

# Augmented Reality, eine Übersicht

Hartmut Rosch

Technische Universität Ilmenau  
Fachgruppe Graphische Datenverarbeitung  
Technischer Bericht vom 17. Sept. 1999

## Inhaltsverzeichnis

1 Einführung .....	3
2 Begriffsbestimmung.....	3
2.1 Virtual Reality (VR) .....	4
2.2 Virtual Environment (VE).....	4
2.3 Augmented Reality (AR).....	4
2.4 Real Reality (RR) .....	4
3 Industrielle Anwendungen .....	4
3.1 Medizin.....	5
3.2 Industrielles Design und Produktion .....	5
3.3 Visualisierung und Anmerkungen .....	6
3.4 Roboterplanung .....	6
3.5 Unterhaltung .....	6
3.6 Militärische Flugzeuge .....	7
4 Technische Geräte.....	7
4.1 Sichtgeräte .....	7
4.1.1 Datenhelm .....	7
4.1.2 3D-Brille.....	8
4.1.3 Augen-Tracker .....	8
4.1.4 Head-Coupled-Display.....	8
4.2 CAVE und Responsive Workbench .....	9
4.3 Akustische Ausgabegeräte.....	9
4.4 Haptische Ausgabegeräte .....	9
4.4.1 Taktile Rückkopplung.....	10
4.4.2 Force Feedback .....	10
4.5 Zeiger.....	11
4.5.1 3D-Maus.....	11
4.5.2 „magische“ Stäbe .....	12
4.5.3 Datenhandschuh .....	12
4.5.4 Datenanzug.....	12
4.6 Erfassen der Position und -orientierung (Tracking) .....	13
4.6.1 Mechanisch.....	13
4.6.2 Magnetisch .....	14
4.6.3 Optisch .....	14

4.6.4 Kameraüberwachung.....	14
4.6.5 Kinematisch.....	15
4.6.6 GPS .....	15
5 Technische Eigenschaften.....	15
5.1 Grundsätzliches zu Augmentation.....	15
5.2 Optisches oder videobasiertes System.....	16
5.2.1 Stärken und Schwächen .....	18
5.2.2 Akkomodation und Kontrast .....	20
5.2.3 Übergang von der virtuellen zur realen Welt .....	20
5.3 Vergleich zu Virtuellen Umgebungen.....	21
6 Übereinstimmung mit der Wirklichkeit.....	21
6.1 Erfassungsprobleme.....	21
6.2 Statische Fehler.....	23
6.2.1 Optische Verzerrungen.....	23
6.2.2 Fehler im Trackingsystem.....	23
6.2.3 Mechanische Ausrichtungsfehler .....	23
6.2.4 Inkorrekte Betrachtungs-Parameter.....	24
6.3 Dynamische Fehler .....	25
6.3.1 Reduzierung der Verzögerung durch die Systemzeit.....	25
6.3.2 Reduzierung der sichtbaren Verzögerung .....	26
6.3.3 Anpassen der zeitlichen Datenströme .....	26
6.3.4 Vorhersage des künftigen Standortes .....	26
6.4 Sichtbasierende Techniken .....	26
7 Mess-Sensoren .....	28
7.1 Größere Eingabeauswahl und Bandbreite .....	28
7.2 Höhere Genauigkeit.....	29
7.3 Größere Entfernungen .....	29
8 Ausblick .....	30
9 Zusammenfassung.....	31
10 Literatur.....	32

# 1 Einführung

Informationsverarbeitende Technologien haben in den vergangenen Jahrzehnten einen bahnbrechenden Einfluß auf die Arbeitsweisen in der Bürowelt gehabt. Kaum ein Schreibtisch verfügt heutzutage nicht über computerbasierte Hilfsmittel, die zunehmend auch miteinander vernetzt sind. Dieser informationstechnische Fortschritt klammert zur Zeit noch große Teile der Arbeitswelt aus - nämlich all jene Berufsfelder, die sich nicht am Schreibtisch erledigen lassen, weil sie sich direkt mit wirklichen Objekten befassen, beispielsweise bei der Montage und Reparatur von Maschinen, aber auch bei der Wartung großer Anlagen, im Baugewerbe, und in der Medizin bei chirurgischen Eingriffen.

Ziel von Augmented Reality (AR) ist es, Informationen dort zur Verfügung zu stellen, wo sie wirklich benötigt werden. Daten sollen AR-Benutzern als drei-dimensionale dynamische Visualisierungen und Vertonungen in ihr Umfeld eingeblendet werden, um so deren Tätigkeit mit adäquaten Informationen direkt zu unterstützen.

Eine wichtige Aufgabe im Produktionsentwicklungsprozess sind Einbauuntersuchungen. Dabei beschränkt sich der Prozess nicht nur auf die reine Gruppierbarkeit von Komponenten anhand von geometrischen Größen, sondern umfaßt darüber hinaus neben der statischen Anordnung auch die Fragestellung ihres Zusammenbaus und Wartung. Während der Wartungsphase kann man mit Hilfe der AR Automechaniker unterstützen, neue Montagevorgänge beispielsweise beim Einbau eines Schlosses in eine Autotür zu erlernen, ohne immer wieder auf Handbücher zurückgreifen zu müssen.

Man erkennt, welch großes Potential Augmented Reality als realitätsbasierte, drei-dimensionale Benutzerschnittstelle in sich birgt. Oder ein Beispiel aus der Architektur: In allen Phasen eines Bauprojektes kann AR unterstützend eingesetzt werden: während der Projektakquisition, um mit dem Kunden zu besprechen, wie sich das neue Gebäude in die Umgebung einfügt; während des Bauvorgangs, um Arbeitern die nächsten durchzuführenden Schritte direkt "vor Augen" zu führen; und nach Abschluß der Bauphase während Service- oder Umbaumaßnahmen, wenn mit AR anhand von Bauplänen gezeigt werden kann, welche Apparaturen sich in Wänden verbergen.

Damit eine solch realitätsbasierte drei-dimensionale Benutzerschnittstelle in der Industrie angewendet werden kann, müssen Techniken und Praktiken in vielen Bereichen der Informatik weiterentwickelt werden. In dem vorliegenden Bericht werden einige generellen Forschungsbereiche, die für eine breite Akzeptanz der Technologie relevant sind, angerissen. Von großer Bedeutung sind Echtzeitansätze der Visualisierung und der Signalverarbeitung, um die Position des Benutzers schnell und präzise zu bestimmen, die abgerufenen Daten dreidimensional in sein Blickfeld einzublenden und auf mögliche multimediale Benutzereingaben zu reagieren.

## 2 Begriffsbestimmung

Virtual Reality ist eine Mensch-Computer-Schnittstelle, um extrem komplexe Daten zu visualisieren, diese zu manipulieren und mit ihnen direkt zu interagieren. Die Virtual Reality erweitert die reale Welt um eine nahezu grenzenlose virtuelle Welt des Rechners. Die Benutzer empfinden die virtuelle Welt als real, tauchen in sie ein. Die Visualisierung beschränkt sich nicht nur auf den Gesichtssinn des Menschen, sondern steht hier als Überbegriff für die verschiedenen Wahrnehmungssinne des Menschen. So fallen auch akustische Ausgaben unter diesen Begriff. Der Begriff Interaktion setzt angemessene Antwortzeiten voraus (Echtzeit - Interaktion).

Je nach Anwendungszweck und Aufgabe haben sich noch weitere Begriffe gebildet, die andere Eigenschaften der „virtuellen Welt“ beschreiben.

## **2.1 Virtual Reality (VR)**

Mit dem Begriff "Virtual Reality", bzw. "Virtuelle Realität" (kurz: VR) wird eine Technik beschrieben mit der es einem Benutzer durch bestimmte Geräte ermöglicht wird, sich interaktiv in einer vom Computer generierten Umgebung zu bewegen.

Lanier [LMFD89] definiert "Virtual Reality" kurz als eine "interaktive Simulation von realistischen oder imaginären Umgebungen". Der Begriff Cyberspace, der häufig in den Medien genannt wird, geht auf den Schriftsteller William Gibson [Gibs87] zurück und kennzeichnet in dessen Science-Fiction-Erzählung "Neuromancer" ein globales halluzinatorisches Computer-Informationssystem.

## **2.2 Virtual Environment (VE)**

Heutige Definitionen sehen den Begriff „Virtual Reality“ untrennbar mit Datenhelmen, Handschuhen und ähnlichem verbunden. Will man sich jedoch lösen von überwiegend visuellen Anforderungen, zieht also noch audiotives und taktiles Fühlen im Raum ein, so spricht man eher von virtuellen Umgebungen (Virtual Environment).

Hierbei stehen die Benutzer in einem kleinen Raum. Auf die Wände und den Boden wird die von mehreren Computern erzeugte virtuelle Umgebung projiziert. Der drei-dimensionale Eindruck wird mit Hilfe einer 3D-Brille erzeugt. Manipulationen innerhalb der virtuellen Umgebung werden mit einer Art Zeigestock oder dem Datenhandschuh ausgeführt.

## **2.3 Augmented Reality (AR)**

Augmented Reality<sup>1</sup> beschäftigt sich mit der Integration von virtuellen Informationen wie z.B. virtuellen Objekten oder Informationen in reale Umgebungen, die beispielsweise in Form von Bildsequenzen gegeben sein können. Der Rechner ergänzt somit die reale Welt durch Modelle der virtuellen Welt. Die Interaktion zwischen virtueller und realer Welt ist stark eingeschränkt, da es sich dabei häufig nur um visuelle und auditive Projektionen handelt.

## **2.4 Real Reality (RR)**

Im Gegensatz zur Virtual oder Augmented Reality steht bei Real Reality der Rechner im Hintergrund. Die Benutzer können mit bekannten gegenständlichen Objekten agieren und müssen nicht eine Abstraktion der Realität durchführen. In [ScBB97] wird ausgeführt „... verschwindet der Rechner im Hintergrund, man ist unabhängig von der Bildschirmausgabe und kann sich auch mit mehreren Teilnehmern konzentriert den Inhalten zuwenden“.

# **3 Industrielle Anwendungen**

Die Anwendungsgebiete der virtuellen Realität sind sehr vielfältig. Ob es nun der industrielle Einsatz oder der Einsatz im Computerspielebereich ist, überall wo man etwas visualisieren will oder kann, können VR- oder AR-Systeme benutzt werden um diese Visualisierung zu verbessern oder überhaupt erst zu ermöglichen.

---

<sup>1</sup>Augmented Reality, im Sinne von vermehrter, vergrößerter, auch angereicherter Realität

### **3.1 Medizin**

Im medizinischen Bereich werden AR-Systeme eingesetzt, um z.B. den Chirurgen in eine neue Operationstechnik einzuweisen. Außerdem sollen durch die virtuelle Realität auch "Fernoperationen" möglich gemacht werden. Bei diesen Tele-Operationen soll einem Chirurgen die Möglichkeit gegeben werden mittels eines VR-System von einem anderen Ort aus einen Patienten zu operieren. Mittlerweile existieren Prototypen von solchen Tele-Operations-Systemen (Tele-Manipulatoren) und die ersten Trainingssysteme für chirurgische Eingriffe sind im Einsatz [Völt95].

Ein weiteres Gerät welches in manchen Krankenhäusern sogar schon regelmäßig eingesetzt wird, ist das 3D-Video-Endoskop. Dieses Endoskop vermittelt dem Chirurgen mittels einer 3D-Brille ein "räumliches" Bild, welches durch einer Mini-Stereokamera auf einem Monitor dargestellt wird. Der Vorteil liegt darin, daß keine großen Operationsschnitte gemacht werden müssen (Minimal Invasive Chirurgie). Ein kommerzielles System ist das PHANToM-System, welches es ermöglicht, neben der Darstellung des virtuellen Operationsbestecks, auch eine haptische Rückkopplung zu erhalten, die es dem Chirurg ermöglicht, ein Gefühl für den Operationsvorgang zu erhalten [Oak199].

### **3.2 Industrielles Design und Produktion**

Schon heute werden in der Automobil- und Flugzeugindustrie die Möglichkeiten der Augmented Reality, vor allem im Bereich des Prototypen-Designs und der Produktion, genutzt. Man wird durch die Verwendung von AR-Systemen in die Lage versetzt, schon während der Designphase das Produkt, z.B. ein Fahrzeug, betreten zu können oder Simulationen durchzuführen. Beispielsweise lassen sich sehr gut ergonomische Aspekte des Produktes untersuchen. Durch die Technik der AR können Planungsfehler schon sehr früh erkannt werden. Neue Ideen und Verbesserungen können schneller implementiert und direkt getestet werden ohne das ein physikalischer Prototyp erstellt werden muß (Rapid Prototyping), was je nach Produkt sehr Kostenintensiv sein kann [Ullm99]. Diese kostenintensive Herstellung von Modellen zu verringern oder ganz zu eliminieren ist das Ziel der an AR interessierten Firmen. Aber auch Produktionsabläufe lassen sich testen, bevor es überhaupt einen realen Prototypen gibt. Auch ein Einsatz in der Produktion ist möglich, die Firma Boeing z.B. plant den Einsatz von portablen VR-Systemen um es Technikern und Ingenieuren während der Arbeit am Flugzeug zu ermöglichen sich mit Hilfe eines Datenhelms Baupläne und sonstige Information "einblenden" zu lassen.

Ein weiteres aktuelles Beispiel für den industriellen Einsatz von AR-Systemen besteht im Zusammenhang mit dem Weltraumteleskop "Hubble" [Hanc93]. Das Teleskop wurde im April 1990 gestartet, aber die Bilder die Hubble lieferte waren nicht optimal, also wurde von der Firma Lockheed eine Korrekturereinheit für die Optik des Teleskops entwickelt. Da sich das Teleskop ja schon im Weltraum befand, hatte man keine Möglichkeit die Korrekturereinheit bezüglich Funktionalität und Installierbarkeit zusammen mit Hubble zu testen. Durch den Einsatz von AR-Systemen ist diese Möglichkeit, zumindest virtuell, gegeben. Somit wurde die Korrekturereinheit nicht nur mit Hilfe von AR-Systemen entwickelt, sondern auch der Einbau in das Hubble-Teleskop konnte simuliert werden ohne das ein reales Modell erstellt werden mußte. Im Dezember 1993 wurde die Korrekturereinheit erfolgreich in das Hubble-Teleskop eingebaut.

### **3.3 Visualisierung und Anmerkungen**

Augmented Reality kann auch benutzt werden, um Objekte oder Umgebungen mit privaten oder öffentlichen Anmerkungen zu versehen. Anwendungen, die öffentliche Anmerkungen benutzen, setzen auch einen öffentlichen Datenbestand voraus [Fitz93].

Im European Computer Industry Research Centre wurde ein AR-System entwickelt, in dem der Benutzer auf einen Teil eines Automotors zeigen konnte und das System die Bezeichnung des angezeigten Teils darstellte [RBAC94]. Alternativ können dies private Anmerkungen eines Benutzers sein, der diese an bestimmte Teile angebracht hat.

Einen anderen Weg wurde von Wissenschaftlern der Columbia Universität in New York, USA beschritten. Hier wird ein Fenster mit einer normalen Benutzungsoberfläche benutzt, die den Gegenständen oder den handelnden Personen transparent überlagert werden. Der Benutzer trägt dazu ein Positionsgeber (Tracking Device), um das Menü und die Anmerkungen dem Benutzer folgen lassen zu können [FIHS93].

### **3.4 Roboterplanung**

Die Führung eines Roboters aus der Ferne ist häufig recht schwierig. Anstatt eine direkte Überwachung zu installieren, ist es sinnvoll eine virtuelle Version des Roboters zu benutzen. Der Benutzer kann nun in einer virtuellen Umgebung die Steuerung des Roboters als Echtzeit-Anwendung simulieren. Die Ergebnisse liegen dann direkt vor und können sofort verarbeitet werden. Ist die Roboterbewegung ausgetestet kann der Roboter in den Einsatz gehen. In [MZDG93] wird gezeigt, dass stereoskopisches AR besser geeignet und schneller bedienbar ist als traditionelle Roboterplanung.

### **3.5 Unterhaltung**

Die Unterhaltungsindustrie ist aber mit eine der größten treibenden Kräfte in der Entwicklung von VR-Systemen. Insbesondere VR-Systeme werden für den Spielhallenbereich entwickelt und sind teilweise schon seit mehreren Jahren im Einsatz. Schon vor einigen Jahren wurde in Chicago der erste VR-Park eröffnet in dem man auch schon in Gruppen gegeneinander spielen konnte [Borm94]. Auch verschiedene andere Simulatoren sind seit mehreren Jahren in Betrieb, z.B. das Expality, ein Spielhallengerät für Flug- und Raumschiffsimulationen.

Ein anderes großes Projekt welches die Filmgesellschaft Paramount Pictures Corp. und die Computerspielfirma Spectrum Holobyte planen ist die Errichtung von mehreren Star Trek Spielhallen. Hierbei soll es dem Spieler möglich sein, virtuell durch die Enterprise (das Raumschiff der Star Trek Serie) gehen zu können und dabei mit den Charakteren der Enterprise zu interagieren. Sperlich [Sper96] berichtet über die ersten Star Trek Spielhallen, die in Amerika schon eröffnet worden sind.

Aber auch VR-System für den Heimanwender sind im Vertrieb einiger Firmen. Die Computerfirma SEGA [Borc94] beispielsweise hat ein Low-End-VR-System auf dem Markt. Die Konkurrenzfirma Nintendo arbeitet zusammen mit der Firma Silicon Graphics an einer 64-Bit 3D-Spielekonsole, welches ebenfalls zu kaufen gibt [Schm98].

## 3.6 Militärische Flugzeuge

Seit Jahren werden in militärischen Flugzeugen und Hubschraubern sogenannte Head-Up-Displays (HUD) [McFo94], bzw. Helmet-Mounted Sights (HMS) eingesetzt. Mit diesen Geräten wird eine Vektorgrafik in das normale Sichtfeld des Piloten eingespiegelt. Neben den Navigationsdaten und den Flugüberwachungsinformationen werden auch Grafiken dargestellt, welche militärische Ziele in der Umgebung darstellen könnten. In den ersten Ausführungen der HMS-Helme wurde angenommen, dass durch Drehen des Kopfes das Ziel aufgenommen wird und dann so mit der Bewaffnung koppeln, das man durch Kopfdrehung auf das Ziel die Bewaffnung ausrichten kann. Neuere Entwicklungen benutzen dazu Augen-Tracker<sup>2</sup> (eye-tracker), welche weiter unten beschrieben werden.

## 4 Technische Geräte

Im Folgenden soll ein Überblick über die technischen Geräte gegeben werden, welche zur Zeit im Bereich der "Augmented Reality" zu finden sind. Ein grafik-orientiertes VR-System besteht im allgemeinen aus einem Display für den Benutzer, einem Bewegungssensor um Interaktion zu ermöglichen, einem Computer zur Generierung der virtuellen Umgebung, einer Datenbank für die 3D-Objekte und natürlich die Software. Man benötigt einerseits Geräte um dem Benutzer ein dreidimensionales Bild der virtuellen Umgebung zu vermitteln und andererseits Geräte um eine Interaktion mit der virtuellen Umgebung zu erlauben.

### 4.1 Sichtgeräte

Um das räumliche Sehen beim Menschen zu ermöglichen, ist es notwendig, dass das rechte und linke Auge ein horizontal verschobenes Bild sieht, welche aus 2 leicht unterschiedlichen Richtungen dargestellt werden. Aufgabe von visuellen VR-Ausgabegeräten ist es, diese Verschiebung nachzubilden und den Augen stereoskopische Bilder anzubieten.

#### 4.1.1 Datenhelm

Das klassische VR-Ausgabegerät ist das Head Mounted Display (HMD), das als eine Art Helm aufgesetzt wird. In dem Helm sind noch zwei kleine Bildschirme integriert. Es arbeitet typischerweise mit zwei Bildern, die entweder von kleinen **LCDs** (Liquid Crystal Display) oder von **CRTs** (Cathode Ray Tube) erzeugt werden. Jeder Bildschirm zeigt dabei wiederum die virtuelle Umgebung aus einer etwas anderen Richtung, um einen dreidimensionalen Effekt zu erzeugen.

Dabei ist eine hohe Auflösung der erzeugten stereoskopischen Bilder notwendig, um ein brauchbares weites Sichtfeld realisieren zu können. Das Gesichtsfeld liegt bei den heute erhältlichen Modellen zwischen 60° und 90°. Solche Geräte gibt es auch in einer Ausführung bei der die Bildschirme transparent sind, somit bleibt auch die reale Umgebung noch sichtbar und eignen sich besonders für Augmented Reality Anwendungen. Die beschränkte Auflösung der meisten HMDs ist bisher eine der Haupteinschränkungen der VR-Technologie. Neuere Entwicklungen haben auch schon höhere Auflösungen, beispielsweise das Glasstron von Sony [Sony99] oder den HMD von Kaiser Electro-Optics [Kais98].

---

<sup>2</sup> Tracking: Von einem Objekt, die Erfassung seiner Position und zugehöriger Orientierung im Raum.

### 4.1.2 3D-Brille

Eine Alternative zu HMDs sind Shutterglasses, bei denen der Betrachter eine Brille aufsetzt und damit Monitor oder Leinwand betrachtet [Ster97]. Vorteile der Shutterglasses gegenüber den HMDs liegen im Anwendungskomfort und in der grafischen Qualität. Auf dem Monitor werden abwechselnd Bilder für das rechte und linke Auge angezeigt. Durch die Synchronisation von Brille und Computer wird nun erreicht, dass jedes Auge ein etwas anderes Bild auf dem Bildschirm sieht. Hierdurch entsteht die Illusion einer dreidimensionalen Ansicht des Computerbildes.

Der so entstandene stereoskopische Effekt hat aufgrund der Darstellungsweise, die Bilder werden sequentiell und nicht parallel erzeugt, eine halb so hohe Bildwiederholrate wie HMDs. Ein Vorteil von Shutterglasses ist, gerade im Hinblick auf AR-Anwendungen, dass der Benutzer seine reale Umgebung ständig sehen kann und nicht in die virtuelle Umgebung „abtaucht“<sup>3</sup>. Der Effekt ist auch darin begründet, dass bei HMDs das virtuelle Objekt vorne und bei Shutterglasses es hinten dargestellt wird.

### 4.1.3 Augen-Tracker

Der Augen-Tracker kann grundsätzlich nach 3 Verfahren realisiert werden. Bei dem ersten Verfahren können mit Hilfe von Elektroden, die auf die das Auge umgebende Haut aufzusetzen sind, Potentialveränderungen zwischen Horn- und Netzhaut gemessen werden. Mit diesen Werten kann dann auf die Blickrichtung geschlossen werden. Dieses Verfahren eignet sich jedoch besser für die relative denn zu Ermittlung der absoluten Position.

Das zweite Verfahren ist eine genauere Methode. Hier bedient man sich einer Kontaktlinse, die der Wölbung der Pupille angepaßt ist. Durch die Messung der Linsenbewegung läßt sich dann die Blickrichtung ermitteln.

Die dritte Methode besteht in der ständigen Bildanalyse eines bestimmten Merkmals des Auges, um Aufschluß über die Blickrichtung zu erhalten. Ein solches Merkmal kann der Umriß der Pupille oder ein Lichtreflex auf der Hornhaut dienen. Die beiden letzten Verfahren benötigen eine Videokamera und kommen nicht ohne eine Lichtquelle aus und können dadurch der Immersion entgegenwirken.

### 4.1.4 Head-Coupled-Display

Bei manchen Anwendungen (insbesondere im industriellen Bereich) ist es aber zu umständlich sich immer den Datenhelm aufzusetzen und den Handschuh anzuziehen, deshalb wurde eine Kombination aus beiden entwickelt. Sogenannte BOOMs (Binocular Omni Oriented Monitor) sind sowohl visuelle Ausgabegeräte mit den Eigenschaften von HMDs, als auch haptische Ausgabegeräte, da sie ein force feedback über einen entsprechenden Hebel ermöglichen.

In einem handlichen Kasten ist dabei die Optik des Systems untergebracht, die an einem Gestell befestigt ist. Der Vorteil dieser Systeme ist ihre Handlichkeit, da sie einen hohen Grad an Immersion erreichen, wobei in Sekundenschnelle ein Eintauchen in die virtuelle Umgebung durch Heranziehen des BOOM und ein ebenso schnelles Verlassen möglich ist. Der lästige Kabelbaum von HMDs und Kraftrückkopplungs-Geräten ist im Gestell integriert.

---

<sup>3</sup> abtauchen im Sinne von immersive Welt



Aufgrund dieser Eigenschaften sind BOOM's gut für den professionellen Einsatz geeignet und mit dem Nachteil der eingeschränkten Bewegungsfreiheit.

## **4.2 CAVE und Responsive Workbench**

Eine ganz andere Möglichkeit ist der CAVE (Cave Automatic Virtual Environment) [Pape96] mit dessen Hilfe man komplett in einer virtuellen Umgebung arbeiten kann. Hierbei steht der Benutzer oder kleinen Gruppen, in einem kleinen Raum dessen Kantenlänge ca. 3 m beträgt. Die Immersions-Illusion der oder die Benutzer entsteht über die auf die Wände, der Decke und den Boden von mehreren Computern projizierten Stereobilder. Der drei-dimensionale Eindruck wird mit Hilfe der schon erwähnten 3D-Shutterbrille erzeugt. Ergänzend zu den virtuellen Bildern kann der Einsatz von Klängen und taktilen Reizen (z.B. über einen Schwingboden) den Grad der Immersion erhöhen. Manipulationen innerhalb der virtuellen Umgebung werden mit einer Art Zeigestock oder Datenhandschuh ausgeführt.

Um dem Benutzer eine realistische Perspektive zur Verfügung zu stellen, ist es notwendig, mindestens die Kopfposition und -richtung aufzunehmen. Wünschenswert wäre sogar die Verfolgung der Blickrichtung. Benutzt man diese Möglichkeiten, so sind Gruppenarbeiten im CAVE nur noch eingeschränkt möglich.

Anstatt eines CAVE läßt sich auch ein große halbrunde Leinwand aufstellen, die einen Blickwinkel von 180° aufspannt. Man agiert hier nicht in, sondern vor dem virtuellen Bild. Man erreicht das indem man 3 Projektoren benutzt und für jeden Projektor ein eigenes Bild generiert, welches an den Überlappungsstellen übereinstimmen muß. Diese Art eignet sich hervorragend für Fahr- und einfache Flugsimulatoren.

Bei der Herstellung von mechanischen Bauteilen bietet sich durch die Benutzung von CAD-Programmen die Metapher „Werkbank“ an. Schon 1991 stellte M. Krueger in einer Kooperation mit der Stanford University das Konzept vor und nannte es „Videoplace“ [Krue91]. Daraus entwickelte sich die Responsive Workbench. Sie hat eine neigbare Projektionsfläche (von einer Tisch- bis zur vertikaler Leinwandstellung). Auf diese wird ein stereoskopisches Bild projiziert und durch die Shutterbrille erhält man einen dreidimensionalen Eindruck [KrFr94]. Versieht man dies Brille mit Positionsgebern, so kann man um den Tisch herumgehen und bekommt das Bild immer korrekt dargestellt.

## **4.3 Akustische Ausgabegeräte**

Akustische Ausgabe verstärkt die realistische Wirkung von VR-Systemen. Dadurch wird Ortung von Geräuschen, die mit virtuellen Objekten verknüpft sind, und die Kommunikation ermöglicht. Als Ausgabegeräte dienen Kopfhörer oder Lautsprecher Arrays, die mit speziellen 3D Audiodisplays angesteuert werden. Die vom menschlichen Ohr wahrgenommenen Geräusche sind abhängig von dem persönlichen Hörvermögen, Filtereigenschaften der Ohrregion und des Körpers und von den akustischen Reflektionseigenschaften der Objekte, die sich in seiner Umgebung befinden. Die hier herrschenden Zusammenhänge sind aber noch nicht vollständig erforscht, deshalb wird die Generierung von 3D Akustik stark approximiert.

## **4.4 Haptische Ausgabegeräte**

Die Möglichkeit, virtuelle Objekte zu fühlen, kann zu einer merklichen Erhöhung der Effektivität von VR-Anwendungen führen. Objekte können über den Tastsinn identifiziert und ihre Position und Orientierung bestimmt werden. Über Kraftrückkopplung ist das Manipulieren und Bewegen von Objekten möglich.

Nach der Art, wie haptische Ausgabegeräte auf den Menschen wirken unterscheidet man Berührungs- (Touch Feedback) und Krafrückkopplung (Force Feedback).

#### 4.4.1 Taktile Rückkopplung

Anwendung findet haptische Ausgabe z.B. beim Training von Ärzten, die ihre virtuellen Patienten fühlen können, in wissenschaftlichen Visualisierungen oder in Anwendungsgebieten, die in völliger oder teilweiser Dunkelheit liegen. Heutige haptische Ausgabegeräte beschränken sich meist nur auf die Hand und teilweise auf den Arm der Benutzer, denn hier gibt es die höchste Anzahl von Tastrezeptoren im Körper. Systeme, die auch andere Körperteile berücksichtigen sind extrem komplex.

Das sanfte Berühren einer Tischfläche spricht z.B. nur die Tastsensoren in der Hand an und wird vom Touch Feedback simuliert.

Bei der Krafrückkopplung hingegen, wird beim Schlag mit der Hand auf den Tisch dieser eine Kraft entgegengesetzt, um so den Widerstand des Tisches zu simulieren. Ein wesentlicher Aspekt bei haptischen Ausgabegeräten sind die Anzahl der Freiheitsgrade (Degree of Freedom (DOF)), der dem Benutzer gelassen wird.

Touch Feedback Systeme versuchen, die Tastsinne der Hand über pneumatische Kissen, elektrische Stimuli oder über Nadeln anzusprechen. Die früher angedachten Stimulation der Hand durch elektrische Impulse wurde ebenso wie die direkte Erregung von Nervenbahnen aus Sicherheitsgründen verworfen.

Pneumatisches Touch Feedback ist dagegen völlig sicher, denn hier wird ein Handschuh verwendet, der mit kleinen Airbags (Lufttaschen) versehen ist, die sich bei Bedarf schnell aufblähen und zusammenziehen können. Ungefähr 20 dieser Taschen sind auf der Handfläche verteilt, die über ein Control Interface gesteuert werden. Von hier führen zu jeder Tasche zwei kleine Röhrchen, eine davon für das Aufblasen, die andere für das Absaugen der Taschen.

Einen anderen Weg bieten Mikropin Arrays, die auf einem Handschuh unter den Tastsensoren der Hand sitzen. Sie sind matrixartig angeordnet und lassen sich elektrisch ansprechen. Wenn durch diese ein kleiner Strom geschickt wird, werden winzige Stachel angeregt und richten sich auf. Fließt kein Strom mehr, nehmen sie ihre Ausgangsposition wieder ein. Bisher waren diese Matrizen aufgrund der verwendeten Materialien enorm schwer. Durch den Einsatz von neueren Materialien sind diese Systeme nun leichter und lassen sich einfacher tragen.

Um nicht nur den Tastsinn zu stimulieren, sondern auch Temperatur Feedback einzubinden, wird momentan an einem Enhanced Tactile Feedback gearbeitet. Damit läßt sich neben räumlichen Eigenschaften auch die Temperatur von virtuellen Objekten ermitteln. Das steigert den Immersionsgrad noch um einiges.

#### 4.4.2 Force Feedback

Der Force-Feedback Arm wurde ursprünglich für die Steuerung von Robotern entwickelt. Heute dient er dazu, in von Rechnern generierten Welten physikalische Kräfte zu simulieren, die auf die Benutzer einwirken.

Force-Feedback Arme bestehen aus einem gelenkigen Arm, der beliebige Orte innerhalb eines Radius erreichen kann. An diesem Arm ist ein Display angebracht (s.a. BOOM), über das die virtuelle Umgebung und der darin integrierte Arm dargestellt wird. Das System wird mit einem Handgriff gesteuert und das Force Feedback über Motoren realisiert.

Mit diesem Arm werden vier Freiheitsgrade ermöglicht. Leider sind Force Feedback Arme noch sehr teuer. Außerdem bereitet ihre Benutzung teilweise Probleme, so z.B. an den Systemgrenzen zwischen realem und virtuellen Sehen oder wenn der Arbeitseinsatz über dem Kopf liegt.

Für den alltäglichen Gebrauch eignen sich eher Desktop basierte Feedback-Systeme. Hier bieten Joysticks mit Force-Feedback eine unkomplizierte und leichte Lösung, die sich auch schon kommerziell durchgesetzt hat. Wenn auch gewöhnungsbedürftig, ermöglichen Joysticks eine relativ genaue Interaktion mit der virtuellen Umgebung.

Der Rutgers Portable Master ist ein Handschuh, bei dem pneumatische Mikrozyylinder Force Feedback auf die Finger ausüben. Mit Klemmen werden die Finger an die Zylinder angeschlossen, welche über kleine Schläuche von einem Control Interface mit Druck versorgt werden. Somit entfallen lästige Kabel und Elektronik in der Hand des Benutzers. In Kombination mit Datenhandschuhen kann das Greifen und Interagieren von und mit virtuellen Objekten möglich werden.

Den ganzen menschlichen Arm bezieht der Force Arm Master mit ein. Die an der Hand angewandte Technik wird hier auf Schulter, Ellenbogen und Unterarm ausgedehnt.

## **4.5 Zeiger**

Für die Interaktion mit VR-System werden neben diesen klassischen Geräten zur besseren Steuerung in 3D-Umgebungen multidimensionale Interaktionsgeräte eingesetzt. Mit diesen ist direkte Interaktion möglich, das heißt, die Steuerung von Bewegungen in einem dreidimensionalen Raum kann direkt durch ähnliche Bewegungen am Eingabegerät vorgenommen werden. Es handelt sich meist um sogenannte 6D<sup>4</sup>-Geräte, wobei die Translation in alle drei Richtung und die Rotation um eine der drei Hauptachsen als jeweils ein Freiheitsgrad gesehen werden. Die Steuerung eines 3D-Objektes kann mit einer herkömmlich Maus dagegen nur indirekt ausgeführt werden, weil die zweidimensionalen Bewegungen der Maus in den 3D-Raum abgebildet werden muß.

### **4.5.1 3D-Maus**

Multidimensionale Interaktionsgeräte lassen sich grob in zwei Kategorien aufteilen. Zur ersten Kategorie zählen Eingabegeräte, die relative Positionsänderungen liefern, wie Spacemouse und Spaceball und diejenigen Geräte, welche absolute Positionsänderungen zur Verfügung stellen.

Auf der Oberseite der Spacemouse befindet sich ein Puck, der ein kleines Stück in allen drei Dimensionen aus der Ruhe ausgelenkt und um alle drei Achsen etwas gedreht werden kann. Es ist damit möglich, ein Objekt in allen drei Dimensionen gleichzeitig zu rotieren und zu bewegen. An der Vorderseite befinden sich noch neun Tasten, die andere Funktionen übernehmen können.

Das Funktionsprinzip des Spaceballs entspricht dem der Spacemouse. An der Oberseite des Gerätes ist eine Kugel angebracht, mit der auch Rotations- und Translationskoordinaten eingegeben werden können.

---

<sup>4</sup> Anzahl der Freiheitsgrade

## 4.5.2 „magische“ Stäbe

magische Stäbe (im englischen Wands genannt) sind einfache und billige und sehr wirkungsvolle Interaktionsinstrumente. Es sind Trackingsensoren mit frei programmierbaren Funktionstasten, die im Gegensatz zu isometrischen Sensoren frei im Raum beweglich sind.

Zum einen sind es „Flying Joystick“: Wands, die mit einem Pistolengriff mit mehreren Knöpfen, die über Daumen bzw. Zeigefinger betätigt werden können oder auch Poolballs: Hohle Kugeln, die mit einem elektromagnetischen Trackingsensor ausgestattet sind. Ein oder zwei Knöpfe erlauben das Ausführen von Drag-and-Drop-Operationen.

## 4.5.3 Datenhandschuh

Um eine Interaktion mit virtuellen Objekten zu ermöglichen wird im allgemeinen ein Datenhandschuh benutzt. Dieser besteht einerseits aus Sensoren um die Bewegung der Finger zu erfassen und andererseits aus einem Positionssensor (entweder elektromagnetisch, optisch oder Ultraschall) um die Position der Hand zu ermitteln. Es gibt auch ein System, bei dem sich der ganze Körper in einem "Datenanzug" befindet, um so alle Bewegungen des Körpers zu erfassen und in die virtuelle Umgebung umzusetzen.

Ein Datenhandschuh ist ein Handschuh aus Lycra, der die Position der Hand sowie die Krümmung der Finger erfassen kann. Er erlaubt wie kein anderes Eingabegerät, ein intuitives Agieren in virtuelle Welten. Es ist möglich, durch sehr einfache Gesten in virtuellen Räumen zu navigieren und Objekte zu selektieren. Elektromagnetische Sensoren im Handschuh regeln die Positionsbestimmung.

Zur Messung der Fingerkrümmung haben sich verschiedene Methoden herausgebildet. **Fiberoptisch:** Für jedes Gelenk an der Oberseite eines jeden Fingers wird ein Glasfaserkabel in einer Schlaufe angebracht. Die Glasfaser läßt je nach Grad der Krümmung des Fingers mehr oder weniger Licht durch. Mit Hilfe einer Leuchtdiode und eines optischen Sensors kann dies gemessen werden und in den Krümmungswinkel der einzelnen Gelenke umgerechnet werden. Obwohl dieses Prinzip relativ einfach und zuverlässig funktioniert, sind die Glasfaserleitungen für viele Einsatzzwecke zu bruchempfindlich. Außerdem erfordert dieses Meßverfahren bei längerer Anwendung, spätestens jedoch bei Wechsel des Benutzers, eine Rekalibrierung des Handschuhs.

**Elektrisch:** Dieses Verfahren arbeitet mit Metallbändern, die oberhalb der Finger angebracht werden und proportional mit dem Krümmungswinkel ihren Widerstand ändern. Dieses Verfahren ist genauer als das fiberoptische Verfahren.

**Mechanisch:** Ein oberhalb des Handschuhs angebrachtes mechanisches Exoskelett erfaßt mit hoher Präzision die Krümmung eines jeden einzelnen Fingergelenks. Für den Normaleinsatz sind solche Geräte zu teuer und zu umständlich in der Handhabung.

Obwohl der Datenhandschuh ein sehr leichtes und angenehmes Eingabegerät darstellt, ist die Navigation manchmal sehr ermüdend. Außerdem kann er wegen der Hygiene und Passgenauigkeit nur schlecht von mehreren Benutzern verwendet werden. Weiterhin kann je nach Messverfahren die Navigation sehr langsam und ungenau sein.

## 4.5.4 Datenanzug

Diese Geräte übertragen die Fähigkeiten des Datenhandschuhs auf den ganzen Körper. Mit dem fiberoptischen Verfahren werden die Krümmungswinkel der wichtigsten Gelenke des Körpers gemessen. Elektromagnetische Sensoren bestimmen die Position wichtigster Körperpartien im Raum. Die fiberoptischen Sensoren sind leider sehr empfindlich und somit besteht ständig die Gefahr, dass sie durch Bruch nicht mehr benutzbar sind. Außerdem kann ein Verrutschen des Anzugs enorme Meßfehler verursachen. Der Benutzer hat auch durch die schwere Kabeln, die zur Übermittlung der Sensordaten dienen, eine eingeschränkte Bewegungsfreiheit. Datenanzüge wurden meistens in der Vergangenheit für Motion Tracking bei der Charakter-Animation eingesetzt.

#### 4.6 Erfassen der Position und -orientierung (Tracking)

Zur zweiten Kategorie zählen Interaktionsgeräte, mit denen absolute Positionen und Richtungen direkt eingegeben werden. Mit Hilfe von Trackingsystemen wird die Position und Richtung des Benutzers oder von Objekten erfaßt. Ein ideales Trackingsystem gibt es nicht, da jedes System Vor- und Nachteile aufweist. Die Leistung solcher Systeme ist von verschiedenen Faktoren abhängig [Henn97]:

- **Timelag:** Verzögerungszeit zwischen Sensorbewegung und resultierendem Signal:  
Gängige Werte: 4-100 ms  
Idealwerte: <5 ms
- **Update-Rate:** Die Frequenz, in der die aktuellen Werte vom Sensor geliefert werden.  
Gängige Werte: 10 - 60 Hz  
Idealwerte: >60 Hz
- **Genauigkeit:** Meßtoleranz des Trackingverfahrens.  
Gängige Werte:  $\pm 0,125$  bis  $\pm 3,2$  mm /  $\pm 0,05^\circ$  bis  $\pm 0,5^\circ$   
Idealwerte: < 0,5mm / < 0,005°
- **Auflösung:** Genauigkeit des Sensors in Positionierung und Orientierung.  
Gängige Werte: 0,25 bis 6,35 mm /  $0,1^\circ$ -  $1,0^\circ$   
Idealwerte: < 1 mm / <  $0,01^\circ$
- **Reichweite:** Maximaler Abstand zwischen Sender und Empfänger bei drahtlosen Verfahren  
Gängige Werte: 45 - 500 cm  
Idealwerte: 500 cm je nach Anwendung
- **Interferenz:** Abhängigkeit des Meßverfahren gegenüber äußeren Einflüssen, wie z.B. elektromagnetischen Feldern, Metallen, Störsignalen und weiteren Sensoren des gleichen Typs.

Hierzu zählen folgende Eingabegeräte:

##### 4.6.1 Mechanisch

Bei diesem Trackingverfahren besteht immer eine mechanische Verbindung zwischen dem verfolgten Objekt und einem fixen Referenzpunkt, der Basis. Die Basis ist fest an der Wand oder Decke montiert und mehrere Gelenkstücke gehen von ihr ab. Die freien Enden erlauben die Bewegung in vielen Freiheitsgraden. Die Winkel der einzelnen Gelenke werden mit Hilfe

von Sensoren gemessen und die daraus berechnete Position bildet die Grundlage für die Positionsbestimmung des darauffolgenden Gelenks.

#### 4.6.2 Magnetisch

Vom einem Sender werden durch drei Spulen in sequenzieller Folge drei senkrecht aufeinanderstehende Magnetfelder erzeugt. Diese induzieren bei den drei in ähnlicher Weise ausgerichteten Spulen des Empfängers Strom. Dadurch entstehen neue Meßwerte, die von einem Controller in räumliche Positionen und Ausrichtungen umgerechnet werden können. Der Bezugspunkt aller Messungen wird durch die Position des Empfängers bei der Initialisierung der Messung festgelegt.

Wegen möglicher Interferenzen muß häufig für jeden Empfänger eine getrennte Messung durchgeführt werden, was die Verzögerungszeit enorm vervielfacht. Ein Vorteil dieser Tracker ist jedoch, dass sie berührungslos sind und dadurch eine große Bewegungsfreiheit ermöglichen. Weiterhin sind die Empfänger besonders klein und können überall befestigt werden. Sie reagieren aber auf elektromagnetische Felder wie auch auf Metalle in der näheren Umgebung, so dass Genauigkeit und Geschwindigkeit durch diese äußeren Störfaktoren negativ beeinflußt werden.

#### 4.6.3 Optisch

Unter dem Begriff „optisches Tracking“ sind mehrere Messverfahren gemeint.

##### *Aktive Trackingverfahren:*

An dem Objekt sind Infrarot-Dioden angebracht, die von einem Empfänger, der meistens an der Decke montiert wird, erfaßt werden. Die Kamera ist am verfolgten Objekt fixiert und sie registriert ebenfalls mehrere gleichmäßig an der Decke angebrachte Leuchtdioden, die zur besseren Erkennbarkeit in regelmäßigen Abständen blinken. Anhand der Lage einiger Lichter kann die Position der Kamera, ermittelt werden.

##### *Passive Trackingverfahren:*

Auf dem Objekt sind reflektierende oder auch nur farbige Markierungen angebracht, die von der Kamera erfaßt werden und über Bilderkennungsalgorithmen in Position und Ausrichtung des verfolgten Objekts umgesetzt werden. Position und Ausrichtung des verfolgten Objekts werden anhand von besonderen optischen Merkmalen der Marker bestimmt.

Aktive Trackingverfahren werden wegen des hohen technischen Aufwands nur in bestimmten Fällen eingesetzt, wohingegen die passiven Methoden sehr flexibel und preisgünstig sind und sich zu einer angenehmen Positionsverfolgung, die berührungs- und kabellos ist, entwickeln können. Da die entsprechende Hardware und Software für solche Systeme noch nicht richtig ausgereift ist, entstehen durch die Bildanalyse enorme Verzögerungszeiten und Ungenauigkeiten, die aber bei einer Weiterentwicklung der eingesetzten Komponenten minimiert werden können und somit eine sehr effizientes und intuitives Interface schaffen.

#### 4.6.4 Kameraüberwachung

Mit Hilfe von 3 CCD-Kameras, welche im Raum verteilt sind, läßt sich eine Positionsbestimmung des Körpers durchführen. Eine Auswertung der 3 Bilder mit Hilfe von Bildauswertungsverfahren liefert dann den genauen Ort des Objektes.

## 4.6.5 Kinematisch

Kinematische Trackingverfahren basieren auf der Messung von Rotationsbewegungen anhand der Trägheit der Masse. Die gängigsten Verfahren arbeiten mit einem kleinen Kreisel, der auf eine hohe Rotationsgeschwindigkeit gebracht wird. Sensoren erfassen Änderungen der Rotationsgeschwindigkeit und lassen so Rückschlüsse auf die Lage zu. Solche Kreisel driften aber mit der Zeit gegenüber der exakten Position und müssen deshalb ständig nachgeführt werden. Ein direktes Messen von Translationen ist mit diesem System nicht möglich. Allerdings kann durch Inertialmessung eine Translation ermittelt werden.

## 4.6.6 GPS

Eine Positionsbestimmung kann auch mit Hilfe des satellitengebundenen Global Positioning Systems (GPS) und kleinen, leichten GPS-Empfängern durchgeführt werden. Dieses System setzt Satelliten ein und jeweils 4 werden zur Standortbestimmung benötigt. Eine Genauigkeit kann so auf wenige Zentimeter genau sein. Durch Differential-GPS (ein Sender ist fest montiert und dadurch in der Position genau bekannt) kann die Genauigkeit noch erhöht werden. Durch die Ausbreitungscharakteristik der Sender kann GPS nicht in Innenräumen verwendet werden. Gegen eine Benutzung spricht auch, dass die Satelliten und deren Betrieb störanfällig sind und die erwartete Genauigkeit nicht immer und zu jeder Zeit gewährleistet werden kann.

# 5 Technische Eigenschaften

In diesem Kapitel sollen die besonderen Eigenschaften beschrieben werden, mit der man ein Augmented Reality System aufbauen kann.

## 5.1 Grundsätzliches zu Augmentation

Neben den grundsätzlichen Anforderungen, Objekte den realen Umgebungen hinzuzufügen, muss Augmented Reality das Potential haben, Objekte zu entfernen. Die gegenwärtige Arbeit konzentriert sich darauf, den realen Umgebungen virtuelle Objekte hinzuzufügen. Allerdings kann man auch mit Bitmaps einen Teil der realen Welt vor dem Benutzer „verstecken“. Diese Art benutzt man, um z.B. in Dokumentarfilmen Artefakte neueren Datums zu überdecken. Diese Anwendung interaktiv in der AR (d.h. in Echtzeit) durchzuführen, ist aber sehr viel schwieriger.

Augmented Reality kann auf alle Sinne angewendet werden, nicht nur auf das Sehvermögen. Bisher hat man sich hauptsächlich um das Vermischen von Pixelbildern und Grafiken beschränkt. Akustische Ausgabe ist jedoch entscheidend für mehr Realismus von AR-Systemen. Durch das zur Verfügungstellen von wichtigen akustischen Informationen wird Ortung von Geräuschen, die mit virtuellen Objekten verknüpft sind, und Kommunikation ermöglicht. Wie weiter oben schon beschrieben, nimmt das menschliche Ohr nur subjektive Geräusche wahr und ist abhängig von dem persönlichen Hörvermögen und Filtereigenschaften der Ohrregion und des Körpers sowie den akustischen Reflektionseigenschaften der Objekte, die sich in seiner Umgebung befinden. Ein Mikrofon nimmt dazu die Geräusche von außen auf. Die hier herrschenden Zusammenhänge sind aber noch nicht vollständig verstanden, deshalb wird die Generierung von 3D Akustik stark vereinfacht.

Ein anderes Beispiel zur Erhöhung Realität ist die Möglichkeit, virtuelle Objekte zu fühlen und kann zur einer markanten Erhöhung der Effektivität von VR-Anwendungen führen. Objekte können über den Tastsinn identifiziert und ihre Position und Orientierung bestimmt werden. Über Kraftrückkopplung ist das Manipulieren und Bewegen von Objekten möglich.

Ein anderes Beispiel zur Erhöhung der Effektivität von VR-Systemen ist die Abstimmung von Bildern und Bewegung. Diese Abstimmung bewirkt eine gegenseitige Verstärkung und Plausibilität der Virtual Reality.

## 5.2 Optisches oder videobasiertes System

Eine grundsätzliche Entwurfsentscheidung zur Erstellung von AR-Systemen ist, wie erfüllt man die Kombination von der realen und der virtuellen Welt. Zwei Technologien stehen zur Auswahl: ein optisches<sup>5</sup> oder ein Video System. Beide haben ihre speziellen Vor- und Nachteile.

Eine Möglichkeit ist ein HMD mit einem halbdurchlässigem Spiegel. Damit lassen sich die reale und virtuelle Welt hervorragend kombinieren, wobei die virtuelle Welt der realen Welt überlagert wird. Dies kann zunächst durch optischen oder ein Videobasierten Systemen geschehen.

Bei sogenannten optischen HMD mit halbdurchlässigem Spiegel wird der Spiegel vor die Augen gesetzt. Damit kann man ganz normal die Umwelt erkennen, andererseits ist die Möglichkeit gegeben, auf diese Spiegel ein vom Rechner generierte Grafik zu überlagern. Damit sieht der Benutzer beides, die Umwelt sowie das generierte Computerbild. Der Ansatz ist ähnlich dem Head Up Display (HUD) in militärischen Flugzeugen außer, das anders wie beim Flugzeug keine Frontscheibe zur Darstellung benutzt wird. Abb. 1 zeigt ein Entwurfsmuster für ein HMD mit einem halbdurchlässigem Spiegel. Ein reales Head Mounted Display mit Tracker zeigt Abb. 2.

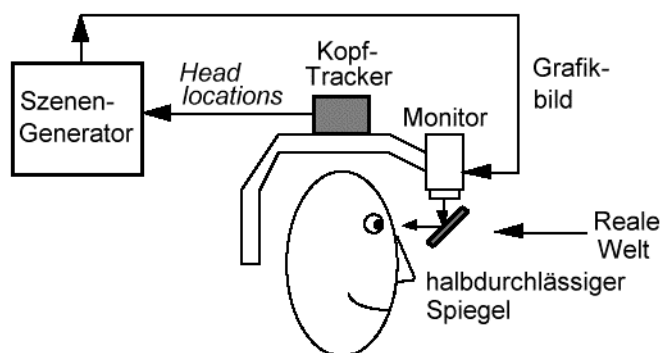


Abb. 1: Schemadarstellung

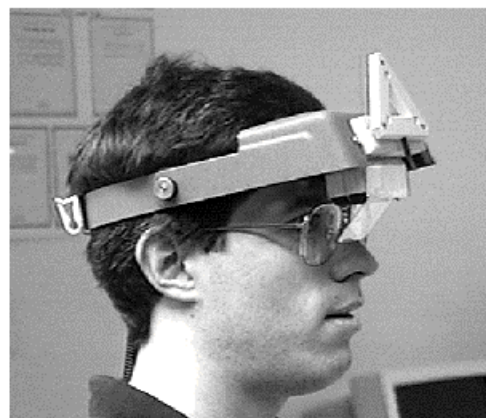


Abb. 2: Aktuelle Ausführung

Die halbdurchlässigen Spiegel haben allerdings den Nachteil, nur einen Teil des Lichtes durchzulassen. Andererseits muß die eingespiegelte Grafik hell genug sein, um gegen die Umgebung erkannt zu werden. Man kann sich einen HMD vorstellen, der eine Abdeckung von 70% erzielt, d.h. nur 30% des einfallenden Lichtes werden durch die Spiegel weitergegeben. Der Grad der Überblendung ist ein Entwurfsproblem. Es ist vorstellbar, dass weiterentwickelte HMD die Anteile der jeweiligen Bilder auf Grund der Wellenlänge variieren können. Zum Beispiel könnte so ein Spiegel so gesetzt sein, dass er Licht einer bestimmten Wellenlänge reflektiert, während das Licht der übrigen Wellenlängen unberührt bleibt. Dies wäre ideal für monochrome Monitore und man könnte bestimmte Wellenlängen komplett reflektieren und die übrigen dämpfen. Allerdings dämpfen alle z.Zt. existierenden HMD das gesamte Licht zu stark für Innenräume.

<sup>5</sup> In diesem Zusammenhang wird im Englischen auch „see-through“ Systemen gesprochen



Im Gegensatz dazu arbeiten geschlossene HMD mit Videokameras um die Umwelt aufzunehmen. Das Videosignal der Kamera wird den Bildern der virtuellen Welt überlagert, welches von einem Programm generiert wird. Das Ergebnis wird über die beiden Monitore ausgegeben. Abb. 3 zeigt ein Prinzipbild und in Abb. 4 wird ein aktuelles Head Mounted Display abgebildet.

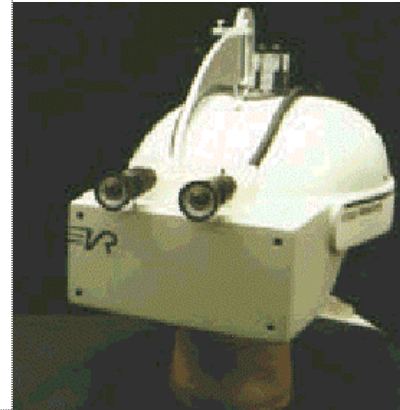
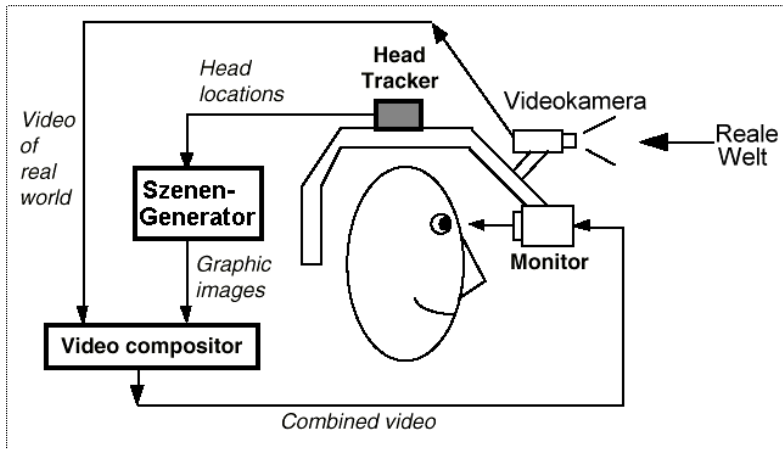


Abb. 3: Schema eines videobasierten HMD

Abb.4: Realisierte Ausführung

Die Mischung der Videosignale kann auf verschiedenen Arten geschehen. Das Chroma-key oder auch Blue Screen Verfahren, ist eines der mächtigsten Verfahren mit dem man Echtzeit Videoaufnahmen mit Computer generierten Hintergrundbildern mischen kann. Das Verfahren sucht nach einer bestimmten Farbe und tauscht dieses gegen das Hintergrundbild aus. Meistens nimmt man dazu die Farbe blau, weil „blau“ als reine Farbe in den menschlichen Hauttönen kaum vorkommt, siehe dazu auch [Bric97]. Da man bei diesem Verfahren mit 2 Bildern arbeiten muß, bietet es sich für AR-Systeme an. Bei geschickter Wahl von einzusetzenden Lichtquellen, kann das resultierende Bild sehr real erscheinen.

Eine weiterentwickelte Methode benutzt die Tiefeninformation der Szene. In diesem Verfahren erhält jedes Pixel eine Tiefeninformation der aufgenommenen Szene. Kombiniert man nun und vergleicht Pixel für Pixel, so kann man schnell entscheiden, welche Teile vorne oder hinten liegen.

Eine preiswerte Alternative zu HMDs sind Shutterglases, bei denen der Betrachter eine Brille aufsetzt und damit Monitor oder Leinwand betrachtet, siehe Abb. 5. Auf dieser werden

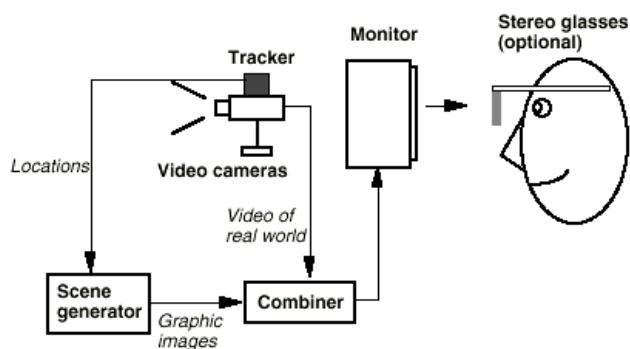


Abb. 5: Prinzipieller Aufbau mit Shutterglases

Abb. 6: Realisierte Ausführung

abwechselnd Bilder für das rechte und linke Auge angezeigt. Synchron dazu schaltet die Shutterbrille jeweils ein Augenglas transparent und das andere undurchsichtig. Der so

entstandene stereoskopische Effekt hat aufgrund der Darstellungsweise eine meist halb so hohe Bildwiederholrate wie HMDs. Ein erheblicher Vorteil von Shutterglases ist, dass der Benutzer seine reale Umgebung ständig sehen kann und nicht in die virtuelle Umgebung „abtaucht“ Abb. 6 zeigt eine reale Umgebung mit einer Shutterbrille. Solche Anwendungen sind ideal für CAD-Applikationen, man sieht das Objekt gleich in der 3D-Umgebung.

### 5.2.1 Stärken und Schwächen

Wie jedes System gibt es auch hier erhebliche Unterschiede. Im folgenden werden die Vor- und Nachteile der optischen, bzw. Videosysteme aufgeführt und eine Vergleich der Systeme vorgenommen.

#### *Einfachheit:*

Optisches Mischen ist einfacher und billiger als das Mischen mit Videosignalen und man braucht nur einen Videokanal zu betrachten. Die reale Welt wird direkt gesehen und der Zeitverzug ist im allgemeinen vernachlässigbar. Der Videokanal hingegen muß 2 unterschiedliche Eingabeströme bearbeiten, nämlich einen für das virtuelle Bild und den anderen für das Bild von der Außenwelt. Beide Ströme haben einen inhärenten Verzug von mehreren Millisekunden. Das Digitalisieren von Videobildern fügt mindestens eine Verzögerung von einer Bildwiederholrate<sup>6</sup> dem Videostrom hinzu. Ein Monitor mit einer Bildwiederholrate von 75 Hz hat aber eine Bildwiederholrate von 13.3 msec. Die beiden Eingabekanäle müssen nun genau synchronisiert werden, damit keine zeitliche Verzerrungen auftreten. Außerdem haben optische HMD mit kleinen Blickwinkeln (Field of View (FOV)) auch nur geringe Verzerrungen zur realen Welt. Videokameras hingegen haben durch ihr Linsensystem immer Verzerrungen, die kompensiert werden müssen. Außerdem sind optische HMD preisgünstiger, weil keine zusätzliche Videokamera und keine elektronische Verknüpfung benötigt werden.

#### *Auflösung:*

Elektronisches Mischen der Videosignale begrenzt die Auflösung für beide Systeme. Für die heutigen Monitore ist die Auflösung von 1024 Bildpunkten in horizontaler Richtung gut, aber immer noch wesentlich größer als das Auflösungsvermögen des Auges. Die optischen Durchsicht-Geräte zeigen das virtuelle Bild in der Auflösung des Anzeigegerätes wobei der Blick des Benutzers durch den Spiegel nicht eingeschränkt ist.

#### *Sicherheit:*

HMD mit Kamerasystemen sehen die Außenwelt nur durch diese Kameras. Sollte der Strom ausfallen, sind die Träger der Helme praktisch blind. Tritt dies bei HMD mit halbdurchlässigem Spiegel auf, so kann der Benutzer die Welt noch ganz normal sehen ohne die Unterstützung der virtuellen Welt. Dies kann zu Sicherheitsbedenken bei einigen Anwendungen führen. Andererseits kann man mit Kamerasystemen viel eher eine Lichtanpassung und ein Ausblenden von Objekten erreichen

#### *Augenversatz:*

Mit Video-HMD wird die Sicht auf die Welt durch Videokameras durchgeführt. Im wesentlichen bedeutet das, dass mit den heutigen Helmsystemen dem Benutzer die virtuellen

---

<sup>6</sup> Bildwiederholrate ist die Zeit, die das System braucht um ein komplettes Bild neu zu zeichnen.

Augen zu seinen eigenen versetzt sind. Gleichzeitig paßt der menschliche Augenabstand nicht zum Augenabstand des Helmes. Die Differenzen in der Kameraposition und den Augen verursachen Verschiebungen zu dem was der Benutzer sieht, bzw. zu dem, was der Benutzer erwartet. Ist zum Beispiel die Kamera oberhalb seiner Augen, sieht er die Welt von einem etwas höherem Standpunkt aus als dem, den er gewohnt ist. Mit Hilfe von Spiegeln kann man die unterschiedlichen Augenpositionen wieder in Übereinstimmung bringen. Das fügt allerdings dem Helm einiges an Komplexität hinzu.

#### *Flexibilität in der Zusammensetzungsstrategie*

Ein grundlegendes Problem ist die nicht vorhandene Überdeckung der virtuellen Welt über die reale. Mit einem optischen HMD ist es schwierig, selektiv Licht von der Außenwelt auszuschließen. Wie in allen optischen Systemen wird auch hier das Bild in einem Punkt fokussiert, nämlich auf das Auge des Benutzers. Jeder Filter, welches ein partielles Abdunkeln bewirken soll, muß aber auf das Abbild fokussiert sein. Somit erhält man 2 Punkte auf der die Bilder fokussiert werden müssen. Dies führt dazu, dass die optischen HMD sehr viel schwieriger und komplexer konstruiert werden müssen. Kein bekanntes System arbeitet in dieser Weise. D.h., die virtuellen Objekte erscheinen schemenhaft und halbdurchsichtig. Das stört die Illusion der Realität, denn die Überdeckung von Objekten ist die stärkste Tiefeninformation des Menschen.

Im Vergleich dazu sind die videobasierten HMD hier flexibler, denn beide Bilder liegen in digitaler Form vor. Sie lassen sie somit leichter nach bestimmten Algorithmen kombinieren. Durch diese Flexibilität lassen sich schließlich überzeugendere virtuelle Umgebungen erstellen.

#### *Blickwinkel*

Verzerrungen in optischen Systemen sind eine Funktion des radialen Abstandes von der optischen Achse. Je weiter man vom Bildmittelpunkt entfernt schaut, desto größer ist die Verzerrung. Ein digitales Bild von einem verzerrtem optischen System, kann durch geeignete Bildverarbeitungstechnologien verzerrungsfrei dargestellt werden. Das erhöht allerdings die Rechenzeit, könnte aber in Zukunft keine Beschränkung sein, bedingt aber ein physikalisches Kameramodell.

Es ist schwierig ein optisches System mit einem großen Blickwinkel herzustellen [Kühl98]. Jede Verzerrung in diesem System muss mit optischen Mitteln korrigiert werden. Komplizierte Optik ist aber sehr teuer und auf Grund der Linsensysteme relativ schwer im Vergleich zu Videosystemen. Wünscht man Systeme mit großem Blickwinkel, sind Video-HMD die bessere Wahl.

#### *Verzögerungen im realen und virtuellem Zweig können in übereinstimmen*

Video-HMD erlauben aber einen Ansatz, die unterschiedlichen Verzögerungen zwischen dem realen Videobild und dem virtuellen Videobild zu minimieren oder ganz zu vermeiden. Optische Systeme haben für die reale Welt das sofort erhältliche Bild, aber ein verzögertes für das virtuelle Bild. Das kann bei der Benutzung zu Problemen führen. Bei Video-HMD kann man für beide Bilder die Verzögerungszeiten synchronisieren, liegt aber dafür um eine bestimmte Zeit hinter der Wirklichkeit zurück.

#### *Zusätzliche Registrierungsstrategie*

In einem optischen System kommen die Lage- und Positionsinformationen ausschließlich vom Headtracker. Bei Videosystemen kann man zusätzlich dazu noch das digitale Bild der Umgebung hinzuziehen. Durch Bildverarbeitungsmethoden lassen sich herausragende Punkte erkennen und den vorhandenen Positionsinformationen hinzufügen. Dieses Problem wird in Kapitel 6.1 näher beschrieben.

### *Leichtere Anpassung an die Helligkeit*

Eine nähere Beschreibung ist im folgenden Abschnitt gegeben. Beide Systeme haben ihre Berechtigung. Die Wahl, zu welchem System man sich entscheidet, hängt stark von der Anwendung ab.

## 5.2.2 Akkomodation und Kontrast

Die Fokussierung auf ein Objekt ist distanzabhängig. Beim natürlichen binokularen Sehen ist die Akkomodation direkt an die Konvergenz gekoppelt und spielt so bei der Tiefenwahrnehmung eine entscheidende Rolle. Da bei HMD der Bildschirm immer in gleicher Distanz zum Auge ist, muß der Benutzer lernen, beide Mechanismen zu trennen. Dies ist nicht nur ungewohnt, sondern stört auch die Tiefenwahrnehmung und kann für Ungeübte schnell zu einer Übermüdung des Augenapparates führen.

Da bei allen HMD der Abstand zwischen Auge und Bildebene immer konstant bleibt, wäre allenfalls eine Analyse der Fokussierung der Augen und eine softwaregesteuerte Anpassung der Tiefenschärfe in Echtzeit die einzige zur Zeit praktikable Alternative [Stur91].

Im Gegensatz dazu ist es ein zusätzliches Problem, dass das menschliche Auge einen großen dynamischen Bereich abdecken kann und was das Auge daraus erkennt. Im Idealfall sollte die Helligkeit beider Bilder annähernd gleich sein. Im schlimmsten Fall, muß ein großer Helligkeitsunterschied dargestellt werden. Die Sehnerven arbeitet wie ein logarithmischer Detektor [Snyd88], so kann das hellste Licht um 11 Größenordnungen größer sein als das dunkelste Licht. In einer Adaption kann es noch immer 6 Größenordnungen überdecken. Die meisten Monitore kommen nicht annähernd an diese Größenordnungen. Das ist ein besonderes Problem mit optischen HMD. Hier hat der Benutzer einen direkten Blick auf die Umwelt. Ist die Umwelt zu hell, so ist das virtuelle Bild kaum sichtbar. Ist die Umwelt zu dunkel, dominiert das virtuelle Bild. Bei Videosystemen ist das Problem nicht so groß, weil beide Bilder über einen Videokanal verlaufen. Da auch schon die Videokamera einen eingeschränkten dynamischen Bereich hat, fallen erst gar nicht große Kontrastunterschiede an.

## 5.2.3 Übergang von der virtuellen zur realen Welt

In fast allen virtuellen Umgebungen (VE) wird der Benutzer nicht ermutigt, durch Szene zu gehen. Eher wird eine Art „fliegen“ dem Benutzer abverlangt, ein Laufband zu bewegen oder in einem Mockup zu sitzen. Egal, welche Möglichkeit man ihm bietet, der Benutzer steht auf einem Fleck in der realen Welt und der Rest bewegt sich.

Einige AR-Applikationen benötigen allerdings eine Unterstützung, um in großen Umgebungen umher laufen zu können, beispielsweise wenn man an die Unterstützung der Wartung von Flugzeugen denkt. AR-Systeme erfordern jedoch, dass die Benutzer dort sind,

wo die Aktion stattfinden soll. Aber dorthin „fliegen“<sup>7</sup> ist keine echte Option. Falls die Applikation es erfordert, muß man auch um das virtuelle Objekt herumgehen können.

Deshalb muß ein AR-System ein gewisses Maß an Beweglichkeit gewährleisten, besonders wenn es darum geht, im freien Feld zu agieren. Die Erzeugung der Szene, das Head-Mounted-Display und die Lagebestimmung muß selbstständig und unabhängig sein, um im realen Gelände wirkungsvoll arbeiten zu können. Anwendungen, selbst für das Museumsumfeld sind vorstellbar.

### **5.3 Vergleich zu Virtuellen Umgebungen**

Im Vergleich zu Virtuellen Umgebungen (VE) können die Anforderungen der AR-Systeme auf 3 grundlegende Teilsysteme summiert werden.

#### **Bild- oder Szenen-Generator**

Die Bildwiedergabe ist kein originäres Problem der AR. In VE-Systemen sind die Anforderung ungleich höher, weil sie komplett die reale Welt gegen die virtuelle Welt austauschen. In AR sind die virtuellen Bilder nur eine Ergänzung zur realen Welt. Deshalb werden auch weniger virtuelle Objekte benötigt und diese müssen noch nicht mal aufwendig gerendert werden, um der Anwendung genüge zu tun. Beispielsweise sind für Anmerkungen an den Objekten kein aufwendiges Grafikbild notwendig, häufig reichen hier Drahtgitterdarstellungen.

#### **Monitore**

Auch hier gilt das, ähnlich dem oben beschriebenen. Es gilt ebenfalls, dass AR die reale Welt nicht ersetzt. Beispielsweise könnten schwarz-weiß-Monitore adäquat für einige Anwendungen sein. In VE dagegen, kommt man nicht ohne große Farbtiefe aus. Bei optischen HMD reicht es aus, einen schmalen Blickwinkel zu haben, weil der Benutzer durch den halbdurchlässigen Spiegel auch noch den Rand seines Gesichtsfeldes wahrnehmen kann. Darüber hinaus kann die Auflösung bei optischen Systemen kleiner sein als man im VE-System tolerieren würde.

#### **Lage- und Positionsverfolgung**

In den 2 vorhergegangenen Punkten hatte AR die kleineren Anforderungen im Vergleich zu VE. Dies gilt nicht für die Lage- und Positionsverfolgung. Auf diesem Gebiet sind die Anforderungen an AR wesentlich strenger. Der Hauptgrund ist das Erfassungsproblem, was im folgenden Abschnitt getrennt behandelt wird.

## **6 Übereinstimmung mit der Wirklichkeit**

### **6.1 Erfassungsprobleme**

Das größte Problem in der Augmented Reality ist das Erfassungsproblem (Registration Problem). Die Objekte der realen und der virtuellen Welt müssen genau übereinstimmen oder die Illusion bezüglich einer erweiterten Welt geht verloren. Viel schwerwiegender ist jedoch, dass einige Applikationen genau dies erfordern. Am Beispiel eines Arztes, der gerade eine Nierenoperation übt, wird das deutlich.

---

<sup>7</sup> Fliegen in dem Sinne, dass der Benutzer sich fest an einem Ort befindet und das Objekt sich um ihn herum oder man sich gar durch die Objekte hindurch bewegt.

Ähnliche Probleme existieren auch in virtuellen Umgebungen, sie sind allerdings ungleich schwerer zu entdecken. Die Diskrepanz zwischen physikalischer Bewegung und der simulierten Bewegung in einer VR-Umgebung wird als „kinästhetische Dissonanz“ bezeichnet [Wals91]. Sie tritt beispielsweise dann auf, wenn der Benutzer sitzt, sein Computermodell aber steht. Weil der Mensch auf die kinästhetische Dissonanz weniger empfindlich reagiert wie gerade bei visuellen Gegensätzen sind solche Ungenauigkeiten tolerierbar.

Beispielsweise erwarten Benutzer mit einem geschlossenem HMD eine eigene, selbst erhobene Hand in ihrem virtuellem Gesichtsfeld. Die virtuelle Hand sollte schon so aussehen, auch von der Handhaltung her, wie man es auch ohne Helm erwartet. Ist diese Hand um einige Millimeter verschoben, so fällt das nicht weiter auf, außer man sucht nach solchen Fehlern. Tritt der diese Diskrepanz in einem optischem System auf, fällt dieser Fehler sofort auf. Visuelle Fehler haben ein deutlich höhere Signifikanz.

Allgemein kann gesagt werden, dass die Fähigkeit des Menschen, „sensorische Lücken“ bei der Wahrnehmung eigenständig zu füllen, um sich somit eine „stabile Wirklichkeit“ [vonF73] zu geben, die geringere Leistungsfähigkeit der anderen menschlichen Sensoren kompensieren. Beispielsweise glaubt man ja auch im Kino, dass genau der Schauspieler spricht, dessen Mund sich bewegt, obwohl man weiß, die Töne kommen aus einem Lautsprecher. Ähnliches passiert mit der Hand, die man sieht und nur glaubt, sie sei an dieser Stelle. Dieser Effekt erhöht die Toleranzgrenzen bei Erfassungsproblemen in virtuellen Umgebungen um einiges. Ist der Fehler systematisch und ist der Benutzer für einige Zeit in der virtuellen Umgebung, kann man davon ausgehen, dass eine Anpassung stattfindet.

Augmented Reality hat wesentlich höhere Anforderungen auf möglichst fehlerfreie Erfassung als die Virtual Environments [Azum93]. Nimmt man das gleiche Beispiel wie oben, dann wird deutlich, dass bei einer Benutzung optischer HMD die Differenz zwischen der realen und virtuellen Hand sofort und deutlich auffällt. Solche Widersprüche sind auf Grund der höheren Auflösung und höherer Empfindlichkeit des Auges einfacher zu entdecken. Selbst kleinste Unterschiede sind erkennbar.

Aus Untersuchungen des Luftfahrtbundesamtes [fsm81] ist bekannt, dass das menschliche Auge nur eine begrenzte Auflösung hat. Welche Auswirkung hat das für die Erfassung von Objekten ? An einem Beispiel soll die Bedeutung verdeutlicht werden: Ein 10-Pfennig-Stück halte man ungefähr in einem Meter Abstand so, dass das Geldstück als Kreis zu sehen ist. Der Durchmesser des Geldstücks überstreicht dann gerade mal 1.2 bis 2.0 Bogenminuten. Im Vergleich dazu überdeckt der Mond 0.5 Bogenminuten. Nun soll ein virtuelles Objekt auf ein reales Objekt überlagert werden und wird um den Durchmesser des Mondes versetzt. So etwas würde man sofort entdecken, also kann man daraus schliessen, dass die geforderte Genauigkeit kleiner als einen Bruchteil einer Bogenminute sein sollte. Die untere Grenze wird durch das Auflösungsvermögen des Auges beschränkt. In [fsm81] wird ausgeführt, dass das Auge nur eine Auflösung von einer halben Bogenminute ermöglicht. Existierende HMD haben aber nur eine Auflösung von mehr als einer Bogenminute, so dass die erreichbare Genauigkeit schlechter als die des Auges ist. In der Praxis sind Fehler von wenigen Pixel bei modernen HMD erkennbar.

Die Erfassung und Lokalisierung von realen und virtuellen Objekten ist nicht nur auf AR beschränkt. Im Film und im Fernsehen werden virtuelle Objekte und reale Schauspieler häufig verknüpft. Der Unterschied liegt in der Art der Kontrolle, die ein Regisseur hat. Hier kann sorgfältig geplant werden und die Akteure proben einige Stunden bevor eine Aufnahme

fertig ist. Als interaktives System kann man in einem AR-System nicht warten bis alles bereit ist. Hier muß schnell und akkurat gehandelt werden.

Erfassungsfehler sind wegen der vielen Genauigkeitsanforderungen und den vielen Fehlerquellen schwierig angemessen zu kontrollieren. Man kann diese Fehler in zwei Typen aufspalten, statische und dynamische Fehler. Statische Fehler sind solche, die auftreten auch wenn das System in Ruhe ist. Dynamische Fehler werden nur entdeckt, wenn sich die Blickrichtung ändert oder die Objekte sich bewegen.

Für heutige HMD-basierten Systeme läßt sich feststellen, dass dynamische Fehler bei weitem den größten Anteil am Erfassungsproblem haben, doch darf man statische Fehler nicht außer Acht lassen.

## **6.2 Statische Fehler**

Im folgenden sollen die wichtigsten 4 Quellen für statische Fehler aufgeführt werden, welche in den nächsten Kapiteln behandelt werden.

### **6.2.1 Optische Verzerrungen**

Optische Verzerrungen existieren in allen optischen Geräten und natürlich auch in den Geräten, die die reale Welt, bzw. die virtuelle Welt ausgeben. Typischerweise sind solche Fehler eine Funktion des Abstandes von der Sehlinie, so dass bei großen Blickwinkeln die Gefahr für Verzerrungen groß ist. Um den Mittelpunkt des Sehstrahles sind die Verzerrungen klein, je weiter aber die Objekte vom Mittelpunkt entfernt sind, desto größer können die Verzerrungen sein. In einem optischen HMD mit kleinen Blickwinkeln werden die Sehfehler auch über die Spiegel recht klein sein, so dass die Sicht auf die reale Welt kaum verzerrt ist. Jedoch kann die Optik zum Fokussieren, bzw. zum Vergrößern des Abbildes Verzerrungen hinzufügen. Das Abbilden der verzerrten virtuellen Sicht auf die unverzerrte reale Welt führt zu statischen Erfassungsfehlern.

Optische Verzerrungen sind häufig systematische Fehler, so dass sie erkannt und kompensiert werden können. Diese Kompensation ist nicht immer einfach, kann aber durchaus ausgeführt werden. Eine Möglichkeit besteht darin, dies elektronisch zu tun. Nach einem Prinzip von Howlett [Hayw93] werden beide Bilder, das virtuelle und das reale vorverzerrt, und können dann optisch einwandfrei dargestellt werden. Bryson [Bry92a] gibt jedoch zu bedenken, die Vorverzerrung erhöhe die Rechenzeit so sehr, dass man besser darauf verzichten sollte.

### **6.2.2 Fehler im Trackingsystem**

Fehler in der Ausgabe von Lage- und Positionssensoren führen zu den schwerwiegendsten Fehlern in der statischen Erfassung. Diese Störungen sind nicht leicht zu messen und auszuschließen, denn das bedeutet, dass man ein noch genaueres Kalibrierwerkzeug benutzen muß, um genau diese Abweichungen festzustellen. Dazu sind diese Fehler zu unsystematisch und schwierig sie genau zu charakterisieren. Um kommerziell erfolgreich zu sein, muß man Abstriche in den Genauigkeitsforderungen machen. Das führt dann dazu, dass fast alle erhältlichen Systeme zur Lage- und Positionserfassung für AR-Anwendungen viel zu ungenau sind.

### **6.2.3 Mechanische Ausrichtungsfehler**

Mechanische Ausrichtungsfehler sind Unterschiede zwischen dem Modell oder der Spezifikation der Hardware und den aktuellen physikalischen Eigenschaften des realen Systems. Die optischen Systeme, wie die HMD mit halbdurchlässigen Spiegeln oder die Monitore sind durchaus nicht immer richtig justiert. Ist der Rahmen nicht ausreichend steif, können alle anderen Teile sich im Verhältnis zueinander verschieben, wenn der Benutzer sich durch die Szene bewegt. Auch diese Lagefehler müssen in Betracht gezogen werden. Solche Fehler verursachen kleine Änderungen in der Position und Lage des virtuellen Bildes, welches schwierig zu kompensieren ist. Einige Ausrichtungsfehler können kalibriert werden, für viele andere muss von Anfang an ausreichend steif gebaut werden.

#### 6.2.4 Inkorrekte Betrachtungs-Parameter

Betrachtungs-Parameter, auch Viewing-Parameter, beschreiben, wie man die erhaltenen Daten der Kopf- oder Kameraposition in die Viewing-Matrizen einsetzen kann, die vom Szenengenerator benutzt werden. Für ein HMD-basiertes System müssen folgende Parameter vorhanden sein:

- Mittelpunkt der Projektion und die Grösse des Bildausschnitts
- Verschiebung zwischen dem Auge und der Kameraposition
- Versatz in der Lageorientierung zwischen dem Tracker und dem Auge
- Blickwinkel

Falsche Betrachtungs-Parameter sind systematische statische Fehler. Bei beispielsweise falscher Einstellung zwischen dem Auge und dem Tracker können alle virtuellen Gegenstände grösser oder auch kleiner erscheinen. Für HMD mit halbdurchlässigem Spiegel kann man unterschiedliche Verfahren anwenden, um die virtuelle Welt mit der realen in Übereinstimmung zu bringen.

Bei einigen Systemen werden die Parameter manuell eingegeben in dem man mehrere Punkte benutzt um das System zu justieren. Das führt häufig zu unsystematischen Fehlern, weil der Benutzer versucht ist, das Bild überall „gut“ aussehen zu lassen. Führt diese Anpassung später eine andere Person durch, können nun Fehler ganz anderer Art vorliegen.

Eine andere Möglichkeit besteht darin, alle Parameter auszumessen und diese einzusetzen. Zum einen sind das die Methoden der Optiker, z.B. Augenabstand und ähnliche Abstände zum anderen nimmt man alles auf und wertet es aus, was die Kamera sieht. Aus beiden Methoden versucht man die Betrachtungs-Parameter zu extrahieren. Dieses Verfahren scheint fehlerträchtig und langsam zu sein, in der Literatur ist darüber nichts wesentliches berichtet.

Zur Kalibrierung können auch sichtbezogene Ansätze gewählt werden. Nach diesem Ansatz wird der Benutzer gefragt, bestimmte Dinge zu tun um damit die geometrischen Bedingungen zu berechnen. Führt man unterschiedliche Anwendungen aus, so kann man davon ausgehen, genügend Informationen gesammelt zu haben um die Viewing Parameter festzulegen. Oishi und Tachi [OiTa96] bewegen dazu einen virtuellen Cursor auf die Spitze eines realen Turmes. Aus den Benutzerangaben werden dann die Parameter berechnet. Alle sichtbezogene Ansätze vertrauen darauf, dass der Benutzer die Eingaben sauber und ordentlich vornimmt.

Zur Kalibrierung video-basierter HMD steht eine umfangreiche Literatur zur Verfügung, beispielsweise aus der Bildverarbeitung, siehe z.B. bei Jähne [Jähn93] oder aus der Roboteranwendung. [LeTs88]. Gerade in neuerer Zeit sind auf dem Robotersektor einige neue Arbeiten erschienen, so z.B. von Wunsch et.al. [WKAH96] oder auch Jägersand [JäNe96]. Diese Techniken berechnen die Viewing Parameter der Kamera aus unterschiedlichen Bildern eines bekannten Objektes. Dazu muss die Kamera an verschiedenen Orten aufgestellt werden. Aus den erkannten Punkten der 2D-Bilder mit den korrespondierenden 3D Punkten der realen



Welt, können so die Viewing Parameter berechnet werden. Einige AR-Systeme benutzen diese Kalibrierungstechnik mit einigem Erfolg [ARGOS94].

## 6.3 Dynamische Fehler

Dynamische Fehler treten bei Zeitverzögerungen des System auf. Als Gesamtverzögerungszeit wird die Zeit bezeichnet, die benötigt wird von der Erfassung der Position und Orientierung des Betrachtungspunktes bis zu dem Zeitpunkt in der das System das neue korrespondierende Bild berechnet und mit den neuen Werten dargestellt hat. Diese Verzögerungen sind für alle Komponenten vorhanden und auch bekannt. Als wichtigste Komponenten seien genannt: Die Verzögerung im Tracking-System, die Verzögerung im Kommunikationssystem, die Zeit, die der Szenengenerator braucht um das passende Bild in den Frame Buffer zu schreiben und die Zeit, um den Frame Buffer zu lesen und auf dem Bildschirm darzustellen. Eine Gesamtverzögerungszeit von ca. 100 msec sind typisch für existierende Systeme. Einfache Systeme haben eine kürzere Verzögerungszeit als z.B. Systeme in einem Netzwerk. Im Netzwerk sind 250 msec oder mehr eher typisch.

Eine Gesamtverzögerungszeit kann nur Auftreten, wenn Bewegung im System ist. Ist das System in Ruhe, fällt es nicht auf, wenn die neuen aber dennoch die selben Bilder, mit einer Zeitverzögerung dargestellt werden. Anders verhält es sich, wenn der Benutzer seinen Kopf bewegt. Hierzu ein Beispiel, ein Benutzer mit einem durchsichtigem HMD bewegt den Kopf. Der Tracker auf dem Helm misst die Lage zu einem Anfangszeitpunkt  $t_0$ . Das Bild, welches zur Zeit  $t_0$  gehört, wird aber erst zur späteren Zeit  $t_1$  dargestellt. Während dieser Verzögerung bewegt der Benutzer seinen Kopf aber weiterhin, so dass das, wenn das Bild für die Zeit  $t_0$  erscheint, der Benutzer in eine andere Richtung schaut als die zum Zeitpunkt  $t_0$ . Demzufolge ist das Bild zum Zeitpunkt der aktuellen Sicht falsch.

Systemverzögerung verhindert die Illusion, dass die reale und die virtuelle Welt koexistiert, weil dies große Erfassungsfehler verursacht. Mit einer typischen Verzögerungszeit von 100 msec und einer mäßigen Kopffrotation von  $50^\circ / \text{sec}$  beträgt der Fehlerwinkel  $5^\circ$ . In einer Armeslänge von 68 cm liegt der Erfassungsfehler bei 59 mm. Die Hauptverursacher dieser Erfassungsfehler können in 4 Gruppen zusammengefasst werden und werden im folgenden analysiert.

### 6.3.1 Reduzierung der Verzögerung durch die Systemzeit

Der einfachste Weg wäre es, die Verzögerung durch die Systemzeit zu eliminieren, denn dann kann es auch keine dynamischen Fehler mehr geben. Moderne Grafiksysteme sind aber auf maximalen Durchsatz optimiert und nicht auf kurze Wartezeiten [FDFH90]. Manchmal ist es möglich, das Grafiksystem zu Lasten des Durchsatzes neu zu konfigurieren. Olano [OCMB95] beschreibt 2 Möglichkeiten die Latenzzeit zu verringern, kommt aber zu dem Schluss, dass diese Methoden für Head-Mounted Displays gut geeignet seien aber der Durchsatz an Polygonrendering stark abnimmt.

Es ist nicht zu erwarten, dass die System-Wartezeiten in absehbarer Zukunft eliminiert werden wird. Der technische Fortschritt erhöht zwar die Schnelligkeit des Gesamtsystems, verringert aber nicht notwendigerweise den Zeitverbrauch des Grafik-Subsystems. Es sei noch mal vermerkt, dass die Erfassungsfehler möglichst klein sein sollen. Bei einer moderaten Kopfbewegung von  $50^\circ / \text{sec}$  sollte der Fehlerwinkel kleiner als  $0.5^\circ$  sein. Damit darf die System-Wartezeit 10 msec nicht wesentlich überschreiten. Dies kann man nur über neuartige aufwendige Grafiksysteme erreichen.

### 6.3.2 Reduzierung der sichtbaren Verzögerung

Imagedeflection ist eine Methode, um die sichtbare Verzögerung zu verkleinern, und kann nur zur Orientierung des Kopfes benutzt werden [RePo94]. Dazu nutzt man die Möglichkeit, die neusten Werte der Lagemessung in die letzten Stufen der Rendering Pipeline einzugeben. Der Szenengenerator rendert üblicherweise ein Bild, welches wesentlich grösser ist als zur Darstellung auf dem Display benötigt wird. Dann wird, kurz vor der Darstellung des Bildes auf dem Schirm die neuste Kopflage abgefragt. Die Werte der Kopforientierung werden benutzt, um nur einen kleinen Teil aus dem Frame Buffer zum Display zu schicken denn zur Korrektur erhält man kleine Werte und die sind dann äquivalent zum Schieben des Frame Buffers um wenige Pixel.

Dieses Verfahren kann man aber nicht bei Positionsänderungen, also Translationen verwenden [McBi95]. Dazu wird das Verfahren etwas modifiziert. Die Arbeitsweise ist ähnlich dem oben beschriebenen Verfahren. Aber diese Technik setzt voraus, dass man die Tiefe zu jedem Pixel kennt und das Bild entsprechend verzerrt. Das bedeutet aber, dass der Verzerrungsalgorithmus wesentlich schneller laufen muss als der des Rendern.

### 6.3.3 Anpassen der zeitlichen Datenströme

In einem videobasiertem AR-System liefert die Videokamera und die zugehörige Digitalisierungs-Hardware eine inhärente Verzögerung bei der Darstellung der realen Welt. Das hört sich für die Reduzierung von dynamischen Fehlern wie ein potentieller Vorteil an weil es erlaubt die zeitlichen Datenströme der realen und der virtuellen Bilder anzupassen. Zusätzliche Verzögerungszeit wird dem Videodatenstrom hinzugefügt, um eine Anpassung an die Verzögerungen aus dem Szenengenerator zu gewährleisten. Auf Grund der unterschiedlichen Komplexität der zu rendernden Szene, kann diese zusätzliche Verzögerungszeit nicht konstant bleiben. Deshalb muss das System die beiden Datenströme permanent dynamisch synchronisieren.

Zu beachten gilt, das die o.a. Methode den Konflikt zwar zwischen dem realen und dem virtuellen Teil reduziert, andererseits aber nun beide Ströme zeitverzögert sind. Für kurze Verzögerungszeiten ist das nicht besonders störend aber in Telepräsenz-Systemen kann das zu großen Fehlern führen.

### 6.3.4 Vorhersage des künftigen Standortes

Eine Methode um die Verzögerungszeiten zu verringern besteht darin, den künftigen Standort und die Orte der Objekte vorab zu bestimmen. Sind die künftigen Orte bekannt, dann kann man die Szene genau an diesen Orten rendern anstatt an den gemessenen Orten. Wenn die Szene dann dargestellt wird, hat sich der Benutzer sowie die Objekte auf die vorher berechneten Orte zubewegt und das Bild ist zum Darstellungszeitpunkt korrekt. Für kurze Verzögerungszeiten (kleiner als 80 msec) kann dieses Verfahren den dynamischen Fehler um eine Größenordnung verringern [AzBi94]. Für fehlerfreie Vorhersagen benötigt man allerdings für die Messungen und den Berechnungen ein System mit Echtzeit-Anforderungen. Einige wenige Vorhersageverfahren sind für AR-Systeme entwickelt [ZCSJ94], die meisten sind jedoch für VE-Umgebungen entworfen worden [AzBi94]. Verlässliche Vorhersagemethoden für unterschiedliche Verfahren um die aktuelle Kopfbewegung zu bestimmen müssen noch entwickelt werden.

## 6.4 Sichtbasierende Techniken

In [MüSt96] wurde darauf hingewiesen, dass Mausbewegungen und Rückkopplung nur über den visuellen Kanal, die man in der Regelungstechnik als „offene Schleife“ bezeichnet. Probleme verursachen. Ähnlich ist es in der AR-Technik. Wenn die Erfassung nur über die Erfassungssensoren stattfindet, hat das System ebenfalls keine Rückkopplung über die Übereinstimmung zwischen der virtuellen und realen Welt. Ohne diese Rückkopplung können aber keine Aussagen über die Übereinstimmung der beiden Bilder gemacht werden. Allerdings können in einem videobasiertem System Methoden aus der Bildverarbeitung hinzugezogen werden um eine Art Rückkopplung dem Erfassungssystem zu ermöglichen. Mit Hilfe der Bildverarbeitung lassen sich beispielsweise Merkmale in dem virtuellem Bild identifizieren, welche mit der realen Welt übereinstimmen. In diesem Fall hätte man einen geschlossenen Regelkreislauf.

Dies ist keine einfache Aufgabe. Dieses Erkennen und Anpassen muss in einer Echtzeit Anwendung laufen. Das erfordert oftmals spezielle Hardware und Software.

In einigen AR-Anwendungen kann man akzeptieren, dass bestimmte Marken in der Umgebung angebracht werden. Diese Marken können LED's [BaNe95] oder auch spezielle Markierungen [Mell95] [KABC97] sein. Die Orte oder die Muster werden als bekannt vorausgesetzt. Die Bildverarbeitung erkennt die Orte der Marken, diese werden dann benutzt um eine Übereinstimmung der virtuellen und der realen Welt herzustellen.

Diese Methoden erwarten, dass mindestens eine Marke zu jeder Zeit zu sehen sein muss. Kann das nicht gewährleistet werden, ist die Übereinstimmung gestört. Sind diese Marken aber sichtbar, kann das Ergebnis bis auf einen Pixel genau sein. Dies entspricht dann ungefähr der Genauigkeit, die mit Videotechniken erreichen können. Klinker [Klink96] demonstriert in ihrer Arbeit, wie man mit einfachen Marken eine gute Übereinstimmung erzielen kann.

Die Methoden, die mit Hilfe von Marken als einzige Erfassungsquelle arbeiten, ermitteln die relative Beziehungen zwischen den Objekten in der Umgebung und der Kamera. Dies reicht aus, die Erfassung zu gewährleisten, kann aber zu wenig sein in manchen AR-Applikationen. Nämlich dann, wenn absolute, keine relativen, Werte benötigt werden. Absolute Werte werden dann benötigt, wenn die realen Objekte und die virtuellen Objekte nicht von der Videokamera verfolgt werden, beispielsweise bei einem 3D-Zeiger oder anderen virtuellen Objekten, die nicht direkt mit realen Objekten verbunden sind.

Eine andere Möglichkeit besteht darin, den Benutzer in die Erfassungsverarbeitung mit einzubeziehen. Normalerweise sollte ein AR-System automatisch ablaufen, aber genau dann, wenn zweideutige Messungen vorliegen, kann es angebracht sein, einige zusätzliche Fragen zu beantworten. Die Kalibrierungstechnik von Tuceryan [TGWB95] basiert auf Bildverarbeitungsmethoden, aber unter bestimmten Umständen wird der Benutzer gebeten manuell einzugreifen.

Die Erfassungsanforderungen für AR-Systeme sind nur schwer zu erfüllen, nur einige wenige Systeme erfüllen die in sie gesetzten Erwartungen. Ullmann [Ullm99] zeigte ein solches System mit offener Regelschleife, welches allerdings ohne die realen Welt auskommt. Entwickelt man jedoch ein echtes AR-System, dann versprechen System mit geschlossenem Regelkreis die besten Erfolge [Mell95] und auch [Neum96].

Das Erfassungsproblem ist bis heute noch nicht zufriedenstellend gelöst. Viele der Systeme arbeiten mit statischen Betrachtungsstandorten oder erlauben nur statische Objekte. Selbst dann werden auch nur kurze Entfernungen gestattet. Die Erfassung wird häufig nur mit einer kleinen Zahl von realen Objekten oder dem System vorab bekannten Gegenständen in überschaubaren Umgebungen gezeigt. Beispielsweise könnte ein System nur mit markierten

Objekten arbeiten und kann demzufolge die unmarkierten nicht erkennen. Verdoppeln dieser Erfassungsverfahren mit sogenannten Hybridmethoden ist durch die Erhöhung der Komplexität und der dann erforderliche zusätzlichen Hardware keine triviale Aufgabe. Einfache aber auch effektive Methoden müssen noch erarbeitet werden, um die Akzeptanz der Benutzer für AR-Systeme zu erhöhen.

## 7 Mess-Sensoren

Eine fehlerfreie Erfassung und Positionierung von virtuellen Objekten im realen Umfeld erfordert eine fehlerfreie Verfolgung der Kopfposition und das Vermessen von anderen Objekten in der Umwelt. Das größte Hindernis zur Erstellung einer effektiven AR-Umgebung ist die Forderung nach genauen, auf größeren Distanzen arbeitende Sensoren und Tracking-Geräten, welche den Ort des Benutzers und der umgebenden Objekte in der realen Umwelt verfolgen kann. Kommerziell erhältliche Tracker zielen meist auf Anwendungen für die virtuellen Umgebungen oder auf solche, welche Bewegungsabläufe festhalten können. Im Vergleich der beiden eben genannten Anwendungsfälle erfordern AR-Anwendungen hingegen einen viel höheren Grad der Genauigkeit und erfordern einen viel größeren Arbeitsraum. Zur Zeit ist kein Tracker erhältlich, der über einen längeren Zeitraum die Daten in der erforderlichen Genauigkeit liefert. In diesem Bereich sind Entwicklungen durchzuführen, welche die strengeren Anforderungen im AR-Bereich erfüllen.

Speziell sind es 3 Bereiche, in der einige Arbeit erwartet wird:

- Größere Eingabeauswahl und Bandbreite
- Höhere Genauigkeit
- Größere Entfernungen

### 7.1 Größere Eingabeauswahl und Bandbreite

Bilder die dargestellt werden, Sound, der erzeugt wird, erwarten von einem VE-Systemen, dass die ganze Bandbreite an Ausgaben beherrscht wird. Die Eingabebandbreite ist hingegen nur schmal, Ort und Lage vom Kopf des Benutzers und dessen Hände, der Eingang einige Knöpfe an den Zeigegeräten und noch andere Eingabegeräte. AR-Systeme erwarten aber eine größere Auswahl an Eingabesensoren [FiIB95]. Es existiert eine größere Auswahl an möglichen Eingabesensoren als an Ausgabemöglichkeiten. Die Ausgabe ist begrenzt auf die 5 menschlichen Sinne. Die Eingabe kann von jeder beliebigen Quelle kommen, vorausgesetzt, ein Sensor kann sie erkennen. Zur Erinnerung sei noch einmal die medizinische Anwendung erwähnt, in diesen Fällen die Computertomographie, MRI<sup>8</sup> oder Ultraschall als Eingabe benutzt. Für künftige Applikationen kann man sich vorstellen, dass das Spektrum des sichtbaren Lichtes in Richtung Infrarot oder Ultraviolett für den AR-Benutzer erlebbar wird. Konzeptionell ist alles denkbar, was maschinelle Messfühler aufzunehmen in der Lage sind und auch übertragen können, für den Benutzer fühlbar dargestellt werden kann.

Geographische Daten als eine besondere Eingabe kann für manche AR-Anwendungen entscheidend sein [BWRT96]. Normalerweise sind dem AR-System die Größe sowie die Entfernungen der einzelnen Objekte bekannt weil diese Objekte mit dem System erstellt worden sind. Aber es ist durchaus nicht sichergestellt, dass dem AR-System bekannt ist, wo die Objekte in der realen Welt plaziert sind. Ein System kann z.B. annehmen dass das die Umgebung im Anfangsstadium initialisiert wird und anschließend es als statisch annimmt. Dennoch wird es einige Anwendungen geben, die der die Objekte sich bewegen und

---

<sup>8</sup> Magnetic resonance imaging, oder Magnetresonanztomographie

demzufolge auf verfolgt werden müssen. Für einige Anwendungen mag es ausreichend sein, eine räumliche Karte zu haben, in der die Tiefe des Raumes abgebildet ist. Das würde erlauben, dass reale Objekte virtuelle Objekte durch Vergleich der Pixel im Z-Buffer verdecken. Diese Tiefeninformation zu erhalten ist keine triviale Aufgabe. Viele Bilderkennungstechniken sind bisher untersucht worden, z.B. das Erkennen der Gestalt durch Stereobilder oder durch ihren Schatten. Es ist schwierig die Tiefeninformation durch existierende Bilderkennungsverfahren in Realzeit wiederzuerlangen.

Einige Anwendungen die mit Anmerkungen oder Erläuterungen an den Objekten arbeiten erfordern u. U. einen Zugang zu einer Datenbank der zugehörigen Umgebung. In einer Architekturapplikation ist es durchaus wichtig „in die Wand zu sehen“, weil dort Wasserleitungen, Elektrokabel und andere Dinge verbaut sein können. Solch eine Datenbank ist nicht frei verfügbar und wenn es denn so wäre, ist sie häufig in einem Format, die nicht zum System paßt. Sollte trotzdem eine Verwendung gewünscht sein, so muss man in Betracht ziehen, dass die Daten sehr unterschiedlich sein können und demzufolge einen hohen Rechenbedarf erfordern, wenn die Daten für das AR-System angepaßt werden müssen.

## **7.2 Höhere Genauigkeit**

Die Genauigkeitsanforderungen der Tracker und Sensoren werden durch die hohe Genauigkeit der visuellen Erfassung bestimmt. Für viele Ansätze ist die Erfassung so genau wie der Tracker sie liefern kann. Das heißt, AR-Systeme benötigen Tracker, die eine Genauigkeit von 1 mm und eine Winkelgenauigkeit von  $0.05^\circ$  über den gesamten Arbeitsbereich haben.

Nur wenige Tracker können diese Spezifikation erfüllen und jede Technologie hat ihre Stärken und Schwächen. Mechanische Tracker sind üblicherweise genau genug, haben aber nur ein begrenztes Arbeitsvolumen. Magnetische Tracker reagieren sehr empfindlich auf metallische Störungen, welche in den Umgebungen mancher gewünschter AR-Applikationen enthalten sind. Ultraschallsensoren leiden unter Umgebungsgeräusche und es ist schwierig auf große Entfernungen genaue Angaben zu erhalten, wenn die Umgebungstemperatur auf dieser Entfernung stark schwankt. Optische Tracker zeigen Verzerrungen und haben Kalibrierungsprobleme [JZMB94]. Tracker auf Trägheitsbasis driften mit der Zeit.

Das zukünftige Trackingverfahren wird vermutlich ein hybrides System sein, weil man so die strengen Anforderungen von AR-Systemen am ehesten erfüllen kann. Hier seien als erstes Trägheitssysteme genannt, die mit optischen Methoden kombiniert oder anstatt der Optik eine Kombinationen auf der Basis von GPS-Systemen hat. Mit der Nutzung der unterschiedlichen Techniken, kann man die systemimmanenten Schwächen übergehen und deren Stärken besser ausschöpfen.

Es sind Versuche unternommen worden, einen allgemein erhältlichen Tracker zu kalibrieren [Bry92b]. Die größten Störungen konnten beseitigt werden, aber die Genauigkeitsanforderungen von AR-Systemen wurden nicht erreicht. Die erreichbare Genauigkeit lag immer noch bei 2.5 cm.

Die Anforderungen an andere Sensortypen sind nicht so stringent. Beispielsweise kann der Mensch tiefe Töne nur schlecht orten, so dass es unwichtig ist, hier hohe Anforderungen anzulegen.

## **7.3 Größere Entfernungen**

Nur wenig Tracker werden für große Distanzen entwickelt, weil die meisten VE-Anwendungen keine großen Entfernungen überdecken. Bei Motion Capturing<sup>9</sup> wird die Bewegung einer Referenzperson über optische oder elektromagnetische Trackingsysteme in Echtzeit auf den künstlichen Akteur übertragen. Damit lassen sich Bewegungsanalysen der Referenzperson durchführen. Das reicht aus, um die Positionen der aufgenommen und verfolgten Punkte zu erhalten, aber nicht deren Lage im Raum. Die Lage wird aus den erhaltenen Positionen berechnet. Selbst kleine Fehler in diesen Positionen führen dann zu Lagefehler von einigen Grad, was für AR-Systeme nicht akzeptabel ist.

Wünschenswert sind skalierbare Trackingsysteme. Als ein skalierbares System wird ein solches bezeichnet, welches beliebige Entfernungen überdecken kann, einfach durch Aneinanderreihen von modularen Komponenten. Dies kann man erreichen durch ein zellenartiges Trackingsystem in dem nur die dem Benutzer nächsten Sensoren aktiv sind, die ihn verfolgen [WABG92]. Bewegt sich Benutzer in der Umgebung, bewegt sich der Satz von Sensoren mit und der Benutzer kann so große Entfernungen überbrücken ohne aus der Reichweite der Sensoren zu kommen. Während skalierbare Tracker sehr effektiv sein können, sind sie doch komplex und benötigen durch Modularität viele Komponenten, was diese Systeme relativ teuer macht.

Das Global Positioning System (GPS) wird benutzt, um den Ort von Flugzeugen oder Fahrzeugen auf der Erde zu bestimmen. Es bietet sich an, dieses System als Tracker für große Reichweiten zu benutzen. Das System als solches ist aber nicht ausreichend. Die besten Werte wurden erzielt, als noch ein fester, örtlich genau bekannter Sender als Referenz mit hinzugezogen worden ist. Dieses sogenannte Differential-GPS erreicht Genauigkeiten bis in den Zentimeterbereich [KüKo95]. Für AR-Anwendungen ist das nicht genau genug.

Außerhalb geschlossener Räume sind AR-Anwendungen bisher noch nicht gezeigt worden, denn auch hier spielen die Genauigkeitsanforderungen eine große Rolle.

## 8 Ausblick

In diesem Abschnitt sollen die Erweiterungen zusammengefasst werden, welche notwendig sind, die AR-Systeme zu verbessern und für die Benutzer handhabbarer zu machen.

Hybride Ansätze: Zukünftige Trackingsysteme werden vermutlich hybride Systeme sein, weil man so auf diese Weise am besten die Schwachstellen umgehen kann. Das kann für andere Gebiete der Augmented Reality genau so gut gelten. Beispielsweise verlassen sich die heutigen Systeme auf nur eine Erfassungsstrategie. Zukünftige Systeme könnten sich wesentlich robuster verhalten, wenn mehr als nur eine Technik benutzt wird. Eine Möglichkeit ist die Benutzung einer auf Bilder basierend Methode plus zusätzlicher Vorhersage. Sind keine Marker vorhanden, kann das System auf eine offene Regelstrecke ausweichen, um die Erfassungsfehler zu minimieren anstatt keine Werte zu erhalten. Der vorhergesagte Betrachtungsort ist im Gegenteil ein idealer Anfangswert für bildverarbeitende Methoden.

Echtzeit Berechnungen: Fast alle VE-Systeme arbeiten nicht in Echtzeit. Üblicherweise wird ein solches System auf einem Unix-Rechner implementiert, der möglichst schnell sein soll. Das mag für manche Systeme ausreichend sein, denn wenn alles im virtuellen Raum stattfindet gibt es keine Synchronisationsprobleme mit anderen Objekten. Augmented Reality arbeitet aber mit der realen Welt. Nun müssen reale und virtuelle Welten synchronisiert werden, nur – die reale Welt ist immer in Echtzeit. Deshalb müssen AR-Systeme erstellt

---

<sup>9</sup> Motion Capturing überträgt Bewegungen einer Referenzperson auf virtuelle Personen.

werden, in der die Echtzeit-Leistung im Vordergrund steht. Dazu müssen dann genaue Zeitstempel verfügbar sein an der die unterschiedlichen Komponenten sich abgleichen können. Das Betriebssystem darf nicht beliebig den AR Softwareprozess gegen einen anderen Prozess tauschen und den AR-Prozess beliebig lang im Hintergrund halten. Das System muss garantieren, dass die Anwendung in einem gewissen Zeitrahmen unterbrechungsfrei läuft. Diese Charakteristika sind typisch für Flugsimulatoren. Das Entwickeln eines Echtzeit-Rechners ist nicht ganz einfach und die Fehlerbeseitigung mag schwierig sein, aber die Anforderungen der AR-Systeme erfordern eine Echtzeit-Leistung.

Wahrnehmbarkeitsstudien: Im Augmented Reality Bereich fehlt es an Untersuchungen, welche psychomotorischen Wahrnehmungen der Benutzer erwartet. Wie lang darf eine Verzögerung sein, bis der Benutzer sie entdeckt? Wie groß darf der Erfassungsfehler sein, wenn man den Kopf bewegt? Neben den Fragen der Wahrnehmung sind auch psychologische Experimente von Interesse um das spezielle Verhalten herauszufinden. Wie wird bei einer Vorhersage der Kopfbewegung das Verhalten des Benutzers verbessert, beispielsweise durch verbesserte Orientierung im Umfeld. Darf der geduldete Fehler bei Bewegungen grösser sein als im Stand? Grundsätzlich stellt der Simulatorschwindel, also die nicht perfekte, bzw. erwartete Übereinstimmung, aber ein ernstzunehmendes Problem dar. Ohne dessen Beseitigung wird die Akzeptanz von AR-Systemen sicherlich erheblich erschwert bleiben.

Tragbares System: Einige zukünftig geplante AR-Applikationen benötigen allerdings eine Unterstützung, um in großen Umgebungen umher laufen zu können, oder gar außerhalb von Gebäuden. Das erfordert, dass die benutzbaren Geräte in sich abgeschlossen und tragbar sind. Vorhandene Trackingsysteme, die mit geforderten Genauigkeit arbeiten, sind zur Zeit dazu noch nicht in der Lage.

Multimodale Monitore: Fast die gesamte Arbeit im Augmented Reality Bereich fokussiert sich auf das menschliche Sehfeld und virtuelle grafische Objekte. Natürlich sollten in einer AR-Umgebung alle Sinne angesprochen werden können. Im besonderen sollte die Möglichkeit bestehen, 3D-Sound hinzuzufügen. Für einige Anwendungen kann das nützlich sein.

Soziale und politische Aspekte: Den zahlreichen Anwendungsbereichen mit all ihren Chancen stehen Gefahren gegenüber, die man nicht ausschließen kann. Dabei soll keine Technikfolgeabschätzung abgegeben werden, nicht weil dies den Rahmen dieses Übersichtsartikels sprengen würde, sondern auch weil die AR-Technik aufgrund ihres Konvergenzcharakters und des vielfältigen uneinheitlichen Anwendungsspektrums eine umfassende Beurteilung extrem erschwert. In diesem Zusammenhang sei auf das Buch von Bormann [Borm94] hingewiesen. Es sei aber auch darauf hingewiesen, dass soziale und politische Aspekte bei der Entwicklung neuer AR-Techniken eine wichtige Rolle spielen können.

## **9 Zusammenfassung**

Es wurde versucht, von einem Überblick ausgehend die Stärken und Schwächen der einzelnen Systeme aufzuzeigen. Augmented Reality ist in Bezug auf ihre Entwicklungsreife weit hinter den virtuellen Umgebungen zurück. Es gibt keinen Hersteller, der HMD-basierende Augmented Reality Systeme herstellt. Demzufolge sind die meisten AR-Anwendungen in der universitären oder industriellen Forschung angesiedelt.

In den Kapiteln 5, 6 und 7 wurde aufgezeigt, wo die Schwachstellen in einem VR-System liegen. Aus diesen Abschnitten läßt sich herleiten, dass es „das“ VR bzw. AR-System nicht

gibt. Vielmehr besteht es aus einer Anzahl von Komponenten, die für die jeweilige Aufgabe ausgesucht und kombiniert werden müssen.

Allgemein kann festgestellt werden, dass in der Medizin die meisten Anwendungen im Einsatz sind. Diese arbeiten zwar nicht mit einem Datenhelm sondern mit grossen Monitoren und Shutterbrillen und benutzen haptisches Feedback um Operationsbestecke zu simulieren. Hier wird verstärkt im Bereich AR und Operationsroboter Forschung betrieben.

AR-Anwendungen im Produktionsbereich sind nur wenige bekannt. Die meisten Forschungen sind im Virtual Reality Bereich angesiedelt. Hier fehlt eindeutig die Möglichkeit, sich in Räumen der Größe einer Fabrikhalle mit AR-Anwendungen bewegen zu können. Gleichzeitig braucht man hier neuartige Trackingsysteme die sich von meist dort vorhandenen Metallteilen nicht beeinflussen lassen.

Anwendungen in der freien Natur können eine grosse Zukunft haben. Sieht man einmal von den militärischen Forderungen einmal ab, so sind Anwendungen im Architekturbereich durchaus heute schon in Ansätzen erkennbar. Hier liegt ein großes Potential, nicht nur für die Ersteller solcher virtuellen Modelle, den Architekten, sondern auch für dessen Kunden. Kann man doch mit AR-Methoden die Funktionsfähigkeit einzelner Module zeigen.

Auch für unser Verständnis der Vergangenheit können AR-Methoden von Vorteil sein. Man stelle sich vor, das Forum Romanum im alten Rom existiert und man könne an einer Senatssitzung unter der Leitung von Cäsar teilnehmen. Aufbereitet für die heutige Generation kann es grossen Beitrag zum Geschichtsverständnis leisten.

Die möglichen Anwendungen, die man sich vorstellen kann, sind sehr vielfältig. Die einzelnen Techniken sind alle bekannt, die Kombinationen der Techniken untereinander nur in Ansätzen. Sind diese Interaktionsmöglichkeiten bekannt, dann sind Augmented Reality Anwendungen nicht mehr aufzuhalten.

## 10 Literatur

[ARGOS94] ARGOS Virtual Pointer Camera Calibration Procedure, 1994, [vered.rose.utoronto.ca/people/david\\_dir/POINTER/Calibration.html](http://vered.rose.utoronto.ca/people/david_dir/POINTER/Calibration.html)

[AzBi94] R. Azuma, G. Bishop: Improving Static and Dynamic Registration in a See-Through HMD, Proc. of SIGGRAPH '94, Orlando, 24.-29.July, 1994

[Azum93] R. Azuma: Tracking Requirements for Augmented Reality, Communication of the ACM, 36, No 7, July 1993, S. 50-51

[Azum95] R. Azuma: Predictive Tracking for Augmented Reality, TR95-007, Dep. of Computer Science, University of North Carolina, Chapel Hill, 1995

[BaNe95] M. Bajura, U. Neumann: Dynamic Registration Correction in Video-Based Augmented Reality Systems, IEEE Computer Graphics and Applications, Vol 15, No 5, 1995

[Borc94] D. Borchers: Die Nackten und die Toten, Multimedia auf der Herbst-Comdex ' 93, c' t 1/1994, Seite 132

[Borm94] S. Bormann: Virtuelle Realität, Genese und Evaluation, Bonn, Addison-Wesley, 1994



- [BWRT96] D. Breen, R. Whitaker, E. Rose, M. Tuceryan: Interactive Occlusion and Automatic Object Placement for Augmented Reality, Proc. of Eurographics, Poitiers, 26.-30. Aug., 1996
- [Bric97] R. Brice: Multimedia & Virtual Reality Engineering, Oxford, UK, Butterworth-Heinemann, 1997
- [Bry92a] S. Bryson: Measurements and Calibration of Static Distortion of Position Data from 3D Tracker, Proc. of SPIE Vol. 1669, Stereoscopic Displays and Applications III, San Jose, USA, 12.-13. Feb., 1992
- [Bry92b] S. Bryson: Survey of Virtual Environment Technologies and Techniques, In: SIGGRAPH ' 92, 19th Intl. Conf. on Computer Graphics and Interactive Techniques, Course No 9, Chicago, USA, 26.-31. July 1992
- [FDFH90] J. D. Foley, A. van Dam, S. Feiner, J. Hughes: Computer Graphics, Principles and Practice (2. Auflage), Reading, Addison-Wesley, 1990
- [FIHS93] S. Feiner, B. MacIntyre, M. Haupt, E. Solomon: Windows on the World: 2D Windows for 3D Augmented Reality, Proc. of UIST '93, Atlanta, GA, USA, 3.-5. Nov., S. 145-155, 1993
- [FiIB95] G. Fitzmaurice, H. Ishii, B. Buxton: Bricks, Laying the Foundation for Graspable User Interfaces, CHI 95 Mosaic of Creativity, Denver, 7.-11. Mai, 1995
- [Fitz93] G. Fitzmaurice: Situated Information Spaces, Spatially Aware Palmtop Computers, CACM 36, No. 7(July 93), s. 38-49, 1993
- [fsm81] Flugsicherheitsmitteilungen des Luftfahrtbundesamtes fsm 1/81, Luftfahrtbundesamt, Braunschweig, 1981
- [Gibs97] W. Gibson; Neuromancer, München, Heyne Verlag, 1987
- [Hanc93] D. M. Hancock: Optomechanical Verification of COSTAR using Computer Graphics Simulation, in SPIE Proc. Vol. 1996, Optical Alignment IV, (Hrsg. M. C. Ruda), San Diego, S. 144-148, 11.7.-16.7.1993
- [Hayw93] T. Hayward: Adventures in Virtual Reality, Carmel, QUE Corp., 1993
- [Henn97] A. Hennig: Die andere Wirklichkeit, Bonn, Addison-Wesley-Longman, 1997
- [JZMB94] A. Janin, K. Zikan, D. Mizell, M. Banner, H. Sowizral: A Videometric Head Tracker for Augmented Reality, SPIE Proc. Vol 2351, Telemicroscopy and Telepresence Technology, Boston, 31. Oct.-4. Nov., 1994
- [Jähn93] B. Jähne: Digitale Bildverarbeitung, 3. Aufl., Berlin, Springer, 1993
- [JäNe96] M. Jägersand, R. Nelson: On-line Estimation of Visual-Motor Models using Active Vision, Proc. ARPA Image Understanding Workshop, 1996
- [Krue91] M. Krueger: Artificial Reality II, Reading, Addison-Wesley, 1991

- [KrFr94] W. Krüger, B. Fröhlich: Responsive Workbench, in: Warnecke, Bullinger (Hrsg.), Virtual Reality '94 – Anwendungen und Trends, Berlin, Springer-Verlag, 1994
- [Kühl98] D. Kühlke: Optik, Grundlagen und Anwendungen; mit zahlreichen Tabellen, Beispielen und Aufgaben mit Lösungen, Thun, Deutsch, 1998
- [KüKo95] E. Kühmstedt, R. Kopischke: Genaues Navigieren mit LW-Realtime-DGPS, Bundesamt für Kartographie und Geodäsie, Außenstelle Potsdam, 1995, [http://www.potsdam.ifag.de/potsdam/dgps/dgps\\_1.html](http://www.potsdam.ifag.de/potsdam/dgps/dgps_1.html)
- [KABC9.] G. Klinker, K. Ahlers, D. Breen, P. Chevalier, C. Crampton, D. Greer, D. Koller, A. Kramer, E. Rose, M. Tuceryan, R. Whitaker: Confluence of Computer Vision and Interactive Graphics for Augmented Reality, Presence: Teleoperations and Virtual Environments, Vol. 6, No. 4, Aug 1997
- [LMFD87] J. Lanier, M. Minsky, S. Fisher, A. Druin; Virtual Environments and Interactivity: Windows to the Future, In: Computer Graphics, Vol. 23, No. 5, 1989, S. 7-18
- [LeTs88] R.K. Lenz, P.Y. Tsai: Techniques for Calibration of the Scale Factor and Image Center for High Accuracy 3-D Machine Vision Methodology, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol 10, No 5, (September) 1988
- [McFo94] R. S. McCann, D. C. Foyle: Superimposed Symbology: Attentional Problems and Design Solutions, SAE Transactions: Journal of Aerospace, Vol. 103
- [McBi95] L. McMillan, G. Bishop: Plenoptic Modeling, An Image-Based Rendering System, Proc. of SIGGRAPH '95, Los Angeles, 6.-11. August, 1995
- [MüSt96] S. Münch, M. Stangenberg: Intelligent Control for Haptic Displays, Eurographics 96, Vol. 15, No 3, 1996
- [MZDG93] P. Milgram, S. Yhai, D. Drascic, J. Grodski: Applications of Augmented Reality for Human-Robot Communication. Proc. of International Conference on Intelligent Robotics and Systems, Yokohama, Japan, July 1993, S. 1467-1472
- [Oakl99] I. Oakley: Comparing Haptic Effects in a GUI, Proc. The First PHANTOM Users Research Symposium, Deutsches Krebsforschungszentrum, Heidelberg, 21./22. Mai, 1999
- [OiTa96] T. Oishi, S. Tachi: Methods to Calibrate Projection Transformation Parameters for See-Through Head-Mounted Displays, In: Presence: Teleoperators and Virtual Environments, No. 5, 1 (Winter) 1996
- [OCMB95] M. Olano, J. Cohen, M. Mine, G. Bishop: Combatting Rendering Latency, Proc. of Symposium on Interactive 3D Graphics, Monterey, 9.-12. April, 1995
- [Pape96] D. Pape: Performer 2.0 in the CAVE, <http://www.evl.uic.edu/ape/CAVE/pf/Seminar.Feb96/>
- [RBAC94] E. Rose, D. Breen, K. Ahlers, C. Crampton, M. Tuceryan, R. Whitaker, D. Greer: Annotating Real-World Objects Using Augmented Reality, Proc. of Computer Graphics International '95, Leeds, UK, 25-30.June, S. 357-370, 1995

[Rege94] H. Regenbrecht: "Virtuelle Realität und Design", Diplomarbeit, Bauhaus-Universität Weimar 1995, [www.uni-weimar.de/architektur/InfAR/public/DIPLHR/diplzus.htm](http://www.uni-weimar.de/architektur/InfAR/public/DIPLHR/diplzus.htm)

[RePo94] M. Regan, R. Pose: Priority Rendering with a Virtual Reality Address Recalculation Pipeline, Proc. of SIGGRAPH '94, Orlando, FL, 24-29 July, 1994

[ScBB97] K. Schäfer, V. Brauer, F.W. Bruns; Real Reality: Eine Synthese aus Simulation und Virtual Reality mit Anwendungsbeispielen aus der Produktionstechnik, In: Tagung der deutschen VR-Anwender, Köln, 10-12. Juni 1997

[Schm98] J. Schmidt: Spezialisten für 3D, Die Hardware-Architektur von Konsolen, c' t 18/1998, Seite 100

[Sony99] Sony Firmeninformation, <http://www.ita.sel.sony.com/products/av/glasstron/>

[Sper96] T. Sperlich: Höllischer Cyber-Ritt, Virtuelle Action in High-Tech-Spielzentren, c' t 2/1996, Seite 84

[Snyd88] H. Snyder: Image Quality, In: Helander, Martin (Hrsg), Handbook of Human Computer Interaction, Amsterdam, North-Holland, 1988

[Ster97] The Stereographics Developers Handbook, StereoGraphics Corp., San Rafael, 1997

[Stur91] D. Sturm: Spürbar real? Virtuelle Wirklichkeit und menschliche Wahrnehmung; In: M. Waffender (Hrsg), Cyberspace - Ausflüge in virtuelle Wirklichkeiten, Reinbek, Rowoldt, 1991, S. 108

[TGWB95] M. Tuceryan, D. Greer, R. Whitaker, D. Breen, C. Crampton, E. Rose, K. Ahlers: Calibration Requirements and Procedures for Augmented Reality, IEEE Transactions on Visualization and Computer Graphics 1, No. 3, September, 1995

[Völt95] S. Völter: Virtual Reality in der Medizin I – Stand, Trends, Visionen, Mannheim, GeSI, 1995

[vonF73] H. von Förster: Das Konstruieren einer Wirklichkeit, In: P. Watzlawik (Hrsg) Die erfundene Wirklichkeit - Wie wissen wir, was wir zu wissen glauben? Beiträge zum Konstruktivismus, München, Piper, 1973

[Wals91] R. Walser: Elements of a Cyberspace Playhouse. Proc. of International Conference on Intelligent Robotics and Systems, Yokohama, Japan, July 1993, S. 1467-1472

[WABG92] M. Ward, R. Azuma, R. Bennett, S. Gottschalk, H. Fuchs: A Demonstrated Optical Tracker with Scalable Work Area for Head Mounted Display Systems, Proc. of Symposium on Interactive 3D Graphics, Cambridge, USA, 29.Mar. – 1.Apr., 1992

[WKAH96] P. Wunsch, G. Kögel, K. Arbter, G. Hirzinger: Kalibrierung eines nichtlinearen, binokularen Hand-Auge Systems, Tech. Report IB-515-96-27, Institute of Robotics and System Dynamics, German Aerospace Center, 1996

[ZCSJ94] K. Zikan, W. D. Curtis, H. Sowizral, A. Janin: A Note on Dynamics of Human Head Motions and on Predictive Filtering of Head-Set Orientations, Proc. of SPIE Vol. 2351: Telemanipulator and Telepresence Technologies, Boston, 31. Oct-4.Nov., 1994