

## **Interaktionstechniken für große Darstellungsflächen**

Joris Ijsselmuiden, Timo Körner, Alexander Schick & Rainer Stiefelhagen

### **Zusammenfassung**

Große Darstellungsflächen finden in unserem Alltag immer größere Verbreitung, wie in Kontroll- und Tagungsräumen, aber auch im privaten Umfeld zu sehen ist. Eine direkte Interaktion mit Objekten auf der Darstellungsfläche ist in der Regel jedoch nicht möglich, sondern nur eine indirekte mittels zusätzlicher Geräte, wie Maus und Tastatur. Doch selbst der Einsatz aktueller Touchscreen-Technologien birgt Probleme in sich. Besonders bei großen vertikalen Darstellungsflächen, wie Videowänden, können manche Objekte außerhalb der Reichweite der Benutzer sein. Benutzer müssen viel entlang der Wand laufen, und bestimmte Objekte können durch ihre Höhe unerreichbar sein.

Wir stellen ein ausschließlich auf Videokameras basierendes System vor, welches eine Interaktion mit großen vertikalen Darstellungsflächen sowohl durch Berührung als auch durch Zeigegesten erlaubt. Für dieses System wurde eine Vielzahl von Interaktionstechniken implementiert. In einer Benutzerstudie wurde evaluiert, welche Interaktionstechniken besonders geeignet sind, um Objekte zu manipulieren. Eine Schlussfolgerung dieser Arbeit ist, dass mit wachsender Komplexität der Aufgabe die Interaktionen mit beiden Händen besser bewertet wurden, während bei einfachen Aufgaben Interaktionen mit einer Hand besser abschnitten. Des Weiteren beinhaltet diese Arbeit empirisch validierte Kriterien für den Entwurf von Interaktionstechniken für große vertikale Darstellungsflächen. Diese können als Grundlage für zukünftige Arbeiten dienen.

### **1 Einleitung**

In zunehmendem Maße halten Benutzeroberflächen, die mittels Gestik bedient werden, Einzug in unser Alltagsleben. Viele Ticketautomaten, Touristik-Informationssysteme und Spielautomaten werden bereits per Touchscreen bedient. Diese Systeme benötigen keine weiteren Eingabegeräte als die eigenen Hände. Obwohl die Technologie für diese Art von Interaktion seit den 1970er Jahren existiert und fortwährend erforscht und weiterentwickelt wurde, ist sie erst seit wenigen Jahren marktfähig geworden. So setzten Firmen wie IBM oder Palm bereits zu Beginn der 1990er Jahre auf Smartphones, die per Touch-Eingabe gesteuert werden. Einen größeren Bekanntheitsgrad erreichte diese Form der

Interaktion aber erst durch Apples iPhone aus dem Jahre 2007, das einen wahren Hype auslöste.

Anwendungsmöglichkeiten für solche Gesten-Interaktionen beschränken sich jedoch nicht nur auf kleine und mittelgroße Displays. Gerade auch in Lage- oder Leitzentren mit großen Bildschirmen und Projektionsflächen hat diese Form der Interaktion ein großes Potential. Statt eine Projektionsfläche von einem Arbeitsplatzrechner aus zu steuern, könnten mehrere Personen direkt mit den Objekten auf der Projektionsfläche interagieren. Dadurch wandelt sich eine Videowand von einer reinen Darstellungsfläche hin zu einem interaktiven Gerät, welches gleichzeitig und kooperativ von mehreren Benutzern verwendet werden kann.

Aufgrund der Größe von gängigen Videowänden stoßen herkömmliche Touchscreen-Technologien schnell an ihre Grenzen. Zu hoch platzierte Objekte können unerreichbar sein und über die Wand verteilte Objekte erfordern viel Bewegung von den Benutzern. Viel sinnvoller ist es in solchen Fällen, eine Interaktion im dreidimensionalen Raum zu erlauben, so dass zusätzlich auch Zeigegesten für die Interaktion verwendet werden können. Eine Schaltfläche kann durch eine Zeigegeste aktiviert werden, ein Fotoalbum kann mittels Wischgesten durchblättern oder Objekte mit Hilfe von Wurfgesten verschoben werden.

Im Folgenden wird die Entwicklung einer von Gesten gesteuerten Benutzeroberfläche für vertikale große Videowände vorgestellt. Mögliche Gesten wurden in diesem Kontext untersucht und implementiert. In einer Benutzerstudie mit einem heterogenen Benutzerkreis wurde das System geprüft, sowohl auf objektive Leistungsmerkmale als auch auf subjektives Empfinden. Diese Arbeit ist wie folgt aufgebaut. In Abschnitt 2 werden zunächst einige verwandte Arbeiten vorgestellt. Abschnitt 3 beschreibt das Gestenerkennungs-System, das eingesetzt wurde, und die spezifischen Gesten, die für die Interaktion benutzt wurden. In Abschnitt 4 werden die durchgeführten Versuche erläutert. Über deren Ergebnisse wird in Abschnitt 5 berichtet, und Abschnitt 6 enthält die entsprechende Diskussion.

## **2 Verwandte Arbeiten**

Die Entwicklung von gestenbasierten Systemen reicht bis in die 1970er Jahre zurück. Zwei besonders wegweisende gestenbasierte Systeme waren „Video Place“ (Krüger et al., 1985) und „Put-That-There“ (Bolt, 1980). Im Folgenden werden kurz die gängigen Technologien vorgestellt (siehe Saffer, 2008, und Buxton, 2009, für einen ausführlichen Überblick). Für kleine Darstellungsflächen wie zum Beispiel Smartphones werden resistive Systeme, Oberflächenwellen-Systeme, akustische Systeme oder kapazitive Systeme eingesetzt, um Berührung zu registrieren. Diese Systeme können nur für kleine Bildschirmgrößen benutzt werden, und sie benötigen eine sichtbare Umrandung der Anzeigefläche, welche die technischen Komponenten beherbergt. Die meisten Handheld-Geräte benutzen entweder ein resistives oder ein kapazitives System. Das bekannteste Beispiel ist das kapazitive Multitouch-System des Apple iPhone (siehe [apple.com/iphone](http://apple.com/iphone)).

Infrarotsysteme sind besser geeignet für große Darstellungsflächen. Sie arbeiten üblicherweise mit einer von zwei Methoden, die beide auch Multitouch erkennen können. Bei der „frustrated total internal reflection“ wird eine Kamera hinter einer infrarotdurchleuchteten Acrylglas-Scheibe platziert. Wenn eine Berührung ausgeführt wird, entstehen Lichtablenkungen, die unterhalb der Glasscheibe empfangen und interpretiert werden können (siehe [perceptivepixel.com](http://perceptivepixel.com) und [microsoft.com/surface](http://microsoft.com/surface)). Bei den gestreuten Beleuchtungs-Systemen wird eine Glasplatte von hinten mit Infrarotlicht bestrahlt. Berührt nun ein Finger oder ein Objekt die Platte, wird an dieser Stelle wesentlich mehr Licht reflektiert. So können Berührungspunkte erkannt werden (Matsushita & Rekimoto, 1997).

Induktive Systeme brauchen einen Stift, Puck oder ähnliches Hilfsmittel. Diese Zeigergeräte enthalten Spulen, um ein elektromagnetisches Feld zu erzeugen. Zusätzlich können Informationen wie Druckstärke oder gedrückte Funktionstasten direkt vom Zeigergerät gemessen werden (siehe [wacom.com](http://wacom.com)). Das DiamondTouch-System von Dietz & Leigh (2001) benutzt ein elektromagnetisches Konzept. Berührungspunkte auf dem von oben anprojizierter Tisch werden dadurch erkannt, dass der Nutzer durch Berührung des Tisches einen Teil der auf dem Tisch vorhandenen Ladung abgreift, die von einem personenbezogenen Empfänger gemessen wird (z.B. im Stuhl). Da jeder Nutzer seinen eigenen Empfänger hat, werden Mehrbenutzer-Interaktion, nutzerspezifische Arbeitsflächen und Rechtevergabe möglich. Nachteile sind der Schattenwurf der Topprojektion und die aufwändige Verkabelung.

Bei videobasierten Systemen werden eine oder mehrere Kameras oder Stereo-Kameras um das Display herum angebracht. Durch Nutzung des Wissens um die relative Position der Kameras zueinander kann eine 3D-Rekonstruktion des Raums vor den Kameras durchgeführt werden. Dadurch lassen sich Informationen über Zeigerichtung, Abstand vom Bildschirm und Haltung der Hand gewinnen. Videobasierte Systeme haben viele Vorteile gegenüber den oben beschriebenen Techniken. So kann man mit dieser Technik nicht nur Zeigegesten erkennen, sondern auch beliebige andere Handgesten. Zusätzlich ist die Methode an Oberflächen unbegrenzter Größe einsetzbar, die Projektionsfläche muss keine Umrandung besitzen und man braucht keinerlei Zeigergeräte. Außerdem hat man die Möglichkeit, neben der Interaktion durch Berührung auch aus der Entfernung mit dem Bildschirm zu interagieren (Bader et al., 2009; siehe [gesturetek.com](http://gesturetek.com)). Für die hier vorgestellte Arbeit wurde ein videobasiertes System eingesetzt, das in der Lage ist, mittels Berührung und Zeigen eine Videowand von 4 x 1,5 m Größe zu bedienen (Bild 1). Aus einer gewissen Entfernung zur Videowand hat man als Nutzer einen besseren Überblick, man muss nicht ständig zwischen Objekten hin und her laufen, und es entsteht nicht die Gefahr dass hoch platzierte Objekte außerhalb der Reichweite liegen. Man kann aber auch aus kurzer Distanz, mittels Berührung, sehr präzise mit der Videowand interagieren.

Seit es Hardware zur Gestenerkennung gibt, werden auch mögliche Gesten zu deren Bedienung erdacht. Einen guten, kompletten Überblick über die möglichen Bedien-Gesten bietet Saffer (2008). Andere Arbeiten, die sich mit dem Thema beschäftigen, findet man in Hancock et al. (2006), Cao et al. (2008) und Wu & Balakrishnan (2003). Eine generische Benutzerschnittstelle, die geeignet für möglichst viele Anwendungen ist, braucht eine große Teilmenge folgender Fähigkeiten: Objektauswahl, Auswahl mehrerer Objekte, Deselektion, Translation, Rotation, Skalierung, Zoomen und Scrollen. Die für diese Arbeit ausgewählten Gesten werden in Abschnitt 3.3 erklärt.

### 3 Methoden

Das hier vorgestellte System stellt eine logische Erweiterung der Interaktion mit berührungssensitiven Flächen dar. Es soll zum einen die Nachrüstbarkeit an bereits bestehenden Projektionsflächen gewährleisten, ohne dass bauliche Veränderungen an einem Display vorgenommen werden müssen. Zum anderen wurde es auch für sehr große Videowände entwickelt, die mit einem herkömmlichen Berührungssystem allein auf Grund der Höhe nicht zu bedienen wären. Das System ist eine Weiterentwicklung des in Schick & Stiefelhagen (2009) und Schick et al. (2009) vorgestellten Systems.

#### 3.1 3D-Rekonstruktion und Gestenerkennung

Zunächst wird eine 3D-Rekonstruktion der Personen vor der Videowand erstellt. Um deren visuelle Hülle zu berechnen, wird eine Technik angewendet, die sich Voxel Carving nennt. Voxel sind Pixel im dreidimensionalen Raum, genauer gesagt kleine Würfel mit konstanter Seitenlänge. Bild 1 zeigt das Ergebnis des Voxel Carvings vor der Videowand. Für die Erkennung sowohl der Berührungs- als auch der Zeige-Interaktion ist es ausreichend, den Arm einer Person zu finden und die Hand als vordersten Teil des Arms anzusehen. Die Zeigerichtung eines Arms wird durch eine Regressionsgerade durch den detektierten Arm bestimmt. Da die 3D-Position der Videowand im Raum bekannt ist, berechnet sich der resultierende Interaktionspunkt durch den Schnitt der Ausgleichsgeraden mit der Videowand-Ebene.



Bild 1: Kameraansicht und Ergebnis des Voxel Carvings; mittels der 3D-Rekonstruktion kann die Zeigerichtung der Arme bestimmt werden.

### 3.2 Implementierte Gesten

Das System kann nicht nur die Zeigerichtung der Arme detektieren, sondern auch seine Entfernung zur Videowand. Somit kann unterschieden werden zwischen Berührung, Hovering und Zeigen aus mehreren Entfernungen. Diese Fälle könnten dann unterschiedlichen Funktionen zugeordnet werden, um so eine mächtigere Interaktion zu schaffen. In dieser Arbeit wird hingegen der Ansatz verfolgt, die Entfernung der Arme zur Videowand möglichst frei von Bedeutung zu halten, damit der Nutzer aus jeder Entfernung dieselben Möglichkeiten hat, wobei jederzeit ein stufenloser Wechsel zwischen Berührungs- und Zeige-Interaktion möglich ist. Falls zum Beispiel die Berührung notwendig wäre, um bestimmte Aktionen auszulösen, würde das zu Problemen führen bei sehr hoch platzierten Objekten und sehr weit auseinander platzierten Objekten.

Es wurden folgende Gesten implementiert, um die Selektion, Manipulation und Deselektion der Objekte auf der Videowand zu ermöglichen. Als Selektier-Geste wurde ein Dwelling-Ansatz verfolgt. Während der Benutzer auf ein Objekt zeigt, schließt sich ein Rahmen um das Objekt. Hat der Rahmen das Objekt erreicht und umschließt er es vollständig, so ist es selektiert. Die Dwelling-Zeit wurde dabei auf einen experimentellen Wert von 0,3 s gesetzt. Um nach der Manipulation ein selektiertes Objekt freizugeben, muss der Benutzer seinen Arm schnell in einer geraden Linie zurückziehen. Ein einfacher Drag-and-Drop Vorgang, bestehend aus oben beschriebener Selektier-Geste, einem Ändern der Zeigerichtung und oben beschriebener Deselektier-Geste, erlaubt das Verschieben der Objekte über die Videowand. Für die Skalierung und Rotation der Objekte wurden jeweils mehrere Varianten implementiert (siehe Abschnitte 4.1 und 4.2). Ziel war es, diese unterschiedlichen Gesten zu evaluieren und zu vergleichen.

## 4 Versuchsaufbau

Im Rahmen dieser Arbeit wurde eine Benutzerstudie durchgeführt, um verschiedene Interaktionstechniken zu evaluieren, aber auch um die Genauigkeit des Systems zu testen. In zwei Versuchsgruppen von jeweils 20 Personen waren insgesamt 15 Versuchspersonen weiblich und 25 männlich. Die Probanden hatten ein Durchschnittsalter von 35,6 Jahren, der jüngste Proband war 21 Jahre alt und der älteste 54. Die Probanden hatten wenig Erfahrung mit Touchscreens ( $\mu_{\text{Touchscreens}} = 2,05$ , mit 1 = keine Erfahrung, 2 = wenig Erfahrung, 3 = viel Erfahrung) und wenig bis keine Erfahrung mit Videowänden und Gestensteuerung ( $\mu_{\text{Videowände}} = \mu_{\text{Gesten}} = 1,525$ ). 30 der Probanden hatten einen technischen Berufshintergrund. Die Körpergröße bewegte sich zwischen 1,59 m und 1,90 m, bei einer Durchschnittsgröße von 1,75 m und dem Median 1,78 m.

Die Versuche sind beschrieben in Abschnitt 4.1 (Versuchsgruppe I) und Abschnitt 4.2 (Versuchsgruppe II). Pro Versuchsperson wurden die Teilversuche jeweils drei Mal durchgeführt, mit zehn Objekten in jedem Durchlauf. Bei der Bestimmung der minimalen Objektgröße reichte ein einziger Durchlauf mit 32 Objekte pro Person. Bevor die Versuchspersonen unter Datenaufzeichnung und

Zeitdruck arbeiteten, hatten sie stets zuerst die Möglichkeit, die unterschiedlichen Interaktionstechniken in einer Einlernphase kennenzulernen. Die Teilversuche wurden in verschiedener Reihenfolge durchgeführt, um eventuelle Lerneffekte auszugleichen. Von jeder Versuchsperson wurde ein Fragebogen ausgefüllt, basierend auf einer Likert-Skala von 1 bis 5: 1 = Aussage trifft eindeutig nicht zu, 2 = Aussage trifft nicht zu, 3 = neutral, 4 = Aussage trifft zu, 5 = Aussage trifft eindeutig zu. Auch wurden die Ausführungszeiten mitprotokolliert und ausgewertet. Um zu untersuchen, ob sich zwei Mengen hinsichtlich ihres Durchschnitts signifikant unterscheiden, wurden der t-Test für normalverteilte bzw. der Wilcoxon Rangsummentest für nicht-normalverteilte oder multivariate Daten verwendet.

#### 4.1 Versuchsgruppe I

In Versuchsgruppe I sollte neben der Evaluation verschiedener Interaktionstechniken eine minimale Objektgröße bestimmt werden, mit denen gerade noch interagiert werden kann. Da aus vergangenen Studien bekannt war, dass die Genauigkeit des Systems von der Position der Objekte auf der Videowand abhängt (Schick et al., 2009), wurde die Wand in 32 Segmente unterteilt und jedes für sich untersucht. In jedem dieser 32 Segmente wurde zufällig ein quadratisches Objekt mit einer Kantenlänge von 100 Pixeln erzeugt (Bild 2). Aufgabe des Benutzers war es, zwei Sekunden lang auf eines der Objekte zu zeigen, damit es verschwindet. Hatte der Proband das Objekt dreimal zwar getroffen, aber nicht lange genug gehalten, so wurde das Objekt innerhalb des gleichen Segments zufällig neu erzeugt, diesmal allerdings 10 px breiter und höher. Hatte der Benutzer alle 32 Objekte der ersten Runde entfernt, folgten noch zwei weitere Runden mit Startgrößen von 50 px bzw. 20 px. Um einzuschätzen, welche Objektgröße für die anderen Teilversuche geeignet war, wurde dieser Genauigkeitstest bereits im Vorfeld mit fünf Versuchspersonen durchgeführt.



Bild 2: Bestimmung der minimalen Objektgröße (Versuchsgruppe I)

Des Weiteren bezogen sich die in Versuchsgruppe I untersuchten Interaktionstechniken auf Skalierung und Rotation. Für die Skalierung wurden Single-Scale, Multi-Scale und 3D-Scale in drei Teilversuche evaluiert (Bild 3). Objekte sollten

in dieselbe Größe wie das Musterobjekt gebracht werden, wobei die 10 Objekte in jeder Runde hinsichtlich der Größe und Position zufällig verteilt wurden. Dabei steuerten jeweils der Abstand zum Objektmittelpunkt (Single-Scale), der Abstand der Hände zueinander (Multi-Scale) oder der Abstand der Hand zur Bildschirmoberfläche (3D-Scale) den Skalierungsfaktor.



Bild 3: Skalierungs-Test (Versuchsgruppe I): Ziel war es, die Objekte in die gleiche Größe zu bringen wie das Muster-Objekt. Von links nach rechts: Single-Scale, Multi-Scale und 3D-Scale.

Für die zwei Rotations-Teilversuche in Versuchsgruppe I bestand die Aufgabe darin, zehn zufällig gedrehte und positionierte Objekte in die gleiche Ausrichtung wie das waagerechte Musterobjekt zu bringen (Bild 4).

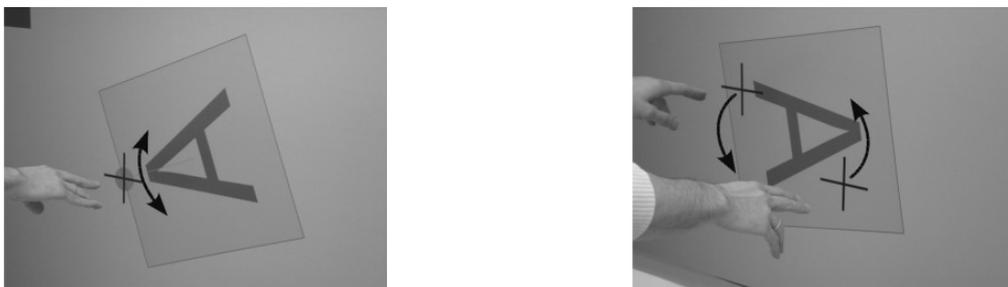


Bild 4: Rotations-Test (Versuchsgruppe I): Ziel war es, die Objekte waagrecht auszurichten. Links: Single-Rotate, rechts: Multi-Rotate.

#### 4.2 Versuchsgruppe II

In Versuchsgruppe II wurde ein Puzzle-Experiment durchgeführt, wobei sowohl das Verschieben von Objekten als auch das Skalieren und Rotieren möglich waren (Bild 5). Zehn verschiedene Objekte mit drei verschiedenen Symbolen und in zwei verschiedenen Farben wurden in zufälliger Größe, Rotation und Position auf der Videowand verteilt. Ziel des Versuches war es, die A, B und C-Objekte in die entsprechenden Ziele an den Rändern zu verschieben. Die Objekte mussten dazu passend rotiert, skaliert und auf die Ziele verschoben werden. Sobald die Objekte, unter Berücksichtigung einer gewissen Toleranz, auf die Ziele verschoben wurden, verschwanden sie sofort. In einem Teilversuch standen die Aktionen Drag-and-Drop (siehe Abschnitt 3.2) und Singletouch-Tools zur Verfügung. Mit den Singletouch-Tools konnte man die Objekte gleichzeitig skalieren und rotieren. Verschieben musste man sie mittels Drag-and-Drop. Im

anderen Teilversuch standen Drag-and-Drop und Multitouch-Tools zur Verfügung. Die Multitouch-Tools erlaubten es der Versuchsperson, Objekte zur gleichen Zeit zu skalieren, rotieren und verschieben. Zusätzlich stand hier noch das Drag-and-Drop zur Verfügung, für den Fall, dass der Benutzer ein Objekt nur verschieben wollte.

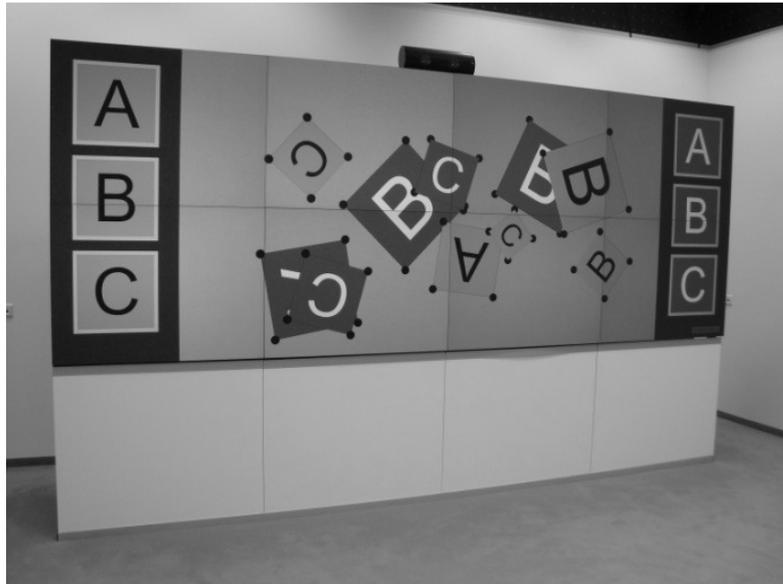


Bild 5: Puzzle-Experiment (Versuchsgruppe II): Die Objekte mussten mit passender Skalierung und Rotation in die entsprechenden Ziele verschoben werden. Unten, von links nach rechts: Drag-and-Drop, Singletouch-Tools und Multitouch-Tools.

## 5 Ergebnisse

Die technische Evaluation soll erstens dazu dienen, das Benutzerverhalten während der Versuche besser zu verstehen, um Benutzeroberflächen entsprechend optimieren zu können. Zweitens soll die Genauigkeit des Systems bestimmt werden, um Aussagen über die Präzision treffen sowie Rückschlüsse auf die minimale Größe von Interaktionsflächen ziehen zu können. Drittens sollen Fehlfunktionen aufgedeckt werden, um Benutzeroberflächen robuster zu gestalten oder gar manche Entwurfs-Entscheidung kritisch zu überdenken.

### 5.1 Systemgenauigkeit

Der in Abschnitt 4.1 beschriebene und in Bild 2 veranschaulichte Genauigkeitstest hatte als Ziel, eine minimale Objektgröße zu bestimmen, die noch präzise

bedient werden kann. Dessen Ergebnisse, bezogen auf die aufgeteilte Fläche der Videowand, sind in Bild 6 visualisiert. Die minimale Objektgröße (Breite und Höhe in Pixeln) wurde hier pro Segment berechnet als Mittelwert vom kleinsten noch getroffenen Objekt über die 20 Versuchspersonen. Da es manchen Versuchspersonen nicht gelang, im oberen Teil der Videowand überhaupt ein Objekt zu treffen, oder bei der Versuchsdurchführung technische Probleme auftraten, berechnet sich in der oberen Reihe der Mittelwert jeweils aus 17 bis 19 Messungen statt 20.

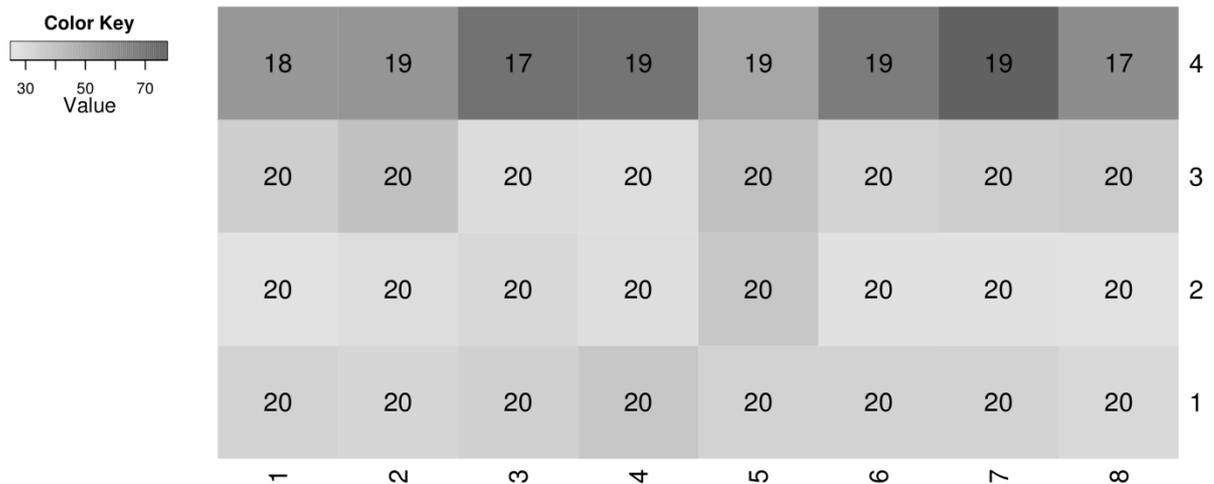


Bild 6: Ergebnisse des Genauigkeitstests: Die minimale Objektgröße für jedes Videowand-Segment in Pixeln lässt sich anhand des Farbschlüssels ablesen; die Anzahl der Messungen pro Segment ist als Zahl hinterlegt.

Bild 6 veranschaulicht, dass mit einer quadratischen Objektgröße von 25 bis 40 Pixeln noch gut bis maximal auf Schulterhöhe interagiert werden kann, also bis einschließlich Reihe 3. In Reihe 4 ist die Interaktion nur mit einer durchschnittlichen Objektgröße von 50 bis 75 Pixeln möglich. Begründet liegt diese Ungenauigkeit in einem Zittern des Zeige-Kreuzes im oberen Bereich. Mit dem jetzigen Stand des Systems scheint daher eine Minimalgröße von 100 Pixeln (10 x 10 cm) sinnvoll. Das heißt aber nicht, dass Objekte immer so groß dargestellt sein müssen. Hier kommt das *Tip-of-the-Iceberg*-Prinzip zum Einsatz, wobei die Größe des Interaktionsbereichs größer ist als das entsprechende Objekt selbst. Des Weiteren lässt sich folgern, dass Elemente, die ein genaues Arbeiten erfordern, am besten im mittleren und unteren Bereich platziert werden sollten. Und schließlich lässt sich feststellen, dass die visuelle Darstellung des Interaktionspunktes durch ein Kreuz entscheidend zur genauen Interaktionsfähigkeit beitrug.

## 5.2 Ausführungszeiten

In Tabelle 1 sind die durchschnittlichen Ausführungszeiten und die entsprechenden Standardabweichungen für die einzelnen Teilversuche veranschaulicht (siehe auch Abschnitt 4). Die Ausführungszeiten des Skalierungstests (Single-Scale, Multi-Scale und 3D-Scale) unterscheiden sich nicht signifikant. Die Standard-

abweichungen und damit die Streuungen der Teilversuche liegen ebenfalls auf dem gleichen Niveau. Und der Vergleich der beiden Rotationstests (Single-Rotate und Multi-Rotate) zeigt das gleiche Muster. Prüft man zum Beispiel mit einem t-Test die Hypothese, dass man mit Single-Rotate die Rotationsaufgabe signifikant schneller erledigen konnte als mit Multi-Rotate, erhält man mit  $p = 0.1$  eine negative Antwort. Beim Puzzle-Experiment aber (Versuchsgruppe II), ergibt sich bei Multitouch-Tools eine erheblich schnellere Ausführungszeit gegenüber Singletouch-Tools. Dieser Unterschied ist mit  $p = 0.015$  beim t-Test statistisch signifikant. Auch die Standardabweichung und damit die Streuung ist kleiner bei den Multitouch-Tools.

Tab. 1: Ausführungszeiten der einzelnen Teilversuche

Teilversuch	$\mu$ (s)	$\sigma$ (s)
Single-Scale	37,1	6,4
Multi-Scale	39,3	12,1
3D-Scale	36,1	9,4
Single-Rotate	56,1	15,6
Multi-Rotate	59,4	14,7
Puzzle Singletouch-Tools	416,1	231,2
Puzzle Multitouch-Tools	278,8	139,6

### 5.3 Auswertung der Fragebögen

Nach jedem Teilversuch füllten die Probanden einen Fragebogen aus, um die ausprobierten Methoden subjektiv zu bewerten. Zum einen sollten Aussagen über die Intuitivität, die Zuverlässigkeit, die Präzision, die Schnelligkeit, die körperliche Anstrengung, das visuelle Feedback sowie die Eignung für nahe und ferne Objekte einer Interaktionstechnik auf einer Likert-Skala von 1 bis 5 bewertet werden (1 = Aussage trifft eindeutig nicht zu, 2 = Aussage trifft nicht zu, 3 = neutral, 4 = Aussage trifft zu, 5 = Aussage trifft eindeutig zu). Zum anderen sollten die Probanden ihren persönlichen Favoriten nennen, wobei die unterschiedlichen Interaktionstechniken auch als gleich gut bewertet werden konnten. Die favorisierten Interaktionstechniken bei jedem Teilversuch sind in Bild 10 visualisiert. Schließlich hatte jeder der Probanden noch die Möglichkeit, die Stärken und Schwächen der einzelnen Methoden per Freitext zu nennen sowie einige allgemeine versuchsübergreifende Anregungen und Bemerkungen zu schreiben. Im Folgenden werden nur die statistisch signifikanten Unterschiede besprochen. Alle anderen Ergebnisse kann man den Bildern entnehmen.

#### 5.3.1 Auswertung des Skalierungs-Tests

Bild 7 zeigt die Bewertungen für die Eigenschaften von Single-Scale, Multi-Scale und 3D-Scale. 3D-Scale schneidet im Schnitt zwar am besten ab, jedoch sind die Unterschiede nicht immer signifikant.

3D-Scale ist die zuverlässigste der drei Methoden. Bezogen auf Multi-Scale ist der Unterschied mit einem p-Wert von  $p_{\text{zuverlässig},3D\text{-Scale}>Multi\text{-Scale}} < 0.01$  hoch signifikant, mit einem p-Wert von  $p_{\text{zuverlässig},3D\text{-Scale}>Single\text{-Scale}} < 0.1$  bezogen auf Single-Scale allerdings lediglich grenzwertig. Bei der Präzision ist nur der Unterschied zwischen 3D-Scale und Multi-Scale mit  $p_{\text{präzise},3D\text{-Scale}>Multi\text{-Scale}} < 0.05$  signifikant.

Bei der Schnelligkeit schneidet 3D-Scale wieder am besten ab. Signifikant ist dabei jedoch nur der Unterschied zwischen Single-Scale und 3D-Scale mit  $p_{\text{schnell},3D\text{-Scale}>Single\text{-Scale}} < 0.05$ .

Bei der körperlichen Anstrengung schneiden Single-Scale und 3D-Scale, welche jeweils nur einen Arm benötigen, besser ab als die beidhändige Technik Multi-Scale. Der Unterschied zwischen 3D-Scale und Multi-Scale ist signifikant mit  $p_{\text{anstrengend},3D\text{-Scale}<Multi\text{-Scale}} < 0.05$ . Der Unterschied zwischen Single-Scale und Multi-Scale mit  $p_{\text{anstrengend},Single\text{-Scale}<Multi\text{-Scale}} < 0.1$  nur grenzwertig signifikant. Bei der körperlichen Anstrengung ist zu beachten, dass im Gegensatz zu allen anderen abgefragten Eigenschaften, eine niedrige Bewertung besser ist. Dies gilt für Abschnitte 5.3.1, 5.3.2 und 5.3.3 und die Bilder 6, 7 und 8.

Für ferne Objekte wird wiederum 3D-Scale bevorzugt. Eine statistische Signifikanz gibt es jedoch nur zwischen 3D-Scale und Multi-Scale mit einem p-Wert von  $p_{\text{fern},3D\text{-Scale}>Multi\text{-Scale}} < 0.05$ .

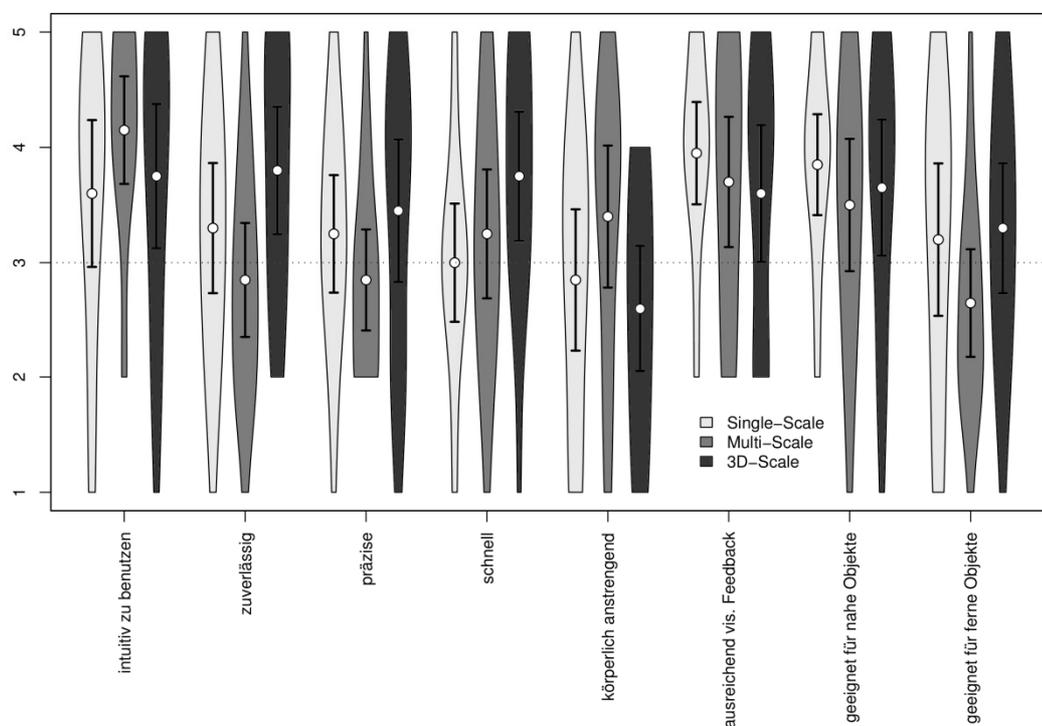


Bild 7: Bewertung der Interaktionstechniken für die Skalierungsaufgabe: Single-Scale, Multi-Scale und 3D-Scale.

Insgesamt geht 3D-Scale klar als Sieger aus der kategorischen Bewertung hervor: Sowohl bei der Zuverlässigkeit als auch bei Präzision, Schnelligkeit und der Eignung für ferne Objekte erhielt die Methode signifikant bessere Bewertungen. Multi-Scale wurde signifikant als körperlich anstrengendste Methode bewertet. Diese Ergebnisse spiegeln sich auch in der Wahl der favorisierten Technik wider (Bild 10, links).

### 5.3.2 Auswertung des Rotations-Tests

Bild 7 zeigt die Bewertungen für die Eigenschaften von Single-Rotate und Multi-Rotate.

Bei der Zuverlässigkeit liegt Single-Rotate klar vor Multi-Rotate. Mit  $p_{\text{zuverlässig, Single-Rotate} > \text{Multi-Rotate}} < 0.01$  ist dieser Unterschied hoch signifikant. Auch bei der Präzision wird Single-Rotate hoch signifikant besser bewertet als Multi-Rotate mit  $p_{\text{präzise, Single-Rotate} > \text{Multi-Rotate}} < 0.01$ . Und Single-Rotate ist auch bezogen auf die Schnelligkeit die signifikant bessere Methode mit  $p_{\text{schnell, Single-Rotate} > \text{Multi-Rotate}} < 0.05$ . Außerdem siegt Single-Rotate auch bei der körperlichen Anstrengung:  $p_{\text{anstrengend, Single-Rotate} < \text{Multi-Rotate}} < 0.05$ . Trotzdem liegen beide Bewertungen der körperlichen Anstrengung knapp unter bzw. etwas über der neutralen Position. Leicht von der Hand geht also keine der beiden Techniken. Schließlich liegt Single-Rotate auch für die Bedienung ferner Objekte vorne:  $p_{\text{fern, Single-Rotate} > \text{Multi-Rotate}} < 0.05$ .

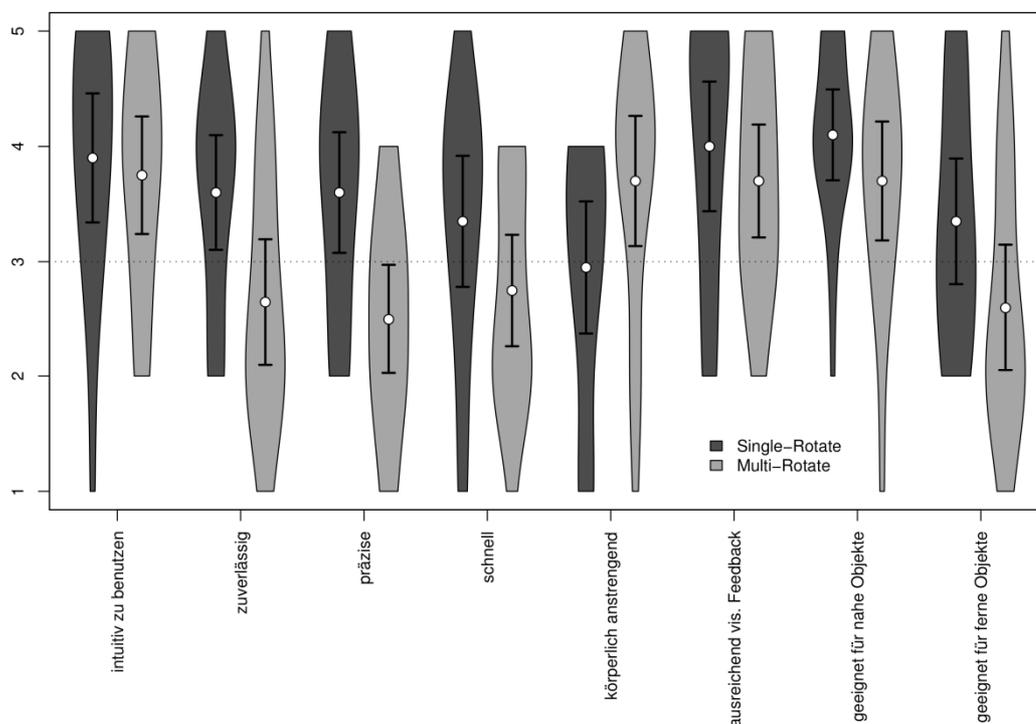


Bild 8: Bewertung der Interaktionstechniken für die Rotationsaufgabe: Single-Rotate und Multi-Rotate.

Single-Rotate kann also in der Zuverlässigkeit, der Präzision, der Schnelligkeit, der körperlichen Anstrengung und der Eignung für ferne Objekte signifikante Unterschiede herausarbeiten und ist somit die bessere Methode. Dies deckt sich mit der Auswahl der favorisierten Methode in Bild 10, Mitte.

### 5.3.3 Auswertung des Puzzle-Experiments

Bild 9 zeigt den Vergleich zwischen Singletouch-Tools und Multitouch-Tools. Die Intuitivität der Multitouch-Tools wurde grenzwertig signifikant höher eingestuft als die Intuitivität der Singletouch-Tools:

$P_{\text{intuitiv, Multitouch-Tools} > \text{Singletouch-Tools}} < 0.1$ . Bei der Schnelligkeit schnitt Multitouch-Tools signifikant besser ab als Singletouch-Tools mit  $P_{\text{schnell, Multitouch-Tools} > \text{Singletouch-Tools}} < 0.05$ .

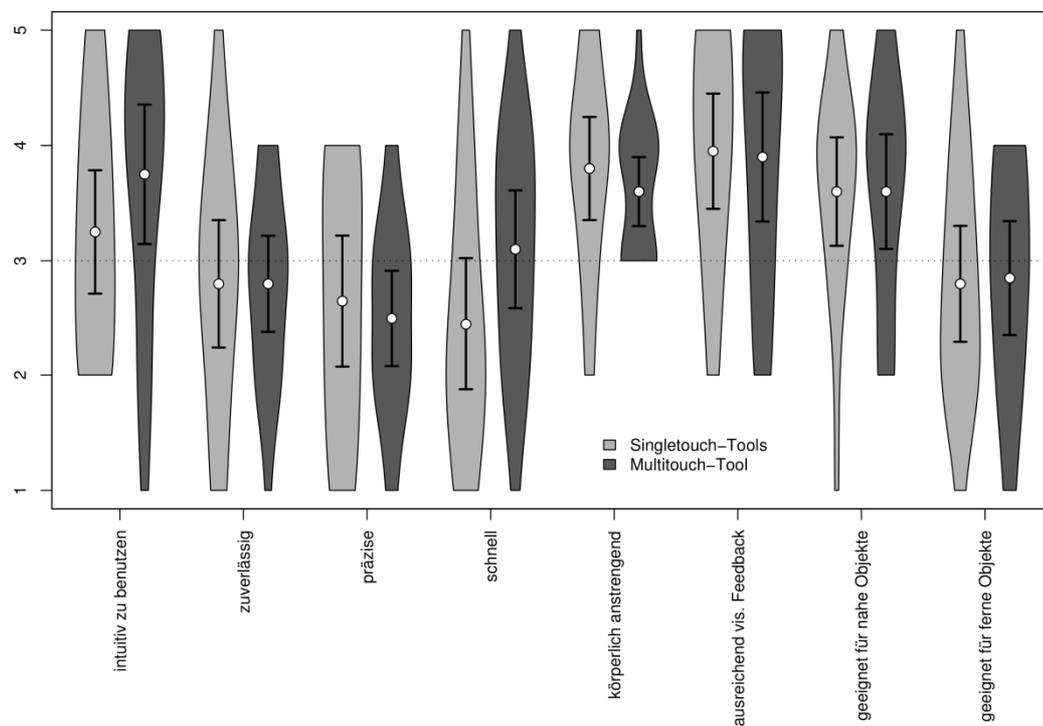


Bild 9: Bewertung der Interaktionstechniken für das Puzzle-Experiment: Singletouch-Tools und Multitouch-Tools.

Wie auch aus Bild 10, rechts, ersichtlich wird, ist Multitouch-Tools die favorisierte Technik.

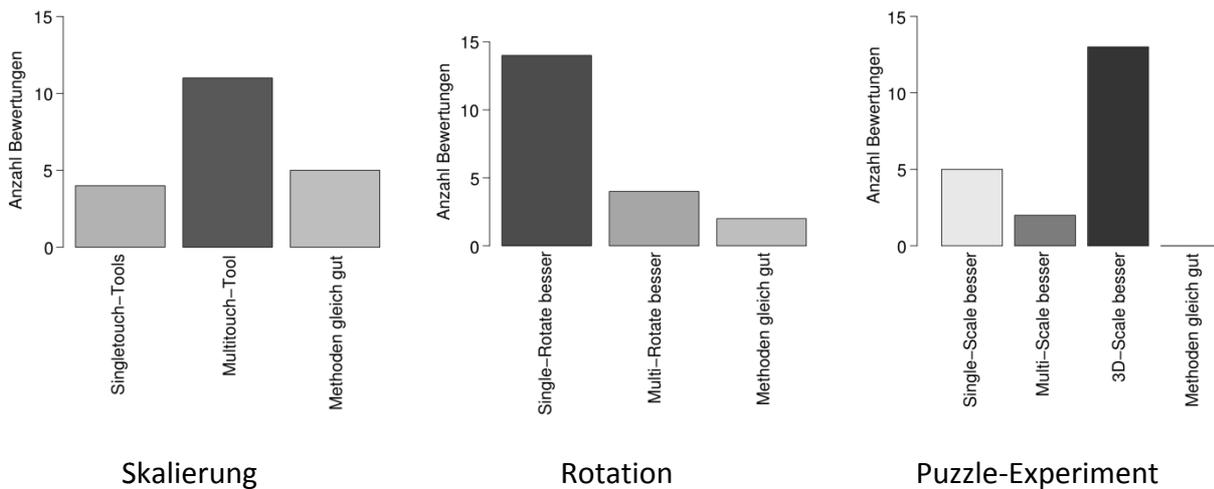


Bild 10: Favorisierte Interaktionstechniken für die Teilversuche

## 6 Diskussion

Die Wahl geeigneter Interaktionstechniken ist schwierig. Wie die Benutzerstudie zeigt, gibt es dabei keine Patentlösung. Und doch zeichnet sich eine Tendenz ab: Je einfacher die Aufgabenstellung ist, desto mehr gewinnen Singletouch-Techniken an Bedeutung. So liegen 3D-Scale und Single-Scale bei der subjektiven Bewertung des Skalierungs-Tests vor der Multitouch-Technik Multi-Scale. Im Rotationsversuch zeichnet sich ein ähnliches Bild ab, denn Single-Rotate wurde besser bewertet als Multi-Rotate. Wird die Aufgabenstellung jedoch komplexer, liegen Multitouch-Methoden in der Benutzergunst deutlich vor den Singletouch-Varianten, wie die Auswertungen des Puzzle-Experiments zeigen: die Methode Multitouch-Tools erhielt hier die besseren Bewertungen.

Oft bewerteten die Teilnehmer die Interaktionstechniken als körperlich anstrengend. Um lange Interaktionsphasen zu unterstützen, muss zum einen die Erkennung der Interaktionstechniken verbessert werden, um fehlerbedingte Wiederholungen zu vermeiden. Zum anderen kann die Oberfläche den Benutzer unterstützen und somit das Arbeiten vereinfachen. Auch der Einsatz zusätzlicher Modalitäten wie Spracheingabe kann das Arbeiten an der Wand vereinfachen. Aus dem Vergleich der Geschwindigkeiten und der bevorzugten Interaktionstechnik ist ersichtlich, dass die jeweils schnellste Technik gleichzeitig auch der Favorit war. Dieser Zusammenhang war jedoch nur für das Puzzle-Experiment statistisch signifikant. Beim Puzzle-Test konnten mit dem Multi-Tool mehrere Aufgaben gleichzeitig gelöst werden, wie das gleichzeitige Rotieren und Verschieben, wodurch der Geschwindigkeitsvorteil gegenüber den Single-Tools besonders deutlich wurde.

## Danksagung

Diese Arbeit wurde gefördert unter dem Fraunhofer Attract-Programm, Projektnummer 692 026. Wir bedanken uns bei Florian van de Camp für die technische

Unterstützung und hilfreiche Ideen. Wir danken auch Michael Voit und Jürgen Geisler für die rege Diskussion und das Schaffen der entsprechenden Kontext und Infrastruktur.

## Literatur

- Bader, T., Räßle, R. & Beyerer, J. (2009). Fast Invariant Contour-Based Classification of Hand Symbols for HCI. In *Computer Analysis of Images and Patterns* (Lecture Notes in Computer Science, Volume 5702, p 689). Berlin, Heidelberg: Springer.
- Bolt, R.A. (1980). "Put-that-there": Voice and gesture at the graphics interface. In *Proceedings of the 7th annual conference on Computer graphics and interactive techniques*, Seattle, USA, pp 262-270.
- Buxton, B. (2009). Multi-Touch Systems that I Have Known and Loved, Bill Buxton, Microsoft Research, <http://www.billbuxton.com/multitouchOverview.html>, November 2009
- Cao, X., Wilson, A.D., Balakrishnan, R., Hinckley, K. & Hudson, S.E. (2008): ShapeTouch: Leveraging Contact Shape on Interactive Surfaces. In *Proceedings of the IEEE International Workshop on Horizontal Interactive Human-Computer Systems*, pp 129-136.
- Dietz, P. & Leigh, D. (2001). DiamondTouch: A Multi-User Touch Technology. In *Proceedings of the 4th Annual ACM Symposium on User Interface Software and Technology*, Orlando, USA, pp 219- 226.
- Hancock, M.S., Vernier, F.D., Wigdor, D., Carpendale, S. & Shen, C. (2006). Rotation and Translation Mechanisms for Tabletop Interaction. In *Proceedings of the First IEEE International Workshop on Horizontal Interactive Human-Computer Systems*, pp 79-88.
- Krüger, M.W., Gionfriddo, T. & Hinrichsen, K. (1985). VIDEOPLACE - An Artificial Reality. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, pp 35-40.
- Matsushita, N. & Rekimoto, J. (1997). HoloWall: Designing a Finger, Hand, Body, and Object Sensitive Wall. In *Proceedings of the 10th annual ACM Symposium on User Interface Software and Technology, Banff, Kanada*, pp 209-210.
- Saffer, D. (2008). *Designing Gestural Interfaces*. Sebastopol, Kanada: O'Reilly Media.
- Schick, A. & Stiefelhagen, R. (2009). Real-Time GPU-Based Voxel Carving with Systematic Occlusion Handling. In *Proceedings of the 31st Symposium of the German Association for Pattern Recognition (DAGM)*, Jena, Deutschland.
- Schick, A., van de Camp, F., Ijsselmuiden, J. & Stiefelhagen, R. (2009). Extending Touch: Towards Interaction with Large-Scale Surfaces. In *Proceedings of Interactive Tabletops and Surfaces 2009 (ITS)*, Banff, Kanada.
- Wu, M. & Balakrishnan, R. (2003). Multi-Finger and Whole Hand Gesture Interaction Techniques for Multi-User Tabletop Displays. In *Proceedings of the 16th annual ACM Symposium on User Interface Software and Technology*, Vancouver, Kanada, pp 193-202.

**Autoren**

J. Ijsselmuiden, M.Sc.  
Dipl.-Inform. T. Körner  
Dipl.-Inform. A. Schick  
Prof. Dr. R. Stiefelhagen

Fraunhofer-Institut für Optronik, Systemtechnik  
und Bildauswertung IOSB  
Karlsruhe  
Karlsruher Institut für Technologie (KIT)  
Institut für Anthropomatik  
Karlsruhe

Kontakt:

[joris.ijsselmuiden@iosb.fraunhofer.de](mailto:joris.ijsselmuiden@iosb.fraunhofer.de)