

Interaktion und Kommunikation in virtuellen Umgebungen

Virtual Environments

(aus Bailenson, J.N., Wiggins, B. , Blascovich, J. & Beall, A.C. (2004, in press))

Der Begriff Virtuelle Realität ist weit gefasst und durch Science-Fictionfilme auch schon mit vielen Fantasien ausgeschmückt worden (man denke z.B. an die Matrixfilme). Außerdem werden viele Begriffe synonym gebraucht wie z.B. Virtuelle Realität, Virtuelle Umgebungen, Artificial Reality, Cyberspace.

Blascovich et al. definieren: „Virtual environments (VEs) as synthetic sensory information that leads to perceptions of environments and their contents as if they were not synthetic”.

Zur Herstellung dieser synthetischen Sensorinformationen dienen im allgemeinen Computer, welche möglichst in Echtzeit mit dem Benutzer interagieren.

Ein gängiges einfaches Beispiel für ein solches System ist liefern (Bente, Krämer & Petersen 2002 S. 10): „Desktop-VR System : Der Computerbildschirm eröffnet ein Fenster oder Portal zu einer interaktiven, dreidimensionalen virtuellen Welt.”

Was bezogen auf Computerspiele den meisten bereits bekannt sein sollte. In vielen Computerspielen wird eine Figur gesteuert, die man im Fachkontext Avatar nennt. Wikipedia gibt folgende Definition für Avatare : "künstliche Person oder ein grafischer Stellvertreter einer echten Person in der Virtuellen Welt. Ein Avatar ist also eine Repräsentation einer Person in der virtuellen Umgebung. Wird ein Avatar von einem Computer gesteuert nennt man das einen Agenten, in der Literatur ist aber auch der Begriff agent-avatar gebräuchlich.

Realer wirkt die virtuelle Umgebung natürlich, wenn man sie als einen umgebend wahrnehmen kann. „An immersive virtual environment (IVE) is one that perceptually surrounds the user of the system.”

In einer solchen IVE trägt man ein Head Mounted Display (HMD) auf dem Kopf, das 2 kleine Displays enthält, die einen die 3D Umgebung aus der eigenen Blickrichtung zeigen. Natürlich braucht ein HMD einen Blickrichtungssensor, damit man bei einer Blickrichtungsveränderung auch dementsprechend die neue virtuelle Sichtrichtung angezeigt bekommt. Des weiteren werden noch Positionsbestimmungskameras benötigt, damit man sich auch durch die virtuelle Umgebung bewegen kann und ein Rechner, der das ganze immer Situationsgerecht berechnet und rendert. Zur auditiven Interaktion trägt man ein Headset. Je genauer die virtuelle Umgebung die eigenen Bewegungen repräsentieren soll, desto genauere Daten über den Körper müssen erfasst werden. Einen wichtigen Beitrag liefert ein Datenhandschuh, der präzise Handbewegungen in der virtuellen Umgebung ermöglicht,

er ermöglicht uns auch Dinge virtuell greifen zu können. Auf Abbildung 1 ist ein solches System zu sehen.



Abbildung 1: System zum Betreten einer IVE

„Often times, multiple people interact with one another inside of the same IVE or VR; this arena is called a Collaborative Virtual Environment (CVE)“.

Ein Beispiel, in dem sich mehrere Personen in einer kollaborativen virtuellen Umgebung befinden, ist ein Chatroom im Internet. Die im Folgenden verwendeten Collaborative Virtual Environments sind via Netzwerk verbundene IVEs.

co-presence

(aus Bailenson, J.N., Beall, A.C., Blascovich, J. (2002))

„Co-presence, also called social presence, occurs when a person uses a virtual environment that contains one or more virtual agents and behaves as if he or she were interacting with other vertiable human beeings.“ Also der subjektive Eindruck der Präsenz des Gegenübers ähnlich wie in der realen Welt. Wenn uns zum Beispiel das Gegenüber vorkommt wie eine Maschine, werden wir ihr weder freundlich guten Tag sagen, noch eine Konversation anstreben. In den Untersuchungen mit CVEs ist es aber meistens gewünscht, dass man so „normal“ oder real mit dem Gegenüber umgeht wie möglich. Deswegen ist eine hohe co-presence in den meisten Versuchsbedingungen erstrebenswert.

Gaze and task performance in shared virtual environments

(aus Bailenson, J.N., Beall, A.C., Blascovich, J. (2002))

Normale Face-to-Face Kommunikation (FTF) ist reich an non-verbalem Informationsgehalt wie Gestik, Mimik. Aber auch nicht verbale Aspekte der Sprache wie die Sprechmelodie und die akustische Qualität der Sprache spielen eine Rolle. Da eine reine Audio-Verbindung als nicht äquivalent zu einer FTF-Kommunikation gesehen wird, versuchen Forscher gezielt

den Informationsgehalt einzelner Kanäle zu ermitteln, z.B. hinzufügen einer Videoverbindung mittels eines Videotelefons. Theoretisch sollte dann der Informationsaustausch besser ablaufen.

Der Einsatz von virtuellen Umgebungen wird vermutlich in der Zukunft zunehmen und eine wichtige Rolle in der weltweiten Kommunikation spielen. Daher ist es wichtig zu wissen, welche Dinge man bei der Entwicklung einer virtuellen Umgebung zu beachten hat. Sozialpsychologen nutzen virtuelle Umgebungen zur Untersuchung von speziellem non-verbalem Verhalten.

Wir konzentrieren uns hierbei auf die Auswirkung von Kopfbewegungen in virtuellen Umgebungen. In der realen Umgebung liefert die Kopfbewegung einer Person einige interessante Informationen:

- Fokus der Aufmerksamkeit:
Die Kopfhaltung liefert gute Anhaltspunkte für den Punkt, auf den die Person gerade ihre Aufmerksamkeit lenkt. Sicherlich muss man auch beachten, dass die Bewegung der Augen auch eine gewisse Rolle spielt. Aber es lässt sich feststellen, dass ein starker Zusammenhang zwischen Augen- und Kopfbewegungen besteht.
- Signalisieren von Zustimmung/Ablehnung:
Durch Kopfbewegungen kann man seine Zustimmung zu bestimmten Sachverhalten durch einfaches Kopfnicken zum Ausdruck bringen bzw. durch Kopfschütteln seine ablehnende Haltung verdeutlichen.

Durch die oben genannten Gründe wird deutlich, dass es also wünschenswert wäre über Kopfbewegungen, Gestik, Mimik usw. innerhalb virtueller Umgebungen zu verfügen. Zum heutigen Zeitpunkt stellt dies aber noch ein technologisches Problem dar. Daher ist es wichtig zu wissen, welche Auswirkungen Kopfbewegungen in virtuellen Umgebungen haben. Also ob sich der zusätzliche Aufwand bei der Entwicklung des virtuellen Systems später auch positiv bemerkbar macht. Diese positive Veränderung kann man z.B. durch gestiegene Task-Performance oder aber auch durch die Rückmeldung der Benutzer über eine höhere co-presence messen.

Überblick über das Experiment:

- Jeweils drei Teilnehmer spielen zusammen in einer virtuellen Umgebung das 20-Fragen-Spiel. Ziel des Spiels ist es durch geschickte Fragen (die mit Ja/Nein zu beantworten sind) einen Begriff zu erraten. 2 Personen müssen die Fragen stellen und die verbleibende Person gibt die entsprechenden Antworten.
- Die Teilnehmer halten sich in 3 getrennten Räumen auf und betreten dann gemeinsam eine virtuelle Umgebung.
- Unabhängige Variable:
Es wurde zwischen drei Bedingungen innerhalb der virtuellen Umgebung unterschieden. Diese waren folgende

- low behaviour: Avatar wird überhaupt nicht dargestellt (reine Audio-Verbindung), vergleichbar mit einer Telefonverbindung
- medium behaviour: Avatar wird dargestellt, zusätzlich bewegen sich die Lippen eines Avatars synchron zur Stimme
- high behaviour: Zusätzliche Darstellung der über Bewegungssensoren wahrgenommenen Kopfbewegung der Probanden innerhalb der virtuellen Umgebung

Durch das Experiment ließen sich folgende Werte feststellen:

- In der high behaviour Bedingung werden (nicht signifikant) weniger Fragen benötigt (Low 9,82 (4,23), Medium 9,77 (5,87), High 8,73 (4,56)).
- Es wird prozentual weniger gesprochen von low nach high Bedingung, wobei die Menge an gesprochenem gleich bleibt. Es gibt also mehr Pausen und es wird insgesamt mehr Zeit benötigt
- Mehr horizontale und weniger vertikale Kopfbewegungen in der high-behaviour Bedingung.

Besonders der letzte Punkt ist interessant. Mögliche Gründe für die Umstellung der Kopfbewegung sind, dass die Teilnehmer versuchen den Aufmerksamkeitsfokus der anderen zu ermitteln oder Kopfbewegungen dazu benutzt wurden, um auf etwas zu zeigen. Die Studie kann hierüber aber keine genaue Aussage treffen.

Entgegen der ursprünglichen Hoffnung, durch zusätzliche non-verbale Informationsquellen die Problemlösung zu vereinfachen, wurde festgestellt, dass die Zeit zur Problemlösung zunimmt. Die Autoren versuchen dies dadurch zu erklären, dass möglicherweise weniger Zeit für die eigentliche Aufgabenstellung verwendet wird.

Die subjektiven Rückmeldungen der Teilnehmer machten jedoch deutlich, dass sie die Kommunikation in der high behaviour Bedingung als am angenehmsten empfunden haben.

Transforming Social Interaction (TSI)

(aus Bailenson, J.N., Beall, A.C., Loomis, J., Blascovich, J. (2004))

Möchte man CVEs z.B. zur Kommunikation zwischen Personen benutzen eröffnen sich ganz neue Möglichkeiten. Man nähert sich damit der realen face to face Kommunikation stark an, mehr als es mit Telefon oder Videotelefon möglich wäre, unterliegt aber nicht allen physischen Gesetzmäßigkeiten der Realität.

Bailenson, Beall & Turk : „ However, unlike a videoconference, a CVE operating system can be designed to render a carefully chosen subset of interactants’ nonverbal behaviors, filter or amplify that subset of behaviors that interactants may not have performed.”

Transformed social interaction (TSI) bietet Systementwicklern die Möglichkeit (non-verbale) Kommunikation zu stärken oder abzuschwächen (z.B. Blickkontakt, Mimik, Körperbewegungen) und im Gegensatz zu Videokonferenzen, kann man bei einem CVE System entscheiden, was gerendert werden soll. Bailenson, Beall & Turk haben einige Ideen für sinnvolle Nutzung von TSI vorgeschlagen:

- Verändern der Selbstdarstellung
 - Veränderung des Avatars um sympathischer zu wirken. „We know that people treat avatars that look like themselves more intimately than avatars that look like others ... ”
 - Anpassung der Körpersprache(Körperhaltung) an das Gegenüber („chameleon effect“). „participants subsequently report that they liked the experimenter more and smoother conversation flowed.“
 - Avatar so verändern, dass er Blickkontakt mit jedem anderen Avatar hat, unabhängig von der eigentlichen Blickrichtung. Damit könnte man in einer virtuellen Lehrereinheit die Aufmerksamkeit der einzelnen Studenten erhöhen, da sich alle angeschaut fühlen.

- Verändern der Sensorinformationen
 - Namen einblenden.
 - Anzeige der Blickkontaktverteilung pro Avatar. Damit hat man selbst eine Rückmeldung ob man einen Avatar bevorzugt behandelt und kann dies wenn gewünscht dann selbst besser auf alle Avatare verteilen (z.B. ein Lehrer in einer virtuellen Lehrereinheit)
 - Automatische Auswertung der non-verbale Signale (Kopfnicken,lächeln,..).
 - Filtern von Signalen (Pokerface, ablenkende Gesten)

- Verändern der Situation
 - Positionen von Gesprächspartnern tauschen.
 - Sich selbst mit den Augen des Gesprächspartners sehen.
 - Aufzeichnen, um sich Erklärungen nochmal anzuhören, während der Veranstaltung
 - Bei Nichtanwesendheit ersetzen des Avatars durch einen intelligenten „Autopiloten“ der die Veranstaltung aufzeichnet und einfache Fragen zum Thema stellen kann. Außerdem abspielen von zuvor aufgezeichneten Teile.

Virtuelle Gesten

Zum Einsatz von VR in der nonverbalen Kommunikationsforschung

(aus Bente, G., Krämer, N.C. (2002))

Da es für die nonverbalen Inhalte einer Kommunikation kein verbindliches Lexikon gibt, also auch keine Grammatik, stellen sich bei der Untersuchung dieser Inhalte diverse Probleme:

- (1) umfassende und detailgenaue Deskription der Daten
- (2) Evaluation zur Bestimmung der Wirkung

Der Messvorgang sollte nicht durch theoretische Vorannahmen beeinflusst werden. Ein solch eingeschränktes Messverfahren setzt die Kenntnis von relevanten Verhaltenselementen voraus. Des Weiteren verhindert ein solcher Ansatz möglicherweise das Auffinden neuer Zusammenhänge.

Man kann nun verschiedene Ansätze der Datenerhebung unterscheiden: restriktiv, generisch und evaluativ. Diese unterscheiden sich in ihrer Zielsetzung, entweder die Wirkung des Verhaltens zu kodieren (evaluativ) oder das Verhalten selbst (generisch und restriktiv). Hierbei stellt sich nun das Problem, dass beide Aspekte von Bedeutung sind. Restriktive oder generische Verfahren führen zu speziellen Informationsverlusten entweder bezüglich der Breite oder des Detailreichtums der Verhaltensprotokolle. Evaluative Ratingverfahren ermitteln zwar die Wirkung des Verhaltens, sie verzichten aber gänzlich auf die Ermittlung des entsprechenden Verhaltensmuster. Bei der Lösung dieses Problems können VR-Technologien eine wichtige Hilfestellung sein:

- Automatische Verhaltensregistrierung durch z.B. Datenhandschuh oder sogar komplette Anzüge) oder aber auch 3D-Animationen als Hilfestellung innerhalb interaktiven Verfahren.
- Realistische Computersimulationen zur Analyse der Wirkung eines Verhaltensmuster

Zunächst einmal kann man auf der Seite der deskriptiven Ansätze noch eine weitere Unterteilung vornehmen, nämlich in die Bereiche direkte Messung und indirekte Messung auch Transkription genannt. Bei indirekten Messverfahren wird sozusagen der Mensch als Messinstrument verwendet und sind somit normalerweise mit einem erhöhten Aufwand verbunden. Die direkten Messverfahren setzen heutzutage oftmals noch das Anbringen von Sensoren voraus. Dies kann dazu führen, dass das natürliche Kommunikationsverhalten verändert wird. Bis die direkten Verfahren ohne diese Sensoren auskommen (also beispielsweise die Daten direkt aus einem Video extrahiert werden können) behalten beide Verfahren ihre Existenzberechtigung.

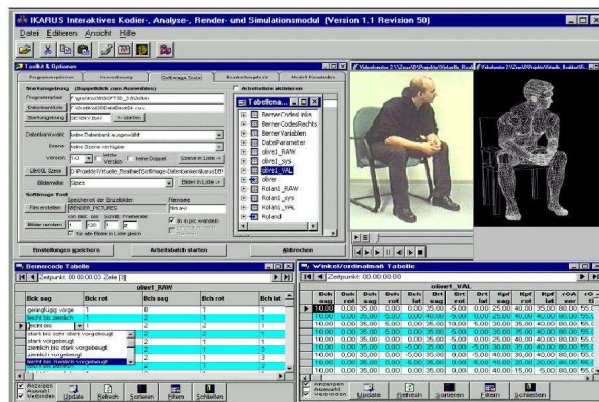


Abbildung 2: Oberfläche des Programms IKARUS

Als hochauflösendes Transkriptionsverfahren im Bereich der Bewegungstranskription hat sich das „Berner System zur Zeitreihennotation nonverbalen Verhaltens“ etabliert. In diesem Verfahren werden Bild für Bild die Positionen der Körperteile kodiert. Durch Weiterentwicklungen dieses Systems unter Verwendung computergestützter Technologien konnte der zur Kodierung nötige Aufwand deutlich reduziert werden. Ein Beispiel für eine solche Entwicklung stellt das Programm IKARUS (Interaktives Kodier-, Analyse, Rendering, und Simulationsmodul) dar. In dem Programm ist es möglich die zur Kodierung nötigen Daten sowohl in der Notation des Berner-Systems als auch in der für Animationsprogramme üblichen Art und Weise einzugeben. Quasi in Echtzeit stellt das Programm ein aus den Daten generiertes 3d-Modell zur Verfügung anhand dessen die Eingaben visuell überprüft werden können und gegebenenfalls punktgenau korrigiert werden können. Abbildung 2 zeigt die Oberfläche des Programms.

Auch im Bereich der direkten Messverfahren konnten durch VR-Technologien bedeutende Fortschritte erzielt werden, insbesondere durch die so genannten Motion-Capturing-Systeme. Neben optischen Systemen zur Messung haben sich vor allen Dingen magnetische Verfahren durchgesetzt. Bei diesem Verfahren werden magnetische Felder von Sensoren erfasst und in Positionsdaten umgewandelt. Ein Vorteil ist die Echtzeitfähigkeit der Systeme. Nachteilig bemerkbar macht sich der durch die Verkabelung eingeschränkte Bewegungsradius.

Für die Aufnahme der Handbewegungen stehen spezielle Methoden zur detaillierten Erfassung der Fingerbewegungen zur Verfügung. Ein leistungsfähiges System dieser Art ist der so genannte CyberGlove (siehe Abbildung 3). Ein Problem, dass viele der direkten Datenverfahren haben, stellt der zeitweise Ausfall der Datenaufzeichnung, z.B. durch Verdecken der Sensoren dar.

Für die Datenerfassung im Bereich der Mimik stehen ebenfalls diverse Systeme zur Verfügung, z.B. durch das Aufbringen von reflektierenden Markierungspunkten im Gesicht deren Bewegungen auf ein 3D-Modell übertragen werden. Nachteilig ist hier, dass spezielle



Abbildung 3: CyberGlove zum Abgreifen der Handbewegung

Markierungen im Gesicht angebracht werden müssen, die evtl. das Verhalten stören könnten. Abhilfe verspricht hier ein System namens Essa (1995). Mit dessen Hilfe es möglich ist, die Gesichtsbewegungen direkt aus einer Videoaufnahme zu extrahieren. Da das System die Daten aus den aus der Bewegung resultierenden Helligkeitsunterschieden gewinnt, muss speziell auf eine gleich bleibende Beleuchtung geachtet werden. Des weiteren ist es hier auch nicht möglich die Bewegungen „frei“ aufzunehmen, denn die Gesichter müssen stets von vorne aufgezeichnet werden.

Zusammenfassend kann man sagen, dass die direkten Verfahren auf Grund ihrer nötigen sensorischen Apparaturen oftmals nur eingeschränkt zur Analyse von spontanem Interaktionsverhalten eingesetzt werden können. Einen möglichen Ausweg stellt hierbei die Möglichkeit der direkten Datenextraktion aus Videoaufzeichnungen dar.

Im Bereich der experimentellen Wirkungsuntersuchungen stellt sich das Problem, dass die gezielte Wirkungsanalyse lange Zeit äußerst begrenzt und erschwert. Besonders beim nahe liegenden Einsatz von instruierten Schauspielern stellt sich das Problem, dass diese nur sehr schwer ihre Aufmerksamkeit auf nur einen speziellen nonverbalen Aspekt konzentrieren können. Eine mögliche Lösung, wenn auch nicht zufrieden stellende, stellt die Verwendung von z.B. mit Fotoretouchetechniken bearbeitetem Bildmaterial dar.

Einen möglichen Ausweg stellt hierbei die Nutzung von VR-Technologien dar, insbesondere die Verwendung der realistischen 3d-Animationen eröffnet viele Möglichkeiten. So wurden solche Systeme beispielsweise schon 1981 bei einer Untersuchung von Cutting und Proffitt verwendet.

Der Vorteil der Verwendung von VR-Technologien zur Wirkungsuntersuchung liegt zum einen in der Möglichkeit Verhalten und Erscheinungsbild frei zu wählen und zum anderen auch in der enormen Anzahl der möglichen Freiheitsgrade durch den gezielten Eingriff.

Virtuelle Realitäten in der psychologischen Forschung und Praxis

(aus Bente, G., Krämer, N.C. & Petersen, A. (2002))

Virtuelle Umgebungen sind sowohl Forschungsgegenstand als auch Forschungsmethode der Psychologie. Um das ganze als Übersicht zu formulieren, lassen sich fünf stark miteinander in Wechselbeziehungen stehende Aufgabenkomplexe herausstellen :

- Grundlagenforschung: Wahrnehmungs- und kognitionspsychologische Untersuchungen zur Auswirkung einzelner Darstellungsparameter in virtuellen Umgebungen (z.B. 3D-Darstellung und Raumwahrnehmung, psychologische Grundlagen der Realitäts-Fiktionsunterscheidung, sozial- und kommunikationspsychologische Analysen der Funktionsprinzipien der Humankommunikation als Grundlage der Modellimplementierungen)
- Realisationsforschung: Umsetzung von kognitions-, emotions- und kommunikationspsychologischen Erkenntnissen in der Entwicklung VR basierter Informations- und kommunikationstechnologien (Implementation von Dialogmanagementsystemen und Erzeugungsalgorithmen für menschliches Interaktionsverhalten im Rahmen von konversationellen Interfaces, Aufbau VR-basierter Kommunikationssysteme - etwa Avatar-Konferenzsystem - etwa zu Testzwecken.
- Evaluationsforschung: Untersuchung der Akzeptanz, Effizienz und Effektivität von implementierten bzw. in der Entwicklung befindlichen VR-Systemen in unterschiedlichen Anwendungskontexten (Kommunikation, Information, Unterhaltung). Bestimmung des Wirkungsgrads (sozial, kognitiv, emotional) einzelner Systemkomponenten und im Vergleich mit anderen Medien
- Einsatz von VR als Forschungsmethode: Herstellung immersiver experimenteller Laborsituationen in der Grundlagenforschung (Erhöhung der ökologischen Validität bei gleichzeitig vollständiger experimenteller Bedingungskontrolle); Computersimulation kognitiver und emotionaler Prozesse; experimentelle Simulation menschlichen Kommunikationsverhaltens.
- Praktische Anwendung von VR: Einsatz von virtuellen Umgebungen im Bereich Wissensvermittlung, Diagnostik, Training und Therapie (virtuelle Seminare, VR-basierte Visualisierung, immersive Simulationsspiele, interaktive Lehr-Lerneinheiten, gestufte Reizexposition in der Angstbehandlung)

Literatur

Bente, G., Krämer, N.C. & Petersen, A. (2002) Virtuelle Realität als Gegenstand und Methode in der Psychologie. In G. Bente N.C. Krämer & A. Petersen (Hrsg.), Virtuelle Realitäten (S.1-31). Göttingen: Hogrefe.

Bente, G., Krämer, N.C. (2002) Virtuelle Gesten: VR-Einsatz in der nonverbalen Kommunikationsforschung. In G. Bente N.C. Krämer & A. Petersen (Hrsg.), Virtuelle Realitäten (S.81-107). Göttingen: Hogrefe.

Bailenson, J.N., Beall, A.C., Blascovich, J. (2002). Mutual gaze and task performance in shared virtual environments. *Journal of Visualisation and Computer Animation*, 13, 1-8.

Bailenson, J.N., Beall, A.C., Loomis, J., Blascovich, J. (2004). Transformed Social Interaction: Decoupling Representation from Behavior and Form in Collaborative Virtual Environments. 428-441 [von <http://www.stanford.edu/group/vhil/>]

Bailenson, J.N., Wiggins, B. , Blascovich, J. & Beall, A.C. (2004, in press). Courtroom applications of virtual environments, immersive virtual environments, and shared virtual environments. *Law and Policy*. [von <http://www.stanford.edu/group/vhil/>]