

Untersuchung von Gestaltungsvarianten blickgestützter Mensch-Computer-Interaktion

Katharina Seifert, Jörn Hurtienne, Thorb Baumgarten
TU Berlin, Fachgebiet Mensch-Maschine-Systeme /
Heinrich-Hertz-Institut für Nachrichtentechnik GmbH Berlin

Zusammenfassung

Im Fokus informationstechnischer Entwicklung stehen zunehmend neue, multimodale Interaktionskonzepte zwischen Mensch und Maschine, die sich in der Gestaltung der Anwendungsschnittstelle niederschlagen. Dabei sind Lösungen gesucht, die Benutzern mit unterschiedlichsten Vorkenntnissen einen unkomplizierten, annähernd natürlichen und intuitiven Umgang mit Computern ermöglichen. Die Frage, wie der breitere Informationsaustausch in der Mensch-Maschine-Interaktion gestaltet werden soll, wird in einem Projekt des Heinrich-Hertz-Instituts Berlin (HHI) untersucht, in dem ein experimentelles Multimodalsystem entwickelt wird. Die vorgestellte empirische Untersuchung zur Blickinteraktion sollte mittels Leistungsdaten (Bearbeitungszeit und Fehler), subjektiver Einschätzung der Beanspruchung und subjektiver Präferenz zur Auswahl einer Gestaltungsvariante dienen. Die Ergebnisse zeigen, dass bei blickgestützter Mensch-Computer-Interaktion mit zuverlässiger Blickortbestimmung eine Rückmeldung der gemessenen Blickkoordinaten nicht erforderlich ist.

1 mUltimo-3D: Ein experimentelles multimodales Computersystem

Die multimodale Mensch-Maschine-Interaktion ist ein Entwicklungsansatz, bei dem der Informationsaustausch parallel über mehrere Kommunikationskanäle stattfindet, wie dies in der zwischenmenschlichen Interaktion typischerweise geschieht. Ein multimodales System ist zudem in der Lage, den Inhalt der Information der verschiedenen Input-Modalitäten automatisch auf einem höheren Abstraktionsniveau zu modellieren (Nigay & Coutaz 1993).

Das am Heinrich-Hertz-Institut für Nachrichtentechnik entwickelte System mUltimo-3D ermöglicht die Interaktion mittels Sprache, Geste und Blick. Es enthält folgende Komponenten, um Benutzerverhalten aufzunehmen, zu interpretieren und systemwirksam umzusetzen:

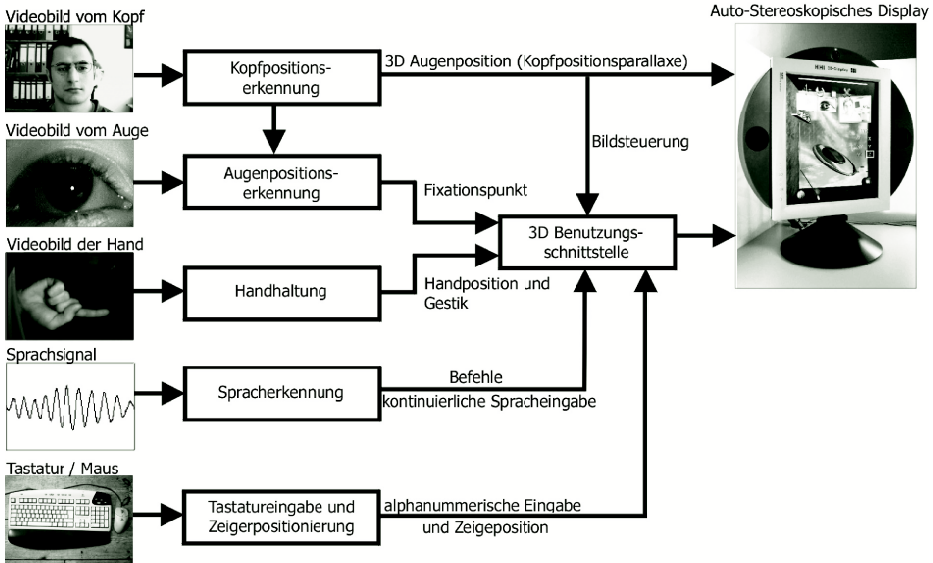


Abb. 1: Diagramm der multimodalen Komponenten von mUltimo-3D.

Abbildung 2 veranschaulicht die basalen Informationsflüsse der multimodalen Mensch-Computer-Interaktion. Dabei sind eine externe, vom Computer vermittelte Rückmeldungsschleife und eine personinterne Kontrolle der motorischen Aktionen zu unterscheiden. Als Modalitäten werden hier die sensorischen Kanäle des Computers bezeichnet, die Verhaltensäußerungen des Benutzers aufnehmen sowie zur weiteren Verarbeitung und Interpretation im Computer weiterleiten. Über die Darstellungsfunktionen (Medien) des Computers wird dem Benutzer die systeminterne Interpretation der Benutzerintentionen aus den erfassten Verhaltenskomponenten zurückgemeldet, indem entsprechende Funktionen ausgelöst werden. Im mUltimo-3D-System können die Informationen auf der grafischen Benutzungsoberfläche dreidimensional mit Hilfe eines autostereoskopischen Displays dargestellt werden.

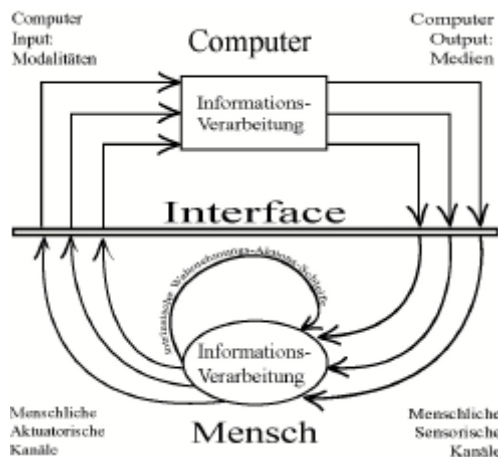


Abb. 2: Modell der Mensch-Computer-Interaktion (nach Schomaker 1995)

Die Computer-Input-Modalitäten sind, unabhängig von ihrer technischen Realisierung, an den Sinnesmodalitäten des Menschen angelehnt. Im mUltimo-3D-System beinhalten sie mit einem Spracherkenner gekoppelte Mikrofone, Maus, Tastatur und Kamerasysteme zur Kopf- und Augenpositions- sowie zur Gestenerkennung. Für jede der zu realisierenden Modalitäten ist zu klären, welche spezifischen Merkmale sie innerhalb der natürlichen Kommunikation aufweisen, um sie in der Mensch-Computer-Interaktion angemessen einzusetzen. Oviatt (1999) verweist darauf, dass Zeitparameter, Informationsinhalte und die interpersonelle Flexibilität in verschiedenen Modalitäten nicht übereinstimmen. Für die blickgestützte Interaktion war zu klären, welche Gestaltungsalternative dem Ziel natürlicher und effizienter Mensch-Computer-Interaktion (Oviatt & Wahlster 1997) am meisten entsprach.

2 Experimentelle Untersuchung von drei Gestaltungsvarianten der blickgestützten Interaktion für mUltimo-3D

Für die Gestaltung der Anwendungsschnittstelle mit blickgestützter Interaktion wurden Varianten der Blickinteraktion untersucht, die zwei Konzeptionen mit unterschiedlichen Zielformulierungen entspringen: dem software-ergonomischen und dem multimodalen Gestaltungsansatz. Es wurde der Frage nachgegangen, ob eine Gestaltungsanforderung nach Transparenz des Systemzustandes gestaltungsleitend sein sollte oder die Natürlichkeit, die sich aus den Gegebenheiten der physikalischen Welt ableitet. Für die Gestaltung von Graphical User Interfaces (GUIs) ist *Rückmeldung* bzw. *Feedback* ein wichtiges Gestaltungskriterium (z.B. Wandmacher 1993, Shneiderman 1992, Norman 1988, Nielsen 1993, Mayhew 1992, Oppermann et al 1992, Preece et al 1994, Smith & Mosier 1986). Meist wird darin allgemein die Rückmeldung des aktuellen Systemzustandes und der vom Benutzer ausgelösten Aktionen gefordert. So fassen die Autoren Rückmeldung unter verschiedenen Begriffen zusammen: Nielsen diskutiert sie unter „Sichtbarkeit des Systemzustandes“, Mayhew unter „Responsiveness“ des Systems. Wandmacher und Oppermann et al stellen die Verbindung zur ISO-Norm 9241-10 her, indem sie die Rückmeldung unter dem Normkriterium Selbstbeschreibungsfähigkeit diskutieren. Feedback wird von ihnen auch mit den Kriterien Durchschaubarkeit oder Transparenz assoziiert. Doch wie sollte die Rückmeldung der Blickpositionen erfolgen? Das Blickverhalten wird im Gegensatz zu Tastatureingaben kontinuierlich vom System interpretiert, allerdings werden nur Ausschnitte der interpretierten Datenmenge systemwirksam umgesetzt. Die Guidelines von Smith und Mosier (1986) schlagen zu diesen Fragen im Kapitel „Eingabe von Positionen“ vor, um Positionseingaben zurückzumelden, einen beweglichen Cursor mit klar erkennbaren visuellen Eigenschaften zu benutzen. Wenn sich die Positionseingabe allerdings nur auf die Auswahl von auf dem Bildschirm dargestellten Objekten beschränkt, genügt es dagegen, die selektierten Objekte durch ein Aufleuchten (Highlight) zu kennzeichnen.

Im Rahmen multimodaler Systemgestaltung soll dem Benutzer eine natürliche Interaktion mit dem Computer ermöglicht werden. Natürlichkeit bedeutet bei der blickgestützten Interaktion, dass durch den Blick des Betrachters keine Veränderung am betrachteten Objekt ausgelöst wird. Der Blick auf ein Objekt gilt als Voraussetzung, gezielte Handlungen am Objekt ausführen zu können (Neumann 1992), manipuliert Gegenstände jedoch nicht eigenständig. Nicht jede Veränderung des Blickortes steht unter der bewussten Kontrolle des Betrachtenden, da Merkmale der Umgebung wie Bewegung oder Auftauchen neuer Objekte unwillkürlich Blickzuwendungen auslösen können. Dadurch wird blickgesteuerte Mensch-Computer-Interaktion erschwert und führt zu Problemen wie dem vielzitierten „Midas-Touch“-Problem (Jacob 1995). In der zwischenmenschlichen Kommunikation wird der Blick vom Interaktionspartner als Indikator für den Ort der visuellen Aufmerksamkeit interpretiert, der jedoch verglichen mit einem Fingerzeig nur vage umschrieben ist. Der Blick wird in natürlichen Kommunikationssituationen als zusätz-

liche nonverbale Verhaltensäußerung interpretiert, die üblicherweise von anderen verbalen und nonverbalen Ausdrucksformen begleitet wird. Die Rückmeldung des Blickes durch eine Veränderung des betrachteten Objektes findet also keine Parallele in der physikalischen Welt. Sie dient aber in der blickbezogenen Mensch-Computer-Interaktion dazu, dem Benutzer die Erkennung seines Blickortes zurückzumelden und ihm zu signalisieren, dass weitere Operationen auf dieses vom Blick selektierte Objekt angewendet werden. Mit Hilfe eines Experimentes sollte darüber entschieden werden, welches Gestaltungsziel für die blickgestützte Mensch-Computer-Interaktion zu bevorzugen ist.

2.1 Experimentalbedingungen und Aufgabe

In einer experimentellen Untersuchung wurden dem Benutzer drei unterschiedliche Varianten der Blickinteraktion angeboten (vgl. Tabelle 1):

1. der Blickort wird kontinuierlich durch einen Blickcursor auf dem Display angezeigt (Cursor-Bedingung),
2. ein interaktives Interface-Element wird optisch hervorgehoben, sobald das System den Blick als darauf ausgerichtet interpretiert (Highlight-Bedingung),
3. der Blickort wird nicht angezeigt, er wird nur systemintern zur Zuweisung eines Tastendrucks zum angeklickten interaktiven Interface-Element genutzt (Bedingung ohne Blickrückmeldung).

Um festzustellen, wie nah die einzelnen Versionen der Blicksteuerung dem natürlichen Verhalten der Benutzer in diesem Experiment sind, gab es eine Tastaturbedingung ohne Blicksteuerung. In jeder der vier Versuchsbedingungen mit mehreren Durchgängen war die Aufgabenstellung für eine Versuchsbedingung grundsätzlich gleich. Auf blauem Bildschirmhintergrund wechselte an drei festen Positionen jeweils ein Buchstabe aus der Menge [T,A,C,U,F,E,H,P,S,L] mit annähernd konstanter Frequenz von 1,7 Sekunden. Die Erstaufgabe der Versuchsperson bestand darin, einen auditiv vorgegebenen Zielbuchstaben („H wie Heinrich“) auf allen drei Positionen zu stoppen. Wurde auf einer Buchstabenposition ein falscher Buchstabe angehalten, dann blieb dieser solange stehen, bis die Probandin ihn durch nochmaliges Anschauen und Tastendruck wieder zum Wechseln brachte. Als Zweitaufgabe tauchte innerhalb eines zeitlich begrenzten Intervalls zufällig ein zum momentanen Blickort peripherer Stimulus auf, den es mit einer definierten Taste zu quittieren galt. Die Verweildauer der peripheren Stimuli betrug zwei Sekunden. Ein Durchgang war beendet, wenn der Proband auf allen drei Positionen den Zielbuchstaben angehalten hat.

2.2 Abhängige Variablen und Hypothesen

Gemessen wurden für alle Bedingungen

1. die Performanz in der Erstaufgabe (Reaktionszeiten, falsche Alarmer und Missings),
2. die Performanz in der Zweitaufgabe (Reaktionszeiten, falsche Alarmer und Missings),
3. die Dauer eines Durchgangs,
4. die subjektiv empfundene mentale Belastung (NASA-TLX, SEA-Skala) und
5. die subjektive Präferenz für eine der Bedingungen.

Für jede dieser Variablen leiten die zuvor geschilderten Gestaltungsziele unterschiedliche Hypothesen ab. Beide Positionen gehen davon aus, dass sie die effiziente und zuverlässige Mensch-Computer-Interaktion unterstützen. Das Gemeinsame der Positionen ist, dass sie die Bedingung

- mit den kürzesten Reaktionszeiten,
- der geringsten Anzahl verpasster Zielbuchstaben (Missings) und
- den wenigsten Fehlern, d.h. Drücken der Zielbuchstaben-Taste bei nicht vorhandenen Zielbuchstaben (falsche Alarmer) bei der Erst- und der Zweitaufgabe sowie

- die Bedingung mit der kürzesten Dauer eines Durchgangs favorisieren.

Die ideale Version ruft auch

- die geringste mentale Beanspruchung hervor und
- wird von den Versuchsteilnehmern bevorzugt.

Ausgehend von der Bedeutung von Rückmeldung für die Mensch-Computer-Interaktion ist anzunehmen, dass die Bedingungen mit Rückmeldung zu kürzeren Zeiten und weniger Fehlern bei der Erstaufgabe führen werden. Bei der Erstaufgabe unterstützt eine Rückmeldung des Blickes die Sichtbarkeit des Systemstatus und ermöglicht einen zügigen Abgleich von interner und externer Rückmeldeschleife im Informationsfluss zwischen Mensch und Computer. Dies hat kürzere Reaktionszeiten und weniger Fehler zur Folge. Bei der Highlight-Bedingung wird ohne Umwege rückgemeldet, welche Buchstabenposition angeschaut wird. Die Cursor-Bedingung dagegen verlangt einen zusätzlichen Kontrollschritt, der prüft, ob sich der Cursor genau auf dem Zielbuchstaben befindet oder nicht. Deshalb sollten in der High-light-Bedingung die kürzesten Zeiten und Fehler auftreten. Da die Blickrückmeldung die Anforderungen bei der Erstaufgabe vereinfacht, sollte mehr kognitive Kapazität für die Zweitaufgabe (Quittieren der peripheren Stimuli) vorhanden sein und die Zeiten und Fehlerraten entsprechend niedriger als bei der Bedingung ohne Rückmeldung liegen. Gleiche Effekte werden für die Gesamtdauer eines Durchgangs erwartet. Die niedrigere Beanspruchung durch die Blickrückmeldebedingungen und hier besonders durch die Version mit Highlight, sollte sich auch in den subjektiven Beanspruchungsdaten und der Präferenz niederschlagen. Tabelle 2 fasst diese Hypothesen zusammen.

Die Unterstützung des Benutzers durch eine natürliche, den Erfahrungen in der physikalischen Welt entsprechende Mensch-Computer-Interaktion lässt andere Annahmen über Leistungen, Beanspruchung und subjektive Präferenzen der Benutzer plausibel erscheinen. Da die Gegenstände der realen Welt nicht auf den Blick reagieren, sollte es für die Nutzer eines blickgesteuerten Systems ungewöhnlich oder gar irritierend wirken, eine Rückmeldung ihrer Blickposition zu erhalten. Vielmehr lenkt die Rückmeldung des Blickes die Benutzer stark auf die Kontrolle des eigenen Blickverhaltens, so dass die aktive Steuerung von Highlight oder Cursor auf die gewünschte Buchstabenposition zusätzliche Anforderungen stellt. Dies lässt die Vorhersage zu, dass Fehler und Zeiten in den Bedingungen mit Blickrückmeldung ansteigen werden. Die permanente Kontrolle des Cursors sollte sich in stärkerem Maße auswirken, als der diskret aufleuchtende sensitive Bereich, wenn das Messsystem Blicke in dieser Region misst. Innerhalb der Bedingungen mit Blickrückmeldung sollten daher die Zeiten und Fehler für die Erstaufgabe in der Cursorbedingung am höchsten sein. Für die Zweitaufgabe gilt entsprechendes für den Vergleich Blickrückmeldung gegenüber fehlender Rückmeldung. Die bewusste Steuerung des Cursors oder des Highlights stellen zusätzliche Anforderungen dar, so dass die Performanz in der Zweitaufgabe darunter leiden sollte. Da die Durchgangsdauer entscheidend von der Performanz in der Erstaufgabe abhängt, gelten hier die gleichen Vorhersagen wie sie für die Erstaufgabe getroffen wurden. Ähnliches gilt für die subjektiven Daten. Die Bedingungen mit Blickrückmeldung werden demzufolge als anstrengender und irritierender wahrgenommen als die Bedingung ohne Rückmeldung. Dabei wird die Cursorbedingung als die forderndste und unangenehmste beurteilt.

Beide theoretische Positionen unterscheiden sich nicht in ihren Vorhersagen zum Verhältnis der Bedingungen mit Blickinteraktion gegenüber der Tastaturbedingung ohne Blickinteraktion. Die blickgestützte Interaktion stellt eine zusätzliche Handlungskomponente (Regulierung des Blickverhaltens) dar, die mit subjektiven und objektiven Kosten verbunden ist. Hinzu kommt eine technische Komponente: Für die Blicksteuerung muss eine komplizierte Sensor- und Interpretationstechnik dazu geschaltet werden, die zu Verzögerungen in der Interpretation des Blickortes führen kann und deswegen Fehleranzahl und Zeiten sowie die subjektive Beanspruchung erhöht.

Tab. 1: Aus den theoretischen Positionen abgeleitete Hypothesen für die Variation der Rückmeldebedingungen. H = Highlight, C = Cursor, O = ohne Rückmeldung, T = Tastaturbedingung ohne Blicksteuerung

Variable	Vorhersage traditioneller Software-Ergonomie	Vorhersage Multimodalitäts-Ansatz	
Erstaufgabe	Reaktionszeit	$((H < C) < O) > T$	$((H < C) > O) > T$
	missings	$((H < C) < O) > T$	$((H < C) > O) > T$
	false alarms	$((H < C) < O) > T$	$((H < C) > O) > T$
Zweitaufgabe	Reaktionszeit	$((H < C) < O) > T$	$((H = C) > O) > T$
	missings	$((H < C) < O) > T$	$((H = C) > O) > T$
	false alarms	$((H < C) < O) > T$	$((H = C) > O) > T$
Dauer eines Durchgangs		$((H < C) < O) > T$	$((H < C) > O) > T$
subjektive Daten	NASA-TLX	$((H < C) < O) > T$	$((H < C) > O) > T$
	SEA	$((H < C) < O) > T$	$((H < C) > O) > T$
	Präferenz (% Nennungen)	$(H > C) > O$	$(H > C) < O$

2.3 Methode

2.3.1 Versuchsaufbau

Die Versuchsperson sass in ca. 80 cm Abstand vor einem hochkant gestellten 18-Zoll-Monitor mit einer Bildschirmauflösung von 1024 x 1280 Pixel. Am Monitor waren eine Videokamera (Head-Tracker), eine Infrarotkamera (Gaze-Tracker), ein Lautsprecherpaar angebracht. Eine Tastatur diente als Eingabegerät (Abbildung 3).

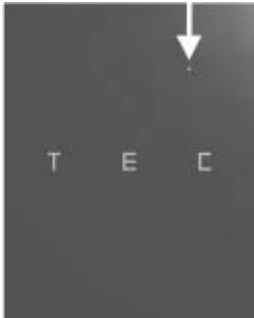


Abb. 3: Versuchsaufbau

Über das Videobild der Kamera wurde die Kopfposition der Versuchsperson verfolgt und der Gaze-Tracker nachgeführt. Der auf dem Cornea-Reflex-Verfahren basierende Gaze-Tracker nahm die Blickrichtung berührungslos und in Echtzeit auf. Mittels Transformationstechnik wurde der momentane Blickort auch bei Kopfbewegungen ermittelt (Liu 1998). Eine detailliertere Darstellung des multimodalen Systems findet sich bei Pastoor et al. (1999). Um einen Buchsta-

ben auf einer spezifischen Position zu stoppen, musste die Versuchsperson in allen drei Blickinteraktionsbedingungen auf die entsprechende Position schauen und gleichzeitig eine der vereinbarten Tasten drücken. Der Blickcursor war eine im Durchmesser 44 Pixel große Scheibe, die kontinuierlich am berechneten Blickort der Versuchsperson auf dem Monitor dargeboten wurde. Ein Highlight hingegen war immer nur dann auf dem Bildschirm zu sehen, wenn der berechnete Blick der Versuchsperson auf ein sensibles Feld einer Buchstabenposition fiel. Dieser Buchstabe wurde dann gelb hinterlegt. Die sensitiven Felder der drei Buchstabenpositionen waren 246x246 Pixel groß, das Highlight war 81x101 Pixel groß und die Buchstaben haben eine Größe von 53x73 Pixel. Im Vergleich dazu hatte der periphere Stimulus einen Durchmesser von 11 Pixel. Bei der Bedingung ohne Rückmeldung gab es keine allein blickbedingte Veränderung der Darstellung auf dem Monitor. In der Kontrollbedingung wurden die Buchstaben nicht durch die Referenzierung des Blickes auf spezifischen Positionen angehalten, sondern immer auf der letzten aktuellen Position mit Zielbuchstaben, sobald die vereinbarte Taste gedrückt wurde. Die Zweitaufgabe blieb über die vier Bedingungen des Experimentes gleich. Die Abfolge der Versuchsbedingungen ist abgesehen von der Kontrollbedingung, mit der immer begonnen wird, vollständig permutiert.

Tab. 2: Variationen der Rückmeldung bei der Blickinteraktion

	Highlight	Blickcursor	Ohne
Beschreibung	Blickinteraktion mit Rückmeldung über das Einfärben der spezifischen Position	Berechneter Blickort wird durch die runde graue Scheibe kontinuierlich rückgemeldet	Berechneter Blickort wird nicht rückgemeldet. Ein peripherer Stimulus ist zu sehen
Abbildung			

2.3.2 Ablauf der Versuche und Datenerhebung

Die 18 bezahlten Versuchspersonen wurden von den Versuchsleitern über den Zweck der Untersuchung informiert und nahmen Platz. Der Stuhl wurde so eingestellt, das die Head-Kamera den Kopf-Schulterbereich der Versuchsperson gut erfassen konnte. Die Versuchspersonen füllten einen Fragebogen mit demografischen Angaben und zur Vorerfahrung in der Computernutzung aus. Dann wurden Head- und Gaze-Tracker konfiguriert und kalibriert. Die Anzahl der Durchgänge pro Version richtet sich nach der Anzahl der quitierten peripheren Stimuli. Es sollten mindestens 20 periphere Stimuli quitiert werden, so dass eine Durchgangsanzahl von 12 bis 18 erreicht wurde. Die ersten drei Durchgänge wurden als Übungsdurchgänge gewertet und von den Datenanalysen, in welche die folgenden neun Durchgänge eingingen, ausgenommen. Nach Beendigung einer Variation wurde den Probanden der NASA-TLX (Hart & Staveland 1988) und der SEA-Fragebogen (Eilers et al. 1986) zur Einschätzung der subjektiv empfundenen Beanspruchung vorgelegt. Zuletzt wurde durch ein strukturiertes Interview die Präferenz der Blick-

rückmeldung erhoben. Die Koordinaten der berechneten Blickorte, die gedrückten Tasten und Reaktionszeiten werden zusammen mit den Informationen über die momentanen Buchstaben auf den Buchstabenpositionen in getrennte Logfiles gespeichert. Die einzelnen Logfiles wurden automatisch personen- und variationsspezifisch sortiert, zusammengefasst, geschnitten und in eine SPSS lesbare Datenmatrix exportiert. Nach der Ausreißerdetektion und –eliminierung wurden die für die Erfüllung der Voraussetzungen der Varianzanalyse nötigen Transformationen durchgeführt und anschließend nach dem Allgemeinen Linearen Modell Varianzanalysen für wiederholte Messungen berechnet.

2.4 Ergebnisse

Rückblickend auf die in Tabelle 1 postulierten Hypothesen kann mit Hilfe der Daten in Tabelle 3 entschieden werden, welche theoretische Position durch die Empirie gestützt wird. Die Spalte „Kontraste“ in der Tabelle 3 zeigt, dass die empirischen Daten in diesem Punkt – außer bei den falschen Alarmen in der Erstaufgabe – die rückmeldungsfree Blickinteraktion unterstützen.

Tab. 3: Mittelwerte der betrachteten Variablen für die einzelnen Versuchsbedingungen und Angaben zur Signifikanz des Einflusses der vier Versuchsbedingungen. *** $p < .005$; ** $p < .01$; * $p < .05$; H = Highlight, C = Cursor, O = ohne Rückmeldung, T = Tastaturbedingung ohne Blicksteuerung; Die Daten zur Präferenz blieben statistisch ungetestet.

Variable		H	C	O	K	F-Wert	Signifi- kanz	Kontraste
Erst- aufgabe	Reaktionszeit (ms)	926	924	792	649	$F(3/51) = 93,537$	***	$((H = C) > O) > T$
	Missings (%)	27,6	23,9	17,5	2,6	$F(3/51) = 22,735$	***	$((H = C) > O) > T$
	falsche Alarme je Durchgang	1,18	0,78	0,91	0,12	$F(2,7/44,7) = 24,361$	***	$((H = C) = O) > T$
Zweit- aufgabe	Reaktionszeit (ms)	841	862	777	703	$F(2,2/37,0) = 14,961$	***	$((H = C) > O) > T$
	Missings (%)	9,8	9,2	4,1	1,8	$F(1,9/32,0) = 6,670$	***	$((H = C) > O) > T$
	falsche Alarme je Durchgang	0,13	0,12	0,05	0,06	$F(3/49) = 3,027$	*	$((H = C) > O) = T$
Durchgangsdauer (s)		22,4	20,5	18,7	13,0	$F(3/50) = 15,722$	***	$((H = C) > O) > T$
subjektive Daten	NASA-TLX	4,8	4,6	3,8	3,5	$F(3/51) = 8,348$	***	$((H = C) > O) > T$
	SEA-Skala	109	112	78	52	$F(3/50) = 22,941$	***	$((H = C) > O) > T$
	Präferenz (% Nennungen)	0,0	22,2	66,7	-			$(H < C) < O$

Entgegen den Vorhersagen für die beiden gestaltungsleitenden Ziele: Systemtransparenz und Natürlichkeit lassen die Ergebnisse keine Differenzierung zwischen der Cursor- und der Highlightbedingung zu. Die Vorhersagen zum Unterschied Blicksteuerung – keine Blicksteuerung bilden sich in den Daten recht gut ab.

2.5 Diskussion

Die Untersuchungsergebnisse zeigen, dass die aus der software-ergonomischen Forschung entsprungene Gestaltungskonzepte im Bereich multimodaler Mensch-Computer-Interaktion nicht ungeprüft übernommen werden sollten. Die Daten sprechen dafür, dass bei ausreichender technischer Zuverlässigkeit der blickgestützten Interaktion, der Benutzer besser mit einer der physikalischen Welt ähnlicheren Gestaltungslösung arbeiten kann. Zwischen den Rückmeldebedingungen mit Cursor oder Highlight konnten keine quantitativen Unterschiede gefunden werden, was darauf schließen lässt, dass sowohl die bewusste Steuerung des Blickcursors als auch des Highlights zusätzliche Kosten für die Aufgabenbearbeitung nach sich zogen. Die Blickinteraktionsbedingungen unterschieden sich nicht gegenüber den falschen Alarmen in der Erstaufgabe und in der Zweitaufgabe. Dies lässt sich damit erklären, dass diese Ereignisse sehr selten auftraten und eine statistische Differenzierung deshalb nicht möglich war. Die meisten Probanden bevorzugten in ihrer subjektiven Beurteilung die Bedingung ohne Rückmeldung.

Für die weitere Entwicklung des multimodalen Systems im Heinrich-Hertz-Institut wird eine blickgestützte Interaktion ohne Rückmeldung bevorzugt. Die überlegene Effizienz der Tastaturbedingung gegenüber allen Blickinteraktionsvarianten legt es nahe zu prüfen, für welchen Aufgabenkontext der Einsatz der blickgestützten Interaktion unter den gegebenen technischen Voraussetzungen sinnvoll ist. Die Unterschiede in den Leistungsmaßen zwischen Tastatur- und den Blickinteraktions-Bedingungen lassen sich vorerst schwierig interpretieren. Ob sie auf technische Faktoren oder Verhaltensänderungen der Probanden zurückzuführen sind, bedarf weiterer Analyse des Prozesses der Aufgabenbearbeitung. Die Betrachtung verschiedener Blickbewegungsparameter können dazu dienen, unterschiedliche Blickstrategien bei der Tastatur- und den Blickinteraktions-Bedingungen aufzuzeigen. Blickgestützte Interaktion sollte nicht zu einer zusätzlich zu bewältigenden Aufgabe des Benutzers werden, sondern ein schnelles Reagieren des Computers auf eine bewusst vorgenommene Eingabe des Benutzers ermöglichen. Die Ergebnisse belegen keineswegs, dass die Natürlichkeit der Interaktion in multimodalen Systemen für alle realisierbaren Modalitäten auf dem Verzicht einer kontinuierlichen Rückmeldung beruht. Welche Merkmale für andere Modalitäten in der Mensch-Computer-Interaktion als natürlich gelten, ist eine empirisch zu beantwortende Frage. Untersuchungen am weiterentwickelten mUltimo-3D-System mit zusätzlichen Modalitäten, die Gesten- und Spracheingabe unterstützen, sollen zukünftig das integrierte multimodale Interaktionskonzept bewerten. Die Bewertungskriterien leiten sich dabei aus den verfolgten Zielstellungen multimodaler Systemgestaltung ab, die Natürlichkeit und Intuitivität der Mensch-Computer-Interaktion zu erhöhen. Für die Gestaltung der multimodalen Mensch-Computer-Interaktion bedarf es einer kritischen Berücksichtigung der etablierten software-ergonomischen Kriterien. Sie müssen daraufhin geprüft werden, wie sie in neuartigen Systemen zu integrieren sind, die durch das neue Konzept des intelligenten Kommunikationspartners Computer dem Benutzer den Umgang mit komplexer Technik erleichtern.

3 Danksagung

Die in diesem Artikel beschriebenen Arbeiten wurden durch Zuwendungen des Bundesministeriums für Bildung und Forschung (Förderkennzeichen 01BK410 und 01BK802) und des Landes Berlin finanziert. Wir danken dem gesamten Team des mUltimo 3D-Projektes für seine engagierte Unterstützung, ohne die die empirische Untersuchung zur Blickinteraktion nicht möglich gewesen wäre. Tanja Köhler ist für ihre Hilfe bei der Versuchsdurchführung zu danken. Die Verantwortung für den Inhalt des Artikels tragen die Autoren.

4 Literatur

- DIN EN ISO 9241 (1996-07) „Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten“, Teil 10: Grundsätze der Dialoggestaltung.
- Eilers, K., Nachreiner, F. & Hänecke, K. (1986). Entwicklung und Überprüfung einer Skala zur Erfassung subjektiv erlebter Anstrengung. *Zeitschrift für Arbeitswissenschaft*, 40, 215-224.
- Hart, S.G. & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock & N. Meshkati (eds.), *Human Mental Workload*. Amsterdam: Elsevier.
- Jacob, R.J.K. (1995). Eye tracking in advanced interface design. In W. Barfield & T.A. Furness (eds.), *Virtual environments and advanced interface design*. New York, Oxford: Oxford University Press.
- Liu, J. (1998). Determination of the point of fixation in a head-fixed coordinate system. *14th Intern. Conf. on Pattern Recognition, Brisbane, 16-20 Aug. 1998*.
- Mayhew, D.J. (1992). *Principles and Guidelines in Software User Interface Design*. Englewood Cliffs, NJ: Prentice Hall.
- Neumann, O. (1992). Theorien der Aufmerksamkeit: von Metaphern zu Mechanismen. *Psychologische Rundschau*, 43, 83-101.
- Nielsen, J. (1993). *Usability Engineering*. Boston: Academic Press.
- Nigay, L. & Coutaz, J. (1993). A design space for multimodal systems – concurrent processing and data fusion. In *INTERCHI'93 – Conference on Human Factors in Computing Systems*. Amsterdam: Addison Wesley.
- Norman, D. A. (1988). *The Design of Everyday Things*. New York u.a.: Doubleday/Currency.
- Opperman, R., Murchner, B., Reiterer, H. & Koch, M. (1992). *Software-ergonomische Evaluation. Der Leitfaden EVADIS II. 2. Auflage*. Berlin New, York: Walter de Gruyter.
- Oviatt, S. & Wahlster, W. (1997). Introduction to *This Special Issue on Multimodal Interfaces, Human-Computer Interaction*, 12, 1-5.
- Oviatt, S. (1999). Ten myth of multimodal interaction. *Communications of the ACM*, 42 (11), 74-81.
- Pastoor, S. Liu, J. und Renault, S. (1999). An Experimental Multimedia System Allowing 3-D Visualization and Eye-Controlled Interaction Without User-Worn Devices. *IEEE Trans. Multimedia*, vol. 1, no. 1, 1999.
- Preece, J. et al. (1994). *Human-Computer-Interaction*. Harlow: Addison Wesley.
- Schomaker, L. et al. (1995). *A Taxonomie of Multimodal Interaction in the Human Information Processing System: A Report of the ESPRIT PROJECT 8579 MIAMI*.
- Shneiderman, B. (1992). *Designing the User Interface. Strategies for Effective Human-Computer Interaction*. 2. Auflage. Reading MA u.a.: Addison-Wesley.
- Smith, S.L. and Mosier, J.N. (1986). *Guidelines for designing user interface software. Technial Report MTR-10090*. Bedford MA: MITRE Corporation.
- Wandmacher, J. (1993). *Software-Ergonomie*. Berlin, New York: Walter de Gruyter.

Adressen der Autoren

Katharina Seifert
Technische Universität Berlin
FG Mensch-Maschine-Systeme
Jebensstr. 1
10623 Berlin
seifert@zmms.tu-berlin.de

Thorb Baumgarten / Jörn Hurtienne
Heinrich-Hertz-Institut für
Nachrichtentechnik Berlin GmbH
Einsteinufer 37
10587 Berlin
baumgarten@hhi.de / hurtienne@hhi.de