

Neue Interaktionstechniken auf Grundlage videobasierter Handgestenerkennung für eine Multi-Display-Umgebung

Sebastian Maier & Thomas Bader

Zusammenfassung

Interaktive Umgebungen entwickeln sich in vielen Anwendungsbereichen zunehmend weg vom Personal-Computer hin zu interaktiven Multi-User Multi-Display-Umgebungen. Für diese Systeme müssen neue Techniken zur Interaktion entwickelt werden, da sich die klassischen Interaktionstechniken nicht einfach übertragen lassen. In diesem Artikel werden neue Interaktionstechniken vorgestellt und diskutiert, welche auf Basis eines videobasierten Systems zur Handgestenerkennung umgesetzt wurden, sowie erste Ergebnisse zur Evaluierung dargestellt.

Als Experimentierplattform dient der Digitale Lagetisch (Bader et al., 2008), ein Multi-Display-Arbeitsplatz, der einem Team Daten zur gemeinsamen Lösung eines Problems interaktiv zur Verfügung stellt und ein neu entwickeltes System zur Handgestenerkennung (Bader et al., 2009), das über dem Tisch angebracht ist. Die vorgestellten Interaktionstechniken wurden exemplarisch entlang eines Szenarios zur Handhabung von Geodaten entwickelt, können aber auf viele andere Anwendungsbereiche übertragen werden.

1 Motivation

Die Variantenvielfalt an kostengünstig zur Verfügung stehenden mobilen und stationären Displays führt zunehmend zu einer verteilten Informationsvisualisierung. Damit einher geht die Notwendigkeit, nicht nur mit einzelnen Displays, sondern mit Multi-Display-Umgebungen interagieren zu können. Die videobasierte Erfassung des Benutzers bietet die Möglichkeit, Eingaben berührungslos und unabhängig von an bestimmte Displays gebundenen Eingabegeräten (z.B. Touchscreen) zu erfassen. Die sinnvolle Umsetzung von Interaktionstechniken für Multi-Display-Umgebungen auf Basis solcher Eingabegeräte ist in großen Teilen noch ungelöst und ist Gegenstand dieses Artikels.

Als Beispiel für eine Multi-Display-Umgebung dient hierbei der Digitale Lagetisch (Bader et al., 2008), welcher in Abschnitt 3 näher beschrieben wird. Zur berührungslosen und displayunabhängigen Erfassung von Eingaben wird ein videobasiertes Handgestenerkennungssystem verwendet, welches im Vergleich zu

anderen Eingabesystemen wie z.B. kapazitiven Touchscreens oder der Benutzung von Hilfsmitteln wie verschiedenen Stiften insbesondere folgende Vorteile bietet:

- Da bei der Gestenerkennung im Gegensatz zu Touchscreens die gesamte Hand und nicht nur einzelne Druckpunkte erfasst werden, ist es möglich Hand-Shortcuts für bestimmte Funktionen zu definieren.
- Durch die Erfassung der Hand im Raum sind Zeigegesten ohne zusätzliche Hilfsmittel möglich.
- Ein Bildschirm muss nicht berührt werden, um Aktionen auszulösen. Dies ermöglicht unter anderem die Interaktion mit entfernten Anzeigen.

In der diesem Artikel zugrunde liegenden Arbeit wurde der Frage nachgegangen, wie sich diese Eigenschaften für die Entwicklung sinnvoller Interaktionstechniken in Multi-Display-Umgebungen nutzen lassen.

2 Stand von Forschung und Technik

In Bader et al. (2010) wurde gezeigt, dass die Erfassung der Hand im Raum Vorteile gegenüber der Erfassung von Eingaben nur bei Berührung einer Displayoberfläche beim Verschieben von Objekten zwischen zwei horizontal angeordneten Displays bietet. Die in diesem Artikel vorgestellten Interaktionstechniken wurden entwickelt, um den durch die videobasierte Handgestenerkennung aufgespannten Designraum weiter zu explorieren, insbesondere was die Interaktion auf entfernten Anzeigen anbelangt.

Der Stand der Forschung, welcher für die Gestaltung von natürlichen Interaktionstechniken relevant ist, wird im Folgenden in zwei verschiedene Bereiche unterteilt: Zum einen der Bereich der Analyse menschlicher Aktivitäten, bei dem untersucht wird wie Menschen mit ihren Händen in natürlichen Umgebungen arbeiten. Zum anderen der Bereich der Mensch-Maschine-Interaktion.

2.1 Analyse menschlicher Interaktion

In (Guiard, 1987) werden die grundlegenden Aktionen beschrieben, die Menschen mit den Händen durchführen. Sie werden in drei Gruppen unterteilt:

- Unimanual: Aktionen, die nur mit einer Hand durchgeführt werden wie z.B. Zähneputzen.
- Bimanual und asymmetrisch: Beide Hände werden zur Aktion benötigt, z.B. Spielen eines Streichinstrumentes.
- Bimanual und symmetrisch: Beide Hände führen die gleiche Aktion aus, z.B. Heben eines schweren Gegenstandes.

Dabei fallen bei bimanualen asymmetrischen Aufgaben den Händen verschiedene Rollen zu. Es kann als gesichert angesehen werden, dass die starke Hand sowohl in zeitlicher als auch in räumlicher Hinsicht genauer arbeitet als die schwächere (Peters, 1980). Dabei hat allerdings die schwächere Hand oft eine stabilisierende Rolle oder dient als Referenzpunkt für die Bewegungen der stärkeren. Der

Bewegungsradius ist dabei bei der schwächeren Hand größer. Für viele asymmetrische Aktionen gilt, dass die Aktion mit der Bewegung der schwächeren Hand beginnt und die stärkere folgt. In Kabbash et al. (1994) wurde untersucht, welchen Einfluss die Verwendung der zweiten Hand auf die Arbeitsgeschwindigkeit hat. Dabei erhöhte sich die Geschwindigkeit nur, wenn bereits beim Design der Interaktionstechnik Fähigkeiten der bedienenden Person aus der realen Welt berücksichtigt wurden.

Hinckley et al. (1997) experimentierten mit beidhändigen Aufgaben, bei denen jeder Hand eine spezielle Rolle zugeteilt war. Die Leistung war dann am besten, wenn sich die starke Hand relativ zu einem „frame-of-reference“ der schwachen Hand bewegte. Ist die Aufgabe einfach genug, spielt es keine Rolle mehr, mit welcher Hand sie gelöst wird.

2.2 Interaktionstechniken

Die Interaktion mittels Touch, Multitouch oder per Gestik wurde für einzelne Displays schon in zahlreichen Arbeiten untersucht. Im Folgenden werden verschiedene daraus abgeleitete Design-Prinzipien vorgestellt, auf deren Basis die Interaktionstechniken für den Digitalen Lagetisch entworfen wurden.

In Wu et al. (2006) werden drei Phasen von Handgesten unterschieden: Startposition, Dynamische Phase und Endposition. Eine wichtige Erkenntnis ist das Handsymbol während der dynamischen Phase: Um ein Verkrampfen der Hand und eine zu hohe körperliche Belastung des Benutzers zu vermeiden, sollte die Handgeste nur in der Erkennungsphase von Bedeutung sein. Ein weiteres Prinzip ist die Anzahl der verwendeten Handsymbole. Gesten sollten soweit wie möglich wiederverwendet werden. Ein weiterer wichtiger Punkt ergab sich aus der Evaluation der Techniken. Eine einmal begonnene Aktion muss ohne Seiteneffekte abgebrochen werden können, um den Benutzer nicht zu verwirren und ohne unerwünschte Aktionen auszulösen.

Utsumi und Ohya (1999) beschreiben ein System, um Handgesten im Raum zu verfolgen. Sie beschreiben weiter eine einfache Anwendung um virtuelle Objekte zu erstellen und zu bearbeiten. Dabei verwenden sie die Übergänge von verschiedenen Gesten, um Aktionen auszulösen. Es wird immer der Übergang zwischen einer neutralen Geste, hier der flachen Hand mit gespreizten Fingern, und einer Zielgeste zum Auslösen von Aktionen verwendet. Wie schon von Utsumi und Ohya (1999) angemerkt, können nicht alle Personen alle Gesten ausführen, was die Menge an ergonomisch sinnvollen Gesten einschränkt.

Wobbrock et al. (2009) gehen den umgekehrten Weg, um Interaktionstechnik zu entwerfen. In einer Studie konnten die Versuchspersonen selbst die Technik bestimmen, die sie für eine Interaktion verwenden würden. Den Versuchspersonen wurde dabei eine Aufgabe gestellt, z.B. Verschieben eines Rechtecks, und dabei wurde beobachtet, welche Aktion sie zum Lösen der Aufgabe ausführten. Das Ergebnis waren 27 verschiedene Interaktionstechniken für den

Einsatz an tischartigen Anzeigen. Die Testpersonen arbeiteten meist nur mit einer Hand, wobei die Anzahl der Finger für sie keine Rolle spielte.

Vogel und Balakrishnan (2005) zeigen ein System zum Zeigen und Klicken auf einem großen entfernten Display. Um die Handposition zu erfassen, verwenden sie ein Motion-Tracking-System, das verschiedene Sensoren an einem Handschuh benötigt. Durch die Handbewegung und kleinere Bewegungen der Finger können Elemente auf dem entfernten Display ausgewählt werden.

3 Der Digitale Lagetisch

Beim Digitalen Lagetisch (Bild 1) handelt es sich um einen Multi-Display-Arbeitsplatz, der für die kooperative Teamarbeit entwickelt wurde (siehe u.a. Bader et al., 2008). Mit Hilfe eines großflächigen tischartigen Displays lassen sich geographische Informationen für ein Team zur Verfügung stellen. Diese Darstellung, die an den klassischen Kartentisch angelehnt ist, ermöglicht es mehreren Personen, ein Problem gemeinsam zu bearbeiten und zu lösen. Der Tisch selbst besteht dabei aus einer Scheibe, auf die über einen Spiegel das Bild mit einer Auflösung von 1400x1050 Pixel projiziert wird. Als Anzeige für die Sekundärinformationen kommt ein 50“ Full-HD Display zum Einsatz. Auf ihm werden Informationen dargestellt, die aus Gründen der Übersicht nicht auf dem Tisch angezeigt werden können. Dabei handelt es sich um verschiedene Medien wie Texte, Bilder, Videos, aber auch 3D-Modelle, für die die vertikale Anzeige besser zur Darstellung geeignet ist.

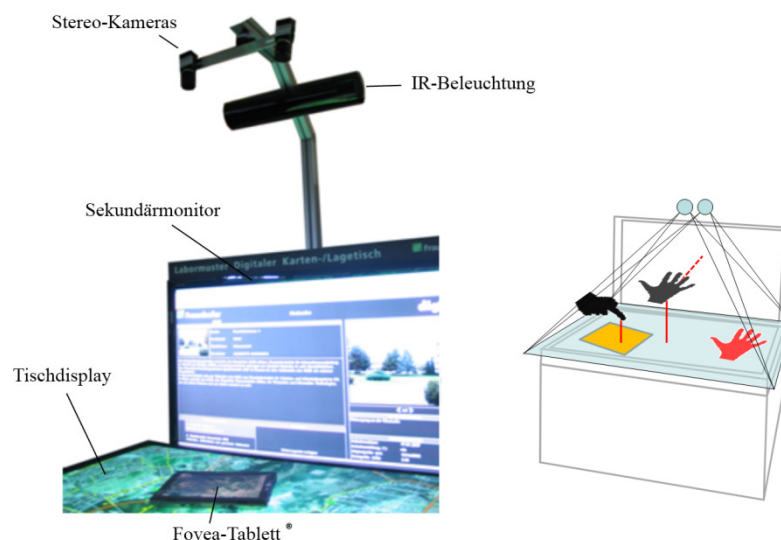


Bild 1: links: Aufbau des Digitalen Lagetisches mit Kameras und Beleuchtungssystem; rechts: schematische Darstellung der Handgestenerkennung

Eine weitere Komponente des Digitalen Lagetisches stellt das so genannte Fovea-Tablett[®] dar. Dabei handelt es sich um Tablet PCs die auf den Tisch gelegt werden können. Die Position und Ausrichtung der Tablett wird dabei von einem videobasierten Trackingsystem erkannt, welches unter dem Tisch angebracht ist. Die Tablett können über einen speziellen Stift bedient werden.

4 Die Gestenerkennung

Das eingesetzte Gestenerkennungssystem wurde am Fraunhofer IOSB entwickelt (Bader et al, 2009). Es setzt auf einem Stereo-Kamera-System auf, das über dem Lagetisch angebracht ist (Bild 1), und ist in der Lage, mehrere Hände gleichzeitig zu erkennen und zu verfolgen.

Durch den Einsatz von Tageslichtsperrfiltern und einer angepassten Infrarot-Beleuchtung ist es möglich, das Gestenerkennungssystem unabhängig von äußeren Lichtbedingungen zu betreiben.

Bei der Klassifizierung der Hände werden von der Gestenerkennung 16 verschiedene Handsymbole unterschieden (Bild 2). Die Anzahl erhöht sich, wenn auch linke Hände betrachtet werden. Alle Gesten werden im Folgenden mit einer einfachen Zeichenkette bezeichnet. Ein führendes „R“ oder „L“ gibt die rechte oder linke Hand an, darauf folgen fünf 0er oder 1er, wobei eine 1 für einen ausgestreckten, eine 0 für einen eingezogenen Finger steht. Sind zwei benachbarte Finger voneinander abgespreizt, wird dies durch einen Unterstrich „_“ verdeutlicht.

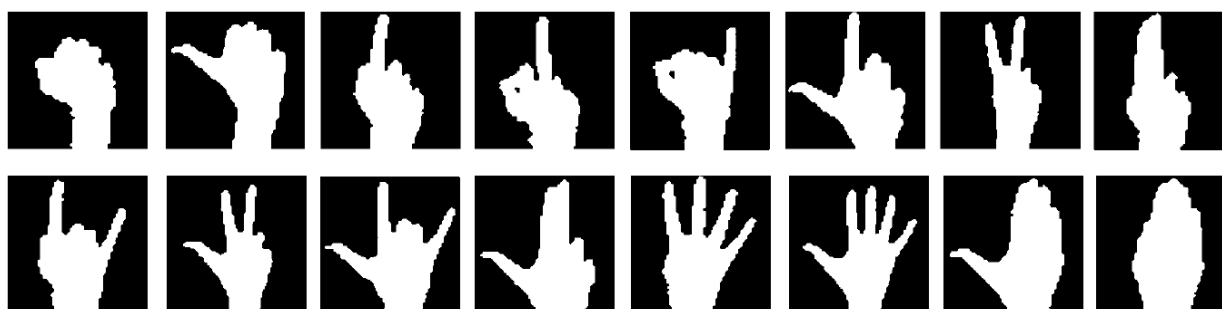


Bild 2: Vom Gestenerkennungssystem erkannte Handsymbole

Eine Faust (Bild 2, links oben) erhält also die Bezeichnung R00000. Die abgespreizten drei Finger (Bild 2, zweite Zeile, zweites von links) werden als R1_1_100 bezeichnet.

5 Interaktionstechniken

5.1 Grundlegende Interaktionstechniken

Aufgrund der Auslegung des Digitalen Lagetisches als Team-Arbeitsplatz musste bei der Entwicklung der Interaktionstechniken ein besonderes Augenmerk auf die bisherige Arbeitsweise eines Teams gelegt werden. Natürliche Gesten zur zwischenmenschlichen Kommunikation sollten beispielsweise nicht zur Interaktion verwendet werden, um keine Aktionen unbeabsichtigt auszulösen. Dennoch sollte eine möglichst intuitive Bedienung des Digitalen Lagetisches ermöglicht werden.

Da der Digitale Lagetisch in Teilen eine Kartenanwendung darstellt, orientieren sich die verschiedenen Metaphern an Interaktionen, die auch an einer Papierkarte auf einem Tisch durchgeführt werden können. Die grundlegenden Interaktionen

stellen dabei Verschieben und Drehen dar. Ergänzt werden diese durch Interaktionen wie Zoomen und das Auswählen von Symbolen auf der Karte, die bei einer Papierkarte technisch nicht möglich sind.

5.1.1 Shortcuts

Durch die Fähigkeiten der Gestenerkennung, Handsymbole zu erkennen, ist es möglich, ähnlich zu Tastatur-Shortcuts am PC, bestimmte Funktionen durch bestimmte Handsymbole direkt aufzurufen. Wird z.B. ein bestimmtes Handsymbol in einem bestimmten Bereich für kurze Zeit gezeigt und gehalten, öffnet sich ein Menü zum Weiterschalten des angezeigten Szenarios. Auf diese Weise lassen sich beliebige Funktionen mit Handsymbolen verknüpfen.

5.1.2 Verschieben

Das Verschieben der digitalen Karte ist an das Verschieben einer Papierkarte angelehnt. Durch einfaches Auflegen einer Hand mit gespreizten Fingern (Geste R1_1_1_1_1 oder L1_1_1_1_1) wie in Bild 3, oder durch das Auflegen der geschlossenen Hand (Geste R1_1111 oder L1_1111) und anschließendes Bewegen der Hand, kann die Karte in die entsprechende Richtung verschoben werden. Diese Technik ist für Links- sowie Rechtshänder gleichermaßen nutzbar.



Bild 3: Verschieben der Karte mit zwei unterschiedlichen Handsymbolen

Durch die Verwendung der recht speziellen Handgesten kann im Gegensatz zu einem Touchscreen das Verschieben der Karte durch unbeabsichtigtes Aufstützen oder Gestikulieren über dem Tisch verhindert werden. Die Handgeste kann während des Verschiebens aufgegeben werden, um – wie auch in Wu et al. (2006) gefordert – ein Verkrampfen zu verhindern.

5.1.3 Rotieren

Für das Rotieren der Karte wurde eine beidhändige Interaktionstechnik gewählt. Das Rotieren der Karte mit nur einer Hand wäre weit weniger genau möglich. Die Interaktionstechnik orientiert sich dabei an Erfahrungen aus Hinckley et al. (1997), indem die schwächere Hand einen „frame-of-reference“ bildet. Wird eine Hand mit gespreizten Fingern (Geste R1_1_1_1_1 oder L1_1_1_1_1) für einige Zeit ruhig auf den Tisch gelegt, wird an dieser Stelle ein Kompass eingeblendet

und der Rotations-Modus damit aktiviert (Bild 4). Das Handsymbol kann nach dem Auslösen aufgegeben werden. Mit dem Zeigefinger der zweiten Hand (Geste R01000 oder L01000) kann nun die Karte gedreht werden. Der Kompass stellt dabei den Rotationsmittelpunkt und gleichzeitig die Referenz für die Bewegungen der zweiten Hand dar. Welche Hand hierbei die Rotation und damit den Kompass auslöst, spielt keine Rolle. Der Abstand der Zeigegeste zum Rotationsmittelpunkt hat dabei Einfluss auf die Geschwindigkeit, mit der die Karte gedreht wird.

Bild 4 zeigt schematisch die Berechnung des Drehwinkels. Die rechte Hand wird dabei von Position a nach Position b bewegt. Dabei dreht sich die Karte mit dem Winkel β um den Mittelpunkt des Kompasses. Wird die Bewegung näher am Zentrum des Kompasses ausgeführt, genügt eine wesentlich kleinere Bewegung, um den gleichen Drehwinkel zu erreichen. Ein Nachteil dieser Technik ist die Überdeckung der Hände. Da die Gestenerkennung keine 3D-Rekonstruktion durchführt, können übereinanderliegende Hände nicht voneinander unterschieden werden. Wird also die Hand, die den Modus auslöste, von der zweiten Hand überdeckt, wird das Rotieren beendet.

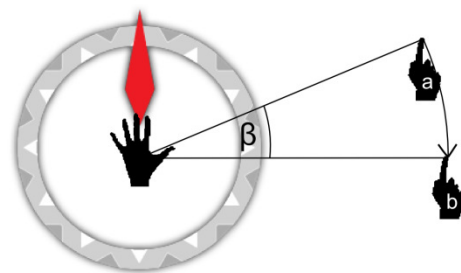
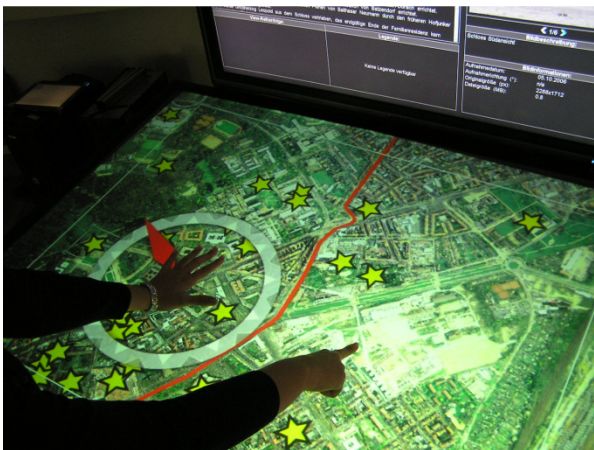


Bild 4: Rotieren der Karte mit einer zweihändigen Interaktionstechnik

Das Arbeiten an einer nicht nach Norden ausgerichteten Karte ist in einem professionellen Umfeld eher ungewöhnlich. Wurde die Orientierung der Karte aus irgendwelchen Gründen geändert, muss sie wieder nach Norden ausgerichtet werden. Man spricht bei diesem Vorgang von Einnorden. Dies ist zwar mit der vorgestellten Technik möglich, aber zeitaufwändig und ungenau. Um das Norden zu verbessern, wurde die bei der Gestenerkennung bestehende Möglichkeit der Shortcuts genutzt. Befindet sich das System im Rotations-Modus, d.h. der Kompass ist eingeblendet, kann mit einer Zwei-Finger-Geste (R01_100 oder L01_100) die Karte schnell wieder nach Norden ausgerichtet werden. Dazu werden die Finger der Geste für kurze Zeit auf die Tischplatte gelegt. Die Karte dreht sich dabei um den Mittelpunkt des Kompasses nach Norden zurück.

5.1.4 Zoomen

Die Interaktionstechnik zum Zoomen der Karte ist an die unter anderem von Appels iPhone bekannte Zoomgeste angelehnt. Auf Grund der Größe des Tisch-

displays werden zwei Hände mit jeweils einer Zwei-Finger-Geste (R01_100 und L01_100) verwendet (Bild 5, links). Das Handsymbol wurde gewählt, da es während natürlicher zwischenmenschlicher Interaktion am Tisch nicht vorkommt. Bewegen sich zwei Hände mit dieser Geste aufeinander zu, wird die Karte herausgezoomt, bewegen sie sich voneinander weg, zoomt die Karte hinein. Der Abstand der beiden Hände zueinander bestimmt dabei den Zoomfaktor.

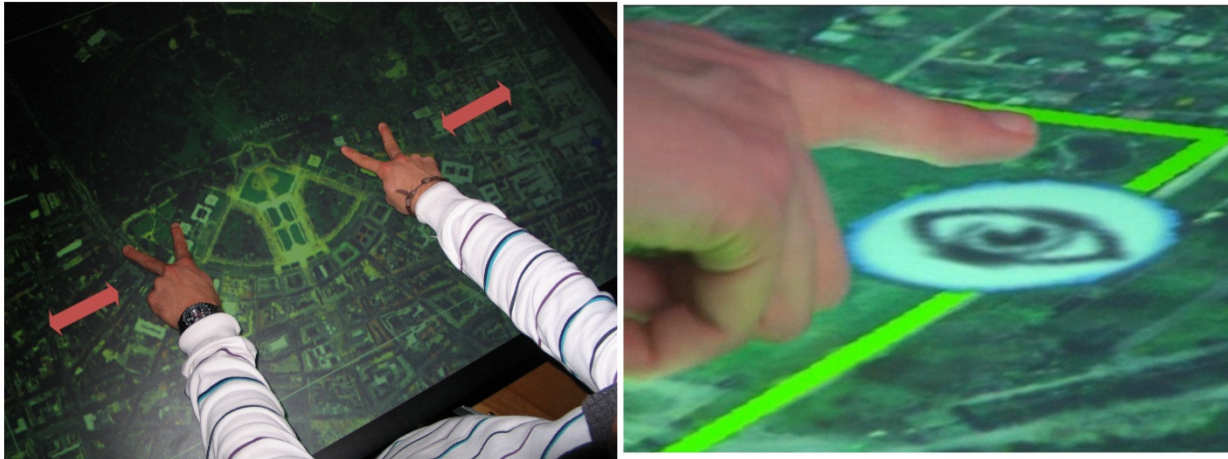


Bild 5: links: Zoomen mit Zwei-Finger-Geste rechts: Auswahl eines POIs

5.1.5 Auswahl

Zu den Grundfunktionen des Digitalen Lagetisches gehört das Anzeigen von verschiedenen Points-of-Interests auf der Karte, die entweder über das Fovea-Tablett oder über die Maus angeklickt werden können, um Zusatzinformationen abzurufen und auf dem Sekundärmonitor darzustellen. Die einfachste und natürlichste Möglichkeit ist das Auswählen mit dem Zeigefinger der Zeigegeste R01000 oder L01000 (Bild 5, rechts). Um die kleinen Symbole besser zu treffen, werden sie automatisch vergrößert, wenn die Zeigegeste in ihre Nähe kommt.

5.2 Displayübergreifende Interaktion

5.2.1 Anzeigen im Greifraum

Durch die geräteunabhängige Erfassung von Eingaben über die videobasierte Gestenerkennung ist es möglich, beliebige Geräte im Greifraum des Benutzers direkt zu bedienen. Beispielsweise lassen sich vorhandene Eingabemodalitäten von Geräten (z.B. Stifteingabe auf Tablett-PCs) ergänzen oder displayübergreifende Interaktionstechniken realisieren (Bader et al., 2010).

5.2.2 Entfernte Anzeigen

Für die Interaktion mit entfernten Anzeigen wurde zunächst ein im Folgenden beschriebener Ansatz verfolgt: Die Elemente die auf dem entfernten Display manipuliert werden sollen, werden in den Greifraum, d.h. auf das Tischdisplay, geholt und können hier direkt bedient werden. Am Digitalen Lagetisch kann dies genutzt werden, um bei einem längerem Text nach unten zu scrollen (Schieberegler) oder Bilder vor und zurück zu schalten (Schaltflächen). Diese Technik lässt sich einfach auf weitere GUI-Elemente übertragen.

In Bild 6 ist ein Menü aus einer Kombination aus Schaltflächen und einem Schieberegler dargestellt. Es dient zur Steuerung eines Video Players inklusive Zeitleiste und nächstes bzw. vorheriges Video auf dem entfernten vertikalen Display.



Bild 6: links: Menü auf horizontalem Display; rechts: Videoplayer auf vertikalem Display

Um die entsprechenden Elemente auf dem entfernten Display auszuwählen, wurden zwei Möglichkeiten in Betracht gezogen, welche beide die besonderen Fähigkeiten der Gestenerkennung ausnutzen. Die erste Möglichkeit ist die Nutzung von Shortcuts, um Bedienelemente auszuwählen. Durch Ausführen und Halten eines bestimmten Handsymbols (Shortcut) wird das entsprechende Element in den Greifraum geholt und kann manipuliert werden. Da hierbei jedem Bedienelement eine andere Handgeste zugeordnet werden muss und diese für Anfänger schwer zu merken sind, wird ein entsprechendes Piktogramm an jedem Element eingeblendet. Bei Experten, die bereits mit dem System vertraut sind, kann auf diese Piktogramme verzichtet werden.

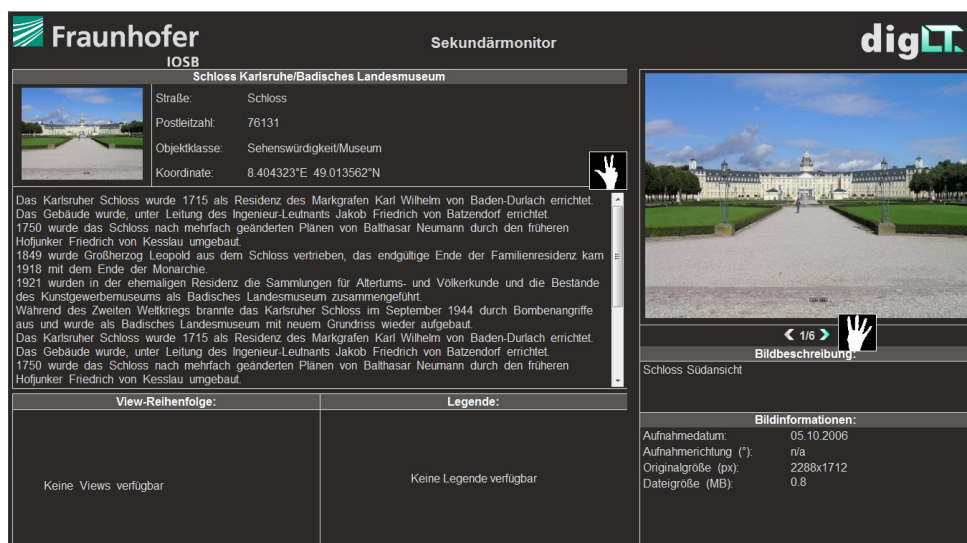


Bild 7: Selektion von zu manipulierenden Elementen auf dem vertikalen Display per Handsymbol

Bild 7 zeigt die Verwendung von R1_1_100 für den Schieberegler und R01_1_1_1 für die Schaltflächen zum Weiterschalten der Bilder. Die so ausgewählten Elemente werden direkt an der Position eingeblendet, an der die Handgeste über dem Tischdisplay ausgeführt wird.

Die zweite Möglichkeit, Elemente in den Greifraum zu holen, ist das direkte Zeigen darauf. Hierbei wird die Fähigkeit der Gestenerkennung zur Erfassung der Position von Hand und Fingerspitzen in 3D genutzt. Es ist sinnvoll, nur bestimmte Bereiche zur Auswahl anzubieten und diesen Bereich direkt zu markieren. Bild 8 zeigt den markierten Textbereich. Wird diese Zeigegeste für kurze Zeit gehalten, wird das entsprechende Bedienelement – in diesem Fall der Schieberegler – auf dem Tisch eingeblendet.



Bild 8: Auswahl von zu manipulierenden Elementen auf dem vertikalen Display per Zeigegeste

Die Bedienung der eingeblendeten Elemente erfolgt dabei wie gewohnt mit einer Art „Maus-Emulation“ bei welcher der Zeigefinger der Zeigegeste den Mauszeiger darstellt. Bei beiden hier vorgestellten Lösungen muss das entfernte Display nicht zwingend wie am Digitalen Lagetisch positioniert sein. Beim Zeigen auf das Display muss allerdings darauf geachtet werden, dass die Zeigegeste im Erfassungsbereich der Gestenerkennung durchgeführt wird.

6 Benutzerstudie

Die in Abschnitt 5 vorgestellten Interaktionstechniken wurden in einer Benutzerstudie evaluiert und durch einen Vergleich mit etablierten Techniken bewertet. Die durchzuführende Aufgabe orientierte sich dabei stark an bestehenden Arbeitsabläufen am Digitalen Lagetisch. Das Drehen war kein Teil der Studie, da diese Technik eher selten am Digitalen Lagetisch eingesetzt wird.

6.1 Aufgabenstellung und Ablauf

Die zu lösende Aufgabe wurde so entworfen, dass ein ständiger Wechsel zwischen der Bedienung des Tischdisplays und des Sekundärmonitors notwendig war. Auf dem Tisch mussten hierzu immer drei zusammengehörende Symbole eines Typs in jeweils drei Farben ausgewählt werden. Die Symbole wurden dabei auf einer Karte verteilt. Auf dem Sekundärmonitor musste immer ein Textfeld nach unten bis zu einem bestimmten Bereich gescrollt werden, um lesen zu können, welches Symbol als nächstes auszuwählen ist.

Um die Interaktionstechniken vergleichen zu können, wurde die Aufgabe mit drei verschiedenen Eingabemethoden durchgeführt. Ein Mal mit der Maus, rein mit der Gestenerkennung und schließlich mit einer Kombination aus Maus und Gestenerkennung, bei der das Tischdisplay mit der Gestenerkennung und der Sekundärmonitor mit der Maus bedient werden musste. Bei der ersten Technik war das System als Doppelmonitorsystem konfiguriert, und der Mauszeiger musste zwischen horizontaler und vertikaler Anzeige hin und her verschoben werden. Zur Manipulation der Darstellung auf dem Tisch wurden bei den beiden letzteren Ansätzen die in Abschnitt 5.1 beschriebenen Techniken verwendet. Bei der reinen Gestensteuerung wurde die Zeigegeste verwendet, um einen Scrollbalken in den Greifraum zu holen.

Um die Interaktionen mit Maus und Gestenerkennung aneinander anzugleichen, wurde das Musrad für das Scrollen am Sekundärmonitor deaktiviert. Dadurch musste bei jeder Interaktion der Schieberegler selbst bewegt werden.

Um einen Einfluss von Lerneffekten auf das Ergebnis der Studie zu verhindern, wurden die 24 Versuchspersonen in verschiedene Gruppen unterteilt. Jede Gruppe fand eine unterschiedliche Reihenfolge der Durchgänge (Maus, Maus + Gestenerkennung, reine Gestenerkennung) vor, so dass alle Permutationen abgedeckt wurden. Keine der Versuchspersonen hatte praktische Erfahrungen mit dem Digitalen Lagetisch. Um allen Personen die gleichen Voraussetzungen zu geben, wurde eine schriftliche Einführung verfasst und vor der Durchführung ausgehändigt.

Bei jedem Durchlauf wurden vom System automatisch Zeiten erfasst und gespeichert sowie eine subjektive Bewertung durch die Testperson mit Hilfe eines Fragebogens vorgenommen. Nach Abschluss aller Durchgänge folgte ein weiterer Fragebogen mit allgemeinen Fragen und der Möglichkeit, Anmerkungen zu jeder Technik zu machen.

6.2 Auswertung

Für die Bewertung der einzelnen Durchgänge wurden sowohl objektive als auch subjektive Kriterien betrachtet.

Bild 9 zeigt die Verteilung bei der Frage nach der besten Eingabemethode. Der hohe Anteil der Personen (48%), welche die kombinierte Eingabemethode bevorzugten, ergab sich durch einen kleinen Vorteil in der Aufgabenstellung. Da die

Maus immer auf dem Tisch lag, war ein sehr schneller Wechsel von der Gestensteuerung auf die Maus möglich. Ein weiterer Vorteil war die Positionierung des Mauscursors auf dem Sekundärmonitor. Da die Maus nur zum Scrollen verwendet wurde, musste die Position des Cursors immer nur geringfügig geändert werden.

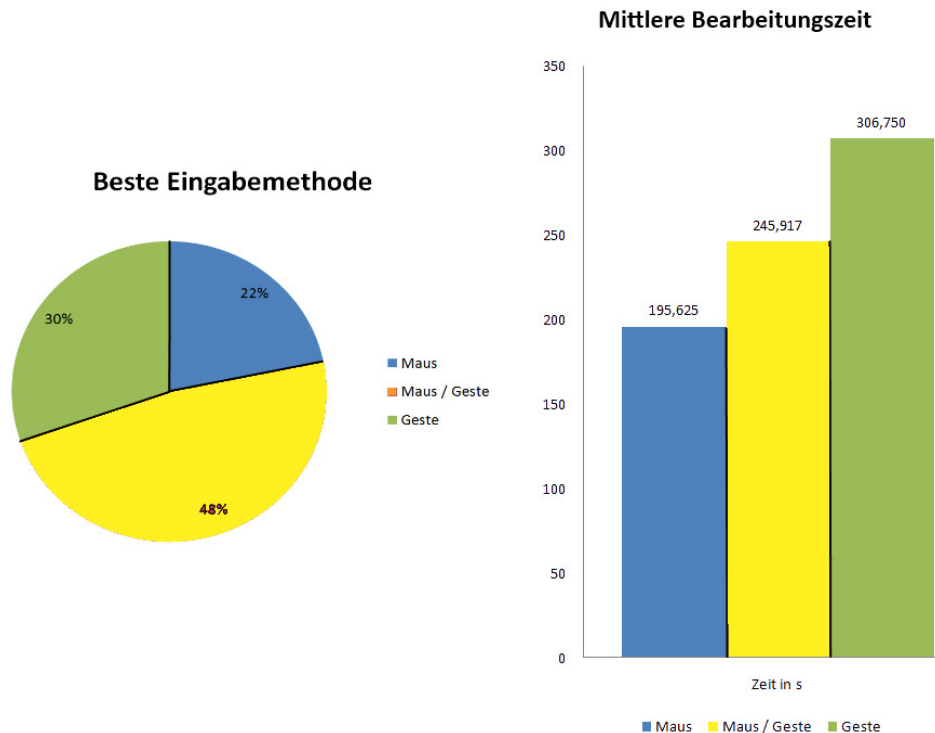


Bild 9: links: Bewertung der Ansätze als beste Interaktionstechnik; rechts: Mittlere Bearbeitungszeit mit einzelnen Interaktionstechniken

Insgesamt zeigte sich, dass 78% der Versuchspersonen eine Interaktion mit der Gestenerkennung gegenüber einer reinen Maussteuerung bevorzugen.

Die rechte Seite von Bild 9 stellt die Zeit dar, die zum Lösen der gesamten Aufgabe benötigt wurde. Hier zeigt sich ein deutlicher Geschwindigkeitsvorteil der Maussteuerung, der auf verschiedene Ursachen zurückzuführen ist. Da sich die Maus als Standard-Eingabegerät etabliert hat, ist es nahezu unmöglich, Versuchspersonen zu finden, die noch keine Erfahrung mit der Maus besitzen. Alle Probanden stammen aus dem IT-Umfeld und können daher als Experten bezeichnet werden. Eine weitere Ursache, die erheblichen Einfluss auf die Bearbeitungszeit hatte, war die indirekte Bedienung über die Maus. Während beim Verschieben der Karte mit der Hand große Strecken zurückgelegt werden mussten, genügte mit der Maus eine sehr kleine Bewegung, um das gleiche Ergebnis zu erzielen. Die technisch perfekte Arbeitsweise und die Möglichkeit die Symbole pixelgenau auszuwählen, waren weitere Umstände, die der Maus einen Geschwindigkeitsvorteil verschafften. Ein Problem der Zeigegeste war die Zeit, bis der Scrollbalken auf dem Tisch angezeigt wurde. Dadurch wurde die Interaktion verlangsamt. Diese Zeit kann durch die Verwendung einer speziellen „Auslöse“-Geste verbessert werden.

7 Fazit

Trotz des Geschwindigkeitsvorteils der Maus wurde die Arbeit mit der Gestenerkennung der Arbeit mit der Maus von den Versuchspersonen vorgezogen. Die Gestenerkennung bietet weitere Vorteile, die nicht in die Benutzerstudie einbezogen wurden und die über die Fähigkeiten von Touchscreens hinaus gehen (Tab. 1). Dazu zählt unter anderem die Möglichkeit der Unterstützung der Interaktion in Multi-User-Umgebungen sowie die Trennung von Gestik zur zwischenmenschlichen Kommunikation von solcher zur Mensch-Maschine Interaktion.

Tab. 1: Vor- und Nachteile der verschiedenen Techniken

| | Vorteile | Nachteile |
|--------------|---|--|
| Maus | <ul style="list-style-type: none"> • etabliert • Benutzer geübt • hohe Genauigkeit • technisch robust • schnell | <ul style="list-style-type: none"> • indirekt • zusätzliches Gerät • nicht für komplexe Multi-Display-Umgebungen geeignet • nicht Multi-User-fähig |
| Geste | <ul style="list-style-type: none"> • intuitiv • kein zusätzliches Gerät • displayunabhängige Erfassung • Multi-User-fähig • Shortcuts • Kein Berührung notwendig • Display / System übergreifend | <ul style="list-style-type: none"> • geringere Genauigkeit • geringere Robustheit |
| Maus / Geste | <ul style="list-style-type: none"> • Kombination Vorteile Geste + Maus | <ul style="list-style-type: none"> • wechsel zwischen Eingabegeräten notwendig • nicht Multi-User-fähig |

Die längere Bearbeitungszeit mit der gestenbasierten Eingabe lässt sich durch weniger Übung der Benutzer, die etwas geringere technische Robustheit, aber auch durch größere räumliche Bewegungen, welche vom Benutzer während der Interaktion durchzuführen sind, begründen. Letztere führen auch zu größerer körperlicher Ermüdung.

In zukünftigen Arbeiten soll untersucht werden, wie sich dies unter anderem durch Kombination mit anderen Modalitäten wie z.B. Blick verbessern lässt.

Literatur

Bader, T., Heck, A. & Beyerer, J. (2010). Lift-and-Drop: Crossing Boundaries in a Multi-Display Environment by Airlift. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, pp. 139-146.

- Bader, T., Meissner, A. & Tscherney, R. (2008). Digital Map Table with Fovea-Tablet®: Smart Furniture for Emergency Operation Centers. In *Proceedings of the 5th International Conference on Information Systems for Crisis Response and Management*, S. 679-688, Springer, 2008.
- Bader, T., Räßle, R. & Beyerer, J. (2009). Fast Invariant Contour-Based Classification of Hand Symbols for HCI. In *Proceedings of Computer Analysis of Images and Patterns (CAIP) 2009*, pp 689-696, Springer.
- Guiard, Y. (1987). Asymmetric division of labor in human skilled bimanual action: the kinematic chain as a model. *Journal of motor behavior*, 19(4), 486-517.
- Hinckley, K., Pausch, R., Proffitt, D., Patten, J. & Kassell, N. (1997). Cooperative bimanual action. In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '97)*, pp 27-34.
- Kabbash, P., Buxton, W. & Sellen, A. (1994). Two-handed input in a compound task. In *Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence*. New York: ACM.
- Peters, M. (1980). Why the preferred hand taps more quickly than the non-preferred hand: Three experiments on handedness. *Canadian Journal of Psychology/Revue canadienne de psychologie*.
- Utsumi, A. & Ohya, J. (1999). Multiple-hand-gesture tracking using multiple cameras. In *Proceedings 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, Band 1. IEEE Comput. Soc, pp. 473-478.
- Vogel, D. & Balakrishnan, R. (2005). Distant freehand pointing and clicking on very large, high resolution displays. In *Proceedings of the 18th annual ACM symposium on User interface software and technology - UIST '05*, pp. 33.
- Wobbrock, J.O., Morris, M.R. & Wilson, A.D. (2009). User-defined gestures for surface computing. In *Proceedings of the 27th international conference on Human factors in computing systems (CHI '09)*, pp. 1083.
- Wu, M., Shen, C., Ryall, K., Forlines, C. & Balakrishnan, R. (2006). Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces. In *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems*, 2006. TableTop 2006 Ieee, 2006.

Autoren

S. Maier, M.Sc. Fraunhofer-Institut für Optronik, Systemtechnik und
Dipl.-Inform. T. Bader Bildauswertung IOSB
Karlsruhe

Kontakt:

sebastian.maier@iosb.fraunhofer.de