

Gesten als neue Eingabemethode für Mobile Geräte

Thomas Kraus

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
thomas.kraus@stud.ifi.lmu.de

Zusammenfassung Mobile Geräte gewinnen immer mehr Bedeutung im Bereich der Informationsverarbeitung und Kommunikation. Aktuelle Technologien ermöglichen mittlerweile komplexe Anwendungen auf Geräten wie Mobiltelefonen oder PDAs. Die Bildschirme sind zwar hochauflösend aber viel kleiner als bei Notebooks oder PCs. Ebenfalls fehlen auf Grund der Gerätegröße und des mobilen Einsatzes Eingabegeräte wie Maus oder Tastatur. Momentan werden solche Geräte durch wenige Tasten oder berührungssensitive Bildschirme bedient. Jedoch benötigt diese Art der Interaktion eine ständige visuelle Aufmerksamkeit des Benutzers. In dieser Arbeit werden einige innovative Eingabemethoden durch Gesten beschrieben. Gesten sind zeichenhafte Bewegungen verschiedenster Körperteile und dienen zur nonverbalen Kommunikation [1]. Wichtiger Bestandteil solcher Eingabesysteme ist die Erkennung der Gesten. Dazu werden unterschiedlichste Sensoren und technische Hilfsmittel verwendet. Einige Beispiele hierfür sowie Prototypen werden im Folgenden erläutert und deren Nutzen für den Anwender dargestellt.

1 Einleitung

Mobiltelefone, PDAs (Personal Digital Assistant) und MP3-Musikspieler gehören mittlerweile zum ständigen Begleiter vieler Menschen. Gründe hierfür sind der wachsende Fortschritt im Bereich Prozessorleistung, persistenter sowie flüchtiger Speicher und die bessere Energieversorgung mobiler Endgeräte. Daher werden auch immer komplexere Anwendungen für mobile Endgeräte möglich. Jedoch sind die Eingabetechologien, wie sie im PC-Bereich verwendet werden aus nachfolgenden Gründen ungeeignet. Diese Geräte verfügen über sehr kleine Bildschirme und kaum Platz für eine Tastatur, Maus oder Ähnliches. Deshalb verwendet man häufig berührungssensitive Bildschirme, die mit einem Plastikstift (engl.: Stylus) bedient werden können. Dadurch erreicht man eine variable Tastenerzeugung und -belegung. Auf dem mobilen Gerät können Bedienelemente grafisch flexibel auf dem Bildschirm erzeugt werden, welche dann durch Anklicken ausgewählt werden können und so für jede Anwendung unterschiedliche Reaktionen ergeben. Jedoch braucht man bei dieser Methode immer beide Hände, eine um das Gerät zu halten und eine um mit dem Stift oder einem anderen

Hilfsmittel die Eingabe zu tätigen. Diese Tatsache fordert außerdem vom Anwender die volle visuelle Aufmerksamkeit für das Gerät. Das Gleiche gilt für die Bedienung von Endgeräten durch physische Knöpfe, wie bei einem Mobiltelefon. Aber genau diese starke visuelle Aufmerksamkeit auf das Gerät kann in manchen Situationen sehr unerwünscht und ablenkend sein, teilweise sogar ein Sicherheitsrisiko für den Benutzer darstellen. Man denke dabei z. B. an die Navigation eines Fahrzeuges. Dabei können schon kurze visuelle Ablenkungen schwere Folgen für den Fahrer bedeuten. Ein aktuelles Eingabegerät im Auto wäre das Bedienkonzept iDrive von BMW [2]. Dieses System der Bedienung der Fahrzeugelektronik ist zwar sehr benutzerfreundlich konzipiert, jedoch auch hier kann die nötige visuelle Aufmerksamkeit auf den Hauptsteuerknopf und vor allem den Bildschirm in der Mitte der Fahrzeugkonsole während der Fahrt lebensbedrohliche Folgen haben. Dies wird deutlich, wenn man bedenkt, dass ein Fahrer bei einer Geschwindigkeit von 100 Stundenkilometern eine Strecke von 55 Metern zurücklegt, falls dieser nur zwei Sekunden auf den Bildschirm blickt.

Außerdem kommt es durch den mobilen Einsatz der Geräte zu häufigeren Eingabefehlern, da viele Tasten eng zusammen liegen müssen, ein teilweise nur kurzer Augenkontakt möglich ist und man sich dadurch leicht vertippt. Auch das Durchblättern von Dateien (Bilder, Text, Web, Landkarten usw.) ist momentan noch sehr umständlich realisiert. Bildlaufleisten nehmen auf den bereits kleinen Bildschirmen viel Platz ein und werden häufig sowohl in vertikaler als auch horizontaler Richtung benötigt. Außerdem muss man bedenken, dass wie bereits erwähnt, Anwender häufig in Bewegung sind und dem Bildschirm nur wenig Aufmerksamkeit schenken können und daher die Benutzung eines Stylus zum Durchblättern oder Auswählen sehr unvorteilhaft und fehleranfällig ist. Ein gutes Beispiel hierfür wäre ein mobiles Navigationssystem auf einem PDA mit einer Ansicht eines durchsuchbaren Stadtplans. Damit ist gemeint, dass man den Bildschirm wie eine Ausschnittsansicht benutzt, vergleichbar einer Lupe als Schaufenster über dem Stadtplan.

Musikwiedergabegeräte hingegen sind heute noch mit physischen Tasten belegt. Teilweise sind diese nur steuerbar, wenn man Sichtkontakt zum Gerät bzw. den Tasten hat. Stellt man sich vor, dass die meisten Menschen Musik hören wollen während sie andere Aktivitäten ausführen, wird dieses Prinzip der Tastensteuerung nutzlos. Ein Jogger möchte nicht stehen bleiben müssen, um das nächste Lied auszuwählen oder die Lautstärke zu regulieren. Außerdem sollen Geräte in Zukunft kleiner werden (z. B. Radio im Ohrstöpsel integriert). Dann fallen momentan gültige Eingabemethoden weg, wie z. B. Knöpfe oder berührungssensitive Bildschirme, da man diese nicht auch noch verkleinern kann, da sie dann unbenutzbar wären für den Menschen. Grundsätzlich kann man also festhalten, dass die Eingabe mit den aktuellen Methoden in Bewegungssituationen sehr viel anstrengender, fehleranfälliger und Geräte grundsätzlich schwerer zu bedienen sind als im Stillstand des Benutzers. Außerdem sind aktuelle Eingabemethoden auf Grund ihrer Beschaffenheit, wie z. B. der Größe des mobilen Geräts, nicht mehr anwendbar.

Innovative Eingabemethoden sind also wichtig für die neue Generation der Eingabegeräte speziell im Bereich mobiler Endgeräte. Metaphern aus anderen Bereichen, wie z. B. Arbeitsplatz-PCs, können nicht weiter verwendet werden, da diese unter anderen Voraussetzungen entstanden sind (andere Ein- und Ausgabemethoden wie z. B. PC-Tastatur und Bildschirm, andere Benutzungssituationen). Fukumoto behauptete in [3], dass es drei Faktoren gäbe für tragbare Geräte und ihre Eingabesysteme: "Tragbarkeit", "Benutzerfreundlichkeit" und "Konstanz". Es solle also komfortabel zu tragen sein, leicht im Gewicht sein, trotz der kleinen Größe einfach zu bedienen sein und ständig bereit sein, um mit dem Anwender zu interagieren. Letzteres ist auch unter dem Begriff der Kontinuität von Interaktion bekannt. Da man mobile Geräte ständig mit sich trägt ist eine kontinuierliche, nebenläufige Kommunikation erwünscht. Man möchte also einen ständigen, indirekten Kontakt zum Endgerät, damit man spontan darauf zugreifen kann und eine Steuerungsphase vom Gerät erkannt wird.

Ein wichtiger Nebenansatz um aufgeführte Veränderungen zu bewältigen, ist der Einsatz von Gesten und Zeichen zur Steuerung mobiler Endgeräte. Als Geste (lat. gerere = tragen, ausführen) bezeichnet man zeichenhafte Bewegungen bestimmter Körperteile (v. a. Hand und Kopf) zum Zwecke der nonverbalen Kommunikation [1]. Um welche Art von Gesten und Zeichen es sich handelt wird im folgenden Text behandelt und zusammengefasst. Dabei wird in diesem Rahmen nicht auf die Eingabe von Text eingegangen, sondern nur auf die Steuerung von Funktionen und Anwendungen des mobilen Geräts. Bezüglich der Texteingabe durch Gesten möchte ich auf die Arbeit von Mihail Tsvyatkov hinweisen mit dem Titel "Texteingabe für mobile Endgeräte" [4]. Die vorgestellten Gesten beschränken sich auf Berührungspunkte und -bewegungen mit Fingern oder Kopf, dem Kippen von Geräten, dem Messen von Signalen (optisch sowie akustisch), die durch Gesten entstehen.

2 Eingabemöglichkeiten durch Gesten

2.1 Beschleunigungsbasierende Gesten

In folgendem Abschnitt wird ein System vorgestellt, das für die Gestenerkennung Beschleunigungssensoren (engl.: Accelerometer) verwendet. Dabei werden Neigung oder Beschleunigung eines Geräts gemessen, um die Lage oder Bewegungsbahnen festzustellen. Diese Sensoren werden dabei in Richtung der verschiedenen Achsen am Objekt angebracht und die Ergebnisparameter im Millisekundenbereich wiederholt gemessen und ausgewertet. Damit können sowohl Stillstand als auch Bewegungsbahnen bzw. Veränderungen der Lage, wie z. B. Schütteln, Neigen oder Drehen ermittelt werden. Joanna Lumsden und Stephen Brewster erforschten Techniken zur Interaktion mit mobilen Geräten mit minimaler visueller Aufmerksamkeit, um den Wert bzw. Nutzen solcher neuen Eingabemethoden zu ermitteln [5]. Sie bauten daher einen Prototypen, der darauf abzielte, Kopfgesten zu erkennen. Dabei wurde ein Kopfhörer verwendet, auf dessen Trägerbügel ein Beschleunigungsmesser angebracht war. Dieser Sensor bestimmte die Orientierung und Bewegung des Kopfes. Typische Gesten des Kopfes sind Nicken

oder Schütteln. Diese Art der Bedienung wird bereits zur Steuerung von Rollstühlen durch körperlich behinderte Menschen verwendet, die ihre Arme nicht mehr benutzen können. Lumsden und Brewster sahen Kopfgesten auch als eine Bereicherung für die Bedienung mobiler Endgeräte. Es wurde für die Steuerung mit dem Kopf eine dreidimensionale Geräuschumgebung um den Kopf des Benutzers geschaffen. Der Kopfhörer simulierte diese Geräuschkulisse als Menüauswahl durch verschiedene Töne im Raum um den Kopf herum verteilt. Es sollte jedoch vermieden werden, dass der Anwender durch diese akustischen Menüpunkte zu sehr abgelenkt würde bzw. die Geräusche aus seiner momentanen Umgebung gestört oder sogar geschluckt würden. Um eine Auswahl zu ermöglichen wurde ein Mechanismus zur Erkennung von Kopfnicken implementiert. Dieser musste sehr stabil sein gegenüber Täuschungen durch die normalen Kopfbewegungen, z. B. beim Gehen. Der Algorithmus sah wie folgt aus: Alle 200 ms wurde eine Messung durchgeführt über die momentane Orientierung des Kopfes. Falls eine Veränderung erfolgte von mehr als sieben Grad vom letzten Ruhepunkt des Kopfes aus gesehen, wurde der Beginn eines Nickens erkannt. Falls nun innerhalb der nächsten 600 ms nochmals eine Veränderung von weiteren sieben Grad in die entgegengesetzte Richtung gemessen werden konnte, wurde ein komplettes Nicken registriert. So wurde vermieden, dass ein einfaches auf den Boden Gucken sofort als Nickbewegung eingestuft wurde.

Die Geräuschkulisse wurde in vier Kategorien eingeteilt: Wetter, Nachrichten, Sport und Verkehr. Es wurden dabei Geräusche gewählt, mit denen man diese Themen in Assoziation bringen würde, wie z. B. Regentropfen, Gewitter und Vogelgezwitscher für den Bereich Wetter. Die Geräusche wurden um den Kopf auf einer Ebene angeordnet (siehe Abbildung 1). Es wurden drei unter-

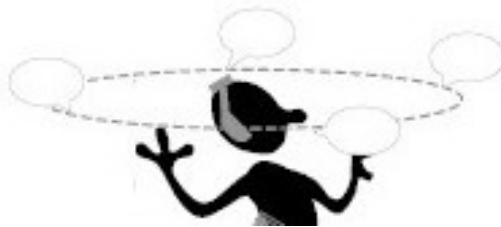


Abbildung 1. Geräuschobjekte werden um den Kopf auf einer Ebene platziert, um ein Auswahlmenü zur Steuerung mobiler Geräte durch Nicken zu simulieren. [5]

schiedliche Optionen der Geräuschumgebung entwickelt. Bei der egozentrischen Anordnung wurden die vier Objekte jeweils alle 90 Grad um den Kopf platziert. Die einzelnen Geräusche bzw. Menüpunkte wurden immer zwei Sekunden lang nacheinander abgespielt. Eine Auswahl fand statt, indem man in die gewünschte Richtung nickte. Diese einfache Methode hatte jedoch den Nachteil, dass es häufig nötig war, nach hinten zu nicken, was als sehr unnatürlich wahrgenommen wurde und anstrengend war. Bei der exozentrischen, konstanten Variante

wurden die Objekte in einer Linie vor dem Benutzer angereiht. Man wählte ein Geräusch aus, indem man sich mit dem Kopf leicht hindrehte und dann nickte. Dabei verwendete man nur das natürlichere Vorwärtsnicken. Die Geräusche wurden dabei alle gleichzeitig und konstant abgespielt. Nur das aktuelle Objekt genau vor einem wurde lauter wahrgenommen, damit man wusste, wo man sich im Menü momentan befindet. Bei der periodischen Variante ertönten die Geräuschobjekte wieder hintereinander. Dadurch wurde eine zu hohe simultane Geräuschbelastung vermieden, jedoch war hier eine schnelle Auswahl schwieriger. Man musste teilweise warten bis das gewünschte Symbol ertönte und man wusste, wohin man nicken musste. Durch diese Kombination von Kopfgesten und einem Auswahlmenü durch Geräusche konnte man eine neue Eingabemethode schaffen, die nur minimale visuelle Aufmerksamkeit benötigte zur Steuerung mobiler Geräte. Weitere interessante Eingabesysteme und Anwendungsszenarien mit Hilfe von Beschleunigungssensoren werden in den Arbeiten [6], [7], [8] und [9] beschrieben.

2.2 Optische Erfassung von Gesten

Hierbei geht es darum, Gesten optisch z. B. durch Kameras oder Infrarotsensoren zu erfassen. Man versucht verschiedene Gesten eindeutig zu unterscheiden, um so Abbildungen zu schaffen von gewissen menschlichen Aktionen auf Ereignisse am Endgerät (wie z. B. eine Menüauswahl). Als ein Beispiel kann man hier das System FreeDigiter [10] aufführen. Dieses System benutzte einfache Fingergesten zur Steuerung von mobilen Geräten, wie z. B. eines MP3-Musikspielers. Bei diesem Ansatz ging es u. a. darum, so wenig visuelle Aufmerksamkeit wie möglich vom Anwender zu fordern. Das Grundprinzip bestand darin, dass man ein bis vier Finger (Hand ohne Daumeneinsatz) vor oder zurück durch einen Sensor bewegte. Dieser Sensor war in der Lage durch Ausstrahlen und Messen des reflektierten Infrarotlichtes festzustellen, wieviele Finger bei einer Vorbeibewegung in einer Reichweite von 10 bis 60 cm ausgestreckt waren. Durch die Lücken zwischen den ausgestreckten Fingern entstanden Pausen bei der Reflektion auf den Sensor und die Fingerzahl konnte ermittelt werden. Dabei lag die frequentielle Messrate der Lichtreflektion bei 6,4 ms, was den Vorteil hatte, dass die Abstände zwischen den Fingern nicht besonders groß sein mussten, um fehlerfrei zu funktionieren. Es ergab also einen Impuls pro Finger, falls etwas Abstand zwischen den einzelnen Fingern gelassen wurde. Ein Beispiel für eine Gesteneingabe ist in Abbildung 2 zu erkennen. Man benutzte dabei maximal vier Finger, da der zusätzliche Einsatz des Daumens zu Nachteilen führte, da dieser nicht ganz parallel zu den Anderen liegt und viel kleiner ist. Die Bewegung der Hand hätte dann kreisförmiger ausfallen müssen und hätte den Anwender im Handgelenk zu sehr angestrengt und Fehler verursacht. Um mehr als vier Zustände zu schaffen und daher auch mehr als vier unterschiedliche Eingabemöglichkeiten zu geben war es nötig, in einer fließenden mehrmaligen Hin- und Herbewegung der Hand durch den Sensor eine Summe über seine Eingabe zu bilden. Dabei ist egal wie die Summe für eine Geste gebildet wurde. Für die Geste '12' als Beispiel konnte man drei mal vier Finger oder auch vier mal drei Finger benutzen. Falls circa



Abbildung 2. Der Benutzer von FreeDigiter bewegt einmal zwei Finger durch den Sensor am Ohr vorbei (Bewegung siehe Pfeil im Bild) für die Eingabe von Option '2'. Dabei ist keine visuelle Aufmerksamkeit für die Eingabe nötig. [10]

nach einer Sekunde kein weiteres Signal in Form eines vorbeigeführten Fingers am Sensor einging, wurde die Geste für das System abgeschlossen und auf eine bestimmte Reaktion abgebildet. Dabei wurde bei dem Prototypen auf einen Microcontroller zurückgegriffen, der die Messwerte des Sensors einlas und via eines Bluetooth-Moduls paketweise an einen Linux-Rechner sendete, der diese Signale dann verarbeitete und auswertete.

Das vorgestellte System könnte in etlichen mobilen Geräten zum Einsatz kommen, welchen eine kontaktlose Interaktion durch einfache Kommandos ausreicht. In dem beschriebenen Projekt FreeDigiter [10] wurde zu Testzwecken an einem MP3-Wiedergabegerät und einer Mobiltelefonanwendung experimentiert. Bei dem Musikspieler wurde der Titel auf dem Speicher abgespielt, entsprechend der Anzahl der Finger, die ein Anwender durch den Sensor bewegte. Bei dem Mobiltelefonansatz gab es drei Modi: "Anruf annehmen" (ein Finger), "Schnellwahl-Menü" (zwei Finger) und "wähle Telefonnummer" (drei Finger). Danach befand man sich in einem Untermenü, von wo aus man neue Eingaben tätigen musste, um eine gewünschte Aktion zu erreichen. War man im Schnellwahl-Menü gelandet, dann konnte man nun eine neue Geste starten. Um den Eintrag '6' als Beispiel in einer Schnellwahlliste auszuwählen, wurde durch zwei mal drei Finger eine Geste aktiviert und der entsprechende Teilnehmer wurde automatisch angerufen. Da Geräte immer kleiner werden ist diese Art der Steuerung sehr innovativ. So könnten Eingabesysteme direkt in z. B. Ohrstöpsel integriert werden, in denen für Knöpfe oder Bildschirme kein Platz wäre.

Bei der MP3-Spieleranwendung mussten die Teilnehmer bei einer Versuchsdauer von 100 Titeln nacheinander zufällig gewählte Lieder von '1' bis '20' auswählen. Dabei wurden 95 % der Gesten korrekt erkannt. Es wurden also fast alle gewollten bzw. ausgeführten Fingerbewegungen durch den FreeDigiter [10] richtig interpretiert. Bei dem Mobiltelefonsystem benötigte man nur Zahlen von '0' bis '9' ('0' dargestellt durch die Handfläche) zur vorherigen Auswahl des oben beschriebenen Menüsystems und zum Wählen einer Telefonnummer. Dabei wurde eine Erkennungsquote von 99 % bei 150 durchgeführten Testgesten ermittelt. Dieses Ergebnis zeigte, dass es für den Benutzer einfacher war, kleine-

re Summen mit den Fingergesten zu bilden und so das Eingabesystem stabiler würde. Eine weitere Erkenntnis war, dass das System völlig robust war gegenüber unterschiedlichen Lichtverhältnissen. Ein Problem ergab sich jedoch beim Einsatz in der realen Welt. Dort wurde festgestellt, dass durch Gegenstände wie Türrahmen, Wände und vorbeigehende Personen positive Fehlergesten ausgelöst wurden. Gesten, die zwar für das System vorhanden waren, jedoch vom Benutzer nicht gewollt. Dies könnte vermieden werden durch Verringerung der Sensorreichweite oder durch eine Funktion im Gerät zum Ein- und Ausschalten des Sensors. Das wurde dann durch einen Beschleunigungssensor (siehe Punkt 2.1) realisiert, der ein lebhaftes Nicken des Kopfes nach rechts wahrnahm und dies als das Einleiten einer Geste erkannte.

Ein anderer optischer Erkennungsansatz von Gesten wurde in dem Paper "The WristCam as Input Device" [11] beschrieben. Hierbei spielte auch wieder die Anordnung der Finger eine Rolle. Es befand sich eine Kamera unter dem Handgelenk des Nutzers, welche die nach unten hängenden Finger im Blickfeld hatte. Zu erkennen ist diese Konstruktion auf Abbildung 3. Ziel der Arbeit war

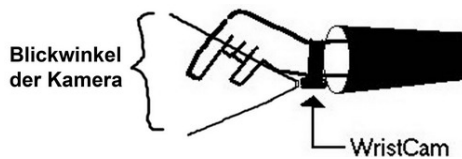


Abbildung 3. Handgelenksposition der Kamera mit angezeichnetem Blickwinkel. Dabei wurden unterschiedliche Fingerpositionen erkannt und auf gewisse Aktionen am Endgerät abgebildet. [11]

es, eine Interaktionsmöglichkeit zu schaffen für tragbare Endgeräte mit wenigen einfachen Gesten. Es wurden keine speziellen Anwendungen getestet, sondern nur allgemein das Eingabesystem WristCam untersucht. Wieder wurde hier eine Abbildung von einzelnen Zeichen auf gewollte Reaktionen eines beliebigen tragbaren Geräts entworfen. Es sollten die Eigenschaften Robustheit, Bedienfreundlichkeit und Geschwindigkeit erfüllt werden. Außerdem war bei dieser Methode von Vorteil, dass sie für Außenstehende nahezu unsichtbar war und der Benutzer bei der Eingabe bzw. Nicht-Eingabe von Gesten immer beide Hände frei hatte. Es wurden mit dieser Kamera bestimmte Bewegungen und Gesten erkannt, indem die Finger von dem variablen Hintergrund und der unterschiedlichen Beleuchtung getrennt wurden. Ein mobiler Nutzer möchte vermutlich nicht seine Hand bei der Eingabe still halten oder auf Hintergrund- oder Lichtverhältnisse achten müssen (Robustheit). Für ein Eingabegerät ist es sehr wichtig, kurze Antwortzeiten zu bekommen, da es sonst vom Nutzer nicht akzeptiert würde (Geschwindigkeit). Benutzerfreundlichkeit sollte bei jeder Entwicklung eines Eingabegeräts eine hohe Priorität haben. Jedoch war in diesem Fall speziell gemeint, dass die Gesten für den Benutzer einfach zu realisieren und vor allem für ihn oder sie angenehm

in der Durchführung waren. Finger- und Handbewegungen können teilweise sehr anstrengend und unangenehm erscheinen und daher ist die Auswahl der vordefinierten Zeichen (siehe Abbildung 4) sehr wichtig (Benutzerfreundlichkeit). Dabei startet man eine Eingabe mit der sog. Ruheposition (engl.: Rest Position)

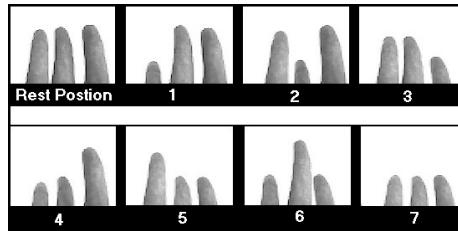


Abbildung 4. Die acht Möglichkeiten einer 3-Finger-Gestenauswahl bei der Wrist-Cam. Die Kamera unterschied dabei die Hautfarbe von anderen Hintergrundfarben und entschied so, welche der acht Gesten aktuell ausgeführt wurde. [11]

worauf dann eine der sieben zur Verfügung stehenden Fingerpositionen folgen sollten. Danach ging man wieder in die Ruheposition, um dann eventuell nach dem gleichen Prinzip noch weitere Gesten zu einem zusammengesetzten Kommando zu addieren. Man wählte Zeige-, Mittel- und Ringfinger, da diese für den Menschen am angenehmsten zu bewegen sind und außerdem parallel zueinander verlaufen. Die Ruheposition wurde gebildet durch Ausstrecken der drei Finger, da so eine Eingabe an z. B. einer Tischkante erleichtert wurde. Man legte die drei Finger auf und führte Gesten aus durch Anheben der Finger und anschließendem Absenken auf den Tisch. So konnte die Hand immer wieder ausruhen und entspannen. Das System sollte trotz dieser Überlegung auch völlig ohne Auflagefläche für die Finger funktionieren. Das Gerät wurde für eine Gesteneingabe aktiviert, indem das Handgelenk abgelenkt wurde und dadurch die Kameralinse völlig verdunkelt wurde. Danach war das System bereit und wartete auf eine korrekt ausgeführte Ruheposition, um die Eingabe zu starten. Hatte man eine Gestenkombination beendet, verdunkelte man wieder die Linse und das System ging in den Leerlauf-Modus. Die Erkennung der Finger wurde durch einen Algorithmus realisiert, der es durch Farbenerkennung und -trennung ermöglichte, die Finger vom Hintergrund zu unterscheiden. Die genauere Funktionsweise des Erkennungsalgorithmus und alle technischen Details der Kamera werden in dem Paper von Andrew Vardy et al. [11] genauer ausgeführt.

2.3 Auditive Messung von Gesten

Bei der auditiven Messung will man durch die Unterscheidung von Geräuschen unterschiedlicher menschlicher Gesten eine Abbildung auf Steuerelemente eines mobilen Endgeräts erwirken. Brian Amento, Will Hill und Loren Terveen experimentierten an einem Prototypen, welcher mit einem kleinen piezoelektronischen

Mikrofon verschiedene Fingerbewegungen am Geräusch unterscheiden sollte [12]. Diese Geräusche waren für den Menschen nicht hörbar. Das Mikrofon, welches am Handgelenk angebracht war, nahm Vibrationen auf der Haut auf. Diese entstanden durch Schwingungen beim Bewegen der Knochen der Hand oder der Finger. Das Mikrofon war so konzipiert, dass es keine Nebengeräusche aufnehmen konnte, welche in der Umgebung des Nutzers durch die Luft übertragen wurden. Das war eine wichtige Voraussetzung, um möglichst genaue Messwerte der Knochenschwingungen zu erlangen. Der Prototyp beinhaltete zusätzlich noch einen Mikrocontroller, der die am Mikrofon ankommenden Spannungen aufnahm, interpretierte und einer vordefinierten Gestenklassen zuordnete. Dies geschah, indem das Audiosignal mit einer Rate von 8000 Messungen pro Sekunde abgetastet wurde und jeder Abtastpunkt dann durch eine der zehn vordefinierten Stufen quantisiert wurde. Dabei wurde immer das Maximum der gemessenen Spannung aus dem Zeitraum einer Abtastung verwendet. Diese Stufen wurden von einer Zustandsmaschine benutzt, um einen Rückschluss auf die vollzogenen Gesten zu ziehen. Die Unterscheidungen im entstandenen Signal werden auf Abbildung 5 deutlich. Weiterhin wurden Pufferzustände eingeführt, um eine kontinuierliche

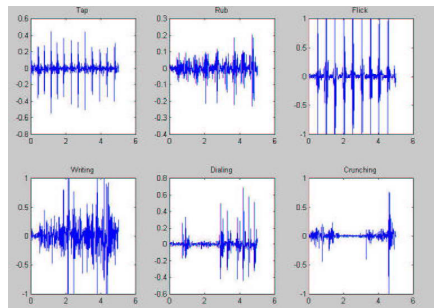


Abbildung 5. Beispielsignale der erkannten Vibrationen auf der Haut des Benutzers: Von links nach rechts in oberer Reihe beginnend: 6 sek. lange Tipp-, Reib- und Schnipp-Gesten, Schreiben mit einem Stift, Wählen einer Telefonnummer und Greifen in eine Tüte mit Chips. [12]

Eingabe zu ermöglichen. Die Gesten-Klassen wurden unterteilt in Schnippen (engl.: flick), Reiben (engl.: rub) und einfaches sowie doppeltes Zusammentippen (engl.: tab) zweier Finger. Auch hier wurde besonders darauf geachtet die Fingerbewegungen für den Anwender so einfach, angenehm und ermüdungsfrei wie möglich zu gestalten. Ein Anwendungsprototyp wurde entwickelt für eine gestengesteuerte Uhr/Pager. Die bereits bestehende Timex Internet Messenger [13] Uhr ist fähig kleine Informationsmengen wie Sportergebnisse, Börsenkurse oder Termine zu empfangen. Man wählt die gespeicherten Daten durch einen Knopfdruck aus. Hier wurde das Mikrofon an der Unterseite des Armbandes der Uhr angebracht und eine Verbindung zu einem oben erläuterten Signalerkenner hergestellt. Ein Tippen der Finger wurde dem Hauptknopf zugeordnet,

welcher eingehende Nachrichten auswählte und ältere Mitteilungen durchsuchte. Der andere Knopf wurde durch ein Doppeltippen belegt und ist zuständig für das Zurückgehen in die Zeitansicht der Uhr. Die Steuerung durch derart auditive Gesten ermöglichte dem Anwender eine einhändige Benutzung ohne Gebrauch der Augen. In diesem Beispiel wurde der Knopfdruck noch durch einen Umleitungsmechanismus simuliert. Aber in naher Zukunft soll es möglich sein diese Techniken z. B. in einer Uhr zu integrieren und somit wäre die Software auf dem Gerät durch Gesten direkt steuerbar.

2.4 Gesten auf einem Berührungsbildschirm

Bei dieser Variante geht es um die Erkennung von Fingergesten auf einem berührungssensitiven Bildschirm. Man kennt diesen Ansatz bereits von vielen PDAs oder Mobiltelefonen. Dabei wird ein Plastikstift (engl.: Stylus) benutzt, um auf dem Bildschirm eingeblendete virtuelle Tasten auszuwählen, Texteingaben zu simulieren wie mit einem echten Stift oder um durch Datenseiten zu navigieren (Blättern). Nachteil dieses Systems ist jedoch, dass man immer beide Hände für die Benutzung braucht, einen Stylus zur Verfügung haben muss und die volle visuelle Aufmerksamkeit dem Bildschirm schenken muss. Es gibt jedoch auch Anwendungen auf mobilen Endgeräten, die für die Steuerung weit weniger Eingabemöglichkeiten benötigen als komplexere Applikationen (z. B. auf PDAs). Einfache Musikspieler wären ein Beispiel hierfür. Diese benötigen Steuerbefehle wie Auswahl des nächsten oder vorherigen Liedes, die Lautstärke regulieren oder Start/Stop-Anweisungen. Diese Anwendungsform wird sehr häufig von Menschen benutzt, die nebenher andere Tätigkeiten erledigen und sich durch die Musik im Ohr unterhalten lassen. Somit wollen sie nicht ihre visuelle Aufmerksamkeit ständig der Steuerung widmen und können es teilweise auch nicht. Man denke dabei an das obige Beispiel eines Fahrradfahrers oder Joggers, der sich schon durch einen kurzen Blick auf das Gerät einem großen Sicherheitsrisiko aussetzen würde. Daher machte sich ein Team von Wissenschaftlern Gedanken über eine gestengesteuerte Musikspieler-Anwendung, die auf Basis des Windows-Media-Player Pocket auf einem PDA laufen sollte [14]. Erster Schritt war es ein Design für die Kontrollgesten zu erarbeiten. Sie verwendeten für ihren sog. TouchPlayer einen iPAQ 3630 Pocket PC und einen Stereokopfhörer. Es wurden für die o. g. Funktionalitäten folgende Gesten entwickelt. Von links nach rechts mit dem Finger über den Bildschirm streifen bedeutete das nächstes Lied und andersherum das vorheriges Lied auswählen (für Linkshänder genau umgekehrt). Von unten nach oben um die Lautstärke zu erhöhen und in die Gegenrichtung um diese zu verringern. Start und Stopp wurden mit einem einfachen Tippen auf den Bildschirm ausgeführt. Die Registrierung der Gesten funktionierte genauso wie bei Eingaben mit dem Stylus, der üblicherweise für dieses Gerät benutzt wird. Die Erkennung war sehr einfach, da es sich nur um Linien und einfache Punkte handelte. Für die Geste "nächstes Lied" wurde z. B. einfach der Startwert in Form der X- und Y-Koordinate des Bildschirms verglichen mit dem Endpunkt der Fingerbewegung. Dabei musste der Anwender mindestens 40 Pixel in X-Richtung und maximal zehn Pixel in der Y-Richtung

mit dem Finger zurückgelegt haben, damit diese Geste als gültig erkannt wurde. Die Länge von 40 Pixeln sollte ungewollte Gesten durch zufällige Berührungen vermeiden. Wiederum sollten die zehn Pixel eine gewisse Toleranzgrenze darstellen, da es für den mobilen Benutzer schwierig war gerade und parallele Linien zum Bildschirmunterrand zu ziehen. Es wurden zwei unterschiedliche Benutzerstudien durchgeführt. Bei der ersten wollte man eine möglichst reale Umgebung imitieren, aber trotzdem genügend Messwerte erhalten, um eine Endbewertung des Tests zu ermöglichen. Daher bauten sie einen acht Meter langen Parcours (siehe Abbildung 6) und ließen die 15 Teilnehmer (Studenten der Universität Glasgow) diesen insgesamt 17-mal durchlaufen. Unter diesen Umständen war

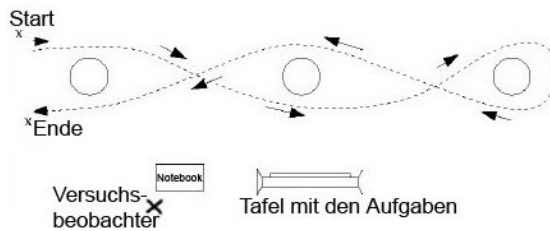


Abbildung 6. Test-Parcours aufgebaut in einem Gang der Universität von Glasgow. Dieser diente zur Ermittlung von Verbesserungen durch die Berührungsgesteneingabe am TouchPlayer gegenüber der am MyMediaPlayer. [14]

es sogar möglich, dass die Testpersonen durch andere Menschen, die den Gang benutzten, gestört wurden. Das lag aber im Sinne der Wissenschaftler, da diese Situation im Alltag ständig passieren würde. Die Aufgabe war, während dem Laufen eine Anzahl von Steuergesten auszuführen. Diese Aufgaben wurden beim Start jeder Runde an einer Tafel angezeigt und waren innerhalb einer Runde leicht schaffbar. Es waren Befehle, wie z. B. ein spezielles Lied suchen oder zwei Titel vorspringen und dann die Lautstärke erhöhen. Dabei wurde die Gesamtzeit genommen, die Zeit für die einzelnen Aufgaben, aufgetretene Fehler bei der Eingabe und der sog. Workload (kognitive Belastung bzw. Ablenkung durch die Steuerung). Dieser Wert bezeichnet in wie weit ein Benutzer durch die Bedienung eines Endgeräts abgelenkt ist bzw. eine kognitive Belastung durch die Belegung der Eingabekanäle stattfindet. Kann dieser Faktor reduziert werden ist ein Gerät sehr gut geeignet für den mobilen Einsatz. Ein weiterer Maßstab war, wie stark die Testpersonen unter diesen Bedingungen von ihrer normalen und bevorzugten Gehgeschwindigkeit abweichen würden. Das sollte zeigen in wie weit die Gesteneingabe den Anwender beeinflusst. Die Studenten mussten das Experiment sowohl über die Steuerung des Mediaplayers via Stylus durchführen als auch den TouchPlayer benutzen durch Gesteneingabe auf dem Berührungsbildschirm. Die Ergebnisse werden im folgenden angeführt: Mit dem TouchPlayer waren die Benutzer durchschnittlich 21 Sekunden schneller als bei der Stylus-Eingabe. In allen Fällen lagen die Anwender näher an ihrer normalen Gehgeschwindigkeit als bei

Stylus-Bedienung (siehe Abbildung 7), was als sehr positiv anzusehen ist. Die

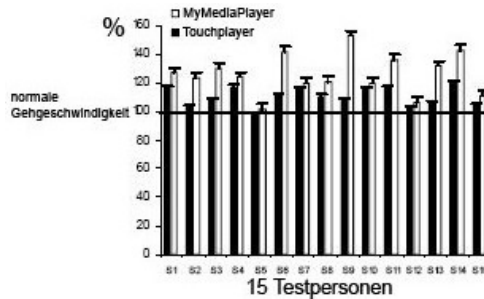


Abbildung 7. Prozentuale Abweichung von der normalen Gehgeschwindigkeit pro Testperson bei der Anwendung von Berührungsgesten am Bildschirm des TouchPlayers gegenüber der Stylus-Eingabe am MyMediaPlayer. [14]

Fehlerquote unterschied sich kaum in beiden Versuchen. Jedoch war der Nutzer schneller fertig mit den Aufgaben und hatte weniger Workload. Er brauchte also viel weniger Aufmerksamkeit als bei einer normalen Stylus-Interaktion. Die meisten Fehler wurden bei Start/Stopp Gesten festgestellt, da es zu einfach war, versehentlich einmal zu tippen. Das wurde bei einem weiteren Test durch ein Doppeltippen verbessert. Insgesamt wurde damit gezeigt, dass es sehr wichtig ist, an neuen Eingabemöglichkeiten für mobile Endgeräte zu arbeiten und dass eine Verbesserung der Benutzerfreundlichkeit speziell im mobilen Einsatz durch solche Gesteneingaben erreicht werden kann.

2.5 Gestenerkennung durch Impulse durch den Körper

In diesem Abschnitt werden zwei Prototypen beschrieben, die beide mit einer Technik arbeiten, die Leitfähigkeitsmessung genannt wird. Dabei handelt es sich um eine Methode, die Distanz zwischen zwei leitfähigen Objekten zu messen. Eine Transmitterelektrode sendet eine Welle (Frequenz im Kilohertzbereich) aus, die dann von einer Empfangselektrode aufgenommen wird. Die Stärke des Signals ist proportional zur Frequenz und der Spannung bzw. der Amplitude der ausgestrahlten Schwingung. Weiterhin verändert die Kapazität von Objekten zwischen den Elektroden die empfangene Amplitude, also die Stärke des Signals. Liegt ein leitendes Objekt nahe der beiden Elektroden, so wird der Amplitudenausschlag verstärkt. Andersherum wird durch ein leitendes, aber geerdetes Objekt der Ausschlag der empfangenen Signalamplitude geschwächt, da Teile des Signals durch die Erdung abgeleitet werden. Diese Eigenschaft nutzte der Wissenschaftler Jun Rekimoto bei seiner Arbeit GestureWrist aus, um die Beschaffenheit des Handgelenks zu messen und unterschiedliche Fingerhaltungen festzustellen [15]. Der Transmitter wurde unter einer Uhr angebracht und die Empfangssensoren an

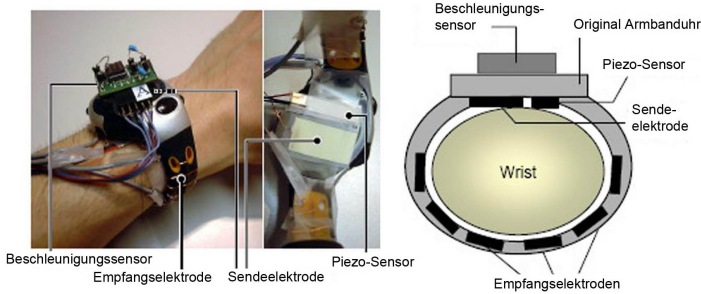


Abbildung 8. links: Prototyp des GestureWrist mit seinen Transmitter-, Empfangselektroden sowie dem Beschleunigungssensor an einer Armbanduhr - rechts: Querschnitt eines Handgelenks mit angelegtem GestureWrist und den integrierten Sensoren. Dabei wurde ein Signal durch das Handgelenk geleitet und erreichte je nach der aktuellen Beschaffenheit des Handgelenks schwächer oder stärker die Empfangselektroden. [15]

der Unterseite des Handgelenks auf der Innenseite des Armbandes (siehe Abbildung 8). Der Vorteil dieses Systems gegenüber z. B. eines Datenhandschuhs zur Erkennung von Finger- bzw. Handstellungen ist klar. Der Benutzer benötigte dabei nur eine gewöhnliche Armbanduhr und jegliche unnatürlichen Messsensoren und -materialien an den Fingern fielen weg. Er konnte dieses System in jeder Situation unauffällig tragen und somit hatte dieses Eingabegerät keinerlei Probleme sozial anerkannt zu werden. Bei dem Prototypen von Rekimoto sollte unterschieden werden, ob die Hand zur Faust geballt ist oder zwei Finger ausgestreckt werden (also die Hand offen ist). Bei diesen beiden Zuständen befand sich der Querschnitt bzw. das Innenleben des Handgelenks in unterschiedlicher Konsistenz. Die Sehnen zogen sich zusammen oder wölbten sich nach außen, je nachdem wie man die Hand bzw. Finger hielt (siehe Abbildung 9). Und genau diese Gegebenheit wurde ausgenutzt und in Kombination mit Kapazitäts- bzw. Leitfähigkeitsmessung gebracht. Die ankommende Amplitude schwankte bezüglich dieser Veränderungen im Handgelenksinneren und konnte somit verwendet werden, um diese Handgesten zu unterscheiden. Es musste berücksichtigt werden, dass die empfangenen Signale oft Nebensignale enthielten, die von umliegenden Stromquellen ausgingen. Daher war es nötig das gesendete Signal zu modifizieren, um es von Störsignalen trennen zu können. Technisch wird hier jedoch nicht weiter darauf eingegangen. Dies kann in der Arbeit von Rekimoto nachgelesen werden [15].

Zusätzlich zur Kapazitätsmessung wurden beim GestureWrist noch Vorderarmbewegungen erkannt mit Hilfe eines Geschwindigkeitssensors, ähnlich wie es in Punkt 2.1 beschrieben wurde. Nun konnten mehrere Gesten mit dem Vorderarm durchgeführt werden, die jeweils nochmals unterschieden wurden durch den Zustand der geschlossenen oder offenen Hand. Mit der Methode des Kapazitätsmessens wäre es wohl möglich, noch mehr Handgesten zu unterscheiden, jedoch beschränkte sich Rekimoto auf die genannten zwei. Er ergänzte diese durch sechs

unterschiedliche Vorderarmstellungen (Handfläche hoch, runter, rechts und links, Vorderarm hoch und runter). Die zwei Zustände der Handstellung wurden verwendet, um aufeinander folgende Gesten zu trennen und so zusammengesetzte Kommandos zu erreichen. Damit wurde eine Möglichkeit geschaffen, ein tragbares Gerät durch Gesten ohne jegliche visuelle Aufmerksamkeit zu steuern.

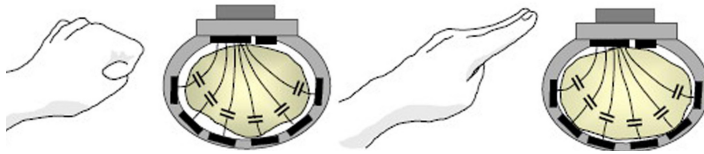


Abbildung 9. Unterschiedliche Querschnitte des Vorderarms basierend auf der Stellung der Hand bzw. der Finger in Bezug auf die Kapazitätsmessung des GestureWrist. [15]

Bei dem zweiten Ansatz von Rekimoto wurde gewöhnliche Kleidung verwendet, um diese als Eingabeobjekte für mobile Endgeräte zu nutzen. Dieses System wurde GesturePad genannt und verwendete kleine Felder, die in die normale Alltagskleidung eingesetzt wurden [15]. Diese waren völlig unauffällig, da sie an der Innenseite der Kleidung angebracht waren. Der Benutzer bediente diese Einheiten durch Handberührungen von Außen. So wurde wieder garantiert, dass das Eingabegerät in jeder sozialen Umgebung getragen werden konnte. Der Anwender wurde daher nicht von einem ungewöhnlich aussehenden Gerät abgehalten, dieses zu tragen und zu akzeptieren. Dieser Aspekt spielt bei der Entwicklung von technischen Geräten immer eine große Rolle. Der Prototyp bestand aus einem rechteckigen Feld mit einem Gitternetz aus Kapazitätssensoren (siehe Abbildung 10 links). Nun wurden mehrere mögliche Konfigurationen des GesturePad

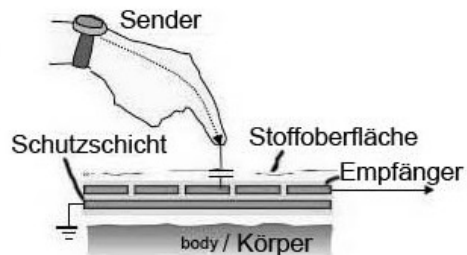


Abbildung 10. links: Prototyp GesturePad mit seinen Gittersensoren zur Kapazitätsmessung - rechts: Variation B des GesturePad, welche in Zusammenhang mit GestureWrist verwendet wurde. [15]

für unterschiedliche Anwendungsmöglichkeiten entworfen (siehe Abbildung 11).

Typ A bestand aus vertikalen Gitterlinien als Transmitter- und aus horizontalen als Empfangselektroden. Das Senden und Empfangen der Signale wurde durch ein Zeitmultiplexing synchronisiert. Das bedeutet, dass die Elektroden nacheinander aktiviert wurden, um eventuelle gegenseitige Mess-Störungen an den Schnittpunkten der Empfangs- und Transmitterelektrode zu vermeiden. Kam nun ein Finger in die Nähe (etwa innerhalb eines cm) des Feldes, dann erkannte das Gitternetz diesen und dessen Position über dem GesturePad. Der Finger nahm als leitendes Objekt die ausgestrahlten Wellen auf und gab diese zurück an die Empfangssensoren. Damit der Körper dieses System nicht beeinflusste, lag unter den Sensoren eine Schutzschicht. Das Feld könnte auch konstante Bewegungen über die Fläche des GesturePad feststellen und somit für die Steuerung von mobilen Geräten verwendet werden. Denkbar wäre eine ähnliche Bedienung eines MP3-Spielers, wie es in Punkt 2.4 beschrieben wurde. Das Feld könnte dabei z. B. im Ärmel einer Jacke oder in einer Hosentasche eingebettet sein.

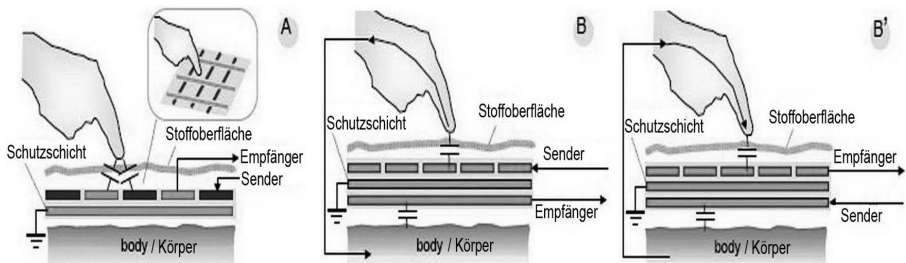


Abbildung 11. Unterschiedliche Konfigurationen des GesturePad. [15]

Typ B und B' stellen eine andere Möglichkeit der Konfiguration dar. Dabei wurden diesmal die Empfangs- und die Transmittersensoren durch eine Schutzschicht getrennt. Auch hier sendete die Transmitterschicht ein Signal aus, welches dann durch den menschlichen Körper geleitet wurde. Schloss der Benutzer den Kreis, indem er das Feld mit dem Finger berührte bzw. nah an das GesturePad kam, dann gelangte das Signal an die Empfangselektrode. B und B' unterschieden sich nur unwesentlich in der Anordnung der Sensoren. Die Position des Fingers bzw. Bewegungen auf dem Feld könnte bei beiden Ansätzen festgestellt werden. Ein großer Vorteil an dieser Konfiguration war, dass nun keine außenstehenden Personen oder andere leitenden Objekte auf das System ungewollt oder gewollt Einfluss nehmen konnten.

Bei allen Konfigurationen waren bis jetzt Empfangs- sowie Sendeelektroden auf dem Feld selbst platziert. Man könnte auch eine einzigen Sendeelektrode anbringen und die Empfangsfelder an den verschiedensten Stellen unter der Kleidung des Nutzers befestigen. Als Beispiel könnte der Sender am Armband des GestureWrist [15] positioniert sein, wie es in Abbildung 10 rechts zu sehen ist. Dieses System könnte man dann in folgendem Szenario verwenden: Ein Redner auf einer Konferenz bedient seine Präsentation durch Berührung bestimmter

Stellen seiner Kleidung mit der Hand. So wäre es vorstellbar, dass er durch Tasten auf seine Ärmelvorderseite eine Folie voran geht und auf der Innenseite eine Folie zurück. Integrierte Videos könnte er vor- oder zurückspulen und falls von Nöten den Projektor an- und ausschalten durch vorher festgelegten Gesten. Er steuert also damit seine mitgebrachten mobilen Geräte, wie Notebook und Projektor, auf eine unauffällige und einfache Art.

2.6 Mechanische Messung von Gesten

In allen bisherigen Gliederungspunkten wurden die Gesten indirekt durch Sensoren erfasst. Damit ist gemeint, dass immer nur bestimmte Sachverhalte und Gegebenheiten zur Interpretation einer Geste verwendet wurden. Folgende Methoden wurden verwendet: welche Geräusche eine Geste verursachte, wie oft ein Finger an einem Sensor vorbei geführt wurde oder durch Impulse, die durch den Körper geleitet wurden und zur Kapazitätsmessung dienten. In diesem Abschnitt wird ein System vorgestellt, das Fingergesten als tatsächlich getätigte Bewegung erkannte. Koji Tsukada und Michiaki Yasumura arbeiteten an einem Prototyp mit dem Namen Ubi-Finger [16]. Dieser sollte zur Steuerung von PDAs oder anderen technischen Geräten im mobilen Einsatz dienen. Es handelte sich um ein kompaktes Eingabegerät, das am Finger angebracht wurde. Vergleichbar mit einem Datenhandschuh, jedoch nur an einem Finger platziert und daher weniger störend und auffällig. Ziel dieser Technik war es ein Eingabegerät zu erschaffen, das zur Steuerung vieler unterschiedlicher Geräte im mobilen Einsatz genutzt werden könnte. Ubi-Finger realisierte dabei durch imitierende Gesten der Finger Operationen von realen Geräten (siehe Abbildung 12). Man versuchte bekannte



Abbildung 12. Umsetzung realer Operationen auf Gesten mit dem Ubi-Finger. [16]

Bewegungen aus dem Alltag auf das System zu übernehmen. Dies erleichterte

den Anwendern sich diese Gesten zu merken und die bekannten Metaphern anzuwenden. Ein Beispiel wäre das Erhöhen der Lautstärke eines beliebigen Gerätes durch intuitives Drehen eines imaginären Knopfes nach rechts, wie man es von Stereoanlagen kennt (siehe Abbildung 12 unten). Die Architektur des Systems teilt sich in vier Bereiche auf. Dem Eingabegerät Ubi-Finger, einer Ansammlung zu bedienender Geräte, einem Steuergerät, welches die Daten des Ubi-Finger verarbeitet und einer zentralen Einheit, um die erkannten Gesten am jeweiligen Gerät auszuführen. Der Ubi-Finger hatte einen Beuge-, Beschleunigungs- und Berührungssensor (siehe Abbildung 13). Außerdem war ein Infrarotsensor daran angebracht, um das zu bedienende Gerät auszuwählen. Der Beuge- und Beschleunigungsmesser waren zur Gestenerzeugung bzw. -erkennung da und der Berührungssensor, um ungewollte Gesteneingaben zu vermeiden. Um eine Eingabe zu starten musste dieser Sensor vorher mit dem Daumen berührt werden, dann konnte ein gewünschtes Gerät durch Daraufzeigen ausgewählt werden, um dann eine Steuerungsgeste zu beginnen. Der genaue Ablauf des Datenaustausches zwischen den vier genannten Komponenten und die Technik der Gestenerkennung können in der Arbeit von Tsukada und Yasumura nachgelesen werden [16].

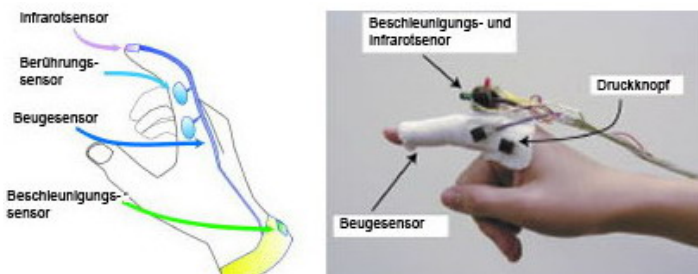


Abbildung 13. links: Basiskonzept des Ubi-Finger mit seinen Sensoren zur mechanischen Gestenerkennung - rechts: Prototyp Ubi-Finger. [16]

Der Vorteil dieses Systems war, dass der Benutzer nun für die Bedienung seiner umliegenden mobilen Geräte nur ein Eingabegerät benötigte. Dieses konnte er intuitiv benutzen, da die Eingabegesten aus der realen Welt übernommen waren und vorhandene Metaphern verwendet wurden. Zu Versuchszwecken mussten Testpersonen eine Stereoanlage und ein Tischlicht kontinuierlich bedienen. Diese Objekte mussten zufällig an- und ausgeschaltet werden, die Licht- bzw. Lautstärke erhöht und verringert werden. Danach wurden diese Personen befragt, ob die verwendeten Gesten aus der realen Welt einfach zu verstehen waren und ob die Auswahltechnik des gewünschten Gerätes durch Zeigen darauf gut gewählt war. Jeweils circa 90 % der Befragten bejahten diese Fragen nach dem praktischen Test. Bei den Gesprächen wurden folgende Kritikpunkte genannt: "Kabel sind sehr störend an dem Gerät", "Ubi-Finger sollte individuell größenverstellbar sein", "Benutzer sollten selbst Gesten hinzufügen bzw. modifizieren können",

"Kontrolle des Berührungssensor fällt etwas schwer" und "Es gab Schwierigkeiten nah zusammenliegende Geräte auszuwählen". Trotz dieser Probleme kann gesagt werden, dass diesem System eine gute Möglichkeit bietet verschiedene mobile sowie stationäre Endgeräte einfach und intuitiv zu bedienen mit nur einem Eingabegerät.

3 Feedback zur Interaktion und Eingabebestätigung

Problem bei fast all diesen Eingabemöglichkeiten für mobile Endgeräte ist, dass der Anwender teilweise auf die bekannten Rückmeldungen über den Erfolg oder Misserfolg seiner Eingabe verzichten muss. Das liegt daran, dass dieses Feedback oft über den visuellen Kanal übermittelt wird, welcher auf Grund der geforderten Eigenschaft mobiler Geräte wenig genutzt werden kann. Mobile Geräte sollen, wie in dieser Arbeit mehrmals beschrieben, wenig visuelle Aufmerksamkeit vom Nutzer fordern, da dieser beweglich ist und sein Augenmerk der Navigation oder anderen Tätigkeiten widmet. Bei Eingaben an Arbeitsplatz-PCs sieht man auf dem Bildschirm sofort, ob die gewünschte Interaktion geklappt hat oder bekommt Warnhinweise bei Fehlern. Dies geschieht entweder bei Erfolg durch eine Aktion, die der Nutzer erreichen wollte oder durch fehleranzeigende Systemnachrichten, wie z. B. aufspringende Fenster mit Textnachrichten.

Rückmeldungen an den Benutzer über seine Eingabe sind sehr wichtig um Eingabefehler zu verringern und zu vermeiden. Durch Feedback direkt nach der Eingabe kann sich der Nutzer erinnern, wie er bestimmte Gesten vollzogen hat und dies unterstützt den Lernprozess. Außerdem kann sich der Anwender durch die Eingaberückmeldung des Geräts sicher sein, ob er die richtige Geste ausgeführt hat oder nicht. Voraussetzung hierfür ist, dass dem Benutzer alle möglichen Rückmeldungen und deren Bedeutung bekannt sind. Als Option für die Rückmeldung des Systems stehen der taktile sowie der auditive Kanal zur Verfügung. Bei dem Musikspieler TouchPlayer in [14] wurde ein auditives Feedback im Kopfhörer erzeugt. Dem Benutzer wurde durch einen kurzen Piep-Ton mitgeteilt, ob er ein Lied vor oder zurück gesprungen ist bzw. ob die Geste nicht erkannt werden konnte. So konnte es vermieden werden diese Gegebenheiten am Bildschirm überprüfen zu müssen. In dem Text von Lumsden und Brewster [5] wurde ebenfalls eine Gestenerkennung auf einem berührungssensitiven Bildschirm eines PDA beschrieben. Sie benutzten als Feedback auditive Signale. Dabei war der Bildschirm in eine drei mal drei Matrix mit 9 Rechtecken aufgeteilt. Jedes der Rechtecke spielte bei Berührung einen Ton basierend aus dem C-Durakkord ab. Dabei wurde die Tonhöhe bezüglich C-Dur von links nach rechts je Spalte erhöht und je Reihe eine Oktave für den jeweiligen Ton aus C-, E- und G-Dur erhöht (siehe Abbildung 14). So konnte der Anwender nach vorangegangener Trainingsphase gut feststellen, wo er sich momentan auf dem Bildschirm befand, wodurch die Eingabe erheblich erleichtert wurde. Das GestureWrist aus [15] gab eine fühlbare Rückmeldung am Handgelenk durch einen piezoelektrischen Sender. Der Anwender spürte also auf der Haut am Uhrarmband eine Berührung nach einer Gesteneingabe und konnte diese interpretieren.

1	2	3
4	5	6
7	8	9

C ₆	E ₆	G ₆
C ₅	E ₅	G ₅
C ₄	E ₄	G ₄

Abbildung 14. links: 3 x 3 Matrix des Berührungsbildschirms - rechts: Feedback Töne angeordnet an dem Raster. Die Buchstaben standen für den Grundton und die kleinstgestellten Zahlen danach bildeten die jeweilige Oktave des Tons [5]

4 Diskussion

Alle vorgestellten Techniken zur Gesteneingabe an mobilen Geräten zeigen, dass großes Potential in dieser Forschung steckt. Grund für die Suche nach neuen Eingabemethoden ist die Beschaffenheit von mobilen Endgeräten und die Situationen in denen sie benutzt werden. In der Vergangenheit zeigte sich, dass es sich lohnt Kraft in solche Forschung zu stecken. Als die ersten Mobiltelefone auf den Markt kamen, hatte noch keiner erwartet, dass der Versand von Kurznachrichten (SMS) zu einer Hauptaufgabe des Geräts werden würde. Man versuchte die Texteingabe durch die Verteilung der Buchstaben auf die zehn Zifferntasten zu lösen. Jedoch war diese Eingabe sehr mühsam, da man bis zu viermal auf eine Taste drücken musste um den gewünschten Buchstaben zu treffen. Eine bessere Lösung war schließlich ein Wörterbuch (T9), das sich den richtigen Buchstaben einer Zifferntaste aus der Kombination aller gewählten Tasten für ein Wort suchte. Dies zeigt, dass technische Fortschritte auch immer wieder neue Probleme an die Interaktion mit den Geräten stellen. Mobile Geräte haben mittlerweile schon so komplexe Funktionen und Anwendungsgebiete, dass man mit der aktuellen Tasteneingabe und mit berührungssensitiven Bildschirmen nicht mehr auskommt. Diese Eingabetechniken fordern außerdem eine zu hohe visuelle Aufmerksamkeit vom Anwender, was teilweise den Nutzen des Geräts einschränkt oder den Benutzer stört und sogar gefährdet.

Jedoch darf man nicht vergessen, dass auch Gesteneingaben ihre Grenzen haben. Die eindeutige Erkennung ist oft schwer und schränkt daher das Operationspotential ein. Bei allen vorgestellten Arbeiten wurden immer nur wenige unterschiedliche Gesten definiert, damit eine gewisse Eingabesicherheit gewährleistet war. Der Anwender muss sich nicht mehr einfach nur eine Taste merken, sondern sich Gesten für gewisse Funktionen einprägen und diese auch üben, um eine korrekte Steuerung zu erreichen. Es gibt also keine Beschriftung mehr für die Funktionen von Gesten, wie es bei Knöpfen am Mobiltelefon z. B. der Fall ist. Außerdem sind zusätzliche Sensoren nötig, die am Körper oder Endgerät angebracht werden müssen. Dies kann die Akzeptanz des Tragens des Geräts beim Anwender beeinflussen und macht Endgeräte zusätzlich teurer. Daher sollte immer abgewogen werden, in wie weit neue Eingabemethoden dem Benutzer Erleichterung verschaffen und ab wann eine Überforderung durch komplizierte oder vielzählige Gesten stattfindet. Meiner Meinung nach ist es daher wichtig sich nicht nur in eine Richtung zu orientieren, sondern viele unterschiedliche

Methoden zu kombinieren. So könnten zusätzliche Eingabehilfen bei häufig wiederkehrenden Grundfunktionen durch Gesten erzeugt werden und seltenere Anforderungen durch z. B. einen Berührungsbildschirm realisiert werden.

Literatur

1. Wikipedia Enzyklopädie: Suchbegriff Geste. <http://de.wikipedia.org/wiki/geste> (2005)
2. Wikipedia Enzyklopädie: Suchbegriff IDrive. <http://de.wikipedia.org/wiki/IDrive> (2005)
3. Fukumoto, M.: Can you input on 24-hours!? - wearable interfaces. In: IPSJ Magazine. Volume 41. (2000) 123–126
4. Institut für Medieninformatik LMU München: Hauptseminar "Mobile Interaktion und mobile Medien". <http://www.medien.ifi.lmu.de/lehre/ss2005/hs.html> (2005)
5. Lumsden, J., Brewster, S.: A paradigm shift: alternative interaction techniques for use with mobile & wearable devices. In: CASCON '03: Proceedings of the 2003 conference of the Centre for Advanced Studies on Collaborative research, IBM Press (2003) 197–210
6. Joel F. Bartlett: Rock 'n' scroll is here to stay. In: IEEE Computer Graphics and Applications. Volume 20. (2000) 40–45
7. Crossan, A., Murray-Smith, R.: Variability in wrist-tilt accelerometer based gesture interfaces. In: Mobile HCI. (2004) 144–155
8. Rekimoto, J.: Tilting operations for small screen interfaces. In: UIST '96: Proceedings of the 9th annual ACM symposium on User interface software and technology, New York, NY, USA, ACM Press (1996) 167–168
9. Strachan, S., Murray-Smith, R., Oakley, I., Ängeslevä, J.: Dynamic primitives for gestural interaction. In: Mobile HCI, Glasgow, Springer Verlag (2004)
10. Metzger, C., Anderson, M., Starner, T.: Freedigiter: A contact-free device for gesture control. In: ISWC '04: Proceedings of the Eighth International Symposium on Wearable Computers (ISWC'04), IEEE Computer Society (2004) 18–21
11. Vardy, A., Robinson, J., Cheng, L.T.: The wristcam as input device. In: ISWC. (1999) 199–202
12. Amento, B., Hill, W., Terveen, L.: The sound of one hand: a wrist-mounted bio-acoustic fingertip gesture interface. In: CHI '02: CHI '02 extended abstracts on Human factors in computing systems, New York, NY, USA, ACM Press (2002) 724–725
13. [presetext.at: Timex stellt Internet-Uhr vor. http://www.presetext.at/pte.mc?pte=010108008](http://www.presetext.at/pte.mc?pte=010108008) (2005)
14. Pirhonen, A., Brewster, S., Holguin, C.: Gestural and audio metaphors as a means of control for mobile devices. In: CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems, New York, NY, USA, ACM Press (2002) 291–298
15. Rekimoto, J.: Gesturewrist and gesturepad: Unobtrusive wearable interaction devices. In: ISWC '01: Proceedings of the 5th IEEE International Symposium on Wearable Computers, IEEE Computer Society (2001) 21
16. Tsukada, K., Yasumura, M.: Ubi-finger: Gesture input device for mobile use. In: Proceedings of APCHI 2002. Number 1 (2002) 388–400