

Augmented and Virtual Reality for Mobile Devices

Sebastian Boring

LFE Medieninformatik
Ludwig-Maximilians-Universität München
Amalienstraße 17, 80333 München, Germany
boring@cip.ifi.lmu.de

Zusammenfassung. Heutige mobile Endgeräte besitzen mehr und mehr Leistung, wodurch das Verlangen nach grafikintensiven Diensten über „erweiterte Realität“ bereits jetzt einen hohen Stellenwert einnimmt. Dazu gehören neben mobilen und personalisierten Navigationssystemen und industriellen Wartungssystemen auch Multiplayer-Spiele. Im Folgenden wird zunächst auf die Verfahren der Augmented Reality und Virtual Reality eingegangen. Danach werden die heute eingesetzten Endgeräte und zusätzliche Beispiele aktueller Forschungsbereiche aufgezeigt. Zum Schluss wird eine Zusammenfassung mit anschließendem Zukunftsausblick die ab heute möglichen Systeme und Dienste darstellen.

1 Einleitung

Im Zuge der heutigen Entwicklung der Mensch-Maschine-Interaktion spielen virtuelle Realität (virtual Reality, VR) und erweiterte Realität (Augmented Reality, AR) eine sehr große Rolle. Dies liegt vor allem daran, dass viele Dinge, Gegenstände oder Beschaffenheiten entweder nicht ohne großen Aufwand oder gänzlich nicht dargestellt werden können. Das trifft zum Beispiel auf antike Stätten wie den Circus Maximus in Rom zu, der nur noch als großer Platz existiert. Aber auch kleinere Szenen, wie zum Beispiel ein Möbelstück in einem leeren Raum oder ganze Gebäude in Straßenzügen konnten bis vor kurzem nur mit Hilfe von virtueller Realität nachgebildet werden, aber diese Darstellungsweise vermittelt jedoch nicht das Gefühl von Realität.

Hinzu kommt der Aspekt, dass viele Gegebenheiten in der realen Welt nicht ohne weiteres für den Benutzer verständlich sind. So ist eine zweidimensionale Karte für viele Menschen nicht auf die dreidimensionale Welt übertragbar [15], was bei papierbasierten Karten häufig zum Drehen der Karte führt. Aber auch dreidimensionale virtuelle Navigation bereitet dem Menschen immer wieder Probleme, da der Bezug zur realen Welt nicht hergestellt werden kann.

Genau hier setzt die noch junge Technologie der Augmented Reality an. Die Grundidee basiert auf der „Erweiterung“ der realen Welt durch virtuelle Objekte. Sie ist eine Echtzeit-Technologie und erfordert dadurch sehr effiziente Geräte und Algorithmen zur Darstellung. Durch AR können sekundenschnell Zusatzinformationen auf die reale Welt projiziert werden, wie zum Beispiel den zu fahrenden virtuellen Weg auf ein Bild der realen Welt, welches meist durch eine Kamera aufgenommen wird.

Diese Technologie stellt somit unterstützende Funktionalitäten bereit, die dem Benutzer erlauben, Zusammenhänge bestimmter Daten mit der realen Welt besser nachvollziehen zu können.

Aber auch die Unterhaltungsindustrie, vornehmlich im Bereich der Computerspiele, hat diese Technik für sich entdeckt. Heutige Spiele basieren noch auf virtuellen, von Designern gestalteten Levels, die zwar meist grafisch sehr gut ausgearbeitet wurden, aber auch wieder jeglichen Bezug zur Realität vernachlässigen. Manch Spieler möchte aber gerade in seiner realen Umgebung das gleiche Spiel spielen, und das möglichst mit seinen Freunden.

Doch all diese Szenarien scheinen noch in weiter Ferne, solange AR nicht auch auf mobilen Endgeräten eingesetzt wird. Der erhöhte Mobilitätsgrad der Bevölkerung und vor allem das Verlangen nach „realen“ Navigationssystemen sagen damit auch das gleiche aus. So werden zunehmend mobile Endgeräte als Fokus für AR- und VR-Anwendungen angesehen. Zusätzlich dazu haben diese Geräte gerade durch Mobiltelefone mit Grafikdisplays einen enormen Verbreitungsgrad in der Bevölkerung und würden so die Entwicklung und Anschaffung neuer Geräte vermindern.

Durch den Einsatz auf leistungsfähigen, mobilen Endgeräten könnten ohne weiteres auch alle zuvor genannten Szenarien umgesetzt werden. Zusätzlich wären der hohe Mobilitätsgrad und die große Flexibilität der Benutzer gewahrt, die dies nicht mehr stationär an einem PC betrachten müssen. Ein Raumplaner beispielsweise könnte mit einem PDA ohne weiteres durch Anwendung von AR dem Kunden demonstrieren wie ein Möbelstück an einer gegebenen Stelle aussieht, ohne es vorher erst in die Wohnung bringen zu müssen (siehe Abbildung 1).



Abbildung 1: Hier wird ein Sofa in virtuell in einer realen Umgebung angezeigt. Der Marker (unter dem Fernseher) hilft der Software auf dem PDA zu erkennen wo sie sich im Raum befindet. Dadurch kann das Sofa korrekt gezeichnet werden. [2]

Im Folgenden soll nun erläutert werden, welche Kriterien und Anforderungen für ein mobiles Augmented Reality System benötigt werden. Des Weiteren sollen alle Begrifflichkeiten geklärt werden, die im Bezug auf AR von Bedeutung sind.

2 Augmented and Virtual Reality

Wie bereits erwähnt erhält man mit Augmented Reality die Möglichkeit, zusätzliche Informationen bereitstellen zu können. Zu diesen Informationen gehören nicht sichtbare oder gar nicht existierende Objekte sowie virtuelle Wege. Um diese Technik genauer einordnen zu können, wird AR zunächst einmal definiert und damit von VR abgegrenzt, bevor die grundlegenden Verfahren für den Einsatz erläutert werden.

2.1 Definition

Da Virtual Reality einerseits von Augmented Reality abgegrenzt werden soll, andererseits ein Bestandteil dieser ist, wird VR zunächst definiert und erläutert.

Wie der Name schon sagt, beschreibt sie eine vollständig virtuelle Welt, wie sie zum Beispiel in Computerspielen oder in Kinofilmen (zum Beispiel „Shrek“) zu finden ist. Durch VR können alle Gegebenheiten der realen Welt nachgebildet werden, das heißt eine komplette virtuelle Szene wird am Computer erstellt und kann dann berechnet und veranschaulicht werden. Diese Technik bietet jedoch einen weiteren großen Vorteil. Sie ermöglicht die Darstellung von vollständig irrealen Szenen, die es so nicht mehr gibt, oder nicht geben würde. Daher erhält sie zudem großen Einzug in die Medizin und in die Forschung in diesem Bereich.

Wie jedoch schon angesprochen, beinhaltet VR das große Problem, dass sie vollständig virtuell ist, das heißt, dass kein absolut realer Bezug vorhanden ist. Dadurch können Benutzer die reale Welt um sie herum nicht mehr wahrnehmen [9]. Dies soll nun durch die „erweiterte“ Realität gelöst werden. Hier werden virtuell erzeugte Objekte direkt auf das reale Bild, entweder durch Projektion auf das reale Bild, oder durch Projektion auf ein real aufgenommenes Bild durch eine Kamera zusätzlich dargestellt. AR wird daher auch als eine Abweichung oder Wechsel der virtuellen Realität bezeichnet [9]. Abbildung 2 zeigt eine Szene der Augmented Reality, die eine Überlagerung der realen Welt mit virtuellen Objekten verdeutlichen soll.

Um den Bereich der AR nicht auf spezifische Technologien zu beschränken, definiert eine Studie von Ronald T. Azuma im Jahre 1997 ein AR-System als ein System, das die folgenden drei Kriterien erfüllen muss:

- Es verbindet reale und virtuelle Welt (Inhalte),
- es ist interaktiv in Echtzeit, und
- es ist im dreidimensionalen Raum registriert [9].

Durch diese Definition ist es möglich, sich nicht auf spezielle Hardware beschränken zu müssen, wobei die grundlegenden Komponenten der AR nach wie vor gewährleistet sind.



Abbildung 2: Realer Tisch mit virtueller Lampe und zwei virtuellen Stühlen. Lampe und Stühle werden korrekt im dreidimensionalen Raum gezeichnet (Verdeckung der Stühle durch den Tisch und Verdeckung des Tisches durch die Lampe) [9].

Zusammenfassend kann man also sagen, dass die virtuelle Realität eine rein synthetisch erzeugte Welt ist, wohingegen die Augmented Reality als Technologie genau zwischen der realen und der virtuellen Welt anzuordnen ist.

2.2 Angewandte Trackingverfahren

Um nun eine solche Szene, wie in beispielsweise in Abbildung 1 dargestellt, möglichst exakt visualisieren zu können muss der Benutzer (bzw. das Endgerät) relativ zur realen Welt bestimmt werden. Hierzu gibt es zwei Verfahren, die Kalibrierung und das Tracking, die zunächst kurz erläutert werden.

Unter Kalibrierung versteht man die exakte Ausmessung der realen Welt und unveränderlicher Parameter, wie zum Beispiel die Kamerabrennweite oder Raumabmessungen. Zusätzlich beinhaltet dieses Verfahren die Positionierung fester Sensoren im Raum, um später bewegliche Objekte orten zu können. Tracking ist die Bestimmung der zeitlich variablen Parameter, wie zum Beispiel die Position von Benutzern oder Gegenständen. Da Tracking mit nicht festen Daten arbeitet, werden an diese Technologie die von AR geforderten Echtzeitkriterien gestellt, die meist nur komplex umzusetzen sind. Folgende Eigenschaften müssen von einem Tracking-System eingehalten werden:

- Geringe Latenzzeit (bis zu maximal 100 ms),
- Genauigkeit bei Positions- und Orientierungsmessung,
- Handliche Geräte im mobilen Einsatz, und
- Multi-User Fähigkeiten.

Erst wenn diese Kriterien eingehalten werden, kann ein erfolgreicher Einsatz einer AR-Anwendung garantiert werden. Um nun einen Einblick zu erhalten, welche Tracking-Verfahren es gibt, und in wie weit sie diese Kriterien unterstützen, werden die zwei am häufigsten eingesetzten Verfahren, Optisches Tracking und GPS, erläutert.

Unter optischem Tracking versteht man die Erkennung von spezifischen Gegebenheiten der realen Umgebung. Dies beinhaltet sowohl speziell in der realen Welt registrierte Marker, als auch andere Umwelteinflüsse, wie zum Beispiel Gebäude. Abbildung 3 zeigt den Einsatz der gerade beschriebenen Marker für ein Multi-User Spiel. Diese Marker werden von der Kamera (und damit von der Software) erkannt, woraufhin deren Lage und Orientierung relativ zur Kamera bestimmt werden kann. Daher eignet sich dieses Tracking-Verfahren besonders zur Erkennung der Position und zusätzlich der Orientierung ohne weitere Techniken einsetzen zu müssen.

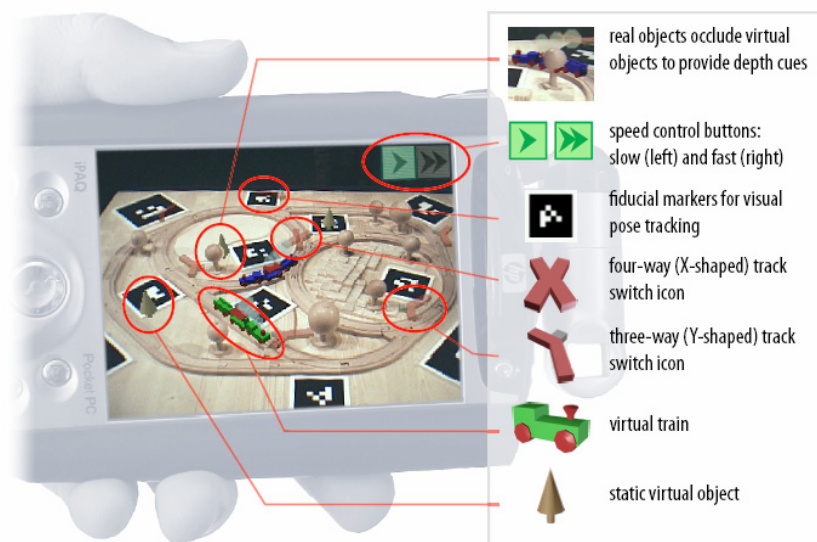


Abbildung 3: Hier wird das Tracking durch visuelle Marker, die eine, in der realen Welt, feste Position besitzen verdeutlicht [4].

Ein weiteres, häufig eingesetztes Tracking-Verfahren ist das Global Positioning System (kurz: GPS). Dieses System stützt sich auf 24 gleichmäßig im All verteilte Satelliten, die alle untereinander synchronisiert sind. Durch Messung der Laufzeit, bzw. die Unterschiede der verschiedenen Laufzeiten, der Signale von (mindestens 4) Satelliten kann die Position relativ genau (ca. 10 m) bestimmt werden. Dieses Verfahren eignet sich jedoch nur für das Erkennen von Benutzern außerhalb geschlossener Räume, da die Satellitensignale hier nicht empfangen werden können. Ein weiterer Nachteil von GPS ist, dass die Orientierung dadurch nicht erkannt werden kann. Dies wird jedoch umgangen, indem man über zwei aufeinander folgende Messungen die Bewegung des Benutzers bestimmen kann. Es eignet sich also für sich bewegende Nutzer, jedoch nicht im stationären Fall, wenn es auf die Orientierung ankommt.

Zusammenfassend lässt sich hier sagen, dass die optischen Tracking-Verfahren im Bereich der Genauigkeit und der einfachen Orientierungsbestimmung einen großen Vorteil gegenüber GPS bieten. Das große Problem dieser ist jedoch, dass die reale Welt demzufolge mit Markern ausgestattet werden muss, bzw. dass alle zu erkennenden Gegenstände genau registriert werden müssen.

2.3 Darstellung der virtuellen Objekte in realer Umgebung

Wenn nun die Daten über Position und Orientierung bekannt sind, können die virtuellen Objekte auf das reale Bild projiziert werden. Hierzu werden zwei grundlegende Verfahren unterschieden, das See-Through und das Video-Blending. Beide werden nun kurz mit ihren Eigenschaften und Vorteilen erläutert.

Unter See-Through versteht man ein Gerät (zum Beispiel eine Brille), durch das man die reale Welt nach wie vor sehen kann. Zusätzlich werden durch einen halbreflektierenden Spiegel virtuelle Daten in dieses Bild projiziert. Geräte mit Video-Blending hingegen nehmen das reale Bild mit Hilfe einer Kamera auf, projizieren danach die virtuellen Objekte auf dieses Bild und zeigen dann das Ergebnis an. Abbildung 4 zeigt zwei typischer Geräte (HMDs) beider Verfahren.

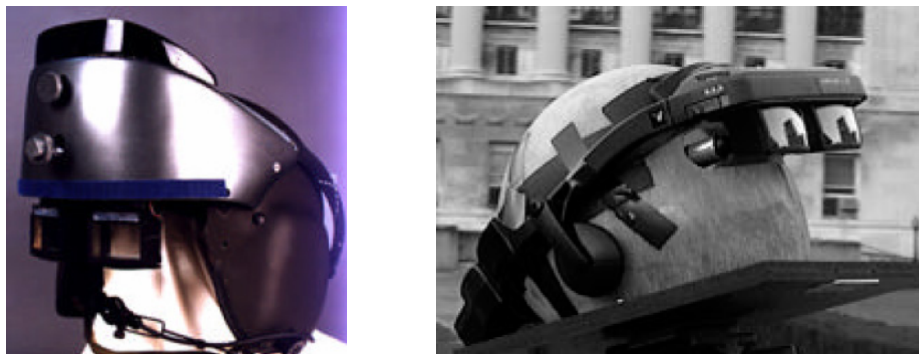


Abbildung 4: Links ist ein Gerät mit See-Through Technologie abgebildet [9], rechts ist ein Gerät mit Video-Blending zu sehen [16]. Bei dem rechten Gerät ist die Kamera für die Aufnahme, die an der Brille befestigt ist, deutlich zu erkennen.

Beide Gerätetypen bringen bestimmte Vorteile gegenüber dem anderen mit sich. Nach Azuma [9] hat der Ansatz des See-Through vier wesentliche Vorteile gegenüber dem Ansatz des Video-Blending. Zunächst ist dort die sehr günstige und effektive Implementierung zu nennen, da man hier weder zwei Videoströme, noch kleine Kameras benötigt. Zusätzlich dazu ergibt sich durch die Tatsache, die reale Welt mit „eigenem“ Auge zu sehen auch der Vorteil der extrem hohen Auflösung dieses „Videobildes“. Außerdem ergibt sich ein großer Sicherheitsvorteil für den Benutzer, wenn zum Beispiel die Stromversorgung der Kamera oder der Displays in der Brille unterbrochen wird. Der Benutzer einer Video-Blending Lösung könnte dann die reale Welt nicht mehr sehen. Als vierter wichtiger Punkt muss noch die Tatsache genannt werden, dass

es zu keiner „Augenverschiebung“ kommt, wie es bei einer Video-Blending Lösung der Fall ist, da hier die Augen die Linse der Kamera sind.

Azuma [9] definiert jedoch auch fünf Vorteile der Video-Blending Technik gegenüber der See-Through Technik. Ein großer Vorteil ist die Flexibilität der Kompositionsstrategien, da hier keine „echte“ Überlagerung, sondern eine digitale Berechnung der beiden Videodaten (real und virtuell) stattfindet. Zusätzlich ergibt sich die Möglichkeit das digitale Bild zu verändern, um die Störungen im erweiterten Bildbereich zu minimieren. Dies wäre mit See-Through Methoden auch möglich, jedoch erhöhen sie das Gewicht und die Kosten eines solchen Gerätes. Ein extrem wichtiger Vorteil bietet sich in der Fähigkeit die digital aufgenommenen Bilder und die virtuell erzeugten Objekte genau zu synchronisieren. Hierzu wird die geringe Latenz der Berechnung des Bildes ausgenutzt. Die letzten beiden Vorteile basieren auf Techniken der Bildbearbeitung, die mit See-Through Techniken nur sehr schwer realisierbar sind. Neben den zusätzlichen Tracking-Möglichkeiten über das digitale Bild können auch Helligkeitswerte realer und virtueller Objekte genau erkannt und korrigiert werden.

Nachdem die Geräte sowie deren Vor- und Nachteile behandelt und definiert wurden, können Anforderungen an Displays gestellt werden. Zu den geeigneten Anzeigegeräten gehören neben Head-Mounted-Displays auch Handhelds und Mobiltelefone. Diese unterstützen gerade den mobilen Einsatz, haben jedoch auch Einschränkungen (vgl. Kapitel 3). Zu den weniger geeigneten Geräten gehören Displays mit sehr hohem Gewicht, wie zum Beispiel Rucksäcke, die einen Computer beinhalten. Diese Gewichte machen eine freie Bewegung nur bedingt möglich und sind damit für mobile AR-Systeme nur eher ungeeignet. Generell lässt sich sagen, dass Geräte, die nicht an einen externen Computer (PC, Notebook) angeschlossen werden müssen, sondern selbst einen internen Computer besitzen, deutlich besser geeignet sind.

3 Mobile Endgeräte für Interaktion

Wie aus dem vorigen Kapitel bereits ersichtlich wurde, spielt die Mobilität der Benutzer eine große Rolle. Gerade aus diesem Grund werden immer häufiger mobile Endgeräte eingesetzt. Noch vor wenigen Jahren waren diese Geräte unhandlich und hatten ein hohes Gewicht. Im Zuge der schnellen Entwicklung können jedoch immer kleinere und handlichere Geräte eingesetzt werden. Abbildung 5 verdeutlicht den Trend der Entwicklung und den Unterschied beider Geräte in der Handhabung.

Ein weiterer gewünschter Aspekt ist die Multi-User Fähigkeit mobiler Endgeräte. Eine solche Architektur lässt sich jedoch heute nur über eine netzwerk-basierte Infrastruktur lösen (vgl. Abschnitt 3.2).

Mobile Endgeräte sind zwar erwünscht, sie bringen aber auch große Probleme mit sich. Gerade Mobiltelefone sind heute im Vergleich mit größeren Systemen noch sehr langsam, was besonders auf die Prozessorleistung zurückzuführen ist. Häufig sind zudem auch optische Trackingverfahren nicht direkt durchführbar, sondern müssen von einem Computer im Netzwerk bearbeitet und zurückgeschickt werden. Dies führt zu einer Erhöhung der Latenzzeit, die gerade bei AR ein enorm wichtiger Faktor ist.



Abbildung 5: Gegenüberstellung von HMD mit Rucksack (links) und einem handlichen PDA (rechts) [4].

3.1 Geräte und deren Eigenschaften für Augmented and Virtual Reality

Im Folgenden werden nun verschiedene Geräte mit ihren jeweiligen Eigenschaften für mobile AR vorgestellt. Im Großen und Ganzen werden drei Klassen unterschieden, und zwar vollständig mobile Geräte mit eingebautem Computer, mobile Geräte mit externem Computer und stationäre Geräte mit externem Computer aber in mobilen Maschinen (z.B. Autonavigation).

Die erste Klasse beinhaltet hauptsächlich Mobiltelefone und PDAs. Deren Vorteile liegen klar in der günstigen Anschaffung, sowie in der handlichen Bauweise. Zusätzlich besitzen sie die Möglichkeit, entweder mit eingebauter oder angeschlossener Kamera, Video-Blending günstig umzusetzen. Bei den Tracking-Verfahren spezialisieren sich diese Geräte heute noch auf GPS und Cell Tracking. Letzteres ist das Tracking durch zelluläre Funknetze, um einen Standort relativ zu mehreren Sendemasten bestimmen zu können. Beide Gerätetypen besitzen jedoch auch ihre Schwächen, die vor allem in der schwachen Rechnerleistung, der zu kleinen Auflösung der Videodaten (schlechte Markererkennung) und der optischen Verzerrung günstiger Kameras zu finden sind [12]. Abhilfe kann hier durch angepasste Algorithmen, das heißt Portierung eines bestehenden Systems für einen Desktop-PC auf ein mobiles Endgerät [1] erzielt werden. Eine weitere Möglichkeit besteht darin, dass ein Client/Server-System eingesetzt wird, wobei das Endgerät lediglich die Bilder aufnimmt, diese an den Server zur Auswertung sendet und dann die fertigen Daten zurückbekommt [2].

In der Klasse der mobilen Geräte mit externem tragbarem Computer findet man vor allem HMDs, die ihre Videodaten von einem Computer der im Rucksack des Benutzers ist erhalten. Durch diesen Computer ist gegenüber den PDAs und Mobiltelefonen eine deutlich höhere Rechenleistung möglich. Dadurch können neben dem gängigen Trackingverfahren GPS auch optische Trackingverfahren eingesetzt werden, um dem Benutzer so einen erhöhten Mobilitätsgrad zu ermöglichen (z.B. innerhalb und außerhalb von Gebäuden). Diese Geräte erlauben sogar beide Display-Verfahren, Video-Blending und See-Through. Doch auch sie haben Nachteile, die nicht zu vernachlässigen sind. Neben der Tatsache, dass sie extrem unhandlich und dadurch un-

komfortabel sind steht vor allem der extrem hohe Preis im Vordergrund. Alle diese Nachteile werden jedoch im Laufe der Zeit reduziert werden, da der Computer der heute noch in einem Rucksack untergebracht ist schon bald ein PDA sein könnte.

Als dritte Geräteklasse werden die Autonavigationssysteme beschrieben. Diese sind nur eingeschränkt mobil, da sie selbst fest in einem Auto eingebaut sind, dieses aber wieder mobil ist. Hier bieten sich analog zur zweiten Klasse ähnliche Vorteile, wie die hohe Rechnerleistung und das Tracking mittels GPS. Zusätzlich dazu kommt noch ein sehr gutes Display (Fernsehqualität), das eine genaue Darstellung ohne weiteres zulässt. Neben der bereits erwähnten eingeschränkten Mobilität sind diese Geräte zudem verhältnismäßig teuer.

Abschließend kann man sagen, dass die Tendenz ganz klar in Richtung einfachen und handlichen mobilen Geräten geht. Doch selbst hier ist der Trend von PDAs auf Mobiltelefone bereits erkennbar, da man für 2005 schätzt, dass mehr als 50% der verkauften Mobiltelefone bereits Kameras besitzen [12].

3.2 Infrastruktur für mobile Endgeräte

Wie bereits erwähnt müssen für bestimmte Szenarien Infrastrukturen bereitgestellt werden. Diese Infrastrukturen bieten mehrere Vorteile. Zum einen können in Client/Server-Systemen die angesprochenen, aufwendigen Berechnungen von einem Server anstelle des Endgerätes durchgeführt werden [2,10], zum anderen können mehrere Endgeräte in Multi-User Umgebungen synchronisiert und mit gleichen Daten beliefert werden [4]. Abbildung 6 verdeutlicht noch einmal das erste Szenario.

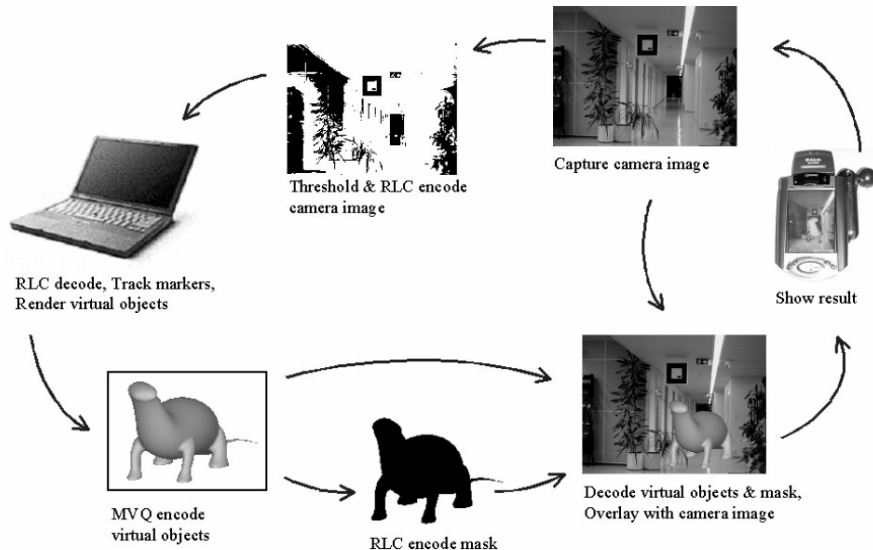


Abbildung 6: Client/Server-Architektur zur Markererkennung und Bildbearbeitung durch den Server

Der Erfolg einer solchen Infrastruktur hängt jedoch maßgeblich von der Geschwindigkeit der Netzwerke sowie von der Größe und Frequenz der zu übertragenden Daten ab. Doch gerade im Freien gibt es massive Probleme mit den Bandbreiten und Übertragungsgeschwindigkeiten. WLAN bietet zwar noch hohe Bandbreiten an, dafür nimmt aber die Reichweite ab. Ein flächendeckendes Netz wie beispielsweise GSM oder UMTS bietet zwar die nahezu lückenlose Anbindung des Gerätes an, dafür sinken die Bandbreiten.

Der Trend wird also dahin gehen, dass die mobilen Endgeräte zunehmend selbst die Berechnung der visuellen Daten vornehmen. Dies ist sicherlich auch nur eine Frage der Zeit, bis die technische Entwicklung so weit fortgeschritten ist, dass mobile Endgeräte genug Rechenleistung für diese Operationen bereitstellen.

Eine interessante Frage der Infrastruktur stellt sich im Bereich der Computerspiele. Da diese bereits rein virtuell am Computer existieren, ist der gewünschte Effekt, die vorhandenen Spieler am Computer mit neuen reellen Spielern zu kombinieren. Dazu muss eine Infrastruktur geschaffen werden, die es ermöglicht „normale“ Spieler an einem Rechner mit Daten der „reellen“ Spieler in der realen Umgebung zu versorgen und umgekehrt. Weiterhin müssen zwei verschiedene Darstellungen der Spielumgebung vorhanden sein, denn lediglich die Spieler an einem Desktop-PC müssen mit vollständig virtuellen Daten beliefert werden, wohingegen die Spieler in der realen Umgebung lediglich Gebäudekoordinaten benötigen. Diese Darstellungen müssen nun auch in beiden Systemen vorhanden sein [13].

3.3 Interaktion für mobile Endgeräte

Sobald einmal virtuelle Objekte in dem Bild der realen Umgebung angezeigt wurden, sollen dem Benutzer selbstverständlich auch Möglichkeiten der Interaktion mit diesen gegeben werden. Dabei werden die verschiedensten Interaktionsformen erforscht. Ein Autonavigationssystem kann beispielsweise mittels Sprachkommando oder auch mittels Joystick (BMW iDrive) bedient werden [8]. Etwas schwieriger gestaltet sich schon die Interaktion mit „echten“ mobilen Systemen, wie zum Beispiel PDAs oder Mobiltelefonen. Bei Mobiltelefonen hat der Benutzer zwar eine Tastatur, diese ist aber auf 10 Eingabetasten und weitere gerätespezifische Tasten beschränkt. PDAs hingegen besitzen eine solche Tastatur nicht, haben dafür aber einen digitalen Stift (Pen), der das Klicken auf dem Bildschirm übernimmt.

Eine neue Eingabeform für diese Geräte besteht nun in der Form von Sprachkommandos [5]. Dadurch können Benutzer einfache, vorgegebene Fragen an das Gerät stellen, zum Beispiel wer sich gerade in einem vom Gerät erkannten Raum befindet. Das Gerät selbst kann mittels der Text-to-Speech Technologie auch antworten.

Eine weitere Möglichkeit besteht in dem Ansatz für optische Trackingverfahren, zusätzliche Objekte, z.B. Stifte oder Pinsel, mit Markern zu versehen. Durch die Erkennung dieser Objekte erhält der Benutzer die Möglichkeit mit der virtuellen Welt zu kommunizieren [11]. Diese Technik eignet sich auch für Multi-User Umgebungen, da das optische Tracking auf beiden Systemen gleichermaßen eingesetzt werden kann. Abbildung 7 verdeutlicht diese Interaktionsform für ein einfaches Mal- und Zeichenprogramm [11].



Abbildung 7: Ein Benutzer, der mit einem optisch erfassten Stift (Pinsel) arbeitet. Mit diesem Pinsel kann sowohl eine Farbe ausgewählt werden (links), als auch das Objekt mit gewählter Farbe bemalt werden (rechts) [11].

Als Alternative kann noch die Möglichkeit genannt werden, reale Objekte mit Computern zu versehen. Diese können problemlos mit dem eigenen Endgerät kommunizieren und vermitteln dem Benutzer ein Gefühl der Realität, da es sich letztlich nicht mehr um virtuelle Objekte handelt [14]. Weiterhin kann man den Benutzer auch direkt in die Interaktion mittels Kombination von Gesten und Sprache einbeziehen. So ist es möglich, auf ein virtuelles Gebäude in einem Navigationssystem zu zeigen, und das Gerät gleichzeitig zu fragen, um welches Gebäude es sich handelt [7].

Zusammenfassend lässt sich sagen, dass die Interaktionsformen noch nicht sehr weit entwickelt wurden, da heute viele Systeme nur eine bestimmte Interaktion erfordern. Dies dürfte sich im Zuge der weiteren Entwicklung von AR sehr bald ändern.

4 Anwendungsbeispiele

Nachdem alle technischen Voraussetzungen für ein erfolgreiches AR-System definiert wurden, werden jetzt aktuelle Entwicklungen in diesem Bereich vorgestellt. Die Systeme unterscheiden sich nach den Gesichtspunkten Tracking und eingesetzten Endgeräten sowie Infrastrukturen. Die schlüssigste Unterteilung ist jedoch, sie nach der Art der Anwendung bzw. des Einsatzgebietes zu klassifizieren.

4.1 Navigationssysteme

Da Navigationssysteme heute immer beliebter werden, beschäftigt sich auch die Forschung der AR intensiv mit diesen. Doch nicht nur die relativ einfachen Autonavigationssysteme stehen hier im Mittelpunkt, sondern auch immer mehr Systeme für den Fußweg, die auf mobilen Endgeräten basieren.

Ein erster Ansatz ist das System INSTAR (Information and Navigation Systems Through Augmented Reality) [8]. Grundlegend basiert es auf einem Autonavigationssystem, das GPS für das Tracking einsetzt. Das Display des Systems ist jedoch ein PDA, der ohne Probleme auch für Wege zu Fuß eingesetzt und mitgeführt werden

kann. Die Neuerung liegt gerade in diesem Punkt, denn ein einfaches Tracking bei Fußwegen ist nur dann möglich, wenn der Weg nicht innerhalb eines Gebäudes ist. Genau hier versucht dieses System eine möglichst einfache Lösung zu präsentieren. Die Idee besteht darin, dass das System zu jedem Zeitpunkt GPS-Daten erwartet, was innerhalb von Gebäuden über eine weitere Instanz ermöglicht wird. In einem geschlossenen Raum wird der PDA räumlich erfasst. Ein weiteres Gerät, der „Indoor Tracking Server“ wertet diese Daten aus und sendet sie zu einem generischen „Tracking Supplier“, der die Daten in GPS-Koordinaten, durch genaue Kenntnis der Gebäudedaten und der relativen Position des PDAs zu diesen, umwandelt. Diese GPS-Koordinaten können dann ohne weitere Zwischenschritte zum INSTAR-System übertragen werden und ermöglichen so wieder eine, dem Gerät vertraute Positionierung. Abbildung 8 zeigt noch einmal die wesentlichen Architekturen der Tracking-Verfahren im INSTAR-System.

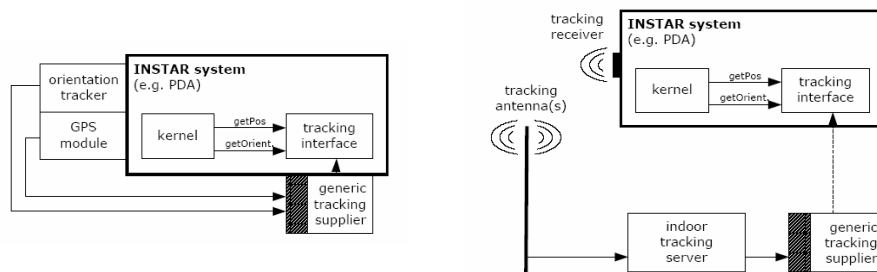


Abbildung 8: Trackingverfahren für INSTAR. Direkte Positionierung mit GPS im Freien (links), indirekte Positionierung innerhalb von Gebäuden (rechts) [8]

Die Visualisierung der Daten erfolgt bei INSTAR durch Anzeigen der virtuell gesuchten Wege direkt auf dem realen Bild. Die realen Bilddaten werden im Auto mit Hilfe einer fest installierten Kamera, auf dem PDA mit Hilfe einer angeschlossenen Kamera erzeugt. Abbildung 9 stellt die zwei Visualisierungen genauer dar.

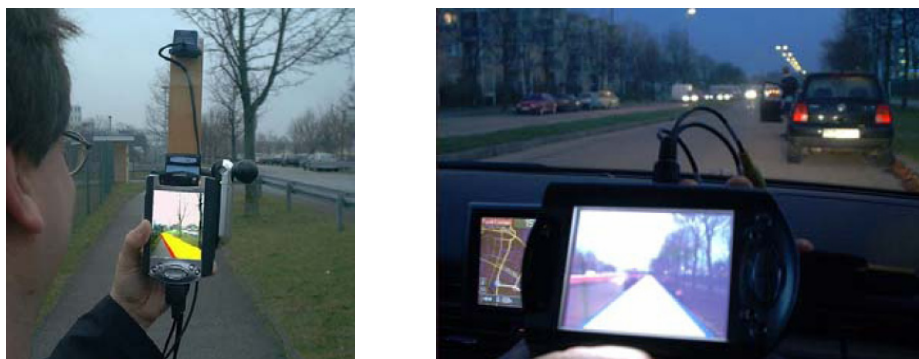


Abbildung 9: Visualisierung von INSTAR auf einem PDA für freie Navigation zu Fuß (links), oder im fest installiert im Auto (rechts) [8]

Ein weitere Anwendung im Bereich der Navigationssysteme ist der BPN (**B**MW **P**ersonal **N**avigator) [7]. Dieses System verbindet mehrere Geräte und Techniken zu einer einzigen Navigationsanwendung: Ein Desktop-PC für die Routenplanung, Ein PDA zur Navigation zu Fuß und ein Autonavigationssystem. Diese Instanzen spielen über eine klare Struktur der Daten (XML) zusammen. Die Routenplanung, die an einem Desktop-PC erfolgt, verhält sich ähnlich wie heutige Planungssysteme. Das mobile Endgerät ist hier sprachgesteuert und erkennt einfache vordefinierte Regelgrammatiken nach dem „command-and-command“ Prinzip. Dadurch kann der Benutzer mit Kommandos wie „zoom in“ mit dem Gerät interagieren. Zusätzlich wurde eine Medienfusions-Komponente eingefügt, die es zusätzlich zur Spracheingabe ermöglicht, Gesten zu erkennen und mit der Sprache zu kombinieren. Dadurch erhält der Benutzer mehr Flexibilität in seinen Anfragen an das System.

Das Tracking wurde hier mittels GPS, über Bluetooth an den PDA angeschlossen, umgesetzt. Selbst wenn das Signal schwach ist, kann noch navigiert werden, in dem ein so genannter „Map-Matching-Algorithmus“ angewendet wird.

4.2 Informations- und Tourguide-Systeme

Eine weitere wichtige Rolle spielen die Informations- und Tourguide-Systeme. Diese Systeme, die eng verbunden mit Navigationssystemen sind, unterstützen den Benutzer mit zusätzlichen Informationen zu Gebäuden, antiken Stätten oder anderen Sehenswürdigkeiten. Ein System aus diesem Bereich stellt ARCHEOGUIDE (**A**ugmented **R**eality-based **C**ultural **H**eritage **O**n-site **G**UIDE) dar [3], mit Hilfe dessen es möglich ist, antike Stätten des alten Griechenlands zu erkunden und zu sehen, wie z.B. Wettkämpfe in damaliger Zeit stattgefunden haben könnten (siehe Abbildung 10).

Dieses System setzt eine Client/Server-Architektur ein, in der ein zentraler Server die mobilen Geräte mit Positionsdaten beliefert und alle wichtigen multimedialen Daten, wie z.B. 2D- und 3D-Bilder, Video, Audio und Text, enthält. Die mobilen Endgeräte bestehen noch aus Head-Mounted-Displays und „Rucksack-Computern“, die mit Hilfe von WLAN mit dem Server kommunizieren (siehe Abbildung 10). Der Grund für eine Netzwerklösung liegt hier zum einen in der hohen Anzahl mobiler Endgeräte (bis zu 50), zum anderen an der Abdeckung aller begehbaren Bereiche.



Abbildung 10: Das ARCHEOGUIDE-System mit den mobilen HMD und Computer (links), sowie einer Szene eines antiken Wettkampfes (rechts) [3]

Eine interessante Variante wird für das Tracking verwendet. Neben GPS für die Position und einer Magnetfeldstärkemessung in drei Achsen (Genauigkeit ca. $0,2^\circ$) für die Orientierung wird zusätzlich optisches Tracking eingesetzt. Das letztgenannte Verfahren basiert auf einem Vergleich des aktuellen Blickfeldes mit Referenzbildern.

4.3 Wartungssysteme

Auch in Industrieanlagen erhalten Virtual Reality und Augmented Reality im Bereich der Wartung Einzug. Ein zu erwähnendes System ist PARIS [6], bei dem Wartungstechniker mit Hilfe eines PDAs versteckte Teile einer Maschine überwachen können. Hierzu wurden mehrere Bereiche einer Fabrikanlage mit Markern (ähnlich zum AR-ToolKit) ausgestattet, um die Lokalisierung des Benutzers zu erfassen. Sobald dem Gerät der Standort bekannt ist, können verborgene Teile, wie zum Beispiel Röhren mit ihren Verbindungen, angezeigt werden. Zusätzlich dazu werden Informationen über diese Gegenstände, wie zum Beispiel die Temperatur, o. ä., bereitgestellt (siehe Abbildung 11).

Die Interaktion mit dem Benutzer findet durch eine sprachgesteuerte Software (VoiceXML) statt. Der Benutzer kann sowohl einfache Fragen an das Gerät stellen und erhält außerdem kontextbezogene Sprachausgaben durch das Gerät. Diese Dialoge entstehen direkt nach dem Erkennen eines Markers.

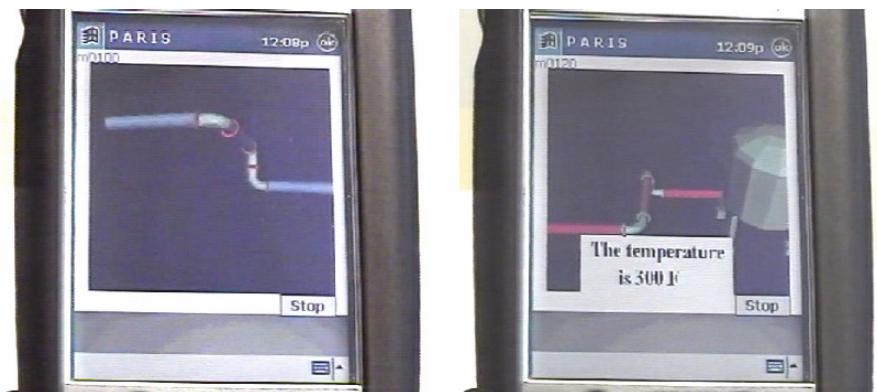


Abbildung 11: Visualisierung des PARIS-Systems: 3D Ansicht eines Rohrs mit den Verbindungen in rot (links), Ansicht zweier Rohre in rot markiert mit einem Dialog für die Temperatur (rechts) [6]

Dieses System setzt VoiceXML nun direkt auf dem PDA ein, ohne eine komplexe Client/Server-Architektur zu verwenden. Die subjektiven Ergebnisse der Entwickler sind zunächst zwar gut, eine ausführliche Fallstudie zu diesem System existiert jedoch nicht. Es bleibt also abzuwarten, ob ein PDA mit einer eigenen Sprachsteuerung bereits in der Form genutzt werden kann.

4.4 Multiplayer-Spiele

Eine der interessantesten Gruppen der Anwendungsbeispiele findet sich im Bereich der Spiele. Gleichzeitig ist hier ein hohes Maß an Multi-User Fähigkeiten gefragt, die so nicht ohne weiteres umsetzbar sind. In dieser Klasse werden nun zwei verschiedene Spiele vorgestellt. Beide Beispiele verwenden einen zentralen Server im Netzwerk, um Applikations- und Szenedaten zu teilen.



Abbildung 12: Zwei PDAs mit laufendem Spiel [4]

Die erste Applikation ist „The Invisible Train“ [4], deren Idee es ist, virtuelle Spielzeugzüge auf realen Holzgleisen fahren zu lassen (siehe Abbildung 12). Dabei können Benutzer ihren Zug schneller oder langsamer fahren lassen und die Weichen steuern (siehe Abbildung 13) und müssen aufpassen, dass keine zwei Züge zusammenstoßen. Das Spiel setzt einen PDA mit angeschlossener Kamera ein. Auf dem PDA selbst wird das ARToolKit verwendet, um die Marker, die eine feste räumliche Anordnung zwischen den Gleisen haben, erkennen zu können. Relativ zu diesen Markern können dann die Züge gezeichnet werden. Die Interaktion mit der Anwendung wird mit Hilfe eines Stifts auf dem PDA realisiert, durch den der Benutzer alle genannten Aktionen durchführen kann.

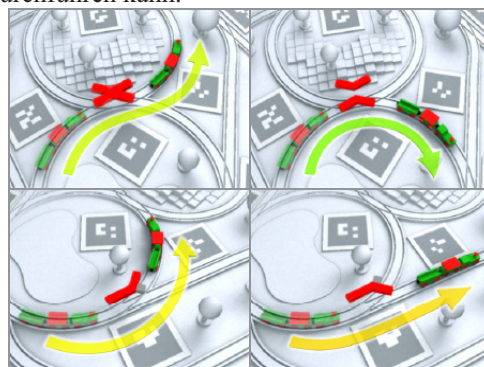


Abbildung 13: Weichensymbole und deren Effekte auf Zugrouten [4]

Dieses System wurde stabil und robust über mehrere Wochen getestet. Dazu hatten untrainierte Benutzer immer wieder die Möglichkeit das Spiel zu testen und zu spielen. Schätzungsweise fünf- bis sechstausend Menschen haben dieses System während insgesamt vier Auswertungszyklen getestet.

Ein weiteres Spiel basiert auf dem bereits existierenden Multiplayer-Spiel Quake und heißt demzufolge ARQuake [13]. Die Zielsetzung des Systems ist es ein Level-basiertes Spiel, das sowohl innen als auch außen einsetzbar ist. Die Benutzer sollen sich dabei frei bewegen können und verwenden ein HMD mit angeschlossenem Computer im Rucksack. Zusätzlich wird ein weiteres Gerät benötigt, das in diesem Fall eine Spielzeugpistole, die auch Geräusche erzeugen kann, ist. Dadurch erhält der Benutzer eine realere Vorstellung des Spiels.



Abbildung 14: Geräte von ARQuake. Ein tragbarer Computer auf dem Rücken (links), sowie das HMD und die „echte“ Pistole (rechts) [13]

Das eingesetzte Tracking-Verfahren basiert sowohl auf optischen Techniken durch Markererkennung mit dem ARToolKit, als auch auf lernfähigen Algorithmen, die zusätzlich optisch die Position feststellen können. Bewegt sich der Benutzer weit entfernt von Gebäuden (mehr als 50 m) kommt GPS zum Einsatz, da hier die Genauigkeit nicht von größter Relevanz ist. Kommt er näher an das Gebäude, werden große Marker (ca. 1 m Seitenlänge) des ARToolKits verwendet, die logisch mit den Gebäuden verknüpft sind. Innerhalb von Gebäuden werden dann ebenfalls große Marker verwendet, jedoch fällt die Hybridlösung mit GPS weg, da dies innerhalb von geschlossenen Räumen nicht möglich ist.

Eine weitere Besonderheit ist die Möglichkeit, dass dieses Spiel sowohl von „realen“ Spielern, als auch von Spielern an einem Desktop-PC gleichzeitig verwendet werden kann. Die Synchronisation wird auch hier von einem zentralen Server übernommen, der mit den mobilen Geräten über WLAN kommuniziert.

Die Interaktion zwischen „realen“ Spielern und Desktop-PC-Spielern wird den Benutzern mit Spracherkennung zur Verfügung gestellt. So gibt es die Möglichkeit, durch Sprachkommandos auf Objekte zu deuten. Danach wird vom System ein großer Pfeil über der Darstellung des Spielers in der virtuellen Welt angezeigt. Spielen alle Spieler mit tragbaren Geräten entfällt diese Option selbstverständlich, da sie physikalisch deuten können.

5 Zusammenfassung und Diskussion

Nachdem nun einige ausgewählte Systeme betrachtet wurden sollen diese noch einmal nach ihren Möglichkeiten verglichen werden. Während Navigationssysteme heute schon einen großen Schritt in Richtung der Tracking-Verfahren (GPS) gemacht haben, bereitet gerade das Auswerten und Darstellen für komplexere Szenen noch große Probleme im Bezug auf handliche, mobile Endgeräte. Daher ist zu beobachten, dass zum Beispiel im Bereich der Spiele der Einsatz kleinster Geräte (PDAs) immer mehr gewünscht ist. Dadurch müssen allerdings weitaus höhere Leistungsfähigkeiten der einzelnen Geräte erreicht werden.

Auch beim Tracking sind genauere Daten applikationsabhängig. Während GPS für ein Navigationssystem durchaus ausreicht, gestaltet sich dies innerhalb von Gebäuden schon schwieriger. Hier werden sehr genaue Daten verlangt, da Unterschiede zwischen verschiedenen Räumen sehr viel geringer sind. Die hier vorgestellten Ansätze, auch im Zuge der Kompatibilität, sind durchaus viel versprechend und werden in den nächsten Jahren sicherlich weiter entwickelt werden. Das optische Tracking hingegen ist bereits sehr präzise, kann jedoch nicht immer von kleinen mobilen Geräten selbstständig durchgeführt werden. Dies ist besonders abhängig von der Komplexität der zu erkennenden Marker oder Gegebenheiten.

Bei der Darstellung hingegen stehen die PDAs anderen Systemen kaum noch nach. Die Anzeige ist zwar durch die Displaygröße beschränkt, liefert aber dennoch ein sehr gutes und scharfes Bild. Wie allerdings an den vorgestellten Systemen zu sehen ist werden Mobiltelefone so gut wie nicht eingesetzt. Das liegt vor allem an deren zu geringer Displaygröße, den eingebauten, günstigen (und damit schwachen) Kameras und ihrer Prozessorleistung. Gerade ihre Kameras eignen sich kaum für das optische Tracking, da die Bilder mit einer zu geringen Auflösung erzeugt werden, und so nicht mehr genau ausgewertet werden können.

Die Interaktionsformen sind heutzutage noch der am wenigsten erforschte Bereich. Während Mobiltelefone noch eine Eingabemöglichkeit über Tastatur bieten, die jedoch auf 10 Tasten beschränkt ist, haben PDAs kaum eine solche Möglichkeit. Sie bedienen sich eher der Eingabemethode mit einem Pen der auf das Gerät zugeschnitten ist. Tragbare Computer haben gegenüber den anderen zwei Geräten den Vorteil, dass sehr viele verschiedene Peripherie-Geräte angeschlossen werden können. Heute Interaktionsformen die vom Desktop-PC her bekannt sind, wären also eine denkbare Alternative.

Zusammenfassend lässt sich also sagen, dass sich der Bereich der Virtual and Augmented Reality in den kommenden Jahren massiv verbessern wird. Dies hat letztlich damit zu tun, dass die benötigten Technologien und auch die Endgeräte erheblich verbessert werden dürften. Zusätzlich werden neue Algorithmen auf der Softwareseite, sowie verbesserte Tracking-Verfahren einen erheblichen Beitrag für den Erfolg dieser Systeme beitragen.

Bleibt noch zu betrachten, wie Benutzer eine solche Technologie annehmen. Werden in Zukunft nun viele Touristen mit Head-Mounted-Displays durch die Straßen einer Großstadt laufen, oder werden sich doch die kleineren Geräte, wie PDAs und Mobiltelefone durchsetzen können. Eine Frage, die sicherlich erst in einigen Jahren geklärt werden kann.

Literatur

1. Wagner D., Schmalstieg, D., "ARToolKit on the PocketPC Platform", Wien, 2003
2. Pasmán, W., et al. "Augmented Reality with Large 3D Models on a PDA - Implementation, Performance and Use Experiences", In Proceedings of the International Conference on Virtual-Reality Continuum & Its Applications in Industry (VRCAI, Singapore, June 16-18), pp. 344-351, 2004
3. Vlahakis, V., et al. "3D Interactive, On-Site Visualization of Ancient Olympia", In Proceedings of the First International Symposium on 3D Data Processing Visualization and Transmission (3DPVT'02), IEEE, 2002
4. Wagner, D., et al. "Towards Massively Multi-User Augmented Reality on Handheld Devices", In Proceedings of the 3rd International Conference on Pervasive Computing (Pervasive 2005), pp. 208-219, Munich, Germany, 2005
5. Goose, S., et al. "Augmented Reality in the Palm of your Hand: A PDA-Based Framework Offering a Location-based, 3D and Speech-Driven User Interface", In Proceedings of the Workshop on "Wearable Computing" tcmc2003, Graz, Austria
6. Goose, S., et al. "Mobile 3D Visualization and Interaction in an Industrial Environment", In HCI International, Crete, Greece, June 2003
7. Krüger, A., et al. "The Connected User Interface: Realizing a Personal Situated Navigation Service", in Proceedings of IUI 2004, January 13-16, 2004, Madeira, Funchal, Portugal
8. Narzt, W., et al. "Pervasive Information Acquisition for Mobile AR-Navigation Systems", In Proceedings of the 5th IEEE Workshop on Mobile Computing Systems & Applications, Monterey, California, USA, October 2003
9. Azuma, T., "A Survey of Augmented Reality", Teleoperators and Virtual Environments 6, 4, pp. 355-385, 1997
10. Ebbesmayer, P., et al. "AR-PDA: Innovative Product Marketing for Innovative Products", In Proceedings of the International Status Conference Virtual and Augmented Reality, Leipzig, 2004
11. Reitmayr, G., Schmalstieg, D., "Mobile Collaborative Augmented Reality", in Proceedings of ISAR, 2001, New York, USA
12. Möhring, M., et al. "Optical Tracking and Video See-Through AR on Consumer Cell-Phones", In proceedings of International Symposium on Augmented and Mixed Reality (ISMAR'04), pp. 252-253, 2004
13. Thomas, B.H., et al. «First Person Indoor/Outdoor Augmented Reality Application: ARQuake», In: Journal of Personal and Ubiquitous Computing
14. Cheok, A.D., et al. Human Pacman : "A Mobile Entertainment System with Ubiquitous Computing and Tangible Interaction over a wide outdoor area", Lecture Notes of Computer Science, Vol. 2795, pp. 209-224
15. Burigat, S., Chittaro, L. "Location-aware Visualization of VRML Models in GPS-based Mobile Guides", Proceedings of Web3D 2005: 10th International Conference on 3D Web Technology, ACM Press, New York, pp. 57-64, April 2005
16. Feiner, S., et al. "A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment", In Proceedings of the 1st International Symposium on Wearable Computers (ISWC '97), October 1997