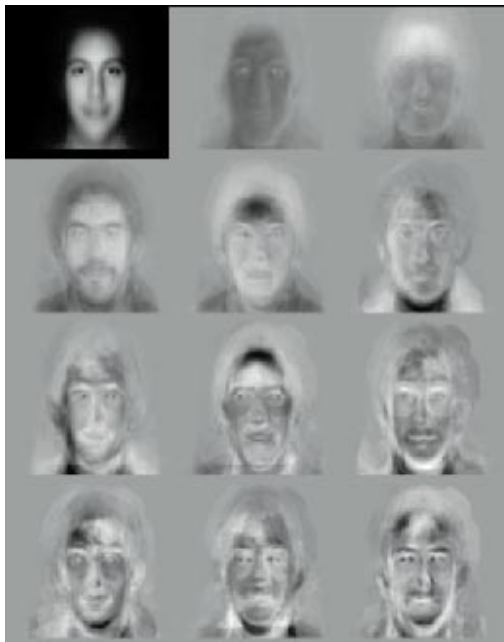


# Zimmer denkt mit

Schuhe lenken, Jacken denken, Zimmer fühlen: Am Massachusetts Institute of Technology versuchen Wissenschaftler, Alltagsgegenständen und der menschlichen Umgebung Intelligenz, Anteilnahme und Fürsorglichkeit beizubringen.



**Erkennungsdienst:** Ein „Smart Room“ identifiziert jeden Besucher, den er einmal erfasst hat, anhand spezifischer „Eigenfaces“

**Herr und Cyberdog:** Die Rückprojektionswand erschließt virtuelle Räume, in denen sich Mensch und computergenerierte Wesen treffen



**D**ie Dinge um uns herum – Räume, Tische, Autos, Schuhe, Mäntel – sind taub und stumm und dumm. Total unfähig, auf die Krone der Schöpfung angemessen einzugehen. IQ null und keine Lust, dazuzulernen, die ganze tumbe Materie. Daran wird kein vernünftiger Mensch zweifeln.

Außer Alex Pentland. Der angesehene Wissenschaftler (über 180 Veröffentlichungen) ist Spezialist für Künstliche Intelligenz, Mensch-Maschine-Interaktion und Computergrafik. Außerdem Chef einer hochkarätigen 50köpfigen Crew, die sich am Media Lab des Massachusetts Institute of Technology (MIT) mit „Percep-

tual Computing“ befasst. Mit Computern also, die etwas wahrnehmen können. Und die sollen dazu beitragen, aus Schuhen, Räumen und den anderen toten Gebrauchsgegenständen einfühlsame, hilfreiche und intelligente Begleiter des Menschen zu machen.

„Diese Objekte sind allesamt unfähig, unsere Bedürfnisse zu erkennen und darauf zu reagieren“, rügt Alex Pentland. „Wir haben deshalb Computersysteme entwickelt, die menschliche Aktionen verfolgen sowie Gesichter, Gesten und Mimik sowohl erkennen als auch interpretieren können.“

Das Ergebnis einer längeren Reihe von Experimenten sind „Smart Rooms“ und „Smart Clothes“. Ein „intelligenter Raum“, so Pentland, sei mit einem Butler zu vergleichen, der still und aufmerksam in einer Ecke stehe und immer darauf achte, wie und wann er helfend eingreifen dürfe.

## Mensch oder Materie – das ist die Frage

Der Aufwand, mit dem diese Art von Science-Fiction in die Realität umgesetzt wird, ist beeindruckend. Als erstes entwickelten die Wissenschaftler ein System namens „Pfinder“ („Person Finder“). Es bezieht seine optischen Informationen von Videokameras und hat zunächst einmal die Aufgabe, einen Menschen von der Umgebung und vom Hintergrund zu unterscheiden. Dazu bedient es sich statistischer Modelle, in denen Schatten und Farben erfasst sind.

Das klingt simpler, als es ist. Schließlich bewegen sich nicht nur die Personen, sondern manchmal auch die ihn umgebenden Gegenstände. Der Pfinder vergleicht also ständig ein Videobild mit dem nächsten, analysiert charakteristische Farbflächen („Blobs“) und interpretiert sie dann als Mensch oder Materie.

Für die Lokalisierung und die Äußerungen einer Person gleichermaßen wichtig ist die akustische Erfassung. Ein konventionelles körpernahes Mikrofon kommt nicht in Frage. Bleibt als Alternative ein Richtmikrofon, das exakt auf den Mund des Menschen ausgerichtet ist.

Wo sich der befindet, kann der Pfinder ziemlich genau feststellen. Die präzise Ortung im dreidimensionalen Raum übernehmen Mikrofonketten. Aus der Laufzeitdifferenz der Schallsignale ermittelt dann der PC den Sprecher und steuert das Richtmikrofon. Alle anderen Geräusche werden auf diese Weise weitgehend ausgeblendet. ○

Bis jetzt geht es noch um einigermaßen banale Technologien. Der schlaue Raum weiß nun, wo sich eine oder mehrere Personen befinden und ob sie etwas zu sagen haben.

Doch als nächstes stellt sich ihm die ungleich schwierigere Frage: „Wer ist das?“ Unserem sagt das ein Blick ins Gesicht des Gegenübers. Dem Computer auch. Die Videokamera liefert dem Rechner das Porträt. Der vergleicht es mit seiner gespeicherten Porträtgalerie, die pro Kopf ein Set ganz spezifischer Erscheinungsformen bereithält. Sie wurden mit dem mathematischen Verfahren der Eigenvektor-Analyse komponiert; Pentland nennt die nur entfernt humanen Anlitze deshalb „Eigenfaces“. Diese Personenerkennung funktioniert einwandfrei, behauptet er. Bei der Identifikation mehrerer hundert Personen in Theatern und Museen irrte sich der Computer nur in einem Prozent der Fälle.



Wenn ein Mensch erst einmal identifiziert ist, kann mit diesem Verfahren auch gleich der Gesichtsausdruck interpretiert werden. Das sei überaus hilfreich, so Pentland: „Ein Auto sollte eigentlich sehen, wenn der Fahrer müde wird, und ein Lernprogramm darauf Rücksicht nehmen, wenn der Student ratlos guckt.“ Der Rechner registriert also die Bewegung der Gesichtszüge und gleicht sie mit gespeicherten Mustern ab.

„Die Fähigkeit, Gestik, Gesicht und Sprechen einer Person zu erkennen, kann nur ein Anfang sein“, so Pentland. „Um Häuser oder Computer wirklich nützlich zu machen, müssen auch die Absichten eines Menschen zu identifizieren sein.“ Schwierig zu unterscheiden, ob sich einer bloß reckt oder auf etwas zeigt.

Vor ähnlichen Problemen stehen die Sprachforscher, wenn sie von bloßen Lauten über Phoneme und schließlich Phonemketten zu bedeutungstragenden



Wörtern gelangen wollen. Die Linguisten bedienen sich dazu des sogenannten Markov-Modells. Es dient dazu, die statistisch wahrscheinlichste Interpretation herauszufinden.

Pentland gibt zu, daß sich dieses Projekt noch im Anfangsstadium befindet. Das heißt, daß sich Gestik vorerst nur zum Beeinflussen virtueller Umgebungen eignet. In einem seiner Experimente übernimmt die Rolle des virtuellen Raumes eine überdimensionale Rückprojektionsleinwand, auf der sich der Akteur inmitten computergenerierten künstlichen Lebens wiederfindet.

**Gebärdensprache: Der Computer versucht, mittels einer Bewegungsanalyse auf die Absichten eines Menschen zu schließen**

konnte der Computer vierzig Gesten mit fast hundertprozentiger Genauigkeit und in Echtzeit analysieren.

Der erste smarte Raum entstand bereits 1991. Inzwischen existieren weitere in Japan, Frankreich, England und mehreren Orten der USA. Sie sind per ISDN miteinander verbunden, Wissenschaftler können sich so virtuelle Umgebungen teilen und darin experimentieren.

Das Potential, das in diesen eben entstehenden Technologien steckt, ist schier unerschöpflich. Als nächstes, so steht zu befürchten, kommt der Computer zum Anziehen. Bei dieser Vorstellung gerät Pentland vollends ins Schwärmen:

**Wenn der Schuh weiß, wohin er laufen soll**

„Bauen Sie eine Kamera in Ihre Brille, und die Gesichtserkennungssoftware identifiziert die Person, auf die Sie blicken, die Sprachausgabe flüstert Ihnen den Namen ins Ohr!“

Im Jackett eingebaute Mikrofone hören ständig mit, der PC in der Jackentasche liefert auf bestimmte Stichworte hin blitzartig zusätzliche Informationen an den Minibildschirm am unteren Brillenrand. Ein GPS-Sensor im Schuh stellt der Navigationssoftware alle Daten zur Verfügung, die nötig sind, um dem Träger des Schuhs den rechten Weg zum Ziel einzuflüstern oder einzublenden. Die Forscher denken derzeit an Biosensorsysteme, die Herzschlag und Muskelaktivitäten auswerten.

An Versuchskaninchen fehlt es Pentland nicht. Mehrere Angehörige seines Instituts haben sich mit den Prototypen smarter Kleidung schon so angefreundet, daß sie den sensorgespickten Dress samt Headmounted Display ständig tragen, auch wenn sie damit „eher wie Cyberpunks“ (Pentland) aussehen. Immerhin sind sie auf diese Weise drahtlos und ununterbrochen mit dem gesamten Forschungsnetz verbunden. So werden Surferträume wahr. **Hans Schmidt** □

**Versuchskaninchen: Am MIT werden „Smart Clothes“ auf ihre Alltagstauglichkeit getestet. Der Computer zum Anziehen soll völlig neue Kommunikationsformen ermöglichen.**

Viel Spaß hatten die Forscher mit einem virtuellen Hund, der recht spontan und gelegentlich ungehörlich auf die Gebärden und Befehle seines echten Herrchens reagiert.

Zu den ernster zu nehmenden Zielen der Gestikanalyse zählt die Erfassung der komplizierten Amerikanischen Zeichensprache (ASL, American Sign Language), mit der sich Taubstumme verständigen.

Pentland scheint es gelungen, mit Hilfe von Pfänder (Bewegungsanalyse) und Markov-Modell (Bedeutungszuweisung) die Basis für eine praktikable Übersetzung von der Gebärdensprache in die gesprochene Sprache (und umgekehrt) geschaffen zu haben. Im Experiment