

# Experten- und Agentensystemtechniken für intuitivere Benutzungsschnittstellen

*Ipke Wachsmuth*  
Technische Fakultät, Universität Bielefeld

## Kurzfassung

Aus der Perspektive benutzerfreundlicher Schnittstellen werden im Beitrag sog. Interface-Agenten thematisiert, die im Bereich der Mensch-Maschine-Interaktion derzeit zunehmend Aufmerksamkeit finden. In der Informatik versteht man unter "Agenten" eigenständige, adressierbare Computerprogramme, die Aktivitäten in ihrer Software-Umgebung beobachten und initiieren können und die mit anderen solchen Agenten kommunizieren können. Zusammen mit Expertensystem-Techniken läßt sich mit Agentensystemen sowohl Wissen über technische Einzelheiten wie auch Wissen über Benutzer ausnutzen, um Menschen von der mühsamen Bedienung komplexer Kommando-Schnittstellen zu entlasten. Der vorliegende Artikel greift grundsätzliche Aspekte dieser Techniken auf und beschreibt dann neuere Forschungsarbeiten in Bielefeld, wo seit 1993 hochinteraktive 3D-computergrafische Konstruktions- und Design-Umgebungen entwickelt werden. In jüngsten Arbeiten werden verstärkt auch multimodale Eingabemöglichkeiten einbezogen, um natürlichere, anthropomorphe Benutzungsschnittstellen für MultiMedia- und Virtual Reality-Systeme zu entwickeln.

## 1 Einleitung und Motivation

*Alltägliches Szenario für viele: Knopf MENU drücken, um das Bildschirm-Menü zu aktivieren, Knopf PR- drücken, um Timerprogramm zu aktivieren, Knopf DIS drücken, um Eingabetabelle zu aktivieren, Kanal eingeben: zwei Knopfdrücke, Datum eingeben: vier Knopfdrücke, Startzeit eingeben: vier Knopfdrücke, Endzeit eingeben: vier Knopfdrücke, Knopf MENU zweimal drücken — Vertippen, Korrektur, VPS und Aufzeichnungsgeschwindigkeit nicht eingerechnet, sind in dem einfachsten Fall bald zwanzig Knopfdrücke nötig, um den Videorecorder zu programmieren, im wahrsten Sinne eine Schnitt-Stelle zwischen Mensch und Maschine. Wieviel einfacher wäre: "Nimm morgen um zwanzig Uhr fünfzehn die Sendung im ersten Programm auf", damit ist alles gesagt. Unvergleichlich komplexer ist die Bedienung eines modernen CAD-Systems: auf jeder Bildschirmseite eine Randmenüleiste mit 20 Feldern, oben und unten noch ein paar Felder, zwischen drei oder vier solchen Bildschirmen kann hin- und hergeschaltet werden, jedes Menüfeld läßt weitere Unter- und Unter-unter-Menüs aufspringen, die mehrere tausend Funktionen aktivieren und kombinieren können, mit Hilfe der Maus sind feinmotorisch anspruchsvolle Manipulationen auf vier Teilbildschirmen vorzunehmen... es ist*

leicht vorstellbar, daß die Einarbeitungszeit in ein solches System ein halbes Jahr oder mehr erfordert, bevor es Nutzern im gewünschten Maße nützlich wird.

Technische Systeme beruhen heute in immer größerem Maße auf Funktionen integrierter Rechnerleistung; damit vervielfacht sich das Angebot an Leistungsmerkmalen, zugleich aber die Bedienungskomplexität, was potentielle Nutzer durchaus überfordern kann. Je mehr Leistungsmerkmale eine Anwendung bietet, desto unzureichender sind die Möglichkeiten klassischer WIMP (Windows, Icons, Menus & Pointers)-Schnittstellen für den unkomplizierten Umgang mit der Technik. Als Herausforderung stellt sich die Schaffung intuitiverer Schnittstellen, die robust und flexibel sind und Nutzer von technischen Details ihrer Hilfsmittel entlasten.

Die Leichtigkeit und Robustheit in der Kommunikation zwischen Menschen gab Anlaß, auch in der Wechselwirkung mit maschinellen Systemen menschengerechtere Modalitäten anzustreben. Hier sind natürlichsprachliche Schnittstellen zu nennen, die sich beim Zugriff auf Datenbanken und bei der Steuerung komplexer technischer Systeme zu bewähren beginnen. In Ergänzung dazu finden in jüngerer Zeit besonders im Multimedia-Bereich Gesten und Zeichensprachen als Mittel der Informationsübermittlung an maschinelle Systeme starkes Interesse, bis hin zu mehrmodalen Schnittstellen. Zum sich entwickelnden 'state of the art' gehören schließlich Unterstützungssysteme, die mit Hilfe intelligenter Techniken Mittler- und Assistenzaufgaben in der Mensch-Maschine-Kommunikation übernehmen.

Das zentrale Thema des vorliegenden Beitrags sind sog. Interface-Agenten, die in der Mensch-Maschine-Kommunikation ein vielversprechendes neues Paradigma bilden. In Verbindung mit Expertensystem-Techniken läßt sich mit solchen Schnittstellen-Systemen sowohl Wissen über technische Einzelheiten der Anwendung wie auch Wissen über Benutzerpräferenzen ausnutzen, um Menschen von der mühsamen Bedienung immer komplexerer Kommando-Schnittstellen zu entlasten. In dem folgenden Abschnitt werden grundsätzliche Aspekte derartiger intelligenten Techniken aufgegriffen. Mit dem Fokus auf Multimedia-Systemen werden im dritten Abschnitt Einsatzbeispiele aus jüngeren Forschungsarbeiten an der Universität Bielefeld vorgestellt. Wir schließen mit einem Ausblick auf zukünftige Entwicklungen.

## **2 Experten- und Agentensystemtechniken**

In zahlreichen Bereichen stellt sich heute die Aufgabe, menschengerechte Mensch-Maschine-Schnittstellen für alltäglichen Einsatz zur Verfügung zu stellen, die aus kontextbezogenen, selbst unscharfen Eingaben von Benutzern verwertbare Informationen beziehen können. Denken wir zurück an das Einleitungsbeispiel der Programmierung eines Videorecorders oder der Bedienung eines modernen CAD-Systems und stellen wir uns vor, daß ein "Experte" in der Benutzung solcher Systeme den sprachlich vorgetragenen Wunsch eines unerfahrenen Nutzers in geeignete Kommandofolgen umsetzen und damit die gewünschte Systemfunktion auslösen kann. Der im folgenden ausgeführte Grundgedanke ist es, durch die Verwertung derartigen "Expertenwissens" die Systemschnittstelle selbst intelligenter zu machen, das heißt, sie in die

Lage zu versetzen, zwischen Benutzer und technischem System zu vermitteln, und zwar dadurch, daß sie Anteile der Lösungsfindung eigenständig erbringen kann.

Hierfür lassen sich Techniken der Künstlichen Intelligenz und Expertensysteme gut einsetzen (GÖRZ, 1995). Das grundsätzliche Vorgehen dabei ist es, die Anwendungsdomäne geeignet zu *konzeptualisieren*, d.h. die in der Benutzerinteraktion beeinflussbaren Stellgrößen begrifflich zu fassen, und ihre Zusammenhänge mit der Änderung von Systemparametern beispielsweise in der Form von Wenn-Dann-Regeln zu beschreiben. Eine derartige Regel könnte für das obige Videorecorder-Beispiel ähnlich lauten wie folgt:

WENN Zeit="morgen" DANN setze Aufnahme datum=heutiges\_Datum + 1, etc.

Derartige Regeln werden einem Interpreter zur Verfügung gestellt, dessen Aufgabe es ist, die mehr oder weniger qualitativ geäußerten Benutzerwünsche in Folgen technischer Kommandos umzusetzen. Die Anwendung von Expertensystemtechniken für das Interfacing mit einem technischen System ist dabei einfacher als bei entscheidungsunterstützenden Systemen, da die systemseitigen Stellgrößen heute in der Regel über Mikroprozessortechnik angesteuert werden, d.h. bereits in digitalisierter, computergerechter Form vorliegen.

Im Kontext der Schnittstellen-Entwicklung werden für derartige Aufgaben zunehmend häufig *Agentensysteme* eingesetzt. Informatisch betrachtet sind unter "Agenten" aktive, adressierbare Computerprogramme zu verstehen, die als mehr oder weniger unabhängige Einheiten agieren, welche Aktivitäten in ihrer Umgebung beobachten und initiieren und mit anderen Agenten kommunizieren können (VON BECHTOLSHEIM, 1993). Dazu sind die einzelnen Agenten durch Kommunikationskanäle vernetzt, verfügen über eine interne Funktionalität, die eine bestimmte Aufgabe erbringen kann, und ferner über die Fähigkeit zu Kooperation bzw. Wettbewerb, mit der sich unterschiedliche Formen der Wechselwirkung bis hin zu Angebot und Nachfrage in sog. Vertragsnetzen modellieren lassen.

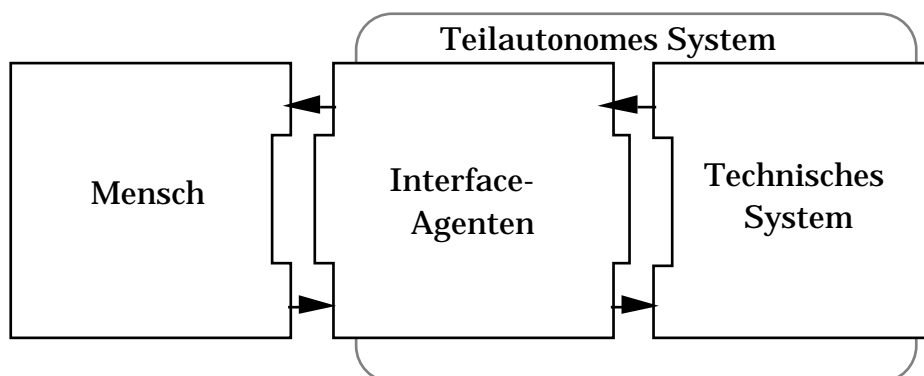


Abbildung 2-1 : Agenten als Bindeglied in der Mensch-Maschine-Kommunikation

Im Bereich der Schnittstellen-Entwicklung werden derartige Agenten-Techniken häufig unter dem Stichwort "Interface-Agenten" diskutiert (LAUREL, 1990), (SHNEIDERMAN & MAES, 1997). Interface-Agenten fungieren als Bindeglied in der Mensch-Maschine-Interaktion; sie vereinfachen die Handhabung komplexer Systeme insofern, als statt der direkten Bedienung über eine komplexe Kommandoschnittstelle oder Maus und Menüs eine *Kommunikation mit dem technischen System* erfolgt. Es ist üblich geworden, auch in dem Zusammenhang von Informationsübertragungen zwischen Mensch und maschinellen Systemen den Begriff der Kommunikation zu gebrauchen, was in dem Sinne gerechtfertigt ist, daß Bedeutungsinhalte durch Zeichenfolgen übermittelt werden, deren Bedeutung empfängerseitig zu rekonstruieren ist. Die Symbolverarbeitungsmöglichkeiten von KI-Systemen bilden dabei das Potential der Realisierung von Zeichenübermittlungen, die beispielsweise natürliche Sprache und Gesten einbeziehen. Technisches System und Interface-Agenten verbinden sich – wie in Abbildung 2-1 veranschaulicht – zu einem *teilautonomen System*, welches Anteile seiner Funktion unabhängig von direkter Steuerung durch Benutzer erbringen kann ("indirektes Management"). Hinzu können auch Techniken des maschinellen Lernens (WEISS, 1997) kommen, beispielsweise zur automatischen Benutzeranpassung von Schnittstellen. Der hier abstrakt beschriebene Ansatz ist Grundlage unserer Arbeiten über Multimedia-Benutzungsschnittstellen, die im Folgeabschnitt vorgestellt werden.

### **3 Intelligente Interfaces für Multimedia und Virtual Reality**

Ein Spitzenbedarf in der Multimedia-Forschung besteht in der Verbesserung der Interaktionsmöglichkeiten mit Multimedia-Systemen, welche bislang überwiegend auf die reine Präsentation von multimedialen Informationen ausgelegt sind. Die Mensch-Maschine-Schnittstelle soll sich intuitiver gestalten; der Mensch soll nicht nur in einseitiger Weise seine perzeptiven, sondern auch seine kommunikativen Fähigkeiten entfalten können. Eine neue Forschungsrichtung der Multimedia-Benutzungsschnittstellen – *Multi Media User Interfaces (MMUI)* – ist im Aufbruch; der 'state of the art' in der Interface-Technik schickt sich an, über Kommandoschnittstellen und die allgegenwärtigen Graphical User Interfaces (GUIs) hinauszugelangen (VAN DAM, 1997). Dazu müssen die traditionellen Eingabemöglichkeiten (Maus, Menüs, Kommandosprachen) um geeignete Mittel erweitert werden. Mit Hilfe intelligenter Funktionen soll es ermöglicht werden, gesprochene Sprache, textuelle und visuelle Eingaben zu verarbeiten und zur Manipulation von Text, Grafik, Bildern und Geräuschen bis hin zur Interaktion mit Virtual-Reality-Umgebungen (ASTHEIMER ET AL., 1994), (BARFIELD & FURNESS, 1995) zu mobilisieren.

Fokus neuerer Forschungsarbeiten im Multimedia-Bereich in Bielefeld sind hochinteraktive 3D-computergrafische Konstruktions- und Design-Umgebungen. Ein Leitmotiv unserer Arbeiten ist die verbesserte Nutzungsmöglichkeit von Multimedia-Systemen durch intelligente Schnittstellen, die verschiedene Eingabemodalitäten (Text, Sprache, Gestik) in einer multimedialen Präsentation umsetzen. Bislang hatte der Computer wenig oder gar kein Verständnis davon, was er dem Nutzer präsentiert. Hieraus erwachsen Aufgaben zur intelligenten Unterstützung der Systemfunktionen mit wissensbasierten Techniken, insbesondere semantisch verarbeitbaren

Datenrepräsentationen der synthetischen Darstellungen. Dazu verwenden wir Agentensysteme in Verbindung mit Expertensystemtechniken, durch die sich beispielsweise die semantischen Beschreibungen von Objekten während der Visualisierung fortschreitend geänderter Modelle anpassen lassen. In verschiedenen Forschungsprojekten wurden hierfür Beispielanwendungen grafisch-interaktiver Systeme im Bereich des Designs und der Konstruktion erarbeitet, die im folgenden kurz vorgestellt werden; weiterführende Literatur ist angegeben.

### 3.1 VIENA – Virtuelle Entwurfsumgebung und Agenten

Im VIENA-Projekt (WACHSMUTH & CAO, 1995) ist es im Anwendungsbeispiel Raumdesign gelungen, eine intelligente Mensch-Maschine-Kommunikation mit einer sprachlich gesteuerten computergrafischen Entwurfsumgebung in größerem Umfang zu verwirklichen. Anstelle der wenig intuitiven Manipulation eines CAD-Tools werden die Einrichtungsgegenstände eines virtuellen Innenraums, der in fotonaher Qualität dargestellt ist (siehe Abbildung 3-1), aufgrund verbaler Eingaben des Benutzers auf diverse Weise umkonfiguriert und farblich verändert. Die entwickelten Verfahren beziehen kognitive Faktoren der Raumwahrnehmung und implizite Erwartungen des Menschen bei der Manipulation der Umgebung ein; sie rekonstruieren etwa im Effekt physikalische Gesetze, um die gegenseitige Durchdringung massiver Objekte im virtuellen Raum zu verhindern. Während dieses Prozesses wird dem Benutzer als Reaktion auf seine Eingaben die Szene als berechnetes 3D-Bild unter Berücksichtigung seines momentanen Sichtfeldes gezeigt; Änderungen des Designs lassen sich unmittelbar beobachten und beurteilen. Der Benutzer kann vom System visualisierte Vorschläge durch Bezug auf die letzte Eingabe einfach korrigieren. Um größtmögliche Annäherung an die situationseingebettete Informationsverarbeitung des Menschen zu erreichen, wurde unter Einbindung eines Spracherkenners und eines Datenhandschuhs im ersten Ansatz eine multimodale Mensch-Maschine-Schnittstelle realisiert, die einfache Formen der gestenunterstützten Spracheingabe gestattet.

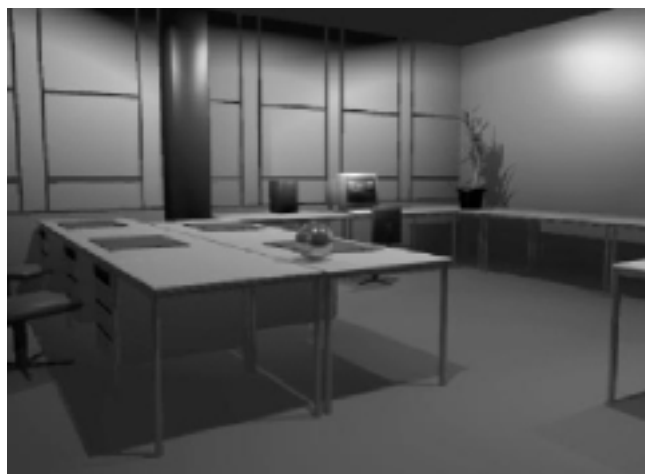
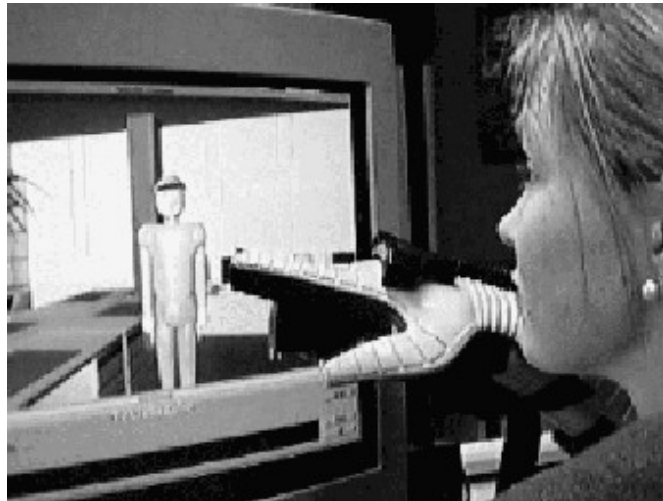


Abbildung 3-1 : Manipulierbarer virtueller Raum im VIENA-Projekt (Original farbig)

Die Benutzerkommunikation mit dem VIENA-System wird durch ein System wissensbasierter Agenten unterstützt, die kontextbezogen agieren. Sie nehmen die Instruktionen des Benutzers entgegen und kooperieren untereinander in dem Sinne, daß jeder einzelne Agent – wo nötig mit Unterstützung von weiteren Agenten – seine Teilaufgabe ausführt und sich damit an einem Kooperationsprozeß zur Berechnung einer koordinierten Gesamtlösung beteiligt. Dabei können die Interface-Agenten die aktuell sichtbare Szene durch die Inspektion der Szenenbeschreibung (bezogen auf das momentane Sichtfeld des Benutzers) als Informationsquelle heranziehen und so auch unterspezifizierte Eingaben verarbeiten. Durch die Kooperation mehrerer Agenten in der Eingabeschnittstelle werden die Modalitäten von Sprache und Gestik integriert, um dadurch komplementäre Eingabeinformation zu verwerten. Die realisierten Funktionen werden zum Teil durch adaptive Verfahren unterstützt, um Benutzer von technischen Details zu entlasten und eine kurzfristige Anpassung an unterschiedliche oder wechselnde Benutzerpräferenzen zu erreichen (LENZMANN & WACHSMUTH, 1996; 1997).



*Abbildung 3-2 : Kommunikation mit einem virtuellen Interface-Agenten*

Mit Hilfe eines in der virtuellen Umgebung platzierten menschenähnlichen Interface-Agenten (Abbildung 3-2) wurde schließlich die Möglichkeit geschaffen, daß Benutzer ihre Eingaben an ein personalisiertes Gegenüber richten wie auch auf die Sicht dieser Figur umschalten können (JÖRDING & WACHSMUTH, im Druck). Auf diese Weise lassen sich in der Kommunikation mit dem VIENA-System verschiedene räumliche Referenzsysteme ausnutzen.

### **3.2 CODY – Virtueller Konstrukteur**

Ein weiteres Beispiel einer wissensunterstützten Multimedia-Anwendung ist der Virtuelle Konstrukteur, ein mit Maus- und Spracheingaben instruierbares Montagesimulationssystem. Der Virtuelle Konstrukteur ermöglicht auf einer bildschirmpräsentierten virtuellen Werkbank die

interaktive Montage komplexer Aggregate aus einfachen Grundbausteinen. Durch eine Stereo-Brille kann der Betrachter die dreidimensionale Struktur der dargestellten Objekte erfassen und durch Anweisungen in natürlicher Sprache sofort interaktiv verändern. Die Spracheingaben können auf alles Sichtbare – Farben, Formen, Positionen, aber auch Namen von Teilen und Baugruppen – bezugnehmen; mit der Maus kann das gezeigte Objekt gedreht und verschoben werden, oder es können die Teile von Aggregaten bewegt werden. Die Aktionen des Systems, beispielsweise beim Zusammenstecken von Teilen, werden durch gesampelten Stereo-Sound hörbar gemacht, wodurch sich der Realitätseindruck verstärkt. Noch unmittelbarer ist die Erfahrung bei der Manipulation solcher virtuellen Objekte auf der *Responsive Workbench* (Abbildung 3-3); für Betrachter mit Stereobrille entfaltet sich das im Bild nur flach sichtbare Objekt dreidimensional über bzw. auf der tischartigen Rückprojektion und kann von allen Seiten inspiziert und manipuliert werden. Bei diesen Arbeiten besteht eine enge Kooperation mit dem Institut für Medienkommunikation der GMD.



*Abbildung 3-3 : Interaktive Montage komplexer Aggregate auf einer virtuellen Werkbank*

Der Kern des Virtuellen Konstrukteurs ist ein wissensbasiertes Repräsentationssystem, das im DFG-Sonderforschungsbereich 360 (RICKHEIT & WACHSMUTH, 1996) in dem Projekt CODY entwickelt wurde. Durch eine integrierte "Konzept-Dynamik" kann das hier realisierte System zeitveränderliche konzeptuelle Strukturbeschreibungen dargestellter Objekten und Aggregate maschinell verarbeiten. Unter Anwendung von Inferenzregeln werden die die einzelnen Objekte und Aggregate beschreibenden Konzepte dynamisch an den jeweiligen Montagezustand angepaßt; ausgewertet werden dabei auch die aus den CAD-Modellen erschließbaren geometrischen und räumlichen Informationen (JUNG & WACHSMUTH, 1996). Mit der Integration und Nutzung von Objektwissen zielt diese Anwendung auf die intelligente Unterstützung von Entwurfsaktivitäten ab und erschließt die Herstellung "virtueller Prototypen" von CAD-basierten Konstruktionen.

### 3.3 SGIM – Sprach/Gesten-Interface für Multimedia

Ein wesentlicher Ausgangspunkt für neuartige Schnittstellentechniken ist der umfassende Einbezug intuitiver Kommunikationsformen wie Gesten und Sprache. Mit "Gesten" werden hier Körperbewegungen bezeichnet, die benutzt werden, um Information von einer Person zu einer anderen zu übermitteln. Wenn Menschen miteinander kommunizieren, benutzen sie verschiedene Typen von Gesten zu vielfältigen Zwecken und in aller Regel in Verbindung mit Sprache. Die Kombination dieser beiden Modalitäten kann in vielen Fällen Mehrdeutigkeiten auflösen, das heißt dazu beitragen, unscharfe, vage Informationen der einzelnen Modalitäten zu konkretisieren.

Ein solches Szenario ist Gegenstand des 1996 begonnenen SGIM-Projektes; in diesem werden verstärkt multimodale Eingabemöglichkeiten einbezogen, um natürlichere, "anthropomorphe" Benutzungsschnittstellen zu entwickeln. Die neuartigen Techniken sollen den Benutzern eines Multimediasystems den Einsatz grober, auf Armzeigen und Kopfstellung basierender gestischer Interaktion ermöglichen. Damit werden erweiterte Interaktionsmöglichkeiten zum Einsatz mit Groß-Displays (Wandprojektionen, Workbenches) vorbereitet, die ein möglichst komfortables, freistehendes Interagieren erlauben. Die Laboreinrichtungen wurden dazu um eine Tracker-Umgebung für die berührungsfreie Erfassung von Hand- und Körpergestik ergänzt, die an die vorhandene Stereo-Großbildprojektion gekoppelt ist. Durch einfache Mikrofon-Spracheingabe, mit der Objekttypen oder Positionen spezifiziert werden, soll die Interaktion unterstützt werden, um bedeutete Objekte oder Richtungen auch dort analysieren zu können, wo direkte Gestik unnatürlich ist oder an technische Grenzen stößt. Am Beispielsatz *"Montiere dieses Rohr links an den Rahmen"* wird deutlich, daß zur Auflösung der Referenzen weitere Informationen notwendig sind. Eine Zeigegeste auf ein Objekt vom Typ Rohr sowie Kenntnis über den Standort des Benutzers im virtuellen Raum können hier die notwendigen Informationen liefern (siehe Abbildung 3-4).

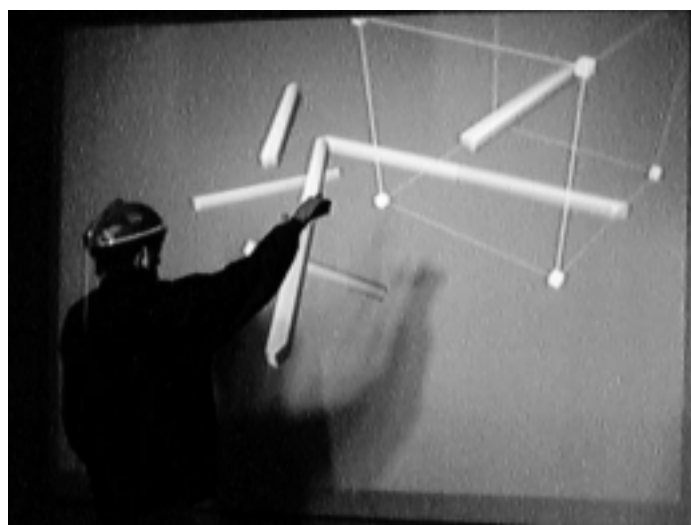


Abbildung 3-4 : Gebrauch von Zeigegestik zur Objektselektion in virtuellen Umgebungen



In den bisherigen Arbeiten wurde zunächst die technische Infrastruktur zur Rohdatenerfassung der Eingaben geschaffen, und es wurden Techniken entwickelt, welche über die Auswertung von elektromagnetischen wie auch elektrischen Sensoren Informationen über die räumliche Bewegungsrichtung der oberen Extremitäten und der Positionierung eines Benutzers ermitteln. Hierdurch ist dem System zu jeder Zeit der Standort des Betrachters bekannt, so daß bei Spracheingaben zum Beispiel Begriffe wie *links* oder *vorn* intuitiv benutzt werden können. Bereits ausgearbeitet wurden erste Verfahren für die Erkennung von entfernter Zeigegestik (LATOSCHIK & WACHSMUTH, im Druck) und Übersetzung registrierter Körperbewegungen in symbolisch verarbeitbare Darstellungen (FRÖHLICH & WACHSMUTH, im Druck). Aufgrund der in Interaktionsszenarien der Virtuellen Realität erforderlichen geringen Latenzzeit für die Gesten- und Spracherkennung wird im SGIM-Projekt eine hochparallele integrierte Verarbeitung, ermöglicht wiederum durch eine Agentenarchitektur, eingesetzt. Hiermit wurde ein erster Prototyp eines lokalen Zeigegestenerkenners fertiggestellt, der testweise mit dem Virtuellen Konstrukteur (siehe Abschnitt 3.2) gekoppelt wurde. Laufende Arbeiten, die zum Teil in Kooperation mit Partnern im Multimedia-Verbund NRW durchgeführt werden, betreffen unter anderem die Einbindung verbesserter Spracherkennungssysteme und die Realisierung einer exemplarischen Anwendung aus dem Bereich des virtuellen Prototyping.

## 4 Ausblick

Sicherlich werden komfortable Schnittstellen der Zukunft stark durch die Multimedia-Technik bis hin zur "virtuellen Umgebung" geprägt sein, unterstützt einerseits durch intuitive Eingabehilfsmittel, die den Gebrauch der natürlichen Ausdrucksformen des Menschen (direkter Zugriff mit den Händen, Sprache, Gestik, Mimik) erlauben, und andererseits durch intelligente, zum Teil auch personifizierte Helfer-Agenten, die über Wissen der Anwendung verfügen und mit Hilfe von Expertensystemtechniken assistierende Funktionen übernehmen. Die Realisierung intelligenter Interfaces als Mehragentensystem bietet – über die geschilderten exemplarischen Anwendungen hinaus – Vorteile der Modularisierung und inkrementellen Entwickelbarkeit, Flexibilität und Robustheit, parallelen Verarbeitbarkeit, Multimodalität und Adaptivität. Ein wichtiger Gesichtspunkt bei unseren Arbeiten mit Agentensystemen ist die Möglichkeit einer kontrollierten Modellierungstiefe: Opake (nur auf die Erbringung von Schnittstellenfunktionen hin betrachtete) Agenten können durch – wiederum als Agentensystem realisierte – "Agenturen" ersetzt werden, wodurch sich zum Beispiel adaptive Fähigkeiten transparent modellieren lassen. Ferner läßt sich durch Kopplung von Agentensystemen in den Datenraum der Anwendung Situationsinformation auswerten, und es lassen sich schließlich innerhalb ein und desselben Systems sowohl ereignis- wie auch zeitlich getriebene Verarbeitungsmuster installieren.

Die damit unternommenen Anstrengungen haben vordringlich mit der 'Anthropomorphisierung' von Schnittstellen, also der besseren Anpassung an den Menschen zu tun. Hier gibt es noch vielfältige Herausforderungen für die Forschung über Mensch-Maschine-Kommunikation. So ist menschliches Kommunikationsverhalten nicht allein von der Informationsübermittlung geprägt, sondern auch von äußerlichen Spezifika in der Nachrichtenübertragung. Etwa erhärten

sich derzeit Forschungserkenntnisse in den Humandisziplinen, die zeigen, daß menschliche Kommunikation inhärent durch rhythmische Muster getragen ist und daß solche Rhythmen, beispielsweise in der Koordination von Sprache und Gestik oder der Synchronisation bidirektionaler Kommunikation (turn taking), hochsignifikant sind. In den theoretischen und praktischen Ansätzen der Mensch-Maschine-Kommunikation haben derartige Erkenntnisse bislang so gut wie keinen Eingang gefunden. Hier ist die Künstliche Intelligenz gefordert, erweiterte Techniken zu entwickeln, die auch das zeitliche Kommunikationsverhalten des Menschen besser berücksichtigen und damit zur Ergonomie von Schnittstellen beitragen. Auf die weiteren Fortschritte darf man gespannt sein.

## Dank und Hinweise

Die hier dargestellten Arbeiten profitieren wesentlich von Forschungsbeiträgen der Mitarbeiter und Doktoranden: Bernhard Jung, Marc Latoschik, Britta Lenzmann, Martin Fröhlich, Martin Hoffhenke, Tanja Jörding, Katy Börner, Yong Cao, unterstützt durch die studentischen Projektmitarbeiter und Diplomanden: Timo Sowa, Marko Merkler, Peter Dawabi, Carla Intrup, Ralf Nolte, Stefan Kopp, Björn Knafla, Sebastian Hübner, Kay Kampeter, Liudger Franzen, Ivo Schnüll, Christian Pütter, Martin Arlt. Ihnen allen sei herzlich gedankt.

Das VIENA-Projekt wurde von 1993 bis 1996 im Forschungsverbund "Anwendungen der Künstlichen Intelligenz in Nordrhein-Westfalen" (KI-NRW) vom Wissenschaftsministerium Nordrhein-Westfalen gefördert und mit dem Ende des Jahres 1997 abgeschlossen. Das CODY-Projekt wird von der Deutschen Forschungsgemeinschaft im Sonderforschungsbereich 360 seit 1993 gefördert. Das SGIM-Projekt wird seit 1996 im Forschungsverbund "Multimedia NRW: Die Virtuelle Wissensfabrik" vom Wissenschaftsministerium Nordrhein-Westfalen gefördert.

## Literatur

ASTHEIMER, P., BÖHM, K., FELGER, W., GÖBEL, M. & MÜLLER, S. (1994). Die Virtuelle Umgebung – Eine neue Epoche in der Mensch-Maschine-Kommunikation; Teil I: *Informatik-Spektrum* 17(5), 281-290; Teil II: *Informatik-Spektrum* 17(6), 357-367.

BARFIELD, W. & FURNESS, T.A. (Eds.) (1995). *Virtual Environments and Advanced Interface Design*, Oxford University Press, 1995.

FRÖHLICH, M. & WACHSMUTH, I. (im Druck). Gesture Recognition of the Upper Limbs – from Signal to Symbol. In I. WACHSMUTH & M. FRÖHLICH (Hrsg.): *Gesture and Sign Language in Human-Computer Interaction*. Berlin: Springer (LNAI 1371).

GÖRZ, G. (Hrsg.) (1995): *Einführung in die Künstliche Intelligenz, 2. Auflage*. Bonn: Addison-Wesley.

- JÖRDING, M. & WACHSMUTH, I. (im Druck). An Anthropomorphic Agent for the Use of Spatial Language. In P. OLIVIER & W. MAASS (eds.): *Representations between Vision and Language*. Berlin: Springer (LNAI).
- JUNG, B. & WACHSMUTH, I. (1996). Ein wissensbasiertes System für die 3D-computergraphische Montage-Simulation. In D. RULAND (Ed.): *Verteilte und intelligente CAD-Systeme: Tagungsband CAD '96 (pp. 107-119)*. Bonn: Ges. für Informatik; Kaiserslautern/Saarbrücken: DFKI.
- LATOSCHIK, M. & WACHSMUTH, I. (im Druck). Exploiting Distant Pointing Gestures for Object Selection in a Virtual Environment. In I. WACHSMUTH & M. FRÖHLICH (eds.): *Gesture and Sign Language in Human-Computer Interaction*. Berlin: Springer (LNAI 1371).
- LAUREL, B. (1990). Interface agents: Metaphors with character. In B. LAUREL (ed.): *The art of human-computer interface design*. Addison-Wesley.
- LENZMANN, B. & WACHSMUTH, I. (1996). Eine multimodale Eingabearchitektur, abstract in M. THIELSCHER & S.-E. BORNSCHEUER (Hrsg.): *Fortschritte der Künstlichen Intelligenz/Workshops KI-96 (p. 93)*, Dresden University Press.
- LENZMANN, B. & WACHSMUTH, I. (1997). Contract-net-based learning in a user-adaptive interface agency. In G. WEISS (ed.): *Distributed Artificial Intelligence Meets Machine Learning - Learning in Multi-Agent Environments (pp. 202-222)*. Berlin: Springer (LNAI 1221).
- RICKHEIT, G. & WACHSMUTH, I. (1996). Collaborative Research Centre "Situated Artificial Communicators" at the University of Bielefeld, Germany. *Artificial Intelligence Review* 10(3-4), 65-170.
- SHNEIDERMAN, B. & MAES, P. (1997). Direct manipulation vs. interface agents – Excerpts from debates at IUI 97 and CHI 97. *Interactions* IV(6), 42-61.
- VAN DAM, A. (1997). Post-WIMP user interfaces. *Communications of the ACM* 40(2), 63-67.
- VON BECHTOLSHEIM, M. (1993). *Agentensysteme – Verteiltes Problemlösen mit Expertensystemen*. Braunschweig: Vieweg.
- WACHSMUTH, I. & CAO, Y. (1995). Interactive Graphics Design with Situated Agents. In W. STRASSER & F. WAHL (eds.): *Graphics and Robotics (pp. 73-85)*. Berlin: Springer.
- WEISS, G. (ed.) (1997). *Distributed Artificial Intelligence Meets Machine Learning – Learning in Multi-Agent Environments*. Berlin: Springer (LNAI 1221).