

Tracking einer Person im urbanen Umfeld mit Positions- und Blickrichtungsbestimmung auf Basis eines 3D-Stadtmodells



Vom Fachbereich Informatik
der Technischen Universität Darmstadt
genehmigte

DISSERTATION

zur Erlangung des akademischen Grades eines
Doktor-Ingenieurs (Dr.-Ing.)
von

Dipl.-Ing. Ursula Kretschmer

Referenten der Arbeit:

Prof. Dr. J. L. Encarnação

Prof. Dr. R. Bill

Tag der Einreichung:

10. Januar 2004

Tag der mündlichen Prüfung:

27. Februar 2004

Dank

Diese Arbeit entstand während meiner Tätigkeit als wissenschaftliche Mitarbeiterin am Fraunhofer-Institut für Graphische Datenverarbeitung.

Mein besonderer Dank gilt Herrn Prof. Dr. h.c. Dr.-Ing. José L. Encarnação, der mir in seinem Institut die Möglichkeit gab, mit engagierten und kreativen Wissenschaftlern zu arbeiten, und der mit wertvollen Anregungen und Hinweisen zum Gelingen dieser Arbeit beitrug.

Herrn Prof. Dr. Bill danke ich herzlich für die freundliche Übernahme der Aufgabe des Zweitgutachters und für die schnelle, sorgfältige und wissenschaftlich kritische Durchsicht der Arbeit.

Mein Dank gilt allen Mitarbeiterinnen und Mitarbeitern der Abteilung Graphische Informationssysteme am Fraunhofer-IGD für die freundliche und kreative Atmosphäre und die kollegiale Hilfe bei der täglichen Arbeit. Besonders danke ich meinem ehemaligen Abteilungsleiter, Herrn Dr. Uwe Jasnoch, für das Schaffen des nötigen Freiraums und die ständige Ermutigung.

Ich möchte mich auch bei allen Studentinnen und Studenten bedanken, die unter meiner Betreuung im Rahmen von Diplomarbeiten, Praktika und sonstigen Tätigkeiten in der Abteilung ebenfalls zum Gelingen dieser Arbeit beigetragen haben.

Von ganzem Herzen danke ich meinem Freund Jochen Ackermann sowie meinen Eltern und meinen Geschwistern Matthias und Wolfgang, auf deren Hilfe und Rückhalt ich mich auch in schwierigen Zeiten verlassen konnte. Sie haben mich stets dabei unterstützt, dieses Projekt zu einem positiven Ende zu bringen.

Inhaltsverzeichnis

1	Einleitung.....	6
1.1	<i>Motivation</i>	6
1.2	<i>Problemstellung</i>	8
1.3	<i>Ziel und Aufbau der Arbeit</i>	9
2	Bestimmung von Position und Orientierung einer Person	12
2.1	<i>Infrarot</i>	13
2.2	<i>Netzbasierte Positionierungsmöglichkeiten.....</i>	14
2.3	<i>GPS.....</i>	18
2.4	<i>GLONASS</i>	23
2.5	<i>GALILEO</i>	23
2.6	<i>Inertiale Sensoren.....</i>	24
2.7	<i>Magnetkompass und Neigungssensoren</i>	25
2.8	<i>Videobasierte Verfahren</i>	26
2.9	<i>Fusion verschiedener Sensoren.....</i>	27
2.10	<i>Zusammenfassung</i>	28
3	Outdoor AR-Systeme	29
3.1	<i>Systeme basierend auf GPS und Orientierungssensoren</i>	29
3.2	<i>Stützung der Orientierungssensoren mit Störungskarten</i>	30
3.3	<i>Offline-Stützung von Orientierungssensoren mit Hilfe von Bildverarbeitungstechniken</i>	30
3.4	<i>Horizontsilhouetten</i>	31
3.5	<i>Einbeziehung von Daten eines 3D-GIS.....</i>	32
3.6	<i>Matching des Kamerabildes mit Referenzphotos</i>	32
3.7	<i>Matching des Kamerabildes mit einem 3D-Modell.....</i>	33
3.8	<i>Zusammenfassung</i>	34

4	3D-Stadtmodelle.....	37
4.1	<i>Photogrammetrische Erfassung.....</i>	38
4.2	<i>Erfassung mit Hilfe des Laserscanning-Verfahrens.....</i>	40
4.3	<i>Photorealistische Erfassung.....</i>	44
4.4	<i>Manuelle Modellierung.....</i>	45
4.5	<i>Beispiele für 3D-Stadtmodelle.....</i>	45
4.6	<i>Zusammenfassung</i>	50
5	Der mobile Anwender im urbanen Umfeld.....	51
5.1	<i>Anwendungsszenarien.....</i>	53
5.2	<i>Dynamische Auswahl eines Dienstes.....</i>	56
5.3	<i>Anforderungen an die Trackingtechnik.....</i>	60
5.4	<i>Aktuelle Sensordaten.....</i>	66
5.5	<i>Parametrisierung der Trackingtechniken.....</i>	70
5.6	<i>Zusammenstellung der Parameter.....</i>	74
5.7	<i>Zuordnung der Trackingtechnik.....</i>	77
5.8	<i>Zusammenfassung des Verfahrens</i>	79
6	Hochgenaues Trackingverfahren mit Hilfe eines 3D-Modells	81
6.1	<i>Anforderungen an ein urbanes AR-System.....</i>	81
6.2	<i>Überblick über das Verfahren</i>	83
6.3	<i>Bildbasierte Informationsgenerierung.....</i>	85
6.4	<i>Korrespondenzfindung im 3D-Modell und Kamerabild</i>	93
6.5	<i>Geometrische Beziehung zwischen Kamerabild und 3D-Modell.....</i>	101
6.6	<i>Erweiterung des Verfahrens.....</i>	114
6.7	<i>Zusammenfassung</i>	116
7	Praktische Umsetzung.....	117
7.1	<i>Übersicht</i>	117
7.2	<i>Verbindung verschiedener Trackingtechniken.....</i>	118
7.3	<i>Tracking auf Basis eines 3D-Modells</i>	122

7.4	<i>Grenzen der umgesetzten Algorithmen</i>	126
7.5	<i>Zusammenfassung</i>	126
8	Zusammenfassung und Ausblick	128
8.1	<i>Zusammenfassung der Ergebnisse</i>	128
8.2	<i>Ausblick</i>	131
9	Literaturverzeichnis	132

1 Einleitung

1.1 Motivation

Mobilität ermöglicht es, unseren Tätigkeitsbereich auszudehnen, und dabei sowohl im Berufsleben als auch in unserer Freizeit fremde Städte und Regionen zu erkunden. Informationen zu diesen neuen Zielen vor der Reise aber auch vor Ort können mit Hilfe von Büchern und Karten, aber auch über interaktive Informationsvermittlungsdienste am Computer abgefragt werden. GIS (Geoinformationssysteme) bieten die Visualisierung geographischer Informationen. Dafür werden topographische und thematische Karten, statistische Graphen, aber auch virtuelle Oberflächen und sogar Animationen angeboten [Kret00], wobei die automatische Verknüpfung zu zusätzlichen Informationen, die über den geographischen Bezug hinausgehen, die Grundlage dieser Möglichkeiten ist. Das Haupteinsatzgebiet vor allem kommerzieller GIS-Werkzeuge ist die Datenanalyse. Für den Zugang eines großen Nutzerkreises zu diesen Daten werden webbasierte Lösungen angeboten, deren Funktionalitätsumfang auf bestimmte Aufgaben zugeschnitten ist. Um diese Werkzeuge auch vor Ort nutzen zu können, gibt es heutzutage Kiosksysteme, an denen sich ein Tourist oder Geschäftsreisender am Bahnhof, vor dem Fremdenverkehrsamt oder an einem touristisch bedeutsamen Ort informieren kann.

Durch die rasanten Entwicklungen im Bereich mobiler IT-Lösungen steigt das Interesse an Diensten für Nutzer von PDAs (Personal Digital Assistants), sich die entsprechenden Informationen auf ihr persönliches Gerät zu laden und immer dann zu nutzen, wenn es ihnen hilfreich oder interessant erscheint. Abbildung 1 zeigt eine Vorhersage des Marktvolumens für verschiedene Informationsdienste, die auf räumlichen Daten basieren.

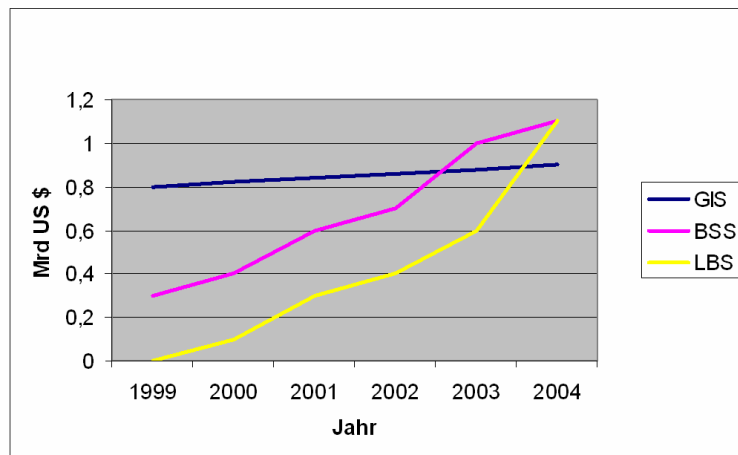


Abbildung 1: Prognosen Marktvolumen GIS, BSS, LBS [Oste01]

Daraus geht hervor, dass GIS den Marktanteil, den sie im Moment haben, leicht vergrößern werden. Business Support Systeme (BSS) nehmen in etwas stärkerem Maße zu. Sie erlauben die Verknüpfung von Geschäftsinformationen zu räumlichen Daten. Dies bedeutet, dass der Einsatz von Werkzeugen für räumliche Daten durch die Übertragung in andere Bereiche zu einer Aufwertung führt. Dies trifft ganz besonders für Location Based Services (LBS) zu, für welche die Vorhersage zeigt, dass sie auch in nächster Zukunft sehr stark zunehmen werden. LBS sind Dienste, die ortsbezogene Informationen auf einem Mobiltelefon oder PDA anbieten. Damit sind sie nicht nur ein Werkzeug zur Analyse räumlicher Daten, sondern bieten auch die Möglichkeit, verschiedene Dienste mit der aktuellen geographischen Position zu verbinden. So werden zahlreiche Anwendungen für einen viel größeren Benutzerkreis attraktiv. LBS benötigen eine Positionierungstechnik, die beispielsweise durch eine GPS-Karte (Global Positioning System) ermöglicht wird. Dann sind nicht nur Autonavigation oder Flottenmanagement, sondern auch Informationen über Restaurants in der Umgebung, den nächsten Bankautomat oder kleine Stadtführungen mit einem mobilen Gerät möglich. Diese Geräte sind damit nicht nur handlicher sondern auch adaptiver als die meisten Desktop-Lösungen, mit denen man weiterhin Informationen über eine bestimmte Stadt abrufen kann.

Die Zunahme der Verknüpfung von Anwendungen mit Techniken, die einen bequemeren Gebrauch erlauben, passt sich sehr gut in die derzeitigen Fortschritte innerhalb der Informationsgesellschaft ein. In vielen Bereichen geht man dazu über, bestimmte Anwendungen nicht nur Experten anzubieten, sondern möglichst viele Menschen mit einer neuen Art von Infrastruktur und daraus abgeleiteten Diensten zu versorgen. Zu diesen Entwicklungen zählt ebenfalls die Kopplung von virtuellen Welten mit Geoinformationen. Diese ermöglichen dem Nutzer bereits heute eine innovative Form der Interaktion mit Geodaten [Bill02]. Aber auch die Zunahme von mobilen, kleinen und leistungsstarken Geräten sowie von Systemen, die das Sammeln und Verarbeiten von Anwenderdaten und Informationen zur Umgebung in hohem Maße automatisieren, ist Teil dieser Entwicklungen.

Wenn zusätzlich zur Position noch die Blickrichtung für verschiedene Aufgaben detektiert wird, kann das Leistungsspektrum von AR-Systemen in Anspruch genommen werden. AR (Augmented Reality) bedeutet, dass reale Objekte mit rechnergenerierten virtuellen Objekten im Sichtfeld des Betrachters überlagert werden [Coor00a]. Damit ist es möglich, Pfeile, Textfelder, Links oder 3D-Modelle im Sichtfeld des Benutzers zu platzieren, um dem Anwender Erklärungen oder Veranschaulichungen zu vermitteln. In diesem Bereich spielen neue Formen der Interaktion ebenfalls eine große Rolle. Die Verbindung von LBS und AR bietet eine Möglichkeit, den Bedarf zu decken, den mobile Menschen heutzutage fordern und ihnen die Mittel anzubieten, die ihnen die Technik heute gestattet. Abbildung 2 zeigt den Einsatz solch kleiner mobiler Geräte. Über den PDA kann beispielsweise eine Karte der Umgebung gezeigt werden, mit Hilfe des tragbaren Displays können reale Objekte mit Informationen annotiert werden.

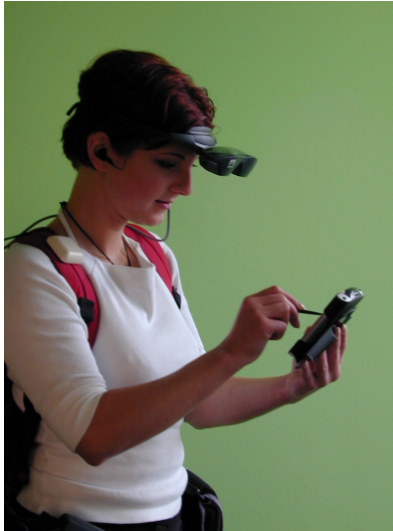


Abbildung 2: Informationen zur Umgebung mit mobilen Geräten

Es ist heute selbstverständlich, dass Techniken entwickelt werden, welche die Wünsche des modernen Menschen nach Mobilität, Information aber auch nach Sicherheit befriedigen. Besonders ortsbezogene Dienste können dazu einen Beitrag leisten. Der Zugriff auf entfernt gespeicherte Informationen von nahezu jedem Ort aus ist schon heute durch kleine, leistungsstarke Geräte möglich. Sie sind für einen großen Nutzerkreis finanziell erschwinglich. Diese Vorteile nutzen LBS und dies begründet ihren Erfolg. Im Zuge der Verknüpfung von weiteren Techniken aus dem Bereich der mobilen Informationsvisualisierung mit diesen Diensten können AR-Systeme den Schritt vollziehen, nicht nur als separate Systeme in spezialisierten Umgebungen eingesetzt zu werden, sondern von einem großen Benutzerkreis angenommen und für viele Probleme als hilfreich angesehen zu werden.

1.2 Problemstellung

AR-Systeme werden weltweit an vielen verschiedenen Forschungszentren entwickelt und finden sich in weit verbreiteten Einsatzgebieten wieder. Vor allem Trackingmethoden für den Innenbereich haben sich in der Medizin als Assistenten für Chirurgen oder bei der Montage von Maschinenteilen an entfernten Einsatzorten bewährt.

Diese Art des hochgenauen AR wurde von verschiedenen Seiten bereits sehr erfolgreich in einer Umgebung integriert, in der Personen und Geräte zu jeder Zeit an jedem Ort notwendige Hilfen angeboten bekommen oder problemlos Zugang zur gerätetechnischen Infrastruktur erhalten. Eine Person bewegt sich beispielsweise durch ein Gebäude und der nächstgelegene Drucker wird ihr angezeigt, sobald die zentrale Einheit erkennt, dass diese Person drucken möchte. Genauso kann ein Telefongespräch automatisch auf das nächstgelegene freie Telefon gelegt werden. Diese Systeme wissen immer, wo sich jemand befindet und kennen die Profile dieser Personen oder der entsprechenden Geräte. Abhängig davon können sie situationsbezogen reagieren. In Gebäuden, die speziell auf die notwendige Infrastruktur angepasst werden, sind solche Systeme bereits im Einsatz [Addl01].

Im Außenbereich gibt es Ansätze, welche die Positionsinformation mit Hilfe von Sensordaten ermitteln. Diese sind allerdings nicht sehr genau [Fein97] [Piek99]. Um diese Daten zu stützen oder zu verbessern, wurden verschiedene Systeme entwickelt. Ihnen ist die Verbindung zu einer Referenz gemeinsam, die die Realität repräsentiert. Die Referenz kann ein digitales Höhenmodell sein [Behr99], ein 3D-Modell [Ribo02], ein 3D-GIS [Chen99] oder ein Referenzphoto [Stri01]. Alle diese Verfahren haben den Nachteil, dass sie entweder nicht zur Bestimmung von Objekten eingesetzt werden können, die sich nahe in Bezug auf den Anwender befinden, oder dass eine manuelle Vorverarbeitung der Daten erfolgen muss, oder dass sie nicht zulassen, dass sich der Benutzer frei bewegt.

Sobald eine Referenz notwendig ist, handelt es sich um Inzellösungen. Aber auch Systeme, die nur Sensordaten auswerten, sind durch die Bedingungen der Sensoren auf bestimmte Bereiche eingeschränkt. Diese verschiedenen Systeme können zwar durch Rahmenwerke für AR-Umgebungen verknüpft werden [Reit01] [Baue01], allerdings bieten diese nur die allgemeine Möglichkeit an, unterschiedliche Trackingmethoden aneinander zu fügen. Es gibt im Moment keine Lösungen, eine Person im ganzen Bereich einer Stadt zu tracken und abhängig von den Anforderungen die beste Trackingmethode automatisch anzusprechen, so dass von einer einfachen Positionsbestimmung bis hin zum hochgenauen Tracking alle Möglichkeiten abgedeckt werden und dabei möglichst ressourcenschonend vorgegangen wird.

1.3 Ziel und Aufbau der Arbeit

Wenn einer Person, die sich in einer Stadt bewegt, mit Hilfe mobiler Geräte Informationen über die Umgebung angeboten werden sollen, so gibt es zurzeit keine Möglichkeit, mit Hilfe verschiedener Sensoren eine Person mit unterschiedlichen Genauigkeitsstufen so zu tracken, dass sie für verschiedene Aufgaben das beste Ergebnis automatisch angeboten bekommt. Das hier aufgezeigte Verfahren ermöglicht diese Zuordnung zwischen einem dem Anwender angebotenen mobilen Informationsdienst und der Trackingtechnik, indem sowohl Anforderungen an die Trackingtechniken abhängig von der Art des Dienstes formuliert werden, als auch die zu diesem Zeitpunkt zur Verfügung stehenden Trackingtechniken parametrisiert werden. Über die Berechnung eines Index der möglichen Zuordnung wird die ideale Kombination bestimmt. Zu diesen Diensten zählen nicht nur Informationen, die nur die Position des Nutzers benötigen. Oft ist zusätzlich die Bestimmung der Blickrichtung notwendig. Wenn eine Rekonstruktion der Umgebung vorliegt, kann diese beispielsweise über die reale Umgebung überblendet werden. Um das Anwendungsfeld Stadt auch für diese Art der Informationsvermittlung bestmöglich zu erschließen, wurde ein Verfahren entwickelt, das es erlaubt, die Position und Blickrichtung eines Anwenders in einem urbanen Umfeld hochgenau zu bestimmen.

Kapitel 2 dieser Arbeit stellt Sensoren und Techniken vor, mit deren Hilfe die Position und Blickrichtung eines mobilen Anwenders bestimmt werden kann. Dabei wird als Anwendungsgebiet der Außenbereich fokussiert. Es werden die Vorteile aber auch die Problembereiche aufgezeigt. Das Kapitel schließt mit der Betrachtung von Möglichkeiten, die Ergebnisse verschiedener Sensoren zu koppeln, wobei Verfahren vorgestellt werden, die diese Ergebnisse so verbinden, dass dem Anwender jeweils die für den Einsatzbereich genaueste Möglichkeit angeboten wird.

Kapitel 3 zeigt den Stand der Technik für AR-Systeme, die im Freien eingesetzt werden. Die einzelnen Lösungen werden beschrieben und es wird aufgezeigt, welche Daten notwendig sind, um ein solches System einsetzen zu können. Die Verfahren, die genauere Ergebnisse liefern, benötigen eine Referenz der Umgebung, die meist in einem videobasierten Ansatz zum Vergleich mit einem Kamerabild führen. Durch die Notwendigkeit von Referenzdaten sind die Verfahren meist Einschränkungen ausgesetzt, auf die ebenfalls näher eingegangen wird.

Kapitel 4 zeigt auf, wie die räumlichen Daten für 3D-Stadtmodelle erfasst werden. Das im Kern dieser Arbeit aufgezeigte videobasierte Trackingverfahren basiert auf den Daten eines Modells der Umgebung. 3D-Stadtmodelle finden weltweit immer mehr Zuspruch, da sie für verschiedenste Anwendungen eine alternativlose Datengrundlage bilden. Die für diese Bereiche erfassten Daten können für dieses Verfahren eingesetzt werden. Am Schluss dieses Kapitels werden Beispiele für 3D-Stadtmodelle aufgezeigt, wobei die verschiedenen Detaillierungsgrade deutlich werden, die die unterschiedliche Datenerfassung reflektieren.

Kapitel 6 bildet den ersten Teil des Schwerpunktes dieser Arbeit. Anhand von Beispielen für Informationsdienste, die ein sich in einer Stadt bewegendem Anwender auf mobilen Endgeräten angeboten bekommen kann, wird eine Klassifizierung abgeleitet. Sie führt zu einer Formulierung von Anforderungen, die Tracking-techniken, die diese Dienste unterstützen, erfüllen sollten. Hierbei werden neben Anwendungen für PCs auch Dienste betrachtet, die auf einem PDA oder Mobiltelefon offeriert werden. Anschließend wird eine Parametrisierung dieser Anforderungen und der Trackingtechniken aufgezeigt. Mit diesen Werten kann eine anforderungsadaptive Zuordnung zwischen einem angebotenen Dienst und den zur Verfügung gestellten Trackingtechniken bestimmt werden, die den Schluss von Kapitel 5 bildet.

Kapitel 6 stellt den zweiten Teil des Schwerpunktes dieser Arbeit dar. Hier wird ein videobasiertes Verfahren aufgezeigt, mit dessen Hilfe die Position und Blickrichtung eines mobilen Anwenders, der sich in einer Stadt im Freien bewegt, bestimmt werden kann. Das Verfahren erlaubt es, dass sich der Anwender frei bewegt und es ist so genau, dass nahe stehende Objekte mit Informationen positionsgenau überlagert werden können. Das Verfahren benötigt als Referenz Daten eines 3D-Modells der Umgebung. Da es sich um den Einsatzbereich Stadt handelt, bieten 3D-Stadtmodelle die entsprechenden Daten. Aus diesem Grund werden beide Begriffe im Folgenden parallel verwendet. Die Daten des 3D-Modells werden den Daten des Kamerabildes zugeordnet. Darüber wird eine dreidimensionale Translation und Rotation zwischen den Näherungsdaten ermittelt. Auf der Grundlage dieser Näherungsdaten wurden die Daten aus dem 3D-Modell extrahiert. Damit ist die Bestimmung des tatsächlichen Standpunktes und der Blickrichtung des Anwenders möglich. Letztere bilden das Endergebnis dieses Verfahrens.

Kapitel 7 zeigt auf, welche Komponenten im Rahmen des Forschungsprojektes GEIST realisiert wurden, die auf den in dieser Arbeit entwickelten Verfahren basieren. Die Verknüpfung von AR mit mobilen, ortsbezogenen Diensten erlaubt es, dass sich der Anwender frei in einer Stadt bewegt und im Freien unterschiedliche Informationen abfragen kann oder automatisch angeboten bekommt. Dazu wurde

eine Komponente implementiert, die die Zuordnung der idealen Trackingtechnik zu einem ausgewählten oder angebotenen Dienst ermöglicht. Gleichzeitig wurden in diesem Projekt Komponenten entwickelt, die auf dem in Kapitel 6 beschriebenen videobasierten Trackingverfahren basieren.

Kapitel 8 fasst die Ergebnisse der Arbeit zusammen und zeigt einen Ausblick auf weitere Forschungsmöglichkeiten in diesem Bereich.

2 Bestimmung von Position und Orientierung einer Person

Die Basis jeden AR-Systems ist die Bestimmung von Position und Blickrichtung des Anwenders. Hierbei werden zwei Gruppen von Einsatzgebieten unterschieden:

- Einsatz im Innenbereich
- Einsatz im Außenbereich

Für den Einsatz im Innenbereich gibt es verschiedene Lösungsansätze, wobei im Allgemeinen eine Infrastruktur in der Umgebung aufgebaut werden muss, in der die Position und die Blickrichtung von Personen bestimmt werden sollen. Das bedeutet, dass Sensoren in der Umgebung befestigt werden, um die Laufzeit von Signalen zu bestimmen oder Marker eingesetzt werden, die genau eingemessen sind und mit Hilfe von Bildverarbeitungsalgorithmen in einem Kamerabild gefunden werden können. Für den Einsatz im Freien, besonders in einer Innenstadt, ist der Aufbau einer solchen Infrastruktur meist nicht möglich. Hier müssen andere Sensoren eingesetzt werden.

Azuma et al. [Azum98] haben verschiedene Trackingtechniken auf ihren Einsatz in AR-Systemen überprüft. Dabei kamen sie zu dem Ergebnis, dass für den Einsatz im Außenbereich ein Sensor allein nicht ausreicht, sondern eine Kombination verschiedener Sensoren entscheidend ist [Azum99b]. Diese Systeme werden als hybride Systeme bezeichnet.

Im folgenden werden zunächst verschiedene Verfahren aufgezeigt, deren Basis aktive Quellen sind. Sie werden zumeist im Innenbereich angewendet. Allerdings ist es denkbar, in abgeschlossenen Gebieten auch diese Techniken einzusetzen. Dazu gehören Infrarotsysteme (Abschnitt 2.1). Sie ermöglichen es, in einer abgeschlossenen, kleinen Umgebung die Position zu bestimmen. Dabei werden von in der Umgebung angebrachten Sendern Signale empfangen. Stellvertretend für netzbasierte Positionierungsmöglichkeiten werden Beispiele für Verfahren vorgestellt, die auf einem WLAN basieren (Abschnitt 2.2.1) und damit ebenfalls, hauptsächlich in Innenräumen, Anwendung finden. Eine weitere Möglichkeit, die hauptsächlich im Außenbereich genutzt wird, sind Techniken, die das Netz eines Mobilfunkanwenders benötigen (Abschnitt 2.2.2). GPS, GLONASS und das zukünftige europäische Pendant GALILEO, die anschließend beschrieben werden (Abschnitte 2.3, 2.4 und 2.5), basieren auf dem Empfang von Satellitendaten. Bei den Betrachtungen wird auf den Einsatz dieser Techniken im Außenbereich fokussiert. Inertiale Sensoren werden daran anschließend erläutert (Abschnitt 2.6). Sie werden im Außenbereich nur für die Detektion der Blickrichtung eingesetzt, da die Translationsergebnisse zu ungenau sind. Die Rotation in horizontaler Richtung ergibt sich ebenfalls mit Hilfe eines Kompasses oder eines Neigungssensors (Abschnitt 2.7). Sehr genaue Ergebnisse sowohl für die Position als auch für die Blickrichtung bekommt man mit Hilfe videobasierter Verfahren (Abschnitt 2.8). Sie werden oft nicht nur für die Stützung von Sensorergebnissen sondern gerade im Innenbereich auch als einzige Trackingtechnik eingesetzt. Die Fusion verschiedener

Trackingtechniken, wenn verschiedene Verfahren nebeneinander meist in verschiedenen Bereichen eingesetzt werden, bilden vor der Zusammenfassung den Abschluss dieses Kapitels (Abschnitt 2.9).

2.1 Infrarot

Eines der bekanntesten Infrarotsysteme ist das Active Badge System [Hart94]. Hier sind Infrarotsensoren in einem Gebäude angebracht, die die periodischen Signale von einer Plakette aufnehmen und an eine zentrale Station zur Auswertung weitergeben. Damit können Personen und Geräte in einem Gebäude lokalisiert werden. Die Technik bietet keine Übertragung von echten Positionierungsinformationen sondern nur die Möglichkeit die Aussage zu treffen, ob die entsprechende Person oder das zu detektierende Gerät den Aktivitätsradius des Sensors betreten oder verlassen haben. Außerdem muss beim Einsatz dieser Technik die Infrastruktur der Umgebung, in der eine Positionierung möglich sein soll, sehr stark verändert werden. Je genauer die Positionierungsdaten ermittelt werden sollen, desto größer ist der Aufwand.

Diese Technik wird häufig in Museen eingesetzt. Besucher bekommen am Eingang eine Audio-Einheit als tragbares Gerät ausgehändigt. Sobald sie sich einem bestimmten Ausstellungsstück nähern, wird die entsprechende, vorher aufgenommene Erklärung abgespielt. Sie wird gestoppt, wenn der Besucher sich wieder entfernt. Dafür wird in der Nähe jeden Ausstellungsstücks, z.B. an der Decke darüber, ein kleiner Infrarotsender angebracht, der eine eindeutige Identifizierung ausstrahlt. Das tragbare Gerät empfängt diese. Damit ist die ungefähre Position des Besuchers bestimmt und die entsprechende Erklärung kann abgespielt werden. Abbildung 3 veranschaulicht das Verfahren eines solchen Audiosystems für ein Museum.

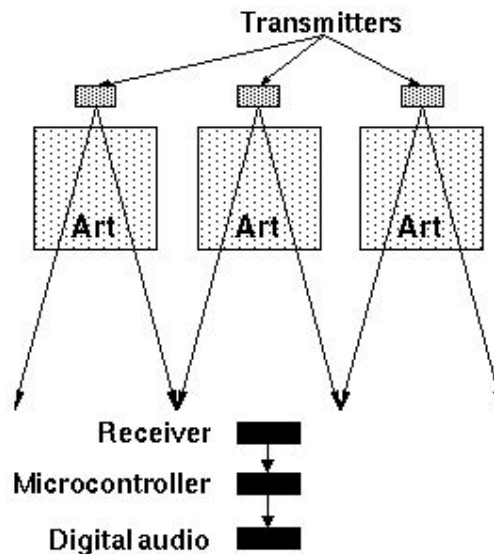


Abbildung 3: Audiosystem für ein Museum [Bede95]

Mehrere Besucher können gleichzeitig die selbe Stelle in der Audiosequenz abspielen lassen, wenn sie zur gleichen Zeit den Senderadius des selben Ausstellungsstückes betreten.

2.2 Netzbasierte Positionierungsmöglichkeiten

Es gibt verschiedene Systeme, die mobile Geräte in einem aus Zellen bestehenden Netz orten. Neben den gerade für die Positionierung oft eingesetzten GSM-Netzen (Global System for Mobile Communications) ist die WLAN-Technik (Wireless Local Area Network) für eine ähnliche Form der Positionsbestimmung anwendbar. Die gebräuchlichste Technik für WLAN-Netze ist der IEEE 802.11 Standard. Eine Zelle für ein WLAN ist etwa 50 m groß, in Gebäuden, abhängig von der verwendeten Bausubstanz, meist kleiner. Durch Techniken wie die Einbeziehung der Antwortzeit eines Signals oder durch Triangulierung zu mehreren Basisstationen kann die Genauigkeit gesteigert werden. Diese Techniken finden auch bei der Positionierung in Mobilfunknetzen Anwendung, in denen die Zellen eine Größe von mehreren Kilometern erreichen können [Kalk02]. Verfahren, die für die Positionsbestimmung ein WLAN oder das GSM benötigen, sind die zur Zeit am weitesten verbreiteten Techniken. Auf diese wird im folgenden (Abschnitt 2.2.1 und 2.2.2) näher eingegangen.

2.2.1 Positionsbestimmung basierend auf der WLAN-Technik

Positionsbestimmungen auf Basis der WLAN-Technik werden zumeist in Innenräumen eingesetzt. Im folgenden werden zwei Beispiele für Positionierungsverfahren vorgestellt, die auf dieser Technik basieren.

Das RADAR-System von Microsoft Research ermittelt die zweidimensionale Position eines mobilen Gerätes über eine Triangulierung [Bahl00]. Dabei wird sowohl eine empirische als auch eine Methode angewendet, die die Signalverbreitung berücksichtigt. Das zuerst genannte Verfahren liefert eine Genauigkeit von 4.3 Metern (die Berechnung erfolgt über den Median), allerdings ist es einfacher zu handhaben als das zweite, das bis zu drei Meter genau ist.

Ein weiteres Beispiel für die Positionierung basierend auf einem WLAN ist ein hybrides System, das die Informationen mit Hilfe eines Datenbank-Management-Systems administriert [Peyr00]. Das System ist hybrid, da es sowohl aus einem Leitungsnetz als auch aus einem drahtlosen Netz besteht. Es werden die folgenden drei Einheiten unterschieden:

- Feste Stationen, die mit dem normalen Netz verbunden sind
- Mobile Geräte, die über Radiowellen kommunizieren
- Zugangspunkte (Access Points), die sowohl an das normale als auch an das drahtlose Netz angeschlossen sind. Diese entsprechen nicht den handelsüblichen Zugangspunkten, da dann keine spezielle Software integriert werden könnte, die die Datenbank bedient.

Die mobilen Geräte senden regelmäßig einen Identifizierungscode an einen der Zugangspunkte. Nach einer festgelegten Zeit muss das mobile Gerät eine Antwort des Zugangspunktes erhalten. Ist dies nicht der Fall, nimmt das System an, dass sich das mobile Gerät außerhalb des Abdeckungsradius des Zugangspunktes befindet. Wenn eine Nachricht eines mobilen Gerätes einen der Zugangspunkte erreicht, wird dieser Vorgang in der Datenbank registriert. Dies trifft ebenfalls zu, wenn ein Paar, bestehend aus mobilem Gerät und Zugangspunkt, keine Verbindung

innerhalb des festgelegten Zeitrahmens mehr aufweist. Dann wird auch die Historie des mobilen Gerätes in der Datenbank an die aktuelle Konstellation angeglichen. Der Abdeckungsbereich eines mobilen Gerätes innerhalb des drahtlosen Netzes der Zugangspunkte wird somit für die Positionierung genutzt.

Diese Verfahren sind ähnlich den im nächsten Abschnitt (2.2.2) erläuterten Methoden, die die Position mit Hilfe eines Mobilfunknetzes berechnen. Im Gegensatz zu diesen ist der Einsatzbereich von WLAN-Netzen meist auf Innenräume beschränkt. Für eine Person, die sich in einer Stadt bewegt, sind diese Verfahren deshalb nur eingeschränkt anwendbar, auch wenn es bereits abgegrenzte Außenbereiche gibt, in denen diese Technik angeboten wird. Stand der derzeitigen Forschung sind allerdings Verfahren für WLAN-Netze, die Personen und Geräte in Innenräumen tracken. Da es aber bereits Überlegungen gibt, in Innenstädten flächendeckend ein WLAN anzubieten, wird diese Positionierungstechnik in der Zukunft sicher an Bedeutung für den Einsatz im Außenbereich gewinnen.

2.2.2 Positionsbestimmung mit Hilfe des GSM-Netzes

Das GSM-Netz umfasst alle Spezifikationen und Schnittstellen, die für ein funktionierendes Mobilfunknetz nötig sind. Es wird von Mobiltelefonen oder PDAs mit einer entsprechenden Karte genutzt, um Telefongespräche zu führen oder Zugriff auf das Internet zu bekommen. Die Infrastruktur dieses Netzes kann für eine Positionierung der entsprechenden Geräte eingesetzt werden. Dabei werden grundsätzlich zwei Arten der Positionsbestimmung unterschieden:

- Netzwerk-basierte
- Geräte-basierte.

Die Unterscheidung besteht darin, ob die Berechnung der Position oder eine Näherung derselben direkt das mobile Gerät oder zentral das Netzwerk des Betreibers vornimmt. Der Vorteil der Netzwerk-basierten Verfahren ist es, dass die einzelnen Geräte nicht mit spezieller Technik ausgestattet werden müssen, da alle Berechnungen von zentralen Rechnern geleistet werden. Das Endergebnis, die Position, wird dann an das Gerät übermittelt. Zusätzlich können besondere Dienste angeboten werden, da die zentralen Rechner direkten Zugriff auf große Datenmengen zur Verfügung stellen können. Das Gerät muss dann allerdings in Betrieb sein, um Positionierungen durchführen zu können. Außerdem ist der Datenschutz für das Wissen des Aufenthaltsortes des Geräte-Anwenders nicht mehr gewährleistet. Das Angebot eines solchen Dienstes bedeutet meist, dass die zentralen Rechner mit zusätzlicher Funktionalität ausgestattet werden müssen. Dies verursacht oft hohe Kosten.

Im folgenden werden zu den Netzwerk-basierten Verfahren auch Methoden gezählt, bei denen die eigentliche Messung am Gerät durchgeführt wird, die Auswertung allerdings an eine zentrale Stelle geschickt und dort ausgewertet wird. Dazu gehören Verfahren, bei denen die Messberichte ausgewertet werden, die konstant vom Gerät zum Netz gesendet werden, wenn das Gerät in Betrieb ist.

Eine Geräte-basierte Bestimmung der Position verschiebt die zusätzliche Funktionalität auf die Seite des einzelnen Gerätes. Damit sind die Datenschutzprobleme vermindert. Allerdings ist zumindest die Nummer der Zelle, in der sich das Gerät im aktiven Zustand, z.B. bei einem Anruf oder beim Schicken einer SMS (Short Message Service), befindet, dem Netzbetreiber bekannt. Das Abfragen einer genaueren Position von außen als über die Zellennummer ist bei dieser Art der Bestimmung erschwert. Der Nutzer kann selbst entscheiden, wann er lokalisiert werden will. Die Verfahren, die hierzu zählen, sind präziser als die Netzwerk-basierten Techniken.

Im folgenden sind die sechs bekanntesten Lokalisierungsmethoden aufgeführt.

Zellen-Nummer

Das GSM-Netz besteht aus Basisstationen, die gemeinsam das gesamte Gebiet abdecken, in dem ein Nutzer einen Anruf tätigen oder empfangen kann. Jede dieser Basisstationen deckt einen bestimmten Bereich ab. In ländlichen Gebieten sind diese Bereiche größer, in Städten ist die Dichte an Basisstationen höher. Diese Gebiete, die einer bestimmten Basisstation zugeordnet sind, werden als Zelle bezeichnet. Die Nummer der Zelle kann zur Positionierung genutzt werden. Die Genauigkeit schwankt mit der Größe der Zelle. Sie kann zwischen etwa 150 Metern und 35 Kilometern liegen. Damit erreicht man kein sehr genaues Positionierungsergebnis. In Städten reicht diese Form der Lokalisierung aber oft aus, um einfache Dienste anbieten zu können. Der Vorteil dieser Methode ist es, dass das Ergebnis sehr schnell vorhanden ist (etwa drei Sekunden) und keine zusätzliche Funktionalität notwendig ist. Das Verfahren kann als Netzwerk-basierte oder als Geräte-basierte Lösung realisiert werden.

Signalstärke

Der Kontrollkanal steht zur ständigen Datenübertragung zur Verfügung. Über die Messung der Signalstärke dieses Kanals von mindestens vier verschiedenen Basisstationen aus kann der Abstand des Gerätes zu diesen ermittelt werden. Da sich die Signale vor allem in urbanen Gebieten nicht frei ausbreiten können, muss ein der Umgebung angepasstes Verbreitungsmodell für die Beziehung zwischen empfangenem Signal und der Distanz zwischen Gerät und Basisstation eingesetzt werden. Die Auswertung der Messungen kann entweder direkt am Gerät oder an einer zentralen Stelle vorgenommen werden. Im aktiven Zustand werden die Berichte kontinuierlich vom Gerät zum Netz übertragen. Das Gerät kann aber so verändert werden, dass diese Berichte auch gesendet werden, wenn das Gerät ausgeschaltet ist. Dieses Verfahren erlaubt es, die Genauigkeit zu erhöhen, die mit Hilfe des Verfahrens über die Zellennummer erhalten wurde. [Lait01]

Timing Advance (TA)

Dieses Verfahren nutzt einen Wert des GSM-Netzes, der die Zeitdifferenz angibt, die notwendig ist, damit das Gerät das Datenpaket so abschickt, dass es innerhalb des Empfangsfensters eintrifft. Dieser Wert liegt zwischen 0 und 63. Eine Einheit entspricht einer Entfernung von etwa 550 m zur Basisstation. Das bedeutet, dass eine Zelle einen maximalen Radius von 35 Kilometern haben kann. Bestimmte Geräte bieten meist eine Kombination aus dieser Technik und dem Verfahren über die Zellen-Nummer an.

Enhanced Observed Time Difference (EOTD)

Dieses Verfahren, das auch unter dem Begriff Down-Link bekannt ist, ist eine Kombination aus einem Netzwerk- und einem Geräte-basierten bzw. Geräte-assistierten Verfahren. Es beruht darauf, über die Zeitdifferenz zu drei Basisstationen die Position des Gerätes zu bestimmen. Da dies voraussetzt, dass die Zeit sehr genau bestimmt wird, müssen die Signale zusätzlich von einer sogenannten Location Measurement Unit (LMU) empfangen werden. Dies ist ein stationärer Empfänger, dessen Position bekannt ist und der mit hochpräzisen Uhren ausgestattet ist. Für diese hochgenaue Synchronisation wird häufig ein GPS-Empfänger eingesetzt. Dieses Verfahren erfordert meist nicht nur zusätzliche Funktionalität in Form von Software. Auch die Hardware muss dieses Verfahren unterstützen. Die Berechnung dauert etwas länger als beim oben genannten Verfahren über die Zellen-Nummer. Allerdings sind Genauigkeiten zwischen 50 und 125 Metern möglich.

Time of Arrival (ToA)

Genauso wie beim EOTD-Verfahren muss auch bei diesem Verfahren das Signal von mindestens drei verschiedenen Basisstationen empfangen werden. Allerdings übernimmt die Berechnung der Position das Netz und nicht das Gerät. Es wird auch Up-Link genannt. Der Nachteil dieses Verfahrens ist, dass die Messung nur erfolgen kann, wenn das Gerät in Betrieb ist. Außerdem können Kapazitätsprobleme auftreten. Das Verfahren ist nur möglich, wenn das Netz synchron ist. Im Falle des GSM-Netzes, wird versucht, dies damit zu erreichen, dass jede Basisstation mit einer hochpräzisen Uhr ausgestattet ist. Die Berechnung der Position dauert etwa zehn Sekunden. Die Uhrendifferenzen zwischen den Basisstationen werden nicht, im Gegensatz zum EOTD-Verfahren, über einen Referenzempfänger bestimmt, der an einer bekannten Position steht und kontinuierlich die beobachteten Zeitdifferenzen misst. Damit ist diese Lösung aufwendiger und weniger ökonomisch als die Lösung über einen Referenzempfänger. [Lait01]

Bei beiden Verfahren beeinflussen die Bandbreite, Multipath-Effekte, Sichtbarkeit zwischen den Basisstationen und dem Gerät die Genauigkeit und Zuverlässigkeit der Ergebnisse. Multipath spielt auch bei der Positionierung mit GPS eine Rolle und bedeutet, dass das Signal nicht den direkten Weg zurückgelegt, sondern einen längeren, verfälschten Weg, da es beispielsweise an Gebäuden „abprallt“. Dieser wird vom Empfänger registriert. Außerdem beeinflussen die Anzahl der Messungen und die Geometrie der Basisstationen in Bezug auf das Gerät die Zuverlässigkeit und damit die Genauigkeit des Positionierungsergebnisses.

Angle of Arrival (AoA)

Über den Winkel, unter dem ein Signal an mehreren Basisstationen ankommt, kann die Position des Gerätes bestimmt werden. Für die Bestimmung des Winkels ist es allerdings notwendig, dass die Basisstationen mit besonderen Antennen ausgestattet werden. Aus diesem Grund findet diese Form der Positionierung kaum Berücksichtigung in den existierenden mobilen Netzen. Die Genauigkeit dieses Verfahrens verringert sich, wenn keine direkte Sicht zwischen der Basisstation und dem Gerät besteht oder wenn das Signal durch andere Einflüsse gestört ist. Die Genauigkeit hängt von der Anzahl der möglichen Messungen, der Geometrie der Basisstationen um das Gerät und von Multipath-Einflüssen ab. Die Ungenauigkeit der Positionierung vergrößert sich, wenn der Abstand zwischen Gerät und Basisstation zunimmt. Das Verfahren funktioniert bei freiem Sichtkontakt zu zwei Basisstationen. Aus diesem Grund ist es besonders in ländlichen Gebieten anwendbar. In Städten liefert es meist zu große Positionsfehler.

Zusammenfassung

Eine Positionierung über die Zellen-Nummer bietet sehr ungenaue Ergebnisse. Wenn TA angewendet wird, ist eine Auflösung von bis zu 554 Metern möglich. In Verbindung mit der Messung der Signalstärke kann eine Verbesserung erreicht werden. Eine Erweiterung dieses Verfahrens beruht auf der Berechnung des Medians mehrerer Messungen über einen längeren Zeitraum.

Die Berechnung der Position über den Winkel zu mindestens zwei Basisstationen (AoA) wird kaum in derzeitigen Systemen eingesetzt, da die Basisstationen über besondere Antennen verfügen müssen. Allerdings wird es von manchen Firmen als E911 (Enhanced 911) Lösung angeboten. Hierbei handelt es sich um in den USA von der FCC (Federal Communications Commission) aufgestellte Bedingungen und Zeitpläne, die die Verfügbarkeit von Lokalisierungsinformationen für Nutzer von Mobiltelefonen für Notrufdienste verbessern sollen. Der Einsatz von AoA ist nur in ländlichen Regionen ratsam, da freier Sichtkontakt zu den berücksichtigten Basisstationen vorhanden sein muss.

Die Verfahren, die die Zeitdifferenz zu mehreren Basisstationen berücksichtigen, sind am weitesten bei der Positionierung von Mobiltelefonen verbreitet. Dabei hat EOTD den Vorteil, dass keine Synchronisierung über hochgenaue Uhren an den Basisstationen erfolgen muss, sondern dass über Location Measurement Units ein Abgleich erfolgt, der höhere Genauigkeiten zulässt.

2.3 GPS

2.3.1 Funktionsweise

GPS ist eine Technik, die es erlaubt mit Hilfe eines Gerätes Satellitensignale zu empfangen. Das System wird vom amerikanischen Verteidigungsministerium unterhalten. Über diese Signale kann die Entfernung zu den Satelliten und damit die Position des Empfängers bestimmt werden. Das GPS besteht aus drei Segmenten. Zum Weltraumsegment zählen die Satelliten, die um die Erde kreisen und ständig Signale aussenden, die vom Nutzersegment, den Satellitenempfängern, aufgenommen und verarbeitet werden. Die Frequenz liegt im Gigahertzbereich, so dass keine Richtantennen notwendig sind. Mit Hilfe des Bodensegmentes werden die Satelliten überwacht, d.h. Laufbahnen und Uhren korrigiert. Abbildung 4 gibt eine Übersicht über die drei Segmente.

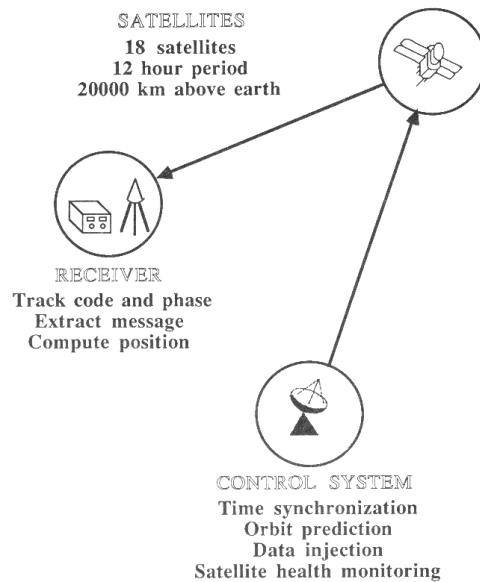


Abbildung 4: Segmente des GPS [IVVM99]

Die Satelliten senden auf zwei verschiedenen Frequenzen: L_1 mit der Frequenz von 1227,60 MHz und L_2 mit einer Frequenz von 1575,42 MHz. Die Grundfrequenz beträgt 10,23 MHz. Die beiden Trägerfrequenzen sind durch die Multiplikation mit 154 bzw. mit 120 davon abgeleitet worden. Grundsätzlich werden die beiden Frequenzen des GPS-Signals L_1 und L_2 mit drei Binärcodes ausgestattet:

- C/A-Code (Coarse Acquisition) moduliert die L_1 -Trägerwelle. Er ist für jeden Satelliten verschieden und die Grundlage für zivile Ortsbestimmungen.
- P-Code (Precise) moduliert sowohl die L_1 - als auch die L_2 -Welle. Mit ihm können sehr genaue Positionierungen vorgenommen werden, allerdings steht er nur bestimmten, von der Regierung der USA autorisierten Nutzern zur Verfügung. Diese müssen bestimmte kryptographische Schlüssel sowie speziell ausgestattete Empfänger besitzen.
- Die Navigations-Nachricht moduliert das L_1 -C/A Code-Signal. Die Nachricht beschreibt die Satellitenbahn, Uhrenkorrekturen und weitere Parameter.

Für die Berechnung der Position spielen in der Praxis zwei Verfahren eine Rolle. Wenn die auf die Trägerfrequenzen aufmodulierten Codes ausgewertet werden, spricht man von Code-Phasen-Messung, wenn die Trägerwelle zur Positionsbestimmung herangezogen wird, von Träger-Phasen-Messung. Das zuletzt genannte Verfahren wird hauptsächlich im Vermessungsbereich eingesetzt und spielt für den Einsatz in einem mobilen Informationssystem keine Rolle. Unter der Annahme, dass die Uhren im Gerät und im Satellit exakt gleich sind, kann die Entfernung zwischen Gerät und Satellit bestimmt werden. Da diese Annahme nicht der Realität entspricht, nennt man diese Distanzen Pseudo-Entfernungen. Um die tatsächliche Entfernung berechnen zu können, muss die Empfängeruhr, bei der es sich nicht wie in den Satelliten um eine sehr präzise Atomuhr handelt, korrigiert werden. Wenn das Signal

von einem vierten Satelliten in die Messung einfließt, kann dieser Uhrenfehler als vierte Unbekannte neben den drei Positionskoordinaten eingeführt und berechnet werden. Aus diesem Grund ist es notwendig, dass die Signale von mindestens vier Satelliten gleichzeitig vom Gerät empfangen werden.

Grundsätzlich können bei der Qualitätsbewertung von GPS-Messungen Eigenschaften des Gerätes berücksichtigt werden, wie das Rauschen, das mit Hilfe des Parameters SNR (Signal to Noise Ratio - Signalrauschverhältnis) ausgedrückt werden kann, oder der Uhrenfehler. Abweichungen von einem Frequenznormal können zusätzlich zu geräteabhängigen Einflüssen durch klimatische Einflüsse, Alterung der Wellen oder Schwankungen in der statistischen Verteilung hervorgerufen werden [IVVM99].

Ein wichtiger Einfluss auf die Qualität der GPS-Ergebnisse ist die Satellitenkonstellation zum Zeitpunkt der Messung. Eine ideale Satellitenkonstellation besteht beispielsweise, wenn ein Satellit genau im Zenit des Beobachtungspunktes und die anderen drei Satelliten in einem Winkel von 120° nahe zum Horizont stehen. Dies wird über den Parameter GDOP (Geometric Dilution of Precision) beschrieben. Dieser kann als Gewicht in die Berechnung des Positionsfehlers eingesetzt werden. Damit besteht eine lineare Beziehung zwischen der Messgenauigkeit und der Satellitenkonfiguration. Zusätzlich zum Parameter GDOP können über die folgenden Werte weitere Aussagen zu einer GPS-Messung abgegeben werden:

- PDOP – Position Dilution of Precision (3D)
- HDOP – Horizontal Dilution of Precision
- VDOP – Vertical Dilution of Precision
- TDOP – Time Dilution of Precision

Sie stehen durch folgende Formeln miteinander in Verbindung:

$$GDOP = \sqrt{(PDOP)^2 + (TDOP)^2}$$

$$PDOP = \sqrt{(HDOP)^2 + (VDOP)^2}$$

2.3.2 Differential GPS

Eine Möglichkeit zur Genauigkeitssteigerung ist es, Korrektursignale zu verarbeiten. Dies bedeutet, dass Relativmessungen mit zwei gleichzeitig laufenden Empfängern durchgeführt werden. Das Verfahren nennt sich Differential GPS (DGPS). Referenzstationen, deren Standorte bekannt sind, berechnen laufend Entfernungen zu den Satelliten und damit Verfälschungen des Satellitensignals. Von den Verfälschungen abgeleitete Korrekturwerte werden im Falle von Echtzeit-DGPS direkt an den Empfänger übermittelt. Dies kann durch die folgenden Möglichkeiten realisiert werden:

- GSM-Mobiltelefon oder Betriebsfunk
- UKW
- Langwelle.

Abbildung 5 zeigt, wie Korrekturwerte an einer Referenzstation gemessen und an das Nutzersegment weitergegeben werden.

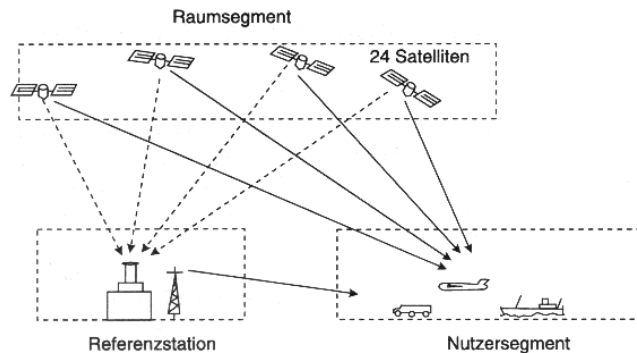


Abbildung 5: Korrekturwerte über eine Referenzstation [IVVM99]

Je näher die Korrekturempfänger der Referenzstation am Einsatzgebiet stehen, desto ähnlicher sind die Empfangsbedingungen der Signale und desto genauer sind die Ergebnisse. Die Verbreitung von Korrektursignalen ergibt bis zu einer Distanz von etwa 1000 Kilometern einen Sinn [Kell94]. Mit diesem Verfahren sind Genauigkeiten im Meter-, Dezimeter-, Zentimeter und sogar Millimeterbereich möglich.

Die deutschen Vermessungsverwaltungen unterhalten Referenzstationen im Rahmen des deutschen Satellitenpositionierungsdienstes SAPOS. Das bedeutet, dass ein Korrektursignal für ganz Deutschland angeboten wird. Innerhalb von SAPOS werden vier verschiedene Dienste angeboten:

- Echtzeit Positionierungs-Service
- Hochpräziser Echtzeit Positionierungs-Service
- Geodätischer präziser Positionierungs-Service
- Geodätischer hochpräziser Positionierungs-Service

Für den Einsatz in einem mobilen Informationssystem ist der zuerst genannte Dienst sinnvoll. Mit diesem einfachen Echtzeitdienst ist es möglich, die ermittelten Positionen auf eine Genauigkeit von ein bis drei Metern zu korrigieren. Das Signal wird durch Zusammenarbeit mit den Rundfunkanstalten der ARD über UKW, mit der Deutschen Telekom über Langwelle und über Sender der Landesvermessung gesendet. Zusätzlich zum GPS-Empfänger ist deshalb ein UKW/LW-Empfänger oder ein 2m-Band-Empfänger mit Decoder notwendig. Auch der zweite Dienst bietet Echtzeitergebnisse an.

Um die GPS-Position in Echtzeit mit einer Genauigkeit von ein bis fünf Zentimetern zu erhalten, ist es notwendig, die Trägerphasenkorrekturdaten zu verarbeiten. Diese Daten werden über die Sender der Landesvermessung im Sekundentakt versendet oder sind über Telefon abrufbar. Es ist geplant, diesen Dienst durch die Vernetzung von diversen Referenzstationen zu erweitern. Dann ist die Erfassung von ortsabhängigen Fehlereinflüssen möglich.

Das Prinzip der Verbesserung der GPS-Werte über Korrekturwerte, die von Referenzstationen in den USA ermittelt werden, wird beim WAAS (Wide Area Augmentation System) eingesetzt [LoAk01]. Die Referenzstationen übermitteln ihre aktuellen Daten, die sie bei der Überwachung der für sie sichtbaren Satelliten ermittelt haben, an eine von zwei Master-Stationen. Die Korrektursignale werden dann über geostationäre Satelliten auf den Empfänger übertragen. WAAS liefert nicht nur Korrekturwerte sondern auch zusätzliche Entfernungen zur Bestimmung der Position. Die Werte werden über das L_1 -Frequenzband übertragen, wenn der Empfänger umgerüstet worden ist. Damit benötigen sie keinen zusätzlichen Empfänger, wie es für DGPS notwendig ist. Das System wurde zunächst zur Verbesserung der Positionierung von GPS für Anwendungen im Flugverkehr entwickelt. Durch die Abdeckung der USA mit nur zwei geostationären Satelliten ist die Reichweite dieses Systems stark eingeschränkt. Aus diesem Grund wurde an der Stanford University der Einsatz von Loran zur Übertragung der WAAS-Information getestet [LoAk01]. Hierbei handelt es sich um ein terrestrisches Navigationssystem, das in der Seenavigation der USA eingesetzt wird. Es hat eine sehr hohe Reichweite, d.h. von bis zu 800 Kilometern. Ein dem WAAS vergleichbarer Dienst wird in Japan in Form von MSAS (Multi-Functional Satellite Augmentation System) und in Europa in Form von EGNOS (Euro Geostationary Navigation Overlay Service) angeboten. EGNOS gilt als erste Stufe auf dem Weg zum gemeinsamen europäischen Satelliten-Navigationssystem GALILEO. Sowohl WAAS als auch EGNOS laufen zurzeit als Test. Wie von Abousalem et al. [Abou00] mit einem Ashtec-Empfänger gezeigt werden konnte, sind bei EGNOS Genauigkeiten in der Lage von besser als vier Metern und in der Höhe von besser als sechs Metern möglich.

Eine weitere Entwicklung im Bereich der GPS-Technologie zielt darauf ab, das Signal auch in Gebieten zu empfangen, die momentan noch Schwierigkeiten bereiten. So genannte Pseudolites senden GPS-ähnliche Signale aus und können in Gegenden eingesetzt werden, in denen die GPS-Abdeckung nicht ausreicht [Monda03]. Sie können nahe der Erdoberfläche angebracht sein oder sich in der Luft befinden [Tsuji01].

2.3.3 Zusammenfassung

Mit einem GPS-Gerät wird nur die Position nicht die Orientierung des Benutzers registriert. Grundsätzlich bietet GPS eine Genauigkeit, die für einfache Trackingaufgaben ausreicht. Bei Nutzung einer Referenzstation oder eines Positionierungsdienstes können zusammen mit einem einfachen Orientierungstracker nicht nur Informationen auf einem PDA über die unmittelbare Umgebung sondern auch blickrichtungsabhängige Informationen in ein tragbares Display eingeblendet werden. Die geodätischen Positionierungs-Dienste spielen für das Tracking einer Person im Außenbereich keine Rolle. Diese Auswertungen sind nicht zur Echtzeit möglich, sondern werden nachträglich durchgeführt.

2.4 GLONASS

Das russische Pendant zu GPS nennt sich GLONASS (Global Navigation Satellite System). Auch die GLONASS-Satelliten senden auf zwei Frequenzen L_1 und L_2 . Ebenso gibt es hier einen C/A- und einen P-Code, wobei letzterer genauer ist und für militärische Anwendungen angeboten wird. GLONASS hat den Vorteil gegenüber GPS, dass das Signal durch SA nicht beeinträchtigt wird. Allerdings wurden durch den Zusammenbruch des Ostblocks die finanziellen Möglichkeiten für viele Bereiche eingeschränkt. Daraus resultierte, dass sich GLONASS zeitweise in einem sehr schlechten Zustand befand und der reibungslose Betrieb nicht sichergestellt werden konnte. Die Satelliten haben außerdem eine verhältnismäßig kurze durchschnittliche Lebensdauer von drei Jahren. Die geplante Konstellation der GLONASS Satelliten war nur für einen relativ kurzen Zeitraum vorhanden [Baue03]. Aus diesem Grund ist eine Positionierung nur mit diesem System kaum möglich. Eine Kombination von GPS und GLONASS kann zu einer Verbesserung der Positionierung führen. Der Markt für kombinierte GLONASS/GPS-Empfänger für geodätische Anwendungen ist allerdings kaum entwickelt [Baue03].

2.5 GALILEO

Das geplante europäische Satellitennavigationssystem GALILEO ist der Beitrag Europas zur globalen Navigationsinfrastruktur GNSS (Global Navigation Satellite System). Es soll die Möglichkeiten von Positionierungsmöglichkeiten über Satellitendaten besonders für europäische Nutzer im zivilen Bereich erweitern. Im Moment ist GPS neben dem russischen GLONASS bisher das einzige Satellitennavigationssystem, das im Einsatz ist. Durch den oben erläuterten schlechten Zustand von GLONASS wird weltweit hauptsächlich GPS genutzt.

Bei beiden derzeit existierenden Systemen steht der militärische Aspekt im Vordergrund. Damit sind viele Bereiche, die auf Positionierungsdaten angewiesen sind, indirekt vor allem von den Entscheidungen des amerikanischen Verteidigungsministeriums abhängig. Das Signal kann nicht nur jederzeit blockiert oder gestört werden, wenn eines der Länder seine Interessen verletzt sieht, auch dauert die Informationsvermittlung bei Problemen oft sehr lang. Dadurch ist es für viele Anwendungen nicht zuverlässig genug. Aus diesem Grund stellt GALILEO ein von Drittländern unabhängiges System dar, das die hochgenaue Bestimmung von Standort und Zeit auf einfache Weise sowie die Kombination dieser Informationen ermöglichen soll. Es ist geplant, die Reichweite für die Satellitennavigation zu erhöhen, indem lokale Sender auch an bisher für die Signale nicht zugänglichen Stellen wie Tunneln oder Parkhäusern die Signale verbreiten. Innerhalb des GALILEO-Systems sollen 30 Satelliten sowie Bodenstationen eingesetzt werden. Abbildung 6 zeigt die geplante Architektur von GALILEO.

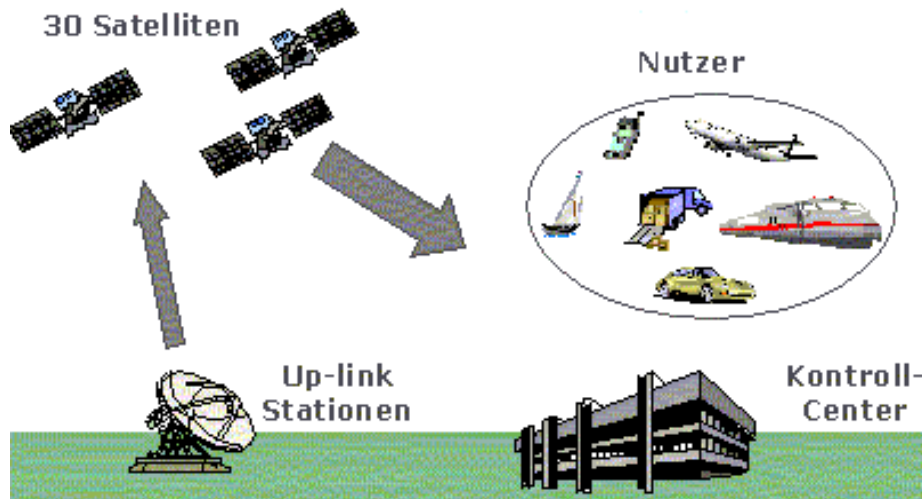


Abbildung 6: GALILEO-Architektur [Gali03]

GALILEO versteht sich nicht nur als Ergänzung zum amerikanischen GPS, indem sich die beiden Systeme auf alle Fälle ergänzen sollen, sondern durchaus auch als Konkurrenz. Damit erhoffen sich die beiden Initiatoren, die Europäische Union und die Europäische Weltraumorganisation ESA ab dem Jahr 2008 die Versorgung von Ortungsinformationen für verschiedene Bereiche. Diese Informationen sollen dann weltweit mit hoher Genauigkeit, Zuverlässigkeit und Sicherheit zur Verfügung stehen.

2.6 Inertiale Sensoren

Das Messprinzip inertialer Sensoren beruht auf dem Trägheitsgesetz der Newtonschen Axiome. Dieses besagt, dass ein Körper im Zustand der Ruhe oder der gleichförmigen geradlinigen Bewegung verharrt, sofern er nicht durch einwirkende Kräfte gezwungen wird, seinen Zustand zu verändern. Mit dem dynamischen Grundgesetz wird ausgedrückt, dass die Beschleunigung der einwirkenden Kraft proportional ist und in Richtung der Kraft erfolgt. Über die Trägheit können demnach die Änderung von Position und Orientierung berechnet werden. Dazu muss die lineare oder rotatorische Beschleunigung bzw. Geschwindigkeit gemessen werden.

Sogenannte inertielle Navigationssysteme (INS) werden bereits seit den Fünfziger Jahren in Schiffen, U-Booten und Flugzeugen eingesetzt. Diese sind sehr schwer. Erst mit dem Einzug von MEMS (Microelectronic Mechanical Systems) in den Neunziger Jahren sind sie auch für mobile Anwendungen einsetzbar. Die meisten der heutzutage üblichen Systeme bestehen aus der Kombination von drei Gyroskopen zur Bestimmung der Winkeländerungen und aus drei linearen Beschleunigungsmessern zur Ermittlung der Position. Gyroskope, auch Messkreisel genannt, werden herangezogen, um Winkeländerungen zu bestimmen. Das Ergebnis erhält man durch Integration der Messwerte. Beim Zusammenspiel beider Techniken wird zunächst über die Ergebnisse der Gyroskope integriert. Damit erhält man nicht nur die Winkeländerungen sondern kann die Ergebnisse der Beschleunigungsmesser, die Beschleunigungsvektoren, in das entsprechende Bezugssystem drehen. Nach Berücksichtigung der Schwerkraft des vertikal angebrachten Beschleunigungsmessers kann über zweifache Integration die Position berechnet werden. Sowohl Gyroskope als auch Beschleunigungsmesser passen heutzutage, wie oben erwähnt, auf Chips [Welc02a].

Das Hauptproblem dieser Sensoren ist die Drift, ein rampenförmiger Fehler, der durch Integration des Nullpunktfehlers entsteht. Dieser Fehler wächst mit der Zeit an. Bei einer erneuten Integration wird aus der Rampe eine Parabelform, die einer quadratischen Funktion entspricht. Sie ist damit der dominierende Fehler. Genauere Sensoren, bei denen dieser Fehler nicht so schnell anwächst finden hauptsächlich Anwendung in Applikationen, die nur eine relativ kurze Messzeit erlauben, d.h. beispielsweise im Minutenbereich. Damit wird der Zeitpunkt, an dem der Fehler auftritt, verschoben. Die für einen Fußgänger angebotenen Geräte sind zwar klein aber eher ungenau.

Bei Gyroskopen werden mechanische und optische Kreisel unterschieden. Für den Einsatz in AR-Systemen finden hauptsächlich Vibrationskreisel Anwendung, die zu den mechanischen Kreiseln zählen. Diese haben den Vorteil, dass die bewegten Teile keinem mechanischen Verschleiß unterliegen. Das eigentliche Sensorelement dieser Kreisel ist eine mechanisch vibrierende Struktur, die aufgrund von Corioliskräften sensitiv auf Drehbewegungen reagiert [Grif00]. Durch den Einsatz der Mikrotechnologie sind sie sehr klein und verbrauchen wenig Strom. Die Werte entsprechen relativen Ergebnissen Sie müssen regelmäßig gestützt werden. Ein Beispiel für diese Form der Korrektur ist in [You99] beschrieben. Es stellt eine Realisierung für ein AR-System im Außenbereich dar.

Zusammenfassend kann gesagt werden, dass für die Orientierungsbestimmung in AR-Systemen inertielle Sensoren eingesetzt werden können. Bei der Positionsberechnung treten allerdings Genauigkeitsprobleme auf. Auch in Zukunft ist nicht abzusehen, dass inertielle Sensoren zur Bestimmung der Position realisiert werden können.

2.7 Magnetkompass und Neigungssensoren

Mit Hilfe eines Magnetkompasses kann die Ausrichtung zu magnetisch Nord bestimmt werden, die für viele Aufgaben anstelle zu geographisch Nord ausreicht. Für manche Aufgaben muss allerdings eine Ergebniskorrektur angebracht werden. Die Abweichung von der geographischen Nordrichtung wird als magnetische Deklination oder im Seefahrerwesen als Missweisung bezeichnet. Diese verändert sich nicht nur innerhalb eines Jahres, wobei dies vom Umlauf der Erde um die Sonne beeinflusst wird, sondern auch täglich. In letzterem Fall wird dies von der Erdrotation bewirkt. Aber auch Änderungen im Laufe von Jahren sind möglich. Es können ebenfalls sprunghafte Variationen, so genannte magnetische Gewitter, auftreten. Sie sind meist nicht wahrnehmbar, können aber durch die Auswertung der Messungen und Auskünfte von magnetischen Observatorien festgestellt werden [Deum88]. Zusätzlich sollten örtliche Störmassen innerhalb des Einsatzbereiches bekannt sein. Auch Gleise, Eisenbetonbauten, Gleichstromleitungen oder Computermonitore sind Störquellen bei der Bestimmung der Richtung mit einem Magnetkompass. Untersuchungen haben gezeigt, dass selbst in einer idealen Umgebung Fehler mit zwei bis drei Grad auftreten. In der Umgebung von Metallobjekten kann die Verzerrung auf 20 – 30° ansteigen. Die Empfindlichkeit hängt vom eingesetzten Sensor ab. Als elektronische Geräte werden so genannte elektronische Magnetfeldsensoren eingesetzt, die auf ein externes magnetisches Feld reagieren. Mit Hilfe eines Microcontrollers können die Ergebnisse dieser als Sensorbausteine angebotenen Bauteile verarbeitet werden.

Neigungsmesser berechnen den Vertikalwinkel direkt mit Hilfe von mit Flüssigkeit gefüllten Messkammern oder rückgestellten Pendelanordnungen. Die Genauigkeit entspricht der von Beschleunigungsmessern, allerdings sind sie preisgünstiger. Sie haben eine niedrige Messbandbreite und oft einen eingeschränkten Messbereich, so dass sie nicht alle Rotationen detektieren können. Die Sensoren sind sehr bewegungsempfindlich. Das liegt daran, dass Neigungsmesser auf dem Prinzip der Schwerkraft beruhen. Auch kleine horizontale Beschleunigungen des Kopfes werden dann detektiert. Ein System der Firma Intersense geht auf dieses Problem ein, indem drei Mikrogyroskope die Kopfbewegungen messen und die Drift korrigiert wird, wenn mit Hilfe eines Neigungsmessers und eines Kompasses festgestellt wird, dass sich der Kopf nicht bewegt. Der Vorteil dieses Systems liegt in der hohen Auflösung, der geringen Zeitverzögerung, niedrigen Kosten und einer Messanordnung, die nicht von Einflüssen in der Umgebung gestört wird. Kombinationen aus einem elektronischen Kompass und Gyroskopen, die die Kopfbewegung in Abhängigkeit der beiden anderen Richtungen bestimmen, werden bei verschiedenen Herstellern oft als Magnetometer bezeichnet. Viele günstige HMDs (Head mounted displays) haben einen elektronischen Kompass und zwei Neigungssensoren eingebaut, so dass sie für AR-Anwendungen eingesetzt werden können.

2.8 Videobasierte Verfahren

Videobasierte Verfahren nutzen passive Quellen zur Detektion von Position und Blickrichtung einer Person. In einer Umgebung, in der Marker angebracht werden können, werden diese mit Hilfe einer am Display der Person montierten Kamera erkannt. Der Einsatz dieser Art an videobasierten Verfahren beschränkt sich damit meist auf Innenräume oder eingeschränkte Bereiche.

Ein bekanntes System hierfür ist die AR-Toolkit-Bibliothek. Dabei handelt es sich um eine Sammlung von Klassen, die selbstdefinierte Marker mit einer Kamera erkennt. Sobald dies der Fall ist, wird beispielsweise ein VRML-Objekt auf dem Marker platziert. Abbildung 7 zeigt das Ergebnis dieses Vorgehens. Ein blauer Würfel erscheint im Kamerabild genau an der Stelle, an der sich der zuvor detektierte Marker befindet.

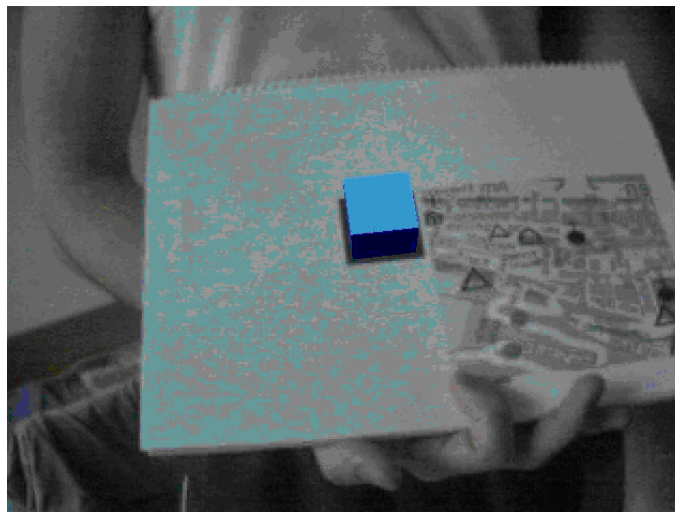


Abbildung 7: Platzierung eines blauen Würfels auf einem Marker

Diese Marker können nicht nur für innovative Interaktionstechniken sondern auch für die Positionierung eingesetzt werden.

Neben Markern können auch Merkmale in der Umgebung wie z.B. die Sonne, Sterne oder andere Objekte zur Orientierung der Kamera im Raum herangezogen werden. Kapitel 3 gibt einen Überblick über bestehende Outdoor AR-Systeme, in denen zum Großteil Bildverarbeitungsverfahren eingesetzt werden. Genauere Verfahren basieren fast ausschließlich auf der Zuordnung des Kamerabildes zu einer Referenz.

2.9 Fusion verschiedener Sensoren

Es gibt verschiedene Verfahren, die sich darauf spezialisiert haben, verschiedene Trackingsensoren miteinander zu kombinieren. Einige der neueren Verfahren sehen ihr Hauptanwendungsgebiet im Bereich AR, ältere kommen aus dem Bereich der Virtuellen Realität (VR).

Die OpenTracker Software-Architektur bietet den Zugriff auf Sensoren, verschiedene Filtertechniken und das Zusammenführen verschiedener Signale an [Reit01]. Sie ermöglicht eine einfache Integration weiterer Trackingsensoren, die ebenfalls konfiguriert werden können. Es wird davon ausgegangen, dass Trackingdaten zumeist von einer Hardware generiert, von Treibern ausgelesen sowie anschließend transformiert werden, um Bedingungen zu genügen und an die entsprechende Stelle über ein Netzwerk verschickt zu werden. Oft gibt es Unterschiede in der genauen Handhabung, allerdings sind die grundsätzlichen Schritte die gleichen. Aus diesem Grund nutzt OpenTracker diese Schritte, um ein Netzwerk an Transformationen anzubieten. Das gesamte System kann durch einen Datenflussgraph dargestellt werden, in dem jede Transformation einen Knoten repräsentiert. So werden vorgegebene Arten von Ereignissen verarbeitet. Damit können eine Position, eine Orientierung, der Status eines Knopfes, ein Zeitpunkt oder ein Vertrauenswert für die Beschreibung der Datenqualität dargestellt werden. Diese vorgegebenen Formate können erweitert werden. Verschiedene Sender und Empfänger können hierbei asynchron miteinander kommunizieren. Das System wurde entwickelt, da es bis dahin keine vergleichbare Plattform für die Vernetzung verschiedener Komponenten gab. Der Schwerpunkt des Systems liegt im Bereich der Wiederverwendung, allerdings in Bezug auf das Testen neuer Eingabeparadigmen und der Konfiguration der Trackingkomponente.

Am Institut für Informatik der TU München wurde die Idee der TU Wien zur Wiederverwendbarkeit soweit weiterentwickelt, dass ein beliebiges AR-System bestehend aus verschiedenen Diensten zusammengesetzt werden kann, das zur Laufzeit dynamisch miteinander verbunden wird [Baue01]. Auch mit diesem System können unterschiedliche Trackingsensoren integriert werden. Abhängig von der Position werden verschiedene Sensoren angesprochen. Das gesamte System besteht aus drei Komponenten. Die Softwaredienste bieten dem Nutzer oder anderen Diensten verschiedene Fähigkeiten an. Die Mittelschicht kombiniert diese Dienste dynamisch und sorgt für die Kommunikation unter ihnen. Damit kann während der Laufzeit die Konfiguration geändert werden. Die Architektur beschreibt die grundlegende Struktur von AR-Systemen, die mit diesem System zusammengestellt werden können. Die Dienste sind unabhängig davon, wann welche Sensorergebnisse genutzt werden. Die Mittelschicht hat diese abhängig von der Verfügbarkeit dynamisch ausgewählt. Damit ist es beispielsweise möglich, ein Tracking von außen in Innenräume hinein durchgehend durchzuführen. Im Außenbereich funktioniert GPS, so dass ein Kompass für die Orientierung

dazugeschaltet wird. Wenn der Benutzer in ein Haus geht, sind die GPS-Signale nicht mehr erreichbar. Aus diesem Grund werden dort die Ergebnisse anderer Sensoren benutzt. Es ist ebenfalls möglich, einen Sensor von vornherein auszuschließen, da er beispielsweise die Bewegungsfreiheit einschränkt und nicht so hohe Genauigkeitsanforderungen notwendig sind. Dieses Ausschließen, aber auch ein Hinzufügen von Sensoren, ist während der Laufzeit möglich.

Eine weitere hybride Technik, nur für die Bestimmung der Position, sind die Kombination von GPS-Ergebnissen und Positionierungsdaten mit Hilfe des Mobilfunknetzes. Diese Technik wird oft als A-GPS (Assisted GPS) bezeichnet. Sie ermöglicht es, dass Positionierungswerte auch in Bereichen möglich sind, in denen GPS nicht funktioniert, beispielsweise in Innenräumen oder in engen Straßen. Bourdeau et al. [Bour02] berechnen die Position eines mobilen Endgeräts in zwei Schritten. In einem ersten Schritt wird über das GSM/GPRS-Netzwerk und über GPS/GNSS eine erste Annäherung berechnet. Das bedeutet, dass hierbei EOTD bzw. Pseudo-Entfernungen genutzt werden. Da hierbei auch Fahrzeugnavigations-Funktionalität unterstützt wird, können diese Ergebnisse gegebenenfalls mit einem Fahrzeugnavigationsystem verbessert werden. Der zweite Schritt besteht aus dem Hinzufügen der Zellennummer und der Information mit Hilfe von TA. Die Ergebnisse des zweiten Schrittes werden über das Netzwerk bestimmt. Die Kombination aus GPS-Daten und Positionierungsinformationen des Mobilfunknetzes erlauben mit ihrer Hybridisierungs-Strategie, dass auch die Information zu nicht genügend Satelliten oder Basisstationen ausreicht, da die beiden eigentlich unabhängigen Messungen kombiniert werden. Dann müssen nicht unbedingt vier Satelliten oder vier Basisstationen erreicht werden, sondern es genügt die Bestimmung mit Hilfe von fünf gemischten Referenzen.

Die Firma PointResearch hat ein System entwickelt, das ebenfalls eine Positionierung erlaubt, wenn der GPS-Empfänger nicht genügend Signale erhält [Point03]. Dabei werden die Ergebnisse eines magnetischen Kompasses, eines elektronischen Schrittzählers und eines barometrischen Höhenmessers miteinander kombiniert. Zusammen mit den GPS-Werten gehen sie in einen Kalmanfilter ein. Die Messungen mit Hilfe von GPS erlauben die Stützung der Werte in einem absoluten System.

2.10 Zusammenfassung

Jede der hier beschriebenen Techniken weist gravierende Nachteile bezüglich des Einsatzes in einem AR-System auf, das einen Nutzer unterstützen soll, der sich frei in einer Stadt bewegt. GPS kann beispielsweise beim Einsatz in einer Stadt in engen Straßen Empfangsprobleme haben. Passive Systeme leiden oft an einer Verschlechterung des Signals bei ungünstigen Lichtverhältnissen. Beim Einsatz eines Kompasses können Erdmagnetfeldschwankungen eine Messung stark beeinflussen. Bei der doppelten Integration über die Beschleunigung führen Rauschen, Kalibrierungsfehler und die Erdbeschleunigung zu falschen Ergebnissen, die sich mit der Zeit aufaddieren [You99]. Aktive Quellen werden durch künstliche und natürliche Interferenzquellen und durch die Reichweite der Signalaufnahme beschränkt. Aus diesem Grund führt nur ein so genanntes hybrides System zu realistischen Ergebnissen.

3 Outdoor AR-Systeme

Für den Einsatz im Außenbereich müssen Sensoren eingesetzt werden, die keine aufwendige Infrastruktur in der Umgebung erfordern und die daher beispielsweise physikalische Gesetzmäßigkeiten nutzen. Wie im vorherigen Kapitel 2 erläutert, kann das GPS zur Positionsbestimmung eingesetzt werden. Zur Ermittlung der Blickrichtung können beispielsweise Kompass und Gyroskope genutzt werden. Sie müssen regelmäßig innerhalb des Einsatzgebietes gestützt werden, um grobe Fehlerfortpflanzungen zu verhindern. Dazu werden meist Bildverarbeitungsverfahren eingesetzt. Im Folgenden werden die wichtigsten der derzeit bekannten AR-Systeme für den Außenbereich vorgestellt. Diese nutzen verschiedene Sensoren, die miteinander kombiniert als Ergebnis die Position und die Blickrichtung einer Person berechnen. Dabei wird besonders auf die Verfahren eingegangen, die Bildverarbeitungsalgorithmen einsetzen, um Signalausfälle zu überbrücken oder die Ergebnisse zu stützen.

3.1 Systeme basierend auf GPS und Orientierungssensoren

Die Touring Machine, eine Entwicklung der Columbia University in New York [Fein97] [Juli00] sowie das AR-System von Piekarski von der University of South Australia [Piek99], setzen Sensoren zur Bestimmung von Position und Blickrichtung einer Person im Außenbereich ein. Die Systeme sind weder sehr genau, noch stabil, aber sie sind problemlos in Echtzeit anwendbar. Allgemein kann man sagen, dass Systeme, die nur die Ergebnisse eines DGPS-Gerätes und eines Gyroskops nutzen, beim Einsatz in einer urbanen Umgebung wegen der relativen Ergebnisse von Gyroskopen und der instabilen GPS-Signale Probleme aufweisen. Letztere werden durch enge Straßen mit hohen Häusern und Bäumen, sowie durch fahrende Autos auf stark befahrenen Straßen verursacht [Chen99]. Für viele Aufgaben sind die hiermit möglichen Genauigkeiten allerdings ausreichend.

Die Touring Machine der Columbia University blendet automatisch den Namen des Gebäudes ein, das sich im Zentrum des tragbaren Displays befindet. Darüber können weitere Informationen beispielsweise über den derzeitigen Standpunkt mit Hilfe eines Webbrowsers abgefragt werden. Wenn ein Gebäude durch längeres Betrachten ausgewählt wurde, wird eine Kompassnadel in dessen Richtung eingeblendet, die dort bleibt, wenn der Anwender den Kopf dreht. Dies ermöglicht ein einfaches Wiederfinden dieses Gebäudes. Der Hauptaspekt dieses Systems lag zu Beginn nicht auf der Entwicklung von hochgenauen Trackingsensoren sondern auf der Entwicklung einer adäquaten Benutzungsschnittstelle für diese Form der Informationspräsentation.

Die University of South Australia entwickelt ein AR-System für militärische Zwecke, in dem einzelne Informationen einer VR-Umgebung für militärische Simulationen zusammen mit realer Information eingeblendet werden. Das System wird auch eingesetzt, um den Anwender bei der Navigation zu unterstützen. Durch den Einsatz von GPS bekommt er die Positionsinformation automatisch mitgeteilt. Dies erleichtert das Zurechtfinden in einer fremden Umgebung und der Anwender kann sich auf andere Aufgaben konzentrieren.

3.2 Stützung der Orientierungssensoren mit Störungskarten

Es gibt verschiedene Ansätze, die eine Referenz zur Verbesserung der Ergebnisse nutzen. In den HRL Laboratorien in Malibu entwickelten Azuma et al. ein AR-System [Azum99a]. Darin werden ein DGPS-Empfänger, ein Kompass mit Neigungssensor und drei Gyroskope kombiniert. Der Kompass zeigt nicht nur Deklinationsabweichungen sondern auch das beschriebene Problem, dass er auf lokale Störungen des Erdmagnetfeldes reagiert. Diese Abweichungen können als Spitzenwerte 20-30° betragen. Aus diesem Grund wurden Karten angefertigt, aus denen diese Feldstörungen abgelesen werden können. Diese werden zu Beginn jedes Einsatzes zur Sensorkalibrierung herangezogen. Nach dem Lauf des Systems über mehrere Stunden zeigte sich, dass das Hauptproblem die Abweichung des Kompasses war. Aus diesem Grund wurden auch während des Einsatzes regelmäßig Kalibrierungen durchgeführt. Für eine weitere Stabilität des Systems sorgen die Gyroskope. Mit Hilfe der Ergebnisse dieser Sensoren können ebenfalls Vorhersagen über die Kopfrichtung gemacht werden. Trotzdem sind die Ergebnisse nicht genau genug, so dass nur Objekte, die mehr als 500 m vom Benutzer entfernt liegen, mit ausreichender Genauigkeit erkannt werden. Diese Objekte überlagernd werden im tragbaren Durchsicht-Display Textschilder eingeblendet.

3.3 Offline-Stützung von Orientierungssensoren mit Hilfe von Bildverarbeitungstechniken

Bildverarbeitungsverfahren werden in verschiedenen AR-Systemen zur Bestimmung von Position und Blickrichtung herangezogen. Sie basieren auf den gleichen Informationen wie sie die Person sieht, die das System anwendet. Mit der Einbeziehung eines Kamerabildes besitzt das System eine Referenz, die absolut der tatsächlichen Umgebung, der Realität des Nutzers, entspricht. Diese muss ausgewertet werden. Hier liegen auch die Nachteile dieser Technik gegenüber aktiven und anderen passiven Methoden, die allerdings auch Problembereiche aufweisen. Die Verfahren sind meist nicht sehr robust und beanspruchen hohe Rechenleistung. You et al. [You99] stellten 1999 ein System der HRL Laboratories in Malibu vor, das Trägheitssensoren, Kompass und Bildverarbeitung kombiniert. Das System basiert auf den beiden folgenden Grundsätzen:

1. Trägheitsgyroskope können sowohl die Robustheit als auch die Leistungsfähigkeit des Systems erhöhen, indem sie eine relative Voraussage der Kameraorientierung ermöglichen.
2. Ein visuelles System kann die akkumulierte Drift der Trägheitssensoren korrigieren.

Das System ist für den Einsatz in Umgebungen entwickelt worden, in denen die zu detektierenden Objekte, verglichen mit der Brennweite, relativ weit entfernt liegen. In diesem Fall beeinflusst eine rotatorische Bewegung der Kamera die Änderung des Inhaltes des Sichtfeldes stärker als eine translatorische. Dann stützen sich die Bildverarbeitungs-komponente und die Richtungskomponente des Systems gegenseitig. Die Signale aller Sensoren – einschließlich der Signale der Kamera – gehen gemeinsam in die Auswertungskomponente ein. Hier werden die

Unterschiede im Signaleingang zwischen den beiden Orientierungssensoren berücksichtigt. Die Bildverarbeitungssoftware arbeitet nicht in Echtzeit. Deshalb werden sowohl die Ergebnisse der Orientierungssensoren als auch die Videobilder für die Offline-Bearbeitung benutzt.

Um Aussagen über die Genauigkeit der Gyroskope zu machen, wurden manuell Merkmale in einem Bild bestimmt, während sowohl Kamera als auch die Gyroskope still standen. Dann wurden die Geräte bewegt. In dieser Zeit verfolgte die Bildverarbeitungssoftware die Merkmale. Die daraus gewonnenen Ergebnisse werden mit denen der Gyroskope verglichen. Die Abstände der Pixel sind proportional zu den aufaddierten Fehlern der Gyroskope.

Das vorgestellte System ermöglicht es, dass einzelne Objekte automatisch verfolgt und mit einem Pfeil und dem Objektnamen versehen werden. Dafür werden Messungen von Orientierungssensoren mit Bildverarbeitungstechniken zusammengebracht, um ein robustes Orientierungsergebnis zu generieren. Das System arbeitet nach dem Vorhersage-Korrektur-Prinzip. Mit Hilfe der Daten der Trägheitssensoren wird die ungefähre 2D-Bewegung der Bildinhalte hervorgesagt. Das Bildverarbeitungssystem korrigiert und verbessert die Vorhersagen. Zum Schluss werden die 2D-Bewegungsresiduen in 3D-Orientierungskorrekturen für die Gyroskope umgewandelt. Mit der Einschränkung des Suchraumes im Bild und der Unterstützung des Systems bei Bildstörungen wird die Effizienz der Bildverarbeitungssoftware erhöht.

3.4 Horizontsilhouetten

Behringer vom Rockwell Science Center behob den Mangel an Genauigkeit von GPS-Gerät, magnetischem Kompass und Neigungssensoren durch die Entwicklung eines hybriden Systems, das auf der Auswertung von Videobildern basiert [Behr99]. Dieses soll eingesetzt werden um eine visuelle Echtzeitregistrierung durchzuführen, während sich Kamera und Beobachter bewegen. Der Einsatzbereich sind Navigation und Gefahrenanzeige. Das System läuft auch mit einer nicht-kalibrierten Kamera und kann dabei die Kalibrierungsparameter ermitteln. Hierbei werden inertielle Tracking-systeme mit statischen Positionsmessungen und Bildverarbeitungstechniken kombiniert. Als Referenz der tatsächlichen Umgebung dienen Landschaftsilhouetten. Sie sind die sichtbare Grenze zwischen Erdoberfläche und Himmel. Im vorliegenden Fall handelt es sich um die natürliche Erdoberfläche. Man könnte aber auch Umriss von Gebäuden zu Hilfe nehmen. Diese können zur Bestimmung von Position und Orientierung eines Beobachters genutzt werden. Man benötigt dafür ein digitales Höhenmodell. Mit dieser Datengrundlage kann die Horizontsilhouette für einen vollständigen Umkreis von 360° berechnet werden. Dazu muss die Position jedes Datenpunkts in ein kartesisches Koordinatensystem mit dem augenblicklichen Standpunkt als Ursprung transformiert werden.

In diesem Verfahren wird zunächst die Position mit Hilfe eines GPS-Geräts ermittelt. Die sichtbare Silhouette erkennt man normalerweise an einem steilen Grauwert-Gradienten im Kamerabild, das das Sichtfeld des Benutzers repräsentiert. Sie kann über ein Kantenerkennungsverfahren mit anschließender Extremwertbestimmung detektiert werden. Für das anschließende Korrespondenzverfahren werden Hypothesen, die den Orientierungswinkel und den Neigungswinkel sowie die Kameraparameter unter der Annahme einer Lochkamera-Abbildung definieren, aufgestellt und gegeneinander abgewogen. Die Hypothese mit der minimalen Distanz wird herangezogen, um einen weiteren Minimierungsalgorithmus für das

eigentliche Matching anzuschließen. Die beschriebenen Ergebnisse zeigen, dass das Verfahren bei einem scharfen Kontrast zwischen Himmel und Erdoberfläche sehr viele Hypothesen schnell ablehnen kann. Die Ergebnisse können dann für die weiteren Schritte näher untersucht werden. Der Korrespondenzfindungs-Algorithmus verlief nicht sehr genau, da die Datengrundlage nicht engmaschig genug war. Das Verfahren ist zum aufgezeigten Zeitpunkt für den Gebrauch in einem Fahrzeug zum Einsatz im Außenbereich vorbereitet worden. Bis dahin lief es auf einem 200 MHz PC. Echtzeitergebnisse sind währenddessen nicht möglich gewesen. Das Verfahren läuft nicht zufrieden stellend, wenn die Silhouetten weniger als drei bis fünf Kilometer nahe sind, da dann nicht genügend Details ausgemacht werden können.

3.5 Einbeziehung von Daten eines 3D-GIS

Das Center for Spatial Information Science der University of Tokyo integriert ihr AR-System in ein GIS [Chen99]. 3D-Modelle aus einer Datenbank sollen in das Kamerabild eingeblendet werden. Dazu nutzt es zwei Entwicklungen. Einerseits entwickeln sich GIS von 2D- zu 3D-Informationssystemen, andererseits bewegen sie sich von Systemen, die man als Desktoplösung anwendet hin zu mobilen Systemen, die auch im Außenbereich anwendbar sind. Die Gruppe setzt ein DGPS-Gerät, ein INS, ein Gyroskop, einen elektrischen Kompass und Bildverarbeitungstechniken ein.

Wenn für die GPS-Signale kein Empfang möglich ist, werden den Bildern einer CCD-Kamera Linien besonders auffälliger Gebäude, so genannter Landmarks, aus 3D-Modellen zugeordnet. Dabei gehen die Ergebnisse des Gyroskops in dieses Verfahren mit ein. Die 3D-Modelle sowie eines der aufeinander folgenden Bilder, das zuvor georeferenziert worden ist, liegen in einer Datenbank. Mit Hilfe dieser georeferenzierten Daten können Position und Richtung der Kamera für die weiteren Bilder einer zuvor bestimmten Sequenz bestimmt werden. Für diesen Vorgang muss eine ausreichende Anzahl an Linien aus dem 3D-Modell in zwei aufeinander folgenden Bildern gefunden werden. Diese Bedingung wird oft nicht erfüllt, da der Abstand der Kamera zu den realen Objekten meist nicht groß genug ist und aus diesem Grund zu wenige oder keine herausragenden Gebäude gefunden werden können.

Wenn deshalb keine Linien aus der Datenbank in zwei aufeinander folgenden Bildern erkannt werden, wird die relative Orientierung mit Hilfe von korrespondierenden Punkten und Linien nur dieser beiden Bilder bestimmt. Dazu sind fünf verschiedene korrespondierende Punkte in dem Stereobildpaar notwendig. Das Ergebnis ist die relative Translation und Rotation der Kamera im Verhältnis zur Kameraposition des Bildes zuvor. Das Verfahren läuft solange, bis für alle Bilder die Position und Ausrichtung der Kamera bestimmt worden ist. Die Ergebnisse der Positionierungs- und Orientierungssensoren werden nur herangezogen, um den Suchbereich in der räumlichen Datenbank einzuschränken und um nicht korrespondierende Gebäudelinen festzustellen.

3.6 Matching des Kamerabildes mit Referenzphotos

Innerhalb des ARCHEOGUIDE-Projektes entwickelte Stricker ein System, das als Referenz zuvor erstellte Photos nutzt [Stri01]. Bilder, die von einer an einem tragbaren Display befestigten Kamera aufgenommen werden und das Sichtfeld des Nutzers repräsentieren, werden mit den Referenzphotos verglichen. Bei jedem Vergleich wird ein Korrelationskoeffizient berechnet. Das Paar mit dem besten

Ergebnis wird zur Berechnung der 2D-Transformation herangezogen. Dafür wurde ein Ansatz entwickelt, der auf der Fourier-Mellin-Transformation basiert. Mit dieser können als Parameter allerdings nur die Rotation, die Translation und der Maßstab berechnet werden. Das System kann nur in der unmittelbaren Umgebung der zuvor definierten Viewpoints, an denen die Referenzphotos erstellt wurden, für bestimmte Kamerabewegungen eingesetzt werden. Für den Einsatz in einer Echtzeit-Umgebung wird die Fouriertransformation für die Referenzbilder im Vorhinein berechnet. Nach Ermittlung der zweidimensionalen Transformation kann die virtuelle zweidimensionale Information in den Videostrom integriert werden. Abbildung 8 zeigt die Einblendung einer solchen Rekonstruktion von einem festen Viewpoint aus in Echtzeit. Das System arbeitet auch bei Verdeckungen bis zu 60 % und bei unterschiedlichen Lichtverhältnissen fast in Echtzeit.



Abbildung 8: Einblendung einer Rekonstruktion von einem Viewpoint aus [Arch03]

3.7 Matching des Kamerabildes mit einem 3D-Modell

An der TU Graz wird ein System entwickelt, das die Signale eines inertialen Trackers bestehend aus einem Beschleunigungsmesser und einem Gyroskop, der eine Eigenentwicklung darstellt, und eine Bildverarbeitungs-komponente kombiniert [Ribo02]. Das System beherrscht damit sechs Freiheitsgrade (drei rotatorische und drei translatorische Richtungen).

Grundlage für das videobasierte Tracking sind hochgenaue, photorealistische 3D-Modelle, die über einen Bildverarbeitungsalgorithmus aus einzelnen nacheinander aufgenommenen Bildern erstellt werden. Aus diesen werden offline auffällige Gebäude herausgenommen. In diesem neuen, verkleinerten Modell werden Ecken gesucht, die auch in dem entsprechenden Bild des Videostroms detektiert werden. Um die Suche zu erleichtern, wird eine Vorauswahl an Zielpunkten in der Modelldatenbank durchgeführt. Das Herausfinden dieser herausstechenden Punkte geschieht vor dem Einsatz des Systems. Das entsprechende Modul benötigt genauso wie das Trackingmodul Informationen über die aktuelle Position. Zu Beginn eines Einsatzes muss das System initialisiert werden. Anschließend ist die Position bekannt, genauso wie die Merkmale, die getrackt werden sollen und der Suchbereich

für jedes Bild des Videostroms. Der Beschleunigungsmesser liefert Ergebnisse mit einer Frequenz von 200 Hz und mehr. Die Frequenz der Bildverarbeitungs-komponente ist nicht so hoch, allerdings liefert diese verlässlichere Ergebnisse. Aus diesem Grund stützt diese die Ergebnisse des Beschleunigungsmessers. Das System wird eingesetzt, um sowohl einfache Objekte zu verfolgen als auch einfache Graphiken einzublenden.

3.8 Zusammenfassung

Der vorliegende Abschnitt zeigte verschiedene Verfahren auf, wie Position und Blickrichtung einer Person im Außenbereich bestimmt werden können. Die als mobile Systeme implementierten Verfahren benötigen alle ein Laptop, auf dem das System läuft. Tragbare Computer wie beispielsweise der Xybernaut werden dafür nicht eingesetzt. Für einige der Anwendungen ist dieser noch nicht leistungsfähig genug. In der folgenden Tabelle 1 sind die erläuterten Verfahren übersichtlich dargestellt. Die Bezeichnungen der Orientierungssensoren sind der in den entsprechenden Abschnitten angegebenen Literatur entnommen worden.

Verfahren	Positonsensoren	Orientierungssensoren	Unterstützung der Sensoren	Referenz	Ziel des Verfahrens
Touring Machine	Realtime-kinematic GPS	Orientierungstracker (InterSense IS300Pro)	keine	keine	Entwicklung einer Benutzungsschnittstelle
University of Southern Australia	GPS	Digitaler Kompass in drei Achsen (TCM2)	keine	keine	Militärische Simulationen
HRL Laboratories	DGPS	Kompass und Neigungssensor (TCM2), drei Gyroskope (Systron Donner GyroChip)	Kalibrierung	Karten mit Störungen des Erdmagnetfeldes	AR für den Außenbereich
Rockwell Science Center	GPS	Magnetometer / Neigungsmesser (Cybertrack CT-3.2)	Kamerabild	Digitales Höhenmodell	Navigation und Gefahrenanzeige
University of Tokyo	GPS und INS	Magnetometer und Gyroskop	Kamerabild	Georeferenziertes Bild einer Sequenz und 3D-GIS-Daten	Mobiles 3D-GIS
Stricker	keine	keine	Kamerabild	Referenzbild	Überlagerung von historischen 3D-Modellen
TU Graz	Eigenentwicklung eines inertialen Trackers		Kamerabild	Signifikante Ecken eines 3D-Stadtmodells	Einblenden von einfachen Graphiken

Tabelle 1: Übersicht über die erläuterten Verfahren zur Positions- und Blickrichtungsbestimmung im Außenbereich

Zusätzlich zu den generellen Genauigkeitsproblemen von GPS und Orientierungssensoren können beim Einsatz in einer Stadt die GPS-Signale durch eng stehende Gebäude behindert und verfälscht werden. Aus diesem Grund ist die Entwicklung eines Bildverarbeitungsverfahrens entscheidend, das die Bildfolgen der Kamera mit einer Referenz vergleicht.

Am performantesten scheint der Ansatz von Stricker zu sein. Allerdings weist er eine Beschränkung auf so genannte Viewpoints auf sowie die Überblendung von Informationen nur in 2D. Bei der Methode der TU Graz werden Zielpunkte aus Bildfolgen extrahiert, die in den Videobildern erkannt werden. Diese Methode funktioniert demnach nur nach einer Bearbeitung der Datengrundlage vor dem Einsatz und ist dadurch nur bei Vorliegen von geeignetem Datenmaterial und unter Einsatz des eigens produzierten Beschleunigungsmessers möglich.

4 3D-Stadtmodelle

3D-Stadtmodelle werden schon sehr lange von Architekten genutzt. Sie helfen beispielsweise bei der Veranschaulichung, wie sich ein geplantes Gebäude in die Umgebung einfügt. Diese Modelle werden oft von Hand erstellt. Wenn sie flächendeckend für eine Stadt vorliegen, handelt es sich meist um einfache Holzmodelle. Abbildung 9 zeigt ein solches Modell für einen Ausschnitt von Berlin.

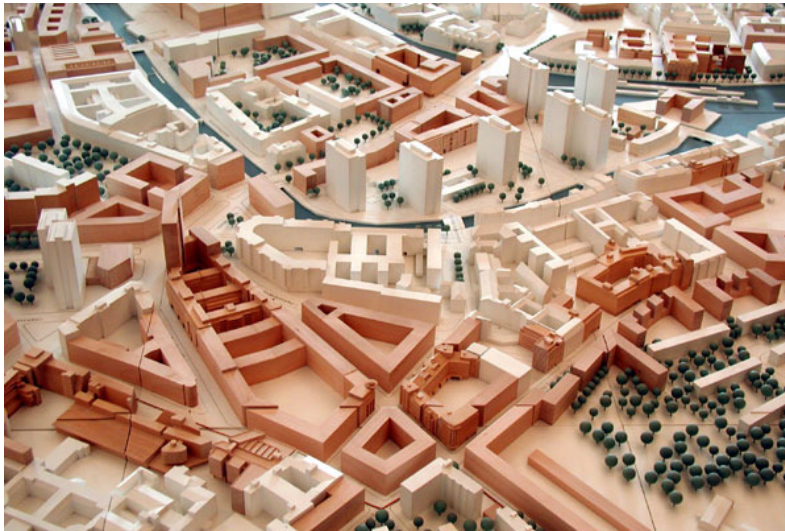


Abbildung 9: Holzmodell eines Ausschnittes von Berlin im Maßstab 1:500 [SenB03]

Die Daten für die dritte Dimension wurden meist nur für diese Anwendung erhoben, da ansonsten bei der Erfassung von Geodaten, beispielsweise für die Katasterverwaltungen oder für Karten, Gebäude und Verkehrswege ausschließlich 2D-Daten eine Rolle spielen [Krau00]. Die einzigen flächenhaften Darstellungen von Geodaten in der dritten Dimension, im digitalen Bereich, waren lange Zeit dreidimensionale Geländemodelle. Diese Einstellung hat sich in den letzten Jahren gewandelt. Besonders seit dem Aufschwung in der mobilen Telekommunikationsbranche sind es diese Firmen, die digitale 3D-Stadtmodelle erfassen. Sie benötigen die Daten im innerstädtischen Bereich für die Mobilfunknetzplanung [Bill99a]. Dies führte dazu, dass Entwicklungen im Bereich Laserscanning-Messungen und Verfahren zur Objektrekonstruktion aus digitalen Bildern entstanden sind [Bill99b]. Die Befliegung zur Erstellung, die Auswertung und auch die Speicherung der Daten für 3D-Stadtmodelle wird heutzutage von verschiedenen Dienstleistungsfirmen angeboten.

Bei der Datenerfassung von 3D-Stadtmodellen finden zwei Techniken verbreiteten Einsatz: Die digitale Luftbildphotogrammetrie, auf die im anschließenden Abschnitt 4.1 eingegangen wird, und die Laserscanning-Messung aus der Luft. Sie wird in Abschnitt 4.2 näher erläutert. In beiden Fällen werden Daten vom Flugzeug aus erfasst. Die beiden danach, in den Abschnitten 4.3 und 4.4 aufgeführten Methoden, die photorealistische Erfassung und die manuelle Modellierung, werden meist nur punktuell für ausgewählte Gebäude eingesetzt, wobei die Grundgeometrien der

Gebäude häufig aus den beiden zuerst genannten Erfassungsmethoden stammen. Aus diesem Grund wird auf sie nicht ausführlich eingegangen. Vor der Zusammenfassung in Abschnitt 4.6 werden in Abschnitt 4.5 Beispiele für 3D-Stadtmodelle genannt.

4.1 Photogrammetrische Erfassung

Bei der photogrammetrischen Erfassung werden Photos aufgenommen, mit denen nach einem Orientierungsprozess ein dreidimensionales virtuelles Gelände erzeugt werden kann. Abbildung 10 zeigt ein solches Luftbild in Grauwertdarstellung, das zusammen mit einem weiteren überlappenden Bild für die 3D-Auswertung herangezogen werden kann.



Abbildung 10: Digitales Luftbild zur photogrammetrischen Auswertung

Durch die Verwendung von so genannten Passpunkten ist das Messen von absoluten Koordinaten möglich. Gebäude in Luftbildern können aufgrund der Intensitätsunterschiede der Grauwerte erkannt werden. Dies führt zur Detektion von Punkten, Linien oder homogenen Gebieten. Eine automatische Detektion wird aufgrund der folgenden Eigenschaften von Luftbildern erschwert [Förs99]:

- Die Form der Objekte kann nicht bestimmt werden, da Bilder nur zweidimensionale Informationen enthalten.
- Bei Verdeckungen tritt ein Informationsverlust auf.
- Es sind Daten vorhanden, die für die Auswertung bedeutungslos sind.
- Die Extraktion der Bildmerkmale wird durch Bildrauschen, geringen Kontrast, eingeschränkte Modellierung und störende Objekte erschwert.
- Die Objektsemantik fehlt.

Aufgrund dieser Probleme ist die manuelle photogrammetrische Datenerfassung weiterhin die verbreitete Vorgangsweise der Luftbildauswertung. Ein Operateur erstellt ein Gebäudemodell aus Stereoluftbildern. Darin werden zunächst die Traufen digitalisiert. Meist genügt dies, wenn der Detaillierungsgrad der Gebäude relativ grob

ist. Aus den Traufflinien werden dann die Gebäudeumrisslinien durch Verschneidung mit dem Gelände abgeleitet. In diesem Fall werden Dachüberstände ignoriert. Um diese in die Auswertung einzubeziehen, können sie manuell erfasst werden. Alternativ können die Dachüberstände terrestrisch vor Ort mit Hilfe eines Dachlots, eines elektronischen Tachymeters oder durch die Kombination mit 2D-GIS-Datenbeständen erfasst werden. Anschließend werden die Dächer digitalisiert. Der Operateur ist derjenige, der entscheidet, um was für ein Objekt es sich handelt und wie er das Gebäude erfassen soll. Der theoretische Detaillierungsgrad der Objekte ist von der Überfliegungshöhe und damit vom Bildmaßstab abhängig. Es können Genauigkeiten in der Größenordnung von 0.2 % der Flughöhe über Grund erreicht werden [Arbe00].

Um 3D-Daten möglichst kostengünstig und schnell zu erfassen, lag der Schwerpunkt der photogrammetrischen Forschung in den letzten Jahren im Entwickeln von automatischen Verfahren. Dabei sind Verfahren zur Erhebung eines Rastermodells durch Korrelation und Interpolation eine effiziente Erhebungsmethode in der offenen Flur. Mit diesen sind annähernd die gleichen Höhengenaugigkeiten erreichbar wie mit der manuellen Auswertung, d.h. 0.2 - 0.3 % der Flughöhe über Grund [Arbe00]. Allerdings sind vollautomatische Methoden zur Rekonstruktion von Gebäuden aus Stabilitäts- und Zuverlässigkeitsgründen im Moment noch nicht praktikabel. Es existieren verschiedene Ansätze für die automatische Gebäudeextraktion aus Luftbildern. Heuel hat ein Verfahren zur automatischen Extraktion von Gebäuden aus mehreren Luftbildern entwickelt, das von einem konsistenten Polyedermodell für die Gebäudeoberfläche ausgeht [Heue00]. Mayer untersuchte sieben verschiedene Ansätze und stellte fest, dass sie für begrenzte Gebiete dem Einsatz in der Praxis nahe kommen [Maye99]. Förstner behauptet, dass diese Methoden wohl auch in den nächsten Jahrzehnten nicht die Erfolgsraten erlangen werden, die mit einer manuellen Auswertung möglich sind [Förs99]. Die bestehenden Ansätze auf Basis von Luftbildern können derzeit noch nicht in eng bebauten Gebieten eingesetzt werden oder die Ergebnisse sind bis jetzt noch nicht mit einer großen Anzahl von Gebäuden überprüft worden.

Aus diesem Grund werden heutzutage semi-automatische Methoden in der Praxis eingesetzt. Sie unterstützen entweder den Operateur beim Digitalisieren oder es wird ein Werkzeug angeboten, das die vom Operateur ermittelten Ergebnisse anzeigt und automatisch korrigiert. Zu dem zuerst genannten Fall der semi-automatischen Auswertung zählt die topologiegestützte Datenerfassung. Das bedeutet, dass eine Wissensbasis zu Hilfe genommen wird, in der die Topologie von Standardflächen gespeichert ist. Dazu ist eine Auswahl an Dachformen vorgegeben, die an das entsprechende Gebäude angepasst werden. Das Verfahren, das in Zusammenarbeit mit der Universität Bonn entstanden ist, benötigt als Eingangsdaten digitale Luftbilder sowie die Parameter der inneren und äußeren Orientierung [Gülc01]. Im Gegensatz zu ähnlichen Verfahren sind keine zusätzlichen 2D-Daten notwendig, soweit es sich nicht um eine Einbild-Auswertung handelt. Das Verfahren bietet verschiedene Gebäudeprimitive an, die für die meisten Haus- und Dachtypen anwendbar sind. Grundsätzlich können weitere Primitive zusätzlich definiert werden. Im Gegensatz zur rein manuellen Auswertung, bei der jeder Punkt im Stereobild einzeln aufgemessen wird, werden hierbei Parameter der Grundprimitive an das vorliegende Gebäude angepasst. Aufwendigere Gebäude können durch eine Kombination von Grundprimitiven dargestellt werden. Der Vorteil des Systems liegt darin, dass Bildverarbeitungsmethoden den Anwender beim Aufsuchen unterstützen. Aus diesem Grund kann in den einzelnen Bildern gearbeitet werden, die zusammen ein Stereobild darstellen. Das Auffinden von Merkmalen in einem Monobild ist für einen Operateur sehr viel angenehmer und dadurch genauer und schneller. Die Arbeit in

einem dieser Bilder wird automatisch in den anderen angezeigt. Dies ist eine einfache und effektive Kontrollmöglichkeit. Das Verfahren kann auch für die Auswertung von Laserscanning-Messungen und Satellitenbilder angewendet werden. Der zweite genannte Fall, die Verbesserung der vom Operateur erfassten Daten, findet sich im CyberCity Modeler, einer Software, wieder, die von einem Spin-Off der ETH Zürich vertrieben wird [Grue01]. Hier werden die Objektidentifikation sowie die Messung vom Operateur durchgeführt. Die Strukturierung der Punktwolken, d.h. die Entscheidung, welche Punkte zu einer Ebene gehören, wird automatisch von der Software vorgenommen, da davon ausgegangen wird, dass das Messen in Stereobildpaaren grundsätzlich notwendig ist, aber oft Schwierigkeiten bereitet. Auch Oberflächen, die nicht in einer Ebene liegen, werden von der Software erfasst. Auf die Dächer und die Landschaft wird automatisch Textur aus den Luftbildern aufgebracht. Fassaden können über terrestrisch erfasste Photos mit Textur versehen werden.

Die herkömmliche Form der photogrammetrischen Stereobildauswertung geht davon aus, dass ein Gebiet von einem Flugzeug überflogen wird und zwei sich überlappende Bilder für die Stereoauswertung notwendig sind. Die HRSC-A (High Resolution Stereo Camera) erlaubt es, einen Objektpunkt mit fünf verschiedenen Kanälen zu erfassen, wodurch eine höhere Abdeckrate gerade in urbanen Gebieten erreicht wird [Neuk99]. Tests haben gezeigt, dass diese Form der Datenerfassung ± 18 cm in vertikaler und ± 13 cm in horizontaler Richtung an Abweichung aufweist. Damit können nur mit einer Überfliegung bessere Daten zur Stereoauswertung erfasst werden, als es mit herkömmlichen photogrammetrischen Informationen möglich ist. Gerade durch die erwähnten Vorteile im urbanen Bereich ist davon auszugehen, dass sich diese Entwicklungen der digitalen photogrammetrischen Datenerfassung für die Erfassung von 3D-Stadtmodellen in Zukunft durchsetzen werden.

4.2 Erfassung mit Hilfe des Laserscanning-Verfahrens

Neben Luftbildern hat sich für die flugzeuggestützte Erfassung der Gelände- und Objektfläche das Laserscanning-Verfahren etabliert [Acke99]. Die Einsatzbereiche von aus Laserscanning-Messungen abgeleiteten Daten reichen von Überschwemmungsvorhersagen über Sichtbarkeitsanalysen, Schallausbreitungen, Volumenänderungen, beispielsweise von Mülldeponien, und dem so genannten Corridor Mapping, d.h. der Beobachtung von Hochspannungsleitungen oder Autobahnen, bis hin zu 3D-GIS Daten. Kommerziell wird diese Art der 3D-Datenerfassung seit 1995 angeboten [Vegt01]. Dabei wird die Laufzeit des ausgesendeten Laserstrahls gemessen. Abbildung 11 zeigt das Scannen der Geländeoberfläche von einem Flugzeug aus.

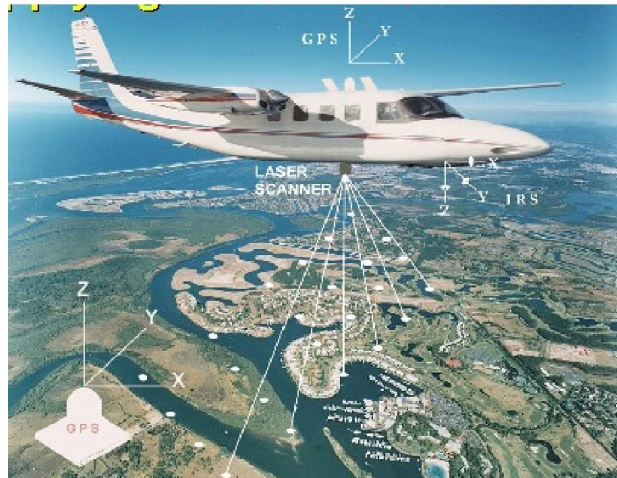


Abbildung 11: Laserscanning-Messung der Geländeoberfläche vom Flugzeug aus [Knor00]

Die Messungen decken eine Fläche ab, die über den Scanwinkel und die Bewegung des Flugzeuges definiert sind. Daraus ergeben sich Abstandswerte, die allerdings die Erfassung von Flächen, beispielsweise unter Brücken, ausschließen. Die Ergebnisse beziehen sich zunächst auf ein sich bewegendes Koordinatensystem mit Ursprung im Flugzeug. Zur Transformation in das globale Koordinatensystem werden DGPS und INS für die Bestimmung von Position und Orientierung eingesetzt. Diese Daten liefern Oberflächenmodelle mit hoher Genauigkeit. Oberflächenmodelle sind Daten, die die Oberfläche der Erde als Punktwolke wiedergeben. Abbildung 12 zeigt eine solche Punktwolke, die die Rohdaten einer Laserscanning-Messung darstellt, zusammen mit Gebäuden, die mit einem von der Universität Bonn entwickelten Werkzeug halbautomatisch extrahiert worden sind. Die Darstellung eines Gebietes als Oberflächenmodell ist wegen der hohen Datenmenge sehr rechenintensiv.

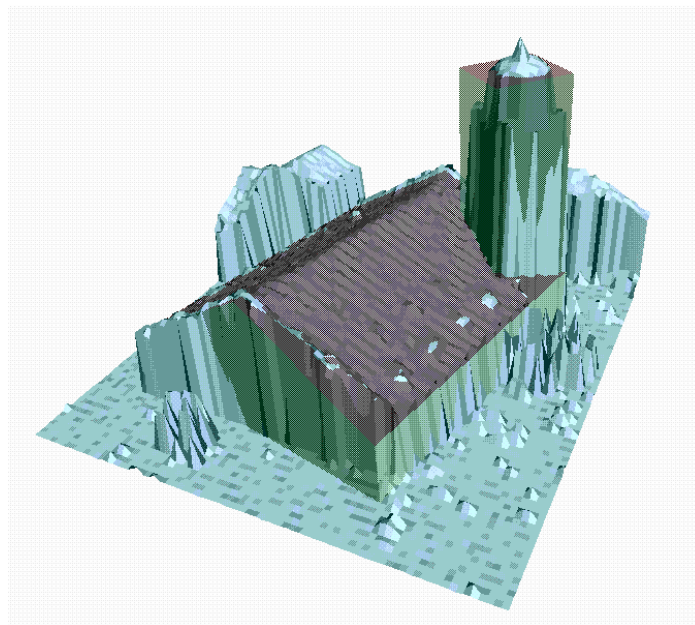


Abbildung 12: Laserscanning-Messung als Punktwolke mit extrahierten Gebäuden [IPB03]

Bei den Rohdaten handelt es sich um eine rein ikonische Präsentation, die nicht objektbezogen vorliegt. Mit Laserscanning-Verfahren können daher keine Geländekanten erfasst werden. Digitale Geländemodelle (DGMs) und dreidimensionale Gebäudedaten müssen erst aus den aus der Luft erfassten Daten ermittelt werden. Nach Bill umfasst ein DGM sowohl ein Digitales Höhenmodell (DHM) als auch ein Digitales Situationsmodell (DSM) [Bill99b]. Das DHM umfasst digital gespeicherte Höhenwerte, das DSM umfasst die als Grundriss bezeichnete Information.

Die reflektierten Signale erlauben allerdings Aussagen über die Bodenbedeckung und lassen auch Messungen in Waldgebieten zu. Dabei werden sowohl der First Pulse als auch der Last Pulse des ausgesendeten und wieder empfangenen Signals ausgewertet. Im urbanen Gebiet enthält der First Pulse beispielsweise Daten zu Gebäuden und Bäumen entlang von Straßen. Der Last Pulse liefert Höheninformationen zu den Straßen. Die Vegetation ist hierbei fast ganz verschwunden. Über die Differenzbildung können Kronenmodelle von Bäumen, sowie zusammen mit der Oberfläche Baumhöhenbestimmungen mit Genauigkeiten zwischen einem halben und einem Meter und sehr präzise Volumenbestimmungen durchgeführt werden. Die Voraussetzung dafür sind Auflösungen von einem Punkt pro Quadratmeter. Es ist eine ganzjährige Befliegung möglich. Allerdings ist der Zeitpunkt der Befliegung meist vom Anbieter abhängig. Laserscanning-Verfahren sind sowohl im Wald als auch in eng bebauten Gebieten bei der Erfassung von Höhenmodellen den herkömmlichen photogrammetrischen Methoden überlegen. Der Vorteil entsteht dadurch, dass bei der photogrammetrischen Auswertung ein Punkt in mindestens zwei Bildern sichtbar sein muss. Außerdem sind genügend Licht und eine ausreichende Textur Grundvoraussetzungen für akzeptables Bildmaterial.

Es sei erwähnt, dass Wolken auch für Laserscanning-Verfahren ein Hindernis darstellen und die klare Wasseroberfläche durch Laserscanning-Verfahren gar nicht erfasst werden kann. Die Verknüpfung der Rohdaten bei Laserscanning-Verfahren basiert auf der Messung von Position und Orientierung des Flugzeuges mit zusätzlichen Sensoren. Dadurch ist eine Kontrolle dieser Zuordnung nicht so einfach durchzuführen wie bei der Verknüpfung von Luftbildern für die Stereoauswertung. Hier basiert die Verknüpfung direkt auf der Auswertung der Rohdaten. Bei photogrammetrischem Bildmaterial wird über die so genannte Bündelblockausgleichung ein Verbund der Ergebnisse erreicht. Der Flugaufwand für die Erfassung photogrammetrischen Bildmaterials ist geringer als der für Laserscanning-Daten. Besonders bei hohen Befliegungen können nur Luftbilddaten erfasst werden.

Schmidt untersuchte die Genauigkeiten verschiedener Methoden zur Erfassung von DGMs und stellte dabei fest, dass die Punktgenauigkeit im offenen Gelände ± 0.1 m bei hoher Messpunktdichte von 1-3 m beträgt [Schm03]. Er weist auch darauf hin, dass die Kosten für Laserscan-Befliegungen sehr hoch sind und aus diesem Grund die von der Landesvermessung erfassten DGMs zukunftssträftig sind. Diese werden allerdings meist nur in einem 10 m-Raster angeboten.

In einer weiterführenden Verarbeitung werden aus der großen Menge von Laserpunkten dreidimensionale CAD-Modelle für jedes Gebäude abgeleitet. Bei diesen Methoden können bei einer hohen Punktdichte gute Ergebnisse auch bei automatisierten Verfahren erzielt werden, die nur auf digitalen Oberflächenmodellen basieren. Dabei ist die Genauigkeit im Allgemeinen nicht so gut wie bei Messungen aus Luftbildern, da die Punktdichte nicht hoch genug ist [Grue01]. Für die automatische Generierung von Dachlandschaften sind 3-5 Punkte pro Quadratmeter notwendig um Bruchkanten geeignet detektieren zu können. Für Dachgauben sind 4-5 Punkte Voraussetzung, praktische Anforderungen erreichen Dichten von 20

Punkten pro Quadratmeter. Dichten von 4-5 Punkten bieten aber nur wenige Anbieter an. Bei nicht ausreichenden Punktdichten führt die Abweichung zwischen dem tatsächlichen Gebäude und der Punktwolke zu Abweichungen in der Orientierung des Hauses und in der Dachneigung. Semi-automatische Verfahren dagegen ermöglichen im Prinzip Erfolgsraten von 100 %. Es ist davon auszugehen, dass die Punktdichte zunehmen wird, da die Sensoren in Bezug auf die Steigerung der Wiederholrate des Lasers sich ständig verbessern [Flo01].

Brenner und Haala entwickelten am Institut für Photogrammetrie der Universität Stuttgart sehr viel versprechende Ergebnisse einer automatischen Rekonstruktion von Bebauung und Vegetation auf Grundlage von Laserscanning-Messungen [BrHa00] [Bren00]. Diese weitestgehend automatisierte Ableitung von 3D-Stadtmodellen ermöglicht die Kombination der Auswertungen mit vorhandenen Grundrissdaten eines 2D-GIS. Zusammen mit einem durch Laserscanning-Messungen erfassten DHM erfolgt die Ableitung des 3D-Stadtmodells ohne weiteren Eingriff von außen. Die zu rekonstruierenden Gebäude werden durch die Kombination einfacher 3D-Grundelemente, so genannter Primitive, beschrieben. Als mögliche Dachformen der Gebäudeelemente sind Flach-, Giebel- und Walmdach vorgesehen. Durch eine Kleinste-Quadrate-Schätzung werden die Differenzen zwischen der zu rekonstruierenden Dachfläche und dem DHM minimiert. Dieses Verfahren eignet sich auch zur interaktiven Korrektur und Verbesserung der dabei generierten Ergebnisse sowie zur semi-automatischen Gebäudeerfassung. Für ein beschriebenes Testgebiet mit einem mittleren Punktabstand der Laserscanning-Messung von einem Meter lagen die Grundrisse der Gebäude in der Stadtgrundkarte im Maßstab 1:500 vor bzw. wurden direkt aus Einmessungen abgeleitet. Das Verfahren verlief zufrieden stellend für den Großteil der Gebäude. Grobe Fehler entstanden bei einzelnen Gebäuden in der Rekonstruktion der Dachform. Die Hauptursache dafür bestand in den GIS Daten, in denen die Innenhöfe der Gebäude teilweise nicht enthalten waren. Eine manuelle Nachbearbeitung über den visuellen Eindruck ist mit einem entsprechenden Programmpaket möglich. Dafür können die vorhandenen Grundrisse mit einem Orthophoto, einer Grauwertdarstellung des DHM oder einer gescannten Karte überlagert werden. Weitere Verfahren, die ein 2D-GIS einsetzen, nutzen beispielsweise die Stockwerkszahl zur Bildung von 3D Hypothesen.

Gülch et al. [Gül00] haben mit AMOR ein System entwickelt, das eine semi-automatische Gebäuderekonstruktion aus verschiedenen Datenquellen wie Satellitenbildern, Laserscanning-Messungen und Luft- bzw. Nahbereichsbildern erlaubt. Zusätzlich zu den Basiswerkzeugen interaktiver Methoden sind hierbei die automatische Messung von Höhen sowie die automatische Bestimmung von Primitivparametern möglich. Als Primitive werden hier 3D-Grundelemente bezeichnet. Nach Messung der Erfassungszeit pro Primitiv wurde festgestellt, dass die Geschwindigkeit bei diesem Verfahren etwa 20 Sekunden pro Primitiv betragen kann. Diese Zeit ist vergleichbar mit den klassischen photogrammetrischen Verfahren.

Ein Problem bei der Erfassung von Daten aus der Luft ist, dass Daten nur periodisch erfasst werden können. Das bedeutet, dass ein Gebiet regelmäßig überflogen, ausgewertet und mit dem älteren Datenbestand verglichen werden muss. Einzelne Gebäude können nicht nachträglich eingemessen werden, die Auswertung erfolgt gebietsweise unabhängig von aktuellen Baumaßnahmen. Dies ist auch der Grund, warum für die Erfassung und Auswertung von 3D-Stadtmodellen erst das Interesse bei den entsprechenden Vermessungsverwaltungen geweckt werden muss. Deren Hauptaufgabe ist die Katastermessung, das bedeutet, dass einzelne Gebäude nach einem Neu- oder Umbau gezielt in das Katasterwerk aufgenommen werden. Der

Einsatz von photogrammetrischen Methoden für die Auswertung von Gebäudedaten kann daher nur als Zusatzangebot von den Katasterbehörden gesehen werden. Die Datenerfassung bei Mobilfunkfirmen basiert allerdings ausschließlich auf Überfliegungen. Aus diesem Grund ist eine automatisierte Fortführung des Datenbestandes ein wichtiges Forschungsgebiet.

Ein sehr viel versprechender Ansatz dafür ist es, eine Aktualisierung durch einen automatisierten Vergleich der Rohdaten durchzuführen. Damit muss nicht der gesamte Bestand pauschal ersetzt werden. Das Ergebnis des Vergleichs ist es, Differenzen zwischen dem aktuellen Zustand und dem der letzten Erfassung festzustellen und nur diese Differenzen zu übernehmen. Besonders hier ist eine eindeutige Identifizierung der Objekte unumgänglich um einzelne Objekte problemlos ersetzen zu können. So kann die Nachbearbeitung auf bestimmte Gebiete beschränkt werden.

Aus einem Vergleich der photogrammetrischen Luftbilddatenauswertung und der Auswertung von Laserscanning-Messungen kann zusammenfassend geschlossen werden, dass letztere eine einfachere Ableitung der Geländeoberfläche ermöglichen. Für die Gebäuderekonstruktion sind Punktdichten von 1-2 m einigermaßen ausreichend. Allerdings genügt das nicht, wenn höhere Genauigkeitsanforderungen bestehen. Aus diesem Grund werden Daten aus Laserscanning-Messungen noch nicht in größerem Umfang auf kommerzielle Weise zur Gebäudedatenerfassung eingesetzt. Generell ist die Datenerfassung mit Hilfe von Laserscanning-Messungen etwas teurer als über Luftbilder, da die Abdeckung geringer ist. Dies resultiert in schmaleren Flugstreifen, weshalb die Flugdauer länger ist.

4.3 Photorealistische Erfassung

Bei der photorealistischen Erfassung werden die Fassaden mit Hilfe terrestrischer Photos unter Zuhilfenahme von Techniken der Bildverarbeitung und der Photogrammetrie modelliert. Oft liegt ein einfaches Blockmodell aus der Auswertung von Überfliegsdaten als Basis vor, allerdings kann theoretisch auch die Geometrie soweit ausgewertet werden, wie sie aus den terrestrischen Photos ersichtlich ist. Bei dieser Art der Modellierung entsteht eine sehr große Datenmenge. Wenn das 3D-Modell nicht nur zur Abgabe an Planungsbüros sondern auch für Echtzeitvisualisierungen herangezogen werden soll, muss das gesamte Modell darauf angepasst werden. Aus diesem Grund müssen intelligente Algorithmen eingesetzt werden, um die Anzahl der sichtbaren Dreiecke, aus denen das 3D-Stadtmodell besteht, zu reduzieren. Dazu gehören Algorithmen wie Level of Detail (LOD), bildbasiertes Rendern und die Bestimmung nicht sichtbarer Bereiche.

Grundlage für die Datenerfassung ist das Auffinden korrespondierender Punkte in Bildern, die gleiche Partien des Gebäudes abbilden. Die Punkte werden genutzt, um die relative Orientierung der Bilder zu bestimmen. Dies führt zur 3D-Geometrie der Inhalte der Bilder und die automatische Modellierung ist somit möglich. Die Texturen werden meist aus allen Photos erstellt, die diesen Bereich abbilden. Auf diesem Weg führen Verdeckungen nicht dazu, dass Texturen falsch dargestellt werden. Gleichzeitig kann eine Qualitätskontrolle der Texturen durchgeführt werden.

4.4 Manuelle Modellierung

Eine weitere, allerdings sehr zeitaufwendige und damit teure Methode stellt die Modellierung von Hand dar. Dazu können Ergebnisse terrestrischer Photogrammetrieauswertungen herangezogen werden. Dieses Verfahren kann allerdings nicht flächendeckend eingesetzt werden, sondern nur für einzelne Gebäude oder für Fassaden. Abbildung 13 zeigt einen Ausschnitt aus einem 3D-Stadtmodell der Darmstädter Innenstadt. Grundlage ist ein photogrammetrisch erfasstes 3D-Stadtmodell als Klötzchenmodell mit Dachlandschaft. Für einzelne hervorstechende Gebäude wurden Fassaden sowie beispielhaft die Vegetation in einer größeren Grünfläche manuell auf die bestehende Datengrundlage modelliert.



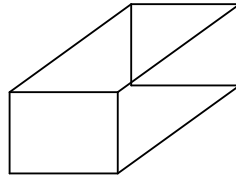
Abbildung 13: Detaillierte Fassaden in einem 3D-Stadtmodell

Wenn externe Bearbeiter diese Daten im entsprechenden Koordinatensystem einbinden können, ist es möglich, die Wünsche aus dem Planungsbereich zumindest partiell zu berücksichtigen. Für externe Übernahmen sind Schnittstellen zur Verfügung zu stellen und Vorgaben zu entwickeln. Damit können entstandene Modelle in ein einfaches 3D-Stadtmodell problemlos integriert werden.

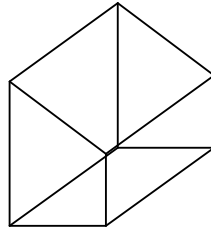
4.5 Beispiele für 3D-Stadtmodelle

Der deutsche Mobilfunkanbieter T-Mobile erfasst Geodaten für die Planung kleinzellulärer Strukturen seiner Mobilfunknetze. Dabei sind alle räumlichen Informationen wichtig, die zwischen einer Mobilfunkstation und einem potentiellen Nutzer liegen. Das bedeutet, dass alle Objekte erfasst werden, die sich ab der durchschnittlichen Höhe über Grund eines Mobiltelefons während der Benutzung befinden. Es werden deshalb nicht nur Gebäude, sondern auch Mauern oder Vegetation erfasst, wenn sie eine bestimmte Höhe überschreiten. Für die Datenerfassung werden die Kernbereiche aller größeren deutschen Städte überflogen und die Luftbilder anschließend photogrammetrisch bearbeitet. Die Gebäude werden als Klötzchen mit bestimmten Dachformen ausgewertet, d.h. es wird zwischen den folgenden Standarddachformen unterschieden:

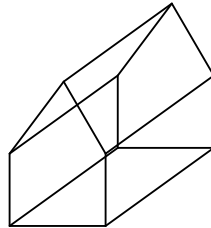
Flachdach



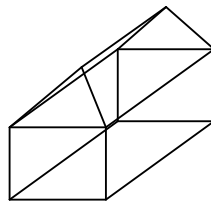
Pulldach



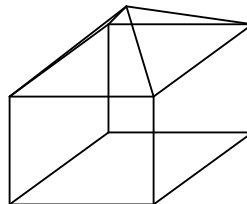
Satteldach



Walmdach



Zeltdach



Bei der Auswertung wird die Gebäudetraufe erfasst und soweit nach unten abgelotet, dass sie mit der Oberfläche eines Geländemodells, das ebenfalls für die Auswertgebiete vorliegt, verschnitten werden kann. Zusätzlich wird je nach Dachform eine unterschiedliche Anzahl an Firstpunkten digitalisiert. Dabei liegt die Lagegenauigkeit für die so erhaltenen Punkte bei unter ± 1.0 m.

Die Höhengenaugigkeit beträgt ± 0.5 m für Gebäude, deren längste Seite kleiner als 5 m ist. Ansonsten beträgt sie 10 % der Gebäudehöhe, allerdings maximal ± 2 m [TMob99].

Zusätzlich zu der geometrischen Information liegen weitere attributive Informationen für die einzelnen Objekte vor. Dazu zählen beispielsweise die Firstrichtung als Azimut sowie die relative und die absolute Gebäudehöhe, wobei die letztere als Höhe über NN bezogen auf die Grundrisspunkte behandelt wird.

Die Stadt Stuttgart ist bereits seit einigen Jahren dabei, ein 3D-Stadtmodell als digitale Information zu den Gebäuden neben der digitalen Stadtkarte und einem digitalen Geländemodell zu erfassen. Dafür wurde eine Befliegung durchgeführt und die entsprechenden Gebiete photogrammetrisch erfasst. Um eine Übereinstimmung mit den bereits vorhandenen Datenbeständen zu gewährleisten, wurden die Höhen der Punkte der Gebäudegrundrisse aus der digitalen Stadtkarte sowohl mit Hilfe des DGMs als auch über die Dachoberfläche in den orientierten Luftbildern berechnet. Auch hier wurden während der Erfassung standardisierte Dachformen herangezogen. Diese Gebäude-Kantenmodelle bilden zusammen mit dem Digitalen Geländemodell das 3D-Stadtmodell [Land98].

Die Stadt Graz in Österreich hat nach Prag die größte zusammenhängende Altstadt Mitteleuropas. Sie ist UNESCO-Welterbe. Hierfür wurde von der TU Graz auf Basis der Daten des Stadtvermessungsamtes ein photorealistisches 3D-Stadtmodell entwickelt [Zach01]. Die Altstadt ist ungefähr 3 km auf 2.5 km groß und besteht aus ungefähr 1700 Gebäuden. Zusätzlich stand ein DHM mit einer Rasterweite von 30 m, das auch Bruchkanten enthält, zur Verfügung. Abbildung 14 zeigt einen Ausschnitt aus diesem Modell, in dem Dachformen enthalten sind. Es wurde das DHM sowie Textur integriert.



Abbildung 14: 3D-Modell mit Textur auf DHM und Dächern [VRVI03]

Die geometrische Auflösung des 3D-Stadtmodells soll etwa 30 cm betragen, die Textur-Auflösung der Fassaden liegt bei 2 auf 2 cm. Dies führt insgesamt zu einem Datenvolumen von 260 Gigabyte. Das zugrunde liegende Blockmodell besteht aus Trauf- und Firstrinien. Abbildung 15 zeigt dieses Blockmodell für den Innenstadtbereich, das auf das DHM aufgebracht wurde.

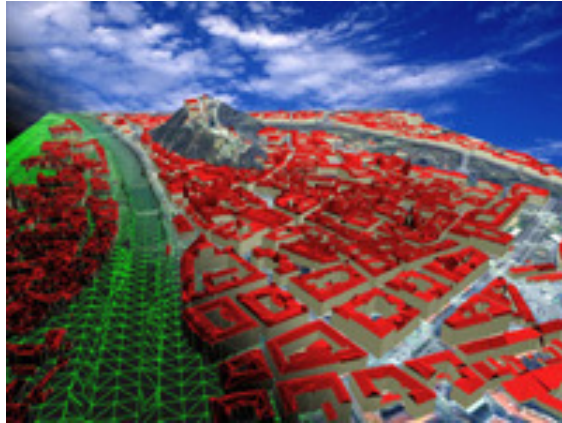


Abbildung 15: Zugrundeliegendes Blockmodell der Innenstadt von Graz [VRVI03]

Für die Reduzierung der Datenmenge bei Echtzeitvisualisierungen wurde ein LOD-Algorithmus an die Daten angepasst. Dieser geht davon aus, dass Objekte, die weiter entfernt von der aktuellen Blickposition sind, nicht so detailliert dargestellt werden müssen. Um eine gleichmäßige Bildrate zu erreichen, sorgt das System dafür, dass die Anzahl der Dreiecke gleich bleibt.

In Berlin existiert ein sehr detailliertes und durch regelmäßige Begehung verhältnismäßig aktuelles 3D-Modell der Innenstadt (vgl. Abbildung 16).



Abbildung 16: Ausschnitt aus dem 3D-Modell der Berliner Architekturwerkstatt [SenB03]

Es wurde manuell erstellt und hat im Moment keinen Bezug zu den amtlichen Daten. Das bedeutet, dass die Grundrisse in den Katasterwerken des Senats nicht mit den Grundrissen dieses 3D-Modells übereinstimmen. Noch etwas aufwendiger gestaltete sich die Modellierung der 3D-Gebäude der Wirtschaftsförderung Berlin. Abbildung 17 zeigt einen Ausschnitt daraus.



Abbildung 17: 3D-Modell der Wirtschaftsförderung Berlin [Wirt01]

Es wurden bekannte Gebäude wie der Reichstag oder der Fernsehturm sehr detailgetreu modelliert, Straßen und Wege sind sichtbar und auch Vegetation wurde in Form einzelner Bäume modelliert. Das 3D-Modell wird zusammen mit Branchendaten, Immobilienangeboten, Arbeits- und Strukturinformationen, Flächenplänen und Förderprogrammen zur Beratung von zukünftigen Investoren präsentiert. Aus diesem Grund steht die Anschaulichkeit im Mittelpunkt. Dies führt zur Visualisierung von zukünftigen Innenräumen, wie z.B. dem Lehrter Stadtbahnhof. Ein Ausschnitt daraus ist in Abbildung 18 dargestellt.

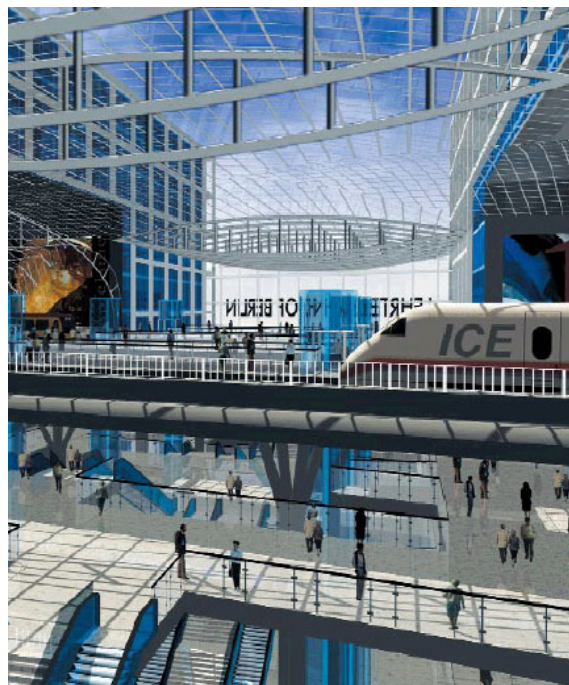


Abbildung 18: Innenansichten des 3D-Modells der Wirtschaftsförderung Berlin [Wirt01]

4.6 Zusammenfassung

Einfache 3D-Stadtmodelle existieren bereits fast flächendeckend für alle größeren deutschen Städte. Sie sind über Laserscanning-Messungen oder mit Hilfe herkömmlicher photogrammetrischer Methoden erfasst worden und liegen meist georeferenziert vor. Das bedeutet, dass sie in das übergeordnete Koordinatensystem der Landesvermessung eingepasst sind. Einzelne Bereiche können dann von verschiedenen Stellen für spezielle Nutzungen herangezogen und detaillierter modelliert werden. Diese detaillierten Modelle benötigen beispielsweise die Bereiche Tourismus und Stadtmarketing. Eine Grundlage dafür könnten die teilweise mehrfach von verschiedenen Mobilfunkfirmen ermittelten Geodaten bilden. Zusätzliche Daten können durch Umrechnungen von den dort eingesetzten Koordinatensystemen in das Landesvermessungs-System über mathematische Gleichungen integriert werden.

5 Der mobile Anwender im urbanen Umfeld

In Kapitel 3 wurden Systeme für die Bestimmung von Position und Blickrichtung eines Anwenders im Außenbereich beschrieben, die alle Insellösungen darstellen. Sie bieten für eine bestimmte Aufgabe und unter bestimmten Voraussetzungen die Möglichkeit eine Person sehr genau zu tracken. Das bedeutet, dass für die Umgebung, in der das System eingesetzt werden soll, bestimmte Daten wie beispielsweise Höhenmodelle, Referenzbilder oder 3D-GIS Informationen vorhanden sein müssen. Dieser Ansatz lässt allerdings außer Acht, dass es oftmals ausreicht, nur ungefähr zu wissen, wo sich eine Person befindet. Damit können Rechenzeit und -leistung gespart werden [Schu03].

Zusätzlich zur oft ausreichenden Genauigkeit von Trackingtechniken für bestimmte Aufgaben kommt dazu, dass der Anwender in vielen Situationen gar nicht möchte, dass die derzeitige Position festgehalten und eventuell sogar anderen Personen zur Verfügung gestellt wird. Feiner weist darauf hin, dass man als Nutzer einer Komponente, die die derzeitige Position trackt, genaue Informationen über den derzeitigen Standort nur einer ausgewählten Gruppe an Freunden zugänglich machen möchte [Fein99]. Zurzeit werden Video- und Audioaufnahme nur von einzelnen Personen oder an einzelnen Plätzen vorgenommen. Feiner bemerkt, dass es in der Zukunft üblich sein wird, dass nicht nur stationär sondern von verschiedenen Stellen und Personen solche Informationen gespeichert werden, gerade in Hinblick auf den Einsatz videobasierter Trackingverfahren. Wenn dies als störend empfunden wird, sollte es demnach möglich sein, auf höchste Genauigkeiten zu verzichten, sofern diese nur mit einem videobasierten Verfahren möglich sind, und dafür die Privatsphäre der Mitmenschen zu achten.

Um ein flächendeckendes mobiles Informationssystem in einer Innenstadt anbieten zu können, kann die AR-Anwendung mit anderen mobilen Diensten gekoppelt werden. Letztere können auch an Stellen in Anspruch genommen werden, an denen das hochgenaue AR-Tracking nicht eingesetzt werden kann. Hinzu kommt, dass die in Kapitel 2 vorgestellten Trackingtechniken unter bestimmten Bedingungen arbeiten und zeitweise außer Betrieb gesetzt werden können. Im Fall von GPS können beispielsweise keine Satellitensignale in engen Straßen empfangen werden. Das bedeutet, dass an diesen Stellen ebenfalls keine Positionsbestimmung mit GPS möglich ist. Es müssen deshalb gegebenenfalls weitere Trackingtechniken in das System integriert werden.

Inhalt dieses Kapitels 5 ist die Entwicklung eines anforderungsadaptiven Zuordnungsverfahrens, das es erlaubt, einer Person, die sich zu Fuß in einer Stadt bewegt, verschiedene Dienste anzubieten. Es wird davon ausgegangen, dass die Präsentation von hochgenauen AR-Informationen nur an bestimmten Stellen möglich ist, andere Dienste aber überall angefordert werden können. Außerdem können verschiedene Trackingtechniken integriert werden. Abbildung 20 zeigt den Aufbau dieses Zuordnungsverfahrens.

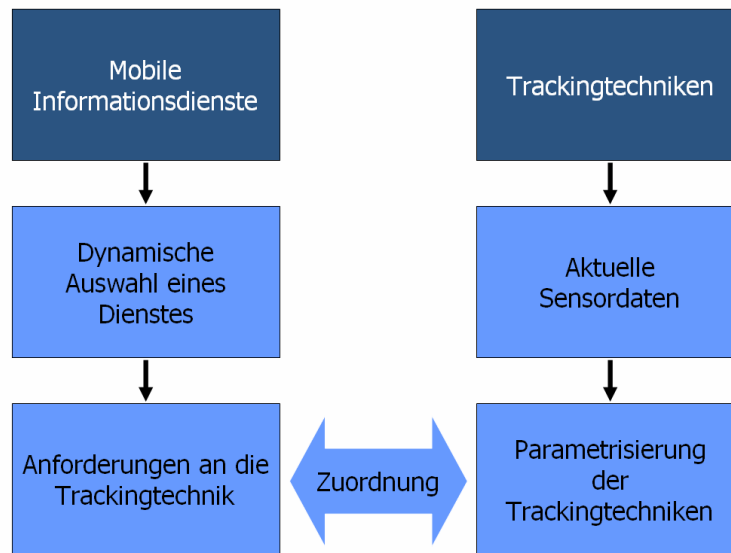


Abbildung 19: Anforderungsadaptive Zuordnung

Die zur Verfügung stehenden Dienste werden angeboten, nachdem das Verfahren festgestellt hat, welche Dienste für den Benutzer möglich sind und welche von der Umgebung her anwendbar sind. Aus dieser Auswahl entscheidet sich der Nutzer für eine Anwendung oder das System bietet ihm automatisch eine an. Diese dynamische Auswahl eines Dienstes aus verschiedenen mobilen Informationsdiensten führt zur Aufstellung von Anforderungen an die möglichen Trackingtechniken. Für diese Trackingtechniken wiederum müssen die aktuellen Daten berücksichtigt werden, die zu einer Aufstellung der Parameter führen. Diese gehen zusammen mit den Anforderungen in das Zuordnungsverfahren ein. Ziel dieser Zuordnung und damit des Verfahrens ist es, je nach Anforderung eine geeignete Kombination an Sensoren und Trackingtechniken anzusprechen. Die Zuordnung reagiert anforderungsadaptiv, d.h. alle zur Verfügung stehenden Einflüsse werden berücksichtigt. Außerdem bietet sie die entsprechende Technik automatisch an.

Die Auswahl der Parameter ist generisch, so dass jede weitere Trackingtechnik mit Hilfe dieser Parameter beschrieben werden kann. Es wird davon ausgegangen, dass LBS, die bereits zum jetzigen Zeitpunkt einer großen Zahl von Kunden zur Verfügung stehen, auch in Zukunft nachgefragt werden. Damit werden die Aufgaben steigen, für die die Information zumindest über die Position des Nutzers entscheidend ist. Aus diesem Grund ist das Verfahren so flexibel, dass weitere Anforderungen problemlos in die Methode integriert werden können.

In diesem Kapitel werden in Abschnitt 5.1 zunächst Anwendungsszenarien aufgezeigt. Anschließend wird in Abschnitt 5.2 erläutert, welche Faktoren die Auswahl eines bestimmten Dienstes beeinflussen. Dazu werden Angaben über den Nutzer, aber auch Einflüsse in der Umgebung, näher betrachtet. In Abschnitt 5.3 wird aufgezeigt, wie mit Hilfe einer aus den Anwendungsszenarien abgeleiteten Klassifikation und der Aufstellung von Parametern Anforderungen der Dienste an die möglichen Trackingtechniken formuliert werden können. Der Begriff Trackingtechniken umfasst sowohl die Bestimmung von Position und Orientierung anhand verschiedener Techniken wie sie in Kapitel 2 vorgestellt wurden als auch von Lösungen des AR-Problems im Außenbereich, die Entwicklungen an Forschungsinstituten darstellen, wo umfangreiche Verfahren entwickelt wurden, die beispielsweise auf der Auswertung von Bildern einer Kamera basieren. Diese Verfahren

wurden in Kapitel 3 erläutert. Damit sind die Eingangsdaten in das Zuordnungsverfahren, die Dienste betreffend, formuliert. Die Informationen der Tracking-techniken müssen zunächst auf die zurzeit möglichen Sensoren reduziert und daraus die Rohdaten ausgelesen werden. Dies ist in Abschnitt 5.4 näher erläutert. Die Eingangsgrößen der Trackingsensoren, die in das Zuordnungsverfahren eingehen, werden in Abschnitt 5.5 beschrieben und sind in Abschnitt 5.6 für mögliche Trackingtechniken beispielhaft aufgeführt. Abschnitt 5.7, die Zuordnung der Trackingtechnik zu einem ausgewählten Dienst, bildet vor der Zusammenfassung in Abschnitt 5.8 den Abschluss dieses Kapitels.

5.1 Anwendungsszenarien

Wie in den Kapiteln 2 und 3 gezeigt, gibt es verschiedene Möglichkeiten, die Position einer Person zu bestimmen. Diese Trackingtechniken müssen mit dem ausgebenden Gerät verbunden oder das AR-System muss auf diesem installiert sein, damit eine orts- und vielleicht sogar blickrichtungsabhängige Informationsdarstellung möglich ist. Im Folgenden werden nun Szenarien aufgezeigt, die darstellen, welche Informationen auf mobilen Endgeräten angeboten werden können. Einige der Informationen könnten auch zu Hause auf dem PC im Internet abgerufen werden, wenn der Nutzer die geographische Position eingegeben hat. Andere sind nur vor Ort vermittelbar, da der Einfluss von Umgebungsfaktoren oder die tatsächliche Realität vor Ort unabdingbar sind. Die Informationsvermittlung vor Ort hat damit viele Informationen oder Informationstechniken erst möglich gemacht. In den beiden folgenden Abschnitten 5.1.1 und 5.1.2 werden Anwendungen aus den Bereichen LBS bzw. AR aufgezeigt, die einem Anwender im urbanen Umfeld angeboten werden können.

5.1.1 LBS für den Fußgänger

Mobiltelefone und PDAs bieten die Möglichkeit, auf einem verhältnismäßig kleinen Display Informationen anzuzeigen, die wenig Rechenleistung benötigen. Diese Geräte haben den Vorteil, dass sie sehr klein sind und damit überallhin mitgenommen werden können. PDAs bieten Steckplätze an, die das Gerät etwa mit erweitertem Speicher aber auch einer GSM- oder GPS-Karte ergänzen.

Im Fall von begrenzten Gebieten wie beispielsweise Museen werden heutzutage oft Führungen über eine tragbare Audio-Ausgabe angeboten. Der Vorteil von mobilen Endgeräten wie PDAs mit Multimedia-Ausgaben gegenüber diesen herkömmlichen Audio-Führern ist, dass die Dienste sich nicht auf die Audioausgabe beschränken müssen und verschiedene Sensoren angeschlossen werden können. Dadurch ist diese Art der Informationsvermittlung nicht mehr auf ein abgeschlossenes Gebiet beschränkt, in dem die Ortsbezogenheit oft mit Hilfe von Infrarot-Marken hergestellt wird, sondern der Einsatz als individuelle Stadtführung ist damit möglich. Die Dienste

sind dann nicht nur weiträumiger einsatzbereit, sondern sie können sowohl bezogen auf den Inhalt als auch bezogen auf die Art der Ausgabe auf persönliche Vorzüge des Nutzers zugeschnitten werden. Diese Art der Darstellung ist gegenüber reinen Audioausgaben ansprechender und ermöglicht ein vollständigeres Erlebnis. [Bell02]

Die für PDAs oder Mobiltelefone konzipierten LBS werden meist für die Bereiche Tourismus und Reisen entwickelt. Die Dienste sollen den Reisenden oder Touristen mit aktuellen Informationen versorgen, wobei diese sowohl einfach zu lesen als auch einfach zu verstehen sein sollen [Coor03]. Die am weitesten verbreiteten Informationen sind Antworten auf die Frage nach der derzeitigen Position, die in einer 2D-Karte dargestellt werden können, oder auch nach interessanten Orten in der näheren Umgebung, wie Restaurants, Geldautomaten oder der nächsten Haltestelle. Das Anzeigen der derzeitigen Position und interessanter Stellen in der Umgebung wird oft durch ein Routing erweitert. Dieses wird meist als Information mit Hilfe einer Karte oder als Textausgaben realisiert.

Auch 3D-Informationen werden in Zukunft auf mobilen Endgeräten eine immer größere Rolle spielen. Heutzutage ist es mit 3D-Grafikkarten in PDAs möglich, bis zu 100.000 Dreiecke zu rendern. Die Entwicklung der Grafikmöglichkeiten wird einen ähnlichen Verlauf nehmen wie sie auf PCs und Laptops bereits begonnen hat. Mit UMTS wird es möglich sein, bis zu 384 Kbits/s in Städten zu übertragen. Zusammen mit 3D-Kompressionstechniken werden 3D-Grafiken in einem kabellosen Netzwerk auch auf mobilen Geräten möglich sein. [Coor01] Die Reduktion der Übertragungszeiten von 3D-Modellen wird ein 3D-Routing auf einem PDA, der die Informationen von einem Server zur Verfügung gestellt bekommt, erlauben. Allerdings ist der Informationsgehalt von vollständigen 3D-Welten für eine Routingaufgabe nicht immer die ideale Lösung. Eine 2D-Karte angereichert mit 3D-Gebäuden bietet meist mehr Information, die vor allem schneller ablesbar ist. Damit kann sich eine Person anhand herausragender Objekte in der Umgebung orientieren.

Des Weiteren existieren Dienste, die es erlauben, mit Freunden in Kontakt zu bleiben um sich zu verabreden, aber auch über das Netz miteinander zu spielen. Wenn die Position anderer Personen auf dem Display angezeigt werden soll, muss nicht nur die Position des Anwenders, sondern auch die der anderen registriert und übertragen werden.

Ein weiteres Einsatzgebiet von LBS sind die Bereiche Verkehr und Transport. Hierbei handelt es sich allerdings meist um in Fahrzeugen eingebaute Systeme. Dazu zählen Notrufdienste, ortsabhängige Verkehrsnachrichten, die entweder angefordert oder automatisch verschickt werden, Navigationsunterstützung in Fahrzeugen aber auch das Flottenmanagement. Diese gehen über die ausgeführten Betrachtungen von Diensten für Fußgänger hinaus.

LBS ermöglichen es, ausgehend von der derzeitigen Position des Anwenders, Informationen in Echtzeit zusammengestellt zu bekommen. Meist können sie auf verschiedene Eingaben des Nutzers reagieren. Grundvoraussetzung für diese Dienste ist die Kenntnis der Position des Nutzers, die automatisch erfasst werden muss.

5.1.2 AR im urbanen Umfeld

Wenn ein Laptop als mobiles Gerät eingesetzt wird, kann ein reichhaltigeres Informationsangebot präsentiert werden als beispielsweise auf einem PDA. Diese Geräte ermöglichen eine dynamische und oft auch adaptive Informations-

präsentation. Allerdings ist das umständliche Installieren von Software der Hauptgrund, warum solche Dienste nicht genutzt werden. Hier muss auf Leihgeräte zugegriffen werden, die die gewünschte Funktionalität bieten.

Gleichzeitig können weitere Informationen zur Umgebung mit in die Generierung der Information einfließen. So muss davon ausgegangen werden, dass sich Personen vor Ort bewegen und nicht allzu lange an einer Stelle stehen bleiben. Wenn beispielsweise Erklärungen zu Objekten in der Realität präsentiert werden, sollte der Benutzer diese möglichst zügig verstehen ohne nachfragen zu müssen. Das Abfragen von zusätzlicher Information sollte nicht nur möglich sein, sondern vom System automatisch unterstützt werden. Auf kleinen mobilen Geräten wie einem Mobiltelefon oder einem PDA ist ein solcher Automatismus technisch nicht möglich. Das bedeutet, dass es mit zusätzlichem manuellem Aufwand verbunden ist und stellt daher eine Form von Benutzerunfreundlichkeit dar. Auf einem PC ist eine personalisierte Benutzerführung möglich, die ebenfalls auf Einflüsse aus der Umgebung reagieren kann.

Als mobile PCs werden Laptops aber auch so genannte Wearable Computer bezeichnet. Letztere können beispielsweise an der Hüfte mit einem Gürtel befestigt werden und sind vollständig darauf ausgerichtet, einen Anwender zu unterstützen, der sich bewegt, während der Rechner in Betrieb ist. Er unterstützt den Anwender, indem er ihm beispielsweise zeigt, wie eine bestimmte Maschine zu reparieren ist. Auch in AR-Umgebungen sind diese Geräte verbreitet. Für die Anwendung dort sind die Rechner mit einem am Kopf tragbaren Display verbunden, so dass die Bildschirmausgabe direkt im Sichtfeld erkennbar ist. Interaktion ist meist mit einer am Arm befestigten Tastatur und einer Maus, die keine glatte Unterfläche benötigt, möglich. Der Nachteil dieser Geräte ist, dass sie nicht so leistungsstark sind wie die neueste Laptopgeneration. Aus diesem Grund sind sie, wie in Abschnitt 3.8 erwähnt, für den Einsatz in manchen AR-Systemen nicht ausreichend. Deshalb werden gerade in Forschungsinstituten eher Laptops für mobile Anwendungen eingesetzt, die einen PC für die Berechnungen benötigen. Diese werden meist in einem Rucksack getragen und sind an Sensoren und an ein tragbares Display angeschlossen. Die Hauptprobleme der Hardware, die für diese Art von mobilen Anwendungen eingesetzt werden, sind zurzeit die Größe, das Gewicht, die Stromversorgung, aber auch die Qualität der Displays [Azum99b]. AR ist eines der Entwicklungsrichtungen für mobile Systeme im Außenbereich.

Haupteinsatzgebiet für AR-Anwendungen ist der Tourismus. Es gibt allerdings weitere große Bereiche wie die Instandsetzung von Maschinen, Notfalldienste, Navigation sowie der militärische Bereich und die Unterhaltungsindustrie, für die AR-Systeme entwickelt werden. Dabei werden Objekte, die sich im Blickfeld des Anwenders befinden, automatisch vom System identifiziert. Im Blickfeld des Benutzers können dann die Bezeichnungen als virtuelle Informationen eingeblendet werden. Wie in Kapitel 3 gezeigt, können die Objekte Berggipfel aber auch Gebäude beispielsweise auf einem Universitätscampus sein. Pfeile helfen bei der Orientierung, wenn Informationen zu nicht sichtbaren Objekten angefragt werden oder eine Benutzernavigation möglich ist. Wie ebenfalls in Kapitel 3 gezeigt, ist eine weitere Anwendung das Einblenden von positionsgenauen 3D-Rekonstruktionen in das Sichtfeld des Anwenders. Diese Aufgabe erfordert eine sehr genaue Detektion der Position und Blickrichtung des Anwenders. Diese Art der Informationsvermittlung kann dazu dienen, Touristen zu zeigen, wie eine Umgebung in der Vergangenheit aussah. In der Stadtplanung bietet sich die Möglichkeit, geplante Änderungen einer

Umgebung Experten aber auch Bürgern besser zu vermitteln. Der Vorteil dieser Art der Darstellung ist es, dass die momentane Realität und die Dimensionen dem Anwender als Basis zur Verfügung stehen. Die Veränderungen der Umgebung sind direkt mit der jetzigen Situation verbunden.

Auch diese leistungsstärkeren Systeme, die auf einem Laptop eingesetzt werden, bieten die Möglichkeit, Navigationsinformationen anzubieten. In diesen werden mit hoher Genauigkeit Position und Blickrichtung des Anwenders bestimmt. Damit kann eine Karte in Blickrichtung ausgerichtet werden. Für ein 3D-Routing sind die Trackingergebnisse bei den meisten AR-Verfahren mit der erforderlichen Genauigkeit vorhanden.

5.2 Dynamische Auswahl eines Dienstes

Ausgangspunkt der anforderungsadaptiven Zuordnung ist ein Angebot an mobilen Informationsdiensten, das zur Verfügung steht. Dazu zählen Dienste, die, wie in Abschnitt 5.1.1 gezeigt, auf einem Mobiltelefon oder PDA, angeboten werden können, aber auch Dienste aus dem AR-Bereich, für die in Abschnitt 5.1.2 Beispiele aufgeführt wurden. Aus dem Angebot wird dem Anwender ein einzelner Dienst oder eine Untermenge dynamisch angeboten. Abbildung 20 zeigt diesen Schritt innerhalb des Ablaufschemas des Zuordnungsverfahrens.

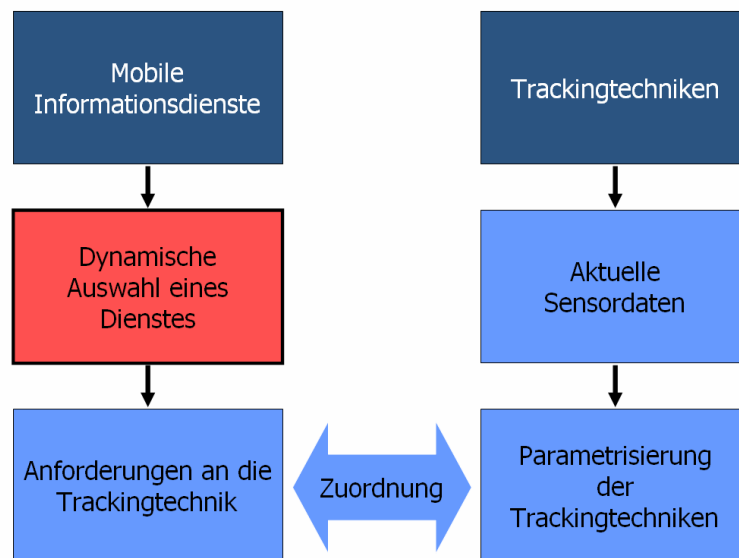


Abbildung 20: Dynamische Auswahl im Ablaufschema des Zuordnungsverfahrens

Die Einschränkung oder Spezialisierung dieser Dienste erfolgt, wie im Folgenden dargestellt, durch persönliche Einstellungen und durch Informationen über die Umgebung. Dadurch sind sie sowohl auf den Anwender als auch auf die Umgebung zugeschnitten.

Zu dieser Einschränkung zählt der Ausschluss einer bestimmten Art der Präsentation oder einer bestimmten Aufgabe oder genau gegensätzlich die Bevorzugung einer Form. Diese persönlichen Informationen werden über Voreinstellungen, Präferenzen und ein Benutzerprofil berücksichtigt und werden in Abschnitt 5.2.1 erläutert. Den

gleichen Einfluss können Informationen zur Umgebung ausüben. Diese hängen davon ab, welche zusätzlichen Sensoren mit dem Verfahren verbunden sind und werden in Abschnitt 5.2.2 dargestellt. Damit ergeben sich die in Abbildung 21 dargestellten Einflüsse auf die Auswahl des Dienstes.

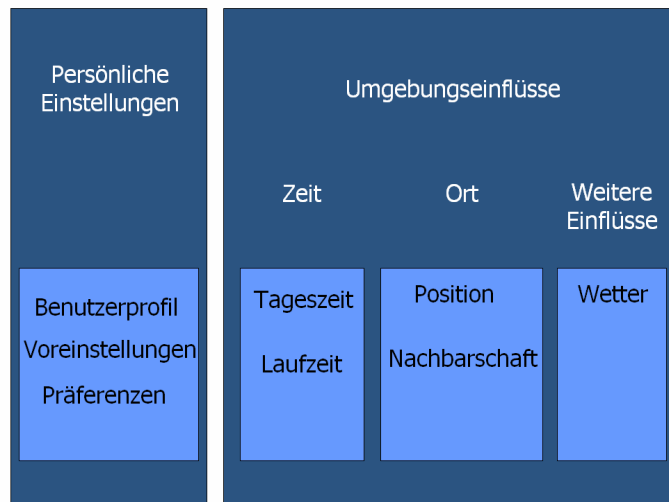


Abbildung 21: Einflüsse auf die möglichen Informationssdienste

5.2.1 Persönliche Einstellungen

Geräte, die auf Einstellungen und Reaktionen des Nutzers reagieren, sollen ihm möglichst intuitiv helfen. Dies bedeutet, dass das Gerät über den Nutzer möglichst viele Informationen benötigt. Dies führt zu einer Personalisierung der Anwendung.

Informationen zur Person des Anwenders können in einem Benutzerprofil gespeichert werden. Bei älteren Menschen oder Behinderten kann mit einer solchen Personalisierungskomponente dafür gesorgt werden, dass sie besondere Strecken geleitet werden, in denen sie möglichst einfach zu ihrem Ziel gelangen. Innerhalb der Anwendungen kann der Inhalt der Information auf den Anwender zugeschnitten werden. Ein Beispiel ist die Präsentation historischer Informationen. Hierbei ist das Alter des Anwenders eine wichtige Information, aus der hervorgeht, in welcher Form welche Fakten präsentiert werden sollten. Um die Angabe der Information über das Alter des Anwenders etwas weniger persönlich zu gestalten, ist auch eine Unterscheidung in die Kategorien Kinder, Jugendliche und Erwachsene möglich. In einem Benutzerprofil kann eingestellt werden, ob man eine ausführliche Stadtführung über das mobile Endgerät angeboten bekommen möchte, und demnach an jedem interessanten Platz, an dem man vorbeikommt, Erklärungen vermittelt bekommt, oder ob man nur Dienste nutzen möchte, die wenig Zeit benötigen. Das Benutzerprofil enthält damit Angaben, die der Anwender bereit ist, preiszugeben. Das bedeutet auch, dass er sich darüber im Klaren ist, dass eventuell andere Personen Zugriff auf diese Daten, z.B. die aktuelle Position, haben. Solch ein Eingriff in die Privatsphäre des Anwenders sollte deshalb so realisiert werden, dass diese Einstellung für den Anwender sichtbar ist.

Jede Form einer Benutzungsoberfläche, die über einen längeren Zeitraum hinweg einem Anwender zur Verfügung steht und von diesem genutzt wird, löst das Problem der Anpassung an den Anwender über Voreinstellungen. Eine Alternative zu Voreinstellungen sind Reaktionen des Systems auf häufig hintereinander ausgeführte Befehle in Form von Nachfragen, ob Grundeinstellungen automatisch vom System geändert werden sollen. Letztere führen nicht bei jedem Nutzer zu mehr Sympathie gegenüber dem Gerät. Oft wird erst bei solch einer Abfrage klar, dass das System Verhaltensweisen registriert und darauf reagiert. All dies geschieht automatisch, ohne Eingriff des Nutzers. Das hat meist keine großen Konsequenzen bei Software, die in einer Desktop-Umgebung angewendet wird, oder die Informationen speichert, die nur dem System und dem Benutzer angezeigt werden. Wenn sich der Benutzer aber in einem Netz befindet, beispielsweise in einem WLAN, weiß er nicht immer, welche Informationen nach außen weitergegeben werden. Dies gilt besonders für Anwendungen, die kein Fachwissen erfordern.

Ein Anwender, der an dem eigenen Laptop mit einer Software arbeitet, mit der er sehr gut vertraut ist, fühlt sich meist imstande, das Programm derart einzustellen, dass es wirklich auf die eigene Arbeitsweise so reagiert, wie er es erwartet. Er hat automatische Abfragen entweder deaktiviert oder sieht diese wirklich als Unterstützung an. LBS sind dagegen meist keine Fachanwendungen sondern sind so konzipiert, dass sie möglichst wenig Einarbeitung bedürfen. Dadurch sind dem Anwender aber auch keine Hintergründe vermittelt worden. Er kann meist nur die Oberfläche bedienen. Für die Bedienung der über LBS angebotenen Dienste reicht dieses eingeschränkt Wissen meist aus. Im Fall von größeren Anwendungen, wie beispielsweise im AR-Bereich, steht die intuitive Interaktion mit dem System im Vordergrund. Die Geräte hierfür müssen ausgeliehen werden. Aus diesem Grund sind Voreinstellungen bei der Übergabe des Gerätes vorzunehmen und sollten anschließend nicht mehr manuell verändert werden können. Dadurch wird dem Nutzer deutlich, welche Informationen er dem System bekannt gibt.

Zusätzlich können Präferenzen angegeben werden. Diese ermöglichen es dem System bei der Auswahl verschiedener Alternativen nur diejenigen anzubieten oder bei der Entscheidung zu berücksichtigen, die der Benutzer zuvor angegeben hat. Experten, die mit geographischen Karten sehr gut umgehen können, wünschen eventuell keine Ausrichtung der Karte auf Grund der momentanen Blickrichtung, sondern möchten jeden Kartenausschnitt so ausgerichtet präsentiert bekommen, dass sich am oberen Rand der Karte Norden befindet. Außerdem finden sich diese Personen oft besser in der Umgebung zurecht, wenn nur ab und zu eine Information über den derzeitigen Standpunkt in der Karte mit einem einfachen Symbol eingeblendet wird als wenn ihnen beispielsweise ein 3D-Routing angeboten wird. Die Verbindung zwischen virtueller Information zur tatsächlichen Umgebung wird von diesen Anwendern automatisch durch Erfahrung hergestellt. Andere Nutzer möchten, dass diese Aufgabe das System übernimmt. In diesem Fall wird dies aus den Informationen über die derzeitige Position und Blickrichtung des Anwenders hergestellt und dementsprechend in einem tragbaren Display angeboten. Virtuelle Information und Realität überlappen sich bei hochgenauem Tracking. Wenn 3D-Informationen auf einem PDA angeboten werden, muss der Anwender die Verbindung noch selbst herstellen. Dies kann durch einzelne detailliert modellierte herausragende Gebäude erleichtert werden. Im Falle von AR-Systemen wird diese Beziehung automatisch hergestellt, indem die Informationen die Realität positionsgenau überlagern.

Damit sind die folgenden Informationen wichtig, die Einfluss auf das Angebot an Diensten haben:

- Benutzerprofil
- Voreinstellungen
- Präferenzen

5.2.2 Einflüsse der Umgebung

Im Fall von virtuellen Informationsausgaben zu Objekten in der Realität müssen die wahren Objekte weiterhin im Vordergrund bleiben, die Erklärungen sollen nur als Unterstützung zum Verständnis dienen. Sie wecken Interesse oder erweitern den Blick, aber die Informationen dürfen das Objekt an sich nicht überdecken. Aus diesem Grund muss darauf geachtet werden, dass die Erfahrungen, die der Benutzer ohne das Gerät machen würde, möglichst aufgefangen werden und in die Informationen mit einfließen. Diese Informationen betreffen die Art der Informationsaufnahme durch den Nutzer und können durch die Beobachtung der Umgebung gesteuert werden. Dafür sind meist verschiedene Sensoren notwendig. Die Auswertung von Ort und Zeit ermöglicht es dem System beispielsweise, objektive Kriterien abzuleiten, die darauf hinweisen, mit welchen Inhalten sich der Anwender im Moment beschäftigt. Damit kann automatisch dafür gesorgt werden, dass das System noch besser auf die Bedürfnisse des Anwenders eingeht. Die Tageszeit kann genutzt werden, um Informationen über Öffnungszeiten auf die derzeitige Situation anzupassen. Außerdem kann das System tageszeitabhängige Empfehlungen für die Freizeitgestaltung einbringen, einen Zoobesuch, eine Einkaufstour oder einen bestimmten Film in einem Kino vorschlagen. Auch die Laufzeit ist eine Information, die genutzt werden kann.

Wenn Informationen über das Wetter vorliegen, kann das System weitere Einschränkungen bei Vorschlägen berücksichtigen oder eine Information wegen Regens an einer Position abbrechen und an einer trockeneren Stelle in abgewandelter Form anbieten.

Weitere Informationen können durch die Auswertung von Bildern einer Videokamera ermittelt werden. Hiermit können nicht nur statische sondern besonders dynamische Informationen zur Umgebung bestimmt werden. Der Nachteil dieser Technik ist, dass sie meist sehr rechenaufwändig ist.

Die Position kann nicht nur als Ergebnis verwaltet werden, das an eine Komponente weitergegeben wird, sondern sie kann auch als Möglichkeit gesehen werden, bestimmte Dienste auszuwählen, die an diesem Ort besonders angebracht sind oder über die Information Ort weitere Daten zur Umgebung abzufragen. Dies setzt voraus, dass Wissen zu diesen Positionen gespeichert ist, das dann abgefragt werden kann.

5.3 Anforderungen an die Trackingtechnik

Die in Abschnitt 5.1 aufgeführten Anwendungsszenarien gehen in die Klassifizierung ein, die im folgenden Abschnitt 5.3.1 hergeleitet wird. Sie ermöglicht die Ableitung von Parametern und die Beschreibung der Anforderungen (Abschnitt 5.3.2). In Abschnitt 5.3.3 werden die Anforderungen beispielhaft für Anwendungen im urbanen Umfeld parametrisiert. Abbildung 22 zeigt, dass diese Anforderungen an die Trackingtechnik in die Bestimmung der Zuordnung eingehen.

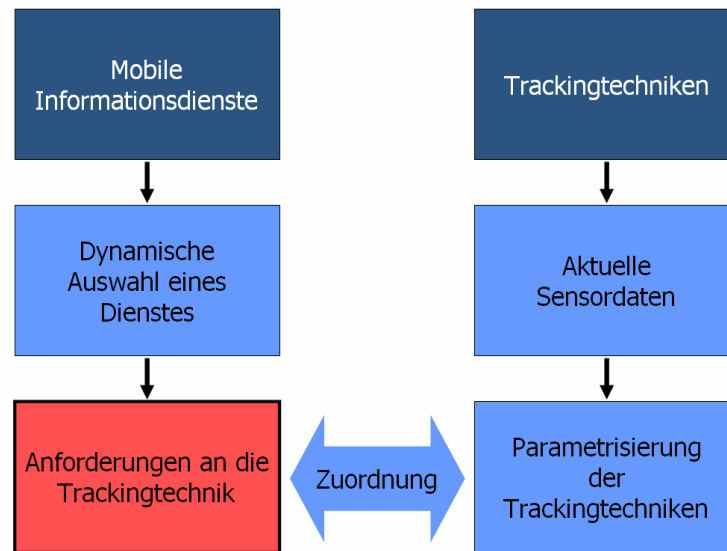


Abbildung 22: Bestimmung der Anforderungen an die Trackingtechnik im Ablaufschema des Zuordnungsverfahrens

5.3.1 Klassifizierung der Dienste

Für die Herleitung der Anforderungen ist es notwendig, dass die in den Anwendungsbeispielen in Abschnitt 5.1 aufgezeigten Dienste generisch zusammengefasst werden, um eine Klassifizierung vornehmen zu können. Das grundsätzliche Unterscheidungsmerkmal der aufgezeigten Anwendungsbeispiele ist es, ob Informationen in das Blickfeld des Anwenders eingeblendet werden sollen. Dann handelt es sich um eine AR-Aufgabe, wofür die Bestimmung von Position und Blickrichtung notwendig ist. Es kann aber auch ausreichen, die derzeitige Position des Anwenders zu bestimmen, um ihm Informationen zur Umgebung anbieten zu können. In diesem Fall kann der Informationsdienst auf einem herkömmlichen Display präsentiert werden. Aus diesem Grund wird zunächst unterschieden, ob für den Dienst ein transparentes Display notwendig ist, da für viele Aufgaben die genaue Bestimmung der Blickrichtung nicht entscheidend ist.

Die reine Information, die an den Anwender weitergegeben wird, kann zu verschiedenen Präsentationen führen. Eine Charakterisierung der Zielsetzung ermöglicht es aber, grundsätzlich festzustellen, welche Anforderungen abgeleitet werden müssen, wenn der Anwender eine bestimmte Information benötigt. Unter der Voraussetzung schneller Trackingergebnisse und dem Vorhandensein einer Karte, die der Nutzer für andere Aufgaben bereits angewendet hat, reicht es ihm oftmals aus, wenn er sich anzeigen lassen kann, wo er sich selber befindet, wie weitere

Objekte in Bezug auf seine derzeitige Position lokalisiert sind und wie er dort hinkommt. Diese Aufgaben können als Navigation betrachtet werden. Damit kann zunächst grundsätzlich unterschieden werden, ob der Anwender Unterstützung bei der Navigation benötigt oder ob er Erklärungen zu einzelnen Objekten in der Umgebung angeboten bekommen möchte. Für beide Aufgaben können grundsätzlich sowohl ein herkömmliches Display als auch ein tragbares Display verwendet werden, das sich der Anwender vor die Augen hält. Um Erklärungen zu Objekten in der Umgebung anzubieten, kann eine Karte eingeblendet oder Bilder mit Erklärungen dargestellt werden. Im Fall eines AR-Systems können virtuelle Information direkt in das Sichtfeld eingeblendet werden, die abhängig von der realen Umgebung sind. Das bedeutet meist, dass auf Objekte in der Umgebung in Form von Pfeilen, Beschriftungen oder durch Überblenden von 3D-Modellen hingewiesen wird.

In der Übersicht in Abbildung 23 ist die Einteilung verschiedener in Kapitel 5.1 aufgeführter Anwendungsbeispiele abhängig davon dargestellt, ob ein transparentes oder ein herkömmliches Display vorliegt und ob dem Anwender Erklärungen zu Objekten angeboten werden oder ob er bei der Navigation unterstützt wird.

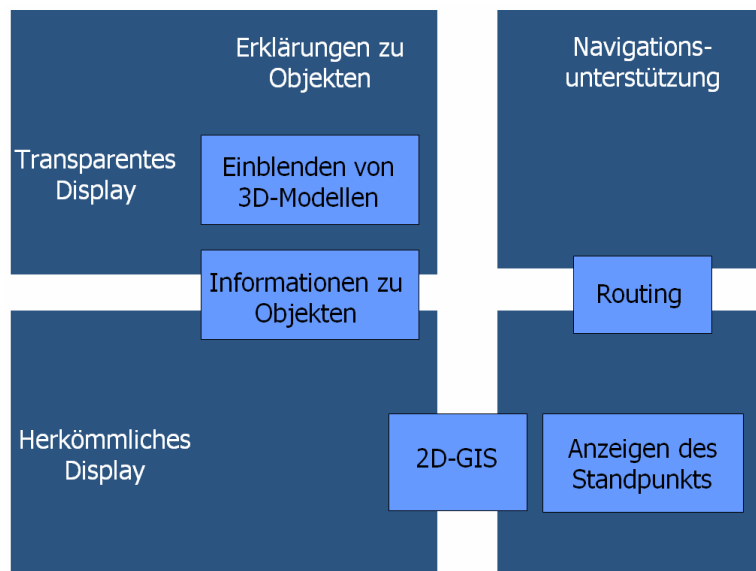


Abbildung 23: Einfache Einteilung von Diensten

Diese grundsätzliche Einteilung der Zielsetzung in die Bereiche Erklärung zu Objekten und in Aufgaben zur Navigationsunterstützung wird im Hinblick auf eine Ableitung von Anforderungen einer genaueren Einteilung unterzogen. Um Erklärungen zu einzelnen Objekten angeben zu können, werden mobile Endgeräte folgendermaßen eingesetzt:

- Informationsdarstellung über ein GIS
- Text- und Bildinformationen zu Objekten in Blickrichtung
- Informationen zu einzelnen Objekten einschließlich Gebäudeteilen direkt im Blickfeld
- Einblenden von positionsgenauen 3D-Modellen.

Für die Navigationsunterstützung ergeben sich damit folgende Möglichkeiten:

- Anzeige des derzeitigen Standpunkts in einer nach Norden / in Blickrichtung ausgerichteten Karte
- Darstellung des Weges vom derzeitigen Standpunkt zu einem angegebenen Ziel in einer nach Norden / in Blickrichtung ausgerichteten Karte
- Routing über Textausgaben oder eine Karte
- 3D-Routing.

5.3.2 Parametrisierung der Anforderungen

Im letzten Abschnitt 5.3.1 wurde eine Klassifizierung in Bezug auf das eingesetzte Display und die Zielsetzung erläutert. Im Folgenden werden nun Parameter zur Beschreibung der Anforderungen an die Trackingtechnik abgeleitet. Dazu werden zunächst zwei Untersuchungen aufgeführt, die zu einer ersten Ableitung von Anforderungen führen. Azuma teilt AR-Aufgaben ein, Darken hat sich mit der Handhabung von Karten beschäftigt.

Zunächst wird deshalb auf Anwendungen eingegangen, die zur Gruppe der AR-Aufgaben zählen. Damit benötigen sie das Einblenden von positions- und blickrichtungsabhängigen Informationen. Azuma zeigt auf, dass sich Aufgaben, die mit Hilfe von AR im Freien gelöst werden können, in drei verschiedene Bereiche einteilen lassen [Azum99b]:

- Anzeigen von einzelnen virtuellen Objekten im Sichtfeld
- Visualisierung des Blickfeldes wie es in der Vergangenheit aussah oder in der Zukunft aussehen wird
- Zusammenbringen von Informationen in Personengruppen.

Wenn einzelne Objekte im Blickfeld angezeigt werden, kann es sich um Textfelder handeln, die den Namen des Gebäudes angeben. Aber auch Pfeile, die eine Erklärung darstellen oder längere Erklärungen beispielsweise in Form von Avataren, die über Audioausgabe Sachverhalte erläutern, sind denkbar. Diese virtuellen Objekte können in Form eines Links auf weitere Informationen verweisen.

Im Gegensatz dazu können in AR-Anwendungen auch vollständige 3D-Szenen eingeblendet werden. Dafür muss der Ausschnitt in der Umgebung, den der Anwender im Moment betrachtet, zunächst bestimmt werden. Jede Änderung der Position oder der Blickrichtung bedeutet, dass ein neuer Ausschnitt berechnet wird. Hierbei handelt es sich also nicht mehr nur um die Identifizierung einzelner Objekte, die mit Erklärungen versehen werden, sondern um den Aufbau einer Szene über die Feststellung der Position und Blickrichtung des Anwenders.

Wenn sich verschiedene Personen in der Realität bewegen, sieht jeder die Umgebung in einem anderen Blickwinkel. Wenn diese Personen zusammen arbeiten sollen, kann es hilfreich sein, dass sie Informationen mit den anderen Personen austauschen. Dies kann dazu führen, dass die Daten, die an die anderen weitergegeben werden, auf den Blickwinkel des einzelnen Anwenders zugeschnitten sind. In diesem Fall wird dann beispielsweise nicht mehr in einer Karte die Position

eines Gebäudes eingezeichnet, sondern das Gebäude wird direkt im Blickfeld des einzelnen Anwenders markiert. Hierzu ist es nicht nur notwendig Position und Blickrichtung einer einzelnen Person, sondern aller Anwender, die zu einer Gruppe gehören, zu bestimmen und diese Informationen, auf den einzelnen Anwender zugeschnitten, weiterzugeben.

Alle diese Dienste aus dem AR-Bereich benötigen Informationen zur Position und zur Blickrichtung des Anwenders. Die Genauigkeit dieser Ergebnisse hängt davon ab, ob Informationen zu Objektteilen oder vollständigen großen Gebäuden sowie zu nahe gelegenen oder sich weiter entfernt befindlichen Gegenständen in das Sichtfeld des Benutzers eingeblendet werden sollen. Die Information, für welche Anzahl an Personen die Position und Blickrichtung bestimmt werden müssen, hat keine Auswirkung auf die Anforderungen an die Trackingtechniken.

Zusätzlich zu diesen AR-Anwendungen können einem mobilen Anwender Informationen zu seiner Umgebung auf einem herkömmlichen Display angeboten werden. Dazu zählt die Handhabung von geographischer Information, die herkömmlich in Form von Karten durchgeführt wurde. Nach Darken, der die Orientierung von Karten in virtuellen Umgebungen untersuchte, können Navigationsaufgaben in zwei Gruppen unterteilt werden [Dark99]:

- Suche nach einem Ziel
- Erforschung der Umgebung.

Damit unterscheidet Darken den Umgang mit einer Karte dahingehend, ob der Anwender sich frei bewegen oder geführt werden möchte. Anwendungen, die zu diesen Gruppen zählen, benötigen keine hochgenauen Trackingergebnisse, allerdings erfüllen nicht alle Techniken, die für den Einsatz im urbanen Umfeld möglich sind, die entsprechenden Anforderungen an die Genauigkeit. Ungenauere Positionierungsergebnisse bedeuten für die hier berücksichtigten Techniken, dass die Position nur etwa Straßenblockweise ermittelt wird. Diese Information reicht aus, wenn der Benutzer mit der Handhabung von herkömmlichen Karten vertraut ist, sich verlaufen hat, und nach dieser oberflächlichen Information seinen Weg ohne elektronische Hilfe fortsetzen kann.

Um eine einfache Orientierung zu ermöglichen, kann die Orientierung nur in einer Ebene bestimmt werden. So kann eine Karte in Blickrichtung ausgerichtet werden, anstatt sie ohne diese Information allgemein mit der Nordrichtung nach oben anzuzeigen. Eine Navigation des Nutzers ist aus seiner Perspektive heraus damit ebenfalls möglich.

Die Anforderungen beinhalten damit zunächst, welche Art von Ergebnis notwendig ist. Dies bedeutet, dass unterschieden wird, ob nur die Position oder die Position und die Blickrichtung bestimmt werden sollen. Da es sich sowohl bei AR-Anwendungen als auch bei LBS um ortsbezogene Dienste handelt, ist immer die Position als Information notwendig. Das Verfahren ist aber grundsätzlich so ausgerichtet, dass Dienste eingebracht werden können, bei denen nur Informationen zur Blickrichtung notwendig sind.

Eine weitere Anforderung an mögliche Trackingtechniken, die sich von Dienst zu Dienst unterscheiden, ist die Genauigkeit des Ergebnisses. Für das positionsgenaue Einblenden von nahen 3D-Modellen in das Sichtfeld des Anwenders sind höhere Genauigkeiten erforderlich als beim Verabreden von Freunden. Wie bereits erwähnt, hängt die Anforderung an die Genauigkeit der Trackingergebnisse für Aufgaben aus

dem Bereich AR davon ab, wie nahe die Objekte in Bezug zum Benutzer liegen, zu denen Informationen eingeblendet werden sollen, und ob Informationen zu, aus der Perspektive des Nutzers her, klein erscheinenden Gebäudeteilen angeboten werden. Grundsätzlich ist davon auszugehen, dass bei der Einblendung von 3D-Modellen höchste Genauigkeitsanforderungen verlangt werden. Wie bereits erläutert, reicht es für einige Dienste aus dem Bereich LBS aus, dass die Genauigkeit der Position sich auf einen Straßenblock bezieht. Für Navigationsaufgaben sind höhere Ansprüche an die Genauigkeit notwendig.

Zusätzlich zu diesen beiden Anforderungen ist die Taktfrequenz wichtig. Besonders wenn Informationen im Blickfeld des Anwenders eingeblendet werden sollen, ist eine hohe Taktfrequenz notwendig. Diese Anforderung umfasst Echtzeitanwendungen, die eine Bildwiederholrate von bis zu 25 Bildern pro Sekunde ermöglichen können. Eine nicht so hohe Taktfrequenz ist beim Routing auf einem PDA gefordert, da der Anwender sich in der Umgebung zurechtfinden muss und nur gelegentlich auf die Ausgabe blickt. Bei der einmaligen Anzeige der derzeitigen Position ist es nicht notwendig, das Trackingergebnis zu aktualisieren.

Eine weitere Gruppe an mobilen Informationsdiensten benötigt keine automatisch generierten Trackinginformationen. Dazu zählen Aufgaben, die von der Position unabhängig sind oder der manuellen Eingabe der Position bedürfen. Diese werden im Folgenden nicht weiter betrachtet.

Damit ergeben sich die folgenden Parameter, die die Anforderungen an ein Trackingergebnis bilden:

- Sensorklasse
- Genauigkeit
- Taktfrequenz.

Die Kombination dieser Anforderungen führt dazu, dass eine bestimmte Trackingtechnik, die diesen Anforderungen genügt, vom System ausgewählt wird.

5.3.3 Formulierung der Anforderungen

Für alle der in Abschnitt 5.3.2 abgeleiteten Anforderungen werden nun die Anforderungen formuliert.

Die Sensorklasse hängt davon ab, welche Trackingergebnisse notwendig sind. Dies führt zu folgender Einteilung der in Abschnitt 5.3.1 aufgeführten Informationen:

1. Position: Anzeige des derzeitigen Standpunktes und Darstellung des Weges vom derzeitigen Standpunkt zu einem angegebenen Ziel jeweils in einer Karte, in der sich am oberen Rand Norden befindet, Routing über Textausgaben oder eine Karte, Informationsdarstellung über ein GIS
2. Blickrichtung: Keine der angegebenen Dienste benötigt Trackinginformationen nur zur Blickrichtung

3. Position und Blickrichtung: Anzeige des derzeitigen Standpunktes und Darstellung des Weges vom derzeitigen Standpunkt zu einem angegebenen Ziel jeweils in einer in Blickrichtung ausgerichteten Karte, 3D-Routing, Informationen zu Objekten in Blickrichtung oder im Blickfeld, Einblenden von positionsgenauen 3D-Modellen

Damit treten nur die Sensorklassen *Position* sowie *Position und Blickrichtung* auf. Letztere wird bei der eigentlichen Zuordnung in einzelne Komponenten getrennt.

Die erforderlichen Genauigkeiten hängen davon ab, wie ausführlich der Anwender informiert werden möchte oder wie nahe er zu den Objekten positioniert ist. Alle Dienste können in Gruppen eingeteilt werden, die eine Genauigkeitsanforderung repräsentieren. Die Werte für die Genauigkeit werden in Metern bei den Positionstechniken und in Altgrad zur Bestimmung der Blickrichtung ausgedrückt. Dies führt zur Bildung der im Folgenden aufgelisteten Gruppen:

- Für die Position: Submeterbereich, 1-10 Meter, mehr als 10 bis 100 Meter, mehr als 100 Meter
- Für die Blickrichtung: Bis 10°, 10-30°, bis 90°

Die zuvor genannten Anwendungen können demnach Genauigkeitsgruppen zugeordnet werden. Dabei wird, wie oben erwähnt, eine Anwendung beiden Sensorklassen zugeordnet, wenn sowohl Position als auch Blickrichtung bestimmt werden müssen. Dies führt zu der im Folgenden dargestellten Auflistung.

Position

- Submeterbereich: Einblenden von positionsgenauen 3D-Modellen
- Bis 10 Meter: Anzeige des derzeitigen Standpunktes und Darstellung des Weges vom derzeitigen Standpunkt zu einem angegebenen Ziel in einer Karte, Routing über Textausgaben oder eine Karte, 3D-Routing, Informationen zu einzelnen Objekten direkt im Blickfeld
- Bis 100 Meter: Informationsdarstellung über ein GIS, Text- und Bildinformationen zu Objekten in Blickrichtung
- Mehr als 100 Meter: keine der aufgeführten Beispielanwendungen benötigt Positionsinformationen in diesem Genauigkeitsbereich.

Blickrichtung

- Bis 10°: Einblenden von positionsgenauen 3D-Modellen
- Bis 30°: 3D-Routing, Informationen zu einzelnen Objekten direkt im Blickfeld,
- Bis 90°: Anzeige des derzeitigen Standpunktes und Darstellung des Weges vom derzeitigen Standpunkt zu einem angegebenen Ziel jeweils in einer in Blickrichtung ausgerichteten Karte, Text- und Bildinformationen zu Objekten in Blickrichtung.

Ob das Ergebnis in einer regelmäßigen Frequenz ermittelt werden soll, hängt davon ab, ob die Anwendung nur einmal auf Anforderung (automatisch oder manuell) die Trackingergebnisse benötigt, oder ob die Anwendung eine regelmäßige Aktualisierung erfordert. In letzterem Fall entspricht die Frequenz bei einfachen

Sensoren der Taktrate, bei aufwendigeren Algorithmen entspricht sie dem Kehrwert der kompletten Rechenzeit in Sekunden. Die Frequenz als Anforderung ist im Folgenden den Beispieldiensten zugeordnet.

- Keine Aktualisierung notwendig: Anzeige des derzeitigen Standpunktes und Darstellung des Weges vom derzeitigen Standpunkt zu einem angegebenen Ziel in einer nach Norden / in Blickrichtung ausgerichteten Karte
- Aktualisierung wird automatisch angefordert: Routing über Textausgaben
- Regelmäßige Aktualisierung mit Sekundenfrequenz oder weniger: Routing über Textausgaben, 3D-Routing, Informationsdarstellung über ein GIS, Text- und Bildinformationen zu Objekten in Blickrichtung
- Regelmäßige Aktualisierung mit Frequenz bis zu 25 Hz: Informationen zu einzelnen Objekten einschließlich Gebäudeteilen direkt im Blickfeld, Einblenden von positionsgenauen 3D-Modellen.

5.4 Aktuelle Sensordaten

In diesem und im nächsten Abschnitt wird beschrieben, wie die Trackingtechniken charakterisiert werden können. Dazu wird davon ausgegangen, dass von den in Kapitel 2 erläuterten Techniken zur Zeit die folgenden für einen Einsatz im urbanen Umfeld bedeutsam sind: GPS, Positionierung mit Hilfe des Mobilfunknetzes, Magnetkompass und Gyroskope. Zusammen mit den videobasierten Techniken aus Kapitel 3 muss für diese festgestellt werden, welche Techniken momentan verfügbar sind und wie verlässlich diese Ergebnisse liefern. Um die Rohdaten möglichst rauschfrei auslesen zu können oder verschiedene Trackingergebnisse zu koppeln, wird anschließend der Einsatz eines Kalmanfilters vorgestellt. Diese Schritte werden als Bestimmung der aktuellen Sensordaten zusammengefasst und sind in Abbildung 24 als vorverarbeitender Schritt zur Parametrisierung der Trackingtechniken aufgezeigt.

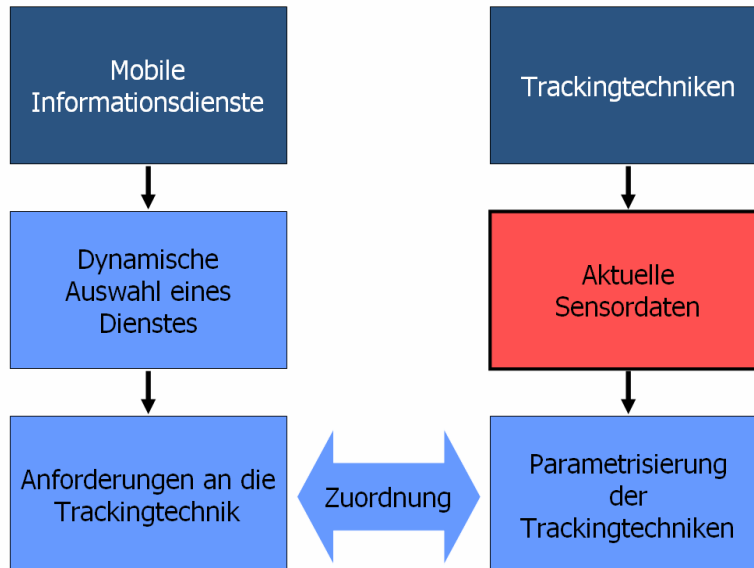


Abbildung 24: Bestimmung der aktuellen Sensordaten im Ablaufschema des Zuordnungsverfahrens

In die am Schluss des Verfahrens durchzuführende Zuordnung gehen nur die Trackingtechniken ein, die unter den aktuellen Bedingungen auch eingesetzt werden können. Dazu muss, abhängig von Einflüssen der Umgebung, zunächst geprüft werden, welche Trackingtechniken grundsätzlich zur Verfügung stehen und welche zu diesem Zeitpunkt an diesem Ort möglich sind. Dabei handelt es sich meist um Umgebungseinflüsse, die die Leistung eines Gerätes oder einer Technik vermindern. Aus diesem Grund muss festgestellt werden, wie robust und beständig sie für den Einsatz in einer Stadt sind und welche Einflüsse in der Umgebung zu einer Leistungsminderung führen können. Diese Verfügbarkeit kann von der Position abhängen, an der sich der Anwender befindet. Deshalb können Problembereiche ausgewiesen werden, in denen Signale oder Informationen geblockt werden. Diese Bereiche beeinflussen das Einsatzgebiet einer Trackingtechnik, das abgefragt werden kann. Wenn dieses beschränkt ist, muss während des Einsatzes ein Ortsvergleich durchgeführt werden. Über dieses Kriterium können Trackingtechniken ausgeschlossen werden. Nur von den möglichen Trackingtechniken werden die notwendigen Parameter der momentanen Ergebnisse abgefragt.

Wenn sofort beim Start des Systems ein Dienst angefordert wird, muss die Initialisierungszeit abgefragt werden. Zur Überbrückung der Initialisierungszeit kann dann gegebenenfalls eine weitere Technik angesprochen werden, die mit zu den möglichen Techniken zählt.

Für manche Trackingtechniken kann nicht nur festgestellt werden, wann sie grundsätzlich zur Verfügung stehen sondern auch, wie zuverlässig die Ergebnisse sind. Die Zuverlässigkeit des Ergebnisses, die beispielsweise bei GPS über die DOP-Faktoren charakterisiert werden kann, gibt an, wie stark der Anwender einem Ergebnis trauen kann. Dieser Wert muss bei einzelnen Techniken dynamisch abgefragt werden. Aufgrund dieser drei Werte

- Einsatzgebiet
- Initialisierungszeit
- Zuverlässigkeit

kann festgestellt werden, welche der Trackingtechniken momentan verfügbar sind und wie verlässlich diese Ergebnisse liefern.

Das Verfahren soll es erlauben, dass theoretisch alle eingesetzten Trackingtechniken Ergebnisse liefern und dass anforderungsadaptiv eine oder eine Kombination zugeordnet werden. Dafür müssen die Techniken einzeln oder in Kombination angesprochen werden. Um eine Kombination verschiedener Ergebnisse zu ermöglichen, wird hier auf einen Kalmanfilter zurückgegriffen. Dieser kann ebenfalls zur Ausgleichung der Rohdaten nur für einen einzelnen Sensor angewendet werden. Die Besonderheit des Kalmanfilters liegt darin, dass die reinen Messwerte bei der Ermittlung des Ergebnisses direkt berücksichtigt werden. Prädiktionen erlauben die Vorhersage von zukünftigen Werten. Damit kann die vorhergehende Messung bei der Ermittlung des Ergebnisses zu einem bestimmten Zeitpunkt berücksichtigt und die Möglichkeit genutzt werden, den vorhergesagte Zustand mit den tatsächlichen Messwerten zu vergleichen. Damit sind aktuelle Informationen über die Güte der derzeitigen Messung möglich. Bei der Berechnung eines Ergebnisses über verschiedene Messungen können die verschiedenen Eigenschaften der Trackingtechniken einbezogen werden. Der Vorteil des Kalmanfilters ist eine hohe Genauigkeit und eine große Effizienz.

Voraussetzung für die Verwendung eines Kalmanfilters ist, dass die Schätzwerte normalverteilt sind, und dass das Prozess- und das Messrauschen, die beide in die Schätzung eingehen, unabhängig von einander, weiß und ebenfalls normalverteilt sind [Welc02b]. Für die Schätzung müssen Anfangswerte vorhanden sein oder berechnet werden können.

Im Allgemeinen benötigt der Kalmanfilter lineare Beziehungen. Dann beschreibt die folgende Formel die Veränderung eines Zustandes x eines diskreten zeitabhängigen Prozesses zwischen zwei Zeitpunkten [Welc02b]:

$$x_k = A_k x_{k-1} + B u_k + w_{k-1}.$$

Eine reine Zustandsänderung wird durch die Matrix A ausgedrückt, wenn weder eine Einwirkung von außen noch ein Prozessrauschen vorhanden wären. Diese Einwirkung ist durch die Matrix B , das Prozessrauschen durch die Variable w_k berücksichtigt. Die Matrix B beinhaltet Kontrolleinflüsse, die durch Beeinflussungen des Messvorganges hervorgerufen werden können. Die Werte für w_k gehen in die Matrix Q beim Ablauf des Kalmanfilters ein. Die erläuterte Formel ermöglicht die Prädiktion innerhalb des Kalmanfilters.

Die entsprechende Messung z_k zu diesem Zustand kann über die folgende Gleichung bestimmt werden:

$$z_k = H x_k + v_k.$$

Hierbei wird die reine Abbildung des Zustandes auf die Messung über die Matrix H und das Messrauschen über die Variable v_k formuliert. Von letzterem wird angenommen, dass es nicht vom Prozessrauschen abhängig ist. Die Werte dieser Variablen gehen in die Matrix R beim Ablauf des Kalmanfilters ein.

Die Werte für das Rauschen werden in den Kovarianzmatrizen Q für das Prozessrauschen und R für das Messrauschen ausgedrückt, die im Allgemeinen mit jeder zeitlichen Veränderung oder mit jeder Messung geändert werden.

Durch die Prädiktion ist ein Vergleich der Vorhersage mit den Messwerten möglich. Darüber können Korrekturen angebracht und Genauigkeitsaussagen gemacht werden. Diese führen beispielsweise dazu, dass das Gewicht einer Messung von deren Genauigkeit abhängt.

Der erweiterte Kalmanfilter ermöglicht die Berücksichtigung von Sachverhalten, die nicht linear sind. Dafür wird eine Linearisierung über den aktuellen Mittelwert und die entsprechende Kovarianzmatrix benötigt.

Die oben genannten Formeln werden so angewendet, dass zunächst eine a posteriori Abschätzung des Zustandes als Linearkombination einer a priori Abschätzung und einer gewichteten Differenz zwischen der tatsächlichen Messung und der Messungsvoraussage Berücksichtigung finden. Damit lässt sich die Differenz zwischen der vorhergesagten Messung und der tatsächlichen Messung ermitteln. Die Eingangsmatrizen werden für jede Trackingtechnik aufgestellt.

Dafür müssen zwei Fehlerkovarianzen berechnet werden. P_k^- ist die a priori Schätzung, P_k ist die a posteriori Schätzung der Kovarianzmatrix. Damit können sowohl a priori als auch a posteriori Schätzungen der Messwerte \hat{x} durchgeführt werden, d.h. \hat{x}_k^- bzw. \hat{x}_k .

Aus den Anfangswerten für \hat{x} und P wird zunächst eine Vorhersage berechnet. Diese wird bei [Welc02b] durch einen hochgestellten Strich nach der Variablen ausgedrückt.

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k$$

$$P_k^- = AP_{k-1}A^T + Q$$

Mit Hilfe der so ermittelten a priori Werte kann eine verbesserte a posteriori Schätzung durchgeführt werden:

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$$

$$P_k = (I - K_kH)P_k^-$$

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1}$$

wobei K der Kalman Gain ist, der die a posteriori Fehlerkovarianz P_k minimiert. Diese Werte gehen in die Ermittlung der nächsten Vorhersage ein.

Ein Überblick über den Ablauf des Kalmanfilters ist in Abbildung 25 enthalten.

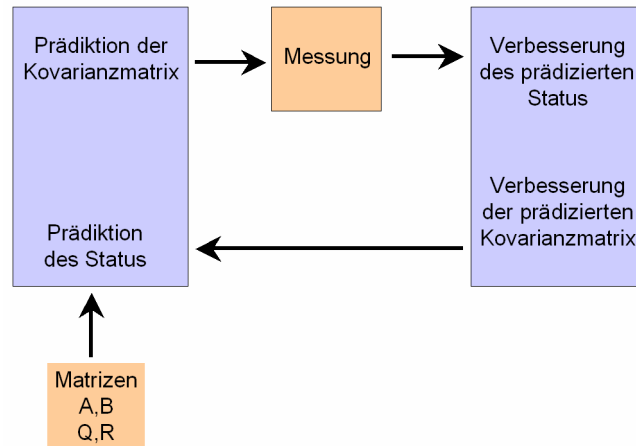


Abbildung 25: Überblick über den Ablauf des Kalmanfilters

Die Matrizen A und Q werden zu Beginn des Ablaufs einmal aufgestellt. Für jede Trackingtechnik werden dann zusätzlich die Matrizen H und R , sowie die Startwerte für die Matrix P und den Status x festgelegt. Mit Hilfe dieser Werte werden die Vorhersagen bestimmt. Das bedeutet, dass der a priori Messstatus sowie die Kovarianzmatrix berechnet werden. Mit Hilfe der tatsächlichen Messergebnisse können Korrekturen an den Werten angebracht werden. Diese gehen in die Vorhersage des nächsten Zustandes ein. Bei jeder neuen Initialisierung werden die entsprechenden Matrizen dynamisch aufgebaut. Es muss jedes Mal mit einer Einschwingphase gerechnet werden, deren Dauer und Größe von der Abweichung der Startwerte von den tatsächlichen Ergebnissen abhängt.

5.5 Parametrisierung der Trackingtechniken

5.5.1 Überblick über die Trackingtechniken

Die zurzeit verfügbaren Trackingtechniken, die zum momentanen Zeitpunkt verlässliche Ergebnisse liefern, müssen in einem nächsten Schritt bezüglich der Güte ihrer Ergebnisse charakterisiert werden. Abbildung 26 zeigt auf, dass nach der im Folgenden beschriebenen Parametrisierung der Trackingtechniken die Ergebnisse in die Berechnung der Zuordnung eingehen.

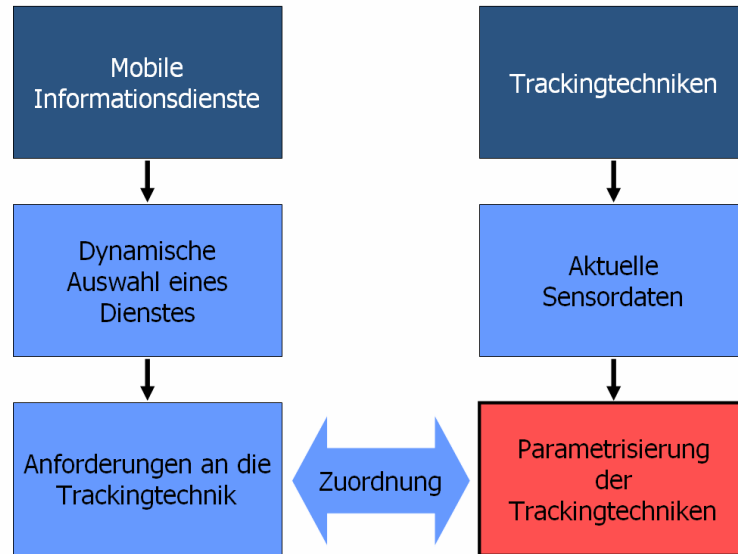


Abbildung 26: Parametrisierung der Trackingtechniken im Ablaufschema des Zuordnungsverfahrens

Die zur Beschreibung der Ergebnisse notwendigen Parameter werden durch die bestehenden Anforderungen des gewählten Dienstes an die Trackingtechnik bestimmt, die in Abschnitt 5.3.2 hergeleitet wurden. Aus diesem Grund müssen neben der Art des Trackingergebnisses die Genauigkeit aber auch die Ergebnisfrequenz bzw. die Rechenzeit der Technik berücksichtigt werden. Die Genauigkeit hängt nicht nur von den äußeren Bedingungen des Messvorganges an sich ab, sondern auch vom Sensor. Die Ergebnisfrequenz gibt an, wie oft eine Messung an den Rechner weitergegeben wird.

Die Art des Trackingergebnisses, das die Technik liefert, die Sensorklasse, spielt eine wichtige Rolle bei einer Zuordnung der Trackingtechnik zu einem bestimmten Dienst. Diese erste Einteilung der möglichen Trackingtechniken erlaubt eine erste Annäherung an die Anforderungen, die die Dienste bei der Zuordnung stellen. Grundsätzlich können diese Trackingtechniken eingeteilt werden abhängig davon, welcher Art das Trackingergebnis ist:

- Position
- Blickrichtung
- Position und Blickrichtung

Damit ergibt sich für die in Abschnitt 5.4 festgelegten, als für im Folgenden näher zu betrachtende Techniken, die in Abbildung 27 dargestellte Einteilung.

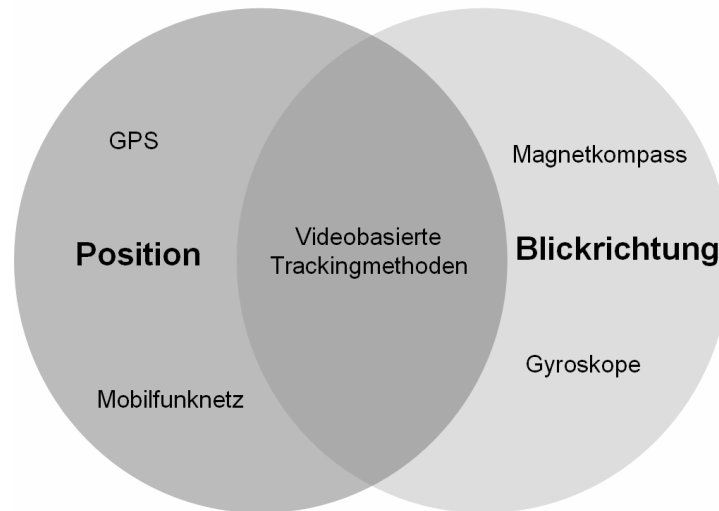


Abbildung 27: Einteilung der Trackingtechniken nach Art des Ergebnisses

GPS und Ergebnisse, die mit Hilfe des Mobilfunknetzes bestimmt wurden, ermöglichen Informationen zur Position des Anwenders. Magnetkompass und Gyroskope ermitteln die Orientierung einer Person. Videobasierte Verfahren liefern entweder die eine oder andere Art des Ergebnisses. Einige dieser Verfahren liefern darüber hinaus sowohl die Position als auch die Blickrichtung.

Nach der Aufstellung der Anforderungen für einen der mobilen Informationsdienste muss innerhalb des Systems die Zuordnung einer Trackingtechnik durchgeführt werden. Die Parametrisierung der Trackingtechniken erfolgt getrennt für die Bestimmung der Position und der Blickrichtung. Dies bedeutet, dass eine Technik, die beide Ergebnisklassen liefert, zweimal parametrisiert wird. Damit ist es möglich, eine Technik einzusetzen, auch wenn nur eines der Ergebnisse in diesem Moment benötigt wird. Allerdings ist zu berücksichtigen, dass nicht unnötiger Rechenaufwand angefordert wird. Es macht deshalb keinen Sinn, eine Orientierungsbestimmung für alle drei Achsen durchzuführen, wenn die Information anschließend auf dem Display eines Laptops ausgegeben wird. Hier würde es ausreichen, wenn der horizontale Winkel, das ist der Winkel um die Körperachse, bestimmt wird.

Es sollte deshalb unterschieden werden, ob der Anwender eine nicht transparente Ausgabe benutzt oder ob er ein transparentes Display einsetzt, das er sich direkt vor die Augen halten kann, so dass virtuelle Informationen ins Sichtfeld eingeblendet werden können und AR-Anwendungen denkbar sind. Abbildung 28 gibt eine Übersicht der Techniken, die für das Tracking im Außenbereich eingesetzt werden können, wenn der Anwender ein Gerät mit einem herkömmlichen Bildschirm oder einer ähnlichen Ausgabe einsetzt.

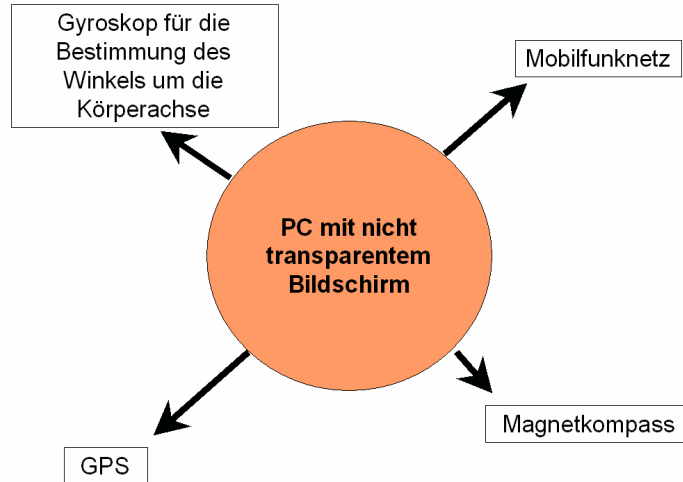


Abbildung 28: Trackingtechniken zur Nutzung mit einem mobilen Endgerät ohne transparentes Display

In diesem Fall kann die Position mit Hilfe von GPS bzw. des Mobilfunknetzes oder aus einer Kombination der beiden berechnet werden. Die Orientierung des Anwenders in Bezug zur Körperachse kann dann über ein Gyroskop bzw. einen Kompass berechnet werden.

Wenn ein transparentes Display an einen PC angeschlossen vorhanden ist, das es ermöglicht, die Information direkt in das Blickfeld als Überlagerung des sichtbaren Bereichs einzublenden, ist es notwendig, die Blickrichtung in allen drei Achsen zu bestimmen. Abbildung 29 zeigt, welche Trackingtechniken dann angeschlossen werden können.

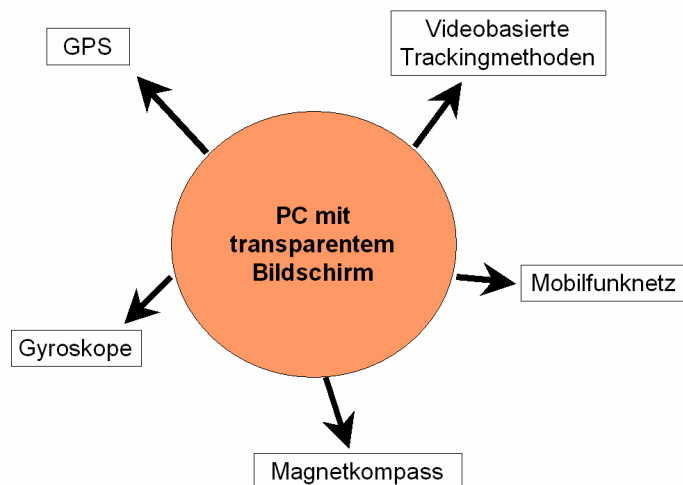


Abbildung 29: Trackingtechniken zur Bestimmung von Position und Blickrichtung bei Vorhandensein eines transparenten Displays

Dies bedeutet, dass zusätzlich zu den bereits möglichen Trackingtechniken videobasierte Trackingmethoden möglich sind. Außerdem sind Gyroskope für die Bestimmung der Neigung und der Drehung um die optische Achse notwendig, wenn die Blickrichtung für alle drei Achsen bestimmt werden soll. Videobasierte Methoden liefern unterschiedliche Arten von Ergebnissen, die vom umgesetzten Algorithmus abhängen. Meist ist es möglich, dass auch nur ein Teil der Ergebnisse weitergegeben wird. Da diese Verfahren aber sehr rechenaufwendig sind, empfiehlt es sich, diese nur anzuwenden, wenn auch tatsächlich ein hochgenaues Ergebnis, das aus der Position und dem Blickwinkel um drei Achsen besteht, verlangt wird.

Wenn es ausreicht, die Position und Blickrichtung mit geringeren Genauigkeiten zu bestimmen, können ungenauere Trackingtechniken eingesetzt werden. Bei Informationen zu sehr nahen Objekten oder Einzelheiten in der Umgebung, müssen diese Ergebnisse durch ein genaueres Verfahren wie beispielsweise eine videobasierte Methode verbessert oder ersetzt werden.

5.6 Zusammenstellung der Parameter

In diesem Abschnitt wird gezeigt, wie die Parameter zur Charakterisierung der Trackingtechniken erhalten werden können, auf die im vorhergehenden Abschnitt bezüglich der Sensorklasse bereits eingegangen worden ist. Dazu werden die in den beiden vorausgegangenen Abschnitten 5.4 und 5.5 erläuterten Parameter Einsatzgebiet, Zuverlässigkeit, Initialisierungszeit, Sensorklasse, Genauigkeit und Ergebnisfrequenz für jede der hier betrachteten Trackingtechniken beschrieben.

5.6.1 GPS

Wie in Abschnitt 2.3 näher erläutert, kann mit Hilfe von GPS die Position bestimmt werden. Die Zuverlässigkeit von GPS kann über die Satellitenkonstellation, die zur Berechnung herangezogen wird, bewertet werden. Sie ist in dem standardisierten NMEA-Format (National Marine Electronics Association) enthalten, das zur Übertragung der Daten bei GPS-Messungen verwendet wird. Die Genauigkeit hängt vom eingesetzten GPS-Sensor und der Art der Messung ab. Durch Kontrollmessungen kann die Standardabweichung bestimmt werden. Die Frequenz ist ebenfalls vom Gerät abhängig.

Die Parameter zur Charakterisierung des GPS-Ergebnisses können damit wie folgt beschrieben werden:

- Einsatzgebiet: Freie Sicht zu mindestens vier Satelliten nötig
- Zuverlässigkeit: Über PDOP bestimmbar
- Initialisierungszeit: Geräteabhängig
- Sensorklasse: Position
- Genauigkeit: Geräte- und technikabhängig
- Ergebnisfrequenz: Geräteabhängig

5.6.2 Positionierung auf Basis des GSM-Netzes

Wie in Abschnitt 2.2.2 beschrieben, gibt es verschiedene Verfahren, um mit Hilfe des GSM-Netzes die Position einer Person zu bestimmen. Dort wird ebenfalls erwähnt, dass im urbanen Umfeld die Methode basierend auf der Zellen-Nummer und das EOTD-Verfahren die verlässlichsten Ergebnisse liefern. Die Parameter zur Charakterisierung eines Ergebnis, das mit Hilfe des Mobilfunknetzes bestimmt wird, können damit wie folgt beschrieben werden:

- Einsatzgebiet: Im Freien überall
- Zuverlässigkeit: Von den Basisstationen und der Umgebung abhängig
- Initialisierungszeit: Im Vergleich zur Frequenz vernachlässigbar klein
- Sensorklasse: Position
- Genauigkeit: Technikabhängig
- Ergebnisfrequenz: Abhängig von der Technik und der Anzahl der weiteren Mobiltelefone pro Zelle

5.6.3 Magnetkompass

Mit Hilfe eines Kompasses kann die Orientierung in der Horizontalebene bestimmt werden. Eventuell muss eine Reduktion auf die geographische Nordrichtung angeschlossen werden. Diese Geräte können praktisch überall eingesetzt werden. Wie in Abschnitt 2.7 erläutert, können die Ergebnisse teilweise in Gebieten, in denen das Erdmagnetfeld gestört ist oder sich beispielsweise entsprechende Bauwerke befinden, sehr stark verfälscht werden. Zusammenfassend können die Parameter zur Charakterisierung eines Ergebnisses, das mit einem Magnetkompass bestimmt wird, wie folgt beschrieben werden:

- Sensorklasse: Orientierung in der Horizontalebene
- Einsatzgebiet: Überall
- Zuverlässigkeit: Verfälschungen bei Erdmagnetfeldstörungen oder Magnetfeldern in der Umgebung
- Genauigkeit: Geräteabhängig
- Ergebnisfrequenz: Geräteabhängig
- Initialisierungszeit: Vernachlässigbar

5.6.4 Gyroskope

Der Vorteil von inertialen Systemen ist, dass sie vollkommen unabhängig von weiteren Sensoren eingesetzt werden können. Aus diesem Grund sind sie in fast allen AR-Systemen für den Außenbereich zu finden. Wie in Abschnitt 2.6 beschrieben, liefern sie allerdings nicht über einen längeren Zeitraum hinweg zuverlässige Ergebnisse. Der wichtigste Grund dafür ist, dass die Ergebnisse von reinen Gyroskopen relativ sind. Sie sollten deshalb gestützt werden. Außerdem muss

zusätzlich zu Beginn des Einsatzes von Gyroskopen ein Startwert bestimmt werden oder die absoluten Ergebnisse werden über einen weiteren Sensor ermittelt. Die zusätzliche Zeit, die dafür beansprucht wird, geht hier in die Initialisierungszeit der Gyroskope ein. Die Genauigkeit und die Frequenz von Gyroskopen sind geräteabhängig. Die Parameter zur Charakterisierung eines Ergebnisses, das mit Hilfe von Gyroskopen bestimmt wird, können damit wie folgt beschrieben werden:

- Einsatzgebiet: Überall
- Zuverlässigkeit: Durch relatives Aufsummieren abhängig von Messzeit
- Initialisierungszeit: hängt von dem Aufwand zur Ermittlung absoluter Ergebnisse ab
- Sensorklasse: Orientierung
- Genauigkeit: Geräteabhängig
- Ergebnisfrequenz: Geräteabhängig

5.6.5 Videobasierte Verfahren

Videobasierte Verfahren, die im vorliegenden Fall alle als AR-Systeme für den Außenbereich entwickelt wurden, sind auf den Einsatz an bestimmten Stellen beschränkt und nur unter bestimmten Bedingungen einsetzbar. Es handelt sich damit um Insellösungen. Dies ist ein grundsätzlicher Unterschied beispielsweise zur Positionsbestimmung mit Hilfe des GSM-Netzes. Dieses funktioniert theoretisch überall in einer Stadt. Bei einer automatischen Detektion der räumlichen Verfügbarkeit ist damit die Bestimmung der Position über einen weiteren Sensor und die Überprüfung dieser Ergebnisse notwendig. Über die Auswertung eines Kamerabildes wird ein Vergleich mit Daten der Umgebung hergestellt und so die Position und Blickrichtung des Benutzers bestimmt. Bei der Initialisierung der entsprechenden Verfahren muss festgelegt werden, ob Eingabeparameter notwendig sind. So ist bei der Bestimmung von Position und Blickrichtung über Referenzphotos die Information, welche Referenzphotos aus der Datenbank zum Vergleich herangezogen werden sollen, entscheidend für ein effektives Tracking.

Die betrachteten videobasierten Verfahren bieten, abhängig von äußeren Einflüssen der Umgebung, verschieden gute Ergebnisse. Alle Verfahren basieren auf dem Vergleich einer Referenz mit der realen Umgebung, die durch Bilder einer Kamera repräsentiert wird. Die Qualität dieser Bilder hängt von der eingesetzten Kamera ab. Je besser die zur Korrespondenzfindung notwendigen Merkmale oder Eigenschaften aus dem Bild zu denen der Referenz passen, desto besser funktioniert das Verfahren. Dies kann durch veränderte Lichtbedingungen, oder zwischen Kamera und Referenz befindlichen Objekten wie Autos, Fußgängern oder Bäumen erschwert werden. Nicht alle davon sind vorhersagbar. Aus diesem Grund können sie nicht als Eigenschaften aufgenommen werden. Lichtverhältnisse können über den Sonnenstand, der über die Tageszeit berechnet werden kann, und Wetterangaben berücksichtigt werden.

Zusammenfassend kann gesagt werden, dass die videobasierten Trackingmethoden sehr verschiedene Eigenschaften aufweisen. Sie können alle nur in solchen Gebieten eingesetzt werden, in denen Referenzdaten vorhanden sind. Dies führt dazu, dass sie meist nur sehr eingeschränkt anwendbar sind. Je nach

implementierten Algorithmen können die Position, die Orientierung oder beide Ergebnisse bestimmt werden. Die Zuverlässigkeit hängt ebenfalls von den eingesetzten Verfahren ab. Wenn aus den Kamerabildern Merkmale extrahiert werden, sind die Verfahren meist sehr abhängig von Helligkeitsänderungen. Die erreichten Genauigkeiten und Frequenzen hängen sehr stark von der Implementierung ab. Aus diesem Grund sind diese Größen durch Testmessungen abzuschätzen. Die Initialisierungszeit hängt ebenso von der Wahl des eingesetzten Verfahrens ab und ist zu überprüfen.

Zusammenfassend können die Parameter zur Charakterisierung eines Ergebnisses, das mit Hilfe von videobasierten Verfahren bestimmt wurde, wie folgt beschrieben werden:

- Einsatzgebiet: Datenabhängig
- Zuverlässigkeit: Verfahrensabhängig
- Initialisierungszeit: Verfahrensabhängig
- Sensorklasse: Position oder Orientierung oder Position und Orientierung
- Genauigkeit: Verfahrensabhängig
- Ergebnisfrequenz: Verfahrensabhängig

Dies zeigt, dass die einzelnen Verfahren durch Testmessungen beurteilt werden müssen.

5.7 Zuordnung der Trackingtechnik

In die Zuordnung der Trackingtechnik zu einem bestimmten Dienst gehen die Anforderungen an die Trackingtechnik sowie Parameter ein, die die zur Verfügung stehenden Trackingtechniken charakterisieren. Dieser letzte Schritt des Verfahrens, die Berechnung der Zuordnung, ist innerhalb des Ablaufs des gesamten Verfahrens in Abbildung 30 dargestellt.

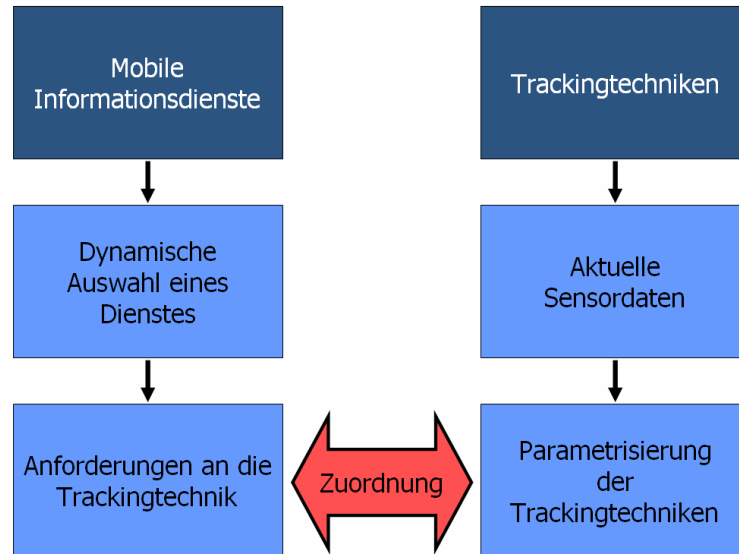


Abbildung 30: Berechnung der Zuordnung im Ablaufschema des Zuordnungsverfahrens

Alle verfügbaren Techniken, die der angefragten Sensorklasse entsprechen, gehen im Anschluss an die Ermittlung der aktuellen Sensordaten in die Zuordnungsberechnung ein. Bei den vorliegenden Verfahren wird nicht davon ausgegangen, dass automatisch die genaueste, schnellste und robusteste Technik angesprochen wird, sondern es soll diejenige zugeordnet werden, deren Parameter zur Charakterisierung der Trackingtechnik am besten zu den Anforderungen des mobilen Informationsdienstes an die Trackingtechnik passen [Kret03]. Dafür wird ein statistisches Verfahren angepasst.

Aufgrund der Datengrundlage handelt es sich um ein multivariates Modell, da die beeinflussenden Faktoren sich nicht durch eine Eigenschaft, sondern durch verschiedene Einflüsse erklären lassen. Die Trackingtechniken entsprechen Klassen, denen die Anforderungen zugeordnet werden sollen. Dazu wird ein Index aus den Parametern, die für die Zuordnung entscheidend sind, berechnet. Um eine standardisierte Zuordnung zu ermöglichen gehen die Standardabweichungen in die Berechnung ein. Der Index wird über die z-Werte berechnen, die auf einer t-Verteilung basieren:

$$z_i = \frac{(\mu_x - X_i)}{s_x}$$

Dabei entspricht X dem Anforderungswert, μ dem Parameterwert der Klasse und s der Standardabweichung der Stichprobe. Bei einer neuen Anforderung wird für jede Klasse ein Indexwert berechnet.

Die Bedeutung, die der einzelne Parameter bei der Zuordnung spielt, geht über einen Faktor in die Zuordnungsberechnung ein. Bei der vorliegenden Zuordnung wird davon ausgegangen, dass der Faktor Genauigkeit bedeutsamer für die Zuordnung ist als der Faktor Ergebnisfrequenz. Aus diesem Grund wird die Ergebnisfrequenz mit dem Faktor 0.8 gewichtet. Aufgrund der umgekehrten Eigenschaften der Ergebnis-

frequenz muss das Vorzeichen beim Einbringen dieser Komponente geändert werden. Dann berechnet sich der Indexwert über die folgende Formel:

$$\text{Index} = z_1 - 0.8 z_2,$$

wobei z_1 dem Wert für den Parameter Genauigkeit, z_2 dem Wert für den Parameter Ergebnisfrequenz entspricht. Aufgrund des Vorzeichens des Indexwertes können Aussagen über die grundsätzliche Ausrichtung der Güte der Zuordnung gemacht werden. Ein negativer Indexwert gibt an, dass die Trackingtechnik die Anforderung übertrifft, ein positiver sagt aus, dass die Anforderung nicht erreicht werden konnte.

Das vorliegende Verfahren geht allerdings davon aus, dass die der Anforderung ähnlichste Trackingtechnik zugewiesen werden soll. Aufgrund der möglichen Dienste ist davon auszugehen, dass eine teilweise Verschlechterung der Trackingergebnisse nicht zu einem vollständigen Ausfall der Information führt. Aus diesem Grund wird der Betrag des Indexwertes als Ergebnis verwendet. Dann ergibt sich die Gleichung zu:

$$\text{Index} = |z_1 - 0.8 z_2|.$$

Für jede verfügbare Trackingtechnik wird der Index bestimmt und durch einen Vergleich der verschiedenen Indizes die Trackingtechnik bestimmt, die den Anforderungen am nächsten liegt.

5.8 Zusammenfassung des Verfahrens

Die Charakterisierung der Eigenschaften sowie das Zuordnungsverfahren werden nun kurz in einem Überblick dargestellt. Dabei wird davon ausgegangen, dass die Parameterwerte, die statisch sind, d.h. die sich während der Laufzeit nicht ändern, zu Beginn an zentraler Stelle abgelegt werden. Wenn die Möglichkeit besteht, einen neuen Dienst anzubieten, werden die folgenden Schritte abgearbeitet:

1. Die momentan möglichen Dienste werden ermittelt. Dabei werden sowohl die vom Nutzer abhängigen als auch von der Umgebung veranlassten Einschränkungen berücksichtigt.
2. Aus dieser Auswahl wird ein Dienst automatisch oder durch den Nutzer bestimmt.
3. Die Anforderungen für diesen Dienst werden abgerufen.
4. Bestimmung der zurzeit verfügbaren Trackingtechniken: Dieser Schritt wird regelmäßig durchgeführt. Er erlaubt das Ab- und Zuschalten von Techniken sowie die Bildung von Kombinationen.
5. Parametrisierung der Trackingtechniken: Die Techniken werden mit Parametern beschrieben, die in den Zuordnungsschritt eingehen. Die Werte der Parameter entsprechen den aufgestellten oder werden gegebenenfalls aktualisiert.
6. Anforderungsadaptive Zuordnung einer Trackingtechnik: Die Anforderungen entsprechen denen des ausgewählten mobilen Informationsdienstes.

Grundsätzlich muss zu Beginn ein Dienst initialisiert werden. Eine neue Berechnung der Zuordnung wird dadurch hervorgerufen, dass eine neue oder keine weitere Information dem Anwender präsentiert wird. Ein mobiler Informationsdienst wird entweder automatisch vom System angeboten, da eine Änderung der Bedingungen erfolgt ist oder der Anwender hat diesen Dienst veranlasst. In einem solchen Fall ist zunächst abzuklären, welche Bedingungen die Techniken erfüllen müssen. Auch jede neue Präsentation der Information startet den Ablauf neu. Anschließend wird über die Umgebungsbedingungen abgefragt, welche Techniken im Moment angesprochen werden können. Aus diesen wird anforderungsadaptiv eine Trackingtechnik ausgewählt.

Wenn kein Dienst vom Benutzer angefragt oder automatisch vom System angeboten wird, ist ein Grundtracking empfohlen, das es erlaubt, regelmäßig Werte zur Verfügung zu haben. Dies ist besonders bei der Neuinitialisierung des Kalmanfilters zweckmäßig, da dadurch die Einschwingphase zu Beginn verkürzt wird.

Mobile Informationsdienste können zu Klassen zusammengefasst werden. Dann gelten für alle Dienste einer Klasse die gleichen Anforderungen. Damit ändert sich die Zuordnung nicht, wenn verschiedene Dienste hintereinander aufgerufen werden. Wechselt die Klasse jedoch, muss eine neue Zuordnung berechnet werden.

6 Hochgenaues Trackingverfahren mit Hilfe eines 3D-Modells

Kapitel 6 zeigt ein Verfahren auf, das es ermöglicht, in einer Stadt dreidimensionale Gebäudemodelle positionsgenau im Blickfeld des Anwenders zu platzieren. Grundbedingung des Verfahrens ist es somit, dass es im Außenbereich einsetzbar ist. Wie in Kapitel 3 aufgezeigt, gelten bei AR-Systemen im Außenbereich einschränkende Maßgaben. Darüber hinaus liefert der Anspruch, 3D-Modelle einzublenden, zusätzliche Bedingungen. Die Anforderungen an dieses Verfahren werden im Folgenden zunächst erläutert.

Aufbauend auf diesen Anforderungen wird dann ein Überblick über das Verfahren gegeben, dessen einzelne Komponenten im Weiteren ausführlich dargestellt werden. Es handelt sich um ein videobasiertes Verfahren, das als Grundlage den Vergleich des Kamerabildes mit einem 3D-Modell der Umgebung herstellt.

6.1 Anforderungen an ein urbanes AR-System

Die positionsgenaue Platzierung dreidimensionaler Gebäudemodelle ist so konzipiert, dass eine Person, die sich frei in der Stadt bewegt, an bestimmten Stellen Informationen angeboten bekommt. Diese Informationen können beispielsweise die Namen der Gebäude sein, vor denen sich der Anwender befindet und die über ein tragbares Display als Überlagerung über die entsprechenden Objekte ins Sichtfeld eingeblendet werden. Bei den Informationen kann es sich auch um Pfeile handeln, die die Richtung einer automatisch generierten Navigationsroute anzeigen. Die genaueste Form des Trackings ist notwendig, wenn 3D-Modelle positionsgenau eingeblendet werden sollen. Dies bedeutet, dass beispielsweise dort, wo sich ein Gebäude in der Vergangenheit befand die entsprechende Rekonstruktion eingeblendet werden soll oder zukünftige bauliche Erweiterungen der Objekte als Veränderungen im Blickfeld des Nutzers dargestellt werden. Besonders wenn es sich nur um Teile von Objekten handelt, die passgenau die bestehende Umgebung ergänzen sollen, ist ein hochgenaues Tracking notwendig. Der Anwender kann in diesem Fall sehr genau feststellen, wie präzise das Verfahren ist. Diese Methodik ist Teil des in Kapitel 5 erläuterten Verfahrens, das das gesamte Stadtgebiet abdeckt, allerdings nur an bestimmten Stellen ein hochgenaues Tracking ermöglicht und die Verbindung zu anderen Trackingverfahren und damit anderen Diensten neben dem positionsgenauen Einblenden von virtuellen Informationen erlaubt.

6.1.1 Einsatzgebiet Stadt

Für ein AR-System im Freien, das ein hochgenaues Tracking benötigt, reicht es nicht aus, sich nur auf handelsübliche Sensoren zu verlassen. Wie in Kapitel 2 erläutert, reichen die damit gewonnenen Ergebnisse nicht aus, um die Genauigkeiten zu erreichen, die beispielsweise für das Einblenden von 3D-Modellen ins Blickfeld des Benutzers notwendig sind. Die erhaltenen Ergebnisse müssen auf jeden Fall verbessert werden, um die notwendigen Genauigkeiten einhalten zu können. Zusätzlich ist dafür zu sorgen, dass kontinuierlich vertretbare Ergebnisse geliefert werden. Dies ist bei fast keinem der üblichen Sensoren möglich, die für den Einsatz in AR-Systemen denkbar sind (siehe Kapitel 2). Die Sensoren liefern Werte, die sich

abhängig von äußeren Einflüssen verschlechtern. Es kann ebenfalls vorkommen, dass Ergebnisse insgesamt ausfallen. Aus diesem Grund müssen die Resultate manchmal verbessert, sicherlich aber gestützt und Ausfälle überbrückt werden.

Auf jeden Fall muss es sich um ein hybrides System handeln. Dies ist die Grundlage für alle AR-Systeme, die im Außenbereich lauffähig sein sollen. Das Vertrauen auf nur einen Sensor für ein bestimmtes Ergebnis über einen längeren Zeitraum führt zu falschen Ergebnissen, mangelnder Mobilität oder Ausfällen.

Um höhere Genauigkeiten zu erhalten, ist ein videobasiertes Verfahren notwendig. Die in Kapitel 3 aufgezeigten Verfahren weisen allerdings alle Nachteile auf, die in dem hier entwickelten Verfahren überwunden werden. Das Verfahren soll in einer Stadt eingesetzt werden. Das bedeutet, dass die Objekte, die mit dem Verfahren detektiert werden, verhältnismäßig nahe zum Benutzer stehen. Als Referenz müssen Daten zur Verfügung stehen, die einen Vergleich von Merkmalen auch innerhalb eines einzelnen Objektes erlauben.

6.1.2 Erhalt der Bewegungsfreiheit

Das im Folgenden beschriebene Verfahren geht davon aus, dass in bestimmten Gebieten ein hochgenaues Trackingsystem eingesetzt werden soll. Es ist nicht notwendig, eine detaillierte Modellierung des gesamten Stadtgebietes durchzuführen, sondern es genügt, wenn in den Gebieten, in denen hochgenau getrackt werden soll, eine solche Datenbasis vorhanden ist.

Diese soll genutzt werden, um einem Anwender Informationen anzubieten, der sich frei in der Stadt bewegt. Er soll nicht nur an einem bestimmten Standort Informationen eingeblendet bekommen, sondern sich frei bewegen können. Aus diesem Grund sind einzelne Photos als Referenz hier nicht denkbar. Wenn Daten eines 3D-Modells der Umgebung direkt verwendet werden, ist ein freies Bewegen problemlos möglich.

Um zu bewirken, dass das System möglichst schnell einsatzbereit ist, sollte das Hinzufügen von Sensordaten möglich sein. Diese liefern sehr schnell Ergebnisse, die es erlauben, die Daten der Referenz so einzuschränken, dass nur ein Teil davon zum Vergleich herangezogen wird.

6.1.3 Einblenden von 3D-Modellen

Wenn Informationen zu Objekten in das Sichtfeld einer Person mit Hilfe eines tragbaren Displays eingeblendet werden sollen, muss nicht unbedingt die Position und Blickrichtung des Anwenders bestimmt werden. Es gibt Lösungsmöglichkeiten für eingeschränkte Informationen, für die weniger umfangreiche Berechnungen ausreichend sind. Grundsätzlich können die beiden folgenden Verfahren unterschieden werden.

1. **Bildverschiebung:** Der einfachere Weg ist es, nicht die Position und Blickrichtung des Anwenders zu bestimmen, sondern festzustellen, wie sich das Kamerabild im Vergleich zu den Referenzdaten, die in 2D vorliegen, verschoben und gedreht hat. Ergebnis ist demnach ein Translationsvektor und eine Rotationsmatrix jeweils im zweidimensionalen Raum. Damit ist kein Zugriff auf einen großen externen Datenbestand notwendig. Alle weiteren Informationen müssen auf den Koordinaten basieren, die in der 2D-Referenz verwendet werden. Da es sich hierbei um abgeleitete Bildkoordinaten

handelt, ist eine Anbindung beispielsweise an ein GIS nicht ohne Umwege möglich. Die eingeblendeten Informationen basieren immer auf 2D-Koordinaten.

2. **Tiefeninformation:** Für viele Anwendungen bei der Einblendung von Informationen in das Sichtfeld eines Anwenders reicht die Bestimmung einer zweidimensionalen Verschiebung und -drehung nicht aus, sondern die Bestimmung der Tiefeninformation ist notwendig. Dies bedeutet, dass eine 3D-Koordinatenbestimmung durchgeführt werden muss. Diese erlaubt es, Informationen nicht nur in einer zweidimensionalen Ebene sondern in 3D in das Sichtfeld des Anwenders einzublenden. Die Positions- und Blickrichtungsbestimmung des Anwenders ist die Grundlage für diese Berechnungen. Je nachdem, welches Koordinatensystem zu Grunde liegt, kann direkt oder über eine einfache Koordinatentransformation auf weitere Daten, die beispielsweise in einem GIS enthalten sind, zugegriffen werden.

Die Wahl des Verfahrens ist abhängig von der Art der gewünschten Information. Im vorliegenden Fall sollen auch dreidimensionale Gebäude akkurat platziert werden können. Hier ist auf jeden Fall eine dreidimensionale Berechnung notwendig. Das bedeutet, dass auch die Tiefeninformation und nicht nur eine Bildverschiebung zu berechnen ist.

6.1.4 Schlussfolgerungen

Die oben genannten Anforderungen ergeben zusammenfassend, dass ein videobasiertes Verfahren entwickelt werden muss, um die Ansprüche eines AR-Systems im Außenbereich in einer Stadt zu erfüllen. Um das Verfahren in einer Stadt anwenden zu können, besteht die Anforderung, dass auch nahe Objekte erkannt werden müssen. Die Ergebnisse müssen nicht nur sehr genau sondern auch zuverlässig über einen längeren Zeitraum hinweg ermittelt werden. Der Anwender soll nicht auf einen bestimmten Standpunkt festgelegt sein, sondern sich frei in der Umgebung, für die Referenzdaten vorhanden sind, bewegen können. Für das Einblenden von 3D-Modellen reicht die Berechnung einer zweidimensionalen Verschiebung zwischen Kamerabild und Referenz nicht aus, sondern es muss zusätzlich die Tiefeninformation als Ergebnis vorliegen.

6.2 Überblick über das Verfahren

Grundlage des Verfahrens zum positionsgenauen Einblenden dreidimensionaler Gebäudemodelle im Blickfeld eines Anwenders ist die Bestimmung von Position und Blickrichtung der Person im urbanen Umfeld. Dafür wird ein Kamerabild mit Inhalten eines 3D-Modells verglichen [Coo00b]. Das in diesem Kapitel ausführlich erläuterte Verfahren zeigt Abbildung 31 anschaulich auf.

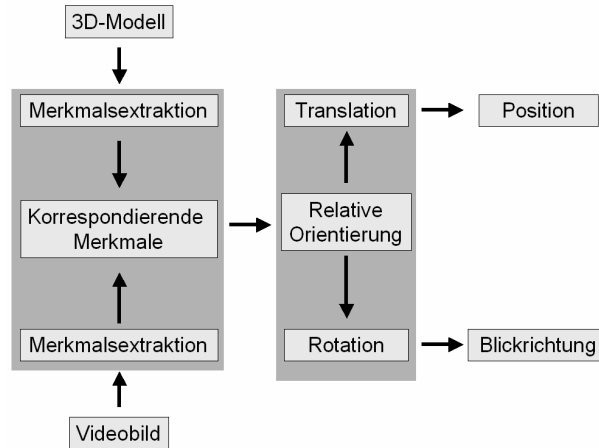


Abbildung 31: Übersicht über die Bildverarbeitungsschritte

Ausgangsdaten für die Bestimmung der Position und Blickrichtung ist das Kamerabild. Dieses repräsentiert das Blickfeld des Benutzers. Das 3D-Modell der Umgebung wird als Referenz herangezogen. Diese beiden Datengrundlagen werden, um die Position und Blickrichtung des Anwenders zu bestimmen, über die folgenden Schritte in Beziehung zueinander gesetzt.

1. **Generieren der Datengrundlage:** Um den Vergleich mit dem Kamerabild zu ermöglichen, wird eine Ansicht aus einem 3D-Modell der Umgebung generiert. Position und Ausrichtung dieser virtuellen Kamera werden über die Sensordaten bestimmt. Gleichzeitig wird ein Bild aus dem Videostrom ausgelesen.
2. **Bildbasierte Merkmalsextraktion:** Über Bildverarbeitungsalgorithmen werden Merkmale aus dem Kamerabild und der generierten Ansicht extrahiert. Merkmale sind in diesem Fall Kanten, die über die Extraktion von Punkten und Helligkeitsunterschieden von Flächen abgeleitet werden können. Ergebnis dieses Schrittes ist eine Vektorliste auffälliger Kanten, die im Kamerabild und in der generierten Ansicht des 3D-Modells enthalten sind.
3. **Korrespondenzfindung:** Das Verfahren der Zuordnung von zwei Bildern aus verschiedenen Blickwinkeln mit dem Ziel der dreidimensionalen Objektbestimmung findet sich in Abläufen der digitalen Photogrammetrie wieder. Es nutzt die Methode der Korrespondenzanalyse, für deren Umsetzung es verschiedene Ansätze gibt [Luhm00]. Aus diesem Grund müssen die in beiden Datenquellen extrahierten Merkmale einander zugeordnet werden. Mit Hilfe der Vektorlisten werden Punkte detektiert, die das gleiche Merkmal darstellen.

4. **Beziehung zwischen Kamerabild und 3D-Modell:** Über die Zuordnung von einzelnen Merkmalen kann die geometrische Zuordnung der Projektionszentren der beiden Kameras berechnet werden. Hierbei handelt es sich um die geometrische Beziehung zwischen dem tatsächlichen Standpunkt des Benutzers, der entsprechenden Blickrichtung, die durch das tatsächliche Kamerabild repräsentiert wird, und der mit Hilfe der Sensorergebnisse ermittelten Situation. Ausgedrückt wird diese Beziehung durch einen dreidimensionalen Translationsvektor und eine dreidimensionale Rotationsmatrix. Hierbei müssen die inneren Parameter der Kamera berücksichtigt werden. Die Kamera-Anordnung muss zum Schluss in das übergeordnete Koordinatensystem, das im folgenden Landeskoordinatensystem genannt wird, eingepasst werden. Dann sind Position und Blickrichtung des Anwenders bestimmt.

Abhängig davon können entsprechende Informationen oder positionsgenaue 3D-Rekonstruktionen in das Sichtfeld eingeblendet werden [Kret02a]. Mit Hilfe von Position und Blickrichtung kann beispielsweise aus dem 3D-Stadtmodell das Gebäude ermittelt werden, welches die Person im Moment betrachtet, und Daten zu diesem Objekt angeboten werden.

Im folgenden Abschnitt 6.3 werden verschiedene Verfahren zur Merkmalsextraktion vorgestellt, wobei besonders auf den SUSAN- und den Canny-Detektor eingegangen wird. Daran anschließend wird in Abschnitt 6.4, nach einer Übersicht über Methoden zur Korrespondenzanalyse, der Algorithmus von Gros et al. zur Bestimmung von korrespondierenden Punkten in zwei Bildern vorgestellt. Mit Hilfe der korrespondierenden Merkmale kann die in Abschnitt 6.5 aufgezeigte Beziehung zwischen Kamerabild und 3D-Modell berechnet werden, aus der die Position und Blickrichtung des Anwenders ermittelt werden kann. Abschnitt 6.6 zeigt vor der Zusammenfassung in Abschnitt 6.7, wie das Verfahren durch Modifikationen abgeändert werden kann.

6.3 Bildbasierte Informationsgenerierung

Für das Einblenden von Informationen in das Blickfeld des Nutzers, die sich auf sichtbare Objekte, wie beispielsweise ein bestimmtes Gebäude, beziehen, würde es eventuell genügen, wenn die realen Objekte im Sichtfeld des Nutzers automatisch erkannt werden und an der entsprechenden Stelle im Display eine Information eingefügt wird. Diese Aufgabe kann durch eine Objekterkennung im Kamerabild erfolgen, das das Blickfeld des Anwenders repräsentiert. Wenn ein Objekt erkannt wird, können die entsprechenden Koordinaten in das Koordinatensystem des Displays umgewandelt werden und so blickpunktgenaue Informationen in das Sichtfeld des Anwenders eingeblendet werden.

Die Bestimmung von Position und Blickrichtung des Nutzers in einem Landeskoordinatensystem kann über die Erkennung von Merkmalen und die Zuordnung zu den entsprechenden Merkmalen, die in einem 3D-Modell vorliegen, erfolgen. Diese Zuordnung basiert dann allerdings nicht auf der Erkennung des Objektes oder seiner Umrisse, sondern auf der Identifizierung von einzelnen, korrespondierenden Merkmalen in beiden Datenbeständen: Kamerabild und 3D-Modell. Diese Merkmalerkennung geschieht auf einem niederen Abstraktionslevel als die richtige Objekterkennung, allerdings ist die Merkmalerkennung oft ein vorbereitender Schritt für die Erkennung von Objekten.

Im vorliegenden Fall müssen Merkmale sowohl im Kamerabild als auch im 3D-Modell erkannt werden. Für diese besondere Form der Merkmalsextraktion müssen die Kamerabilder und die Ansichten aus dem 3D-Modell in einem Bildkoordinatensystem vorliegen. Im Folgenden wird das in Abbildung 32 aufgezeigte Koordinatensystem verwendet.

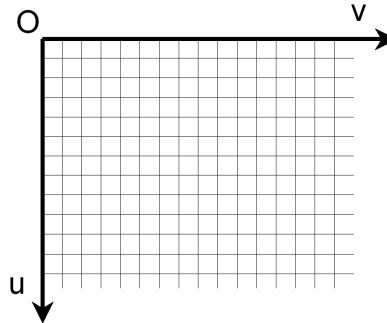


Abbildung 32: Bildkoordinatensystem

Es handelt sich um ein linksdrehendes Koordinatensystem, dessen Ursprung in der linken oberen Ecke liegt. Die Koordinatenachsen werden im Folgenden mit u und v bezeichnet.

Ziel einer Merkmalsextraktion ist es, unnötige Informationen aus einem Bild zu reduzieren. Aus diesem Grund werden geometrische Merkmale detektiert und nur diese weiter betrachtet. Vor dieser Informationsextraktion wird mögliches Rauschen im Bild durch eine Glättung herausgefiltert. Bildrauschen kann u.a. durch die Aufnahmetechnik, durch ein nicht richtig fokussiertes Bild oder durch Unregelmäßigkeiten in der Objektstruktur hervorgerufen werden. Eine Stufenkante wie in Abschnitt 6.3.1 beschrieben, weist dann keinen geraden Helligkeitsverlauf, sondern eine unregelmäßige Kurve auf. Grundsätzlich wird bei der Rauschunterdrückung ein Durchschnittswert benachbarter Helligkeitswerte ermittelt. Diese Operation wird mit einem Filter durchgeführt, der die Durchschnittsberechnung für jedes Pixel durchführt.

6.3.1 Überblick über Verfahren zur Merkmalsextraktion

Die Merkmalsextraktion liefert einen reduzierten Informationsgehalt des Bildes. Dabei wird die Information herausgefiltert, die für die weiteren Schritte der Bildbearbeitung benötigt wird. Fuchs und Heuel unterscheiden, ob Punkte, Kanten oder Regionen mit Hilfe des Algorithmus gefunden werden sollen [Fuch98].

Punkte

Bei Punkten unterscheidet man, ob es sich um von Kanten unabhängige Punkte, um Endpunkte von Linien oder um Schnittpunkte von zwei oder mehr Linien handelt. Eine Gruppe von Punkterkennungsalgorithmen sind Methoden des Template Matchings. Hier wird eine Maske über dem Bild verschoben, die den gesuchten Punkt repräsentiert. Bei hoher Übereinstimmung liegt ein gesuchter Punkt an einer bestimmten Stelle vor. Die Erfolgsrate hängt allerdings von der Art der Muster ab, außerdem sind diese Verfahren abhängig vom Maßstab und von der Rotation, solange es sich nicht um punktsymmetrische Muster handelt.

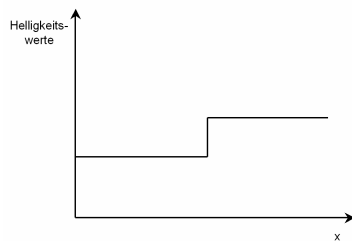
Die Daten eines Bildes können in einer Bildfunktion abgespeichert werden. Diese ist von den zweidimensionalen Koordinaten des Bildes abhängig. Eine weitere Gruppe von Punktverfahren ist dann diejenige, die die Krümmung dieser Bildfunktion nutzt [Fuch98]. Die Krümmung kann durch die Hesse-Matrix berechnet werden. Für die Bestimmung der Hesse-Matrix müssen die zweiten partiellen Ableitungen entlang der beiden Koordinatenachsen ermittelt werden. Nach der Berechnung der Krümmung für jedes Pixel können die detektierten Punkte festgestellt werden.

Interestoperatoren sind v.a. in der Photogrammetrie weit verbreitet. Sie bilden die letzte Gruppe an Verfahren zur Punktdetektion. Dafür werden in einer Maske die Abweichungen zwischen den Pixeln berechnet. Diese Nachbarschaftsbeziehungen in einer Maske bilden die Grundlage für diese Art von Operatoren. Je nach Art des Operators können Aussagen über die Gestalt des Punktes getroffen werden.

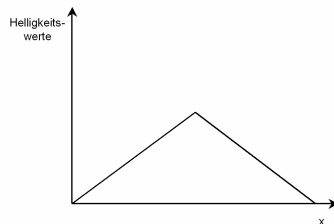
Kanten

Im Gegensatz zu Linien repräsentieren Kanten die Grenze zwischen zwei Regionen. Als Regionen werden Bereiche in einem Bild bezeichnet, die verschieden konstante Helligkeitswerte aufweisen. Bei den Kanten handelt sich um Pixel mit einem starken Helligkeitswechsel. Sie können ebenfalls die Grenzlinie von Objekten beschreiben und zur Segmentierung oder Identifizierung von Objekten dienen. Es gibt verschiedene Kantenarten:

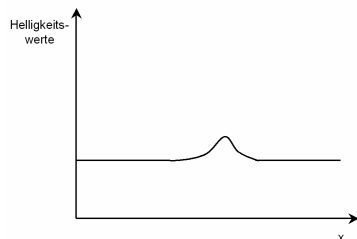
- Stufenkanten



- Dachkanten



- Dornenkanten



Der meistverbreitetste Typ von Kanten ist die Stufenkante. Sie wird durch einen besonders abrupten Helligkeitsunterschied charakterisiert. Die Helligkeitsänderung kann durch die Berechnung des Differentials ermittelt werden. Dafür wird sie als Variable über die beiden Koordinaten festgelegt. Für die Berechnung des Differentials ist eine partielle Ableitung über die beiden Koordinaten notwendig. Die Richtung der Helligkeitsänderung wird mit einem Winkel beschrieben, der über die ersten Ableitungen bestimmbar ist. Der Gradient zeigt in Richtung der stärksten Helligkeitsänderung. Die gesuchten Kanten liegen dann senkrecht dazu. Der Gradient wird ebenfalls über Ableitungen bestimmt. Welche Ableitung für die Erkennung von Kanten angewendet wird, hängt von der Art der Kante ab, die detektiert werden soll. Grundsätzlich ergeben sich der Gradient der Funktion $g(u,v)$ und die Richtung des Gradienten ψ als Winkel zur u -Achse an der Stelle mit den Koordinaten u und v aus folgenden Formeln [Sonk99]:

$$|\text{grad } g(u, v)| = \sqrt{\left(\frac{\partial g}{\partial u}\right)^2 + \left(\frac{\partial g}{\partial v}\right)^2} \text{ und}$$

$$\psi = \arg\left(\frac{\partial g}{\partial u}, \frac{\partial g}{\partial v}\right)$$

Wenn nicht die Richtung, sondern nur der Wert des Gradienten benötigt wird, können die zweiten Ableitungen herangezogen werden. Dies führt zum Laplace-Operator, der richtungsunabhängig ist. Er wird mit folgender Formel definiert [Sonk99]:

$$\nabla^2 g(u, v) = \frac{\partial^2 g(u, v)}{\partial u^2} + \frac{\partial^2 g(u, v)}{\partial v^2}$$

Kantenoperatoren können grundsätzlich in drei Klassen eingeteilt werden [Sonk99]:

1. Operatoren, die die Ableitung über Differenzen annähern: An der Stelle, an der ein großer Helligkeitsunterschied besteht, wird eine Kante detektiert. Da ein Bild mit diskreten Werten und nicht mit Funktionen beschrieben wird, wird dieser Punkt in einem Helligkeitsdiagramm durch eine breite Spitze lokalisiert. Ziel dieses Schrittes ist deshalb die Bestimmung der Position der Spitze. Dies kann entweder dadurch erreicht werden, dass die Spitze verdünnt oder das Maximum hochgenau detektiert wird. Die meisten der existierenden Kantenoperatoren werden dieser Gruppe zugeordnet. Es gibt Differentialoperatoren, mit denen die Richtung des Gradienten bestimmt werden kann. Allerdings muss dann das Bild in mehrere Richtungen gefiltert werden. Die bekanntesten Operatoren sind der Roberts-, der Laplace-, der Prewitt- und der Sobeloperator. Die Größe des Filterkerns für die Differenzierung sollte so klein gewählt werden, dass eine gute Näherung der Ableitung erreicht wird aber gleichzeitig so groß, dass unwichtige Änderungen in der Helligkeitsfunktion unterdrückt werden.
2. Operatoren, die auf der Nullpunktsbestimmung der zweiten Ableitung basieren: Eine hochgenaue Bestimmung dieser Stelle wird durch eine Interpolation erreicht. Die Genauigkeit hängt von dem Signal-Rausch-Verhältnis ab. Hierzu zählt das Verfahren von Marr-Hildreth.

- Operatoren, die die Übereinstimmung mit einem parametrischen Kantenmodell berechnen.

Regionen

In einem Bild werden Regionen durch einen ähnlichen Pixelwert in nächster Umgebung repräsentiert. Um Pixel zu Regionen zu gruppieren, müssen Nachbarschaften und Ähnlichkeiten bestimmt werden. Eine Klassifizierung der Verfahren kann beispielsweise mit Hilfe der Anzahl an Pixeln vorgenommen werden, die zur Gruppierung zu Regionen eingesetzt werden. In diesem Fall unterscheidet man lokale und globale Techniken. Ein anderer Ansatz unterteilt die Verfahren danach, ob Nachbarschaften und Ähnlichkeiten direkt mit Hilfe benachbarter Pixeleigenschaften in einem gemeinsamen Schritt bestimmt werden oder ob die Bestimmung der Ähnlichkeit und der Zusammengehörigkeit in zwei verschiedenen Schritten berechnet werden. Dafür werden über Schwellwerte oder das Zusammenfassen von Pixeln verschiedene Klassen gebildet. Alle Pixel, die zu einer Klasse gehören, können dann gemeinsam betrachtet werden und es ist nicht mehr notwendig, über die einzelnen Pixel im Bild vorzugehen.

Für das hier beschriebene Verfahren der Bestimmung von Position und Blickrichtung einer Person durch den Vergleich eines Kamerabildes mit den Daten eines 3D-Stadtmodells reicht es aus, wenn Punkte und Kanten detektiert werden, die als Ergebnis in die Korrespondenzfindung eingehen. Aus diesem Grund werden im Folgenden je ein Punkt- und ein Kantendetektor kurz vorgestellt. Tiefergehende Informationen sind der angegebenen Literatur zu entnehmen.

6.3.2 SUSAN-Detektor

Der SUSAN-Detektor (Smallest Univalued Segment Assimilating Nucleus) zählt zur Gruppe der Interestoperatoren. Er untersucht die Unterschiede der Helligkeitswerte eines jeden Pixels innerhalb einer kreisförmigen Maske zu der des Maskenmittelpunktes [Smit95]. Daraus kann eine Region abgeleitet werden, die gleiche oder ähnliche Helligkeitswerte wie der Mittelpunkt hat. Diese Region wird USAN (Univalued Segment Assimilating Nucleus) genannt. Hiermit können Ecken gefunden werden, die Geradenschnittpunkte oder Endpunkte sind. Abbildung 33 zeigt die Funktionalität des SUSAN-Detektors entlang eines dunklen Rechtecks.

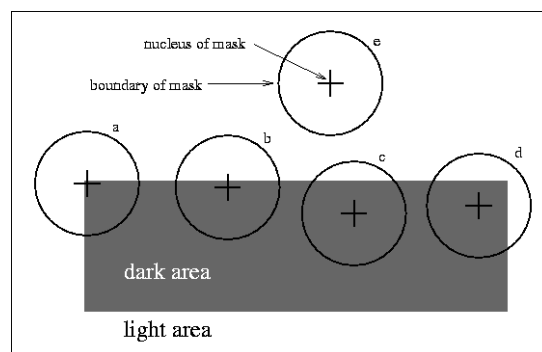


Abbildung 33: SUSAN-Eckendetektor [Smit95]

Die Berechnung findet direkt im Bild statt ohne Strukturen bestimmen zu müssen. Außerdem ist es nicht nötig Ableitungen zu berechnen. Die Maske besteht meist aus 37 Pixeln. Jedes der Pixel in der Maske wird mit dem Mittelpunkt verglichen. Das Ergebnis geht in die im Folgenden aufgeführten Formeln ein [Smit95]:

$$c(\vec{r}, \vec{r}_0) = \begin{cases} 1 & \text{falls } |I(\vec{r}) - I(\vec{r}_0)| \leq t \\ 0 & \text{falls } |I(\vec{r}) - I(\vec{r}_0)| > t \end{cases}$$

$$n(\vec{r}_0) = \sum_{\vec{r}} c(\vec{r}, \vec{r}_0)$$

Hier sind I die Helligkeitswerte der Pixel, r bezeichnet die Lage des zu vergleichenden Pixels, r_0 ist das Pixel im Mittelpunkt, t ist der Grenzwert und c bezeichnet das Ergebnis des Vergleiches. Durch Aufsummieren der Werte für c wird die Größe n des USAN bestimmt.

Um eine Ecke festzustellen wird ein zweiter Grenzwert g festgelegt. Er liegt unterhalb der Hälfte der Maskengröße.

$$R(\vec{r}_0) = \begin{cases} g - n(\vec{r}_0) & \text{falls } n(\vec{r}_0) < g \\ 0 & \text{sonst} \end{cases}$$

Alle so bestimmten Pixel, d.h. diejenigen, deren USAN unterhalb des Grenzwertes liegen, sind nun Kandidaten für eine Ecke. In einem nachfolgenden Schritt werden aus dieser Menge diejenigen entfernt, die nur zufällig als Eckpunkt detektiert worden sind, in Wirklichkeit aber keine Ecke darstellen. Dies kann dadurch verursacht werden, dass dünne Linien durch einen runden Bereich verlaufen, sowie dadurch, dass Rauscheffekte oder komplizierte geometrische Strukturen als Eckpunkte angesehen werden. Das Aufdecken dieser falschen Eckpunkte wird dadurch erreicht, dass die Mitte der Maske und die Mitte des USAN durch eine Linie verbunden werden. Wenn die Länge dieser Linie sehr klein ist oder nicht alle Punkte dieser Linie vollständig zum USAN gehören, wird auch dieses Pixel aus den Kandidaten ausgegliedert.

Zum Schluss werden mit Hilfe aller übrig gebliebenen Kandidaten die Positionen der Eckpunkte bestimmt. Dafür wird die Tatsache genutzt, dass der USAN-Wert abnimmt je größer die Wahrscheinlichkeit ist, dass das entsprechende Pixel einem Eckpunkt entspricht. Alle Regionen, die in mehreren USANs vorkommen, werden pixelweise mit dem USAN-Wert attribuiert. Je näher ein Pixel dem tatsächlichen Eckpunkt ist, desto kleiner ist dessen Wert. Das Pixel mit dem kleinsten Wert innerhalb einer Region wird dann als tatsächlicher Eckpunkt detektiert. Über die Betrachtung der Achter-Nachbarschaft eines Eckpunktkandidats wird so eine Non-Maxima-Suppression durchgeführt. Dies bedeutet, dass ein Kandidat ausgeschlossen wird, wenn der Wert $R(r_0)$ dieses Pixels, der in der oben genannten Formel berechnet wurde, kleiner ist als der eines Nachbarn.

6.3.3 Canny-Detektor

Eine Möglichkeit, Stufenkanten zu detektieren, die durch weißes Rauschen entstellt wurden, ist die Anwendung des Canny-Detektors [Sonk99]. Er basiert auf der Tatsache, dass eine Stufenkante über ihre Position, ihre Orientierung und ihre Stärke ausgedrückt werden kann und besteht aus der Abfolge der folgenden vier Schritte:

1. Zunächst wird auf das Bild ein Gauß-Filter angewendet. Dieser kann durch folgende Formel beschrieben werden [Sonk99]:

$$G(u, v) = e^{-(u^2+v^2)/(2\sigma^2)}$$

Dabei ist σ die Standardabweichung der Wahrscheinlichkeitsverteilung, die proportional zur Größe der Filtermaske ist. Pixel, die weiter entfernt vom Zentrum des Operators sind, haben weniger Einfluss auf das Ergebnis. Die Koordinaten des Pixels, an deren Stelle der Filter angewendet wird, werden wie oben beschrieben mit u und v bezeichnet.

2. Anschließend wird eine gradientenbasierte Ableitung auf das Bild angewendet. Der Wert des Gradienten und der Winkel werden für die weitere Berechnung separat abgespeichert. Dafür werden der Gradient $g(u, v)$ und die Richtung des Gradienten $\psi(u, v)$ mit folgenden Formeln berechnet [Sonk99]:

$$g(u, v) = \sqrt{\Delta u^2 + \Delta v^2}$$

$$\psi(u, v) = a \tan \frac{\Delta u}{\Delta v}$$

mit

$$\Delta u = I(u+1, v) - I(u-1, v)$$

$$\text{und } \Delta v = I(u, v+1) - I(u, v-1)$$

3. Im nächsten Schritt werden die Pixel eliminiert, die lokal keinen maximalen Wert haben. Damit verdünnt man die Kanten auf eine Breite von einem Pixel. Dafür werden die Pixel auf Null gesetzt, deren Gradient kleiner als der Gradient des vorhergehenden Pixels ist. Die Punkte werden in Gradientenrichtung untersucht. Die Methode wird hier ebenfalls Non-Maxima-Suppression-Verfahren genannt, wobei der Vergleich mit zwei aufeinander folgenden Gradientenwerten in Gradientenrichtung auf verschiedenen Seiten des zentral zu untersuchenden Pixels durchgeführt wird [Enca97].
4. Als letzter Schritt des Canny-Detektors werden die bis zu diesem Schritt aufgefundenen Kantenpixel zu Kanten zusammengefasst. Dafür werden zwei Grenzwerte festgelegt. Wenn der Gradientenbetrag eines Pixels über dem höheren Grenzwert liegt, bedeutet das, dass der Punkt eine neue Kontur beginnt. Liegt der Gradientenbetrag eines Pixels dagegen über dem unteren Grenzwert, dient der Punkt der Fortsetzung [Enca97]. Damit gehören zu einer Kante alle Pixel, deren Gradientenwerte über dem oberen Grenzwert oder zwischen den beiden Grenzwerten liegen, wobei letztere gleichzeitig in einer Achter-Nachbarschaft Verbindung zu einer Kante haben müssen.

Der Canny-Detektor liefert damit zusammengehörende Kantenteilstücke. Um diese zu einer Kante zusammenzufassen, wird der Douglas-Peucker-Algorithmus verwendet [Doug73]. Dieser besteht aus den folgenden Schritten:

1. Verbindung von zwei weit auseinander liegenden Endpunkten von Kantenteilstücken durch eine Linie
2. Berechnung der Abstände aller Kantenteilstücke zu der im vorherigen Schritt bestimmten Linie
3. Festlegung des Schnittpunktes als Endpunkt zweier neuer Linien, wenn Abstände größer als ein zuvor bestimmter Grenzwert sind.

Für jede der beiden Linien werden die Schritte 1 bis 3 erneut durchgeführt. Liegen alle in Schritt 2 berechneten Abstände unter dem Grenzwert, werden die Kantenteilstücke zu einer Kante zusammengefasst, deren Endpunkte denjenigen aus Schritt 1 entsprechen.

6.3.4 Übertragung auf das Einsatzgebiet Stadt

Der Algorithmus soll auf Datensätze angewendet werden, die zum einen aus einem 3D-Stadtmodell generiert werden und zum anderen durch die Bilder einer Kamera repräsentiert werden, wobei sich der Anwender in einer urbanen Umgebung aufhält. Die in diesem Bereich wichtigen Objekte sind Gebäude. Aus diesem Grund werden diese für das Tracking genutzt. Sie zeichnen sich durch rechtwinklige Merkmale aus, die mit Hilfe eines Kantendetektors erkannt werden können. Dafür kann beispielsweise der Canny-Kantendetektor gefolgt vom Douglas-Peucker-Algorithmus eingesetzt werden.

Diese Ergebnisse werden in dem in Abschnitt 6.4.3 beschriebenen Verfahren zur Merkmalszuordnung verwendet und sollten deshalb bestmöglich vorbereitet werden. Um die Ergebnisse des Kantendetektors für den Einsatz im urbanen Umfeld zu optimieren, wurden zwei Verfahren entwickelt.

Oft werden Linien bei der Bearbeitung eines Bildes mit Kantendetektoren in Einzelteile zerlegt. Besonders bei Gebäuden kann davon ausgegangen werden, dass sich ein Teil der Linien schneidet. Diese Tatsache wird genutzt, um Linienteilstücke über Schnittpunkte mit anderen Linien zu verbinden. Der Nachteil dieses Verfahrens besteht darin, dass Artefakte entstehen können.

Aus diesem Grund wurde ein zweites Verfahren entwickelt, das die Ergebnisse des Kantendetektors mit denen eines Eckpunkt-Detektors koppelt. Dabei werden Punkte gesucht, die sich in Richtung einer Kante befinden. Es wird davon ausgegangen, dass ein Großteil der Kanten nicht bis zum Eckpunkt detektiert wird. Diese Lösung ermöglicht es, die Kanten bis zu den Eckpunkten zu verlängern.

6.4 Korrespondenzfindung im 3D-Modell und Kamerabild

Die Zuordnung von Merkmalen eines Objektes in zwei verschiedenen Datenquellen ist eine Funktionalität, die für viele Aufgaben der Bildverarbeitung notwendig ist. Im vorliegenden Fall werden korrespondierende Punkte benötigt, um die geometrische Beziehung zwischen dem Kamerabild und dem 3D-Modell herzustellen. Abbildung 34 zeigt das Bildmaterial, aus dem korrespondierende Punkte bestimmt werden sollen. Links ist das Bild einer Kamera, rechts die generierte Ansicht aus einem 3D-Stadtmodell zu sehen. Die über eine bildbasierte Erkennung bestimmten Merkmale in beiden Bildern dienen als Grundlage der Zuordnung von einzelnen Punkten. Dies führt zur Detektion übereinstimmender Merkmale, die mit dem menschlichen Auge verhältnismäßig einfach aufzufinden sind.



Abbildung 34: Vergleich Kamerabild mit Modell

6.4.1 Überblick über Matching-Verfahren

Die detektierten Merkmale müssen im Schritt der Korrespondenzfindung einander zugeordnet werden. In [Lang95] wird zwischen gebietsbasierten und merkmalsbasierten Verfahren unterschieden.

Gebietsbasierte Verfahren versuchen eines der Bilder so zu verschieben und zu verzerren, dass die Helligkeitswerte des einen Bildes genau mit denen des anderen übereinstimmen. Dafür werden meist entweder die Ähnlichkeit oder die Distanz gemessen. Zu diesen Verfahren zählen die Kreuzkorrelation und die Kleinste-Quadrate-Zuordnung.

Für die merkmalsbasierten Zuordnungsverfahren muss eine symbolische Beschreibung der Bilder vorliegen. Die Berechnung findet demnach nicht mit Hilfe der Pixelwerte, sondern mit Hilfe detektierter Merkmale statt. Hierzu gehören robuste Schätzverfahren aber auch die Zuordnung über Merkmalslisten sowie relationale Zuordnungen.

Bei letzteren werden die Inhalte der Bilder, wie Kanten oder Regionen, mit Hilfe von Primitiven beschrieben und zusätzlich die topologischen und geometrischen Relationen zwischen den Primitiven in die Darstellung der Bildeigenschaften aufgenommen. Bei einer so beschriebenen Relation handelt es sich um die Beziehung zwischen bestimmten Primitiven, die mit Attributen ergänzt werden können. Die relationale Zuordnung ermöglicht es, eine Abbildung zu berechnen, die die Relationen und Primitive der einen Beschreibung den entsprechenden der

anderen Beschreibung zuordnet. Die relationale Beschreibung lässt sich mit Hilfe eines relationalen Graphen darstellen. Hier stellen die Knoten die Primitive und die Kanten die Beziehungen zwischen zwei Primitiven dar. Ziel dieser Verfahren ist es, eine optimale Abbildung zwischen den Beschreibungen der beiden korrespondierenden Bilder zu finden. Bei der Korrespondenzfindung werden dann entsprechende Suchmethoden angewendet. Der Einsatz eines relationalen Zuordnungsverfahrens für den Bereich Tracking wurde von Borotschnig et al. vorgestellt [Boro97].

Im Gegensatz zur Zuordnung über Merkmalslisten werden bei der relationalen Zuordnung keine räumlichen Beziehungen benötigt, da die topologischen und geometrischen Relationen zwischen einzelnen Primitiven bekannt sind. Wenn, wie in Abbildung 34 beispielhaft gezeigt, ein Kamerabild mit der Ansicht eines Modells verglichen wird, das aufgrund von Näherungen der tatsächlichen Position und Blickrichtung generiert wurde, stimmen die beiden Graphen meist nicht überein. Der Grund sind die abweichende Position und die unterschiedliche Blickrichtung, die zu einem ungleichen Blickwinkel oder zu Verdeckungen führen können. Außerdem hängt ein Großteil der Genauigkeit des Matchingergebnisses von der zuvor durchgeführten Merkmalsextraktion ab. Dadurch können bei verschiedenen Datensätzen auch verschiedene Merkmale erkannt werden. Auch führt der unterschiedliche Blickpunkt zu Relationen, die nur für diese Abbildung gelten aber nicht mit der Realität übereinstimmen. Dazu zählen zwei Linien, die in der Realität weit entfernt voneinander sind, sich in den beiden Abbildungen aber in einem Punkt schneiden.

Die allgemeine Hough-Transformation ist ebenfalls ein Beispiel für ein merkmalsbasiertes Verfahren. Das in Abschnitt 6.4.3 näher beschriebene Verfahren nach Gros et al. [Gros95] ist ein Beispiel für eine Hough-ähnliche Transformation zur Korrespondenzfindung mit Hilfe merkmalsbasierter Zuordnungsverfahren.

6.4.2 Anforderungen an die Korrespondenzfindung

Ziel des aufgezeigten Verfahrens ist die Bestimmung von Position und Blickrichtung eines Anwenders, der sich mit einer AR-Ausrüstung in einer Stadt bewegt. Für diese hochgenaue Bestimmung sind Sensordaten als Näherungswerte, eine Kamera, die die reale Situation, wie sie der Anwender sieht, wiedergibt, und ein 3D-Modell der Umgebung vorhanden. Damit werden die in Abschnitt 6.1 aufgeführten Anforderungen an ein urbanes Trackingsystem erfüllt. Allerdings führen diese unterschiedlichen Datensätze – Kamerabild und 3D-Modell – zu weiteren Anforderungen, die an ein stabiles Verfahren gestellt werden müssen.

Mit Hilfe der Korrespondenzfindung muss es möglich sein, Daten eines 3D-Modells mit einem Kamerabild zu vergleichen. Da es sich um das Einsatzgebiet Stadt handelt, werden hauptsächlich Häuser auf den Kamerabildern abgebildet. Auch in den Daten des 3D-Stadtmodells sind vornehmlich Gebäude enthalten. Hier werden bei weiteren Objekten wie Straßenlaternen oder Bäumen meist nur die Koordinaten des Objektmittelpunktes und die Ausmaße abgespeichert. Mit diesen Informationen kann dann bei der Visualisierung aus einer Objektdatenbank an der entsprechenden Stelle eine Instanz gebildet werden, die allerdings allgemeinen Visualisierungsregeln entspricht. Die eigentliche Modellierung, die für die Korrespondenzfindung notwendig ist, wird nur für Gebäude durchgeführt. In den Daten des 3D-Modells können die zugrunde liegenden Details in der Regeln nicht, manchmal aber auch in abgewandelter Form vorhanden sein. Aus diesem Grund muss ein Korrespondenzverfahren gewählt werden, das Merkmale verwendet, die in beiden Datensätzen

extrahiert werden können. Gebäudetexturen erleichtern in vielen Fällen die Erkennung von Objekten. Der sichtbare Eindruck verändert sich aber leicht unter einem veränderten Blickwinkel oder unterschiedlichen Lichtverhältnissen.

Aus diesem Grund sollte ein Verfahren gewählt werden, das auf den geometrischen Strukturen der Gebäude basiert. Allerdings können diese Strukturen auch bei ähnlichen Datensätzen bei der Merkmalsextraktion in unterschiedlich große Segmente zerlegt werden, so dass zusätzlich darauf zu achten ist, dass das Verfahren unterschiedliche Ergebnisse in zwei Datensätzen bei der Merkmalsextraktion berücksichtigt.

Das Ergebnis der Korrespondenzfindung sollte zuverlässig sein. Mit einzelnen Ausreißern ist zu rechnen, allerdings ist der anschließende Schritt der Bestimmung der Position und Ausrichtung der Kameras zueinander nur möglich, wenn ein Großteil der Ergebnisse aus dem Korrespondenzfindungs-Schritt korrekt ist.

Mit Hilfe der Sensoren sollten die Näherungswerte so genau wie möglich ermittelt werden. Dann stimmen der Inhalt des Kamerabildes und der des Datensatzes, der aus dem 3D-Modell generiert wurde, möglichst weitgehend überein, so dass ein bestmögliches Matchingergebnis erreichbar ist. Wenn der Unterschied des Blickwinkels zu groß wird, ist es schwieriger eine genügend große Anzahl an korrespondierenden Punkten zu finden. In diesem Fall wird empfohlen, mehrere Ansichten mit unterschiedlicher Position und veränderndem Blickwinkel zu generieren. Dies führt allerdings zu einer Zunahme des Berechnungsaufwandes.

6.4.3 Überblick über die Korrespondenzfindung aufgrund von Segmentkonfigurationen

Die in Abschnitt 6.4.2 genannten Anforderungen erfüllt das von Gros et al. [Gros95] entwickelte Matching-Verfahren. Es geht davon aus, dass alle korrespondierenden Merkmale in zwei zugeordneten Bildern durch eine geometrische Transformation aufeinander abgebildet werden können. Die Berechnung dieser Transformation erfolgt mit Hilfe einer modifizierten Hough-Transformation. Das Verfahren nähert die Transformation zwischen den beiden Bildern durch eine Ähnlichkeitstransformation an und ist aus diesem Grund nicht anwendbar, wenn zwischen den beiden Bildern große projektionsbedingte Effekte auftreten. Es weist allerdings die folgenden Vorteile auf:

- Robustheit gegenüber Bildrauschen
- Ermittlung von nur wenigen falschen Korrespondenzen
- Berücksichtigung der Verschiebung von erkannten Merkmalen, da keine minimalen Abstände zwischen den korrespondierenden Merkmalen benötigt werden
- Weniger Rechenaufwand im Gegensatz zu Verfahren mit normaler Hough-Transformation oder relationalen Zuordnungsverfahren.

Bei einer solchen Transformation kann vorausgesetzt werden, dass sich bestimmte Größen, die in den Bildern gemessen werden, nicht verändern. Da es sich um eine mathematisch einfach zu beschreibende Transformation handelt, die jedoch nur eine Näherung der tatsächlichen Transformation darstellt, werden diese unveränderten Größen als Quasi-Invarianten bezeichnet.

Die Ähnlichkeitstransformation kann durch vier Parameter beschrieben werden:

- Skalierung s
- Rotation R um den Winkel β
- Translationsvektor mit den Komponenten t_u und t_v

Mit Hilfe dieser Parameter wird ein Punkt P' in dem einen Bild auf den Punkt P'' in dem anderen Bild mit Hilfe der folgenden Gleichung zugeordnet:

$$P'' = s(R \cdot P') + \begin{pmatrix} t_u \\ t_v \end{pmatrix}$$

Es sei noch einmal darauf hingewiesen, dass es sich bei der Ähnlichkeitstransformation nur um eine Annäherung an die Realität handelt. Die tatsächliche Transformation würde alle Merkmale des einen Bildes auf die korrespondierenden Merkmale des anderen Bildes transformieren. Tests von Gros et al. haben aber gezeigt, dass mit Hilfe der Ähnlichkeitstransformation eine ausreichende Anzahl an Punktkorrespondenzen gefunden werden kann. Im Gegensatz zu einer affinen oder projektiven Transformation hat die Ähnlichkeitstransformation den Vorteil, dass sie unempfindlicher gegenüber Bildrauschen ist. Außerdem ist der in Abschnitt 6.4.5 beschriebene Schritt zur Berechnung der Quasi-Invarianten einfacher und die anschließende Suche im Parameterraum durch die reduzierte Anzahl der Parameter leichter.

6.4.4 Vorbereitung der Datensätze

Als erster Schritt werden in dem Algorithmus nach Gros zunächst so genannte Segmentkonfigurationen aus den extrahierten Kanten berechnet (siehe Abbildung 35). Diese bestehen aus zwei Linien, die sich in einem Punkt schneiden.

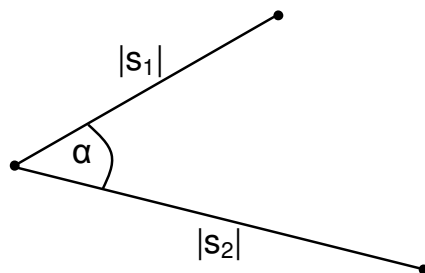


Abbildung 35: Segmentkonfiguration nach Gros et al.

Die Linien werden als s_1 und s_2 mit den Längen $|s_1|$ und $|s_2|$ der entsprechenden Vektoren bezeichnet. Der Winkel α liegt am Schnittpunkt zwischen diesen beiden Linien. Die Vektoren, die in dem zuvor durchgeführten Schritt der Merkmalsextraktion bestimmt wurden, sind für eine Bildung der hier benötigten Segmentkonfigurationen einer Vorverarbeitung zu unterziehen. Vor allem durch Bildrauschen, aber auch durch die Begrenztheit der Merkmalsextraktion werden lange Kanten in Teilkanten zerlegt oder zu kleine Kanten als signifikant angesehen. Auch eine Nachbearbeitung mit Hilfe des Douglas-Peucker-Algorithmus und der Kombination mit einem Punktdetektor führen nicht bei allen Vektoren zu einem Ergebnis, das in dem hier beschriebenen Zuordnungsverfahren genutzt werden kann. Für das Zuordnungsverfahren können Konfigurationen, die aus einem langen und einem kurzen Segment bestehen, wichtig sein. Aus diesem Grund ist eine Bewertung nicht der Vektoren an sich, sondern einzelner Segmentkonfigurationen von Vorteil. Es empfiehlt sich, diese Signifikanz über die Summen der beiden Vektorenlängen, die eine Segmentkonfiguration bilden, zu ermitteln. Dadurch werden Konfigurationen mit kurzen Vektoren höher bewertet, die zusammen mit einem langen Vektor eine Konfiguration bilden. Segmentkonfigurationen, die aus zwei kurzen Kanten bestehen, werden als nicht signifikant bewertet und gehen in das Korrespondenzverfahren nicht ein.

Um zu entscheiden, ob eine Segmentkonfiguration signifikant ist, muss ein Grenzwert festgelegt werden. Es ist ratsam, dafür nicht das arithmetische Mittel heranzuziehen, da eine einzige sehr lange Kante den Wert zu sehr verfälschen kann. Es wird empfohlen, dafür das p -Quantil für alle ermittelten Signifikanzwerte des Bildes zu berechnen. Dieser Wert sagt aus, dass mindestens $p \cdot 100$ % aller anderen Signifikanzwerte nicht größer und mindestens $(1-p) \cdot 100$ % dieser Werte nicht kleiner als dieser Grenzwert sind. Der Wert des p -Quantils hängt von der Qualität der extrahierten Kanten und somit vom Bildrauschen und der vorhandenen Perspektive ab. Er beeinflusst die Anzahl der im Matchingverfahren zu verwendenden Segmentkonfigurationen. Diese sollen auf signifikante Konfigurationen eingeschränkt werden, aber gleichzeitig nicht dazu führen, dass zu wenige Konfigurationen zur Verfügung stehen. Aus diesem Grund sollte die Anzahl der nach dem Signifikanztest über das p -Quantil übrig gebliebenen Konfigurationen daraufhin getestet werden, ob die Anzahl über einem Mindestwert und unter einem Höchstwert liegt. Genauso wie die Größe des Parameters p sollten die beiden Grenzwerte über Testdatensätze abhängig von den Verfahren für die Merkmalsextraktion und das Matching bestimmt werden.

6.4.5 Berechnung der Quasi-Invarianten

Für jede der Segmentkonfigurationen werden die Quasi-Invarianten berechnet. Dies sind Größen, die sich nach der Transformation von einem Bild in ein anderes nicht ändern. Sie werden als Quasi-Invarianten bezeichnet, da es sich nur um eine Näherung der tatsächlichen Transformation handelt. Die folgenden Größen ändern sich nach einer Transformation für kleine Bereiche nicht (vergleiche auch Abbildung 35):

- Der Winkel α zwischen den beiden Linien s_1 und s_2
- Das Längenverhältnis r zwischen den beiden Linien s_1 und s_2

Die Invarianten werden mit Hilfe der folgenden Formeln berechnet:

$$\alpha = \arccos \frac{\vec{s}_1 \cdot \vec{s}_2}{|\vec{s}_1| \cdot |\vec{s}_2|}$$

$$r = \frac{|s_1|}{|s_2|}$$

Der Winkel α wird so gewählt, dass er spitz und positiv ist. Die Quasi-Invarianten für jedes Bild werden nach diesem Winkel sortiert.

Das Bildrauschen beeinflusst die Länge und den Winkel der extrahierten Merkmale. Aus diesem Grund sind Schwankungsbereiche für die Zuordnung der Quasi-Invarianten zulässig. Diese werden durch Grenzwerte für den Winkel α und die Beziehung der Längenverhältnisse des ersten Bildes r' zu denen des zweiten Bildes r'' und umgekehrt ausgedrückt. Sie orientieren sich am Bildrauschen. Zu jeder Quasi-Invariante des ersten Bildes werden Quasi-Invarianten aus dem zweiten Bild bestimmt, die einen ähnlichen Wert aufweisen. Dazu werden die folgenden drei Abfragen durchgeführt:

- Liegt die Differenz der beiden Winkel für die Segmentkonfiguration in zwei Bildern unterhalb des festgelegten Grenzwertes?
- Liegt das Verhältnis $\frac{r''}{r'}$ unterhalb des festgelegten Grenzwertes?
- Liegt das Verhältnis $\frac{r'}{r''}$ unterhalb des festgelegten Grenzwertes?

Wenn alle drei Bedingungen für ein Paar an Segment-Konfigurationen erfüllt sind, liegen zwei potentiell korrespondierende Segmentkonfigurationen vor.

6.4.6 Bestimmung der Transformationsparameter

Für alle diese Paare werden im darauf folgenden Schritt die Parameter der ebenen Ähnlichkeitstransformation bestimmt. Wie oben erwähnt, stellt die Ähnlichkeitstransformation eine Approximation der tatsächlichen Transformation dar.

Einen Überblick über die im Folgenden verwendeten Größen gibt Abbildung 36.

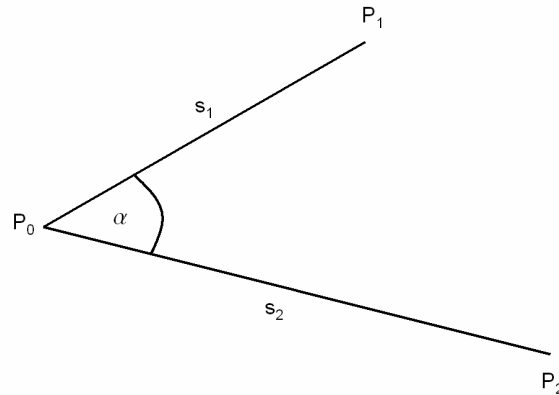


Abbildung 36: Größen der Segmentkonfigurationen

Die Segmentkonfigurationen umfassen allgemein drei Punkte, die mit P₀, P₁ und P₂ bezeichnet werden. Diejenigen für das erste Bild werden dann durch die Bezeichnungen P₀' , P₁' und P₂' von denen des zweiten Bildes mit P₀'' , P₁'' und P₂'' unterschieden. Das Gleiche gilt für die Linien s₁ und s₂, die im ersten Bild mit s₁' und s₂' und im zweiten Bild mit s₁'' und s₂'' bezeichnet werden, sowie für den Winkel alpha.

Um die korrespondierenden Merkmale eindeutig zuordnen zu können, müssen die Parameter der Ähnlichkeitstransformation ermittelt werden. Diese in Abschnitt 6.4.3 erwähnten Parameter lassen sich mit Hilfe der folgenden Formeln berechnen:

$$\text{Mittelung der Streckenverhältnisse } s = \frac{1}{2} \left(\frac{|s_1''|}{|s_1'|} + \frac{|s_2''|}{|s_2'|} \right)$$

$$\text{Rotationswinkel } \beta = B(s_1', s_1'') + \frac{1}{2} [\alpha'' - \alpha']$$

Dabei bezeichnet B(s₁', s₁'') den gerichteten Winkel zwischen den Vektoren s₁' und s₁''.

$$\text{Translationsvektor } \begin{pmatrix} t_u \\ t_v \end{pmatrix} = \begin{pmatrix} s \cdot (\cos \beta \cdot u_0 - \sin \beta \cdot v_0) - u_0 \\ s \cdot (\sin \beta \cdot u_0 + \cos \beta \cdot v_0) - v_0 \end{pmatrix}$$

Der Schnittpunkt P₀ der Linien, die eine Segmentkonfiguration bilden, ist leichter mit der Merkmalsextraktion zu detektieren als die beiden Endpunkte P₁ und P₂. Dies hängt damit zusammen, dass längere Vektoren durch Bildrauschen in einzelne Segmente aufgeteilt werden. Die Länge der Linien unterscheidet sich daher in vergleichbaren Bildern. Im Gegensatz dazu betrifft dies den Schnittpunkt der Linien nicht. Deshalb wird die Translation nur über diesen Punkt berechnet.

Diese vier Transformationsparameter werden für alle potentiellen Matchingpaare berechnet.

6.4.7 Bestimmung der Ähnlichkeitstransformation

Aus allen diesen potentiellen Zuordnungen müssen in dem hier beschriebenen Schritt diejenigen Paare ermittelt werden, die die tatsächliche Korrespondenz zwischen zwei Segmentkonfigurationen in zwei verschiedenen Bildern darstellen. Dabei wird davon ausgegangen, dass alle tatsächlichen Paare mit ungefähr der gleichen Ähnlichkeitstransformation ineinander überführt werden können. Potentielle Paare, die in Wirklichkeit allerdings gar keinen korrespondierenden Punkten entsprechen, weisen Parameter für die Transformation auf, die sich vom Großteil der Parameter für die anderen Paare sehr stark unterscheiden.

Aus diesem Grund wird jeder Parametersatz für ein potentielles Paar, der aus vier Werten besteht, als Punkt im Parameterraum R^4 ausgedrückt. Dieser besitzt die Koordinaten t_u , t_v , β und s , wobei diese Werte für jedes potentielle Paar berechnet werden. Für alle diese Punkte wird nun die Menge der Punkte ermittelt, die sich in nächster Nachbarschaft dieses Punktes befinden. Dafür werden für alle vier Parameter Grenzwerte festgelegt. Diese spannen eine Region um jeden Punkt auf. Mit Hilfe der Überprüfung aller anderen Punkte wird festgestellt, ob sich weitere Punkte innerhalb dieser Region befinden. Ist dies der Fall, d.h. liegen alle Koordinaten eines Punktes innerhalb der festgelegten Grenzen um den Punkt, wird die Menge, die die Punkte in der Umgebung eines Punktes festlegt, um eins erhöht.

Der Punkt mit der zahlenmäßig größten Menge stellt die beste Näherung der Ähnlichkeitstransformation dar. Das bedeutet, dass die Koordinaten dieses Punktes die Parameter der Ähnlichkeitstransformation repräsentieren. Punkte, die nicht in direkter Nachbarschaft zu diesem Punkt liegen, können als Ausreißer betrachtet werden und entsprechen keinen korrespondierenden Punkten. Die Segmentkonfigurationen aller anderen Punkte können als korrespondierende Punkte gewertet werden. Dies zeigt, dass die vier Grenzwerte so gewählt werden müssen, dass die Anzahl an korrespondierenden Segmentkonfigurationen groß genug ist. Allerdings sollte die Wahl nicht dazu führen, dass Ausreißer nicht erkannt werden.

Dieser Schritt entspricht der Einschränkung des Parameterraumes über eine Hough-Transformation. Dabei wurde davon ausgegangen, dass die Transformationsparameter der meisten Paare ähnliche Werte annehmen, die Ausreißer dagegen weit davon entfernt liegen. Alle Punktepaare, die in diesem eingeschränkten Gebiet liegen, werden als richtig angenommen. Für diese Punkte liegt eine Merkmalszuordnung vor.

6.4.8 Zuordnung der Bildmerkmale

Im Folgenden werden die Paare weiter betrachtet, die sich innerhalb der Region um den Punkt mit der größten Menge an Nachbarpunkten befinden. Jedes Paar einer Segmentkonfiguration bildet drei mögliche korrespondierende Punktepaare: (P_0', P_0'') , (P_1', P_1'') und (P_2', P_2'') . Diese Punktepaare können auch in weiteren Segmentkonfigurationen auftreten. Aus diesem Grund werden für jeden Punkt des ersten Bildes alle korrespondierenden Punkte des zweiten Bildes bestimmt. Wenn Punkte mehrfach auftreten, werden sie mit ihrer Häufigkeit gewichtet. Der Punkt mit dem größten Gewicht wird als korrekter korrespondierender Punkt des zweiten Bildes bestimmt. Die daraus resultierenden nicht korrekten Punktepaare werden nicht weiter betrachtet, ebenso wie die dazugehörigen Segmentkonfigurationen.

Anschließend können die korrespondierenden Kanten ermittelt werden. Dafür werden die korrespondierenden Endpunkte für jede Kante einzeln ermittelt. Wenn für einen der Kanten-Endpunkte kein korrespondierender Punkt gefunden werden

konnte, liegt keine korrespondierende Kante vor. Ansonsten ermittelt man die korrespondierende Kante über die Suche nach einer Kante im zweiten Bild, die die korrespondierenden Endpunkte verbindet.

6.5 Geometrische Beziehung zwischen Kamerabild und 3D-Modell

Um die tatsächliche Blickrichtung zu bestimmen, muss eine geometrische Beziehung zwischen dem Kamerabild und dem 3D-Modell aufgebaut werden. Ausgangspunkt für den Aufbau der geometrischen Beziehung ist die Bestimmung von Position und Ausrichtung von zwei Kameras mit Hilfe der entsprechenden Bilder, die beide auf den gleichen Ausschnitt der Umgebung gerichtet sind. Diese Bestimmung geschieht über die korrespondierenden Bildpunkte, die zuvor für jedes Bild ermittelt wurden. Das Verfahren der Korrespondenzfindung ist im vorhergehenden Abschnitt 6.4 ausführlich erläutert worden. Abbildung 37 zeigt die geometrische Beziehung.

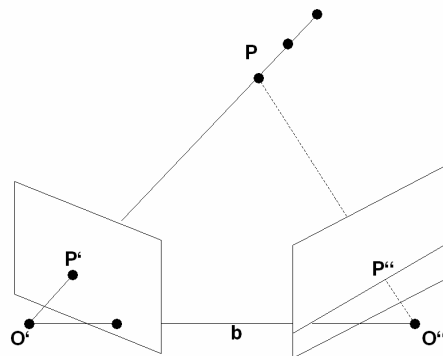


Abbildung 37: Epipolargeometrie korrespondierender Punkte

Hier sind die beiden Kameras als Punkte dargestellt und mit O' und O'' bezeichnet. Die Strahlen zu einem Objektpunkt P schneiden die Bildebenen in den Punkten P' und P'' . Für das vorliegende Verfahren ist es notwendig, die geometrischen Beziehungen mit Hilfe dieser korrespondierenden Punkte P' und P'' wiederherzustellen. Dadurch werden die Positionen und Ausrichtungen der beiden Kameras zueinander und in Bezug zu den wirklichen Objekten berechnet. Ergebnis ist dann die Translation und die Rotation, um die Perspektive der einen Kamera in die der anderen überführen zu können. Daraus ergibt sich die Bestimmung von Position und Blickrichtung des Anwenders, die durch die Position und Ausrichtung einer der Kameras repräsentiert werden.

Die geometrische Beziehung zwischen den korrespondierenden Punkten zweier Bildebenen und den entsprechenden Objektpunkten wird als Epipolargeometrie bezeichnet. Sie ist unabhängig vom Szenenaufbau und hängt nur von den inneren Parametern der Kameras und der relativen Orientierung zwischen den beiden Kameras ab [Hart02]. Um diese Geometrie mit Hilfe der beiden Bildebenen herzustellen und dadurch Position und Blickrichtung des Anwenders zu bestimmen,

müssen zunächst die inneren Parameter der Kameras bestimmt werden. Dies ist in Abschnitt 6.5.1 dargestellt. Anschließend wird die relative Orientierung der Kameras in Abschnitt 6.5.2 berechnet. Mit diesen Ergebnissen, kann dann die Lage der Kameras im übergeordneten Landeskoordinatensystem ermittelt werden (Abschnitt 6.5.3).

6.5.1 Bestimmung der inneren Kameraparameter

Eine Bestimmung der inneren Kameraparameter ist erforderlich, da in die Berechnung der Geometrie Abbildungsgleichungen eingehen, die die Beziehung zwischen der realen Umgebung und den Bildebenen repräsentieren. Diese Abbildungsgleichungen werden durch verschiedene Kameraparameter beeinflusst. Die Parameter der inneren Orientierung sind der Bildhauptpunkt, die Brennweite sowie Parameter von Bildfehlern beschreibenden Funktionen [Luhm00]. Diese Parameter können anhand des Lochkameramodells erklärt werden [Hart02].

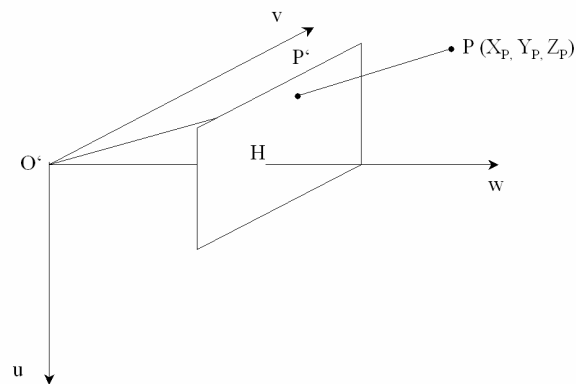


Abbildung 38: Lochkameramodell

Hierbei werden die dreidimensionalen Objektkoordinaten X_P , Y_P und Z_P auf die Bildpunkte $u_{P'}$ und $v_{P'}$ abgebildet. Dies kann über folgende Formeln ausgedrückt werden:

$$u_{P'} = \frac{X_P f}{Z_P}$$

$$v_{P'} = \frac{Y_P f}{Z_P}$$

Die Variable f entspricht der Brennweite, dem Abstand vom Kamerazentrum der ersten Kamera O' zur Bildebene. Bei der Verwendung von homogenen Koordinaten können diese Abbildungsgleichungen durch die im folgenden aufgeführte Gleichung ausgedrückt werden:

$$\begin{pmatrix} u_{p'} \\ v_{p'} \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X_p \\ Y_p \\ Z_p \\ 1 \end{pmatrix}$$

Die Matrix, die den 4x1-Vektor in einen 3x1 Vektor überführt wird als intrinsische Kameramatrix bezeichnet. Die linke Seite kann mit einem Skalierungsfaktor multipliziert werden ohne dass sich die Gleichung ändert. Dieser ist in der oben angegebenen Gleichung mit eins gleichgesetzt.

Der Durchstoßpunkt der optischen Achse durch die Bildebene wird als Hauptpunkt H bezeichnet. Er stimmt in der Realität nicht mit dem Bildmittelpunkt überein. Abbildung 39 zeigt die ideale Lage des Hauptpunktes H mit den Koordinaten u_0, v_0 in der Mitte des Kamerabildes.

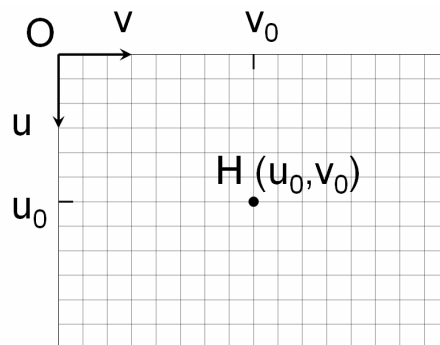


Abbildung 39: Bild- und Kamerakoordinatensystem

Wie bei [Hart02] erläutert, wird beim Lochkamera-Modell der Eindruck vermittelt, dass die Koordinaten in beide Richtungen gleich skaliert sind. Im Fall von CCD-Kameras ist es möglich, dass die Pixel nicht quadratisch sind. Wenn dies berücksichtigt wird, ergibt sich die Kameramatrix A mit den Parametern der Hauptpunktlage u_0, v_0 zu folgender Form:

$$A = \begin{pmatrix} s_u & 0 & u_0 & 0 \\ 0 & s_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Die Parameter s_u und s_v entsprechen den Skalierungsfaktoren in u- und v-Richtung. Sie beschreiben die Vergrößerung des Linsensystems in u- und v-Richtung. Diese Größen werden zumeist in der Einheit Pixel/mm angegeben. Sie lassen sich direkt aus der Brennweite f und der Größe eines Pixels k_u und k_v auf dem Sensorchip berechnen:

$$s_u = f k_u$$

$$s_v = f k_v$$

Mit Hilfe dieser Parameter kann man den Bildpunkt im Bildkoordinatensystem nach folgenden Formeln berechnen:

$$u_{p'} = \frac{s_u \cdot u'}{w'} + u_0$$

$$v_{p'} = \frac{s_v \cdot v'}{w'} + v_0$$

Zusätzlich zu diesen Parametern spielen Verzerrungen bzw. Abberationen eine Rolle. Diese Fehler können durch Translationsvektoren beschrieben werden [Brow71]. Dabei handelt es sich um einen radialen Translationsvektor \vec{t}_R und um einen tangentialen Translationsvektor \vec{t}_T . Die Komponenten der Vektoren ergeben sich mit Hilfe der Parameter k_1 , k_2 , k_3 und k_4 zu

$$\vec{t}_R = \begin{pmatrix} u_{p'}(k_1(u_{p'}^2 + v_{p'}^2) + k_2(u_{p'}^2 + v_{p'}^2)^2) \\ v_{p'}(k_1(u_{p'}^2 + v_{p'}^2) + k_2(u_{p'}^2 + v_{p'}^2)^2) \end{pmatrix}$$

$$\vec{t}_T = \begin{pmatrix} 2k_3 u_{p'} v_{p'} + k_4((u_{p'}^2 + v_{p'}^2) + 2u_{p'}^2) \\ 2k_4 u_{p'} v_{p'} + k_3((u_{p'}^2 + v_{p'}^2) + 2v_{p'}^2) \end{pmatrix}$$

Alle Parameter können mit Hilfe einer Kamerakalibrierung bestimmt werden. Diese wird für die eingesetzte Kamera durchgeführt, wobei davon auszugehen ist, dass die mit Hilfe des 3D-Modells gerenderten Bilder auf künstlich erzeugten projektiven Abbildungen basieren. Deshalb ist hierfür keine Kalibrierung notwendig. Die Kalibrierung beruht auf der Berechnung der intrinsischen und teilweise der gleichzeitigen Berechnung der extrinsischen Parameter. Letztere geben Auskunft über die geometrische Beziehung zwischen verschiedenen Kameras, die den gleichen Ausschnitt der Realität abbilden, und spielen für die Kalibrierung als Ergebnis keine Rolle. Die grundsätzlichen Algorithmen, auf denen diese Berechnungen basieren, finden bei dem hier entwickelten Verfahren der Berechnung der Position und Blickrichtung einer Person bei der relativen Orientierung Verwendung. Auf sie wird in Abschnitt 6.5.2 ausführlich eingegangen. Zur Kalibrierung von Kameras können die folgenden Gruppen an Kalibrierverfahren unterschieden werden:

1. **Photogrammetrisch-passive Verfahren:** Die zu dieser Gruppe gehörenden Verfahren benötigen einen Kalibrierkörper. Dieser wird von der Kamera aufgenommen. Die Merkmale werden mit Hilfe von Bildverarbeitungs-algorithmen extrahiert. Die so gewonnenen Bildkoordinaten werden dann den entsprechenden Objektkoordinaten zugeordnet. Diese Verfahren können so weit geführt werden, dass auch die radiale Linsenverzerrung und die Berechnung der Verzerrung des Bildverarbeitungssystems berücksichtigt werden. Dazu zählen der Einfluss technischer Eigenschaften des CCD-Chips und die Übertragung des Bildes von der Kamera in den Bildspeicher des verwendeten Rechners. Das Verfahren läuft in zwei Schritten ab. Zunächst werden die als bekannt vorausgesetzten dreidimensionalen Objektkoordinaten in dreidimensionale Kamerakoordinaten umgewandelt. Anschließend werden diese in zweidimensionale Bildkoordinaten transformiert. Dabei werden die intrinsischen Kameraparameter bestimmt.

Dies geschieht durch die Aufnahme eines einzigen Bildes. Es ist damit ein passives Verfahren, das die Aufnahme des Kalibrierkörpers voraussetzt. Jede neue Aufnahme kann zu neuen Kalibrierwerten führen. Diese können über einen Ausgleichungsansatz zusammengefasst werden.

2. **Aktive halbautomatische Verfahren:** Auch bei dieser Gruppe von Verfahren wird ein Kalibrierkörper oder ein Kalibriermuster verwendet. Allerdings sind hier verschiedene Aufnahmen dieses Kalibrierkörpers notwendig. Dazu ändert man entweder die Brennweite der Kamera, dreht die Kamera um einen bekannten Winkel oder verschiebt den Kalibrierkörper um einen bekannten Vektor. Die Basis dieser Verfahren bildet der Algorithmus nach Zhang [Zhan00], der anstelle von dreidimensionalen Objektkoordinaten, wie sie bei photogrammetrisch passiven Verfahren vorausgesetzt werden, nur zweidimensionale Objektkoordinaten benötigt. Dadurch können Kalibriermuster verwendet werden. Eine Möglichkeit besteht darin, sie in Form von geometrischen Mustern auf einer planaren Fläche aufzubringen. Die hierbei verwendeten Algorithmen führen zu einer Minimierung des Projektionsfehlers zwischen den berechneten und den tatsächlich aufgenommenen Bildpunkten.
3. **Verfahren zur Selbstkalibrierung:** Im Gegensatz zu den zuvor beschriebenen Verfahren wird hierbei kein Kalibriermuster benötigt. Eine Möglichkeit besteht darin, Bildaufnahmen einer Szene aus verschiedenen Blickwinkeln aufzunehmen und mit Hilfe dieses Datenmaterials die intrinsischen Kameraparameter zu bestimmen. Ein alternativer Lösungsweg besteht darin, die Kamera um bekannte Winkel zu drehen. Die Algorithmen basieren auf der Bestimmung der so genannten Kruppa-Matrix. Die Schwierigkeit dieser Verfahren besteht in der relativen Ungenauigkeit der Ergebnisse. Darüber hinaus ist die Bestimmung der Parameter sehr stark von der aufgenommenen Szene abhängig. So bereiten sehr ähnliche oder gleiche Bildausschnitte große Probleme. Im Moment gibt es kein zuverlässiges Verfahren zur Selbstkalibrierung, das völlig automatisch funktioniert.

Im Folgenden wird davon ausgegangen, dass durch eine Kalibrierung die Bestimmung der Punkte in den Bildebenen verbessert wird. Dies bedeutet, dass die intrinsischen Parameter berücksichtigt werden müssen, da sie die Eigenheiten der eingesetzten Kamera beinhalten. Dabei handelt es sich im vorliegenden Fall um die folgenden Parameter:

- Skalierungsfaktoren s_u und s_v
- Koordinaten des Hauptpunktes u_0 und v_0 .

Diese gehen in die folgenden Formeln zur Berechnung der tatsächlichen Koordinaten der korrespondierenden Punkte ein:

$$P' = \begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{u'_{unkalibriert} - u_0}{f \cdot s_u} \\ \frac{v'_{unkalibriert} - v_0}{f \cdot s_v} \\ 1 \end{pmatrix}$$

$$P'' = \begin{pmatrix} u'' \\ v'' \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{u''_{\text{unkalibriert}} - u_0}{f \cdot s_u} \\ \frac{v''_{\text{unkalibriert}} - v_0}{f \cdot s_v} \\ 1 \end{pmatrix}$$

Damit können die tatsächlichen Koordinaten der korrespondierenden Punkte berechnet werden, die in die im folgenden Abschnitt erläuterte relative Orientierung eingesetzt werden. Hier wird die geometrische Beziehung zwischen Kameras und Bildebenen, die gleiche Teile der Realität abbilden, über die Zuordnung zwischen korrespondierenden Punkten hergestellt.

Für den Einsatz einer Kamera im Rahmen eines videobasierten Trackingverfahrens eignet sich ein aktives halbautomatisches Kalibrierungsverfahren am besten. Die zu dieser Gruppe zählenden Verfahren benötigen keinen dreidimensionalen Kalibrierkörper. Ein zweidimensionales Muster ist ausreichend. Im Gegensatz zu den Verfahren zur Selbstkalibrierung sind diese Verfahren zuverlässig genug, um hochgenaue Ergebnisse zu erzielen. Vorteil der Verfahren zur Selbstkalibrierung ist es, dass sie während des Einsatzes des Videotrackings ein problemloses Nachkalibrieren ermöglichen, sogar ohne, dass der Anwender dies mitbekommt. Für die hier vorgestellte Anwendung erscheint es allerdings wichtiger zu sein, dass das Verfahren zuverlässige Werte liefert. Aus diesem Grund wird empfohlen, vor jedem neuen Einsatz eine Kontrollkalibrierung durchzuführen.

6.5.2 Bestimmung der relativen Orientierung

Nach der Bestimmung der intrinsischen Kameraparameter, die über einen gewissen Zeitraum bei unveränderten Einstellungen gleich bleiben, muss in einem nächsten Schritt die Beziehung der Kameras zu den zugeordneten Bildebenen berechnet werden. Die hierbei zu bestimmenden extrinsischen Kameraparameter werden nur für eine Konstellation aus Kameraposition, -ausrichtung und Lage der Bildebene ermittelt. Die so genannte relative Orientierung entspricht der Translation und der Rotation zwischen den beiden Kameras. Diese werden über korrespondierende Punkte in beiden Bildern bestimmt. Es sei darauf hingewiesen, dass nicht alle korrespondierenden Punkte, die über das in Abschnitt 6.4.3 beschriebene Verfahren ermittelt wurden, korrekt sind. Noch vorhandene Ausreißer müssen erkannt und ungenaue Zuordnungen ausgeglichen werden.

Bildlich ausgedrückt ist in beiden Ansichten derselbe reale Punkt in Form der korrespondierenden Punkte in den Bildern sichtbar. Das bedeutet, dass es zwei Linien gibt, die jeweils von den Projektionszentren der beiden Kameras ausgehen, durch die entsprechenden Punkte, der Projektion des realen Punktes in den Ansichten, laufen und sich in dem realen Punkt schneiden. Ein Punkt in einem der beiden Bilder wird als Gerade in das andere Bild projiziert. Diese Gerade wird als Epipolarlinie bezeichnet. Wenn der reale Punkt bekannt ist, wird diese Gerade zu einem Punkt. In der Ebene, die durch diese beiden Linien aufgebaut wird, liegt ebenfalls die Verbindungslinie zwischen den beiden Kameras, die als Basis bezeichnet wird. Für jeden Punkt, der in beide Kamerabilder projiziert wird, gibt es eine so genannte Epipolarebene. Diese enthält die Projektion der Punkte, den realen Punkt und die Epipolarlinien. Auf der gemeinsamen Schnittgeraden aller dieser Ebenen liegen die beiden Projektionszentren der Kameras und damit die Basis. Den

Zusammenhang zwischen einem projizierten Punkt P_i' und seinem Raumpunkt P_i stellen die Kollinearitätsgleichungen dar:

$$x_i' = -f \frac{r_{11}(X_i - X_0) + r_{12}(Y_i - Y_0) + r_{13}(Z_i - Z_0)}{r_{31}(X_i - X_0) + r_{32}(Y_i - Y_0) + r_{33}(Z_i - Z_0)}$$

$$y_i' = -f \frac{r_{21}(X_i - X_0) + r_{22}(Y_i - Y_0) + r_{23}(Z_i - Z_0)}{r_{31}(X_i - X_0) + r_{32}(Y_i - Y_0) + r_{33}(Z_i - Z_0)}$$

Hierbei ist f die Brennweite. Die Koeffizienten r sind Funktionen der drei Rotationswinkel. Der Ursprung des lokalen Koordinatensystems ist das Projektionszentrum O mit den Koordinaten X_0 , Y_0 und Z_0 . Der Raumpunkt P_i besteht aus den Koordinaten X_i , Y_i und Z_i , der projizierte Punkt p_i' aus den Koordinaten x_i' und y_i' .

Die Beziehung zwischen zwei Bildpunkten und dem entsprechenden Raumpunkt kann demnach durch ein Gleichungssystem, bestehend aus vier Gleichungen, hergestellt werden.

Wenn die Koordinaten der beiden Bildpunkte mit Hilfe der Kalibrierergebnisse der Kameras ausgedrückt werden, kann ein Punkt des zweiten Bildes in den korrespondierenden Punkt des ersten Bildes mit Hilfe folgender Formel transformiert werden:

$$P' = RP' + t.$$

Hierbei ist R die Rotationsmatrix zwischen den beiden Kameras und t der entsprechende Translationsvektor.

Nach Anwendung des Vektorproduktes ergibt sich daraus:

$$t \times P' = t \times RP' + t \times t.$$

Nach Erweiterung mit P' wird dies zu

$$P'(t \times P') = P't \times RP'.$$

Die drei Vektoren P' , P'' und t liegen alle in der Epipolarebene und sind somit koplanar. Damit ist ihr Spatprodukt 0, und die vorhergehende Gleichung lässt sich umformen zu

$$P'(t \times R)P'' = 0.$$

Durch Einführung der essentiellen Matrix E werden die Rotation R und die Translation t zusammengefasst. Die Gleichung für kalibrierte Kameras zur Beschreibung der Epipolarometrie ergibt sich damit zu

$$P'^T E P'' = 0$$

Die essentielle Matrix kapselt damit die Transformation zwischen zwei Kamerakoordinatensystemen und stellt die Beziehung zwischen korrespondierenden Punkten in zwei Bildern dar. Sie entspricht dem Produkt TR , wobei T eine Matrix ist, die sich aus den Komponenten des Translationsvektors zusammensetzt:

$$T = \begin{pmatrix} 0 & -t_w & t_v \\ t_w & 0 & -t_u \\ -t_v & t_u & 0 \end{pmatrix}$$

Die essentielle Matrix E ist eine 3x3-Matrix. Sie hängt aber nur von fünf Parametern ab: Drei für die Rotation und zwei für die Translation. Der Translationsvektor zwischen den Kameras ist nicht eindeutig bestimmt. Über die korrespondierenden Punkte kann nur die Richtung, nicht aber die Länge ermittelt werden. Eine Skalierung des Translationsvektors führt auch zu einer Skalierung der essentiellen Matrix.

Anschaulich bedeutet dies, dass die Koordinaten der Bildkoordinatensysteme gleich bleiben, wenn alle Seitenlängen des Dreiecks, dessen Eckpunkte durch die Projektionszentren der Kameras und einen Objektpunkt aufgespannt werden, mit dem gleichen Faktor multipliziert werden. Eine Seite verbindet die beiden Projektionszentren, die anderen beiden Seiten werden jeweils durch ein Projektionszentrum, die Abbildung des Objektpunktes in die entsprechende Bildebene und den Objektpunkt definiert.

Um die Fehler bei der Bestimmung der korrespondierenden Punkte und Ungenauigkeiten durch Bildrauschen berücksichtigen zu können, wird die Matrix E mit mehr als der notwendigen Anzahl von Punktkorrespondenzen ermittelt. Damit ergibt sich ein überbestimmtes Gleichungssystem, das mit Hilfe von Ausgleichsmethoden über die beste Schätzung des Ergebnisses gelöst werden kann. Beim Vorliegen von Ausreißern können diese mit Hilfe eines robusten Verfahrens während der Berechnung der relativen Orientierung entfernt werden. Ausreißer können in diesem Fall die nicht korrekten korrespondierenden Punkte sein, die bei der Durchführung der Korrespondenzfindung nicht als fehlerhaft erkannt wurden. Es gibt verschiedene robuste Verfahren, die für die Berechnung der essentiellen Matrix anwendbar sind:

- **Case Deletion Diagnostics (CDD)** sind Methoden, die aus den ermittelten Punktkorrespondenzen schrittweise die potentiellen Ausreißer entfernen. Dafür wird der Einfluss jeder Punktkorrespondenz auf den Fehler bestimmt, indem sukzessiv jedes Punktpaar aus der Berechnung ausgeschlossen wird. Bei Überschreiten des Einflusses über einen Grenzwert wird dieses Paar als Ausreißer eingestuft.
- Bei Verwendung des **M-Estimators** werden Ausreißer nicht entfernt, sondern schwächer gewichtet, indem die Auswirkung von falschen Korrespondenzen auf die Lösung des Gleichungssystems minimiert wird. Es handelt sich hierbei um ein iteratives Verfahren, da für die Berechnung des Gewichtes die Abschätzung des Fehlers bekannt sein muss, dieser jedoch nur berechnet werden kann, wenn die Lösung bekannt ist. Es können verschiedene Gewichtungsfunktionen angewendet werden, allerdings ist nach [Torr96] das Gewicht von Huber für die Berechnung der essentiellen Matrix gut geeignet.
- Das Least Median Squares (LMS) Verfahren und das Random Sample Consensus (RANSAC) Verfahren gehören zu den **Random Sampling** Verfahren. Der Median entspricht dem Wert, bei dem mindestens 50 % der Werte kleiner oder gleich groß sind. Die Verfahren verwenden im Gegensatz zu den zuvor genannten Verfahren keine möglichst große Menge von korrespondierenden Punkten, sondern die Mindestanzahl, um das Gleichungssystem lösen zu können. LMS und RANSAC unterscheiden sich dadurch, dass beim LMS-Verfahren der Median der Fehler, bei RANSAC die

Anzahl der Punktpaare, für die der Fehler unter einem bestimmten Grenzwert liegt, ermittelt wird. Das Abbruchkriterium für die Iteration ist bei LMS das Erreichen einer bestimmten Anzahl von Iterationsschritten, wobei mindestens eine Untermenge ohne Ausreißer entnommen werden muss. Bei RANSAC ist es das Erreichen einer festgelegten Mindestanzahl an Punktpaaren, für die der Fehler über dem Grenzwert liegt. Bei LMS wird demnach die essentielle Matrix über den kleinsten Fehlermedian bestimmt. Bei RANSAC ist es diejenige, bei der das Verfahren abgebrochen wurde.

Diese Verfahren haben alle bestimmte Vor- und Nachteile: Das CDD-Verfahren ist bezüglich der Effizienz das ideale Verfahren, wenn sich unter den korrespondierenden Punkten nur wenige Ausreißer befinden. Das LMS-Verfahren liefert die genauesten Ergebnisse bei der Berechnung der Fundamentalmatrix, die auf der essentiellen Matrix beruht und für unkalibrierte Koordinaten definiert ist [Torr96]. Der M-Estimator ist zwar robust gegenüber Ungenauigkeiten bezüglich der Position der Punktkorrespondenzen, allerdings können falsche Korrespondenzen nicht unterdrückt werden. LMS kann die falschen Korrespondenzen sowohl erfolgreich ermitteln als auch entfernen [Zhan94]. Ein Vergleich zwischen LMS und RANSAC hat gezeigt, dass LMS bei einem Anteil von bis zu 40 % Ausreißern ein besseres Ergebnis liefert als RANSAC. Allerdings scheitert LMS vollständig, wenn der Anteil der Ausreißer über 50 % liegt [Torr93].

Es wird im vorliegenden Fall davon ausgegangen, dass der Anteil der Ausreißer relativ gering ist, dass es sich allerdings nicht nur um einige wenige Ausreißer handelt. Das LMS-Verfahren liefert hierbei die genauesten Ergebnisse und ist sogar in der Lage falsche Korrespondenzen zu entfernen.

Bei der Umsetzung dieses Verfahrens nach [Zhan94] werden zufällig gewählte Untermengen mit jeweils acht Punktkorrespondenzen analysiert. Dazu wird zunächst die Anzahl der acht-elementigen Untermenge ermittelt. Eigentlich müssten alle

$\frac{n!}{((n-8)! \cdot 8!)}$ Untermengen der n korrespondierenden Punktpaare berücksichtigt

werden. Wenn so vorgegangen wird, ist das Verfahren von der Ordnung $O(n! \cdot n \log n)$ und damit für eine Implementierung nicht geeignet. Aus diesem Grund wird nur eine bestimmte Zahl von Untermengen betrachtet. Dabei ist darauf zu achten, dass mindestens eine davon keine Ausreißer enthält. Die dadurch bestimmte Untermenge liefert den kleinsten Median und die beste essentielle Matrix. Um die Anzahl m der Untermengen zu bestimmen, wird die folgende Formel aus der Wahrscheinlichkeitsrechnung herangezogen:

$$m \geq \frac{1-P}{\log(1-(1-\varepsilon)^n)}$$

Dabei steht P für die Wahrscheinlichkeit und wird bei [Zhan94] mit 0.99 gleichgesetzt. Der Parameter ε bezeichnet die Wahrscheinlichkeit, dass ein korrespondierendes Punktpaar ein Ausreißer ist und kann über die folgende Formel berechnet werden:

$$\varepsilon = \frac{\text{Anzahl der Ausreißer}}{\text{Anzahl der korrespondierenden Punktpaare}}$$

Da im vorliegenden Fall die Anzahl der Ausreißer unbekannt ist, muss ϵ abgeschätzt werden.

Gemäß [Zhan94] werden die acht korrespondierenden Punktepaare einer Untermenge aus verschiedenen Bildregionen gewählt, nachdem die Bildebene in horizontale und vertikale, rechteckförmige Regionen unterteilt wurde. Jede dieser Regionen beinhaltet eine bestimmte Anzahl von korrespondierenden Bildpunkten. Aus diesen wird je ein korrespondierendes Punktepaar herangezogen. Für jede Untermenge werden die folgenden Berechnungen durchgeführt:

- Mit Hilfe der Methode der kleinsten Fehlerquadrate wird die essentielle Matrix E für die korrespondierenden Punktepaare berechnet.
- Der Fehler eines korrespondierenden Punktepaares kann über den Abstand zwischen den mit Hilfe der inneren Parameter verbesserten Bildkoordinaten und den dazugehörigen Epipolarlinien der korrespondierenden Punkte ausgedrückt werden. Zu jedem Punktepaar können zwei solcher Abstände berechnet werden, da sie nicht symmetrisch sind. Für die Berechnung des Abstandes kann die essentielle Matrix eingesetzt werden. Der Fehler wird für jedes korrespondierende Punktepaar berechnet, der in einen Fehlervektor eingeht.
- Für jeden Fehlervektor wird der Median berechnet. Dafür müssen die Fehler der Größe nach aufsteigend sortiert werden.

Nach Ermittlung aller Mediane bestimmt man die dem kleinsten Median zugeordnete essentielle Matrix. Ausgehend von ihr wird danach für jedes korrespondierende Punktepaar der Fehler über den Abstand zur Epipolarlinie berechnet. Wenn diese Differenz einen bestimmten Wert übersteigt, handelt es sich um einen Ausreißer. Dieser wird aus der Menge der korrespondierenden Punkte eliminiert. Für die übrigen Paare wird die essentielle Matrix nach der Methode der kleinsten Quadrate ermittelt.

Nach Berechnung der essentiellen Matrix können die Rotationsmatrix R und der Translationsvektor t berechnet werden. Wegen der erwähnten Skalierungsinvarianz kann die Länge des Translationsvektors zu diesem Zeitpunkt noch nicht eindeutig bestimmt werden. Aus diesem Grund wird die Länge mit eins festgelegt. Es gilt also:

$$|t|^2 = 1.$$

Damit ergibt sich die Norm für die essentielle Matrix mit

$$|E|^2 = 2.$$

Durch Berücksichtigung dieser Norm werden der ermittelte Lösungsvektor bzw. die essentielle Matrix zu Beginn der Berechnung auf $\sqrt{2}$ normiert.

Zur Berechnung des Translationsvektors muss das folgende Minimierungsproblem gelöst werden:

$$\min |E^T t| \text{ mit } |t|^2 = 1.$$

Dazu wird der Eigenvektor zum kleinsten Eigenwert der Matrix $E E^T$ bestimmt und auf eins normiert.

Die Rotationsmatrix wird ebenfalls durch ein Minimierungsproblem bestimmt. In diesem Fall handelt es sich um eine Minimierung der Matrixnorm

$$|E - TR|^2.$$

Über die Berücksichtigung multiplikativ verknüpfbarer Matrizen kann die folgende Gleichung hergeleitet werden:

$$\sum_{i=1}^3 |R^T \cdot t_i - e_i|^2.$$

Dabei sind t_i die drei Zeilenvektoren t_1, t_2 und t_3 der Matrix T und e_1, e_2 und e_3 die Zeilenvektoren der essentiellen Matrix E . Mit Hilfe von Quaternionen $q = (q_1, q_2, q_3, q_4)$ führt dies zu dem folgenden Minimierungsproblem:

$$\min qBq^T, \text{ wobei } |q|=1 \text{ und } B = \sum_{i=1}^3 C_i C_i^T \text{ gilt, sowie}$$

$$C_i = \begin{pmatrix} 0 & t_i^u - e_i^u & t_i^v - e_i^v & t_i^w - e_i^w \\ e_i^u - t_i^u & 0 & e_i^w + t_i^w & -e_i^v - t_i^v \\ e_i^v - t_i^v & -e_i^w - t_i^w & 0 & e_i^u + t_i^u \\ e_i^w - t_i^w & e_i^v - t_i^v & -e_i^u - t_i^u & 0 \end{pmatrix}$$

Für die Lösung muss der Eigenvektor zum betragsmäßig kleinsten Eigenwert von B auf eins normiert werden. Das bestimmte Quaternion beschreibt die Rotationsachse und den Rotationswinkel. Daraus kann dann die Rotationsmatrix berechnet werden:

$$R = \begin{pmatrix} q_1^2 + q_2^2 - q_3^2 - q_4^2 & 2 \cdot (q_2 \cdot q_3 - q_1 \cdot q_4) & 2 \cdot (q_2 \cdot q_4 + q_1 \cdot q_3) \\ 2 \cdot (q_2 \cdot q_3 + q_1 \cdot q_4) & q_1^2 - q_2^2 + q_3^2 - q_4^2 & 2 \cdot (q_3 \cdot q_4 - q_1 \cdot q_2) \\ 2 \cdot (q_2 \cdot q_4 - q_1 \cdot q_3) & 2 \cdot (q_3 \cdot q_4 + q_1 \cdot q_2) & q_1^2 - q_2^2 - q_3^2 + q_4^2 \end{pmatrix}$$

Die Länge des Translationsvektors wurde auf eins festgelegt. Nach der Bestimmung der Rotationsmatrix muss die Richtung des Translationsvektors festgesetzt werden. Dazu werden die Objektpunkte der korrespondierenden normierten Bildpunkte mit Hilfe des Translationsvektors t und der Rotationsmatrix R berechnet. Es wird davon ausgegangen, dass beide Bildebenen zum Objektpunkt hin ausgerichtet sind. Dann müssen die Vorzeichen der Tiefenkoordinaten der Kamerakoordinatensysteme übereinstimmen. Es können die folgenden drei Fälle auftreten:

1. Beide Tiefenkoordinaten sind positiv: R und t stellen das Endergebnis dar
2. Beide Tiefenkoordinaten sind negativ: t muss negiert werden, R stellt das Endergebnis dar

- Die beiden Tiefenkoordinaten weisen unterschiedliche Vorzeichen auf: In diesem Fall muss die essentielle Matrix E mit -1 multipliziert werden, t wird beibehalten und R erneut berechnet.

6.5.3 Absolute Orientierung

Als letzter Schritt des videobasierten Trackings müssen die Position und die Blickrichtung des Anwenders bestimmt werden. Dabei ist einerseits zu beachten, dass die Ergebnisse in dem Koordinatensystem vorliegen, in dem die virtuellen Daten georeferenziert sind, die eingeblendet werden sollen, und zusätzlich muss berücksichtigt werden, dass Position und Ausrichtung des tragbaren Displays das eigentliche Ziel der Berechnungen sind.

In den vorhergehenden Schritten wurden die internen Kameraparameter bestimmt sowie die Position und Ausrichtung der beiden Kameras zueinander, wobei eine Skalierungsinvarianz besteht, die sich insofern bemerkbar macht, als die Länge des Translationsvektors zwischen den beiden Kameras unbekannt ist. Das Ergebnis dieser Schritte ist die so genannte relative Orientierung. Der nachfolgende Schritt, die absolute Orientierung, passt die beiden zueinander orientierten Kameras in ein übergeordnetes Koordinatensystem ein. Wie oben erwähnt, handelt es sich hierbei um das Landeskoordinatensystem. Die Beziehungen zu diesem übergeordneten Koordinatensystem zeigen die folgenden Gleichungen.

Transformation vom übergeordneten Koordinatensystem in das Koordinatensystem der ersten Kamera:

$$P' = R'P + T'$$

Transformation vom übergeordneten Koordinatensystem in das Koordinatensystem der zweiten Kamera:

$$P'' = R''P + T''$$

Diese Formel kann umgewandelt werden zu

$$P = R''^{-1}(P'' - T'') = R''^{-1}P'' - R''^{-1}T''$$

Die Transformation eines Punktes vom zweiten in das erste Kamerakoordinatensystem lässt sich aus den vorausgehenden Gleichungen ableiten:

$$P' = R'R''^{-1}(P'' - T'') + T' = (R'R''^{-1})P'' - (R'R''^{-1})T'' + T'$$

Mit Hilfe der Parameter $R''' = R'R''^{-1}$ und $T''' = T' - R'''T''$ lässt sich die Gleichung umwandeln zu

$$P' = R'''P'' + T'''$$

Das Kamerabild gibt die reale Position und Blickrichtung des Anwenders wieder. Allerdings sind die tatsächlichen Werte unbekannt, so dass aus diesem Grund die Modellansicht mit Hilfe von ungefähren Informationen generiert wurde. Neben den Näherungswerten für Position und Blickrichtung sind die Verschiebung und die

Rotation der Kameras zueinander bis auf den Skalierungsfaktor bekannt. Mit Hilfe von Daten des 3D-Modells können damit Position und Ausrichtung der Anwenderkamera bestimmt werden.

In einem ersten Schritt wird deshalb die Transformation vom übergeordneten Koordinatensystem in das Koordinatensystem der ersten Kamera bestimmt. Diese kann mit Hilfe der Transformation des übergeordneten Koordinatensystems in das zweite Kamerakoordinatensystem, das dem der Modellansicht entspricht, und der Transformation zwischen den beiden Kamerakoordinatensystemen berechnet werden.

Aus den vorangegangenen Gleichungen lässt sich die Rotation vom übergeordneten Koordinatensystem in das erste Kamerakoordinatensystem R' herleiten:

$$R' = R''' R''.$$

Die Translation vom übergeordneten Koordinatensystem in das erste Kamerakoordinatensystem T' wird mit folgender Gleichung ausgedrückt:

$$T' = T''' + R''' T''.$$

Da die Translation T''' zwischen den beiden Kamerakoordinatensystemen nur bis auf einen Skalierungsfaktor bekannt ist, wird sie mit folgender Formel ausgedrückt:

$$T''' = s T_s''''.$$

Damit ergibt sich die oben angegebene Gleichung für die Transformation vom übergeordneten ins erste Kamerakoordinatensystem folgendermaßen:

$$T' = s T_s'''' + R''' T''.$$

Das Resultat ist die Transformation vom übergeordneten Koordinatensystem in das erste Kamerakoordinatensystem:

$$P' = R' P + s T_s'''' + R''' T''.$$

Die beiden unbekannt Parameter hierbei sind die Koordinaten P' des Koordinatensystems der ersten Kamera und die Koordinaten P des Objektpunktes. Diese werden in Koordinaten des übergeordneten Systems ausgedrückt. Aus diesem Grund müssen die Koordinaten des Objektpunktes für ein korrespondierendes Punktepaar bestimmt werden. Bekannt sind die Transformationsparameter der zweiten Kamera, die die Kamera der Modellansicht repräsentiert und über Näherungswerte ermittelt wurde. Die Objektpunkte können mit Hilfe des bekannten 3D-Modells bestimmt werden. Dafür wird der Schnittpunkt eines Strahls vom Zentrum der zweiten Kamera durch den Punkt im zweiten Kamerabild P'' mit dem 3D-Modell berechnet. Auf die Bildebene wird der Punkt abgebildet, der im 3D-Modell der Bildebene am nächsten liegt.

Die Gerade vom Kamerazentrum durch die entsprechende Bildebene wird dann in das übergeordnete Koordinatensystem umgewandelt und mit der im übergeordneten Koordinatensystem vorliegenden Modellfläche verschnitten. Bei mehreren Schnittpunkten mit dem 3D-Modell wird der Punkt mit dem kleinsten Abstand zum Bildpunkt ausgewählt.

Im Anschluss daran muss die Skalierung der Translation zwischen den beiden Kamerazentren berechnet werden. Dafür wird der Objektpunkt zu einem korrespondierenden Bildpunktpaar im Koordinatensystem der ersten Kamera benötigt. Er wird in die Gleichung eingesetzt, die die Beziehung zwischen den beiden Kamerakoordinatensystemen und dem Objektpunkt darstellt. Die einzige unbekannte Größe darin ist der Skalierungsfaktor s . Da dieser für alle Bildpunktpaare identisch ist, kann er durch ein überbestimmtes Gleichungssystem ermittelt werden. Es wird mit Hilfe der Methode der kleinsten Fehlerquadrate gelöst und berücksichtigt damit, dass die korrespondierenden Bildpunktpaare nicht vollkommen exakt ermittelt worden sind.

Damit wurden alle Transformationen bestimmt und die Position und Ausrichtung der Anwenderkamera kann über diese Transformationen berechnet werden.

6.6 Erweiterung des Verfahrens

6.6.1 Generierung mehrerer Ansichten

Alternativ zum in Abschnitt 6.4.3 beschriebenen Verfahren nach Gros et al. besteht die Möglichkeit über die Generierung mehrerer Ansichten die dem Kamerabild ähnlichste über einen Ähnlichkeitswert festzulegen und die entsprechende Position und Blickrichtung als genähertes Ergebnis weiter zu verwenden. Damit wird die sehr aufwendige Berechnung der Orientierung der Kameras, wie sie in Abschnitt 6.5 ausführlich erläutert ist, durch die Berechnung eines Ähnlichkeitsmaßes ersetzt.

Die Berechnung des Ähnlichkeitsmaßes schließt sich an die Bestimmung der korrespondierenden Bildmerkmale an, wie sie in Abschnitt 6.4.8 erläutert wurde. Ein Kamerabild und mehrere Datensätze stehen aus dem 3D-Modell zur Verfügung. Es wird demnach nicht ein Paar aus Kamerabild und 3D-Modellansicht betrachtet, sondern die Anzahl der Paare richtet sich nach der Anzahl der generierten 3D-Ansichten. Für alle diese Paare wurden Merkmale extrahiert und die korrespondierenden Segment-Konfigurationen ermittelt.

Der folgende Ansatz geht davon aus, dass das Paar den höchsten Ähnlichkeitswert besitzt, das in Position und Blickrichtung des Betrachters am meisten übereinstimmt. Es zeigt damit einen sehr ähnlichen Bildausschnitt.

Für die Bestimmung des Paares mit dem ähnlichsten Blickwinkel werden die Anzahl der Kanten und Endpunkte der detektierten Segmentkonfigurationen für jedes Paar und die korrespondierenden Kanten und Endpunkte näher betrachtet. Damit lässt sich auf die inhaltliche Übereinstimmung zwischen beiden Bildern eines Paares schließen. Die Ähnlichkeit eines Paares, das aus den zwei Bildern I' und I'' besteht, wird mit folgender Formel berechnet [Gros95]:

$$sim(I', I'') = \frac{2 \cdot AnzahlEndpunkte(I') \cdot AnzahlEndpunkte(I'')}{3 \cdot Endpunktmatches(I', I'')^2} + \frac{1 \cdot AnzahlKanten(I') \cdot AnzahlKanten(I'')}{3 \cdot Kantenmatches(I', I'')^2}$$

Der Parameter $\text{sim}(I, I')$ nimmt den Wert eins an, wenn zwei Bilder völlig übereinstimmen. Wenn gar keine Übereinstimmung existiert konvergiert er gegen unendlich. Dies bedeutet, dass der Wert umso kleiner wird, je mehr die Bilder übereinstimmen.

6.6.2 Reine Bild-zu-Bild-Zuordnung

Das hier beschriebene Verfahren ist durch die Extraktion und Zuordnung von Merkmalen aus zwei grundsätzlich verschiedenen Datenquellen sehr rechen- und damit zeitaufwendig. Diese absolute Bestimmung der Position und Blickrichtung einer Person in einer Stadt liefert allerdings zuverlässige Ergebnisse. Um das Tracking zu beschleunigen, sollte ein ähnliches Verfahren integriert werden. Die Integration besteht darin, dass Ausgangspunkt des alternativen Verfahrens die Ergebnisse des zuvor angewendeten Verfahrens darstellen, die über das Ergebnis der Zuordnung zwischen 3D-Modell und Kamerabild entstanden sind. Für das alternative Verfahren wird dann die relative Orientierung nicht zwischen einer virtuellen und einer realen Kamera berechnet, sondern es werden die Unterschiede zwischen zwei Bildern des Videostroms herangezogen, die zu einer relativen Bestimmung von Position und Blickrichtung führen. Man kann davon ausgehen, dass die rein relative Berechnung nach einer bestimmten Anzahl an Bildern divergiert. Aus diesem Grund sollte dann wieder das Verfahren angewendet werden, das auf der Korrespondenz eines Kamerabildes mit Daten eines 3D-Stadtmodells basiert. Abbildung 40 zeigt den Ablauf des Verfahrens mit Hilfe von zwei Kamerabildern, wobei die Merkmale des ersten Bildes bereits aus der Zuordnung zu den Merkmalen des 3D-Modells bekannt sind.

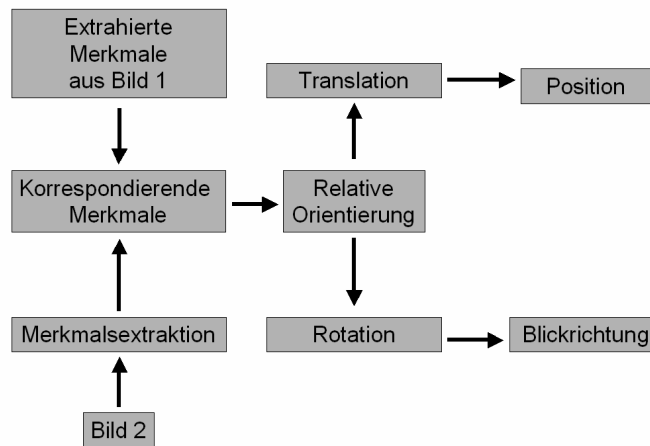


Abbildung 40: Überblick über die Positions- und Blickrichtungsbestimmung mit Hilfe von zwei Kamerabildern

Da die Lage und Ausrichtung dieser Merkmale sich im zweiten Bild nicht sehr stark verändern, ist eine einfachere Berechnung der korrespondierenden Merkmale möglich. Die Ermittlung der relativen Orientierung zwischen den Kameras der beiden Bilder entspricht dem zuvor beschriebenen Ablauf.

6.7 Zusammenfassung

In diesem Kapitel wurde ein Verfahren mit Lösungsansätzen aufgezeigt, das es erlaubt auf Basis eines 3D-Modells im urbanen Umfeld die Position und Blickrichtung einer Person für AR-Anwendungen zu bestimmen. Dazu repräsentiert eine Kamera die tatsächliche Position und Blickrichtung des Anwenders. Die Daten eines 3D-Modells werden so aufbereitet, dass sie ebenfalls eine zweidimensionale Ansicht repräsentieren. Für die Generierung der Daten werden Näherungswerte für Position und Blickrichtung herangezogen. Durch den Vergleich des Kamerabildes mit den Daten des 3D-Modells werden die Korrekturparameter bestimmt, die die Näherung in die tatsächliche Lösung überführen.

Das beschriebene Verfahren extrahiert in einem ersten Schritt Merkmale in beiden Datensätzen. Dies führt zur Lokalisierung von Kanten in einer zweidimensionalen Umgebung, die als Vektoren identifiziert werden. Die Merkmale des Kamerabildes müssen im nachfolgenden Schritt den entsprechenden Daten des 3D-Modells zugeordnet werden. Die dafür entwickelte Lösung basiert auf dem Ansatz von Gros et al., womit die Transformation des einen Datensatzes, der des Kamerabildes, in den anderen, die Ansicht des 3D-Stadtmodells, durch eine Ähnlichkeitstransformation angenähert wird. Dieses Verfahren weist die folgenden Vorteile auf: Es ist robust gegenüber Bildrauschen, und es ermittelt nur wenige falsche Korrespondenzen, wenn sich die beiden Ansichten nicht zu sehr unterscheiden. Dies kann im vorliegenden Fall vorausgesetzt werden. Außerdem können die erkannten Merkmale im Bild verschoben sein, da das Verfahren keine minimalen Abstände zwischen den korrespondierenden Merkmalen benötigt. Im Gegensatz zu einem Verfahren mit normaler Hough-Transformation oder relationalen Zuordnungsverfahren benötigt es weniger Rechenaufwand. Ergebnis dieses Schrittes ist die Zuordnung von Punkten des Kamerabildes zu Punkten der Ansicht des 3D-Stadtmodells.

Diese zugeordneten Punktpaare ermöglichen die Herstellung der geometrischen Beziehungen zwischen den Zentren der Anwenderkamera und der Kamera, die Position und Ausrichtung der Näherungswerte repräsentiert, und deren Werte für die Generierung der Ansicht genutzt wurden. Darüber werden die beiden Bildebenen und, über einen weiteren Schritt, die Objektpunkte des 3D-Modells miteinander in Beziehung gesetzt, wodurch die Position und Blickrichtung der Anwenderkamera bekannt sind. Damit stehen die Daten für eine Überlagerung der Realität mit virtuellen Informationen im Blickfeld einer Person fest. Der Weg dorthin führt über die Berücksichtigung der inneren Parameter der Anwenderkamera, die über eine Kalibrierung bestimmt werden können. Anschließend wird zunächst die relative Ausrichtung der beiden Kameras bestimmt, die über die korrespondierenden Bildpunktpaare berechnet wird. Dies führt zur Rotation und Translation zwischen beiden Kameras. Die Translation kann allerdings nur bis auf einen Skalierungsfaktor berechnet werden. In einem letzten Schritt werden die Positionen und Ausrichtungen der beiden Kameras in Beziehung zum 3D-Modell gesetzt, das eine Referenz der Realität darstellt. Über die Berücksichtigung des Objektpunktes korrespondierender Bildpunktpaare wird der Skalierungsfaktor des Translationsvektors zwischen den beiden Kamerazentren ermittelt.

7 Praktische Umsetzung

Dieses Kapitel der vorliegenden Arbeit stellt die Umsetzung von Komponenten vor, die auf den in dieser Arbeit entwickelten Verfahren basieren. Die Implementierung erfolgte im Rahmen des Forschungsprojektes GEIST, in dem ein historisches Lernspiel für Schüler prototypisch entwickelt wurde, das AR mit Digital Storytelling verbindet [Kret02b]. Die Umsetzung der Komponenten erfolgte in der Programmiersprache Java.

Nach einem Überblick über das Projekt in Abschnitt 7.1 zeigt Abschnitt 7.2 beispielhaft an möglichen Anwendungen für positionsabhängige Dienste auf, wie die anforderungsadaptive Zuordnung von Trackingtechniken umgesetzt wird. Anschließend werden in Abschnitt 7.3 die Realisierungen des Trackingverfahrens auf Basis eines 3D-Modells aufgezeigt. Den Abschluss des Kapitels bildet die Zusammenfassung.

7.1 Übersicht

Im Projekt GEIST werden Schüler ausgestattet mit einem Laptop, einem tragbaren Display, das sie sich vor die Augen halten können um blickrichtungsabhängige Informationen vermittelt zu bekommen, einem GPS-Gerät, einem Orientierungstracker, einer Kamera und einem PDA. Die schweren Geräte sowie die Batterien werden in einem Rucksack verstaut. Abbildung 41 zeigt den Einsatz der entsprechenden Geräte.



Abbildung 41: Geräte für die Trackingkomponente

Die GPS-Antenne ist am Rucksack befestigt. Das Display kann vor die Augen gehalten werden, wenn der Anwender im Blickfeld eingeblendete Informationen sehen möchte. Der Orientierungstracker ist auf dem Display befestigt und ermöglicht damit eine genaue Bestimmung der Blickrichtung. Kopfhörer ermöglichen Audioausgaben.

Die Schüler bewegen sich selbständig in der Altstadt Heidelbergs. An verschiedenen Stellen erscheinen ihnen Geister vor historischer Kulisse [Kret01]. Diese Rekonstruktionen sind im Display erkennbar und überdecken positionsgenau die gleichzeitig sichtbare Realität. Die Geister erzählen ihnen Teile einer Geschichte, deren Inhalt davon abhängig ist, in welcher Reihenfolge die Schüler die verschiedenen Stellen besuchen, an denen Geister sichtbar werden. Für jeden Schüler ergibt sich auf diesem Wege eine individuell zusammengestellte Geschichte, die immer spannend ist und mit einem guten Ende abschließt. Die Geister bitten um Mithilfe, die in Form von Interaktion durch die Schüler ausgeführt werden kann. Ziel des Spieles ist es, Motivation zu wecken, sich mit Regionalgeschichte zu beschäftigen. Aus diesem Grund können mit Hilfe des PDAs zusätzliche Informationen abgerufen werden.

Die Architektur der prototypischen Implementierung ist in Abbildung 42 dargestellt.

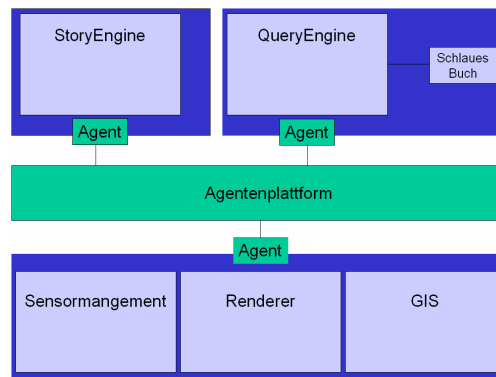


Abbildung 42: Architektur der Implementierung

Über eine Agentenplattform kommunizieren die Komponenten miteinander. Die Datenbasis wird über die QueryEngine zur Verfügung gestellt. Auf diese kann direkt mit dem Schlaues Buch zugegriffen werden. Dieses wird über einen PDA realisiert. Die StoryEngine administriert die Teile der Geschichte und stellt deren Inhalt zusammen. Dieser wird im Renderer präsentiert. Das Sensormangement verwaltet die verschiedenen Sensoren, die Informationen zum Status des Anwenders und zur Umgebung erfassen, und stellt sie den anderen Komponenten zur Verfügung. Das GIS wird sowohl für die Navigation eingesetzt als auch für die Entscheidung, in welchem Spielbereich sich der Anwender befindet.

7.2 Verbindung verschiedener Trackingtechniken

Die automatische Auswahl der idealen Trackingtechnik ist Inhalt des folgenden Abschnittes. Hier wird beispielhaft aufgezeigt, wie der Zugriff auf verschiedene Trackingtechniken über die Implementierung eines Verfahrens so ermöglicht wird, dass für bestimmte Dienste anforderungsadaptiv eine Trackingtechnik zugeordnet wird.

Im Projekt GEIST ist in einer Datenbank das Gebiet definiert, in dem die Anwendung eingesetzt werden soll. Wenn dieser Bereich verlassen wird, wird der Anwender darauf aufmerksam gemacht, und es wird eine einfache Navigation über eine Karte angeboten. Um die gesamte Zeit über, wenn möglich, Positionsinformationen zu erhalten, ist das GPS-Gerät grundsätzlich in Betrieb. Wenn keine Ergebnisse möglich sind, da beispielsweise die Straßen zu eng sind und keine ausreichende Anzahl von Satelliten zur Verfügung steht, wird das zuletzt empfangene Ergebnis solange weiterverwendet, bis neue Ergebnisse vorliegen.

Auf bestimmten Plätzen ist der Einsatz des Trackingverfahrens mit der Hinzunahme von Orientierungsergebnissen möglich. Dieses erlaubt beispielsweise das positionsgenaue Einblenden von Rekonstruktionen in das Blickfeld des Anwenders. Ebenso ist die Einblendung von Informationen zu Gebäuden oder Teilen von Gebäuden möglich. Diese Plätze sind als geographische Informationen ebenfalls in der Datenbank gespeichert. Voreinstellungen für das System werden in einer XML-Datei gehalten. Dazu gehört eine Beschreibung der möglichen Dienste und der angeschlossenen Trackingtechniken.

7.2.1 Mobile Dienste im Projekt GEIST

Die in Kapitel 5 aufgezählten Dienste erfordern teilweise, dass der Benutzer mit Hilfe eines Menüs eine Auswahl trifft oder dass er einen Dienst zu einem bestimmten Zeitpunkt auswählt. Allerdings lässt die manuelle Bestimmung des Dienstes durch den Anwender außer Acht, dass er möglichst ohne klassische Interaktionsmethoden durch die Stadt geführt werden möchte, wie sie die Auswahl aus einem Menü beispielsweise darstellen. Aus diesem Grund wurde das Lernspiel so konzipiert, dass der Benutzer 3D-Rekonstruktionen und damit einen Teil der Geschichte immer dann automatisch angeboten bekommt, wenn dies möglich ist. Zusätzlich kann er sich in einer Karte anzeigen lassen, wo er sich im Moment befindet. Unabhängig von diesen beiden Informationen kann er in einem GIS auf einem PDA Informationen zur Umgebung abrufen.

Sobald der Anwender einen der zuvor definierten Bereiche betritt, werden positionsgenaue 3D-Modelle eingeblendet, vor denen Avatare erscheinen, die ihm Informationen zur Umgebung vermitteln und ihn bitten, Aufgaben zu erfüllen. Der Anwender kann diese Form der Information nur abbrechen, wenn er den entsprechenden Bereich verlässt.

Wenn er seine derzeitige Position wissen möchte, kann er die Anzeige des derzeitigen Standpunktes in einer Karte anfordern, die unabhängig von seiner Blickrichtung ist. Außerdem wird ihm dieser Dienst automatisch angeboten, wenn er den Bereich, in dem das Lernspiel stattfindet und außerhalb dessen es keine Informationen zu einzelnen Objekten gibt, verlässt. Um eine möglichst intuitive Navigation zu ermöglichen, hat der Anwender zusätzlich eine Karte aus Karton in einer mit dem Rechner verbundenen Tasche zur Verfügung. Diese kann er herausnehmen, um sich seinen Standpunkt anzeigen zu lassen. Abbildung 43 zeigt diese Karte zusammen mit der Tasche, die über ein USB-Kabel mit dem Rechner verbunden ist.



Abbildung 43: Interaktive Karte zur Anzeige der derzeitigen Position

Das Herausnehmen der Karte aus der Tasche wird vom System registriert und die derzeitige Position wird im tragbaren Display in Form eines Punktes eingeblendet. Dadurch, dass der

Anwender die Karte aus Karton genau vor seine Augen hält und so mit den Umrissen der Karte im Display in Übereinstimmung bringt, kann er seine derzeitige Position durch die Überlagerung von virtueller Information über eine reale Karte angezeigt bekommen. Alternativ ist die Anzeige der Position im tragbaren Display direkt auf einem digitalen Kartenausschnitt möglich. Damit entfällt die Aufgabe des Anpassens der Karte aus Karton. Die Navigationsunterstützung wird hierbei ebenfalls durch ein kurzes Ziehen der Karte angefordert, allerdings kann die Karte sofort wieder zurückgesteckt werden.

Wenn weitere Informationen auch über die weitere Umgebung notwendig sind, kann der Anwender mit Hilfe des PDAs Informationen in einem GIS abfragen. Die Anforderung der Information wird manuell durchgeführt und ist nicht an die Trackingkomponente angeschlossen. Diese Form der Informationsdarstellung erlaubt dem Anwender sehr viel Einfluss auf Form und Inhalt der Information. Allerdings erfordert dies das Auswählen der Informationen mit Hilfe eines Stifts. Diese Form der Interaktion entspricht herkömmlichen Methoden und stellt nur ein zusätzliches Informationsangebot dar. Die aufgeführten mobilen Informationsdienste werden in einem Zuordnungsverfahren berücksichtigt, das anforderungsadaptiv eine Trackingtechnik zuordnet. Um eine weitere Möglichkeit zu integrieren, die im Moment noch nicht umgesetzt worden ist, wird ein Routingdienst mit in die Berechnung aufgenommen.

Damit werden die folgenden Dienste berücksichtigt:

1. Anzeigen der derzeitigen Position in einer Karte
2. Routing
3. Anzeigen von 3D-Modellen in einem transparenten Display.

Die Anforderungen dieser Dienste an die Trackingkomponente sind beispielhaft für die Sensorklasse Position in der folgenden Tabelle dargestellt:

	Anzeigen der Position	Routing	Anzeigen von 3D-Modellen
Sensorklasse	Position	Position	Position
Genauigkeit	10 m	1 m	0.5 m
Frequenz	0 Hz	0.1 Hz	25 Hz

Wenn die Frequenz keine Anforderung darstellt, wird als Wert Null Hertz angenommen. Dies ist für das Anzeigen der derzeitigen Position in einer Karte der Fall, da der Anwender hier nicht erwartet, dass das Ergebnis automatisch aktualisiert wird. Durch das wiederholte Aufrufen der Funktion aktiviert er die Aktualisierung manuell.

7.2.2 Bestimmung der idealen Trackingtechnik

Am Beispiel der Bestimmung der Position wurde ein Demonstrator entwickelt, der zusätzlich zu den möglichen Positionierungsmethoden GPS und DGPS eine Positionierung über das EOTD-Verfahren einer GSM-Karte berücksichtigt.

Als GPS-Gerät steht ein GeoKombi 12B Plus mit einem Trimble BD112-Board zur Verfügung. Es empfängt Korrekturdaten des SAPOS-Dienstes über das RTCM-Format (Radio Technical Commission for Maritime Services), die zu einer Verbesserung der Positionierungsergebnisse

führen. Damit ist das Gerät DGPS-fähig. Die GPS-Ergebnisse werden laufend über das NMEA-Format ausgelesen. Das GPS-Gerät kann nur Ergebnisse liefern, wenn mindestens vier Satelliten sichtbar sind. Diese Information ist über das NMEA-Format abrufbar, genauso wie die Information, ob die Position mit Hilfe der Korrekturdaten berechnet wurde oder ob ein reines GPS-Ergebnis vorliegt. Diese aktuellen Informationen zum Zustand der GPS-Ergebnisse werden über die Integration des Kalmanfilters berücksichtigt, dessen Umsetzung in Abschnitt 7.2.3 vorgestellt wird.

Für die Zuordnungsberechnung der GPS-Geräte werden die Herstellerdaten verwendet. Wie oben erwähnt, wird für die GSM-Karte angenommen, dass das EOTD-Verfahren genutzt wird. Die Werte dazu wurden der Literatur entnommen. Die Eingangswerte für das Zuordnungsverfahren sind in der folgenden Tabelle aufgeführt:

	DGPS	GPS	GSM
Sensorklasse	Position	Position	Position
Genauigkeit	0.4 m	3.0 m	100 m
Frequenz	5 Hz	5 Hz	0.2 Hz

Mit Hilfe dieser Werte wird eine neue Zuordnung berechnet, sobald sich die Anforderungen durch die Zuweisung eines neuen Dienstes ändern.

7.2.3 Implementierung eines Kalman-Filters für die Positionierung

Nach dem Auslesen des GPS-Gerätes werden die Ergebnisse in einen Kalman-Filter eingelesen. Um die entsprechenden Werte für die Aufstellung der Gleichungen zu erhalten, werden aus dem NMEA-Format der VTG- und der GGA-Satz ausgelesen. Damit ist eine Berücksichtigung des DOP-Wertes möglich. Um die Nichtlinearität der Umwandlung der WGS-84-Koordinaten in Gauß-Krüger-Koordinaten zu umgehen, werden die umgerechneten Koordinaten für die Berechnung genutzt. Für alle im Kalmanfilter verwendeten Daten müssen ein Startwert sowie eine Kovarianzmatrix vorhanden sein. Da die Position während der gesamten Laufzeit des Systems bestimmt wird, stehen dafür immer Startwerte zur Verfügung. Im Folgenden ist der Ablauf des Kalmanfilters für die Bestimmung der Position mit dem GPS-Gerät vorgestellt.

Die Gleichung für die Überführung des Zustandes $x_k = A_k x_{k-1} + B u_k + w_{k-1}$ ergibt sich damit zu

$$x_k = \begin{pmatrix} 1 & t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & t \\ 0 & 0 & 0 & 1 \end{pmatrix} x_{k-1} + w_{Pos}$$

Dabei ist t die Zeit, die zwischen zwei Messungen verstrichen ist. Durch die Einbeziehung der Geschwindigkeit in die Zustandsgleichung sind bessere Aussagen möglich. w_{Pos} ist die Genauigkeit der Positionierung. Sie kann über die Genauigkeit der Punktkoordinaten ermittelt werden. Diese hängt von den Spannungen ab, die innerhalb des Netzes der Landesvermessungspunkte auftreten, in die alle Punkte eingepasst werden. Die Genauigkeiten werden

herangezogen, um die Matrix Q aufzustellen.

Die Gleichung für die Messung $z_k = H x_k + v_k$ wird dabei durch folgende Beziehung ausgedrückt:

$$z_k = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} x_k + v_{GPS}$$

v_{GPS} ist die Genauigkeit der Messungen mit dem GPS-Gerät. Bei dem vorliegenden Sensor wird diese daraufhin unterschieden, ob ein Korrektursignal empfangen wird, also DGPS-Messungen vorliegen. Ist dies nicht der Fall, sinkt die Genauigkeit des Gerätes. Die Genauigkeit der Messungen geht in die Matrix R ein.

Beim Ablauf des Kalmanfilters wird zunächst der Startwert über die aktuelle Messung bestimmt. Danach werden die Vorhersagen für die Messung sowie die Kovarianzmatrix P ermittelt. Nach der Messung werden die Vorhersagen verbessert und die Werte der nächsten Messung vorhergesagt.

Für den Einsatz des Kalmanfilters für die Positionierung wurden zwei Filter implementiert. Der Positionsfilter geht von einem konstanten Systemrauschen aus, der adaptive Filter hingegen merkt sich den letzten Zustandsvektor und bildet aus diesem und dem aktuellen Zustand eine gemittelte Beschleunigung, mit deren Hilfe das Systemrauschen bei jedem Schritt angepasst wird. Im Rahmen einer Messreihe mit dem eingesetzten GPS-Gerät wurde eine Strecke zwischen zwei bekannten Punkten zurückgelegt und anschließend mit den beiden Varianten gefiltert. Die Abweichung bei dieser Messreihe liegt für den Positionsfilter bei 0.60 m und für den adaptiven Filter bei 0.57 m. Die vom Hersteller angegebene Genauigkeit von 0.4 m für DGPS konnte damit nicht nachgewiesen werden. Das Vorhandensein von Bäumen in der Parkanlage führte allerdings zu Abschattungen des Signals. Aus diesem Grund sind die Genauigkeiten vertretbar. Grundsätzlich kann gesagt werden, dass durch den Filter die Messungen geglättet werden. Damit ist eine gleichmäßigere Einblendung von Informationen in das Blickfeld des Anwenders möglich.

7.3 Tracking auf Basis eines 3D-Modells

Um hochgenaue Trackingergebnisse im urbanen Umfeld zu ermöglichen, wurden Komponenten des in Kapitel 6 erläuterten Verfahrens innerhalb des Projektes GEIST umgesetzt. Das Verfahren basiert auf dem Vergleich von Daten einer Kamera, die das Sichtfeld des Benutzers repräsentiert, mit den Daten eines 3D-Modells der Umgebung. Die Umsetzung einer Kalibrierungsmethode in Abschnitt 7.3.1 ist die Voraussetzung für die Berechnung der Orientierung zwischen zwei Kameras. In Abschnitt 7.3.2 wird zunächst die Datengrundlage vorgestellt, auf der die Berechnungen basieren. In Abschnitt 7.3.3 wird dann die Implementierung der Bildverarbeitungsschritte erläutert. Mit Hilfe dieser Berechnungen kann die Position und die Blickrichtung des Anwenders bestimmt werden.

7.3.1 Kalibrierung der Kamera

Die umgesetzte Kalibrierungsmethode basiert auf dem Algorithmus von Zhang, der in Abschnitt 6.5.1 als Basis für aktive halbautomatische Verfahren zur Kamerakalibrierung eingeführt wurde. Im vorliegenden Fall wird dafür ein zweidimensionales Kalibrierungsmuster eingesetzt, das in der folgenden Abbildung 44 dargestellt ist.

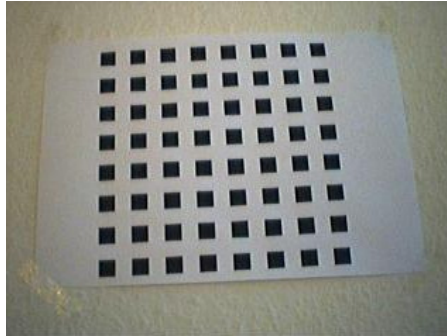


Abbildung 44: Eingesetztes Kalibriermuster

Die Objektpunkte, die für die Kalibrierung herangezogen werden, werden automatisch detektiert. Dafür wurde der Harris-Eckendetektor implementiert [Harr88]. Lokale Änderungen innerhalb des Bildes werden mit Hilfe einer lokalen Autokorrelationsfunktion detektiert. Ein Ausschnitt des Bildes wird mit seiner Umgebung korreliert. Die daraus zu berechnende Autokorrelationsfunktion zeigt einen Hochpunkt, wenn sich in dem entsprechenden Bildausschnitt eine lokale Grauwertänderung befindet. Mit Hilfe der Homographie-Relation wird die Beziehung zwischen dem detektierten Bildpunkt und dessen Objektpunkt hergestellt. Die Homographie fasst die inneren Kameraparameter und die extrinsischen Kameraparameter zusammen. Letztere beschreiben die Transformation eines Objektpunktes in den entsprechenden Kamerapunkt mit Hilfe einer Rotation und einer Translation.

Da es sich um ein aktives halbautomatisches Kalibrierungsverfahren handelt, sind verschiedene Aufnahmen des Kalibrieremusters mit Hilfe der zu kalibrierenden Kamera notwendig. Dabei muss die Stellung der Kamera bezüglich des Kalibrieremusters bei jeder Aufnahme geändert werden. Die Implementierung erlaubt die Bestimmung der folgenden Parameter:

- Hauptpunkt
- Brennweite bzw. berechnete Bildvergrößerung in beiden Koordinatenrichtungen
- Verzerrungsparameter der Linse

Die hier vorgestellte Implementierung umfasst eine Bedienungsfläche, mit der einzelne Parameter geändert und die Ergebnisse sowie Zwischenresultate kontrolliert werden können.

7.3.2 Herstellung der Datenbasis

Um das mit einer Kamera aufgenommene Videobild mit Daten eines 3D-Modells zu vergleichen, muss dieses zunächst erzeugt werden. Im vorliegenden Fall standen für Heidelberg Geometriedaten der Gebäude aus der Auswertung von Laserscanning-Messungen zur Verfügung. Abbildung 45 zeigt einen Ausschnitt aus diesem Modell in der herkömmlichen Version.



Abbildung 45: Ausschnitt aus dem texturierten Modell von Heidelberg

Hier wurden auf die Gebäudefassaden Texturen, die aus entzerrten Photographien stammen, aufgebracht. Der Nachteil dieser anspruchslosen Texturierung ist, dass beispielsweise Bäume in den entzerrten Photos auch auf die Fassaden aufgemappt in das 3D-Modell übernommen werden. Diese Form der Texturierung ist damit verhältnismäßig schnell und bequem, beinhaltet aber Fehlerquellen, die bei einer Nutzung dieser Daten berücksichtigt werden müssten. Aus diesem 3D-Stadtmodell wurden für die Verwendung im Projekt GEIST nur die Standardgeometrien verwendet. Dies bedeutet, dass die Texturierung nicht übernommen wurde. Die räumlichen Daten wurden dann durch manuelle Modellierungen ergänzt. Das bedeutet, dass die einzelnen Bereiche des 3D-Stadtmodells, in denen das hochgenaue Tracking stattfindet, manuell nachbearbeitet wurden.

Um eine Ansicht aus diesem Modell zu generieren werden zunächst Daten für die Position und Ausrichtung der entsprechenden Kamera benötigt. Diese werden mit Hilfe des GPS-Sensors und des Orientierungstrackers als Näherungswerte ermittelt. Über Klassen der Java3D-Bibliothek wird ein Offscreen-Rendering an der durch die Sensordaten festgelegten Position mit der entsprechenden Ausrichtung durchgeführt. Dieses gerenderte Bild wird in Grauwerte umgewandelt und an eine Matrix übergeben. Entsprechend des Formates der Kamerabilder hat diese ebenfalls die Dimensionen 640 x 480.

7.3.3 Bildverarbeitung

Die Kamerabilder werden mit einer USB-Kamera erstellt, wobei der Videostrom kontinuierlich eingelesen wird. Die Bilder werden mit dem Java Media Framework JMF, wie oben erwähnt, in der Größe 640 x 480 Pixel eingelesen. Sie werden dann in Grauwerte umgewandelt und in einer Matrix übergeben.

Abbildung 46 zeigt ein Videobild am Universitätsplatz in Heidelberg, das mit der verwendeten USB-Kamera aufgenommen wurde, im Vergleich mit der aus dem 3D-Modell generierten Ansicht an der Position der Sensordaten.



Abbildung 46: Videobild und an dieser Position generierte Ansicht aus dem 3D-Modell

Beide Bilder werden im Folgenden als in Grauwertbilder umgewandelte Matrizen weiterverwendet. Ein Überblick über die einzelnen Schritte des daran anschließenden Verfahrens wurde in Abschnitt 6.2 gegeben. Ausgangspunkt bildet dabei die Umsetzung der Detektion von herausragenden Merkmalen, die in Abschnitt 6.3.4 vorgestellt wurde. Dies wird durch die Verwendung des Canny-Kantendetektors verwirklicht, der auf beide Bilder angewendet wird. Das Ergebnis ist in Abbildung 47 erkennbar.



Abbildung 47: Bearbeitung des Kamerabildes (links) und der Ansicht (rechts) mit dem Canny-Kantendetektor

Der Canny-Kantendetektor liefert Kanten, die zur Korrespondenzfindung nicht lang genug sind, da sie durch Rauschen in einzelne Teile zerlegt wurden. Die Kombination mit den Ergebnissen des SUSAN-Eckendetektors ermöglicht es die wahren Kanteneckpunkte genauer zu detektieren und so längere Kanten zu erhalten. Diese Kanten werden als Vektoren mit einem Anfangs- und einem Endpunkt weitergegeben. Ergebnis sind zwei Vektorlisten, eine für das Kamerabild und eine für die generierte Ansicht.

Aus diesen beiden Merkmalslisten wird im darauf folgenden Schritt eine Zuordnung zwischen korrespondierenden Merkmalen bestimmt. Dazu wird das in Abschnitt 6.4 dargestellte Verfahren, das auf dem Algorithmus von Gros et al. basiert, angewendet. Ergebnis ist eine Liste korrespondierender Punkte im Kamerabild und in der generierten Ansicht.

Diese Zuordnung ist der Ausgangspunkt für die Berechnung der Orientierung der beiden Kameras, die durch die reale Kamera des Anwenders und die virtuelle, die für die Generierung der Ansicht genutzt wurde, repräsentiert werden. Ergebnis dieses Schrittes ist die Ausrichtung der beiden Kameras zueinander in einem Koordinatensystem, dessen Ursprung das Zentrum einer der Kameras ist. Die Ausrichtung und die Translation der beiden Kameras zueinander können bis auf einen Skalierungsfaktor bestimmt werden. Mit Hilfe der Transformation in das Koordinatensystem, das dem 3D-Modell zugrunde liegt, wird nicht nur die Skalierungsinvarianz sondern auch die Bestimmung der absoluten Orientierung gelöst. Die Position und die Ausrichtung der Kamera und damit die Positionierungsdaten des Anwenders können damit bestimmt werden.

7.4 Grenzen der umgesetzten Algorithmen

Die Ergebnisse haben gezeigt, dass Kanten und Ecken an den Stellen detektiert wurden, an denen sich gerade Linien befinden. Aus diesem Grund funktioniert das Verfahren mit Gebäuden, die keine abweichenden Formen aufweisen. Es fiel auf, dass die Ergebnisse der Merkmalsextraktion von entscheidender Bedeutung sind. Des Weiteren wurde deutlich, dass der Algorithmus dann genügend korrespondierende Punkte findet, wenn die Vektorlisten aus Modellansicht und Kamerabild die gleichen Kanten als Vektoren enthalten. Ansonsten führt das Verfahren besonders beim Vorhandensein von ähnlichen Strukturen zur Zuordnung von falschen Paaren. Es konnte gezeigt werden, dass eine hohe Übereinstimmung zwischen den beiden Datensätzen zum Auffinden von korrespondierenden Merkmalen unabdingbar ist. Die Ansichten dürfen deshalb nicht zu stark unterschiedlich sein, damit auch tatsächlich gleiche Kanten in beiden Datensätzen detektiert werden. Als Grenzwert hat sich gezeigt, dass bei einer Entfernung zum Gebäude von etwa vierzig Metern die Abweichung in der Rotation nicht mehr als acht Grad betragen darf. Ab einer Distanz von drei Metern nehmen die falschen Korrespondenzen zu und verfälschen damit das Ergebnis.

Bei Unterschieden im Kontrast, in der Beleuchtung und in den Grauwerten zwischen den beiden Datensätzen mussten die Parameter der einzelnen Detektoren für die Daten der beiden verschiedenen Quellen separat angepasst werden. Deshalb sind die Anforderungen an die Modellierung besonders hoch. Veränderte Lichtbedingungen führen zu schlechteren Ergebnissen. Aus diesem Grund ist es denkbar, dass in Zukunft bei der Generierung der Ansichten ein dynamisches Beleuchtungsmodell einen Teil der Unterschiede ausgleichen könnte, in dem beispielsweise Helligkeitswerte im Kamerabild berechnet werden.

7.5 Zusammenfassung

Im Rahmen des Projektes GEIST wurden Komponenten für die Implementierung der in dieser Arbeit aufgezeigten Verfahren entwickelt. Für die Zuordnung der idealen Trackingtechnik wird davon ausgegangen, dass zusätzlich zu den beiden zur Verfügung stehenden Diensten

- Einblenden von 3D-Rekonstruktionen in das Sichtfeld des Anwenders und
- Anzeigen der derzeitigen Position in einer Karte

ein Routing angeboten wird. Um diese Dienste den eingesetzten Positionierungsverfahren DGPS und GPS und dem zusätzlich aufgenommenen Positionierungsverfahren EOTD zuordnen zu können, wurden die entsprechenden Abläufe implementiert.

Für das Trackingverfahren auf Basis eines 3D-Modells wurde das Kalibrierungsverfahren nach Zhang umgesetzt. Außerdem wurden Bildverarbeitungsalgorithmen umgesetzt, die die Basis für die Positions- und Blickrichtungsbestimmung eines mobilen Anwenders in einer Stadt bilden.

8 Zusammenfassung und Ausblick

Das Ziel dieser Forschungsarbeit bestand darin, AR-Systeme für den Außenbereich in die Gruppe der ortsbezogenen Dienste zu integrieren.

Es handelt sich bei allen hochgenauen AR-Systemen um Insellösungen, die nur an bestimmten Stellen eingesetzt werden können. Die Verbindung verschiedener Trackingtechniken wird durch erste Ansätze angegangen. Allerdings bieten diese Systeme immer nur die Technik automatisch an, die das genaueste Ergebnis liefert. Dies lässt die Anforderungen außer acht, die an die Trackingkomponente in einem mobilen Informationssystem durch persönliche Einstellungen, Umgebungsbedingungen und vor allem die Art des Dienstes, der dem Anwender angeboten wird, gestellt werden. Oftmals reicht es aus, wenn die Ergebnisse eines ungenaueren Sensors mit einer höheren Frequenz benutzt werden können. Dies führt zu der Forderung nach einem Verfahren, das auf intelligente Weise verschiedene Trackingtechniken miteinander so verbindet, dass ein Anwender während er sich in einer Innenstadt bewegt, verschiedene Dienste angeboten bekommt, die jeweils anforderunsadaptiv ein Trackingergebnis verarbeiten.

Bisherige AR-Verfahren für den Außenbereich haben den Nachteil, dass sie nicht für den Einsatz in einer urbanen Umgebung anwendbar sind. Um höhere Genauigkeiten zu erhalten, muss eine Referenz herangezogen werden, die den Bezug zur Realität ermöglicht. Wenn auch nahe stehende Objekte zufriedenstellend erkannt werden sollen, muss diese Referenz entweder manuell vorbereitet werden oder der Anwender muss sich an einer bestimmten Stelle aufhalten, um blickrichtungabhängige Informationen eingeblendet zu bekommen.

Zum Abschluss werden in diesem letzten Kapitel die Ergebnisse der Arbeit zusammengefasst und ihre Bedeutung für den Einsatz mobiler Informationssysteme im urbanen Umfeld, die AR-Dienste berücksichtigen, aufgezeigt. Anschließend wird ein Ausblick auf zukünftige Arbeiten in diesem Forschungsgebiet gegeben.

8.1 Zusammenfassung der Ergebnisse

Es wurden zunächst Trackingtechniken untersucht, die für den Einsatz im Freien geeignet sind. Außerdem wurde ein Konzept sowohl für die Zuordnung dieser Techniken zu verschiedenen Diensten als auch für eine hochgenaue videobasierte Methode aufgezeigt und beide Ansätze prototypisch realisiert.

Die wichtigsten Ergebnisse dieser Forschungsarbeit sind unter wissenschaftlichen Gesichtspunkten heraus:

- Ein Verfahren zur Bestimmung der bestmöglichen Zuordnung einer Trackingtechnik zu einem bestimmten Dienst
- Ein Verfahren zum hochgenauen Tracking einer Person im urbanen Umfeld basierend auf den Daten eines 3D-Modells der Umgebung
- Die prototypische Realisierung eines Informationssystems, das die Ergebnisse verschiedener Trackingsensoren verwaltet und bestmöglich einsetzt

Zuordnung einer Trackingtechnik zu einem ortsbezogenen Dienst

Die Formulierung von Anforderungen ortsbezogener Dienste (Abschnitt 5.3.2) bildet die Voraussetzung um einem Dienst die bestmögliche Trackingtechnik zuweisen zu können. Anhand von Anwendungen wurden verschiedene Möglichkeiten aufgezeigt, wie ortsbezogene Dienste einen mobilen Anwender unterstützen können, der in einer Stadt zu Fuß unterwegs ist. Diese Dienste wurden klassifiziert und daraus Anforderungen abgeleitet, die über Parameter beschrieben werden können. Dabei spielen die folgenden Größen eine Rolle:

- Sensorklasse
- Genauigkeit
- Ergebnisfrequenz

Um eine Parametrisierung der Trackingtechniken zu ermöglichen wurden verschiedene Techniken auf ihren Einsatz im Freien hin untersucht. Von denjenigen, die für einen Einsatz in einer Stadt zur Verfügung stehen, wurden in Kapitel 2, neben einer technischen Einführung, die Einflüsse vorgestellt, die die Zuverlässigkeit der Sensoren beeinflussen. Zusätzlich zu den Eigenschaften der Sensoren, die für die Formulierung der Anforderungen genutzt werden, muss zur Laufzeit festgestellt werden, ob eine Technik im Moment verfügbar ist. Dafür können die folgenden Parameter genutzt werden:

- Einsatzgebiet
- Zuverlässigkeit
- Initialisierungszeit

Diese Eigenschaften wurden für alle betrachteten Techniken näher untersucht. Für alle zu einem bestimmten Zeitpunkt möglichen Trackingtechniken kann daraufhin ein statistisches Verfahren angewendet werden, das auf der Verwendung eines Index basiert.

Dieses Verfahren der Zuordnung einer Trackingtechnik zu einem ortsbezogenen Dienst ist eine Weiterentwicklung von Verfahren, die verschiedene Trackingtechniken miteinander kombinieren ([Reit01] [Baue01]), um so beispielsweise die Position einer Person zu verfolgen, die sich von einem Gebäude ins Freie bewegt. Hierbei muss ein Wechsel der eingesetzten Trackingtechniken erfolgen. Je automatischer diese Verfahren funktionieren, desto flächendeckender können die Techniken eingesetzt werden.

Das Besondere an dem hier vorgestellten Verfahren ist, dass die Trackingtechnik auf die Anforderungen reagiert, die an das Positions- und Orientierungsergebnis gestellt werden. Damit wird nicht unbedingt die genaueste Technik angesprochen, sondern diejenige, deren Ergebnisse am besten zu den Anforderungen passen.

Das aufgezeigte Verfahren kann für jeden beliebigen Dienst und für jede denkbare Trackingtechnik eingesetzt werden. Dabei können die für die Zuordnung verwendeten Parameter gegebenenfalls verändert und erweitert werden. Die mathematische Lösung für die Zuordnung beruht auf der Berechnung eines Index und ist damit nicht auf den Einsatzbereich der Zuordnung von Trackingverfahren beschränkt.

Trackingverfahren zum Einsatz in einem urbanen Umfeld

Zur Bestimmung von Position und Blickrichtung einer Person werden das Bild einer Kamera, die das Sichtfeld des Anwenders darstellt, und Daten eines 3D-Modells der Umgebung, die für die Generierung einer Ansicht genutzt werden, miteinander verglichen. Das Auffinden korrespon-

dierender Punkte ermöglicht eine geometrische Zuordnung des Bildes zu der Ansicht und damit zwischen dem Projektionszentrum der tatsächlichen Kamera und dem der Kamera, die für die Generierung der Ansicht herangezogen wurde. Über Koordinatentransformationen unter Berücksichtigung der Daten des 3D-Modells kann die Position und Ausrichtung der tatsächlichen Kamera berechnet werden.

Das entwickelte Trackingverfahren auf Basis eines 3D-Modells gliedert sich in Entwicklungen im Bereich videobasierter Trackingverfahren für den urbanen Außenbereich ein, die den Bezug zu einer Referenz herstellen. Diese wird über 3D-Modelle [Ribo02] oder 3D-GIS [Chen99] erstellt, wobei die Daten vorverarbeitet werden. Diese Vorverarbeitung ist bei dem hier vorgestellten Verfahren nicht mehr notwendig. Es wird davon ausgegangen, dass immer mehr Städte flächendeckende 3D-Stadtmodelle anbieten werden, wobei die für ein solches Verfahren interessanten Bereiche manuell detaillierter modelliert werden. In diesem Fall ist darauf zu achten, dass die Georeferenzierung aller Merkmale erhalten bleibt.

Neu an dem aufgezeigten Verfahren ist, dass es die Nachteile bisheriger Lösungen beseitigt:

- Es ist keine Vorverarbeitung der für das Matching eingesetzten Daten notwendig.
- Der Anwender kann sich frei bewegen und ist nicht auf einen Viewpoint angewiesen.
- Es werden tatsächlich die Position und Blickrichtung des Anwenders bestimmt und nicht nur eine 2D-Verschiebung der Referenz. Damit ist das Einblenden von 3D-Modellen möglich.

Der Einsatz von Daten eines 3D-Stadtmodells ermöglicht die direkte Verbindung zu weiteren Informationen, die beispielsweise in einem 3D-GIS gespeichert werden. Auf diesem Weg ist nicht nur ein direkter Datenzugang geschaffen, die Analyse- und Präsentationsfunktionalitäten eines GIS können so ebenfalls vor Ort genutzt werden.

Prototypische Implementierung

Innerhalb des Projektes GEIST wurde eine Trackingkomponente prototypisch realisiert. Sie besteht aus einem Trackingsystem, mit dem im urbanen Umfeld 3D-Rekonstruktionen positionsgenau im Blickfeld des Anwenders platziert werden können. Außerdem ermöglicht sie die Verbindung verschiedener Trackingsensoren, die abhängig von Daten zur Umgebung verschiedene Dienste anbietet, welchen über die Implementierung des Zuordnungsverfahrens anforderungsadaptiv eine Trackingtechnik zugeordnet wird.

8.2 Ausblick

Beide in dieser Arbeit vorgestellten Verfahren bilden Anknüpfungspunkte für weiterführende relevante Forschungsarbeiten. Dabei stehen die Zunahme mobiler Geräte und die damit verbundene Entwicklung innovativer ortsbezogener Dienste im Mittelpunkt.

Bei einer Auslagerung von Komponenten des videobasierten Trackingverfahrens ist es möglich, einzelne Funktionalitäten auch auf einem PDA anwenden zu können. Hier steht dann nicht die Einblendung von Informationen in das Sichtfeld des Benutzers im Mittelpunkt, sondern es ist denkbar, dass mit Hilfe des entwickelten Verfahrens weitere Daten über die Umgebung gesammelt und beispielsweise in einer zentralen Datenbank abgelegt werden können. Dies wird dadurch erleichtert, dass kleine aber meist sehr hochauflösende Kameras in mobilen Geräten in der Zwischenzeit zur Grundausstattung gehören. Eine Datenbankanbindung der 3D-Stadtmodell-daten ermöglicht es, den Schritt des Renderns und der Merkmalsextraktion datenbankseitig erfolgen zu lassen.

Die Herleitung der Anforderungen und die anforderungsadaptive Zuordnung der Trackingtechnik werden bei dem vorgestellten Verfahren jedes Mal neu berechnet. Durch Einsatz eines selbstlernenden Verfahrens können die Initialisierungsphase und die Berechnungszeit verkürzt werden.

Die Akzeptanz mobiler Endgeräte steigt mit der Möglichkeit mit anderen Personen in Kontakt zu treten. Die Berücksichtigung der Erfahrungen anderer Personen in einer Umgebung und das damit verbundene Teilen von Wissen ist ein weiterer Punkt, der für eine Weiterentwicklung von Werkzeugen und Diensten für mobile Endgeräte als sinnvoll erachtet wird. Damit können Verfahren entwickelt werden, die die Trackingresultate der verschiedenen Anwender miteinander koppeln um so zusätzliche Informationen berücksichtigen zu können.

Das vorgestellte videobasierte Verfahren benötigt Referenzen, die nur für ausgewählte Bereiche innerhalb einer Innenstadt vorliegen. Es existieren erste Ansätze, texturierte Gebäude möglichst automatisch in 3D zu erfassen. Forschungsbedarf besteht nicht nur darin, Verfahren zur anschließenden Visualisierung der Daten zu entwickeln sondern auch Methoden für diese Form der Datenerfassung zu finden. Damit betrifft diese Form des hochgenauen Trackings nicht nur den Bereich mobiler Dienste sondern auch den der Geodatenerfassung, wobei sowohl im Forschungs- als auch im Anwendungsbereich Verfahren immer wichtiger werden, die eine berührungslose 3D-Datenerfassung möglichst automatisch erlauben.

9 Literaturverzeichnis

- [Abou00] Abousalem, M., Lusin, S., Tubalin, O. und de Salas, J. Performance Analysis of GPS Positioning Using WAAS and EGNOS. GNSS Conference, Edinburgh, Großbritannien, 2000.
- [Acke99] Ackermann, F. Airborne Laser Scanning - Present Status and Future Expectations. ISPRS Journal of Photogrammetry & Remote Sensing, S. 64-67, 1999.
- [Addl01] Addlesee, M., Curwen, R., Hodges, S., Newman, J., Steggles, P., Ward, A. und Hopper, A. Implementing a Sentient Computing System. IEEE Computer Magazine, Vol. 34, S. 50-56, 2001.
- [Arbe00] Arbeitsgruppe Qualitätssicherung. Leitfaden Qualitätssicherung – Photogrammetrie und DTM-Generierung. Richtlinie der Konferenz der Kantonalen Vermessungsämter, Schweiz, Juli 2000.
- [Arch03] Projekt ARCHEOGUIDE. Internet: <http://archeoguide.intranet.gr/project.htm> am 10.07.2003.
- [Azum98] Azuma, R. T., Hoff, B. R., Neely III, H. E., Sarfaty, R., Daily, M. J., Bishop, G., Vicci, L., Welch, G., Neumann, U., You, S., Nichols, R. und Cannon, J. Making Augmented Reality Work Outdoors Requires Hybrid Tracking. Proceedings of IWAR, 1998.
- [Azum99a] Azuma, R., Hoff, B., Neely III, H. und Sarfaty, R. A Motion-Stabilized Outdoor Augmented Reality System. Proceedings of IEEE VR, S. 252-259, 1999.
- [Azum99b] Azuma, R. T. The Challenge of Making Augmented Reality Work Outdoors. Yuichi Ohta and Hideyuki Tamura (ed.): Mixed Reality: Merging Real and Virtual Worlds, Springer Verlag, Kapitel 21, S. 379-390, 1999.
- [Bahl00] Bahl, P. und Padmanabhan, V. N. RADAR: An In-Building RF-based User Location and Tracking System. INFOCOM 2000.
- [Baue03] Bauer, M. Vermessung und Ortung mit Satelliten – GPS und andere satellitengestützte Navigationssysteme. Herbert Wichmann Verlag Heidelberg, 2003.
- [Baue01] Bauer, M., Bruegge, B., Klinker, G., Mac Williams, A., Reicher, T., Riß, S., Sandor, C. und Wagner, M. Design of a Component-Based Augmented Reality Framework. Proceedings of The Second IEEE and ACM International Symposium on Augmented Reality ISAR 2001.
- [Bede95] Bederson, B. Audio Augmented Reality: A Prototype Automated Tour Guide. ACM Human Computer in Computing Systems Conference CHI, S. 210-211, 1995.
- [Behr99] Behringer, R. Registration for Outdoor Augmented Reality Applications Using Computer Vision Techniques and Hybrid Sensors. Proceedings of IEEE VR 1999.

- [Bell02] Bellotti, F., Berta, R., de Gloria, A. und Margarone, M. User Testing a Hypermedia Tour Guide. IEEE Pervasive Computing, S. 33 – 41, April – June 2002.
- [Bill02] Bill, R. Virtuelle Realität und GIS im kommunalen Umfeld. In: Zeitschrift für Vermessungswesen. Heft 1, 2002.
- [Bill99a] Bill, R. Grundlagen der Geo-Informationssysteme. Band 1. Hardware, Software und Daten. Herbert Wichmann Verlag, Heidelberg, 1999.
- [Bill99b] Bill, R. Grundlagen der Geo-Informationssysteme. Band 2. Analysen, Anwendungen und neue Entwicklungen. Herbert Wichmann Verlag, Heidelberg, 1999.
- [Boro97] Borotschnig, H., Pinz, A. und Sinclair D. Fuzzy Graph Tracking. Proceedings of 5th Symposium for Intelligent Robotics Systems, 1997.
- [Bour02] Bourdeau, N., Geaux, M., Riba, J. und Sansone, F. Hybridised GPS and GSM Positioning Technology for High Performance Location Based Services, 2002.
- [Bren00] Brenner, C. Towards Fully Automatic Generation of City Models. IAPRS Vol. XXXIII, Part B3/1, Comm. III, pp. 85-92, ISPRS Congress, Amsterdam, 2000.
- [BrHa00] Brenner, C. und Haala, N. Erfassung von 3D Stadtmodellen. PFG - Photogrammetrie, Fernerkundung, Geoinformation, Heft 2/2000, S.109 – 118, 2000.
- [Brow71] Brown, D. C. Lens Distortion for Close-Range Photogrammetry. Photometric Engineering, Vol. 37 No. 8, S. 855-866, 1971.
- [Chen99] Chen, T. und Shibasaki R. A Versatile AR Type 3D Mobile GIS Based on Image Navigation Technology. IEEE International Conference on Systems, Man and Cybernetics, S. 1070-1075, 1999.
- [Coor03] Coors, V., Gjesdal, O., Sulebak, J. R. und Lakso, K. 3D Maps for Boat Tourists. International Conference on Information and Communication Technologies in Tourism, S. 298-305, 2003.
- [Coor01] Coors, V. 3D-GIS in Networking Environments. Proceedings of 3D Cadastre, Delft, 2001.
- [Coor00a] Coors, V., Huch, T. und Kretschmer, U. TouriSTAR – ein mobiles AR System in einer urbanen Umgebung. Schmidt, B. et al.: Visualisierung raumbezogener Daten: Methoden und Anwendungen, Bd. II. Solingen, S. 191-206, 2000.
- [Coor00b] Coors V., Huch T. und Kretschmer, U. Matching Buildings: Pose Estimation in an Urban Environment. Proceedings of IEEE and ACM International Symposium on Augmented Reality ISAR, München, 2000.
- [Dark99] Darken, R. P. und Cevik, H. Map Usage in Virtual Environments: Orientation Issues. Proceedings of IEEE VR, S. 133-140, 1999.
- [Deum88] Deumlich, F. Instrumentenkunde der Vermessungstechnik. VEB Verlag für Bauwesen, Berlin, 1988.

- [Doug73] Douglas, D. und Peucker, T. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. The Canadian Cartographer 10(2), S. 112-122, 1973.
- [Enca97] Encarnação, J., Straßer, W., Klein, R. Graphische Datenverarbeitung: 2. Modellierung komplexer Objekte und photorealistische Bilderzeugung. Oldenbourg Verlag, 1997.
- [Fein99] Feiner, S. K. The Importance of Being Mobile: Some Social Consequences of Wearable Augmented Reality Systems. Proceedings of International Workshop on Augmented Reality the IWAR, San Francisco, S. 145-148, 1999.
- [Fein97] Feiner, S., MacIntyre, B., Höllerer, T. und Webster, A. A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment. International Symposium on Wearable Computers. S. 74-81, Cambridge, MA, 1997.
- [Floo01] Flood, M. Laser Altimetry: From Science to commercial Lidar Mapping. Journal of the American Society for Photogrammetry and Remote Sensing, Volume 67, Number 11, 2001.
- [Förs99] Förstner, W. 3D-City Models: Automatic and Semiautomatic Acquisition Methods. Fritsch, D. und Spiller, R. (Hrsg.): Photogrammetrische Woche, Wichmann, Heidelberg, S. 291-304, 1999.
- [Fuch98] Fuchs, C. und Heuel, S. Feature Extraction. Institut für Photogrammetrie, Universität Bonn, Februar 1998.
- [Gali03] GALILEO - Europäisches Satellitennavigationssystem, Internet: http://europa.eu.int/comm/dgs/energy_transport/galileo am 10.12.2003.
- [Grif00] Griffin, B. und Huber, B. Beschreibung inertialer Meßsysteme und Beurteilung der Verwendbarkeit als AR-Tracker. Studie erstellt im Auftrag des Fraunhofer-IGD, GeneSys Elektronik GmbH, 2000.
- [Gros95] Gros, P., Bournez, O. und Boyer, E. Using Geometric Quasi-Invariants to Match and Model Images of Line Segments. Rapport de recherche no 2608, INRIA, Grenoble, France, 1995.
- [Grue01] Gruen, A. Cities from the Sky – Photogrammetric Modeling of CyberCity is coming of Age. GeoInformatics, Ausgabe Dezember 2001.
- [Gülc01] Gülc, E. New Features in Semi-Automatic Building Extraction. Proceedings of ASPRS Conference, St. Louis, 2001.
- [Gülc00] Gülc, E., Müller, H. und Läbe, T. Semi-automatische Verfahren in der photogrammetrischen Objekterfassung, PFG - Photogrammetrie Fernerkundung Geoinformation 3/2000, Schweizerbart'sche Verlagsbuchhandlung, Stuttgart, S. 200-209, 2000.
- [Harr88] Harris, C. G. und Stephens, M. A Combined Corner and Edge Detector. Alvey Vidion Conference, S. 147-151, 1988.

- [Hart02] Hartley, R. und Zisserman, A. Multiple View Geometry in Computer Vision. Cambridge University Press, Reprinted 2002.
- [Hart94] Harter, A. und Hopper, A. A Distributed Location System for the Active Office. IEEE Network 6(1), S. 62-70, Jan-Feb 1994.
- [Heue00] Heuel, S. Zur automatischen Erfassung von Gebäuden aus Luftbildern. PFG - Photogrammetrie Fernerkundung Geoinformation 3/2000, Schweizerbart'sche Verlagsbuchhandlung, Stuttgart, S. 177-188, 2000.
- [IPB03] Institut für Photogrammetrie, Universität Bonn. System zur 3D-Objekterfassung für CAD und Virtual Reality Anwendungen. Internet: <http://www.ipb.uni-bonn.de/ipb/projects/semi-automatic.html> am 10.07.2003.
- [IVVM99] Informations-Verarbeitungs-Versorgungseinheit (IVV) Geowissenschaften. Vorlesungsskript GPS-Anwendungen in den Geowissenschaften. Westfälische Wilhelms-Universität Münster, Internet: <http://ivvgeo.uni-muenster.de> am 24.07.1999.
- [Juli00] Julier, S., Lanzagorta, M., Baillot, Y., Rosenblum, L., Feiner, S., Höllerer, T. und Sestito, S. Information Filtering for Mobile Augmented Reality. Proceedings of IEEE and ACM International Symposium on Augmented Reality ISAR, München, 2000.
- [Kalk02] Kalkbrenner, G. und Köppe, E. Mobile Management of Local Infrastructure. IEEE Softcom 2002.
- [Kell94] Kelly, A. Modern Inertial and Satellite Navigation Systems. Technical Report CMU-RI-TR-94-15, Robotics Institute, Carnegie Mellon University, May, 1994.
- [Knor00] Knorr, F. Airborne-Laser-Scanning. Internet: <http://www.lmv-gmbh.de/aktuell/AirbornLaserScanning9911.pdf> am 23.11.2000.
- [Krau00] Kraus, K. Photogrammetrie Band 3: Topographische Informationssysteme. Dümmler Verlag Köln, S. 275-299, 2000.
- [Kret03] Kretschmer, U. Measuring Positions of a User in Altering Settings. Bieber, Kirste: IMC 2003. Proceedings of the 4th International Workshop of Mobile Computing, Rostock, 2003.
- [Kret02a] Kretschmer, U. Using Mobile Systems to Transmit Location Based Information. Kalliany, R. et al.: International Society of Photogrammetry and Remote Sensing and Spatial Information Sciences (ISPRS), Proceedings of Photogrammetric Computer Vision 2002.
- [Kret02b] Kretschmer, U. und Holweg, D. GEIST – Ein mobiles AR-Informationssystem zum Erleben historischer Zusammenhänge im urbanen Umfeld mit Digital Storytelling. Bundesministerium für Bildung und Forschung: Internationale Statustagung „Virtuelle und Erweiterte Realität“, 2002.
- [Kret01] Kretschmer, U., Coors, V., Spierling, U., Grasbon, D., Schneider, K., Rojas, I. und Malaka, R. Meeting the Spirit of History. Proceedings of Virtual Reality, Archaeology, and Cultural Heritage VAST, Griechenland, 2001.

- [Kret00] Kretschmer, U. und Roccatagliata, E. Common-GIS: A European Project for an Easy Access to Geo-data. Markus, B: Second European GIS Education Seminar. Proceedings. Ungarn, 2000.
- [Lait01] Laitinen, H., Ahonen, S., Kyriazakos, S., Lähteenmäki, J., Mensolascino, R. und Prkkila, S. Cellular Location Technology. Deliverable of the CELLO Project (IST-2000-25382-CELLO), 2001.
- [Land98] Landeshauptstadt Stuttgart, Stadtmessungsamt, Abteilung Geoinformation. 3D-Stadtmodell. Beschreibung, 1998.
- [Lang95] Lang, F. und Förstner, W. Matching Techniques. Institut für Photogrammetrie, Universität Bonn, 1995.
- [LoAk01] Lo, S. C., Akos, D., Houck, S., Normark, P. L. und Enge P. WAAS Performance in the 2001 Alaska Flight Trials of the High Speed Loran Data Channel. IEEE Position Location and Navigation Symposium, Palm Springs, CA, 2001.
- [Luhm00] Luhmann, T. Nahbereichsphotogrammetrie – Grundlagen, Methoden und Anwendungen. Herbert Wichmann Verlag, Heidelberg, 2000.
- [Mayer99] Mayer, H. Automatic Object Extraction from Aerial Imagery - A Survey Focusing on Buildings. Computer Vision and Image Understanding, 1999.
- [Monda03] Monda, E. W., Glenn Lightsey, E. und Key, K. An Investigation of GPS Pseudolite Based Relative Navigation. American Astronautical Society Astroynamics Specialist Conference, 2003.
- [Neuk99] Neukum, G. The Airborne HRSC-A: Performance Results and Application Potential. Photogrammetrische Woche, Stuttgart, 1999.
- [Oste01] Ostensen, O. The Expanding Agenda of Geographic Information Standards. ISO Bulletin, S. 16-21, 2001.
- [Peyr00] Peyrard F., Soutou, C. und Mercier, J. Mobile Stations Localization in a WLAN. IEEE Conference on Local Computer Networks LCN, 2000.
- [Piek99] Piekarski, W., Gunther, B. und Thomas, B. Integrating Virtual and Augmented Realities in an Outdoor Application. Proceedings of IEEE IWAR 1999.
- [Point03] Point Research Corp. DRM-III OEM Dead Reckoning Module for Personnel Positioning. Internet: www.pointresearch.com am 10.07.2003
- [Reit01] Reitmayr, G. und Schmalstieg, D. An Open Software Architecture for Virtual Reality Interaction. ACM VRST 2001.
- [Ribo02] Ribo, M., Lang, P., Ganster, H., Brandner, M., Stock, C. und Pinz, A. Hybrid Tracking for Outdoor Augmented Reality Applications. IEEE Computer Graphics and Applications, November/December 2002.
- [Schm03] Schmidt, F. Hochgenaue Digitale Geländemodelle – Untersuchung zur Erstellung, Analyse und Anwendung in der Landwirtschaft. Dissertation der Fakultät für Ingenieurwissenschaften der Universität Rostock, 2003.

- [Schu03] Schulz, T. und Kretschmer, U. Adaptive mobile Ortsbestimmung. Tagungsband für Mensch und Computer, 2003.
- [SenB03] Senatsverwaltung für Stadtentwicklung, Berlin, Internet: <http://www.stadtentwicklung.berlin.de/> am 10.12.2003
- [Smit95] Smith, S. M. und Brady, J. M. SUSAN – A New Approach to Low Level Image Processing. Technical Report, Defence Research Agency, Farnborough, Hampshire, 1995.
- [Sonk99] Sonka, M., Hlavac, V. und Boyle, R. ImageProcessing: Analysis and MachineVision. Brooks/Cole Publishing Co, PacificGrove, CA, 1999.
- [Stri01] Stricker, D. Tracking with Reference Images: A Real-Time and Markerless Tracking Solution for Out-Door Augmented Reality Applications. Proceedings of Virtual Reality, Archaeology, and Cultural Heritage VAST, Griechenland, 2001.
- [TMob99] DeTeMobil Deutsche Telekom MobilNet GmbH. Geo-Daten für den Mobilfunk 3D-Stadtmodelle. Allg. Datenbeschreibung, 1999.
- [Torr96] Torr, P. H. S und Murray, D. W. The Development and Comparison of Robust Methods for Estimating the Fundamental Matrix. International Journal of Computer Vision, Kluwer Academic Publishers, S. 1-33, Boston, 1996.
- [Torr93] Torr, P. H. S. und Murray D. W. Outlier Detection and Motion Segmentation. Proceedings of Sensor Fusion VI (Editor: P. S. Schenker), Boston, USA, 1993.
- [Tsuji01] Tsujii, T., Rizos, C., Wang, J., Dai, L., Roberts, C. und Harigae, M. A Navigation/Positioning Service Based on Pseudolites Installed on Stratospheric Airships. 5th Int. Symp. on Satellite Navigation Technology & Applications, Canberra, Australia, 2001.
- [Vegt01] van der Vegt, J. W. und Hoffmann, A. Airborne Laser Scanning Reaches Maturity. GeoInformatics, S. 32-39, Ausgabe September 2001.
- [VRVI03] VRVis Zentrum für Virtual Reality und Visualisierung Forschungs GmbH, Graz, Internet: <http://www.icg.tu-graz.ac.at/~Research/vrvis/> am 10.12.2003
- [Welc02a] Welch, G. und Foxlin, E. Motion Tracking: No Silver Bullet, but a Respectable Arsenal. IEEE Computer Graphics and Applications, Ausgabe November/Dezember 2002.
- [Welc02b] Welch, G. und Bishop, G. An Introduction to the Kalman Filter, UNC-Chapel Hill, TR 95-041, 2002.
- [Wirt01] Wirtschaftsförderung Berlin GmbH. Business Location Center. 2001.
- [You99] You, S., Neumann, U. und Azuma, R. Orientation Tracking for Outdoor Augmented Reality Registration. IEEE Computer Graphics and Applications, Ausgabe November/Dezember 1999.
- [Zach01] Zach, C., Klaus, A., Bauer, J., Karner, K. und Grabner, M. Modelling and Visualizing the Cultural Heritage Data Set of Graz. Proceedings of Virtual Reality, Archaeology, and Cultural Heritage VAST, Griechenland, 2001.

- [Zhan00] Zhang, Z. A Flexible New Technique for Camera Calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence, S. 1330-1334, 2000.
- [Zhan94] Zhang, Z., Deriche, R., Faugeras, O. und Luong, Q.-T. A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry. Research Report 2273, INRIA Sophia-Antipolis, 1994.