich begeben mich aus diesem Grund auf die Spur der AkteurInnen, dorthin wo die Technologie hergestellt wird, in die Labore der ComputerwissenschaftlerInnen. Die Entwickler und Entwicklerinnen einer Technologie können nach Akrich als Vermittler und Vermittlerinnen begriffen werden, welche die Bindeglieder zur Benutzung und Anwendung darstellen (Akrich 2006: 415). Neben der Lektüre der wichtigsten Arbeiten zur automatischen Gesichtsausdruckserkennung erhalte ich insbesondere durch ExpertInneninterviews mit den Entwicklern und Entwicklerinnen der Technologien grundlegende detaillierte Informationen über die Konstruktion von Technologien zur Gesichtsausdruckserkennung und somit Einblicke in die *black box* der Technologie.

² Michel Callon definiert den Begriff des Akteurs als „jede Entität, die fähig ist, Texte, Menschen, Nicht-Menschen und Geld zu assoziieren“ (Callon 2006: 318)

In dieser Masterarbeit steht weniger eine *ready made technology* im Zentrum, also eine Technologie, die fix und fertig nach festgelegtem und vorhersehbarem Verhalten funktioniert (Schulz-Schaeffer 2000a:139), sondern vielmehr *technology in the making*. Das bedeutet, die *Entstehung der Technik* soll nachvollzogen werden, um beobachten zu können, wie sie hergestellt wird, welche Signale und Zeichen von der Maschine in welcher Weise erkannt und gedeutet werden, also alles das, was später in der *black box* der Technik unsichtbar wird. Auf was genau achtet die Maschine, wenn es darum geht, Gefühle zu erkennen, was fließt in die Maschine in welcher Weise ein, was sind die kulturellen Annahmen, die dahinter stecken? In einem breiteren Kontext soll diese Fokussierung auf die Mikroebene Aufschluss darüber geben, wie die Vielfältigkeit und Komplexität der menschlichen Gesichtsausdrücke und Emotionen standardisiert und reduziert wird, wenn Maschinen mit der Fähigkeit ausgestattet werden, Emotionen zu erkennen. Die daraus entstehenden Auswirkungen auf die Makroebene werden ebenfalls thematisiert. Ich möchte also den Prozess der Standardisierung von Gesichtsausdrücken und Emotionen verfolgen, welcher aus gewissen Gründen in Kauf genommen wird, wenn daran gearbeitet wird, Emotionen automatisch erkennen zu können.

Aufbau der Arbeit

Im ersten Teil der Arbeit begeben sich auf die Spur von historischen Vorläufern der Technologien zur automatischen Gesichtsausdruckserkennung. Ich spanne dabei den Bogen bis zurück in das dritte Jahrhundert vor Christus, wenn ich auf die Physiognomik vor Lavater eingehe. Lavater selbst kann als der Begründer der wissenschaftlichen Auseinandersetzung mit der Physiognomik angesehen werden, als er im späten 18. Jahrhundert versuchte, den Charakter eines Menschen in seinen Gesichtszügen wissenschaftlich fundiert erkennen zu können. Über Darwin lande ich bei Paul Ekman, dessen Forschungsarbeiten zu Gesichtsausdrücken in der zweiten Hälfte des 20. Jahrhunderts als stärkster Referenzpunkt der zeitgenössischen Technologien zur Gesichtsausdruckserkennung angesehen werden können.

Im zweiten Teil der Arbeit versuche ich meinen theoretischen und konzeptionellen Rahmen darzustellen, indem ich zuerst der Frage nachgehe, welches Verständnis von Technologien wir eigentlich haben und anschließend für mich entscheidende Konzepte und Zugänge zu technowissenschaftlichem Wissen anführe.

Im dritten empirischen Teil setze ich die Geschichte der Sehnsucht, das Innere des Menschen in seinem Äußeren zu erkennen fort, indem ich beschreibe, wie die Vielfältigkeit und Komplexität der menschlichen Gesichtsausdrücke, Emotionen und anderer Signale

standardisiert und reduziert werden, wenn Maschinen mit der Fähigkeit ausgestattet werden, Gesichtsausdrücke und Emotionen automatisch erkennen zu können. Bevor ich genauer auf diese angesprochenen Standardisierungs- und Reduktionsprozesse von Emotionen eingehe, erläutere ich in diesem dritten Kapitel zuerst meine methodische Arbeitsweise, meinen Feldzugang und meine Auswertungsstrategie, um die empirische Vorgehensweise nachvollziehbar zu machen. Daran anschließend stelle ich die Technologien zur automatischen Gesichtsausdruckserkennung in den Kontext ihrer Anwendungen, die sich zwischen Mensch-Maschine Interaktionen und Überwachung bewegen und werfe einen Blick auf die Erwartungen an diese Technologien. Anschließend gebe ich noch einen Überblick darüber, seit wann an welchen Plätzen zu diesem Thema geforscht wird, vor allem um zu zeigen, dass es ein Thema ist, das vor allem in den letzten 20 Jahren große Aufmerksamkeit erfahren hat.

Erster Teil:

Historische Analyse

In der Geschichte der Wissenschaften wurde immer wieder versucht, den Menschen „lesbar“ zu machen. Man konzentrierte sich insbesondere auf den menschlichen Körper und erhoffte sich von verschiedenen Körpervermessungstechniken Aufschluss über dahinter liegende Sinnwelten und Wahrheiten zu bekommen. Die Sehnsucht der Wissenschaften, das Innere des Menschen gerade in seinem Äußeren erkennen zu können, ist in unserer Geschichte stark verankert. Eine historische Analyse von bedeutenden wissenschaftlichen Entwicklungen, die sich darauf konzentrierten, über das Gesicht des Menschen Aufschluss über sein Inneres—sei es seinen Charakter oder sein Befinden—zu erhalten, soll vor allem zeigen, dass diese Art von Wissenschaft eine lange Tradition hat. Sie soll aber auch als analytischer Startpunkt für eine Auseinandersetzung mit zeitgenössischen Technologien zur Gesichtsausdruckserkennung dienen. Die historische Analyse soll keine vollständige Geschichte der Wissenschaften des Gesichts sein, sie soll vielmehr einige markante Punkte in der Wissenschaftsgeschichte aufzeigen. Dennoch denke ich, dass sich selbst in dieser verhältnismäßig kurzen Abhandlung ein Zusammenhang der einzelnen wissenschaftlichen Aktivitäten herstellen und erkennen lässt und dass sich daraus eine Geschichte ergibt, die einem gewissen roten Faden folgt. In diesem Streifzug durch die Geschichte soll darüber hinaus versucht werden, einer zentralen Frage meiner Arbeit näher zu kommen, nämlich wie alltägliche, kulturell geschliffene Annahmen und Konzepte mit wissenschaftlichen Erkenntnissen verwoben waren und sind? Meine Quellen für den historischen Abschnitt speisen sich insbesondere aus Sekundärliteratur, aber auch aus Primärliteratur. Zusätzlich habe ich ein ExpertInneninterview geführt, welches mir Einblicke in die physiognomische Praxis bei Lavater gab.

Ich beginne die kleine Geschichte der Sehnsucht der Wissenschaften, das Innere des Menschen in seinem Äußeren zu erkennen im späten 18. Jahrhundert mit dem Schweizer Theologen Johann Caspar Lavater, welcher versuchte, die Physiognomik als solche wissenschaftlich zu etablieren. Lavater war der festen Überzeugung, dass man den Charakter eines Menschen insbesondere an den festen Teilen seines Gesichts erkennen kann. Im Laufe der folgenden 100 Jahre nahm man, das Gesicht betreffend, zunehmend Abstand von den festen Teilen des Gesichts und konzentrierte sich mehr auf die Pathognomik, also die weichen Teile des Gesichts sowie auf die Mimik. In diesem

Zusammenhang werde ich den Weg von Lavaters Physiognomik hin zur Ausdruckslehre von Charles Darwin in Umrissen nachzeichnen. Darwin veröffentlichte 1872 sein Buch *Der Ausdruck der Gemütsbewegungen bei dem Menschen und den Tieren*. Weniger als 100 Jahre später war es dann der amerikanische Psychologe Paul Ekman, der die Theorien Darwins in den 1960er Jahren mit ins 20. Jahrhundert aufnahm und seitdem Forschung zu Gesichtsausdrücken und Emotionen betreibt.

Lavater und die Physiognomik: Zwischen Wissenschaft und Magie

Omnipräsente Physiognomik

Wer oder was ist Physiognomik? Irgendwie wissen wir es, irgendwie auch nicht. Deshalb zunächst ein Blick in das Lexikon unserer Zeit, *Wikipedia*: Das Wort Physiognomik kommt aus dem Griechischen und setzt sich zusammen aus den Worten φυσίς (physis) = Körper und γνῶμη (gnome) = Wissen, Lehre. Demzufolge ist die Physiognomik das Wissen und die Lehre über den Körper. Gleich anschließend erfahren wir in *Wikipedia*, dass Physiognomik „die „Kunst“ [ist], aus dem unveränderlichen physiologischen Äußeren des Körpers, besonders des Gesichts, auf die seelischen Eigenschaften eines Menschen zu schließen.“³ An anderer Stelle in *Wikipedia* erfahren wir auch, dass Physiognomik der „Versuch [ist], methodisch aus der körperlichen Erscheinung eines Menschen zu lesen, ...“.⁴ Ich will wissen, ob es noch andere Auffassungen darüber gibt und hole aus meinem Bücherregal das *Universal Lexikon* heraus. Dort heißt es zu Beginn: „Physiognomik, Lehre vom menschlichen Gesichtsausdruck im Ruhezustand (Physiognomie), im Gegensatz zum belebten Gesichtsausdruck (Mimik)“. (vgl. Faktum Lexikon Institut: 2000). Das reicht mir noch nicht, ich will mir auch einen fremdsprachigen Eindruck davon verschaffen, was Physiognomik ist. Die Online Ausgabe der *Encyclopaedia Britannica* (www.britannica.com) sagt folgendes über physiognomy⁵: „the study of the systematic correspondence of psychological characteristics to facial features or body structure.“⁶ [Die Lehre der systematischen Übereinstimmung von seelischen, psychologischen Eigenschaften und Gesichtszügen oder dem Körperbau].

Erstaunlicherweise scheinen sich die drei eben zitierten enzyklopädischen Quellen nicht ganz einig zu sein, was denn Physiognomik jetzt wirklich ist. Wir haben zwar eine leise Ahnung, um was es ungefähr dabei gehen soll, doch einerseits setzen die drei Nachschlagewerke unterschiedliche Schwerpunkte in der Auswahl der vielfältigen Aspekte der Physiognomik und andererseits sagen sie einfach auch unterschiedliches aus. Während die beiden zitierten Stellen aus *Wikipedia* von der Physiognomik als eine „Kunst“ und „dem Versuch“ sprechen, findet man sowohl im *Universal Lexikon* als auch in der *Encyclopaedia Britannica* das Wort „Lehre“. Auch was den Gegenstand anbelangt, bezieht sich das *Universal Lexikon* gleich nur auf das menschliche Gesicht, während die anderen beiden

3 <http://de.wikipedia.org/wiki/Physiognomik> [Online verfügbar am 14.04.2009]

4 <http://de.wikipedia.org/wiki/Physiognomie> [Online verfügbar am 14.04.2009]

5 Die Übersetzung von Physiognomik in das Englische ist physiognomy, was insofern missverständlich sein kann, da das deutsche Wort Physiognomie die äußerliche Erscheinung des Menschen allgemein bezeichnet im Gegensatz zu Physiognomik, der Lehre und des Wissens über das äußerliche Erscheinungsbild des Menschen.

6 <http://www.britannica.com/EBchecked/topic/458823/physiognomy> [Online verfügbar am 14.04.2009]

Lexika das Gesicht als nur einen Teil der Fokussierung anführen. Auch was den *modus operandi* der Physiognomik anbelangt, gibt es Unterschiede: Während man in *Wikipedia* lesen kann, dass man aus dem Äußeren des Menschen auf die seelischen Eigenschaften schließt, ist bei der *Encyclopaedia Britannica* von einer systematischen Übereinstimmung von seelischen Eigenschaften und Körper die Rede. Diese gewisse Uneinigkeit der Lexika macht deutlich, dass die Physiognomik keine eindeutige Sache ist, sondern für unterschiedliche Personen in unterschiedlichen Kontexten und Situationen nicht kongruente Bedeutungen und Assoziationen hervorruft. Auch Claudia Schmölders diagnostizierte, dass die physiognomische Kenntnis eine ziemlich komplexe Angelegenheit sei, was sich wissenschaftsgeschichtlich sehr deutlich an der Vielzahl der Disziplinen zeigt, die sich physiognomisch betätigt haben und auch heute noch betätigen. In der Medizin, der Astrologie, der Persönlichkeits- und Ausdruckspsychologie, der Ethnologie, der „Rassenlehre“, der Anthropologie, in Kunst, Literatur und Theater wird physiognomisch geschlossen, und zwar von der äußeren Erscheinung auf ein „inneres“ Schicksal, sei es biographisch, physiologisch oder theatralisch (Schmölders 1997: 11). Diese zusätzliche Erweiterung der Komplexität, was Physiognomik ist, verdeutlicht, dass es eben nicht nur ein Gesicht oder ein Körper sein muss, von dem wir auf etwas Inneres schließen, es kann eben auch wie Schmölders anmerkt eine Biographie sein. Dennoch hat die Physiognomik insbesondere etwas mit dem Körper zu tun, was sich an der Beschäftigung mit dieser unter dem Deckmantel der Wissenschaft verdeutlicht. Ob Lavater's Physiognomik, Gall's Phrenologie, die Anwendung und Auslegung der Physiognomik im NS-Regime oder auch die Ausdruckspsychologie, alle diese speziellen Physiognomiken beziehen sich auf den menschlichen biologischen Körper. Alle diese eben genannten Erscheinungen der Physiognomik müssen sich auch den Vorwurf gefallen lassen, dass sie un-, pseudo- bzw. parawissenschaftlich sind. Zu diesem Streit, ob etwas wissenschaftlich oder nicht wissenschaftlich ist, werde ich später noch genauer eingehen, wenn ich mich auf das Konzept des *boundary work* bei Thomas Gieryn (1999) beziehe.

An dieser Stelle möchte ich zugleich die Aufmerksamkeit auf die Verwobenheit von Wissenschaft und Alltag lenken, denn in diesem Spannungsfeld findet sich gerade die Physiognomik in ihrer historischen Entwicklung wieder. Schmölders geht davon aus, dass, auch wenn die Physiognomik als „unwissenschaftlich“ angesehen wird, sie deshalb nicht gleich irrational wäre. Denn im Alltag hat sie ihren festen rationalen Platz wenn es darum geht mit „laienhaftem Verstand“ physiognomisch zu deuten. Nach Schmölders können wir gar nicht anders, als physiognomisch wahrzunehmen. (Schmölders 1997: 16). Sie brauchen sich einfach nur eine beliebige Interaktion der letzten Zeit vorstellen: Sie sitzen am Frühstückstisch und die Miene Ihres Sohnes verrät, dass er noch nicht ganz wach ist und

lieber noch länger geschlafen hätte, den Zeitungsverkäufer erkennen Sie ohne weitere Probleme als „Person mit Migrationshintergrund“, die Verkäuferin in der Bäckerei nehmen Sie sofort und ohne darüber weiter nachzudenken z.B. als „etwas ältere Frau“ wahr, die Ärztin beim darauffolgenden Arztbesuch als „intelligente Analytikerin“. Dieses Wahrnehmen und „quasi-intuitive“ Zuordnen von Personen in Personengruppen und von Merkmalen in Merkmalsgruppen ist aber nicht nur alleine vom Körper bestimmt, sondern auch vom Kontext der Situation, der Kleidung, der Sprache, ja ganz allgemein, von seinem Habitus. Der Körper und gerade das Gesicht spielen meiner Meinung nach aber eine nicht zu vernachlässigende Rolle. Die Alltagswelt ist eine Welt voll von Physiognomik. Gerade deshalb muss man aus einer soziologischen Betrachtungsweise die Physiognomik ernst nehmen und sie nicht nur beiläufig behandeln, wie Schmölders in Bezug auf *Die gesellschaftliche Konstruktion der Wirklichkeit* von Berger und Luckmann (vgl. Berger/Luckmann 2007 [1977]) anmerkt. Denn dort beschreiben sie die Vis-a-Vis Situation von Angesicht zu Angesicht als den „Prototyp aller gesellschaftlichen Interaktion“, von der jede andere Interaktionsform abgeleitet sei. (ebd.: 31) In diesem alltäglichen Zusammenhang erscheint die Physiognomik, ganz allgemein gesprochen als der Prozess des Schließens von einem äußeren Merkmal auf ein inneres Schicksal, und sei es noch so klein und unbedeutend, allgegenwärtig und weit weniger problematisch als es die (un-, pseudo-, para-)wissenschaftliche Beschäftigung mit der Physiognomik zulässt. Diese (un-, pseudo-, para-)wissenschaftliche Problemgeschichte der Physiognomik mag gerade im deutschen Kontext vor allem historisch mit der exzessiven Verwendung und Auslegung der Physiognomik im Nationalsozialismus assoziiert werden. Schmölders erinnert dabei an den Einsatz von Berufs-Physiognomen im NS-Regime, welche JüdInnen auch im Ausland finden sollten oder an die Propaganda, die im Zusammenhang mit der Ausstellung *Entartete Kunst* im Jahre 1937 stattfand, als im Programmheft der Ausstellung Fotos von entstellten Gestalten den Portraits von modernen KünstlerInnen gegenübergestellt wurden, um diese dadurch „biologisch zu diffamieren“ (Schmölders 1997: 12), d.h. sie grundsätzlich in ihrem Wesen zu verleumden. Auf diese Entwicklungen haben sich bereits viele Forschungsarbeiten bezogen, welche die Motive und Mechanismen der physiognomischen Vorurteilsbildung insbesondere der rassistischen Vorurteile aufgezeigt haben sowie die damit verbundene Bildung von Stereotypen („der Schwarze“, „der Jude“, „die Frau“) als solche enthüllt haben (ebd.: 8).

Der berühmteste Physiognomiker bis heute ist nach Schmölders aber im späten 18. Jahrhundert noch weit vor den nationalsozialistischen Exzessen zu finden. Dabei handelt es sich um Johann Caspar Lavater (ebd.: 29), welcher in meiner Auffassung als zumindest einer der Begründer, wenn nicht der Begründer der damals als wissenschaftlich geltenden Beschäftigung mit der Physiognomik anzusehen ist. Geht es nach der Lavater Biographin

Anne-Marie Jatón ist der „Schöpfer der Physiognomik“ (vgl. Jatón 1988). Nach dieser allgemeinen Einführung zur Physiognomik werde ich jetzt auf ebendiesen Schweizer Theologen und sein Werk genauer eingehen. Dabei stelle ich zunächst ihn selbst und sein Werk vor, gehe dann auf die Physiognomik von Lavater ein, diskutiere die Rezeption und Kritik seines Werks und gebe zuletzt einen Einblick in die physiognomisch-wissenschaftliche Praxis bei Lavater.

Johann Caspar Lavater (1741–1801) – Person und Werk

Johann Caspar Lavater wurde am 15. November 1741 in Zürich als 13. Kind des Arztes Hans Heinrich Lavater und der Mutter Regula Escher in eine gut situierte Familie geboren. Seine schulische Laufbahn führte ihn schnell auf eine der beiden Züricher Lateinschulen, bereits 1748 besuchte er als Sechsjähriger die Schola Carolina und begann dort Latein, Griechisch und Hebräisch zu erlernen. Eine erste Einführung in die Theologie erhielt er in der Zeit von 1754 bis 1756 auf dem Collegium Humanitatis, bevor er schließlich am Collegium Carolinum, dem Vorläufer der Universität Zürich, welche 1833 gegründet werden sollte, das Studium der Theologie aufnahm. Im Frühjahr 1762 legte Lavater als 20-jähriger sein Abschlussexamen und seine kirchliche Aufnahmeprüfung ab. Nach einer ausführlichen Bildungsreise in Norddeutschland bekam Lavater bei seiner Rückkehr nach Zürich wegen des zu dieser Zeit herrschenden Pfarrerüberschusses zunächst auch keine Stelle als solcher, konnte sich aber aufgrund der finanziellen Unterstützung seiner Familie als Autor und Dichter betätigen (Weigelt 1991: 7ff.). Im April 1769 wurde Lavater schließlich zum Diakon an der Waisenhauskirche in Zürich gewählt, im März 1775 erhielt er dort die Pfarrstelle (ebd.: 16). In der Zeit ab 1768 war Lavater literarisch äußerst produktiv und publizierte mehrere bedeutende theologische und philosophische Schriften. Er baute sich mit der Zeit auch ein immer größeres Netzwerk an Bekannten und Freunden auf, darunter so bekannte Zeitgenossen wie Johann Wolfgang von Goethe (1749–1832), den er 1774 zum ersten Mal persönlich antreffen konnte (ebd.: 22ff.). 1778 wechselte Lavater zur Kirche St. Peter und übernahm dort eine Stelle als Diakon. Diese Zurücknahme seiner Tätigkeit ist wohl unter anderem auf seine äußerst intensiven Kontakte, die er ausführlich zu pflegen wusste, zurückzuführen. Im Jahr 1777 kam es sogar zu einer Zusammenkunft mit Kaiser Joseph II., welcher auf seiner Heimreise von Paris nach Wien durch die Schweiz reiste (ebd.: 29). In diese Zeit fällt auch das für diese Arbeit entscheidende und wahrscheinlich bekannteste Werk von Lavater: Zwischen 1775 und 1778 veröffentlichte er die *Physiognomischen Fragmente zur Beförderung der Menschenkenntnis und Menschenliebe*. Lavater hat aber nicht nur physiognomische Schriften verfasst, sondern er hat auch in vielen anderen Bereichen publiziert. Weigelt ordnet Lavaters Werk in fünf Kategorien ein: theologische

Werke, poetische Werke, pädagogische Schriften, physiognomische und psychologische Werke sowie politisch patriotische Schriften (ebd.:74ff.). Den „europäischen Ruhm“ hat Lavater aber seinem physiognomischen Werk zu verdanken, insbesondere die schon erwähnten *Physiognomischen Fragmente* (Pestalozzi 1988: 137). Diese wurden in Leipzig und Winterthur in vier Folianten publiziert, zunächst nur auf Deutsch, dann aber auch in anderen Sprachen, was auch die Verbreitung von Lavaters Werk in ganz Europa erklärt. Neun Jahre nach Lavaters Tod, also 1810, gab es insgesamt 16 deutsche, 15 französische, zwei russische, eine holländische, eine italienische, 20 englische und sogar zwei „amerikanische“ Ausgaben der *Physiognomischen Fragmente* (Weigelt 1991: 99). Die Physiognomischen Fragmente bestanden aus vier Teilen, die zwischen 1775 und 1778 zunächst auf Deutsch publiziert wurden: Im ersten Band nahm Lavater eine physiognomische Analyse von bedeutenden Gestalten der Geschichte vor, im zweiten Band setzte er sich mit der Physiognomie der Tiere auseinander, im dritten Teil konzentrierte er sich auf einzelne Teile des menschlichen Körpers wie z.B. Hände oder Ohren und im vierten Band beschäftigte er sich insbesondere mit Christusbildern und deren physiognomischen Bedeutungen (ebd.: 98). Man kann sich, auch wenn Lavater wohl für ein breites Publikum schrieb, die *Physiognomischen Fragmente*, die sich sprachlich und stilistisch dem Sturm und Drang zuordnen lassen (ebd.: 97) nicht als ein Werk vorstellen, dass man ohne weiteres im örtlichen Buchladen kaufen konnte. Die *Physiognomischen Fragmente*, an welchen sich Goethe auch maßgeblich beteiligte, hatten eine Subskribentenliste von vermögenden Personen, die sich diese Folianten mit aufwendigen opulenten Illustrationen und Goldrücken auch leisten konnten (Schmölbers 1997: 29).

Doch auch schon vor der Veröffentlichung der Physiognomischen Fragmente erschien im Jahr 1772 eine erste Abhandlung zur Physiognomik unter dem Titel *Von der Physiognomik*. Diese veröffentlichte ein naher Freund Lavaters, der Arzt Johann Georg Zimmermann, zunächst anonym und ohne Wissen von Lavater im Hannoverischen Magazin (Pestalozzi 1988: 138). Wenn man vom Werk Lavaters spricht, muss man auf jeden Fall auch die unglaubliche Menge an Zeichnungen, die im Zuge seiner physiognomischen Tätigkeit entstanden sind, erwähnen. Diese lassen sich heute in der Sammlung Lavater, welche sich in der Österreichischen Nationalbibliothek in Wien befindet, vorfinden. Diese Sammlung Lavater „lässt sich wahrscheinlich am besten beschreiben als eine Kollektion sehr unterschiedlicher Arten von physiognomisch ausdruckskräftigen Linienzügen.“ (Interview 0) In der Sammlung lassen sich etwa 22.000 Grafiken von vor allem Gesichtern und Körpern von Tieren und Menschen finden. Etwa zwei Drittel der Sammlung sind Porträts, daneben lassen sich aber auch Karikaturen, christologische Darstellungen, unterschiedliche Formen von Tierstudien, Tableaus mit Stirnen, Nasen, Ohren und anderen physiognomischen Details

finden. Die Sammlung Lavater wurde nach dem Tod von Lavater von dessen Sohn Heinrich zunächst an den Kunstliebhaber Moritz Christian Graf von Fries nach Wien verkauft. Nach dessen Bankrott erwarb sie Kaiser Franz I. und nahm die Sammlung in die kaiserliche Bibliothek in Wien auf (Weigelt 1991: 100).

Bevor ich auf die wissenschaftliche Praxis von Lavater und auf die Reputation und Kritik seines Werkes eingehe, will ich noch einen Blick auf die Physiognomik vor Lavater werfen.

Die Physiognomik vor Lavater

Die Einsicht, dass Physiognomik schon lange vor Lavater betrieben worden ist, darf nicht die Bedeutung von Lavaters Versuch einer wissenschaftlichen Physiognomik schmälern. Diese Einsicht zeigt vielmehr, dass sich die Physiognomik aus einem alltagspragmatischen Bedürfnis der Menschen heraus entwickelt hat, unser Gegenüber—insbesondere die den Menschen fremden Tiere oder fremden Menschen—besser einschätzen zu können.

Man kann, was die überlieferten Quellen anbelangt, bis zurück zu Aristoteles blicken—also in die Zeit über drei Jahrhunderte vor Christi Geburt—, der die Beziehungen zwischen Gesicht und Charakter beobachtete (Jaton 1988: 63), aber auch bis zurück zu den vor-konfuzianischen Gesichtslesern in China (McNeill 2002: 232) um auf Physiognomik zu treffen. Die Physiognomik vor Lavater war hauptsächlich geprägt von Bezugspunkten zur Tierwelt. So schreibt etwa Adamantius, der im frühen 5. Jahrhundert nach Christi Geburt lebte: „[...] diejenigen mit kleinen Kiefern [sind] verräterisch und grausam. Die Schlangen, die ein kleines Gebiß haben, weisen all diese Laster auf.“ (zit. nach Jaton 1988: 64).

Im Mittelalter kam die Physiognomik wieder etwas aus der Mode und erfuhr in der Zeit, die wir heute als Renaissance bezeichnen, insbesondere im späten 16. Jahrhundert eine neue Blüte. In diese Zeit fällt auch das wesentliche physiognomische Werk vor Lavater: *De Humana Physiognomica* vom Neapolitaner *Giambattista della Porta* (1535–1615). In diesem Werk aus dem Jahre 1586 werden Mensch und Tier mithilfe von Stichen systematisch gegenübergestellt, die Übereinstimmung der Charakterzüge der beiden wird unmittelbar durch die vorher bearbeiteten und nahezu entstellten Bilder herausgearbeitet. Die vielen Vergleiche bei della Porta beziehen sich z.B. auf den „Unvorsichtigen“ und den Esel, „den Rohen“, „Groben“ und das Schwein oder „den Geistesgestörten“ und den Affen (zit. nach Jaton 1988: 65). Bei della Porta findet sich auch „eine Entsprechungslehre zwischen Gesichtslinien und anderen Körperteilen“ sowie in dem umfangreichen Werk *Phytognomica* (Porta 1591) eine entsprechende Lehre, um die Ähnlichkeiten zwischen Pflanzen und menschlichen Organen aufzuzeigen (Buser 1973: 25).

Im 17. Jahrhundert sollte der Physiognomik mit dem Kartesianismus ihre erste „wissenschaftliche Weihe“ zuteilwerden, der französische Maler *Charles Le Brun* wurde zu einem Theoretiker einer neuen Physiognomik (Jaton 1988: 66). Heiko Christians sieht in der Physiognomischen Geometrie von Le Brun den Beginn der Vernaturwissenschaftlichung der Physiognomik (Christians 2006: 4). In seinen geometrischen Schemata versuchte Le Brun eine erste mathematische Bewertung des Charakters: Er entwickelte die Theorie des Gesichtswinkels. Aus der Neigung der Gesichtsknochen heraus versuchte er sowohl bei Menschen als auch bei Tieren Klassifikationen abzuleiten (Jaton 1988: 68).

Die physiognomisch-wissenschaftliche Praxis bei Lavater

„Physiognomik ist die Wissenschaft, den Charakter (nicht die zufälligen Schicksale) des Menschen im weitläufigsten Verstande aus seinem Aeüßerlichen zu erkennen; ...“ (Lavater 1991 [1772]). Obwohl die Physiognomik Lavaters im heutigen Sinne als Wissenschaft keine Anerkennung findet und als „para-, pseudo- oder unwissenschaftlich“ bezeichnet wird, muss sie in einem retrospektiven Ansatz dennoch als eine physiognomisch-wissenschaftliche Praxis verstanden werden. Dies ergibt sich aus den Konsequenzen des Thomas-Theorems, einem der Grundpfeiler der Soziologie: Menschen handeln der Situation gegenüber aufgrund der Definition, die sie ihr geben. Die Folgen von etwas sind demnach wirklich, wenn Menschen Situationen als wirklich definieren (vgl. Richter 2002: 70). Schreibt also Lavater, dass es sich bei der Physiognomik um eine Wissenschaft handelt, muss man diesem Ruf folgen und seine physiognomische Praxis am Ende des 18. Jahrhunderts als wissenschaftliche Praxis zu dieser Zeit verstehen. „Der Versuch, die Lehre vom Ausdruck erstmals zu einer selbständigen Wissenschaft zu erheben, kann wohl niemand anderem als J.C. Lavater zugeschrieben werden,[...]“ (Buser 1973: 26). Genau an diesem „Versuch“, etwas zu einer Wissenschaft zu machen, kann man meiner Meinung nach gerade rückblickend sehr gut verdeutlichen, was es bedeutet und welche Kriterien verlangt werden, dass etwas als wissenschaftlich angesehen wird. Die Wissenschaftlichkeit von Lavaters Arbeit wurde schon zu dessen Lebzeiten in Frage gestellt (vgl. folgendes Kapitel *Rezeption und Kritik an Lavaters Physiognomik*). Bevor ich diese Kritik an Lavaters Arbeit anführe, möchte ich einen zumindest kleinen Einblick in die physiognomisch-wissenschaftliche Praxis bei Lavater in aller Kürze geben. Wie kann man sich also Lavaters Arbeit vorstellen, wie versuchte er wissenschaftlich zu arbeiten?

Lavaters Arbeit war vor allem eine Arbeit mit Bildern und Zeichnungen: „[...] die erste, die natürlichste, die sicherste Sprache der Physiognomik“ sei die Zeichnung (Lavater zit. nach Swoboda 2002: 39). Er arbeitete also in erster Linie nicht mit lebendigen Menschen, sondern

mit den Zeichnungen von diesen. Lavaters Physiognomik kann also als eine statische und nicht dynamische Wissenschaft angesehen werden (Jaton 1988: 68). Die Zeichnungen, die der Ausgangspunkt der Arbeit Lavaters waren, wurden größtenteils nicht von ihm selbst angefertigt. An der Produktion von Zeichnungen war eine Vielzahl an anderen Personen beteiligt. Dabei muss man zwischen externen und internen BildproduzentInnen unterscheiden, wobei die meisten der Zeichnungen und Bilder von externen Künstlern angefertigt wurden. Entweder kaufte Lavater Zeichnungen an oder ließ sich diese auch schenken, was ihm als Pastor nicht schwer gefallen sein dürfte. Andererseits gibt es auch Künstler, die Lavater selbst beschäftigte und die teilweise auch in seinem Haus lebten.

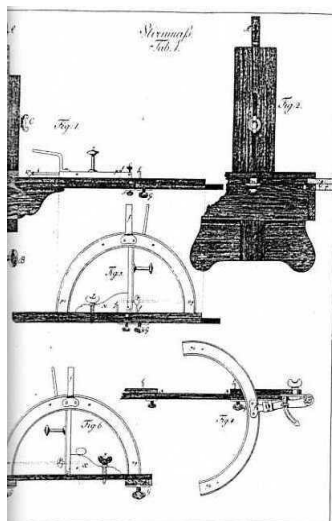
„Da gibt es eine Gruppe von mittelprächtigen Künstlern, die heißen auch Schweizer Kleinmeister, die haben auch einen eigenen Titel bekommen und wohnen teilweise bei ihm [Lavater] im Haus, also da sieht er auch eine soziale Verantwortung. Wie dieser zum Beispiel [...] Schellenberg, der hat mehrere Jahre bei ihm im Haus gewohnt und der war beschäftigt zu aquarellieren und das zu rahmen und umzuzeichnen „ (Interview 0)

Das Umzeichnen kann als ein wichtiger Schritt in der physiognomischen Bildpraxis Lavaters angesehen werden. Denn um das teilweise sehr unterschiedliche, heterogene Bildmaterial der verschiedenen Künstler in irgendeiner Weise wissenschaftlich bearbeiten zu können, musste es in eine Form gebracht werden, welche es erlaubte, die Zeichnungen vergleichen zu können. Swoboda spricht von graphischen Schemata wie z.B. Silhouetten, oder Umrisszeichnungen (Swoboda 2002: 3f.). Dieser Prozess kann als Homogenisierung von heterogenem Bildmaterial bezeichnet werden (ebd.: 63). Diese Homogenisierung kann auch als ein Versuch angesehen werden, die künstlerisch angefertigten Bilder ihrer Funktion als „Kunstbild“ zu entziehen und sie zu wissenschaftlichen Bildern zu machen. Die Kunstbilder werden ihrer Vielfalt und interpretativen Flexibilität entzogen und werden vereinheitlicht. Es wird versucht, eine klare Grenze zwischen Kunstbild und Wissenschaftsbild zu ziehen. Dieses Streben nach einer klaren Abgrenzung zu anderen Bildformaten kann als eine Charakteristik von wissenschaftlichen Bildern angesehen werden (Heßler 2006: 14ff.). Die von Lavater verwendeten Kunstbilder werden nicht nur homogenisiert, indem sie von seinen angestellten Künstlern in standardisierte Bildformate gebracht werden, sie werden auch simplifiziert und konkretisiert, um die Linien—und auf genau diese will Lavater letztendlich hinaus—aus den Bildern herauszuarbeiten.

Neben dem Umzeichnen von vorhandenen Kunstbildern, der Isolation von einzelnen Köpfen oder Gesichtsteilen, der Vereinfachung und Homogenisierung dieser sowie der Herausarbeitung von Stufenleitern und Collagen aus den isolierten Bildern, erstellte Lavater aber auch selbst Bilder: Lavater bevorzugte das Profilgesicht, aus welchem Schattenrisse produziert wurden, da es Ausmessungen leichter ermöglichte. Wenn Lavater von einer

physiognomischen Wissenschaft spricht, meint er vor allem die Anwendung von exakten Messmethoden, die für ihn erforderlich seien, um zu mathematisch gesicherten Ergebnissen zu kommen. Lavater erscheinen die Schattenrisse gerade deshalb als „wissenschaftlich zuverlässig“, weil sie nicht die Unsicherheiten eines Malers enthalten. Darüber hinaus fällt bei den Schattenrissen auch der emotional bedingte Gesichtsausdruck weg, welcher die physiognomische Deutung verfälschen könnte (Brednow 1969: 8f.) Bei der Herstellung der für Lavater so wichtigen Silhouetten verwendete er „Maschinen zur Vermessung und bildlichen Dokumentation realer Gesichtszüge“. (Interview 0) Dabei sind vor allem die Schattenrissmaschine und das Stirnmessgerät (Abbildung 1) zu erwähnen.

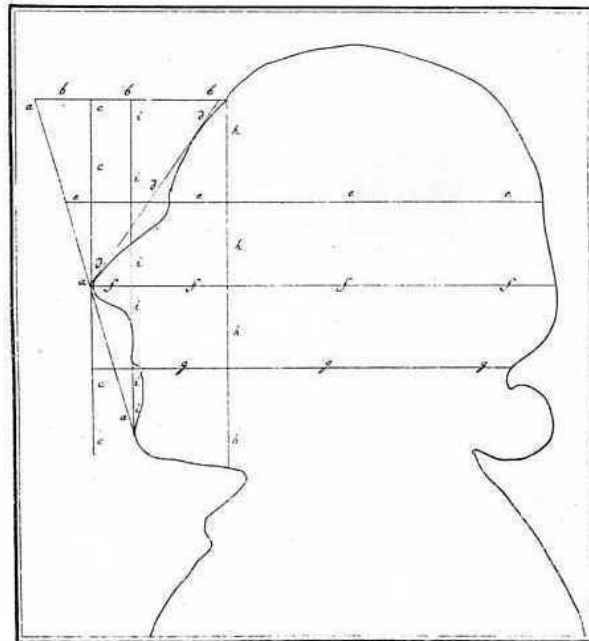
Abbildung 1: J. C. Lavater, Stirnmessgerät



Quelle: Tab. I in: PHF III, 237 in: Swoboda 2002: Abb. 48

Die Verwendung von solchen Maschinen zur Herstellung von Bildern steigert das Empfinden von Objektivität, man kann nach Daston und Galison von einem mechanischen Bild sprechen, das unter Mitwirkung der Maschine entsteht (Daston & Galison 1992:84; Daston 2001:129). In der Folge wird das mechanische Bild (bei Lavater der Schattenriss) mit Linien und daraus resultierenden Winkeln so weiterbearbeitet, dass es mathematische Berechnungen zulässt und in weiterer Folge Klassifikationen und Ordnungen möglich werden. Der zunächst empirisch beschreibende Charakter von Lavaters Zeichnungen wird also immer abstrakter (Swoboda 2002: 63). So entwirft er auch ähnlich wie schon Le Brun ein Gesichtswinkel-System, welches z.B. die Winkel vom äußeren Augenwinkel zur Nasenspitze und zum äußeren Mundwinkel messen sollte (vgl. Abbildung 2) und wovon quantitative Gesetzmäßigkeiten abgeleitet werden konnten, welche wiederum der physiognomischen Wissenschaft zu objektiver Geltung verhelfen sollten (ebd.: 64).

Abbildung 2: Anonym, Vermessungen an der Silhouette Lavaters



Quelle: In: J. C. Lavater, EP II, 201; In: Swoboda 2002: Abb. 50

Rezeption und Kritik an Lavaters Physiognomik

„Er [Lavater] war unglaublich populär. Also das war ein Anziehungspunkt in Zürich. Wer auf sich hielt, wer nach Zürich reiste, musste im Hause Lavater vorbeischauen. Er hat also zusätzlich zu diesen Aktivitäten noch sehr viele Besucher gehabt. Es gibt da Gästebücher, mit der europäischen Prominenz, die bei ihm vorbeigeschaut haben.“ (Interview 0)

Der Erfolg von Lavaters Physiognomik zu seinen Lebzeiten scheint unbestritten zu sein. Die Übersetzung der *Physiognomischen Fragmente* in mehrere Sprachen sowie seine enorme Korrespondenz mit wichtigen Persönlichkeiten insbesondere der deutschsprachigen Aristokratie, aber auch des gebildeten Bürgertums sind Indikatoren für diesen Umstand. Lavater war bis Mitte des 19. Jahrhunderts eine europäische Berühmtheit (Jaton 1988: 61). Sein Werk, die *Physiognomischen Fragmente* wurden von Lavaters Zeitgenossen begeistert aufgenommen, erhoffte man sich mit dieser Lehre doch, Auskünfte über den Charakter von Bekannten, Bediensteten oder Verlobten zu erhalten (Weigelt 1991: 100).

„Wer auf sich hielt, der hatte also seinen kleinen Taschen-Lavater bei sich und hatte die Nase und das Auge und eben das Haar eines Wollüstlings vergleichen können mit dem, was da vorgegeben war.“ (Interview 0)

Ein wichtiger Aspekt ist jedoch, dass Lavaters Werk zu seinen Lebzeiten bedeutend intensiver wirkte und extensiver aufgenommen wurde als nach seinem Tod, was vor allem

auf dem Zauber, der von seiner Person ausgegangen sein muss, beruhte (Weigelt 1991: 111). Weigelt erklärt sich Lavaters Erfolg aus einer Mischung von unglaublicher Ausstrahlungskraft und seinen Bestrebungen, immer neue Beziehungen zu knüpfen (ebd.), in diesem Sinne kann Lavater als ein Vernetzungskünstler angesehen werden (ebd.: 120). Auch Brednow spricht davon, dass Lavater zu seinen Lebzeiten als Wundermann angesehen worden sein muss, der den Menschen „ihre Wesensart von der Stirn ablesen konnte“ (Brednow 1969: 11). Schmölders erweitert diese Erklärung, wenn sie den Physiognomen Lavater als eine „Überfigur aus Wissenschaftler, Künstler und gläubigem Christ“ bezeichnet (Schmölders 1997: 29). Lavater versuchte ja tatsächlich „gegen jede Form der Zufälligkeit“ vorzugehen, was dem Wunsch der Aufklärung entsprach, „jede Form der Kontingenz und Unberechenbarkeit auszugrenzen“ (Gray 1994: 176). Dennoch sind bei Lavater seine Schwärmerei und Überschwänglichkeit kaum wegzudenken. Genau diese dienten als „eine Art rhetorisch-diskursiver Leim, mit dem die Risse im Fundament der Aufklärungsepisteme zugeklebt oder verdeckt werden sollten“ (ebd.). Diese Doppeldeutigkeit in Lavaters Werk ist besonders interessant und verdeutlicht meiner Meinung nach sehr schön, wie der eigentliche Widerspruch zwischen Glauben an die Schöpfung und Glauben an die Wissenschaft zu dieser Zeit vereinbar gewesen sein muss. Der Mensch konnte „gleichzeitig Spiegel der Gottheit und ein Kapitel im großen Buch der Natur sein.“ (Jaton 1988: 82). Lavater versuchte zwar, aus der Physiognomik eine Wissenschaft zu machen und agierte demnach auch nach wissenschaftlichen Kriterien wie Quantifizierung, mathematischen Berechnungen, Klassifikationen usw., der hauptsächliche Grund für seine Physiognomik war allerdings ein christlich motivierter Zugang: Denn Lavater war „immer auf der Suche nach dem wahren Ebenbild Christi. Also dieser christologische Aspekt ist wahrscheinlich sein Motor gewesen.“ (Interview 0). Das große Anliegen Lavaters war also die „manifeste Erfahrbarkeit des Göttlichen“ (Weigelt 1991: 120). Er glaubte, in jedem Menschen die Ebenbildlichkeit Gottes erkennen zu können (Brednow 1969: 1), erkannte den Menschen als das lebendige „Symbol der Herrlichkeit Gottes“ und sah ihn als Mittelpunkt der Natur an (Jaton 1988: 70). Folgt man Gray, so konnte sogar der Theologe Lavater dem Zwang der Verwissenschaftlichung im aufklärerischen Sinne nicht entkommen: Lavater lehnte sich an den szientistischen Positivismus der Aufklärung an, „dessen Ziel es war, nicht nur das sinnliche Leben der Menschen, sondern auch ihr übersinnliches Geistes- und Seelenleben auf rational kalkulierbare Gesetze zurückzuführen“ (Gray 1994: 168). Die Ebenbildlichkeit des Menschen mit Gott sollte also wissenschaftlich fundiert werden. Lavater versuchte meiner Meinung nach, diese Ebenbildlichkeit naturwissenschaftlich zu begründen und bediente sich dabei der gerade aufkommenden wissenschaftlichen Verfahren. Demnach wurde er nicht unbedingt einem Zwang unterworfen, sondern sprang vielmehr auf den mächtigen Zug des aufklärerischen Zeitgeistes auf, um eine Fundierung der göttlichen

Schöpfung in der Naturwissenschaft zu finden. In diesem Sinne wird die Physiognomik „zur Kernwissenschaft, die das (Wieder-)Erlernen der ursprünglichen Natursprache des Menschen ermöglicht“ (ebd.: 170). Diese Natursprache führt den Menschen in das „vorkulturelle Paradies“ zurück (ebd.). Das Wesen und der individuelle Charakter des Menschen werden bei Lavater naturwissenschaftlich am biologischen Körper festgemacht. Der Körper des Menschen, insbesondere sein Gesicht, wird auf eine göttliche Ordnung zurückgeführt. Lavater versucht „Das Antlitz Gottes im Antlitz des Menschen“ (vgl. Pestalozzi & Weigelt 1994) naturwissenschaftlich zu begründen. „Das göttliche Alphabet“ sollte entziffert werden (Jaton 1988: 71). Es scheint nicht besonders schwer zu erraten, wer für Lavater der Inbegriff aller Schönheit gewesen ist: Jesus Christus (Weigelt 1991: 98). In diesem Zusammenhang ist ein Grundsatz der Physiognomik von Lavater besonders kritisiert worden:

„Je moralisch besser, desto schöner.

Je moralisch schlimmer; desto hässlicher.“ (Lavater 1984 [1775]: 53)

Schönheit wird sozusagen mit positiver Moral gleichgesetzt. Das Antlitz Christi wird zum Inbegriff aller Schönheit und gleichzeitig zum Inbegriff aller Moral, an dem sich Lavaters physiognomisches Rasonnement (vgl. Schmölders 1997) orientiert. Es liegt nahezu auf der Hand, dass das Urteil des physiognomischen Meisters Lavater, dessen Expertise zwar einerseits auf genauen Beobachtungen basierte, aber andererseits auch auf einer ihm vielfach unterstellten Intuition, bei „moralisch besseren Personen“, also Personen, die er kannte und schätzte, besser ausfiel, als bei ihm unbekannten Personen. So merkte sein Freund und langjähriger Hausgenosse Ulrich Hegner an, dass Lavater bei der Ausdeutung von Bildmaterial von ihm unbekannten Personen sehr unsicher gewesen sei, bei ihm bekannten Personen aber auch „menschliche Wertmaßstäbe aus beiderseitigem Bekanntsein mit herangezogen habe“ (Brednow 1969: 8). Werte, insbesondere christliche Werte, spielten also in der wissenschaftlichen Arbeit Lavaters eine entscheidende Rolle.

Einer der größten Kritiker von Lavater, der Göttinger Schriftsteller und Physiker Georg Christoph Lichtenberg (1742–1799) bezeichnete die Physiognomik Lavaters in seiner Schrift *Über Physiognomik: Wider die Physiognomen* als „groben Aberglaube“, der „die Maske der Vernunft“ trägt (zit. nach Gray 1994: 167). Die Kritik Lichtenbergs, so polemisch sie auch war, war inhaltlich gesehen auf die Verallgemeinerungen bei Lavater ausgerichtet (Buser 1973: 46). Er kritisierte die Lehre der Analogie zwischen körperlicher und moralischer Harmonie aber insbesondere die Rolle des Körperbaus als Quelle der Erkenntnis (ebd.: 49). Nach Lichtenbergs Auffassung dürfte man nicht Physiognomik betreiben, sondern müsste

sich vielmehr auf die Pathognomik konzentrieren (ebd.: 49; Schmölders 1997: 29)⁷. D.h., dass das Augenmerk nicht auf die festen Körperteile gerichtet werden sollte, sondern vielmehr auf die weichen, beweglichen Teile des menschlichen Körpers, die sich auch verändern können. Doch gerade die weichen, beweglichen Teile waren in Lavaters Augen hinderlich, denn sie verfälschten die physiognomische Deutung des Grundcharakters des Menschen (Brednow 1969: 9). In Lavaters Auffassung sei die Pathognomik unsicher, weil man es bei dieser mit den „Verstellungskünsten“ der Menschen zu tun hätte (Weigelt 1991: 96). Brednow stellte fest, dass sich dieser Trend von der Physiognomik hin zur Pathognomik nach Lavater immer stärker zeigte (Brednow 1969: 13), andererseits setze sich die Physiognomik insbesondere in den Schädelvermessungen von Gall fort (ebd.; Christians 2006: 4; Jatón 1988: 86; Schmölders 1997: 29, Aerni 2003: 80). Auf diese Zeit nach Lavater möchte ich im nächsten Abschnitt *Von Lavater zu Darwin* eingehen.

⁷ Eine ausführliche Beschreibung der Kritik von Lichtenberg an Lavater findet sich bei Buser (vgl. Buser 1973: S. 46 – 56).

Von Lavater zu Darwin

Den Titel für dieses Unterkapitel *Von Lavater zu Darwin* habe ich mir von einem Heft, einem von Walter Brednow verfassten Sitzungsbericht der sächsischen Akademie der Wissenschaften zu Leipzig, das ich glücklicherweise noch in einem Antiquariat finden konnte, entlehnt, welches den gleichen Titel trägt (vgl. Brednow 1969). Dieses Büchlein beschreibt eine ansonsten kaum auffindbare

„Wegstrecke der Wissenschaftsgeschichte vom Menschen, [...] die sich so intensiv und auch so kampfbereit mit dem Wesen des Menschen, seiner Herkunft und vor allem seinen Ausdrucksformen beschäftigt hat wie die mit den Namen Lavater und Darwin gekennzeichnete.“ (ebd.: 1).

Es ist genau dieser Weg ausgehend von Johann Caspar Lavater hin zu Charles Darwin, den ich hier einschlagen und skizzieren möchte. Die Betonung muss hier auf Weg liegen, denn einen direkten Sprung von Lavater zu Darwin hat es nicht gegeben. Zwischen Lavaters erster physiognomischer Veröffentlichung *Von der Physiognomik* (1772) und Darwins Buch *Der Ausdruck der Gefühle bei Mensch und Tier* von 1872 liegen genau 100 Jahre. In diesen 100 Jahren Wissenschaftsgeschichte der Physiognomik lässt sich eine deutliche naturwissenschaftliche Wendung erkennen (Brednow 1969: 27) und obwohl Darwins Buch auch explizit von der traditionellen Physiognomik Abstand hält, muss es gleichzeitig auch als wichtiger Bestandteil der physiognomischen Tradition gelesen werden (Person 2005: 45). Die Wendung aber bezieht sich auf drei wesentliche Punkte: Der von Lavater eingeschlagene Weg wurde zum einen dadurch verlassen, dass die bei ihm anzutreffende „moralisch-sittliche Verbindlichkeit“ bei Darwin fehlt. Des Weiteren verlor auch die auf den beständigen, ruhenden und unveränderlichen Körperteilen aufbauende Physiognomik zu Gunsten der Pathognomik an Bedeutung (Brednow 1969: 27). Und drittens änderte sich auch die Beziehung von Mensch und Tier grundlegend: War bei Lavater der Mensch-Tier-Vergleich noch von göttlicher Ordnung bestimmt, so wurden bei Darwin sowohl tierische als auch menschliche Ausdrucksformen mit den gleichen physiologischen Modellen erklärt (Person 2005: 47). Hier soll aber nicht der Eindruck entstehen, dass die Physiognomik, die auf den beständigen, ruhenden und unveränderlichen Körperteilen, also auf dem Knochenbau, basierte, zunehmend verschwand: Gerade die Phrenologie von Franz Joseph Gall und in der Folge die kraniologischen Arbeiten von Carl Gustav Carus schlossen bei Lavater an, kritisierten diesen zwar, aber setzen diesen Zweig der Physiognomik äußerst erfolgreich fort (Brednow 1969: 13; Christians 2006: 4; Jatón 1988: 86; Schmölders 1997: 29, Aerni 2003: 80). Auf diese Entwicklung möchte ich direkt anschließend eingehen, um

exemplarisch an Gall und Carus die Fortsetzung und Weiterentwicklung der Physiognomik zu verdeutlichen, bevor ich dann die parallel stattfindende Wende von der Physiognomie hin zur Mimik, also wenn man so will den *mimic turn*, beschreiben werde. Hierbei gehe ich insbesondere auf Johann Jakob Engel, Charles Bell und Theodor Piderit ein, wobei im selben Zug auch Georg Christoph Lichtenberg genannt werden sollte. Schließlich komme ich im dritten Teil dieses Unterkapitels bei Charles Darwin und insbesondere bei seinem Buch *Der Ausdruck der Gefühle bei Mensch und Tier* an. Auch wenn sich all diese erwähnten historischen Wissenschaftler von ihren Motiven und Kontexten einmal mehr, einmal weniger unterscheiden, ist ihnen doch gemein, dass sie alle die Sehnsucht hatten, das Innere des Menschen in seinem Äußeren zu erkennen.

Bevor ich mich gleich auf die weiteren Spuren der Physiognomik begeben, will ich aber einen kleinen Moment der Geschichte nicht vorenthalten, welcher die Nähe von Darwin zu Lavater in satirischer Weise verdeutlicht: Darwin trat 1836 seine Weltreise mit der *Beagle* an. Auf dieser Reise sollte er die Materialien und Eindrücke sammeln, welche als Ausgangspunkt für seine spätere Evolutionstheorie gesehen werden können. Dabei hätte ihm beinahe der Kapitän des Schiffes den Zutritt zum Schiff verweigert und damit die Reise vereitelt, da dieser—als Anhänger Lavaters—aus der Nase Darwins ablesen wollte, dass es sich bei Darwin um eine Person handle, die zu energielos und zu wenig entschlossen für eine solche Reise wäre (Pestalozzi & Weigelt 1994: 189).

Auf den Spuren der Physiognomik Lavaters: Franz Joseph Gall und Carl Gustav Carus

Über die Gegebenheit, dass die Phrenologie von Franz Joseph Gall (1758–1828) in unmittelbarer Nähe und Nachfolge zur Physiognomik von Johann Caspar Lavater steht, scheinen sich viele AutorInnen einig zu sein (Brednow 1969: 13; Christians 2006: 4; Jatón 1988: 86; Schmölders 1997: 29, Aerni 2003: 80). Die Phrenologie hatte ähnlich wie die Physiognomik Lavaters ungeheuren Erfolg (Schmölders 1997: 29), wurde aber genauso kontrovers aufgenommen wie diese. So schreibt etwa Wilhelm von Humboldt am 28.04.1833 in einem Brief an eine Freundin:

„Gall [...] habe ich noch persönlich gekannt und seinen Lehrstunden über die Schädellehre in Wien 1797 beigewohnt. Ich habe nie einen Augenblick daran geglaubt. Es war eine der Erfindungen, die, wenn man sie des Charlatanismus entkleidet, der sie umgibt, eine sehr dürftige Wahrheit hinterlassen [...]“ (von Humboldt 1853: 212f. zit. nach Brednow 1969 14).

Die Phrenologie oder auch Schädellehre von Gall ging von der Annahme aus, dass beim ausgewachsenen Menschen die Vertiefungen und Erhöhungen der Schädeldecke auf das

Vorhandensein bzw. Fehlen von Gehirnarealen schließen lassen. Diese Gehirnareale wiederum würden menschliche Fähigkeiten und Dispositionen erzeugen. Die Phrenologie versuchte, die Fähigkeiten, Anlagen und den Charakter eines Menschen aus dessen Schädel abzuleiten (Schmölders 1997: 29). Das Äußere des Schädels sollte also darüber Aufschluss geben, wie das Innere des Menschen, sein Charakter oder seine Fähigkeiten aussehen. Gall kann als einer der Begründer der Schädelkunde, einer „Wissenschaft der Beurteilung geistiger Fähigkeiten nach der Größe bestimmter Hirnregionen“, angesehen werden (Gould 1988: 94). Die Phrenologie Galls erhob großen Anspruch auf Sachlichkeit und Wissenschaftlichkeit, indem sie sich nur auf die harten Kanten und Flächen des Schädels konzentrierte. Dass sie zunächst zum Teil euphorisch auch als ernsthafte Wissenschaft wahrgenommen wurde, verdankt sie wohl dem endgültigen Bruch mit dem Spiritualismus, mit dem die Physiognomik Lavaters noch durchdrungen war (Jaton 1988: 86). Diese Abwendung vom Spiritualismus hin zu einem Materialismus verursachte in Österreich sogar ein Verbot der Vorlesungen und Publikationen Galls, welches von Kaiser Franz I. auf Drängen medizinischer und jesuitischer Kreise 1801 ausgesprochen worden ist (Aerni 2003: 81).

In engem Zusammenhang zu Franz Joseph Gall, Johann Wolfgang von Goethe und Johann Caspar Lavater, auf welche er in seinem Werk *Symbolik der menschlichen Gestalt* zurückgriff, müssen auch die kraniologischen Arbeiten des sächsischen Arztes Carl Gustav Carus (1789–1869) gesehen werden (Brednow 1969: 15). Der Einfluss der Ideen der drei eben genannten Personen führt meiner Meinung nach bei Carus zu einer differenzierteren Ansicht. Das äußert sich darin, dass er aus dem Körper des Menschen dessen Konstitution, das Temperament und die geistigen Anlagen mit Hilfe von Organoskopie, Physiognomik und Pathognomik erkennen will. Dabei soll sich aus bestimmten Zeichen am „Stamme“ und den „Gliedmassen“ des Menschen dessen Konstitution, am Antlitz des Menschen dessen Temperament und am Schädel dessen geistige Anlagen erkennen lassen (Buser 1973: 70). Carus knüpft mehrmals an Lavater an und lobt auch dessen ehrliches Bemühen, kritisiert aber auch die fehlende Systematik, die Lückenhaftigkeit und Unvollkommenheit bei Lavater sowie dessen „mystische Natur“ (ebd.: 71, Brednow 1969: 15). Eine entscheidende Ähnlichkeit zu Lavater zeigt sich darin, dass auch Carus ein menschliches Ideal vorgibt, an dem das Individuum gemessen werden soll. Lavater geht allerdings von einem polaren Verhältnis zwischen Hässlichkeit und Schönheit—dazu in Analogie die Moral—aus, während Carus die Schönheit in das Zentrum stellt, von welchem die Hässlichkeit sowohl in das Minus als auch in das Plus abweicht (Buser 1973: 72). Carus ging von einer engen Verknüpfung von psychischen und physischen Symptomen aus, was er z.B. am Buckel demonstriert, da zwischen Rückenbildung und Geistesverhalten eine besondere Beziehung bestehe. Dabei

sei anzumerken, dass Carus von Rückenbildung spricht, was auf einen Prozess hindeutet: So kann die Form des Rückens als Buckel neben der biologischen Anlage auch durch Gewohnheiten, Krankheiten oder Verletzungen zustande kommen (ebd.: 78).

Die heutige Bedeutung von Carus zeigt sich darin, dass einige wichtige Institutionen nach Carus benannt wurden: das Universitätsklinikum Carl Gustav Carus in Dresden, die Carl Gustav Carus Stiftung für psychosomatische Forschung in Zürich oder die Carl Gustav Carus Akademie in Hamburg.

Der *mimic turn*: Lichtenberg, Engel, Bell und Piderit

Während sich die einen wie Gall oder Carus weiterhin vor allem auf die festen, vermeintlich unveränderlichen Teile des Körpers, auf den Knochenbau konzentrierten und so die Physiognomik in der Phrenologie und Kraniometrie eine Fortsetzung und Umdeutung fand, so hielten andere an der außerordentlichen Bedeutung des Gesichts fest, verschoben aber den Fokus auf die weichen, beweglichen Teile, auf die Pathognomik und Mimik. Dabei meint Pathognomik etwas anderes als Mimik: Pathognomik ist in einem medizinischen Kontext zu sehen, und wird hier als Kunst, Krankheiten an ihren Erscheinungen zu erkennen“ (Herders Conversations Lexikon 1856: 472) bezeichnet. Andererseits kann die Pathognomik aber auch allgemeiner gesehen werden als „die Kunst, die Gemütsbewegungen aus den Veränderungen des Körpers, hauptsächlich der Gesichtszüge, zu erkennen“ (Meyers Großes Konversationslexikon 1908: 502). Wie bei der Kritik an Lavater schon ausgeführt wurde, war es insbesondere Georg Christoph Lichtenberg der hier in unmittelbarer Reaktion auf Lavater dessen Konzentration auf den Körperbau kritisierte und forderte, sich stärker auf die Pathognomik zu konzentrieren.

Wie auch Lichtenberg stellt der deutsche Schriftsteller und Philosoph Johann Jakob Engel (1741–1802) die Physiognomik von Lavater in Frage—ohne dabei aber trotz dessen Lektüre direkt auf Lavater einzugehen—und geht vielmehr von einer ungewissen Grenze von Physiognomik und Mimik aus (Brednow 1969: 13). Er stellt also die Physiognomik nicht grundsätzlich in Frage, geht aber davon aus, dass sie nicht die einzige Erkenntnisquelle ist, sondern daneben die Mimik auch eine bedeutende Rolle spielt. Engel gab zwischen 1785 und 1786 in Form eines Romans in 44 Briefen eine Anleitung zur Mimik für Schauspieler. Gerade auch deshalb ist sein Beitrag problemgeschichtlich von geringerer Bedeutung, da er undogmatisch und auf das Praktische des Schauspiels ausgerichtet war. Das Werk von Engel besteht vor allem aus Beobachtungen, die er versucht klar und deutlich darzulegen (Buser 1973: 79f.) Einen inhaltlichen Beitrag liefert Engel, indem er die Abhängigkeit des Gesichtsausdrucks von sozialen und kulturellen Faktoren erkennt:

„Nur sehen Sie, bitt ich, wenn Sie vollen, kräftigen Ausdruck suchen, nicht auf die vornehmern Stände, nicht auf die sogenannten Leute von Welt, von Lebensart, von Erziehung. Die Erziehung macht den Menschen zu einem zweifachen Lügner; sie lehrt ihn die eine Art von Empfindungen nach ihrer wahren Stärke verbergen, die andere in einer falschen Stärke erheucheln. Alle lebhaften Ausdrücke eigennütziger, und eben so, alle schwächern Ausdrücke geselliger Neigungen und Triebe, wie wahr und angemessen sie übrigens seyn mögen, sind wider den Wohlstand. Jene werden daher unter aller Wahrheit niedergehalten; diese über alle Wahrheit hinaufgeschoben.“ (Engel 1785: 176f. zit. nach Buser 1973: 80).

Die Erziehung und damit zusammenhängend auch der Stand, die Klasse oder Schicht, aus welcher man kommt, spielen nach Engels also eine Rolle, wie intensiv man seinen Gesichtsausdruck zeigen oder verbergen kann.

Mit der Rolle der Gesichtsausdrücke beschäftigte sich aus einer anderen Motivation einige Jahre später auch der schottische Anatom und Physiologe Charles Bell (1774–1842). Er untersuchte 1806 in seinem Werk *The Anatomy and Philosophy of Expression* die charakteristischen Ausdrücke der ruhenden und durch Emotionen bewegten Gesichtszüge bei Menschen und Tieren. Er bezieht sich zunächst auch auf den knöchernen Schädel, geht aber schon bald zu den Muskelbewegungen über, die die verschiedenen mimischen Ausdrucksformen begleiten und sie zur Darstellung bringen. Dabei beschreibt Bell die Gesichtsmuskulatur bei Menschen und Tieren anatomisch detailliert, wobei der Mensch mit seinen spezifischen Ausdrucksformen zwischen Fleisch- und Pflanzenfressern eingeordnet werden kann. Am Beispiel der Wut schildert er die Muskelbeteiligung des Ausdrucks beim Tier (Zähne zeigen und Hinaufziehen der Lippen), welche man aber auch beim Menschen findet.

In stärkerem Zusammenhang mit Lavater und Lichtenberg als Charles Bell muss der Deutsche Theodor Piderit (1826–1912) gesehen werden. Er entfernte sich in seinen beiden Werken *Grundsätze der Mimik und Physiognomik* von 1858 und *Wissenschaftliches System der Mimik und Physiognomik* von 1867 besonders stark vom Knochenbau als Basis der Physiognomik—ist sich hierbei mit Lichtenberg einig—und versuchte die Mimik als Grundlage der Physiognomik zu etablieren. (Buser 1973: 80ff.). Der Leitsatz von Piderit hebt die Bedeutung der Muskeln des menschlichen Körpers hervor:

„Physiognomische Merkmale darf man nur an den Theilen suchen, welche unter dem Einflusse der Geistesthätigkeit stehn. Diese Theile aber sind die Muskeln und vorzugsweise die zahlreichen und beweglichen Muskeln des Gesichts. Die vorübergehenden, mimischen Bewegungen dieser Muskeln, die mimischen Züge werden durch häufige Wiederholung zu bleibenden, zu physiognomischen Zügen, und ein physiognomischer Ausdruck ist anzusehn als ein habituell gewordener mimischer Ausdruck“ (Piderit 1867: 148f. zit. nach Buser 1973: 81).

Das Werk von Piderit ist problemgeschichtlich besonders interessant, da es eine Abkehr von der absoluten und eindeutigen Gültigkeit physiognomischer Befunde einleitet. Bei Piderit spielen auch die individuellen, ontogenetischen Entstehungsgeschichten der Merkmale in der Physiognomie eine Rolle, wie sie z.B. durch Krankheiten, aber auch durch äußere Ursachen wie Wind und Wetter, durch Gewohnheit, Grimassen, Alter oder Temperament verursacht werden (Buser 1973: 83). Obwohl Piderit das Werk Lavaters ablehnt und die Wissenschaftlichkeit bei Lavater in Frage stellt, indem er ihm „Redensarten statt Gesetze“ oder „religiöse Tiraden statt wissenschaftlicher Methodik“ unterstellt (Brednow 1969: 23), will er die Physiognomik wie Lavater zu einer mathematisch bestimmbar Wissenschaft machen, welche streng rationalistischen – positivistischen Gesetzen gehorchen soll. Seine Regeln in Bezug auf die Mimik werden

„dem Künstler ein Mittel an die Hand geben, einen beliebigen verlangten Gesichtsausdruck gleichsam mit mathematischer Bestimmtheit zu construieren und in einem Gesichte darzustellen“. (Piderit 1867: 14f. zit. nach Buser 1973: 80).

Das Zitat zeigt auch die Motivation von Piderit: Er erwartete sich von seiner rationalen Begründung der Mimik einen besonderen Nutzen für die Kunst (Brednow 1969: 23), indem dadurch Künstlern Gesichtsausdrücke exakt gelehrt werden konnten.

Ein völlig anderes Interesse an der Mimik und an Gesichtsausdrücken wie Theodor Piderit hatte einer der berühmtesten Wissenschaftler der Geschichte, Charles Darwin. Dieser lehnte sich eher an Charles Bell an, für ihn war der Gesichtsausdruck der „missing link“, also das fehlende Bindeglied zwischen Mensch und Tier (Person 2005: 52).

Angekommen bei der Mimik: Charles Darwin (1809–1882)

Bevor ich auf die Beschäftigung Darwins mit Gesichtsausdrücken und der Mimik eingehe, möchte ich nicht nur wegen seinem 200. Geburtstag kurz auf dessen Leben und Werk eingehen, sondern ich möchte dies auch tun, weil Darwins Buch *Der Ausdruck der Gemütsbewegungen bei den Menschen und Tieren* von 1872 kaum von seinem Leben und Werk getrennt gesehen werden kann.

Charles Darwin: Leben und Werk

Der britische Naturforscher Charles Darwin, geboren am 12. Februar 1809 in Shrewsbury als fünftes von sechs Kindern, ist neben Albert Einstein der am häufigsten abgebildete Wissenschaftler der Geschichte und ist seit dem Jahr 2000 auf der englischen Zehnpfundnote zu finden, als er dort Charles Dickens verdrängte.

Abbildung 3: Darwin auf der englischen Zehnpfundnote



Darwins Name, sein Bild und seine Theorie sind unauflösbar miteinander verbunden (Voss 2008: 15). Sein Buch *Die Entstehung der Arten* ist heute von der Bedeutung her vergleichbar mit Karl Marx *Das Kapital* oder Adam Smiths *Der Wohlstand der Nationen* (ebd.: 10). Darwins große Theorie ist die Evolutionstheorie, die—in wenigen Worten zusammengefasst—aussagt, dass sich alle Organismen durch winzige Merkmale unterscheiden und dass diese, falls sie erblich sind, zur Grundlage des Artenwandels werden können. Außerdem geht die Evolutionstheorie von einer Verwandtschaft und einem gemeinsamen Ursprung von Mensch und Tier aus (ebd.: 9). Auch wenn die Evolutionstheorie heute zur Allgemeinbildung gehört, kann sie unterschiedlich ausgelegt und instrumentalisiert werden. Wenn wir heute von DarwinistInnen sprechen, kann damit zweierlei gemeint sein: Zum einen kann das ein Mensch mit einem wissenschaftlich aufgeklärtem Weltbild sein, der die Richtigkeit der Evolutionstheorie voraussetzt und die Annahme, dass Gott fortwährend in das Leben eingreift negiert, zum anderen kann es sich um einen Menschen handeln, der die Annahme als „*survival of the fittest*“, also die Anschauung, dass „die Stärkeren“ siegen und „die Schwächeren“ unterlegen, als Ideologie auffasst (ebd.). Nicht als der Beste, aber zumindest als Zehnter von insgesamt 178 Prüfungskandidaten schloss Charles Darwin 1831 sein Studium der Theologie—Naturwissenschaften wurden damals in der Theologie gelehrt—in Cambridge ab, welches er erst nach seinem abgebrochenen Medizinstudium in Edinburgh aufnahm. Im Dezember desselben Jahres 1831 ermöglichte ihm ein anglikanischer Geistlicher namens John Stevens die Mitreise auf der H.M.S. Beagle, dem schon erwähnten Schiff, auf dem er die Materialien und Eindrücke sammeln sollte, die als Ausgangspunkt für seine spätere Evolutionstheorie gesehen werden können. Die Reise führte Darwin um die ganze Welt, unter anderem nach Teneriffa, Brasilien, Patagonien, Galapagos, Tahiti, Australien und Mauritius. Aber erst nach seiner Rückkehr entwickelte er ab 1836 in London die Evolutionstheorie (ebd.: 38f.). So umstritten Darwins Evolutionstheorie war und auch heute noch ist, so erfolgreich war sie auch. Der Erfolg seiner Theorie war aber nicht zufällig, sondern wurde von mehreren gesellschaftlichen Entwicklungen begünstigt: Zum einen bildete sich im 19. Jahrhundert erstmals eine Medienöffentlichkeit heraus, Bücher wurden billiger und eine breitere Öffentlichkeit konnte erreicht werden. Zweitens stand Darwin durch den Aufstieg Englands zur weltweit größten Kolonialmacht ein riesiges globales Netz von

Beamten, Händlern, Plantagenbesitzern und Gärtnern zur Verfügung, welches er nützte, um Informationen zu erhalten. Darwin baute sich, vielleicht ähnlich wie Lavater in Europa, ein riesiges Netzwerk—eine Art Agentensystem—an Menschen auf, mit denen er in regem Briefwechsel stand. Und drittens wurde Darwins Erfolg auch durch einen allgemeinen Trend des politisch- gesellschaftlichen Umbruchs im 19. Jahrhundert begünstigt: Nationalstaaten und Parlamente wurden gebildet, die Sklaverei abgeschafft, das Bürgertum löste den Adel ab und der Kapitalismus entwickelte sich (ebd.: 15ff.). 1859 wird *Die Entstehung der Arten* veröffentlicht, 1871 *Die Abstammung des Menschen*. Nur vier Monate später im Jahre 1872 gibt Darwin dann *Der Ausdruck der Gemütsbewegungen bei den Menschen und Tieren* heraus, welches ursprünglich ein Teil der *Abstammung des Menschen* sein sollte. Allein dieser Sachverhalt zeigt die enge Verbindung dieser beiden Werke (ebd.: 156) und so auch die enge Verbindung der Beschäftigung Darwins mit Gesichtsausdrücken und seiner großen Evolutionstheorie.

Der Ausdruck der Gemütsbewegungen bei Mensch und Tier

Das 1872 veröffentlichte Buch war das Buch Darwins mit den meisten Abbildungen und das einzige Buch, das Fotografien enthielt. Es war auch das einzige Buch, das den Menschen zeigte. Dies mag unter anderem erklären, wieso *Der Ausdruck der Gemütsbewegungen bei Mensch und Tier* zunächst das erfolgreichste Werk Darwins war. Innerhalb von vier Monaten war es 9000mal verkauft. Bis zur Jahrhundertwende erschien es auch in den USA, in Deutschland, Italien, Russland, Holland und Frankreich (Voss 2008: 157).

Wie schon erwähnt lehnte Darwin Piderit und auch Lavater eher ab und schloss an Charles Bell an, arbeitete sich an dessen Buch *The Anatomy and Philosophy of Expression* ab. Vor dem Hintergrund der Evolutionstheorie ist es nicht verwunderlich, dass Darwin die von Bell aufgestellte Hypothese, dass sich die Mimik von Mensch und Tier fundamental voneinander unterscheidet, weil der sog. Brauenrunzler (ein Muskel, mit dem die Partie zwischen Auge und Stirn bewegt wird) nur beim Menschen vorhanden ist, ablehnte. Darwin hatte diese Muskelbewegung bei Affen aber häufig beobachten können (ebd.: 158f.). Darwin ging also von einer grundsätzlichen Verwandtschaft von Tier und Mensch aus. War bei Lavater das Tiergesicht noch als Schablone gedacht, um die verborgenen Eigenschaften des menschlichen Charakters entziffern zu können, änderte sich das bei Darwin durch die Verwandtschaftsbeziehung grundlegend (ebd.). Darwins Schrift markiert einen Wendepunkt und es kommt zu einer entscheidenden Umdeutung der physiognomischen Traditionslinien. War bei Lavater der Mensch-Tier Vergleich von göttlicher Ordnung bestimmt, werden bei Darwin sowohl tierische als auch menschliche Ausdrucksformen mit den gleichen physiologischen Modellen erklärt (Person 2005: 47). Wenn Darwin eine Fotografie von einer Dame zeigt, die ihren Eckzahn durch das Hochziehen der Oberlippe freigibt, erinnert das an

den Ausdruck eines „durchaus hündischen Fletschens“. Und wenn er gleichzeitig die Muskelbewegungen der Mund- und Augenpartie eines Schopfmakaken als ein (bis dato exklusiv) menschliches Lächeln zeigt, so ist der tierische Ausdruck damit „so menschlich wie der des Menschen animalisch“ (Voss 2008: 158).

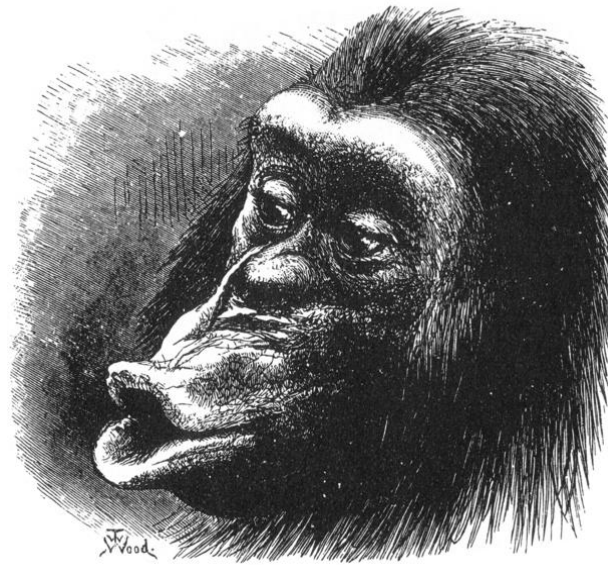
Abbildung 4: Spott und Missachtung



Quelle: Darwin 1948 [1872:]: p.124, Fig. 18– Sneering and Defiance; the canine tooth on one side is uncovered

Die Nähe zwischen Mensch und Tier wurde nicht mehr durch das schwarze Gorillagesicht der Drohgebärde und des Gorillas als Bestie hergestellt, sondern über den sonderlich lächelnden Schopfmakaken namens Adam aus dem Zoologischen Garten Regent's Park in London. In der Naturgeschichte war der lachende Affe eine Premiere (ebd.: 161). Die vielen Bilder in Darwins Buch waren wohl maßgeblich an seinem Erfolg beteiligt. Sie stellten eine gewisse Nähe zur Lebenswelt der Leserinnen und Leser her. Neben dem lächelnden Schopfmakaken aus dem Zoo konnte man u.a. einen Schimpansen sehen, der aus Enttäuschung über eine vorenthaltene Orange schmollend die Lippen nach vorne schob, einen Schwan aus dem Hyde Park, Hunde und Katzen, aber auch schreiende Kleinkinder oder eine von der Liebe enttäuschte junge Frau. Diese Bebilderung ergab „eines der vergnügtesten Bücher der Geschichte der Wissenschaft“ (ebd.: 161). Und während bei anderen Evolutionstheoretikern wie Ernst Haeckel oder Thomas Henry Huxley „der Australier“ bzw. „Afrikaner“ das Bindeglied zwischen Mensch und Tier war, so argumentierte Darwin diese Nähe mit den Bildern von EngländerInnen und deren Haustieren. Es waren also nicht die wilden Tiere oder außereuropäische Völker, sondern die Lebewesen aus der unmittelbaren Umgebung, die man kannte und die man ständig sehen konnte. Die Verbindung zwischen Mensch und Tier schuf also nicht „der animalische Mensch“, sondern „das menschliche Tier“ (ebd.:).

Abbildung 5: Enttäuschter und mürrischer Schimpanse



Quelle: Darwin 1948 [1872]: p.68, Fig. 14 – Chimpanzee disappointed and sulky because an orange had been taken away from him. A similar, though slighter pouting may be seen in sulky children.

So erfolgreich Darwins Buch auch war, so umstritten waren die drei Prinzipien, unter denen er die Entstehung des Gesichtsausdrucks verstanden wissen wollte. Das erste Prinzip war das der zweckdienlich assoziierten Gewohnheiten, das zweite Prinzip das Prinzip des Gegensatzes und das dritte Prinzip

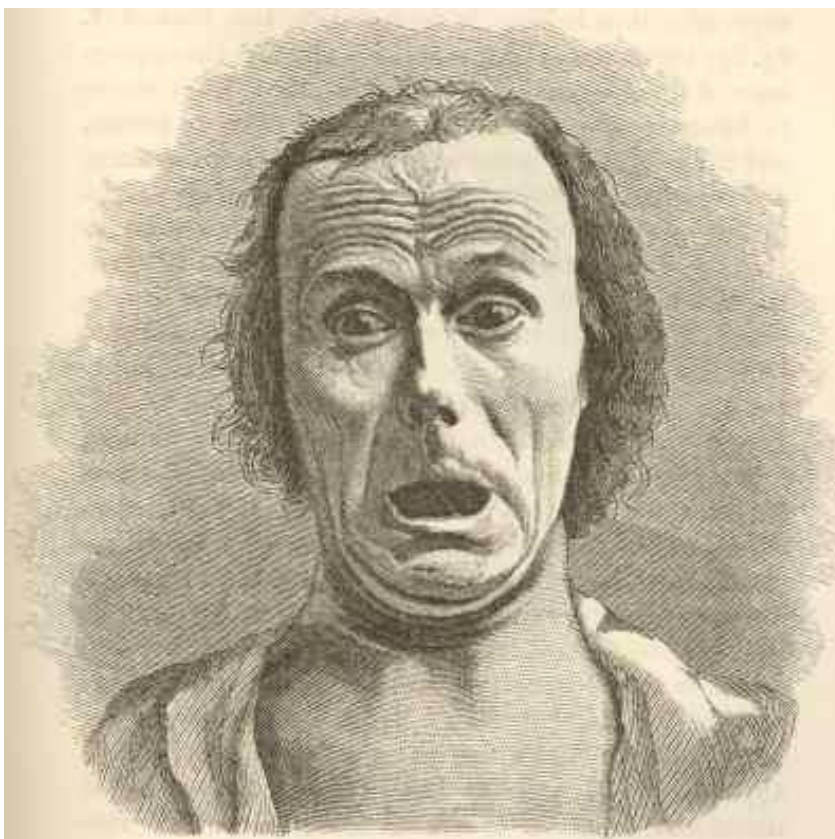
„Das Prinzip, dass Tätigkeiten der Verfassung des Nervensystems entstammen, von Anfang an unabhängig vom Willen und bis zu einem gewissen Grade unabhängig von der Gewohnheit sind.“ (zit. nach Buser 1973: 86).

Gerade das hier zitierte dritte Prinzip stellt das Individuum als „Sklave seines Nervensystems“ dar (ebd.). Die biologische Veranlagung in Form des Nervensystems geht über den Willen des Menschen und manifestiert sich gut sichtbar im äußeren Ausdruck des Menschen. Die Ausdrucksbewegung beim Menschen ist somit ein automatisch gewordener mechanischer Ablauf geworden, welcher über gewohnheitsmäßigen Gebrauch über weite Zeiträume hinweg von den Urfahren ererbt worden ist (Brednow 1969: 25).

Die Annahme, dass zumindest einige Basisemotionen des Menschen angeborenes Verhalten sind und der Gesichtsausdruck genetisch bestimmbar ist, wurde von Darwin mit einem Mix aus verschiedenen wissenschaftlichen Methoden untersucht: Er beobachtete Kinder, weil bei diesen der Gesichtsausdruck noch weitgehend ungehemmt sei. Außerdem beobachtete er „Geistesranke“, weil diese „Ausbrüche der stärksten Leidenschaft“ nicht kontrollieren könnten. Neben den Beobachtungen kamen bei Darwin auch sog. Dekodierungsstudien zum Einsatz. Bei diesen Studien mussten Versuchspersonen

einschätzen, welche Emotionen auf Fotografien von Gesichtern zu erkennen waren. Die dafür verwendeten Fotografien stammten von Duchenne de Bologne, welcher mit Elektroden die einzelnen Gesichtsmuskeln so reizte, dass bestimmte Gesichtsausdrücke gezeigt wurden (vgl. Abbildung 6).

Abbildung 6: Schrecken



Quelle: Darwin 1948 [1872]: p.145, Fig. 21 – Terror. The eyes stare, the nostrils and pupils are dilated; the strong contraction of the Platysma—the muscle of the fright—wrinkles the neck and helps to depress the lower jaw and keep the mouth open. The brow is horizontally furrowed and the general attitude is one of flaccidity and weakness..

Weiters untersuchte Darwin auch Personen aus verschiedenen Kulturen, die keinen Kontakt zueinander hatten. Hier ist die sogenannte Missionarstudie hervorzuheben (Merten 2003: 39). Darwins Missionarstudie wollte die Universalität der Gesichtsausdrücke überprüfen, er stellte sich die Frage, ob die bekanntesten Ausdruckerscheinungen bei allen Menschen in gleicher Weise anzutreffen seien. Zu diesem Zweck versandte er einen Fragebogen an Bekannte, die weltweit verstreut waren und Kontakt zu „fremden“, nicht europäischen Völkern hatten, vornehmlich zu Missionaren, aber auch zu Lehrern und Farmern (Buser 1973: 85). Eine Frage aus Darwins Fragebogen lautete z.B.: „Bewirkt Scham ein Erröten, wenn die Farbe der Haut gestattet, dass dasselbe sichtbar wird? Und besonders: Wie weit am Körper hinab erstreckt sich das Erröten? (Darwin zit. nach Beer 1964: 12 zit. nach Buser

1973: 85). Weitere Fragen erkundigten sich nach den Zusammenhängen von Nachdenken und Stirn runzeln, nach Vergnügen und Augenfunkeln sowie zurückgezogenen Mundwinkeln, nach Verachtung und einem leichten Vorstrecken der Lippen gepaart mit einem Emporheben der Nase usw. (ebd.).

Aus dieser Tradition des interkulturellen Vergleichs zur Überprüfung der Universalitätshypothese entwickelten sich weitere Studien, die methodisch verbessert wurden. Eine der wichtigsten Studien in dieser Tradition ist die sogenannte Fore Studie von Ekman und Friesen (1971). Überhaupt ist Paul Ekman hier als die maßgebliche Person zu erwähnen, die in den 1970er Jahren die Ansätze Darwins wieder aufnahm und sie wiederbelebte (Euler 2009: 407). Der amerikanische Psychologe Paul Ekman stellt für mich den nächsten markanten Punkt in meiner Geschichte der *Sehnsucht der Wissenschaft, das Innere des Menschen in seinem Äußeren zu erkennen* dar. Auch wenn zwischen Darwin und Ekman beinahe 100 Jahre liegen, so muss Ekmans Forschung als direkte Fortsetzung von Darwin gesehen werden. Paul Ekman wiederum stellt den Bezug zum Heute dar, denn bis heute beschäftigt er sich mit Gesichtsausdrücken, lässt sich in Interviews als der „weltbeste Experte für Emotionen“ (SZ, 24.01.2009) bezeichnen und kann wohl als einer der wichtigsten Bezugspersonen für die zeitgenössischen Technologien zur Gesichtsausdruckserkennung angesehen werden.

Paul Ekman und Emotionen: The way to change the world is research

Einen markanten Punkt in der Geschichte der *Sehnsucht der Wissenschaften, das Innere des Menschen in seinem Äußeren zu erkennen* stellt für das 20. Jahrhundert der amerikanische Psychologe Paul Ekman dar. Im Jahr 1965 begann er mit seiner Forschung zu Gesichtsausdrücken. 2001 wurde Ekman von der *American Psychological Association* als einer der einflussreichsten Psychologen des 20. Jahrhunderts erwähnt und seit Anfang 2009 gibt es sogar eine TV – Serie („*Lie to me*“), die auf seiner Person basiert. Auch für die Soziologie stellt Paul Ekmans Forschung einen wichtigen Bezugspunkt im Bereich der nonverbalen Kommunikation dar (vgl. Giddens 2001: 84f.).

In diesem Abschnitt erzähle ich die Geschichte von Paul Ekman, die wohl als eine wahre Erfolgsgeschichte eines Wissenschaftlers des 20. und 21. Jahrhunderts angesehen werden kann. Die Person Paul Ekman ist dabei von seiner Forschung zu Gesichtsausdrücken ebenso wenig zu trennen, wie es schon bei Lavater und der Physiognomik sowie bei Darwin und der Evolutionstheorie der Fall war. Zunächst will ich Paul Ekman selbst und seinen Weg in die Forschung zu Gesichtsausdrücken vorstellen, bevor ich auf einige entscheidende Aspekte seiner Forschung eingehe: Einerseits auf die Frage nach der Universalität von Gesichtsausdrücken und andererseits die Rolle der Basisemotionen und Darbietungsregeln von Emotionen. Gerade für die zeitgenössischen Technologien zur automatischen Gesichtsausdruckserkennung von großer Bedeutung ist das von Ekman und seinem Kollegen Friesen entwickelte *Facial Action Coding System (FACS)*, ein System zur objektiven Messung von Gesichtsausdrücken, welches ich anschließend beschreibe. Abschließend gehe ich noch auf die Rezeption und Kritik an der Forschung Ekmans ein, welche vor allem in der Literatur der Emotionspsychologie zu finden ist.

Paul Ekman (geb. 1934): Person und Werk

Paul Ekman wurde 1934 in Washington D.C. geboren und wuchs in Newark, New Jersey auf. Mit 15 Jahren kam er auf die University of Chicago und studierte auch auf der University of New York. Seinen Doktorgrad (PhD) erhielt er 1958 auf der New Yorker Adelphi University in Klinischer Psychologie. Zuvor war er bereits Praktikant am Langley Porter Neuropsychiatric Institute, wohin er nach einer zweijährigen Anstellung als Psychologe bei der Armee auch wieder zurückkehrte. Schon in seiner Masterarbeit interessierte er sich für nonverbale Kommunikation und auch seine Doktorarbeit beschäftigte sich mit Gesichts- und Körperbewegungen von Menschen (www.paulekman.com/about-ekman/), wobei er zu Beginn in den 1950er Jahren weniger an Gesichtsausdrücken, sondern vielmehr an den

Bewegungen der Hände interessiert war (Ekman 2007: 1). Eigentlich wollte Ekman auch Psychotherapeut werden, da er der Meinung war, dass man mit dieser Arbeit die Welt verbessern könnte. Später erkannte er aber, dass der Weg, mit dem man die Welt verbessern konnte, nicht die Psychotherapie, sondern die Forschung war. Während seiner Zeit bei der Armee bemerkte er, dass seine Arbeit dort wirklich etwas bewirkte und stellte daraufhin fest: „Oh my God, the way to change the world is research.“ (Ekman Interview 2004: 1) Und er bekam auch seine Chance in der Forschung: Durch einen „Glücksfall“ bekam er 1965 von der *Advanced Research Projects Agency (ARPA)* des *U.S. Department of Defense*—die heutige *Defense Advanced Research Projects Agency (DARPA)*—den Auftrag und die Förderung für Grundlagenforschung im Bereich der interkulturellen Studien des nonverbalen Verhaltens. Dabei hatte Ekman diese Förderung nur bekommen, weil ein anderes Projekt wegen eines Skandals gescheitert war und das dafür vorgesehene Geld des Verteidigungsministeriums für nicht kontroversielle Forschung in Übersee ausgegeben werden musste.

„By accident I happened to walk into the office of the man who had to spend the funds. He was married to a woman from Thailand and was impressed by differences in their nonverbal communication.“ (Ekman 2007: 2)

Ekman sollte in diesem Projekt herausfinden, welche Gesichtsausdrücke und Gesten universell und weltweit vorkommen und welche Gesichtsausdrücke kulturspezifisch sind. Er fühlte sich der Herausforderung zunächst nicht gewachsen, weil er nicht als Anthropologe geschult war und er der Meinung war, dass das die Voraussetzung für eine interkulturelle Studie war. „But they [the ARPA], for their own reasons, wrote the proposal and they gave me more money than I could spend in five years.“ (Ekman Interview 2004: 2). Ekman, der ja schon zuvor bei der Armee als Psychologe eine Anstellung hatte, hatte durch die großzügige Forschungsförderung des Militärs plötzlich die Möglichkeit, in aller Ruhe zu forschen. Ekmans Karriere als Forscher wurde wohl wie so vieles zu dieser Zeit vom Vietnam Krieg geprägt (The Guardian 10.07.2003) und der mit dem Krieg verbundenen Sehnsucht der Vereinigten Staaten, zu erkennen, wie das Innere des Feindes aussieht. Die thailändische Frau des Verantwortlichen der ARPA kann als eine harmlose Vorschiebefigur für den fremden, verfeindeten Soldaten des kommunistischen Nordvietnam angesehen werden. Dabei ist die Verbindung von militärischen Interessen und wissenschaftlicher Forschung keine Seltenheit. Seit dem Zweiten Weltkrieg hat Militärforschung etwa 40% der weltweiten finanziellen Aufwendungen für Forschung und Entwicklung verschlungen (Felt, Nowotny, Taschwer 1995: 198). Die Trennbarkeit von reiner Grundlagenforschung und militärisch relevantem und anwendbarem Wissen ist jedenfalls nicht aufrechtzuerhalten (ebd. 204) und am prominenten Beispiel des Manhattan State Projects, also dem Bau der US-

amerikanischen Atombombe, zeigte sich deutlich der dominante Einfluss des Militärs auf das Wissenschaftssystem (ebd.: 199f.).

Dafür, dass Ekmans Forschung direkt in die Kriegsführung im Vietnamkrieg eingeflossen ist, habe ich bisher keine direkten Hinweise gefunden und möchte diese Frage an dieser Stelle auch nicht weiter behandeln. Zunächst sieht es so aus, dass Ekman—und hier sei zu erwähnen, dass er den Großteil seiner Forschung gemeinsam mit Wally Friesen betrieben hat (vgl. Ekman 2007)—tatsächlich reine Grundlagenforschung betreiben konnte. Seine interkulturellen Studien zur Frage der Universalität des Gesichtsausdrucks und die daraus resultierenden Basisemotionen sowie das Facial Action Coding System (FACS), worauf ich inhaltlich gleich anschließend eingehen werde, beantworten grundlegende Fragen, die auch schon Darwin ein Jahrhundert früher in einem anderen Kontext—in der Frage nach der gemeinsamen Abstammung von Mensch und Tier—beschäftigten. Dennoch, und das wird sich an der Rezeption von Ekmans Forschung zeigen, sind seine grundlegenden Forschungsarbeiten, welche 1965 starteten, heute eine Basis für die US-amerikanische Landesverteidigung, welche insbesondere durch den Terroranschlag des 11. September 2001 legitimiert wird.

Die Frage der Universalität von Gesichtsausdrücken

Die grundlegende Forschungsfrage von Paul Ekman war zu Beginn, ob die menschlichen Gesichtsausdrücke weltweit universell vorkommen oder ob sie kulturell unterschiedlich sind und deshalb erst erlernt werden müssen. Es ging um die Frage, ob die unterschiedlichen Gesichtsausdrücke biologisch und evolutionär bedingt sind, also quasi bei jedem Menschen ähnlich schon „von Natur aus“ da sind, oder ob es in unterschiedlichen Kulturen jeweils unterschiedliche Gesichtsausdrücke gibt, die in anderen nicht vorkommen. Am Anfang seiner Forschung war Ekman fest davon überzeugt, dass Gesichtsausdrücke nicht universell vorkommen, somit kulturell verschieden sind (Ekman 1988: 18) und sowohl Ausdrücke als auch Gestik erst sozial erlernt werden müssen. Er war sogar so davon überzeugt, dass er zunächst nicht einmal Darwins berühmtes Buch *Der Ausdruck der Gemütsbewegungen bei den Menschen und Tieren* von 1872 gelesen hat (Ekman 2007: 2). Retrospektiv kann aber Charles Darwins Buch als einer der wichtigsten Einflüsse für Paul Ekman gesehen werden. Neben Darwin nennt Ekman auch den im Kapitel über Darwin erwähnten Neurologen Duchenne de Boulogne und den Psychologen und Philosophen Silvan Tomkins als seine wichtigsten Ideengeber (Ekman Interview 2004: 2). Tomkins kommt dabei in pragmatischer Hinsicht eine besondere Bedeutung zu, denn er hatte unmittelbar vor Ekman 1962 und 1963 zwei Bücher über Emotion geschrieben und war der Auffassung, dass Gesichtsausdrücke

universell vorkommen und angeboren sind, er hatte aber dafür keine Evidenz und keine Beweise (Ekman 2007: 2). Tomkins hatte aber eine außergewöhnliche, quasi-intuitive Fähigkeit, Gesichter zu lesen, „Tomkins may have been the best face reader there ever was“ (Gladwell 2002: 41). Von Silvan Tomkins wurden unglaubliche Geschichten erzählt: Als Doktorand in Harvard lebte er einige Zeit sehr gut in Manhattan von Pferdewetten und hatte dort bei den Pferderennen auf der Tribüne den Spitznamen „Professor“. Tomkins beobachtete mit seinem Fernglas die Pferde und konnte auf der Grundlage der emotionalen Beziehungen der Pferde voraussagen, wie die Pferde abschneiden werden. Tomkins konnte aber nicht nur die Gesichter der Pferde „lesen“, auch die Gesichter von Verbrechern auf den „Wanted“-Postern in den Postämtern sagten ihm, welches Verbrechen diese Personen begangen hatten (ebd.). Diese intuitive Fähigkeit von Tomkins spornte Ekman unter anderem dazu an, sich genauer mit dem Gesicht zu beschäftigen, da es eine „Goldmine an Informationen“ bereitstellte (ebd.). Zunächst aber war die Streitfrage der Universalität oder Kulturalität von Gesichtsausdrücken zu beantworten. Wie bereits erwähnt, hatte schon Darwin die Vorstellung von Emotionen als Instinkte, die universell vorkommen und somit angeboren sind. Diese Vorstellung der Emotionen als Instinkte blieb zunächst bis ins 20. Jahrhundert erhalten, wurde dann aber nach dem 1. Weltkrieg mit dem neuen „milieutheoretischen Zeitgeist“ in Psychologie, Anthropologie und Soziologie degradiert. Instinkt wurde zum fachlichen „Unwort“ und Emotionen zum „undifferenzierten Plural“, einem Störfaktor in Experimenten und zu einem Hindernis im vernünftigen sozialen Umgang. Eine Wiederbelebung erfuhr diese „evolutionäre Emotionstheorie“ erst wieder mit Paul Ekman in den 1970er Jahren (Euler 2008: 407), wobei hier sicher auch Tomkins schon in den 1960er Jahren erwähnt werden muss. Ekman hatte aber insbesondere durch die Förderung der ARPA die Möglichkeit, „to settle it once and for all: Are expressions universal, or are they, like language, specific to each culture?“ (Ekman 2007: 3). Auf der einen Seite sah Ekman die UniversalistInnen, die von Vererbung ausgingen: Neben Darwin seien hier Floyd Allport (1924) und eben Silvan Tomkins (1962, 1963) zu erwähnen. Auf der anderen Seite waren die RelativistInnen, die einen kulturspezifischen Standpunkt einnahmen: Hier bezieht sich Ekman auf den eher gemäßigten Klineberg (1938), auf La Barre (1947) und vor allem auf Ray Birdwhistell (1970), den bekanntesten Relativisten (Ekman 1988: 16). Weitere KulturrelativistInnen zu dieser Zeit waren Margaret Mead, Gregory Bateson und Edward Hall (Ekman 2007: 2).

Ekman versuchte die Frage der Universalität auf zwei Wegen zu lösen: In Kulturvergleichsstudien und in Entwicklungsstudien (vgl. Ekman 1988). In einer ersten Studie mussten in einem Experiment Personen aus fünf unterschiedlichen Kulturen (Chile, Argentinien, Brasilien, Japan und USA) Gesichtsausdrücke auf Fotografien erkennen und

zuordnen. Zur Überraschung Ekmans wurden die Gesichtsausdrücke von allen erkannt (Ekman 2007: 3). Dies war aber noch kein Beweis für die Universalität für Gesichtsausdrücke. Die untersuchten westlichen Kulturen könnten die Gesichtsausdrücke über Filme, Magazine und Fernsehserien erlernt haben. Deshalb entschied sich Ekman, eine visuell isolierte Kultur zu untersuchen, die keinen Fernseher, keine Filme und keine Magazine kannte. Ekman zog zunächst umfassendes Filmmaterial einer ebensolchen visuell isolierten Kultur—genauer gesagt von den Fore und von den Kukukuku—heran, das der Neurologe Carleton Gajdusek im Hochland von Papua Neu Guinea aufgenommen hatte (ebd.: 4f.).

Abbildung 7: Ekel dargestellt von einem South Fore



Quelle: Nature, Vol 452, p. 413

Nach interessanten Erkenntnissen nach der Sichtung und Analyse des Materials ging Ekman aber selbst nach Neu Guinea und besuchte die Fore im südöstlichen Hochland Papua Neu Guineas. Da die Fore keine Schriftsprache hatten musste ihre Sprache über den Umweg von Pidgin ins Englische übersetzt werden. Ekman zeigte Kindern und Erwachsenen der Fore Fotos mit Emotionsgesichtern. Diese mussten sie in einer kurzen Schilderung einer dazugehörigen Geschichte zuordnen, welche beschrieb, was vor und nach dem Gesichtsausdruck passiert ist (ebd.: 6ff.). In einem weiteren Experiment wurden den Fore Geschichten vorgelesen, welche sie mit dem dazu passenden Gesichtsausdruck kommentieren sollten. Von diesen Gesichtsausdrücken wurden Fotos gemacht, welche später amerikanische Studierende beurteilen sollten (ebd.: 11). In diesen kulturvergleichenden Studien wurden bis auf Angst alle Emotionen im oberen Bereich zwischen 64% und 100% wiedererkannt. Angst wurde dabei häufig mit Überraschung

verwechselt. In der Folge untersuchte Ekman weitere zehn Kulturen, wobei die Hypothese der Universalität der Gesichtsausdrücke gestützt wurde. Freude wird in allen untersuchten Kulturen in 90 – 98% der Fälle erkannt, eine Ausnahme bildet Sumatra mit nur 69%. Die niedrigste Wiedererkennungsrates wies aber mit 60% Ekel in Japan auf (Merten 2003: 47f.).

Die Kulturvergleichsstudien von Ekman brachten also zwei grundlegende Ergebnisse: Erstens belegen Beobachter unabhängig von ihrer Kultur bestimmte Formen des Gefühlsausdrucks durch das Gesicht mit den gleichen Bezeichnungen (Ekman 1988: 117) und zweitens zeigen Angehörige verschiedener Kulturen den gleichen Gesichtsausdruck, wenn sie das gleiche Gefühl empfinden. Hier können aber kulturspezifische Darbietungsregeln dazwischentreten (ebd.: 118).

Neben den Kulturvergleichen beschäftigte sich Ekman auch mit Entwicklungsstudien und fand u.a. heraus, dass die Gesichtsmuskulatur schon bei Geburt voll ausgebildet und einsatzbereit ist (ebd.: 122), und dass charakteristische Formen des Gesichtsausdrucks bereits im frühen Säuglingsalter auftreten (ebd.: 123) und Säuglinge im Alter von drei bis vier Monaten differenzierte Reaktionen auf Gesichtsausdrücke zeigen (ebd.: 124). Kinder im Vorschulalter wissen bereits, wie die gebräuchlichsten mimischen Ausdrucksformen aussehen, was sie bedeuten und von welchen Situation sie abhängen (ebd.: 125). In einem weiteren Experiment in diesem Zusammenhang konnte gezeigt werden, dass auch Blindgeborene ähnliche Gesichtsausdrücke haben (Ekman 2007: 14).

Von Basisemotionen und Darbietungsregeln

Im Anschluss an die Ergebnisse der Kulturvergleichsstudien eröffneten sich für Ekman weitere Fragen: Wie viele Ausdrücke sind im Gesicht möglich? Sind auch richtige oder irreführende Informationen in Gesichtsausdrücken zu finden? Entspricht jede Bewegung einer Emotion? (Ekman 2007: 14).

Nach Ekman gibt es sechs bzw. sieben Basisemotionen: Angst, Ärger, Ekel, Freude, Trauer, Überraschung und als siebte Emotion Verachtung. Eine Basisemotion ist dadurch definiert, dass sie bestimmte Kriterien erfüllt. Die Basisemotionen Angst, Ärger, Ekel, Freude, Trauer und Überraschung erfüllen die neun von Ekman postulierten Kriterien für Basisemotionen (Ekman zit. nach Alpers, Mühlberger & Pauli 2009: 424):

- emotionsspezifisch unterschiedliche Zeichen,
- die auch bei anderen Primaten beobachtbar sind
- emotionsspezifische Physiologie

- emotionsspezifische und universell auslösende Ereignisse
- Kohärenz zwischen emotionalen Reaktionssystemen z.B. zwischen Ausdruck und Physiologie
- Schneller Beginn
- Kurze Zeitdauer
- automatische Bewertung
- unkontrollierbares Auftreten ohne eigenes zutun

Nur wenn diese neun Kriterien erfüllt sind, kann man von einer Basisemotion sprechen. Je weniger von diesen Kriterien erfüllt sind, desto größer ist der Anteil von sozialen Konstruktionen am Zustandekommen der kulturell geteilten Emotion (Merten 2003: 61).

Dieser Punkt weist darauf hin, dass Emotionen nicht ausschließlich biologisch bedingt sind, sondern dass sowohl kulturelle als auch biologische Prozesse Emotionen formen und verändern (ebd.: 59). Die kulturellen Einflüsse auf Emotionen zeigen sich vor allem über sog. Darbietungsregeln (*display rules*). Darbietungsregeln sind „Regeln für den Ausdruck von Emotionen in Abhängigkeit von der Situation, in der man sich gerade befindet“ (ebd.: 61). Die Darbietungsregeln sind abhängig von der jeweiligen Kultur und sind über soziale Lernprozesse vermittelt. Es können vier Arten von Darbietungsregeln unterschieden werden: 1. Deintensivierung, 2. Übertreibung, 3. Affektlosigkeit, 4. Maskierung. Welche der jeweiligen Regel wiederum angewendet wird, hängt von individuellen Eigenschaften der Person, dem Status, der sozialen Rolle, dem Geschlecht und dem sozialen Kontext ab (ebd.). Darbietungsregeln sagen ganz einfach aus, „wer wann welches Gefühl wem gegenüber zeigen darf“ (Ekman 1988: 30). Die Regeln sind dabei teilweise so gut gelernt, dass sie automatisch ablaufen und nur bei Nichtbeachtung wirklich auffallen. Ein Beispiel für eine Darbietungsregel kann man bei einem Schönheitswettbewerb finden: Hier ist es die Siegerin, die weint vor Glück und es sind die Verliererinnen, die lachen. Auch auf einer Beerdigung finden wir Darbietungsregeln von Emotionen: Ekman spricht von einer „Hackordnung der Trauer“, die vorgibt, wer in welcher Weise Emotionen zeigen darf. Neben kulturell geschliffenen Darbietungsregeln gibt es auch persönliche Darbietungsregeln. Als Beispiel hierfür erwähnt Ekman, dass eine Frau generell Kummer unterdrücken sollte, aber in einer Situation als Witwe wird Kummer von ihr verlangt (ebd.: 31).

Ein weiterer kultureller Aspekt von Emotionen ist das Bewältigungshandeln (*coping*). Unter Bewältigungshandeln können alle Bemühungen verstanden werden, mit einem Gefühl und der Ursache des Gefühls fertig zu werden. Dabei werden kognitive und organisierte Verhaltensweisen wie Flucht oder Angriff voneinander unterschieden.

Bewältigungstechniken werden erlernt und automatisiert, und sind kulturell und individuell spezifisch (ebd.: 32f.).

Gesichtsausdrücke messen: Das Facial Action Coding System (FACS)

Auch wenn bei Ekman sowohl kulturelle als auch biologische Einflüsse auf Emotionen zu finden sind, hat er sich in der Folge vor allem mit den physiologischen Aspekten von Emotionen beschäftigt. 1978 entwickelte Ekman gemeinsam mit Wally Friesen das sog. *Facial Action Coding System*, welches 2002 grundlegend überarbeitet wurde (Ekman, Friesen & Hager 2002). Dieses System (FACS) ist eine Methode, mit der die über 10.000 möglichen Muskelbewegungen im menschlichen Gesicht kodiert werden können. Von diesen Codes ausgehend kann dann unter anderem auf die sechs bzw. sieben Grundemotionen geschlossen werden. Die Bewegungseinheiten der Gesichtsmuskulatur werden als *action units* bezeichnet, also verschiedene Muskelbewegungen werden zu Gruppen zusammengefasst. Ekman und seine Kollegen entwickelten eine genaue Liste der einzelnen *action units*. AU05 (d.h. *action unit* fünf) ist z.B. „das Heben des oberen Lides wie beim erstaunten Blick.“⁸ Das oberste Ziel bei der Entwicklung von FACS war es, ein umfassendes System zu schaffen, mit welchem sich sämtliche mit dem Auge unterscheidbaren Muskelbewegungen im Gesicht bestimmen lassen. Das Ziel war Vollständigkeit (Ekman 1988: 183). FACS sollte ein umfassendes Kategoriensystem darstellen, das alle visuell unterscheidbaren Gesichtsbewegungen erfassen kann (Kaiser & Wehrle 2009: 525). Bei der Entwicklung des *Facial Action Coding Systems* wiesen Ekman und Friesen ein induktives Vorgehen zurück und leiteten es deshalb aus einer Analyse der anatomischen Grundlagen der Gesichtsbewegungen ab. Jede Bewegung des Gesichts ist demnach das Ergebnis von Muskelanspannungen und –entspannungen. Die Gesichtsbewegungen wurden in die minimalen Einheiten, somit in *action units* zerlegt (Ekman 1988: 183f.). Wichtig ist auch, dass es sich bei FACS um ein beschreibendes Messsystem handelt, welches zwischen Beschreibung und Interpretation unterscheidet. Mit FACS kann das mimische Verhalten zunächst nur beschrieben, aber nicht interpretiert werden. Das Kodieren ist interpretationsfrei, erst nach der Beschreibung und Festlegung der Codes kann eine Bedeutungszuweisung erfolgen (ebd.: 184f.).

Mit Hilfe von FACS ist es aber möglich, aus vorher aufgenommenem Bildmaterial die Emotionen von Personen erkennen zu können. Dazu müssen aber speziell ausgebildete KodiererInnen die Muskelbewegungen zunächst messen und klassifizieren, um später eine

⁸ Eine genauere Beschreibung von AU05 und den anderen action units lässt sich auf einer Homepage der Uni Saarland finden: http://www.uni-saarland.de/fak5/orga/Kurs/Seiten/facs/au5_test.htm [Online verfügbar am 18.11.08]

Emotion zuweisen zu können. Grundsätzlich können sowohl Fotografien als auch Videoaufnahmen verwendet werden. Im Falle von Videomaterial wird es zunächst in Realzeit abgespielt, bis eine mimische Bewegung erkennbar ist. Anschließend wird die erkannte mimische Bewegung mehrmals vor- und zurückgespielt, um die Bewegung genauer analysieren zu können. Die Anfangs- und Endzeit der Muskelbewegung wird festgelegt, was bei Filmmaterial mit 25 Bildern pro Sekunde auf 4/100 Sekunden bestimmbar ist. Bei manchen *action units* werden auch einzelne Intensitätsgrade auf einer fünfstufigen Skala kodiert (Kaiser & Wehrle 2009: 525). Der Kodierungsprozess ist sehr mühsam und langwierig: eine Minute Videomaterial auszuwerten dauert etwa eine Stunde. Deshalb wurde das sog. EMFACS entwickelt, das nur die emotionsrelevanten action units kodiert (ebd.: 526). Deshalb arbeiten ComputerwissenschaftlerInnen weltweit daran, dieses Messen zu automatisieren und dadurch um ein Vielfaches schneller zu machen (Ekman 2007: 14).

Derzeit gibt es weltweit etwa 500 ExpertInnen, die das Facial Action Coding System erlernt haben und anwenden. Nach Ekman kann man es selbst mit Hilfe von Lernmaterialien lernen und ist deshalb nicht auf seine Person angewiesen. FACS wird in vielen unterschiedlichen Bereichen verwendet: Von der Forschung zu Mutter-Kind Beziehungen, über die Evaluierung von Werbung bis hin zur Erforschung des Einflusses von Drogen auf Emotionen (Ekman Interview 2004: 3).

Rezeption und Kritik an Ekmans Forschung

Wie in der Einleitung zu diesem Abschnitt über Paul Ekman schon angeführt, kann Paul Ekman heute als einer der weltweit wichtigsten ForscherInnen für Gesichtsausdrücke und Emotionen angesehen werden. In einem Interview in der *Süddeutschen Zeitung* ließ er sich als der „weltbeste Experte für Emotionen“ (SZ, 24.01.2009) bezeichnen. 2001 wurde Ekman von der *American Psychological Association* als einer der einflussreichsten Psychologen des 20. Jahrhunderts erwähnt und seit Anfang 2009 gibt es eine TV – Serie („*Lie to me*“), die auf seiner Person basiert. „The truth is written all over our faces“ heißt es im Untertitel der Serie, die zunächst in den USA, später auch in einigen europäischen Ländern zu sehen sein wird. Ekman selbst spricht von Horror Geschichten, was mit der Wissenschaft in Hollywood angestellt wird, es macht ihm zwar großen Spaß als Berater der Serie tätig zu sein, „aber die haben natürlich von Wissenschaft wirklich keine Ahnung.“ (SZ, 24.01.2009). Eine an anderer Stelle vorzunehmende Analyse der TV Serie „*Lie to me*“ könnte zeigen, was wirklich passiert, wenn die Forschung zu Gesichtsausdrücken im Fernsehen popkulturell verarbeitet wird. Der Untertitel der Serie verrät aber schon, dass das Gesicht die Quelle für „Wahrheit“ ist.

Neben seiner Beratertätigkeit in Hollywood ist Ekman außerdem Direktor einer kleinen Firma, die Lernmaterialien für emotionale Fähigkeiten herstellt (www.paulekman.com/about-ekman). Darüber hinaus agiert er als „Lügenexperte“ für das FBI und die CIA (SZ, 24.01.2009). Unter anderem konsultierte, bewertete und befürwortete Ekman das SPOT (*Screening Passengers by Observational Techniques*) Programm, welches 2006 von der *Transportation Security Administration* auf 14 US amerikanischen Flughäfen eingeführt wurde, um potentielle TerroristInnen erkennen zu können. TerroristInnen sollen nicht etwa über ihre Physiognomie oder ihre Schädelform entlarvt werden, sondern über auffälliges Verhalten in ihren Gesichtsbewegungen.

„Within the next year or two, maybe sooner, it will be possible to program surveillance cameras hooked to computers that spit out FACS data to identify anyone whose facial expressions are different from the previous two dozen people in line.“ (Ekman 2006).

Jede Person, die andere, nicht als normal definierte und daher auffällige Gesichtsbewegungen zeigt, gilt als verdächtig. Man kann, in Anlehnung an Simon Coles Buch *Suspect Identities*, das die Geschichte der Fingerabdrücke und Kriminalidentifikation erzählt (vgl. Cole 2001), von „*Suspect Expressions*“ sprechen. Potenzielle TerroristInnen und andere kriminelle Personen sollen an verdächtig erscheinenden Gesichtsausdrücken erkannt werden können. Eine Kontroverse über den von Ekman empfohlenen Einsatz von Gesichtsausdruckserkennung, um Terroristen erkennen zu können gibt es bereits: In einem FOCUS Online Artikel (FOCUS Online 17.08.2007) wird Wolfgang Miltner, Professor für biologische und klinische Psychologie an der Universität Jena zitiert, welcher mögliche Fehlerquellen u.a. in den unterschiedlichen menschlichen Gesichtszügen („George Bush zum Beispiel sieht immer so aus, als würde er grinsen.“), aber auch in den unterschiedlichen Gründen für auffälliges Verhalten sieht. Dabei geht Ekman selbst davon aus, dass die eigentliche Intention eines gezeigten Gesichtsausdrucks nicht erkannt werden kann, was er als den „*Othello Fehler*“ bezeichnet hat (Ekman 2003: 218). Othello erkannte Desdemonas ängstlichen Ausdruck als eine Reaktion auf Angst vor Verrat. Dabei verrät der Gesichtsausdruck eben noch nichts über die Ursache des Ausdrucks. Die Angst, dass einem nicht geglaubt wird sieht genauso aus wie die Angst, bei etwas erwischt zu werden. So erzählt Ekman in seinem Artikel in der *Washington Post* von einem Mann, der wegen seiner Gesichtsausdrücke am Logan International Airport in Boston als verdächtig eingestuft wurde, es sich aber bei weiterem Nachfragen herausstellte, dass er auf dem Weg zu einer Beerdigung war (Ekman 2006).

So hoffnungsvoll Paul Ekmans Wissen und seine Expertise im US-amerikanischen Kampf gegen den Terror eingesetzt werden, so vorsichtig muss man damit sein, inwiefern die wissenschaftlichen Ergebnisse zu Gesichtsausdrücken auch als gefestigt und

abgeschlossen angesehen werden können. Denn wissenschaftsintern gibt es durchaus Kritikpunkte an Paul Ekmans Theorie, die in vielen Büchern zur Emotionspsychologie gelesen werden können. So teilt etwa Fridlund nicht den *emotions view* von Gesichtsausdrücken, wonach mimische Verhaltensweisen den emotionalen Zustand einer Person wiedergeben. Nach Fridlund, der einen *verhaltensökologischen Standpunkt* einnimmt, sind mimische Verhaltensweisen lediglich als Signale an InteraktionspartnerInnen zu verstehen. Diese Signale werden intentional gezeigt, um soziale Motive realisieren zu können, sie sagen jedoch nichts über den inneren Zustand des Zeichengebers aus. Die relevanten Signale sind in einer Koevolution von SenderIn und EmpfängerIn entstanden, d.h. das Verstehen und Wiedergeben von Signalen hat sich parallel zueinander entwickelt. Überhaupt treten Emotionen häufiger in sozialen Situationen auf, also wenn mindestens zwei Personen miteinander interagieren bzw. wenn man sich implizit einen Empfänger oder eine Empfängerin vorstellt (Fridlund 1991a, 1991b, 1994 zit. nach Merten 2003: 62). Es gibt auch eine Reihe von Befunden, die zeigen, dass der soziale Kontext auf das Zeigen von mimisch-emotionalem Verhalten Einfluss hat (Merten 2003: 62), wobei hier erwähnt werden muss, dass auch Ekman in Form der Darbietungsregeln den kulturellen und sozialen Kontext mitgedacht hat. Generell, so folgert Merten, wird durch die Kritik die kulturübergreifende Verankerung der Basisemotionen nicht grundsätzlich in Frage gestellt, „sondern lediglich eine Überschätzung des biologischen Anteils gefolgert...“ (ebd.).

Eine andere Linie von Kritik betrifft die Reduktion von Emotionen auf ihre körperlich sichtbaren Gesichtsausdrücke. So führen Averill et al. (1990) am Beispiel der eigentlich typischen Emotion Hoffnung vor, dass Hoffnung weder ein typisches Ausdrucksmuster noch spezifische physiologische Reaktionen hat (Averill et al. 1990 zit. nach Merten 2003: 61). Hoffnung kann also bei einer Emotionstheorie, die nur die Gesichtsausdrücke beachtet, nicht berücksichtigt werden, da sie keine typischen Ausdrucksmuster hat. Auch Bottenberg & Dassler gehen in ihrer *Einführung in die Emotionspsychologie* (2002) von einem weiten Verstehens-Panorama von Gefühl aus und öffnen es in acht verschiedene Aspekte, wobei der Ausdruck nur ein Aspekt davon ist (Bottenberg & Dassler 2002: 31). Auch innerhalb des Ausdrucks Aspekts (ebd.: 51ff.) muss zunächst zwischen verbalen und nonverbalen Ausdrücken unterschieden werden. Innerhalb des nonverbalen Ausdrucks gibt es neben dem Gesichtsausdruck weitere „körpermotorische Komponenten“ (ebd.: 64ff.): Den Blick, das Gestikulieren mit Händen und Armen, das Kopfnicken sowie die sensorische Komponenten wie Hautsinne, Geruch oder Geschmack. Weiters die posturale Komponente, also die Stellung und Haltung des Kopfes, auch der Habitus wird hier angeführt sowie ökoergente oder ambientale Komponenten wie der Umweltbezug oder das Setting der Situation.

Zweiter Teil:

Theorien und Konzepte

Der Blick zurück in die Geschichte der Sehnsucht, das Innere des Menschen in seinem Äußeren zu erkennen, sollte aufzeigen, dass Technologien zur automatischen Gesichtsausdruckserkennung nicht einfach nur so aus dem Nichts entstehen, sondern eine kulturelle Vorgeschichte besitzen. In der Folge soll es nun darum gehen, einen genaueren Blick auf die gegenwärtigen Technologien zur Gesichtsausdruckserkennung zu werfen. Doch welchen Blick wirft man darauf? Unter welchen Aspekten untersucht man eine Technologie? Genauso wie es eine historische Verankerung gibt, gibt es auch eine theoretische und konzeptionelle. Diese muss man nicht unbedingt erzwingen, man könnte sie in der Welt des impliziten Wissens zurücklassen. Doch denke ich ist es angemessener, einen Blick auf meinen theoretischen Blickwinkel zu werfen, um zu verdeutlichen, wie ich an die Technologien zur automatischen Gesichtsausdruckserkennung herangegangen bin. Dieses Kapitel gibt dem darauf folgenden empirischen Kapitel sozusagen die Rahmung, es verdeutlicht in gewisser Weise mein Vorwissen, mit welchem ich das empirische Material konfrontiert habe.

Ich möchte hier zunächst die Frage stellen, welches Verständnis von Technologie wir eigentlich haben? Was bedeutet Technologie für uns, welche Geschichte hat dieses Verständnis und welche Möglichkeiten gibt es, Technologien zu untersuchen? Dabei spanne ich den Bogen von einem deterministischen Technikverständnis, über ein Technikverständnis der sozialen Formung und Konstruktion bis hin zu einem Technikverständnis, das die Asymmetrie von menschlichen und nicht-menschlichen AkteurInnen in einem Netzwerk überwinden will, und verorte mich schließlich in diesem letztgenannten Bereich.

Ist diese Frage des Technologieverständnisses dem Umfang dieser Arbeit angemessen beantwortet, so möchte ich auf Konzepte und Zugänge eingehen, wie die Produktion von technowissenschaftlichem Wissen verstanden wird. Zunächst einmal soll geklärt werden, was eigentlich wissenschaftliches Wissen ist, was es von anderem Wissen unterscheidet (*Grenzziehungen*) und wie diese Unterscheidung aufrechterhalten und gefestigt wird. Die Begriffe Objektivität, Expertise und Quantifizierung spielen hier eine entscheidende Rolle. Daran anschließend thematisiere ich die besondere Rolle von technowissenschaftlichen Bildern, die in ihrer digitalen Ausprägung realer erscheinen als jemals zuvor. Bevor ich dann

den Sprung zu den Technologien zur automatischen Gesichtsausdruckserkennung mache, gehe ich noch auf das Konzept des Impliziten Wissens von Michael Polanyi und dessen Weiterführung beim deutschen Soziologen Werner Rammert ein.

Ich bin mir dabei durchaus bewusst, dass ich mit dieser Auswahl des theoretischen und konzeptionellen Rahmens nicht alles abdecken kann, was ich in meinem Hinterkopf hatte, als ich mich mit der automatischen Gesichtsausdruckserkennung beschäftigt habe. Dennoch denke ich, dass der Leser oder die Leserin so besser manche Herangehensweisen, Schwerpunktsetzungen oder Analysen verstehen kann. Ich will somit meine Handlungen nachvollziehbar machen, ich will meine Aussagen dadurch selbst reflektieren, denn nur so kann Wissenschaft gelingen (Richter 2001: 15f.)

Welches Verständnis von Technologie haben wir?: Definitionen, Konzepte und Zugänge

Für die einen bedeutet Technologie Segen, für die anderen Fluch. So oft wir beispielsweise das Mobiltelefon lobpreisen, so oft könnten wir es auch verwünschen, wenn wir kein Netz haben oder der Akku leer ist. Der neue Laptop erfreut uns mit Multimedia-Erlebnissen von höchster Qualität, bringt uns aber auch zur Weißglut, wenn er mal wieder abstürzt. Der Fahrkartenautomat ist für die technikversierte junge Person eine Zeitersparnis, für den älteren Menschen aber vielleicht eine undurchschaubare und unsichere Angelegenheit. Dieses einfache, alltagspraktische Verständnis von Technologie ist uns wohl allen bekannt, denn Technik ist mittlerweile in allen Bereichen unseres alltäglichen Lebens vorzufinden und ist unverzichtbarer Bestandteil sozialer Wirklichkeit. Doch wenn man sich ernsthaft mit Technologien zur automatischen Gesichtsausdruckserkennung auseinandersetzen will, so ist ein Technikverständnis, das sich das technische Paradies oder den kulturellen Untergang vorstellt, zu kurz gedacht (Dollhausen 1997: 96). Technik hat viele verschiedene Konnotationen und es scheint eigentlich keinen Bereich mehr zu geben, in dem wir in unserem Leben nicht mit Technik konfrontiert sind. Der Begriff Technik wandelt sich aber im Laufe der Zeit. In der Antike war mit dem Begriff *techne* die Kunst des Machens gemeint. Durch Kunstfertigkeit wurde etwas hergestellt, das eine nützliche Wirkung hatte (Rammert 2007: 15). Krohn geht von einer Abfolge von vier Schritten aus, wie der Technikbegriff in verschiedenen historischen Konstellationen idealtypisch angesehen werden konnte. Das griechische Verständnis von Technik als eine subjektive Könnerschaft durch Erkenntnis von Handlungsregeln, also als eine Kunstfertigkeit, wurde bereits angesprochen. In der Renaissance und der frühen Neuzeit entwickelte sich der Begriff der Technik hin zu einem in Experimenten arrangierten Naturprozess, welcher mittels der Erforschung von Naturgesetzen erzeugbar war. Im 19. Jahrhundert wurde der Technikbegriff als Maschine verstanden, im 20. Jahrhundert entsteht das Muster eines soziotechnischen Systems, welches maschinelle und organisationale Komponenten umfasst und auf der Erkenntnis von deren Kopplungsbedingungen beruht (Krohn 1989: 17). Der Technikbegriff verändert sich also in unterschiedlichen Gesellschaften und den jeweils vorherrschenden Techniken, große technische Revolutionen bringen wiederum Veränderungen in der Gesellschaft mit sich, denken wir nur an Begriffe wie *Maschinenzeitalter*, *Industriegesellschaft*, *Atomzeitalter*, *Risikogesellschaft*, *Computerzeitalter* oder *Netzwerkgesellschaft* (Rammert 2007: 19).

Verschiedene Autoren und Autorinnen haben versucht, den Begriff Technik zu definieren. Vitti sieht beispielsweise Technik als „a system based on the application of knowledge, manifested in physical objects and organizational forms for the attainment of specific goals“

(Volti 1995:6 zit. nach Degele 2002: 18). In diesem soziologischen Technikbegriff wurden sowohl die Materialität von Technik, das verwendete Wissen und das Vorhandensein einer Handlung berücksichtigt (Degele 2002: 19). Ziele und gewünschte Effekte spielen auch in der Definition von Rammert eine Rolle:

„Unter Technik verstehen wir demnach die Gesamtheit der in der Gesellschaft kreativ und künstlerisch eingerichteten Wirkzusammenhänge, die aufgrund ihrer Form, Funktionalität und Fixierung in verschiedenen Trägermedien zuverlässig und dauerhaft erwünschte Effekte hervorbringen.“ (Rammert 2007: 17).

Technik wird hier recht allgemein aufgefasst, man kann Technik aber auch unter bestimmten Zugängen verstehen, wie etwa Max Weber, der Technologie als Ausbildung eines neuen Rationalitätstypus verstanden hat oder Niklas Luhmann, der die Hauptfunktion von Technik in der Simplifikation und Reduktion von Komplexität sah (vgl. Mayer 2001: 14f.). Grundsätzlich lassen sich drei Zugangsweisen finden, wie Technik und damit in Verbindung das Verhältnis Technik/Gesellschaft konzeptualisiert werden: Im Technikdeterminismus ist es die Technik, die die Gesellschaft beeinflusst. Im Sozialkonstruktivismus dreht sich dieses Verhältnis, die Gesellschaft ist es, die die Technik formt und konstruiert. Die Akteur – Netzwerk Theorie (ANT) und im deutschsprachigen Raum der Technopragmatismus überwinden diese einseitige Kausalität, indem sie die Unterscheidung von Technik und Gesellschaft entweder ganz auflösen (ANT), oder keine eindeutige Kausalität, sondern eine Verzahnung von Technik und Gesellschaft in hybriden Ordnungen postulieren (Rammert 2007: 21). Im Folgenden möchte ich kurz auf diese unterschiedlichen Herangehensweisen eingehen, um mein Verständnis von Technologie verdeutlichen zu können.

Technologischer Determinismus

Wie schon erwähnt sehen technikdeterministische Herangehensweisen eine kausale Verknüpfung von technischen Innovationen mit sozialen Anpassungen und Wandlungen. Beispiele dafür sind etwa bei Daniel Bell zu finden, der „in der Entstehung ‚intelligenter‘ elektronischer Technologien die materiale Grundlage heraufziehender postindustrieller Gesellschaften“ (Bell 1975 zit. nach Degele 2002: 23) sah oder auch bei Ulrich Beck, der „die Figur einer ausgreifenden Herrschaft der Technik“ herauf beschworen hatte, wenn er „die Politik als Exekution technischer Rationalität“ brandmarkte (Degele 2002: 23). In einem technikdeterministischen Verständnis ist Technik etwas, das sich außerhalb der Gesellschaft vorfinden lässt, das vielmehr die Umwelt der Gesellschaft bildet und somit eine außersoziale Tatsache darstellt. Degele charakterisiert drei Annahmen als den technikdeterministischen Kern (ebd.: 24f.): Die erste Annahme ist, dass technischer Wandel unverursacht geschieht,

eine gewisse Eigendynamik hat und somit als autonom angesehen werden muss. In Anlehnung an Arnold Gehlen und Jürgen Habermas steht die Zweckorientierung von Technik im Vordergrund, die das Mängelwesen Mensch zur technischen Entwicklung treibt. Die zweite Annahme technikdeterministischer Sichtweisen ist, dass technische Entwicklungen vom Willen des Menschen unabhängig sind und somit außer Kontrolle geraten können. Ein einmal eingeschlagener Weg scheint nicht mehr rückgängig machen zu sein. Die dritte Annahme ist schließlich, dass technischer Wandel sozialen Wandel verursacht bzw. ihn determiniert. Das soziale System muss sich nach dieser Logik dem technischen Wandel anpassen.

Social Shaping and Social Construction of Technology

Der Sozialkonstruktivismus entfernt sich entschieden von einem technikdeterministischen Ansatz, weil er Technik nicht mehr als der Gesellschaft extern ansieht, sondern sie in die Gesellschaft integriert (Rammert 2007: 37ff.). Technik wird als ein historisch-gesellschaftliches Projekt, als Produkt und Resultat der Gesellschaft (Jokisch 1982; Rammert 1983) und als sozialer Prozess begriffen (Weingart 1989). In den späten 1980er und frühen 1990er Jahren entwickelten sich in dieser Richtung mehrere Forschungsprogramme, die sich unter dem Begriff Sozialkonstruktivismus vereinen lassen. Im deutschsprachigen Raum entstand die *Technikgeneseforschung* (vgl. Dierkes & Hoffmann 1992; Rammert 1993), international bekannter sind aber *Social Shaping of Technology* und *Social Construction of Technology*. Unter dem Titel *The Social Shaping of Technology (SSI)* stellen Donald MacKenzie und Judy Wajcman zunächst 1985, dann in einer Überarbeitung 1999 fest, dass die soziale Gestaltung von Technik ein Prozess ist, in dem es nicht eine einzige dominante Gestaltmacht gibt (MacKenzie & Wajcman 1999: 16). Deshalb muss man den Begriff soziale Gestaltung in weitere Begriffe wie ökonomische, kulturelle, politische oder organisatorische Gestaltung aufteilen, welche analytisch gesehen auch nicht mehr von der technischen Gestaltung unterschieden werden sollen. Neben der Abwendung vom Technikdeterminismus ist der SSI Ansatz auch durch ein Abwenden von einzelnen ErfinderInnen bzw. dem *Genie* hin zu der beteiligten sozialen Interessensgruppe geprägt (Felt, Nowotny, Taschwer 1995: 189). Generell sehen SSI Ansätze „Innovationen als Prozesse der Auseinandersetzung und des Lernens, in denen verschiedene Interessen artikuliert werden“ (ebd.: 190) an. Während in diesem Ansatz die Formung der Technologie durch verschiedene soziale Prozesse im Vordergrund steht, geht der Ansatz der *Social Construction of Technology (SCOT)* darüber hinaus, indem er auch die Formung der Gesellschaft durch die Technologie mit einbezieht (ebd.: 191). Trevor Pinch und Wiebe Bijker übernahmen für ihre konstruktivistische Theorie Konzepte der Wissenschaftsforschung

und wendeten sie auf die Technikforschung an: SCOT sah den Erfolg einer Technologie von der Stärke und Größe der Gruppen abhängig, welche die Technologie begründeten und sie unterstützten. Es waren also weniger die technischen Prinzipien, die zum Erfolg einer Technologie beitrugen, sondern vielmehr soziale Faktoren, die darüber entschieden, ob sich eine Technologie durchsetzen werde. In diesem Sinne muss Technologie als das Resultat der Interpretationen der *relevanten sozialen Gruppen* definiert werden (Sismondo 2008: 16). SCOT Studien untersuchten Technologien zu einem frühen Zeitpunkt ihrer Entwicklung und verfolgten empirisch den Prozess, wie neue Artefakte sich von interpretativer Flexibilität über Schließungsmechanismen hin zu Stabilität entwickelten (Oudshoorn & Pinch 2008: 544). In diesem Satz sind neben der schon erwähnten *relevanten sozialen Gruppe* drei weitere Grundbegriffe von SCOT enthalten, die diesen theoretischen Zugang sehr gut charakterisieren. Mit *interpretativer Flexibilität* ist gemeint, dass unterschiedliche soziale Gruppen Technologien radikal andere Bedeutungen zuweisen können. So gaben ältere Personen und Frauen dem Fahrrad eine völlig andere Bedeutung, wie es zuvor junge Sportler gemacht hatten. Das Fahrrad entwickelte sich vom gefährlichen Hochrad als Sportgerät, das insbesondere jungen Männern die Lust an der Geschwindigkeit befriedigen sollte, zu einem sicheren und stabilen Fortbewegungsmittel für alle. Das zunächst zu bewältigende Problem der Geschwindigkeit wurde also im Laufe der Entwicklung verstärkt zu einem Sicherheitsproblem, das gelöst werden musste. Unter *Schließungsmechanismen* (*closure mechanisms*) können diejenigen sozialen Prozesse verstanden werden, in denen die *interpretative Flexibilität* einer Technologie eingedämmt wird. Schließlich kommt es zur Stabilisierung einer Technologie, was dadurch charakterisiert ist, dass eine vorherrschende Bedeutung und ein vorherrschender Nutzen einer Technologie einsetzen (ebd.).

Mit SCOT verfolgte man eine dichte Beschreibung, um die Komplexität einer technologischen Entwicklung erfassen zu können. Durch die Generalisierung und den Vergleich von Fallstudien wollte man in Richtung einer *Theorie des soziotechnischen Wandels* gelangen (Bijker 1995: 97). Dies stellte sich jedoch als schwierig heraus, da durch SCOT Fallstudien deutlich wurde, dass sich technologische Entwicklungen eher selten linear durch eindeutig abgegrenzte Stadien oder Bahnen—wie SCOT es verfolgte—beschreiben lassen, sondern dass sie sich vielmehr unregelmäßig und zeitlich verschieden entwickeln. Eine in der Technikentwicklung in früheren Modellen angenommene lineare Abfolge von Innovations-, Entwicklungs- und Diffusionsstadien war nur sehr schwer wiederzuerkennen (Felt, Nowotny, Taschwer 1995: 192f.).

Akteur – Netzwerk Theorie und die pragmatistische Technik- und Sozialtheorie

Die Beziehung von Technologie und Gesellschaft wurde in den bisherigen Ansätzen entweder von technologischem Determinismus (Technik beeinflusst Gesellschaft) oder vom Sozialkonstruktivismus (Gesellschaft beeinflusst Technik) charakterisiert. Ein dritter Ansatz, der diese kausale Gegenüberstellung überwindet ist die *Akteur – Netzwerk Theorie (ANT)* bzw. im deutschsprachigen Raum die technopragmatische Technik- und Sozialtheorie von Werner Rammert. In beiden Ansätzen sollte die Gegenüberstellung von Technologie und Gesellschaft vermieden werden, die technologische und soziale Ordnung entstehen vielmehr gleichzeitig in Ko-Produktion (Rammert 2007: 29f.) oder um es in den Worten von Sergio Sismondo zu formulieren:

“The process of adjusting pieces of technoscience and their environments to each other, or of simultaneously creating both knowledge and institutions is a process of co-production (Jasanoff 2004) or co-construction (Taylor 1995) of the natural, technical and social orders.” (Sismondo 2008: 17).

Die Diffusion und Entstehung einer Technologie kann nicht aufgrund der Erforschung der relevanten sozialen Gruppen erklärt werden, denn diese untersuchten sozialen relevanten Gruppen wurden von der technologischen Innovation selbst transformiert. Kurz gesagt: Sowohl die relevanten sozialen Gruppen, als auch die Technologien verändern sich (Latour 2006: 382). Interessen der sozialen relevanten Gruppen haben in diesem Sinne eine Geschichte, sie können nicht einfach vorausgesetzt werden, sie müssen genauso erklärt werden wie die Entstehung der Technologie selbst (Jasanoff 2004: 20).

In der *Akteur – Netzwerk Theorie (ANT)* wird Technologie als das Ergebnis der von WissenschaftlerInnen und IngenieurInnen geleisteten Arbeit gesehen, die Interessen einer sehr breit gefassten Gruppe von einzelnen AkteurInnen—nicht von relevanten sozialen Gruppen—zu übersetzen, wobei diese AkteurInnen sowohl menschlicher als auch nicht-menschlicher Art sein können (Sismondo 2008: 16). ANT geht also davon aus, dass es keine sozialen Gruppen von Anfang an gibt, demnach sollten SoziologInnen die Idee aufgeben, soziale Gruppen am Beginn einer Analyse zu definieren und festzulegen, wie sie aussehen. Die soziale Gruppe muss erst von den beteiligten AkteurInnen geformt und „*performt*“ werden. Die Aufgabe der SoziologIn soll vielmehr sein, den AkteurInnen zu folgen, sie zu beobachten, und den Prozess der möglichen Gruppenbildung festzuhalten und zu beschreiben. In diesem Prozess muss bedacht werden, dass die SoziologInnen auch eine aktive Rolle spielen können, sie können genauso wie andere AkteurInnen Teil der Gruppenbildung werden (Latour 2007: 50ff.).

Die *Akteur – Netzwerk Theorie (ANT)* ist ein Konzept zur Erklärung wissenschaftlicher und technischer Innovationen, das in den 1980er Jahren in Frankreich vor allem durch Schriften von Bruno Latour und Michel Callon entstanden ist und sich über den angelsächsischen Sprachraum mittlerweile auch in deutschsprachige Länder ausgebreitet hat. Vornehmliches Ziel der ANT ist es, traditionelle Unterscheidungen zwischen Gesellschaft und Natur bzw. Technik oder von Menschen und Nicht-Menschen aufzubrechen. An diese Stelle tritt der Begriff des Netzwerks (Schulz-Schaeffer 2000b: 187), durch welchen Wissenschafts- und Technikentwicklung erklärt werden sollen:

„Wissenschafts- und Technikentwicklung, so die zentrale These, ist das Resultat der Verknüpfung heterogener Komponenten zu Netzwerken, ein Prozess, der in dem Maße erfolgreich ist, in dem es gelingt, die beteiligten Komponenten dazu zu bringen, sich in einer aufeinander abgestimmten Weise zu verhalten. Solche Prozesse des Netzerkbildens beruhen stets auf einer doppelten Innovation: der Einrichtung oder Veränderung von Beziehungen zwischen den Komponenten des entstehenden Netzwerks und der Konstruktion oder Veränderung der Komponenten selbst.“ (ebd.: 188).

Die *Akteur – Netzwerk Theorie (ANT)* geht im Gegensatz zu konventionellen soziologischen Ansätzen davon aus, dass Performanz im Gegensatz zu Ordnung die Regel ist (Latour 2007: 63). Für konventionelle SoziologInnen, die Latour die *SoziologInnen des Sozialen* nennt, existiert die soziale Welt bereits von vornherein. Für SoziologInnen der Assoziationen, also SoziologInnen, die sich der ANT verschreiben, kann weder die soziale Welt, die Gesellschaft, noch das Soziale von vornherein bestehen (ebd.: 65). In der ANT bewegt sich immer etwas, es wird immer etwas gemacht. An diesem Punkt sind wir an den Wurzeln der ANT angelangt, welche sich in den *Labor Studien* finden lassen, die Latour zusammen mit Steve Woolgar Ende der 1970er Jahre durchführte. Der Ausgangspunkt der *Akteur – Netzwerk Theorie* ist die Konstruktion von wissenschaftlichen Tatsachen. In diesem Zusammenhang vermittelte das Wort „konstruiert“ fälschlicherweise, dass etwas nicht wahr sei. Insbesondere NaturwissenschaftlerInnen fühlten sich dadurch angegriffen, da sie die soziale Konstruktion so verstanden, als dass ihre naturwissenschaftlichen Ergebnisse falsch seien. Bruno Latour versucht dieses Missverständnis aufzuklären: Für ihn und seine KollegInnen war es nicht von Bedeutung, ob etwas wahr oder falsch sei, sondern vielmehr war interessant, ob eine vorhandene wissenschaftliche Tatsache gut oder schlecht konstruiert sei. Vielleicht sollte man hierbei anmerken, dass das Wort konstruiert nicht als hergestellt im Sinne von nicht wirklich verstanden werden sollte (ebd.: 155ff.). Latour führt weiters an, dass man den von ihm vertretenen Konstruktivismus nicht mit Sozialkonstruktivismus verwechseln sollte. Er erklärt eine konstruierte Tatsache nämlich folgendermaßen: „Die solide objektive Realität wird erklärt, indem wir verschiedene Entitäten

mobilisieren, deren Zusammensetzung auch scheitern könnte.“ (ebd.: 158). Es handelt sich also trotz der Konstruiertheit um etwas Wahres, das eben unter bestimmten Bedingungen in einer bestimmten Art und Weise zusammengesetzt wurde. Im Gegensatz dazu kritisiert Latour den Sozialkonstruktivismus scharf: In diesem wird die Realität durch einen anderen Stoff erklärt bzw. ersetzt. Bei diesem anderen Stoff handelt es sich um das Soziale im konventionellen Verständnis, welches erklärt, wie die Realität „in Wirklichkeit aussieht“. (ebd.: 158). Latour stellt hiermit soziale Erklärungen überhaupt in Frage. An anderer Stelle bezeichnet er sie sogar als „unfair, blind, obsessiv“ (ebd.: 178). Im Gegensatz zu nicht-menschlichen AkteurInnen spricht Latour dem Sozialen die Handlungsfähigkeit sogar ab. Er fragt sich, wieso das Soziale die Hauptrolle spielt, indem es alles Mögliche erklärt, ohne dabei aber selbst etwas zu tun. Er kritisiert damit, wenn Soziologen oder Soziologinnen sagen, dass eine bestimmte soziale Kraft hinter diversen Erscheinungen liegt, also z.B. hinter der Kunst, der Religion, der Wirtschaft, der Technologie usw. In seiner Auffassung ist das Soziale „noch mysteriöser als die Heilige Dreifaltigkeit“ (ebd.: 178). Was Latour hingegen fordert ist Naivität und Kurzsichtigkeit gegenüber dem Sozialen. Nach seiner Auslegung scheint jeder zu wissen, was denn das Soziale genau ist und was es kann und tut. Er beruft sich dabei auf sozialgeschichtliche Literatur, welche eine Vielzahl an Dingen, wie etwa die Entstehung des modernen Staates oder den Aufstieg des Kleinbürgertums durch das Soziale erklärt und fragt sich dabei, ob es wirklich eine einzige Ursache geben kann, „von der eine Million Wirkungen ausgehen“. Die SozialtheoretikerInnen würden diese Aussage, dass die Ursache „*Soziales*“ alles andere erklärt natürlich verneinen, aber Latour fragt sich, welche Rolle sie denn tatsächlich sozialen Kräften geben (ebd.: 179f.).

Um Aufschluss darüber zu bekommen geht Bruno Latour auf die Geburtsstunde der Akteur – Netzwerk Theorie bzw. der „Soziologie der Übersetzung“ zurück. Für ihn war „der Rubikon überschritten, als Verbindungen von ehemals nicht-sozialen Objekten berücksichtigt werden mussten“, also wo in etwa Mikroben, Muscheln und Korallen mit in die Untersuchung und Analyse eingebunden werden mussten. An diesem Punkt musste man sich fragen, ob all diese Objekte aus der Sozialtheorie verstoßen werden sollten, weil sie ja nicht den Anschein hatten, sozial zu sein oder ob sie in ihr aufgenommen werden würden. Die Folge des letzteren ist gewesen, dass das Konzept des Sozialen tiefgehend verändert werden musste. Die Akteur – Netzwerk Theorie war die Folge (ebd.: 183).

Bisher schienen die „konventionellen Soziologen“ zu viel Respekt vor naturwissenschaftlichen Tatsachen und den Objekten zu haben, was nach Meinung von Ludwig Fleck falsch war. Latour wirft die Frage auf, wie man davon ausgehen kann, dass eine Tatsache so verfestigt und solide ist, wo doch schon das Wort *Tat* - *Sache* selbst aussagt, dass es um etwas geht, dass mit dem *Tun* in Verbindung gebracht werden muss.

Tatsachen können aus dieser Sicht also nicht mehr als gefestigt, sondern müssen als umstritten angesehen werden (ebd.: 195). Die Akteur – Netzwerk Theorie sollte dafür sorgen, dass die natürlichen Gegenstände und Objekte von den unbestreitbaren Tatsachen gelöst werden und vielmehr als umstrittene Tatsachen angesehen werden sollten. Gleichzeitig müsse man aber auch die menschlichen AkteurInnen vom Sozialen befreien. D.h., das Soziale als Erklärung für ein bestimmtes Verhalten von Menschen darf nicht länger bestimmend sein (ebd.: 199).

Für mich stellt die *Akteur - Netzwerk Theorie (ANT)* ein passendes Verständnis von Technologie zur Verfügung, da sie die empirische Analyse und die dichte Beschreibung dessen, was passiert und wahrgenommen werden kann, in den Vordergrund stellt und nicht mit vorgefertigten Mustern agiert. Sie bietet mir die notwendige Offenheit, um mich den Technologien zur Gesichtsausdruckserkennung und damit dem Maschinensehen annähern zu können, ohne dabei in die Falle voreiliger theoretischer Schlüsse zu tappen. Dennoch wäre es vermessen zu sagen, dies ist eine ANT Studie. Ich habe zwar durchaus ein ANT Verständnis, aber die folgende empirische Analyse ist ebenso von anderen theoretischen und konzeptionellen Strömungen beeinflusst, von denen ich einige weitere in der Folge anführen werde, wenn ich frage, wie Wissen denn überhaupt produziert wird.

Wie wird Wissen produziert? Konzepte und Zugänge

Die Feststellung, dass Tatsachen umstritten sind kann wohl am Anfang dieses Unterkapitels stehen, wenn ich der Frage nachgehen möchte, wie denn Wissen und somit auch Tatsachen und Erkenntnis produziert (vgl. Knorr Cetina 1984) werden. Knorr Cetina und KollegInnen hatten in den frühen 1980er Jahren die Absicht, durch ethnographische Mikrostudien zu beschreiben, wie wissenschaftliches Wissen in der alltäglichen Laborpraxis produziert wird (vgl. Knorr Cetina 1983). „Anstatt Wissen als eine Repräsentation von Wirklichkeit zu analysieren, kann man es als aus dieser Wirklichkeit fabriziert ansehen.“ (Knorr 1980: 227). Wissenschaftliches Wissen stellt nach dieser Ansicht also nicht die Wirklichkeit selbst dar, sondern es ist nur ein Teil der Wirklichkeit, die im Laufe der Forschung fabriziert wird. Wie wird also wissenschaftliches oder besser: technowissenschaftliches Wissen hergestellt? Wie wissen wir eigentlich, was wir wissen und wie wird dieses Wissen und das Wissen über das Wissen legitimiert? Was unterscheidet wissenschaftliches Wissen von nicht wissenschaftlichem Wissen, wo und wie wird da die Grenze gezogen? Wer besitzt die Expertise, um sagen zu können, dieses Wissen ist wissenschaftlich oder einfach nur richtig? Besitzen vielleicht sogar Zahlen diese Expertise, denn Zahlen gelten allgemein als objektiver als menschliche Meinungen? Doch woher kommen die Zahlen überhaupt? Und wer interpretiert sie? Diese Vielzahl an Fragen kann an dieser Stelle nicht ausführlich beantwortet werden, dennoch möchte ich einen kleinen Einblick in Themen der Wissenschaftsforschung geben, die sich mit Fragen wie diesen beschäftigen. Wenn ich frage, welchen Standardisierungen und Reduktionen Wissen unterliegt, um in eine Maschine, die sehen kann, integriert werden zu können und welche Auswirkungen dies wiederum auf das Wissen selbst hat, so muss man darauf achten, wer überhaupt mit welchen Argumenten sagen kann, dass es sich bei dem verwendeten Wissen um valides Wissen handelt und dass es in richtiger Art und Weise angewendet wird. In diesem Zusammenhang verdienen technowissenschaftliche Bilder eine eigene Unterüberschrift, da sie gerade im Zusammenhang mit dem Maschinensehen eine entscheidende Rolle spielen: Die Referenz der Maschinen ist nicht die Realität selbst, sondern die digitalen Bilder als scheinbare Repräsentanten und Beweise dieser Realität. Schließlich ist auch das Konzept des impliziten Wissens in diesem Zusammenhang des Maschinensehens von großer Bedeutung, denn wir Menschen wissen mehr, als wir zu sagen wissen (Polanyi 1985: 14). Wie können wir also den Maschinen sagen, was wir wissen, damit sie dann genauso viel wissen, wie wir wissen?

Was ist technowissenschaftliches Wissen? Grenzziehungen, Objektivität, Expertise und Quantifizierung

Was ist Wissen und was ist Nichtwissen? Diese Frage der Grenzziehung ist auch eng an die Unterscheidung von wissenschaftlichen Fakten und gesellschaftlichen Werten, an die Unterscheidung von ExpertInnen und Laien sowie an die Unterscheidung von objektiv und subjektiv gebunden (Pinkl 2009: 104f.) Thomas Gieryn versucht sich der Frage, was Wissen und was Nichtwissen ist mit der Hilfe von Landkarten zu nähern: „We learn about science by seeing what is far from it, or near“ (Gieryn 1999: 10). Im rechten unteren Teil einer von ihm verwendeten Landkarte ist der *Mount Science* angesiedelt (ebd.: 8f.), um welchen sich Kämpfe um Glaubwürdigkeit finden lassen. In Regionen nahe dem Mount Science befinden sich etwa die *College Mountains*, der *Intelligence River* oder der *State of Intelligence*, weiter entfernt liegen Plätze wie die *Blind Island*, also diejenigen Orte, von denen sich das wissenschaftliche Feld gerne abgrenzen will (ebd.: 8f. zit. nach Pinkl 2009: 110). Dennoch kann diese Grenzziehung nicht als universell angesehen werden, da für die einen etwas als wissenschaftliches Wissen gilt, für die anderen es aber pseudowissenschaftlich ist, was Gieryn am Beispiel der *Phrenologie* festmacht, die auch heute noch sowohl für Wahrheit als auch für Irrtum stehen kann (Gieryn 1999: 121 zit. nach Pinkl 2009: 115).

Dennoch glauben wir zu wissen, dass es für bestimmte Gebiete Experten oder Expertinnen gibt, die objektiv sagen können, dass etwas wahr oder falsch ist. Expertise ist, wie man an Synonymen wie Fachwissen, Sachkunde oder Expertenwissen unschwer erkennen kann, sehr stark an Wissen gekoppelt. Im Gegensatz zu einem Experten oder einer Expertin steht traditionell der Laie, dem es an Wissen mangelt (Evans & Collins 2008: 609). Doch wie kann man klar zwischen Experten und Laien trennen? Um technowissenschaftliche Glaubwürdigkeit zu erreichen und aufrechtzuerhalten, müssen Grenzen geschaffen und diese im Zaum gehalten, kontrolliert werden (ebd.:610). Die Wissenschaftsforschung konnte aufzeigen, wie diese Grenzen aufrechterhalten wurden und welche Schwierigkeiten sich daraus ergeben haben, als man das wissenschaftliche Expertenwissen aus dem Labor in die der Wissenschaft außen stehende Welt geholt hat. Die wissenschaftliche Expertise reichte nicht mehr aus, die Kontrolle wurde reduziert, neue Komplexitäten entstanden. Wissenschaftliche Expertise sollte deshalb durch die Expertise von denjenigen ergänzt werden, die in dem Bereich, wo das wissenschaftliche Laborwissen quasi draußen in der Welt angewendet wird, bereits ausreichend Erfahrungen gemacht haben (ebd.: 611). Diese Kritik der Wissenschaftsforschung hat meiner Meinung nach aber das Problem, dass gerade individuelle Erfahrungen einen subjektiven Charakter haben und sie damit nicht vereinbar mit Objektivität zu sein scheinen, die für die Wissenschaft von außerordentlicher Bedeutung ist.

Objektivität wird heute als Synonym für das Empirische, für das Faktische, für das Wissenschaftliche, für Unparteilichkeit oder für das Rationale gebraucht (Daston 2001: 127).

Objektivität ist aber nicht gleich Objektivität und vor allem ist sie nicht gleich Wahrheit. Denn Objektivität hat eine Geschichte (vgl. Daston & Galison 2007; Daston 2001). Im Begriff objektiv haben sich im Laufe der Zeit verschiedenste Dinge abgelagert. So muss man zwischen verschiedenen Arten von Objektivität unterscheiden:

„Bei der ontologischen Objektivität geht es um die Welt, bei der mechanischen Objektivität darum, den universellen menschlichen Hang zum Beurteilen und Ästhetisieren auszuschalten. Dagegen zielt die aperspektivische Objektivität darauf, individuelle Eigenarten zu eliminieren.“ (Daston 2001: 129)

Objektivität bedeutet also nicht immer dasselbe, obwohl wir das Wort so selbstverständlich gebrauchen. Gerhard Fasching spricht sogar von einer *Objektiven Illusion* der naturwissenschaftlichen Wirklichkeit und will damit das scheinbar sichere Fundament der Naturwissenschaften in Frage stellen (vgl. Fasching 2007). Fasching beschreibt in einem Essay den Prozess der Quantifizierung, welcher sich einerseits durch das Kommensurabilitätsprinzip, also dem Prinzip, dass etwas mit gleichem Maß messbar und vergleichbar ist und andererseits durch Metrisierungsregeln, entwickelt hat (ebd. 22ff.). Entscheidend für die Frage nach der Objektivität ist, dass es eine Entwicklung, einen Prozess gibt, welcher gerne übersehen wird und welcher z.B. verdeutlicht, dass es keine Art gibt,

„...die Zeit zu messen, die richtiger ist, als eine andere; die, die allgemein angewendet wird ist nur bequemer. Wir haben nicht das Recht, von zwei Uhren zu sagen, dass die eine richtig gehe und die andere falsch, wir können nur sagen, dass die eine vorteilhafter ist, sich nach den Angaben der ersteren zu richten.“ (ebd. 26).

Wenn also etwas gemessen wird, so ist das nicht die einzige gültige Wahrheit, sondern es gäbe die Möglichkeit, auch anders zu messen. Man muss sich nur vorab darauf einigen (Porter 1995: 33). Quantifizierungen und Objektivität haben dennoch ein enges Verhältnis, dies sei aber nach Theodore Porter nicht etwa darauf zurückzuführen, dass Formen der Quantifizierung die Wirklichkeit genauer wiedergeben würden, sondern weil sie besser übermittelbar waren, insbesondere dort, wo die Barrieren der Entfernung und des Misstrauens überwunden werden mussten (vgl. Porter 1995). Zahlen waren also vor allem zu Zwecken der Kommunikation in der wissenschaftlichen Gemeinschaft nützlich (Porter 1995: viii). Dennoch kann man kaum darüber hinwegsehen, dass sehr oft Behauptungen und Annahmen gerade durch Quantifizierungen die entsprechende Objektivität bzw. Evidenz erhalten. Zahlen und andere Methoden der Quantifizierung basieren auf einer allgemein

anerkannten Sprache, einer Sprache der harten Fakten, die scheinbar zuverlässiges Wissen erzeugen (Pinkl 2009: 129ff.). „Zahlen signalisieren Unbestreitbarkeit und Objektivität. Die Objektivität von Zahlen ist jedoch kein Sachverhalt, sondern eine Zurechnung.“ (Heintz 2007: 80). Diese Zurechnung hat sehr viel damit zu tun, dass man bei Zahlen die Handschrift der Autorinnen oder Autoren nicht mehr erkennen kann, die sie produziert haben. Zahlen werden somit als interpretationsfreier Raum angesehen und können so, ähnlich wie wissenschaftliche Bilder überzeugen und Akzeptanz einer oder über eine Tatsache schaffen. Diese Überzeugungskraft wird wiederum durch den bekräftigenden Charakter von Zahlen geschaffen, denn im Gegensatz zu einem Satz können Zahlen nicht negiert werden (ebd.: 80f.). „What is counted usually counts.“ (Miller 2001: 386 zit. nach Heintz 2007: 81). Und was als Zahl übermittelt oder dargestellt wird, hat eine besondere Beweiskraft, denn meistens sind Zahlen zu einem hohen Maße aggregiert, wie sie zustande gekommen sind kann wegen der hohen Komplexität der Produktion aber kaum noch nachvollzogen werden (Heintz 2007: 81).

Die besondere Rolle von technowissenschaftlichen Bildern

Auch technowissenschaftlichen Bildern wird eine besondere Beweiskraft zugeschrieben. Dahinter steht die Annahme, dass wissenschaftliche Bilder der Inbegriff von Objektivität seien und dass ihnen Ästhetik fern sei. Dabei muss man sich aber eher fragen, wie viel Ästhetik erlaubt sei, um den Objektivitätsstatus aufrecht erhalten zu können. Gerade durch das Einsetzen der Fotografie bekam man eine Möglichkeit, die Schwächen des menschlichen Auges wettzumachen und in etwa bei der kriminalistischen Spurensicherung die Spuren eines Tatorts dauerhaft und objektiv zu sichern. Die Fotografie beerbte dabei die „Augenscheinseinnahme“ und hatte somit auch den Anspruch der „unmittelbaren Erfahrung von Wahrheitsmomenten“ (Heßler 2006: 26).

An dieser Stelle zeigt sich die Wissenschaftsforschung skeptisch, ob wissenschaftliche Bilder überhaupt Bilder sind (ebd.: 18f.). Der Bildstatus per se wird in Frage gestellt, da sich die Referenten der Bilder verschoben haben. Die vornehmliche Referenz von wissenschaftlichen Bildern ist nicht mehr die Wirklichkeit, welche sie einfach nur abbildet. Vielmehr stehen die „Konstruktions- und Inszenierungsprozesse“ der Bilder im Mittelpunkt. Die Bilder sollen etwas sichtbar machen, was zumindest das menschliche Auge so nicht sehen kann. Die Bilder werden also konstruiert und zeigen nicht mehr „nur“ die Realität, so wie wir sie sehen, sondern so, wie andere Referenten (z.B. ein Mikroskop oder eine Kamera) sie sehen. Im Fokus der Wissenschaftsforschung steht deshalb nicht (mehr) die Spezifik von Bildern,

sondern die Herstellungs- und Konstruktionsprozesse dieser. Dass Bilder dabei eindeutig sind wird ebenso wie der Abbildcharakter von Bildern in Frage gestellt.

Der britische Kunsthistoriker Martin Kemp verweist darauf, dass moderne Wissenschaften größtenteils eine hochgradig visuelle Angelegenheit sind, „...und niemals so sehr wie im Zeitalter der Computergraphik“ (vgl. Kemp 1997 zit. nach Borck 2001: 383). Der Einsatz von Computertechnik führt in der Herstellung von wissenschaftlichen Bildern nach Cornelius Borck zu einem „Paradox der Transparenz der Technik“ (Borck 2001: 388). Einerseits wird immer mehr Technik in der Herstellung der Bilder verwendet, andererseits erscheinen sie gerade dadurch als besonders real und die Wirklichkeit abbildend. Nach Borck wird hier eine Traditionslinie im Umgang mit wissenschaftlichen Bildern seit dem 19. Jahrhundert fortgeführt: Wissenschaftliche Bilder und Darstellungen sind mit der Rhetorik der Selbstevidenz propagiert worden, also damit, dass sie „die Natur selbst, ohne das störende Dazwischenkommen menschlicher Aktivitäten“, sondern nur durch den Einsatz störungsfreier technischer Gerätschaften, repräsentieren (ebd.: 389). Nach wie vor erscheint die visuelle Evidenz der Bilder entscheidend, obwohl und gerade weil noch mehr Technik verwendet wird. Technik wird hier in dieser spezifischen Konzeption von Objektivität mit Neutralität, Bilder werden mit Evidenz gleichgesetzt (Daston & Galison 1992 zit. nach Borck 2001: 389). Gerade wissenschaftliche Bilder führen zur intuitiven Einsicht, sie stellen eine Augenscheinlichkeit (*lat. evidentia*) dar und ermöglichen einen „interpretationsfreien Einblick“, welcher aber nur durch eine „Einpassung in sozial etablierte, seit langer Zeit existierende Sehgewohnheiten“ entsteht (ebd.: 390). Gerade die Gewohnheit, das Selbstverständliche, das „Normale“ interessiert mich an diesem Punkt am meisten. Insbesondere diese Vorannahmen scheinen die Art und Weise maßgeblich zu beeinflussen, wie WissenschaftlerInnen mit ihren Bildern umgehen und wie diese Bilder letztlich auch Verbreitung finden. Und insbesondere diese Vorannahmen sind es, die letztlich den „interpretationsfreien Einblick“ in wissenschaftliche Bilder ermöglichen. Sie wirken im Endeffekt als objektiv, obwohl eine Vielzahl an kollektiv geteilten Vorannahmen an der Betrachtung und Verwendung beteiligt sind, welche man auch als kulturelle Sehkonventionen bezeichnen könnte. Daneben werden wissenschaftliche Bilder erst im Prozess erzeugt. Technische Geräte spielen dabei eine entscheidende, aber scheinbar neutrale Rolle. Doch auch hier darf man nicht vergessen, dass die technischen Geräte selbst vom Menschen hergestellt sind und dass sich Vorannahmen und Interpretationen selbst in diesen finden lassen.

Heßler spricht in diesem Zusammenhang auch von der Herstellung von Sichtbarkeiten. Diese Herstellungsprozesse geraten aber im historischen Prozess in der Regel in Vergessenheit. Heßler betont hier, was Latour und Woolgar schon Ende der 1970er Jahre

feststellten: Die Konstruktion von wissenschaftlichen Tatsachen erscheint jedem im Nachhinein als unkonstruiert (vgl. Latour & Woolgar 1979). Obwohl unzählige Transformationen, Eingriffe und Interpretationen im Herstellungsprozess von Sichtbarkeiten vollzogen werden, erscheinen wissenschaftliche Bilder schon bald als selbstverständlich. Sie haben einen Normalisierungsprozess durchlaufen, ihre hergestellte Sichtbarkeit erscheint schon bald als ein normales Phänomen, der Herstellungsprozess ist unsichtbar (Heßler 2006: 30f.).

Implizites und nicht-explizites Wissen

An das zunehmende Interesse der Wissenschaftsforschung, wissenschaftliche Bilder und visuelle Repräsentationen zu erforschen, schließt sich die Erkundung des visuellen Wissens an, das diese erzeugen. Hier ergeben sich Fragen nach der Spezifität des Visuellen als Wissensform, also ob das Visuelle eine spezielle Form des Wissens darstellt (Burri & Dumit 2008: 297)? Bisher ist die Form des visuellen Wissens noch unzureichend bestimmt worden, doch gibt es durchaus Ähnlichkeiten zu körperlichem Wissen, wenn man sich etwa auf körperliche Gesten bezieht (vgl. Schnettler & Pötsch 2007). Wenn man Michael Polanyis Verständnis von implizitem Wissen folgt, so erkennt man im impliziten Wissen eine doch nahe Verwandtschaft zum visuellen Wissen. Wenn er passend zu dieser Arbeit schreibt, dass wir Stimmungen vom menschlichen Gesicht ablesen und doch außerstande sind, genauer anzugeben, welchen Zeichen wir dieses Wissen entnehmen (Polanyi 1985: 15), so wird diese Verwandtschaftsbeziehung von körperlichem, visuellem und implizitem Wissen doch recht klar. Wir können all diese Formen des Wissens nicht richtig benennen, wenn wir ihnen begegnen. Michael Polanyi betrachtet das menschliche Erkennen ausgehend von der Tatsache, „dass wir mehr wissen, als wir zu sagen wissen“ (ebd.: 14). Parallel zum Erkennen von Gesichtsausdrücken gibt Polanyi auch das Beispiel des Wiedererkennens von Gesichtern. Er geht davon aus, dass wir, obwohl wir Gesichter wiedererkennen können nicht sagen können, wie wir ein uns bekanntes Gesicht schließlich wiedererkannt haben. Wir können zwar wiedererkennen, aber dieses Wiedererkennen kann nicht in Worte gefasst werden. Die Kenntnis einer Physiognomie kann aber letztlich durchaus mitgeteilt werden, aber nur, wenn wir über angemessene Ausdrucksmittel verfügen würden. Allerdings kann eine explizite Integration die implizite nicht ersetzen, der Prozess der Formalisierung von Wissen, die Polanyi als eine Ausschließung aller Elemente impliziten Wissens versteht, würde sich selbst zerstören (ebd.: 27).

Aufgrund der von Polanyi gegebenen Beispiele kann abgeschätzt werden, wie weitreichend implizites Wissen in unser Leben integriert ist. Da dieses zum großen Teil unterbewusst

operiert und viele Entscheidungen und Vorgänge auf Basis des impliziten Wissens auch gut funktionieren, sind wir normalerweise nicht damit konfrontiert, das implizite Wissen, das wir in einer Situation angewendet haben, um z.B. einen alten Freund wiedererkennen zu können, explizieren zu müssen. Durch den zunehmenden Einsatz von Informationstechnologien auf der Basis von Computern werden aber immer mehr Bereiche des Lebens dem Druck des Explizierens ausgesetzt. Haben sich bisher Entscheidungen oder Arbeitsvorgänge am impliziten Wissen der beteiligten Personen orientiert, so ist man immer öfter damit konfrontiert, dass für die Übertragung menschlicher Arbeitsfunktionen auf informationstechnische Systeme dieses implizite Wissen expliziert werden muss (Rammert 2007: 147). Die einer Handlung oder einer Entscheidung zugrunde liegenden Regeln müssen explizit gemacht werden. Computerprogramme müssen diese Regeln in formalisierter und modellierter Form empfangen, ansonsten können sie nicht damit arbeiten. Rammert sieht diesen Prozess als ein Indiz für eine zunehmende Verwissenschaftlichung und Informatisierung der Gesellschaft (ebd.:). Rammert schließt zwar direkt an Polanyis Konzept des impliziten Wissens an, er stellt aber in der Folge fest, dass implizites Wissen nur eine bestimmte Form von nicht-explizitem Wissen ist. Demnach kann man implizites Wissen nicht als das Gegenteil von explizitem Wissen anerkennen, sondern findet dies im nicht-expliziten Wissen (ebd.: 163). Dennoch nimmt das implizite Wissen in der Diskussion die weit wichtigere Rolle ein. Nach Mead gibt es drei Formen, wie man sich implizites Wissen aneignen kann, was er am Beispiel des Boxens demonstriert. Erstens durch Teilnahme, zweitens durch Realexperimente und drittens durch Videoanalyse der Bewegungen (vgl. Mead 1968 zit. nach Rammert 2007: 150). Garfinkel stellte in einem seiner Krisenexperimente fest, dass es eine Grenze der Explizierung gibt. Er forderte Probanden dazu auf, selbstverständlich erscheinende Worte immer weiter zu explizieren. Da dies aber nur bis zu einem bestimmtem Zeitpunkt und Explizierungsgrad möglich war, verzweifelte sie schon sehr bald an dieser Aufgabe und wurden deshalb zornig (vgl. Garfinkel 1967 zit. nach Rammert 2007: 150). Die Probanden bei Garfinkel waren es einfach nicht gewohnt, etwas explizieren zu müssen. Explizites und exaktes Wissen spielt im alltäglichen Handeln mit anderen Menschen oder der restlichen Umwelt und auch im Umgang mit sich selbst eine untergeordnete Rolle, implizites Wissen hingegen ist von zentraler Bedeutung (Rammert 2007: 151). Aber nicht nur im Alltag, auch in scheinbar sehr expliziten Bereichen wie der Mathematik gibt es implizites, quasi „ungesagtes“ und nicht vollständig explizierbares Wissen, das zum selbstverständlichen Grundbestand des mathematischen Know-hows gehört (Heintz 2000: 175).

Das Verhältnis von implizitem bzw. nicht-explizitem Wissen zu explizitem Wissen ist wie gerade gezeigt wurde kompliziert. Um Maschinen ein für die meisten Menschen quasi

selbstverständliches, implizites Wissen vermitteln zu können, damit sie gewisse Aufgaben übernehmen oder Entscheidungen treffen können, muss dieses Wissen explizit gemacht werden. Nach Collins sind Maschinen und Computer aber überhaupt nicht in der Lage, Wissen verarbeiten zu können (Collins 1990: 8). Rammert folgert daraus:

„Wenn Maschinen wie Computer trotzdem funktionieren, dann tun sie 1.) selten dieselbe Arbeit, die Menschen tun, oder es sieht 2.) manchmal nur so aus, als ob sie wirklich funktionieren, weil die Menschen mit ihnen interagieren, um ihre Schwächen zu kompensieren, oder sie wirken 3.) in dem großen Bereich menschlichen Verhaltens, der Maschinen nachahmt und den daher Maschinen perfekt nachahmen können.“ (Rammert 2007: 158).

Dies hat zur Folge, dass die explizierte Form von implizitem Wissen niemals das gleiche sein kann, wie das implizite Wissen selbst. Man kann diesen Schritt gehen, doch muss man damit rechnen, dass dadurch eine andere Form von Wissen erreicht wird.

Dritter Teil:

Empirische Analyse der automatischen Gesichtsausdruckserkennung

In unserem alltäglichen Leben sind wir ständig mit Emotionen und Gefühlen konfrontiert. Dabei spielt nicht nur der eigene emotionale Zustand eine Rolle, auch unsere zwischenmenschlichen Beziehungen sind ständig von Emotionen und deren Vermittlung betroffen. Dabei stellt sich die Frage, inwiefern man den emotionalen Zustand eines Gegenübers erkennen kann, um dann gegebenenfalls darauf reagieren zu können. Bei dieser Frage besteht die Annahme, dass der emotionale Zustand eines Mitmenschen insbesondere in dessen Gesichtsausdruck erkannt werden kann. Viele Artikel zur automatischen Gesichtsausdruckserkennung stützen diese Annahme auf die Untersuchungen von Mehrabian (vgl. Mehrabian 1968), wo festgestellt wurde, dass nur 7% einer Nachricht von einer Person zur anderen über das gesprochene Wort, 38% über andere vokale Anteile wie z.B. die Sprachintonation, aber 55% der Nachricht über den Gesichtsausdruck übermittelt werden. Die Schlussfolgerung daraus ist, dass der Gesichtsausdruck einen erheblichen Einfluss auf das Empfangen einer Nachricht hat (Chibellushi & Bourel 2002: 1) bzw. sogar die hauptsächliche Modalität in der menschlichen Kommunikation darstellt (Pantic & Rothkrantz 2000: 1424). Was den Gefühlszustand eines Menschen betrifft, kann angenommen werden, dass der Einfluss des Gesichtsausdrucks einen noch stärkeren Einfluss auf die Wahrnehmung dessen ausübt. Folgt man Paul Ekman, so wissen Kinder bereits im Vorschulalter, wie die gebräuchlichsten mimischen Ausdrucksformen aussehen, was sie bedeuten und von welchen Situationen sie abhängen (Ekman 1988: 125).

Da sich immer mehr Interaktionen unseres alltäglichen Lebens mit Maschinen und Computern abspielen, besteht unter IngenieurInnen und ComputerwissenschaftlerInnen die große Sehnsucht, dass Maschinen, Computer, Roboter, Agenten usw. in einer selben Art und Weise wie Menschen die Gesichtsausdrücke und somit in weiterer Folge den inneren Gefühlszustand eines Menschen erkennen und wahrnehmen können. „Der Trend geht zu emotionalen Maschinen“ erklärt der Computerwissenschaftler Bernd Kleinjohann von der Universität Paderborn⁹. Maschinen sollen Emotionen in den Gesichtsausdrücken der

⁹ Quelle: <http://www.wissenschaft.de/wissenschaft/hintergrund/275969.html> [Online verfügbar am 11.08.09]

Menschen erkennen können. Die Annahme besteht, dass Menschen das beste System zur Gesichtsausdruckserkennung besitzen (Pantic & Rothkrantz 2000: 1425) und gleichsam mühelos und ohne Zeitverzögerung Gesichtsausdrücke erkennen können (Chibelushi & Bourel 2002: 1). Aber wie funktioniert das menschliche System, um Gesichtsausdrücke erkennen zu können, wie funktioniert das menschliche Sehen überhaupt? So gut der Mensch scheinbar darin ist, in Interaktionen Gesichtsausdrücke deuten zu können, so schwer tut er sich damit sagen zu können, wieso er zu einem bestimmten Zeitpunkt einen gewissen Gesichtsausdruck wahrnimmt, an welchen Punkten oder Bewegungen er dieses Erkennen fest macht. Versuchen Sie doch einfach einmal, genau zu beschreiben, an welchen Merkmalen Sie einen bestimmten Gesichtsausdruck und somit einen bestimmten Gefühlszustand oder eine Nachricht festmachen und erkennen, wenn Sie das nächste Mal mit einer anderen Person interagieren. Wahrscheinlich haben sich die meisten Personen diese Frage bisher noch nicht gestellt, da es gar nicht notwendig erscheint, sich diese Frage überhaupt zu stellen. Das liegt wohl insbesondere an der Annahme, dass dem Menschen diese Fähigkeit, einen Gesichtsausdruck zu erkennen und ihn richtig zu deuten angeboren ist oder wir es sehr früh lernen (vgl. Ekman 1988). Es ist wohl kaum zu leugnen, dass es einen gewissen biologischen Anteil beim Zeigen von Gesichtsausdrücken und beim Erkennen und Deuten von Gesichtsausdrücken gibt, was die Forschungsarbeiten von Darwin (vgl. Kap. *Von Lavater zu Darwin*) und in weiterer Folge von Ekman (vgl. Kap. *Paul Ekman und Emotionen: The way to change the world is research*) verdeutlichen, allerdings gibt es auch eine erhebliche kulturelle Komponente, die selbst bei Ekman in Form der Darbietungsregeln von Gesichtsausdrücken zum Tragen kommt. Was die Fähigkeit betrifft, einen Gesichtsausdruck richtig deuten zu können wissen wir allerdings sehr wenig, obwohl die Forschungsarbeiten von sowohl Darwin als auch Ekman und seinen KollegInnen gezeigt haben, dass gewisse universell vorkommende Gesichtsausdrücke, die sogenannten Basisemotionen (vgl. Kap. *Von Basisemotionen und Darbietungsregeln*) in verschiedenen Kulturen gleichsam erkannt werden können.

Bei einer Untersuchung von Zucker zeigte sich, dass Menschen nicht so gut beim Erkennen der Basisemotionen sind, wie vielfach angenommen wird. Im Schnitt wurden die sechs Basisemotionen in 64% der Fälle erkannt, wobei es deutliche Unterschiede zwischen einzelnen Emotionen gab. Am besten konnten die befragten 250 Personen Fröhlichkeit (90,5%) erkennen, am schlechtesten in nur 27,9% der Fälle Angst (vgl. Zucker 2007). Dieses Ergebnis ist ein Hinweis dafür, dass Menschen Gesichtsausdrücke nicht so mühelos und schnell erkennen können, wie angenommen wird (vgl. Chibelushi & Bourel 2002). Oder doch? Ich denke, dass es einen Unterschied macht, einen Gesichtsausdruck erkennen zu können und ihn *richtig* erkennen zu können. Aber was heißt hier überhaupt richtig? Im Falle

der Untersuchung von Zucker wurde die Erkennungsrate der Gesichtsausdrücke in Bezug zur sogenannten *ground truth* berechnet, es gab also eine bestimmte ExpertInnenvorgabe, was richtig ist. Diese ExpertInnenvorgabe ergab sich aus der genauen Codierung mit Hilfe des Facial Action Coding System (vgl. Kap. *Gesichtsausdrücke messen: Das Facial Action Coding System (FACS)*), das ja beansprucht, die richtige Interpretation der beschriebenen Gesichtsausdrücke treffen zu können. Wieso gibt es aber eine solche Diskrepanz zwischen der Erkennungsrate von Personen und der „richtigen“ Vorgabe der FACS KodiererInnen? Diese Diskrepanz könnte damit zu tun haben, dass Personen Gesichtsausdrücke in einer je spezifischen Situation der Situation angemessen interpretieren. Erst durch eine kommunikative Erschließung in verschiedenen Situationen ergibt sich die richtige, weil geteilte Interpretation eines Gesichtsausdruckes. Erst in der Praxis einer bestimmten Situation entwickelt sich eine Dynamik, wie man einen Gesichtsausdruck liest und deutet, also den visuell wahrgenommenen Gesichtsausdruck mit Bedeutung anfüllt, sei es auf der von Emotionen oder von anderen Signalen. Hier muss auf den Unterschied zwischen der Erkennung des bloßen Gesichtsausdrucks und auf die Emotionserkennung, die eine Interpretationsleistung im Kontext einer Situation darstellt hingewiesen werden (Fasel & Luettin 2002). An dieser Stelle möchte ich auch das Konzept der *situated actions* von Lucy Suchman (vgl. Suchman 1987; 2007) anführen, um meinem Argument theoretische Anbindung und somit Bodenhaftung zu geben. Nach ihrer Auffassung sagt der Term der *situated action* aus, dass jede Entwicklung einer bestimmten Aktion in essentieller Weise vom damit verbundenen Material und der sozialen Kontextbedingungen abhängt (ebd.: 50). Nach dieser Auffassung kann man also nicht von einem Muster ausgehen, das sich ständig wiederholt und so ohne weiteres abrufbar und erkennbar ist, man muss sich vielmehr auf die direkten Umstände einer Situation konzentrieren, und wieso Personen diese Situation so deuten und dementsprechend handeln bzw. dementsprechend einen Gesichtsausdruck erkennen.

Gesichtsausdrücke und Emotionen betreffend möchte ich dieses Argument an einem Beispiel illustrieren, das in ähnlicher Weise vielfach gefunden werden kann. Bei den Recherchen und beim Schreiben dieser Arbeit sind mir immer wieder Sportveranstaltungen aufgefallen, bei denen von den FernsehmoderatorenInnen ständig Gesichtsausdrücke von SportlerInnen, TrainerInnen oder Fans interpretiert werden. So geschehen auch bei der Übertragung der 17. Etappe der Radrundfahrt Tour de France 2009 von Bourge-Saint-Maurice nach Le Grand-Bornand. Zum Zeitpunkt der 17. Etappe befand sich die Tour de France in einer entscheidenden Phase, die 17. Etappe war die Königsetappe der dreiwöchigen Rundfahrt, da fünf Bergpässe der Savoyer Alpen zu überqueren waren. Am letzten Anstieg des Tages bildete sich eine vierköpfige Spitzengruppe, in welcher sich unter

anderem der Führende in der Gesamtwertung Alberto Contador und sein ärgster Verfolger und Konkurrent Andy Schleck, der allerdings in der Gesamtwertung schon erheblichen Rückstand auf Contador hatte, befanden. Im letzten Drittel stellte der Moderator der ARD fest, dass man Schleck die Anstrengung deutlich in seinem Gesicht anerkenne, wohingegen Contador keine Miene zeigte. Da Contador überlegen in der Gesamtwertung führte, war es nun an Schleck gelegen, sich von Contador losreißen zu können. Doch er schaffte es nicht, er war schließlich der unterlegene Fahrer. So war die Interpretation des Moderators für mich zu diesem Zeitpunkt absolut plausibel: Schleck zeigte Schwächen in Form von Anstrengung, Contador nicht. Die Interpretation und vor allem die Plausibilität und „Richtigkeit“ der gezeigten Gesichtsausdrücke war also nicht nur von dem wirklich gezeigten äußerlichen Ausdruck beeinflusst, sondern ein erheblicher Anteil war auch der Situation zuzuschreiben, die so individuell war, wie jede andere Situation auch. Erst durch das Wissen der Situation, durch das Wissen des Kontextes und der Geschichte der Situation¹⁰ wurde eine plausible Deutung des Gesichtsausdrucks geschaffen. Hier kann man sich die Frage stellen, wie ein System zur automatischen Gesichtsausdruckserkennung die Gesichtsausdrücke von Contador und Schleck zum gegenwärtigen Stand der Technik gedeutet hätte? Contador 100% neutral, Schleck 80% traurig? Wie kann ein System all den Kontext und die spezifische Situation deuten, wenn es nur auf das Gesicht eines Menschen blickt und alles andere außen vor lässt? Wie kann es das richtige Wort für den Gesichtsausdruck finden—wie in diesem Fall „angestrengt“, wenn es nur die Bezeichnungen von höchstens sechs Basisemotionen kennt? Auch wenn das System erkennen könnte, dass es sich um Radsportler handle, wie könnte es wissen, dass Contador der Führende und Schleck der Verfolger ist?

Was an diesem Beispiel deutlich wird ist die extreme Komplexität einer Situation, nicht nur hinsichtlich der Objekte und Gesichtsausdrücke, sondern insbesondere hinsichtlich des Bedeutungskontextes. Zum jetzigen Standpunkt der Entwicklung ist die automatische Gesichtsausdruckserkennung hochgradig und notwendigerweise von Komplexitätsreduktion beeinflusst¹¹. Das hat zur Folge, dass die Technologien nur unter klar definierten und abgegrenzten Bedingungen funktionieren, komplexe *real world* Phänomene aber unzureichend verarbeitet werden können.

Im nächsten Kapitel möchte ich eine Geschichte erzählen, wie die Vielfältigkeit und Komplexität der menschlichen Gesichtsausdrücke, Emotionen und anderer Signale, die in

10 Cicourel spricht im Zusammenhang der Geschichte einer Situation von retrospektiv-prospektivem Ereignissinn. Etwas kann also nicht isoliert betrachtet werden, es muss in eine Geschichte eingeordnet werden, um verstanden werden zu können (Richter 2002: 86).

11 Reduktion ist in den Augen von Introna & Wood (2004: 185) bei Systemen zur Gesichtserkennung (facial recognition) ebenso ein zentraler Punkt, der vor allem deshalb in Kauf genommen wird, um effizientere und schnellere Systeme zu produzieren.

unterschiedlichen je spezifischen Situationen auftreten und verschiedensten Bedeutungsverschiebungen unterliegen, standardisiert und reduziert werden, wenn Maschinen mit der Fähigkeit ausgestattet werden, Gesichtsausdrücke und Emotionen zu erkennen. Ich möchte den Prozess der Standardisierung und Reduktion von Gesichtsausdrücken und Emotionen verfolgen, welcher aus gewissen Gründen in Kauf genommen wird, wenn daran gearbeitet wird, diese automatisch erkennbar zu machen.

Bevor ich genauer auf diese angesprochenen Standardisierungs- und Reduktionsprozesse von Emotionen eingehe, möchte ich zunächst meine methodische Arbeitsweise, meinen Feldzugang und meine Auswertungsstrategie darstellen und somit die empirische Vorgehensweise nachvollziehbar machen. Daran anschließend stelle ich die Technologien zur automatischen Gesichtsausdruckserkennung in den Kontext ihrer Anwendungen, die sich zwischen Mensch Maschine Interaktionen und Überwachung bewegen und werfe einen Blick auf die Erwartungen an diese Technologien. Anschließend gebe ich noch einen Überblick darüber, seit wann an welchen Plätzen zu diesem Thema geforscht wird, vor allem um zu zeigen, dass es ein Thema ist, das vor allem in den letzten 20 Jahren große Aufmerksamkeit erfahren hat.

Methode, Feldzugang, Auswertung

In meiner Arbeit geht es insbesondere darum, nachvollziehen zu können, welches Wissen in die Anwendungen und Technologien zur automatischen Gesichtsausdruckserkennung einfließt, wie dieses Wissen verarbeitet und transformiert wird, sodass es letztendlich in der technischen Anwendung nach gewissen erstrebenswerten Kriterien funktioniert. Ich möchte die *black box* der automatischen Gesichtsausdruckserkennung öffnen und beleuchten, welche Transformationen und Übersetzungen von unterschiedlichen AkteurInnen in die *black box* eingeflossen sind und weiterhin einfließen. Ich begeben mich aus diesem Grund auf die Spur der AkteurInnen, dorthin wo die Technologie hergestellt wird, in die Labore der ComputerwissenschaftlerInnen. Ein zentraler Akteur ist aber auch die Technologie selbst. Bruno Latour begreift Objekte bzw. nicht-menschliche AkteurInnen als aktiv an der Netzbildung Beteiligte und befreit sie in gewisser Weise von ihrem asozialen Charakter, welcher sie zu bloßen Artefakten macht. Wie schon im Theorienteil bei der *Akteur – Netzwerk Theorie* genauer ausgeführt, müssen auch nicht-menschliche AkteurInnen zu Beteiligten an den Handlungen gemacht werden, wodurch neue Assoziationen geschaffen werden können und erst dadurch das Soziale erklärt werden kann (Latour 2007: 136). Freilich kann es zu Verbindungen und Allianzen zwischen menschlichen und nicht-menschlichen AkteurInnen kommen, wie Latour eindrucksvoll am Beispiel von Louis Pasteur und dem Milchsäureferment gezeigt hat (Latour 2002: 137ff.). Doch mit der Radikalisierung des Symmetrieprinzips ergibt sich auch ein grundlegendes Problem: Die beiden zentralen AkteurInnen können in einer sozialwissenschaftlichen Untersuchung niemals gleichberechtigt sein, da sich der Zugang zum nicht-menschlichen Akteur niemals primär wie beim menschlichen Akteur über verschiedene sozialwissenschaftliche Methoden (Interviews, teilnehmende Beobachtung, Textanalyse) vollziehen lässt, sondern nur sekundär (Kneer 2008: 281). Schon allein aus forschungsökonomischen Gründen muss sich der Zugang also über jene menschlichen Akteure ergeben, welche am direktesten mit der Technologie verbunden sind. Nur über diese ergibt sich ein zumindest sekundärer Zugang und nur über diese können auch die nicht-menschlichen AkteurInnen beschrieben werden. Die EntwicklerInnen einer Technologie können als VermittlerInnen begriffen werden, die die Bindeglieder zur Benutzung und Anwendung darstellen (Akrich 2006: 415). Aus dieser theoretischen Vorannahme ergibt sich als Konsequenz, dass jene menschlichen AkteurInnen befragt werden müssen, welche in engem Kontakt zur Technologie stehen, die Entwicklerinnen und Entwickler der Technologie. Über sie ergibt sich der direkteste Zugang zur *black box* der Technologie, über sie erhält man Einblicke in die *technology in the making*. Sie sind Experten und Expertinnen im Umgang mit der Technologie. Aus diesem Grund liegt

es auf der Hand, die EntwicklerInnen und ExpertInnen im Umgang mit Technologien zur Gesichtsausdruckserkennung zu interviewen, die hauptsächliche Methode meiner Arbeit ist also das ExpertInneninterview. Auch wenn es auf den ersten Blick als logisch und schlüssig erscheint, was damit gemeint ist, muss man aufpassen und genau bedenken, welche Funktion ein ExpertInneninterview haben kann und welche verschiedenen Ausprägungen es gibt. Im Folgenden werde ich theoretisch auf diese „methodologische Mehrdeutigkeit“ des ExpertInneninterviews eingehen.

Das ExpertInneninterview

Bogner und Menz unterscheiden grundsätzlich drei Formen des ExpertInneninterviews: das explorative, das systematisierende und das theoriegenerierende ExpertInneninterview (Bogner & Menz 2005: 37f.).

Das explorative ExpertInneninterview dient primär zu einer Herstellung einer ersten Orientierung in einem Feld, welches thematisch neu oder unübersichtlich ist. Sekundär soll es dabei helfen, das Problembewusstsein zu schärfen und das Untersuchungsgebiet thematisch zu strukturieren. Der inhaltliche Schwerpunkt liegt also bei der thematischen Sondierung. Dabei ist es durch Offenheit geprägt.

Das systematisierende ExpertInneninterview zielt auf die Teilhabe an exklusivem Wissen ab. Ziel ist die systematische und lückenlose Gewinnung von Informationen, welche nur über den Experten oder die Expertin zugänglich sind. Die dabei gewonnenen Informationen können als objektive Tatbestände oder als Fachwissen charakterisiert werden. Im Gegensatz zum explorativen, zielt das systematische ExpertInneninterview auf eine thematische Vergleichbarkeit der Daten ab.

Das theoriegenerierende Interview geht über die Funktionen der beiden bereits genannten Ausprägungen hinaus. Es ist nicht nur ein Katalysator des Forschungsprozesses bzw. zur Gewinnung von Faktenwissen und Informationen angedacht, sondern zielt auf die kommunikative Erschließung und die analytische Rekonstruktion der „subjektiven Dimension“ des ExpertInnenwissens ab. Implizite Wissensbestände, Weltbilder und Routinen sollen theoretisch gehaltvoll konzeptualisiert werden.

Neben den verschiedenen Formen des ExpertInneninterviews gibt es auch unterschiedliche Zugänge zum Begriff des Experten oder der Expertin: Bogner und Menz unterscheiden auch hier wieder zwischen drei verschiedenen Zugängen: ein voluntaristischer, ein konstruktivistischer und ein wissenssoziologischer ExpertInnenbegriff (ebd.: 40ff.).

Der voluntaristische Begriff von ExpertInnenwissen geht davon aus, dass prinzipiell alle Menschen als Experten oder Expertinnen angesehen werden können, da sie über je

spezifische Lebenserfahrungen und –geschichten verfügen. Somit wird man zum Experten oder zur Expertin seines eigenen Lebens.

Der konstruktivistische Zugang geht davon aus, dass der Experte oder die Expertin ein Konstrukt eines je spezifischen ForscherInneninteresses ist und es zu einer Zuschreibung der ExpertInnenrolle durch diese kommt. ExpertInnen sind also jene, die zu ExpertInnen gemacht werden.

Schließlich begreift der wissenssoziologische Zugang das ExpertInnenwissen als Sonderwissen und grenzt es damit vom Allgemeinwissen ab. Der Experte oder die Expertin sind nicht durch real existierende Kompetenzdifferenzen, sondern durch die soziale Relevanz ihres Sonderwissens charakterisiert. Besonders interessant ist dabei neben dem (technischen) Fachwissen und dem Prozesswissen über Handlungsabläufe und Interaktionsroutinen vor allem das Deutungswissen. Das Deutungswissen bezieht sich auf subjektive Relevanzen, Regeln, Sichtweisen und Interpretationen der ExpertInnen und macht dadurch das ExpertInnenwissen zu einem heterogenen Wissen, welches die subjektive Dimension anspricht und dabei auch den Experten oder die Expertin als Privatperson mit integriert. Von besonderem Interesse sind dabei „Alltagsweisheiten“ oder Metaphern aus dem privaten Bereich. Im wissenssoziologischen Zugang werden Experten und Expertinnen gerade deswegen befragt, weil ihr Wissen eine gewisse soziale Relevanz besitzt. Diese soziale Relevanz äußert sich darin, dass die Handlungsorientierungen, das Wissen und die Einschätzungen der ExpertInnen die Handlungsbedingungen anderer AkteurInnen in entscheidender Weise (mit-)strukturieren. Hierbei wird klar, dass nicht die Exklusivität des ExpertInnenwissens von besonderem Interesse ist, sondern die Wirkmächtigkeit. Dabei kann das Wissen der ExpertInnen in einem bestimmten praktischen Kontext hegemonial werden (ebd.: 45f.).

Wendet man den theoretischen Zugang auf meine Arbeit an, so wird es schwierig, sich eindeutig für eine Funktion des ExpertInneninterviews zu entscheiden. Meiner Meinung nach muss es in diesem Zusammenhang alle drei Funktionen erfüllen können, weshalb ein starrer kategorisierter Begriff des ExpertInneninterviews wohl kritisch zu sehen ist, da jedes Interview anders verlaufen kann und es sich um einen dynamisch-situativen Kommunikationsprozess handelt. In meiner Arbeit geht es zunächst einmal darum, Orientierung im neuen Feld der Technologien der Gesichtsausdruckserkennung zu gewinnen und so die spannenden Fragen aufmachen zu können. Neben dieser explorativen Funktion des ExpertInneninterviews sollte aber auch die systematisierende Funktion nicht vergessen werden. Denn um die Technologien genau beschreiben zu können, um Einblicke in die *black box* der Technologie zu bekommen, muss ich zumindest einen grundlegenden, besser, einen möglichst detaillierten Zugang zum Fachwissen haben. Hier stellt sich also insbesondere die Frage, wie die Technologien technisch funktionieren, was—lapidar gesagt—alles in ihnen

drinnen steckt, dass sie so funktionieren wie sie funktionieren. Nur ein solches Wissen verhindert spekulative Interpretationen über verschiedene konstitutive Funktionszusammenhänge. Dabei muss davon ausgegangen werden, dass es unterschiedliche Modelle gibt, die den Technologien zur Gesichtsausdruckserkennung zugrunde liegen. Schließlich ist auch das theoriegenerierende ExpertInneninterview von Bedeutung für mein Forschungsvorhaben, da es auch auf implizite Wissensbestände, Weltbilder und Routinen abzielt, deren Erforschung ich in meinen Fragen fixiert habe.

Deutlicher kann ich meinen Zugang zur Art des ExpertInnenwissens zuordnen. Zwar definiere ich in gewisser Weise auch konstruktivistisch, wer für mich als Experte oder Expertin gilt, jedoch lässt sich gehaltvoller mit dem wissenssoziologischen Zugang und dem Begriff des ExpertInnenwissens als Deutungswissen arbeiten. Das Deutungswissen ist für mich in Beziehung zur Ethnomethodologie zu sehen, da beide Ansätze die alltägliche und subjektive Dimension nicht nur mit integrieren, sondern auch besonders hervorheben. Das Deutungswissen offenbart eine Möglichkeit, die gegenseitig (stumm und selbstverständlich) geteilten Deutungen über die Welt und deren Sinn zu erschließen. Also jenen Punkt, wo sich alltägliche, kulturell geschliffene Annahmen mit den technowissenschaftlichen Konzepten verbinden. Laudel und Gläser gehen allerdings davon aus, dass qualitative Interviews nicht die richtige Methode sind, um implizites Wissen und selbstverständliche Annahmen erforschen zu können. Sie verweisen dabei auf die teilnehmende Beobachtung als passende Methode. Allerdings führen sie auch an, dass in besonderer Weise das wissenschaftlich informierte Interview (*scientifically informed interviewing*) gefährdet ist, implizites Wissen und selbstverständliche Annahmen zu verpassen, weil die interviewten WissenschaftlerInnen davon ausgehen, dass die wissenschaftlich informierten InterviewerInnen sowieso wissen, was gemeint ist und darauf verzichten, etwas weiter oder banaler auszuführen (Laudel & Gläser 2007: 109). Obwohl ich den Einwand von Laudel und Gläser als durchaus berechtigt halte, würde ich davon abraten, ihn zu generalisieren, vor allem wenn man bedenkt, dass das Interview nicht mehr als „archäologisches“ Modell angesehen werden kann, sondern vielmehr ein Interaktionsmodell darstellt. Das ExpertInneninterview ist geprägt von unterschiedlichen Interaktionssituationen, welche sich erst in der Praxis ergeben, welche man aber auch strategisch einsetzen kann (Bogner & Menz 2005: 48). Der Interviewer oder die Interviewerin kann verschiedene Rollen einnehmen, welche das Interview verschieden mitbestimmen können. Die InterviewerIn kann als Co-ExpertIn oder als ExpertIn einer anderen Wissenskultur fungieren. Die InterviewerIn kann aber auch als Laie, als Autorität, als potenzielle KritikerIn oder als KomplizIn auftreten (ebd.: 50ff.). Als WissenschaftsforscherIn oder als SoziologIn befindet man sich in den meisten Fällen nicht in der Lage, den Status einer Co-ExpertIn erheben zu können. Vielmehr muss man davon ausgehen, dass man als ExpertIn einer anderen Wissenskultur wahrgenommen wird. Meiner

Meinung nach hat das auch zur Folge, dass man im untersuchten Bereich als Laie wahrgenommen wird. Durch entsprechende Vorbereitung kann dies natürlich verhindert werden. Dennoch bietet für mich die Kombination aus Experte oder Expertin einer anderen Wissenskultur, allein schon um eine gewisse Kompetenz zu demonstrieren, welche die Hierarchie nicht zu stark ausprägt und aus InterviewerIn als Laie die sinnvollste strategische Kombination. Denn gerade die „naiven“ Fragen, die ein Laie stellen kann und darf, scheinen unter dem Hintergrund meines Erkenntnisinteresses die interessantesten und erfolgreichsten Antworten zu liefern. Ziel muss es sein, sich die Sicht- und Handlungsweisen von Grund auf erklären zu lassen. Es gibt natürlich auch Nachteile dieser Herangehensweise, welche vermieden werden sollten: endlose Monologe und Lehrbuchwissen sind dabei zwar teilweise gewünscht, dürfen aber nicht den Großteil des Interviews ausmachen. Außerdem läuft man Gefahr, dass fachlich anspruchsvolle Fragen nicht geklärt werden können. Nachfragen werden auch eher ignoriert und überhaupt kann es passieren, dass der oder die Befragte die Kontrolle über das Interview übernimmt (ebd.: 56).

Interviewleitfaden

Allein schon aus der Perspektive von demonstrativer Kompetenz habe ich es als notwendig erachtet, einen Interviewleitfaden zu konstruieren und bei den Interviews zu verwenden. Ich wollte nicht, dass der Leitfaden im Zentrum der Interviews steht, er sollte den Gesprächsverlauf nicht diktieren (Hopf 2005: 358). Der Leitfaden sollte vielmehr einerseits narrative Strukturen zulassen, andererseits aber auch genügend Fragen zur Verfügung stellen, um die Forschungsfragen befriedigend beantworten zu können. Der Leitfaden wurde immer an die je spezifische Interviewsituation und an die unterschiedlichen Befragten angepasst und adaptiert. Der von mir verwendete Interviewleitfaden befindet sich im Anhang dieser Arbeit.

Feldzugang

Zu Beginn meiner Arbeit ging es vor allem darum, geeignete ExpertInnen zu finden, die bereit waren, mir über Interviews Einblicke in die Technologien zur Gesichtsausdruckserkennung zu geben. Es war schnell klar, dass ich auch aus einer gewissen Ortsgebundenheit heraus meine geographische Umgebung nach ExpertInnen absuchen musste. So konzentrierte ich mich auf der Suche nach InterviewpartnerInnen auf den deutschsprachigen Raum, also Österreich, Deutschland und die Schweiz. Es ging auf jeden Fall aber auch nicht darum, eine vollständige Erhebung aller an Technologien zur Gesichtsausdruckserkennung beteiligten ExpertInnen vorzunehmen. Vielmehr wollte ich ja

Einblicke in die Forschungstätigkeit und in die *black box* der Technologie erhalten, um die spannenden Fragen aufmachen zu können. Es ging also weniger um die Quantität der Interviews, sondern vielmehr um die Qualität, ich wollte möglichst gehaltvolles Interviewmaterial und brauchte deshalb ExpertInnen, die mir dieses auch liefern konnten. So habe ich mich auf die Spuren der AkteurInnen begeben, die sie auf ihrem Weg in Form von z.B. Zeitungsartikeln oder Internetseiten hinterlassen haben. Zunächst suchte ich nach einem Computerwissenschaftler vor Ort, der mir Einblicke in die Praxis der Gesichtserkennung geben sollte, sowohl von *face recognition* als auch von *face detection*, da beides nicht voneinander getrennt werden kann. Wie die folgenden InterviewpartnerInnen auch, kontaktierte ich den Computerwissenschaftler von der *Technischen Universität Wien* per e-Mail und organisierte mir einen Interviewtermin. Nachfolgend wollte ich auch genauere Einblicke in die wissenschaftliche Auseinandersetzung mit Gesichtsausdrücken erhalten und konnte ein Interview mit einer Verhaltensforscherin von der *Universität Wien* führen. Da es in Österreich keine Forschungstätigkeit im Bereich von automatischer Gesichtsausdruckserkennung gibt, musste ich auf Deutschland ausweichen und konnte mir zwei weitere Interviews an der *Technischen Universität München* (dort mit zwei Computerwissenschaftlern) und am *Fraunhofer Institut für Integrierte Schaltungen* in Erlangen organisieren. An beiden Institutionen werden Systeme zur automatischen Erkennung von Gesichtsausdrücken entwickelt, sie waren also höchst relevant für meine Arbeit. Eine Übersicht über alle geführten Interviews befindet sich im Anhang dieser Arbeit.

Auswertung

Bei der Auswertung von ExpertInneninterviews war es naheliegend, sich auf eine Auswertungsmethode zu stützen, die sich speziell auf ExpertInnenwissen konzentriert. Eine systematische Analyse bietet die Auswertungsstrategie von ExpertInnenwissen von Meuser und Nagel an (Meuser & Nagel 2005: 80ff.). Dieser Ansatz erkennt jedes einzelne Interview als spezifische und besondere Interaktion und Kommunikation an und versteht sich insgesamt als Entdeckungsstrategie, was meiner Vorgehensweise sehr entgegen kommt. Das Ziel der Auswertung sollte die Gewinnung von empirischem Wissen sein und somit wird auch die „dichte Beschreibung“ des Materials angestrebt. Bruno Latour sieht im Aufzeichnen und Beschreiben die beiden „ehernen Gesetze“ der Soziologie, welche der Soziologe oder die Soziologin dem Herausfiltern und der Disziplinierung vorziehen sollte (Latour 2008: 97). Die Auswertung startet mit den Phasen von Transkription und Paraphrase. Die Paraphrase erfolgt hauptsächlich hinsichtlich der leitenden Forschungsfragen und sollte das Problem der Reduktion von Komplexität betreffend weder Textstellen unterschlagen, noch hinzufügen oder verzerren. In der Folge sollte das Material in Form von Überschriften zusammengefasst

und konzentriert, somit weiter verdichtet werden. Anschließend kommt es zum thematischen Vergleich aller Interviews. Erst dann kann man den reinen Text verlassen und zu einer soziologischen Konzeptualisierung übergehen, was ich als die Gegenüberstellung des empirischen Materials mit Theorie und Literatur verstehe.

Qualitative Forschung ist oft mit der Frage konfrontiert, welche Kriterien für die Wissenschaftlichkeit, Güte und Geltung der Ergebnisse herangezogen werden können (Steinke 2005: 319). Ich will mich hier in der postmodernen Ablehnung von klaren Kriterien verorten und folge damit vor allem Denzin, welcher postmoderne Ethnographie dadurch definiert,

„dass Forscher ihre Texte in der ersten Person Singular schreiben, wodurch die Kluft zwischen der beobachtenden Person und der beobachteten Realität überwunden wird und Fragen nach Reliabilität oder Validität sich nicht mehr stellen“ (Denzin 1990: 231 zit. nach Steinke: 2005: 321).

Subjektivität kann also sehr wohl ein Teil der Wissenschaft sein, der eigene Standpunkt wird so zum Erkenntnishorizont. Diese nach Denzin neue Sicht der Ethnographie versucht, die Realität interpretierend darzustellen (Richter 2002: 82f.)

Kontextualisierung:

Anwendungen zwischen Mensch-Maschine Interaktion und Surveillance

Zum ersten Mal wurde ich auf Technologien zur Gesichtsausdruckserkennung durch einen Artikel der Süddeutschen Zeitung (SZ Nr.195, 25./26. August 2007, S.45) über die Forschungsaktivitäten des Fraunhofer Instituts für Integrierte Schaltungen in Erlangen aufmerksam. Dort wurde ein System entwickelt, das neben Gesichtsausdrücken auch das Geschlecht einer Person automatisch erkennen kann.

Abbildung 8: Echtzeit Gesichtsausdruckserkennung



Quelle: Süddeutsche Zeitung Nr.195, 25./26. August 2007, S.45, Foto: Kurt Fuchs

Auf dem zum Artikel gehörigen Bild erkennt man zwei Personen, die vor einem großen Bildschirm stehen und mit unterschiedlichen Gesichtsausdrücken in diesen blicken. Auf dem Bildschirm „spiegeln“ sich die Gesichter wieder, haben aber zusätzlich jeweils um den Bereich von Augen, Nase und Mund einen rechteckigen weißen Kasten platziert bekommen sowie weitere zwei solche Kästen um die Nase und um den Mund. Auf den Pupillen der Augen befinden sich jeweils weiße Kreuze und unterhalb des Gesichts finden wir eine Skala

mit den englischen Worten *für ärgerlich, fröhlich, traurig, überrascht*, sowie den Ausdrücken *Beide Augen* und *Gender*. Hinter den ersten vier Worten befinden sich jeweils Balken, die mit roter Farbe mehr oder weniger angefüllt sind. Bei der rechten Person ist der Balken von ärgerlich leer, von glücklich und traurig jeweils zu einem Viertel angefüllt und der Balken von überrascht ganz mit roter Farbe angefüllt. Beide Augen sind: Offen und bei Gender handelt es sich um männlich. Die Balken der linken Person sind bis auf fröhlich ganz leer, fröhlich ist voll mit roter Farbe angefüllt. Beide Augen sind ebenso offen und neben Gender steht weiblich. Schön und gut, jetzt wissen wir mehr über den Gefühlszustand des Mannes, der überrascht und gleichzeitig ein bisschen fröhlich und traurig ist und auch, dass die Frau neben ihm einfach nur fröhlich und sonst nichts ist, aber wieso das alles? Handelt es sich etwa um einen neuartigen digitalen Spiegel, der uns sofort sagen kann, wie wir uns gerade fühlen und welche Geschlechtsidentität wir gerade haben, also eine Art persönlicher Assistent für das alltägliche Leben, mit dem man seine Emotionen besser kontrollieren kann? Man kann viele Fantasien haben, was man alles mit einer automatischen Gesichtsausdruckserkennung zukünftig anstellen könnte, die Interessen daran sind unterschiedlich. Die größten Erwartungen legt man aber in eine Verbesserung der Mensch – Maschine bzw. der Mensch – Computer Interaktion.

Mensch – Maschine Interaktion

Maschinen, Computer und Roboter sollen in die Lage versetzt werden, sich emotional verhalten zu können, dies beinhaltet auch das Erkennen von Gesichtsausdrücken. Ein prominentes Vorbild könnte HAL 9000 sein, der berühmte Computer aus Stanley Kubricks Film *2001: Odyssee im Weltraum* (Originaltitel: *2001: A Space Odyssey*). HAL kann im Film nicht nur Emotionen erkennen, er kann auch Emotionen wie Angst ausdrücken, etwa in der Szene, als das Besatzungsmitglied Dave beginnt, ihn zu demontieren. Erstaunlicherweise erweckt sich im Film beim Zuseher oder der Zuseherin der Eindruck, dass die Maschine HAL stärkere Emotionen ausdrückt als die Menschen (Picard 1997: 279). HAL ist deshalb eine zutiefst sympathische Maschine, man kann mit ihm relativ „natürlich“ interagieren, was auch das Ziel von ComputerwissenschaftlerInnen ist (Cheon 2009; Deng et al. 2005; Chibelushi & Bourel 2002; Cohen et.al. 2002). Computer und ihre Nutzer sollen enger zusammenrücken (Dornaika 2008), die Kommunikation zwischen Mensch und Maschine soll verbessert werden (Ioannou et al. 2007), indem der Computer oder die Maschine dem Menschen ähnlicher werden sollen. Computer sollen deshalb die menschlichen Kommunikationsfähigkeiten annehmen (Wimmer et al. 2008; Cerezo et al. 2007), damit der Mensch in Zukunft weniger Probleme im Umgang mit ihnen hat. In Form eines auf den Menschen ausgerichteten Designs soll stärker auf eben diesen eingegangen werden (Zeng et al. 2007). Dass ein

Computer oder eine Maschine angemessen auf einen Menschen reagieren kann, sollen sie in die Lage versetzt werden, den Zustand oder die Befindlichkeit eines Menschen erkennen zu können (Bashyal & Venayagamoorthy 2008; Wong et al. 2006). Die Annahme besteht, dass Menschen von Maschinen erwarten, dass sich diese ähnlich wie Menschen verhalten. Ist dieses Verständnis gestört, so kann auch die Interaktion zwischen Mensch und Maschine misslingen (Wimmer et al. 2008). Mit dem Erkennen von Gesichtsausdrücken erwarten sich die ComputerwissenschaftlerInnen einen direkteren, „natürlicheren“ Zugang, als über andere Umwege. Dies kommt in folgender Interviewpassage eines von mir befragten Computerwissenschaftlers ausführlich und deutlich zum Ausdruck:

„Es geht darum, dass die Maschinen, die wir konstruieren können immer komplizierter und immer besser werden, die Leistung wird besser, die Rechner werden schneller, man kann immer mehr damit machen, aber was sich eigentlich seit Jahrzehnten nicht verändert hat ist die Art und Weise, mit denen man die Maschinen handhabt. Klar, die grafischen Oberflächen werden etwas bunter, werden etwas schöner, aber letztendlich hat man immer nur Tastatur, Maus, Bildschirm, um mit dem Computer in Interaktion zu treten. Und wenn man an einfachere Maschinen denkt wie z.B. Videorekorder, dann hat man noch nicht einmal Maus und Bildschirm, dann hat man eigentlich nur ein paar Tasten. Auf der anderen Seite wissen wir aber, dass Menschen untereinander sehr schnell und effizient miteinander kommunizieren können, wie wir ja hier gerade auch. Also die Informationen, die wir hier sehr schnell weitergeben, die würden halt mit diesen einfachen Methoden wie z.B. e-Mail wesentlich mehr Zeit in Anspruch nehmen. Es ist viel komplizierter dieselbe Informationsdichte zu erreichen. D.h. es gibt schon seit mehreren Jahren die Idee, dass man versucht, die menschliche Interaktion nachzuempfinden und eben diese Effizienz zu übertragen, um Maschinen einfacher bedienbar zu machen. Menschen nützen eben verschiedene Kommunikationskanäle, natürlich Sprache, aber eben auch den visuellen Kanal wie Körpersprache, Gesten, Kopfgesten oder eben Mimik. Das vom Gegenüber zu ermitteln und das in der Kommunikation zu verwenden“ (Interview 3).

Die Kommunikation soll nicht länger nur über den Umweg von Bildschirm, Tastatur und Maus erfolgen, sie soll direkt quasi von Angesicht zu Angesicht möglich sein. Bei diesem Prozess spielt unter anderem das Erkennen des Gesichtsausdrucks eine wichtige Rolle. Das Kommunikationsspektrum von Tastatur und Maus soll nach Auffassung der Wissenschaftler erweitert werden:

„Wir denken eben, dass man über die Möglichkeit eine Kamera zu haben, eben auch Möglichkeiten schaffen können, die den Mehrwert schaffen, dass der Computer mehr als halt nur die Tastatur und die Mauseingaben vom Menschen erfährt“ (Interview 4).

Computer oder Maschinen sollen also mit Hilfe einer Kamera und einer damit verbundenen Software Gesichtsausdrücke direkt erkennen und interpretieren können, ohne dass ein

Nutzer oder eine Nutzerin es der Maschine über Mausklick oder Tastatureingabe „erzählen“ muss. Dies soll einen Mehrwert für die Mensch – Maschine Interaktion darstellen.

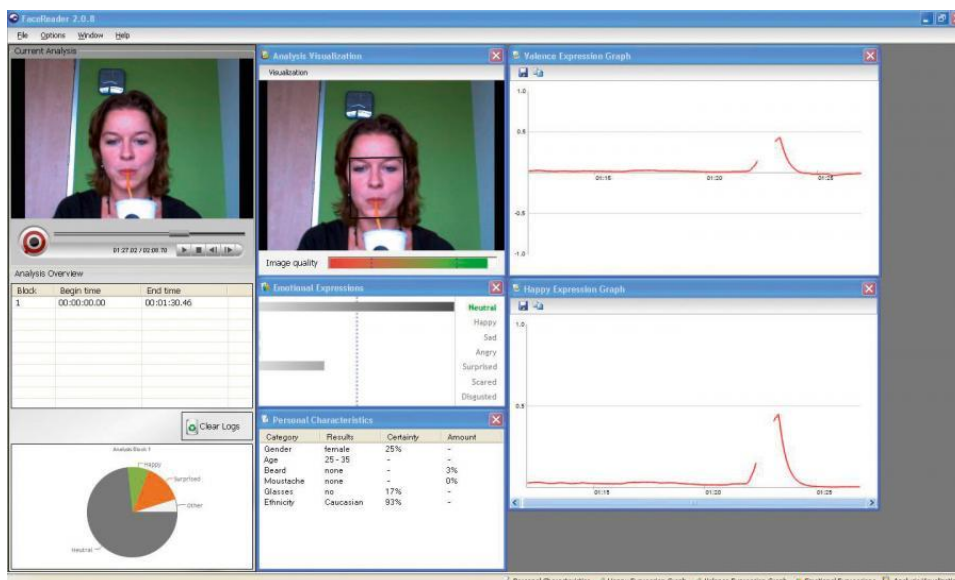
Surveillance

Neben der Verbesserung der Mensch - Maschine Interaktion gibt es aber auch eine große Bandbreite an anderen möglichen Anwendungen: Anfang 2008 veröffentlichte die britische Zeitung *The Times* (January 16, 2008) einen Artikel über einen Patentantrag von Microsoft für eine Software zur Arbeitsplatzüberwachung. Das „einzigartige Überwachungssystem“ könne mit Hilfe von drahtlosen Sensoren den Herzschlag, die Hauttemperatur, Hirnsignale, Körpertemperatur, Blutdruck als auch Gesichtsausdrücke von ArbeitnehmerInnen erkennen. Würden unregelmäßige Muster auftreten, so könnten ArbeitgeberInnen der Person Unterstützung anbieten. Ähnliche Ideen gibt es bei Lernsoftware:

„Das Potenzial in so einer Lernsoftware, wenn man erkennen kann, ob der Benutzer gelangweilt ist oder gestresst ist beim Lernen, dann könnte man natürlich super mit einer Software reagieren“ (Interview 4)

Weitere Anwendungen, die in den Bereich Beobachtung gehen, existieren bereits: Im niederländischen Wageningen werden im *Restaurant der Zukunft*, einem Labor zur Erforschung von Essgewohnheiten in Restaurants, die Gesichtsausdrücke der Restaurantbesucher automatisch mit Hilfe des *FaceReader*™ erkannt und mit in die Analysen einbezogen.

Abbildung 9: Bildschirmbild der Software FaceReader™



Quelle: <http://www.noldus.com/content/facereader> [Online verfügbar am 15.08.2009]

Überhaupt scheint KonsumentInnenverhalten eine große Rolle zu spielen, denn auch beim Fraunhofer Institut in Erlangen will man mit der dort entwickelten Software SHORE erkennen können, wie Personen auf Werbeplakate im öffentlichen Raum reagieren.

„...also was sich für uns als Hauptmarkt anbietet oder der Bereich, wo wir das größte Interesse sehen, ist der Bereich Werbewirkungsforschung, d.h. wie wirkt Werbung auf Menschen. Das ist ein Thema, das die Leute, die Werbung machen, brennend interessiert. Welche Anzahl von Leuten schauen eine Werbung an, das hat man früher oder löst man bis jetzt noch im Fernsehen mit Geräten, die dann bei ausgewählten Fernsehnutzern stehen. Nachdem ja Werbung mehr und mehr im öffentlichen oder halb öffentlichen Raum auch platziert wird, besteht da natürlich auch das Interesse. Und bei Plakatwerbung wusste man es also lange Zeit nicht. Und jetzt kommt also mehr und mehr digitale Werbung, die schon gezeigt wird und da kann man natürlich jetzt mit einer Kamera analysieren zunächst mal wie viele Leute bleiben stehen und wie lange bleiben sie stehen, wie ist die Verweildauer pro Person und dann kann man sagen, sind das Männer oder Frauen die stehenbleiben, das geschlechtsspezifisch analysieren. Und man kann die Emotionen natürlich auch als interessantes Feedback verwenden. Das ist eigentlich momentan der Hauptanwendungsbereich, den wir momentan sehen.“ (Interview 4)

Eine Kamera beobachtet, wie eine Person auf einen Auslöser wie in diesem Fall auf ein Werbeplakat reagiert. Die an die Kamera gekoppelte Software kann in der Folge die Reaktion interpretieren, sie einer bestimmten Emotion zuordnen. Man braucht also nicht mehr danach fragen, wie einer Person ein Werbeplakat gefällt, man kann sie einfach mit Hilfe einer Kamera beobachten und anschließend bzw. in Echtzeit mit einer Software feststellen, ob die Person fröhlich, traurig, ärgerlich oder überrascht auf das Plakat reagiert hat. Das Gesicht bietet die benötigte Information—wie reagieren Personen auf Werbung?—offensichtlich auf seiner Oberfläche an, man braucht sie nur noch automatisch von dort ablesen.

Ein Geheimnis gelüftet hat auch der niederländische Computerwissenschaftler Nicu Sebe von der Universität Amsterdam, welcher bereits 2005 spektakulär auf die Gesichtsausdruckserkennung aufmerksam machen wollte, indem er den Gesichtsausdruck der Mona Lisa dechiffrierte: 83% fröhlich, 9% empört, 6% ängstlich und 2% ärgerlich. Jetzt wissen wir es. Endlich. Die Wahrheit. Gesichtsausdruckserkennung nimmt langsam Einzug in unser Leben und so könnte es auch bald möglich sein, TerroristInnen oder VerbrecherInnen an ihren verdächtigen Gesichtsausdrücken vor dem Einsteigen in Flugzeuge identifizieren zu können, wie es sich Paul Ekman vorstellen könnte (vgl. Kapitel *Rezeption und Kritik an Paul Ekman*).

Die Überwachung und das Beobachten von Personen, was man im Englischen unter dem Begriff *surveillance* zusammenfassen könnte, ist neben der Verbesserung der Mensch-

Maschine Interaktion sicherlich ein nicht zu vernachlässigender möglicher Anwendungsbereich der Technologien zur automatischen Gesichtsausdruckserkennung. Computerwissenschaftler weltweit beziehen sich direkt in ihren Forschungsarbeiten auf den Bereich *surveillance*, wenn sie darauf eingehen, in welchen Bereichen die automatische Gesichtsausdruckserkennung eingesetzt werden könnte (Gomathi et al. 2009; Khan & Bhuiyan 2009; Chen et al. 2008; Bashyal et al. 2007; Wong et al. 2006; Kuilenburg et al. 2005; Essa 1999; Zhang 1998).

Sonstige Anwendungsfelder

Es wäre zu einfach zu sagen, dass zukünftige Anwendungen der automatischen Gesichtsausdruckserkennung nur im Bereich der Mensch-Maschine Interaktion oder zu Beobachtungs- und Überwachungszwecken eingesetzt werden könnten.

So kann man etwa Anwendungsbeispiele in der Verhaltensforschung und der psychologischen Forschung finden, wie z.B. die Erforschung der Unterscheidung von echtem und gespieltem Schmerz mit Hilfe von automatischer Gesichtsausdruckserkennung (vgl. Littlewort et al. 2009).

Desweiteren könnte man die Technologie auch bei digitalen Bilddatenbanken einsetzen, um etwa Bilder mit bestimmten Gesichtsausdrücken automatisch finden zu können (vgl. Essa 1999). Einen anderen Anwendungsbereich findet man in der Arbeit mit autistischen Kindern:

„Our research is linked to a project that aims at studying social competency and emotion awareness in autistic children. The project [...] is developing a gamelike tool that engages children in a variety of activities in order to assess their social skills“ (Arya et al. 2009:1)

In eine ähnliche Richtung gehen Computerprogramme, die sich im Bereich *Edutainment* verorten lassen. Hier wäre beispielhaft die Software *Emotional Flowers* zu erwähnen: Mit ihrem Gesichtsausdruck können Kinder eine virtuelle Blume durch ein Lächeln wachsen oder andererseits verwelken lassen¹².

Eine starke Verbindung der Gesichtsausdruckserkennung ist auch zu Anwendungen in der Virtuellen Realität, zu Animationsfilmen und vor allem zu Avataren zu erkennen, also Bereiche, wo ein Gesichtsausdruck virtuell nachgebildet wird. So sieht die von mir interviewte Verhaltensforscherin die nähere Zukunft auch eher in diesem Bereich:

„Weniger das Erkennen. Eben diese virtuellen Agenten werden unser Leben zunehmend prägen. Ein „Maxel“, der uns beim Automaten bedient, bei der Auskunft, bei Web Applikationen. Und da glaub ich wird sich einiges tun.“ (Interview 2)

12 <http://www.icts.sbg.ac.at/media/pdf/pdf1362.pdf> [Online verfügbar am 15.08.2009]

In dieser Richtung geht es also weniger darum, Gesichtsausdrücke automatisch erkennen zu können. Vielmehr sollen virtuelle Agenten Gesichtsausdrücke zeigen können und so diese menschliche Eigenart nachstellen können.

Forschungsaktivitäten

Seit über 30 Jahren gibt es das Interesse unter Wissenschaftlern und Wissenschaftlerinnen aus verschiedenen Disziplinen, menschliche Emotionen automatisch erkennen zu können. Dabei hat man sich noch vor der automatischen Gesichtsausdruckserkennung um die automatische Emotionserkennung mit Hilfe der Stimme konzentriert (vgl. Williams & Stevens 1972). Den ersten Ansatz, um Gesichtsausdrücke automatisch erkennen zu können lieferten Suwa und Kollegen (vgl. Suwa et al. 1978) Ende der 1970er Jahre (Zeng et al. 2007).

Ein neuer Schub in der automatischen Gesichtsausdruckserkennung setzte zu Beginn der 1990er Jahre mit der Pionierarbeit von Mase und Pentland (vgl. Mase & Pentland 1991) ein, was unter anderem an den Entwicklungen in verwandten Bereichen wie Gesichtsdetektion, Gesichtstracking und Gesichtserkennung sowie an dem relativ günstigen Zugang zu schnellen Computern gelegen haben könnte (Fasel & Luetin 2002). Zeng et al. (2007) datieren das zunehmende Interesse zwar erst auf die späten 1990er Jahre, erwähnen aber auch die frühen Arbeiten von Mase & Pentland (1991) sowie von Kobayashi & Hara (1991).

Zu Beginn des 21. Jahrhunderts entstanden bereits einige Überblicksartikel über die Technologien zur automatischen Gesichtsausdruckserkennung (Chibelushi & Bourel 2002; Fasel & Luetin 2002; Pantic & Rothkrantz 2000), auf welche sich die meisten der folgenden Arbeiten beziehen. Bis heute finden wir eine Vielzahl an unterschiedlichen Artikeln und Ansätzen zur automatischen Gesichtsausdruckserkennung, die versuchen, die Aufgabe zu lösen. In einer globalen Welt ist es lange nicht mehr der einzelne Wissenschaftler oder Wissenschaftlerin als Genie, sondern vielmehr eine kaum zu überblickende Anzahl an WissenschaftlerInnen aus unterschiedlichen Kontexten, die über den gesamten Globus verteilt an ähnlichen Problemen arbeiten. In folgender Tabelle (nächste Doppelseite) möchte ich einen Überblick geben, an welchen universitären sowie außeruniversitären Forschungsinstituten und in welchen Ländern Forschungsaktivitäten im Bereich der automatischen Gesichtsausdruckserkennung festzustellen sind und waren. Dieser Überblick hat nicht den Anspruch auf Vollständigkeit, es soll einzig die Vielzahl an Orten, wo man an automatischer Gesichtsausdruckserkennung arbeitet, verdeutlicht werden.

Tabelle: Forschung zu automatischer Gesichtsausdruckserkennung

Institution	Country
School of Computer Science and Software Engineering, University of Wollongong	Australia
Department. of Computer Science & Engineering, Chittagong University of Engineering and Technology	Bangladesh
Deptartment of Computer Science & Technology, Jahangirnagar University	Bangladesh
Department of Psychology, Carleton University Ottawa	Canada
Human Development and Applied Psychology, University of Toronto	Canada
School of Information Technology, Carleton University Ottawa	Canada
School of Interactive Arts and Technology, Simon Fraser University Surrey	Canada
Department of Automation, Tsinghua University	China
Department of Computer Science and Engineering, Fudan University	China
Motorola China Research Center, Shanghai	China
School of Communication and Control Engineering, Jiangnan University	China
School of Electronic and Information Engineering, University of Technology Guangzhou	China
Heudiasyc Mixed Research Unit, CNRS/UTC	France
Institut Géographique National, Laboratoire MATIS	France
ORSYP, Paris	France
Chair for Image Understanding and Knowledge-Based Systems, Technical University of Munich	Germany
C-Lab, University of Paderborn	Germany
Fraunhofer Institute for Integrated Circuits IIS, Erlangen	Germany
Department of Electrical and Computer Engineering, National Technical University of Athens	Greece
Department of Informatics, Aristotle University of Thessaloniki	Greece
Image, Video and Multimedia Systems Laboratory, National Technical University of Athens	Greece
Centre for Signal Processing, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University	Hong Kong
Department of Computer Science and Engineering, The Hong Kong University of Science and Technology	Hong Kong
Lo Kwee-Seong Medical Image Analysis Laboratory, The Hong Kong University of Science and Technology	Hong Kong
College of Engineering Kidangoor	India
Marian Engineering College	India
National Engineering College, Kovilpatti	India
ATR IRC/MIS Labs, Keihanna Science City, Kyoto	Japan
Canon Research Center	Japan
Institute of Electronics, Information and Communication Engineers, Tokyo	Japan
The University of Tokyo	Japan
Department of Computer Science and Engineering, Pohang University of Science and Technology	Korea
Faculty of Information Technology & Quantitative Science, Universiti Teknologi MARA	Malaysia
Department of Media, Engineering and Mathematics, Delft University of Technology	Netherlands

Leiden Institute of Advanced Computer Science	Netherlands
Man-Machine Interaction Group, Delft University of Technology	Netherlands
Noldus Information Technology	Netherlands
Utrecht University	Netherlands
Vicar Vision, Amsterdam	Netherlands
Institute of Information Sciences and Technology, Massey University	New Zealand
Eastern Mediterranean University	Northern Cyprus
Department of Electrical and Computer Engineering, National University of Singapore	Singapore
Department Ciencias de la Computacion, Universidad Rey Juan Carlos	Spain
Department Inteligencia Artificial, Universidad Politecnica Madrid	Spain
Departament de Matemàtiques i Informàtica, Universitat de les Illes Balears	Spain
Departamento de Informática e Ingeniería de Sistemas, Instituto de Investigación en Ingeniería de Aragón, Universidad de Zaragoza	Spain
Universitat Politecnica de Catalunya, Barcelona	Spain
Department of Computer Science and Engineering, University of Moratuwa	Sri Lanka
aIDIAP—Dalle Molle Institute for Perceptual Artificial Intelligence, Martigny	Switzerland
Ascom Systec AG, Applicable Research and Technology, Mägenwil	Switzerland
Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan	Taiwan
Electrical Engineering Department, National Tsing-Hua University, Hsin-Chu	Taiwan
Centre for Vision Speech and Signal Processing, University of Surrey, Guildford	United Kingdom
Computer Laboratory, University of Cambridge	United Kingdom
Department of Computer Science, Queen Mary University London	United Kingdom
Department of Computing Science, University of Glasgow	United Kingdom
Department of Computing, Imperial College London	United Kingdom
School of Computing, Staffordshire University	United Kingdom
Beckman Institute, University of Illinois at Urbana-Champaign	USA
Computer Science Department, University of California, Santa Barbara	USA
Department of Computer Science, University of Nevada, Reno	USA
Department of Electrical, Computer and System Engineering, Rensselaer Polytechnic Institute	USA
Electrical & Computer Engineering Department, Florida International University	USA
Institute for Advanced Computer Studies - UMIACS, University of Maryland	USA
Machine Perception Lab, Institute for Neural Computation, University of California	USA
Massachusetts Institute of Technology, Cambridge	USA
Microsoft Research	USA
Real-Time Power and Intelligent Systems Laboratory, Department of Electrical and Computer Engineering, Missouri University of Science and Technology	USA
Robotics Institute, Carnegie Mellon University, Pittsburgh	USA
School of Interactive Computing, College of Computing, Georgia Institute of Technology	USA
University of Illinois at Urbana-Champaign	USA

Quelle: Eigene Recherche

Auf dem Weg zum Maschinensehen: Eine Geschichte von Standardisierungen und Reduktionen

Auf dem Weg zum Maschinensehen wird Maschinen unter anderem beigebracht, wie man Gesichtsausdrücke erkennt. Auf diese „Aufgabe“ konzentriert sich das hauptsächliche Interesse dieser Arbeit. Diese spezielle Aufgabe des Maschinensehens gliedert sich in eine Forschungsrichtung ein, die das Gesicht in das Zentrum der Aufmerksamkeit stellt. Die bekannteste Anwendung in diesem Zusammenhang ist die Gesichtserkennung (*face recognition*), also das automatische Wiedererkennen von individuellen Personen. Beide Anwendungen, sowohl Gesichtserkennung als auch Gesichtsausdruckserkennung basieren auf der Gesichtsdetektion (*face detection*), also dem Erkennen eines Gesichts auf einem digitalen Bild. Die Gesichtsdetektion ist auch der erste von drei Schritten im Prozess einer automatischen Gesichtsausdruckserkennung (Pantic & Rothkrantz 2000: 1425). Nachdem ein Gesicht auf einem digitalen Bild erkannt wurde, müssen die für einen Gesichtsausdruck relevanten Informationen und Gesichtszüge automatisch erkannt und extrahiert werden. Hier kann man entweder auf das Gesicht als Ganzes fokussieren (holistischer Zugang) oder auf einzelne charakteristische Bereiche des Gesichts achten (lokaler Zugang). (Fasel & Luetttin 2002: 262). Nachdem zunächst das Gesicht und die relevanten Gesichtselemente auf einem digitalen Bild erfasst wurden, muss in einem dritten Schritt der Gesichtsausdruck erkannt und klassifiziert werden. In diesem Fall geht man über eine reine Beschreibung von Gesichtsausdrücken hinaus und benennt Kategorien von Gesichtsausdrücken. Hierbei kann man sich entweder auf „prototypische Ausdrücke“ wie Emotionsausdrücke (z.B. Freude, Ärger etc.) oder auf „nicht prototypische“ Ausdrücke wie „hochgezogene Augenbrauen“ einigen. Die meisten Ansätze von automatischer Gesichtsausdruckserkennung beziehen sich auf die Emotionskategorien von Ekman (Pantic & Rothkrantz 2000: 1426; Zeng et al. 2007), obwohl es nicht abgesichert ist, dass Gesichtsausdrücke nur Emotionen abbilden. Eine automatische Gesichtsausdruckserkennung, welche sich nur auf die sechs Basisemotionen bezieht, stellt einen ersten Reduktionsschritt der vielfältigen menschlichen Gesichtsausdrücke dar, da sie alle anderen Gesichtsausdrücke und ihre vielfältigen, situativ unterschiedlichen Bedeutungen ausblendet (z.B. Anstrengung bei einem Radrennen).

Auf konzeptioneller, technischer und prozessualer Ebene lassen sich weitere Standardisierungen und Reduktionen erkennen, welche ich in der Folge genau beschreiben werde. Zunächst einmal stellt die Fokussierung auf das Gesicht und somit auf den biologischen Körper genauso wie der universelle Anspruch eine Reduktion dar. Auch die fehlende Einbindung des Kontextes reduziert die tatsächlich vorhandenen Informationen um ein Weiteres. In der Trainingsphase eines Systems zur automatischen Gesichtsausdruckserkennung kommt es zu weiteren Standardisierungen und Reduktionen:

Vor allem wenn es darum geht, eine Grundwahrheit (*ground truth*) zu bilden, muss man sich auf einen gewissen Zugang einigen, der die Komplexität reduziert. Auch die für die Trainingsdatenbank verwendeten Bilder von Gesichtsausdrücken müssen bearbeitet und transformiert werden (*Normalisierung*), um in weiterer Folge bearbeitet werden zu können. Sobald das System einmal arbeitet ist es meistens mit einer *Zwei-Klassen Entscheidung* konfrontiert, also der Frage, ob ein bestimmter Gesichtsausdruck besteht oder nicht. Grundlage dafür ist wiederum die Übersetzung des visuellen Bildmaterials in Zahlen als Repräsentanten. Und schließlich werden zu guter Letzt die Ergebnisse einer automatischen Gesichtsausdruckserkennung gefiltert und geglättet, um ein klares Ergebnis präsentieren zu können. Diesen hier knapp umrissenen Prozess werde ich in der Folge genau beschreiben und als Grundlage für eine abschließende *Conclusio* dieser Arbeit verwenden.

Biologischer Körper: Gesicht

Es klingt auf den ersten Blick vielleicht komisch, dass die Fokussierung auf das Gesicht bei einer Gesichtsausdruckserkennung eine Reduzierung darstellt, denn mit was sonst als dem Gesicht sollte sich eine Gesichtsausdruckserkennung beschäftigen. Ich will damit auch gar nicht die Gesichtsausdruckserkennung per se in Frage stellen, dennoch vertrete ich den Standpunkt, dass sich die Gesichtsausdruckserkennung nicht nur ausschließlich im Gesicht und noch weniger im biologischen Gesicht abspielen sollte, da dadurch eine Reduktion der menschlichen Kommunikation und Interaktion bzw. der Mensch-Maschine Interaktion begangen wird. Was Emotionen betrifft, obwohl immer wieder darauf hingewiesen wird, dass es einen Unterschied zwischen Emotionserkennung und Gesichtsausdruckserkennung gibt (z.B. Fasel & Luettin 2002: 259), werden Emotionen, in einem solchen Ansatz, der sich nur auf das biologische Gesicht bezieht, reduziert. Einerseits gibt es typische Emotionen wie etwa Hoffnung, die kein spezifisches Ausdrucksmuster besitzen (Averill et al. 1990 zit. nach Merten 2003: 61), andererseits muss der Ausdruck als nur ein Aspekt von Emotion angesehen werden (Bottenberg & Dassler 2002: 31), der Gesichtsausdruck wiederum ist neben anderen Ausdrücken wie der Stimme, dem Blick, dem Gestikulieren mit Händen und Armen, Kopfnicken, Geruchs- oder Geschmackssinn und vielen anderen Komponenten nur ein Ausdruck von vielen (ebd.: 51ff.). Schließlich werden auch Gesichtsausdrücke, die andere Funktionen der Interaktion und Kommunikation einnehmen (*verhaltensökologischer Standpunkt*), nicht nur anhand des spezifisch gezeigten Gesichtsausdrucks gedeutet, sondern werden in der Interaktion in einer bestimmten Situation und in einem spezifischen Kontext kommunikativ erschlossen. In ihrem Überblicksartikel fordern Fasel & Luettin eine Integration von anderen Kommunikationskanälen in Systeme zur automatischen Gesichtsausdruckserkennung (Fasel & Luettin 2002: 271). Auch Pantic & Rothkrantz

erwähnen die Vielfältigkeit der menschlichen Kommunikationskanäle (Pantic & Rothkrantz 2000: 1424).

Man kann davon ausgehen, dass bestimmte Gesichtsausdrücke angeboren sind, universell auf der ganzen Welt vorkommen und Emotionen zugeordnet werden können, dennoch bedeutet diese Fokussierung auf das biologische Gesicht eine Reduktion von sowohl Emotionen als auch von Gesichtsausdrücken. Meiner Auffassung nach ist ein Gesichtsausdruck und gegebenenfalls auch die damit verbundene Emotion nicht nur bloß im Gesicht ablesbar, sondern muss in einer situativ spezifischen Bedeutungszuweisung erst erschlossen werden. Die Wahrheit ist also nicht nur bloß im Gesicht zu finden, sondern auch in den zwischenmenschlichen Beziehungen, in welchen die Bedeutung kommunikativ erschlossen wird oder um es lapidar zu sagen, wenn Menschen miteinander reden werden sie sich schon darauf einigen, wie jetzt ein bestimmter Gesichtsausdruck zu deuten oder einzuordnen ist. Will man also ein System zur automatischen Gesichtsausdruckserkennung im Rahmen einer Mensch-Maschine Interaktion so gestalten, dass es für den Menschen „natürlich“ wirkt, so sollte dieser Aspekt in der Entwicklung eigentlich von Grund auf berücksichtigt werden.

Universeller Anspruch nicht haltbar

Ein funktionierendes System, das Gesichtsausdrücke automatisch erkennen soll, muss Gesichtsausdrücke von allen Menschen in gleicher Weise erkennen können. Gesichtsausdruckserkennung hat also einen universellen Anspruch, den es aber aufgrund der Komplexität der individuellen menschlichen Gesichter kaum einhalten kann.

„Ja, das Erkennen von Mimik ist extrem schwierig, weil eben der Status quo, von dem man ausgeht bei jedem menschlichen Gesicht sehr unterschiedlich ist. Sie kennen das, dass z.B. manche Leute diese klassisch hängenden Mundwinkel haben. Ein sehr berühmtes Beispiel bei Leuten, die sich mit Gesichtsausdruckserkennung auseinandersetzen ist Wally Friesen, der FACS Experte; Paul Ekman hat mehr so den horizontalen Mund, also da gehen die Mundwinkel jetzt nicht nach oben, aber Wally Friesen der hat einen Mund, das ist wirklich ein verkehrtes U. Und das ist für mich ein wunderbares Beispiel dafür, dass es wahnsinnig schwierig ist, Gesichtsausdrücke automatisch zu erkennen, weil natürlich die Muskeln mit dem arbeiten müssen, was schon da ist. Also wenn es da eine Möglichkeit gibt, irgendwie zu einem Erfolg zu gelangen, dann ist es wahrscheinlich die Dynamik, also was verändert sich.“ (Interview 2)

Die Physiognomik eines Gesichts ist bei jedem Menschen individuell, d.h. jeder Mensch hat eine andere Voraussetzung dafür, wie seine Muskeln, die ja für Gesichtsausdrücke verantwortlich sind, bewegt werden können. Die Konsequenz muss sein, dass auch jeder

Mensch individuell behandelt werden muss, wenn es darum geht, Gesichtsausdrücke automatisch erkennen zu können. Systeme, die mit statischen Bildern arbeiten, müssen aus diesem Grund kritisch gesehen werden. Doch auch bei Systemen mit dynamischen Bildsequenzen wird es kaum möglich sein, die jeweils individuelle physiognomische Grundvoraussetzung, die nach Faktoren wie Alter, Ethnizität, Geschlecht oder Lebenssituation unterschiedlich sein wird, ausreichend in die automatische Interpretation des Gesichtsausdrucks integrieren zu können. Ein weiterer Aspekt ist, dass sich die Bedeutungsinterpretation von Gesichtsausdrücken auch nach der individuellen Physiognomik unterscheidet:

„Die Physiognomik spielt auch eine große Rolle, wie wir die Bedeutung von Gesichtsausdrücken interpretieren. Z.b. ein weibliches Gesicht wird immer freundlicher und mit positiven Assoziationen besetzter Ausdruck wahrgenommen als das männliche Gesicht. Da gibt es sehr interessante Forschung, die wir teilweise gemacht haben aber vor allem Ursula Hess in Kanada macht, solche Geschichten mit Avataren, wo sie weibliche und männliche Avatare hat, wo man standardisiert Gesichtsausdrücke herstellen kann und findet da immer wieder, dass die Frauen einfach freundlicher wirken als die Männer.“ (Interview 2)

Neben dem Geschlecht können hier sicherlich auch andere Umstände in jeweils spezifischen Situationen eine Rolle spielen. Dabei ist nicht nur das, was quasi in der Physiognomik verankert ist von Bedeutung, sondern auch, wie wir gelernt haben, Menschen mit einer bestimmten Physiognomik einzuschätzen. Die persönliche Beziehung spielt sicherlich auch eine Rolle: Jemanden, den ich kenne, mit dem ich schon Erfahrungen gemacht habe, werde ich aus diesem Grund besser einschätzen können und werde einen bestimmten Gesichtsausdruck in einer bestimmten Situation besser deuten können, als den Gesichtsausdruck einer fremden Person.

Ein weiterer Aspekt, wieso es schwierig ist, dass automatische Gesichtsausdruckserkennung verallgemeinert werden kann ist die Möglichkeit von fehlenden Muskeln im Gesicht:

„Dann kommt noch dazu, FACS, super Instrument, ganz toll, aber nicht alle Menschen haben alle Gesichtsmuskeln. Ein sehr bekanntes Beispiel, oder weniger bekannt aber sehr aufregend ist, dass manchen Leuten dieser AU12 fehlt, der die Mundwinkel nach oben zieht, also der klassische „Lächelmuskel“, der fehlt manchen Leuten und die ersetzen dann dieses Lächeln durch einen anderen Muskel, ich hab jetzt die Nummer vergessen und da entsteht dann, man kann es am Besten als Joker Lächeln bezeichnen, das ist etwas komisch, eine andere Formung, die die da erreichen...“ (Interview 2)

Ein System, das die Gesichtsausdrücke von allen Menschen in gleicher Weise erkennen kann, egal wie alt sie sind oder woher sie kommen, muss auf die individuellen Voraussetzungen von Menschen eingehen können. Da die Grundvoraussetzungen aber

derart komplex sind und es kaum möglich ist, alle in ein System integrieren zu können, muss sich eine weitere Standardisierung und Reduktion ergeben, nämlich, dass ein System nur unter bestimmten Umständen für bestimmte Personen gültig ist. Ein System, das vorgibt, Gesichtsausdrücke universell erkennen zu können, dazu aber gar nicht in der Lage ist, kann je nach Anwendungsbereich gefährliche Konsequenzen haben. Sobald dieser Mangel des Systems einmal in Vergessenheit geraten ist, so ist die Wahrscheinlichkeit sehr hoch, dass man dem hochentwickeltem technischen System trotzdem mehr Vertrauen schenkt, gerade in Situationen der Unsicherheit würde man dem System vielleicht mehr Autorität zuschreiben als einem Menschen (Introna & Wood 2004: 183). Bei einer Evaluation eines Gesichtserkennungsalgorithmus (*facial recognition*) konnte man systematische Verzerrungen feststellen: So wurden Männer im Gegensatz zu Frauen, Asiaten und Afro-Amerikaner im Gegensatz zu Weißen und Ältere im Gegensatz zu Jüngeren besser wiedererkannt (ebd.: 190). Eine Folge davon könnte sein, dass diejenigen, die besser erkannt werden, mit höherer Wahrscheinlichkeit z.B. einen Alarm auslösen würden (ebd.: 192). Für die automatische Gesichtsausdruckserkennung bedeutet dies, dass systematische Verzerrungen aufgedeckt werden sollten bzw., dass eine Begrenzung der Aussagekraft eines Systems ausdrücklich betont werden und es deshalb nur ergänzend zu anderen Systemen oder menschlichen Einschätzungen verwendet werden sollte.

Fehlende Einbindung von Kontextinformationen

So unterschiedlich die Menschen sind, so unterschiedlich ist auch der Kontext von Situationen, in denen Gesichtsausdrücke automatisch gemessen werden können.

“What I may add here is that humans rely also on context information for facial expression estimation. So if you just see an image with a facial expression and you have no idea about the context of the image then you are not as good as recognizing the facial expression.” (Interview 3)

Wenn etwa eine Person an einem Werbeplakat vorbeigeht und vom System ein gewisser Ausdruck erkannt und festgehalten wird, kann die Ursache für einen bestimmten Gesichtsausdruck eine ganz andere sein. Ich kann mich über das schöne Wetter freuen und lächle deshalb, ich kann aber auch an meine hohe Stromrechnung denken und habe deshalb eher einen wütenden Gesichtsausdruck. Das beste Beispiel für die notwendige Integration von Kontextinformationen liefert wohl Paul Ekman, der—wie schon vorher angesprochen—davon ausgeht, dass die eigentliche Intention eines gezeigten Gesichtsausdrucks nicht erkannt werden kann, was er als den *Othello Fehler* bezeichnet hat (Ekman 2003: 218). Othello erkannte Desdemonas ängstlichen Ausdruck wie bereits erwähnt als eine Reaktion auf Angst vor Verrat. Dabei verrät der Gesichtsausdruck eben noch nichts über die Ursache

des Ausdrucks. Die Angst, dass einem nicht geglaubt wird sieht genauso aus wie die Angst, bei etwas erwischt zu werden. Hier möchte ich noch einmal an Ekmans Artikel in der *Washington Post* erwähnen, in dem er von einem Mann erzählt, der wegen seiner Gesichtsausdrücke am Logan International Airport in Boston als terroristisch verdächtig eingestuft wurde, es sich aber bei weiterem Nachfragen herausstellte, dass er auf dem Weg zu einer Beerdigung war und deshalb einen auffälligen Gesichtsausdruck zeigte (Ekman 2006).

Die Kontextinformationen lassen sich aber nicht nur in der Ursache von Gesichtsausdrücken finden, sie können auch in der direkten visuellen Umgebung gelegen sein. Ein Mensch mit einer Pistole am Kopf, der von der PistolenhalterIn zum Lächeln gezwungen wird, würde wohl von einem automatischen System auch als lächelnd erkannt werden, da es nur auf das Gesicht und nicht auf die Umgebung achtet. Einfache Systeme wären wohl auch nicht in der Lage, die „Echtheit“ des Gesichtsausdrucks erkennen zu können, obwohl diesbezüglich eine neuere Forschungsarbeit aus der Verhaltensforschung gezeigt hat, dass es wohl grundsätzlich möglich erscheint, echten Schmerz von gespielter Schmerz automatisch unterscheiden zu können (vgl. Littlewort et al. 2009). Dies würde aber voraussetzen, dass mit sehr exakten Systemen gearbeitet wird, die in einem Kontext außerhalb des Labors aber kaum zum Einsatz kommen würden.

Generell muss man an dieser Stelle aber festhalten, dass automatische Gesichtsausdruckserkennung bisher nur unter kontrollierten und standardisierten Bedingungen funktionieren kann, da sie die Vielfalt des Kontextes von Situationen nicht erkennen kann. Algorithmen zur Gesichtsausdruckserkennung sind bisher nicht in der Lage, die ganze Komplexität einer situierten Aktion (vgl. Suchman 1987; 2007) erfassen zu können, sie operieren also nur mit einem kleinen Teil einer Situation, die Komplexität der Gesamtsituation wird auf diesen kleinen Teil reduziert.

Bearbeitung von Bildern (Normalisierung)

Auf technischer Ebene können gewisse Kontextinformationen wie die Neigung des Gesichts oder die Beleuchtung „normalisiert“ werden. Idealerweise sollte ein System unter festgelegten, einheitlichen Bedingungen arbeiten. Sobald man diese aber verlässt und man mit geneigten Gesichtern oder mit dunklen Umgebungen konfrontiert ist, muss man diesen Umstand technisch ausgleichen.

„Und da ist immer noch das Problem mit der Beleuchtung. Also Beleuchtung hat einen starken Einfluss auf die Textur. Das Ganze funktioniert unter kontrollierten Beleuchtungsbedingungen

besser, es gibt mittlerweile auch Modelle, die Beleuchtung explizit modellieren, d.h. das ist ein eigener Parameter in dem Modell und den kann ich dann sozusagen rausrechnen.“ (Interview 1)

Wenn man also nicht die Beleuchtung an sich kontrollieren kann, muss man im digitalen Bild etwas ändern. Deshalb wird, wie in diesem Fall von einem Computerwissenschaftler beschrieben, die Beleuchtung modelliert und sozusagen auf eine einheitliche Beleuchtung standardisiert. Auch die Neigung des Gesichts wird normalisiert, um es weiter verarbeiten zu können:

„Man nimmt also das gesamte Gesichtsbild, versucht das irgendwie geometrisch zu normalisieren, und wenn man viele verschiedene Gesichter hat, die geometrisch normalisiert wurden über drei Punkte, z.B. Augen, Pupillen und vielleicht den Mund, dann liegen die Merkmale der Gesichter übereinander und dann kann man die Hauptkomponenten berechnen und kriegt dadurch sozusagen die Hauptmodi der Variationen, d.h. wie verändern sich Gesichter am stärksten, also das ist die Richtung der größten Varianz, kann z.B. sein, verursacht, dass manche einen Bart haben, manche nicht, dass da sehr große Varianzen verursacht, oder wenn andere Dinge, andere Gesichtsmerkmale große Varianz verursachen.“ (Interview 1)

„Also ich suche dann in dem Bereich des Gesichts, wo sind hier die zwei Stellen mit den Augen. Das mach` ich, weil die Augen nicht immer auf gleicher Höhe sind, das mach` ich um das Gesicht gerade zu rücken kann, damit ich eine möglichst normale, oder normalisierte Form des Gesichts habe.“ (Interview 4)

Vor der Analyse von Gesichtsausdrücken sollten die Gesichtsbilder also vor allem bezüglich der Gesichtsneigung und der Beleuchtung normalisiert werden, was man auch an anderer Stelle in der Literatur finden kann (Fasel & Luetten 2002: 261). Dieser Normalisierungsprozess der Bilder bedeutet nichts anderes, als dass die zu verarbeitenden Bilder standardisiert und reduziert werden. Ein weiterer Aspekt ist hier, dass insbesondere mit Grauwertbildern (schwarz/weiß Bildern) gearbeitet wird.

„Genau, wenn es nicht notwendig ist, wenn man weiß, dass schwarz/weiß Bilder, oder wesentliche Merkmale, die in einem schwarz/weiß Bild auch zu erkennen sind, Farbe ist nicht so wichtig, dann ist das natürlich ein Vorteil.“ (Interview 1)

Durch die Verwendung von Grauwertbildern kann die Datenmenge der zu verarbeitenden Bilder reduziert werden, was für eine schnellere Verarbeitung der Bilder von Vorteil ist.

„Wir verwenden ja auch keine Farbe, das ist also alles in Grauwertbildern. Das machen wir deshalb, weil wir festgestellt haben, dass Farbe, Farbauswertung nur dann funktioniert, wenn man eine definierte Farbtemperatur hat, die Farbtemperatur der Beleuchtung spielt eine wesentliche Rolle. Wenn ich beliebige Umgebungsverhältnisse habe, einmal Sonne, dann eine

Neonröhre, dann kann ich mit Farbe keine Gesichts-, oder keine Hautfarbe detektieren zur Einstellung der Regionen, die man untersucht.“ (Interview 4)

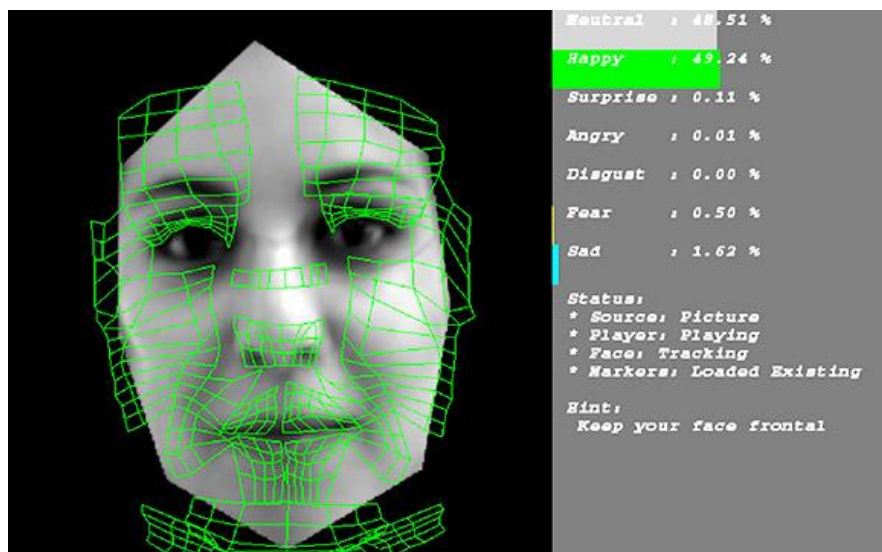
Auch diese Interviewpassage zeigt, dass man bei der Detektion von Gesichtern keine Verwendung für Farbe finden kann, da die unterschiedlichen Beleuchtungsbedingungen die Farben verfälschen. In der nächsten Interviewpassage wird deutlich, dass man zwar schon gerne mit Farbbildern arbeiten würde, aber die Voraussetzungen dafür nicht gegeben sind:

“It depends. Of course we like to have colour images. But for instance the Cohn Kanade database consists of black and white images. There is one pre-processing step in the facial expression approach that we use. That’s a skin colour detection.” (Interview 3)

Obwohl Farbbilder von wenig Nutzen sind, weil die verwendete Bilddatenbank aus schwarz/weiß Bildern besteht, wird in diesem Ansatz die Hautfarbe detektiert, es gibt also durchaus Verwendungsbereiche für Farbbilder.

Digitale Bilder sind die Voraussetzung für automatische Gesichtsausdruckserkennung. Damit diese aber auch verarbeitet werden können, um letzten Endes Gesichtsausdrücke feststellen zu können, müssen sie einerseits unter gewissen Gesichtspunkten standardisiert und vereinheitlicht werden, andererseits müssen sie auch auf die notwendigen Informationen reduziert werden, um sie möglichst schnell bearbeiten zu können. Letzten Endes wird also nicht mehr mit dem gefilmten oder fotografierten Gesicht gearbeitet, sondern mit dem standardisierten und reduzierten digitalen Bild, welches normalisiert wurde. Das Bild wird aber nicht nur standardisiert und reduziert, es werden zusätzliche Raster und Linien hinzugefügt (siehe Abbildung 10).

Abbildung 10: Bearbeitetes Gesichtsbild mit Raster



Was für die AnwenderIn verborgen bleibt ist der Transformationsprozess des Bildes, da sich das Outputbild direkt auf das Inputbild bezieht und die technische Verarbeitung und Normalisierung somit verschleiert (vgl. dazu Heßler 2006: 30f.). Wenn die Verarbeitung in Echtzeit passiert und man vielleicht sogar sein eigenes Gesicht auf einem Bildschirm sieht, so bekommt man den Eindruck, dass man es mit seinem Spiegelbild zu tun hat, also mit der Realität. Hier wird noch stärker vergessen, dass zahlreiche technische Transformationen zwischen Gesicht und dargestelltem und analysiertem Gesicht liegen. Das Ergebnis der Gesichtsausdruckserkennung—also dass man z.B. traurig ist—bezieht sich dann nicht mehr auf das eigentliche Gesicht, sondern auf andere Referenten wie Kameras, Algorithmen, Zahlen, geometrische Normalisierungen, Beleuchtungsmodellierungen usw. Daraus kann man schließen, dass sich das Ergebnis nicht auf die Natur oder die Wirklichkeit selbst, sondern auf eine andere, transformierte Realität bezieht, die durch den Einsatz von Computertechnik hergestellt wurde.

Die Grundwahrheit (ground truth) - Eine umstrittene Tatsache

Die Grundlage, um Maschinen das Sehen zu lernen ist das Anfertigen einer so genannten „Grundwahrheit“ (*ground truth*) oder „Grundwirklichkeit“ (*ground reality*). Aber was bedeutet das, wenn man zunächst eine Grundwahrheit anfertigen muss, handelt es sich dabei nicht um etwas, was einfach schon da ist? In der folgenden Interviewpassage erklärt ein Computerwissenschaftler was unter Grundwahrheit gemeint ist:

“...And maybe back to the ground truth question. That is, I said all our algorithms are based on machine learning and for machine learning you supervise machine learning that means that you give the machine example data to train from. So for instance if you want a machine to recognise a specific person then you show the machine images of this person and you tell the machine that this image shows that person. You give the correct answer already in the training phase. If you want to recognize laughing or fear or whatever you show the machine images of laughing or afraid persons and you tell the machine these images show laughing or afraid persons. And so the machine can recognize it later. But in the training phase this information has to be given and this is called ground truth, I hope this clarifies.” (Interview 3)

Die Grundwahrheit ist nicht etwa schon da, sie muss zunächst konstruiert werden. Bevor die Maschine eine Aufgabe überhaupt selbständig angehen und lösen kann, muss sie zunächst vom Computerwissenschaftler oder von der Computerwissenschaftlerin lernen. Diese/r erklärt der Maschine, wie die Grundwahrheit aussieht, zum Beispiel wie ein ängstlicher Gesichtsausdruck aussieht. In einer anderen Passage wird der Vergleich des Maschinenlernens mit dem menschlichen Lernen als ein Beispiel angeführt:

“... but it's pretty much to human learning. If you learn vocabulary then you have been given vocabulary. You have to match the German word to the English word. If you don't know the vocabulary and you hear the English word, you know the German word, you don't have to see it anymore. But during learning, of course, you have to match it. That's what the machine does.”
(Interview 3)

Zwei Dinge, die vermutlich dasselbe bedeuten müssen zunächst zusammengeführt werden. Und so besteht die Annahme, dass so wie ein deutsches Wort im Englischen ein Äquivalent hat und umgekehrt, so hat auch eine Emotion wie z.B. Angst ihr Äquivalent in einem spezifischen Gesichtsausdruck, der auf einem digitalen Bild dargestellt ist. Aber wie sieht dieses Äquivalent aus, wer zeigt der Maschine oder dem Computer überhaupt, wie dieses Äquivalent aussieht? In meinen Interviews habe ich zwei unterschiedliche Zugänge in dieser Frage gefunden: Einerseits einen *Facial Action Coding System (FACS)* ExpertInnen Zugang und andererseits einen laienhaften ComputerwissenschaftlerInnen Zugang. Zunächst werde ich den FACS ExpertInnen Zugang an einer Interviewpassage verdeutlichen:

a) *Facial Action Coding System (FACS) ExpertInnen Zugang*

I3a: Cohn-Kanade. That's a really standard database. Many, many people are working with that.

I3b: These databases are generated on the basis of ground realities. Cohn-Kanade facial expressions database is connected with the Facial Action Coding System.

C.M.: What does it mean ground reality?

I3b: For facial expressions there is a full coding of a face. That if you move your one eye up and if you are smiling so your lips are going up. Databases are generated by persons sitting in front of a camera.

C.M.: But people are said to do facial expressions?

I3a: In Cohn-Kanade they are advised to give an extreme facial expression. So if they have to smile, in the first streams they are neutral and in the ending they are really smiling. So they generate smile, it is not natural as when I am talking with you, but I am really forced to laugh.

C.M.: Is any expert controlling these expressions? Like in an experiment if anybody tells me to smile and is there anybody who says yes that's a correct smile or too much fake?

I3a: It depends on the database. In the Cohn-Kanade database it is like that. There is an annotation which tells you in which ways the person is smiling and it has been annotated by an expert to give evidence.

C.M.: And do you know more about these experts, from which profession do they come from? Are they psychologists?

I3a: Usually they are FACS experts, usually they annotate such data.

Die Grundwahrheit, mit dieser hier operiert wird, wird aufgrund von experimentellen und artifiziellen Daten geschaffen, welche zunächst von ExpertInnen annotiert werden müssen. Die Versuchspersonen im Labor werden dazu aufgefordert, einen bestimmten extremen Gesichtsausdruck vor einer Kamera zu zeigen. Und auch wenn es sich bei diesem gezeigten Gesichtsausdruck nicht um einen natürlichen, spontanen Gesichtsausdruck aus dem alltäglichen Leben handelt, werden diese Ausdrücke von FACS ExpertInnen annotiert so als wären sie natürlich und spontan verursacht (vgl. Ekman). In diesem Zugang wird die Grundwahrheit, die später die Basis für das Maschinensehen darstellen wird also in Koproduktion von LaiendarstellerInnen und FACS ExpertInnen in einem Labor geschaffen. Was in diesem Ansatz schließlich als Emotion gelten wird, also zum Beispiel als Angst, ist das Koprodukt von künstlichen—weil auf Anordnung dargestellt—Gesichtsausdrücken und einer Expertise von FACS ExpertInnen, die auf natürlichen und spontanen Gesichtsausdrücken basiert.

b) Laienhafter Zugang der ComputerwissenschaftlerInnen

Neben dem FACS Experten Zugang konnte ich in meinen Interviews einen zweiten Zugang zur Schaffung einer Grundwahrheit identifizieren, welcher sich aus der Alltagserfahrung der ComputerwissenschaftlerInnen speist:

C.M.: Woher kommen die Bilder, die sie haben...?

I4: Das ist eine Mischsammlung aus allen möglichen Quellen, die man so auftut. Und das fängt an bei irgendwelchen Datenbanken, geht weiter bei Bildern, die wir selbst fotografiert haben, das sind Bilder, die wir aus dem Internet sammeln und endet bei Bildern, die wir aus dem Fernsehen gewinnen. Das heißt, wir haben die Möglichkeit, mit einer Fernseekarte Gesichtserkennung laufen zu lassen und jedes Mal, wenn ein Gesicht gefunden wird, das entsprechende Bild abzuspeichern und damit kann man natürlich Gesichter sammeln...

... bei uns und bei dem Verfahren, wie wir diese Bibliothek trainiert haben ist ein völlig laienhafter Zugang, d.h. die Leute sind nicht trainiert, das sind einfach nur Menschen wie du und ich, die einfach aus ihrer Erfahrung heraus die Annotation durchführen. Und das ist natürlich auch mit ein Grund, dass das Ergebnis, das die trainierte Software liefert, nicht an Ergebnisse, also einfach relativ grob ist würde ich einfach einmal sagen. Keine feine Differenzierung zwischen verschiedenen Ausdrücken zulässt. Um in der Richtung weiterzuentwickeln muss man dann speziell annotierte Gesichter verwenden, die dann auch mit FACS oder von FACS Codern entsprechend konnotiert wurden. Da führt sonst kein Weg vorbei."

Dieser laienhafte Zugang basiert auf praktischem und implizitem Wissen, das bei den Personen, die die einzelnen Bilder, die mit in die Grundwahrheit kommen, vorhanden ist. Diese Personen, die die Daten annotieren, also zum Beispiel sagen, dass man auf einem bestimmten digitalen Bild Angst erkennt, sind die ComputerwissenschaftlerInnen und ihre Studierenden selbst. Sie haben die Gesichtsausdruckserkennung betreffend keine spezielle Ausbildung oder kein spezielles Wissen. In diesem Ansatz basiert die Grundwahrheit auf einer Bilderbibliothek, die sich aus unterschiedlichsten Quellen speist. Das Ziel der Bilderbibliothek ist es, möglichst viele unterschiedliche Bilder aus möglichst vielen unterschiedlichen Quellen zu sammeln. So wird vermieden, dass sich in der Datenbank nicht nur Bilder befinden, welche unter kontrollierten Bedingungen im Labor und unter der Annotation von ExpertInnen entstanden sind. Der laienhafte Zugang ist laut einem Computerwissenschaftler auch dafür verantwortlich, dass die Ergebnisse nur grobe Schätzungen sind. Andererseits nimmt man diese Schwäche in Kauf, damit das System in Echtzeit arbeiten kann. *“...es ist für uns essenziell für die Echtzeitfähigkeit, dass wir wirklich einfache Merkmale berechnen”* (Interview 4). Ich lese diese Essentialität zur Echtzeitfähigkeit des Systems als eine Demonstration dessen, was man als die „Natürlichkeit“ eines System bezeichnen könnte. Das System sollte so schnell wie Menschen Gesichtsausdrücken erkennen und deuten können. Exaktes detailliertes Expertenwissen könnte hier die Echtzeitfähigkeit beeinträchtigen.

Was wir an diesen beiden Zugängen erkennen können ist, dass die Konstruktion einer Grundwahrheit, welche die Basis für das automatische Erkennen von Gesichtsausdrücken darstellt, eine umstrittene Tatsache ist. Die Wahrheit ist nicht einfach von Natur aus da, sie muss erst produziert werden. Diese beiden unterschiedlichen Zugänge, um zu einer Grundwahrheit zu gelangen zeigen, dass es notwendig ist sich darauf zu einigen, was das „richtige“ Wissen ist, mit dem man operieren kann bzw. auf das man sich bezieht. Das Wort Grundwahrheit selbst deutet jedenfalls darauf hin, dass die Wahrheit im Gesicht und insofern im biologischen Körper gefunden werden kann. Aber wer hat nun Recht? Ist es einerseits die Wirklichkeit des Expertenwissens und der ExpertInnen, die genau wissen, wohin sie blicken müssen, um eine echte Emotion in der Expertenrealität erkennen zu können oder ist es das praktische Alltags- und Erfahrungswissen von Laien, die daran gewöhnt sind, Gesichtsausdrücke und Emotionen im alltäglichen Leben erkennen zu können? Am Ende könnte es sogar das System der automatischen Gesichtsausdruckserkennung selbst sein, welches eine spezifische Emotion auf Basis von statistischen Modellen definiert. Für die sechs Basisemotionen scheint es ExpertInnen zu geben, denen zugeschrieben wird, dass sie diese richtig erkennen können. Was ist aber mit all den anderen nicht prototypischen Gesichtsausdrücken? Wer sind dafür die ExpertInnen, wem wird zugeschrieben, dass er

oder sie richtig benennen kann, was man in einem bestimmten Gesichtsausdruck erkennen kann. Was passiert z.B., wenn man anhand des Gesichtsausdrucks die Müdigkeit abschätzen will. Reicht es dabei aus, den Lidschlag zu verfolgen? Wer sind in dieser Frage die ExpertInnen und wie wird man zu so einem Experten oder so einer Expertin?

Klassifikation der Kategorien

Der letzte Schritt einer automatischen Gesichtsausdruckserkennung ist die Klassifikation von Gesichtsausdrücken (Pantic & Rothkrantz 2000: 1437). Die Kategorisierung wird von einem Klassifikator vollzogen und besteht oft aus Musterverteilungsmodellen, die an bestimmte Entscheidungsprozeduren gekoppelt sind (Chibelushi & Bourel 2002: 4). Hier gibt es unterschiedliche Zugänge, die Klassifikatoren basieren entweder auf Schablonen (*template-based*), auf neuronalen Netzwerken (*neural-network-based*) oder auf Regeln (*rule-based*) (Pantic & Rothkrantz 2000: 1437). Das Ziel ist aber immer eine Einordnung des gezeigten Gesichtsausdrucks in eine der vorher festgelegten Kategorien.

„Wir haben in der Regel ein Zwei Klassen Problem, d.h. wir haben zum Beispiel ein Gesicht von der Klasse nicht-Gesicht zu unterscheiden oder die Klasse Mann von der Klasse Frau zu unterscheiden oder Klasse fröhlich von nicht-fröhlich. Das System ist so trainiert, dass wir viele Merkmale auch von dem Gesichtsausschnitt berechnen. Also wir haben zur Zeit 30.000 Merkmale zur Verfügung, die wir im Gesicht auf unterschiedlichste Art in unterschiedlichsten Skalierungen berechnen können. Dann brauchen wir eben eine große Menge an Beispielen, weil wir dann eine Merkmalsauswahl treffen und dieses Verfahren, das funktioniert so, dass wir versuchen, jene Merkmale herauszubekommen, die möglichst signifikant unterscheiden können zwischen den beiden Klassen. Der Ansatz der da dahinter steckt, geht von starken und schwachen Klassifikatoren aus. Jedes einzelne Merkmal, das verwendet wird, z.B. der Kontrast an irgendeiner Stelle am Mundwinkel ist ein schwacher Klassifikator, d.h. mit diesem einzelnen Merkmal ist man in der Lage, zwei Merkmale ein bisschen zu trennen. D.h. man kann nicht eine 50:50 Entscheidung herbeiführen, sondern z.B. eine 48% zu 52% Entscheidung. Wenn man jetzt aber viele dieser Merkmale geschickt kombiniert, in Gewichtung und in Auswahl, dann ist man in der Lage, eben einen starken Klassifikator zu bauen, der dann eben eine starke Erkennungsleistung von z.B. 95% hat.“ (Interview 4)

Wie in dieser Interviewpassage deutlich wird, ist die Klassifikation eine Unterscheidungsleistung zwischen zwei möglichst signifikanten Klassen, eine entweder-oder-Entscheidung, ob es sich jetzt um die eine oder die andere Klasse handelt. Abstufungen sind aber grundsätzlich auch möglich:

„Wobei wir bei der Klassifikation schon so eine Art Konfidenzmaß bekommen, das dann noch darüber Aussage gibt, ob das jetzt ein starkes Lachen ist oder ein leichtes Lachen. Aber es ist

nicht speziell trainiert, das kann man dann sicher mit anderen Methoden exakter feststellen.“
(Interview 4)

Die Auswahl der Klassen und die möglichen Abstufungen (stark/schwach) müssen aber schon vorab festgelegt werden, in der *ground truth* fixiert werden. Um überhaupt ein Ergebnis erhalten zu können, muss die gesamte beobachtete Situation, in welcher sich ein Gesichtsausdruck zeigt, in Kategorien eingeteilt und klassifiziert werden. Das Ergebnis kann sich also nur in ein schon vorgegebenes Schema einordnen, alle anderen Informationen gehen verloren bzw. können überhaupt nicht verarbeitet werden. Wenn man den Anspruch stellt, dass das Maschinensehen dem menschlichen Sehen nachgebildet wird, so muss man sich fragen, ob das Einordnen des Wahrgenommenen in Kategorien einen quasi-natürlichen Prozess darstellt, den wir als sinnvoll erachten oder der für uns eher befremdlich ist. In unserem alltäglichen Leben ist es kaum möglich, Klassifikationen und Standards zu entkommen: Der Gang auf die öffentliche Toilette ist nach Geschlechtern klassifiziert, die Bücher in einer Bibliothek sind nach verschiedenen Klassen geordnet oder die Passkontrolle am Flughafen ist nach verschiedenen Nationalitäten organisiert (Bowker & Star 2000: 3). Nur: Woher kommen diese Kategorien und Klassifikationen? Sind sie schon immer dagewesen oder sind sie nicht erst mit der Zeit entstanden, sind sie „natürlich“ oder vielmehr historisch und politisch gewachsene Artefakte (ebd.: 285)? Wenn Systeme zur automatischen Gesichtsausdruckserkennung mit feststehenden Kategorien oder Klassen arbeiten, muss immer auch hinterfragt werden, wo diese Kategorien herkommen, wie sie entstanden sind und was das für Konsequenzen auch für das menschliche Sehen und Wahrnehmen hat. Ein System, das nur sechs oder sogar weniger Emotionen erkennen kann stellt somit automatisch eine Reduzierung der wahrgenommenen Wirklichkeit dar.

Repräsentation in Zahlen

Der schon eben angesprochene Vergleich der *ground truth* und des zu analysierenden digitalen Bildes passiert notwendigerweise über den Umweg von Zahlen. Die Bilder können nicht direkt miteinander verglichen werden, der Computer oder die Maschine benötigt Zahlen, deshalb muss das visuelle Datenmaterial in Form von Zahlen repräsentiert werden.

“What you try to do is to reduce the image which is a lot of data to some smaller data, represent the image by a couple of numbers that really represent the image content. And from this you try to train some machine learning technique. If you just want to recognize the facial expression then you can use classification, so you classify those describing numbers to laughing, surprise, fear, whatever or you can alternatively use regression. That means that you get a number out of it, then you can say the person is laughing with 50% or is afraid by 25% That are two different techniques. And also it depends on the context. If you have to know how much a person is

laughing you should train some regression algorithm and if you want to know just if a person is laughing than you can use classification.” (Interview 3)

Es gibt also ein vorgegebenes Set an Zahlen, die einen gewissen Gesichtsausdruck bzw. eine gewisse Emotion (z.B. Angst), also eine gewisse Klasse oder Kategorie, beschreiben und repräsentieren. Dieses Set an Zahlen wird mit dem Set an Zahlen des wahrgenommenen Gesichtsausdrucks verglichen und es sind nicht Bilder, die miteinander verglichen werden.

“That is what the computer has about the face and then it tries to fit the model into the image. It tries to find model parameters that describe the face by a couple of numbers and it tries to estimate the numbers that fit to the image, that the model really describes what the face in the image looks like. With a couple of numbers this is something the computer can handle. And these couple of numbers it depends on the model how many that are they are put into some machine learning algorithm and the computer learns what numbers represent which facial expression. For instance, maybe it's less abstract, one number could be the opening of the mouth. And one number could change this to the model. So if the computer gets an image with a widely opened mouth then the corresponding parameters should be estimated that the model also is a widely opened mouth. And then maybe the machine learning algorithm could find out if the opened mouth represents surprise for instance. This is how it works.” (Interview 3)

Mit Zahlen kann der Computer umgehen. Der Computer lernt im Grunde nicht von Bildern, sondern von Zahlen, die das Bild bzw. die benötigte Bildinformation in Form einer Kategorie repräsentieren. Ohne diese Übersetzung in Zahlen hätte der Computer größte Schwierigkeiten, die visuelle Information verarbeiten zu können. Visuelles Wissen muss in Zahlen umgewandelt werden, dass es für einen Computer erkennbar wird, was vermittelt werden soll. Der Computer kann diese Zahlen aber nicht ohne zuvor erhaltene Informationen deuten. Er braucht schon vorab ein Bündel an Zahlen, die ein gewisses Objekt oder eine gewisse Aktion repräsentieren. Diese Zahlen sind aber nicht schon vorab da, man muss sich erst auf eine gewisse Art der Quantifizierung einigen (vgl. Porter 1995; Fasching 2007). Schließlich müssen Informationen auch in Zahlen übertragen werden, um Kommunikation über ein Datennetz oder Vergleichbarkeit der Daten zu ermöglichen.

Quantifizierung—und damit auch Standardisierung—stellt bei den Technologien zur automatischen Gesichtsausdruckserkennung eine entscheidende Übersetzungsleistung dar. Zunächst wird das visuelle Wissen, welches Informationen über den aufgezeichneten Gesichtsausdruck liefert, in Zahlen übersetzt, woraufhin diese Zahlen wiederum in sprachliche Kategorien wie z.B. eine Emotionskategorie übersetzt werden. Zwischen Visuellem und Sprachlichem stehen also Quantifizierungen. Dass dadurch wertvolle Informationen, die vielleicht Bestandteil einer umfassenden sprachlichen Beschreibung der

Situation gewesen sein könnten, verloren gehen (vgl. Daston 2001: 147), scheint bei diesem Zwischenschritt der quantifizierenden Übersetzung sehr wahrscheinlich zu sein.

Filtern und Glätten

Um ein sauberes und eindeutiges Ergebnis erreichen zu können, muss das beobachtete und wahrgenommene Material gefiltert und geglättet werden. Dazu werden von den ComputerwissenschaftlerInnen Schwellenwerte oder Grenzen festgelegt, etwa ab wann eine bestimmte Klasse eintritt. Diesen Prozess des Filterns und Glättens möchte ich anhand des Geschlechts beschreiben. Ich weiche hier bewusst von Gesichtsausdrücken ab, da ich in meinem Interviewmaterial eine enge Verbindung von Gesichtsausdruckserkennung und Geschlechtserkennung festgestellt habe, welche ich nicht vorenthalten will. Gerade unter diesem Punkt des Filterns und Glättens finde ich es besonders spannend, diesen Exkurs zum Geschlecht zu machen, da wir in unserem alltäglichen Leben das Geschlecht nicht als eine Kategorie wahrnehmen, die—wahrscheinlich weniger als Gesichtsausdrücke—gefiltert und geglättet werden muss, um es erkennen zu können. Das Geschlecht oder besser die Geschlechterdifferenz erscheint in diesem Zusammenhang als ein Ordnungsinstrument, das wir als natürlich ansehen und das anscheinend immer wieder hergestellt werden muss, auch von der Wissenschaft. So kam es, dass in den 1990er Jahren die Hauptaufgabe der Computerwissenschaft die Detektion von Gesichtern in digitalen Bildern war. In der folgenden Interviewpassage beschreibt ein Computerwissenschaftler, wie sich der Weg von der Gesichtsdetektion zur Gesichtsausdruckserkennung, aber auch zur Geschlechtserkennung entwickelte.

„In den letzten Jahren haben wir dann weiter überlegt, was kann man denn, wenn man jetzt ein Gesicht schon so robust detektieren kann, was kann man dann weiter mit diesen Gesichtern anstellen. Und die naheliegendste Variante war dann, man kann natürlich Teile im Gesicht finden, man kann nach den Augen suchen, man kann nach der Nase suchen, man kann nach dem Mund suchen. Wir haben da versucht die Algorithmen anzupassen, das Training anzupassen und haben dann weiterhin noch versucht z.B. zu erkennen, ob die Augen geschlossen sind, ob der Mund geschlossen oder offen ist. Und das hat dann auch ganz gut funktioniert und der nächste Schritt, den wir gemacht haben und der dann aus den Überlegungen heraus entstanden ist. Kann man nicht vielleicht auch anhand des Erscheinungsbildes des Gesichts auch das Geschlecht unterscheiden, kann man nicht auch Gesichtsausdrücke erkennen. In dem Gebiet ging es dann weiter und da haben wir dann auch für beide Themen Geschlechtserkennung und Gesichtsausdruckserkennung erste Ergebnisse bekommen, die dann auch ganz vielversprechend waren. Und so ging es halt in die Richtung, und jetzt geht es halt weiter z.B. kann man auch das Alter schätzen anhand des Gesichts. Eine spannende Sache.“
(Interview 4)

Nach der einfachen Beschreibung von Gesichtsteilen und einfachen Bewegungen wie dem Augenschließen oder Mundöffnen, waren die Gesichtsausdruckserkennung und die Geschlechtererkennung das offensichtlichste, was man als nächstes in einem Gesicht automatisch erkennen könnte. Dies impliziert, dass sowohl Gesichtsausdrücke als auch Geschlecht schon da sind und man sie nur erkennen muss. Aber wie kann ich einem Computer oder einer Maschine beibringen, das Geschlecht in einem Gesicht zu erkennen? Der Computer kann das jedenfalls nicht alleine, es muss auch hier eine Grundwahrheit (*ground truth*) gebildet werden. Jedes neue Gesichtsbild, das für das Training verwendet wird und somit in die Bilddatenbank aufgenommen wird muss vorher handannotiert werden. Auf Basis des praktischen alltäglichen Wissens entscheidet der Computerwissenschaftler, ob auf einem Bild ein Mann oder eine Frau zu erkennen ist. Es ist eine einfache Entscheidung und es kann nicht nachvollzogen werden, auf welchen Kriterien diese Entscheidung basiert. Bei den meisten Bildern hat der Computerwissenschaftler mehr Informationen als das biologische Gesicht zur Verfügung: Er kann seine Entscheidung aufgrund der Frisur, der Kleidung, dem Make-up usw., also einer Vielzahl an kulturellen und sozialen Zeichen basieren. In der Folge muss aber der Computer mit der biometrischen Form auskommen und erkennt diese als entweder männlich oder weiblich an, also so, wie es ihm vorher gesagt wurde. Schließlich kann man kaum mehr sagen, ob der Computer auf Basis des biologischen Geschlechts (*sex*) oder des sozialen Geschlechts (*gender*) arbeitet und man muss wohl eher von einer hybriden Form der beiden ausgehen. Auch was die klare Unterscheidung von männlich und weiblich betrifft, gibt es ein Durcheinander:

“...That’s what I was saying in morphing faces they have a slider and you can move the slider from more man to more woman. So it’s getting man man man and in the final end he’s a man. So you’re right, in the middle there are some confusions. Like you can on the slider make my face as a female. Then somebody shows you my face with different sliders. More man less man.”
(Interview 3)

Jetzt versteht man vielleicht besser, wieso auch das Geschlecht gefiltert und geglättet werden muss, um schließlich zwei klare Kategorien erreichen zu können. Bei einer einzelnen Person kann es in einer kurzen Situation zu Verwirrungen kommen, was unter anderem auch an Gesichtsbewegungen und Kamerapositionen liegt. In Abbildung 11 sind zwei verschiedene Fotos des Fußballspielers Cristiano Ronaldo zu sehen, die ich für eine Demonstration ausgewählt habe. Im linken Bild detektiert die Demoversion der Software SHORE ihn „fälschlicherweise“ als weiblich, im rechten Bild als männlich.

Abbildung 11: Fließende Geschlechtsidentität



Quelle: Eigene Verwendung der Software SHORE

Während ein Gesicht von der Kamera wahrgenommen und von der Software verarbeitet wird, kann das Geschlecht einer Person durchaus fließend sein, aber am Ende ist das Ergebnis, also das, was wir am Bildschirm angezeigt bekommen, entweder männlich oder weiblich. Die fließende Grenze der Geschlechter muss bei manchen Personen gefiltert werden:

“... wir berechnen auf Einzelbildern, was wir aber dann machen, wir filtern die Ergebnisse. Filtern heißt, wenn wir jetzt, und das bezieht sich durchaus auf Geschlecht, wenn wir jetzt über die Zeit hinweg, ein Gesicht klassifizieren, das jetzt vielleicht relativ nah an der Schwelle zwischen männlich und weiblich ist, dann kann das durch Kopfdrehen oder andere Störungen durchaus auch rüber springen in die andere Richtung und dann versuchen wir natürlich schon das zu glätten und rauszufiltern. Sowohl bei Geschlecht, als auch bei Emotionen.

C.M.: Was heißt glätten oder herausfiltern genau?

„Das wird über die Zeit einfach gemittelt. Und wenn jetzt bei zehn Folgebildern neunmal ein Mann und einmal eine Frau erkannt wurde, dann wird dieses als Frau erkannt werden nicht gezeigt. Herausgefiltert.“ (Interview 4)

Am Ende gibt es kein 90% Mann und 10% Frau, am Ende gibt es in diesem Fall eine klare Entscheidung dafür, dass es sich um einen Mann handelt. Im Gegensatz zu den Gesichtsausdrücken, wo durchaus gröbere Abstufungen zwischen schwacher und starker Ausprägung gemacht werden können, gibt es beim Geschlecht ein eindeutiges Ergebnis, das aber zuvor gefiltert und geglättet wurde.

Conclusio

Maschinen, Computer, Automaten, Roboter und Agenten sollen mit der Fähigkeit zum Sehen und Erkennen ausgestattet werden. Wie wir Menschen sollen diese neuen AkteurInnen der sozialen Welt, also des Zusammenlebens von verschiedenen menschlichen und nicht-menschlichen Wesen, sehen und somit ihre Umwelt bzw. Mitwelt wahrnehmen können. Auf diesem Weg zum Maschinensehen muss notwendigerweise die Frage gestellt werden, wie denn eigentlich Menschen sehen, erkennen und wahrnehmen. Deshalb fragte ich zunächst, welches implizite Modell vom Sehen und Erkennen in den seit etwa zwei Jahrzehnten zunehmend entstehenden Technologien zur automatischen Gesichtsausdruckserkennung stecken und wie alltägliche, kulturell geschliffene Annahmen über das menschliche Sehen mit den Technologien und Anwendungen dieser speziellen Aufgabe des Maschinensehens verknüpft sind. So sehr auch die Sehnsucht verfolgt wird, das menschliche Sehen und Erkennen im Maschinensehen nachzuahmen, so sehr ist dieser Vorgang ein Prozess der Standardisierung und Reduktion. Dieser Prozess wurde im empirischen Teil dieser Arbeit von mir rekonstruiert. Auf den verschiedenen Ebenen der Konzeption, der technischen Umsetzung und der prozessualen Verarbeitung wurden von mir Schritte der Standardisierung und Reduktion beschrieben. Zusammenfassend lässt sich sagen, dass die technische Komponente des Sehens und Erkennens vorherrschend ist und die interpretative Komponente nur in Ansätzen zur Geltung kommt. Das Sehen und Erkennen wird also auf ein technisches und regelgeleitetes Sehen reduziert. Mit einer solchen Reduzierung geht einher, dass nur andeutungsweise die ganze Komplexität des menschlichen Sehens und Erkennens nachgebildet werden kann und dass dadurch auch nur Hinweise und Tendenzen darüber geschaffen werden können, was gesehen, erkannt und wahrgenommen wird. Es reicht nicht aus, sich nur auf die technische Komponente des Sehens und Erkennens zu konzentrieren, vielmehr müssen auch die komplexen Informationen über den Kontext einer Situation mit einbezogen werden. Denn Information steht nicht für sich selber, sondern sie wird in einem gewissen Kontext interpretiert (Richter 2002: 104). Ebenso sind Handlungen nur in einem bestimmten situativen Zusammenhang verstehbar (ebd.: 109).

Eine Technologie, die sich darauf konzentriert, nur das visuell erkennbare zu sehen und wahrzunehmen blendet die kommunikative Erschließung der Bedeutung des visuell Erkannten aus. Sehen kann nicht der Summe der einzelnen wahrgenommenen und registrierten Komponenten entsprechen, es muss davon ausgegangen werden, dass Sehen im Wandel ist und in je spezifischen Situationen andere Bedeutungen von ein und

demselben Objekt oder von ein und demselben Verhalten hervorgerufen werden. Charles Horton Cooley, ein Vorläufer des Symbolischen Interaktionismus, unterschied zwischen einer physischen Umwelt, also greifbaren Gegenständen, denen dingliches oder räumliches Wissen entspricht und einer sozio-kulturellen Umwelt, der soziales Wissen entspricht. Dieses soziale Wissen entwickelt sich im Gegensatz zum dinglichen Wissen, das in den Naturwissenschaften immer feiner gemessen und quantifiziert wird, erst in der Auseinandersetzung und Kommunikation mit den Denkinhalten von anderen Menschen (Helle 1977: 53f.). In den gegenwärtigen Systemen zur automatischen Gesichtsausdruckserkennung fehlt diese Komponente des bei Cooley genannten sozialen Wissens gänzlich.

Ein System, das Gesichtsausdrücke erkennen kann, ist weiters noch nicht problematisch, solange es verschiedene Gesichtsausdrücke, die sich über Muskelbewegungen ergeben, nur beschreibt und detektiert. Doch sobald wir diesen Gesichtsausdrücken Bedeutung zuweisen wollen und für die Beschreibung notwendigerweise auch Worte brauchen, die in der Folge Klassen und Kategorien von Gesichtsausdrücken bilden sollen, begeben wir uns auf unsichere Pfade, wenn wir dies quasi ein für alle mal klären wollen. Gesichtsausdrücke nur zu beschreiben, etwa in Form von festgelegten Aktionseinheiten (*action units*) bringt noch nicht die Informationen hervor, die wir verarbeiten können, da sie für uns fremd und abstrakt sind. Diese Beschreibungen müssen, um weiter verarbeitet werden zu können, in Worte, Muster, Kategorien usw. übersetzt werden, um darin auch Sinn zu finden. Die Frage ist nur, ob die Einordnung in klare Kategorien alleine ausreicht, um diesen Sinn finden zu können. Dies ist insofern fragwürdig, da der Mensch es gewohnt ist, durchaus flexibel auch Mischformen erkennen und einordnen zu können, ohne sich dabei auf klar abgegrenzte Kategorien zu beziehen.

An diesem Punkt scheinen vorerst zwei Wege möglich zu sein: Einerseits können sich Systeme zur automatischen Gesichtsausdruckserkennung auf die vorgegebenen Kategorien der sechs Basisemotionen Freude, Trauer, Wut, Ekel, Furcht und Überraschung beziehen, andererseits kann daran gearbeitet werden, andere nicht prototypische Übersetzungen zu finden, die dem dynamischem, flexiblem menschlichem Sehen, Erkennen und Einordnen, das nicht unbedingt auf Kategorien angewiesen ist, näher kommen. Im ersten Fall würden Gesichtsausdrücke auf die sechs universell vorkommenden Basisemotionen reduziert werden und alle anderen Emotionen, genauso wie alle anderen Informationen und Signale der vielfältigen menschlichen Gesichtsausdrücke verloren gehen. Im zweiten Fall fehlt allerdings größtenteils noch das Beschreibungs- und Interpretationsvokabular, mit dem wir Gesichtsausdrücke in all die anderen möglichen nicht prototypischen Übersetzungen transformieren könnten.

Unter diesen Voraussetzungen erscheint es zum gegenwärtigen Zeitpunkt folgeschwer, wenn sich Technologien zur automatischen Gesichtsausdruckserkennung ausbreiten und Anspruch auf Allgemeingültigkeit erlangen. Denn wir erinnern uns: Die Konstruktion von wissenschaftlichen Tatsachen erscheint jedem im Nachhinein als unkonstruiert (vgl. Latour & Woolgar 1979) und somit als Teil der schon immer vorhandenen Natur. Technik ist stabilisierte Gesellschaft (Latour 2006). Breiten sich also Technologien zur Gesichtsausdruckserkennung aus, so wird vergessen, dass sie eigentlich unter bestimmten Voraussetzungen hergestellt wurden und eine bestimmte Komponente des Sehens und Erkennens —die technische—bevorzugt worden ist. Der Prozess der Standardisierung und Reduktion gerät in Vergessenheit. Die Folgen können schwerwiegend sein: Ein System, das vorgibt, Gesichtsausdrücke wie ein Mensch erkennen zu können, dazu aber gar nicht in der Lage ist, kann je nach Anwendungsbereich gefährliche Konsequenzen haben. Sobald dieser Mangel des Systems, also nur eine reduzierte Fähigkeit des menschlichen Sehens zu besitzen, einmal in Vergessenheit geraten ist, so ist die Wahrscheinlichkeit sehr hoch, dass man dem hochentwickelten technischen System trotzdem mehr Vertrauen schenkt. Gerade in Situationen der Unsicherheit würde man dem technischen, neutralen System vielleicht mehr Autorität zuschreiben als der scheinbar beeinflussten Meinung eines Menschen (Introna & Wood 2004: 183). Aus diesem Grund dürfen Systeme der automatischen Gesichtsausdruckserkennung keine allgemeine Gültigkeit erlangen. Sie können unter kontrollierten Bedingungen in Labors angewendet werden und tendenzielle Aussagen hervorbringen. Sie sind aber nicht verallgemeinerungsfähig. Dazu sind sowohl die Menschen, als auch die Situationen, in denen sie sich befinden, genauso wie die Bedeutungen, die die Menschen den Situationen zuweisen, einfach zu verschieden. Um diesem Umstand gerecht werden zu können, müssten automatische Systeme individuell über einen langen Zeitraum trainiert werden, die individuellen Eigenheiten von Personen und ihren Gesichtsausdrücken müssten erlernt werden. Dieser Lernvorgang müsste auch berücksichtigen, dass das Wissen, mit dem diese Systeme operieren, nicht statisch, sondern dynamisch ist. Das Wissen verändert sich und ist außerdem auch vom Kommunikationskontext der Situation abhängig.

Eine weitere Konsequenz von verallgemeinerten, sich still und heimlich ausbreitenden Systemen zur automatischen Gesichtsausdruckserkennung könnte bei Menschen sogar der Verlust der Fähigkeit bzw. die Entwertung der menschlichen Fähigkeit sein, Gesichtsausdrücke bzw. Emotionen erkennen zu können. Wenn der Soziologe Richard Sennett von den BäckerInnen in Boston erzählt, die gar nicht mehr Backen können, weil sie es durch das Vorhandensein von automatischen Backsystemen gar nicht mehr können müssen und es deshalb verlernt haben (Sennett 2006: 87), so können wir uns auch fragen,

was passiert, wenn wir nicht mehr unserem eigenen Urteil vertrauen können, wenn es darum geht, Gesichtsausdrücke und Emotionen erkennen zu können, sondern dem Expertenurteil der emotionalen Maschinen, die ein Expertensystem (*expert system*) (vgl. Forsythe 2001 zit. nach Suchman 2008: 144) darstellen, mehr Einfluss zumessen. Sobald wir dem Urteil der Expertenmaschine mehr vertrauen als dem eigenen, geraten unsere eigenen Fähigkeiten in Vergessenheit. Unsere eigenen Fähigkeiten liegen nicht unbedingt darin, einen Gesichtsausdruck aufgrund der einzelnen Muskelbewegungen im Gesicht eines Menschen genau identifizieren und zuordnen zu können, sondern insbesondere in der kommunikativen Erschließung in einer spezifischen Situation. Ein spezifischer Gesichtsausdruck lässt genügend Interpretationsspielraum frei, um im Prozess der Bedeutungszuweisung in verschiedene Richtungen gehen zu können. Ein von einem automatischen System erkannter trauriger Gesichtsausdruck könnte beispielsweise im Falle eines Radrennens als ein Anzeichen für Anstrengung, aber im Falle eines Liebesfilms als Anzeichen für Liebeskummer gewertet werden. Gesichtsausdruckserkennungsalgorithmen dürfen zum gegenwärtigen Zeitpunkt immer nur in Zusatz oder als Ergänzung zu anderen Aktivitäten eingesetzt werden, da ihre Ergebnisse nur einen reduzierten und transformierten Teil einer spezifischen komplexen Gesamtsituation, ja im Endeffekt der jeweils erlebten Wirklichkeit darstellen. Ähnlich wie die automatische Spracherkennung (Mayer 2001: 163) ist auch die automatische Gesichtsausdruckserkennung eine für den Einzelnen kaum noch durchschaubare, komplexe Technologie, die durch Standardisierungen und Transformationen Komplexität reduziert, dabei aber vergisst, dass die Welt durch komplexe Situationen, die durch Kommunikation entstehen und dadurch bewältigt werden, aufgebaut ist. Schließlich wird die Komplexität der menschlichen Kommunikation nicht nur transformiert und reduziert, sie wird durch diese Reduktion auf maschinenähnliche Vorgehensweisen auch kontrolliert, indem sie besser rationalisierbar, berechenbar und vorhersagbar wird (vgl. Ritzer 1995).

Warum die Sehnsucht, das Innere des Menschen in seinem Äußeren zu erkennen, eine Sehnsucht bleibt und trotzdem in Erfüllung geht

Die Sehnsucht, das Innere des Menschen in seinem Äußeren zu erkennen, wurde in dieser Arbeit insbesondere an den gegenwärtigen Technologien zur automatischen Gesichtsausdruckserkennung verdeutlicht, doch wurde sie zunächst im ersten Teil dieser Arbeit historisch verankert. Es wurde gezeigt, dass sich die Physiognomik aus einem alltagspragmatischen Bedürfnis der Menschen heraus entwickelt hat, unser Gegenüber—insbesondere die den Menschen fremden Tiere oder fremden Menschen—besser einschätzen zu können. Die Verwissenschaftlichung der Physiognomik durch Johann Caspar Lavater scheiterte, als er die vielfältigen Charaktereigenschaften des Menschen auf

verallgemeinerte, rational kalkulierbare Gesetze, die am menschlichen Gesicht abzulesen waren, zurückführen wollte. Obwohl er unter anderem wissenschaftliche Methoden wie Quantifizierungen anwendete, wurden die Interpretationen seiner Messungen als subjektiv entlarvt. In ähnlicher Weise musste sich auch Franz Joseph Gall für seine Schädellehre „Charlatanismus“ vorwerfen lassen. Charles Darwin, der in der physiognomischen Traditionsliste gesehen werden muss, wendete sich aber von den festen Körperteilen ab und den weichen, beweglichen Körperteilen zu. Darwin wollte über den Gesichtsausdruck die noch fehlende Verwandtschaftsbeziehung zwischen Menschen und Tieren herstellen. In der Folge wurde der Gesichtsausdruck als angeboren und vom Willen des Menschen unabhängig angesehen. Darwin stellte auch die These auf, dass Gesichtsausdrücke universell unter allen Menschen vorkommen. Diese These nahm Paul Ekman in den 1960er Jahren wieder auf, und als er im Zuge des Vietnamkrieges die Möglichkeit durch umfassende Finanzierung der *Advanced Research Projects Agency (ARPA)* des U.S. Verteidigungsministeriums erhielt, wollte er diese Frage ein für alle mal klären. In diesem Zusammenhang stand wohl die Frage im Vordergrund, ob der feindliche Soldat Nordvietnams andere oder ähnliche Gesichtsausdrücke zeigt wie der Amerikaner. Ekman und seine KollegInnen kamen zu der Annahme, dass es sechs Basisemotionen gibt, die universell vorkommen und damit angeboren sind, aber auch kulturelle Darbietungsregeln das Zeigen und Erkennen von diesen Emotionen beeinflussen. Die Emotionserkennung wird in der Folge insbesondere von PsychologInnen und VerhaltensforscherInnen in vielfältigen wissenschaftlichen Kontexten zur Anwendung gelangen. Das von Ekman und KollegInnen entwickelte *Facial Action Coding System (FACS)* gilt bis heute als ein „objektives“ Instrument zur Gesichtsausdruckserkennung. Das Argument für diese scheinbare Objektivität ist, dass das Verfahren zwischen Beschreibung und Interpretation trennt. Die Bedeutung eines Gesichtsausdrucks wird erst in der Interpretation von einem Experten oder einer Expertin zugewiesen, die auch die Kontextbedingungen der Situation kennen und einschätzen können. Allerdings ist fraglich, ob nicht auch schon in der „reinen“ Beschreibung Interpretationen vorgenommen werden.

Die neue Sehnsucht, um insbesondere Emotionen im Gesicht erkennen zu können, basiert vor allem auf der zunehmenden Verflechtung von Menschen und Maschinen in sozio-technischen Konstellationen. Maschinen sollen besser auf den Menschen eingehen können, indem sie sich wie ein Mensch verhalten. Deshalb besteht die Annahme, dass Maschinen unter anderem menschliche Gesichtsausdrücke erkennen sollten, weil der nonverbalen Kommunikation eine außerordentliche Rolle in der menschlichen Kommunikation zugestanden wird. Zwischen Mensch und Maschine sollen einerseits keine Eingabegeräte wie Computermouse oder Druckknöpfe mehr stehen, andererseits soll auch die menschliche

Meinung, vermittelt über die Sprache nicht mehr zwischen ihnen stehen, wenn es etwa darum geht, den eigenen Gefühlszustand zu beschreiben. Der direkte Weg der nonverbalen Kommunikation wird bevorzugt. Die Annahme besteht, dass aufgrund der erkennbaren und detektierbaren Zeichen in einem Gesicht, Gesichtsausdrücke und Emotionen erkannt werden können. Das „Innere“ eines Menschen kann einfach nur am „Äußeren“ abgelesen werden. Dass dazwischen aber eine Unmenge an Festlegungen, Standardisierungen, umstrittenen Tatsachen, Transformationen und Reduktionen liegen, bleibt den AnwenderInnen aber verborgen. Damit einher geht, dass wir es letztendlich mit einer ganz anderen Realität zu tun haben, die auf andere Referenten verweist, als auf die alltägliche Lebenswirklichkeit. Die Sehnsucht, das Innere des Menschen in seinem Äußeren erkennen zu können bleibt deshalb eine Sehnsucht.

Die Sehnsucht, das Innere des Menschen in seinem Äußeren zu erkennen, wird dennoch eingelöst. Es passiert ständig, jeden Tag, in vielfacher Weise. Ob Charakter oder Gefühlsleben, wir schließen andauernd physiognomisch vom Äußeren auf das Innere. Dieses Schließen vollzieht sich aber nicht nach eindeutigen Mustern, Regeln oder Gesetzen, sondern ist individuell und situationsabhängig sehr unterschiedlich. Wenn die FernsehmoderatorIn einem Sportler oder einer Sportlerin einen angestregten Gesichtsausdruck zuweist, können wir das glauben oder nicht. Wenn eine mutmaßliche angeklagte Mörderin als Engels Gesicht bezeichnet wird, können wir das als plausibel erachten oder es zurückweisen. Was passiert aber, wenn eine scheinbar neutrale Technologie eine Aussage über einen Gesichtsausdruck trifft? Können wir dieser Aussage nicht glauben?

Literatur

Aerni, Fritz (2003): Lehrbuch der Menschenkenntnis: Einführung in die Huter'sche Psychophysiognomik und Kalligraphie. Zürich / Waldshut-Tiengen: Carl-Huter-Verlag.

Alpers, W. Georg, Mühlberger, Andreas & Pauli, Paul (2009): Psychophysiologie der Emotionen. In: Brandstätter, Veronika & Otto, Jürgen H. (Hrsg.): Handbuch der Allgemeinen Psychologie - Motivation und Emotion: Handbuch der Psychologie Band 11: BD 11. Wien: Hogrefe Verlag. S. 412 – 421.

Arya, Ali et al. (2009): Perceptually Valid Facial Expressions for Character-Based Applications. In: International Journal of Computer Games Technology. Volume 2009, Article ID 462315, 13 pages.

Averill, J.R., Catlin, G., & Chon, K.K. (1990): The rules of hope. New York: Springer.

Bashyal, Shishir & Venayagamoorthy, Ganesh K. (2008): Recognition of facial expressions using Gabor wavelets and learning vector quantization. In: Engineering Applications of Artificial Intelligence. Volume 21, Issue 7, October 2008, Pages 1056-1064

Beer, U. (1964): Vorwort des Herausgebers. In: Darwin, Charles: Der Ausdruck der Gefühle bei Mensch und Tier. Düsseldorf: Rau.

Bell, Daniel (1975): Die nachindustrielle Gesellschaft. Frankfurt/ New York: Campus.

Berger, Peter L. & Luckmann, Thomas (2007 [1977]): Die gesellschaftliche Konstruktion der Wirklichkeit. Frankfurt am Main: Fischer Taschenbuch Verlag

Bijker, Wiebe (1995): Of Bicycles, Bakelites, and Bulbs. Toward a Theory of Sociotechnical Change. The MIT Press.

Borck, Cornelius (2001): Die Unhintergebarkeit des Bildschirms. In: Heintz & Huber (Hrsg.): Mit dem Auge denken. Springer. S. 383 – 394.

Bottenberg, Ernst Heinrich & Daßler, Henning (2002): Einführung in die Emotionspsychologie. Regensburg: Roderer.

Bowker, Geoffrey C & Leigh Star, Susan (2000): *Sorting things out. Classification and its consequences*. Cambridge, London: The MIT Press,

Brednow, Walter (1969): *Von Lavater zu Darwin*. Berlin: Akademie Verlag.

Burri, Regula V. & Dumit, Josph (2008): *Social Studies of Scientific Imaging and Visualization*. In: Hackett et al. (ed.): *The Handbook of Science and Technology Studies*. Third Edition. MIT. p. 297 – 317.

Buser, Remo (1973): *Ausdruckspsychologie. Problemgeschichte, Methodik und Systematik der Ausdruckswissenschaft*. München & Basel: Ernst Reinhardt Verlag.

Callon, Michel (2006): *Techno-ökonomische Netzwerke und Irreversibilität*. In: Belliger, Andréa & Krieger, David J. (Hg.) (2006): *ANThology. Ein einführendes Handbuch zur Akteur-Netzwerk-Theorie*. Bielefeld: transcript. S.309-342.

Cerezo, Eva et al. (2007): *Real-Time Facial Expression Recognition for Natural Interaction*. In: J. Martí et al. (Eds.): *IbPRIA 2007, Part II, LNCS 4478*, pp. 40–47, 2007.

Chen , Hsiuao-Ying et al. (2008): *Hybrid-boost learning for multi-pose face detection and facial expression recognition*. In: *Pattern Recognition* 41 (2008) 1173 – 1185.

Cheon, Yeongjae & Kim, Daijin (2009): *Natural facial expression recognition using differential – AAM and manifold learning*. In: *Pattern Recognition* 42(2009)1340-1350.

Chibelushi, Claude C., & Bourel, Fabrice (2002): *Facial Expression Recognition: A Brief Tutorial Overview*. In: *CVonline: On-Line Compendium of Computer Vision*. R. Fisher (ed). [Online verfügbar am 12.08.2009 unter: http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/CHIBELUSHI1/CCC_FB_FacExprRecCVonline.pdf]

Christians , Heiko (2006): *Sehnsüchte der Physiognomik*. In: Böckelmann, Frank & Seitter, Walter: *Tumult Band 31. Gesichtermoden*. Berlin: Berliner Buchdienst. S. 4 – 12.

Cohen, Ira et al. (2002): *Facial Expression Recognition from Video Sequences..* In: *Multimedia and Expo, 2002. ICME '02. Proceedings. 2002 IEEE International Conference* Volume: 2, pp. 121- 124 vol.

Cole, Simon A. (2001): *Suspect identities. A history of fingerprinting and criminal identification*. Cambridge, London: Harvard University Press.

Collins, Harry M. (1990): *Artificial Experts. Social Knowledge and Intelligent Machines*. Cambridge: The MIT Press.

Darwin, Charles (1948 [1872]): *The Expression of the Emotions in Man and Animals*. London:Watts.

Daston, Lorraine (2001): *Wunder, Beweise und Tatsachen. Zur Geschichte der Rationalität*. Frankfurt am Main: Fischer Taschenbuch Verlag.

Daston, Lorraine & Gallison, Peter (1992): *The Image of Objectivity*. In: *Representations* 40.

Daston, Lorraine & Gallison, Peter (2007): *Objektivität*. Frankfurt am Main: Suhrkamp.

Degele, Nina (2002): *Einführung in die Techniksoziologie*. München: Wilhelm Fink Verlag.

Deng, Hong-Bo et al. (2005): *A New Facial Expression Recognition Method Based on Local Gabor Filter Bank and PCA plus LDA*. In: *International Journal of Information Technology* Vol. 11 No. 11 2005

Denzin, Norman (1990): *Writing the interpretative postmodern ethnography: Review essay of The phyton killer by Vinigi L. Grottanelli and In sorcery`s shadow, by Paul Stoller and Cheryl Olkes*. *Journal of Contemporary Ethnography*, 19, 231-236.

Dierkes, Meinolf & Hoffmann, Ute (eds.) (1992): *New Technology at the Outset. Social Forces in the Shaping of Technological Innovations*. Boulder, COL: Westview Press.

Dollhausen, Karin (1997): *Technik – Konstruktionen. Neue Technologien als soziologisches Theorieproblem*. Pfaffenweiler: Centaurus-Verlags Gesellschaft.

Dornaika, Fadi (2008): *Simultaneous Facial Action Tracking and Expression Recognition in the Presence of Head Motion*. In: *International Journal of Computer Vision* (2008) 76: 257–281.

Fasching, Gerhard (2007): *Objektive Illusionen. Ein Essay über das Wesen der naturwissenschaftlichen Wirklichkeit*. Wien: Dr. G. Fasching (Eigenverlag).

Fasel, B. & Luettin, Juergen (2003): *Automatic facial expression analysis: a survey*. In: *Pattern Recognition* 36 (2003) 259 – 275

Ekman, Paul & Friesen, Wallace V. (1971): *Constants across cultures in the face and emotion*. *Journal of Personality and Social Psychology*, 17, p. 124 – 129.

Ekman, Paul (1999): Basic Emotions. In: T.Dalgleish and M.Power (Eds.): Handbook of Cognition and Emotion. Sussex: John Wiley & Sons Ltd.

Ekman, Paul, Friesen, Wallace V., & Hager, Joseph C. (2002): The Facial Action Coding System. Second Edition. London: Weidenfeld & Nicolson (world).

Ekman, Paul (2003): Darwin, Deception, and Facial Expressions. In: Ann. N.Y. Acad. Sci. 1000: 205 – 221 (2003). New York: Academy of Sciences.

Ekman Interview (2004): Conversations with History; Institute of International Studies, UC Berkeley: Face to Face: The Science of Reading Faces. Interview am 14.01.2004 geführt von Harry Kreissler. [Online verfügbar am 08.05.2009 unter: <http://globetrotter.berkeley.edu/people4/Ekman/ekman-con0.html>]

Ekman, Paul (2006): How to Spot a Terrorist on the Fly. Washington Post, October 29, 2006, B03. [Online verfügbar am 10.05.2009 unter: <http://www.washingtonpost.com/wp-dyn/content/article/2006/10/27/AR2006102701478.html>]

Ekman, Paul (2007): Emotions Revealed. Recognizing Faces and Feelings to Improve Communication and Emotional Life. New York: Owl.

Engel, Johann Jakob (1785/1786): Ideen zu einer Mimik, 2 Theile. Berlin: Auf Kosten des Verfassers und in Commission bey August Mylius.

Essa, Irfan A. (1999): Computers Seeing People. In: AI Magazine Volume 20 Number 2

Euler, Harald (2009): Evolutionäre Psychologie. In: Brandstätter, Veronika & Otto, Jürgen H. (Hrsg.): Handbuch der Allgemeinen Psychologie - Motivation und Emotion: Handbuch der Psychologie Band 11: BD 11. Wien: Hogrefe Verlag. S. 405 – 411.

Evans, Robert & Collins, Harry (2008): Expertise: from attribute to attribution and back again?. In: E. J. Hackett, O. Amsterdamska, L. M. and J. Wajcman (eds.): The Handbook of Science and Technology Studies. The MIT Press

Faktum Lexikon Institut (2000): Universal Lexikon. Gütersloh/München: Bertelsmann Lexikon Verlags GmbH

Felt, Ulrike / Nowotny, Helga / Taschwer, Klaus (1995): Wissenschaftsforschung: eine Einführung. Frankfurt a. Main, New York: Campus Verlag.

Forsythe, Diana (2001): *Studying Those who study us: An Anthropologist in the World of Artificial Intelligence*. Stanford: Stanford University Press.

Fridlund, Alan J. (1991a): Evolution and facial action in reflex, social motive, and paralanguage. *Biological Psychology*, 48, 150 – 170.

Fridlund, Alan J. (1991b): The sociality of solitary smiles: Effects of an implicit audience. *Journal of Personality and Social Psychology*, 60, 229 – 240.

Fridlund, Alan J. (1994): *Human facial expressions: An evolutionary view*. San Diego: Academic Press.

Garfinkel, Harold (1967): *Studies in Ethnomethodology*. Englewood Cliffs, New York: Prentice Hall.

Giddens, Anthony (2001): *Sociology*. 4th Edition. Oxford: Blackwell.

Gieryn, Thomas F. (1999): *Cultural boundaries of science*. Chicago: University of Chicago Press.

Gladwell, Malcolm (2002): *The Naked Face*. Can you read people's thoughts just by looking at them? *The New Yorker*, August 5, 2002.

Gomathi, V. et al. (2009): Human Facial Expression Recognition using MANFIS Model. In: *World Academy of Science, Engineering and Technology* 50 2009.

Gould, Stephen J. (1988 [1981]): *Der falsch vermessene Mensch*. Frankfurt a. Main: Suhrkamp.

Gray, Richard T. (1994): Aufklärung und Anti- Aufklärung: Wissenschaftlichkeit und Zeichenbegriff in Lavaters Physiognomik. In: Pestalozzi, Karl & Weigelt, Horst: *Das Antlitz Gottes im Antlitz des Menschen*. Göttingen: Vandenhoeck & Ruprecht. S.166 – 178.

Heintz, Bettina (2000): *Die Innenwelt der Mathematik. Zur Kultur und Praxis einer beweisenden Disziplin*. Wien: Springer.

Heintz, Bettina (2007): Zahlen, Wissen, Objektivität. Wissenschaftssoziologische Perspektiven. In: Mennicken, Andrea & Vollmer, Hendrik (Hrsg.): *Zahlenwerk. Kalkulation, Organisation und Gesellschaft*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Helle, Hors Jürgen (1977): *Verstehende Soziologie und Theorie der Symbolischen Interaktion*. Stuttgart: Teubner.

Herders Conversations-Lexikon (1856): Band 4. Freiburg im Breisgau. [Online verfügbar am 24.04.2009 unter: <http://www.zeno.org/Herder-1854/A/Pathognomik>].

Heßler, Martina (2006): Einleitung. Annäherungen an Wissenschaftsbilder. In: Heßler(Hrsg.): Konstruierte Sichtbarkeiten. Wissenschaft- und Technikbilder seit der Frühen euzeit. Wilhelm Fink. S.11-37.

Hitzler, Ronald & Honer, Anne (1989): Vom Alltag der Forschung. Bemerkungen zu Knorr Cetinas wissenschaftssoziologischem Ansatz. In: Österreichische Zeitschrift für Soziologie (ÖZS), 14. Jg., 5 H. 4/1989, S. 26-33

Von Humboldt, Wilhelm (1853): Briefe an eine Freundin. 28.4.1833. Leipzig.

Ioannou, Spiros et al. (2007): Robust Feature Detection for Facial Expression Recognition. In: EURASIP Journal on Image and Video Processing, Volume 2007, Article ID 29081, 22 pages.

Jasanoff, Sheila (2004): States of knowledge. The co-production of science and social order. London & New York: Routledge

Jaton, Anne-Marie (1988): Johann Caspar Lavater. Philosoph – Gottesmann. Schöpfer der Physiognomik. Eine Bildbiographie. Zürich: Schweizer Verlagshaus.

Jokisch, Rodrigo (Hrsg.) (1982): Techniksoziologie. Frankfurt am Main: Suhrkamp.

Kaiser, Susanne & Wehrle, Thomas (2009): Methoden der Mimikanalyse und –synthese. In: Brandstätter, Veronika & Otto, Jürgen H. (Hrsg.): Handbuch der Allgemeinen Psychologie - Motivation und Emotion: Handbuch der Psychologie Band 11: BD 11. Wien: Hogrefe Verlag. S. 521 - 531.

Kemp, Martin (1997): Seeing and Picturing: Visual Representation in Twentieth-Century Science. In: Krige, John & Pestre, Dominique (eds.): Science in the Twentieth Century. Amsterdam.

Khan, Mohammad Ibrahim & Bhuiyan, Md. Al-Amin (2009): Facial Expression Recognition for Human-Robot Interface. In: IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.4, April 2009

Knorr, Karin (1980): Die Fabrikation von Wissen. In: Stehr, Nico & Meja, Volker (Hrsg.): Wissenssoziologie (Sonderheft 22 der ‚KZfSS‘). Opladen: Westdeutscher Verlag.

Knorr Cetina, Karin (1983): New Developments in Science Studies: The Ethnographie Challenge. In: Canadian Journal of Sociology 8, H. 2/1983

Knorr Cetina, Karin (1984): Die Fabrikation von Erkenntnis. Frankfurt a. Main: Suhrkamp.

Kobayashi, H. & Hara, F. (1991). The recognition of basic facial expressions by neural network. In: Proceedings of International Joint Conference on Neural Networks, 460- 466.

Krohn, Wolfgang (1989): Die Verschiedenheit der Technik und die Einheit der Techniksoziologie. In: Weingart, Peter (Hrsg.): Technik als sozialer Prozeß. Frankfurt a. Main: Suhrkamp. S.15-43.

van Kuilenburg, Hans et al. (2005): A Model Based Method for Automatic Facial Expression Recognition. In: Machine Learning: ECML 2005, 16th European Conference on Machine Learning, Porto, Portugal, October 3-7, 2005. Proceedings. S. 194 – 205.

Latour, Bruno & Woolgar, Steve (1979): Laboratory Life: the Social Construction of Scientific Facts. Sage: Los Angeles.

Latour, Bruno (2002): Die Hoffnung der Pandora. Frankfurt a. Main: Suhrkamp.

Latour, Bruno (2006): Technik ist stabilisierte Gesellschaft. In: Belliger, Andréa & Krieger, David J. (Hg.) (2006): ANThology. Ein einführendes Handbuch zur Akteur-Netzwerk-Theorie. Bielefeld: transcript. S.369 - 398

Latour, Bruno (2007): Eine neue Soziologie für eine neue Gesellschaft. Frankfurt a. Main: Suhrkamp.

Lavater, Johann C. (1984 [1775]): Physiognomische Fragmente. Stuttgart: Reclam.

Lavater, Johann C. (1999 [1772]): Von der Physiognomik. Und Hundert physiognomische Regeln. Eds.: Karl Riha and Carsten Zelle. Frankfurt: Insel.

Littlewort, Gwen C. et al. (2009): Automatic coding of facial expressions displayed during posed and genuine pain. In: Image and Vision Computing xxx (2009) xxx–xxx (in press).

Mase, K. & Pentland, A. (1991): Recognition of facial expression from optical flow. In: IEICE Trans. E 74 (10) (1991) 3474-3483.

Mayer, Katja (2001): Eine Technologie zwischen Kommerz und Kontrolle. Die soziale Konstruktion der Sprachtechnologie. Universität Wien: Unveröffentlichte Diplomarbeit.

- McNeill, Daniel (2001): *Das Gesicht. Eine Kulturgeschichte*. Wien: Kremayr & Scheriau
- Mead, G. Herbert (1968): *Geist, Identität und Gesellschaft*. Frankfurt a. Main: Suhrkamp.
- Mehrabian, Albert (1968): Communication without words. *Psychology Today*, vol. 2, no. 4, pp.53-56
- Merten, Jörg (2003): *Einführung in die Emotionspsychologie*. Stuttgart: Kohlhammer.
- Meyers Großes Konversations-Lexikon (1908): Band 15. Leipzig. [Online verfügbar am 24.04.2009 unter: <http://www.zeno.org/Meyers-1905/A/Pathogn%C5%8Dmik>]
- Miller, Peter (2001): Governing by Numbers: Why Calculative Practices Matter. In: *Social Research* 28, S. 379 - 390
- Oudshoorn, Nelly, & Pinch, Trevor (2008): User-Technology relationships: Some Recent Developments. In: E. J. Hackett, O. Amsterdamska, L. M. and J. Wajcman (eds.): *The Handbook of Science and Technology Studies*. The MIT Press
- Pantic, Maja & Rothkrantz, Leon J.M. (2000): Automatic analysis of facial expressions: the state of the art. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 12, December 2000
- Person, Jutta (2005): *Der pathographische Blick: Physiognomik, Atavismustheorien und Kulturkritik 1870-1930*. Königshausen & Neumann.
- Pestalozzi, Karl (1988): Physiognomische Methodik. In: Finck, Adrien (Hrg.): *Germanistik aus interkultureller Perspektive: en hommage à Gonthier-Louis Fink*. Strasbourg: Inst. d'Etudes Allemandes. S. 137 – 153.
- Pestalozzi, Karl & Weigelt, Horst (1994): *Das Antlitz Gottes im Antlitz des Menschen*. Göttingen: Vandenhoeck & Ruprecht.
- Picard, Rosalind W. (1997): Does HAL Cry Digital Tears? Emotion and Computers. In: Stork, David G. (ed.): *HAL's Legacy: 2001's Computer as Dream and Reality*. Cambridge, London: The MIT Press.
- Piderit, Theodor (1867): *Wissenschaftliches System der Mimik und Physiognomik*. Detmold: Klingenberg'sche Buchhandlung.

Pinkl, Petra (2009): Alter (ver)messen? Gesellschaftspolitische Anwendungszusammenhänge wissenschaftlicher Vermessungstechniken zur chronologischen Lebensalterbestimmung im österreichischen Kontext. Universität Wien: Dissertation.

Polanyi, Michael (1985): Implizites Wissen. Frankfurt a. Main: Suhrkamp.

Porter, Theodore (1995): Trust in Numbers: The Pursuit of Objectivity in Science and Public Life. Princeton: Princeton University Press

Rammert, Werner (1983): Soziale Dynamik der technischen Entwicklung. Theoretisch-analytische Überlegungen zu einer Soziologie der Technik am Beispiel der „science-based industry“. Opladen: Westdeutscher Verlag.

Rammert, Werner (1993): Technik aus soziologischer Perspektive. Forschungsstand – Theorieansätze – Fallbeispiele. Ein Überblick: Opladen: Westdeutscher Verlag.

Rammert, Werner (2007): Technik - Handeln - Wissen. Zu einer pragmatistischen Technik- und Sozialtheorie. Wiesbaden: VS-Verlag für Sozialwissenschaften.

Richter, Rudolf (2001): Soziologische Paradigmen. Eine Einführung in klassische und moderne Konzepte. WUV.

Richter Rudolf (2002): Verstehende Soziologie. Wien: Facultas.

Ritzer, George (1995): Die McDonaldisierung der Gesellschaft. Frankfurt am Main: Fischer.

Schmölders, Claudia (1997): Das Vorurteil im Leibe: eine Einführung in die Physiognomik. Berlin: Akademie Verlag.

Schnettler, Bernt & Pötzsch, Frederik S. (2007): Visuelles Wissen. Erscheint in: Rainer Schützeichel (Hrsg.): Handbuch Wissenssoziologie und Wissensforschung, Konstanz: UVK. S.472 – 484.

Schulz-Schaeffer, Ingo (2000a): Sozialtheorie der Technik. Frankfurt/Main: Campus

Schulz-Schaeffer, Ingo (2000b): Akteur-Netzwerk-Theorie. Zur Koevolution von Gesellschaft, Natur und Technik. In: Weyer, Johannes (Hrsg.): Soziale Netzwerke. Konzepte und Methoden der sozialwissenschaftlichen Netzwerkforschung. München & Wien: Oldenbourg Verlag. S. 187-209.

Sennett, Richard (2006): Der flexible Mensch. Berliner Taschenbuch Verlag.

Sismondo, Sergio (2008): Science and Technology Studies and an Engaged Program. In: Hackett, Edward, Amsterdamska, Olga, Lynch, Michael and Wajcman, Judy (eds): The Handbook of Science and Technology Studies. Third Edition. Cambridge & London: MIT. p. 13 – 32.

Steinke, Ines (2005): Gütekriterien qualitativer Forschung. In: Flick, Uwe, von Kardorff, Ernst & Steinke, Ines (Hrsg.): Qualitative Forschung. Ein Handbuch: Reinbek bei Hamburg: Rowohlt.

Suchman, Lucy (1987): Plans and Situated Actions. The problem of human-machine communication. Cambridge University Press.

Suchman, Lucy (2007): Human-Machine Reconfigurations. Plans and Situated Actions. 2nd Edition. Cambridge University Press.

Suchman, Lucy (2008): Feminist STS and the Sciences of the Artificial. In: Hackett, Edward, Amsterdamska, Olga, Lynch, Michael and Wajcman, Judy (eds): The Handbook of Science and Technology Studies. Third Edition. Cambridge & London: MIT. p. 139 – 163.

Suwa, M. et al. (1978): A preliminary note on pattern recognition of human emotional expression. In: Proceedings of the Fourth International Joint Conference on Pattern Recognition, Kyoto, Japan, pp. 408-410.

Swoboda, Gudrun (2002): Lavaters Linienspiele. Techniken der Illustration und Verfahren graphischer Bildbearbeitung in einer physiognomischen Studiensammlung des 18. Jahrhunderts. Wien: Dissertation.

Taylor, Peter (1995): Building on Construction: An Exploration of Heterogeneous Constructionism, Using an Analogy from Psychology and a Sketch from Socioeconomic Modelling. Perspectives on Science 3: 66-98.

Volti, Rudi (1995): Society and Technological Change. Third edition. New York: St. Martins' s Press.

Voss, Julia (2008): Charles Darwin zur Einführung. Hamburg: Junius.

Weigelt, Horst (1991): Johann Kaspar Lavater: Leben, Werk und Wirkung. Göttingen: Vandenhoeck & Ruprecht.

Weingart, Peter (Hrsg.) (1989): Technik als sozialer Prozeß. Frankfurt am Main: Suhrkamp.

Williams, Carl E. & Stevens, Kenneth N. (1972): Emotions and Speech: Some Acoustical Correlates. In: The Journal of the Acoustical Society of America 52, 1238.

Wimmer, Matthias et al. (2008): Tailoring Model-based Techniques to Facial Expression Interpretation. *ACHI*, pp.303-308, First International Conference on Advances in Computer-Human Interaction, 2008

Wong et al. (2006): Using a sparse learning Relevance Vector Machine in Facial Expression Recognition. In: Proceedings of EUROMEDIA 2006, p.33-37.

Zeng, Zhihong et al. (2007): A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions. In: IEEE Transactions on Pattern Analysis and Machine Intelligence, accepted. [Online verfügbar am 12.08.2009 unter: <http://www.ifp.illinois.edu/~zhzeng/PAMI-AVemotionSurvey.pdf>]

Zhang, Zhengyou (1998): Feature-Based Facial Expression Recognition: Sensitivity Analysis and Experiments With a Multi-Layer Perceptron. In: International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI). [Online verfügbar am 12.08.2009 unter: <http://research.microsoft.com/en-us/um/people/zhang/Papers/IJPRAI.pdf>]

Zucker, Ursula (2007): Facial Expression Recognition – A Comparison Between Humans and Algorithms. Unveröffentlichtes Systementwicklungsprojekt an der Technischen Universität München. [Online verfügbar am 24.07.2009 unter http://www9-old.in.tum.de/people/wimmerm/lehre/sep_zucker/sep_zucker.pdf]

Internetquellen

http://www.bild.de/BILD/news/vermishtes/2009/01/15/das-engelsgesicht-und-der-sex-mord/der-anklaeger-steht-selber-wegen-amtsmissbrauch-vor_20gericht.html
[Online verfügbar am 07.07.09]

<http://de.wikipedia.org/wiki/Physiognomik> [Online verfügbar am 14.04.2009]

<http://de.wikipedia.org/wiki/Physiognomie> [Online verfügbar am 14.04.2009]

<http://www.britannica.com/EBchecked/topic/458823/physiognomy>
[Online verfügbar am 14.04.2009]

<http://www.paulekman.com/about-ekman/> [Online verfügbar am 18.05.2009]

http://www.uni-saarland.de/fak5/orga/Kurs/Seiten/facs/au5_test.htm

[Online verfügbar am 18.11.08]

<http://www.wissenschaft.de/wissenschaft/hintergrund/275969.html>

[Online verfügbar am 11.08.09]

<http://www.noldus.com/content/facereader> [Online verfügbar am 15.08.2009]

<http://www.icts.sbg.ac.at/media/pdf/pdf1362.pdf> [Online verfügbar am 15.08.2009]

http://www.wired.com/science/discoveries/news/2007/07/expression_research

[Online verfügbar am 15.08.2009]

<http://www.guardian.co.uk/science/2003/jul/10/scienceinterviews.science>

[Online verfügbar am 15.08.2009]

[http://www.focus.de/wissen/wissenschaft/psychologie/verbrechen/tid-](http://www.focus.de/wissen/wissenschaft/psychologie/verbrechen/tid-7131/gedankenforschung_aid_69979.html)

[7131/gedankenforschung_aid_69979.html](http://www.focus.de/wissen/wissenschaft/psychologie/verbrechen/tid-7131/gedankenforschung_aid_69979.html) [Online verfügbar am 15.08.2009]

Zusammenfassung / Abstract

Zusammenfassung auf Deutsch

Im Äußeren des Menschen, vor allem in seinem Gesicht scheinen sich eine Vielzahl an Informationen über sein Inneres, etwa seinen Charakter oder seine Gefühle zu verbergen. Auch die Wissenschaften haben seit dem 18. Jahrhundert das Verlangen, die Informationen und Codes des menschlichen Gesichts zu entschlüsseln. Lavater wollte z.B. die Physiognomik als Wissenschaft etablieren, als er versuchte den Charakter eines Menschen in seinem Äußeren zu erkennen.

Heute versuchen ComputerwissenschaftlerInnen weltweit, Maschinen und Robotern das Sehen und Erkennen von Objekten, Menschen und sogar von menschlichem Verhalten und Gefühlen beizubringen. Ein aktuelles Projekt auf diesem Weg zu menschlichen Maschinen sind Arbeiten, die sich im Bereich automatischer Gesichtsausdruckserkennung finden lassen. Auf Basis von Gesichtsausdruckserkennungsalgorithmen besteht die Möglichkeit, Emotionen einer Person automatisch und in Echtzeit zu erkennen. In dieser Masterarbeit nähere ich mich vorsichtig der großen Frage an, wie unseren Maschinen und Robotern eigentlich das Sehen und Erkennen von Objekten, Menschen, ihrem Verhalten und ihren Gefühlen beigebracht wird. Ich sehe meine Masterarbeit als eine explorative Studie an und habe sie als eine kleine Geschichte der Sehnsucht, das Innere des Menschen in seinem Äußeren zu erkennen, mit besonderem Fokus auf gegenwärtigen Technologien zur automatischen Gesichtsausdruckserkennung konzeptualisiert.

Bei den Technologien zur automatischen Gesichtsausdruckserkennung besteht vielfach die Annahme, dass Gesichtsausdrücke und Emotionen aufgrund der einfach detektierbaren Zeichen in einem Gesicht erkannt werden können, das Innere eines Menschen kann einfach nur am Äußeren abgelesen werden. Dass dazwischen aber eine Unmenge an Festlegungen, Standardisierungen, umstrittenen Tatsachen, Transformationen und Reduktionen liegen, bleibt dem Anwender verborgen. Sehen und Erkennen wird auf ein technisches und regelgeleitetes Sehen reduziert. Mit einer solchen Reduzierung geht einher, dass nur andeutungsweise die ganze Komplexität des menschlichen Sehens und Erkennens nachgebildet werden kann und dass dadurch nur Hinweise und Tendenzen darüber geschaffen werden können, was gesehen, erkannt und wahrgenommen wird.

Abstract English

In the human outside, especially in its face, there seems to be a wide range of information about the inside of human beings, such as the human nature or human feelings. For centuries science has had and still has the desire to be able to read the multiple information and codes the face provides. E.g., in 18th century Lavater tried to establish physiognomy as a scientific discipline. In physiognomy the human character should have been recognized in the outlines of the human face.

Nowadays computer scientists around the world on the way to engineer humanlike machines have to teach machines and robots how to see and how to recognize, not only objects but also humans, human interactions and feelings. Quite a new field of research in this project are technologies of facial expression recognition. On the basis of facial expressions algorithms there is the possibility to detect emotions of an individual automatically and in real time. In this Master's thesis I want to carefully approach the big question of how our machines and robots are being taught to see and recognize objects as well as humans, their movements, interactions and feelings. I understand my Master's thesis as an explorative study and conceptualized it as a story respectively a little history that tells about the desire of mankind and especially of science to see human inside through outside with the current development of facial expression recognition at the temporary end.

There is the assumption that facial expression recognition algorithms are able to recognize facial expressions and emotions simply by detecting signs and movements in the human face. In this view the inside of a person can be seen in the outside directly. But on this way there are to find many standardisations, definitions, transformations and reductions. These are forgotten easily and are inaccessible by users. Seeing and recognizing is reduced to technical and rule-based seeing. Because of this comprehensive reduction, the complexity of human seeing and recognizing can be simulated only in outlines. The consequence is that only a rough trend towards human seeing and recognizing can be simulated.

Anhang

Verzeichnis der geführten Interviews

Interview 0: Kunsthistorikerin, KHM Wien, November 2008

Interview 1: Computerwissenschaftler, TU Wien, März 2009

Interview 2: Verhaltensforscherin, Universität Wien, März 2009

Interview 3: Zwei Computerwissenschaftler (I3a & I3b), TU München, Mai 2009

Interview 4: Computerwissenschaftler, Fraunhofer Institut für Integrierte Schaltungen, Erlangen, Mai 2009

Interviewleitfaden

Der Interviewleitfaden wurde von mir auf die jeweilige Interviewsituation und den professionellen Hintergrund der befragten Personen angepasst, insofern wurden nicht alle in der Folge aufgelisteten Fragen auch an alle interviewten Personen gestellt. Einige Fragen wurden auch umformuliert oder an einer anderen Stelle formuliert.

Einleitende Fragen

Ich würde Sie zu Beginn einfach bitten, sich ganz kurz einmal hinsichtlich ihrer Tätigkeit hier am Institut kurz vorzustellen. Also was haben Sie studiert und welche Profession haben Sie sozusagen jetzt, in welche akademischen Bereiche lassen sie sich genau einordnen?

Also vielleicht könnten Sie mir jetzt einfach am Anfang erzählen, wie sie denn überhaupt zu der Forschung (zur Gesichtsausdruckserkennung) gekommen sind, also einfach mal den Prozess beschreiben, wie es zu der Forschung kam und in welchem Rahmen das alles stattgefunden hat? Also ganz kurz gesagt: Wie sind Sie bei dem Thema der Gesichtsausdruckserkennung überhaupt gelandet?

Und welche Personen waren da außer Ihnen noch beteiligt, von wem kamen quasi Impulse oder gab es spezielle Interessensgruppen, intern oder extern, die das voran getrieben haben?

Wo wurde eigentlich sozusagen der Bedarf gesehen, dass man jetzt eine solche Technologie, (die den Gesichtsausdruck eines Menschen erkennen kann) haben muss oder eben, dass man jetzt daran

forschen könnte? Hat bei der ganzen Sache also der Bedarf oder auch der Nutzen eine Rolle gespielt?

Wie ist oder war dieser Forschungsstrang denn eigentlich hier am Institut eingebettet? Wie steht das im Verhältnis zu den anderen Forschungsschwerpunkten, die es hier gibt, wie hängt das damit zusammen oder wie unterscheidet sich das auch davon?

Was mich auch interessieren würde ist, wie denn sozusagen so ihr Arbeitsalltag aussieht. Wie kann man sich das also vorstellen, wie arbeiten sie? Mit welchen Geräten, mit welchen Leuten, wo findet das alles statt? Vielleicht könnten Sie einfach einmal einen ganz normalen Arbeitstag beschreiben, wenn sie sich jetzt gerade mit der Gesichtsausdruckserkennung beschäftigen!

Gab es da eigentlich oder vielleicht auch, gibt es da vielleicht immer noch in diesem Bereich irgendwelche Kooperationen mit externen Personen oder Einrichtungen, also Berater vielleicht, die herangezogen wurden, um bestimmte Dinge abklären zu können. Also vielleicht auch unter dem Stichwort Interdisziplinarität, gab es das in irgendeiner Form? (vielleicht auch informell?)

Und vielleicht noch ein letzte Frage zu diesem Komplex: Wer finanziert ihre Forschung eigentlich? War das die Industrie oder waren das andere Fördergeber, also gab es eine interne oder eine externe Finanzierung?

Ich würde gerne auf die konkrete Forschung eingehen: Was waren da eigentlich so die Herausforderungen, die Sie am Anfang Ihrer Arbeit, ihrer Forschung zu bewältigen hatten, also was musste vorab alles geklärt werden? (Gab es bestimmte Annahmen in Bezug auf die Mimik? Wo lag die Herausforderung überhaupt? Wo gab es Schwierigkeiten?)

Anwendungsbereiche

Als nächstes würd ich gerne auf die möglichen Anwendungsgebiete der Technologien zur Gesichtsausdruckserkennung eingehen, und zwar würde mich interessieren in welchen Bereichen die Gesichtsausdruckserkennung eigentlich momentan bereits angewendet wird? Gibt es da schon Bereiche, und wenn ja, wo liegen die derzeit?

Und wenn Sie jetzt einmal etwas weiter denken, was sind dann Ihre Vorstellungen darüber, wo könnte die automatische Gesichtsausdruckserkennung zukünftig angewandt, eingesetzt werden? Was sind da so die möglichen Anwendungsgebiete?

Und wenn Sie diese Art von Technologie jetzt vielleicht in einem breiteren Rahmen ihrer Tätigkeiten einordnen, was sind da so ihrer Meinung nach sozusagen die Leitbilder, in welche Richtung geht das, könnte das gehen, oder sollte das auch gehen?

Wie soll die Zukunft aussehen und was können Sie dazu beitragen?

Was würden sich daraus ihrer Meinung nach denn für Möglichkeiten oder vielleicht auch für Probleme ergeben?

Funktionsweise der Technologien

Jetzt würde ich gerne noch auf die Funktionsweise der Gesichtsausdruckserkennung eingehen.

Und zwar würde ich Sie jetzt einfach mal bitten, mir möglichst genau zu erklären, wie denn so etwas eigentlich funktioniert, also eine automatische Gesichtsausdruckserkennung, also wie muss man sich das als technischer Laie eigentlich alles vorstellen? Vielleicht könnten Sie das einfach einmal erklären, wie das klappt, dass jetzt ein Computer, eine Software oder wer oder was auch immer dann wirklich erkennen kann, wie jetzt ein Mensch grade fühlt?

Gibt es da bestimmte Modelle oder einen bestimmten Zugang zu Emotionen, von Mimik oder von Gefühlen, die Sie verwenden?

Was mich auch interessieren würde, was ist denn eigentlich der Unterschied oder wo liegt der Unterschied, wenn man jetzt mit Standbildern oder mit bewegten Bildern arbeitet? Und welche Rolle spielen Bilder überhaupt in dem Prozess, ich nehme ja mal an, dass es sich nur um digitale Bilder handelt! Was passiert mit den Bildern im Laufe des Prozesses, wie werden die bearbeitet, was wird mit denen gemacht?

An welchen Kriterien wird denn eigentlich festgemacht, welche Emotion wie stark zu erkennen ist, welche Parameter bestimmen quasi das Ausmaß der Emotion?

Ab wann kann man nach ihrer Kenntnis denn eigentlich sagen, dass es sich jetzt eindeutig um einen bestimmten Gesichtsausdruck handelt? Welche Zahlen gibt es da? Und was sagen die aus?

Ekman schreibt ja auch immer wieder, dass mit FACS die Gesichtsbewegungen nur beschrieben und klassifiziert werden, aber interpretiert müssen sie dann erst im Nachhinein mit weiteren Informationen werden. Kann das der Computer/die Software/der Algorithmus übernehmen?

Was passiert denn eigentlich, wenn ein Ausdruck nicht erkannt wird? Ich hab da schon öfters jetzt was von Erkennungsraten gelesen, wie hoch sind die denn bzw. wer sagt denn überhaupt, wann eine Emotion richtig erkannt worden ist?

Wenn wir jetzt noch mal auf mögliche Anwendungen schauen: Ab wann handeln denn eigentlich Computer, Maschinen? Gibt es da auch irgendwelche Regeln, also wenn die Emotion einmal gezeigt wurde oder erst wenn sie fünf Mal deutlich gezeigt wurde? Und was passiert dann bzw. könnte passieren?

Probleme & Risiken

In welcher Hinsicht müsste denn noch daran geforscht werden, also wo gibt es möglicherweise noch Schwächen oder Lücken, dass das dann alles eben perfekt funktionieren würde? Wo sind da quasi die

Problemzonen, einerseits auf einer technischen, dann auf einer konzeptuellen aber auch auf einer gesellschaftlichen Ebene?

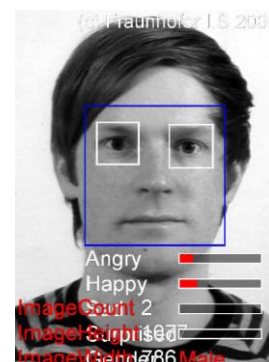
Glauben Sie, dass man in Zukunft Vorteile hat, wenn man seine Mimik bewusst beherrschen kann? Hat man dann Vorteile, wenn man so mit den Maschinen und Computern interagieren kann, dass sie auch erkennen, was man will?

Vielleicht könnten Sie mir jetzt noch als allerletztes noch ein paar Informationen darüber geben, wo es sonst noch Forschung in diesem Bereich gibt? Gibt es da irgendwie bestimmte Institute oder Unis, oder Personen, die da einen großen Namen haben, oder evtl. auch länderspezifisch? Wo auf der Welt wird ihres Wissens überall daran geforscht und wie unterscheiden sich die verschiedenen Ansätze?

Lebenslauf

Persönliche Daten

Christoph Musik
geb. am 30. September 1983
in Rosenheim (Deutschland)



Ausbildung

seit Okt 2007	Masterstudium der Soziologie mit Spezialisierung in Wissenschaftsforschung an der Universität Wien
Okt 2004 – Aug 2007	Bakkalaureatsstudium der Soziologie im geistes- und kulturwissenschaftlichem Zweig an der Universität Wien
Sep 1994 – Jul 2003	Besuch des musisch-humanistischen Ignaz-Günther- Gymnasiums in Rosenheim / Deutschland
Sep 1990 – Jul 1994	Besuch der Grundschule Oberaudorf / Deutschland

Beruflich-wissenschaftliche und berufsbildende Tätigkeiten

seit Sep 2009	Projektmitarbeiter am Institut für Höhere Studien (IHS), Wien im Projekt „tripleB ID - Identifikation von Bedrohungsszenarien in Banken durch Bildanalyse“
Sep 2007 – Aug 2009	Stipendiat am Institut für Höhere Studien (IHS), Wien. Mitarbeit unter anderem bei folgenden Projekten: • Evaluierung der Implacementstiftung Pflegeberufe • Der Beitrag der Wissensinstitutionen zur Wertschöpfung der Region Zürich • Ursache von Studienabbrüchen an Universitäten • Analyse der Arbeitsmarktkarrieren von AbsolventInnen der FH Wien
Sep 2006 – Jan 2008	Tutor für Erstsemestrige am Institut für Soziologie an der Universität Wien
Okt 2006 – Mai 2007	Studentisches Forschungsprojekt zu „Armut und Migration in Guatemala“ mit Feldaufenthalt in Guatemala im Februar 2007

<i>Aug – Dez 2006</i>	Projektpraktikum an der Bayerischen Akademie für Naturschutz und Landschaftspflege (ANL): <i>Evaluation der Aus- und Fortbildung der Naturschutzwacht in Bayern</i>
<i>Mär 2005 – Jan 2007</i>	Shop Assistent im Verkaufsshop des Haus der Musik in Wien
<i>Sep 2003 – Juni 2004</i>	Zivildienst an der onkologischen Klinik Bad Trissl in Oberaudorf
<i>Okt 1998 – Nov 2001</i>	Praktikum beim Regionalfernsehen Rosenheim (RFR) im Rahmen des Schulprojekts „Schülerfernsehen“

Projektberichte und Publikationen

UNGER, Martin, WROBLEWSKI, Angela, LATCHEVA, Rossalina, HOFMANN, Julia, MUSIK, Christoph, ZAUSSINGER, Sarah (2009): Ursache von Studienabbrüchen an Universitäten. IHS Projektbericht, Wien. Im Auftrag des Bundesministeriums für Wissenschaft und Forschung BMWF

UNGER, Martin, ATES, Gülay, MUSIK, Christoph, GASIOR, Kathrin (2008): Wertschöpfung im weiteren Sinne. In: UNGER, Martin et al.: Der Beitrag der Hochschulen zur Wertschöpfung der Region Zürich. IHS Projektbericht, Wien

VOGTENHUBER, Stefan & MUSIK, Christoph (2008): 1. Österreichische Konferenz für Berufsbildungsforschung. Feedback. IHS Forschungsbericht, Wien.

MUSIK, Christoph (2007): Naturschutzwacht in Bayern – Aspekte einer sozialwissenschaftlichen Untersuchung. In: ANLiegen Natur H.1 (2007) – 31.Jg.. Bayerische Akademie für Naturschutz und Landschaftspflege, Laufen a.d.Salzach.