

Blickverhalten bei gestenbasierter Interaktion an großflächigen Anzeigen

Thomas Bader, Edmund Klaus

Zusammenfassung

Eingabegeräte, welche auf Blick oder Handgesten basieren, bergen großes Potential für die Gestaltung von intuitiven und ergonomisch guten Interaktionstechniken für großflächige Anzeigen und Multi-Display Umgebungen. Die Nutzung des menschlichen Blicks als Eingabemedium, insbesondere in Kombination mit anderen Modalitäten, muss allerdings auf einem fundierten Verständnis des natürlichen Blickverhaltens basieren.

In diesem Artikel werden die Ergebnisse einer Studie vorgestellt, in der das Blickverhalten während der Interaktion an einer großflächigen horizontalen Anzeige per Handgesten untersucht wurde. Das aufgetretene Blickverhalten wird anhand eines einfachen Modells klassifiziert und bietet eine Grundlage für weitere Untersuchungen bzw. die Entwicklung von blickbasierten Interaktionstechniken.

1. Einleitung

Im Zuge des zunehmenden Informationsangebots wird in vielen Anwendungsbereichen die Visualisierung auf immer mehr und immer großflächigere Anzeigen verteilt. Beispiele hierfür sind Lagezentren oder industrielle Leitwarten, in denen unterschiedlichste Informationskanäle zusammen fließen und (z.T. mehreren) Menschen interaktiv zur Verfügung gestellt werden müssen. Für die Interaktion mit solchen Anzeigekonglomeraten durch mehrere Benutzer sind neue anzeigeübergreifende Interaktionstechniken erforderlich, da herkömmliche Eingabegeräte wie beispielsweise Maus und Tastatur sich hierfür nicht eignen.

In diesem Artikel werden die Ergebnisse einer Studie zur Kombination von blick- und handgestenbasierten Interaktionstechniken zu einem multimodalen Eingabegerät für die Interaktion in Multi-Display Umgebungen vorgestellt und diskutiert. Die Kombination beider Modalitäten birgt großes Potential für die Gestaltung von ergonomisch günstigen und intuitiven Interaktionstechniken für dieses Anwendungsfeld. Per Handgesten können im Greifraum des Benutzers liegende Objekte direkt und intuitiv manipuliert werden. Bei der Interaktion an großflächigen und/oder räumlich verteilten Anzeigen führt dies allerdings schnell zu Ermüdung, da hierfür große räumliche Distanzen überwunden werden müssen. Per Blick können große Distanzen hingegen schnell überbrückt werden, ohne weiträumige und körperlich anstrengende Bewegung durchführen zu müssen.

Eine ausschließlich blickbasierte Interaktionstechnik ist, unter Anderem aufgrund des sogenannten „Midas-Touch“ Problems (Jacob, 1990), jedoch nicht sinnvoll. Durch geschickte Kombination von Blick und Handgestik zu einer multimodalen Eingabetechnik könnten die Vorteile beider Modalitäten genutzt, sowie deren Nachteile gegenseitig aufgehoben werden. Grundlegendes Problem dabei ist, dass es sich beim Auge im Gegensatz zur Hand um ein Wahrnehmungsorgan handelt. Dessen Bewegungen werden oft unterbewusst beeinflusst und hängen von vielen Parametern ab, insbesondere vom Kontext der Interaktion. Um solche Abhängigkeiten verstehen und formal beschreiben zu können, wurde in der hier vorgestellten Studie das menschliche Blickverhalten während der gestenbasierten Interaktion an einer großflächigen horizontalen Anzeige untersucht. Hierzu wurden Verfahren zur vollautomatischen Registrierung und Klassifikation von Blickbewegungen entwickelt. Beobachtete Blickbewegungsmuster werden anhand ihrer räumlich/zeitlichen Beziehung zu aufgabenrelevanten Landmarken klassifiziert und analysiert. Die Ergebnisse werden in Form eines Modells zusammengefasst, welches das Blickverhalten während der Interaktion beschreibt.

In bisherigen Arbeiten vorgestellte Ergebnisse zum Blickverhalten oder zur Hand-Auge Koordination in natürlichen Umgebungen, beispielsweise beim Autofahren (Land, 1994), beim Nachbauen eines Musters aus Kunststoffwürfeln (Pelz, 2001), oder bei rudimentären Objektmanipulationen (Johansson, 2001) wurde festgestellt, dass sich der Blick überwiegend proaktiv verhält, d.h. dem beobachteten Objekt vorausgeht. In einer Studie zur Hand-Auge Koordination während einer einfachen Objektmanipulationsaufgabe (Johansson, 2001) stellte sich zudem heraus, dass Versuchspersonen fast ausschließlich Landmarken fixierten, welche kritisch für die Durchführung der Aufgabe waren (z.B. Hindernisse) und nie das sich bewegende Objekt oder die eigene Hand. In der in diesem Artikel vorgestellten Studie konnten sowohl reaktive Fixationen auf manipulierten Objekten als auch proaktive Fixationen an aufgabenrelevanten Landmarken beobachtet werden. Beide erlauben, wie weiter unten im Artikel diskutiert, interessante Rückschlüsse auf das Benutzerverhalten während der Mensch-Maschine Interaktion.

Das menschliche Blickverhalten wurde auch anhand zahlreicher Aufgaben im Bereich der Mensch-Maschine Interaktion, zum Beispiel bei der Manipulation virtueller Objekte mittels unterschiedlicher Zeigertechnologien (Smith, 2000) untersucht. Die dort erarbeiteten Erkenntnisse basieren aber fast ausschließlich auf Aufgaben, welche an einem herkömmlichen PC-Arbeitsplatz durchgeführt wurden und lassen sich nicht direkt auf die handgestenbasierte Interaktion mit

großflächigen Anzeigen übertragen. Insbesondere die für solche Systeme typische leicht verzögerte und aufgrund mangelnder Robustheit gelegentlich fehlerhafte Reaktion des Systems führt zu einer systematischen Beeinflussung des Blickverhaltens und ist in bisherigen Untersuchungen nicht berücksichtigt.

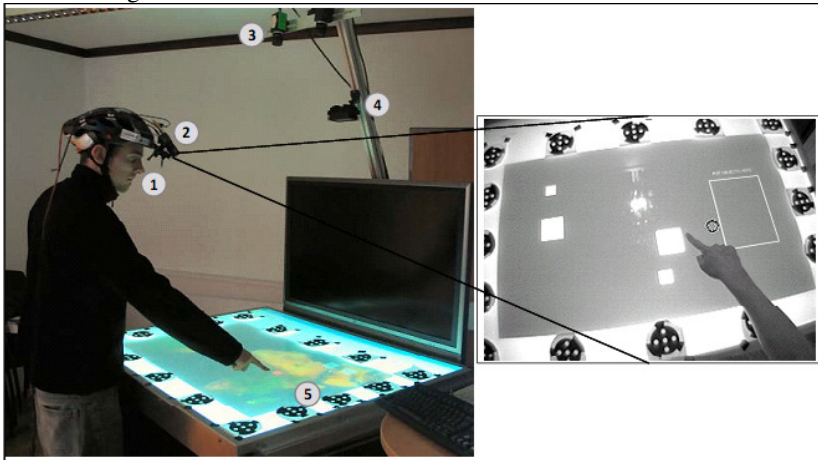


Bild 1: *links*: Versuchsteilnehmer am Digitalen Lagetisch. (1: Eye Tracker, 2: Szenenkamera, 3: Kameras für Handgestenerkennung, 4: Infrarot Beleuchtung, 5: Kodierte Marken) *rechts*: Ausschnitt aus einem Szenenvideo. Der Kreis kennzeichnet die Blickposition im Szenenbild. Auf dem Digitalen Lagetisch ist eine Versuchsaufgabe dargestellt.

2. Methodik

2.1. Aufbau der Versuchsumgebung

Während eines Versuches befand sich jeder Versuchsteilnehmer in zentraler Position vor der horizontalen Anzeige des Digitalen Lagetisches des Fraunhofer Instituts für Informations- und Datenverarbeitung (vgl. Bild 1 und Geisler, 2007). Die Anzeige misst 90 x 120 cm und ist als eine Rückprojektionsanzeige mit einer Auflösung von 1024 x 768 Pixel realisiert. Jedes Pixel hat eine Größe von 1,17 x 1,17 mm. Entsprechend dem Greifraum der Versuchsteilnehmer wurde der

Interaktionsbereich auf ein 824 x 500 Pixel großes Rechteck eingeschränkt, welches zentral auf der horizontalen Anzeigefläche positioniert war.

Als Eingabegerät wurde ein videobasiertes Handgestenerkennungssystem eingesetzt, welches die Pose von Händen oberhalb der horizontalen Anzeigefläche dreidimensional erfasst, sowie unterschiedliche Handsymbole voneinander unterscheiden kann (siehe Bader, 2008). Der Fehler bei der Positionsbestimmung von Hand bzw. Fingerspitzen beträgt 0-7 mm in horizontaler und 1-2 cm in vertikaler Richtung in Nähe der Anzeigefläche. Für die Versuche in dieser Studie wurde nur eine Handgeste eingesetzt, nämlich der ausgestreckte Zeigefinger (siehe Bild 1). Eine Berührung der Anzeigefläche wurde durch das System detektiert, falls der Abstand zwischen der Spitze des Zeigefingers und der Anzeigefläche einen festen Schwellwert unterschritt. Das Handgestenerkennungssystem arbeitet schritthaltend mit der Framerate der Kameras bei einer Frequenz von 25 Hz. Aufgrund von Latenzen während der Bildakquisition und -verarbeitung ist die Reaktion des Systems auf Benutzereingaben konstant um 80 ms verzögert. Während der Versuche konnten zusätzliche, durch Fehlklassifikationen verursachte Latenzen festgestellt werden, welche unregelmäßig in Erscheinung traten.

Die Blickbewegungen der Versuchsteilnehmer wurden mit Hilfe eines SMI iViewX™ HED Eye Trackers aufgezeichnet, welcher die Blickrichtung eines Benutzers mit 50 Hz erfasst. Das System liefert ein Szenenvideo der Versuchsumgebung und die Blickposition in Koordinaten des Szenenvideobildes (vgl. Bild 1).

Um die Blickposition im Szenenvideo in Pixelkoordinaten der horizontalen Anzeigefläche zu transformieren, wurde die Anzeige des Digitalen Lagetisches durch kodierte Marken gekennzeichnet (vgl. Bild 1). Diese können durch eine Tracking-Software im Szenenvideo detektiert, und so die Lage der horizontalen Anzeigefläche im Szenenvideo bestimmt werden. Dies ermöglicht die Transformation der vom Blickbewegungsmessgerät gelieferten Blickposition ins Koordinatensystem der Anzeigefläche. Der durchschnittliche Fehler der Positionsbestimmung des Blicks auf der Anzeigefläche lag zwischen 43,9 und 44,0 mm in x- und 51,2 und 73,2 mm in y-Richtung.

2.2. Datenakquisition- und Verarbeitung

Während der Versuche wurden erfasste Handgesten, Position und Zustand von auf der horizontalen Anzeige dargestellten Interaktionselementen sowie Blickpositionen aufgezeichnet und mit synchronen Zeitstempeln versehen. In den

aufgezeichneten Blickbewegungsdaten wurden durch „Velocity-Threshold Identification“ (Salvucci, 2000) Fixationen detektiert. Die Ergebnisse der automatischen Fixationsdetektion wurden mit manuell ausgewerteten Blickbewegungsdaten erfolgreich verifiziert.

2.3. Versuchsaufgaben

Als Versuchsaufgabe diente eine einfache „Drag-and-Drop“ Aufgabe. Hierbei mussten im linken Teil der Anzeigefläche dargestellte, unterschiedlich große Quadrate in einen Zielbereich im rechten Teil verschoben werden (siehe Bild 1, *rechts*). Die Entfernung der einzelnen Quadrate zum Zielbereich wurde hierbei absichtlich unterschiedlich gewählt, um den Einfluss dieser Größe auf das Blickverhalten zu untersuchen.

Die Position der Quadrate konnte durch einen, aus drei Phasen bestehenden, Interaktionszyklus manipuliert werden. Durch das Berühren eines Quadrats mit dem ausgestreckten Zeigefinger wurde dieses selektiert. In diesem Zustand konnte es durch Bewegung der Hand an eine beliebige Position auf der Anzeigefläche verschoben werden (Manipulation). Durch anheben des Fingers wurde das Quadrat wieder „deselektiert“. Ob sich ein Quadrat im Zustand „selektiert“ befand wurde dem Benutzer durch einen roten Rahmen um das Objekt angezeigt (siehe Bild 2).

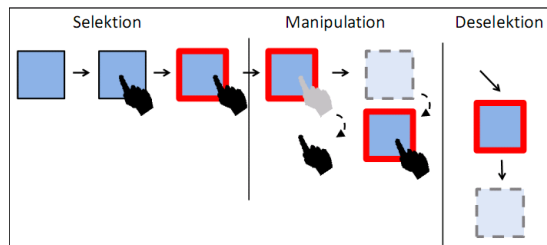


Bild 2: Schematische Darstellung eines handgestenbasierten Interaktionszyklus.

2.4. Versuchsteilnehmer

Elf Personen (2 weiblich, 9 männlich) nahmen an der Studie teil. Alle waren Rechtshänder. Keine der Versuchspersonen benutzte eine Sehhilfe. Alle Personen waren erfahrene Computerbenutzer. Keiner der Teilnehmer hatte Erfahrung mit Handgestenbasierter Interaktion an einer großflächigen Anzeige.

3. Ergebnisse

3.1. Datenanalyse- und Klassifikation

Zum Zweck der Untersuchung des Blickverhaltens von Versuchsteilnehmern wurden die aufgezeichneten Handgesten-, Objektbewegungs- und Blickdaten wie folgt analysiert. Blickbewegungen können in zwei unterschiedliche Kategorien unterteilt werden, Fixationen und Sakkaden. Während Sakkaden schnelle Blickbewegungen sind, die dazu genutzt werden den Blick zu positionieren, bleibt das Auge während einer Fixation fast unbewegt, um die Aufnahme visueller Daten zu ermöglichen.

Wie zuvor erwähnt beginnt ein Interaktionszyklus mit der Selektion eines Objekts und endet mit der Deselektion dieses Objekts im Zielbereich. Erste Untersuchungen der aufgezeichneten Versuchsdaten zeigten, dass Fixationen während eines Interaktionszyklus hauptsächlich von Objektbewegungen und nicht von Handbewegungen beeinflusst werden. Die Hand wurde während der Versuche immer direkt zum Zielbereich hin bewegt, während das Blickverhalten von diesem Schema abwich. Keiner der Versuchsteilnehmer fixierte während einer Objektmanipulation seine Hand. Vermutlich ist dies durch das propriozeptive Feedback zu erklären, welches dem Menschen unter anderem eine grobe Lokalisation der eigenen Gliedmaßen ermöglicht. Im weiteren Verlauf dieses Artikels wird deshalb nur die Relation zwischen Hand- und Objektpositionen ausführlich analysiert.

Im Durchschnitt dauert eine Fixation üblicherweise 150 bis 600 ms (Duchowski, 2003). Während dieser Zeitspanne kann sich die Position eines Objekts verändern. Bei der Analyse wurden daher die Objektpositionen zu Beginn und Ende jeder Fixation untersucht und in Relation zur Position der Fixation gesetzt. In Bild 3 sind die unterschiedlichen räumlichen Relationen zwischen Objekt- und Fixationsposition zu einem bestimmten Zeitpunkt dargestellt. Eine Fixation kann sich auf (a), vor (v) oder hinter (h) dem Objekt befinden. Relationen zwischen Objekt- und Fixationspositionen werden immer in Manipulationsrichtung bestimmt, welche als zweidimensionaler Vektor zwischen der anfänglichen Objektposition und der Position des Objekts bei der Deselektion im Zielbereich definiert ist.

Zur automatisierten Analyse des Blickverhaltens während eines Interaktionszyklus wurde folgende Methodik gewählt. Der Blick verändert seine Position während einer Fixation nicht. Ein Objekt kann während einer Fixation jedoch seinen Zustand (selektiert/nicht selektiert) oder die Position ändern. Um

das Blickverhalten eines Benutzers formal beschreiben zu können wurde die oben eingeführte räumliche Objekt-Fixation Relation am Anfang und Ende jeder Fixation bestimmt.

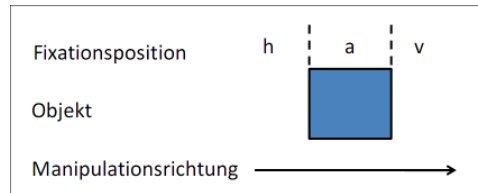


Bild 3: Möglichen Relationen zwischen Objekt- und Fixationsposition zu einem bestimmten Zeitpunkt während der Objektmanipulation.

Im Folgenden sind alle relevanten Tupel von Objekt-Fixation Relationen aufgelistet. Zusätzlich sind diese in Bild 4 illustriert.

- c1: ‚a-a‘ Beginn der Fixation ist auf dem Objekt. Ende der Fixation ist auf dem Objekt. → Objekt hat die Position während der Fixation nicht verändert.
- c2: ‚v-a‘ Beginn der Fixation ist vor dem Objekt. Am Ende der Fixation stimmen Objekt- und Fixationsposition miteinander überein. → Objekt hat sich während der Fixation zur Fixationsposition hin bewegt.
- c3: ‚a-h‘ Beginn der Fixation ist auf dem Objekt. Am Ende der Fixation ist das Objekt vor der Fixationsposition. → Objekt hat sich während der Fixation von der Fixationsposition weg bewegt.

Die Objekt-Fixation Relationen a, v und h beschreiben den Abstand zwischen Objekt- und Fixationsposition qualitativ. Der Ausdruck $dist(x)$ soll im Folgenden diesen Abstand quantitativ beschreiben, wobei $x = \{h, a, v\}$. Damit gilt $dist(a) = 0$ und $dist(v), dist(h) > 0$. Als weitere Objekt-Fixations Relationen ergeben sich damit

- c4: ‚ v_1-v_2 und $dist(v_2) < dist(v_1)$ ‘ Beginn der Fixation ist vor der Objektposition. Am Ende der Fixation ist die Fixationsposition immer noch vor der Objektposition, der Abstand hat sich jedoch verringert. → Objekt hat sich während der Fixation zur Fixationsposition hin bewegt.
- c5: ‚ h_1-h_2 und $dist(h_1) < dist(h_2)$ ‘ Beginn der Fixation ist hinter dem Objekt. Am Ende der Fixation ist die Fixationsposition immer noch hinter dem Objekt und das Objekt ist weiter von der Fixationsposition. → Objekt hat sich von der kurz hinter dem Objekt ausgeführten Fixation weg bewegt.

- c6: ‚v-h und $dist(h) < dist(v)$ ‘ Beginn der Fixation ist vor dem Objekt. Am Ende der Fixation befindet sich die Fixationsposition hinter dem Objekt. Die Distanz zwischen Objekt- und Fixationsposition ist dabei geringer als zu Beginn der Fixation. → Objekt hat sich zum Fixationspunkt hin und ein wenig darüber hinaus bewegt.
- c7: ‚v-h und $dist(h) > dist(v)$ ‘ Beginn der Fixation ist kurz vor dem Objekt. Am Ende der Fixation ist die Fixationsposition hinter dem Objekt und der Abstand zwischen Objekt- und Fixationsposition hat zugenommen. → Objekt hat sich über die in geringer Entfernung davor einsetzende Fixation von der Fixationsposition weg bewegt.

Alle anderen möglichen Kombinationen von Relationen zwischen Objekt- und Fixationspositionen würden eine umgekehrte Manipulationsrichtung des Objekts voraussetzen und finden in dieser Studie daher keine weitere Beachtung.

Die aufgelisteten Tupel von Objekt-Fixation Relationen können bezüglich des zugrundeliegenden Blickverhaltens wie folgt in drei Gruppen unterschieden werden.

- O (c1): Objekt und Blickposition stimmen überein. Die Objektposition ändert sich während einer Fixation nicht.
- P (Proaktiv / c2,c4,c6): Die Fixation geht der Objektbewegung voraus. Zu Beginn der Fixation ist die Fixationsposition vor der Objektposition. Am Ende der Fixation stimmen Objekt- und Fixationsposition überein oder liegen eng beieinander.
- R (Reaktiv / c3,c5,c7): Die Fixation folgt der Objektbewegung. Zu Beginn der Fixation stimmen Objekt- und Fixationsposition überein oder liegen eng beieinander. Am Ende der Fixation befindet sich das Objekt vor der Fixationsposition.

Die Anzahl der Fixationen während eines Interaktionszyklus im Rahmen der Versuche der Studie bewegten sich zwischen 2 und 8. Für weitere Analysen wurden alle Fixationen entsprechend der zuvor erläuterten Objekt-Fixation Relationen klassifiziert. Dies führt zu einer symbolischen Darstellung des Blickverhaltens während eines Interaktionszyklus durch eine Sequenz der Buchstaben ‚O‘, ‚P‘ und ‚R‘. Eine Beispielsequenz ist in Bild 5 dargestellt. Der dort dargestellte Interaktionszyklus besteht aus fünf Fixationen. Anfang und Ende aller Fixationen sind jeweils mit den Buchstaben h, a und v versehen, die Auskunft über die räumliche Relation zwischen Objekt- und Fixationsposition zu

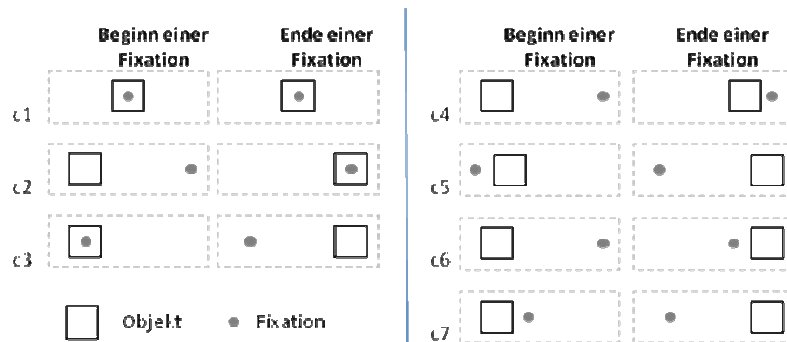


Bild 4: Mögliche Relationen zwischen Objekt- und Fixationspositionen.

diesem Zeitpunkt geben. Des Weiteren ist jede Fixation durch einen der Buchstaben O, P und R gekennzeichnet, die das Verhalten während der gesamten Fixation beschreiben. Daraus ergibt sich im Beispiel die Sequenz ‚RPPPP‘, die eine Charakterisierung des Blickverhaltens während des gesamten Interaktionszyklus zulässt.

3.2. Proaktives und reaktives Verhalten

In den Experimenten der hier beschriebenen Studie starteten alle Interaktionszyklen entweder mit einer O- oder mit einer R-Fixation. Folglich wurde zu Anfang jedes Interaktionszyklus ein Objekt vor der Selektion fixiert. Die darauf folgenden Blickbewegungsmuster lassen sich grob in drei unterschiedliche Verhaltenskategorien einteilen: proaktive, reaktive und gemischte Verhaltensmuster.

Proaktives Verhalten liegt dann vor, wenn der initialen O-Fixation ausschließlich proaktive Fixationen folgen. Analog liegt reaktives Verhalten vor, wenn der initialen O-Fixation ausschließlich reaktive Fixationen folgen. Die symbolische Repräsentation beider Muster kann in Form von regulären Ausdrücken dargestellt werden: $[OR]R^+$ für reaktives (z.B. ‚ORR‘) und $[OR]P^+$ für proaktives Verhalten (z.B. ‚OPP‘).

Reaktives Blickverhalten ($[OR]R^+$) und das proaktive Muster OP mit einer initialen Fixation die direkt von einer proaktiven Fixation im Zielbereich gefolgt wird, wurden bereits in anderen Studien beobachtet (Smith, 2000; Law, 2004).

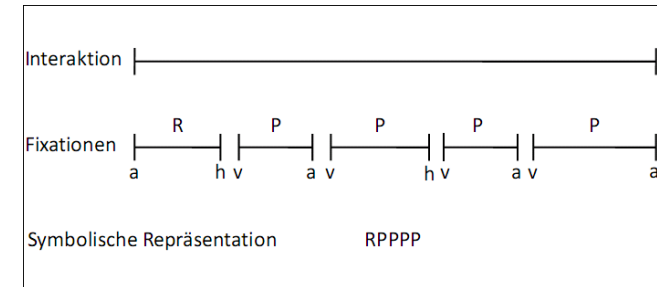


Bild 5: Beispiel eines Interaktionszyklus, mit dazugehörigen Fixationen und deren Klassifikationen.

Im Folgenden werden die Begriffe „direkt zum Ziel“ für das Muster OP und „Objektverfolgung“ für das Muster $[OR]R^+$ verwendet. In der hier vorgestellten Studie konnte das Auftreten weiterer Muster festgestellt werden, die mehrere proaktive Fixationen innerhalb eines Interaktionszyklus beinhalten. Uns sind keine Veröffentlichungen bekannt, in der diese Muster beschrieben wurden. Es werden zwei Gründe für das Auftreten dieses sogenannten „schrittweise proaktiven“ Verhaltens vermutet. Im Vergleich zu Studien, wie Smith (2000) oder Law (2004), in denen die Größe der Anzeige sehr eingeschränkt ist (Desktop-PC), sind im vorliegenden Fall die Ausgangspositionen von Objekten und ihrer Endzustände im Zielbereich über eine größere räumliche Distanz verteilt. Die in Bild 6 dargestellten Ergebnisse unterstreichen den Einfluss der Objekt-Ziel Distanz auf die Anzahl der während einer Manipulation durchgeführte Fixationen. Ein weiterer möglicher Grund für das Auftreten des „schrittweise proaktiv“ Musters ist die leicht verzögerte Systemreaktion durch die von der Handgestenerkennung verursachten Latenzen. Dies kann dazu führen, dass ein Benutzer öfter verifizieren muss, ob das System korrekt auf seine Eingaben reagiert. (vgl. Abschnitt 4).

3.3. Gemischte Verhaltensmuster

Die zuvor beschriebenen Blickbewegungsmuster „Objektverfolgung“, „direkt zum Ziel“ und „schrittweise proaktiv“ traten nicht durchgehend während aller Versuche auf. Es konnten Interaktionszyklen beobachtet werden, die nicht ausschließlich P- oder R-Fixationen enthielten. Die beiden am häufigsten aufgetretenen Muster werden im Folgenden diskutiert.

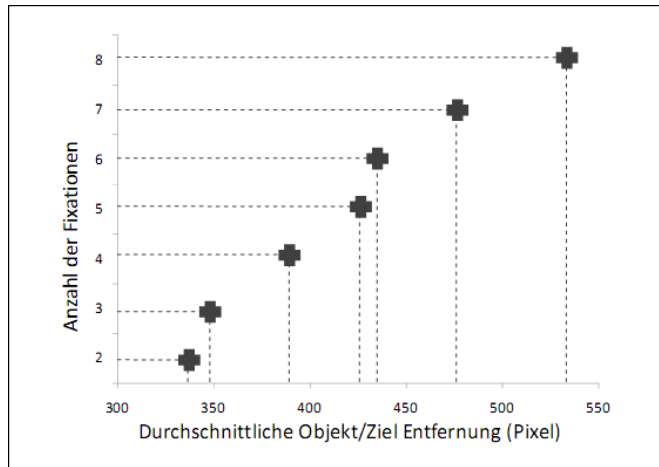


Bild 6: Zusammenhang zwischen der Anzahl von Fixationen während eines Interaktionszyklus und dem Abstand zwischen initialen Objektpositionen und den Endpositionen der Objekte im Zielbereich.

- [OR]R+P+ oder R→P: Versuchsteilnehmer ändert sein Blickverhalten von Reaktiv zu Proaktiv.
- [OR]P+R+ oder P→R: Versuchsteilnehmer ändert sein Blickverhalten von Proaktiv zu Reaktiv.

Bild 7 vermittelt einen Eindruck, wie oft die unterschiedlichen Muster aufgetreten sind. Das häufigste Muster war „schrittweise proaktiv“, gefolgt von „Objektverfolgung“, „R→P“ und „direkt zum Ziel“.

4. Diskussion

In Bild 8 ist ein einfaches Modell dargestellt, welches die in den vorherigen Abschnitten erläuterten Ergebnisse zusammenfasst. Des Weiteren zeigt es eine mögliche Interpretation der Ergebnisse auf und bildet die Basis für weitergehende Diskussionen.

Während der Interaktion hat der Benutzer eines Systems normalerweise ein bestimmtes Ziel vor Augen. In der hier vorgestellten Studie war dies die Positionierung von Objekten in einem bestimmten Bereich, wie in Abschnitt 2.3

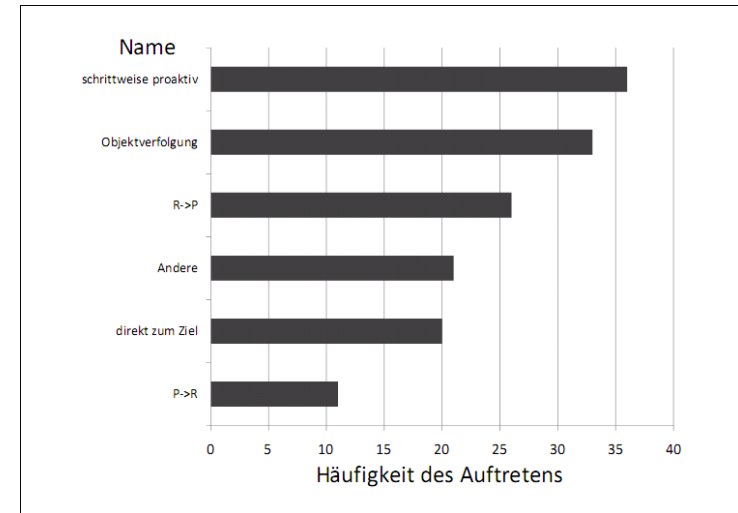


Bild 7: Anzahl des Auftretens von Blickbewegungsmustern innerhalb der durchgeführten Studie.

beschrieben. Um ein bestimmtes Ziel oder Teilziel (z.B. Selektion eines zu manipulierenden Objekts) zu erreichen, müssen vom Benutzer bestimmte Aktionen ausgeführt werden. Wenn eine bestimmte Aktion durchgeführt wurde, kann der Benutzer entweder die Konformität der Reaktion des Systems auf die Eingabe mit dem von ihm erlernten Modell des Systems oder aber das Erreichen eines Ziels bzw. Teilziels (z.B. Objekt ist an gewünschter Position) verifizieren. Im Folgenden sind diese unterschiedlichen Verifikationsverhalten mit aktions- und zielorientierter Verifikation bezeichnet.

Das Feedback, welches vom System bezüglich Reaktionen auf Benutzereingaben oder möglichen Zielzuständen gegeben wurde, war in der durchgeführten Studie hauptsächlich im visuellen Kanal kodiert. Durch die Messung und Auswertung von Blickbewegungen können daher Rückschlüsse gezogen werden, wo und wann welches Feedback vom Benutzer wahrgenommen wurde bzw. wo und wann Feedback erwartet wurde.

Bei reaktivem Blickverhalten kann lediglich verifiziert werden, ob sich ein bestimmtes Objekt von einer Stelle weg bewegt hat, möglicherweise auch in welche Richtung (im peripheren Sichtfeld). Bei proaktivem Blickverhalten wird

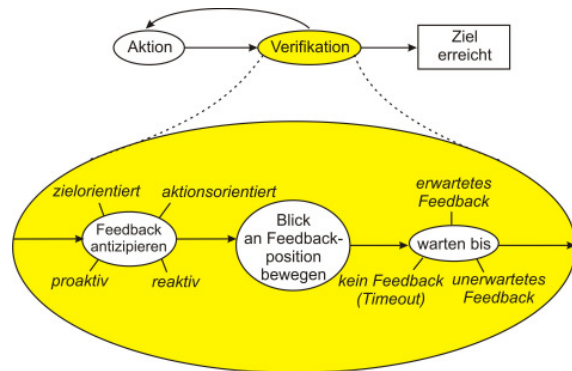


Bild 8: Einfaches Modell zur Charakterisierung von Blickverhalten während der Interaktion an einer großflächigen Anzeige.

unter Verwendung eines erlernten Systemmodells eine bestimmte Systemreaktion bzw. ein Objektzustand (z.B. Position) vom Benutzer antizipiert und mit der tatsächlichen Systemreaktion bzw. dem tatsächlichen Objektzustand verglichen. Das erlaubt z.B. während einer Bewegungsphase die gleichzeitige Verifikation ob ein Objekt sich überhaupt bewegt hat (wie auch mit reaktivem Verhalten möglich) *und* ob das Objekt sich an die gewünschte Position bewegt hat. Der Informationsgewinn für den Benutzer ist daher größer mit proaktiver als mit reaktiver Verifikation. Andererseits ist für proaktive Verifikation mehr Wissen über das System notwendig, da dessen Zustand auf Basis der durchgeführten Eingaben prädiziert werden muss. Im Folgenden wird diskutiert, wie die im vorherigen Abschnitt dargestellten Ergebnisse dieses Modell und die damit verknüpfte Interpretation stützen.

Während der Bewegungsphase der durchzuführenden Manipulationsaufgabe wurde sowohl aktions- als auch zielorientiertes Verifikationsverhalten beobachtet. Während das Blickbewegungsmuster „direkt zum Ziel“ proaktivem zielorientiertem Verifikationsverhalten entspricht, enthalten „schrittweise“ pro- und reaktive Blickbewegungsmuster sowohl aktions- als auch zielorientierte Verifikationsschritte (Fixationen). In Übereinstimmung mit (Law, 2004) wurde das Blickbewegungsmuster „direkt zum Ziel“ häufiger mit zunehmender Erfahrung der Benutzer festgestellt. Dies unterstützt obige Vermutung, dass mehr Wissen für diese Art von Verifikationsverhalten nötig ist. Dasselbe gilt für Wechsel von proaktivem zu reaktivem Verifikationsverhalten während einem

Interaktionszyklus („P→R“). Während die Dauer von P-Fixationen über alle Experimente und Versuchspersonen ungefähr 320 ms beträgt, war die letzte P-Fixation vor einem Wechsel zu reaktivem Verhalten im Muster „P→R“ um ungefähr 200 ms länger (durchschnittlich 520 ms). Dies deutet darauf hin, dass verzögertes, d.h. unerwartetes Feedback mittelbar zum Wechsel von proaktivem zu reaktivem Verifikationsverhalten geführt hat. Da unerwartetes Feedback gleichzusetzen ist mit zu dem vom Menschen erlernten Systemmodell nicht konformer Systemreaktion, führt eine solche unerwartete Reaktion des Systems zwangsläufig zu einem Vertrauensverlust in dieses Modell. Der damit einhergehende Verlust an menschlicher Prädiktionsfähigkeit von Systemreaktionen führt schließlich zum Wechsel zu reaktivem Verifikationsverhalten bei dem diese Fähigkeit, wie oben bereits erwähnt, nur in geringerem Ausmaß benötigt wird.

5. Zusammenfassung

In diesem Artikel wurden die Ergebnisse einer Studie zu menschlichem Blickverhalten bei handgestenbasierter Interaktion an einer großflächigen horizontalen Anzeige vorgestellt. Ein Verfahren zur automatischen Analyse und Klassifikation von Blickverhalten basierend auf Fixation-Objekt Relationen wurde dargestellt. Die beobachteten Blickbewegungsmuster konnten direkt in Beziehung zu Verifikationsverhalten während der Interaktion gesetzt werden. Dieses wurde entlang zweier Dimensionen als ziel- bzw. aktionsorientiertes und proaktives bzw. reaktives Verhalten charakterisiert. Systemreaktionen welche aufgrund des mentalen Modells des Systems vom Benutzer erwartet wurden führten zu zielorientiertem und proaktivem Verifikationsverhalten, während unerwartetes verspätetes Feedback zu aktionsorientiertem reaktivem Verhalten führte.

Die dargestellten Ergebnisse sind interessant für die Gestaltung von blickbasierten oder multimodalen Interaktionstechniken für tischartige oder großflächige Anzeigen. Außerdem bilden sie die Grundlage für eine blickbasierte Intentionsschätzung und können zur Evaluierung und Verbesserung nicht-blickbasierter Interaktionstechniken genutzt werden.

In zukünftigen Arbeiten soll der Einfluss von unerwartetem und verzögertem Feedback auf das Blickverhalten eines Benutzers weiter untersucht werden. Außerdem sollen neue, blickbasierte Interaktionstechniken für Multi-Display Umgebungen auf Basis der gewonnenen Erkenntnisse entwickelt werden.

Literatur

- Bader, T. (2008). Videobasierte Handgestenerkennung: Anforderungen und Umsetzung für die Interaktion an horizontalen Anzeigen, In *tm - Technisches Messen* 75 Nr. 7-8, S. 429-436
- Duchowski, A. (2003). *Eye Tracking Methodology, Theory and Practice*. Springer
- Flanagan, J. & Johansson, R. (2003). Action plans used in action observation. *Nature*, 424:769–771
- Geisler, J., Eck, R., Rehfeld, N., Peinsipp-Byma, E., Schütz, C., & Geggus, S. (2007). Fovea-Tablett: A New Paradigm for the Interaction with Large Screens. *Proceedings of HCII 2007, Human Interface, Part I*, S. 278-287
- Gesierich, B., Bruzzo, A., Ottoboni, G. & Finos, L. (2008). Human gaze behaviour during action execution and observation. *Acta Psychologica*, 128:324–330
- Jacob, R. (1990). What you look at is what you get: eye movement-based interaction techniques. In *CHI '90: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 11–18. ACM
- Johansson, R., Westling, G., Bäckström, A. & Flanagan, J. (2001). Eye-hand coordination in object manipulation. *The Journal of Neuroscience*, 21(17):6917–6932
- Land, M. & Lee, D. (1994). Where we look when we steer. *Nature*, 369:742–744
- Law, B., Atkins, M., Kirkpatrick, A. & Lomax, A. (2004). Eye gaze patterns differentiate novice and experts in a virtual laparoscopic surgery training environment. In *ETRA '04: Proceedings of the 2004 symposium on Eye tracking research & applications*, pages 41–48. ACM
- Pelz, J., Hayhoe, M. & Loeber, R. (2001). The coordination of eye, head, and hand movements in a natural task. *Exp Brain Res*, pages 266–277
- Salvucci, D. & Goldberg, J. (2000). Identifying fixations and saccades in eye-tracking protocols. In *ETRA '00: Proceedings of the 2000 symposium on Eye tracking research & applications*, pages 71–78. ACM
- Smith, B., Ho, J., Ark, W., & Zhai, S. (2000). Hand eye coordination patterns in target selection. In *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications*, S. 117-122

Autoren

Dipl.-Inform. Thomas Bader Fraunhofer Institut für Informations- und
Dipl.-Inform. (FH) Edmund Klaus Datenverarbeitung IITB, Karlsruhe