

# Darstellung physikalischer Objekte in Immerviven Head-Mounted Display Umgebungen

Annika Busch, Marius Staggenborg, Tobias Brix,  
Gerd Bruder, Frank Steinicke, Klaus Hinrichs

Arbeitsgruppe Visualisierung und Computergrafik  
Institut für Informatik  
Westfälische Wilhelms-Universität Münster  
Einsteinstr. 62, 48149 Münster

E-Mail: {a\_busc12|m\_stag02|t\_brix02|g\_brud01|fsteini|khh}@uni-muenster.de

**Zusammenfassung:** Head-Mounted Displays (HMDs) isolieren in der Regel den Benutzer von visuellen Eindrücken der realen Umgebung und erlauben somit das vollständige Eintauchen in eine virtuelle Realität (VR). Insbesondere ist es den Trägern von HMDs nicht möglich den eigenen Körper, physikalische Objekte oder andere Menschen, die sich in der realen Umgebung befinden, visuell wahrzunehmen. Einige HMDs sind mit ein bis zwei Kameras ausgestattet, welche die Darstellung von realen Objekten in der Sicht einer virtuellen Welt erlauben. Meist werden solche *Video-see-through* HMDs benutzt, um die reale Welt mit virtuellen Informationen zu überlagern.

In dieser Arbeit beschreiben wir, wie sich innerhalb des Sichtbereiches der Kamera befindliche reale Objekte in die Darstellung einer virtuellen Welt integrieren lassen. Wir stellen eine Benutzerschnittstelle vor, die es ermöglicht beliebige Objekte in Echtzeit aus Kamerabildern zu segmentieren und in der virtuellen Welt darzustellen. Wir stellen in dieser Arbeit ebenfalls eine Evaluierung der entwickelten Konzepte vor.

**Stichworte:** Virtuelle Realität, Augmentierte Virtualität, Head-Mounted Displays

## 1 Einleitung

Ivan Sutherland stellte 1965 das theoretische Konzept eines *ultimativen* Displays vor, welches dem Anwender erlauben sollte vollständig in eine virtuelle Welt einzutauchen [Sut65]. Drei Jahre später wurden die ersten Head-Mounted Displays (HMDs) präsentiert, mit denen es möglich war, dem Benutzer eine virtuelle Welt aus der *Ich*-Perspektive immersiv darzustellen, während der Anwender visuell von der realen Umgebung isoliert wurde. Heutzutage werden solche Displays mit Tracking-Systemen kombiniert, durch die die Sicht auf die virtuelle Umgebung in Abhängigkeit der Kopfbewegungen des Benutzers aktualisiert wird.

In den meisten Anwendungen der Virtuellen Realität spielt das intuitive Erforschen einer virtuellen Welt, sowie die natürliche Interaktion eine wichtige Rolle. Ein Nachteil von HMDs im Vergleich zur Interaktion in der realen Welt ist, dass Benutzer nicht in der Lage

sind, reale Objekte zu sehen. Insbesondere die Möglichkeit, den eigenen Körper als Referenz für die Interaktion zu verwenden, ist somit nicht inhärent realisierbar. Um die Position und Orientierung von physikalischen Objekten zu bestimmen, werden deshalb häufig optische Marker-basierte Trackingsysteme eingesetzt. Diese Systeme erlauben es, Bewegungen realer Objekte zu detektieren, welche dann mit Bewegungen virtueller Objekte synchronisiert werden. Anwender bekommen auf diese Weise haptisches Feedback durch die realen Gegenstände, während sie visuelles Feedback von den nachmodellierten virtuellen Objekten erhalten.

Dieses Vorgehen ermöglicht die Bereitstellung eines virtuellen Körpers (engl. *Virtual Body*) [SUS95] für den Benutzer, indem die physikalischen Körperbewegungen anhand von einzelnen Feature-Punkten detektiert werden. Diese werden anschließend auf die entsprechenden Stellen eines virtuellen Körpermodells übertragen. In virtuellen Umgebungen lässt sich die Raumwahrnehmung durch die Darstellung eines derartigen virtuellen Körpers stark verbessern, da dieser Aufschlüsse über die Größe von Objekten zulässt und die korrekte Einschätzung von Entfernungen fördern kann [Dra95, MBTCRss]. Darüber hinaus haben Slater u.a. gezeigt, dass das Gefühl der Präsenz eines Benutzers in der virtuellen Welt, d.h. das Empfinden, Teil der virtuellen Welt zu sein, durch einen virtuellen Körper verstärkt werden kann [SUS94, SUS95]. Dies gilt insbesondere dann, wenn Bewegungen des Benutzers in der realen Welt mit virtuellen Körperbewegungen übereinstimmen [UAW<sup>+</sup>06]. Um die Bewegungen des ganzen Körpers eines Benutzers erfassen und in der virtuellen Welt darstellen zu können, muss der Benutzer in der Regel mit vielen Markern oder einem Motion-Capture Anzug instrumentiert werden. Heutige HMDs sind bereits teilweise mit Kameras ausgestattet, welche vom Standpunkt und in Blickrichtung des Benutzers Bilder der realen Umgebung in Echtzeit bereitstellen. Die Bilder, die auf solchen *Video-see-through HMDs* angezeigt werden, lassen sich mit digitalen Informationen erweitern. Virtuelle Objekte, deren Position und Orientierung in Abhängigkeit des HMDs aktualisiert werden, können somit in die reale Welt integriert werden.

In dieser Arbeit beschreiben wir ein Verfahren um physikalische Objekte der realen Umgebung des Benutzers als visuelle Repräsentation in der virtuellen Umgebung darzustellen. Die Sicht auf die virtuelle Welt wird durch eine visuelle Darstellung realer Objekte aus der physikalischen Umgebung des Benutzers, z.B. der eigene Körper eines Benutzers [BSRH09], erweitert. Wir erläutern unser Kamera-basiertes System, bei dem die Kamerabilder in Echtzeit in Vordergrund- und Hintergrund-Pixel segmentiert werden. Somit lassen sich reale Objekte, bspw. der Körper des Benutzers, vom Hintergrund separieren und in die Ansicht der virtuellen Welt integrieren. Die Arbeit gliedert sich wie folgt: Kapitel 2 zeigt den aktuellen Stand der Forschung in diesem Kontext auf. In Kapitel 3 beschreiben wir die notwendigen Segmentierungs- und Verarbeitungsschritte, sowie die Benutzerschnittstelle zur Auswahl der zu segmentierenden Objekte. Kapitel 4 beschreibt eine Benutzerstudie, in der wir unseren Ansatz zur Visualisierung eines virtuellen Körpers evaluiert haben. Kapitel 5 fasst die Arbeit zusammen und gibt einen Überblick über zukünftige Forschungsaktivitäten.

## 2 Verwandte Arbeiten

Im Fokus vieler Forschungsgruppen steht sowohl die Erweiterung der realen Welt, als auch die virtueller Welten mit digitalen beziehungsweise analogen Daten. Die Erweiterung einer Ansicht auf die reale Welt mit computergenerierten Bildern wird als *Augmentierte Realität* (AR), die Erweiterung einer virtuellen Ansicht durch reale Objekte als *Augmentierte Virtualität* (AV) bezeichnet. Der Ausdruck *Gemischte Realität* (engl. Mixed Reality (MR)) [MTY<sup>+</sup>04, MDG<sup>+</sup>95] beschreibt das Kontinuum zwischen AR und AV [JL04, MK94]. Das wesentliche Problem im Bereich der MR ist die Konsistenz realer und virtueller Geometrie und Beleuchtung [JL04, MDG<sup>+</sup>95]. Im Bereich der Augmentierten Virtualität wird die überwiegend virtuelle Szene durch die Darstellung realer Objekte erweitert. Die Interaktion ist in der Regel nicht beidseitig, d.h. reale Objekte erlauben zwar die Interaktion mit virtuellen Objekten, virtuelle Objekte geben jedoch kein physikalisches Feedback. AV-Umgebungen werden häufig im Bereich von Telefonkonferenzsystemen eingesetzt, in denen 2D-Kamerabilder der realen Welt in eine virtuelle Umgebung integriert werden, so dass mehrere Benutzer sich gegenseitig in Echtzeit sehen können (einschließlich Körperhaltung, Gestik und Mimik) [MTY<sup>+</sup>04].

Kamerabilder mehrerer Kameras von realen Objekten können auch genutzt werden, um die dreidimensionale Geometrie der Objekte zu extrahieren und diese als texturierte, virtuelle 3D-Objekte darzustellen [AMR<sup>+</sup>07, GWN<sup>+</sup>03, Lok01]. Dieses Vorgehen erlaubt die Darstellung visuell ansprechender und realistisch aussehender Objekte in virtuellen Umgebungen. Speziell lässt sich dadurch eine virtuelle Nachbildung vom Körper des Benutzers erstellen. Lok u.a. haben in solchen Umgebungen eine leichte, aber nicht signifikante Steigerung der Präsenz des Benutzers festgestellt, wenn diese nicht mit einem abstrakten virtuellen Körper, sondern mit einer realistisch texturierten, virtuellen Nachbildung ausgestattet wurden [Lok02]. Des Weiteren wurde gezeigt, dass realistisch dargestellte virtuelle Körper einen Einfluss auf soziale Interaktion haben [TT]. Ein weiterer Vorteil liegt in der Tatsache, dass Anwender nicht mit zusätzlichen Sensoren oder Markern ausgestattet werden müssen. Das Problem mit den meisten dieser Ansätze zur Rekonstruktion von Geometrie und Materialien realer Objekte ist jedoch, dass dafür statische Konfigurationen mit mehreren Kameras benötigt werden, d.h. die Vorrichtungen können nicht abhängig von der Position und Ausrichtung des Benutzers bewegt werden, was den physikalischen Arbeitsbereich stark einschränkt.

Innerhalb der virtuellen Umgebung kann der Benutzer auf unterschiedliche Arten durch einen Avatar repräsentiert werden. Ein generischer Avatar (meist aus einer Bibliothek) ist hierfür der gebräuchlichste Ansatz [Bro99]. Studien belegen, dass die Nutzung solcher Avatare das Gefühl der Präsenz erheblich verbessern kann [SU93]. Hingegen wird angenommen, dass eine nicht völlig korrekte Repräsentation des eigenen Körpers mit einer Reduzierung des Präsenzgefühls einhergehen kann. Als Alternative zu Avataren haben erste Ansätze aufgezeigt, dass Kameras, die am HMD selbst befestigt sind, verwendet werden können, um die real texturierten Hände des Anwenders in die virtuelle Welt zu übertragen [BSRH09]. Allerdings sind in diesen Ansätzen lediglich die Kamerabilder nach Haut durchsucht worden, während Kleidung und andere Objekte ignoriert wurden.

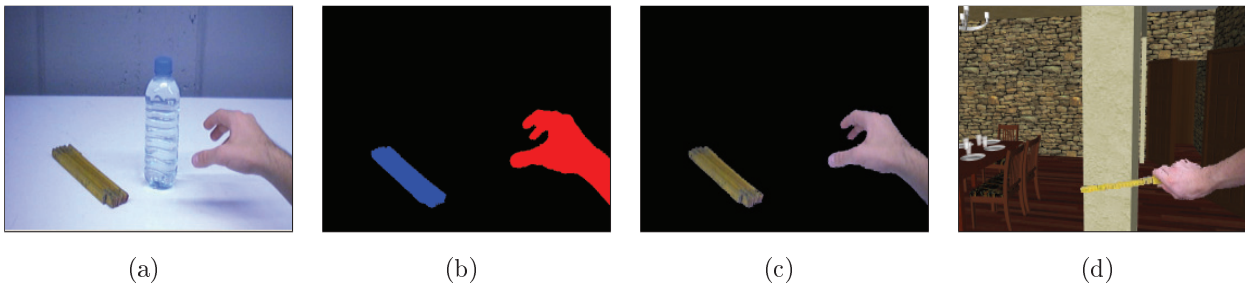


Abbildung 1: (a) Kamerabild des realen Arbeitsbereiches aus der Ich-Perspektive, (b) klassifizierte Pixel identifizierter Objekte, (c) segmentierte Hand und Zollstock, und (d) Verwendung des Zollstocks zur Vermessung eines virtuellen Pfeilers.

### 3 Segmentierung realer Objekte

In diesem Kapitel beschreiben wir ein Verfahren, welches es erlaubt, reale Objekte in die Sicht des Benutzers auf eine virtuelle Umgebung zu integrieren. Über die in Kapitel 2 vorgestellten Ansätze hinaus bietet unser Ansatz nicht nur die Möglichkeit Haut, sondern beliebige, physikalische Objekte zu segmentieren. Dies geschieht mithilfe von in Echtzeit aufgenommenen Kamerabildern aus der Ich-Perspektive (siehe Abbildung 1(a)). Wir erklären im Detail, wie der Benutzer diese Objekte in den Kamerabildern auswählen kann, die wir dann in beliebigen Anwendungen identifizieren und in die Sicht auf die virtuelle Welt übertragen (siehe Abbildung 1(b) und 1(c)). Neben einer oder zwei am HMD angebrachten Kameras ist hierzu keine zusätzliche Instrumentierung des Benutzers notwendig. Das Verfahren erlaubt es, über Farbwerte spezifizierte, reale Objekte im Aufnahmebereich einer Kamera einzeln zu segmentieren und selektiv in eine virtuelle Welt einzublenden. Dadurch können Objekte abhängig von Situation und Aufgabe in der virtuellen Umgebung dargestellt werden (siehe Abbildung 1(d)). Somit ist der Benutzer in der Lage diese Objekte konsistent in der virtuellen Welt zu sehen, d.h. genau dann, wenn sie in der realen Welt in sein Sichtfeld eintreten.

#### 3.1 Video-See-Through System

Wir verwenden ein eMagin Z800 HMD ( $800 \times 600$ , 60 Hz,  $40^\circ$  diagonales Sichtfeld) für die visuelle Präsentation. Dieses HMD ist nativ nicht mit Kameras ausgestattet. Daher befestigen wir eine USB-Kamera mit einer Auflösung von  $640 \times 480$  Pixeln und einer Bildwiederholrate von 30 Bildern pro Sekunde vor dem HMD (siehe Abbildung 2). Die Kamera wird verwendet, um in Echtzeit die reale Welt aus der Ich-Perspektive aufzunehmen. Da das Sichtfeld der Kamera größer als das des HMDs ist, passen wir das Kamerabild so an, dass es mit dem Sichtbereich des HMDs übereinstimmt. Des Weiteren haben wir auf dem HMD eine Infrarot-LED angebracht, deren Position wir im Raum mit einem aktiven, optischen Trackingsystem (WorldViz' Precise Position Tracking), das eine zentimetergenaue Präzision und Genauigkeit garantiert, bestimmen. Die Aktualisierungsrate des Trackingsystems beträgt 60 Hz. Für die Bestimmung der Orientierung des HMDs verwenden wir einen InterSense InteriaCube2 mit einer Updaterate von 180 Hz, der ebenfalls auf das HMD montiert ist. Die virtuelle Welt



Abbildung 2: Foto des verwendeten eMagin Z800 HMD mit einer angebrachten USB-Kamera.

und der virtuelle Körper werden auf einem Intel Computer mit einem Dual-Core-Prozessor, 4 GB Arbeitsspeicher und einer *nVidia GeForce 8800 GTX* Grafikkarte gerendert.

### 3.2 Klassifizierung und Segmentierung

Das zu segmentierende Eingabebild wird von einer an dem HMD angebrachten Kamera geliefert. Wie bereits oben erwähnt, zeigen die Kamerabilder die reale Welt aus Sicht des Benutzers, abhängig von dessen Position und Orientierung. Vom Tracking-System erkannte Kopfbewegungen werden genutzt, um die Sicht der virtuellen Umgebung an die Kopfbewegung anzupassen. Mehrere Verfahren können eingesetzt werden, um Objekte aus den Kamerabildern zu segmentieren. Die einfachste Methode, um die gewünschten Objekte vom Hintergrund zu trennen, wäre ein *Greenbox*-Ansatz. Dieser Ansatz beruht auf der Auswahl einer Hintergrundfarbe, die idealerweise nicht in den zu segmentierenden Objekten vorkommt. Da somit alle Pixel dieser Farbe zum Hintergrund gehören, lässt sich die Segmentierung der Vordergrund-Objekte leicht umsetzen. Allerdings ist in einer realen Laborumgebung meist kein einfarbiger Hintergrund gegeben.

Unser Segmentierungsverfahren ist in zwei Phasen unterteilt. Wie häufig im Bereich der Computer Vision und speziell bei Klassifikatoren, unterscheiden wir zwischen einer Offline- und einer Onlinephase. In der Offlinephase kann der Benutzer einstellen, welche Objekte segmentiert werden sollen. Dazu kann der Benutzer neue, zu segmentierende Objekte anlegen. Diese Objekte können während der Laufzeit aktiviert und deaktiviert werden. Auf diese Weise ist es dem Benutzer möglich, zu entscheiden, welche Objekte in der virtuellen Umgebung dargestellt werden sollen. Beispiele dafür sind der Körper des Benutzers, Hilfsmittel wie ein Zollstock oder auch andere Personen, die sich gerade im Labor aufhalten. Jedem dieser zu segmentierenden Objekte können mehrere Farbinsertalle zugewiesen werden, da sie häufig aus mehreren Farben und Schattierungen bestehen. Um ein neues Farbinsertall hinzuzufügen, markiert der Anwender ein Rechteck im angezeigten Kamerabild, in dem daraufhin die RGB-Werte der Pixel bestimmt werden. Aus diesen wird der Erwartungswert ( $E$ ) jedes Farbkanals, sowie die jeweilige Standardabweichung ( $d$ ) bestimmt. Für die Erwartungswerte



Abbildung 3: Segmentierte Hand (a) ohne und (b) mit aktiver Umgebungskorrektur.

der einzelnen Farbkanäle werden die Intervalle  $[E - d, E + d]$  gespeichert. Anschließend kann der Benutzer die Intervalle manuell vergrößern oder verkleinern, bis ein zufriedenstellendes Ergebnis erreicht ist. Dazu wird dem Anwender visuelles Feedback über die klassifizierten Pixel gegeben, indem diese in einer eindeutigen Farbe für das Objekt eingefärbt werden (vgl. Abbildung 1(b)). Des Weiteren kann der Anwender auswählen, ob die RGB-Werte normiert werden sollen. Da durch die Normierung alle RGB-Werte im Intervall  $[0, 1]$  liegen, lassen sich leichte Schwankungen in der Beleuchtung ausgleichen. Hat der Anwender alle zu segmentierenden Objekte bzw. Farben markiert, ist die Offlinephase abgeschlossen.

In der Onlinephase werden die markierten Objekte bzw. Farben aus den Kamerabildern in Echtzeit segmentiert. Hierfür wird jedes Pixel des Kamerabildes überprüft, ob es eine der Farbwertbedingungen erfüllt. Ist dies der Fall, so wird das Pixel in die virtuelle Umgebung gezeichnet. Um auftretende Lücken in Objekten, die durch Lichtschwankungen entstehen können, zu schließen, werden nach der ersten Segmentierung die Pixel zusätzlich noch einmal bezüglich Nachbarschaftsinformationen in einer  $5 \times 5$ -Umgebung überprüft. Wurde in der Umgebung eines Pixels eine ausreichend große Anzahl an Pixeln eines Objektes erkannt, so wird das Pixel selbst auch dem Objekt zugeordnet. Auf diese Weise können kleinere Lücken aufgefüllt werden. Entsprechend wird die Markierung einzelner Pixel in einem zweiten Durchgang wieder aufgehoben, sofern nicht ausreichend Pixel in der Umgebung markiert sind. Dies erlaubt das Abfangen von Bildstörungen und führt zu einem konsistenteren Bild mit weniger Pixelfehlern (siehe Abbildungen 3(a) und 3(b)). Der Schwellenwert für die Anzahl der Pixel, die in der Umgebung eines Pixels benötigt werden, um hinzugefügt bzw. entfernt zu werden, kann für jedes Objekt separat eingestellt werden.

### 3.3 Komposition realer und virtueller Ansicht

In diesem Abschnitt beschreiben wir, wie die finale Komposition - bestehend aus dem Kamerabild und der gerenderten virtuellen Welt - gewonnen wird. Da das virtuelle Bild sich jeder Bewegung des Kopfes (drehen, kippen, neigen) anpasst, hat es zu jeder Zeit nahezu die gleiche Perspektive wie das Bild der am HMD befestigten Kamera. Somit können die segmentierten Objekte direkt in die virtuelle Welt übertragen werden und passen zu der Blickrichtung des Benutzers. Um Pixel des Kamerabildes in die Sicht der virtuellen Umge-

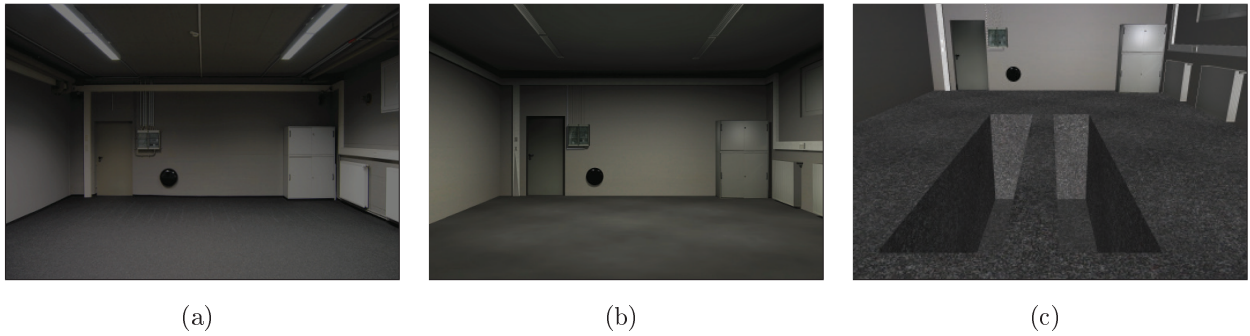


Abbildung 4: (a) Foto des Labors, (b) die gerenderte Sicht des Anwenders vom virtuellen 3D-Modell des Labors und (c) Darstellung der virtuellen Welt mit Abgrund.

zung einzufügen, bedienen wir uns eines einfachen Alpha-Blendings. Dabei kodieren wir im Alphakanal eines jeden Pixels der Kamerabilder die Zugehörigkeit zum Hintergrund oder zu den in der Offline-Phase selektierten Objekten bzw. deren Farbintervalle. Mittels eines einfachen Shaders, der den Alpha-Kanal prüft, werden solche Pixel, die zu ausgewählten Objekten gehören, in den Vordergrund des gerenderten Bildes der virtuellen Welt gezeichnet. Wir blenden das Kamerabild ohne Beachtung der Objektentfernung über die Ansicht auf die virtuelle Szene, da keine Informationen über die Entfernung von realen Objekten aus den Kamerabildern bekannt sind. Auf diese Weise erhalten wir solange ein korrektes Bild, wie die Hand des Benutzers, oder ein anderes segmentiertes Objekt, nicht Objekte in der virtuellen Umgebung durchstößt oder perspektivisch von ihnen verdeckt werden müsste.

## 4 Experiment

In diesem Kapitel beschreiben wir ein Experiment, um das in Kapitel 3 vorgestellte Verfahren zur Darstellung eines Körpers in der virtuellen Welt zu evaluieren.

### 4.1 Experimentelles Design

Als virtuelle Umgebung haben wir ein maßstabsgetreues Modell unseres realen Labors verwendet (siehe Abbildung 4(a) und (b)). Das Labor besteht aus ca. 20.000 Polygonen, welche mit über 100 Fotos texturiert sind. Die Positions- und Orientierungsdaten des getrackten HMDs werden auf die entsprechende Position und Ausrichtung der virtuellen Kamera im 3D-Modell des Labors übertragen. Im Experiment haben wir zwei Bedingungen unterschieden: (Bedingung M) mit und (Bedingung O) ohne Darstellung segmentierter, virtueller Körper aus den Kamerabildern der realen Umgebung.

In diesem Experiment haben wir uns insbesondere auf das subjektive Gefühl der Präsenz, Simulator-Krankheit sowie Präferenzen der Testpersonen fokussiert. Demzufolge haben wir den Slater-Usoh-Steed (SUS) Präsenz-Fragebogen, den Simulator Sickness Questionnaire (SSQ) und einen eigenen Fragebogen (siehe Appendix) verwendet.

Wir haben das Experiment in unserem  $10m \times 7m$  Labor durchgeführt. Die Hardware

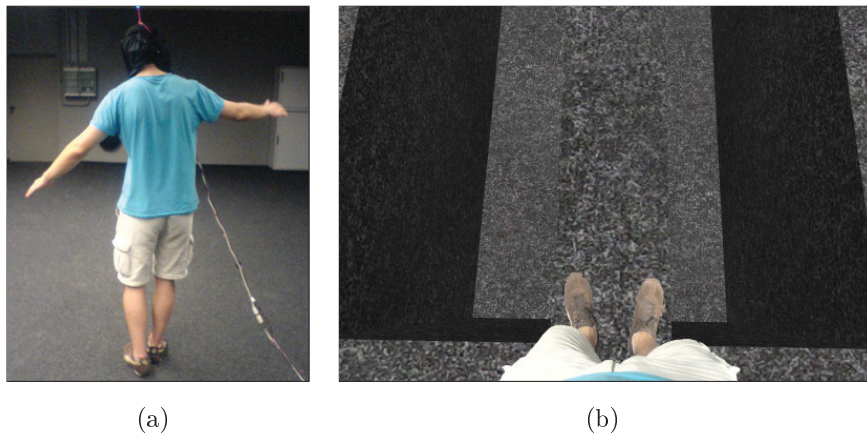


Abbildung 5: (a) Foto eines Teilnehmers während des Experiments und (b) der dargestellte virtuelle Körper bei der Überquerung eines virtuellen Abgrunds.

bestand aus einem eMagin Z800 HMD an dem eine USB-Kamera und Trackingsensoren angebracht wurden (vgl. Abschnitt 3.1). Das HMD und die Kamera wurden über 10m Kabel mit dem Rendering-Rechner verbunden. Die virtuelle Szene wurde mittels unserer eigenen, auf OpenGL-basierenden Software mit einer Bildwiederholfrequenz von 40 – 60 Bildern pro Sekunde dargestellt.

### Experimenteller Ablauf

Zu Beginn des Experiments erläuterte ein experimenteller Beobachter den Teilnehmern die Aufgabe und den weiteren Ablauf. Zunächst erkundeten die Testpersonen das virtuelle Modell unseres Labors durch reales Laufen. Dabei wurden noch keine segmentierten Objekte aus Kamerabildern dargestellt und die Testpersonen konnten sich an die VR-Hardware gewöhnen. Nach drei Minuten haben wir die dargestellte virtuelle Szene derart modifiziert, dass in der Mitte des virtuellen Labors ein virtueller Abgrund dargestellt wurde, der nur über eine schmale Planke überquert werden konnte (siehe Abbildung 4(c) und Abbildung 5(b)). Die Aufgabe der Teilnehmer war es insgesamt 8-mal in der virtuellen Welt über die Planke zu laufen und damit den dargestellten Abgrund in der Mitte der virtuellen Laborumgebung zu überqueren. Das Experiment bestand pro Person aus zwei Durchgängen – mit bzw. ohne Darstellung segmentierter virtueller Körper. Die Reihenfolge wurde randomisiert.

### Teilnehmer

Am Experiment nahmen sieben männliche und drei weibliche Personen im Alter von 21 – 36 ( $\bar{x}$  : 24,7) Jahren teil. Die Testpersonen waren Mitarbeiter oder Studenten an den Fachbereichen Informatik, Mathematik und Psychologie. Fünf der Teilnehmer hatten Erfahrung mit HMDs; sechs hatten Erfahrung mit 3D Computerspielen. Die Gesamtzeit inklusive Trainingsphase, Experiment, Fragebögen und Nachbesprechung betrug pro Person 30 Minuten.



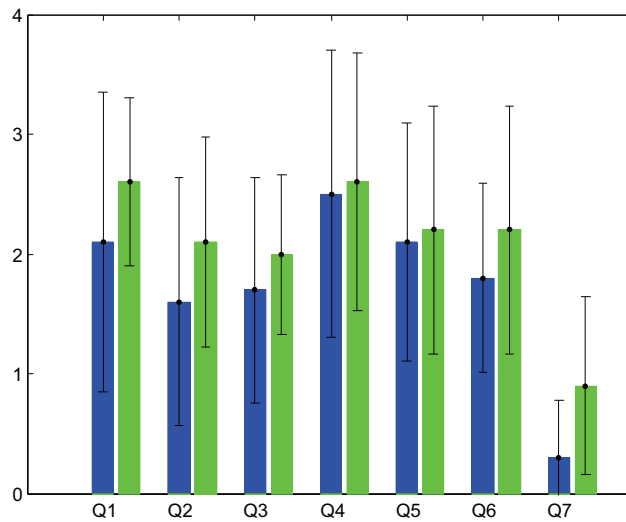


Abbildung 6: Ergebnisse des Fragebogens auf einer 0-4 Skala zu Fragen Q1-Q7 (siehe Appendix). Die blauen Balken repräsentieren die Ergebnisse ohne Segmentierung (O), die grünen Balken die mit Segmentierung (M).

## 4.2 Ergebnisse

Das Experiment hat das Potenzial unseres Ansatzes deutlich aufgezeigt. Die Ergebnisse des SUS-Fragebogens (gemittelte Punktzahl der Präsenz-Fragen auf einer Skala von 1 – 7) weisen auf ein signifikant höheres Präsenzgefühl hin, wenn die Testpersonen ihren eigenen Körper sehen. Der Mittelwert über alle Teilnehmer ist 4,1 unter der Bedingung O (ohne Darstellung des virtuellen Körpers in der virtuellen Welt) und 4,5 unter der Bedingung M (mit Darstellung der segmentierten Körperteile). Der Durchschnittswert der subjektiv eingeschätzten Präsenz der Teilnehmer war 10,25% höher in dem Fall, dass der virtuelle Körper dargestellt wurde. Ein t-Test hat gezeigt, dass die Zunahme der subjektiven Präsenz statistisch signifikant ist ( $p\text{-Wert} = 0,04173$ ). Die Ergebnisse motivieren, dass durch das vorgestellte Verfahren zur Darstellung eines virtuellen Körpers die Präsenz von Anwendern in VR-Umgebungen deutlich verbessert werden kann.

Die Teilnehmer füllten vor und nach beiden Experimentbedingungen jeweils den SSQ-Fragebogen aus. Die Werte des SSQ-Fragebogens in der Bedingung O erreichten im Mittel 7,1 vor dem Experiment und 8,6 nach dem Experiment mit einer relativen Zunahme von etwa 20%. Im Falle der Darstellung virtueller Körper (Bedingung M) erreichten die Angaben der Testpersonen im Mittel 6,7 vor und 9,7 nach dem Experiment, was einer Zunahme von etwa 44% entspricht. Der Unterschied zwischen Bedingung O und Bedingung M ist nicht signifikant. Die Teilnehmer zeigten unter beiden Bedingungen keine starken Anzeichen von Simulator-Krankheit.

Die Ergebnisse der subjektiven Benutzerbefragungen (siehe Anhang) sind in Abbildung 6 dargestellt. Die vertikalen Balken geben die Mittelwerte der Antworten aller Teilnehmer zu den verschiedenen Fragen 1 bis 7 an. Die blauen Balken repräsentieren die Ergebnisse für die Experimentbedingung O (ohne Darstellung segmentierter virtueller Körper); die grü-

nen Balken stellen die Ergebnisse für die Bedingung M (mit sichtbaren Körperteilen) dar. Die schwarzen Balken zeigen die jeweilige Standardabweichung. Obwohl keine der Fragen 1 – 6 eine signifikante Zunahme der subjektiven Beurteilungen der Teilnehmer aufweist, lässt sich ablesen, dass sich durch die Darstellung virtueller Körper die Beurteilungen nicht verschlechtert, sondern sogar leicht verbessert haben (vgl. Abbildung 6). Teilweise auftretende Bildstörungen während der Segmentierung (Frage 7) wurden von den Teilnehmern als nur leicht störend beurteilt (im Durchschnitt mit 0,9).

## 5 Zusammenfassung und Ausblick

In dieser Arbeit haben wir ein einfaches Segmentierungsverfahren vorgestellt, das es erlaubt, eine visuelle Darstellung von physikalischen Objekten in der virtuellen Welt zu erhalten. Wir glauben, dass die beschriebene Technik in Situationen, in denen der Benutzer nicht mit einem Ganzkörper-Trackingsystem ausgestattet werden soll, eine gute Alternative zur kompletten Instrumentierung des Anwenders ist. Im Gegensatz zu aktuellen Ansätzen kann dieses visuelle Körperfeedback leicht in bereits vorhandene HMD-Konfigurationen integriert werden. Der vorgestellte Software-basierte Ansatz benötigt dazu lediglich günstige Kamera-Hardware, die mittlerweile an einigen HMDs bereits vorhanden ist.

Wir haben gezeigt, wie ausgewählte Objekte, wie z.B. der Körper eines Benutzers, aus diesen Kamerabildern aus der Ich-Perspektive segmentiert werden können und wie wir diese Informationen mit dem aktuellen Sichtbereich des Benutzers in der virtuellen Welt kombinieren. Der vorgestellte Algorithmus erlaubt die Echtzeit-Farbsegmentierung von unterschiedlichen Objekten, beispielsweise von Händen, Füßen oder der Kleidung des Benutzers. Die Evaluierung hat gezeigt, dass das subjektive Gefühl der Testpersonen Teil der virtuellen Welt zu sein, die auf einem HMD angezeigt wird, durch unseren Ansatz signifikant gestiegen ist.

In der Zukunft planen wir zusätzliche Tiefeninformationen aus Stereobildern oder einer zusätzlichen Tiefenbildkamera zu extrahieren, die eine effektivere Segmentierung beliebiger Objekte in den Kamerabildern ermöglicht. Die Tiefeninformationen können darüber hinaus verwendet werden, um korrekte Okklusionen sowie natürliche 3D-Interaktion mit der virtuellen Umgebung zu unterstützen. Die vorgestellte Methode hat den Nachteil, dass Interaktion nur mit sich aktuell im Sichtfeld des Benutzers befindlichen Objekten möglich ist. Wir werden untersuchen, inwiefern die Interaktion in praktischen Situationen dadurch eingeschränkt ist, oder ob eine Verbesserung mit zusätzlicher Hardware, etwa weiterer am Kopf angebrachter Kameras, sinnvoll wäre.

## Appendix

### Fragebogen

|  |  |
|--|--|
| 1. „Wie gut konnten Sie Ihre Position im virtuellen Raum einschätzen?“                       | (0 sehr schlecht, 4 sehr gut)<br><input type="checkbox"/> 0 <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4    |
| 2. „Wie schwer war es den Fuß auf das Hindernis zu setzen?“                                  | (0 sehr schwer, 4 sehr leicht)<br><input type="checkbox"/> 0 <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4   |
| 3. „Wie bewerten Sie die Schwierigkeit das Hindernis zu überqueren?“                         | (0 sehr schwer, 4 sehr leicht)<br><input type="checkbox"/> 0 <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4   |
| 4. „Sind Sie sicher, dass Sie das Hindernis korrekt überquert haben?“                        | (0 sehr unsicher, 4 sehr sicher)<br><input type="checkbox"/> 0 <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 |
| 5. „Wie gut konnten Sie die Bewegungen Ihres Körpers im virtuellen Raum einschätzen?“        | (0 sehr schlecht, 4 sehr gut)<br><input type="checkbox"/> 0 <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4    |
| 6. „Wie schwer war es räumliche Zusammenhänge (Distanz zum Boden etc.) korrekt zu erfassen?“ | (0 sehr schwer, 4 sehr leicht)<br><input type="checkbox"/> 0 <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4   |
| 7. „Wie stark sind Sie durch Bildstörungen abgelenkt worden?“                                | (0 gar nicht, 4 sehr stark)<br><input type="checkbox"/> 0 <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4      |

## Literatur

- [AMR<sup>+</sup>07] ALLARD, J., C. MÉNIER, B. RAFFIN, E. BOYER und F. FAURE: *Grimage: Markerless 3D Interactions*. In: *SIGGRAPH (Emerging Technologies)*, 2007.
- [Bro99] BROOKS, F.: *What’s Real About Virtual Reality?* IEEE Computer Graphics and Applications, 19(6):16–27, 1999.
- [BSRH09] BRUDER, G., F. STEINICKE, K. ROTHHAUS und K. H. HINRICHS: *Enhancing Presence in Head-mounted Display Environments by Visual Body Feedback using Head-mounted Cameras*. In: *CyberWorlds Conference*, Seiten 43–50, 2009.
- [Dra95] DRAPER, M.: *Exploring the Influence of a Virtual Body on Spatial Awareness*. Diplomarbeit, Seattle, WA 98195, USA, 1995.
- [GWN<sup>+</sup>03] GROSS, M., S. WÜRMLIN, M. NAEF, E. LAMBORAY, C. SPAGNO, A. KUNZ, E. KOLLER-MEIER, T. SVOBODA, L. VAN GOOL, S. LANG, K. STREHLKE, A. VANDE MOERE und O. STAADT: *Blue-c: A Spatially Immersive Display and 3D Video Portal for Telepresence*. In: 819-827 (Herausgeber): *ACM Transactions on Graphics (TOG)*, Band 22, 2003.
- [JL04] JACOBS, K. und C. LOSCOS: *Classification of Illumination Methods for Mixed Reality*. Technischer Bericht, Eurographics, 2004.

- [Lok01] LOK, B.: *Online Model Reconstruction for Interactive Virtual Environments*. In: *Symposium on Interactive 3D Graphics*, Seiten 69–72, 2001.
- [Lok02] LOK, B.: *Interacting With Dynamic Real Objects in Virtual Environments*. Doktorarbeit, University of North Carolina at Chapel Hill, 2002.
- [MBTCR<sub>ss</sub>] MOHLER, B., H. BÜLTHOFF, W. THOMPSON und S. CREEM-REGEHR: *A full-body avatar improves egocentric distance judgments in an immersive virtual environment*. In: *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization (APGV 08)*, Seiten 194–197, ACM Press.
- [MDG<sup>+</sup>95] MILGRAM, P., D. DRASCIC, J. GRODSKI, A. RESTOGI, S. ZHAI und C. ZHOU: *Merging Real and Virtual Worlds*. In: *Proceedings of IMAGINA '95*, 1995.
- [MK94] MILGRAM, P. und F. KISHINO: *A Taxonomy of Mixed Reality Visual Displays*. In: *IEICE Transactions on Information and Systems*, 1994.
- [MTY<sup>+</sup>04] MAEDA, HIROYUKI, TOMOHIRO TANIKAWA, JUN YAMASHITA, KOICHI HIROTA und MICHITAKA HIROSE: *Real World Video Avatar: Transmission and Presentation of Human Figure*. In: *Proceedings of the IEEE Virtual Reality 2004*, Seite 237, Washington, DC, USA, 2004. IEEE Computer Society.
- [SU93] SLATER, M. und M. USOH: *The Influence of a Virtual Body on Presence in Immersive Virtual Environments*. In: *Proceedings of the Third Annual Conference on Virtual Reality*, Seiten 34–42, London, April 1993. Meckler.
- [SUS94] SLATER, M., M. USOH und A. STEED: *Depth of Presence in Immersive Virtual Environments*. In: *Presence: Teleoperators and Virtual Environments*, Band 3, Seiten 130–144, 1994.
- [SUS95] SLATER, M., M. USOH und A. STEED: *Taking Steps: The Influence of a Walking Metaphor on Presence in Virtual Reality*. In: *ACM Transactions on Computer-Human Interaction (TOCHI)*, Band 2, Seiten 201–219, 1995.
- [Sut65] SUTHERLAND, I.: *The Ultimate Display*. In: *Proceedings of IFIP Congress 2*, Seiten 506–509, 1965.
- [TT] THALMANN, N. M. und D. THALMANN: *The artificial life of synthetic actors*. Transactions of the Institute of Electronics, Information and Communication Engineers D-II, J76D-II(8):1506–14. MIRALab, Geneva Univ., Switzerland.
- [UAW<sup>+</sup>06] USOH, M., K. ARTHUR, M. WHITTON, R. BASTOS, A. STEED, F. BROOKS und M. SLATER: *The Visual Cliff Revisited: A Virtual Presence Study on Locomotion*. In: *International Workshop on Presence*, 2006.