

**People–Tracking–Verfahren
zur autonomen Identifikation
und Begleitung von Personen
durch mobile Roboter**

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch–Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich–Wilhelms–Universität Bonn

vorgelegt von

Andreas Kräußling

aus

Schweinfurt

Bonn (März) 2009

Angefertigt mit Genehmigung
der Mathematisch–Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich–Wilhelms–Universität Bonn

1. Gutachter: Prof. Dr. Armin B. Cremers

2. Gutachter: Prof. Dr. Joachim Hertzberg

Tag der Promotion: 6.8.2009

Erscheinungsjahr: 2009

Zur Erinnerung an Ludwig Lohbrandt

L ÓCÉAN, DEBORAH,
IL Y A L ÓCÉAN

Danksagung

An erster Stelle gilt mein Dank Herrn Professor Dr. Armin B. Cremers, der sich ohne Zögern zur Betreuung dieser externen Dissertation bereit erklärt hat. Auch dem zweiten Gutachter, Herrn Professor Dr. Joachim Hertzberg, und den weiteren Mitgliedern der Promotionskommission, Herrn Professor Dr. Rolf Klein und Herrn Professor Dr. Wolfgang Förstner, soll mein besonderer Dank ausgesprochen werden.

Diese Arbeit wäre aber nicht entstanden ohne die sehr großzügige Unterstützung der Forschungsgesellschaft für Angewandte Naturwissenschaften (FGAN) in Wachtberg. Zu nennen ist hier neben der fachlichen Beratung durch die Herren Professor Dr. Christopher Schlick, Frank E. Schneider, Dr. Wolfgang Koch, Dr. Martin Ulmke und vor allem Dr. Dirk Schulz als dem betreuenden Habilitanden auch die administrative Unterstützung durch meine weiteren Vorgesetzten, Herrn Professor Dr. Ludger Schmidt und Herrn Professor Dr. Jürgen Grosche. Auch die kollegialen Umgangsformen an meiner Arbeitsstelle trugen nicht unwesentlich zum Gelingen dieser Arbeit bei. Stellvertretend möchte ich hier nennen die Herren Dr. Wolf Käßler und Dr. Jörg Schweingruber. Von den Kollegen aus dem Bereich Robotik möchte ich besonders die Herren Dennis Wildermuth und Bernd Brüggemann hervorheben, die mich bei den Experimenten sehr tatkräftig unterstützt haben.

Desweiteren haben mich während der Abfassung dieser Dissertation meine Mutter, Karin Kräußling, und die Herren Bernhard Renner, Hans Münch und Walter Mantel ideell unterstützt.

Bemerkungen

Die in dieser Arbeit wiedergegebenen Abbildungen genügen nicht immer höchsten Ansprüchen. Zum einen stammen einige Graphiken aus englischen Originalveröffentlichungen, so dass Achsenbeschriftungen oder Textkommentare in Englisch sind. Zum anderen haben die verwendeten Tools zur Erzeugung der Bilder ihre Tücken, so dass nicht alle Abbildungen in der zu wünschenden Qualität erzeugt werden konnten. Der geneigte Leser wird gebeten, freundlicherweise über diese kleinen Mängel hinwegzusehen.

Die hintere Umschlagsfalte enthält eine CD mit einem Film. Dieser Film gibt die Videos zu den Experimenten aus Kapitel 11 wieder.

Zusammenfassung

Die vorliegende Arbeit stellt ein Verfahren vor, das es erlaubt, einen mobilen Roboter mit einer Fähigkeit auszustatten, die als eine der Grundfertigkeiten im Kontext einer sinnvollen Interaktion eines Roboters mit Menschen gilt, nämlich die Begleitung einer vorher bestimmten Person aus einer Personengruppe durch die mobile Einheit. Hierzu müssen im Wesentlichen folgende Kriterien erfüllt sein:

- der Roboter muss in der Lage sein, die Positionen aller Einzelpersonen und aller Personengruppen kontinuierlich zu schätzen. Hierbei haben gewöhnliche Verfahren insbesondere dann Probleme, wenn eine Personengruppe sich auflöst und die Spuren der Einzelpersonen wieder separat aufgenommen werden müssen.
- der Roboter muss den Einzelpersonen jederzeit ihre Identitäten zuordnen können, um die Person zu erkennen, der er folgen soll.
- das Verfahren muss soweit optimiert sein, dass es auch auf der mobilen Einheit mit ihrer beschränkten Rechenkapazität in Echtzeit laufen kann.

Zur Gewinnung der für die Positionsschätzung benötigten Messdaten werden im Rahmen der vorliegenden Arbeit Laserscanner verwendet, die aus Laufzeitbestimmungen die Entfernungen zu den reflektierenden Objekten berechnen. Aus diesen Entfernungswerten werden die Positionen von Einzelpersonen und Personengruppen mit zwei unterschiedlichen Verfahren ermittelt. Bei Einzelpersonen werden alle Messungen, die dem Ziel zugeordnet wurden, als ungewichtetes Mittel benutzt, während für Personengruppen ein komplizierterer Ansatz entwickelt wurde, der auf dem Viterbi-Algorithmus beruht und speziell an die Situation angepasst ist, dass sich eine Personengruppe wieder in einzelne Individuen auflöst.

Für die Identifizierung der Personen werden in dieser Arbeit zwei unterschiedliche Ansätze verfolgt. Zum einen kann, wie in der Literatur üblich, die Identität von Menschen aus Farbhistogrammen bestimmt werden, die hier mittels eines Systems aus einer Farbkamera und einer Wärmekamera gewonnen werden. Der zweite Ansatz, der in dieser Arbeit untersucht wird, ist neu und wertet die Reflektivitätsmessungen aus, die durch die Laserscanner ebenfalls angeboten werden. Da die Laserscanner auf Kniehöhe an den Robotern montiert sind, liefern diese Messungen von den Beinbekleidungen der Versuchspersonen. Hosen unterschiedlichen Materials rufen dabei oft Messungen unterschiedlicher Intensität hervor. Diese Intensitätsunterschiede sind in vielen Fällen ausreichend groß, um Personen zuverlässig erkennen zu können.

Die Optimierung der Verfahren geschieht dadurch, dass die Messdaten, die für die höheren, rechenintensiveren Programmebenen zugelassen werden, geeignet vorgefiltert werden, da der Rechenbedarf dieser höheren Ebenen vor allem durch die Anzahl

der zu verarbeitenden Messwerte festgelegt wird. Die zu dieser Vorauswahl verwendeten Verfahren sind dabei ähnlich denen, wie sie für gewöhnlich zum Aufspüren unbekannter Personen verwendet werden, und benutzen charakteristische geometrische Muster, wie sie durch Menschen typischerweise in den Laserentfernungsmessungen hervorgerufen werden.

Basierend auf diesen Ansätzen war es nun möglich, einen mobilen Roboter so zu programmieren, dass er einer bestimmten, vorher festgelegten Person folgen kann. Bei dieser Fähigkeit handelt es sich wie schon angedeutet um eine von vielen Autoren geforderte Grundfertigkeit eines mobilen Roboters, um mit Menschen sinnvoll kooperieren zu können. Im Rahmen der vorliegenden Arbeit konnte zum ersten Mal ein System vorgestellt werden, das diese Eigenschaft besitzt. Die Lauffähigkeit sämtlicher, in der Arbeit vorgestellter Verfahren wurde dabei anhand von aussagekräftigen Experimenten bewiesen.

Inhaltsverzeichnis

1	Einführung	1
2	Der Stand der Forschung	17
2.1	Einfache Verfahren	17
2.2	Die Verfolgung mehrerer Personen	19
2.3	Verfolgung und Identifikation	21
3	Mathematische Grundlagen	25
4	Einfache Algorithmen	31
4.1	Einleitung und experimentelle Rahmenbedingungen	31
4.2	Das verwendete Modell	33
4.3	Die Grundstruktur eines Verfolgungs–Algorithmus	37
4.4	Einfache Verfahren zur Verfolgung	43
4.4.1	Verfahren mit Mittelwertbildung	44
4.4.2	Verfahren ohne Mittelwertbildung	45
5	Kreisförmige Objekte	53
5.1	Motivation	53
5.2	Die geeignete Wahl der Modellparameter	58
5.3	Die Information über die Position des Objektes	60
5.4	Die Komplexität der Verfahren	70
5.5	Die Verhältnisse beim People–Tracking	72
5.6	Das Problem der Lokalisierung	73
5.7	Ausblick	77
6	Zwei interagierende Ziele	79
6.1	Eine Einführung in die Problemstellung	79
6.2	Der Cluster–Sorting Algorithmus	88
6.3	Der Switching Algorithmus	94
7	Beliebig viele Ziele	99
7.1	Einleitung	99

7.2	Das Verfahren	101
7.3	Experimente	105
8	Ein Vergleich des ECSA mit dem SJPDAF	113
8.1	Eine Einführung und kurze Beschreibung des SJPDAF	113
8.2	Der Verfahrensvergleich	118
8.2.1	Genauigkeit der Positionsschätzung und Rechenzeitbedarf . .	118
8.2.2	Zwei interagierende Ziele	120
8.3	Schlussfolgerungen	124
9	Ansätze zur Identifikation von Personen	125
9.1	Einleitung	125
9.2	Die Verwendung von Farbinformation	131
9.3	Die Verwendung von Reflexionsmessungen	142
10	Erweiterungen	161
10.1	Spuraufnahme und Präselektion	161
10.2	Das Problem der Datenfusion	169
11	Experimente in Echtzeit	175
12	Zusammenfassung und Schlussfolgerungen	195
A	Konvergenzeigenschaften des Modells	201
B	Die Berechnung des Schwerpunktes S_O	209
C	Das Simulationsprogramm	215
D	Identifikation mit einem Kamerasystem	219
E	Identifikation mit Reflektivitätsmessungen	225
E.1	Voruntersuchungen	225
E.2	Graphiken zu einigen Experimenten	230
	Literaturverzeichnis	239

Kapitel 1

Einführung

Die Robotik ist die Wissenschaft von der Erfassung und der Manipulation der realen Welt mit Hilfe von Computer-kontrollierten Maschinen. Sie ist wie die Bereiche des maschinellen Lernens, der Bildverarbeitung oder der Expertensysteme ein Teilgebiet der künstlichen Intelligenz. Diese versucht, höhere kognitive Funktionen abzubilden, unter anderem, um dadurch Rückschlüsse auf die Funktionsweise des menschlichen Gehirns zu ermöglichen (Hofstadter, 1979; Klix, 1979; Shapiro, 1987). Bei der künstlichen Intelligenz handelt es sich um ein in einem hohen Maße interdisziplinäres Forschungsgebiet. Beispiele erfolgreicher Systeme im Rahmen der Robotik sind Roboter zur Exploration anderer Planeten, Industrieroboter, Autos, die sich selbst steuern und Manipulatoren, die Chirurgen unterstützen. Robotersysteme befinden sich in der realen Welt, erhalten durch Sensoren Informationen über ihre Umgebung und interagieren mit ihrer Umwelt durch die Ausübung physikalischer Kräfte.

Um diese Ziele zu erreichen, müssen Roboter in der Lage sein, die enorme Unsicherheit zu beherrschen, die in der realen Welt existiert. Es gibt eine Reihe von Faktoren, die zu dieser Unsicherheit beitragen.

Zuerst und vor allem sind die Umgebungen wie Museen, die überwacht werden sollen, oder Krankenhäuser, in denen Patienten umher laufen, sehr dynamisch und kaum vorhersagbar. Die Unsicherheit ist insbesondere sehr groß für Roboter, die mit Menschen interagieren sollen.

Außerdem sind Sensoren beschränkt in ihrer Fähigkeit, Daten zu registrieren. Diese Beschränkungen rühren von verschiedenen Umständen her. So sind die Reichweite und die Auflösung eines Sensors durch physikalische Umstände eingeschränkt. Zum Beispiel können Kameras nicht durch Wände hindurchsehen und haben ein beschränktes räumliches Auflösungsvermögen. Sensoren sind auch immer einem statistischen Rauschen ausgesetzt, das die Messungen in unvorhersagbarer Weise verfälscht und damit die Information limitiert, die von den Sensoren erhalten werden kann. Schließlich können Sensoren auch beschädigt sein. Die Ermittlung eines fehlerhaften Sensors kann sich dabei unter Umständen sehr schwierig gestalten.

Daneben erfordern die Aktionen von Robotern Motoren, deren Verhalten bis zu einem gewissen Grad ebenfalls unvorhersagbar ist. Diese Unsicherheit entsteht durch

das statistische Rauschen der Kontrolleinheit und die Fehler und Ungenauigkeiten der Mechanik.

Eine gewisse Unsicherheit entstammt auch der Software des Roboters, da dessen Modelle über seine Umwelt lediglich Näherungen sind, die die physikalischen Prozesse des Roboters und seiner Umwelt nur ungenau wiedergeben.

Schließlich entsteht Unsicherheit durch die Ungenauigkeit der Algorithmen, die die Roboter benützen. Diese müssen nämlich in Echtzeit ausführbar sein, was die Anzahl der Berechnungen einschränkt, die ausgeführt werden können. Zahlreiche eingesetzte Algorithmen sparen deshalb Rechenzeit auf Kosten der Genauigkeit und erreichen letzten Endes erst dadurch Echtzeitfähigkeit.

Der Grad der Unsicherheit hängt in hohem Maße von der beabsichtigten Anwendung ab. In manchen Fällen, wie zum Beispiel auf dem Gebiet der Industrieroboter, können die Systeme so konstruiert werden, dass die Unsicherheit vernachlässigbar klein ist. Demgegenüber müssen sich Systeme, die in Häusern oder auf fremden Planeten operieren sollen, mit einer großen Unsicherheit auseinandersetzen. Solche Systeme müssen auch dann noch handlungsfähig sein, wenn weder ihre Sensoren noch ihre inneren Modelle ihnen genügend Informationen liefern, um noch mit absoluter Sicherheit die richtigen Entscheidungen treffen zu können. Da sich aber die Anwendungen der Robotik mehr und mehr in Richtung der realen Welt verschieben, wurde das Problem der Unsicherheit eine wesentliche Beschränkung für die Handlungsfähigkeit von autonomen Systemen. Deshalb ist der adäquate Umgang mit Unsicherheit vielleicht der wichtigste Schritt in Richtung von Systemen, die sinnvoll mit der realen Welt interagieren können.

Diese Arbeit benutzt deshalb in sehr starkem Umfang probabilistische Methoden. Die Anwendung probabilistischer Verfahren in der Robotik ist ein vergleichsweise neuer Ansatz, der die Unsicherheit eines Roboters bezüglich seiner Handlungen und seiner Sensorik berücksichtigt. Die Schlüsselidee besteht dabei darin, Unsicherheit explizit unter Verwendung der Wahrscheinlichkeitstheorie zu modellieren. Anstelle einer einzigen besten Wahl für den echten Zustand repräsentieren probabilistische Algorithmen die zur Verfügung stehende Information durch Wahrscheinlichkeitsverteilungen über eine ganze Menge möglicher Zustände. Dadurch können Zweifel und Vermutungen in mathematisch sinnvoller Form wiedergegeben werden. Dies zieht als Folgerung nach sich, dass probabilistische Methoden anderen Techniken in vielen Anwendungen in der realen Welt überlegen sind. Die probabilistische Robotik integriert nahtlos Modelle mit Sensordaten und beseitigt damit gleichzeitig deren Grenzen.

Meilensteine auf dem Gebiet der probabilistischen Robotik sind die Einführung des Kalman-Filters zur Anwendung auf hochdimensionale Messungen (Smith und Cheeseman, 1986) und die Erfindung von sogenannten Belegungs-Gitterkarten (Elfes, 1987). Das letzte Jahrzehnt war dann Zeuge einer wahren Explosion von neuen Techniken: Teilchen- oder Partikel-Filter wurden sehr populär (Dellaert et al., 1999) und einige Forscher haben neue Programmier Techniken entwickelt, die auf Bayesscher Informationsverarbeitung beruhen (Leibel et al., 2004; Thrun et al., 2000). Diese Entwicklung ging Hand in Hand mit der Entwicklung von Robotersystemen, die von

probabilistischen Ansätzen gesteuert werden, wie zum Beispiel Industriemaschinen zur Verladung von Gütern (Durrant-Whyte, 1996), Unterhaltungsroboter in Museen (Burgard et al., 2000) und Roboter im Pflegebereich (Pineau et al., 2003).

Der Kern der probabilistischen Robotik ist die Idee, einen Zustand aus Sensordaten zu schätzen. Unter Zustandsschätzung versteht man dabei allgemein das Problem der Schätzung von Größen, die nicht direkt beobachtbar sind, aus Sensordaten. Bei den meisten Anwendungen in der Robotik ist es relativ einfach, Handlungspläne zu entwerfen, wenn man bestimmte Größen genau kennt. Zum Beispiel ist die Bewegungsplanung für einen mobilen Roboter ziemlich einfach, wenn die genaue Position des Roboters und aller Hindernisse in der Umgebung bekannt sind. Unglücklicherweise sind diese Größen aber im Allgemeinen keiner direkten Messung zugänglich. Anstattdessen muss der Roboter seine Sensoren verwenden, um indirekte Informationen über diese Größen zu erhalten. Aber Sensoren liefern oft nur unvollständige Informationen über die gesuchten Größen und ihre Messungen werden wie schon erwähnt durch ein Rauschen gestört. Die Zustandsschätzung versucht deshalb, Information über die Zustände aus den Messdaten zu gewinnen. Algorithmen zur probabilistischen Zustandsschätzung berechnen nun Wahrscheinlichkeitsverteilungen über mögliche Zustände. Ein Beispiel für einen solchen Ansatz ist die sogenannte Lokalisierung. Darunter versteht man das Problem, dass ein Roboter über eine Karte seiner Umgebung verfügt und mit Hilfe dieser Karte und seiner Sensordaten seine aktuelle Position innerhalb dieser Umgebung schätzen soll (Burgard et al., 1996; Fox et al., 1999b).

Ein Roboter kann auf verschiedene Arten mit seiner Umgebung wechselwirken. Er kann zum Beispiel mit aktiven Handlungen auf die Umgebung einwirken oder mittels Sensoren Informationen über diese gewinnen. Die Umgebung oder Welt ist dabei ein dynamisches System, das einen wohldefinierten inneren Zustand besitzt. Sensoren weisen aber wie bereits erwähnt ein stochastisches Rauschen auf und gewöhnlicherweise gibt es viele Dinge in der Umgebung, die einer direkten Messung nicht zugänglich sind. Konsequenterweise kann ein Roboter deshalb nur Annahmen oder Vermutungen über den wahren Zustand seiner Umgebung entwickeln. Der wahre Zustand der Welt ist die Sammlung aller Eigenschaften des Roboters und seiner Umgebung, die die Zukunft beeinflussen können. Bestimmte Zustandsvariablen ändern sich mit der Zeit, wie z.B. die Positionen der Personen in der Umgebung des Roboters. Andere bleiben konstant, wie z.B. die Lage der Wände in den Gebäuden. Man unterscheidet deshalb dynamische von statischen Zuständen. Ein Beispiel für einen Zustand ist die Position des Roboters, die seine Koordinaten beispielsweise in einem kartesischen System und seine Orientierung einschließt. Ein anderes Beispiel sind Landmarken, unter welchen man stationäre Objekte in der Umgebung versteht, die fortlaufend beobachtet werden können. Als letztes Beispiel sollen sich in der Umgebung bewegende Objekte genannt werden. Diese besitzen eine sich ständig ändernde Position und Geschwindigkeit.

Der Inhalt dieser Arbeit gehört in den Bereich der mobilen Robotik. Dabei handelt es sich um einen multidisziplinären Ansatz zur Entwicklung von Systemen, die in realen und unvorhersagbaren Umgebungen Entscheidungen treffen und autonom han-

deln können, um vorher festgelegte Aufgaben zu erfüllen. Zu diesem Zweck sind sie in der Lage, permanent neue Positionen in ihrer Umgebung einzunehmen, falls dies erforderlich ist.

Im letzten Jahrzehnt ermöglichte der technologische Fortschritt die Entwicklung sehr effizienter Sensoren mit relativ günstigem Preis wie zum Beispiel sogenannte Laserscanner. Diese Tatsache hatte großen Einfluss auf die Forschung in der mobilen Robotik. So konnte es zum Beispiel den Robotern ermöglicht werden, relativ komplizierte Aufgaben wie zum Beispiel das in dieser Arbeit diskutierte Multi-People-Tracking (Schulz et al., 2003a) oder den 6D-SLAM (Nüchter et al., 2007) zu bewältigen. Diese Entwicklung war mit neuen Herausforderungen an Wissenschaftler und Ingenieure verbunden.

In dieser Arbeit wird vor allem das Problem der Sensorik näher untersucht. Eine Messung ist ein Prozess, bei dem der Roboter seine Sensoren nutzt, um Informationen über den Zustand der Welt zu erhalten. Dabei kann der Roboter zum Beispiel ein Kamerabild verwenden oder Entfernungen messen. Zur Entfernungsmessung dienen beispielsweise Laser-Scanner. Ein Laser-Scanner sendet einen Laserstrahl aus, der von den Objekten in der Umgebung reflektiert wird und wieder zum Laser zurückkommt. Dort wird er registriert und die Laufzeit bestimmt. Aus der Laufzeit kann trivialerweise die Entfernung des reflektierenden Objektes berechnet werden. Üblicherweise werden Laser verwendet, welche jeweils ein Sichtfeld von 180 Grad abdecken, so dass auf dem Roboter zwei dieser Laser Rücken an Rücken montiert werden, um die gesamte Umgebung zu erfassen. Die Laser können beispielsweise so eingestellt werden, dass sie die Welt mit einer Frequenz von fünf oder sechs Hertz und einer Winkelauflösung von einem Grad abtasten.

Mit Hilfe dieser Messdaten können verschiedene Probleme aus der mobilen Robotik behandelt werden. Es seien hier stellvertretend die folgenden genannt:

- Das Problem der Kartenbildung. Darunter versteht man die Gewinnung eines räumlichen Modells der Umgebung eines Roboters. Diese Karten werden dann zur Navigation oder zur Lokalisierung verwendet. Einen historischen Abriss und einen Überblick über aktuelle probabilistische Verfahren zu diesem Thema gibt (Thrun, 2003).
- Das Problem der Lokalisierung, das schnell als ein Schlüsselproblem in der mobilen Robotik erkannt wurde (Borenstein et al., 1996; Cox, 1991). Darunter versteht man, dass ein mobiler Roboter versucht, seine Position in einem globalen Koordinatensystem zu finden. Das Problem der Lokalisierung schließt zwei Teilprobleme ein: lokale Lokalisierung oder Positionsschätzung und globale Lokalisierung. Das erste Problem ist bei Weitem das meist Untersuchte. Dabei kennt der Roboter seine Anfangsposition und muss lediglich während seiner Fahrt kleine Fehler in der Odometrie ausgleichen, um seine Position in einem globalen Koordinatensystem zu bestimmen. Unter Odometrie versteht dabei man die Fähigkeit eines mobilen Roboters, seine eigene Bewegung aus der Messung der Bewegung seines Antriebes zu schätzen.

Im Falle der globalen Lokalisierung kennt der Roboter seine Anfangsposition nicht. Somit stellt die globale Lokalisierung eine viel schwierigere Aufgabe dar. Die Fähigkeit zur Lokalisierung spielte eine bedeutende Rolle in zahlreichen Anwendungen in der mobilen Robotik (Burgard et al., 1998; Endres et al., 1998; Kortenkamp et al., 1997).

Im Wesentlichen gibt es drei Methoden zur Lokalisierung, nämlich das sogenannte Dead Reckoning, d.h. die Messung von Richtung, Geschwindigkeit und Zeit mittels der Odometrie, INS (Inertial Navigation Systems) unter Verwendung von Gyrometern oder Beschleunigungsmessern und schließlich die Vermessung von Landmarken (Wu und Yang, 2007). Der Hauptnachteil der ersten beiden Verfahren besteht darin, dass sich Fehler schnell akkumulieren können. Bei der dritten Methode werden Landmarken in der Umgebung platziert oder natürliche Landmarken wie zum Beispiel Bäume benutzt. Der Roboter misst dann fortlaufend seine relative Position zu diesen Landmarken. Da hierbei ähnlich starkes a-priori-Wissen wie in Kapitel 5.6, nämlich über die Position der Landmarken in einer Karte verwendet wird, sollen im Weiteren einige Realisierungen dieses Ansatzes kurz vorgestellt werden.

Während sich die Mehrzahl der frühen Arbeiten auf die lokale Lokalisierung beschränkte, wurden seit 1995 einige sehr erfolgreiche Ansätze entwickelt, die beide Lokalisierungsprobleme lösen können. Diese Ansätze werden unter dem Begriff der Markov-Lokalisierung subsummiert (Burgard et al., 1996; Kaelbling et al., 1996; Nourbakhsh et al., 1995; Simmons und Koenig, 1995). Die zentrale Idee besteht dabei darin, das Wissen des Roboters mit Hilfe einer Wahrscheinlichkeitsverteilung über mögliche Positionen zu repräsentieren. Die Bayessche Regel wird dann angewandt, um das Wissen zu aktualisieren, sobald der Roboter misst oder sich bewegt. Die Idee der probabilistischen Zustandsschätzung geht zurück auf die Anwendung von Kalman-Filtern, die multivariate Gaußverteilungen verwenden, um das Wissen des Roboters zu repräsentieren (Gelb, 1974; Smith et al., 1990). Die Gaußverteilung weist aber gewisse Restriktionen auf, da sie nur eine Hypothese, versehen mit einer gewissen Unsicherheit, wiedergeben kann. Deshalb werden Kalman-Filter gewöhnlicherweise nicht auf das Problem der globalen Lokalisierung angewandt. Markov-Lokalisierung verwendet dagegen diskrete, aber multimodale Verteilungen zur Repräsentation des Wissens eines Roboters und kann somit auch das Problem der globalen Lokalisierung lösen. Unter einer multimodalen Verteilung versteht man dabei eine Verteilung mit mehreren lokalen Maxima, die hier den möglichen Positionen eines Roboters entsprechen. Dieser Ansatz kann das Wissen aber nur approximieren und benötigt sehr viel Rechenzeit und Speicherplatz.

In mehreren Veröffentlichungen wurde danach die sogenannte Monte-Carlo-Lokalisation eingeführt (Dellaert et al., 1999; Fox et al., 1999b). Monte-Carlo-Methoden, die letzten Endes nichts anderes sind als Teilchen-Filter, kamen in den siebziger Jahren des vergangenen Jahrhunderts auf (Handschin, 1970) und

wurden dann unabhängig voneinander in so verschiedenen Gebieten wie Target-Tracking (Gordon et al., 1993), Statistik (Kitagawa, 1996) und Computer-Vision (Isard und Blake, 1998) wiederentdeckt. Sie wurden darüber hinaus auch erfolgreich auf dem Gebiet der dynamischen probabilistischen Netzwerke angewandt (Kanazawa et al., 1995). Die Monte-Carlo-Lokalisierung benutzt schnelle Sampling-Techniken, um das Wissen eines Roboters zu repräsentieren. Wenn der Roboter sich bewegt oder mit seinen Sensoren eine Messung ausführt, kann beispielsweise die Methode des Importance-Resampling (Rubin, 1988) verwendet werden, um die a-posteriori-Verteilung zu ermitteln. Ein adaptives Sampling-Schema, das die Anzahl der Teilchen oder Samples fortlaufend festlegt, kann benutzt werden, um den Bedarf an Rechenkapazität zu reduzieren und die Genauigkeit zu verbessern (Koller und Fratkina, 1998). Damit benutzt Monte-Carlo-Lokalisierung viele Teilchen, wenn dies nötig ist, also während der globalen Lokalisierung, und wenige Teilchen während der anschließenden lokalen Lokalisierung, wenn die Position des Roboters ungefähr bekannt ist.

- Das SLAM-Problem. Dabei handelt es sich um eine der fundamentalsten Fragestellungen in der mobilen Robotik. SLAM steht für Simultaneous Localization and Mapping. Dies bedeutet, dass der Roboter weder eine Karte für seine Umgebung besitzt noch seinen eigenen Zustand, d.h. Ort und Orientierung, kennt. Vielmehr stehen ihm nur seine Messungen zur Verfügung. Aus diesen Messungen baut er sukzessive eine Karte der Welt auf und bestimmt gleichzeitig seine eigene Position innerhalb dieser Karte. Er muss also die a-posteriori-Verteilung sowohl seines Zustandes als auch der Karte bestimmen. Das SLAM-Problem geht der Erfindung der modernen Roboter um viele Jahrzehnte voraus. Das Problem der Exploration der physischen Struktur eines Objektes mit Hilfe einer sich bewegenden Plattform steht im Zentrum einiger Disziplinen wie der Geowissenschaften oder Computer Vision. Viele der mathematischen Techniken, die das Herz des SLAM bilden, waren ursprünglich zur Berechnung von Planetenbahnen entwickelt worden. Die Methode der kleinsten Quadrate beispielsweise kann bis zu Johann Carl Friedrich Gauß (1809) zurückverfolgt werden. SLAM ist letzten Endes ein geographisches Problem. Übertragen auf einen Roboter treten Herausforderungen auf, denen ein menschlicher Landvermesser nur selten gegenübersteht, wie zum Beispiel die Suche nach geeigneten Landmarken.
- Das Problem des Tracking. Unter Tracking versteht man die fortlaufende Positionsschätzung sich bewegender Objekte, d.h. die Verbesserung und Aktualisierung der Positionsschätzung aufgrund konsekutiver Messungen. Eine andere Definition bezeichnet Tracking als die Verwendung von Positionsdaten von Sensoren zur Festlegung der Position und der Geschwindigkeit eines sich bewegenden Objektes als eine Funktion der Zeit. Die Notwendigkeit des Tracking erklärt sich unter anderem daraus, dass es für viele Anwendungen sinnvoll ist, dass ein Roboter mit seiner Umgebung interagieren kann. Dies ist der Fall für

Multi-Roboter-Systeme, wo eine Gruppe von Robotern in einer bestimmten Formation fahren oder gemeinsam eine gegebene Aufgabe lösen soll, wie zum Beispiel beim Robot-Soccer (Schmitt et al., 2002), oder wenn Roboter Menschen unterstützen sollen. In solchen Situationen ist die Kenntnis der Position der menschlichen Partner oder der anderen Roboter von fundamentaler Bedeutung für die mobile Einheit, um mit ihrer Umgebung in sinnvoller Weise kooperieren zu können. Die Bedeutung des Tracking-Problems in der mobilen Robotik unterstreicht auch folgendes Zitat aus einem Review für die Konferenz 'Robotics: Science and Systems 2005': 'The problem of people tracking is very significant for robotics.' Auch auf aktuellen bedeutenden Konferenzen wie 'ICRA 2007' oder 'IROS 2007' werden immer wieder Arbeiten zum Thema Tracking veröffentlicht (Brscic und Hashimoto, 2007; Cho et al., 2007; Cielniak et al., 2007; Derenick et al., 2007; Furukawa et al., 2007; Huang et al., 2007; Ikeda et al., 2007; Liu et al., 2007; Ryu und Huber, 2007; Shao et al., 2007; Tajana et al., 2007; Wang et al., 2007; Zhang et al., 2007).

Tracking zerfällt in die Teilprobleme der Trackinitialisierung oder Objektdetektierung, der Datenassoziiierung, der Positionsschätzung oder dem Tracking im engeren Sinne und der Objektzuordnung (Klassifizierung oder Identifizierung). Unter Klassifizierung versteht man dabei die Einordnung eines Objektes in vorgegebene Klassen, wie zum Beispiel 'Kampffjet' oder 'ziviles Luftobjekt', während Identifizierung die Zuordnung einer genau definierten Identität zu einem Objekt, wie zum Beispiel einer Person beim People Tracking, bezeichnet. Im Falle des Tracking mit mobilen Robotern müssen im Rahmen der Assoziierung sowohl Einzelbeobachtungen zu Clustern zusammengefasst werden, die von einer Einzelperson oder einer Gruppe von Personen stammen, als auch diese Features oder Cluster von Beobachtungen den computerinternen Tracks zugeordnet werden.

Das Problem des Tracking ist gut untersucht im Falle der Luftüberwachung mit Radar einschließlich einer wahren Flut von Veröffentlichungen, zum Beispiel (Bar-Shalom und Fortmann, 1988; Koch, W., 2000; Sea, 1971; Singer und Stein, 1971). Es gehörte im Rahmen des SDI-Programms ('Strategic Defense Initiative') der USA, das Anfang der 80-er Jahre zur Abwehr eines Atomwaffenangriffs initiiert wurde und wo es darum ging, viele hunderte Marschflugkörper und andere Flugobjekte gleichzeitig zu orten und zu tracken, zu den schwer lösbaren Teilaufgaben. Auf dem Gebiet der mobilen Robotik ist Tracking ein gleichfalls gut eingeführtes, wenn auch neueres Forschungsgebiet (Fod et al., 2002; Fuerstenberg et al., 2002; Prassler et al., 1999; Schulz et al., 2001b). Aber auch auf dem Gebiet der Teilchenphysik spielt Tracking eine wichtige Rolle. Dort geht es darum, Kollisionen in Teilchenbeschleunigern zu analysieren (Hall und Llinas, 2001).

Die Bedeutung der Verfolgung oder der Kenntnis der Position von Personen liegt darin begründet, dass bei zahlreichen Anwendungen mobile Service-Roboter in Um-

gebungen operieren sollen, in denen sich auch Menschen aufhalten, und den Raum, der ihnen zur Navigation zur Verfügung steht, mit diesen Menschen teilen sollen (Belotto und Hu, 2007c). Eventuell sollen Roboter Menschen auch begleiten und bei der Ausführung von Handlungen unterstützen. Zum Beispiel könnten Roboter älteren Menschen im Haushalt zur Hilfe gehen. Man spricht in diesem Zusammenhang oft auch von 'Human-Centred Robotics', 'Service-Robotics' oder 'Socially Assistive/Interactive Robotics'. Dabei sollen autonome Roboter in Büros und Kranken- und Wohnhäusern Menschen unterstützen. So war das Motto der Konferenz ICRA 2005 auch 'Robots get closer to humans'. Die Kenntnis der Bahnen der Menschen in seiner Umgebung erlaubt es einem Roboter oder einem anderen autonomen Fahrzeug auch, seine Geschwindigkeit der Bewegung dieser Menschen anzupassen und auf diese Weise Zusammenstöße zu vermeiden, falls die Bahn des Roboters den Weg eines Menschen kreuzt (Rosales und Sclaroff, 1998; Tadokoro et al., 1995). Dadurch können eventuell zukünftige Arbeitsplätze sicherer gemacht werden. Beispiele für mögliche Anwendungen sind unter anderem:

- die Unterhaltung oder die Information von Menschen auf Ausstellungen und öffentlichen Plätzen (Burgard et al., 2002)
- das Melden von und die Überwachung von Eindringlingen in weite Gebiete oder Gebäude, die mit festen Sensoren schwierig zu überwachen sind (Treptow et al., 2005).
- die Verbesserung der Sicherheit und die Vermeidung von Unfällen mit Fußgängern im Straßenverkehr (Broggi, 2008; Ngako Pangop et al., 2008).

Die Bedeutung dieser Fragestellungen wurde auch von der Europäischen Kommission erkannt. Sie unterhält das 'Institute for the Protection and Security of the Citizen'. Dort werden unter anderem Methoden entwickelt, um öffentliche Plätze zu überwachen. Daneben gibt es beispielsweise auch noch die Forschungsinitiative 'Personal Robotic Assistants for the Elderly' der Universitäten von Pittsburgh und der Carnegie Mellon University. Der Themenbereich 'People Tracking with Robots and Intelligent Environments' im Unterprojekt 'Nursebot' wurde dabei maßgeblich von den Universitäten Bonn und Freiburg gestaltet.

Inhalt der vorliegenden Arbeit ist vor allem das Tracking oder die Verfolgung von ausgedehnten Zielen, und hier insbesondere von Personen (People Tracking), mit mobilen Robotern. Ein Resultat der Untersuchungen lässt sich dabei auch auf das Problem der Lokalisierung anwenden, weswegen dieser Aspekt der mobilen Robotik ebenfalls kurz gestreift wird.

Im Falle der Luftüberwachung mit Radar wird das Problem der Verfolgung von Flugzeugen durch folgende Eigenschaften charakterisiert:

1. die Messungen werden mit Hilfe von Radar generiert, d.h. eine Radaranlage, die auf dem Boden oder in einem Flugzeug stationiert ist, sendet über einen gewissen Winkelbereich Radarstrahlen aus, die reflektiert werden können. Die

reflektierten Radarstrahlen kommen wieder zur selben oder einer anderen Radaranlage zurück und werden dort registriert. Aus der Laufzeit eines Strahls kann dann auf die Entfernung des reflektierenden Objektes geschlossen werden.

2. die Objekte, über die Informationen erhalten werden sollen, sind zivile oder militärische Flugzeuge.
3. die Objekte sind punktförmig, d.h. sie generieren zum selben Zeitpunkt maximal eine Messung. Der Fall, dass ein Flugzeug gelegentlich gar keine Messung verursacht, ist also durchaus möglich.
4. es können Störmeldungen oder Falschmeldungen oder sogenannter Clutter auftreten. Darunter versteht man Messungen, die nicht von Flugzeugen herrühren. Verursacher solcher Meldungen können beispielsweise Vogelschwärme, Nebelbänke oder Wolken sein.
5. es können Winkelbereiche auftreten, unter denen die Radarstrahlen überhaupt nicht reflektiert werden, d.h. keine Radarstrahlen zum Empfänger gelangen.
6. halten sich mehrere Flugzeuge im sogenannten Überwachungsgebiet auf, können sich diese einander annähern und anschließend wieder voneinander entfernen. Man spricht in diesem Falle auch von Crossing Targets oder interagierenden Zielen.
7. beschränkt man sich auf den zivilen Bereich, so können die Objekte keine allzu abrupten Manöver ausführen.

Demgegenüber wird das Tracking mit mobilen Robotern in geschlossenen Räumen im Rahmen der vorliegenden Arbeit durch folgende Charakteristika definiert:

1. die Messungen werden mit Hilfe von Laserscannern generiert. Ein Laserscanner sendet dabei beispielsweise in Abständen von einem Winkelgrad Laserstrahlen aus, die von den Objekten in der Umgebung reflektiert werden. Die reflektierten Strahlen werden dann von dem Laserscanner registriert und aus der Laufzeit kann dann wieder die Entfernung der Objekte berechnet werden.
2. die Ziele, über die Information erhalten werden soll, sind andere mobile Roboter oder Menschen.
3. die Objekte sind ausgedehnt, d.h. sie können zum selben Zeitpunkt gar keine Messung, eine Messung oder mehrere Messungen generieren. Der Fall keiner Messung kann beispielsweise dann auftreten, wenn das Objekt sich soweit vom Beobachter entfernt hat, dass es sich zum Zeitpunkt der Messung komplett zwischen zwei benachbarten Laserstrahlen befindet, oder aber wenn es von einem Hindernis oder einem anderen beweglichen Objekt verdeckt wird. Mehrere Messungen erhält man oft deshalb, weil die Winkelauflösung der Laserscanner sehr

hoch ist im Vergleich zur Ausdehnung der Objekte. Die Laserscanner werden nämlich meistens mit einer Winkelauflösung von einem Grad betrieben und die Ziele haben in der Regel eine räumliche Ausdehnung von circa einem halben Meter. Außerdem ist die Entfernung der Objekte zum Beobachter vergleichsweise gering. Deshalb kann auch versucht werden, aus den Messungen auf die Form oder Ausdehnung der Objekte Rückschlüsse zu ziehen. Außerhalb der Robotik spielt die Verfolgung ausgedehnter Ziele kaum eine Rolle. Arbeiten auf diesem Gebiet beschäftigen sich dabei in erster Linie mit Gruppen von Flugzeugen, die nicht mehr in Einzelobjekte aufgelöst werden können und deshalb Ursprung mehrerer Messungen sind. Hier sind in erster Linie die Arbeiten von Koch zu nennen (Dezert, 1998; Koch, 2006; Koch und Saul, 2005; Koch und van Keuk, 1997; Salmond und Gordon, 1999).

4. der überwiegende Anteil der Messungen rührt nicht von beweglichen Objekten, sondern in erster Linie von den Wänden des Gebäudes oder anderen Hindernissen wie Papierkörben oder Möbeln, her. Falschmeldungen, d.h. Messungen von nicht existenten Objekten, treten praktisch nicht auf.
5. da die Untersuchungen sich auf geschlossene Räume in Gebäuden beschränken, wird jeder Laserstrahl auch reflektiert, so dass keine Winkelbereiche vorkommen können, zu denen es keine Entfernungsmessungen gibt.
6. das Problem der interagierenden Ziele kann auch auf dem Gebiet der mobilen Robotik auftreten und ist hier vielleicht noch wichtiger als in der Luftraumüberwachung. Es kommt nämlich häufig vor, dass Personen sich einander nähern, zum Beispiel, um sich miteinander zu unterhalten oder gemeinsam eine Handlung auszuführen. Das Problem der interagierenden Personen ist das zentrale Thema der vorliegenden Arbeit.
7. handelt es sich um Menschen, so können diese unter Umständen sehr abrupte Manöver ausführen und dabei ihre Bewegungsrichtung und Geschwindigkeit in einem sehr kurzen Zeitintervall rapide ändern. Ein Mensch kann zum Beispiel aus einer Vorwärtsbewegung heraus, ohne sich umzudrehen, von einem Augenblick auf den nächsten die Bewegungsrichtung ändern und rückwärts laufen.

Dennoch besteht aber eine wesentliche Gemeinsamkeit darin, dass in beiden Szenarien oft mehrere dicht beieinander liegende Einzelmessungen mit einem und demselben Objekt assoziiert werden müssen. In der Luftüberwachung ist dabei allerdings höchstens eine Messung tatsächlich vom Ziel, während alle anderen Messungen Falschmeldungen sind. Bei ausgedehnten Zielen in der Robotik sind dagegen in der Regel alle Messungen im sogenannten Erwartungsgebiet vom Ziel selbst. Ausnahmen bestehen lediglich dann, wenn sich ein anderes Ziel nähert oder eine Wand oder ein Hindernis in der Nähe ist. Die einzelnen Messungen unterscheiden sich lediglich darin, dass ihr Ursprung unterschiedliche Punkte auf der Oberfläche des Zieles sind. Deshalb

können viele Methoden, die für das Tracking im Falle punktförmiger Ziele in Clutter entwickelt wurden, im Prinzip auf den Fall der ausgedehnten Ziele in der mobilen Robotik übertragen werden, falls man jeweils geeignete Modifikationen vornimmt.

Im Folgenden werden die Begriffe Tracking und Verfolgung synonym verwendet. Die von einem Tracking-Algorithmus berechneten Zustandsschätzungen werden in ihrer zeitlichen Gesamtheit als Track, Spur oder Bahn bezeichnet.

In dieser Arbeit werden vor allem die folgenden drei unterschiedlichen Methodiken angewandt:

1. das Kalkül der Bayes-Statistik. Grundlegendes Konzept der Bayes-Statistik ist der Begriff der bedingten Wahrscheinlichkeit, d.h. die Frage, was sich über die Wahrscheinlichkeit eines bestimmten Ereignisses aussagen lässt, wenn ein anderes Ereignis, das mit dem ersten in irgendeiner Form verknüpft ist, schon eingetreten ist. Im konkreten Fall heißt dies: was lässt sich über den Zustand eines Systems aussagen, wenn man gewisse Messdaten zur Verfügung hat, die man durch die Interaktion geeigneter Sensoren mit den zu charakterisierenden Objekten gewonnen hat. Bei der Bayes-Statistik handelt es sich um eine mathematische Theorie, die der Wahrscheinlichkeitsrechnung zuzuordnen ist und die auch Anleihen aus den Gebieten der multidimensionalen Analysis und der Linearen Algebra macht. Die in dieser Arbeit verwendeten Hilfsmittel aus der Bayes-Statistik betreffen unter anderem den Kalman-Filter, den Kalman-Glätter und den Viterbi-Algorithmus.
2. das Gebiet der ebenen Euklidischen und der algorithmischen Geometrie (Klein, 2005). Die Laserscanner tasten nämlich das Überwachungsgebiet in einer Ebene ab, die sich circa einen halben Meter über dem Fußboden befindet. Die Objekte, zum Beispiel die Beine der Personen, von denen die Messungen deshalb herrühren, besitzen nun gewisse geometrische Eigenschaften, die sich in den Messdaten in charakteristischer Weise widerspiegeln. Die Anzahl und die Anordnung dieser Messungen hängen nämlich sowohl von der Form und der Ausdehnung der Objekte selbst als auch von der Entfernung dieser Objekte vom Laserscanner ab. Aus der Struktur der Messungen kann somit wertvolle Information gewonnen werden. Dieser Umstand hilft bei der Entwicklung geeigneter Algorithmen zur Verfolgung in der mobilen Robotik sehr viel weiter. Beispielsweise sei hier die genaue Positionsschätzung eines kreisförmigen Objektes aus Kapitel 5 genannt. In diesem Fall besitzt das Objekt eine wohldefinierte Form, die a priori bekannt ist, so dass die Messungen in charakteristischer Weise angeordnet sind. Zur Analyse dieses Sachverhaltes werden dort neben der ebenen Geometrie auch Hilfsmittel aus der Analysis benutzt. Außerdem soll die Detektion oder der Spuraufgriff unbekannter Ziele in Kapitel 10.1 erwähnt werden. Daneben werden aber auch bei der Verfolgung interagierender Personen exzessiv geometrische Eigenschaften der Objekte, wie zum Beispiel die beschränkte Ausdehnung eines Menschen, benutzt. Auf diese Weise kann zum Beispiel festgestellt werden, ob sich eine Gruppe von Personen schon wieder in ihre Einzelob-

jekte aufgelöst hat. In letzterem Falle macht eine genaue Analyse der geometrischen Gegebenheiten des Problems dessen Lösung sogar erst möglich. Obwohl in den letzten Jahren in der mobilen Robotik sehr viel Forschung im Bereich des Tracking stattgefunden hat, wurden die geometrischen Aspekte dieser Problematik bisher kaum beachtet. Vielmehr konzentrierten sich die meisten Forscher darauf, Positionsschätzungen mit einer möglichst hohen Genauigkeit zu berechnen, wobei sie ausgeklügelte Bayessche Methoden wie Teilchen-Filter oder den EM (Expectation Maximisation) Algorithmus (Khan et al., 2003; Montemerlo et al., 2002b) verwendeten, und ließen somit wertvolle Information außer Acht.

3. der Bereich des maschinellen Lernens (Michalski et al., 1999; Mitchell, 1997). Maschinelles Lernen ist ein Oberbegriff für die automatische Generierung von Wissen aus Erfahrung. Ein künstliches System lernt aus Beispielen und kann nach Beendigung der Lernphase verallgemeinern, d.h. es lernt nicht einfach die Beispiele auswendig, sondern es erkennt Gesetzmäßigkeiten in den Lerndaten. So kann das System dann auch unbekannte Daten beurteilen. Maschinelles Lernen wurde in der vorliegenden Arbeit in folgenden beiden Bereichen angewandt:
 - (a) das Lernen von Hosenmaterialien zur Erkennung und Identifizierung von Personen. Im Rahmen der Untersuchungen zur Personenerkennung wurden zehn unterschiedliche Materialien betrachtet und jeweils das Reflexionsverhalten in Abhängigkeit von der Entfernung für jeden der beiden Laserscanner des Roboters vermessen. Der Roboter war danach in der Lage, auch Menschen zu unterscheiden, die Hosen trugen, die den Gelernten vom Material her zwar ähnlich waren, aber nicht identisch mit den Hosen, von denen die Lerndaten stammten. So war die Jeanshose, mit der die Echtzeitversuche in Kapitel 11 durchgeführt wurden, nicht identisch mit der Jeanshose, anhand derer die Reflexionseigenschaften einer Jeans gelernt worden waren. In weiteren Experimenten wurde eine schwarze Lederhose verwendet, die ebenfalls richtig zugeordnet wurde, obwohl in der Lernphase eine schwarze Lederjacke benutzt worden war.
 - (b) das Problem des Spuraufgriffs. Dabei geht es um die Aufgabe, diejenigen Personen zu entdecken, die sich neu im Überwachungsbereich aufhalten. Falls der Roboter sich nicht bewegt, können zu diesem Zweck einfach die aktuellen Laser-Scans mit einem gespeicherten Referenzscan verglichen werden. Betritt eine Person den Raum, so zeigen sich dann typische Unterschiede zu dem Referenzscan, die zum Zweck des Spuraufgriffs ausgenutzt werden können. Diesen Ansatz kann man unter dem Oberbegriff des Novelty Detection subsummieren. Unter Novelty Detection versteht man allgemein die Identifikation von neuen oder unbekanntem Daten durch ein intelligentes System, das maschinelles Lernen benutzt. Diese neuen Daten unterscheiden sich dabei signifikant von den Daten, die das System während einer Trainingsphase gelernt hat. Konkret geht es um die Erken-

nung von Sensordaten, die von dem gelernten Wissen über die Umgebung abweichen. Dies können zum Beispiel Lasermessungen von einem Eindringling in ein Gebäude sein, das von einem Robotersystem überwacht wird. Novelty Detection wird in vielen Bereichen der Wissenschaft angewandt, wie zum Beispiel bei der Analyse von Mammographie-Daten (Bishop, 1994; Tarassenko et al., 1995; Taylor und MacIntyre, 1998; Ypma und Duin, 1997). Auch im Bereich der mobilen Robotik ist Novelty Detection ein wohlbekannter Ansatz (Marsland et al., 2005; Neto und Nehmzow, 2007).

Bewegt sich dagegen der Roboter, so ändert sich der Standard ständig, insbesondere wenn der Roboter in noch unbekanntes Terrain vordringt. In diesem Fall können unbekannte Personen nicht über ein Verfahren wie Novelty Detection aufgespürt werden. Allerdings können die Menschen dann dadurch identifiziert werden, dass sie gewisse eindeutig definierte Bedingungen erfüllen. Das Wissen über diese Bedingungen oder das Konzept wird dabei in einem sogenannten symbolischen System gespeichert. Von einem symbolischen System spricht man, wenn das Wissen explizit repräsentiert ist. Genauer wird eine Taxonomie verwendet. Von einer Taxonomie redet man dann, wenn Gegenstände aufgrund von Konzepten oder Hypothesen in Kategorien oder Gruppen eingeteilt werden (Bratko, 2001; Witten und Eibe, 2005).

Da im Falle des gleichzeitigen Trackings und Identifizierens von Personen mit Lasern und Kameras, wie es in Kapitel 9.2 vorgestellt wird, Daten von mehreren Sensoren verwendet werden, spielen am Rande auch noch Aspekte der Bildverarbeitung und der Sensor-Daten-Fusion oder Information-Fusion eine Rolle (Hall, 1992; Hall und Llinas, 2001). Methoden der Sensor-Fusion werden in vielen Bereichen angewandt, wie zum Beispiel in der Medizin. Dort forscht man unter anderem daran, zur besseren Körperdarstellung und Diagnosefindung Daten aus Magnetresonanztomographen, Computertomographen oder Positronen-Emissions-Tomographen zu fusionieren. Das beste Beispiel für Sensor-Fusion sind aber Tiere oder der Mensch. Im Allgemeinen versteht man unter Multisensor-Daten-Fusion die Kombination von Daten von mehreren, möglicherweise unterschiedlichen oder komplementären Sensoren. Die Begründung für ein solches Vorgehen liegt darin, dass ein Sensor oft nicht ausreicht, um alle benötigten Informationen bereitzustellen, zumal ja jeder Sensor auch mit einer gewissen Unsicherheit behaftet ist. Sensor-Fusion ist auch in der Robotik ein wohlbekanntes Werkzeug (Andreasson et al., 2005; Garcia et al., 2008; Liu und Stamos, 2005).

Die zentrale Fragestellung dieser Arbeit ist wie schon angedeutet die Verfolgung interagierender Personen. Darunter versteht man das Phänomen, dass sich Versuchspersonen für eine Zeitlang sehr nahe kommen und danach wieder voneinander entfernen. Dabei treten vor allen Dingen drei Probleme auf:

- gewöhnliche Algorithmen verlieren am Ende einer Interaktion Spuren, d.h. allen

beteiligten Personen wird nach der Interaktion fälschlicherweise dieselbe Bahn zugeordnet. Im Rahmen der vorliegenden Arbeit wurde deshalb ein Algorithmus entwickelt, der jeder Person einen anderen, realen Track zuordnet (Kapitel 7). Zuvor wird das Problem allerdings erst für den Spezialfall zweier Objekte gelöst (Kapitel 6).

- diese Tracks können nach der Interaktion vertauscht sein, d.h. die Zuordnung der Personen zu den computerinternen Tracks ist in der Regel eine andere als vor der Interaktion. Die vorliegende Arbeit stellt deshalb in Kapitel 9 zwei Ansätze vor, mit deren Hilfe dieser Makel beseitigt werden und Personen identifiziert werden können. Dies ist zum einen die Verwendung von Farbinformation, die mit Hilfe eines Kamerasystems gewonnen wird, und zum anderen die Unterscheidung anhand des Reflexionsverhaltens der Hosen bezüglich der Laserstrahlen. Dieses Reflexionsverhalten ist dabei für die verschiedenen Personen und Hosenmaterialien innerhalb gewisser Grenzen charakteristisch.
- die Verfahren müssen onlinefähig sein, d.h. sie müssen so schnell sein, dass sie auch mit den begrenzten Rechenkapazitäten einer Onboard-Einheit in Echtzeit ausgeführt werden können. Auf diese Weise sind Versuche möglich, wie sie in Kapitel 11 vorgestellt werden. Da die meiste Rechenzeit für das Gating, d.h. die Auswahl der Messungen, die den Personen zugeordnet werden sollen, verbraucht wird, ist ein effektives Gating eine essentielle Voraussetzung für ein Echtzeitverfahren. In Kapitel 10.1 werden deshalb Ansätze, zum Beispiel in Form einer Vorfilterung der Messdaten, vorgestellt, die zur Optimierung dieses Gatings beitragen können.

Die Verfolgung mehrerer interagierender Personen ist also ein sehr schwieriges Problem. Demgegenüber ist die Verfolgung von Einzelzielen in der Robotik sehr einfach, da es aufgrund der Abwesenheit von Störmeldungen sehr leicht ist, die Messungen, die vom jeweiligen Ziel stammen, auszuwählen.

Jede Methode muss im Kontext ihrer spezifischen Anwendung betrachtet werden. Die im Rahmen dieser Arbeit vorgestellten Verfahren sind speziell an die Verfolgung mehrerer ausgedehnter Objekte, insbesondere von interagierenden Menschen, mit mobilen Robotern und Laserscannern adaptiert. Dabei wird gezeigt, dass die drei eben skizzierten Probleme von einem Verfahren gleichzeitig gelöst werden können.

Die weitere Arbeit ist wie folgt aufgebaut: in Kapitel 2 werden bisherige Arbeiten anderer Autoren vorgestellt. In Kapitel 3 werden kurz die mathematischen Grundlagen dargestellt. In Kapitel 4 werden die verwendeten Modelle der Objekte und der Sensoren und einige einfache Algorithmen zur Verfolgung ausgedehnter Ziele skizziert. In Kapitel 5 werden das Problem der Verfolgung eines kreisförmigen Objektes diskutiert und anhand dieser Fragestellung die Verfahren aus dem vorhergehenden Kapitel hinsichtlich Genauigkeit und Komplexität miteinander verglichen. Außerdem wird eine mögliche Anwendung auf das Problem der Lokalisierung eines mobilen Roboters vorgeschlagen. In Kapitel 6 wird erstmals das zentrale Problem interagierender Personen

analysiert, allerdings noch beschränkt auf den Spezialfall zweier Probanden. Zunächst werden dabei die grundsätzlichen Schwierigkeiten diskutiert und danach drei Algorithmen vorgeschlagen, die diese umgehen. In Kapitel 7 wird als Verallgemeinerung ein Verfahren vorgestellt, das das Problem beliebig vieler interagierender Personen zufriedenstellend behandelt. In Kapitel 8 wird dieser Ansatz mit dem aus der Literatur bekannten SJPDFAF hinsichtlich Genauigkeit, Komplexität und der Robustheit im Falle zweier interagierender Personen verglichen. In Kapitel 9 werden zwei Ansätze zur Identifikation von Personen, zum einen basierend auf mit einem Kamerasystem gewonnener Farbinformation und zum anderen fußend auf den mit den Laserscannern gemessenen Reflexionswerten, vorgestellt. In Kapitel 10 werden kurz die Probleme der Spuraufnahme und der Präselektion der Messdaten bzw. der Datenfusion bei Verwendung mehrerer Roboter besprochen. In Kapitel 11 werden schließlich Experimente in Echtzeit wiedergegeben, die es einem mobilen Roboter erlauben, eine vorher festgelegte Person zu begleiten. Der Roboter ist dabei in der Lage, sowohl Gruppen von Personen ohne Trackverlust zu verfolgen als auch einzelne Individuen zu identifizieren. Kapitel 12 enthält schließlich die Zusammenfassung und die Schlussfolgerungen. Im Anhang werden dann noch einige mathematische Betrachtungen zusammengefasst, und zwar über die Konvergenzeigenschaften des verwendeten Modells, die Verfolgung eines kreisförmigen Objektes und über das Simulationsprogramm. Außerdem werden die Ergebnisse einiger Versuche zur Identifikation von Personen wiedergegeben, auf deren Darstellung im Hauptteil verzichtet wurde.

Kapitel 2

Der Stand der Forschung

Die Verfolgung von Menschen und anderen beweglichen Objekten mit mobilen Robotern unter Verwendung von Laser-Scannern oder anderen Sensoren wie Ultraschall oder Sonar wurde in den vergangenen ca. zehn Jahren von zahlreichen Autoren untersucht und ist ein wichtiges Problem in der mobilen Robotik. Die Fähigkeit zur Verfolgung erlaubt es einem Roboter beispielsweise, seine Geschwindigkeit der Bewegung der Menschen in seiner Umgebung anzupassen und seine Fähigkeiten zur Kollisionsvermeidung zu verbessern, wenn er die Bahnen dieser Menschen kreuzt (Rosales und Sclaroff, 1998). Besonders schwierig ist hierbei das Verfolgen von mehreren Personen gleichzeitig. Dieses Problemfeld ist Kernpunkt der vorliegenden Arbeit und wird auch als Multi-People-Tracking bezeichnet. Dabei lassen sich bezüglich der Forschung auf diesem Gebiet drei Phasen unterscheiden. Zunächst wurden in einer Art Vorphase zahlreiche einfache Verfahren vorgestellt, die noch nicht speziell auf diese Problematik abzielen, teilweise aber schon mehrere Personen gleichzeitig zu verfolgen versuchen. In einer zweiten Phase wurde die Fragestellung der gleichzeitigen Verfolgung mehrerer Personen dann von einigen wenigen Autoren gezielt untersucht. Schließlich wurde von diesen Autoren auch noch das Problem sowohl der Verfolgung als auch der Identifizierung mehrerer Personen (People Tracking and Identification) eingeführt.

2.1 Einfache Verfahren

Es gibt zahlreiche Arbeiten in der mobilen Robotik, die in der ersten Phase veröffentlicht wurden und einfache Probleme behandeln. Erste Anwendungen von mobilen Robotern als sogenannte Service-Roboter in von Menschen bevölkerten Umgebungen betrafen Bürogebäude (Arras und Vestli, 1998; Asoh et al., 1997; Horswill, 1993; Simmons et al., 1997), Supermärkte (Endres et al., 1998), Krankenhäuser (Engelberger, 1993), die Führung von Personen in Museen (Burgard et al., 2000; Thrun et al., 1999) oder die Unterhaltung von Menschen auf öffentlichen Plätzen (Burgard et al., 2002). Davor bestand die Anwendung häufig darin, es humanoiden Robotern zu ermöglichen, auf Menschen in geeigneter Weise zu reagieren. Frühe Anwendungen von People-

Tracking im Bereich der mobilen Robotik waren zum Beispiel auch die Verbesserung des Kartenbaus (Hähnel et al., 2002) oder von SLAM-Methoden (Wang et al., 2003) durch Berücksichtigung sich bewegender Objekte. Außer auf dem Gebiet der mobilen Robotik wurde People-Tracking zuvor bereits im Bereich Computer-Vision oder Bildverarbeitung ausführlich untersucht (Rosales und Sclaroff, 1998). Es handelte sich dabei im Wesentlichen um die Extraktion von Bewegungen von Menschen aus Bildern. Dabei wurden verschiedene Ansätze vorgeschlagen. In (Darrell et al., 1996) wurde ein fixer Hintergrund abgezogen, um Information über die Kontur einer Person zu gewinnen. In (Wren et al., 1997) werden Farbinformationen verwendet und in (Rasmussen und Hager, 1998) Körperteile unterschieden. Einen umfassenden Überblick über Arbeiten auf diesem Gebiet gibt auch (Aggarwal und Cai, 1999). Im Bereich des People-Tracking in der mobilen Robotik wurde Bildverarbeitung in erster Linie zur Identifizierung von Personen verwendet wie in (Beymer und Konolige, 2001; Kahn et al., 1996; Kortenkamp et al., 1996; Waldherr et al., 1998).

Die Verwendung von genauen Laser-Abstandsmessungen in Verbindung mit verbesserten Algorithmen zur Schätzung der Bahnen hielt um die Jahrtausendwende verstärkt Einzug in der mobilen Robotik (Bennewitz et al., 2002a; Fod et al., 2002; Schulz et al., 2003a). Die Benutzung von Lasern unterscheidet sich dabei grundlegend von der Verwendung von Kameras im Bereich des sogenannten Visual-Tracking und weist einige wesentliche Vorteile auf. Im Bereich Vision stehen nämlich Informationen über die Farbe, die Intensität und die Tiefe zur Verfügung. Dagegen liefern Laser die Information in Form nur einer einzigen Messgröße, nämlich der Entfernung zum jeweils nächsten Hindernis für den jeweiligen Winkel. Laserscanner projizieren die Umgebung also gewissermaßen auf eine Ebene. Diese Entfernungsmessungen weisen aber eine sehr hohe Genauigkeit auf, insbesondere im Vergleich mit anderen Entfernungssensoren wie Ultraschall oder Infrarotwellen. Deshalb haben Laser in der mobilen Robotik in sehr kurzer Zeit sehr große Popularität erlangt und sind heute die bevorzugten Sensoren, zum Beispiel zur Hindernisvermeidung, zur Navigation, zur Lokalisierung, zur Kartenerzeugung, zur Exploration und zum People-Tracking (Burgard et al., 2005; Fox et al., 1999b; Thrun, 1998b). Erste Ansätze in diesem Bereich verwendeten dabei Besetzungsgitter und lineare Extrapolationen von Besetzungskarten, um Trajektorien zu schätzen (Prassler et al., 1999). Auch in (Kluge et al., 2001) wird ein noch sehr rudimentärer Ansatz verwendet. Dort wird ein autonomer Roboter in die Lage versetzt, bewegliche Hindernisse zu detektieren, um diesen auszuweichen. Es wird allerdings kein Bewegungsmodell der Objekte zugrundegelegt und auch die Bahnen werden nicht kontinuierlich geschätzt, so dass gar kein eigentliches Tracking stattfindet. Erste Ansätze zur Verfolgung mehrerer Personen mit einem mobilen Roboter gab es schon in (Lindström und Eklundh, 2001). In (Fod et al., 2002) werden für das gleiche Ziel statische Laser in einem Arbeitsraum verwendet. Die Assoziation der sogenannten Laser-Features zu den Tracks geschieht dabei mit einem Nearest-Neighbour-Ansatz. Die Anwendung bestand bei diesen frühen Arbeiten oft darin, die Roboter so agieren zu lassen, dass die Objekte nicht außer Sicht geraten oder die Beobachtungsmöglichkeiten zu optimieren wie in (Gonzalez-Banos et al., 2002; Lavelle et al., 1997), wo

speziell die Anwesenheit von Hindernissen betrachtet wird, oder in (Jung und Sukhatme, 2004; Murrieta-Cid et al., 2002; Parker, 1997; Pirjanian und Mataric, 2000). In diesen einfachsten Fällen waren in der Regel auch keine Tracks einzelner Personen notwendig. So ging es zum Beispiel lediglich um eine Geschwindigkeitsanpassung des Roboters an die Personen in seiner Umgebung (Tadokoro et al., 1995) oder die Unterscheidung zwischen statischen und dynamischen Objekten (Hähnel et al., 2002). Später wurde die gewonnene Information auch dazu genutzt, um tatsächlich die Position der Ziele zu schätzen, zum Beispiel, um Bewegungen vorherzusagen oder die Bewegungsmuster von Menschen zu lernen (Bennewitz et al., 2002a; Bui et al., 2001; Illmann et al., 2002; Zhu, 1991). In (Tadokoro et al., 1995) wurde ein probabilistisches Modell typischer Bewegungen verwendet, um die zukünftigen Positionen von Menschen vorherzusagen. In (Chakravarty und Jarvis, 2006) wurden Laser und Vision mit Hilfe eines Teilchen-Filters kombiniert, um mehrere Personen zu verfolgen. In (Kruse und Wahl, 1998) wurde schließlich eine an der Decke befestigte Kamera verwendet, um die Gehwege von Personen zu lernen.

2.2 Die Verfolgung mehrerer Personen

Als Vorläufer-Arbeiten zur Verfolgung mehrerer Personen sollen angeführt werden:

- (Beymer und Konolige, 1999). Dort werden mehrere Personen mit einer fixen Plattform mit Hilfe von Vision verfolgt. Es wird dabei für jede Person ein eigener Kalman-Filter implementiert.
- (Bennewitz et al., 2003). Von diesen Autoren wird dasselbe Setting untersucht, allerdings wird als Sensor ein zweidimensionaler Laser verwendet.
- (Montemerlo et al., 2002b). In dieser Arbeit werden Laser und Teilchen-Filter benutzt, um mit Hilfe einer Karte der Umgebung gleichzeitig den Roboter zu lokalisieren und Personen zu verfolgen. Allerdings verwenden die Autoren für die Bewegung der Menschen ein Brownsches Modell, was von Nachteil ist, wenn Interaktionen auftreten.
- (Khan et al., 2003). Hier werden zwanzig Ameisen, die sich teilweise sehr dicht zueinander bewegen, betrachtet. Um die Robustheit zu verbessern, wird die Tatsache berücksichtigt, dass sich zwei dieser Ameisen nicht gleichzeitig am selben Ort aufhalten können. Die geschätzten Trajektorien sind dabei in ca. achtzig Prozent der Fälle korrekt.

Erste grundlegende und bedeutende Arbeiten zur Verfolgung mehrerer Personen mit einer mobilen Plattform wurden von Schulz und Kollegen von der Universität Bonn vorgestellt (Schulz et al., 2001a; Schulz et al., 2001b; Schulz et al., 2003a). Ein Charakteristikum dieses sogenannten Multi-People-Tracking besteht darin, dass sowohl die

Anzahl der Personen als auch die Assoziation (Zuordnung) der Messungen oder sogenannten Features zu den computerinternen Tracks bestimmt werden muss. Im Wesentlichen müssen hierzu drei Schritte hintereinander ausgeführt werden, nämlich zunächst die Aufbereitung der Messungen, zum Beispiel zur Gewinnung der Features aus den Laserabstandsmessungen, die Korrelation oder Assoziation dieser Features zu den computerinternen Objekten und schließlich die Aktualisierung der Zustandsschätzungen dieser Objekte. Die Assoziation setzt sich dabei aus der Hypothesengeneration, der Evaluation der Hypothesen mittels einer Metrik und abschließend der Auswahl geeigneter Hypothesen zusammen. Die Anzahl der Personen wird über einen Bayeschen Filter geschätzt, wobei das Auftauchen und Verschwinden von Probanden mit Hilfe eines Poisson-Prozesses modelliert wird. Die Features werden extrahiert, indem mit Hilfe von Entfernungshistogrammen lokale Minima aus den Laserscans gewonnen werden. Solche lokalen Minima rühren oft von Menschen her, die sich vor einer Wand aufhalten. Die Assoziation der Features zu den Tracks geschieht mit Hilfe des JPDAF (Cox, 1993; Fortmann et al., 1983) in Form einer Bayesschen Schätzung. Eine wesentliche Erweiterung in dem Ansatz von Schulz besteht darin, dass zur Schätzung der Zustände der Objekte Teilchen-Filter (Gordon et al., 1993; Pitt und Shephard, 1999) verwendet werden und nicht wie üblich der Kalman-Filter. Der Hauptvorteil dieser Vorgehensweise ist, dass damit beliebige und insbesondere multimodale Wahrscheinlichkeitsverteilungen, also Verteilungen mit mehreren lokalen Maxima der Teilchendichte, modelliert werden können. Das Verfahren von Schulz wird deshalb auch als SJPDAF bezeichnet (S steht dabei für Sample Based). Bei Schulz wird aufgrund der Komplexität darauf verzichtet, den kombinierten Zustand aller Personen zu schätzen wie in (Isard und Blake, 1996; Koller-Meier und Ade, 2001; MacCormick und Blake, 1999; Meier und Ade, 1999). Vielmehr sind die Tracks unabhängig voneinander angelegt wie in (Ryu und Huber, 2007), d.h. jeweils mit einem eigenen Teilchen-Filter. Experimente zeigten, dass das Verfahren in der Lage ist, die Positionen der einzelnen Personen mit einer Genauigkeit im Bereich von ca. zwanzig Zentimetern zu schätzen. Als Ausblick wird angegeben, dass die Verfolgung von Gruppen, insbesondere im Hinblick auf die Assoziation der Features zu den Objekten, günstiger sein könnte, als das Verfolgen von Einzelpersonen, falls sich diese sehr nahe kommen. Gerade diese Idee ist aber die Grundlage der in dieser Dissertation vorgestellten Ansätze zur Verfolgung interagierender Personen. In (Romera et al., 2004) wurde versucht, das Verfahren von Schulz zu modifizieren, was die Bedeutung dieses Verfahrens unterstreicht. Anstatt für jede Person einen unabhängigen Teilchen-Filter zu verwenden, wurden dort doch alle Personen gleichzeitig mit einem multimodalen Teilchen-Filter verfolgt.

Neuere Arbeiten stammen von Bellotto und Hu (Bellotto und Hu, 2006), einer Gruppe aus Essex, UK. Diese Autoren verwenden sowohl Laserscanner als auch Kameras. Die Messungen des Lasers werden für den Tracking-Prozess selbst verwendet, während die von der Kamera gelieferte Information auch zur Identifikation der Personen benutzt wird. Hierzu werden die Gesichter detektiert und daraus auf die Körpergröße geschlossen. Wie bei Schulz wird auch hier nicht speziell auf den Fall interagierender Personen abgezielt. Da diese Arbeiten von Anfang an auch die Identifikation

der Personen mit einschlossen, werden sie erst im nächsten Abschnitt im Detail besprochen.

2.3 Die Verfolgung und zusätzliche Identifikation mehrerer Personen

Das Gebiet des 'People Tracking and Identification' wurde ebenfalls in erster Linie von Schulz und Bellotto bearbeitet. Langfristiges Ziel dieser Arbeiten ist es, eine längere Interaktion mit einer einzelnen Person zu ermöglichen, zum Beispiel um für diese schwere Lasten zu tragen oder diese zu führen. Ein Vorläufer war (Bennewitz et al., 2003; Bennewitz et al., 2005), wo die Personen mit Lasern lokalisiert wurden und Farbhistogramme zur Unterscheidung herangezogen wurden. Ziel war das Lernen ihres Bewegungsverhaltens. Ein ähnlicher Ansatz wurde in (Cielniak und Duckett, 2003) verfolgt. Dort wurde aber nur der Fall einer Person untersucht. Erwähnt werden sollen auch noch die Arbeiten in (Zajdel et al., 2005), wo mit Hilfe von Vision mehrere Personen verfolgt und identifiziert wurden. Eine starke Einschränkung dieses Ansatzes bedeutete aber das begrenzte Sichtfeld der Kamera. Im Rahmen der im Folgenden vorgestellten Arbeiten wurde den Tracks im Gegensatz zu den Ansätzen in Kapitel 9 auch dann eine Identität zugeordnet, wenn die Objektdichte sehr hoch war.

Die erste Arbeit von Schulz zu diesem Gebiet stammt aus dem Jahre 2003 (Schulz et al., 2003b). Dort wird ein Netzwerk von stationären Sensoren verwendet, um die Positionen und Identitäten von Personen zu schätzen. Diese Sensoren sind dabei fest in einem Gebäude montiert. Man spricht in einem solchen Fall auch von 'smart environments' oder intelligenten Häusern. Zur Positionsschätzung von Personen in einem solchen Szenario werden beispielsweise Kameras, Laser, Infrarotlicht, Radar oder Ultraschall verwendet (Hightower und Borriello, 2001). Kennzeichen des von Schulz vorgestellten Ansatzes ist die simultane Nutzung sowohl von Sensoren lediglich zur genauen Entfernungsmessung und Positionsschätzung in Form von Laser-Scannern und von Identitäts- oder ID-Sensoren, die zwar eine explizite Bestimmung der Identitäten von Personen erlauben, aber nur eine sehr ungenaue Ortsangabe liefern (Priyantha et al., 2000; Want et al., 1992).

Zur Bestimmung der Identitäten wurden dabei sowohl Infrarot- als auch Ultraschallsysteme benutzt, d.h. die Personen trugen entsprechende Sender, die jeweils Signale einer für die Person charakteristischen Frequenz aussandten. Die Sensoren schlugen an, sobald sich eine Person in einer Reichweite von circa vier Metern aufhielt. Zur Verarbeitung der Daten wurde ein sogenannter Rao-Blackwellised Teilchen-Filter verwendet. Es handelt sich dabei um eine Kombination aus einem Teilchen-Filter zur Verwaltung der ID-Daten und einem Kalman-Filter zur Berechnung der Positionsschätzungen. Jedes Teilchen repräsentiert dabei eine Assoziationsgeschichte zwischen den Tracks und den gemessenen Features und zusätzlich die ID-Zuordnungen. Im Rahmen des 'People Tracking and Identification' zerfällt das Assoziierungspro-

blem dabei in zwei Unterprobleme, nämlich wie bisher die Assoziierung der computerinternen Tracks zu den gemessenen Features und zusätzlich die Zuordnung der Identitäten zu den Tracks. In einem Experiment mit sechs Personen und zahlreichen Interaktionen waren von insgesamt 30 Durchläufen 29 erfolgreich, d.h. die Identitäten wurden korrekt zugeordnet. Dabei war der Fall der korrekten Zuordnung aber sehr großzügig definiert. Diese Zuordnungen wurden nämlich, wie schon angedeutet, mit Teilchen repräsentiert. Falls nur eines von 1000 Teilchen die richtige Zuordnung repräsentierte, galt der Durchlauf schon als erfolgreich. Bezüglich der Interaktionen wird auch in dieser Arbeit vorgeschlagen, Personen in Gruppen zusammenzufassen und für Personen innerhalb einer Gruppe keine Identitäten zu bestimmen, ähnlich wie in (Rosencrantz et al., 2003). Dies ist aber wie schon erwähnt gerade der Ansatz, wie er in der vorliegenden Dissertation verfolgt wird.

Ebenfalls von Schulz wurde 2006 ein Ansatz vorgestellt, der Laser-Entfernungsmessungen und Kamerabilder verknüpft (Schulz, 2006a). Dabei werden probabilistische Beispielmuster benutzt, die typische Muster von Personen repräsentieren und aus Laserabstands- und Kameradaten gelernt werden. Genauer werden Konturen in Bildern und Features aus Laserpunkten verwendet. Verarbeitet werden die Daten wieder mit einem Rao-Blackwellised Teilchen-Filter. Die Modelle werden dabei unter Anwendung des EM-Algorithmus (Ghahramani, 2001) durch Clustering aus Trainingsdaten gelernt. Durch die Kombination von Laserabstandsmessungen mit Kamerabildern ist es möglich, die aus den Lasermessungen mit Hilfe des Tracking-Verfahrens gewonnenen Informationen über die Position der Personen dazu zu benutzen, die Versuchspersonen in den Kamerabildern wiederzufinden. Dazu wird die Position relativ zum Roboter auf die Position und Größe im Bild abgebildet. Auch diese Abbildung wird während des Trainings gelernt. Zur Identifikation werden zwei Modelle einer gehenden Person benutzt, nämlich die Abbilder der Beine und deren Bewegungsmuster in den Abstandsmessungen und die Bewegungsmuster der Personen in den Kamerabildern, die bis zu einem gewissen Maße miteinander korreliert sind. Die Modelle bestehen aus Beispielen und deren zeitlichen Übergängen und einer Abstandsfunktion in Form einer quadratischen Chamfer-Distanz. Die Parameter dieser Metrik werden ebenfalls aus den Trainingsdaten gelernt. Damit besteht die gelernte Hintergrundinformation insgesamt aus den Beispielen, den Parametern der Metrik, der Abbildung der mit dem Laser gewonnenen Position ins Kamerabild und den Übergängen zwischen den Beispielen. Die Beispiele werden aus Trainingsdaten gelernt, indem diese in disjunkte Teilmengen, deren Anzahl vorgegeben ist, unterteilt werden (Hofmann und Buhmann, 1997; Roth et al., 2003). Jeder Cluster wird dabei durch ein Beispiel repräsentiert. Zum Lernen der zeitlichen Übergänge zwischen den Beispielen wird der EM-Algorithmus in der Form einer Variante des Baum-Welch-Algorithmus in Verbindung mit Gibbs-Sampling verwendet (Ghahramani, 2001). Für die Trainingsdaten werden Silhouetten durch Subtraktion des Hintergrundes aus den Bildern und zweidimensionale Punktmengen von Beinen aus den Laserdaten gewonnen. In Experimenten mit zwei Probanden konnte das System den beiden Personen in allen Versuchen nach spätestens zehn Sekunden die richtige Identität zuordnen.

In den Arbeiten von Bellotto und Hu (Bellotto und Hu, 2006; Bellotto und Hu, 2007b) geschieht die Identifikation der Personen durch die Bestimmung der Körpergröße mit einer Kamera. Dazu werden, wie schon erwähnt, die Gesichter detektiert. Außerdem werden aus der Beobachtung der Bekleidung Farbhistogramme gewonnen, die mit Hilfe des Bhattacharyya-Koeffizienten miteinander verglichen werden. Auch die Gesichter selbst werden erkannt (Bellotto, 2008). Der Prozess der Identifizierung umfasst dabei vier Schritte. Zunächst wird eine Vorhersage über die Position einer Person berechnet. Dann werden alle Features ausgesondert, die sich zu sehr von dieser Vorhersage unterscheiden (Gating). Die Features werden dabei aus den Laser-scans extrahiert, indem nach typischen Mustern, hervorgerufen durch die Beine einer Person, gesucht wird. Danach wird unter Verwendung sowohl der Laser- als auch der Kamerainformation eine Assoziationsmatrix zwischen den echten und den vorhergesagten Messungen berechnet, wobei sowohl die Beinmessungen der Laser als auch die Gesichtsmessungen der Kamera verwendet werden. Hierbei werden ähnlich wie beim Gating auch die Identitätsmessungen hinzugezogen. Unter Anwendung des Nearest-Neighbour-Algorithmus wird schließlich im letzten Schritt eine Zuordnung zwischen den Features und den Tracks getroffen. Dabei wird versucht, jedes Feature mit dem nächst gelegenen Objekt zu assoziieren. Im Falle von Interaktionen können dabei aber falsche Zuordnungen auftreten, was sich auch in den Experimenten bestätigt hat. Ansonsten waren die Ergebnisse aber recht genau. Durch die Identifizierung kann darüberhinaus die Zuordnung der Features zu den Tracks verbessert werden, insbesondere, wenn sich Gruppen bilden.

In (Bellotto und Hu, 2007a) wurde dann das Problem der gleichzeitigen Verfolgung und Identifikation ('Joint People Tracking and Identification') eingeführt. Gewöhnliche Algorithmen schätzen nämlich zunächst die Positionen und führen danach erst die Identifikation durch. Die Positionsschätzung beruht dabei in der Regel auf Kamera- oder Laserdaten (Bellotto und Hu, 2007b; Burgard et al., 2002; Chakravarty und Jarvis, 2006; Zajdel et al., 2005). In (Bellotto und Hu, 2007a) dagegen werden diese beiden Schritte simultan ausgeführt. Davor war das Problem des Trackings und gleichzeitigen Klassifizierens in erster Linie im Falle von Fahrzeugen und Flugzeugen bekannt (Gordon et al., 2002; Smets und Ristic, 2004), wo die Objekte bestimmten Fahrzeug- oder Flugzeugtypen zugeordnet werden. Mögliche Anwendungen des 'Joint People Tracking and Identification' sind beispielsweise 'Human-Centred Robotics', wo es um eine synergistische Kooperation zwischen Mensch und Roboter geht, oder das neue Gebiet des 'Socially Assistive Robotics', das das Abstand-Halten und die Interaktion mit einer bestimmten Person zum Inhalt hat (Tapus et al., 2007). Da es eine wichtige Fähigkeit von Service-Robotern ist, bestimmte Personen zu begleiten und mit diesen zu interagieren, müssen diese dazu in der Lage sein, Personen in der Umgebung zu Tracken und zu Erkennen. In (Treptow et al., 2005) beispielsweise wurden eine Wärme- und eine normale Kamera miteinander kombiniert, um eine Person aufzuspüren, zu Verfolgen und zu Identifizieren. Da Vision-basierte Methoden aber eher nur für einfache Situationen geeignet sind, wie zum Beispiel dem Fall nur einer Person (Calisi et al., 2007; Minvielle et al., 2005; Zajdel et al., 2005), werden in komplizierten

Fällen, wie zum Beispiel bei hoher Personendichte oder dem Begleiten einer Person in einem bestimmten Abstand in der Regel kombinierte Ansätze (oft eine Kombination aus Laser und uni- bzw. omnidirektionaler Kameras) verwendet (Bellotto und Hu, 2007b; Chakravarty und Jarvis, 2006). Gebräuchliche Ansätze für Personenerkennung werten dabei Informationen über die Farbe der Bekleidung mit Hilfe von Farbhistogrammen aus (Bellotto und Hu, 2007b; Cielniak und Duckett, 2003). Die Vorteile der Kombination von räumlicher Information und Information über die Identität zeigten sich auch in den Arbeiten von (Zajdel et al., 2005; Zhou und Chellappa, 2002). Letztere Autoren benutzten Informationen aus der Beobachtung von Gesichtern. Die wesentliche Neuigkeit in (Bellotto und Hu, 2007a) besteht nun in der simultanen Verarbeitung sowohl der Positions- als auch der Identitätsdaten, während der experimentelle Aufbau, die Sensoren und die verwendeten Daten im Vergleich zu früheren Arbeiten (Bellotto und Hu, 2006; Bellotto und Hu, 2007b) unverändert geblieben sind. Experimente zeigten, dass die Daten von zehn Personen problemlos in Echtzeit verarbeitet werden können.

Die in Kapitel 11 vorgestellten Experimente zur Begleitung einer ausgewählten Person stellen eine völlig neue Funktionalität dar, die aber beispielsweise auf dem gerade erwähnten Gebiet des 'Socially Assistive Robotics' sehr wichtig ist (Tapus et al., 2007). Es gab hierzu lediglich gewisse Vorarbeiten im Rahmen einer Diplomarbeit an der Universität Bonn (Höller, 2006; Hoeller et al., 2007), wo ein mobiler Roboter einer Person hinterher fuhr. Die Schätzung der Positionen der Personen geschah dabei wieder mit dem SJPDAF von Schulz (Schulz et al., 2001b; Schulz, 2002). Es konnte aber nicht zwischen einzelnen Personen unterschieden werden, wenn diese sich zu nahe kamen. Erwähnt werden müssen auch die Arbeiten von Beymer und Konolige bzw. Lindström und Eklundh (Beymer und Konolige, 2001; Lindström und Eklundh, 2001). Im ersten Fall wurde aber lediglich eine Person mit Hilfe von Vision getrackt, der der Roboter dann folgte. Diese Arbeiten zeigten auch, dass es schwierig sein kann, Personen zu verfolgen, wenn sich der Sensor bewegt. Eine weitere Arbeit zu diesem Thema ist (Montemerlo et al., 2002b). Dort werden mit Laser und Teilchen-Filtern zwei Personen verfolgt, wobei der Roboter eine dieser Personen begleiten soll. Auch dort werden allerdings keine Interaktionen oder Auflösungskonflikte zwischen den beiden Personen zugelassen.

Kapitel 3

Mathematische Grundlagen

Ein grundlegendes Konzept der vorliegenden Arbeit ist der Begriff der Wahrscheinlichkeitsverteilung. Darunter versteht man eine Funktion, die jeder Realisierung eines Vektors oder Zustandes x die Wahrscheinlichkeit ihres Eintrittes $p(x)$ zuordnet. Eine solche Funktion heißt auch Wahrscheinlichkeitsdichte. Für Wahrscheinlichkeitsverteilungen sind zwei wichtige Kenngrößen definiert. Dabei handelt es sich zum einen um den Mittelwert oder Erwartungswert μ mit

$$\mu = \int x p(x) dx. \quad (3.1)$$

Dabei wird über den gesamten Zustandsraum, d.h. alle möglichen Realisierungen, also in der Regel über den \mathbb{R}^n integriert. Der Erwartungswert ist ein Maß für den Wert, um den die Wahrscheinlichkeitsverteilung zentriert ist. Die zweite wichtige Kenngröße ist die Kovarianzmatrix oder Kovarianz Σ . Sie ist definiert zu

$$\Sigma = \int (x - \mu)(x - \mu)^\top p(x) dx. \quad (3.2)$$

Die Kovarianzmatrix gibt ein Maß für die Streuung der Verteilung um den Mittelwert μ bzw. das Rauschen. Große Einträge der Kovarianzmatrix bedeuten eine große Streuung und umgekehrt.

Eine sehr bedeutende Wahrscheinlichkeitsverteilung ist die Gaußverteilung oder Normalverteilung. Seien dazu ein Vektor $\mu \in \mathbb{R}^n$ und eine positiv definite Matrix $\Sigma \in \mathbb{R}^{n \times n}$ gegeben. Dann ist die Dichtefunktion der zugehörigen Normalverteilung definiert durch

$$p(x | \mu, \Sigma) = \frac{1}{(2\pi)^{n/2} (\det \Sigma)^{1/2}} e^{-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)}. \quad (3.3)$$

Der Faktor $\frac{1}{(2\pi)^{n/2} (\det \Sigma)^{1/2}}$ garantiert dabei, dass gilt

$$\int_{\mathbb{R}^n} p(x | \mu, \Sigma) dx = 1. \quad (3.4)$$

Somit ist die Verteilung auf Eins normiert, wie es für eine Verteilungsfunktion auch erforderlich ist. Mit Hilfe der sogenannten charakteristischen Funktion kann man nun zeigen, dass die so definierte Normalverteilung den Mittelwert μ und die Kovarianzmatrix Σ besitzt. Eine Normalverteilung ist also durch Angabe ihres Mittelwertes und ihrer Kovarianz bereits vollständig charakterisiert.

Eine Matrix ist positiv definit genau dann, wenn sie symmetrisch ist und nur positive Eigenwerte besitzt. Insbesondere gilt für eine positiv definite Matrix Σ $x^\top \Sigma x > 0$ für alle $x \in \mathbb{R}^n$. Außerdem ist der geometrische Ort aller Punkte mit $x^\top \Sigma x = c$ mit einem beliebigen $c > 0$ ein Ellipsoid (im Zweidimensionalen insbesondere eine Ellipse) mit dem Ursprung als Mittelpunkt. Die Länge der Halbachsen dieses Ellipsoids wird dabei für eine vorgegebene Konstante c durch die Eigenwerte der Matrix Σ und die Lage der Halbachsen durch deren Eigenvektoren bestimmt. Da für eine positiv definite Matrix Σ auch deren Inverses Σ^{-1} positiv definit ist, sind somit die Orte gleicher Wahrscheinlichkeit für eine Gaußverteilung konzentrische Ellipsoide mit Mittelpunkt μ .

Die Bedeutung der Gaußverteilung liegt darin begründet, dass nach dem zentralen Grenzwertsatz viele Prozesse in der Natur oder der Technik einer Normalverteilung unterliegen. Desweiteren sind für gewisse Prozesse, die auf Normalverteilungen beruhen, einige sehr mächtige mathematische Werkzeuge wie der Kalmanfilter bekannt.

Ein weiterer zentraler Begriff ist der der bedingten Wahrscheinlichkeit $p(A|B)$. Diese ist definiert als die Wahrscheinlichkeit für das Ereignis A , falls das Ereignis B schon eingetreten ist. Im Rahmen der vorliegenden Arbeit entsprechen dabei dem Ereignis A gewisse innere Eigenschaften eines Objektes wie Ort und Geschwindigkeit eines Roboters, die allerdings einer direkten Messung nicht zugänglich sind. Man verfügt aber über gewisse Messungen, die ausgewertet werden können und dem Ereignis B entsprechen, wie zum Beispiel Laserabstandsmessungen. Der Begriff der bedingten Wahrscheinlichkeit liegt auch der Bayes-Statistik zugrunde (Koch, K. R., 2000). Mit Hilfe der Bayes-Statistik können sogenanntes Vorwissen und Messwerte miteinander verknüpft werden, um eine sogenannte a-posteriori Schätzung zu generieren. Im Unterschied dazu können klassische Wahrscheinlichkeiten oder Hypothesentests, die auf dem Gesetz der großen Zahlen beruhen, kein Vorwissen berücksichtigen.

Der Anwendung der bedingten Wahrscheinlichkeit und der Bayes-Statistik im Rahmen dieser Arbeit liegen sogenannte versteckte lineare Gauß-Markov-Prozesse zugrunde. Ein Markov-Prozess ist dabei ein Spezialfall eines stochastischen Prozesses. Unter einem stochastischen Prozess versteht man eine Folge von Realisierungen x_k einer Zufallsverteilung. Dabei ist der Index k diskret und entspricht der Zeitachse. Das besondere eines Markov-Prozesses besteht nun darin, dass die Verteilung zum Zeitpunkt k durch die Verteilung zum Zeitpunkt $k - 1$ schon vollständig festgelegt ist. Von einem Gauß-Markov-Prozess spricht man, wenn die zugrundeliegenden Verteilungen Gaußverteilungen sind. Ein Markov-Prozess wird charakterisiert durch die Zustandsübergangswahrscheinlichkeiten $p(x_k|x_{k-1})$. Bei einem Gauß-Markov-Prozess setzt sich der Übergang vom Zeitpunkt $k - 1$ zum Zeitpunkt k zusammen aus einer deterministischen Komponente $f_k(x_{k-1})$, die nur vom Zustand zum Zeitpunkt $k - 1$

abhängt, und einer stochastischen Komponente w_k . Es ist also

$$x_k = f_k(x_{k-1}) + w_k. \quad (3.5)$$

Dabei ist die Verteilung w_k eine Gaußverteilung mit Mittelwert Null und Kovarianzmatrix Q_k . Man spricht in diesem Falle auch von einem weißen Gaußschen Rauschen. Im Falle eines linearen Modells gilt zusätzlich noch

$$f_k(x_{k-1}) = A_{k-1}x_{k-1} \quad (3.6)$$

mit einer geeigneten Matrix A_{k-1} , die man als Zustandsübergangsmatrix bezeichnet. Der Begriff eines versteckten Modells bedeutet nun, dass die Größen x_k einer direkten Beobachtung nicht zugänglich sind, weshalb diese Variablen auch als versteckte oder innere Variablen bezeichnet werden. Gemessen werden können allerdings gewisse Größen in Form des Beobachtungsvektors z_k , die im Falle eines linearen Gauß-Markov-Modells in der Weise vom inneren Zustand x_k abhängen, dass gilt

$$z_k = B_k x_k + v_k \quad (3.7)$$

mit einer geeigneten Matrix B_k , die auch als Beobachtungsmatrix bezeichnet wird. Die Zufallskomponente v_k ist ähnlich wie zuvor eine Gaußverteilung mit Mittelwert Null und Kovarianzmatrix R_k , insbesondere handelt es sich also wieder um ein weißes Gaußsches Rauschen. Da das eben vorgestellte System rekursiv ist, benötigt man zur vollständigen Charakterisierung noch eine Initialisierung, d. h. man muss eine Anfangsverteilung x_0 vorgeben. Hierfür benutzt man eine Gaußverteilung mit Mittelwert μ_0 und Kovarianzmatrix $P(0|0)$. Unter der Annahme der Unkorreliertheit für die Anfangsverteilung und die Rauschterme w_k und v_k und aller Rauschterme untereinander gilt dann, dass alle Variablen x_k und z_k ebenfalls normalverteilt sind.

Als nächstes wird das Konzept der Filterung oder Zustandsschätzung betrachtet. Darunter versteht man den Versuch, aus den Messungen des Beobachtungsvektors z_k einschließlich bis zum Zeitpunkt k auf die inneren Variablen des Objektes oder den Zustandsvektor x_k zum Zeitpunkt k zu schließen. Aus Rohdaten wird also ein Zustandsvektor erzeugt, wobei das Ziel die beste Wahl des Zustandsvektors für die gegebenen Beobachtungen ist. Dieser wird dabei als dynamisch angenommen, d.h. der Zustandsvektor verändert sich mit der Zeit. Zugrundegelegt wird dabei das eben vorgestellte versteckte lineare Gauß-Markov-Modell, das zur Zusammenfassung noch einmal angegeben wird. Es ist also

$$x_k = A_{k-1}x_{k-1} + w_k \quad (3.8)$$

und

$$z_k = B_k x_k + v_k \quad (3.9)$$

mit den Zustandsübergangsmatrizen A_k und den Beobachtungsmatrizen B_k . Außerdem sind die Kovarianzmatrizen Q_k des sogenannten Prozessrauschens w_k und R_k des Beobachtungsruschens v_k vorgegeben. Diese Rauschprozesse werden dabei als

weiß und Gaußsch, d.h. mit Mittelwert Null und verteilt nach einer Normalverteilung, und als paarweise unkorreliert angenommen. Außerdem bezeichnet man mit $Z_T = z_1, \dots, z_T$ die Gesamtheit aller Messungen vom Zeitpunkt $k = 1$ bis zum Zeitpunkt $k = T$. Unter Filterung versteht man nun die Berechnung der bestmöglichen Schätzung für x_k unter Berücksichtigung aller Messungen Z_T . Es wird also die bedingte Wahrscheinlichkeit $p(x_k|Z_T)$ gesucht. Unter den oben schon erwähnten geeigneten Voraussetzungen sind diese bedingten Wahrscheinlichkeiten ebenfalls wieder Normalverteilungen und somit durch Mittelwert und Kovarianzmatrix vollständig charakterisiert. Aufgabe der Filterung ist es also, diese Mittelwerte und Kovarianzmatrizen zu berechnen. Die Mittelwerte werden dabei mit $x(k|T)$ und die Kovarianzen mit $P(k|T)$ bezeichnet. Allgemein bezeichnet $x(k|l)$ die Schätzung für den Zeitpunkt k basierend auf allen Messungen bis zum Zeitpunkt l . Man unterscheidet dabei drei Fälle. Im Falle von $k > T$ spricht man von Vorhersage oder Prädiktion, wobei hier im Allgemeinen nur der spezielle Fall $T = k - 1$ von Interesse ist. Im Falle von $k = T$ spricht man von Filterung im engeren Sinne. Im Falle von $k < T$, d.h. wenn zur Schätzung des Zustandes zum gerade betrachteten Zeitpunkt auch Messungen verwendet werden, die erst zu einem späteren Zeitpunkt gewonnen werden, spricht man von auch Glättung. Die ersten beiden Probleme werden vom Kalman-Filter gelöst, während auf das dritte Problem der Kalman-Glätter anwendbar ist. Der Kalman-Filter ist ein sehr gebräuchliches Werkzeug in der mobilen Robotik (Crowley, 1989; Dissanayake et al., 2001; Leonard und Durrant-Whyte, 1992). Diese Verfahren werden nun im Einzelnen beschrieben.

Der Kalman-Filter (Kalman, 1960) ist ein rekursives oder sequentielles Verfahren, d.h. es wird eine rekursive und sukzessive Integration verrauschter Messdaten vorgenommen, wobei die aktuelle Messung jeweils sofort verarbeitet wird. Die andere Möglichkeit bestünde darin, alle Messdaten auf einmal zu verarbeiten, was natürlich erst dann möglich wäre, wenn der Messvorgang abgeschlossen ist und alle Messdaten vorliegen. Im letzteren Fall spricht man auch vom offline-Betrieb im Gegensatz zum online-Betrieb, für den der Kalman-Filter geeignet ist und der häufig bei Applikationen im Bereich des Tracking benutzt wird. Außerdem ist der Kalman-Filter linear und diskret, d.h. er berechnet den Zustandsvektor nur zu bestimmten Zeitpunkten. Er wurde am Anfang vor allen Dingen auf Probleme der Raumfahrt angewandt und hat letzten Endes viele Missionen in diesem Bereich erst ermöglicht. Auch im Bereich der Robotik wird der Kalman-Filter für probabilistische Fragestellungen wie eben erwähnt sehr häufig eingesetzt. Der Kalman-Filter berechnet mit Hilfe der sogenannten Bayesschen Regel aus der a-priori-Verteilung, die das Vorwissen, zum Beispiel aus früheren Messungen, beinhaltet, und der aktuellen Sensormessung, deren Wahrscheinlichkeitsverteilung auch als Likelihood bezeichnet wird, die a-posteriori-Verteilung. Der Kalman-Filter liefert dabei für lineare Gauß-Markov-Prozesse die optimale Schätzung. Er benutzt Vorwissen in Form von physikalischen Modellen über das Bewegungsverhalten der Objekte und das Messverhalten der Sensoren und statistischer Annahmen über die bei diesen Prozessen auftretenden Fehler, um mit Hilfe der Beobachtungen in einem geeigneten Koordinatensystem einen Zustandsvektor zu

berechnen. Ein Rekursionsschritt des Kalman-Filters sieht dabei wie folgt aus: gegeben ist eine Verteilung $p(x_{k-1}|Z_{k-1})$ mit zugehörigem Mittelwert $x(k-1|k-1)$ und Fehlerkovarianzmatrix $P(k-1|k-1)$. Die Initialisierung erfolgt dabei mit der Anfangsverteilung. Zunächst wird dann die Prädiktion mit zugehörigem Mittelwert $x(k|k-1)$ und Kovarianz $P(k|k-1)$ berechnet zu

$$x(k|k-1) = A_{k-1}x(k-1|k-1) \quad (3.10)$$

und

$$P(k|k-1) = A_{k-1}P(k-1|k-1)A_{k-1}^\top + Q_k. \quad (3.11)$$

Damit ist das Problem der Vorhersage im engeren Sinne, d.h. im Falle $T = k - 1$, bereits gelöst. Der Kalman-Filter betrachtet nun als nächstes die bedingte Wahrscheinlichkeit $p(z_k|Z_{k-1})$, die ebenfalls eine Normalverteilung ist. Der zugehörige Mittelwert, der eine Vorhersage für die Messung darstellt, soll mit $y(k|k-1)$ und die zugehörige Kovarianzmatrix mit $S(k)$ bezeichnet werden. Es ist dann

$$y(k|k-1) = B_k x(k|k-1) \quad (3.12)$$

und

$$S(k) = B_k P(k|k-1) B_k^\top + R_k. \quad (3.13)$$

Als nächstes wird die sogenannte Innovation $\nu(k)$ untersucht. Diese ist definiert zu

$$\nu(k) = z_k - y(k|k-1). \quad (3.14)$$

Die Innovation entspricht also der Abweichung der Messung von der Vorhersage für diese Messung. Sie ist normalverteilt mit Mittelwert Null und Kovarianzmatrix $S(k)$. Man bezeichnet die Matrix $S(k)$ deshalb auch als Innovationskovarianz. Als nächstes wird die sogenannte Kalman-Gain (Gain steht hierbei für den Informationsgewinn durch den Messvorgang) $K(k)$ berechnet zu

$$K(k) = P(k|k-1) B_k^\top [S(k)]^{-1}. \quad (3.15)$$

Damit ergibt sich die aktuelle Schätzung $x(k|k)$ gemäß der sogenannten Update-Gleichung zu

$$x(k|k) = x(k|k-1) + K(k)\nu(k) \quad (3.16)$$

mit zugehöriger Kovarianz

$$P(k|k) = (I - K(k)B_k)P(k|k-1). \quad (3.17)$$

Der Zustandsvektor wird also in Abhängigkeit von der Differenz aus der tatsächlichen Beobachtung und der Vorhersage für diese Beobachtung korrigiert. Diese Vorhersage wird dabei aus der Vorhersage für den aktuellen Zustand und dem Beobachtungsmodell berechnet. Diese Vorgehensweise garantiert, dass der mittlere quadratische Fehler zwischen berechnetem und tatsächlichem Zustandsvektor minimiert wird.

Man spricht deshalb auch davon, dass der Kalman–Filter optimal ist für lineare Systeme mit weißem Gaußschen Rauschen.

Die Zustandsschätzung besteht also im Wesentlichen aus den drei Schritten der Vorhersage der Messung, der Aktualisierung der Schätzung mit der Update–Gleichung und dem Management der Unsicherheit. Der Update–Schritt ist dabei sequentiell, d.h. nur die jeweils aktuelle Messung wird verarbeitet. Dies ist möglich, da die einzelnen Messungen zueinander unkorreliert sind.

Die Kalman–Gain beeinflusst das Ausmaß der Korrektur der Vorhersage. Falls das Prozessrauschen klein gegenüber dem Messrauschen ist, ist die Kalman–Gain ebenfalls klein und die Vorhersage $x(k|k-1)$ erhält bei der Berechnung der Schätzung $x(k|k)$ gegenüber der Innovation ein größeres Gewicht, da die Messung mit einer größeren Unsicherheit behaftet ist im Vergleich zur Vorhersage, und umgekehrt. Die Kovarianzen $P(k|k-1)$ und $P(k|k)$ beschreiben die Qualität der Zustandsschätzungen. Die Innovationskovarianz wird in erster Linie beim Gating, d.h. der Auswahl derjenigen Messungen, die vom Ziel stammen, benötigt.

Der Kalman–Glätter (Rauch et al., 1965; Shumway und Stoffer, 2000) ist ebenfalls ein rekursives Verfahren. Er wird am Ende einer Messreihe auf Messdaten angewandt, die zuvor schon mit dem Kalman–Filter verarbeitet worden sind. Er führt eine rückwärts gerichtete Rekursion aus und wird mit dem Startwert $x(T|T)$ initialisiert, der zuvor mit dem Kalman–Filter berechnet wurde. Dabei markiert der Zeitpunkt $k = T$ das Ende der Messreihe. Außerdem verwendet er die vom Kalman–Filter bereitgestellten Mittelwerte $x(k-1|k-1)$ und $x(k|k-1)$ und Kovarianzen $P(k-1|k-1)$ und $P(k|k-1)$. Da das Verfahren rekursiv ist, ist vor Ausführung des Glättungsschrittes für den Zeitpunkt $k-1$ desweiteren die Verteilung $p(x_k|Z_T)$ mit zugehörigem Mittelwert $x(k|T)$ und Kovarianz $P(k|T)$ gegeben. Damit ergeben sich Mittelwert $x(k-1|T)$ und Kovarianz $P(k-1|T)$ der Normalverteilung $p(x_{k-1}|Z_T)$ zu

$$x(k-1|T) = x(k-1|k-1) + J(k-1)(x(k|T) - x(k|k-1)) \quad (3.18)$$

und

$$P(k-1|T) = P(k-1|k-1) + J(k-1)(P(k|T) - P(k|k-1))J(k-1)^\top. \quad (3.19)$$

Dabei ist $J(k-1)$ die Matrix

$$J(k-1) = P(k-1|k-1)A_{k-1}^\top [P(k|k-1)]^{-1}. \quad (3.20)$$

Die Kovarianzen und alle weiteren Matrizen sind sowohl beim Kalman–Filter als auch beim Kalman–Glätter unabhängig von den Messungen und können somit bereits vor Aufzeichnung der Messdaten berechnet werden. Insbesondere bedingt eine starke Abweichung einer Messung z_k von der Vorhersage $y(k|k-1)$ keine größere Ungenauigkeit in der Schätzung $x(k|k)$.

Die in diesem Kapitel knapp behandelten Grundlagen sind in (Kräußling, 2007a) noch einmal in aller Ausführlichkeit dargestellt.

Kapitel 4

Einfache Algorithmen zur Verfolgung ausgedehnter Ziele

4.1 Einleitung und experimentelle Rahmenbedingungen

Zur Verfolgung der Objekte, d.h. in erster Linie von Menschen oder anderen Robotern, werden im Rahmen dieser Arbeit Laserabstandsmessungen herangezogen. Ältere Arbeiten benutzen dagegen Kameras. Auch die Verwendung von Sonarsensoren ist möglich. Aufgrund ihrer hohen Genauigkeit und ihres relativ günstigen Preises sind Laserscanner eine gute Wahl für verschiedene Anwendungen in der mobilen Robotik wie zum Beispiel Kollisionsvermeidung, Navigation, Lokalisierung und Kartenbau (Fox et al., 1999b; Thrun, 1998b). Laserabstandsmesser haben im Allgemeinen jeweils ein Sichtfeld von 180 Grad. Deshalb werden zwei solche Sensoren Rücken an Rücken in Beinshöhe an die Roboter montiert, um das gesamte Überwachungsgebiet abzudecken und ein Sichtfeld von 360 Grad zu erhalten (Abbildung 4.1).



Abbildung 4.1: Zwei SICK Laserscanner Rücken an Rücken montiert an einem Roboter.

Die Laser senden Strahlen aus, die von den Objekten in der Umgebung reflektiert werden. Die zurückkommenden Laserstrahlen werden dann von den Scannern registriert und dabei Laufzeit und Intensität gemessen. Aus der Laufzeit kann dann auf die Entfernung der reflektierenden Objekte, insbesondere der Beine von Personen, geschlossen werden. Diese Information wird von den Algorithmen zur Verfolgung der Objekte benutzt. In Kapitel 9.3 dagegen wird untersucht, inwieweit die Intensitäts-

messungen zur Unterscheidung von Personen herangezogen werden können. Die Laserscanner besitzen ein sehr hohes räumliches und zeitliches Auflösungsvermögen. Im Rahmen dieser Arbeit werden die Laser mit einer Frequenz von circa drei bis vier Hertz und einer Abtastrate von einem Grad betrieben, was insgesamt bei jedem Scan in 360 Messungen resultiert. In (Guivant et al., 2000a) werden die Eigenschaften der Laserbeams, die von Laserscannern emittiert werden, genauer untersucht. Grundlegende Ausführungen zu Messungen mit Laserscannern finden sich auch in (Burgard et al., 1996; Fox et al., 1999b; Thrun et al., 2005). Da diese Lasermessungen mit einer gewissen Messungenauigkeit behaftet sind, müssen zur Auswertung der Daten probabilistische Verfahren angewandt werden. Abbildung 4.2 zeigt eine typische Messung in der Roboterexperimentalhalle der FGAN. Die Position des beobachtenden Roboters

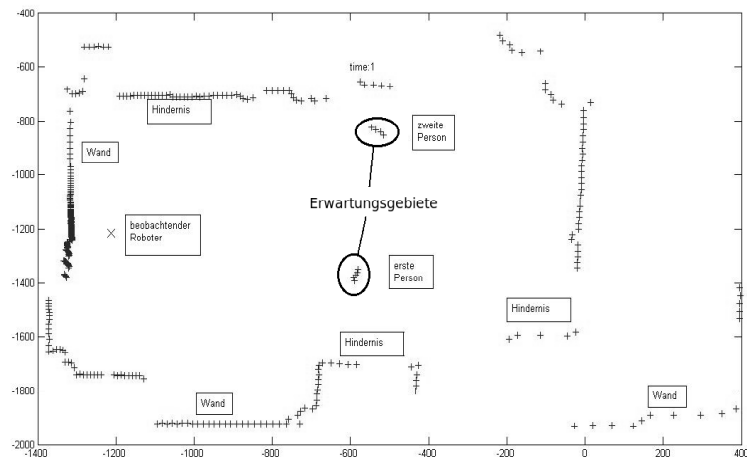


Abbildung 4.2: Messungen eines Laserscans. Als Einheiten wurden Zentimeter verwendet. Für die Messungen von den beiden Personen sind auch mögliche Erwartungsgebiete eingezeichnet.

ist links mit einem Kreuz markiert. Die meisten Messungen stammen von den Wänden des Labors. Daneben befinden sich noch einige längliche Hindernisse und zwei Personen in der Halle. Die beiden Personen erzeugen jeweils einen Cluster von Messungen. Um diese Cluster sind auch zwei sogenannte Erwartungsgebiete skizziert. Diese werden im übernächsten Abschnitt 4.3 eingeführt. Bilder des Labors finden sich in Kapitel 11.

In diesem Kapitel werden einige einfache Algorithmen, die zur Verfolgung ausgedehnter Einzelziele entwickelt wurden, vorgestellt und beschrieben. Die einzelnen Verfahren sind dabei zum Teil eng miteinander verwandt. Zuvor werden aber einige Gemeinsamkeiten wie das verwendete Modell oder das Erwartungsgebiet untersucht.

Gängige Ansätze extrahieren aus den Lasermessungen sogenannte Features als lokale Abstandsminima und bilden danach ein ungewichtetes Mittel aus den zugehörigen

Messungen, das dann mit einem Kalman–Filter oder einem Teilchen–Filter weiterverarbeitet wird (Bellotto und Hu, 2006; Schulz et al., 2001b). Dabei kann aber wichtige Information verloren gehen, was sich insbesondere im Falle interagierender Ziele sehr nachteilig auswirken und zum Verlust von Tracks führen kann. Es wird deshalb nun ausführlich untersucht, auf welche Weise die Messungen, die durch Feature–Extraktion oder mit Hilfe eines Erwartungsgebietes einem Ziel oder einer Gruppe von Zielen zugeordnet wurden, weiterverarbeitet werden können.

4.2 Das verwendete Modell

Bei der Wahl des Zustandsmodells der Bewegung von Personen muss beachtet werden, dass die Bewegungen von Menschen nicht vorhersagbar sind, auch wenn dies gelegentlich, wie in (Bennewitz et al., 2002a; Bennewitz et al., 2002b; Bennewitz et al., 2003), versucht wird. Dort werden Bewegungsmuster in einem Gebäude gelernt, was aber auch noch sehr aufwendig ist. Auch in (Höller, 2006; Hoeller et al., 2007) wird versucht, das Bewegungsverhalten von Personen vorherzusehen. Dieser Umstand macht das Tracking von Personen sehr schwierig. Eine Möglichkeit der Modellierung besteht in der Annahme einer Brownschen Bewegung wie in (Montemerlo et al., 2002a). Oft wird auch eine Bewegung mit konstanter Geschwindigkeit angenommen wie in (Beymer und Konolige, 2001; Schulz et al., 2003a). In (Bellotto und Hu, 2007b) wird ebenfalls von einer Bewegung mit konstanter Geschwindigkeit ausgegangen. Zusätzlich wird aber ein weißes Gaußsches Rauschen aufaddiert und es werden nur Bewegungen nach vorne erlaubt.

Im Rahmen dieser Arbeit werden die Dynamik der zu beobachtenden Objekte und der Beobachtungsprozess selbst mit dem im vorigen Kapitel vorgestellten versteckten linearen Gauß–Markov–Prozess beschrieben. Ein lineares Modell ist deshalb ausreichend, da die Laserscanner mit einer hohen Abtastrate (mehrere Scans pro Sekunde) betrieben werden und somit ausreichend häufig Beobachtungen zur Verfügung stehen. Dabei ändern sich die Zustandsübergangsmatrizen A_k , die Beobachtungsmatrizen B_k und die Kovarianzmatrizen Q_k und R_k des Prozessrauschens und des Beobachtungsrauschens nicht über die Zeit. Auf das Subscript k kann also in diesen Fällen verzichtet werden. Es ist also

$$x_k = Ax_{k-1} + w \quad (4.1)$$

und

$$z_k = Bx_k + v. \quad (4.2)$$

In diesen Gleichungen wird dabei normalerweise zugrundegelegt, dass die Objekte punktförmig sind und der innere Zustand x_k den Zustand eines solchen punktförmigen Objektes beschreibt. Außerdem wird angenommen, dass das Objekt zu jedem Zeitpunkt nur Urheber einer Messung z_k ist. Dennoch ist dieses Modell auch nützlich für die Beschreibung eines ausgedehnten Zieles, denn in diesem Fall kann x_k beispielsweise aufgefasst werden als Zustandsvektor des Schwerpunktes oder eines ande-

ren ausgezeichneten Punktes des Objektes. Zusätzlich kann aus allen Messungen eine repräsentative Beobachtung oder ein repräsentatives Feature z_k gebildet werden, das dann in der Update-Gleichung des Kalman-Filters verwendet wird.

Da die Bewegung der Ziele durch die Lasermessungen auf eine horizontale Ebene projiziert wird, wird ein zweidimensionales kinematisches Modell zugrundegelegt. Deshalb wird als Zustandsvektor gewählt

$$x_k = \begin{pmatrix} x_{k1} & x_{k2} & \dot{x}_{k1} & \dot{x}_{k2} \end{pmatrix}^\top. \quad (4.3)$$

Dabei sind x_{k1} und x_{k2} die kartesischen Koordinaten des Objektes in der Ebene zum Zeitpunkt k und \dot{x}_{k1} und \dot{x}_{k2} die entsprechenden Geschwindigkeiten. Allgemein ist die Wahl des Zustandsvektors abhängig von den beobachtbaren Daten, was in der Kontrolltheorie auch als das Problem der Beobachtbarkeit bezeichnet wird. Gleichzeitig sollte der Zustandsvektor im Idealfalle der minimale Vektor sein, der die Möglichkeit zur Vorhersage aller benötigten Größen liefert.

Für die Beschreibung der zeitlichen Entwicklung des Ortes wird die Gleichung für eine deterministische Bewegung mit konstanter Geschwindigkeit verwandt. Es ist also

$$x_{k+1,j} = x_{kj} + \Delta T \dot{x}_{kj} \quad (4.4)$$

mit $j = 1, 2$. Dabei ist ΔT die Zeitspanne zwischen den Messungen zu den Zeitpunkten k und $k + 1$. Für die zeitliche Entwicklung der Geschwindigkeit wird ein an der FGAN entwickeltes Modell (van Keuk, 1971) mit exponentiell abklingenden Werten und einem weißen Gaußschen Rauschterm angenommen, d.h. es soll gelten

$$\dot{x}_{k+1,j} = e^{-\Delta T/\Theta} \dot{x}_{kj} + \Sigma \sqrt{1 - e^{-2\Delta T/\Theta}} u(k) \quad (4.5)$$

für $j = 1, 2$. Dabei soll $u(k)$ demzufolge ein paarweise unkorreliertes weißes Gaußsches Rauschen mit

$$E [u(m)(u(n))^\top] = \delta_{mn} \quad (4.6)$$

sein. Die Wahl des ersten Terms in Gleichung 4.5 sorgt dafür, dass die Geschwindigkeit exponentiell abfällt, während der zweite Term das Prozessrauschen und die Beschleunigungen abbildet. Damit folgt für die Zustandsübergangsmatrix A

$$A = \begin{pmatrix} 1 & 0 & \Delta T & 0 \\ 0 & 1 & 0 & \Delta T \\ 0 & 0 & e^{-\Delta T/\Theta} & 0 \\ 0 & 0 & 0 & e^{-\Delta T/\Theta} \end{pmatrix} \quad (4.7)$$

und für die Kovarianzmatrix Q des Prozessrauschens

$$Q = \Sigma^2 (1 - e^{-2\Delta T/\Theta}) \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (4.8)$$

In den Rauschterm $\Sigma\sqrt{1 - e^{-2\Delta T/\Theta}}u(k)$ gehen sowohl Störungen und zufällige Fehler in der Bewegung zum Beispiel durch Umwelteinflüsse als auch Ungenauigkeiten im Dynamikmodell selbst und schließlich auch noch die aktiven Beschleunigungsprozesse des verfolgten Zieles ein.

Die Parameter Θ und Σ beschreiben die Dynamik des Zieles. Θ kann dabei als Abklingkonstante angesehen werden. Je größer Θ ist, desto langsamer klingen die Bewegungen des Zieles ab. Σ beschreibt die Beschleunigungen oder den Input des Systems. Ein großes Σ entspricht dabei großen und häufigen Beschleunigungen. Die Einfügung des Faktors $\sqrt{1 - e^{-2\Delta T/\Theta}}$ in den Rauschterm der Gleichung 4.5 geschieht aus Normierungsgründen. Es folgt nämlich deswegen für den Erwartungswert $E[\dot{x}_{kj}]^2$ des Quadrates der Geschwindigkeit (Kräußling et al., 2005b)

$$\lim_{k \rightarrow \infty} E[\dot{x}_{kj}]^2 = \Sigma^2. \quad (4.9)$$

Die Messwerte, die das System liefert, sind zunächst die Entfernungen der reflektierenden Objekte und die dazugehörigen Winkel. Diese Polarkoordinaten werden aber sofort in kartesische Koordinaten umgerechnet. Die Beobachtungsmatrix B extrahiert somit lediglich aus dem Zustand x_k die kartesischen Koordinaten und ist damit eine einfache Projektionsmatrix. Es gilt also

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \quad (4.10)$$

Natürlich könnten auch kompliziertere Modelle der Bewegung und der Form eines Menschen verwendet werden wie in (Arras et al., 2008; Brooks und Williams, 2003; Taylor und Kleeman, 2004; Zhao und Shibasaki, 2005). Diese Autoren haben Modelle für den Gang eines Menschen entwickelt, indem sie die Bewegungen der beiden Beine eines Menschen explizit modelliert haben. Sie verwendeten dabei die Tatsache, dass die Laserscanner für gewöhnlich in der Höhe der Beine der Personen an den Robotern befestigt sind. In dieser Arbeit wird ein solcher Ansatz nicht verfolgt, da der damit verbundene Rechenaufwand sehr hoch wäre. Darüberhinaus scheint es oft schwierig zu sein, die beiden Beine aus den Lasermessungen zu identifizieren, wie zum Beispiel aus Abbildung 4.2 ersichtlich ist. Außerdem kann die Situation gerade bei interagierenden Zielen sehr kompliziert werden und in einem dramatischen Anstieg der Anzahl der Hypothesen des Modells resultieren (Taylor und Kleeman, 2004).

Die Standardabweichung des Fehlers in der Winkelmessung wird mit σ_ϕ bezeichnet, während bei der Entfernungsmessung für die Standardabweichung σ_r geschrieben wird. Damit erhält man nach einer nichtlinearen Transformation von polaren in kartesische Koordinaten für die Kovarianz R des Beobachtungsrauschens (Kräußling et al., 2005b)

$$R = \begin{pmatrix} \sigma_r^2 \cos^2 \phi_k + r_k^2 \sigma_\phi^2 \sin^2 \phi_k & (\sigma_r^2 - r_k^2 \sigma_\phi^2) \sin \phi_k \cos \phi_k \\ (\sigma_r^2 - r_k^2 \sigma_\phi^2) \sin \phi_k \cos \phi_k & \sigma_r^2 \sin^2 \phi_k + r_k^2 \sigma_\phi^2 \cos^2 \phi_k \end{pmatrix}. \quad (4.11)$$

Dabei sind r_k und ϕ_k die Polarkoordinaten der Position des Objektes. Für r_k und ϕ_k setzt man dabei die Werte ein, die der besten Schätzung aufgrund des gerade vorhandenen Wissens entsprechen. Dies sind aber gerade die Werte r_k und ϕ_k , die zu der aktuellen Vorhersage $x(k|k-1)$ gehören. Bei den für unsere Experimente verwendeten Sick Lasern liegt σ_ϕ im Bereich weniger Winkelgrade, während σ_r wenige Zentimeter beträgt (Sick AG, 2006).

Die wesentliche Besonderheit im Zusammenhang mit den in der mobilen Robotik auftretenden Objekten und den verwendeten Sensoren besteht nun gerade darin, dass die Messungen nicht nur aufgrund des Messrauschens, sondern auch wegen der Ausdehnung der Objekte eine gewisse Streuung zeigen und von der vorhergesagten Position des Zieles abweichen. Die von den Sensoren ausgesandten Laserstrahlen treffen die Objekte nämlich in der Regel in unterschiedlichen Oberflächenpunkten. Die Ausdehnung der Objekte muss deshalb an derselben Stelle berücksichtigt werden wie das Messrauschen. Dies geschieht aber bei der Berechnung der Innovationskovarianz $S(k)$. Die herkömmliche Formel zur Berechnung der Innovationskovarianz im Rahmen des Kalman-Filters lautet nach Gleichung 3.13

$$S(k) = BP(k|k-1)B^\top + R \quad (4.12)$$

mit der Kovarianz R des Messfehlers. Bei ausgedehnten Zielen sollte deshalb $S(k)$ berechnet werden nach der Formel

$$S(k) = BP(k|k-1)B^\top + R + E \quad (4.13)$$

mit einer positiv definiten Matrix E , die die Ausdehnung und Oberflächenform der Objekte beschreiben soll. Die Eigenwerte und Eigenvektoren dieser Matrix bilden dabei die Geometrie der Objekte ab. Eine gute Näherung für die Form der Ziele in der mobilen Robotik ist nun ein Kreis, da viele mobile Roboter kreisförmig sind, während sich der Querschnitt von Personen sehr schnell ändern kann, so dass es sehr schwer ist, die Form von Personen besser abzuschätzen als durch einen Kreis. Einem Kreis entspricht aber eine Diagonalmatrix mit gleichen positiven Einträgen. Es wird deshalb

$$E = \delta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (4.14)$$

mit $\delta > 0$ gesetzt. Der Parameter δ sollte dabei umso größer sein, je ausgedehnter ein Ziel ist.

Da die durch die Ausdehnung der Objekte hervorgerufenen Abweichungen von der Vorhersage für die in der mobilen Robotik auftretenden Objekte wie Menschen oder andere mobile Roboter in der Regel erheblich größer sind als die Abweichungen durch den Messfehler, werden die Einträge der Matrix E wesentlich größer sein als die Einträge der Matrix R . Tatsächlich entsprechen die Einträge von E einigen hundert Quadratzentimetern, während sich die Einträge von R lediglich im Bereich von einigen Quadratzentimetern bewegen. Man begeht deshalb keinen allzu großen

Fehler, wenn man bei der Berechnung der Innovationskovarianz die Kovarianz R des Messfehlers nicht berücksichtigt. Für die Berechnung von $S(k)$ wird also im Folgenden die vereinfachte Gleichung

$$S(k) = BP(k|k-1)B^T + E \quad (4.15)$$

zugrunde gelegt. Diese Vorgehensweise hat auch den wichtigen Vorteil, dass die Folge der Kalman-Gains und die entsprechenden Folgen der Kovarianzen sehr schnell konvergieren, so dass lediglich einmal die zugehörigen Grenzwerte berechnet werden müssen und die Verfahren ansonsten keiner Berechnung von Matrizen mehr bedürfen. Die Berechnung dieser Grenzwerte kann dabei auch vor den eigentlichen Experimenten durchgeführt werden. Dieser Sachverhalt wird in Anhang A näher untersucht.

Damit ist die Beschreibung des den Algorithmen zugrundeliegenden Modelles der Bewegung und Beobachtung ausgedehnter Objekte vollständig. Der wesentliche Vorteil bei der Verwendung dieses Modells besteht darin, dass der Kalman-Filter (Kalman, 1960; Maybeck, 1979; Shumway und Stoffer, 2000) zur Durchführung der Berechnung der entsprechenden Positionsschätzungen herangezogen werden kann. Der Kalman-Filter ist dabei der gebräuchlichste Filter im Rahmen von Tracking-Anwendungen. Daneben werden in jüngerer Zeit gelegentlich auch noch Teilchen-Filter eingesetzt (Montemerlo et al., 2002b; Schulz et al., 2001b). Das Modell verwendet drei Parameter, nämlich Θ und Σ , die die Dynamik der Ziele beschreiben, und δ , das die Ausdehnung des Zieles wiedergibt. Die geeignete Wahl dieser Parameter wird in Abschnitt 5.2 diskutiert.

4.3 Die Grundstruktur eines Verfolgungs-Algorithmus

Die im Folgenden verwendeten Tracking-Verfahren basierend auf Gauß-Markov-Modellen sind rekursive Ansätze, die als Zustandsschätzer den Kalman-Filter verwenden. Der Zustandsschätzer spielt dabei eine Schlüsselrolle in jedem Tracking-System. Er integriert die Sensordaten und berechnet daraus die aktuelle Zustandsschätzung. Die Grundstruktur der Algorithmen ist solchen Verfahren entnommen, wie sie in der Luftraumüberwachung üblich sind (Bar-Shalom und Fortmann, 1988). Diese Methoden berechnen eine Schätzung für die Position und die Geschwindigkeit des verfolgten Objektes. Es wurden lediglich einige Modifikationen vorgenommen, wie die Einführung einer positiv definiten Matrix zur Berücksichtigung der Ausdehnung der Ziele.

Die Verfahren benötigen, da sie rekursiv sind, sowohl eine Initialisierung, d.h. eine Anfangsschätzung, als auch eine Vorschrift zur Durchführung des Rekursions-schrittes. Die Initialisierung oder Track-Detektion kann im Falle eines stationären, d.h. sich nicht bewegenden beobachtenden Roboters am Einfachsten wie folgt realisiert werden: die beim ersten Scan ermittelten Entfernungswerte werden für alle Winkel von einem bis 360 Grad gespeichert. Bei jeder folgenden Messung wird jeder Messwert dahingehend geprüft, ob seine Entfernung von der für diesen Winkelwert gespeicherten Entfernung in nennenswerter Weise abweicht. Ist die Entfernung der neuen Messung

größer, so wird die gespeicherte Entfernung durch die neue Entfernung ersetzt. Auf diese Weise wird der Fall behandelt, dass die ursprünglich gespeicherte Entfernung nicht zu Wänden oder anderen festen Hindernissen, sondern zu einem beweglichen Objekt gehört, das seinen Platz in der Zwischenzeit verlassen und dadurch dem Laser die Sicht auf weiter entfernt liegende Ziele freigemacht hat. Diese Vorgehensweise ermöglicht es, auch in den Bereichen des Überwachungsgebietes, die ursprünglich von nicht-stationären Zielen verdeckt waren, neue Objekte aufzuspüren. Ist die neu gemessene Entfernung dagegen kleiner als die gespeicherte Entfernung, so muss die neue Messung von einem sich bewegenden Ziel herrühren, das den Blick auf das ursprünglich beobachtete Objekt versperrt. Lässt sich diese Messung nicht einem bereits bekannten Ziel zuordnen, so muss in diesem Fall ein neues Ziel initialisiert werden. Allerdings muss dabei aufgrund der Messungenauigkeiten mit einer Sicherheitsschwelle gearbeitet werden, d.h. Messungen werden nur dann einem beweglichen Objekt zugeordnet, wenn die neu gemessene Entfernung kleiner ist als die Ursprüngliche abzüglich dieser Toleranz. Eine Sicherheitsschwelle von dreißig Zentimetern hat sich dabei als brauchbar erwiesen. Im Falle eines sich bewegenden beobachtenden Roboters ist die Track-Detektion ungleich schwieriger und wird deshalb erst in Kapitel 10.1 besprochen.

Zur Initialisierung sind ein Zustandsvektor $x(0|0)$ mit Ort und Geschwindigkeit und eine Kovarianzmatrix erforderlich. Als Ort wählt man einfach ein ungewichtetes Mittel der fraglichen Messungen in kartesischen Koordinaten, während die Anfangsgeschwindigkeit gleich Null gesetzt wird. Als initiale Kovarianzmatrix $P(0|0)$ wurde jeweils

$$P(0|0) = \begin{pmatrix} 100 & 0 & 0 & 0 \\ 0 & 100 & 0 & 0 \\ 0 & 0 & 10000 & 0 \\ 0 & 0 & 0 & 10000 \end{pmatrix} \quad (4.16)$$

verwendet. Dies entspricht einer Standardabweichung von zehn Zentimetern in den Ortskoordinaten und von $100 \frac{cm}{s}$ in den Geschwindigkeitskoordinaten. Der innere Zustand des Zieles wird also initial durch eine Gaußverteilung beschrieben. Da das verwendete Modell ein Gauß-Markov-Modell ist, gewährleistet diese Vorgehensweise, dass der innere Zustand des Objektes auch zu allen folgenden Zeiten durch eine Normalverteilung, charakterisiert durch Mittelwert und Kovarianzmatrix, festgelegt wird.

Nun wird der Rekursionsschritt beschrieben. Dieser hat die Berechnung der aktualisierten Werte des Erwartungswertes und der Kovarianzmatrix des Zustandes zum Ziel und wird mit Hilfe des Kalman-Filters umgesetzt. Seien dazu der Mittelwert oder die Schätzung $x(k|k)$ für den inneren Zustand x_k zum Zeitpunkt k und die entsprechende Kovarianzmatrix $P(k|k)$ gegeben. Dann werden der neue Mittelwert $x(k+1|k+1)$ und die neue Kovarianz $P(k+1|k+1)$ für den Zeitpunkt $k+1$ unter Verwendung der Messungen zum Zeitpunkt $k+1$ wie folgt bestimmt. Zunächst wird nach den Gleichungen 3.10 und 3.11 eine Vorhersage $x(k+1|k)$ und $P(k+1|k)$ des inneren Zustandes

gemäß den Formeln

$$x(k+1|k) = Ax(k|k) \quad (4.17)$$

und

$$P(k+1|k) = AP(k|k)A^T + Q \quad (4.18)$$

berechnet. Danach wird mit den Formeln 3.12 und 4.15 eine Vorhersage $y(k+1|k)$ für die Messungen mit zugehöriger Innovationskovarianz $S(k+1)$ mit Hilfe der Gleichungen

$$y(k+1|k) = Bx(k+1|k) \quad (4.19)$$

und

$$S(k+1) = BP(k+1|k)B^T + E \quad (4.20)$$

bestimmt. Dann wird unter Zugrundelegung dieser beiden Größen das sogenannte Erwartungsgebiet festgelegt. Anschließend werden mit dessen Hilfe aus allen Entfernungsmessungen $z_i(k+1)$, $i = 1, \dots, 360$, zum Zeitpunkt $k+1$ diejenigen Messwerte ausgewählt, die dem gerade betrachteten Ziel zugeordnet werden sollen. Hierbei wird die Annahme zugrundegelegt, dass die Laserscanner mit einer Winkelauflösung von einem Grad betrieben werden und somit 360 Entfernungsmessungen vorliegen.

Man spricht beim letzten Schritt auch von der sogenannten Datenassoziiierung (Bar-Shalom und Fortmann, 1988; Hall, 1992). Darunter versteht man allgemein die Zuordnung von realen Messungen zu computerinternen Objekten. Bei vielen Anwendungen geht dieser die sogenannte Feature-Extraktion voraus. Im Rahmen dieser Arbeit wird eine Feature-Extraktion erst in Abschnitt 10.1 eingeführt. Die Datenassoziiierung ist in der mobilen Robotik nicht nur im Bereich des People-Tracking (Cielniak et al., 2007) von Bedeutung, so dass es nicht verwundert, dass es beispielsweise auf der Konferenz 'Robotics: Science and Systems 2008' einen eigenen Workshop mit dem Titel 'Inside Data Association' zu diesem Thema gab. Auch im Bereich der Lokalisierung oder des SLAM zählt die Daten-Assoziiierung zu den schwierigsten und komplexesten Teilproblemen (Durrant-Whyte et al., 2003; Guivant et al., 2004; Montemerlo und Thrun, 2007; Nieto et al., 2003). Darunter versteht man in diesem Fall die Zuordnung der Messungen oder Features zu den computerinternen Landmarken, wobei zusätzlich die Schwierigkeit besteht, dass Messungen von noch unbekanntem Landmarken stammen oder nicht echt sein können (Leonard et al., 2003). Ansätze zur Lösung des Assoziierungs- oder Korrespondenzproblems im Bereich des SLAM benutzen nun Maximum-Likelihood- (Dissanayake et al., 2001; Thrun et al., 2005) oder Sampling-Methoden (Neira et al., 2003), neuerdings auch den EM-Algorithmus (Expectation Maximization) (Thrun et al., 1998) oder Monte-Carlo-Methoden (Arnedo, 2003). Ganz aktuell sind Arbeiten im Bereich des 6D-SLAM. Dort rotiert ein Laserscanner um seine Achse, um dreidimensionale Daten aufzuzeichnen. Auch hier tritt bei der Zuordnung der gemessenen Features zu den computerinternen Landmarken wieder das Korrespondenzproblem auf (Nüchter et al., 2005; Nüchter et al., 2007). Im Bereich Computer-Vision spielt das Korrespondenzproblem ebenfalls eine wichtige Rolle (Fischler und Bolles, 1981). Dabei kann ein schlecht gewählter Ansatz zur

Lösung der Daten–Assoziierung in allen Bereichen in den meisten Fällen zu einem katastrophalen Fehlverhalten des gesamten Systems führen.

Bei der Verfolgung von sich bewegendenden Objekten in Gebäuden liegt nun der Fall vor, dass die Umgebung in Winkelabständen von zum Beispiel einem Grad mit einem Laser abgetastet und dabei der Abstand zum nächsten reflektierenden Objekt gemessen wird. Es stammen aber nur sehr wenige dieser dann 360 Messungen beispielsweise bei der Beobachtung der Bewegung eines Menschen in einem typischen Abstand von einigen Metern von dem zu verfolgenden Ziel selbst, während die weit überwiegende Mehrzahl aller Messungen von Hindernissen sowie den Wänden oder anderen Objekten herrührt. Größenordnungsmäßig liegt die Anzahl der Messungen von einem Ziel in der Regel im Bereich von ca. fünf bis zehn. Deshalb müssen die wenigen Messungen, die von dem zu verfolgenden Objekt stammen, aus der Menge aller Messungen herausgefiltert werden. Andernfalls wird beispielsweise die Position des Zieles bei zu starker Annäherung an eine Wand von den meisten Algorithmen mit der Position dieser Wand geschätzt, da dann die überwiegende Mehrzahl aller Messungen von dieser Wand stammt. Diese Filterung geschieht wie schon angedeutet mit Hilfe eines Erwartungsgebietes (Bar-Shalom und Fortmann, 1988; Singer und Kanyuck, 1971; Sittler, 1964). Dazu wird, ausgehend von der eben berechneten Vorhersage $y(k + 1|k)$ für die Messungen und der Innovationskovarianz $S(k + 1)$, für jeden Messwert $z_i(k + 1)$ ($i = 1, \dots, 360$) des Zeitpunktes $k + 1$ die sogenannte Mahalanobis–Distanz λ (Mahalanobis, 1936) mit

$$\lambda = (z_i(k + 1) - y(k + 1|k))^T \cdot [S(k + 1)]^{-1} \cdot (z_i(k + 1) - y(k + 1|k)) \quad (4.21)$$

ermittelt.

Im Gegensatz zum Euklidischen Abstand berücksichtigt die Mahalanobis–Distanz mögliche unterschiedliche Unsicherheiten in den verschiedenen Richtungen. Um dies zu veranschaulichen, kann man sich vorstellen, dass $S(k + 1)$ eine Diagonalmatrix ist, was mit einer geeigneten Koordinatentransformation in Form einer Drehung immer erreicht werden kann. Eine größere Unsicherheit in einer Richtung geht dann mit einem größeren zugehörigen Diagonaleintrag in der Matrix $S(k + 1)$ und damit einem kleineren zugehörigen Diagonaleintrag in der inversen Matrix $[S(k + 1)]^{-1}$ einher. Damit werden Abweichungen der Messungen $z_i(k + 1)$ von der Vorhersage $y(k + 1|k)$ in dieser Richtung mit einem kleineren Faktor multipliziert und somit bei der Berechnung der Mahalanobis–Distanz weniger stark gewichtet. Genauer gesagt liegen, da mit $S(k + 1)$ auch $[S(k + 1)]^{-1}$ positiv definit ist, alle Punkte mit derselben Mahalanobis–Distanz auf der Oberfläche einer Ellipse mit Mittelpunkt $y(k + 1|k)$, deren große Halbachse zur Richtung mit der größeren Unsicherheit gehört.

Sei beispielsweise $\lambda = 1$ und

$$S = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} \quad (4.22)$$

und damit

$$S^{-1} = \begin{pmatrix} 1/4 & 0 \\ 0 & 1 \end{pmatrix}. \quad (4.23)$$

Alle Punkte mit der Mahalanobis-Distanz $\lambda = 1$ liegen dann auf einer Ellipse mit der großen Halbachse in x -Richtung vom Betrag 2 und der kleinen Halbachse in y -Richtung vom Betrag 1 (Abbildung 4.3).

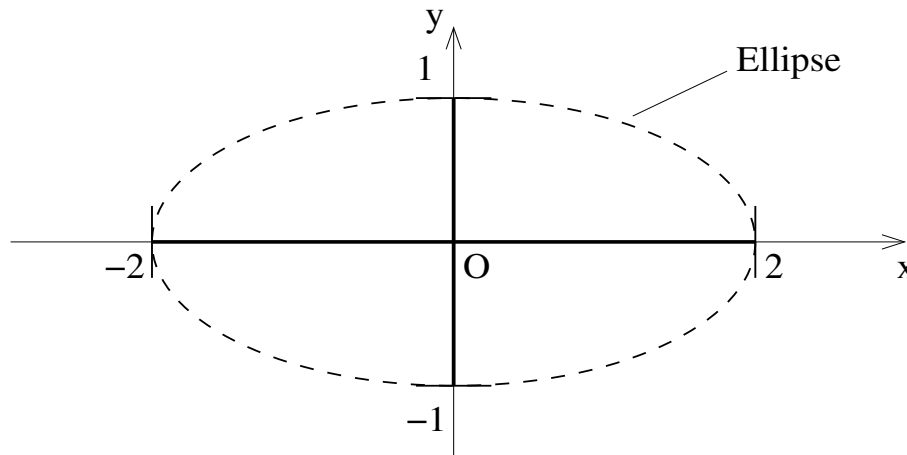


Abbildung 4.3: Illustration der Mahalanobis-Distanz.

Die Mahalanobis-Distanz ist χ^2 -verteilt mit zwei Freiheitsgraden und somit wird ein χ^2 -Test verwendet, um die richtigen Messungen zu finden (Kräußling, 2007a). Alle Messwerte mit einer Mahalanobis-Distanz von der Vorhersage kleiner als eine vorgegebene Schranke λ_{\max} liegen innerhalb einer bestimmten Ellipse und werden mit dem Ziel assoziiert. Dabei wurden in der Regel Werte zwischen $\lambda_{\max} = 4$ und $\lambda_{\max} = 9$ verwendet. Ein Wert von $\lambda_{\max} = 9$ entspricht dabei im Falle eines punktförmigen Zieles einer Wahrscheinlichkeit $p_G = 0.989$ dafür, dass die wahre Messung innerhalb des Erwartungsgebietes liegt (Bar-Shalom und Fortmann, 1988). Im Falle von $\lambda = 4$ ($\lambda = 9$) werden im Rahmen des hier vorgestellten Modells und einer Wahl der Parameter gemäß Abschnitt 5.2 noch Messungen dem Ziel zugeordnet, die von der Vorhersage einen Abstand von ca. sechzig (neunzig) Zentimetern haben. In Abbildung 4.2 ist für die beiden beobachteten Personen jeweils ein denkbare Erwartungsgebiet eingezeichnet.

Letzten Endes werden die zu den einzelnen Tracks gehörigen Vorhersagen mit den Messungen verglichen. Weitere Abstandsmaße sind die Euklidische Metrik und die Bhattacharyya-Metrik, bei der es sich um eine Verallgemeinerung der Mahalanobis-Distanz für den Fall handelt, dass Beobachtung und Vorhersage nach verschiedenen Kovarianzen verteilt sind.

An dieser Stelle erkennt man auch, warum es so wichtig ist, bei der Berechnung der Innovationskovarianz eine genügend große positiv definite Matrix aufzuaddieren, die die Ausdehnung der Objekte modellieren soll. Aufgrund dieser Ausdehnung streuen die Messungen nämlich, insbesondere bei Objekten mit dynamischer Form wie Menschen, sehr stark um die Vorhersage. Es könnte deshalb sonst passieren, dass viele

Messungen, die eigentlich vom Ziel stammen, ausgeschlossen werden oder sogar gar keine Messung innerhalb des dann viel zu kleinen Erwartungsgebietes liegen würde und das Ziel deshalb verloren geht.

Als Ergebnis dieser Daten–Assoziierung, die sowohl geometrische (die Ausdehnung der Objekte) als auch statistische Aspekte in Form des Prozessrauschens berücksichtigt, erhält man eine Menge der Mächtigkeit m_{k+1} von Messungen $\{\tilde{z}_i(k+1)\}_{i=1}^{m_{k+1}}$, deren Mahalanobis–Distanz von der Vorhersage $y(k+1|k)$ kleiner ist als λ_{\max} und die deshalb innerhalb des Erwartungsgebietes liegen. Im Idealfall handelt es sich dabei genau um diejenige Gruppe von Messungen, die zu dem gerade betrachteten Ziel gehören. Ausnahmen hiervon treten eigentlich nur dann auf, wenn sich Objekte einander annähern oder mit statischen Objekten wie Möbeln oder Wänden in Kontakt kommen, da bei der Verwendung von Laserscannern praktisch keine Falschmeldungen auftreten. Bei hoher Objektdichte ist eine zuverlässige Datenassoziiierung in der Regel dagegen nicht mehr möglich. Die damit verbundenen Probleme werden in den Abschnitten 6 und 7 behandelt.

Mit Hilfe dieser Menge $\{\tilde{z}_i(k+1)\}_{i=1}^{m_{k+1}}$ wird nun ein repräsentativer Wert $z(k+1)$ erzeugt, der dann mit Hilfe des Kalman–Filters weiter verarbeitet wird. Die Generierung dieses Wertes oder Features kann im Wesentlichen auf zwei verschiedenen Wegen geschehen. Zum einen können alle Messwerte $\tilde{z}_i(k+1)$ in Form eines ungewichteten oder gewichteten Mittels verwendet werden. Im letzteren Fall wird gesetzt

$$z(k+1) = \sum_{i=1}^{m_{k+1}} \alpha_i \tilde{z}_i(k+1) \quad (4.24)$$

mit geeigneten Gewichten α_i , die sich auf Eins summieren. Die Bestimmung der Gewichte α_i kann dabei auf unterschiedliche Arten geschehen. So kann sie beispielsweise an den PDAF Algorithmus (Bar-Shalom und Birmiwal, 1983; Bar-Shalom und Jaffer, 1972; Bar-Shalom und Tse, 1975) angelehnt werden.

Die zweite Möglichkeit besteht darin, einen möglichst geeigneten Messwert aus der Menge $\{\tilde{z}_i(k+1)\}_{i=1}^{m_{k+1}}$ auszuwählen, der am besten zur Vorhersage $y(k+1|k)$ passt. Ein Beispiel für einen solchen Algorithmus ist der Nächster–Nachbar oder Nearest–Neighbour Algorithmus.

Eine völlig andere und wichtige Möglichkeit besteht schließlich darin, für jeden Messwert $\tilde{z}_i(k+1)$ eine eigene Kalman–Filterung durchzuführen und damit eine separate Hypothese zu berechnen. In diesem Fall erhält man bei Benutzung z.B. des Viterbi–Algorithmus zum Zeitpunkt $k+1$ m_{k+1} Schätzungen des inneren Zustandes.

Mit Hilfe des nun vorliegenden repräsentativen Wertes $z(k+1)$ wird dann der Filterschritt des Kalman–Filters durchgeführt. Dazu werden zunächst die Innovation $\nu(k+1)$ und die Kalman–Gain $K(k+1)$ nach den Formeln 3.14 und 3.15 berechnet gemäß

$$\nu(k+1) = z(k+1) - y(k+1|k) \quad (4.25)$$

und

$$K(k+1) = P(k+1|k)B^T[S(k+1)]^{-1}. \quad (4.26)$$

Danach werden Erwartungswert und Kovarianzmatrix der Filterung gemäß den Update-Gleichungen 3.16 und 3.17 bestimmt zu

$$x(k+1|k+1) = x(k+1|k) + K(k+1)\nu(k+1) \quad (4.27)$$

und

$$P(k+1|k+1) = (I - K(k+1)B)P(k+1|k). \quad (4.28)$$

Damit ist die Beschreibung des Rekursionsschrittes vollständig.

Mit Hilfe des Kalman–Glätters (Shumway und Stoffer, 2000) kann bei Vorliegen der Messungen bis zum Zeitpunkt T noch eine verbesserte Schätzung $x(k|T)$ für weiter zurückliegende Zeitpunkte $k < T$ berechnet werden. Es ist dabei nach den Gleichungen 3.18 und 3.20

$$x(k|T) = x(k|k) + J(k)(x(k+1|T) - x(k+1|k)) \quad (4.29)$$

mit

$$J(k) = P(k|k)A^\top [P(k+1|k)]^{-1}. \quad (4.30)$$

Initialisiert wird diese Rückwärtsrekursion mit der Filterung $x(T|T)$, die der Schätzung des Kalman–Filters für den letzten Zeitpunkt entspricht.

Im nächsten Abschnitt werden nun die speziellen Details der einzelnen Tracking–Algorithmen beschrieben.

4.4 Einfache Verfahren zur Verfolgung

Wie schon in der Einleitung erwähnt, bestehen zwischen den Problemen der Verfolgung punktförmiger Ziele bei Vorliegen von Falschmeldungen im Bereich der Luftüberwachung und ausgedehnter Ziele in der mobilen Robotik strukturelle Gemeinsamkeiten. In beiden Fällen werden in der Regel mehrere Messungen in der Nähe der vorhergesagten Position eines Zieles beobachtet, die mit diesem assoziiert werden müssen. Deshalb lassen sich zahlreiche Verfahren aus der Luftüberwachung unter geeigneten Modifikationen auch in der mobilen Robotik einsetzen. Daneben werden im Folgenden auch einige Ansätze vorgestellt, die in der Luftüberwachung kein Analogon besitzen und nur in der Robotik Sinn machen, wie zum Beispiel der EWA Algorithmus.

Gängige Verfahren zur Verfolgung ausgedehnter Ziele mit Laserscannern gehen so vor, dass sie zunächst aus den Abstandsmessungen sogenannte Features in Form von lokalen Abstandsminima bestimmen (Bellotto und Hu, 2006; Bellotto und Hu, 2007b; Schulz et al., 2001a; Schulz et al., 2001b). Im Rahmen dieser Arbeit wird dagegen auf die nativen Lasermessungen zurückgegriffen, so dass für die weiteren Verarbeitungsschritte der einzelnen Algorithmen mehr Information zur Verfügung steht. Dies kann die Stabilität der Verfahren wesentlich erhöhen, zum Beispiel wenn Personen miteinander interagieren. Wie diese nativen Messungen im Einzelnen verwendet werden können, wird nun beschrieben.

4.4.1 Verfahren mit Mittelwertbildung

Die einzelnen Ansätze unterscheiden sich vor allem darin, wie der repräsentative Wert für die Update-Gleichung des Kalman-Filters aus allen Messungen im Erwartungsgebiet gewonnen wird. Bei den Algorithmen mit Mittelwertbildung wird dieser repräsentative Wert $z(k+1)$ aus den herausgefilterten Messungen $\{\tilde{z}_i(k+1)\}_{i=1}^{m_{k+1}}$ berechnet nach der Formel

$$z(k+1) = \sum_{i=1}^{m_{k+1}} \alpha_i \tilde{z}_i(k+1). \quad (4.31)$$

Für die Berechnung der Gewichte werden hier zwei Möglichkeiten vorgestellt. Zum einen können alle Gewichte gleich gewählt werden, d.h. es soll gelten $\alpha_i = c$ für alle i . Da die Summe aller Gewichte aus Normierungsgründen gleich Eins sein muss, folgt daraus

$$\alpha_i = \frac{1}{m_{k+1}}. \quad (4.32)$$

Es wird also ein ungewichtetes Mittel berechnet und der dadurch definierte Tracking-Algorithmus wird als EWA Algorithmus (Equal Weights Algorithmus) bezeichnet.

Daneben können die Gewichte wie schon erwähnt auch berechnet werden nach dem PDAF (Probabilistic Data Association Filter) Algorithmus (Bar-Shalom und Birmiwal, 1983; Bar-Shalom und Fortmann, 1988; Bar-Shalom und Jaffer, 1972; Bar-Shalom und Tse, 1975) für punktförmige Ziele mit Störmeldungen. Es gilt also mit den Innovationen

$$\nu_i(k+1) = \tilde{z}_i(k+1) - y(k+1|k) \quad (4.33)$$

und der Innovationskovarianz $S(k+1)$

$$\alpha_i = \frac{e^{-\frac{1}{2}[\nu_i(k+1)]^T [S(k+1)]^{-1} \nu_i(k+1)}}{\sum_{j=1}^{m_{k+1}} e^{-\frac{1}{2}[\nu_j(k+1)]^T [S(k+1)]^{-1} \nu_j(k+1)}}. \quad (4.34)$$

Das dadurch festgelegte Verfahren wird als WMA Algorithmus (Weighted Mean Algorithmus) bezeichnet, da ein gewichtetes Mittel verwendet wird. Der PDAF Algorithmus sieht zusätzlich vor, dass die Kovarianzmatrix $P(k+1|k+1)$ jeweils um einen Zusatzterm korrigiert wird, der aus der Streuung der Messwerte $\tilde{z}_i(k+1)$ herrührt und die zusätzliche Unsicherheit in der Schätzung des inneren Zustandes abbildet, die daraus resultiert, dass nicht bekannt ist, welche der Messungen $\tilde{z}_i(k+1)$ tatsächlich diejenige ist, die vom Ziel verursacht wird. Beim WMA Algorithmus wird auf einen solchen Zusatzterm verzichtet, da in diesem Fall die hier durch die Ausdehnung des Objektes hervorgerufene Streuung der Messwerte schon bei der Berechnung der Innovationskovarianz durch die Addition der Matrix E berücksichtigt wurde.

Schließlich soll an dieser Stelle noch ein sehr einfacher Algorithmus vorgestellt werden, der die Schätzung des Zustandes direkt aus den Messungen vornimmt, d.h. es gilt

$$x(k+1|k+1) = \frac{\sum_{i=1}^{m_{k+1}} \tilde{z}_i(k+1)}{m_{k+1}}. \quad (4.35)$$

Allerdings besteht in diesem Fall die Schätzung $x(k+1|k+1)$ für den Zustandsvektor lediglich aus den kartesischen Koordinaten des Objektes. Der zugehörige Algorithmus wird als RDA Algorithmus (Raw Data Algorithmus) bezeichnet.

4.4.2 Verfahren ohne Mittelwertbildung

Die einfachste Möglichkeit, auf die Berechnung eines Mittelwertes zu verzichten, besteht darin, aus der Menge $\{\tilde{z}_i(k+1)\}_{i=1}^{m_{k+1}}$ der herausgefilterten Messungen den am besten geeigneten Wert als repräsentativen Wert auszuwählen. Dadurch erhält man einen Algorithmus, der dem Nächster-Nachbar Algorithmus im Falle punktförmiger Ziele entspricht und deshalb als NNA Algorithmus (Nearest Neighbour Algorithmus) bezeichnet wird. Die Auswahl des besten Wertes geschieht dabei mit Hilfe der Mahalanobis-Distanz. Es wird also diejenige Messung $\tilde{z}_j(k+1)$ ausgewählt, für die der Term

$$\lambda = [\tilde{z}_i(k+1) - y(k+1|k)]^\top [S(k+1)]^{-1} (\tilde{z}_i(k+1) - y(k+1|k)) \quad (4.36)$$

den minimalen Wert annimmt.

Im Falle eines ausgedehnten Zieles besteht ein großer Nachteil dieser Methode darin, dass jeweils nur eine Schätzung zur weiteren Verarbeitung im nächsten Rekursionsschritt zur Verfügung steht. Deshalb kann es vorkommen, dass diejenige vorhergehende Messung, die am besten zu den neuen Messungen passt, zur weiteren Verarbeitung, zum Beispiel zur Berechnung der entsprechenden Vorhersage, gar nicht mehr zur Verfügung steht. Diese Einschränkung besteht bei Verwendung des Viterbi-Algorithmus nicht. Der Viterbi-Algorithmus wurde ursprünglich in (Viterbi, 1967) eingeführt. Eine gute Beschreibung findet sich in (Forney Jr., 1973). Seine Verwendung zur Verfolgung punktförmiger Ziele bei Vorliegen von Störmeldungen im Bereich der Luftüberwachung mit Radar wurde in (Quach und Farooq, 1994) vorgeschlagen. Die Verfolgung ausgedehnter Ziele mit Hilfe dieses Ansatzes erfolgte erstmals in (Kräußling et al., 2004a).

Der Viterbi-Algorithmus berechnet für jede Messung, die mit dem Ziel assoziiert wurde, eine eigene Positionsschätzung und im nächsten Schritt ein separates Erwartungsgebiet. Dieser Umstand ist das wesentliche Kennzeichen dieses Ansatzes. Er benutzt dabei einen gerichteten Graphen zur Bestimmung der optimalen Folgen der Messungen für eine gegebene Menge von herausgefilterten Messwerten. Diese herausgefilterten Messungen entsprechen dabei den Knoten des Graphen.

Die herausgefilterten Messwerte zum Zeitpunkt $k+1$ werden dabei aus den herausgefilterten Messungen zum Zeitpunkt k abweichend von dem in Kapitel 4.3 vorgestellten Verfahren wie folgt konstruiert: für jede herausgefilterte Messung $\tilde{z}_i(k)$ und zugehörige Schätzung $x_i(k|k)$ zum Zeitpunkt k wird das entsprechende Erwartungsgebiet auf die Messungen zum Zeitpunkt $k+1$ angewandt. In Abbildung 4.4 ist dies für die Messung $\tilde{z}_1(k)$ und die dazugehörige Vorhersage $y_1(k+1|k)$ gezeigt. In diesem Fall liegt $z_1(k+1) = P_1$ innerhalb des Erwartungsgebietes und $z_2(k+1) = P_2$ außerhalb.

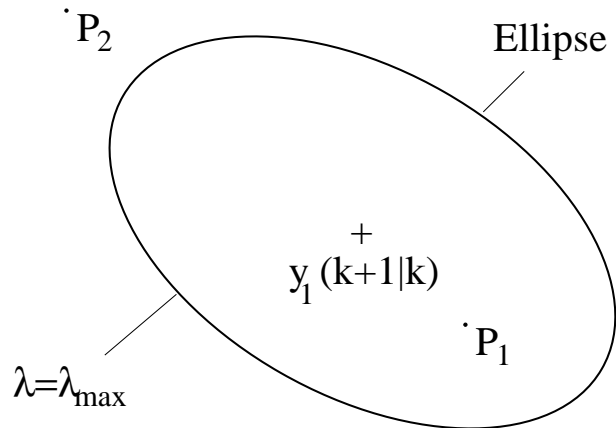


Abbildung 4.4: Anwendung des Erwartungsgebietes für die Vorhersage $y_1(k+1|k)$.

Als Ergebnis erhält man für jedes $i \in \{1, \dots, m_k\}$ eine Menge $Z_i(k+1)$ von Messungen, die innerhalb des zugehörigen Erwartungsgebietes liegen. Im Beispiel ist $Z_1(k+1) = \{z_1(k+1)\}$. Die Menge Z_{k+1} aller herausgefilterten Messwerte zum Zeitpunkt $k+1$, die mit dem Ziel assoziiert werden, ist dann gerade die Vereinigung dieser Mengen, d.h. es ist

$$Z_{k+1} = \cup_{i=1}^{m_k} Z_i(k+1). \quad (4.37)$$

Es handelt sich also gerade um die Menge aller Messungen, die in mindestens einem Erwartungsgebiet liegen. Im Beispiel der Abbildung 4.5 mit zwei Erwartungsgebieten ist $Z_1(k+1) = \{z_1(k+1) = P_1\}$ und $Z_2(k+1) = \{z_2(k+1) = P_2\}$ und damit $Z_{k+1} = \{z_1(k+1), z_2(k+1)\}$. Die Messung $z_3(k+1) = P_3$ liegt in keinem Erwartungsgebiet und wird deshalb nicht weiter berücksichtigt.

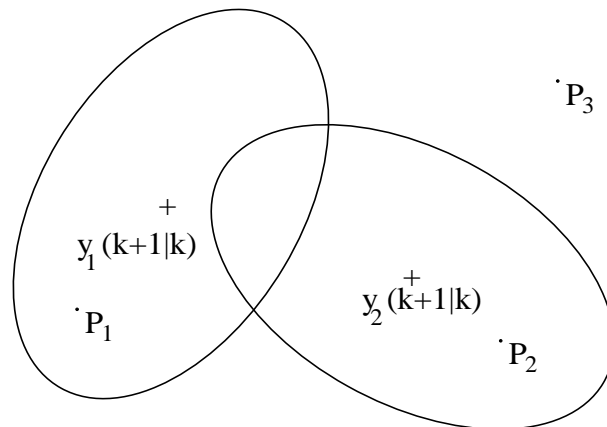


Abbildung 4.5: Bestimmung der Menge Z_{k+1} .

Im nächsten Schritt wird für jede dieser ausgewählten Messungen $\tilde{z}_j(k+1) \in Z_{k+1}$

der sogenannte Vorgänger bestimmt, dem die aktuelle Messung dann zugeordnet wird. Dieser entspricht einer Messung zum vorhergehenden Zeitpunkt k . Dies geschieht mit Hilfe des schon erwähnten gerichteten Graphen. Für diesen Zweck werden nur dieje-

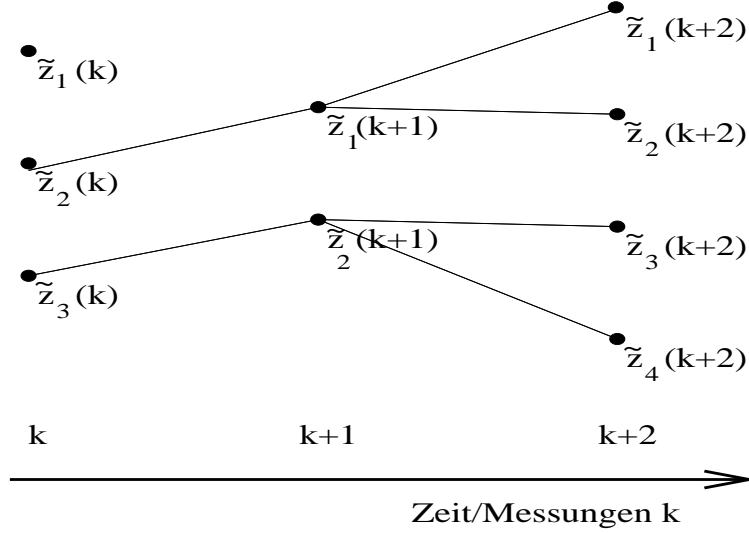


Abbildung 4.6: Bestimmung der Vorgänger mittels eines gerichteten Graphen.

nigen herausgefilterten Messungen des Zeitpunktes k betrachtet, innerhalb deren Erwartungsgebiet die Messung $\tilde{z}_j(k+1)$ liegt (Pulford und La Scala, 1995). Für jede dieser Messungen $\tilde{z}_i(k)$ wird die Länge $d_{i,j}(k+1)$ des Pfades von der Initialisierung $x(0|0)$ nach $\tilde{z}_j(k+1)$ durch $\tilde{z}_i(k)$ berechnet nach der Formel

$$d_{i,j}(k+1) = d_i(k) + a_{i,j}(k+1). \quad (4.38)$$

Dabei ist $d_i(k)$ die Länge des Pfades, der in $x(0|0)$ beginnt und im Knoten $\tilde{z}_i(k)$ endet. $d_i(k)$ wird wie weiter unten im Text beschrieben rekursiv definiert. $a_{i,j}(k+1)$ ist die Distanz zwischen den Knoten $\tilde{z}_i(k)$ und $\tilde{z}_j(k+1)$ und definiert durch

$$a_{i,j}(k+1) = \frac{1}{2} [\nu_{i,j}(k+1)]^\top [S(k+1)]^{-1} \nu_{i,j}(k+1) + \ln \left(\sqrt{|2\pi S(k+1)|} \right). \quad (4.39)$$

Dabei ist die verwendete Innovation $\nu_{i,j}(k+1)$ festgelegt zu

$$\nu_{i,j}(k+1) = \tilde{z}_j(k+1) - y_i(k+1|k) \quad (4.40)$$

und $S(k+1)$ ist wieder die Innovationskovarianz. Außerdem ist $y_i(k+1|k)$ die Vorhersage für die Messungen zum Zeitpunkt $k+1$ basierend auf den Knoten $\{\tilde{z}_{t(l,i)}(l)\}_{l=1}^k$, die zu dem Pfad gehören, der im Knoten $\tilde{z}_i(k)$ endet. Diese Knoten werden dabei über eine Rückwärtsrekursion über die jeweiligen Vorgänger definiert. Die Folge

$$\tilde{Z}_i(k) = \{\tilde{z}_{t(l,i)}(l)\}_{l=1}^k \quad (4.41)$$

heißt auch die zu dem Knoten $\tilde{z}_i(k)$ gehörige Verfolgungsgeschichte oder Messhistorie. Diese Verfolgungsgeschichten definieren eine Zuordnung, die jedem

$$i \in \{1, \dots, m_k\}$$

und jedem $l \leq k$ einen Knoten oder eine Messung $\tilde{z}_{i(l,i)}(l)$ zuweist. Dabei kann eine vergangene Messung mehreren Verfolgungsgeschichten zugeteilt sein und nicht jede alte Messung muss einer aktuellen Messhistorie angehören.

Der Vorgänger des Knotens $\tilde{z}_j(k+1)$ ist nun gerade der Knoten $\tilde{z}_i(k)$, für den die entsprechende Pfadlänge $d_{i,j}(k+1)$ minimal wird. Dabei werden wie schon angedeutet nur diejenigen Knoten $\tilde{z}_i(k)$ berücksichtigt, in deren Erwartungsgebiet $\tilde{z}_j(k+1)$ liegt. Für diesen Vorgänger wird dann $\tilde{z}_{i(k,j)}(k)$ mit zugehörigem Index $i(k,j)$ geschrieben. Außerdem wird die Länge des Pfades, der in $\tilde{z}_j(k+1)$ endet, zu $d_j(k+1) = d_{i(k,j),j}(k+1)$ definiert. In Abbildung 4.6 ist beispielsweise die Messung $\tilde{z}_1(k+1)$ der Vorgänger der Messungen $\tilde{z}_1(k+2)$ und $\tilde{z}_2(k+2)$, während der Knoten $\tilde{z}_1(k)$ kein Vorgänger einer Messung zum Zeitpunkt $k+1$ ist.

Man kann zeigen, dass im Falle punktförmiger Ziele in Clutter die Bestimmung der Verfolgungsgeschichte mit der minimalen Pfadlänge gerade der Bestimmung der Verfolgungsgeschichte mit der maximalen Likelihood entspricht (Kräußling et al., 2005b). Man spricht deshalb dort auch von einer Maximum-a-posteriori (MAP) Schätzung. Im Falle ausgedehnter Ziele trifft dies allerdings nicht zu, da hier die Distanzen zwischen den einzelnen Knoten in erheblichem Maße davon herrühren, dass das Objekt durch den Laserstrahl jeweils an verschiedenen Stellen seiner Oberfläche getroffen wird.

Mit der dem Vorgänger $\tilde{z}_{i(k,j)}(k)$ zugehörigen Schätzung $x_{i(k,j)}(k|k)$ und den entsprechenden Vorhersagen $x_{i(k,j)}(k+1|k)$ und $y_{i(k,j)}(k+1|k)$ kann dann wie in Abschnitt 4.3 beschrieben basierend auf der Messung $\tilde{z}_j(k+1)$ mit der Update-Gleichung eine neue Schätzung $x_j(k+1|k+1)$ berechnet werden. Letzten Endes wird also für jede Messung zum Zeitpunkt $k+1$ rekursiv eine Verfolgungsgeschichte definiert und mit den zugehörigen Messungen der Tracking-Algorithmus aus Kapitel 4.3 durchlaufen. Als Ergebnis hat man zum Zeitpunkt $k+1$ m_{k+1} Schätzungen $x_j(k+1|k+1)$ mit $j = 1, \dots, m_{k+1}$ vorliegen. Diese Schätzungen können nun auf zwei verschiedene Weisen zur Beschreibung des Zustandes des Objektes herangezogen werden.

Zum einen kann nur die aktuelle Schätzung ausgewählt werden, für die die Länge des zugehörigen Pfades minimal wird. Dabei werden die der Messhistorie zugehörigen Schätzungen als Schätzungen für den Zustand und damit die Position des Objektes im Zeitverlauf herangezogen. Das zugehörige Verfahren wird als VBA Algorithmus (Viterbi-Basierter Algorithmus) bezeichnet und entspricht der Vorgehensweise im Falle punktförmiger Ziele in Clutter. Dabei wird aber die Information, die die anderen Schätzungen enthalten, nicht weiter verwendet. Weil aber jede Schätzung einer Messung von einem Punkt auf der Oberfläche des Objektes bedingt durch die Wechselwirkung mit dem Laserstrahl entspricht, sind eigentlich alle Schätzungen gleichwertig. Außerdem ist das Verfahren nicht online-fähig, da ja die Verfolgungsgeschichte mit dem kürzesten Pfad erst am Ende des Datensatzes ausgewählt werden kann.

Zum anderen kann ein ungewichtetes Mittel aller Schätzungen verwendet werden, d.h. es werden auch tatsächlich alle Schätzungen genutzt. Dazu wird

$$x(k+1|k+1) = \frac{\sum_{j=1}^{m_{k+1}} x_j(k+1|k+1)}{m_{k+1}} \quad (4.42)$$

gesetzt. Das dadurch festgelegte Verfahren wird als VAA Algorithmus (Viterbi Average Algorithmus) bezeichnet.

Wenn a-priori-Information vorliegt, zum Beispiel wenn man die Form und die Ausdehnung eines Objektes kennt, kann man auch versuchen, die Schätzungen und die Oberfläche des Zieles zur Kongruenz zu bringen. Wenn dies gelingt, erhält man sofort eine Schätzung für die Position und die Orientierung des Zieles. Mit der Schätzung der Position und der Orientierung bekannter Ziele haben sich auch schon andere Autoren beschäftigt (Dezert, 1998). Wenn die Form aller möglichen Objekte bekannt ist, können mit Hilfe dieser Information auch Tracks den Objekten aus der Datenbank zugeordnet werden oder in der mobilen Robotik zwischen Robotern und Menschen unterschieden werden.

Bei der Verfolgung ausgedehnter Ziele können auch noch zwei Modifikationen vorgenommen werden. Dies beruht auf der Tatsache, dass im Falle ausgedehnter Ziele im Kontext der mobilen Robotik in den meisten Fällen alle Messungen, die in der Menge Z_k liegen, auch vom Ziel selbst stammen. Ausnahmen können lediglich dann auftreten, wenn das Ziel sich einem Hindernis, einer Wand oder einem anderen Objekt nähert. Bei punktförmigen Zielen bei Vorliegen von Störmeldungen im Umfeld von Radaranwendungen stammt dagegen höchstens eine Meldung vom Ziel. Deshalb wählt man dort diejenige Verfolgungsgeschichte aus, die dem Pfad mit der kürzesten Länge entspricht, da diese Messhistorie, wie eben schon angedeutet, am wahrscheinlichsten diejenige ist, die tatsächlich vom Ziel herrührt. Die berechneten Abstände zwischen den Knoten von Störmeldungen entstehen dabei nämlich dadurch, dass die Störmeldungen zufällig über das Überwachungsgebiet verteilt auftreten, so dass die Abstände zwischen diesen Knoten in der Regel größer sind als Abstände zwischen Knoten, die tatsächlich vom Ziel stammen, das ja einer kontinuierlichen Bewegung folgt. Da bei ausgedehnten Zielen dagegen in der Regel alle ausgewählten Messungen vom Objekt selbst stammen, rühren die Abstände zwischen den Knoten dort in erster Linie davon her, dass das Objekt, das im Falle eines Menschen zusätzlich noch seine Form verändern kann, in der Regel immer an unterschiedlichen Stellen von den Laserstrahlen getroffen wird. Es macht deshalb im Falle ausgedehnter Ziele wenig Sinn, zu versuchen, aus der Länge der Pfade Information zu extrahieren. Deshalb sind die folgenden zwei Modifikationen sinnvoll:

- als Vorgänger wird diejenige Messung ausgewählt, die die Mahalanobis-Distanz

$$[\nu_{i,j}(k+1)]^\top [S(k+1)]^{-1} \nu_{i,j}(k+1) \quad (4.43)$$

minimiert. Dabei ist

$$\nu_{i,j}(k+1) = \tilde{z}_j(k+1) - y_i(k+1|k) \quad (4.44)$$

wieder die Innovation zwischen zugehöriger Vorhersage und aktueller Messung und $S(k+1)$ die Innovationskovarianz. Dieses Vorgehen ist ähnlich zu dem bei einem Nearest–Neighbour–Ansatz. Eine Besonderheit des vorgestellten Modells liegt dabei darin, dass die x - und y -Koordinaten gleichberechtigt sind. Vergleiche zum Beispiel Gleichung 4.14. Die Äquipotentiallinien der Innovationskovarianzen $S(k)$ sind deshalb nicht nur Ellipsen, sondern sogar Kreise. Der Vorgänger kann deshalb auch einfach durch Minimierung der Euklidischen Distanz

$$[\nu_{i,j}(k+1)]^\top \nu_{i,j}(k+1) \quad (4.45)$$

zwischen Vorhersage und Messung festgelegt werden.

- falls aus allen Positionsschätzungen am Ende eine Einzige ausgewählt werden soll, kann diese auch zufällig aus allen Schätzungen bestimmt werden. Man kann zum Beispiel einfach die Schätzung mit dem Index Eins benutzen.

Der dem VBA entsprechende Algorithmus wird dann als MVA Algorithmus (Modified Viterbi Algorithmus) bezeichnet. Das dem VAA Algorithmus analoge Verfahren wird im Folgenden dagegen MVAA Algorithmus genannt (Modified Viterbi Average Algorithmus). Ein Vorteil dieses Vorgehens besteht dabei darin, dass die Länge der Pfade weder berechnet noch gespeichert werden muss. Es wird sich aber auch zeigen, dass der MVAA Algorithmus tatsächlich eine präzisere Positionsschätzung liefert als der VAA Algorithmus.

Zum Abschluss werden noch einmal die einzelnen Schritte und insbesondere der komplizierte Gating–Prozess der Verfahren, die den Viterbi–Algorithmus benutzen, zusammengefasst:

1. Gegeben sind m_k Positionsschätzungen $\{x_i(k|k)\}_{i=1}^{m_k}$ zum Zeitpunkt k .
2. Berechne m_k Vorhersagen $\{y_i(k+1|k)\}_{i=1}^{m_k}$ für die Messungen zum Zeitpunkt $k+1$ gemäß den Gleichungen 4.17 und 4.19.
3. Berechne um diese Vorhersagen m_k Erwartungsgebiete.
4. Bestimme zu jedem dieser Erwartungsgebiete die Menge aller neuen Messungen, die in ihm enthalten sind. Die Menge $\{\tilde{z}_j(k+1)\}_{j=1}^{m_{k+1}}$ der m_{k+1} Messwerte des Zeitpunktes $k+1$, die dem Ziel zugeordnet und weiter betrachtet werden, ist dann gerade die Vereinigung dieser Mengen.
5. Bestimme zu jeder dieser neuen Messungen den Vorgänger $\tilde{z}_{i(k,j)}(k)$. Dabei kommen als Vorgänger nur diejenigen alten Messwerte in Frage, in deren Erwartungsgebiet die gerade betrachtete Messung liegt.
6. Berechne für jede dieser Messungen eine eigene Positionsschätzung $x_j(k+1|k+1)$ aus den Vorhersagen des Vorgängers für den Zustand und die Messung und der neuen Messung gemäß den Gleichungen 4.25 und 4.27.

Tabelle 4.1: Übersicht über die Algorithmen.

Bezeichnung des Algorithmus	Abkürzung	Mittelwertbildung
Equal Weights Algorithmus	EWA	ja
Weighted Mean Algorithmus	WMA	ja
Raw Data Algorithmus	RDA	ja
Nearest Neighbour Algorithmus	NNA	nein
Viterbi Basierter Algorithmus	VBA	nein
Viterbi Average Algorithmus	VAA	ja
Modified Viterbi Algorithmus	MVA	nein
Modified Viterbi Average Algorithmus	MVAA	ja

Der Hauptvorteil dieser komplizierten Routine liegt darin, dass die entsprechenden Verfahren in gewisser Weise multimodale Verteilungen wiedergeben können. Dies bedeutet, dass mehrere Orte gleichzeitig angeboten werden, wo sich Ziele aufhalten können, was einen großen Vorteil darstellt, wenn mehrere interagierende Objekte untersucht werden. Im Wesentlichen beruht dieses Verhalten darauf, dass stets eine ausreichende Anzahl von Erwartungsgebieten berechnet wird, um so alle Messungen von allen Zielen erfassen zu können. Vergleiche hierzu auch Kapitel 6.

Am Ende dieses Kapitels gibt Tabelle 4.1 noch einmal eine Übersicht über alle eingeführten Verfahren.

Kapitel 5

Die Verfolgung eines kreisförmigen Objektes

Die in diesem Kapitel vorgestellte analytische Behandlung der Verfolgung eines kreisförmigen Objektes und der darauf basierende Verfahrenvergleich wurden in (Kräußling, 2006a; Kräußling, 2008b) publiziert.

5.1 Motivation

Ziel dieses Kapitels ist eine Bewertung der im letzten Kapitel vorgestellten einfachen Tracking-Verfahren bezüglich des Informationsgehaltes über die wahre Position eines getrackten Objektes und der dabei benötigten Rechenkapazität. Der Vergleich der Güte verschiedener Verfahren, wie zum Beispiel zur Lokalisierung, ist ein gut eingeführter Ansatz in der mobilen Robotik (Gutmann et al., 1998; Gutmann und Fox, 2002). Für eine solche Evaluierung ist die Einführung eines exemplarischen Problems, anhand dessen der Gütevergleich durchgeführt werden soll, erforderlich. Bei diesem Problem soll es sich hier um die Verfolgung eines kreisförmigen Objektes mit vorgegebenem Radius handeln, das sich auf einer Kreisbahn um den Beobachter bewegt. In der mobilen Robotik gibt es nämlich in erster Linie zwei Klassen von Zielen, und zwar andere Roboter oder Menschen. Da viele Service-Roboter, wie sie in der Forschung eingesetzt werden, wie zum Beispiel der B21, aber kreisförmig sind, sind diese Objekte auch von erheblicher praktischer Relevanz. Da kreisförmige Ziele außerdem auch einer analytischen Behandlung zugänglich sind und darüberhinaus leicht Simulationsdaten erzeugt werden können, werden die Betrachtungen in diesem Kapitel im Wesentlichen auf solche Objekte eingeschränkt. Darüberhinaus wird kurz noch auf das Problem des People-Tracking eingegangen.

Um die Güte der im vorhergehenden Kapitel vorgestellten Algorithmen zu vergleichen, ist es zunächst notwendig, sich zu überlegen, welchen Punkt eines ausgedehnten Objektes ein Algorithmus als Schätzung der Position dieses Objektes liefern kann. Zunächst könnte man vermuten, dass dieser Punkt der Schwerpunkt S des Zie-

les, bezogen auf die Höhe des Laserstrahls und berechnet aus allen Punkten in der zugehörigen Ebene, ist, und somit die Distanz der Positionsschätzung zu diesem Schwerpunkt ein valides Kriterium für die Güte eines Algorithmus sein sollte. Dabei tritt aber das Problem auf, dass ein Tracking-Algorithmus nur diejenigen Messpunkte zur Schätzung eines Schwerpunktes verwenden kann, die ihm der Lasersensor oder allgemeiner das Messgerät auch tatsächlich liefert. Der Lasersensor wechselwirkt aber nur mit der Oberfläche des Objektes und hier auch nur mit dem für ihn sichtbaren Bereich. Man erhält somit nur Information über die Oberfläche im sichtbaren Bereich. In die Berechnung des gerade eingeführten Schwerpunktes geht aber auch der nicht sichtbare Bereich ein, der sich jedoch für zwei Objekte mit identischem sichtbarem Bereich sehr stark unterscheiden kann (Abbildung 5.1). Ein Tracking-Algorithmus kann aber für zwei solche Ziele nur identische Schätzungen liefern, da er vom Sensor für beide Objekte die gleichen Messwerte erhält. Deshalb sollte für diese beiden Objekte auch der Punkt, den der Algorithmus schätzen soll, identisch sein. Dies erreicht man, indem man für diesen Punkt den Schwerpunkt, berechnet lediglich aus der Oberfläche im sichtbaren Bereich, wählt. Dieser Punkt wird im Folgenden mit S_O bezeichnet (O für Oberfläche). Er ist identisch mit dem Erwartungswert für den Messwert, den der Sensor vom Objekt misst.

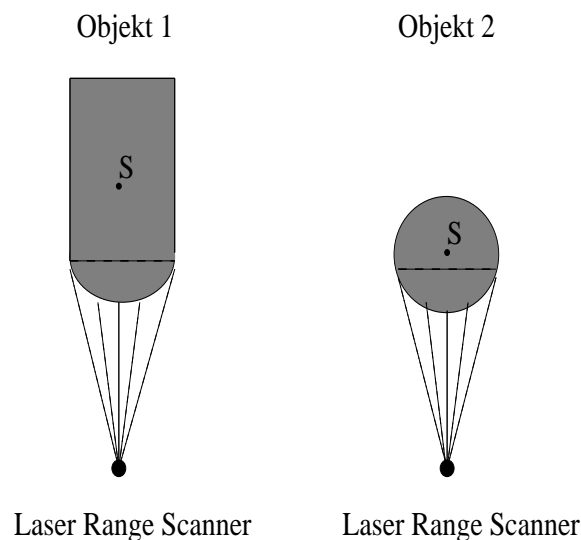


Abbildung 5.1: Zwei Objekte mit verschiedenen Schwerpunkten S , aber identischem sichtbarem Bereich.

Dieser Sachverhalt soll an dem schon erwähnten Beispiel eines kreisförmigen Objektes mit gegebenem Radius näher erläutert werden, da in diesem Fall auf analytischem Wege die Entfernung dieses Schwerpunktes S_O vom Mittelpunkt des Zieles berechnet werden kann. Die Kenntnis der Position des Mittelpunktes des Objektes liefert hier aber bereits die komplette Information. Die Vermutung, die den folgenden Untersuchungen zugrunde liegt, lautet dahingehend, dass die Algorithmen mit Mittel-

Tabelle 5.1: Grenzwinkel ϕ und Abstand x in Abhängigkeit von der Entfernung d .

d/cm	ϕ /Grad	ϕ /rad	x /cm
100	15.6647	0.2734	23.3965
200	7.7578	0.1354	22.3607
400	3.8732	0.0676	21.7837
600	2.5783	0.0450	21.6103
800	1.9366	0.0338	21.4803

wertbildung diesen Punkt S_O , der in diesem Falle im Inneren des Objektes liegt, mehr oder weniger genau schätzen.

Die Bestimmung des Abstandes x des Schwerpunktes S_O vom Mittelpunkt benutzt Methoden aus der Analysis und der ebenen Euklidischen Geometrie und ist in Anhang B dargestellt. Als Endergebnis erhält man:

$$x = \frac{d}{2} + \frac{r}{2d \arcsin \frac{r}{d}} \left(\frac{\pi r}{2} - \sqrt{d^2 - r^2} \right). \quad (5.1)$$

Dabei ist r der Radius des Objektes und d die Entfernung des Mittelpunktes des Objektes von dem als punktförmig angenommenen Sensor.

Tabelle 5.1 zeigt den Effekt des Abstandes d auf den Schwerpunkt oder Erwartungswert für die Werte von d , die im Folgenden verwendet werden. Diese sind im Bereich der typischen Entfernungen zwischen Sensor und Objekt, wie sie in der mobilen Robotik auftreten. Für den Radius r des kreisförmigen Objektes wurde $r = 27$ cm angenommen, ein Wert, der typisch ist für Service-Roboter, wie sie für Forschungszwecke genutzt werden. Neben den Werten für x werden in der Tabelle auch die zugehörigen Grenzwinkel ϕ , die beschreiben, welcher Teil der Oberfläche im sichtbaren Bereich liegt, in Winkelgrad und Bogenmaß wiedergegeben. Zur Definition des Grenzwinkels ϕ vergleiche auch Abbildung B.1. Mit wachsendem Abstand d wird der Grenzwinkel ϕ kleiner und wächst die beobachtbare Oberfläche des Zieles und somit wird auch x kleiner, d.h. der Erwartungswert S_O bewegt sich näher an den Mittelpunkt des Objektes heran. Abbildung 5.2 stellt x noch einmal graphisch in Abhängigkeit von d dar.

Für die Extremwerte von d , d.h. für $d = r$ und $d = \infty$ erhält man:

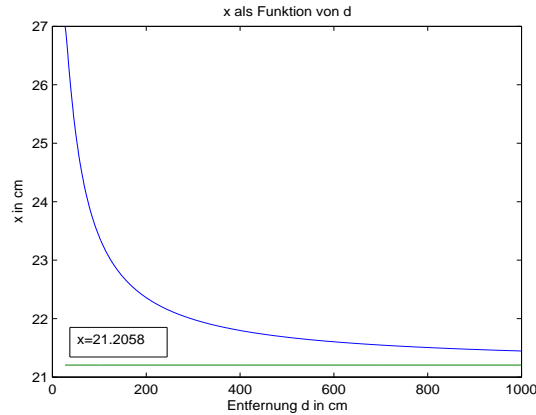
- für den kleinst-möglichen Wert von d , nämlich $d = r$, gilt

$$x = \frac{r}{2} + \frac{r}{2r \arcsin 1} \left(\frac{\pi r}{2} - \sqrt{r^2 - r^2} \right) = \frac{r}{2} + \frac{\pi r}{4 \frac{\pi}{2}} = r. \quad (5.2)$$

Dieser Fall tritt dann ein, wenn der Laserscanner dem Objekt direkt aufsitzt. Man erhält dabei nur Messungen von einem einzigen Punkt.

- für den Grenzwert $\lim_{d \rightarrow \infty} x$ erhält man als Endergebnis (vergleiche Anhang B):

$$x_\infty = \lim_{d \rightarrow \infty} x = \frac{\pi r}{4}. \quad (5.3)$$

Abbildung 5.2: x als Funktion der Entfernung d .

Im Falle von $r = 27$ cm beträgt dieser Grenzwert 21.2058 cm. Dieser ist auch als waagerechte Asymptote in Abbildung 5.2 eingezeichnet. Der Fall $d = \infty$ entspricht einer Beobachtung aus unendlicher Entfernung, wobei gerade die dem Beobachter zugewandte Hälfte des Objektes im sichtbaren Bereich liegt.

Eine Möglichkeit, auf direktem Wege eine Schätzung für den Mittelpunkt eines kreisförmigen Objektes zu erhalten und dabei ebenfalls von Wissen über die Geometrie des Zieles zu profitieren, besteht in der Verwendung der einzelnen Positionsschätzungen, die der VBA oder der MVA liefern. Dies beruht auf dem Umstand, dass diese Positionsschätzungen zu Punkten auf der Oberfläche des kreisförmigen Objektes korrespondieren und der Mittelpunkt eines solchen Objektes bei bekanntem Radius r und Vorgabe zweier Punkte auf der Oberfläche bereits eindeutig festgelegt ist.

Es wird deshalb an dieser Stelle gezeigt, wie in diesem Falle eine Schätzung für den Mittelpunkt berechnet werden kann (Abbildung 5.3). Seien dazu $x_i(k|k)$ und $x_j(k|k)$ ein Paar zweier solcher Schätzungen. Dann gilt zunächst

$$v = \begin{pmatrix} v_x \\ v_y \end{pmatrix} = \frac{x_j(k|k) - x_i(k|k)}{2}. \quad (5.4)$$

Desweiteren ist nach dem Satz des Pythagoras $r^2 = v^2 + w^2$ und damit

$$|w| = \sqrt{r^2 - v^2}. \quad (5.5)$$

Da w senkrecht auf v steht, gibt es zwei Möglichkeiten w_1 und w_2 für w , nämlich

$$w_1 = \frac{|w|}{|v|} \begin{pmatrix} -v_y \\ v_x \end{pmatrix} \quad (5.6)$$

und $w_2 = -w_1$. Somit bestehen für die zugehörige Schätzung $c_{i,j}(k|k)$ des Mittelpunktes ebenfalls zwei Möglichkeiten, nämlich

$$c_{i,j}^1(k|k) = x_i(k|k) + v + w_1 = \frac{x_i(k|k) + x_j(k|k)}{2} + w_1 \quad (5.7)$$

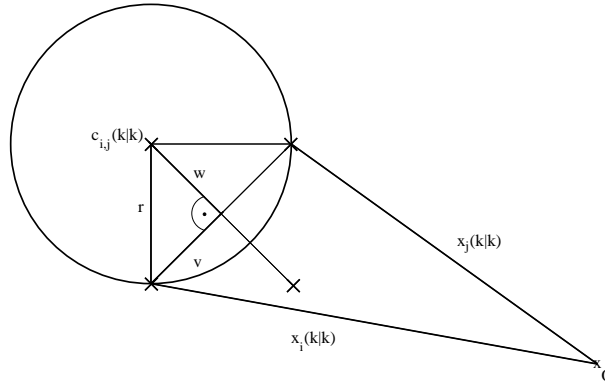


Abbildung 5.3: Berechnung des Mittelpunktes $c_{i,j}(k|k)$ eines kreisförmigen Objektes basierend auf zwei Schätzungen $x_i(k|k)$ und $x_j(k|k)$ des VBA oder des MVA.

und

$$c_{i,j}^2(k|k) = x_i(k|k) + v + w_2 = \frac{x_i(k|k) + x_j(k|k)}{2} - w_1. \quad (5.8)$$

Offensichtlich ist die richtige Schätzung für den Mittelpunkt diejenige, für die der Euklidische Abstand zum Sensor größer ist. Diese Schätzung wird deshalb auch für die weiteren Betrachtungen herangezogen. Falls die Anzahl der Positionsschätzungen $x_i(k|k)$ gleich m_k ist, können diese Berechnungen für alle $\frac{m_k(m_k-1)}{2}$ Paare von Schätzungen $x_i(k|k)$ und $x_j(k|k)$ durchgeführt werden. Danach kann abschließend aus diesen Mittelpunktsschätzungen $c_{i,j}(k|k)$ ein ungewichtetes Mittel $c(k|k)$ bestimmt werden zu

$$c(k|k) = \frac{\sum_{i \neq j} c_{i,j}(k|k)}{\frac{m_k(m_k-1)}{2}}. \quad (5.9)$$

Eine weitere Verbesserung könnte noch in der Verwendung von Hilfspunkten bestehen. Dies ist begründet durch die Tatsache, dass die Punkte auf der Oberfläche des kreisförmigen Objektes, die dem Vorgänger $\tilde{z}_{i(k,j)}(k)$ und der Messung $\tilde{z}_j(k+1)$ entsprechen, in den allermeisten Fällen unterschiedlich sind. Deshalb könnte es sinnvoll sein, anstelle der Vorhersage $y_{i(k,j)}(k+1|k)$ den Punkt auf der Oberfläche des kreisförmigen Objektes zu verwenden, der der Schnittpunkt der Verbindungslinie zwischen einer Vorhersage für den Mittelpunkt und der Messung $\tilde{z}_j(k+1)$ und einem Kreis um die Vorhersage für den Mittelpunkt mit Radius r ist (Abbildung 5.4). Die Vorhersage $c(k+1|k)$ für den Mittelpunkt kann dabei wie folgt berechnet werden. Zunächst wird wie oben eine Schätzung $c(k|k)$ für den Mittelpunkt zum Zeitpunkt k bestimmt. Als nächstes wird eine Schätzung v_{ck} für die Geschwindigkeit des Mittelpunktes zum Zeitpunkt k aus den bekannten Geschwindigkeiten

$$v_{ki} = \begin{pmatrix} \dot{x}_{k1i} \\ \dot{x}_{k2i} \end{pmatrix}, \quad (5.10)$$

die zu den Schätzungen

$$x_i(k|k) = (x_{k1i} \ x_{k2i} \ \dot{x}_{k1i} \ \dot{x}_{k2i})^\top \quad (5.11)$$

gehören, wie folgt berechnet:

$$v_{ck} = \frac{\sum_{i=1}^{m_k} v_{ki}}{m_k}. \quad (5.12)$$

Dabei ist m_k wieder die Anzahl der Positionsschätzungen $x_i(k|k)$ zum Zeitpunkt k . Dann gilt für die Vorhersage $c(k+1|k)$ für die Position des Mittelpunktes zum Zeitpunkt $k+1$

$$c(k+1|k) = c(k|k) + \Delta T v_{ck}. \quad (5.13)$$

Dabei ist ΔT wieder die Zeitspanne zwischen den Messungen zu den Zeitpunkten k und $k+1$. Der zur Messung $\tilde{z}_j(k+1)$ gehörige Hilfspunkt h_j kann dann wie folgt berechnet werden (Abbildung 5.4). Sei

$$w = \tilde{z}_j(k+1) - c(k+1|k). \quad (5.14)$$

Dann ist

$$h_j = c(k+1|k) + \frac{r}{|w|} w. \quad (5.15)$$

Dieser Hilfspunkt h_j wird dann in der Update-Gleichung des Kalman-Filters anstelle der Vorhersage $y_{i(k,j)}(k+1|k)$ verwendet. Analog dazu wird anstelle der Vorhersage

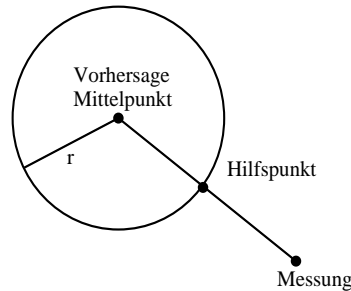


Abbildung 5.4: Konstruktion von Hilfspunkten.

$x_{i(k,j)}(k+1|k)$ für den inneren Zustand der Wert

$$\begin{pmatrix} y_{i(k,j)}(k+1|k) \\ v_{ck} \end{pmatrix} \quad (5.16)$$

eingesetzt.

5.2 Die geeignete Wahl der Modellparameter

Das in Kapitel 4 vorgestellte Modell verwendet drei Parameter, nämlich die beiden Größen Θ und Σ , die die Dynamik der Objekte beschreiben, und den Parameter δ , der die Ausdehnung der Ziele abbildet. Auf empirischem Wege soll nun die Frage geklärt werden, wie die Güte eines Tracking-Algorithmus von der Wahl dieser Parameter abhängt und wie diese gewählt werden sollten, um einen möglichst optimalen

Algorithmus zu erhalten. Dazu wurden Daten verwendet, die mittels eines Simulationsprogrammes (vergleiche Anhang C) gewonnen worden sind. Für den Algorithmus EWA wurden für eine Entfernung von $d = 8$ m und eine feste Standardabweichung σ des Messfehlers die drei Parameter variiert und die jeweilige mittlere Abweichung d_{S_O} der berechneten Positionsschätzungen vom Erwartungswert S_O bestimmt. Die Details dieser Vorgehensweise werden im nächsten Abschnitt geschildert.

Es zeigte sich bei den Untersuchungen, dass die mittlere Abweichung in komplizierter Weise von den Parametern abhängt. Dabei weist diese mittlere Abweichung, falls man die anderen beiden Parameter konstant hält, für einen bestimmten Wert des dritten Parameters ein globales Minimum auf. Als geeignete Werte für Θ und Σ stellten sich $\Theta = 20$ und $\Sigma = 60$ heraus. Diese Werte wurden deshalb auch bei den weiteren Betrachtungen zugrundegelegt. Die Untersuchungen hinsichtlich der optimalen Wahl von δ für diese beiden Werte erbrachten die in Abbildung 5.5 dargestellten Resultate. Die genauere Betrachtung im kritischen Bereich um den Wert $\delta = 750$ führte zu den

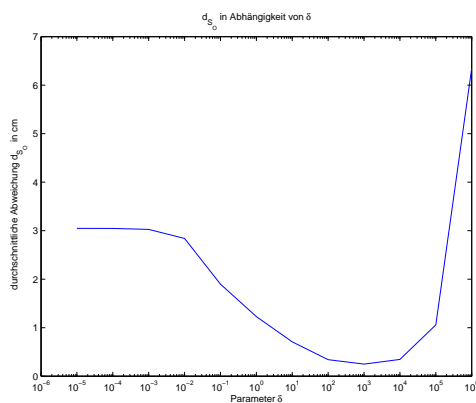


Abbildung 5.5: Mittlere Abweichung d_{S_O} in Abhängigkeit vom Modellparameter δ , einfach logarithmische Übersichtsdarstellung.

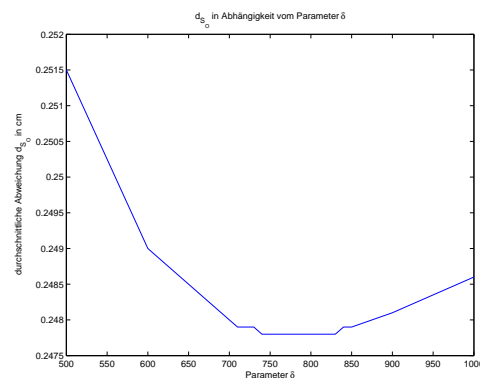


Abbildung 5.6: Mittlere Abweichung d_{S_O} in Abhängigkeit vom Modellparameter δ , feinere Darstellung im Bereich um $\delta = 750$.

Mittlere Abweichung d_{S_O} in Abhängigkeit vom Modellparameter δ .

in Abbildung 5.6 veranschaulichten Ergebnissen.

Für die weiteren Experimente wurde deshalb $\delta = 780$ verwendet. Dieser Wert ist auch physikalisch sinnvoll, da die beweglichen Ziele in der mobilen Robotik oft eine Ausdehnung von ungefähr dreißig Zentimetern besitzen, was einer Varianz von $30^2 = 900 \approx 780$ entspricht. Außerdem war den Simulationen ein Objekt mit Radius $r = 27$ cm zugrundegelegt gewesen.

5.3 Die Information über die Position des Objektes

Im Folgenden wird die Bewegung eines kreisförmigen Objektes um den Sensor auf einer Kreisbahn mit Radius d betrachtet. Der Radius des Objektes selbst wurde dabei wieder mit 27 Zentimetern angenommen, ein Wert, der typisch ist für die Ausdehnung eines mobilen Roboters oder eines Menschen. Um die Güte der Algorithmen anhand dieses Problems zu vergleichen, wurden simulierte Daten verwendet, da die wahre Position des Zieles aufgrund der teilweise sehr geringen Unterschiede zwischen den einzelnen Verfahren mit sehr hoher Genauigkeit bekannt sein musste. Dies ist praktisch nicht mit Daten aus einem realen Experiment zu erreichen. Die Gewinnung von exakten Werten in realen Experimenten ist in der mobilen Robotik generell sehr schwierig. Darauf wurde auch schon von anderen Autoren im Zusammenhang mit anderen Untersuchungsgegenständen hingewiesen (Brooks und Williams, 2003; Zhao und Shibasaki, 2005).

Wie schon bei der Vorstellung der einzelnen Algorithmen in Kapitel 4 dargestellt, gibt es zwei unterschiedliche Klassen von Verfahren, nämlich solche mit und solche ohne Mittelwertbildung. Zur ersten Gattung gehören der EWA, der WMA, der RDA, der VAA und der MVAA, während die zweite Gruppe aus dem NNA, dem VBA und dem MVA besteht. Da im ersten Fall in irgendeiner Form ein Mittel aus allen Messungen weiterverarbeitet wird und der Erwartungswert S_O als Mittel aller Punkte auf der Oberfläche des Zieles im beobachtbaren Bereich berechnet wurde, kann vermutet werden, dass diese Verfahren gerade den Punkt S_O mit unterschiedlicher Genauigkeit schätzen. Diese Fragestellung soll zunächst untersucht werden, wobei sich die Vermutung bestätigen wird. Die Eigenschaften der zweiten Gruppe der Algorithmen werden später analysiert.

Da eine exakte Bewegung auf einem Kreis betrachtet wird, rührt das Prozessrauschen lediglich von der Zentripetalkraft her, die das Objekt auf dem Kreis hält. Die Simulationen wurden für Kreisbahnen mit den unterschiedlichen Radien d aus Tabelle 5.1 ausgeführt. Diese Werte sind im Bereich der typischen Entfernungen zwischen den Objekten und dem Lasersensor, wie sie im Bereich der mobilen Robotik auftreten. Die Werte für die Standardabweichung σ des Messrauschens der simulierten Laserscanner aus Gleichung C.11 betragen 0, 1, 3, 5, 7.5 und 10 cm. Diese Werte sind im typischen Bereich der Fehler der kommerziellen Laserscanner, wie sie üblicherweise in der mobilen Robotik verwendet werden (Sick AG, 2006). Es wurde dabei angenommen, dass lediglich die Entfernungswerte der Messungen fehlerbehaftet sind, während die Winkel exakt bekannt sein sollen. Für jedes Paar von Radius d und Standardabweichung σ wurden zwanzig Durchläufe ausgeführt. Für jeden Zeittakt k wurde dabei die Euklidische Distanz des Ergebnisses des Verfolgungsalgorithmus vom Schwerpunkt S_O berechnet und davon wurde der Durchschnitt dieser Distanzen über den ganzen Versuch ausgewertet. Schließlich wurde aus diesen Durchschnitten der Mittelwert über alle zwanzig Versuche bestimmt. Die Ergebnisse mit Einheit ein Zentimeter sind in den Tabellen 5.2 bis 5.6 wiedergegeben. Die jeweils aus den Mitteln für die zwanzig Versuche berechneten Standardabweichungen sind klein. Sie reichen von ungefähr

Tabelle 5.2: Mittelwert des Abstandes vom Erwartungswert S_O für den EWA in Zentimetern.

d/cm	100	200	400	600	800
$\sigma = 0 \text{ cm}$	0.9689	0.3654	0.2154	0.5597	0.2478
$\sigma = 1 \text{ cm}$	0.9698	0.3654	0.2202	0.5609	0.2575
$\sigma = 3 \text{ cm}$	0.9710	0.3805	0.2962	0.5744	0.3487
$\sigma = 5 \text{ cm}$	0.9853	0.4266	0.4133	0.6531	0.5392
$\sigma = 7.5 \text{ cm}$	1.0497	0.5617	0.5673	0.7845	0.7811
$\sigma = 10 \text{ cm}$	1.1054	0.6317	0.7409	0.9626	1.0119

Tabelle 5.3: Mittelwert des Abstandes vom Erwartungswert S_O für den WMA in Zentimetern.

d/cm	100	200	400	600	800
$\sigma = 0 \text{ cm}$	1.4066	0.8513	0.7847	1.0653	0.8635
$\sigma = 1 \text{ cm}$	1.4048	0.8494	0.7867	1.0655	0.8666
$\sigma = 3 \text{ cm}$	1.3918	0.8391	0.8025	1.0677	0.8569
$\sigma = 5 \text{ cm}$	1.3670	0.8329	0.8124	1.0883	0.9521
$\sigma = 7.5 \text{ cm}$	1.3711	0.8766	0.8968	1.1257	1.0936
$\sigma = 10 \text{ cm}$	1.3533	0.8776	0.9708	1.2484	1.2676

0.01 cm für die kleineren Standardabweichungen des Messrauschens bis zu circa 0.1 cm für die größeren Werte.

Es ist offensichtlich, dass die Ergebnisse aller Verfahren außer des VAA, die einen Mittelwert aus den Messungen benutzen, eine gute Schätzung für den Schwerpunkt liefern, wobei EWA am besten und RDA am schlechtesten abschneiden. Der EWA unterscheidet sich dabei von dem einfachen RDA im Wesentlichen nur dadurch, dass der Kalman-Filter zur Verbesserung der Positionsschätzungen eingesetzt wird. Dies bringt deutliche Vorteile vor allem bei größeren Entfernungen, selbst bei Abwesenheit

Tabelle 5.4: Mittelwert des Abstandes vom Erwartungswert S_O für den RDA in Zentimetern.

d/cm	100	200	400	600	800
$\sigma = 0 \text{ cm}$	0.3173	0.6693	1.3591	2.7099	3.0324
$\sigma = 1 \text{ cm}$	0.3431	0.6961	1.4012	2.7376	3.0804
$\sigma = 3 \text{ cm}$	0.5431	0.8975	1.6576	3.0047	3.3769
$\sigma = 5 \text{ cm}$	0.7792	1.2337	2.0547	3.4259	3.8777
$\sigma = 7.5 \text{ cm}$	1.1518	1.7079	2.6529	4.0475	4.6124
$\sigma = 10 \text{ cm}$	1.5119	2.1838	3.3107	4.7831	5.4437

Tabelle 5.5: Mittelwert des Abstandes vom Erwartungswert S_O für den VAA in Zentimetern.

d/cm	100	200	400	600	800
$\sigma = 0 \text{ cm}$	6.9678	3.0463	3.0767	6.5293	6.4645
$\sigma = 1 \text{ cm}$	11.5044	4.5481	1.7062	5.8976	6.5338
$\sigma = 3 \text{ cm}$	12.2924	5.0634	2.4768	5.2141	7.1276
$\sigma = 5 \text{ cm}$	12.4933	6.6898	3.6525	5.2175	6.7693
$\sigma = 7.5 \text{ cm}$	14.2701	7.6690	4.6502	5.4802	6.7554
$\sigma = 10 \text{ cm}$	14.5553	8.3581	5.7584	6.2627	7.1146

Tabelle 5.6: Mittelwert des Abstandes vom Erwartungswert S_O für den MVAA in Zentimetern.

d/cm	100	200	400	600	800
$\sigma = 0 \text{ cm}$	2.1260	1.2587	1.7045	2.3523	2.6191
$\sigma = 1 \text{ cm}$	1.8316	1.1886	1.6577	2.4278	2.5511
$\sigma = 3 \text{ cm}$	1.5376	1.1199	1.6894	2.6091	2.7851
$\sigma = 5 \text{ cm}$	1.6535	1.3429	1.9600	2.8546	3.1441
$\sigma = 7.5 \text{ cm}$	1.9116	1.7742	2.4516	3.3260	3.6828
$\sigma = 10 \text{ cm}$	2.2212	2.1866	3.0103	3.9349	4.2878

von Messrauschen ($\sigma = 0$). Im letzteren Fall, d.h. bei Vorliegen exakter Messungen, rührt die Abweichung der Schätzungen des RDA vom wahren Wert nur daher, dass der sichtbare Bereich des Objektes lediglich in Ein-Grad-Schritten abgetastet wird und somit keine volle Überdeckung dieses Bereiches, der ja der Berechnung des Schwerpunktes S_O zugrundeliegt, vorliegt. Eine Ausnahme bilden lediglich die Resultate für $d = 100 \text{ cm}$. In diesem Fall liegen sehr viele Messungen und damit eine gute Überdeckung des sichtbaren Bereiches vor, so dass der RDA hier bei den kleineren Standardabweichungen etwas besser abschneidet. Auch die maximalen Abweichungen für einzelne Zeitpunkte sind beim RDA erheblich größer (Kräußling et al., 2005b). Die Verbesserung in der Genauigkeit rechtfertigt somit den geringfügigen Zuwachs in der Komplexität des Verfahrens für den EWA. Vergleicht man den VAA mit dem MVAA, so fällt sofort ins Auge, dass es viel besser ist, den Vorgänger durch Minimierung der Mahalanobisdistanz zu bestimmen wie beim MVAA.

Die hier skizzierten Ergebnisse könnten auch durch einen statistischen Test in Form einer Analysis of Variance (ANOVA) (Backhaus et al., 2000; Montgomery, 2001) untermauert werden. Da die Unterschiede zwischen den einzelnen Algorithmen in der Regel aber erheblich größer sind als die Streuung für ein einzelnes Verfahren, erscheint dies überflüssig. Außerdem würde ein solches Vorgehen den Rahmen dieser Arbeit sprengen. Es soll an dieser Stelle lediglich darauf hingewiesen werden, dass die Genauigkeit der einzelnen Verfahren signifikant von der Entfernung des Sensors vom

Objekt abhängt (Kräußling et al., 2005b).

Nun soll der Frage nachgegangen werden, ob durch eine geeignete Wahl der Gewichte α_i aus Gleichung 4.31 eine Verbesserung gegenüber dem EWA Algorithmus erzielt werden kann. Dazu wird gemäß Gleichung 4.14

$$E = \delta_g \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (5.17)$$

gesetzt und für verschiedene Werte von δ_g der WMA Algorithmus durchlaufen. Dabei wurde der Radius $d = 4$ m und die Standardabweichung $\sigma = 5$ cm zugrundegelegt. Außerdem wurde zur Vereinfachung der Untersuchungen die Innovationskovarianz jeweils gleich E gesetzt. Da bei der Berechnung der Innovationskovarianz nach Gleichung 4.15 der Hauptbeitrag vom Term E herrührt, stellt dies keine qualitative Einschränkung dar.

Um ein Maß für die unterschiedlichen Gewichte in Abhängigkeit vom Parameter δ_g zu erhalten, wurden zwei unterschiedliche Messungen z_1 und z_2 betrachtet. z_1 soll dabei zehn Zentimeter von der Vorhersage $y(k+1|k)$ für die Messung und z_2 zwanzig Zentimeter von dieser Vorhersage entfernt sein, was bei einem kreisförmigen Objekt mit Radius $r = 27$ cm typische Werte sind. Die zugehörigen Innovationen sind dann nach Gleichung 4.25 einfach $\nu_1 = 10$ cm und $\nu_2 = 20$ cm. Für die Verhältnisse q der zugehörigen Gewichte erhält man damit nach Gleichung 4.34

$$q = \frac{\exp(-0.5 \cdot 10 \cdot \delta_g^{-1} \cdot 10)}{\exp(-0.5 \cdot 20 \cdot \delta_g^{-1} \cdot 20)} = \frac{\exp(-50 \cdot \delta_g^{-1})}{\exp(-200 \cdot \delta_g^{-1})} = \exp(150 \cdot \delta_g^{-1}). \quad (5.18)$$

Tabelle 5.7 gibt für die verwendeten Werte von δ_g die Quotienten q und die Werte d_{S_O} wieder. Dabei ist d_{S_O} der Abstand der Positionsschätzung des Algorithmus vom Schwerpunkt S_O . Für kleine Werte von δ_g wird der Quotient q der Gewichte sehr groß, d.h. im Wesentlichen wird nur diejenige Messung, die am nächsten zur Vorhersage $y(k+1|k)$ benachbart ist, berücksichtigt. Dies entspricht aber einem Nearest-Neighbour-Algorithmus. Aus der graphischen Darstellung der Abweichung d_{S_O} in Abhängigkeit vom Gewichtparameter δ_g (Abbildung 5.7) erkennt man nun, dass d_{S_O} für $\delta_g \rightarrow 0$ ungefähr gegen den Wert $d_{S_O} = 21$ konvergiert. Wenn man den NNA anwendet und die mittlere Abweichung d_{S_O} vom Schwerpunkt S_O bestimmt, erhält man jedoch gerade den Wert 21.2197 (Kräußling et al., 2005b). Zur weiteren graphischen Veranschaulichung ist deshalb in Abbildung 5.7 die horizontale Gerade $d_{S_O} = 21.2197$ eingezeichnet.

Für sehr große Werte von δ_g nähert sich der Quotient q dem Wert Eins, d.h. alle Messwerte werden gleich gewichtet. Dies entspricht aber einem EWA Algorithmus. Aus der graphischen Darstellung erkennt man in diesem Fall, dass d_{S_O} sich immer mehr dem zugehörigen Wert $d_{S_O} = 0.4133$ aus Tabelle 5.2 für den EWA Algorithmus annähert. Zur weiteren Erläuterung dieses Sachverhaltes wurde deshalb auch die horizontale Gerade $d_{S_O} = 0.4133$ in der Graphik dargestellt.

Tabelle 5.7: Die mittlere Abweichung d_{S_O} vom Erwartungswert für den WMA Algorithmus in Abhängigkeit vom Gewichtsparameter δ_g .

δ_g	Quotient q	d_{S_O}/cm
1	$1.3937 \cdot 10^{65}$	21.2412
10	$3.2690 \cdot 10^6$	21.4232
100	4.4817	9.7480
1000	1.1618	0.8420
10000	1.0151	0.4331
100000	1.0015	0.4149
1000000	1.0002	0.4135
10000000	1.0000	0.4134
100000000	1.0000	0.4133
1000000000	1.0000	0.4133
10000000000	1.0000	0.4133

Man erkennt aus der Zeichnung auch empirisch, dass d_{S_O} eine streng monoton fallende Funktion des Gewichtsparameters δ_g ist, d.h. je stärker die einzelnen Messungen unterschiedlich gewichtet werden, desto schlechter verhält sich der dazugehörige Algorithmus. Insbesondere kann der EWA Algorithmus durch einen Algorithmus mit unterschiedlicher Gewichtung der Messwerte an Genauigkeit nicht übertroffen werden. Dieser Sachverhalt ist auch unmittelbar einsichtig. Bei punktförmigen Zielen in Clutter, bei denen höchstens eine Messung tatsächlich vom Ziel stammen kann und alle übrigen Messungen Störmeldungen sein müssen, ist es nämlich sinnvoll, Messungen, die weiter von der Vorhersage entfernt liegen, weniger stark zu gewichten, da diese mit geringerer Wahrscheinlichkeit derjenigen Messung entsprechen, die tatsächlich vom Ziel verursacht wird. Außerdem gibt es keine sichere Möglichkeit, festzustellen, welche Messung gerade diejenige ist, die vom Ziel herrührt. Bei ausgedehnten Einzelobjekten dagegen stammen in der Regel alle Messungen im Erwartungsgebiet vom Ziel selbst. Ausnahmen bestehen lediglich bei hoher Objektdichte oder wenn ein Ziel sich einem statischen Objekt in der Umgebung, wie zum Beispiel einem Hindernis oder einer Wand, annähert. Lediglich die Bereiche auf der Oberfläche des Zieles, die, zum Beispiel durch Reflexion des Laserstrahls, Ursprung der einzelnen Messungen sind, sind jeweils verschieden. Alle Messungen enthalten somit die gleiche Information über die tatsächliche Position des Objektes und sollten deshalb auch alle gleich gewichtet werden.

Da der Schwerpunkt S_O durch die Mehrzahl der Algorithmen sehr gut geschätzt wird, kann eine genaue Schätzung für den Mittelpunkt des kreisförmigen Objektes unmittelbar erhalten werden. Dies beruht auf der Tatsache, dass der Euklidische Abstand x dieses Schwerpunktes zum Mittelpunkt des Zieles in Abhängigkeit von der Entfernung d wie oben angegeben berechnet werden kann. Allerdings müsste hierzu strenggenommen der Abstand des Sensors vom Mittelpunkt des Objektes bekannt

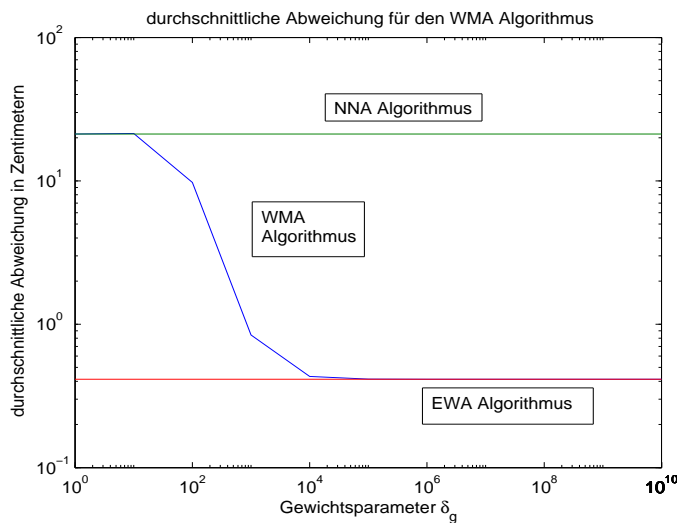


Abbildung 5.7: Die mittlere Distanz d_{S_O} vom Erwartungswert S_O für den WMA Algorithmus in Abhängigkeit vom Gewichtsparameter δ_g , doppelt logarithmische Darstellung.

sein. Die Tracking-Algorithmen liefern aber zunächst nur die Distanz vom Sensor zum Schwerpunkt S_O . Dieser Abstand weicht allerdings wie eben gezeigt nur um ca. zwanzig Zentimeter von der tatsächlichen Entfernung des Beobachters vom Mittelpunkt des Zieles ab und kann somit als Ausgangswert für d zur Berechnung des Abstandes x des Schwerpunktes S_O zum Mittelpunkt des Objektes benutzt werden. Mit diesem Wert x kann dann eine verbesserte Schätzung für den Abstand d des Sensors zum Mittelpunkt des Objektes berechnet werden. Auf diese Weise kann iterativ recht schnell eine sehr genaue Schätzung für den Mittelpunkt des Objektes erhalten werden, da darüberhinaus der Beobachter, der Schwerpunkt S_O und der Mittelpunkt M des Objektes aus Symmetriegründen auf einer geraden Linie liegen, wie es in Abbildung 5.8 gezeigt ist. Somit liefern die Verfahren EWA, WMA, RDA und MVAA eine gute Information über die Position des Objektes.

Die Berechnung des Schwerpunktes S_O mit dem VAA unter Verwendung von Hilfspunkten liefert keine Verbesserung gegenüber dem MVAA (Tabelle 5.8). Immerhin ist dieses Vorgehen aber wesentlich besser als ein reiner VAA.

Als Nächstes wird untersucht, wie geeignet diejenigen Verfahren sind, die direkt eine Schätzung für den Mittelpunkt des kreisförmigen Objektes berechnen. Da der Mittelpunkt bei bekanntem Radius konstruiert werden kann, wenn zwei Punkte auf der Oberfläche des Zieles bekannt sind, kommen hierfür nur Algorithmen in Betracht, die Schätzungen für mehrere Punkte auf der Oberfläche des Objektes liefern. Dies sind aber gerade der VBA, der MVA, der RDA und der VBA mit Hilfspunkten. Die Tabellen 5.9 bis 5.12 zeigen die Ergebnisse. Es fällt auf (Tabelle 5.9), dass es wieder relativ ungünstig ist, den Vorgänger über die Länge des Pfades zu bestimmen. Auch die Re-

Tabelle 5.8: Mittelwert des Abstandes vom Erwartungswert S_O für den VAA mit Hilfspunkten in Zentimetern.

d/cm	100	200	400	600	800
$\sigma = 0 \text{ cm}$	4.4188	3.0382	2.0803	2.7200	3.0784
$\sigma = 1 \text{ cm}$	4.2383	3.0154	2.0878	2.7347	3.0917
$\sigma = 3 \text{ cm}$	3.7013	2.8636	2.1371	2.8515	3.2302
$\sigma = 5 \text{ cm}$	3.2625	2.7463	2.2850	3.0617	3.4696
$\sigma = 7.5 \text{ cm}$	2.8972	2.6996	2.5306	3.4442	3.9371
$\sigma = 10 \text{ cm}$	2.6361	2.8252	3.0408	4.0095	4.6077

sultate bei Verwendung von Hilfspunkten sind nicht wesentlich besser (Tabelle 5.12). Die Bestimmung des Vorgängers über die Mahalanobisdistanz beim MVA und die Verwendung der Rohdaten beim RDA liefern noch relativ gute Ergebnisse. Dabei schneidet der Versuch der Verwendung der Rohdaten bei kleinen Standardabweichungen des Messrauschens besser ab als die Anwendung des Viterbi-Algorithmus, während die Verwendung der Schätzungen des MVA bei großen Standardabweichungen günstiger zu sein scheint. Dies ist auch nicht weiter verwunderlich, da es naheliegend ist, dass eine Filterung nur bei großen Messfehlern Sinn macht.

Vergleicht man jeweils die Genauigkeit der Schätzungen des Erwartungswertes S_O mit derjenigen der direkten Schätzungen des Mittelpunktes, so fällt Folgendes auf: der Viterbi-Ansatz, bei dem der Vorgänger über die Länge des Pfades bestimmt wird, ist in beiden Fällen ziemlich ungenau (Tabelle 5.5 und 5.9). Bestimmt man den Vorgänger dagegen unter Verwendung der Mahalanobisdistanz, so wird der Erwartungswert S_O vor allem für die kleineren Entfernungen genauer geschätzt als der Mittelpunkt (Tabelle 5.6 und 5.10). Verwendet man den Viterbi-Ansatz mit Hilfspunkten, so schneidet die Berechnung des Mittelpunktes bei einer Entfernung von 800 Zentimetern für die Standardabweichungen $\sigma = 5$, $\sigma = 7.5$ und $\sigma = 10$ cm etwas besser ab als die Berechnung des Schwerpunktes S_O (Tabelle 5.8 und 5.12). Dies liegt vermutlich daran, dass die Berechnung des Schwerpunktes S_O umso ungenauer ist, je weniger Messungen von Punkten auf der Oberfläche des Objektes zur Verfügung stehen. Bei einer Entfernung von 800 Zentimetern liegen aber nur noch relativ wenige solcher Messungen (circa drei bis fünf) vor. Im Großen und Ganzen ist es aber bei den Ansätzen, die den Viterbi-Algorithmus benutzen, insbesondere für die kleineren Entfernungen deutlich besser, den Erwartungswert S_O zu schätzen.

Für den RDA sind die Ergebnisse uneinheitlich (Tabelle 5.4 und 5.11). Bei den kleineren Standardabweichungen ist es besser, den Mittelpunkt zu berechnen, während für die größeren Standardabweichungen die Berechnung des Schwerpunktes S_O mit höherer Genauigkeit gelingt. Dies lässt sich wie folgt erklären: die Berechnung des Mittelpunktes ist bei vernachlässigbarem Messfehler schon bei Vorliegen zweier Punkte auf der Oberfläche exakt möglich, während eine exakte Berechnung des Schwerpunktes S_O nur gelingen kann, wenn von jedem Punkt auf der einsehbaren Oberfläche

Tabelle 5.9: Mittelwert des Abstandes vom Mittelpunkt für den VBA in Zentimetern.

d/cm	100	200	400	600	800
$\sigma = 0 \text{ cm}$	7.3433	3.9713	2.6189	3.4526	4.2712
$\sigma = 1 \text{ cm}$	9.6936	4.6554	2.2307	3.2798	4.4182
$\sigma = 3 \text{ cm}$	12.4534	5.2160	2.7105	3.6516	5.7831
$\sigma = 5 \text{ cm}$	12.4986	6.3547	3.5831	4.4436	6.1093
$\sigma = 7.5 \text{ cm}$	13.7370	6.7226	4.5513	5.6088	6.9626
$\sigma = 10 \text{ cm}$	15.4690	7.4177	5.8548	7.1893	8.1105

Tabelle 5.10: Mittelwert des Abstandes vom Mittelpunkt für den MVA in Zentimetern.

d/cm	100	200	400	600	800
$\sigma = 0 \text{ cm}$	6.3106	3.7726	2.1740	1.8275	1.5086
$\sigma = 1 \text{ cm}$	5.4591	3.5677	2.2193	2.0141	1.6513
$\sigma = 3 \text{ cm}$	3.8101	2.4429	2.1590	2.1211	2.1511
$\sigma = 5 \text{ cm}$	3.6390	2.3306	2.3373	2.6855	2.9442
$\sigma = 7.5 \text{ cm}$	4.8879	3.5369	3.3580	3.8252	4.1711
$\sigma = 10 \text{ cm}$	6.4753	4.8892	4.6250	5.1785	5.4488

eine Messung vorliegt. Dies ist aber in den Simulationen und auch in der Realität nie der Fall, da beispielsweise nur eine Messung pro Winkelgrad zur Verfügung steht. Bei kleinen Standardabweichungen liegen die Messungen nun ziemlich exakt auf der Oberfläche, so dass eine Berechnung des Mittelpunktes mit hoher Genauigkeit möglich ist. Im Falle $\sigma = 0 \text{ cm}$ müssten die Ergebnisse des RDA sogar exakt mit den tatsächlichen Werten übereinstimmen. Dass dies nicht der Fall ist, kann nur auf die endliche Rechengenauigkeit von MATLAB zurückgeführt werden. Insgesamt kann festgehalten werden, dass es numerisch wesentlich stabiler ist, den Erwartungswert S_O zu berechnen als direkt den Mittelpunkt des kreisförmigen Objektes.

Für die Verfahren VBA und MVA, die keine Mittelwertbildung benutzen, sind die Verhältnisse wesentlich komplizierter, falls nicht wie eben diskutiert aus jeweils zwei

Tabelle 5.11: Mittelwert des Abstandes vom Mittelpunkt für den RDA in Zentimetern.

d/cm	100	200	400	600	800
$\sigma = 0 \text{ cm}$	0.2309	0.1145	0.0581	0.0369	0.0273
$\sigma = 1 \text{ cm}$	0.7348	0.5754	0.6458	0.7679	0.8620
$\sigma = 3 \text{ cm}$	2.2349	1.7371	1.8530	2.2083	2.4738
$\sigma = 5 \text{ cm}$	3.6831	3.0140	3.0492	3.5738	3.9988
$\sigma = 7.5 \text{ cm}$	5.3939	4.5331	4.5255	5.1457	5.7468
$\sigma = 10 \text{ cm}$	6.9201	5.8355	5.8764	6.7312	7.3384

Tabelle 5.12: Mittelwert des Abstandes vom Mittelpunkt für den VBA bei Verwendung von Hilfspunkten in Zentimetern.

d/cm	100	200	400	600	800
$\sigma = 0 \text{ cm}$	9.5110	6.2426	4.1587	3.6245	3.2424
$\sigma = 1 \text{ cm}$	9.1314	6.1723	4.1457	3.6159	3.2465
$\sigma = 3 \text{ cm}$	8.7604	5.7118	3.9972	3.5822	3.2836
$\sigma = 5 \text{ cm}$	8.7857	5.5033	3.7542	3.5173	3.3834
$\sigma = 7.5 \text{ cm}$	9.3159	5.5242	3.6546	3.6972	3.7291
$\sigma = 10 \text{ cm}$	10.0072	5.9972	4.1485	4.3184	4.4339

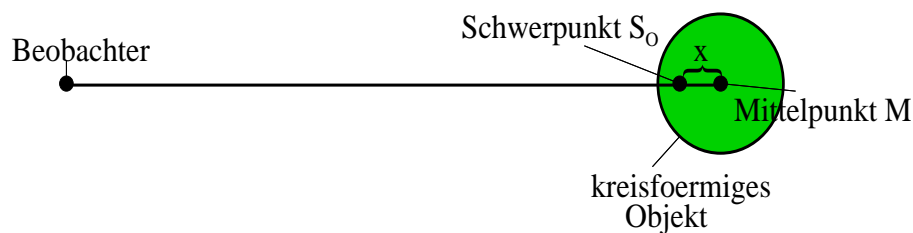


Abbildung 5.8: Bestimmung des Mittelpunktes eines kreisförmigen Objektes.

Positionsschätzungen eine Schätzung für den Mittelpunkt berechnet wird. Da diese beiden Ansätze für jede Sensormessung, die von der Oberfläche des Zieles herrührt, eine eigene Positionsschätzung berechnen, korrespondiert jede dieser Positionsschätzungen zu einem nicht näher bekannten Punkt auf der einsehbaren Oberfläche des Objektes. Die Algorithmen wählen dann eine dieser Schätzungen aus ohne das Wissen, zu welchem Punkt auf der Oberfläche diese Schätzung gehört. Deshalb besteht eine große Unsicherheit über die wahre Position des Zieles.

Dieser Sachverhalt soll in den Abbildungen 5.9 bis 5.14 veranschaulicht werden. In den Graphiken 5.9 und 5.10 sind die Punkte auf der Oberfläche des Objektes, die im Gesichtsfeld des Beobachters liegen, wiedergegeben. Die Bilder zeigen, dass für eine größere Entfernung des Zieles zum Beobachter auch eine größere Menge von solchen Punkten vorliegt. Auf diesen Umstand wurde bereits in Abschnitt 5.1 hingewiesen. Somit besteht für eine größere Entfernung eine größere Unsicherheit darüber, welcher Punkt auf der Oberfläche geschätzt wird. Der Winkel ϕ , der in den Zeichnungen wiedergegeben ist, ist dabei derselbe wie der Winkel ϕ aus Abbildung B.1.

Die Abbildungen 5.11 und 5.12 zeigen die möglichen Lagen des Mittelpunktes des Objektes, falls eine bestimmte Schätzung ausgewählt worden ist. Es besteht wiederum eine größere Unsicherheit für eine größere Entfernung.

Schließlich zeigen die Graphiken 5.13 und 5.14 einige Beispiele für mögliche Positionen des Zieles, falls man sich wieder schon auf eine Schätzung festgelegt hat. Es besteht wieder eine größere Bandbreite für größere Entfernungen.

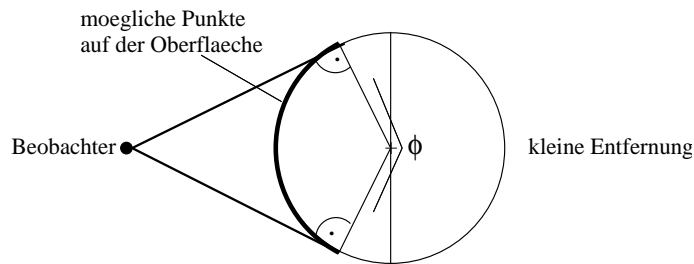


Abbildung 5.9: Mögliche Punkte auf der Oberfläche, von denen eine Messung stammen könnte, kleine Entfernung.

Aus den Bildern 5.13 und 5.14 kann geschlossen werden, dass eine große Unsicherheit in der Positionsschätzung besteht, wenn der VBA oder der MVA angewandt werden. Für große Entfernungen liegt diese Unsicherheit dabei sogar im Bereich des Durchmessers des Objektes. Es existieren darüberhinaus zwei weitere Probleme, die die Situation noch zusätzlich verschlimmern:

- der Punkt auf der Oberfläche des Zieles, der vom Laserstrahl getroffen wird, wechselt in der Regel von Scan zu Scan. Somit korrespondieren Vorhersage und Messung in Gleichung 4.25 nicht miteinander, sondern gehören zu unterschiedlichen Punkten auf der Oberfläche des Objektes.
- es besteht ein zusätzlicher Fehler in Form des Messrauschens, der die Messungen verfälscht.

Zusammenfassend kann gesagt werden, dass der VBA und der MVA nur wenig Information über die genaue Position des Objektes liefern. Ähnliches gilt auch für den NNA. Der VBA bzw. der MVA werden sich aber dennoch als nützlich erweisen. Zusammen mit dem EWA bilden sie die Basisalgorithmen für alle folgenden, komplexeren Verfahren.

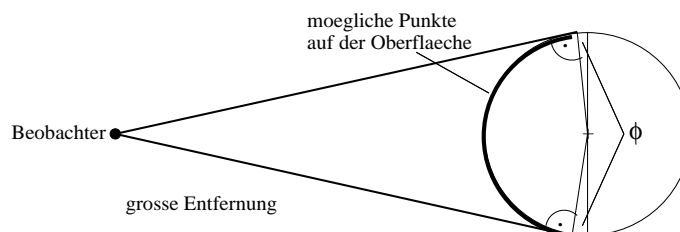


Abbildung 5.10: Mögliche Punkte auf der Oberfläche, von denen eine Messung stammen könnte, große Entfernung.

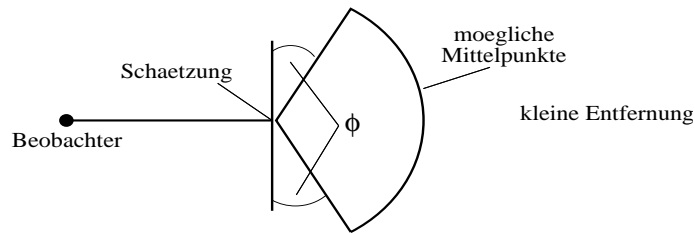


Abbildung 5.11: Mögliche Mittelpunkte, kleine Entfernung.

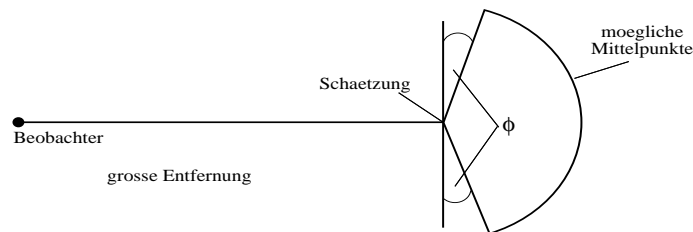


Abbildung 5.12: Mögliche Mittelpunkte, große Entfernung.

5.4 Die Komplexität der Verfahren

Es wird nun als zweites Gütekriterium die Komplexität der Algorithmen untersucht. Als ein Maß für diese Eigenschaft wurde der Bedarf an Rechenzeit für die Ausführung der Berechnungen für einen Zeittakt herangezogen. Der Grund hierfür liegt darin, dass einige der Verfahren sehr komplex sind und es deswegen sehr schwierig wäre, zum Beispiel die Anzahl der Matrizenmultiplikationen zu schätzen. Die Algorithmen wurden in MATLAB implementiert und auf einem Pentium IV mit 2.8 GHz ausgeführt.

Die mit Abstand einfachsten und schnellsten Ansätze sind EWA, WMA, RDA und NNA. Diese Verfahren benötigen ungefähr zwanzig Millisekunden pro Zeittakt für alle Kombinationen von Radius d und Standardabweichung σ des Messrauschens. Die Algorithmen VBA, VAA, MVA und MVAA, die alle den komplexen Viterbi-Algorithmus benutzen, sind sich in ihrem Verhalten ebenfalls sehr ähnlich. Repräsentativ sind hier die Ergebnisse für den VBA in Tabelle 5.13 wiedergegeben. Die Tabelle zeigt, dass die Laufzeiten in Abhängigkeit von der Entfernung d zwischen Objekt und Sensor von ungefähr 80 Millisekunden für die größeren Abstände bis zu circa 1.5 Sekunden für die kleineren Entfernungen variieren. Dies liegt daran, dass die Anzahl der Messungen, die vom Objekt stammen und damit im Erwartungsgebiet liegen, aufgrund des Strahlensatzes mit wachsendem Abstand d abfällt, und zwar ungefähr indirekt proportional. Die Komplexität dieser Verfahren hängt aber sehr stark von dieser Größe ab. Im Wesentlichen ist diese Abhängigkeit quadratisch, da die meiste Rechenzeit auf den Gatingprozess verwandt wird und hier für jedes Paar alter und neuer Messungen die Mahalanobisdistanz zwischen zugehöriger Vorhersage und Messwert berechnet wer-

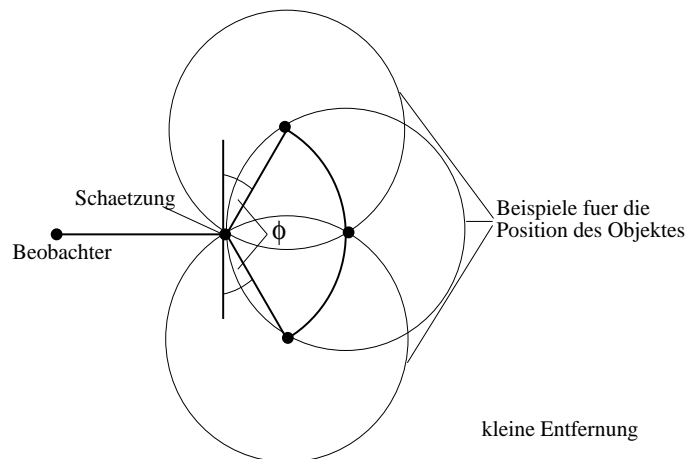


Abbildung 5.13: Beispiele für mögliche Positionen des Objektes, kleine Entfernung.

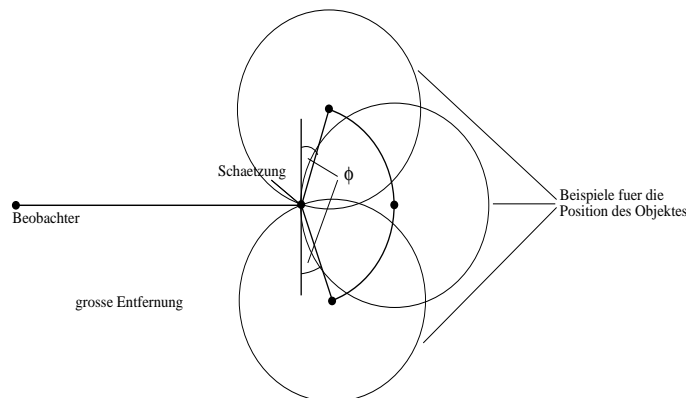


Abbildung 5.14: Beispiele für mögliche Positionen des Objektes, große Entfernung.

den muss. Außerdem muss für jede aktuelle Messung der Vorgänger aus allen Messungen des vorhergehenden Zeittaktes ausgewählt und eine separate Kalman–Filterung durchgeführt werden. Deswegen ist der VBA auch wesentlich schneller, wenn anstelle des Vorgängers Hilfspunkte verwendet werden, wie Tabelle 5.14 belegt. Dabei zeigt sich allerdings auch eine gewisse, wenn auch schwächere Abhängigkeit vom Abstand d . Dies liegt daran, dass hier viel Zeit darauf verwendet wird, den Mittelpunkt des Zieles zu berechnen. Dies geschieht aber für alle möglichen Paare von Messungen, die von der Wechselwirkung des Sensors mit der Oberfläche des Objektes herrühren. Im Falle von n Messungen beträgt die Anzahl dieser Paare aber gerade $\binom{n}{2} = \frac{n(n-1)}{2}$. Wie eben erläutert hängt die Anzahl n dieser Messungen aber sehr stark von der Entfernung des Zieles zum Sensor ab. Typischerweise beträgt die Anzahl der dem Objekt zugeordneten Messwerte für $d = 1$ m ca. dreißig und für $d = 8$ m nur noch ungefähr

Tabelle 5.13: Rechenzeitbedarf für den VBA in Sekunden.

d/cm	100	200	400	600	800
$\sigma = 0\text{ cm}$	1.4500	0.4079	0.1673	0.1023	0.0765
$\sigma = 1\text{ cm}$	1.4615	0.4122	0.1712	0.1043	0.0779
$\sigma = 3\text{ cm}$	1.4581	0.4199	0.1686	0.1026	0.0776
$\sigma = 5\text{ cm}$	1.4739	0.4144	0.1648	0.1048	0.0783
$\sigma = 7.5\text{ cm}$	1.4712	0.4124	0.1682	0.1028	0.0778
$\sigma = 10\text{ cm}$	1.4814	0.4231	0.1652	0.1044	0.0783

Tabelle 5.14: Rechenzeitbedarf für den VBA bei Verwendung von Hilfspunkten und Berechnung des Schwerpunktes S_O in Millisekunden.

d/cm	100	200	400	600	800
$\sigma = 0\text{ cm}$	69.7	34.7	25.5	24.1	22.8
$\sigma = 1\text{ cm}$	69.6	34.7	25.5	23.3	22.7
$\sigma = 3\text{ cm}$	69.5	34.7	25.6	23.5	22.8
$\sigma = 5\text{ cm}$	68.1	34.8	25.6	23.5	22.8
$\sigma = 7.5\text{ cm}$	69.6	34.5	25.6	23.2	22.8
$\sigma = 10\text{ cm}$	69.6	34.8	25.6	23.5	22.7

drei bis vier Messungen.

Die Verhältnisse für diejenigen Verfahren, die direkt eine Schätzung für den Mittelpunkt des Objektes liefern, sind folgendermaßen. VBA und MVA benötigen geringfügig mehr Rechenzeit als die analogen Algorithmen VAA und MVAA, die den Schwerpunkt S_O berechnen. Werden Hilfspunkte benutzt oder der RDA angewandt, so erhält man ähnliche Werte wie in Tabelle 5.14, da wieder die Hauptlast auf der Berechnung des Mittelpunktes liegt.

5.5 Die Verhältnisse beim People-Tracking

Die Verfolgung von Menschen ist wesentlich schwieriger zu analysieren als die Verfolgung von kreisförmigen Objekten, da man hier keine feste Form annehmen kann, sondern es vielmehr mit beliebigen, sich mit der Zeit dynamisch verändernden Oberflächen zu tun hat, die sogar von einem Zeittakt zum Nächsten stark variieren können. Im Gegensatz zu kreisförmigen Objekten, wo man einfach den Mittelpunkt des Zieles als die Position des Objektes verwenden kann und damit dessen Zustand bereits vollständig beschrieben hat, ist es bei Personen bereits schwierig zu definieren, welchen Punkt eines Menschen man als seine Position verwenden soll. Dieses Problem wird zum Beispiel dann noch verstärkt, wenn eines der Beine der Person durch das andere Bein verdeckt ist oder die Messungen nicht eindeutig zwei Beinen zugeordnet

werden können.

Die Möglichkeit der Modellierung der Bilder der Bewegung von Menschen in Laserdaten wird in der Literatur kontrovers diskutiert. Die meisten Autoren lehnen die Verwendung expliziter Modelle ab, da sie das Bewegungsverhalten als zu komplex und unvorhersagbar einschätzen (Zhao und Shibasaki, 2005). Ansätze zur expliziten Modellierung finden sich deshalb nur wenige (Bellotto und Hu, 2006; Brooks und Williams, 2003; Taylor und Kleeman, 2004), wobei in der Regel zwei nahe benachbarte Wolken von Messungen als Beinpaar erkannt werden. In der hier vorliegenden Arbeit wird der Meinung der Mehrheit der Autoren gefolgt und ebenfalls auf ein explizites Modell verzichtet. Dies liegt auch darin begründet, dass der Schwerpunkt dieser Arbeit auf der Interaktion von Personen liegt und es während einer solchen Interaktion aufgrund von Auflösungskonflikten oft völlig unmöglich ist, einzelne Wolken zu identifizieren und Beinen zuzuordnen (Taylor und Kleeman, 2004).

In Analogie zur Definition des Erwartungswertes S_O im Falle eines kreisförmigen Objektes ist es aber möglich, das Mittel aller Punkte auf der eingesehenen Oberfläche eines Menschen als seine Position zu definieren. In diesem Falle wäre wieder der EWA eine geeignete Tracking-Methode. Allerdings wäre es dann zum Beispiel so, dass, falls von einem Bein mehr Messungen als von dem anderen vorliegen, die geschätzte Position sich in Richtung des Beines mit der größeren Anzahl an Messwerten verschieben würde. Um dieses Problem besser zu lösen, könnte man deshalb versuchen, eine gute Überdeckung der Oberfläche einer Person durch die Verwendung mehrerer Sensoren zu erreichen. Näheres hierzu in Kapitel 10.2.

5.6 Eine mögliche Anwendung auf das Problem der Lokalisierung

Der Equal Weights Algorithmus liefert in Verbindung mit der genauen Berechnung des Schwerpunktes S_O eine sehr genaue Schätzung des Mittelpunktes eines mobilen Roboters mit einem Fehler im Bereich zwischen zwei Millimetern und 1.2 Zentimetern (Tabelle 5.2), falls dieser eine kreisförmige Oberfläche mit bekanntem Radius hat, wie in Abbildung 5.8 erläutert. Die letzte Annahme ist aber keine zu starke Einschränkung, da viele Service-Roboter gerade eine kreisförmige Oberfläche besitzen. Wenn man einen solchen mobilen Roboter mit Laser-Scannern beobachtet und dabei die Entfernung des Sensors zu den Punkten auf der Oberfläche misst und außerdem seine Position mit Hilfe dieses Verfahrens kontinuierlich bestimmt, erhält man somit ein sehr schnelles und effizientes Verfahren zur lokalen Lokalisierung oder Positionsschätzung eines mobilen Roboters (Kräußling, 2007b), das unmittelbar den Ort in kartesischen Koordinaten mit einem maximalen Fehler von ca. einem Zentimeter liefert und somit bisherige Methoden an Genauigkeit und Recheneffizienz übertrifft (Burgard et al., 1996; Fox et al., 1998; Fox et al., 1999a). Lokale Lokalisierung oder Tracking meint dabei die fortlaufende Positionsschätzung eines mobilen Roboters, wenn seine Anfangsposition

bekannt ist. Näheres zur Definition von lokaler und globaler Lokalisierung findet sich auch in der Einleitung (Kapitel 1). Der hier vorgestellte Ansatz unterscheidet sich allerdings dadurch wesentlich von gängigen Verfahren, dass die Sensoren sich nicht an dem mobilen Roboter selbst befinden.

Um das Problem der Positionsschätzung vollständig zu lösen, wird allerdings auch noch eine Schätzung für die Orientierung des Roboters benötigt. Diese kann aber innerhalb des hier vorgestellten Ansatzes sehr einfach erhalten werden. Dies beruht auf der Tatsache, dass der innere Zustand x_k gemäß Gleichung 4.3 den Vektor v_k der Geschwindigkeit des Roboters mit einschließt. Genauer ist

$$v_k = \begin{pmatrix} \dot{x}_{k1} & \dot{x}_{k2} \end{pmatrix}^\top. \quad (5.19)$$

Damit erhält man unter Zuhilfenahme einer einfachen Fallunterscheidung Ausdrücke für die Orientierung θ des Roboters. Zum Beispiel gilt im Falle von $\dot{x}_{k1} > 0$ und $\dot{x}_{k2} > 0$ die Beziehung

$$\theta = \arctan \frac{\dot{x}_{k2}}{\dot{x}_{k1}}. \quad (5.20)$$

Um auch noch das Problem der globalen Lokalisierung zu lösen, muss zusätzlich die Anfangsposition des mobilen Roboters bestimmt werden. Dies kann aber sehr leicht mit den in Kapitel 10.1 vorgestellten Methoden geschehen.

Nun wird beschrieben, wie die Methode in der Praxis angewandt werden kann. Zunächst wird eine volle Abdeckung der Umgebung durch Laserscanner benötigt. Dies bedeutet, dass es für jede mögliche Position des Roboters mindestens einen Laserscanner geben muss, in dessen Sichtfeld sich der Roboter vollständig aufhält. Diese Laser können in die Wände des zu überwachenden Gebäudes eingelassen sein, so dass ihre Position sehr genau bekannt ist. Sie müssen Messungen in einer horizontalen Ebene durchführen, und zwar alle in der gleichen Höhe. Der zu lokalisierende Roboter muss in dieser Ebene kreisförmig mit bekanntem Radius sein. Dann ist es möglich, die Position dieses Roboters mit den oben skizzierten Methoden sehr effizient zu schätzen. Dazu werden jeweils die Daten benutzt, die von denjenigen Lasern stammen, die den Roboter vollständig erfassen. Falls dabei mehr als ein Laser den Roboter komplett abdeckt, wird wie folgt verfahren. Diese Laser schätzen jeweils zunächst die Position des Roboters unabhängig voneinander und berechnen dabei eine Schätzung für den Mittelpunkt des Roboters wie beschrieben. Dann wird ein ungewichtetes Mittel dieser Schätzungen bestimmt und als aktuelle Schätzung des Mittelpunktes benutzt. Vergleiche hierzu auch die Erläuterungen in Kapitel 10.2. Für die Orientierung wird ebenfalls ein ungewichtetes Mittel herangezogen.

Der hier vorgestellte Ansatz ist zwar sehr genau und schnell, aber auch sehr teuer. Zum Beispiel würden, falls er in einer Büroumgebung verwendet werden soll, sehr viele Laser benötigt, die in den Wänden der Büroräume eingelassen sein müssten. Für jedes Büro würde dabei mindestens ein Laser benötigt. Die Situation wird dadurch noch zusätzlich verschlimmert, dass Hindernisse wie Schreibtische vorhanden sein können, die das Sichtfeld der Sensoren einschränken. Außerdem laufen in der

Regel Menschen in den Räumen umher. Trotzdem sollte der hier vorgestellte Ansatz nicht einfach verworfen werden, und zwar hauptsächlich aus zwei Gründen:

1. die Verwendung von Sensornetzwerken ist ein sehr gut etablierter Ansatz in der mobilen Robotik oder im Bereich smart environments / intelligente Gebäude, wie die folgenden Beispiele zeigen. Man spricht in diesen Fällen auch von multi-sensor surveillance systems oder distributed networked environment sensing:
 - in (Durrant-Whyte und Stevens, 2001) und (Makarenko und Durrant-Whyte, 2004) wird das Thema kooperierender autonomer Sensorplattformen diskutiert. Als Beispiel für mögliche Anwendungen werden Überwachung und intelligente Gebäude angeführt.
 - in (Brooks und Williams, 2003) wird ein System von verteilten Sensoren vorgestellt, das sich bewegende Personen verfolgt. Dabei werden verschiedene Typen von Sensoren wie Kameras und Laserscanner verwendet. Ergebnisse für eine Büroumgebung werden diskutiert.
 - in (Schulz et al., 2003b) wird ein Netzwerk von Laserscannern und Identitätssensoren in einem Gebäude verwendet, um Menschen genau und eindeutig zu verfolgen.
 - in (Zhao und Shibasaki, 2005) werden Fußgänger in einer Ausstellungshalle mit einem Netzwerk von Laserscannern verfolgt. Diese Laserscanner wurden dabei direkt auf dem Fußboden platziert und führten horizontale Messungen in einer Höhe von circa zwanzig Zentimetern über dem Boden aus.
 - in (Schumitsch et al., 2006a) wird ein Netzwerk von Kameras verwendet, die an der Decke befestigt sind, um Leute, die in einem Labor umherlaufen, zu verfolgen und zu identifizieren.

Die Verwendung von Sensornetzwerken zur Lokalisierung und zum Tracking wird im Übrigen auch in anderen Bereichen, wie zum Beispiel dem Aufspüren und dem Verfolgen von feindlichen U-Booten, eingesetzt (Erding et al., 2008).

2. es gibt Umgebungen, in denen die Anwendung des hier vorgestellten Ansatzes sehr hilfreich sein kann. Als Beispiel sei die Überwachung eines Kunstmuseums während der Schließungszeiten genannt. In diesem Falle müssen weite und leere Räume überwacht werden, die sehr gut von Laserscannern abgedeckt werden können, die in die Wände integriert sind. Abbildung 5.15 zeigt ein typisches Szenario. Da einige der Gemälde sehr berühmt und teuer sein können, sind die Kosten für die Laserscanner und die mobilen Roboter vernachlässigbar im Vergleich zu dem Diebstahl eines dieser Kunstwerke. Wie gezeigt wurde, kann ein Roboter sehr genau durch die Laserscanner lokalisiert werden, so dass er sich in den Räumen bewegen kann, um diese zu überwachen. Dabei kann er fortlaufend seine aktuelle Position, die die Lokalisierung bereitstellt, mit der

Gewünschten vergleichen und gegebenenfalls seine Antriebssteuerung anpassen. Natürlich kann eingewendet werden, dass die Räume auch durch die Laserscanner selbst überwacht werden könnten. Aber die Verwendung eines mobilen Roboters kann großen Nutzen bringen. Zum Beispiel kann ein mobiler Roboter versuchen, möglichen Eindringlingen zu folgen, um von diesen mit einer Kamera Aufnahmen zu machen. Diese Eindringlinge können dabei beispielsweise mit Methoden aufgespürt werden, wie sie in (Moors et al., 2005; Moors und Schulz, 2006; Moors, 2008) vorgestellt wurden.

Ähnliche Systeme gibt es aber auch schon in der Praxis. So vertreibt die Berliner Firma Robowatch den mobilen Sicherheitsroboter MOSRO zur Überwachung von Gebäuden (Schmundt, 2004). Dieser wird heute bereits von Firmen wie Bayer, BMW oder Daimler eingesetzt.

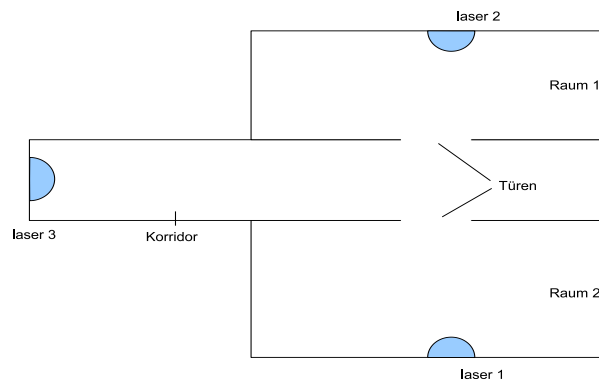


Abbildung 5.15: Skizze einer typischen Anwendungsumgebung.

Nun soll die hier vorgestellte Methode mit einigen der wohlbekanntesten Ansätze verglichen werden. Da der hier vorgestellte Algorithmus ein sehr starkes Vorwissen über die Umgebung, zum Beispiel in Form der Position der Lasersensoren, verwendet, sind hierfür am ehesten Landmarken-basierte Verfahren wie (Betke und Gurvits, 1997; Murillo et al., 2006; Thrun, 1998a) geeignet. Dazu gehören auch die im Folgenden erwähnten Ansätze. Dabei kennt der Roboter die Position einiger Landmarken in der Umgebung aus einer Karte oder lernt diese zunächst und nimmt dann mit Hilfe seiner Sensoren Messungen bezüglich dieser Landmarken vor. In (Fox et al., 1999b) wird angeführt, dass bei Verwendung von Laserdaten für Zellen-basierte Lokalisation und Monte-Carlo-Lokalisation ein Fehler von ungefähr fünf Zentimetern erreichbar ist. Dabei wird der Zustandsraum durch stückweise konstante Funktionen angenähert (Burgard et al., 1996; Fox et al., 1998). Allerdings können die Ergebnisse für die Zellen-basierte Methode nicht in Echtzeit erhalten werden. Demgegenüber erlaubt der hier vorgestellte Ansatz die Lokalisierung des Roboters mit einer Genauigkeit zwischen 0.2 und 1.2 Zentimetern, was um eine Größenordnung besser ist. Darüberhinaus

ist diese Methode sehr schnell und benötigt sehr wenig Speicherplatz im Vergleich zu den anderen Methoden. Zum Beispiel müssen keine Landmarken und auch keine Karte gespeichert werden. Lediglich die Position der statischen Laserscanner muss bekannt sein. Bezüglich des Problems der globalen Lokalisierung benötigt die Zellen-basierte Methode ungefähr 120 Sekunden, während die Monte-Carlo-Lokalisierung ungefähr drei Sekunden für die Verarbeitung eines einzelnen Laserscans verbraucht. Auf der anderen Seite ist die hier vorgestellte Methode in der Lage, einen Laserscan fast augenblicklich zu verarbeiten. Es ist dabei nur nötig, dass sich der Roboter ein wenig bewegt, um die vierte Bedingung für die Trackinitialisierung aus Kapitel 10.1 zu erfüllen.

Ein weiterer Vorteil der hier vorgestellten Methode besteht darin, dass ein wohldefinierter Punkt des Roboters, nämlich sein Mittelpunkt, geschätzt wird. In der gängigen Literatur ist dagegen keine Information zu finden, auf welchen Punkt eines Roboters, der ja eine gewisse räumliche Ausdehnung aufweist, sich die Lokalisierung eigentlich bezieht. Ein gravierender Nachteil der hier vorgestellten Methode ist aber, dass die Anforderungen an das System wesentlich restriktiver sind als bei üblichen Ansätzen, die lediglich voraussetzen, dass der Roboter über geeignete eigene Sensoren verfügt und eine Karte seiner Umgebung kennt.

Der hier vorgestellte Ansatz ist auch gut angepasst an das Problem des entführten Roboters. Falls der Roboter an eine andere Stelle verbracht worden ist, werden nämlich keine Messungen mehr mit der alten Spur assoziiert, da das zugehörige Erwartungsgebiet leer ist. Somit wird die alte Spur gelöscht gemäß Abschnitt 10.1. Darüberhinaus wird aufgrund des gewählten Ansatzes zur globalen Lokalisierung eine neue Spur an der aktuellen Position des Roboters initialisiert.

Schließlich soll noch darauf hingewiesen werden, dass auch der Fall der Multi-Roboter-Lokalisierung in einer sehr natürlichen Art und Weise in den hier vorgestellten Ansatz integriert werden kann. Solange die Roboter nicht interagieren, was bedeutet, dass sie sich sehr nahe kommen, werden sie problemlos unabhängig voneinander lokalisiert. Da gewöhnliche Algorithmen dazu neigen, Spuren zu verlieren, wenn eine Interaktion stattfindet, wie zum Beispiel in (Schumitsch et al., 2006a) festgestellt wurde, sollte andernfalls der in dieser Arbeit vorgestellte ECSA Algorithmus verwendet werden, um auch interagierende Roboter ohne Spurverlust verfolgen zu können. Natürlich kann dieses Verfahren auch dazu verwendet werden, dass sich die Roboter untereinander verfolgen, um die Genauigkeit weiter zu verbessern.

5.7 Ausblick

Die Ausdehnung von Objekten spielt auch in anderen Anwendungsgebieten der mobilen Robotik wie der Lokalisierung oder dem SLAM-Problem (Simultaneous Localization and Mapping) eine bedeutende Rolle. Ein wichtiger Ansatz in diesen Bereichen besteht nämlich wie gerade erwähnt darin, in der Umgebung sogenannte Landmarken zu identifizieren und die relative Position des Roboters zu diesen Landmarken zu vermessen. Auch hier können von einer Landmarke wieder mehrere Messungen erhalten

werden und in der Regel sorgt die Ausdehnung der Landmarken sogar für eine größere Streuung der Messwerte als die Ungenauigkeit der Messgeräte selbst. Diese Tatsachen rücken in den letzten Jahren mehr und mehr in das Blickfeld der Forschung. So wurde in (Tardos et al., 2002) die Hough–Transformation (Ballard und Brown, 1982) benutzt, um punktförmige und ebene Features voneinander zu unterscheiden. In (Guivant et al., 2000b; Guivant und Nebot, 2003) werden Bäume in einem Park als Landmarken oder Features verwendet. Die Bäume selbst werden dabei als kreisförmig angenommen. Der Radius eines bestimmten Baumes wird dann ähnlich wie in Abbildung 10.2 bestimmt.

Kapitel 6

Die Verfolgung zweier interagierender Ziele

6.1 Eine Einführung in die Problemstellung

Bei der Verfolgung mehrerer Personen kann es vorkommen, dass sich zwei dieser Probanden für einige Zeit sehr nahe kommen oder sich treffen, um sich dann vorübergehend zusammen zu bewegen und anschließend wieder in verschiedene Richtungen auseinanderzulaufen. Dies kann zum Beispiel der Fall sein, wenn zwei Personen miteinander reden oder Gegenstände austauschen. Im Falle von mobilen Robotern ist es möglich, dass die Roboter eine Formation bilden, um zusammen eine Aufgabe auszuführen. Schließlich können auch ein Mensch und ein Roboter sich gemeinsam an einem Ort aufhalten. In all diesen Fällen spricht man von einer Interaktion, einer Begegnung, einer Kreuzung, dem Kreuzen oder man verwendet den englischen Begriff *Crossing*. Eine typische Situation skizziert Abbildung 6.1. Zunächst bewegen sich die beiden Objekte getrennt (linke Hälfte des Bildes). In der Mitte findet dann eine Begegnung statt. Anschließend trennen sich die beiden Ziele und bewegen sich in der rechten Bildhälfte wieder separat.

Formal lässt sich ein *Crossing* dahingehend definieren, dass sich die Erwartungsgebiete zweier Objekte überlagern, d.h. dass es Messwerte gibt, die in den Erwartungsgebieten beider Ziele liegen. Dann können die Messungen von den beiden Objekten nicht mehr auseinandergehalten und eindeutig einem Ziel zugeordnet werden, und zwar solange, bis sich die Objekte wieder voneinander getrennt haben. Somit kommt es zu Auflösungskonflikten mit den im Folgenden beschriebenen Schwierigkeiten. Das Phänomen der interagierenden Ziele ist ein zentrales Problem im Bereich des *People-Tracking* (Bellotto und Hu, 2007a; Mucientes und Burgard, 2006; Prassler et al., 1999; Schumitsch et al., 2006a) oder in der Luftraumüberwachung (Fortmann et al., 1983; Reid, 1979). In diesem Kapitel wird zunächst nur der Fall zweier kreuzender Objekte behandelt. Der allgemeine Fall beliebig vieler interagierender Ziele findet sich dann im nächsten Kapitel.

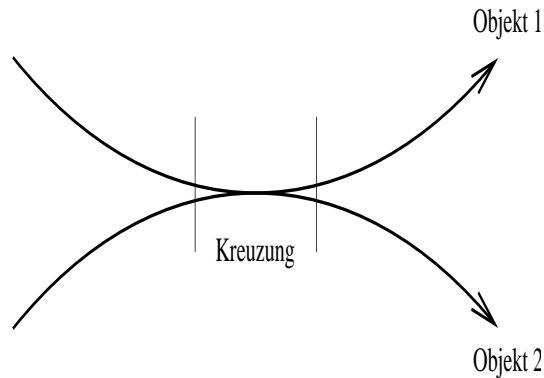


Abbildung 6.1: Zwei kreuzende Objekte 1 und 2.

Der Begriff des Crossings stammt ursprünglich aus der Luftraumüberwachung. Dort versteht man darunter, dass sich zwei Luftfahrzeuge für kurze Zeit sehr nahe kommen und dann wieder getrennt weiterfliegen. Da die Objekte in der mobilen Robotik während einer Begegnung oft gemeinsam eine Aufgabe ausführen oder auf andere Weise miteinander wechselwirken, soll hier auch der deskriptive Begriff der Interaktion verwendet werden. Das wesentliche Charakteristikum des Phänomens der interagierenden Ziele oder allgemeiner von Zielgruppen ist nun, dass einfache Verfahren die einzelnen Objekte nach der Interaktion oft nicht mehr separat verfolgen. Vielmehr wird dann für beide Ziele derselbe computerinterne Track berechnet. Hierfür sind in der mobilen Robotik vor allem folgende Umstände maßgebend:

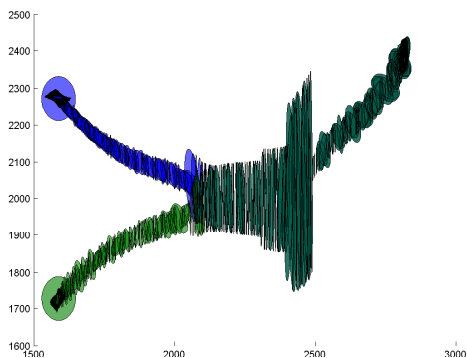


Abbildung 6.2: Zwei interagierende Ziele, verarbeitet mit dem EWA.

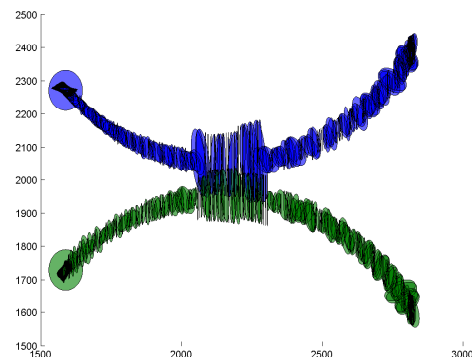


Abbildung 6.3: Zwei interagierende Ziele, verarbeitet mit dem WMA.

Zwei interagierende Ziele, verarbeitet mit dem EWA bzw. dem WMA.

1. die Interaktion kann sich über einen längeren Zeitraum erstrecken.

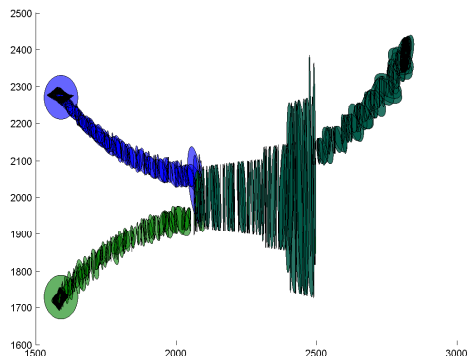


Abbildung 6.4: Zwei interagierende Ziele, verarbeitet mit dem RDA.

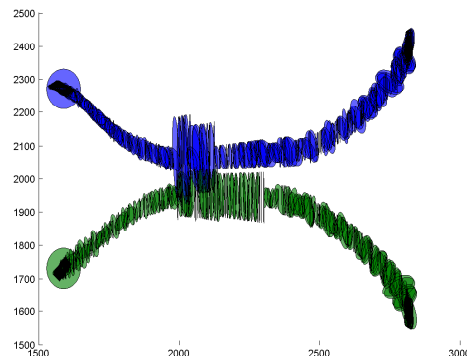


Abbildung 6.5: Zwei interagierende Ziele, verarbeitet mit dem NNA.

Zwei interagierende Ziele, verarbeitet mit dem RDA und dem NNA.

2. eines der beiden Objekte kann von dem anderen Ziel für einige Zeit verdeckt werden. Von dem Objekt, das sich im Schatten des anderen bewegt, werden dann während dieser Zeitspanne keine Messungen registriert.
3. solange keine Verdeckung vorliegt, liefert jedes Objekt in der Regel mehrere Messwerte.
4. wenn es sich um Menschen handelt, können diese während und vor allem am Ende einer Interaktion, also wenn eine Gruppe sich wieder auflöst und die realen Bahnen divergieren, sehr abrupte Manöver ausführen. So können Personen zum Beispiel plötzlich ihre Bewegungsrichtung ändern oder sogar von einem Augenblick zum Anderen, ohne sich umzudrehen, rückwärts laufen.

Im Folgenden wird nun unter Berücksichtigung der eben angeführten vier Punkte erläutert, wie die in Kapitel 4 eingeführten einfachen Tracking-Algorithmien mit dem Problem zweier interagierender Objekte zurechtkommen. Genauer wird untersucht, ob nach der Interaktion für die beiden Ziele zwei verschiedene Tracks weitergeführt werden. Dabei werden zunächst simulierte Daten betrachtet, bei denen die Objekte auch während der Interaktion noch einen gewissen Mindestabstand voneinander einhalten. Ansonsten verläuft die Bewegung der Ziele wie in Abbildung 6.1 skizziert. In den Graphiken sind die Bahnen der Objekte durch Ellipsen wiedergegeben. Die Mittelpunkte dieser Ellipsen entsprechen dabei jeweils den Positionsschätzungen, die das Tracking-Verfahren liefert, während die Ellipse selbst jeweils die aktuelle Form, Orientierung und Ausdehnung oder einfach die Geometrie der Ziele beschreiben soll, wie es in (Koch und Stannus, 2003; Stannus et al., 2004) vorgeschlagen wurde. Zur Berechnung dieser Ellipsen vergleiche (Kräußling, 2007a).

Für den EWA gilt Folgendes (Abbildung 6.2): schon kurz nach Beginn der Interaktion sind in den beiden Erwartungsgebieten der beiden Objekte alle Messungen von

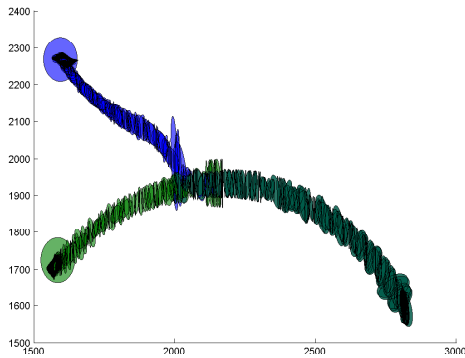


Abbildung 6.6: Zwei interagierende Ziele, verarbeitet mit dem VBA.

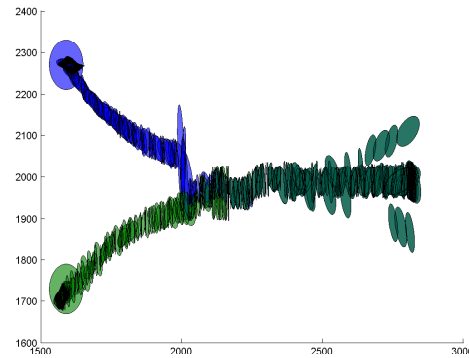


Abbildung 6.7: Zwei interagierende Ziele, verarbeitet mit dem VAA.

Zwei interagierende Ziele, verarbeitet mit dem VBA bzw. dem VAA.

beiden Zielen enthalten, so dass für beide Objekte dieselben Messwerte verarbeitet werden. In den Update-Gleichungen 4.25 und 4.27 des Kalman-Filters geht somit für beide Tracks dasselbe ungewichtete Mittel aller Messungen von beiden Zielen ein. Da das Crossing für eine längere Zeitspanne andauert, geschieht dies für mehrere Zeittakte hintereinander. Somit nähern sich die Positionsschätzungen für beide Objekte immer weiter an, bis sie schließlich am Ende der Begegnung fast identisch sind (Konvergenz der Tracks, vergleiche den weiteren Text). Da das Verfahren deterministisch ist, entscheidet sich der Algorithmus dann am Ende der Begegnung bei beiden Tracks für dasselbe Ziel. Dabei handelt es um das Objekt, zu dem die mit dem Kalman-Filter berechnete Vorhersage besser passt. Für beide Tracks werden dann im Weiteren identische Erwartungsgebiete berechnet, so dass die Messungen eines der beiden Objekte schließlich nicht mehr erfasst werden. Die Information über dieses Ziel, das sich im vorliegenden Fall nach der Interaktion nach rechts unten bewegt, geht deshalb komplett verloren. Beiden Objekten wird also nach der Interaktion dieselbe Position zugeordnet. Man spricht in einem solchen Fall auch von Trackverlust. Da während der Begegnung im jeweiligen Erwartungsgebiet Messungen von beiden Zielen liegen und diese über einen gewissen Bereich verstreut sind, reißen während der Interaktion die Ellipsen, die in der graphischen Darstellung die aktuelle Geometrie der Objekte beschreiben sollen, sehr weit auf.

Für den WMA gilt (Abbildung 6.3): da während der Interaktion die Messungen von dem jeweils anderen Objekt weiter von der eigenen Vorhersage entfernt liegen, haben diese Messwerte bei der Berechnung des gewichteten Mittels nach Gleichung 4.34 ein kleineres Gewicht. Somit wird die geschätzte Position im Vergleich zum EWA weniger stark in die Richtung des anderen Zieles verschoben. Die Positionsschätzungen behalten deshalb bis zum Ende der Interaktion noch einen gewissen Mindestabstand voneinander bei. Dies führt dazu, dass sich jeder der beiden Tracks am Ende der Inter-

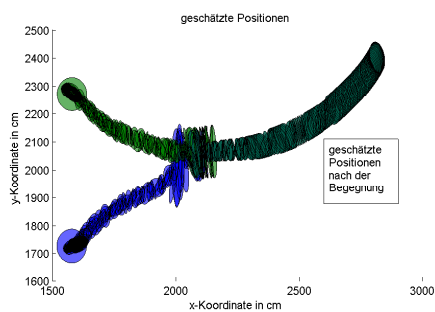


Abbildung 6.8: Zwei interagierende Ziele, verarbeitet mit dem MVA.

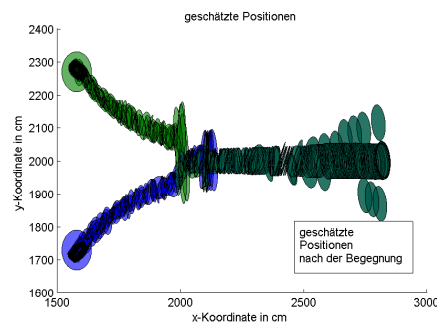


Abbildung 6.9: Zwei interagierende Ziele, verarbeitet mit dem MVAA.

Zwei interagierende Ziele, verarbeitet mit dem MVA und dem MVAA.

aktion für das richtige Objekt entscheidet.

Für den RDA gilt Folgendes (Abbildung 6.4): ähnlich wie beim EWA werden auch hier während der Interaktion für beide Objekte dieselben Messungen verarbeitet. Da hier ein ungewichtetes Mittel aller Messwerte als Positionsschätzung verwandt wird, sind diese Schätzungen während der Interaktion sogar identisch. Am Ende der Interaktion muss sich der Algorithmus dann jeweils wieder für eines der beiden Objekte entscheiden. Da für beide Tracks jeweils dieselbe Vorhersage und dasselbe Erwartungsgebiet berechnet wird, wählen dabei beide Tracks wie im Falle des EWA die Bahn desselben realen Objektes. Es kommt also erneut zum Trackverlust.

Für den NNA gilt (Abbildung 6.5): da die beiden Objekte während der Interaktion noch einen gewissen Mindestabstand voneinander beibehalten, wird bei jedem Zeittakt eine Messung als nächster Nachbar ausgewählt, die jeweils vom richtigen Ziel stammt. Dies gilt auch noch am Ende der Begegnung, so dass beide Tracks sich für das richtige Objekt entscheiden.

Die Verhältnisse für den VBA und den MVA sind sehr ähnlich, so dass diese beiden Verfahren zusammen besprochen werden sollen (Abbildung 6.6 und 6.8). In beiden Fällen stammen die Positionsschätzungen, die zu dem ausgewählten kürzesten Pfad bzw. dem Pfad mit Index Eins gehören, nach der Interaktion von demselben Objekt. Dies liegt daran, dass schon kurz nach Beginn der Interaktion für beide Tracks alle Messungen von beiden Zielen verarbeitet werden. Dies führt dazu, dass die Erwartungsgebiete der beiden Tracks am Ende der Interaktion nahezu identisch sind und die Algorithmen somit auch dann damit fortfahren, für beide Tracks alle Messwerte von beiden Objekten gleichzeitig zu verarbeiten. Diese bilden dann jeweils zwei Wolken oder Cluster. Wie bald gezeigt wird, ist aber gerade dieses Verhalten, das dann noch näher erläutert wird, die entscheidende Grundlage für die Entwicklung von Verfahren, die interagierende Ziele stabil, d.h. ohne Trackverlust, verfolgen können.

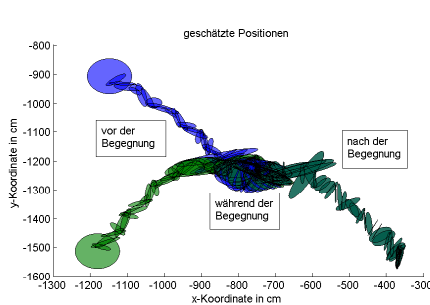


Abbildung 6.10: Zwei interagierende Ziele, verarbeitet mit dem WMA.

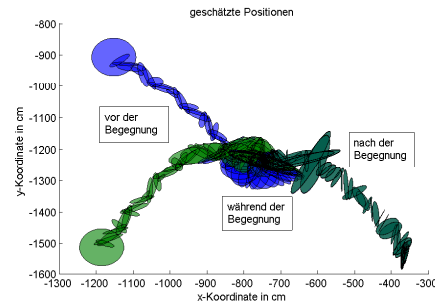


Abbildung 6.11: Zwei interagierende Ziele, verarbeitet mit dem NNA.

Zwei interagierende Ziele, reale Daten, verarbeitet mit dem WMA bzw. dem NNA.

Auch für den VAA und den MVAA gelten ähnliche Verhältnisse (Abbildung 6.7 und 6.9). Nachdem vor der Interaktion die beiden Objekte separat getrackt werden, werden nach der Begegnung die zu den jeweiligen Messungen gehörenden Positionsschätzungen von beiden Objekten dazu benutzt, um ein ungewichtetes Mittel zu berechnen, da der Algorithmus nicht in der Lage ist, die Schätzungen einzelnen Objekten zuzuordnen. Dieses Mittel entspricht dabei einem nicht existierenden Ziel, das sich in der Mitte zwischen den beiden realen Objekten befindet. Der Grund für dieses Verhalten liegt darin, dass diese beiden Verfahren für jede Positionsschätzung vom vorhergehenden Zeitpunkt eine eigene Vorhersage für die Messungen und ein eigenes Erwartungsgebiet berechnen. Auf diese Weise werden nach der Interaktion von jedem Track alle Messwerte von beiden Objekten erfasst und für jede dieser Messungen eine eigene Positionsschätzung bestimmt. Vergleiche zu diesem Verhalten auch wieder die Ausführungen zu Beginn von Abschnitt 6.2. Da schließlich als Position ein ungewichtetes Mittel all dieser Schätzungen berechnet wird, liegt diese Position in der Mitte zwischen den echten Positionen der beiden Ziele. Dabei kommt es gelegentlich vor, dass die Schätzung etwas in Richtung des einen oder anderen Objektes verschoben ist. Dies liegt daran, dass die Anzahl der Messungen, die man erhält, nicht immer für beide Ziele gleich ist. In einem solchen Fall bewegt sich die Positionsschätzung dann in Richtung desjenigen Objektes, das gerade mehr Messwerte liefert. Ein ähnliches Verhalten im Falle punktförmiger Ziele in Clutter ist für das sogenannte PDAF Verfahren, das einen gewichteten Mittelwert aus den Messungen und den Kalman-Filter benutzt, aus der Literatur bekannt (Bar-Shalom und Fortmann, 1988).

Von den aus Kapitel 4 bekannten Algorithmen lösen somit nur der WMA und der NNA das Problem zweier interagierender Ziele im Falle der betrachteten simulierten Daten. Eine wesentliche Voraussetzung dabei war aber in beiden Fällen, dass die bei-

den Objekte während der Interaktion noch einen gewissen Mindestabstand voneinander bewahren. Insofern überrascht es nicht, dass auch diese beiden Verfahren scheitern, wenn sie auf reale Daten angewandt werden, wo sich zwei Menschen sehr nahe kommen und zeitweise während der Interaktion auch noch eine Verdeckung auftritt. Dies bedeutet wie schon erwähnt, dass von einer der beiden Personen eine Zeitlang keine Messungen erhalten werden, da sie sich im Schatten der anderen Person befindet. Deshalb werden für beide Tracks während dieser Zeitspanne dieselben Messwerte verarbeitet. Im Falle des WMA (Abbildung 6.10) werden somit insbesondere aufgrund der Verdeckung die Positionsschätzungen der beiden Tracks analog zum EWA sehr ähnlich. Beim NNA (Abbildung 6.11) stehen dagegen während der Verdeckung für die Auswahl des nächsten Nachbarn nur Messungen von einem der beiden Objekte zur Verfügung. Insgesamt entscheiden sich sowohl beim WMA als auch beim NNA beide Tracks am Ende der Interaktion für dasselbe reale Objekt, da sich die Tracks während der Interaktion zu ähnlich geworden sind. Ein analoges Verhalten im Falle punktförmiger Ziele in Clutter ist für den Nearest-Neighbour-Algorithmus auch schon wieder aus der Literatur bekannt (Bar-Shalom und Fortmann, 1988).

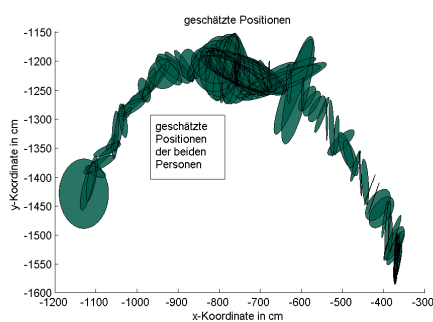


Abbildung 6.12: Zwei interagierende Ziele, verarbeitet mit dem MHT.

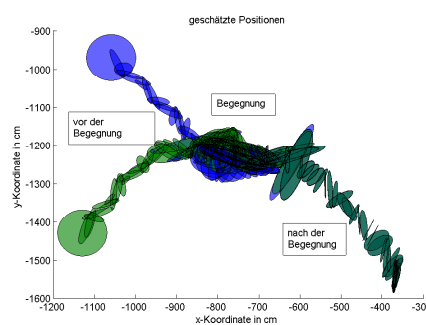


Abbildung 6.13: Zwei interagierende Ziele, verarbeitet mit dem MHT.

Zwei interagierende Ziele, reale Daten, verarbeitet mit dem MHT.

Somit kann keines der einfachen Verfahren das Problem der Interaktion zufriedenstellend, d.h. ohne Trackverlust, lösen, wie auch schon in (Kräußling, 2006a; Kräußling, 2008b) dargestellt wurde. In den allermeisten Fällen wird vielmehr beiden Tracks nach der Interaktion dieselbe Positionsschätzung zugeordnet. Die Verfolgung interagierender Ziele erfordert deshalb speziell angepasste Ansätze. Auf dem Gebiet der Luftüberwachung sind ähnliche Probleme schon länger bekannt. Dort wurden deshalb zwei Verfahren entwickelt, die das Problem des Crossings im Falle punktförmiger Ziele in Clutter lösen sollen:

1. der Multi-Hypothesen-Tracker (MHT), der 1979 von Reid eingeführt wurde (Reid, 1979).

2. der Joint Probabilistic Data Association Filter (JPDAF), der letztendlich 1983 von Fortmann, Bar-Shalom und Scheffe eingeführt wurde (Bar-Shalom, 1974; Bar-Shalom et al., 1980; Fortmann et al., 1980; Fortmann et al., 1983).

Im Folgenden sollen diese beiden Algorithmen, angepasst auf ausgedehnte Ziele, kurz beschrieben und aufgezeigt werden, warum auch diese Ansätze auf dem Gebiet der mobilen Robotik häufig scheitern.

Der MHT berechnet für jede mögliche Folge von Messungen eine eigene Folge von Positionsschätzungen. Hat man beispielsweise zehn Zeittakte hintereinander jeweils fünf Messungen vorliegen, so erhält man 5^{10} unterschiedliche Hypothesen. Der MHT ist somit eine Verallgemeinerung des Viterbi-Ansatzes. Ähnlich wie beim VBA wird nun jede Hypothese mit einer Wahrscheinlichkeit belegt und die Hypothesen werden nach diesen Wahrscheinlichkeiten oder Likelihoods geordnet (Cox, 1993; Cox und Hingorani, 1996). Werden nun zwei Objekte verfolgt, so wählt man am Ende die beiden Hypothesen aus, die die beiden höchsten Wahrscheinlichkeiten aufweisen. Da Hypothesen, die zu Bahnen realer Objekte gehören, aufgrund der Korrelation der Messwerte in der Regel eine höhere Wahrscheinlichkeit aufweisen als Hypothesen, die auf Clutter beruhen, hofft man auf diese Weise gerade die beiden Hypothesen auszuwählen, die den beiden echten Bahnen entsprechen. Da auf dem Gebiet der mobilen Robotik aber beide Objekte in der Regel mehrere Messungen generieren, wird dieser Ansatz dort häufig aus folgenden zwei Gründen scheitern:

1. da ein Objekt Ursprung von ungefähr zehn Messungen sein kann, kann die Anzahl der Hypothesen nach nur zehn Zeittakten schon circa $(2 \cdot 10)^{10}$ betragen. Eine solche Anzahl von Hypothesen ist aber mit gängigen Rechnern, wie sie zur Verwendung auf mobilen Einheiten gedacht sind, nicht mehr darstellbar.
2. da ein Objekt in der Regel Ursprung von mehr als einer Messung ist, können beide ausgewählten Hypothesen zu der Endposition desselben Zieles gehören. Auch die Startposition und alle Positionen dazwischen können Messwerten von demselben Objekt entsprechen. Der Ausschluss der Zuordnung von Messungen desselben Ziels zu beiden Tracks ist aber die entscheidende Voraussetzung für den Erfolg dieses Verfahrens.

Diese Verhältnisse erläutern die Abbildungen 6.12 und 6.13. Im ersten Fall ordnet der MHT sowohl vor als auch nach der Interaktion beide Tracks der Bahn desselben realen Objektes zu. Im zweiten Fall wird das Verfahren von Hand gezwungen, beiden Tracks unterschiedliche Startpositionen zuzuweisen. Die Endpositionen sind trotzdem aber auch in diesem Fall wieder nahezu identisch.

Der JPDAF benutzt ähnlich wie der WMA ein gewichtetes Mittel aller Messungen in der Update-Gleichung des Kalman-Filters. Dabei wird über alle möglichen Aufteilungen aller Messwerte auf die beiden Tracks aufsummiert, wobei jede dieser Aufteilungen oder Assoziationen entsprechend ihrer Wahrscheinlichkeit gewichtet wird (Cox, 1993). Aus folgenden beiden Gründen führt auch der JPDAF im Bereich der mobilen Robotik in der Regel nicht zum Ziel:

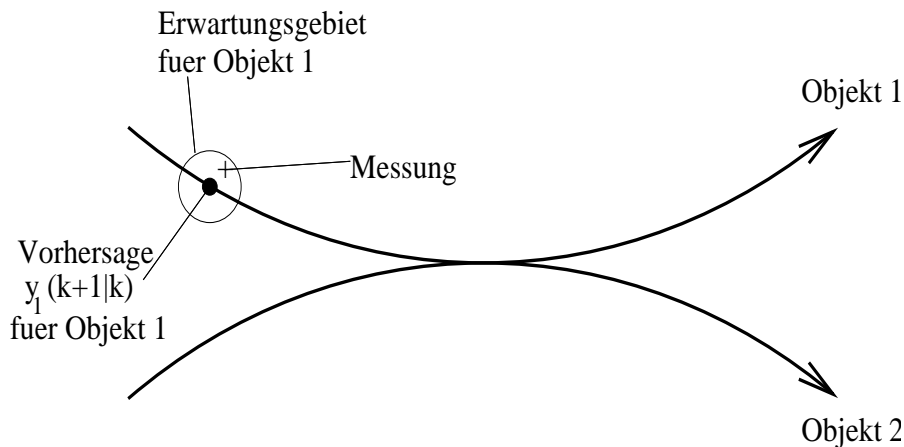


Abbildung 6.14: Zwei kreuzende Ziele vor Beginn einer Interaktion, Anwendung des VBA.

1. da von beiden Objekten jeweils circa zehn Messungen generiert werden können, existieren bei jedem Zeittakt circa $2^{20} \approx 10^6$ mögliche Aufteilungen. Eine solche Anzahl von Assoziationen kann aber von gängigen Rechnern nicht mehr bewältigt werden.
2. dauert die Interaktion längere Zeit an und kommen sich die beiden Objekte dabei sehr nahe, treten dieselben Schwierigkeiten wie beim WMA auf: die Positionsschätzungen für die beiden Ziele werden im Laufe der Interaktion sehr ähnlich (Konvergenz der Tracks) und die Tracks werden sich deshalb am Ende der Interaktion oft beide für die Bahn desselben realen Objektes entscheiden.

Auf den letzteren Umstand der Konvergenz der Tracks wurde im Kontext der Luftüberwachung auch schon von einem sehr namhaften Autor hingewiesen (Blackman, 1991). Vergleiche hierzu auch die Ausführungen in Kapitel 8. Neben diesen beiden Ansätzen könnten, wie in (Kräußling et al., 2004b) aufgezeigt, auch noch weitere analytische Verfahren erdacht werden, die aber ebenfalls nicht praktikabel sind, da sie zum Beispiel zuviel Rechenzeit benötigen oder eine Modellierung der Form der Ziele erfordern, was im Falle von dynamischen Objekten wie Menschen nicht hinreichend genau und stabil möglich ist. Außerdem können auch diese Verfahren in keinem Fall garantieren, dass die beiden Ziele nach der Interaktion wieder separat voneinander verfolgt werden.

Diese Schwierigkeiten sind im Bereich der Mobilien Robotik wohlbekannt, wie die folgenden Zitate belegen:

- Tracking moving objects whose trajectories cross each other is a very general problem . . . Problems of this type cannot be eliminated even by more sophisticated methods¹ . . . (Prassler et al., 1999).

¹Hiermit wird insbesondere auf den JPDAF abgezielt.

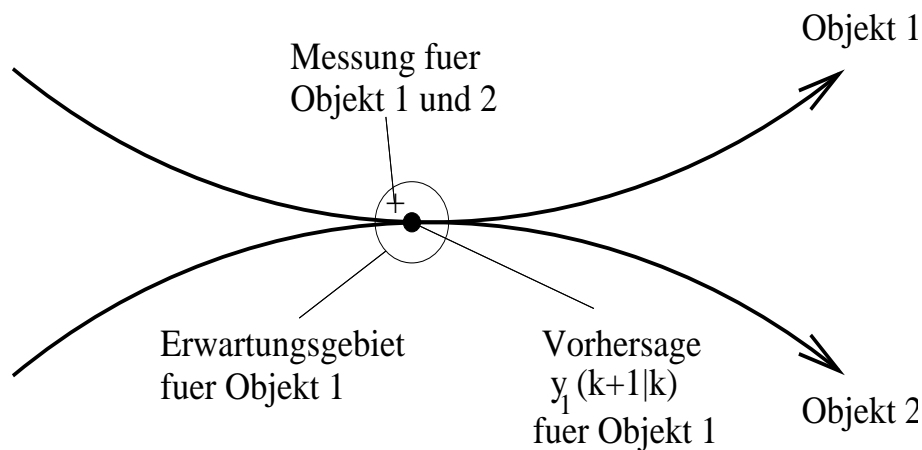


Abbildung 6.15: Zwei interagierende Ziele während einer Interaktion, Anwendung des VBA.

- To solve these problems, existing methods based on tracking each moving object individually are not appropriate (Mucientes und Burgard, 2006).
- Tracks are lost when people walk too closely together ... (Schumitsch et al., 2006a).²

Zusammenfassend lässt sich feststellen, dass das Problem der interagierenden Ziele in der mobilen Robotik eine große Herausforderung für einen Tracking-Algorithmus darstellt.

6.2 Der Cluster-Sorting Algorithmus

Die bisher vorgestellten Ansätze sind also nicht geeignet im Hinblick auf alle drei Kriterien, die von einem Verfahren erfüllt werden sollten, i.e. exakte Information über die Position der Objekte, überschaubare Komplexität und insbesondere stabiles Verhalten bei Crossing Targets. Mit Rücksicht auf diese Umstände ist es gelungen, drei zuverlässige Methoden zu entwickeln, die das Problem der Verfolgung zweier interagierender Ziele im Bereich der mobilen Robotik lösen, d.h. keines der beiden Objekte am Ende der Interaktion verlieren. Diese Algorithmen sind eng miteinander verwandt und benutzen im Wesentlichen die Daten, die der VBA bzw. der MVA liefern und zusätzlich eine einfache Clustering-Routine.³ Vor allem kann der VBA aufgrund seiner komplizierten Gating-Prozedur, die für jede alte Messung eine separate Vorhersage für die neuen Messwerte und ein eigenes Erwartungsgebiet berechnet, ähnlich wie ein

²Zweiter Autor dieser Arbeit ist Sebastian Thrun.

³Einen Überblick über Clustering-Verfahren gibt (Anderberg, 1973).

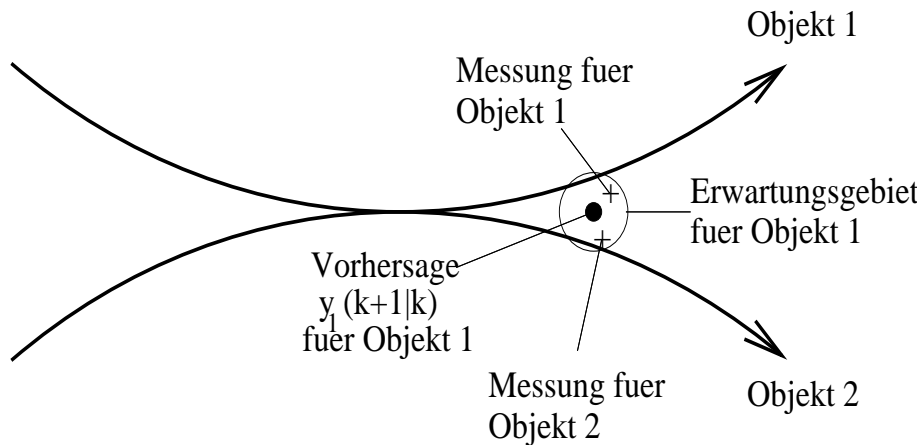


Abbildung 6.16: Zwei kreuzende Ziele am Ende einer Interaktion, Anwendung des VBA.

Teilchen-Filter (Ulmke, 2003) oder speziell der SJPDAF (Schulz et al., 2001b) in gewisser Weise multimodale Wahrscheinlichkeitsverteilungen repräsentieren, was dazu führt, dass sich dieser Ansatz am Ende einer Interaktion nicht für eine Bahn entscheiden muss.⁴ Er kommt dabei auch mit wesentlich weniger 'Teilchen' aus (typischerweise im Bereich von zehn im Vergleich zu einigen hundert bei einem Teilchen-Filter) und außerdem entspricht jedes dieser Teilchen einer Positionsschätzung, die zu einer tatsächlichen Messung gehört. Deshalb enthält diese Positionsschätzung auch noch Information über die Oberfläche und damit die Geometrie des Objektes. Schließlich müssen im Gegensatz zu einem Teilchenfilter keine Gewichte berechnet werden.

Ausgangspunkt der Entwicklung dieser Methoden war die Beobachtung, dass der VBA oder der MVA am Ende einer Interaktion die volle Information über beide Objekte liefern und für beide Tracks alle Messungen von beiden Zielen verarbeiten. Lediglich die Zuordnung erfolgt falsch, d.h. beiden Tracks wird die Bahn desselben realen Objekts zugewiesen. Der erste Punkt beruht auf der komplizierten Gating-Prozedur (vergleiche Kapitel 4), die der VBA benutzt. Insbesondere spielt dabei die Berechnung eines separaten Erwartungsgebietes für jede alte Messung und einer eigenen Positionsschätzung für jede neue Messung die entscheidende Rolle. Demgegenüber berechnen gewöhnliche Verfahren wie der EWA zu jedem Zeitpunkt für jedes Ziel nur ein Erwartungsgebiet.

Diesen Sachverhalt sollen die Abbildungen 6.14 bis 6.17 erläutern. Dabei wird aus Gründen der Übersichtlichkeit angenommen, dass jedes Objekt zu einem bestimmten Zeitpunkt nur der Ursprung einer Messung ist und während der Interaktion beide Ziele nur eine gemeinsame Messung generieren. Außerdem soll lediglich diskutiert werden, wie sich der VBA in Bezug auf einen Track verhält. Vor der Interaktion (Abbil-

⁴Natürlich gilt dies auch für den MVA. Aus Gründen der Übersichtlichkeit der Darstellung wird im Folgenden aber oft nur auf den VBA Bezug genommen.

dung 6.14) liegt also eine Messung für das Objekt 1 vor, die in dem Erwartungsgebiet liegt, das der Vorhersage $y_1(k+1|k)$ entspricht. Außerdem werden die beiden Objekte noch unabhängig voneinander verfolgt. Während der Begegnung liegt eine gemeinsame Messung vor und für das Objekt 1 wird wieder genau ein Erwartungsgebiet berechnet (Abbildung 6.15). In der Realität werden natürlich während des Crossings alle Messungen von beiden Objekten gleichzeitig verarbeitet. Die nächsten beiden Graphiken zeigen dann das entscheidende Phänomen. Am Ende des Crossings, wenn sich die beiden Ziele schon wieder etwas voneinander entfernt haben, liegen zwei Messwerte im Erwartungsgebiet für das Objekt 1, und zwar stammt eine Messung von Objekt 1, die andere dagegen von Objekt 2 (Abbildung 6.16). Für diese beiden Messungen wird nun jeweils ein separates Erwartungsgebiet berechnet, von denen das eine dann mehr auf der Bahn von Objekt 1 liegt, während sich das andere mehr auf der Bahn von Objekt 2 befindet. Deshalb werden im nächsten Schritt, wenn die Distanz zwischen den beiden Zielen schon größer ist, sowohl die Messung von Objekt 1 als auch diejenige von Objekt 2 erfasst (Abbildung 6.17).

In der Realität, wenn für jedes Ziel mehrere Messwerte vorliegen, 'verfolgt' der VBA also im Prinzip nach dem Crossing zunächst beide Objekte gleichzeitig, wobei jedes Objekt durch eine Wolke oder einen Cluster von Positionsschätzungen repräsentiert wird. Der VBA gibt somit eine Art bimodale Verteilung wieder. Er ordnet dann lediglich fälschlicherweise beide Tracks demselben Cluster zu wie in Abbildung 6.6 oder 6.8, d.h. nur die Zuweisung der Cluster zu den Tracks ist falsch.

Deshalb wurden drei stabile Methoden entwickelt, die nach der Interaktion jedem Track einen eigenen Cluster zuordnen. Sie benutzen die Daten, die der VBA bereitstellt, bevor ein bestimmter Pfad ausgewählt wird, und gewisse geometrische Eigenschaften, die das Problem charakterisieren. Sie unterscheiden sich dabei im Wesentlichen nur im Rechenbedarf. Das zunächst entwickelte Verfahren wurde Cluster-Sorting Algorithmus (CSA) genannt, da es wie die anderen beiden Ansätze auch am Ende einer Interaktion eine bijektive Zuordnung der Cluster oder Features zu den computerinternen Tracks erzwingt. Es wurde in (Kräußling et al., 2004b) eingeführt und soll nun beschrieben werden.

Der CSA macht also Gebrauch vom Verhalten des VBA und besteht aus den folgenden drei Schritten. Er nutzt dabei die geometrischen Muster aus, die Lasermessungen auszeichnen, die von zwei dicht benachbarten Personen stammen. Der erste Schritt wird zu jedem Zeitpunkt k ausgeführt, die anderen beiden dagegen nur, falls interagierende Ziele vorliegen:

1. Für jedes Paar von Tracks wird geprüft, ob zwischen den beiden zugehörigen Objekten soeben eine Interaktion begonnen hat. Die Ergebnisse werden dann in einer binären Matrix A in Form einer oberen Dreiecksmatrix gespeichert. Das Element a_{ij} mit $j \geq i$ dieser Dreiecksmatrix wird dabei auf Eins gesetzt, falls die Erwartungsgebiete der beiden Ziele i und j sich überlappen, d.h. mindestens eine gemeinsame Messung in den Erwartungsgebieten beider Objekte liegt (Abbildung 6.18). Dies bedeutet gemäß der Definition auf Seite 79 ja gerade, dass

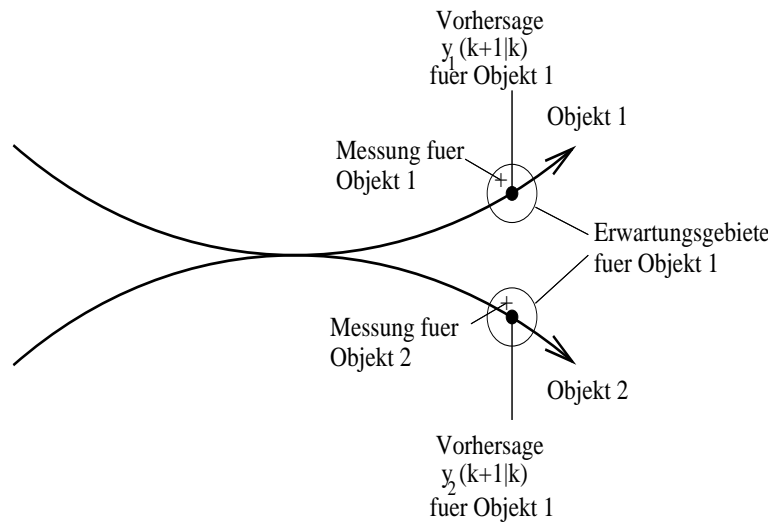


Abbildung 6.17: Zwei interagierende Ziele nach einer Interaktion, Anwendung des VBA.

zwischen den Objekten i und j ein Kreuzungsvorgang begonnen hat. Ansonsten wird das Element a_{ij} auf Null belassen. Sobald ein Kreuzen festgestellt wurde, behält das Element a_{ij} den Wert Eins solange, bis in Schritt zwei das Ende der Interaktion festgestellt wird.

2. Für jedes Paar von interagierenden Zielen, d.h. für jedes Paar von Objekten i und j mit $a_{ij} = 1$, wird Schritt Eins nicht mehr durchgeführt. Vielmehr wird bei jedem folgenden Scan untersucht, ob die Interaktion soeben beendet wurde. Dies geschieht dadurch, dass überprüft wird, ob die Messungen, die durch den VBA mit den Zielen assoziiert wurden, sich bereits wieder in zwei unterschiedliche Cluster mit einem Euklidischen Abstand oberhalb einer gewissen Schranke aufgeteilt haben. Bei der Verfolgung von Menschen kann dies angenommen werden, wenn zwei Positionsschätzungen vorliegen, die mindestens 300 Zentimeter voneinander entfernt sind (Abbildung 6.19). Dieser Wert wurde deshalb gewählt, da angenommen werden kann, dass Messwerte, die zu den zwei Beinen eines gehenden Menschen gehören, deutlich weniger als 150 Zentimeter auseinander liegen. Diese Wahl der Schranke garantiert also, dass nicht in beiden Clustern Messungen von derselben Person enthalten sein können. Bei den experimentellen Untersuchungen zeigte sich darüberhinaus, dass der Wert von 300 Zentimetern sogar so groß ist, dass auch garantiert ist, dass sich die Erwartungsgebiete der beiden Personen nicht mehr überlagern, nachdem das Ende der Interaktion festgestellt worden ist. Auf diese Weise wird der Fall vermieden, dass sich die beiden Ziele noch so nahe sind, dass sofort wieder gemäß Schritt Eins ein Crossing detektiert wird.

3. Sobald das Ende eines Kreuzungsvorganges festgestellt worden ist, müssen die beiden zugehörigen Cluster von Positionsschätzungen getrennt und den beiden Tracks zugeordnet werden (Abbildung 6.20). Dazu wird zunächst eine beliebige Positionsschätzung dem Track mit dem niedrigeren Index zugeordnet. Danach wird für jede weitere Messung bzw. zugehörige Positionsschätzung, die durch den VBA mit den beiden Zielen assoziiert wurde, berechnet, ob der Euklidische Abstand zu dieser ersten Schätzung kleiner oder größer als 150 Zentimeter ist. Auch hier wird der Wert von 150 Zentimetern benutzt, da zwei Messungen mit einem Euklidischen Abstand größer als 150 Zentimeter nicht von derselben Person herrühren können. Im ersten Fall wird die neue Schätzung dann dem Track mit dem niedrigeren Index und im zweiten Fall dem Track mit dem höheren Index zugewiesen. Aufgrund der im zweiten Schritt gewählten Schranke werden auf diese Weise jedem Track nur Messungen von genau einer Person zugeteilt.

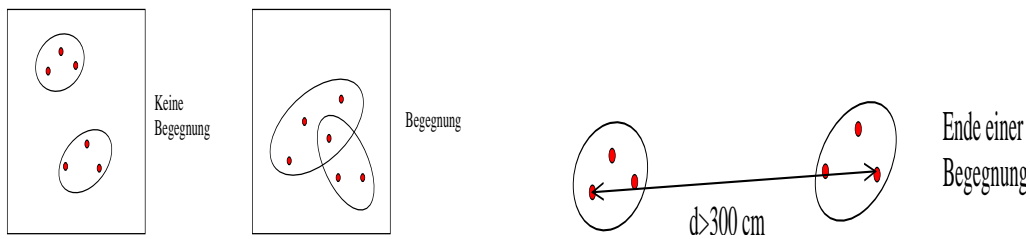


Abbildung 6.18: Beginn einer Interaktion.

Abbildung 6.19: Ende einer Interaktion.

Cluster-Sorting Algorithmus, Schritte 1 und 2.

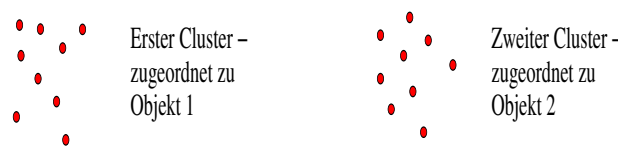


Abbildung 6.20: Cluster-Sorting Algorithmus, Schritt 3: Zuschreibung der Cluster zu den computerinternen Objekten oder Tracks.

Am Ende des Datensatzes wird schließlich ähnlich wie beim ursprünglichen VBA oder MVA für jeden Track der Pfad mit der minimalen Länge oder dem Index Eins bestimmt und anschließend rekursiv die zugehörige Bahn konstruiert. Außerdem kann auch noch der Kalman-Glätter (Shumway und Stoffer, 2000) angewandt werden. Auf

diese Weise erhält man für jedes Objekt eine zusammenhängende Bahn. Dieser Ansatz ist allerdings nicht für Online-Anwendungen geeignet, da der Datensatz zu Beginn des Verarbeitungsprozesses komplett vorliegen muss.

Natürlich kann bei dieser Vorgehensweise, da im dritten Schritt die erste Schätzung frei gewählt wird, der Fall auftreten, dass die beiden Ziele nach der Interaktion miteinander vertauscht sind. Es gibt aber für dieses Problem der richtigen Zuordnung keine generelle Lösung, die lediglich die Abstandsmessungen eines Laser-Scanners verwendet, da die Bewegungen von Menschen unvorhersagbar sind. Eine Verbesserung dieses Ansatzes könnte erreicht werden, wenn zusätzliches Wissen über die Objekte zur Verfügung steht. Zum Beispiel könnte im Falle der Verfolgung eines Menschen und eines Roboters die unterschiedliche Form der Ziele dazu beitragen, die richtige Zuordnung zu treffen. Man könnte sich auch vorstellen, dass der reflektierte Laserstrahl unterschiedliche Intensität besitzt, falls zwei Personen Hosen unterschiedlicher Farbe oder Textur tragen. Diese Fragestellungen werden in Kapitel 9 näher untersucht.

Abbildung 6.21 veranschaulicht, dass der CSA in der Lage ist, das Problem der Interaktion zweier Objekte im Falle der oben eingeführten schwierigeren realen Daten zufriedenstellend zu lösen, d.h. keines der beiden Ziele geht verloren. Außerdem ist in diesem Falle auch noch die Zuordnung nach dem Crossing korrekt. Abbildung 6.22 zeigt dagegen ein Beispiel für eine falsche Assoziierung, wie man an der abrupten, unnatürlichen Richtungsänderung der beiden Tracks erkennen kann. Wiederum werden den beiden Personen nach der Interaktion aber zwei unterschiedliche reale Bahnen zugeordnet.

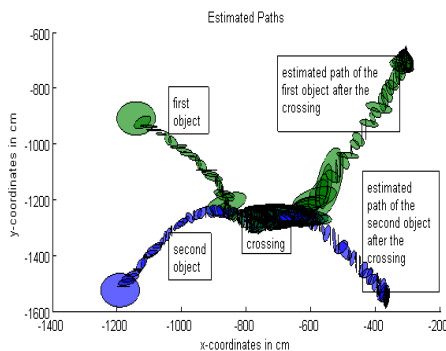


Abbildung 6.21: Zwei interagierende Ziele, reale Daten, verarbeitet mit dem CSA, richtige Zuordnung nach dem Crossing.

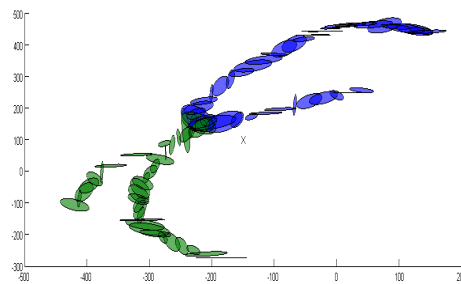


Abbildung 6.22: Zwei interagierende Ziele, reale Daten, verarbeitet mit dem CSA, falsche Zuordnung nach dem Crossing.

Zwei interagierende Ziele, reale Daten, verarbeitet mit dem CSA.

Die darauffolgende Abbildung 6.23 veranschaulicht noch einmal die Verhältnisse beim CSA für insgesamt vier aufeinanderfolgende Interaktionen. Am Beginn eines jeden Kreuzungsvorganges steigt jeweils die Anzahl der Messungen, die mit den beiden

Zielen assoziiert werden, steil an, da auch die Messwerte von dem jeweils anderen Objekt mitverarbeitet werden. Dieses Verhalten veranschaulichen die oberen beiden Graphen. Während der Interaktion verarbeitet der CSA dann für beide Tracks alle Messungen von beiden Zielen, so dass die Differenz dieser Messwerte konstant gleich Null ist (unterste Abbildung). Sobald der CSA die beiden Objekte am Ende des Crossings wieder voneinander trennt, fällt die Anzahl der verarbeiteten Messungen schließlich für beide Tracks steil ab.

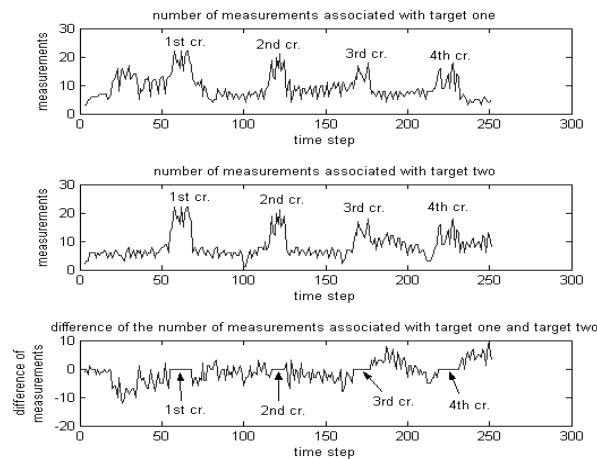


Abbildung 6.23: Zwei interagierende Ziele, reale Daten, vier aufeinanderfolgende Begegnungen, verarbeitet mit dem CSA.

6.3 Der Switching Algorithmus

Da der CSA in der Lage ist, das Problem zweier interagierender Objekte zu lösen, könnte er im Prinzip für den gesamten Tracking-Prozess verwendet werden. Da dieser Algorithmus auf dem VBA basiert, liefert er aber eine ungenauere Information über die Position der Ziele als der EWA, solange keine Interaktion auftritt. Außerdem benötigt er wesentlich mehr Rechenzeit und Speicherplatz, was in Anbetracht der begrenzten Rechenkapazitäten von on-board Einheiten einen gravierenden Nachteil darstellt (vergleiche auch Kapitel 5). Deswegen sollte der VBA bzw. der CSA nicht für die Verfolgung von einzelnen Personen benutzt werden. Deshalb wurde als Verbesserung ein Switching Algorithmus entwickelt, der auf den CSA nur zurückgreift, solange tatsächlich der Fall eines Crossings vorliegt (Kräußling et al., 2005a; Kräußling et al., 2007; Kräußling, 2008a). Während der restlichen Zeit wird der EWA eingesetzt, der ja für jedes Objekt nur eine Kalman-Filterung durchführt und deshalb, wie auch der Verfahrensvergleich in Kapitel 5 gezeigt hat, einer der schnellsten Algorithmen ist und eine sehr genaue Information über die Position der Ziele liefert. Auf diese

Weise werden die Vorteile des EWA und des CSA miteinander vereinigt und Einzelziele sozusagen mit unabhängigen Kalman-Filtern verfolgt. Dabei wird der Beginn einer Interaktion wie beim ersten Schritt des CSA detektiert. Am Ende der Begegnung wird dann jeweils aus dem entsprechenden Cluster eine Schätzung ausgewählt und anschließend rekursiv über die Dauer der Interaktion der zugehörige Pfad bestimmt. Die entsprechenden Schätzungen werden dann als Positionsschätzung für die beiden Objekte während der Interaktion benutzt. Auf diese Weise entstehen auch keine Lücken in den graphischen Darstellungen, wie sie später beim ECSA auftreten. Abbildung 6.24 zeigt das Flussdiagramm dieses Switching Algorithmus (SA), der so genannt wurde, weil er je nach der aktuellen Situation zwischen zwei einfacheren Algorithmen hin und her springt.

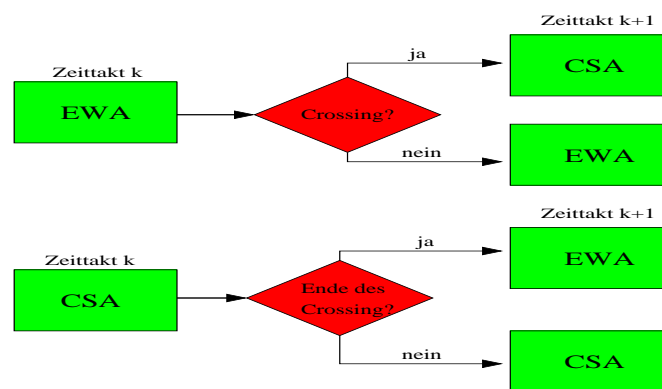


Abbildung 6.24: Flussdiagramm des Switching Algorithmus.

Abbildung 6.25 zeigt unter Verwendung der schon bekannten realen Daten, dass der Switching Algorithmus das Problem zweier interagierender Ziele genauso gut beherrscht wie der CSA. Um die Vorteile des Switching Algorithmus noch ausführlicher zu demonstrieren, wurden weitere Experimente mit realen Daten durchgeführt. Dabei liefen zwei Personen in unserer Experimentalhalle umher. Es wurden insgesamt fünf ähnliche Szenarien aufgezeichnet. In jedem Szenario liefen die beiden Personen am Anfang des Experimentes für ein Zeitintervall t_1 getrennt voneinander. Dann begegneten sich die beiden Personen und liefen für ein Zeitintervall t_2 gemeinsam, so dass eine Interaktion stattfand. Schließlich trennten sich die beiden Personen wieder voneinander und liefen erneut für ein Zeitintervall t_3 getrennt. Dabei betrug die Länge des Zeitintervalls t_2 für alle Szenarien ungefähr dreißig Sekunden. Darüberhinaus waren die Zeitintervalle t_1 und t_3 jeweils ungefähr von derselben Länge, die von ca. dreißig Sekunden in Schritten von dreißig Sekunden auf schließlich 150 Sekunden anstieg. Abbildung 6.26 zeigt ein Beispiel aus Szenario Eins für die geschätzten Bahnen am Ende der Begegnung bei Anwendung des Switching Algorithmus. Man erkennt, dass die beiden Versuchspersonen am Ende der Interaktion wieder getrennt verfolgt werden. Insgesamt scheiterten die einfachen EWA und VBA in allen fünf Szenarien, während

der CSA und der Switching Algorithmus jeweils nach der Interaktion für beide Objekte separate Bahnen berechnen.

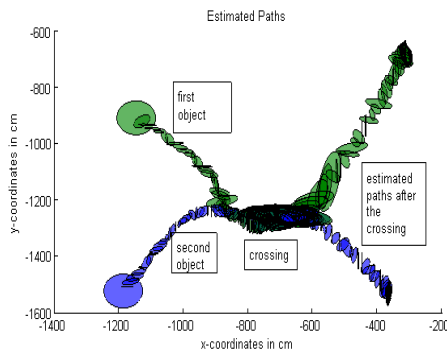


Abbildung 6.25: Zwei interagierende Ziele, reale Daten, verarbeitet mit dem Switching Algorithmus.

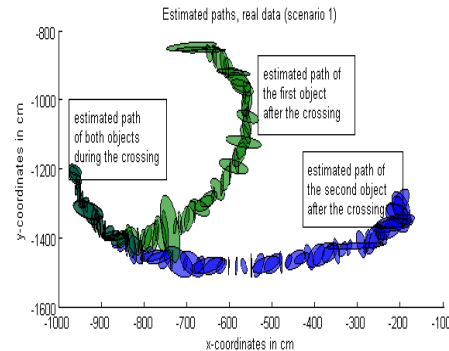


Abbildung 6.26: Zwei interagierende Ziele, reale Daten von Szenario 1 am Ende der Interaktion, verarbeitet mit dem Switching Algorithmus.

Zwei interagierende Ziele, reale Daten, verarbeitet mit dem Switching Algorithmus (entnommen im englischen Original aus (Kräußling et al., 2005a)).

Tabelle 6.1 zeigt die Einsparung an Rechenzeit bei Verwendung des Switching Algorithmus. Sie gibt die durchschnittliche Zeit, die es für die Ausführung der Berechnungen für einen Zeittakt bedurfte, in Millisekunden an. Die Tabelle zeigt die Verbesserung im Vergleich zum CSA, die erreicht werden kann, wenn man den Switching Algorithmus benutzt. Der Gewinn wächst dabei rapide, wenn die Zeitintervalle t_1 und t_3 , während derer der Switching Algorithmus den EWA benutzt, größer werden. Die Tatsache, dass der VBA gemäß der Tabelle mehr Rechenzeit benötigt als der CSA, erklärt sich dabei dadurch, dass der VBA nach einer Interaktion für beide Tracks alle Messungen von beiden Objekten verarbeitet. Der VBA verbraucht nach einer Interaktion also fast doppelt so viel Rechenzeit wie der CSA, der die beiden Ziele ja sofort wieder voneinander trennt.

Tabelle 6.1: Bedarf an Rechenzeit in Millisekunden.

Experiment	1	2	3	4	5
EWA	54.1	53.8	54.6	53.9	54.4
VBA	379.0	355.7	478.6	366.3	469.3
CSA	294.3	258.6	329.5	260.7	320.3
SA	171.3	123.7	119.2	110.8	91.5

Eine weitere Verbesserung könnte noch darin bestehen, den VBA während der Interaktion nur auf einen Track anzuwenden. Da der CSA für beide Tracks während der

Begegnung schon nach wenigen Zeittakten alle Messungen beider Objekte verarbeitet, ist dies möglich, ohne dass dabei Information verlorengeht. Konkret wird dabei so verfahren, dass die Berechnungen nur noch für einen Track durchgeführt werden, sobald der VBA für beide Tracks dieselben Messwerte verarbeitet. Dies ist in der Regel schon nach sehr wenigen Zeittakten der Fall. Abbildung 6.27 zeigt die Anwendung dieses Verfahrens auf die schon benutzten Daten. Auch dieser Algorithmus löst also das Problem zweier interagierender Ziele. Dieser verbesserte Switching Algorithmus ist dem später vorgestellten ECSA für den Spezialfall zweier Objekte auch schon sehr ähnlich.

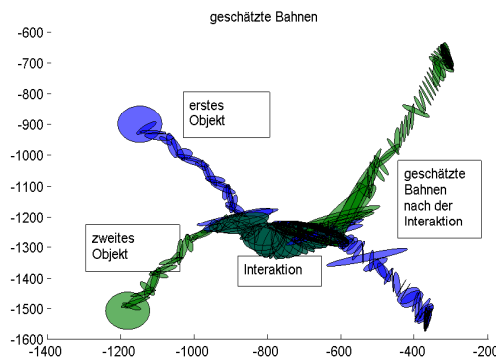


Abbildung 6.27: Zwei interagierende Ziele, reale Daten, verarbeitet mit dem verbesserten Switching Algorithmus.

Abschließend wurde dieser verbesserte Switching Algorithmus noch anhand der fünf Szenarien mit dem ursprünglichen Switching Algorithmus bezüglich der Laufzeit verglichen. Zunächst ist dabei festzustellen, dass auch das verbesserte Verfahren das Problem der interagierenden Ziele in allen fünf Szenarien löst, d.h. keines der beiden Objekte geht am Ende der Interaktion verloren. Die Einsparung an Rechenzeit ist besonders groß in Szenario Eins, da hier der Anteil der Interaktion am gesamten Experiment am größten ist (Abbildung 6.28), und beträgt dort circa 35 Prozent. Da der Switching Algorithmus während einer Interaktion im Wesentlichen alle Berechnungen doppelt ausführt, benötigt der verbesserte Algorithmus in dieser Phase im Vergleich zum SA nur etwas mehr als die Hälfte der Rechenzeit, wie in Abbildung 6.29 dargestellt ist.

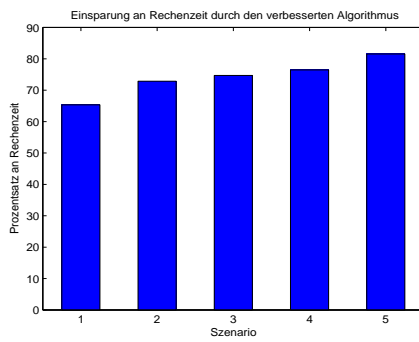


Abbildung 6.28: Prozentsatz des Rechenzeitverbrauchs für den verbesserten Algorithmus im Vergleich zum SA während des gesamten Versuchs.

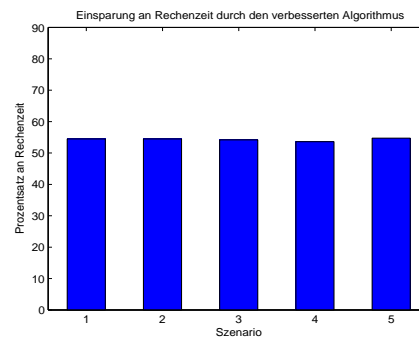


Abbildung 6.29: Prozentsatz an Rechenzeitbedarf für den verbesserten Algorithmus im Vergleich zum SA während der Begegnung.

Zwei interagierende Ziele, Vergleich des Rechenzeitbedarfs des Switching Algorithmus und des verbesserten Switching Algorithmus mit Hilfe der fünf Szenarien.

Kapitel 7

Die Verfolgung beliebig vieler interagierender Ziele — ein erweiterter Cluster–Sorting–Algorithmus

7.1 Einleitung

Bisherige Ansätze zur Verfolgung mehrerer Personen gehen so vor, dass sie aus Lasermessungen sogenannte Features in Form von lokalen Abstandsminima extrahieren, die den Personen entsprechen sollen, und diese Features dann den computerinternen Tracks zuordnen. Diese Assoziierung, die eines der zentralen Probleme im Falle des Multi–People–Tracking darstellt (Frank et al., 2003), kann dabei sowohl allein auf Positions– (Schulz et al., 2003a) als auch zusätzlich auf Identitätsdaten (Bellotto und Hu, 2007a; Bellotto und Hu, 2007b) beruhen und wird auch als Position–Fusion bezeichnet (Hall, 1992). Die gleichzeitige Verwendung von Identitätsmessungen kann insbesondere dann sinnvoll sein, wenn aufgrund einer hohen Objektdichte Auflösungskonflikte bestehen.

Bei dieser Zuschreibung wird mit Hilfe eines Abstands– oder Ähnlichkeitsmaßes eine Metrik definiert und darauf aufbauend eine Assoziationsmatrix berechnet. Dies geschieht durch einen Vergleich der beobachteten Features mit den Vorhersagen, die zu den computerinternen Tracks gehören. Diese Vorhersagen werden dabei aufgrund des Sensormodells und des Dynamikmodells der Personen berechnet. Dieser Vergleich kann auch vor der eigentlichen Assoziierung dazu herangezogen werden, durch ein sogenanntes Gating die Anzahl der zu betrachtenden Paare aus Features und Vorhersagen zu reduzieren. Die Zuteilung kann dabei beispielsweise jeweils unmittelbar nach jeder Messung erfolgen bzw. in Form von nur einer einzigen Hypothese vorliegen. Gelegentlich wird auch eine bijektive Zuordnung angestrebt (Bellotto und Hu, 2006). Ansonsten besteht bei hoher Objektdichte die Neigung, dasselbe Feature mehreren

Tracks zuzuweisen. Dies entspricht einem Verlust derjenigen Ziele, die eigentlich zu anderen Features gehören (Kräußling und Schulz, 2006). Beispiele für mögliche Zuordnungsansätze sind der Nearest-Neighbour-Algorithmus (Bellotto und Hu, 2007b) und der JPDAF (Frank et al., 2003; Schulz et al., 2003a). Es wurde allerdings schon früh für den Bereich der Luftüberwachung von einem namhaften Autoren festgestellt, dass beispielsweise der JPDAF im Falle dichter Bahnen zur Konvergenz dieser Tracks, d.h. zum Trackverlust, führen kann (Blackman, 1991). Vergleiche hierzu auch die Kapitel 6 und 8.

Wie in Kapitel 6 gezeigt wird, ist schon die Verfolgung zweier Personen, die sich sehr nahe kommen, ein sehr schwieriges Problem, das von einfachen Algorithmen nicht mehr zufriedenstellend gelöst werden kann. Die Gefahr des Verlusts von Tracks ist umso mehr im allgemeinen Fall beliebig vieler Personen gegeben, insbesondere dann, wenn die Personen so nahe beieinander sind, dass sie nicht mehr einzeln in Features aufgelöst werden können. So kann es zum Beispiel schwierig sein, zu entscheiden, ob zwei dicht benachbarte Features den Beinen eines Menschen entsprechen oder von zwei verschiedenen Personen herrühren. In diesem Falle ist eine Assoziierung der gemessenen Features zu den computerinternen Objekten im Allgemeinen nicht mehr zuverlässig möglich. Außerdem können bei hoher Objektdichte auch Tracks vertauscht werden, so dass die entsprechenden Identitäten neu zugewiesen werden müssten. Bei Verfahren, die Interaktionen nicht explizit detektieren, fehlt aber dann das Wissen, wann dies der Fall ist. Deshalb müssen die Identitäten in diesem Falle permanent bestimmt werden, was einen hohen Aufwand darstellt. Auch deshalb gilt das Multiple-Target-Tracking als eine sehr schwierige Aufgabe (Hall, 1992; Hall und Llinas, 2001). Erste Anwendungen für dieses Multi-Target-Tracking stammen im Übrigen bereits aus dem Jahre 1795, als Gauß mit Hilfe der Methode der kleinsten Quadrate, die als ein Vorläufer des Kalman-Filters bezeichnet werden kann, die Bahnen von Asteroiden, Kometen und Planeten berechnete.

Der hier vorgestellte Ansatz betrachtet deshalb im Falle einer Interaktion keine Einzelziele oder Features mehr, sondern bildet Gruppen von Beobachtungen oder Personen oder auch Cluster. Diese Vorgehensweise wurde auch schon in (Rosencrantz et al., 2003) oder (Schulz et al., 2003a; Schulz et al., 2003b) vorgeschlagen. In (Frank et al., 2003) wird auch darauf hingewiesen, dass es sinnvoll sein kann, diese Gruppen mit einem stabileren und aufwendigeren Verfahren zu verfolgen als Einzelziele. Im Rahmen der hier vorgestellten Methode wird nun der eigentliche Interaktionsprozess gewissermaßen als eine Art Black Box betrachtet. Während dieser Phase wird also nicht versucht, auf analytischem Wege die exakten Bahnen einzelner Personen zu bestimmen. Es wird vielmehr lediglich Sorge dafür getragen, dass während der Interaktion ausreichend Information erhalten bleibt, so dass am Ende einer Begegnung keine Person verloren geht. Dazu werden alle Messwerte, die der Gruppe zugeordnet werden müssen, in geeigneter Form verarbeitet. Auch die Zuordnung der richtigen Messungen zu einer Gruppe ist dabei ein Problem, das in geeigneter Weise gelöst werden muss. Sobald die Interaktion beendet ist, werden dann mit Hilfe der gewonnenen Information die jetzt wieder separaten Bahnen der Einzelpersonen erneut aufgegriffen. Dazu wer-

den jeweils diejenigen Messwerte, die vermutlich von einem und demselben Objekt stammen, geclustert. Außerdem können die Identitäten der Personen mit den in Kapitel 9 vorgestellten Verfahren neu zugewiesen werden. Insbesondere wird durch diese Vorgehensweise das schwierige Problem der Daten–Assoziierung, d.h. der Zuordnung der gemessenen Features zu den computerinternen Tracks, umgangen. Darüberhinaus kann es, wie schon erwähnt, überhaupt schwierig sein, Features von einzelnen Personen aus den Daten zu extrahieren, wenn Gruppen von Personen vorliegen.

Da der folgende Ansatz eine Verallgemeinerung des im vorhergehenden Kapitel vorgestellten Cluster–Sorting Algorithmus bzw. der beiden Switching Verfahren darstellt, die jeweils speziell für den Fall zweier Objekte konzipiert waren, wird er als erweiterter Cluster–Sorting Algorithmus (ECSA) bezeichnet. Er wurde in (Kräußling, 2006b; Kräußling, 2008c) vorgestellt und benutzt wie der CSA einen adaptierten Viterbi–Algorithmus und geometrische Charakteristika des Problems. Gegenüber den in Kapitel 6 vorgestellten Ansätzen hat das Verfahren auch noch den zusätzlichen großen Vorteil, dass es online–fähig ist und somit in Echtzeitexperimenten eingesetzt werden kann. Es bildet deshalb die Basis für die in Kapitel 11 beschriebenen Versuchsreihen.

7.2 Das Verfahren

Der erweiterte Cluster–Sorting Algorithmus (ECSA) benutzt zwei Klassen von Objekten, nämlich:

- Einzelziele.
- Cluster oder Gruppen von Personen, die mindestens zwei interagierende Menschen oder Objekte repräsentieren. Interagierende Ziele sind hierbei Objekte, die sich sehr nahe zueinander bewegen. In einem Cluster sind also mindestens zwei Einzelpersonen miteinander verschmolzen.

Einzelobjekte werden ähnlich wie beim Switching Algorithmus mit dem EWA verfolgt, da hier keine multimodalen Wahrscheinlichkeitsverteilungen repräsentiert werden müssen. Außerdem benötigt dieser Ansatz wenig Rechenzeit und Speicherplatz und gibt eine sehr genaue Schätzung über die Position des zu verfolgenden Zieles. Cluster werden dagegen mit dem VBA bzw. dem MVA verfolgt, da hier multimodale Wahrscheinlichkeitsverteilungen abgebildet werden müssen, um den Verlust von Tracks zu vermeiden. Auf diese Weise liegt stets die vollständige Information über alle Messungen, die zu der Gruppe gehören, vor. Ein analoges Verhalten der Abbildung von multimodalen Verteilungen zeigen auch gewisse, in der Einleitung beschriebene Verfahren zur Lokalisierung von mobilen Robotern, nämlich Markov– (Burgard et al., 1996; Kaelbling et al., 1996; Nourbakhsh et al., 1995; Simmons und Koenig, 1995) und Monte–Carlo–Lokalisierung (Dellaert et al., 1999; Fox et al., 1999b). Im Übrigen vergleiche zu den Charakteristika des EWA und des VBA auch die Kapitel 4 bis 6. Diese Vorgehensweise garantiert, dass keines der Objekte, die mit einem Cluster assoziiert

sind, verloren geht. Dieser Umstand ist dann von besonderer Bedeutung, wenn die Ziele sich voneinander trennen und damit beginnen, sich wieder unabhängig voneinander zu bewegen.

Drei verschiedene Ereignisse müssen beachtet werden, wenn mehrere interagierende Objekte verfolgt werden sollen (vergleiche dazu auch den Pseudocode in Tabelle 7.1):

1. Die Vereinigung zweier Einzelobjekte zu einem Cluster. Dies bedeutet, dass sich zwei einzelne Personen sehr nahe kommen. Das ist genau dann der Fall, wenn zumindest eine Messung in den Erwartungsgebieten beider Einzelobjekte liegt. Dann hört der ECSA auf, die beiden Einzelobjekte getrennt mit dem EWA zu verfolgen und beginnt dafür, einen Cluster, der beide Einzelobjekte repräsentiert, unter Verwendung des VBA zu verfolgen. Bei dieser Umschaltung werden alle diejenigen Messwerte berücksichtigt, die im Erwartungsgebiet zumindest eines der beiden Einzelobjekte liegen.
2. Die Vereinigung eines Einzelobjektes und eines Clusters zu einem kombinierten Cluster. Dies bedeutet, dass eine einzelne Person und ein Cluster sich sehr nahe kommen. Das ist genau dann der Fall, wenn mindestens eine Messung in den Erwartungsgebieten sowohl des Einzelobjektes als auch des Clusters liegt. In diesem Falle hört der ECSA auf, den Cluster und das Einzelobjekt unabhängig voneinander zu verfolgen. Anstatt dessen beginnt er einen kombinierten Cluster zu verfolgen, wobei ein kombinierter Cluster letztlich nichts anderes ist als ein einfacher Cluster, wie er oben eingeführt wurde. Bei dieser Transformation werden alle diejenigen Messwerte verwendet, die in den Erwartungsgebieten entweder des Einzelobjektes oder des zuvor betrachteten Clusters liegen oder in beiden.
3. Die Vereinigung zweier Cluster zu einem kombinierten Cluster. Dies bedeutet, dass sich zwei Cluster sehr nahe kommen. Das ist genau dann der Fall, wenn zumindest eine Messung in den Erwartungsgebieten beider Cluster liegt. Falls dies gilt, hört der ECSA auf, die beiden Cluster unabhängig voneinander zu verfolgen und beginnt vielmehr, einen kombinierten Cluster zu verfolgen. Dazu benutzt er alle Messwerte, die in den Erwartungsgebieten zumindest eines der beiden zuvor betrachteten Cluster liegen.

Es ist bei all diesen Schritten zu beachten, dass sich das Verfahren immer dann, wenn eine Vereinigung stattfindet, diejenigen Einzelobjekte merkt, die zu dem neu gebildeten Cluster gehören.

Für jeden Cluster wird sodann fortlaufend überprüft, ob er schon wieder in seine Einzelpersonen zerfallen ist. Zu diesem Zweck wurden zwei Bedingungen definiert, die von dem Algorithmus geprüft werden:¹

¹Dabei muss beachtet werden, dass mit einem Cluster mehrere Erwartungsgebiete assoziiert sind, da ja Cluster mit dem VBA verfolgt werden.

<i>Merging of objects</i>	
<i>n single objects, m clusters</i>	
1.	<i>Merging of single objects</i>
1.1	For each pair of single objects
1.2	If a measurement belongs to both single objects
1.3	Merge the two single objects into a cluster
1.4	End
1.5	End
2.	<i>Merging of single objects and clusters</i>
2.1	For each pair of a single object and a cluster
2.2	If a measurement belongs to both the single object and the cluster
2.3	Merge the single object and the cluster into a (combined) cluster
2.4	End
2.5	End
3.	<i>Merging of clusters</i>
3.1	For each pair of clusters
3.2	If a measurement belongs to both clusters
3.3	Merge the two clusters into a (combined) cluster
3.4	End
3.5	End

Tabelle 7.1: Mögliche Interaktionen zwischen den Objekten, Pseudocode.

- Die Positionsschätzungen, die zu den Messungen in den Erwartungsgebieten gehören, werden in Subcluster unterteilt.² Ein Subcluster ist dabei eine Menge von Schätzungen mit einem paarweisen Abstand unterhalb einer gewissen Schranke. Zu diesem Zweck wird die erste Schätzung dem ersten Subcluster zugewiesen. Für alle anderen Schätzungen, die mit dem Cluster assoziiert sind, wird dann der Euklidische Abstand zu dieser ersten Schätzung berechnet. Falls diese Entfernung unterhalb einer gewissen Schranke liegt, wird die jeweilige Schätzung ebenfalls dem ersten Subcluster zugeteilt. In den entsprechenden Experimenten wurde diese Schranke auf 150 Zentimeter festgesetzt, was eine obere Schranke für die maximale Entfernung der beiden Beine einer gehenden Person darstellt. Somit bedeutet diese Vorgehensweise, dass Schätzungen, die in unterschiedliche Subcluster einsortiert werden, nicht von derselben Person stammen

²Den Hinweis, dass an dieser Stelle Clustering-Methoden zum Erfolg führen könnten, verdanke ich meinem damaligen Projektleiter Frank E. Schneider (Schneider, 2006).

können. Danach müssen die Schätzungen betrachtet werden, für die der Abstand zu der ersten Schätzung größer ist als diese Schranke, und die deshalb nicht dem ersten Subcluster zugeordnet wurden. Unter Anwendung derselben Vorgehensweise, mit der der erste Subcluster gebildet wurde, werden dann weitere Subcluster konstruiert, bis schließlich alle Schätzungen genau einem Subcluster zugeordnet sind. Falls die Anzahl dieser Subcluster gleich der Anzahl der Einzelpersonen ist, die mit dem betrachteten Cluster assoziiert sind, ist die erste Bedingung für die Auflösung eines Clusters erfüllt. Dann wird mit Schritt zwei fortgefahren.

2. Es wird nun paarweise der Abstand zwischen den einzelnen Subclustern untersucht. Der Abstand zwischen zwei Subclustern ist dabei definiert als das Maximum aller Euklidischen Distanzen zwischen zwei Schätzungen, wobei eine dieser beiden Schätzungen zu dem ersten Subcluster und die andere zu dem zweiten Subcluster gehören muss. Falls dieser Abstand oberhalb einer geeigneten gewählten Schranke ist, werden die jeweiligen Subcluster als separiert betrachtet. Diese Schranke wurde zu 300 Zentimetern festgesetzt, was deutlich das Zweifache der maximalen Distanz der beiden Beine einer gehenden Person übertrifft. Die Wahl dieser Schranke ist im Übrigen genauso motiviert wie die Wahl der Schranke von 300 Zentimetern in Schritt zwei des CSA. Die zweite Bedingung ist genau dann erfüllt, falls die Anzahl der Paare separierter Subcluster gerade gleich $\frac{n(n-1)}{2}$ ist. Dabei ist n die Anzahl der Einzelobjekte, die mit dem Cluster assoziiert sind. Somit wird überprüft, ob alle Subcluster paarweise voneinander wohlsepariert sind.

Falls diese beiden Bedingungen erfüllt sind, werden die n Subcluster den n Einzelobjekten des Clusters zugeteilt, die dann wieder separat mit dem EWA verfolgt werden. Somit wird der Cluster wieder in seine Einzelpersonen aufgeteilt. Das Ende einer Interaktion wird also dadurch festgestellt, dass die Schätzungen permanent in Subcluster aufgeteilt werden. Vergleiche hierzu auch den Pseudocode in Tabelle 7.2.

Darüberhinaus wird der Cluster kontinuierlich in die maximale Anzahl an Subclustern zerlegt und diese werden in den graphischen Darstellungen wiedergegeben. Dazu wird die im ersten Schritt vorgestellte Methode angewandt. Es ist dabei zu beachten, dass das Verfahren nicht in der Lage ist, anzugeben, wieviele Einzelobjekte mit einem bestimmten Subcluster assoziiert sind.

Da die Zuteilung der Subcluster zu den Tracks rein zufällig erfolgt, kann dabei natürlich ähnlich wie im vorhergehenden Kapitel nicht garantiert werden, dass die Zuordnung der Einzelobjekte zu den Tracks dieselbe ist wie vor der Interaktion. Es ist aber anzunehmen, dass es hierfür keine Lösung gibt, solange nur anonyme Messungen wie Laserabstandsmessungen verwendet werden. In Kapitel 9 wird deshalb diskutiert, wie unter Verwendung von zusätzlicher Attributinformation eine Zuordnung der Identitäten der Personen doch möglich ist.

Der hier vorgestellte Algorithmus enthält wie auch der CSA neben numerischen Komponenten, wie sie ebenfalls für die Daten-Assoziierung oder die Positionsschätz-

ung einer einzelnen Person verwendet werden, auch Techniken aus dem Bereich der Knowledge Based Inference (Hastie et al., 2001) oder des Data Mining (Han und Kamber, 2001; Hand et al., 2001), da Vorwissen, zum Beispiel über den maximalen Abstand der Beine einer Person, benutzt wird, um die Daten zu strukturieren und zu verarbeiten.

7.3 Experimente

Das Verfahren wurde mit realen Daten getestet, die in unserem Labor aufgezeichnet worden waren. Es wurden dabei zwei Versuche durchgeführt, einmal mit drei und einmal mit vier Personen. Bevor diese beiden Experimente ausführlich diskutiert werden, zeigt Abbildung 7.1 die Anwendung der Methode auf die schon oft benutzten Daten der Interaktion zweier Personen aus Abbildung 6.10. In den Abbildungen dieses Abschnittes werden Einzelobjekte durch grüne Ellipsen wiedergegeben, während Cluster und Subcluster mit blauen Rechtecken angezeigt werden. Sobald die Bedingungen 1 und 2 erfüllt sind, werden die beiden Ziele wieder separat verfolgt. Die Lücke in der Bahn vor allem des Objektes, das sich nach der Interaktion nach rechts oben bewegt, rührt daher, dass die beiden Ziele einen gewissen Mindestabstand voneinander einnehmen müssen, damit zwei verschiedene Subcluster vorliegen. In Abbildung 7.1 wird auch wieder die aktuelle Geometrie der Einzelobjekte durch die Ellipsen repräsentiert, während in den restlichen Bildern dieses Kapitels standardisierte Ellipsen und Rechtecke verwendet werden. Genauer werden dabei im ersten Fall nicht Ellipsen sondern Kreise mit konstantem Radius benutzt, die lediglich aufgrund der Maßstabsverzerrung als Ellipsen erscheinen. Analog dazu handelt es sich bei den Rechtecken eigentlich um Quadrate. Die Graphiken dieses Abschnittes sind im Übrigen der Veröffentlichung (Kräußling, 2008c) im englischen Original entnommen.

Im ersten Experiment bewegten sich drei Personen gegen den Uhrzeigersinn um den beobachtenden Roboter, der dabei stationär war, herum. Diese Versuchsanordnung wurde aufgrund des beschränkten Platzangebotes in unserer Experimentalumgebung gewählt. Zunächst gingen die drei Personen separat und somit benutzte der Algorithmus den EWA (Abbildung 7.2). Dann gerieten zwei Personen in enge Nachbarschaft zueinander, so dass sich aus zwei Einzelobjekten ein Cluster bildete. Dieser wurde mit dem VBA verfolgt (Abbildung 7.3). In Abbildung 7.4 näherte sich auch der dritte Proband der Gruppe aus den anderen beiden und wurde deshalb ebenfalls mit dem Cluster assoziiert. Es kam somit zur Bildung eines kombinierten Clusters aus einem Einzelobjekt und einem Cluster. Dieser kombinierte Cluster wurde weiterhin mit dem VBA verfolgt. Danach teilte sich der kombinierte Cluster in zwei Subcluster auf, die in Abbildung 7.5 wiedergegeben sind. Als nächstes zerfiel der Cluster in drei Subcluster (Abbildung 7.6). Allerdings war zu diesem Zeitpunkt nur Bedingung 1, aber noch nicht Bedingung 2 erfüllt, so dass noch keine Einzelobjekte angezeigt werden. Schließlich trat aber auch Bedingung 2 ein, so dass die drei Personen wieder als Einzelobjekte wiedergegeben und mit dem EWA verfolgt wurden (Abbildung 7.7).

Im zweiten Versuch wurden vier Personen verfolgt, die sich ebenfalls gegen den

Uhrzeigersinn um den beobachtenden Roboter herum bewegten. Zunächst liefen alle vier Probanden getrennt. Deshalb wurden sie unabhängig voneinander mit dem EWA verfolgt (Abbildung 7.8). Dann gingen zwei Personen gemeinsam und wurden deshalb als einzelner Cluster mit dem VBA verfolgt, während die anderen beiden Personen weiter alleine umhergingen (Abbildung 7.9). Im nächsten Schritt kamen auch diese anderen beiden Probanden zusammen. Als Ergebnis wurden auch diese beiden Personen als Cluster repräsentiert und somit verfolgt der Algorithmus zu diesem Zeitpunkt zwei verschiedene Cluster, wobei jeder dieser Cluster für zwei Personen steht (Abbildung 7.10). In Abbildung 7.11 schließlich gingen alle vier Personen zusammen. Dies resultiert in der Vereinigung der beiden Cluster zu einem kombinierten Cluster, der alle vier Personen wiedergibt. Als nächstes teilt sich dieser kombinierte Cluster in zwei Subcluster auf (Abbildung 7.12). Jeder dieser Subcluster repräsentiert eine unbekannte Anzahl von Einzelpersonen. In Abbildung 7.13 und 7.14 teilen sich diese Subcluster sukzessive weiter in drei beziehungsweise vier Subcluster auf. Schließlich ist auch wieder Bedingung 2 erfüllt, so dass die vier Personen erneut unabhängig voneinander als Einzelpersonen verfolgt werden (Abbildung 7.15).

In der Versuchsreihe wurde somit gezeigt, dass der ECSA in der Lage ist, mit allen Situationen und allen Objektklassen, die bei der Verfolgung beliebig vieler Ziele auftreten können, zurechtzukommen:

- die Vereinigung zweier Einzelobjekte zu einem Cluster (erstes und zweites Experiment).
- die Verschmelzung eines Einzelobjektes und eines Clusters zu einem kombinierten Cluster (erstes Experiment).
- die Vereinigung zweier Cluster zu einem kombinierten Cluster (zweites Experiment).
- die Aufteilung eines Clusters in Subcluster, die in den Graphiken dargestellt werden können (erstes und zweites Experiment).
- die Auflösung eines Clusters in seine Einzelobjekte (erstes und zweites Experiment).

In jüngster Zeit wurde die Idee, eng benachbarte Personen in Cluster zusammenzufassen bzw. Interaktionen zwischen Personen explizit zu modellieren, auch von anderen Autoren aufgegriffen. In (Mucientes und Burgard, 2006) werden aus Tracks eng benachbarter Personen Cluster gebildet. Außerdem wird die Vereinigung und die Aufteilung von Clustern betrachtet. Allerdings werden auch weiterhin die Tracks einzelner Personen, die sich in einem Cluster befinden, separat berechnet. Das Clustern dient hier lediglich dazu, die Wahrscheinlichkeitsdichtefunktionen dieser Tracks genauer und zuverlässiger berechnen zu können. In (Song et al., 2008) wird bei hoher Objektdichte auch der Fall zugelassen, dass ein Feature mehreren interagierenden Personen zugeordnet werden kann.

	<i>Splitting of clusters</i>
	<i>n clusters</i>
	<i>Cluster i ($1 \leq i \leq n$) represents p_i single objects</i>
1.	For each cluster
2.	<i>Condition 1</i>
2.1	Subdivide the cluster into subclusters
2.2	<i>Definition of a subcluster:</i> <i>group of estimates with a pairwise distance below 150 cm</i>
2.3	<i>Result: q_i subclusters</i>
2.4	If $p_i = q_i$
2.5	Condition 1 is true
2.6	End
3.	if condition 1 is true
4.	<i>Condition 2</i>
4.1	For each pair of subclusters s_i and s_j
4.2	<i>Definition of $d(s_i, s_j)$</i>
4.2.1	<i>Subcluster s_i: k_i estimates x_1, \dots, x_{k_i}</i>
4.2.2	<i>Subcluster s_j: k_j estimates y_1, \dots, y_{k_j}</i>
4.2.3	<i>$d(x_l, x_m)$: Euclidian distance</i> <i>between the points x_l and x_m</i>
4.2.4	<i>$d(s_i, s_j) = \max_{1 \leq l \leq k_i, 1 \leq m \leq k_j} d(x_l, y_m)$</i>
4.3	Calculate $d(s_i, s_j)$
4.4	If $d(s_i, s_j) > 300$ cm
4.5	Subclusters s_i and s_j are separated
4.6	End
4.7	End
4.8	If the number of pairs of separated subclusters is equal to $\frac{p_i(p_i-1)}{2}$
4.9	Condition 2 is true
4.10	End
5.	<i>Splitting of clusters into single objects</i>
5.1	If condition 1 and 2 are true for cluster i
5.2	Assign the q_i subclusters to the p_i single objects
5.3	End
6.	End
7.	End

Tabelle 7.2: Aufteilung eines Clusters in Einzelobjekte, Pseudocode.

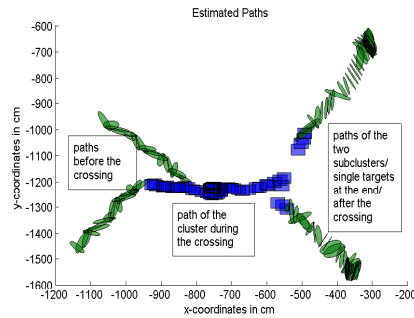


Abbildung 7.1: Zwei kreuzende Objekte, reale Daten, verarbeitet mit dem ECSA.

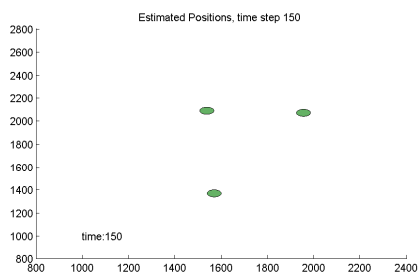


Abbildung 7.2: Der ECSA verfolgt drei verschiedene Einzelobjekte.

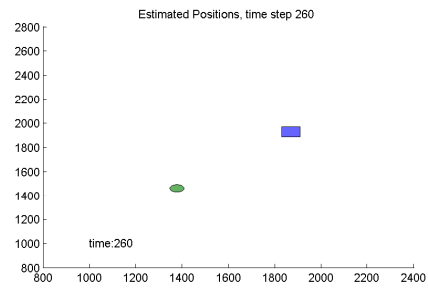


Abbildung 7.3: Zwei der Einzelobjekte kommen sich sehr nahe und werden deshalb durch einen Cluster repräsentiert.

Drei interagierende Ziele, verarbeitet mit dem ECSA.

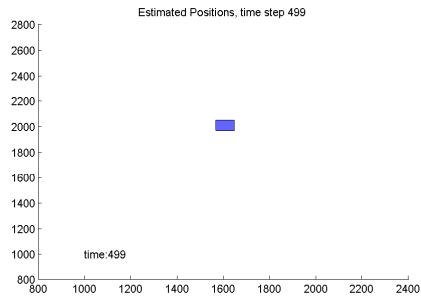


Abbildung 7.4: Das dritte Einzelobjekt fusioniert mit den anderen beiden und wird deshalb ebenfalls mit dem Cluster assoziiert.

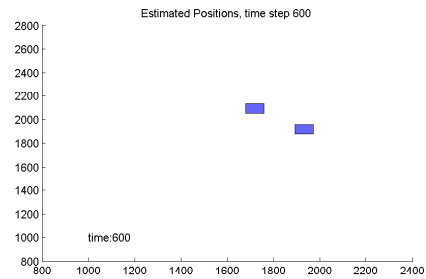


Abbildung 7.5: Eines der Ziele trennt sich von den anderen beiden. Somit werden zwei Subcluster dargestellt.

Drei interagierende Ziele, verarbeitet mit dem ECSA.

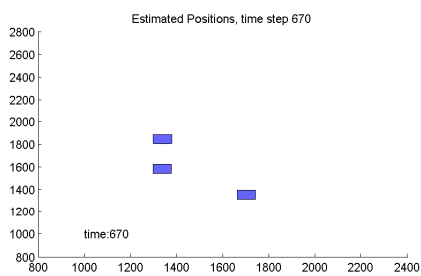


Abbildung 7.6: Die anderen beiden Ziele trennen sich ebenfalls voneinander und somit werden drei Subcluster angezeigt.

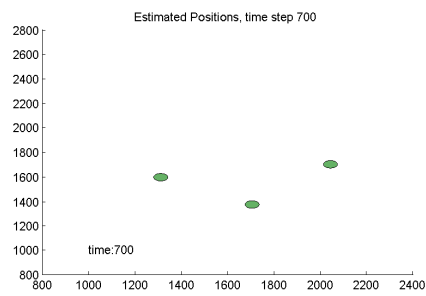


Abbildung 7.7: Sobald die drei Subcluster ausreichend weit voneinander entfernt sind, werden wieder drei Einzelobjekte verfolgt.

Drei interagierende Ziele, verarbeitet mit dem ECSA.

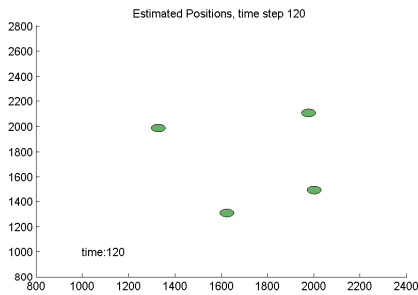


Abbildung 7.8: Der Algorithmus verfolgt vier verschiedene Einzelobjekte.

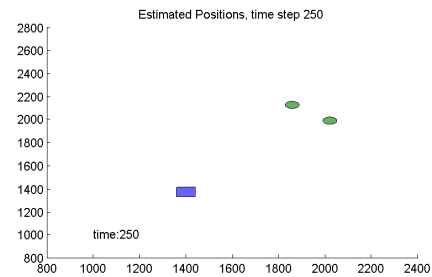


Abbildung 7.9: Zwei der Einzelobjekte kommen sich sehr nahe und werden als Cluster verfolgt.

Vier interagierende Ziele, verarbeitet mit dem ECSA.

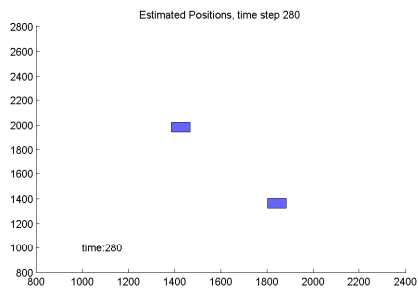


Abbildung 7.10: Auch die anderen beiden Einzelobjekte kommen in enge Nachbarschaft zueinander und werden deshalb ebenfalls durch einen Cluster repräsentiert.

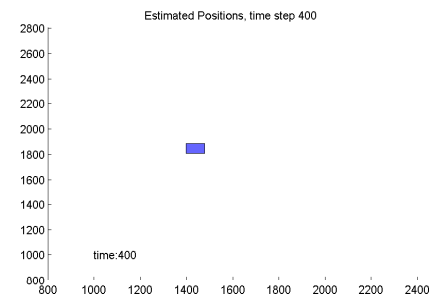


Abbildung 7.11: Die beiden Cluster vereinigen sich und deshalb werden alle vier Ziele mit einem einzigen Cluster assoziiert.

Vier interagierende Ziele, verarbeitet mit dem ECSA.

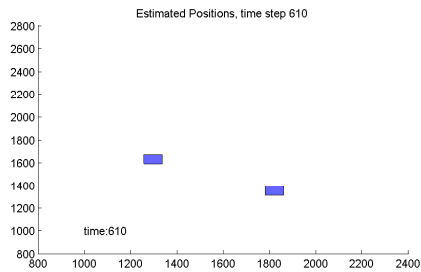


Abbildung 7.12: Der kombinierte Cluster trennt sich nach einer Weile in zwei Subcluster, die jeweils aus einer unbekanntenen Anzahl von Einzelobjekten bestehen.

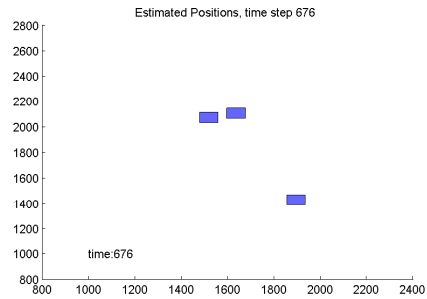


Abbildung 7.13: Einer der beiden Subcluster trennt sich in zwei weitere Subcluster auf.

Vier interagierende Ziele, verarbeitet mit dem ECSA.

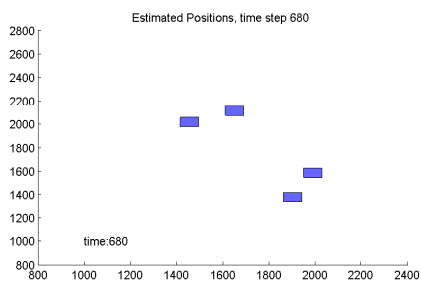


Abbildung 7.14: Die Ziele trennen sich voneinander, so dass vier Subcluster angezeigt werden.

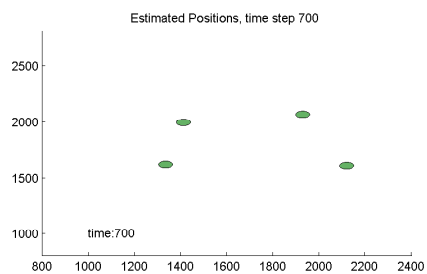


Abbildung 7.15: Schließlich haben sich die Ziele soweit voneinander entfernt, dass wieder vier Einzelobjekte wiedergegeben werden.

Vier interagierende Ziele, verarbeitet mit dem ECSA.

Kapitel 8

Ein Vergleich des ECSA mit dem SJPDAF

8.1 Eine Einführung und kurze Beschreibung des SJPDAF

Wie in Kapitel 2 bereits angedeutet, existiert im Bereich der mobilen Robotik schon seit einigen Jahren ein etabliertes und bedeutendes Verfahren zur Verfolgung mehrerer Personen mit Laserscannern, nämlich der SJPDAF (Sample Based Joint Probabilistic Data Association Filter) von Schulz (Schulz et al., 2001a; Schulz et al., 2001b; Schulz, 2002; Schulz et al., 2003a). Dieser Ansatz kombiniert die Ideen der Teilchenfilter (Gordon et al., 1993; Pitt und Shephard, 1999) mit denen des JPDAF (Joint Probabilistic Data Association Filter) (Fortmann et al., 1983). Ein wesentliches Charakteristikum dieses Verfahrens für das sogenannte Multi-People-Tracking besteht dabei darin, dass sowohl die Anzahl der Personen als auch die Assoziation (Zuordnung) der Features zu den computerinternen Tracks approximiert wird, wobei zusätzlich noch Verdeckungen berücksichtigt werden. Zu diesem Zweck werden drei Schritte hintereinander ausgeführt, nämlich zunächst die Aufbereitung der Messwerte, zum Beispiel zur Gewinnung der Features aus den Laserabstandsmessungen, dann die schon erwähnte Korrelation oder Assoziation dieser Features zu den Tracks, und schließlich die Aktualisierung der Zustandsschätzungen.

In diesem Kapitel werden der ECSA und der SJPDAF bezüglich folgender in Kapitel 5 und 6 eingeführter Kriterien miteinander verglichen:

- die Genauigkeit der Positionsschätzung bzw. der Informationsgehalt über die tatsächliche Position des Objektes.
- der Bedarf an Rechenzeit.
- die Fähigkeit zur separaten Verfolgung zweier sich begegnender Ziele am Ende einer Interaktion und die Stabilität gegenüber dem Verlust von Tracks bei diesem

Problem.

In den ersten beiden Fällen wurden Experimente mit simulierten Daten (die Verfolgung eines kreisförmigen Objektes) und im dritten Fall mit realen Daten zweier Personen durchgeführt. Die vorgestellten Ergebnisse wurden in (Kräußling und Schulz, 2006) publiziert.

Im nächsten Abschnitt werden zunächst die Kriterien der Genauigkeit und des Rechenzeitbedarfs ähnlich wie in Kapitel 5 für den Fall eines einzelnen kreisförmigen Zieles, das sich auf einer Kreisbahn bewegt, betrachtet. Die Genauigkeit wurde jeweils über die Abweichung der Schätzung von der tatsächlichen Position bestimmt. Da dabei nur ein Objekt vorhanden war, reduziert sich der SJPDFAF in diesem Falle auf einen einfachen Teilchenfilter zur Berechnung von Positionsschätzungen, bestehend aus einer Vorschrift zur Initialisierung und zur rekursiven Durchführung des Filterschrittes. Dieser Teilchenfilter wurde in MATLAB implementiert und soll im Folgenden kurz beschrieben werden.

Der kinematische Zustand des Zieles wird wieder mit Hilfe eines vierdimensionalen Vektors beschrieben. Es ist also

$$x = \left(x_1 \quad x_2 \quad \dot{x}_1 \quad \dot{x}_2 \right)^\top. \quad (8.1)$$

x_1 und x_2 sind dabei einfach die kartesischen Koordinaten in der Ebene und \dot{x}_1 und \dot{x}_2 die zugehörigen Geschwindigkeiten. Die Teilchen, die den Zustand des Objektes beschreiben, sind durch einen N -dimensionalen Vektor s mit

$$s = \left(s^{(1)}, \dots, s^{(N)} \right) \quad (8.2)$$

gegeben. Dabei ist N die Anzahl der Teilchen. Hier wurde $N = 400$ benutzt. Weiterführende Untersuchungen haben gezeigt, dass eine allzu große Reduzierung der Anzahl der Teilchen zu einem Verlust des Zieles führen würde. Dies rührt daher, dass es in diesem Falle vorkommen kann, dass keines der wenigen Teilchen nach der Vorhersage des Zustandes nahe genug am gemessenen Feature liegt und somit die Messung zu wenig Einfluss auf die Berechnung des geschätzten Zustandes hat.

Jedes Teilchen $s^{(j)}$ besteht aus zwei Komponenten, nämlich dem Zustand $x^{(j)}$ und dem Gewicht $w^{(j)}$, d.h. es ist

$$s^{(j)} = \left(x^{(j)}, w^{(j)} \right). \quad (8.3)$$

Die Zustände $x^{(j)}$ sind von der Struktur, wie sie in Gleichung 8.1 eingeführt wurde. Orte mit hoher Teilchendichte oder mit Teilchen mit hohen Gewichten entsprechen dabei einer hohen Aufenthaltswahrscheinlichkeit für das Objekt und umgekehrt. Letzten Endes ist eine solche Menge von Teilchen also eine diskrete Approximation einer Wahrscheinlichkeitsverteilung.

Der Hauptvorteil einer solchen Vorgehensweise besteht darin, dass beliebige und insbesondere multimodale Wahrscheinlichkeitsverteilungen modelliert werden können,

d.h. Verteilungen mit mehreren lokalen Maxima. Auf diese Weise können zum Beispiel Hindernisse besser behandelt werden.

Die Initialisierung der Teilchen wurde wie folgt durchgeführt. Zunächst wurde mit den Methoden der Trackinitialisierung eine Schätzung x_0 für den Zustand des Systems am Anfang der Messreihe ermittelt. Damit wurden dann die Zustände der Teilchen generiert. Dazu wurde für jedes Teilchen eine Zufallszahl von einer Gaußverteilung mit Mittelwert Null und Standardabweichung Eins, also einer sogenannten Standardnormalverteilung, gezogen. Diese Zufallszahl wurde anschließend mit der Standardabweichung σ des Messrauschens multipliziert. Als Ergebnis dieses Vorgehens erhält man eine Zufallszahl gezogen von einer Gaußverteilung mit Mittelwert Null und Standardabweichung σ (Renyi, 1970). Diese Zufallszahlen wurden schließlich jeweils zu den Ortskoordinaten des Zustandes x_0 aufaddiert. Dabei wurde dieses Vorgehen getrennt für die x_1 - und x_2 -Koordinate durchgeführt. Die initialen Gewichte der Teilchen schließlich wurden jeweils auf Eins gesetzt.

Nun wird der Filterschritt beschrieben. Dessen Ausführung geschieht bei diesem Teilchen-Filter bei jeder Messung durch die folgenden drei Schritte:

1. Entsprechend des Bewegungs- oder Vorhersagemodells werden Vorhersagen für die neue Position gemacht, d.h. für jedes Teilchen wird ein neuer Zustand berechnet. Da hierbei auch nichtlineare Dynamikmodelle berücksichtigt werden können, kann in schwierigen Fällen oft ein besseres Assignment der Features zu den Tracks ermöglicht werden.
2. Die Gewichte werden aufgrund der Messung unter Berücksichtigung des Sensormodells neu berechnet. Man spricht dabei auch vom Sampling. Auf diese Weise wird eine verbesserte Schätzung der tatsächlichen Wahrscheinlichkeitsverteilung erhalten. Diese wird zu diesem Zeitpunkt sowohl durch die Dichte der Teilchen als auch durch deren Gewichte repräsentiert.
3. Im sogenannten Resampling werden neue Teilchen mit gleichen Gewichten erzeugt. Dabei werden die im ersten Schritt berechneten Zustände der Teilchen übernommen und die Wahrscheinlichkeit dafür, dass ein Teilchen gezogen wird, hängt von den im zweiten Schritt berechneten Gewichten ab. Dadurch werden tendenziell Teilchen mit hohen Gewichten öfter gezogen, während Teilchen mit sehr geringem Gewicht gelöscht werden. Die auf diese Weise gezogenen Teilchen erhalten schließlich alle das Gewicht Eins. Die Wahrscheinlichkeitsverteilung wird folglich nach dem Resampling wieder nur noch durch die Dichte der Teilchen wiedergegeben, da erneut alle Gewichte gleich sind.

Im Folgenden wird erläutert, wie diese Schritte im Rahmen des hier vorgestellten Ansatzes von Schulz realisiert sind.

Zunächst wird also eine Vorhersage $x_{\text{prä}}^{(j)}$ ¹ für den Zustand $x^{(j)}$ eines jeden Teilchens j berechnet. Dabei wird für die Bewegung des Objektes das Modell von van

¹prä für Prädiktion.

Keuk benutzt, wie es in Abschnitt 4.2 besprochen wurde.² Es ist also

$$x_{\text{prä}}^{(j)} = \begin{pmatrix} 1 & 0 & \Delta T & 0 \\ 0 & 1 & 0 & \Delta T \\ 0 & 0 & e^{-\Delta T/\Theta} & 0 \\ 0 & 0 & 0 & e^{-\Delta T/\Theta} \end{pmatrix} x^{(j)} + \Sigma \sqrt{1 - e^{-2\Delta T/\Theta}} \begin{pmatrix} 0 \\ 0 \\ r_1 \\ r_2 \end{pmatrix}. \quad (8.4)$$

Dabei sind r_1 und r_2 Zufallszahlen, die unabhängig voneinander von einer Gaußverteilung mit Mittelwert Null und Standardabweichung Eins, d.h. wieder einer Standardnormalverteilung, gezogen worden sind.

Als nächstes wird ein Erwartungsgebiet (vergleiche Abschnitt 4.3) berechnet, um die m Messungen $\{\tilde{z}_i\}_{i=1}^m$, die vom Ziel stammen, aus allen Messwerten herauszufiltern. Dann wird wie beim EWA als Feature z der Mittelwert dieser Messungen zu

$$z = \frac{\sum_{i=1}^m \tilde{z}_i}{m} \quad (8.5)$$

berechnet.

Nun werden in einem zweiten Schritt die Gewichte der Teilchen durch Integration der Messung aktualisiert. Dazu wird ein Modell des Sensors in Form einer Gaußverteilung zugrunde gelegt und nach Gleichung 3.3

$$w^{(j)} = \frac{1}{2\pi\sqrt{|T|}} \exp \left\{ -\frac{1}{2} \left(z - \begin{pmatrix} x_{1,\text{prä}}^{(j)} \\ x_{2,\text{prä}}^{(j)} \end{pmatrix} \right)^\top T^{-1} \left(z - \begin{pmatrix} x_{1,\text{prä}}^{(j)} \\ x_{2,\text{prä}}^{(j)} \end{pmatrix} \right) \right\} \quad (8.6)$$

gesetzt. Dabei ist

$$T = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix} \quad (8.7)$$

mit der jeweiligen Standardabweichung σ des Messrauschens die Kovarianzmatrix des Messprozesses. $x_{1,\text{prä}}^{(j)}$ und $x_{2,\text{prä}}^{(j)}$ entstammen der Vorhersage $x_{\text{prä}}^{(j)}$ für den Zustand des Teilchens j mit

$$x_{\text{prä}}^{(j)} = \begin{pmatrix} x_{1,\text{prä}}^{(j)} & x_{2,\text{prä}}^{(j)} & \dot{x}_{1,\text{prä}}^{(j)} & \dot{x}_{2,\text{prä}}^{(j)} \end{pmatrix}^\top. \quad (8.8)$$

Schließlich müssen zum Abschluss des Filterschrittes wieder N neue Teilchen mit Gewicht Eins generiert werden, die den Zustand des Objektes repräsentieren. Diesen Vorgang bezeichnet man als Resampling. Bei der hier verwendeten Variante spricht man auch vom Bootstrap-Resampling. Dazu werden zunächst die Gewichte gemäß der Gleichung

$$w_n^{(j)} = \frac{w^{(j)}}{\sum_{i=1}^N w^{(i)}} \quad (8.9)$$

²Ursprünglich benutzte Schulz ein anderes Bewegungsmodell. Um einen fairen Verfahrenvergleich zu ermöglichen, der nicht von dem verwendeten Bewegungsmodell abhängt, wird hier aber dasselbe Bewegungsmodell wie in der übrigen Arbeit zugrundegelegt.

auf Eins normiert. Dann setzt man $i = 1$, $j = 0$ und $sw = w_n^{(1)}$ und zieht eine Zufallszahl p von einer Gleichverteilung in $[0, 1]$. Nun müssen, solange $j < N$ gilt, zwei Fälle unterschieden werden:

1. Es ist $sw > p$. In diesem Falle wird $j = j + 1$, $x_n^{(j)} = x_{\text{prä}}^{(i)}$ und $p = p + 1/N$ gesetzt.
2. Es ist $sw \leq p$. In diesem Fall wird $i = i + 1$ und $sw = sw + w_n^{(i)}$ gesetzt. Dabei wird außerdem, falls i soeben den Wert N überschritten hat, $i = i - N$ gesetzt, und erst danach wird $w_n^{(i)}$ zu sw addiert.

Nun können die N neuen Teilchen generiert werden. Dazu setzt man einfach $x^{(j)} = x_n^{(j)}$ und $w^{(j)} = 1$. Damit ist die Beschreibung des Rekursionsschrittes abgeschlossen.

Als letztes wird noch eine Schätzung x_k für die neue Position des Zieles gemäß der Formel

$$x_k = \frac{\sum_{i=1}^N x^{(i)}}{N} \quad (8.10)$$

berechnet.

Damit ist die Beschreibung des Teilchenfilters zur Verfolgung eines unabhängigen Einzelzieles abgeschlossen. Die Verfolgung mehrerer Ziele erfordert dagegen im Prinzip die Berechnung der Verbundwahrscheinlichkeit oder des kombinierten Zustandes aller Objekte. Dies ist aber in der Praxis bereits für eine kleine Anzahl von Zielen undurchführbar, da dabei die Komplexität exponentiell mit der Zahl der Objekte anwächst. Deshalb besteht ein häufig gebrauchter Ansatz darin, die einzelnen Ziele unabhängig voneinander zu verfolgen. Dann müssen aber die Messungen oder Features mit den zugehörigen computerinternen Objekten oder Tracks assoziiert werden. Zu diesem Zweck kann das JPDAF-Verfahren (Cox, 1993; Fortmann et al., 1983) verwendet werden, das letztendlich eine Bayessche Schätzung darstellt. Der SJPDFAF benutzt dabei zur Repräsentierung der Zustände unabhängige Teilchenfilter, während klassischerweise der Kalman-Filter angewandt wird. Der SJPDFAF wird nun kurz erläutert.

Sei $X(k) = \{x_1^k, \dots, x_M^k\}$ der Zustand der M verfolgten Ziele zum Zeitpunkt k . Dabei ist jedes x_i^k eine Zufallsvariable über den Zustandsraum eines einzelnen Zieles. Außerdem sei $Z(k) = \{z_1^k, \dots, z_{m_k}^k\}$ eine Messung zum Zeitpunkt k , wobei z_i^k ein Feature einer solchen Messung ist. Um die beobachteten Features einem einzelnen Objekt zuzuordnen, werden nun punktförmige Ziele angenommen. Dies wird dadurch erreicht, dass man zur Generierung dieser Features wie in Gleichung 8.5 ein ungewichtetes Mittel der Messungen verwendet, die einem Feature zugewiesen wurden.

Im Rahmen des JPDAF versteht man nun unter einem Verbundereignis θ einer Assoziation eine Menge von Paaren

$$(j, i) \in \{0, \dots, m_k\} \times \{1, \dots, M\}. \quad (8.11)$$

Dabei entspricht die 'Messung' $j = 0$ dem Fall, dass ein Objekt, zum Beispiel aufgrund einer Verdeckung, vom Sensor nicht registriert wird. Jedes solche Verbundereignis legt fest, welches Feature welchem Objekt zugeordnet wird. Außerdem sei Θ_{ji}

die Menge aller Verbundereignisse θ , die das Feature j dem Objekt i zuteilen. Damit ergibt sich die a-posteriori-Wahrscheinlichkeit β_{ji} , dass das Feature j durch das Objekt i hervorgerufen wurde, zu

$$\beta_{ji} = \sum_{\theta \in \Theta_{ji}} P(\theta | Z(k)). \quad (8.12)$$

Der SJPDFAF benutzt nun eine Menge von N Teilchen, um die Wahrscheinlichkeitsdichte $p(x_i^k | Z(k))$ zu repräsentieren. Die Aktualisierung der Gewichte $w^{(i)}$ basiert dabei auf den Wahrscheinlichkeiten $p(z_j^k | x_i^k)$ der Beobachtung der Features z_j^k und den zugehörigen Wahrscheinlichkeiten der Assoziationen β_{ji} . Die Gewichte der Teilchen werden also gemäß der Beziehung

$$w^{(i)} = \alpha \sum_{j=0}^{m_k} \beta_{ji} p(z_j^k | x_i^k) \quad (8.13)$$

aktualisiert. Dabei ist α ein Normierungsfaktor, der garantiert, dass die Summe über alle Gewichte Eins ergibt. Es wird also letzten Endes eine gewichtete Summe über alle möglichen Zuordnungen berechnet. Schließlich erhält man durch das oben geschilderte sogenannte Bootstrap-Resampling N neue Teilchen mit Gewicht Eins.

8.2 Der Verfahrensvergleich

Es werden nun die Ergebnisse des Vergleichs des ECSA mit dem SJPDFAF dargestellt. Dabei wird bezüglich der ersten beiden Kriterien (Genauigkeit der Positionsschätzung und Rechenzeitbedarf) der zunächst geschilderte einfache Teilchenfilter zugrundegelegt, auf den sich der SJPDFAF im Falle eines Einzelobjektes reduziert, während bezüglich des Kriteriums der Robustheit bei zwei interagierenden Zielen der SJPDFAF mit der Berechnung der Gewichte gemäß Gleichung 8.13 zugrundegelegt wurde. Im letzteren Fall wurde dabei auf die Original-Software von Dirk Schulz zurückgegriffen.

8.2.1 Genauigkeit der Positionsschätzung und Rechenzeitbedarf

Hier wird wieder das Problem eines kreisförmigen Objektes, das sich auf einer Kreisbahn um den beobachtenden Sensor herum bewegt, zugrunde gelegt. Da der SJPDFAF gemäß Gleichung 8.5 ein ungewichtetes Mittel aller Messungen eines Features verwendet, liegt die Vermutung nahe, dass auch er ähnlich wie der ECSA oder der EWA, auf den sich der ECSA im Falle eines Einzelzieles ja reduziert, den Schwerpunkt S_O mit hoher Genauigkeit schätzt. Deshalb gibt Tabelle 8.1 den Euklidischen Abstand der Schätzung x_k aus Gleichung 8.10 von diesem Erwartungswert an. Wegen Gleichung 8.6 kann hierbei der Fall $\sigma = 0$ nicht betrachtet werden, da in diesem Falle die Inverse der Nullmatrix berechnet werden müsste. Für Details dieses Vorgehens vergleiche

Abschnitt 5.3. Tabelle 8.2 gibt die zugehörigen Standardabweichungen an. Die Ergebnisse zeigen, dass der SJPDFAF ebenfalls mit guter Genauigkeit den Schwerpunkt S_O schätzt, allerdings nicht mit der gleichen Präzision wie der EWA oder der ECSA (Tabelle 5.2). Die auftretenden Standardabweichungen sind auch hier wieder relativ klein.

Tabelle 8.1: Durchschnittliche Abweichung der Schätzung vom Schwerpunkt S_O für den SJPDFAF in Zentimetern.

d/cm	100	200	400	600	800
$\sigma = 1\text{ cm}$	0.5284	0.7302	1.4468	2.2417	3.0989
$\sigma = 3\text{ cm}$	0.8261	0.7681	1.4852	1.7295	2.0373
$\sigma = 5\text{ cm}$	1.2667	1.0773	1.6216	1.8382	2.2059
$\sigma = 7.5\text{ cm}$	1.8217	1.4752	1.8846	2.1225	2.4957
$\sigma = 10\text{ cm}$	2.4399	1.9176	2.2290	2.4866	2.8916

Tabelle 8.2: Durchschnittliche Abweichung der Schätzung vom Schwerpunkt S_O für den SJPDFAF, Standardabweichungen in Zentimetern.

d/cm	100	200	400	600	800
$\sigma = 1\text{ cm}$	0.0627	0.0956	0.0348	0.1151	0.1639
$\sigma = 3\text{ cm}$	0.0547	0.0494	0.0378	0.0310	0.0344
$\sigma = 5\text{ cm}$	0.0857	0.0609	0.0477	0.0474	0.0470
$\sigma = 7.5\text{ cm}$	0.1054	0.1064	0.0839	0.0586	0.0768
$\sigma = 10\text{ cm}$	0.1878	0.1386	0.0711	0.0821	0.1072

Nun wird das zweite Kriterium betrachtet, nämlich die Komplexität oder der Rechenbedarf. Als Maß hierfür wird wie in Kapitel 5 der Bedarf an Rechenzeit zur Ausführung der Berechnungen für einen Zeitschritt herangezogen. Der SJPDFAF wurde dabei ebenfalls in MATLAB auf einem Pentium IV mit 2.8 GHz implementiert. Dabei benötigte der ECSA ungefähr zwanzig Millisekunden pro Zeittakt für alle Kombinationen von Entfernung d und Standardabweichung σ , während der SJPDFAF jeweils ungefähr 89 Millisekunden benötigte. Vergleiche hierzu auch Abschnitt 5.4. Somit ist der ECSA dem SJPDFAF bezüglich dieses Kriteriums überlegen. Natürlich hängt der Rechenzeitbedarf des SJPDFAF sehr stark von der Anzahl der Teilchen ab, die verwendet werden. Dabei wurde in den Experimenten jeweils eine Zahl von 400 Teilchen zugrunde gelegt. Wie schon erwähnt, würde eine zu starke Reduktion der Anzahl der Teilchen in einem Verlust des Ziels resultieren. Auch von anderen Autoren (Bellotto und Hu, 2007c) werden einige hundert Teilchen als Untergrenze dafür angegeben, dass keine Ziele verloren gehen.

8.2.2 Zwei interagierende Ziele

In einer zweiten Serie von Experimenten wurden der ECSA und der SJPDAF bezüglich der Frage verglichen, wie robust die Algorithmen gegenüber des Verlusts von Objekten im Falle von zwei sich begegnenden Zielen sind. Dazu wurden die fünf Datensätze aus Kapitel 6.3 verwendet. Die Begegnungen dauerten also jeweils ca. dreißig Sekunden. Wie Schulz in seinen Arbeiten ausführt, benötigt es längere Zeit, bis genügend Teilchen oder Samples, die ja eine gewisse Tendenz zur räumlichen Ausbreitung aufweisen, von einer Person auf die andere übergesprungen sind. Deshalb wären bei kurzen Interaktionen wenig Probleme zu erwarten.

Der ECSA als Verallgemeinerung des CSA verfolgt, wie aus Kapitel 6.3 bekannt ist, beide Personen in allen fünf Szenarien nach der Interaktion getrennt, d.h. kein Objekt geht verloren. Für den SJPDAF sind die Verhältnisse wesentlich komplizierter. Da dieses Verfahren Zufallszahlen benutzt, ist der Ausgang eines einzelnen Experimentes nicht deterministisch. Deswegen wurden hier für jedes Szenario außer dem fünften jeweils zwanzig Versuchsdurchläufe ausgeführt. Für das fünfte Szenario dagegen wurden nur zehn Versuche durchgeführt. Tabelle 8.3 gibt jeweils den Prozentsatz der erfolgreichen Durchläufe, d.h. der Versuche, bei denen nach der Begegnung beide Personen getrennt verfolgt wurden, an. Es zeigt sich, dass das Verhalten des SJPDAF in den Szenarien eins und zwei noch einigermaßen passabel ist, während das Verfahren bei den restlichen drei Szenarien sehr schlecht abschneidet und im fünften Szenario sogar immer scheitert. Szenario vier konnte dabei überhaupt nicht ausgewertet werden, da es hier zu einer Interaktion mit einem statischen Objekt kam. Eine der beiden Personen lief nämlich sehr nahe an diesem Objekt vorbei, so dass das Verfahren diese Person mit dem Objekt verwechselte und als Position des Probanden danach stets die Position dieses statischen Objektes angab. Da der ECSA immer erfolgreich ist, zeigt diese Methode somit in allen fünf Szenarien ein besseres Verhalten als der SJPDAF. Der Grund für das insgesamt sehr schlechte Abschneiden in den Szenarien drei und fünf liegt vermutlich darin, dass eine der beiden Personen hier am Ende der Interaktion, d.h. als sie sich von der anderen Person trennte, ein sehr abruptes Manöver im Sinne einer scharfen Richtungsänderung ausführte. Dies führte dazu, dass zu wenige Teilchen eine geeignete Positionsschätzung repräsentierten, da das verwendete Modell geradlinige Bewegungen bevorzugt.

Tabelle 8.3: Verhalten des SJPDAF bei zwei interagierenden Personen, Prozentsatz der erfolgreichen Versuchsdurchläufe.

Szenario	1	2	3	4	5
Prozentsatz	80	90	15	—	0

Dieses Verhalten soll an dieser Stelle noch etwas näher betrachtet werden. Da sich die beiden Personen während der Interaktion längere Zeit gemeinsam bewegen und somit für eine längere Zeitspanne in beiden Fällen dieselben Messungen verarbeitet wer-

den, sind die Tracks der beiden Probanden am Ende der Begegnung nahezu identisch (Konvergenz der Tracks, (Blackman, 1991)). Lediglich durch die Verwendung von Zufallszahlen ergeben sich geringgradige Unterschiede. Der Einfachheit halber wird nun angenommen, dass sich die beiden Personen am Ende der Begegnung gemeinsam auf einer geradlinigen Bahn bewegen wie in Abbildung 8.1 angedeutet. Die Wahrscheinlichkeit für den SJPDF, die Person eins, die sich nach oben bewegt, weiter zu verfolgen, sei dann p , und die analoge Wahrscheinlichkeit dafür, die Person zwei, die sich nach unten bewegt, weiter zu verfolgen, sei dann q . Aufgrund des verwendeten Dynamikmodells werden nun aber Bewegungen bevorzugt, die eine geradlinige Fortsetzung der bisherigen Bewegung darstellen, d.h. deren Bahn unmittelbar nach der Begegnung einen möglichst kleinen Winkel θ (zur Definition des Winkels θ vergleiche Abbildung 8.2) mit den Bahnen am Ende der Begegnung einschließt. Somit erhält man für $\theta = 0^\circ$

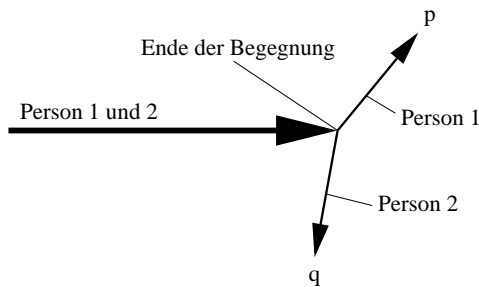


Abbildung 8.1: Zwei kreuzende Personen 1 und 2 am Ende einer Begegnung.

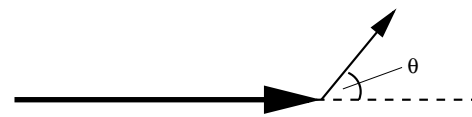
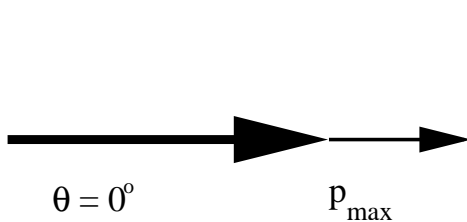
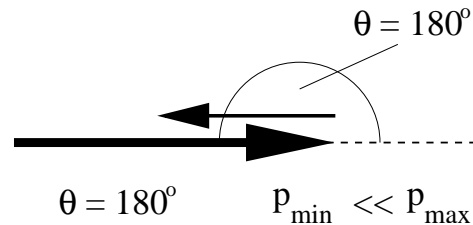


Abbildung 8.2: Definition des Streuwinkels θ .

Zwei interagierende Ziele.

(vergleiche Abbildung 8.3) die maximale Wahrscheinlichkeit p_{max} , während man für $\theta = 180^\circ$ (vergleiche Abbildung 8.4) eine sehr kleine minimale Wahrscheinlichkeit p_{min} dafür erhält, dass der SJPDF einen Track mit der Person, die die betreffende Bewegung ausführt, assoziiert. Dieses Verhalten hat eine gewisse Ähnlichkeit mit Streuexperimenten in der Physik, wie zum Beispiel der Rutherfordstreuung. Dort wurden Alphateilchen auf eine Goldfolie geschossen und die Anzahl der gestreuten Teilchen in Abhängigkeit vom Streuwinkel gemessen. Auch dort war diese Anzahl für größere Streuwinkel geringer, was einer geringeren Wahrscheinlichkeit für die Alpha-Teilchen entspricht, in Raumbereiche mit größerem Streuwinkel gestreut zu werden. Mit dem Rutherfordschen Streuexperiment wurden um 1911 elementare Kenntnisse über den Aufbau von Atomen, wie zum Beispiel den Kerndurchmesser, gewonnen.

Anhand der beiden möglichen Fälle eines symmetrischen und eines asymmetrischen Verhaltens der beiden Personen am Ende der Begegnung sollen die Verhältnisse nun noch etwas genauer erläutert werden. Zunächst gilt allgemein mit den Bezeich-

Abbildung 8.3: Maximale Wahrscheinlichkeit für $\theta = 0^\circ$.Abbildung 8.4: Minimale Wahrscheinlichkeit für $\theta = 180^\circ$.

Zwei interagierende Ziele.

nungen aus Abbildung 8.1

$$p + q = 1. \quad (8.14)$$

Diese Beziehung rührt daher, dass sich der SJPDF für eine der beiden realen Bahnen entscheiden muss. Die Wahrscheinlichkeit dafür, dass sich der SJPDF beide Male für die Bahn der Person eins entscheidet, ist dabei p^2 . Die Wahrscheinlichkeit dafür, dass er sich einmal für die Bahn der Person eins und das andere Mal für die Bahn der Person zwei entscheidet und dabei auch die Zuordnung richtig wählt, ist $p \cdot q$, während eine falsche Zuordnung ebenfalls mit der Wahrscheinlichkeit $p \cdot q$ realisiert wird. Insgesamt beträgt also die Wahrscheinlichkeit dafür, dass beide Objekte nach der Begegnung getrennt verfolgt werden, $2 \cdot p \cdot q = 2 \cdot p(1 - p)$. Die Wahrscheinlichkeit dafür, dass der SJPDF beide Male den Track der Bahn der Person zwei zuordnet, ist schließlich q^2 . Die Wahrscheinlichkeiten p und q hängen nun von den Winkeln θ_1 und θ_2 ab, die die Bahnen der beiden Personen mit der bisherigen gemeinsamen Bewegung am Ende der Begegnung einschließen.

Im symmetrischen Fall, d.h. falls gilt $\theta_1 = \theta_2$ (Abbildung 8.5), ist auch $p = q$ und damit folgt mit Gleichung 8.14

$$p = q = \frac{1}{2}. \quad (8.15)$$

Die Wahrscheinlichkeit, dass beide Objekte nach der Begegnung getrennt verfolgt werden, ist damit in diesem Fall $2 \cdot p \cdot q = \frac{1}{2}$, wobei in genau der Hälfte der Fälle, d.h. in insgesamt 25 Prozent, auch die Zuordnung richtig ist. Die Wahrscheinlichkeit, dass beide Objekte jeweils mit der Bahn der Person eins oder zwei assoziiert werden, beträgt ebenfalls jeweils 25 Prozent.

Im asymmetrischen Fall, d.h. falls gilt $\theta_2 \gg \theta_1$ (Abbildung 8.6), ist $p \gg q$. Somit ist der Fall am wahrscheinlichsten, dass beide Tracks der Bahn der Person eins zugewiesen werden, während die beiden Fälle, dass beide Personen nach der Begeg-

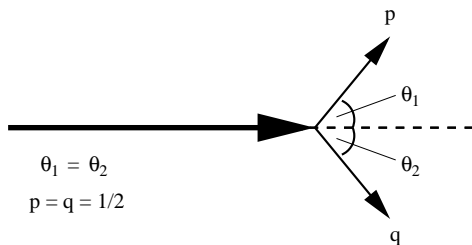


Abbildung 8.5: Symmetrisches Verhalten der beiden Personen am Ende der Begegnung.

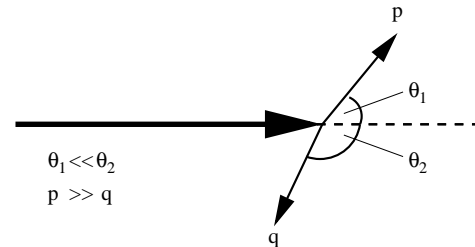


Abbildung 8.6: Asymmetrisches Verhalten der beiden Personen am Ende der Begegnung.

Zwei interagierende Ziele.

nung getrennt verfolgt werden, und zwar jeweils mit oder ohne richtige Zuordnung, eine mittlere Wahrscheinlichkeit haben. Schließlich hat der Fall, dass beide Tracks der Bahn der Person zwei zugeteilt werden, nur eine sehr kleine Wahrscheinlichkeit.

Abbildung 8.7 zeigt abschließend die Wahrscheinlichkeit für die getrennte Verfolgung beider Objekte nach der Begegnung in Abhängigkeit von der Wahrscheinlichkeit p .

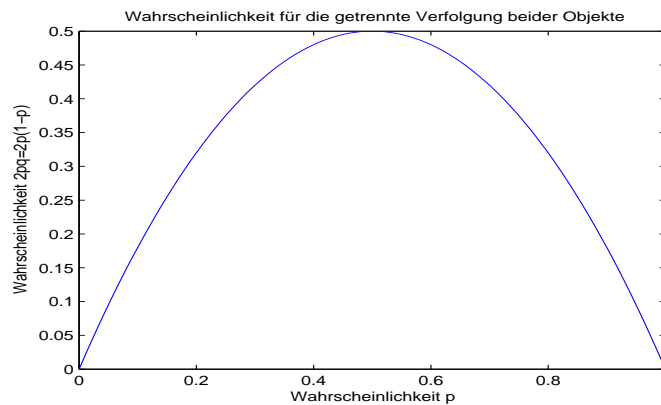


Abbildung 8.7: Die Wahrscheinlichkeit für die getrennte Verfolgung beider Objekte nach der Begegnung in Abhängigkeit von der Wahrscheinlichkeit p .

8.3 Schlussfolgerungen

Der ECSA wurde mit dem SJPDAF bezüglich der Kriterien der Genauigkeit und der Komplexität bei der Positionsschätzung eines kreisförmigen Objektes verglichen. Außerdem wurde als drittes Kriterium die Fähigkeit, zwei Personen am Ende einer Interaktion wieder getrennt zu verfolgen, betrachtet. Der ECSA wies dabei bezüglich aller dreier Kriterien Vorteile auf und kann somit als eine Verbesserung des Standes der Forschung, den der SJPDAF bisher darstellt, angesehen werden.

Es gibt aber noch weitere Unterschiede zwischen den beiden Verfahren. So versucht der ECSA über die in Kapitel 10.1 vorgestellte Methode stets die genaue Anzahl der Personen zu ermitteln. Der SJPDAF dagegen schätzt die Anzahl der Personen mit einem Bayesschen Filter, so dass die Anzahl der computerinternen Tracks nicht notwendigerweise mit der Anzahl der realen Personen oder auch der Features übereinstimmt. Auf diese Weise sollen Fehler bei der Detektion der Features geglättet werden. Letzten Endes wird durch dieses Vorgehen mit dem SJPDAF immerhin in ca. neunzig Prozent aller Scans die korrekte Anzahl an Personen geschätzt, wie experimentelle Untersuchungen gezeigt haben. Die Features werden dabei wie auch bei dem in Abschnitt 10.1 vorgestellten Verfahren durch die Bestimmung von lokalen Minima in den Abstandsmessungen der Laserscanner detektiert. Dazu werden in (Schulz, 2002) Entfernungshistogramme benutzt. Um statische Objekte auszuschließen, wurden, ähnlich wie in der vorliegenden Arbeit (Bedingung 4 in Kapitel 10.1), auch noch die Unterschiede in aufeinanderfolgenden Scans mit Hilfe von Besetzungskarten (Moravec und Elfes, 1985) berücksichtigt. Sowohl beim ECSA als auch beim SJPDAF ist die gewählte Methode aber, insbesondere wenn der Roboter sich bewegt, nicht absolut zuverlässig. Einige Gedanken zu grundsätzlichen Unterschieden zwischen dem ECSA und dem SJPDAF bei der Verfolgung von mehreren Personen, insbesondere von Personengruppen, finden sich auch in der Einleitung zu Kapitel 7.

Vergleicht man aber die Ergebnisse des SJPDAF mit den Resultaten der Verfahren aus Kapitel 4, wie sie in Kapitel 5 und 6 dargestellt sind, so erkennt man, dass auch der SJPDAF noch ein sehr guter Ansatz ist, der allen einfachen Algorithmen aus Kapitel 4 überlegen ist, wenn man alle drei Kriterien berücksichtigt.

Kapitel 9

Ansätze zur Identifikation von Personen

9.1 Einleitung

Da bisher nur von Laserscannern generierte Entfernungsmessungen analysiert wurden, kann die Zuordnung der Personen zu den computerinternen Tracks nach einer Interaktion vertauscht sein, d.h. die Kenntnis über die Identitäten geht verloren. Dieses Problem tritt auch bei anderen Verfahren auf, die nur Laserabstandsmessungen als Messdaten verwenden (Bellotto und Hu, 2006; Schulz, 2006a), da bei hoher Objektdichte die Analyse des Bewegungsverhaltens oft nicht ausreicht, um Messungen zuverlässig Tracks zuordnen zu können. Die Personen, die am Ende einer Interaktion durch den ECSA mit den Tracks assoziiert werden, müssen deshalb neu identifiziert werden (Schulz, 2006b). Da die Probanden sich, außer bei der Methode von Schulz (Schulz et al., 2003b), wo sie Signale mit einer charakteristischen Frequenz ausstrahlen, nicht selbst identifizieren können, spricht man hierbei auch von Surveillance im Gegensatz zu Control, wo die Objekte sich selbst identifizieren, wie zum Beispiel über Funk in der Luftfahrt.

Das Wissen über die Identitäten, die den Tracks zugeordnet sind, kann im Prinzip auch das Wissen über die Positionen verbessern. Wenn Personen interagieren, können auf diese Weise nämlich eventuell den Personen ihre aktuellen Tracks zugeordnet werden (Bellotto und Hu, 2007b). Zur Identifizierung der Personen werden nun neben den Positionsdaten, die mit den Laserscannern gemessen werden, auch Identitätsdaten benötigt. Ein Beispiel für ein System, das sowohl Positionsdaten als auch Identitätsdaten misst, ist der Mensch, der üblicherweise keine Probleme hat, ihm bekannte Personen in seiner Umgebung zu lokalisieren und zu identifizieren.

Allgemein benötigt ein System zur Erkennung von Personen sowohl geeignete aktuelle Messungen als auch a-priori-Wissen über charakteristische Eigenschaften der Menschen in seiner Umgebung, so dass die aktuellen Messungen im Rahmen des Identifizierungsprozesses jeweils mit dem a-priori-Wissen verglichen werden können. Das

a-priori-Wissen wird dabei in der Regel in einer Datenbank zusammengefasst oder in einer frühen Phase zu Beginn des eigentlichen Versuches gelernt. Die Fähigkeit zur Unterscheidung von Personen ist deshalb von besonderer Bedeutung, weil ein Roboter in vielen Fällen mit einer bestimmten Person interagieren soll, zum Beispiel bei der Betreuung älterer Menschen in Altenheimen.

Die Zuordnung von Identitäten geschieht im Allgemeinen in mehreren Schritten. Zunächst müssen aus den Sensor-Daten sogenannte Features extrahiert werden, d.h. die Rohdaten müssen in diese Features transformiert werden. Die Features repräsentieren dabei den charakteristischen Anteil der ursprünglichen Information und erlauben damit einen schnelleren Vergleich. In der Literatur wird gelegentlich auch zwischen Features und Attributen unterschieden. Attribute können dabei nur diskrete Werte annehmen (Bar-Shalom und Chen, 2007). Die Wahl der Features hängt von dem physikalischen Wissen über die Objekte und den verfügbaren Sensoren ab. In Experimenten muss dabei jeweils überprüft werden, ob die Features geeignet sind und eine ausreichende zeitliche Konstanz auch in Abhängigkeit von den Umweltbedingungen (z.B. Beleuchtungsverhältnisse bei Verwendung einer Kamera) gegeben ist (Reliabilität). Außerdem muss eine ausreichend kohärente Beziehung zwischen den Features und den Identitäten bestehen (Validität). Aus diesen Features müssen dann Muster erkannt werden (Pattern Recognition), beispielsweise durch Vergleich mit Objekten oder Bereichen im sogenannten Feature-Raum. Die Features müssen dabei auf geeignete Weise repräsentiert werden, zum Beispiel indem aus Bildsegmenten Information über die Form und das Lichtspektrum von Personen und deren Bekleidung gewonnen wird. Andere Charakteristika zur Personenerkennung können auch das Bewegungsmuster oder die Stimme einer Person sein. Der Feature-Raum muss in der Regel in einer Trainingsphase gelernt werden. Diese Trainingsphase muss entweder dem eigentlichen Versuch vorgeschaltet sein oder findet statt, solange die Objekte wohlsepariert sind. Im letzteren Fall kann dieser Prozess in den eigentlichen Tracking-Vorgang integriert sein. Dabei müssen sowohl die Features ausgewählt als auch ein Abstandsmaß oder eine Metrik definiert werden. Die eigentliche Klassifikation der Personen geschieht dann aufgrund der Distanzen zu den gelernten Objekten.

Neben einem Feature-Raum kann das Vorwissen auch mit Hilfe von neuronalen Netzen repräsentiert werden, die dann auch die Klassifikation während des eigentlichen Versuchs vornehmen können. Neuronale Netze wurden beispielsweise dazu verwendet, um mit Hilfe von Radardaten Schiffe zu identifizieren (Preve und Marchette, 1987). Eine weitere Möglichkeit besteht in der Verwendung von physikalischen Modellen, aufgrund derer Vorhersagen über die zu erwartenden Messungen von den einzelnen Objekten gemacht werden. Die Folgerungen über die Identitäten können auch auf Wissen beruhen (knowledge-based inference). Dabei werden Regeln verwendet, die beispielsweise in einem Expertensystem abgebildet sind. Probleme können bei Zugrundelegung eines Expertensystems aus den Unsicherheiten in den Regeln und den Daten entstehen. Ein Expertensystem kann seine Folgerungen, die meist auf einem höheren Niveau stattfinden, z.B. mit Hilfe der Fuzzy-Logik ausführen. Es führt dabei nicht in erster Linie analytisch-numerische Berechnungen durch, sondern verwendet

zum Beispiel Ursache–Wirkungs–Beziehungen oder Vorhersagen. Expertensysteme werden oft auch zur Lagebeurteilung, vor allem im militärischen Kontext, verwendet.

Werden mehrere Sensoren zur Bestimmung der Identitäten eingesetzt, so können die Informationen auf der Ebene der Rohdaten, auf der Ebene der Features oder auf der Ebene der jeweils schon zugeordneten Identitäten zusammengeführt werden. Im letzteren Fall können z.B. klassische Wahrscheinlichkeiten, die Formel von Bayes oder die Theorie von Dempster–Shafer benutzt werden. Ein Beispiel für die Fusionierung von Daten auf dieser Ebene (Decision–Level) im Bereich der Robotik findet sich in (Luo und Lin, 1988). Die Fusionierung auf der Ebene der Daten setzt voraus, dass Sensoren vorliegen, die ähnliche Daten liefern.

Für einen mobilen Roboter sind nun verschiedene Ansätze denkbar, die diese Identitätsdaten bereitstellen könnten und eine automatische Objekterkennung (Automatic Target Recognition) erlauben. Sie unterscheiden sich bezüglich der verwendeten (zusätzlichen) Sensoren und der betrachteten Eigenschaften der Personen:

1. die Hosen, die die Probanden tragen, haben oft verschiedene Farben oder Oberflächenbeschaffenheiten, da sie aus unterschiedlichen Materialien bestehen können. Dies könnte zu unterschiedlichen Intensitäten der reflektierten Laserstrahlen führen. Diese Intensitätsunterschiede können aber mit den Laserscannern gemessen werden. Dieser Umstand beruht auf der Tatsache, dass die Laserscanner üblicherweise so an die Roboter montiert werden, dass die ausgesandten Laserstrahlen die Personen in der Höhe ihrer Hosen treffen.
2. die Personen können Sender tragen, die Infrarot– oder Ultraschallsignale charakteristischer Frequenz ausstrahlen, die von statischen Sensoren in der Umgebung registriert werden (Schulz et al., 2003b).
3. zusätzliche Information könnte auch mit einer Kamera oder einem Kameranetzwerk (Vision) generiert werden. Denkbar wäre es zum Beispiel, die Personen aufgrund
 - (a) ihrer Physiognomie wie Körperform und –größe sowie des Bewegungsverhaltens wie in (Schulz, 2006a) oder
 - (b) der Farbe ihrer Bekleidung wie in (Cielniak et al., 2007; Munoz-Salinas et al., 2007; Schumitsch et al., 2006a) zu unterscheiden.

Außer im ersten Fall muss dabei mit Methoden der Sensor–Daten–Fusion die räumliche Information der Laserabstandsmesser zur Positionsschätzung mit den Attributinformationen zur Identifikation kombiniert werden.

Die Verfahren zur Identifikation von Personen durch Messung von Attributinformationen können nach den folgenden Kriterien klassifiziert werden:

1. müssen die Attribute vor dem eigentlichen Versuch gelernt werden?

2. handelt es sich um natürliche oder künstliche Attribute, d.h. Attribute, über die eine Person normalerweise nicht verfügt?
3. wird zur Messung der Information über die Attribute ein zusätzlicher Sensor benötigt (Problem der Datenregistrierung)?
4. handelt es sich um aktive oder passive Sensoren?
5. erlaubt das Verfahren eine exakte oder nur eine probabilistische Zuordnung der Identitäten, d.h. werden mehrere Möglichkeiten mit ihren zugehörigen Wahrscheinlichkeiten angeboten?

Falls die Attribute vor dem eigentlichen Versuch in einer Trainingsphase gelernt werden, so kann die zugehörige Information in einer Datenbank gespeichert werden. Während des Versuches können dann die aufgespürten Probanden den aus der Datenbank bekannten Personen zugeordnet werden, und zwar auch schon vor einer Interaktion. Methoden, die eine Datenbank benötigen, können aber beispielsweise nicht für Überwachungszwecke in Bahnhöfen oder auf Flughäfen eingesetzt werden, da hier permanent bisher unbekannte Personen erscheinen.

Mit dem Begriff des Data-Alignment oder der Datenregistrierung bezeichnet man das Problem, ein gemeinsames räumliches und zeitliches Koordinatensystem für die Messungen verschiedener Sensoren zu finden bzw. die Koordinatensysteme verschiedener Sensoren aufeinander abzubilden. Außerdem müssen die Objekte, die jeweils mit den verschiedenen Sensoren detektiert werden, einander zugeordnet werden. Es handelt sich dabei um einen essentiellen Vorverarbeitungsschritt, um Messungen von unabhängigen Sensoren miteinander in Beziehung zu setzen. Hierfür werden in der Regel Korrelationen zwischen den Daten von den verschiedenen Sensoren ausgenutzt. Das Problem der Datenregistrierung spielt auch auf anderen Gebieten eine große Rolle, wie zum Beispiel in der Medizin. So gibt es an der Universität Bonn eine Kooperation zwischen Mathematikern und Neurochirurgen. In deren Rahmen wird versucht, Daten aus einem Operationsmikroskop mit denjenigen aus einem Computertomographen abzugleichen, um bei Gehirnoperationen eine höhere Genauigkeit zu erreichen (Luerweg, 2007).

Ein aktiver Sensor wechselwirkt mit dem Objekt über ausgesandte Signale, während ein passiver Sensor lediglich die natürlichen Emissionen der Objekte beobachtet. Ein Beispiel für einen aktiven Sensor ist ein Laserscanner, während eine Kamera einen passiven Sensor darstellt.

Die einzelnen Verfahren weisen bezüglich dieser Kriterien folgende Charakteristika auf:

- Verfahren 1: Da das Reflexionsverhalten von Hosen von der Entfernung abhängt und außerdem sehr starke Schwankungen auftreten, muss vor den eigentlichen Versuchen eine ausreichende Menge von Trainingsdaten gewonnen werden, aus denen dann Eichkurven bestimmt werden. Diese sollen die Abhängigkeit des

Reflexionsverhaltens von der Entfernung widerspiegeln. Da die meisten Menschen Hosen tragen oder ansonsten auch die Haut betrachtet werden kann, handelt es sich um natürliche Attribute. Da die Reflexionswerte ebenfalls mit dem Laserscanner gemessen werden, ist im Gegensatz zu den anderen Verfahren kein zusätzlicher Sensor nötig. Es entfällt somit das Problem der Datenregistrierung. Wie schon erwähnt, sind Laserscanner aktive Sensoren, da sie Laserstrahlen aussenden, die von den Hosen der Personen reflektiert werden. Die Zuordnung ist allerdings nur probabilistisch, d.h. eine bestimmte Zuordnung ist nur mit einer gewissen Wahrscheinlichkeit die Richtige.

- Verfahren 2: Da die Frequenzen, unter denen die einzelnen Sender Signale aussenden, bekannt sind, müssen die Attribute vor dem Versuch lediglich bestimmt, aber nicht gelernt werden. Da Menschen normalerweise keine Infrarot- oder Ultraschallsender tragen, handelt es sich allerdings im Gegensatz zu den anderen Verfahren um künstliche Attribute. Da zahlreiche zusätzliche Sensoren zur Messung der Infrarot- und Ultraschallsignale nötig sind, tritt auch das Problem der Datenregistrierung auf. Darüberhinaus können die Sensoren nicht einmal an einem einzelnen Roboter befestigt werden, sondern müssen über das gesamte Gebäude verteilt angebracht sein. Da die Sensoren lediglich Signale registrieren, die von den Personen ausgesandt werden, handelt es sich um passive Sensoren. Da die Sender jeweils Signale einer genau definierten Frequenz aussenden, ist aber eine exakte Zuordnung der Identitäten möglich, es sei denn mehrere Personen halten sich in der Reichweite desselben Sensors auf.
- Verfahren 3a: Die jeweilige Körperform und -größe muss zusammen mit anderen Kenngrößen vor dem Versuch in einem komplizierten Verfahren aus Trainingsdaten gelernt werden. Es handelt sich aber wieder um natürliche Attribute. Es ist ein zusätzlicher Sensor in Form einer Kamera nötig, so dass wieder das Problem der Datenregistrierung auftritt. Da Kameras verwendet werden, handelt es sich um passive Sensoren. Das Verfahren liefert auch wieder nur eine probabilistische Zuordnung.
- Verfahren 3b: Die Farbhistogramme der Bekleidung, bei denen es sich um natürliche Attribute handelt, können zu Beginn des eigentlichen Versuchs gelernt werden, da hier schon sehr wenige Messungen genügen. Zur Gewinnung der Information ist aber wieder ein zusätzlicher Sensor in Form einer Kamera nötig, der somit auch wieder ein passiver Sensor ist. Außerdem ist die Zuordnung gleichfalls nur probabilistisch.

Bei der Verwendung von Kameras können Probleme auch aufgrund des begrenzten Sichtfeldes auftreten, solange keine omnidirektionale Kamera oder ein Kameranetzwerk verwendet wird. Tabelle 9.1 fasst diese Erörterungen nochmals zusammen (vergleiche auch (Kräußling und Schulz, 2008)).

Tabelle 9.1: Charakterisierung der eingeführten Verfahren zur Personenidentifikation.

Verfahren	1: Reflexion	2: Sender	3a: Körperform	3b: Farbhistogramme
Trainingsphase	+	-	+	-
natürliche Attribute	+	-	+	+
zusätzlicher Sensor	-	+	+	+
aktiver Sensor	+	-	-	-
exakte Zuordnung	-	+	-	-

Daneben kann man auch noch versuchen, Personen über ihre Gesichter (Zhou und Chellappa, 2002) oder ihre Stimmen zu identifizieren. Außerdem kann das Bewegungsverhalten der Probanden mittels Laserscannern analysiert werden, wie in (Song et al., 2008). Dort wurde unter anderem die durchschnittliche Geschwindigkeit, die Schrittlänge und die Gehfrequenz einer Person betrachtet. Schließlich wurden in jüngster Zeit auch Versuche unternommen, Personen, von denen ein Gefährdungspotenzial ausgehen kann, zu identifizieren und zu verfolgen. Beispielsweise ist es mit chemischen Sensoren möglich, Menschen, die Sprengstoff mit sich führen, aufzuspüren (Becher et al., 2008; Wienecke et al., 2008a; Wienecke et al., 2008b). Dies könnte beispielsweise den Flugverkehr sicherer machen.

Zusammenfassend kann festgestellt werden, dass es keinen perfekten Sensor bzw. kein perfektes System aus Sensor und Algorithmus zur Feststellung der Identitäten von Personen gibt, der dieses Problem vollständig und zuverlässig behandelt. Am ehesten lässt sich eine geeignete Lösung durch die Kombination mehrerer Verfahren erreichen (Munoz-Salinas et al., 2007). Vermutlich deshalb verwenden ja auch biologische Systeme wie der Mensch mehrere Ansätze gleichzeitig, wie binokulares Sehen und binaurales Hören, und ein sehr komplexes Verarbeitungssystem in Form mehrerer verschalteter Neuronensysteme und jahrelange Übung, um ihre Artgenossen anhand zahlreicher Attribute wie deren Gesichter, Größe, Körperbau, dem Bewegungsverhalten und der Stimme zu erkennen. Außerdem können die vorgestellten Charakteristika, wie zum Beispiel die Bekleidung, insbesondere bei einer größeren Anzahl von Personen, teilweise so ähnlich sein, dass eine Unterscheidung nicht mehr möglich ist. Im Folgenden werden zwei der eingeführten Methoden näher untersucht, nämlich die Erkennung mittels der Farbhistogramme der Bekleidung und des Reflexionsverhaltens der Hosen.

Die Identifikation von Objekten im Falle von Auflösungskonflikten spielt auch in anderen Bereichen der Wissenschaft eine große Rolle, so zum Beispiel bei Teilchenexperimenten in der Physik. Dort wird beispielsweise versucht, in Streuexperimenten neue, theoretisch vorhergesagte Teilchen, zu erzeugen und nachzuweisen. Hierzu müssen die Teilchen, die bei einem solchen Streuexperiment entstehen, identifiziert werden. Dazu werden unter anderem elektrische und magnetische Felder verwendet. Aufgrund der Lorentzkraft bewegen sich nämlich geladene Teilchen in solchen elektromagnetischen Feldern auf einer Schraubenlinie. Aus der Ganghöhe und dem Ra-

dius einer solchen Schraubenlinie kann man dann auf die Masse und die Ladung der Teilchen schließen und diese so identifizieren. Weitere mögliche Systeme sind Kalorimeter, Photomultiplier, Szintillationszähler und Ionisationskammern (Gruben, 1993; Kleinknecht, 2005). Auch hier gibt es aber keinen idealen Detektor, mit dem alle verschiedenen Teilchen bestimmt werden könnten. Deshalb werden in Teilchenbeschleunigern, wie dem großen Hadronen-Beschleuniger, der sich im Europäischen Kernforschungszentrum CERN in Genf befindet, sogenannte Großdetektoren verwendet. Diese setzen sich aus zahlreichen einzelnen Detektoren unterschiedlichen Typs zusammen.

9.2 Die Verwendung von Farbinformation mit Hilfe eines Kamerasystems

Die gemeinsame Verwendung von Bilddaten und räumlichen Daten von verschiedenen Sensoren kann gegenüber Einzelsensoren eine signifikante Verbesserung bezüglich der Detektion und der Klassifizierung von Objekten und der Situationserfassung aufweisen. Im englischen Sprachraum spricht man bei einer solchen Vorgehensweise auch von *multiple sensor imagery*. Hierbei treten wieder die Probleme des Alignment und der Datenregistrierung auf. Unter Alignment versteht man wie schon erwähnt die Zuordnung der Daten von verschiedenen Sensoren in ein gemeinsames räumliches und zeitliches Koordinatensystem. Datenregistrierung bezeichnet die Zuordnung von Objekten zu Daten. Grundvoraussetzung für das Alignment ist eine Korrelation zwischen ähnlichen Sensoren. Die hier verwendete Kombination von Laser und Vision wird in der Robotik auch in anderen Bereichen, wie zum Beispiel auf dem Gebiet der Outdoor-Robotik, eingesetzt (Douillard et al., 2008).

Das hier vorgestellte kombinierte Verfahren zum Tracking und zur Identifikation von Personen wurde in (Kräußling et al., 2008) vorgestellt. Um die benötigten Informationen aus den von der Kamera erzeugten Bildern zu extrahieren, müssen die Daten zunächst vorverarbeitet werden. Es gibt dabei verschiedene Möglichkeiten, eine bestimmte mit einer Kamera gemessene Farbe in einem geeigneten Raum darzustellen, wie zum Beispiel das RGB-Schema (Red, Green, Blue) und die HSV-, HSB- und HLS-Farbräume (Burger und Burge, 2005; Gonzalez und Woods, 1992). Im Rahmen dieser Arbeit wird dem HSV-Farbraum der Vorzug gegeben. Da die Kamera die Daten aber im RGB-Format liefert, müssen diese zunächst umgerechnet werden. Die zugehörigen Formeln finden sich in der gerade angegebenen Literatur.

Im Kontext des HSV-Raums steht H für das englische Wort Hue, das Farbton bedeutet. S steht für Saturation (Farbsättigung) und V für Value (Helligkeit). Da beispielsweise die Helligkeit oder auch die RGB-Werte im RGB-Raum mit der Beleuchtung variieren, wird hier nur der Farbton im HSV-Raum, der eine stabile Eigenschaft ist, zur Unterscheidung der Personen herangezogen. Dies ist gerade der wesentliche Vorteil bei Verwendung des HSV-Raums. Der Farbton kann dabei Werte von 0° bis

360° annehmen. Tabelle 9.2 zeigt die Zuordnung einiger aus dem Alltag bekannter Farben zu diesen Gradwerten.

Tabelle 9.2: Farbton und Winkelgrade im HSV-Farbraum.

Farbe	Rot	Gelb	Grün	Cyan	Blau	Magenta
Winkelgrade	0°	60°	120°	180°	240°	300°

Zur Gewinnung der Farbinformation wurde eine Farbkamera (Vision) mit einer Infrarot- oder Wärmekamera kombiniert, wie in Abbildung 9.1 dargestellt. Die thermische Kamera hatte dabei den Zweck, diejenigen Bereiche im Bild der Farbkamera zu finden, die von Menschen herrühren. Dabei handelt es sich um einen gebräuchlichen Ansatz (Treptow et al., 2005).

Bei der Farbkamera handelte es sich um eine CCD Kamera (Charge Coupled Device) mit einem F 2.0 Objektiv zur Messung der Farbinformation bezüglich der Bekleidungen der Versuchspersonen. Da die beiden Kameras parallel und unmittelbar benachbart montiert waren, war die Assoziation bzw. das Alignment zwischen diesen beiden Sensoren sehr leicht. Da die beiden Koordinatensysteme aufgrund dieser Versuchsanordnung praktisch zusammenfallen, konnte angenommen werden, dass jeder Punkt in beiden Koordinatensystemen identische Koordinaten hat. Deshalb war die Zuordnung der Bereiche im Bild der Wärmekamera, die aufgrund der höheren Temperatur im Vergleich zum Hintergrund wie den Wänden oder dem Fußboden vermutlich Menschen zuzuordnen sind, zu den entsprechenden Regionen im Bild der CCD Kamera sehr leicht. Man spricht in einem solchen Fall im englischen Sprachgebrauch auch von *co-measurable sensors*. Im Bereich Computer Vision, wo statische Kameras verwendet werden, ist das Herausfiltern der Personen aus dem Bild dagegen wesentlich einfacher (Choo und Fleet, 2001; Sigal et al., 2004).



Abbildung 9.1: Ein B21-Roboter mit den Laserscannern und dem Kamerasystem (oberer Bereich). Die linke Kamera ist die Wärmekamera.

Die Sensoren liefern nun nur Rohdaten. Kompliziertere Sensoren können auch Features oder Entscheidungen, zum Beispiel über Positionen oder Identitäten, liefern. Da Menschen in geschlossenen Räumen in der Regel eine höhere Temperatur als ihre Umgebung aufweisen, kann man sie, wie schon angedeutet, mit einer Infrarotkamera oder

thermischen Kamera, die ein zweidimensionales Wärmeemissionsbild der Umgebung liefert, von der Umgebung unterscheiden (Bestimmung sogenannter Hot Spots), wie zum Beispiel in (Cielniak et al., 2007). Damit ist es möglich, zu entscheiden, welche Bereiche des von der Farbkamera aufgenommenen Bildes von Menschen stammen. Die Farbpixel dieser einzelnen Bereiche werden dann entsprechend Tabelle 9.2 zugeordnet. Dabei wird der 360° -Raum in sechs Bereiche von je 60° unterteilt und die Pixel diesen Bereichen, die den sechs Grundfarben aus Tabelle 9.2 entsprechen, zugewiesen. Auf diese Weise erhält man sogenannte Farbhistogramme.

Im Rahmen der Experimente wurden drei Personen betrachtet. Der erste Proband trug einen blauen Pullover und eine blaue Hose. Die zweite Person trug einen gelben Pullover und ebenfalls eine blaue Hose. Die dritte Person schließlich trug eine rote Hose und ein rotes Hemd. Abbildung 9.2 zeigt die Farbverteilungen der drei Versuchspersonen. Die erste Person zeigt ein Maximum im blauen Bereich, die zweite Person zeigt ein Maximum im gelben Bereich und die dritte Person zeigt ein Maximum im roten Bereich. Die gemessenen Werte stimmen also gut mit den tatsächlichen Farben überein.

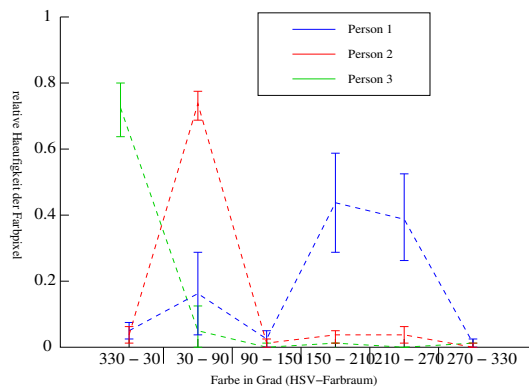


Abbildung 9.2: Farbverteilungen der drei Versuchspersonen.

Ein Problem bei dieser Versuchsanordnung besteht darin, die von dem Kamerasystem detektierten Bilder den Tracks, die mit Hilfe der Entfernungsmessungen der Laserscanner berechnet werden, zuzuordnen (Datenregistrierung), da das horizontale Koordinatensystem des Lasers nicht auf das Bild des Kamerasystems projiziert werden kann. Dabei ist hilfreich, dass das Kamerasystem eine Information darüber liefert, in welcher Reihenfolge die detektierten Personen im Uhrzeigersinn angeordnet sind. Auf der anderen Seite kann nämlich aus den mit dem ECSA geschätzten Positionen der Probanden und der Position und Orientierung des beobachtenden Roboters genug Information gewonnen werden, um die Tracks ebenfalls im Uhrzeigersinn anzuordnen. Dazu werden einfach die Winkel berechnet, denen die Positionsschätzungen in dem Koordinatensystem entsprechen, in dem sich der Roboter in der Mitte befindet. Durch Vergleich der beiden Reihenfolgen können dann die Farbhistogramme den Tracks zu-

geordnet werden, da nur der vordere Laser verwendet wird und dieser in dieselbe Richtung wie die Kamera orientiert ist.

Ein Vorteil dieses Settings besteht darin, dass die Farbcharakteristika der einzelnen Personen zu Beginn des eigentlichen Versuchs gelernt werden können. Dazu wird zunächst angenommen, dass die relativen Häufigkeiten der Farbtöne ungefähr nach einer Normalverteilung, die durch Mittelwert und Kovarianz charakterisiert wird, verteilt sind. Tatsächlich handelt es sich allerdings um Dirichlet-Verteilungen. Mit Rücksicht auf die Komplexität des Verfahrens wird diese kleine mathematische Ungenauigkeit aber zunächst in Kauf genommen. Später wird dann auch noch der Fall einer Dirichlet-Verteilung untersucht.

Sei nun der zugehörige empirische Mittelwert für die Person i $\mu_i \in \mathbb{R}^6$ und die zugehörige Kovarianz $\Sigma_i \in \mathbb{R}^{6 \times 6}$. Seien außerdem $y_{i,j}$, $j = 1, \dots, k$ die k Beobachtungen des Farbvektors für die Person i bis zum Zeitpunkt k . Dann gilt für den empirischen Mittelwert $\mu_{i,k}$ nach k Messungen

$$\mu_{i,k} = \frac{\sum_{j=1}^k y_{i,j}}{k} \quad (9.1)$$

und für die empirische Varianz $\Sigma_{i,k}$

$$\Sigma_{i,k} = \frac{\sum_{j=1}^k (y_{i,j} - \mu_{i,k})(y_{i,j} - \mu_{i,k})^\top}{k-1}. \quad (9.2)$$

Die Werte $\mu_{i,k+1}$ und $\Sigma_{i,k+1}$ können nun online iterativ aus den Werten $\mu_{i,k}$ und $\Sigma_{i,k}$ berechnet werden, ohne dass die alten Messungen $y_{i,j}$ mit $j \leq k$ gespeichert werden müssen. Es ist nämlich

$$\mu_{i,k+1} = \frac{\sum_{j=1}^{k+1} y_{i,j}}{k+1} = \frac{k}{k+1} \frac{\sum_{j=1}^k y_{i,j}}{k} + \frac{y_{i,k+1}}{k+1} = \frac{k}{k+1} \mu_{i,k} + \frac{y_{i,k+1}}{k+1} \quad (9.3)$$

und mit

$$\Delta_{k+1} = \mu_{i,k+1} - \mu_{i,k} \quad (9.4)$$

gilt auch

$$\begin{aligned} \Sigma_{i,k+1} &= \frac{\sum_{j=1}^{k+1} (y_{i,j} - \mu_{i,k+1})(y_{i,j} - \mu_{i,k+1})^\top}{k} = \\ &= \frac{\sum_{j=1}^{k+1} (y_{i,j} - \mu_{i,k} - \Delta_{k+1})(y_{i,j} - \mu_{i,k} - \Delta_{k+1})^\top}{k} = \\ &= \frac{1}{k} \left[(y_{i,k+1} - \mu_{i,k} - \Delta_{k+1})(y_{i,k+1} - \mu_{i,k} - \Delta_{k+1})^\top + \right. \\ &\quad \left. + \sum_{j=1}^k (y_{i,j} - \mu_{i,k})(y_{i,j} - \mu_{i,k})^\top - \right. \end{aligned}$$

$$\left. - \sum_{j=1}^k (y_{i,j} - \mu_{i,k}) \Delta_{k+1}^\top - \Delta_{k+1} \sum_{j=1}^k (y_{i,j} - \mu_{i,k})^\top + k \cdot \Delta_{k+1} \Delta_{k+1}^\top \right]. \quad (9.5)$$

Aufgrund der Definition von $\mu_{i,k}$ ist nun aber

$$\sum_{j=1}^k (y_{i,j} - \mu_{i,k}) = 0 \quad (9.6)$$

und außerdem gilt aufgrund der Definition von $\Sigma_{i,k}$

$$\sum_{j=1}^k (y_{i,j} - \mu_{i,k}) (y_{i,j} - \mu_{i,k})^\top = (k-1) \Sigma_{i,k}. \quad (9.7)$$

Damit folgt insgesamt

$$\Sigma_{i,k+1} = \frac{(y_{i,k+1} - \mu_{i,k+1}) (y_{i,k+1} - \mu_{i,k+1})^\top}{k} + \frac{k-1}{k} \Sigma_{i,k} + \Delta_{k+1} \Delta_{k+1}^\top. \quad (9.8)$$

Die Lernphase findet dabei während der Zeitspanne statt, in der die Personen separat gehen und die Anzahl der vom Laserscanner und dem Kamerasystem detektierten Personen identisch ist. Sobald durch den ECSA ein Crossing festgestellt wird, endet die Lernphase und die aktuellen Werte $\mu_{i,k}$ und $\Sigma_{i,k}$ werden den Personen als Farbcharakteristika μ_i und Σ_i zugeschrieben. Wenn dann die Interaktion beendet ist und der Laserscanner und das Kamerasystem die Personen wieder unabhängig voneinander detektieren, wird mit Hilfe der gelernten Information eine neue Zuordnung zwischen Tracks und Personen bestimmt. Dazu werden die gelernten Farbstatistiken mit den aktuellen Messungen verglichen.

Sei dazu m die Gesamtzahl der Probanden und mit $y_{j,k}$, $j = 1, \dots, m$ werden die Farbmessungen bezeichnet, die zum Zeitpunkt k zum Track j gehören. Da für die Farbverteilungen eine Normalverteilung angenommen wurde, ist dann die Wahrscheinlichkeit $p_{i,j,k}$ dafür, dass der Track j aufgrund der Messung zum Zeitpunkt k zu der vor der Interaktion gelernten Person i gehört, gemäß Gleichung 3.3 gleich

$$p_{i,j,k} = \frac{1}{(2\pi)^3 (\det \Sigma_i)^{1/2}} \exp \left\{ -\frac{1}{2} (y_{j,k} - \mu_i)^\top \Sigma_i^{-1} (y_{j,k} - \mu_i) \right\}. \quad (9.9)$$

Da die Messungen mit der Farbkamera zu unterschiedlichen Zeitpunkten als unabhängig voneinander angenommen werden können, gilt dann für die Wahrscheinlichkeit $p_{i,j,k_1:k_2}$, dass der Track j vom Zeitpunkt k_1 bis zum Zeitpunkt k_2 von der Person i stammt

$$p_{i,j,k_1:k_2} = \prod_{k=k_1}^{k_2} p_{i,j,k}. \quad (9.10)$$

Die Tracks j sind den Personen i , die aufgrund ihrer Farbcharakteristika definiert sind, nach der Interaktion nun in der Regel anders zugeordnet als vor der Interaktion, d.h.

sie werden einer Permutation unterworfen. Sei deshalb σ eine solche Permutation der Zahlen von 1 bis m . Dann ist die Wahrscheinlichkeit Π_σ dafür, dass durch die Interaktion die Identitäten gemäß der Permutation σ vertauscht worden sind, gleich

$$\Pi_\sigma = \prod_{l=1}^m p_{l,\sigma(l),k_1:k_2}. \quad (9.11)$$

Dabei werden wieder die Messungen von k_1 bis k_2 zugrundegelegt.

Die beste Wiederzuordnung $\hat{\sigma}$ der Tracks j zu den Personen i ist nun diejenige, für die die Wahrscheinlichkeit Π_σ maximal wird. Um diese Permutation zu finden, geht man folgendermaßen vor. Zunächst berechnet man jeweils den negativen Logarithmus der Wahrscheinlichkeiten, da beispielsweise die Werte Π_σ sehr klein werden können. Außerdem erhält man so wegen der Beziehung $\log(a \cdot b) = \log(a) + \log(b)$ die Logarithmen der Produktwahrscheinlichkeiten nicht durch Multiplikation, sondern durch Addition.

Hier zeigt sich auch ein wesentlicher Unterschied zum Kalman–Filter. Bei diesem enthalten länger zurückliegende Messungen weniger Information, da sie bei jedem Zeittakt durch das Prozessrauschen korrumpiert werden. Deshalb werden länger zurückliegende Messwerte bei der Berechnung der aktuellen Schätzung $x(k|k)$ weniger stark gewichtet als aktuellere Messungen. Genauer ist

$$x(k|k) = \sum_{j=1}^k a_j y_j + a_0 x_0. \quad (9.12)$$

Dabei sind die y_j die Messungen zum Zeitpunkt j und die a_j sind die entsprechenden Koeffizienten oder Gewichte. Außerdem ist x_0 die Initialisierung. Die Folge der a_j ist dabei monoton steigend mit $\sum_{j=0}^k a_j = 1$.

Bei dem hier vorgestellten Filter, der aus zusätzlichen Attributmessungen, im hier vorliegenden Falle aus Farbhistogrammen, Zuordnungswahrscheinlichkeiten berechnet, verhält es sich dagegen so, dass länger zurückliegende Messungen den gleichen Informationsgehalt besitzen wie aktuellere Messwerte. Deshalb werden die Logarithmen einfach addiert, ohne diese unterschiedlich zu gewichten. Ähnliche Ansätze wurden in (Schumitsch et al., 2006a; Schumitsch et al., 2006b) verfolgt, wo die Autoren ihr Verfahren auch als Information–Form Data Association Filter oder Identity Management Kalman Filter bezeichnen. Auch dort wurden Farbinformationen benutzt, um Tracks Personen zuzuordnen. Wie schon erwähnt, wurden dort aber Tracks verloren. Eine kurze Übersicht über einige mögliche Ansätze zur Verwaltung von Information über Identitäten gibt auch (Guivant, 2008).

Die negativen Logarithmen $-\log(p_{i,j,k_1:k_2})$ werden nun in einer $m \times m$ -Matrix M angeordnet. Da der Logarithmus eine streng monoton wachsende Funktion ist, entspricht dabei der Maximumsuche bei den Wahrscheinlichkeiten die Minimumsuche bei den negativen Logarithmen, wobei hier nicht die Produkte Π_σ , sondern die entsprechenden Summen betrachtet werden müssen. Man kann hierbei auch von einer Minimierung von Kosten sprechen.

Es existieren im Wesentlichen zwei verschiedene Möglichkeiten, die optimale Permutation $\hat{\sigma}$ zu finden. Diese Suche bezeichnet man dabei auch als Bipartite Matching (West, 2000). Eine stochastische Variante ist der Metropolis–Algorithmus, der aber oft nicht die exakte Lösung findet. Diese liefert dagegen immer der Algorithmus von Munkres (Munkres, 1957), der auch als Ungarischer Algorithmus bezeichnet wird und lediglich den Nachteil hat, dass er bei großen Matrizen sehr viel Rechenzeit benötigt. Da bei der hier gewählten Versuchsanordnung aber nur eine geringe Anzahl von Versuchspersonen vorliegt und m deshalb relativ klein ist, wurde im Rahmen dieser Arbeit auf den Algorithmus von Munkres zurückgegriffen.

Im Prinzip gewichtet der hier vorgestellte Ansatz Sensormessungen mit dem Inversen einer gelernten Kovarianz und summiert, um Score–Funktionen zu erhalten (Weighted Sum Score). Ein anderer Ansatz wird in (Schulz et al., 2003b) angewandt, wo eine Identifizierung mittels Ultraschall– und Infrarotsensoren versucht wurde. Auch dort beklagten die Autoren dabei den Verlust von Tracks. Die mathematische Verarbeitung der Sensorinformation zur Objekterkennung geschah dort mit einem Bayesschen Ansatz, der die Bayessche Regel anwendet, um bedingte Wahrscheinlichkeiten zu kombinieren. Dieser Bayessche Ansatz wurde mit einem sogenannten Rao–Blackwellised Teilchen–Filter realisiert. Alternativ könnte man auch die Belief–Funktionen der Sensoren unter Anwendung der Dempster–Shafer–Theorie zusammenführen oder Fuzzy–Variablen unter Benutzung der Fuzzy–Logik.

Eine andere Möglichkeit zur Verarbeitung der Farbinformation besteht wie schon erwähnt in der Verwendung von Dirichlet–Verteilungen. Da hier relative Häufigkeiten von Farbpixeln in den Bildern von Personen untersucht werden, die sich zu Eins aufsummieren, ist dies eigentlich auch die mathematisch korrekte Vorgehensweise (Schulz, 2008). Auch hier wird ein einfaches Farbmodell der Versuchspersonen gelernt, solange diese getrennt sind. Dieses Modell besteht aus einer Verteilung über die relativen Häufigkeiten von Farben in den Bildern der einzelnen Personen. Dazu wird für jede Person i in jedem Bild k ein Vektor $y_{i,k} = (v_1, \dots, v_6)^\top$ der relativen Häufigkeiten der einzelnen Farben bestimmt. Dabei ist $0 \leq v_l \leq 1$ und die Summe der v_l ergibt Eins.

Dem Modell liegt die Annahme zugrunde, dass die Pixels einer Person in einem Bild von einer Multinomialverteilung gezogen werden, die das Aussehen des Probanden in diesem Kamerabild beschreibt. Da sich die Form einer Person von Bild zu Bild ändert und die Farbverteilung diskretisiert wurde, ist diese Multinomialverteilung für die einzelnen Bilder unterschiedlich. Es wird nun angenommen, dass diese Unterschiede selbst für jede Person charakteristisch sind und von einer Verteilung über Multinomialverteilungen gezogen werden. Im Rahmen der Bayes–Statistik ist aber die Dirichletverteilung die zur Multinomialverteilung konjugierte a–priori Verteilung (Gelman et al., 2003). Deshalb wird angenommen, dass die Vektoren $y_{i,k}$ der relativen Häufigkeiten der Farbpixel einer Person i gemessen zum Zeitpunkt k ungefähr verteilt sind nach einer Dirichletverteilung mit Parametern $\alpha_{i,k} = (\alpha_{i,k,1}, \dots, \alpha_{i,k,6})^\top$, d.h. es ist

$$y_{i,k} \sim \text{Dir}(y_{i,k}, \alpha_{i,k}). \quad (9.13)$$

Dabei ist

$$\text{Dir}(y_{i,k}, \alpha_{i,k}) = \frac{\Gamma(\sum_{l=1}^6 \alpha_{i,k,l})}{\prod_{l=1}^6 \Gamma(\alpha_{i,k,l})} \prod_{l=1}^6 v_l^{\alpha_{i,k,l}-1} \quad (9.14)$$

mit der Gammafunktion Γ die Wahrscheinlichkeitsdichteverteilung der Dirichletverteilung.

Sobald eine Interaktion beginnt, wird für jede Person i die Maximum-Likelihood Schätzung $\hat{\alpha}_i$ des Parameters der Dirichletverteilung, beruhend auf den gespeicherten Vektoren der relativen Häufigkeiten, berechnet.¹ Zur Wiedererkennung der Personen nach einer Interaktion wird ähnlich wie im Falle der Normalverteilung angenommen, dass die relativen Häufigkeiten der Farbpixel für einen Track j in aufeinanderfolgenden Bildern unabhängig voneinander und Dirichlet-verteilt sind, falls die Zuordnung zu einer Person i vorgegeben ist. Falls nach der Interaktion die Farbhäufigkeiten $y_{j,k_1}, \dots, y_{j,k_2}$ für einen Track j beobachtet werden, kann deshalb die zugehörige Wahrscheinlichkeit der Zuordnung dieses Tracks zur Person i berechnet werden gemäß

$$p_{i,j,k_1:k_2} = \prod_{k=k_1}^{k_2} \text{Dir}(y_{j,k}, \hat{\alpha}_i). \quad (9.15)$$

Das weitere Vorgehen entspricht dem bei Verwendung der Gaußverteilung.

Die gelernten Verteilungen sollen auch noch graphisch dargestellt werden. Da es schwierig ist, sechsdimensionale Dirichlet-Verteilungen abzubilden, werden jeweils nur die Randverteilungen wiedergegeben. Diese Randverteilungen sind Betaverteilungen, die einen guten Überblick darüber geben, wie die Verteilungen der sechs Farben für eine einzelne Person während des Trainings variieren. Die Abbildungen 9.3 bis 9.5 stellen diese Betaverteilungen jeweils für alle sechs Farben und für jede Person dar. Dabei bedeutet ein Peak einer Kurve im rechten Teil des Diagramms eine größere relative Häufigkeit der entsprechenden Farbe und ein Peak in der linken Hälfte dementsprechend eine kleinere Häufigkeit. Genauer gibt die Lage des Peaks an, welche relative Häufigkeit der Farbpixel die größte Wahrscheinlichkeit hat. Aus Abbildung 9.3 geht hervor, dass die erste Person am wahrscheinlichsten Kleidung der Farben Cyan und Blau trägt. Dabei muss darauf hingewiesen werden, dass Cyan Blau sehr ähnlich ist (Tabelle 9.2). Auch die gelbe Farbe tritt noch mit einer gewissen Häufigkeit auf. Aus Abbildung 9.4 erkennt man, dass die wahrscheinlichste Farbe der Bekleidung des zweiten Probanden Gelb ist, während dies für die dritte Person die rote Farbe ist (Abbildung 9.5). Somit stimmen die gelernten Verteilungen sehr gut mit den tatsächlichen Farben der Bekleidung der einzelnen Versuchspersonen überein.

Das Verfahren wurde mit realen Daten, aufgezeichnet in unserer Experimentalumgebung, getestet. Dabei wurde zur Berechnung der Positionsschätzungen wie auch im

¹Dazu wird Thomas Minkas fastfit MATLAB Toolbox verwendet:
<http://research.microsoft.com/~minka/software/fastfit/>

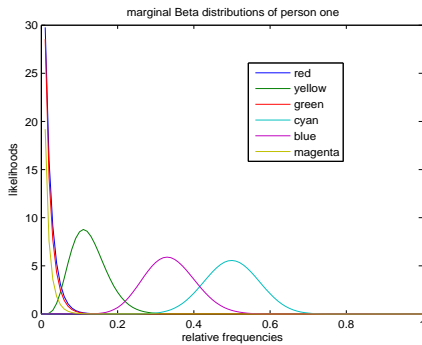


Abbildung 9.3: Beta-Randverteilungen von Person Eins.

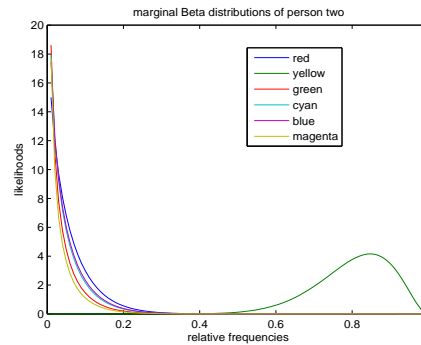


Abbildung 9.4: Beta-Randverteilungen von Person Zwei.

Beta-Randverteilungen der Personen Eins und Zwei.

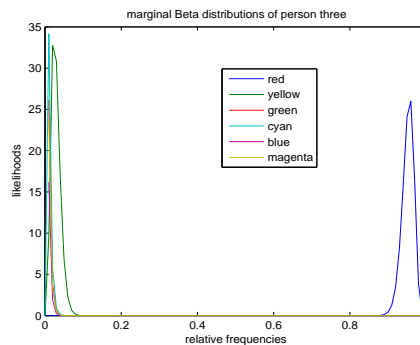


Abbildung 9.5: Beta-Randverteilungen von Person Drei.

folgenden Abschnitt 9.3 der in Kapitel 7 vorgestellte ECSA mit dem jeweiligen Ansatz zur Identifikation verknüpft. Da die verwendeten Kameras nur einen kleinen Winkelbereich von ca. sechzig Grad abdecken, wurden aus Platzgründen nur Experimente mit den oben beschriebenen drei Personen durchgeführt. Die Anzahl der Pixels war 640 in der Horizontalen und 480 in der Vertikalen. Für die Anfangs- und Endpositionen der Versuchspersonen waren drei verschiedene Bereiche in der Experimentalhalle fest vorgegeben. Zunächst wurden dabei alle möglichen sechs Permutationen der Endpositionen bezüglich einer festen Anfangsposition durchprobiert. Damit wurden insgesamt Permutationen getestet, die entweder aus gar keiner, aus einer oder aus zwei Transpositionen bestehen. Eine Transposition ist dabei eine Permutation, die genau zwei Elemente vertauscht. Da Permutationen, die nur aus einer Transposition bestehen, mathematisch nicht äquivalent sind zu Permutationen, die aus zwei Transpositionen bestehen, können damit unterschiedliche Fehler in der Implementierung des Algorithmus ausgeschlossen werden. So spielt es zum Beispiel bei der Anwendung von Munkres'

Algorithmus im Falle einer Permutation, die nur aus einer Transposition besteht, keine Rolle, ob man den Algorithmus auf M oder M^T anwendet, wohl aber im Falle einer Permutation, die aus zwei Transpositionen besteht.

Zunächst wurden den Untersuchungen Gaußverteilungen zugrunde gelegt. Als Position 1 wurde in den graphischen Darstellungen der rechte obere Teil des Überwachungsgebietes, als Position 2 der linke obere und als Position 3 der untere mittlere Bereich definiert. Die Anordnung der drei Versuchspersonen zu Beginn dieser sechs Versuche, d.h. ihre Startpositionen, waren, wie schon angedeutet, immer gleich. Vergleiche hierzu auch Abbildung 9.6. Dabei werden in den Graphiken, solange die Identitäten bekannt sind und Einzelpersonen vorliegen, die entsprechenden Ziffern an der Stelle eingezeichnet, wo der ECSA die Personen lokalisiert. Cluster und Subcluster werden dagegen wieder mit fest vorgegebenen blauen Rechtecken wiedergegeben.

Für die erste Permutation $123 \mapsto 123$, die auch als identische Permutation bezeichnet wird, soll der Versuchsablauf genauer beschrieben werden. Zunächst halten sich die drei Probanden separat voneinander im Überwachungsbereich auf (Abbildung 9.6). Der Algorithmus lernt in dieser Phase die Farbcharakteristika der drei Versuchspersonen. Die Zuordnung der Nummern, die in den Abbildungen wiedergegeben werden, erfolgt dabei in der Reihenfolge, in der die Personen initialisiert werden, d.h. den Überwachungsbereich betreten. Dann vereinigen sich Person eins und zwei zu einem Cluster (Abbildung 9.7). Als nächstes stößt Person drei zu den anderen beiden und

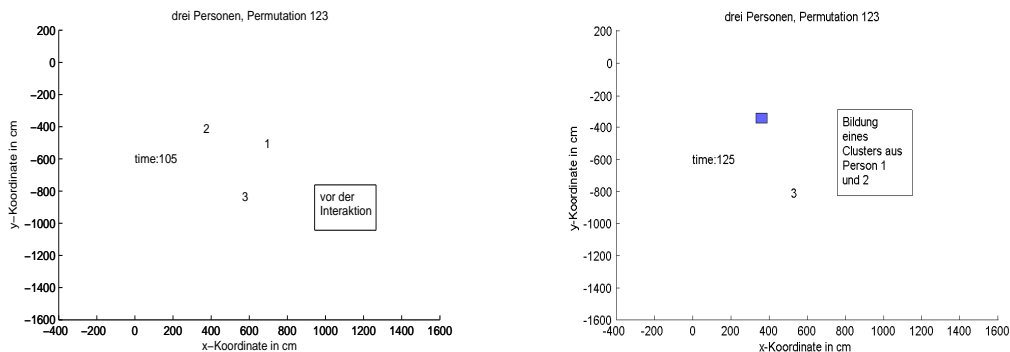


Abbildung 9.6: Drei Einzelpersonen vor einer Interaktion. Abbildung 9.7: Ein Cluster aus zwei Probanden hat sich gebildet.

Drei Personen, Permutation $123 \mapsto 123$.

es bildet sich ein kombinierter Cluster (Abbildung 9.8), der dann in zwei Subcluster zerfällt (Abbildung 9.9). Schließlich zerfällt der Cluster zunächst in drei Subcluster, die jeweils eine Person repräsentieren (Abbildung 9.10), und danach wieder in seine Einzelpersonen, die der Algorithmus sofort korrekt identifiziert (Abbildung 9.11).

Die Abbildungen D.1 bis D.6, die in Anhang D wiedergegeben sind, zeigen die Verhältnisse für die drei Permutationen, die aus einer Transposition bestehen, d.h. bei

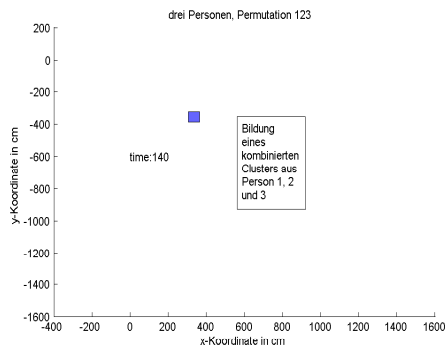


Abbildung 9.8: Der Cluster und die dritte Einzelperson vereinigen sich zu einem kombinierten Cluster.

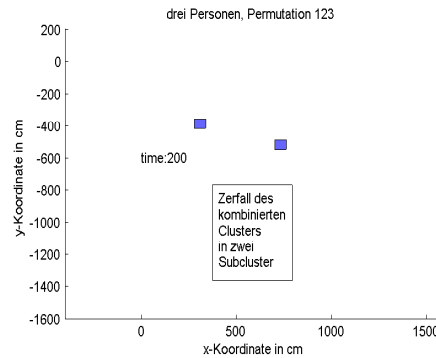


Abbildung 9.9: Zwei Subcluster haben sich gebildet.

Drei Personen, Permutation 123 \mapsto 123.

denen während der Interaktion genau zwei Personen vertauscht werden. Dabei wurden jeweils nur die Anfangs- und Endpositionen mit den zugeordneten Identitäten abgebildet. Auch hier werden den Probanden die Identitäten nach der Interaktion sofort korrekt zugeordnet.

Schließlich werden die beiden Permutationen betrachtet, die aus zwei Transpositionen bestehen. In diesem Fall nehmen alle drei Personen nach der Interaktion eine neue Position ein. Auch bei diesen Experimenten werden die Identitäten nach der Interaktion unmittelbar korrekt zugewiesen (Abbildungen D.7 bis D.10).

Nun wird untersucht, ob das Verfahren auch erfolgreich ist, wenn die drei Personen eine andere Startposition einnehmen. Genauer wurde die Anfangsposition der Probanden Eins und Zwei vertauscht (Abbildung D.11). Auch in diesem Falle wurden die Personen nach der Interaktion sofort korrekt identifiziert (Abbildung D.12).

Am Ende wurde noch der Fall analysiert, dass mehrere, in diesem Falle zwei Interaktionen, hintereinander stattfinden (Abbildungen D.13 bis D.15), so dass die Identitäten mehr als einmal neu zugeordnet werden müssen. Dabei wurde wieder die ursprüngliche Startposition zugrundegelegt (Abbildung D.13). Sowohl nach der ersten (Abbildung D.14) als auch nach der zweiten Interaktion (Abbildung D.15) erfolgte die Zuordnung wieder sofort korrekt. Eine Besonderheit des verwendeten Algorithmus ist hierbei, dass nach einer Interaktion zunächst die Zuordnung der Identitäten aufgrund der gelernten Farbverteilungen erfolgt. Danach werden die Farbhistogramme wieder neu gelernt. Auf diese Weise wird beispielsweise versucht, den Fall zu beherrschen, dass eine Person einen Pullover oder eine Jacke überstreift, so dass sich ihr Farbhistogramm ändert. Voraussetzung hierfür ist, dass dies nicht während einer Interaktion geschieht.

Das vorgestellte Verfahren ist also in der Lage, mehrere Personen ohne Trackver-

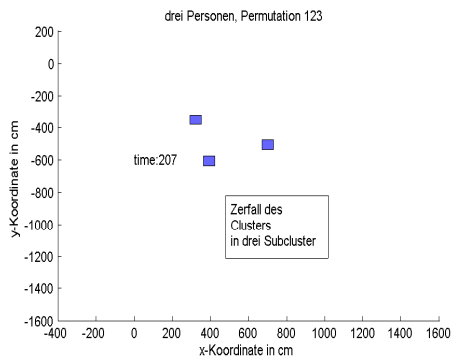


Abbildung 9.10: Der Cluster ist in drei Subcluster zerfallen.

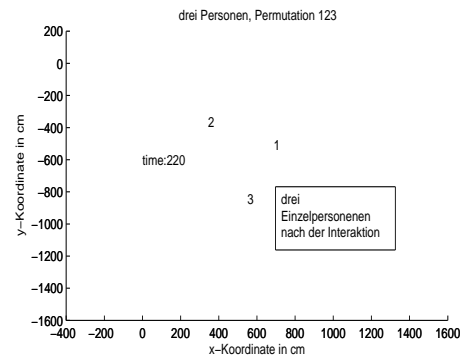


Abbildung 9.11: Die Einzelpersonen werden wieder unabhängig voneinander mit der korrekten Identität verfolgt.

Drei Personen, Permutation 123 \mapsto 123.

lust zu verfolgen und nach einer Interaktion die Personen korrekt wiederzuerkennen.

Bei Verwendung der Dirichletverteilung wurden die Identitäten für die ersten sechs Versuche, bei denen die sechs Permutationen angewandt wurden, richtig zugeordnet (Kräußling et al., 2010). Bei dem Versuch mit der veränderten Startposition waren zwei Personen nach der Interaktion vertauscht. Beim letzten Experiment, bei dem zwei Interaktionen hintereinander betrachtet wurden, waren ebenfalls nach einer Interaktion zwei Probanden vertauscht. Im Rahmen der hier betrachteten Versuchsreihe ist es also günstiger, Gaußverteilungen zu verwenden, da in diesem Falle die Zuordnung stets richtig war.

9.3 Die Verwendung der Reflexionsmessungen der Laserscanner

Die Laserscanner, die im Rahmen dieser Arbeit zur Positionsschätzung von Menschen verwendet werden, senden etwa in Höhe der Knie der Personen Laserstrahlen aus. Diese werden von den Beinen dieser Personen reflektiert. Die reflektierten Laserstrahlen gelangen dann teilweise zurück zum Laser und dieser misst dabei sowohl die Laufzeit als auch die Intensität.

Bis vor kurzem wurde in diesem Zusammenhang angenommen, dass Laser-Scanner keine ausreichende Information zur Unterscheidung und Identifikation von Personen bereitstellen:

- However, the sensor does not directly provide information that allows to distinguish between persons (Schulz, 2006a).

- As laser range scanners only provide proximity information, they cannot be used to reliably identify or distinguish between persons during tracking (Schumitsch et al., 2006a).

Dennoch ist es im Rahmen dieser Arbeit gelungen, ein Verfahren zu entwickeln, das es erlaubt, mit Hilfe der von einem Laserscanner gemessenen Reflexions- oder auch Remissionswerte die Identitäten von Personen festzustellen. Die hierbei durchgeführten Voruntersuchungen sind in Anhang E.1 dargestellt. Die Ergebnisse dieser Voruntersuchungen können dahingehend zusammengefasst werden, dass die Intensität des registrierten Laserstrahls von folgenden Faktoren abhängt:

- der Entfernung der Probanden vom Laserscanner. Für größere Entfernungen erhält man dabei kleinere Intensitäten und umgekehrt.
- den Reflexionseigenschaften des Hosenmaterials.

Das Reflexionsverhalten einer Hose hängt dabei sowohl von der Farbe der Hose als auch von ihrem Material bzw. der Oberflächenbeschaffenheit oder Textur ab. Die Remissionswerte enthalten also Information, die zur Erkennung von Personen herangezogen werden kann. Zusätzlich kann die registrierte Intensität auch noch vom Messverhalten oder der Kalibrierung des Laserscanners selbst abhängen. Darüberhinaus ist das Messverhalten für einen bestimmten Laserscanner zeitlich konstant. Diese beiden Umstände werden im Folgenden noch näher ausgeführt.

Dieses Verhalten soll das folgende Modell für die Intensitäten, die vom Laserscanner registriert werden, näher erläutern. Die Intensität des registrierten Laserstrahls, genauer der Anteil des emittierten Laserstrahls, der nach der Reflexion an der Hose wieder zum Laserscanner zurückkehrt, wird in erster Linie durch drei Prozesse beeinflusst:

- die Ausbreitung vom Laserscanner zum Ziel. Die Divergenz des Laserstrahles während dieses Zeitraumes ist äußerst gering, da sich der Laserstrahl, der vom Scanner emittiert wurde, praktisch nur in eine Raumrichtung ausbreitet. Dies führt dazu, dass das Signal, wenn es auf die Hose trifft, eine Flächenausdehnung von nur wenigen Quadratzentimetern besitzt. Meistens liegt deshalb der Fall vor, dass das Signal komplett auf die Hose fällt und von dieser reflektiert wird. Ausnahmen bestehen lediglich dann, wenn die Hose nur am Rande eines Beines gestreift wird. Dieser Fall wird später noch etwas genauer erörtert.
- die Reflexion der Photonen durch die Moleküle auf der Oberfläche des Zieles. Dabei wird nur eine bestimmte Menge der Photonen des Laserstrahls reflektiert, während der Rest absorbiert wird. Der Anteil der reflektierten Photonen hängt nun gerade vom Material der Hose einer Person ab. Somit besteht für die Intensität I_b des Laserstrahls vor der Reflexion und die Intensität I_n danach der Zusammenhang

$$I_n = \alpha I_b. \quad (9.16)$$

Dabei ist $0 \leq \alpha \leq 1$ und α ist abhängig vom Material der Hosen.

- die Ausbreitung vom Objekt zurück zum Laserscanner. Da das Signal im Moment der Wechselwirkung mit der Hose eine Flächenausdehnung von nur wenigen Zentimetern aufweist, kann die Ausdehnung des Streuzentrums bei der Reflexion als vernachlässigbar im Vergleich zu der Entfernung dieses Zentrums vom Sensor angenommen werden. Dieses Streuzentrum kann deshalb in guter Näherung als punktförmig betrachtet werden. Da die Photonen dabei von den Molekülen der Hosen zufällig in alle Raumrichtungen gestreut werden, gilt somit für die Intensität I_1 , die vom Laserscanner registriert wird, die Gleichung

$$I_1 = I_n r^{-2}. \quad (9.17)$$

Dabei ist r die Entfernung vom Ziel zum Beobachter.

Insgesamt gilt also die Beziehung

$$I_1 = \alpha I_0 r^{-2} \quad (9.18)$$

für die Intensität I_1 des registrierten Laserstrahles. Dabei ist I_0 die Intensität des vom Laserscanner emittierten Signals. Abbildung 9.12 zeigt die Abhängigkeit der Intensität des registrierten Laserstrahls von der Entfernung zwischen Beobachter und Ziel. Es besteht eine gute Übereinstimmung zu den Abbildungen E.11, E.13, E.15 und E.17 oder den entsprechenden Graphiken dieses Abschnittes, wie zum Beispiel Abbildung 9.13 oder 9.15.

Um das Reflexionsverhalten des Hosenmaterials zur Identifizierung von Personen heranziehen zu können, wurde nun folgendermaßen vorgegangen. Da die Intensität des registrierten Laserstrahls stark vom Abstand zu einer Person abhängt, muss zunächst für jeden möglichen Entfernungswert der zu erwartende Reflexionswert ungefähr bekannt sein, da die Personen in der Zuordnungsphase nach einer Interaktion beliebige Distanzen zum Laserscanner einnehmen können. Aus demselben Grund werden aber in der Regel vor einer Interaktion, also während der möglichen Lernphase, nicht alle möglichen Entfernungen abgedeckt. Außerdem müssen, da die Abhängigkeit nicht durch eine einfache Funktion beschrieben werden kann und ein erhebliches statistisches Rauschen vorliegt, für alle Abstandswerte ausreichend viele Messdaten vorliegen, um jeweils einen hinreichend aussagekräftigen Mittelwert berechnen zu können. Deshalb wurde das Reflexionsverhalten der verschiedenen Materialien in einer Trainingsphase vor dem eigentlichen Experiment gelernt. Um einen möglichst großen

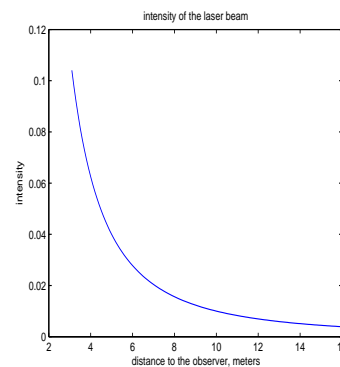


Abbildung 9.12: Intensität des registrierten Laserstrahles in Abhängigkeit von der Entfernung.

Entfernungsbereich abdecken zu können, wurde dazu der beobachtende Roboter in der einen Ecke des Labors platziert, während die zu vermessenden Personen in der diagonal gegenüberliegenden Ecke starteten und dann auf der Diagonalen auf einer annähernd geraden Linie langsam auf den Roboter zu liefen. Die Probanden wurden dabei jeweils mit dem EWA verfolgt. Auf diese Weise konnten die Messungen, die von den Versuchspersonen stammen, von den restlichen Messwerten separiert werden. Für jede dieser Einzelmessungen wurde dann aus den Rohdaten die Entfernung zum beobachtenden Laser berechnet und der zugehörige Remissionswert bestimmt. Dieser Versuch wurde für jedes Material zehnmal wiederholt. Die möglichen Entfernungen wurden nun in Zellen zu je zehn Zentimetern unterteilt und alle Remissionsmessungen der der zugehörigen Entfernung entsprechenden Zelle zugeteilt. Auf diese Weise erhält man für jede Zelle, solange sie nicht zu nahe am oberen oder unteren Ende liegt, einige hundert Remissionswerte. Daraus wurde schließlich für jede Zelle wie beim Lernen der Farbverteilungen der entsprechende Mittelwert und die Standardabweichung berechnet. Diese Werte können dann in Abhängigkeit von der Entfernung in einer sogenannten Look-Up Tabelle zusammengefasst und auch graphisch aufgetragen werden.

Es wurde nun zunächst die Frage untersucht, ob die auf diese Weise bestimmten Eichkurven und damit das Messverhalten eines Laserscanners zeitlich konstant sind. Dazu wurden im Abstand von sechs Tagen für zwei unterschiedliche Materialien dieselben Messungen durchgeführt. Es wurde dabei eine blaue Jeans und eine orange Stoffhose verwendet. Die Graphiken 9.13 und 9.14 zeigen, dass die Sensorkalibrie-

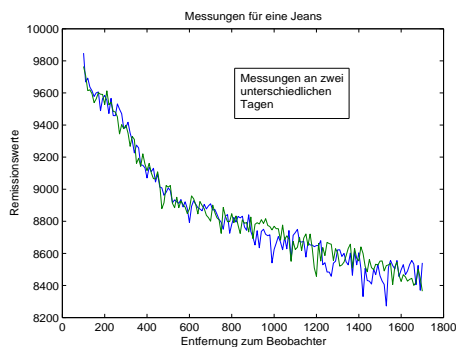


Abbildung 9.13: Messungen für eine blaue Jeanshose.

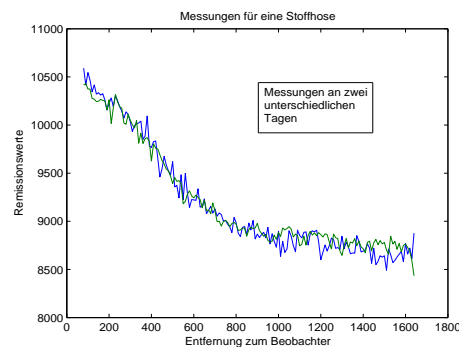


Abbildung 9.14: Messungen für eine orange Stoffhose.

Messverhalten eines Laserscanners an zwei unterschiedlichen Tagen am Beispiel einer Jeans und einer Stoffhose.

ung als zeitlich konstant angesehen werden kann. Darüberhinaus erkennt man bei genauerer Betrachtung der Kurven, dass die Fluktuationen dieser Graphen nicht zeitlich konstant sind. Diese Abweichungen von einer glatten Kurve entsprechen somit un-

ter anderem einem statistischen Rauschen. Daneben muss auch noch die oben schon erwähnte Flächenausdehnung des Lasersignals, wenn es mit der Hose wechselwirkt, in Erwägung gezogen werden. Dies kann, wenn das Signal die Hose in einem Randbereich trifft, dazu führen, dass ein Teil des Strahles gar nicht von der Hose selbst, sondern beispielsweise von der dahinterliegenden Wand reflektiert wird. In diesem Fall erhält man also einen Mischwert für die Remission, der sowohl von den Reflexionseigenschaften der Wand als auch der Hose beeinflusst wird. Auch in diesem Fall ist mit einer Abweichung des gemessenen Wertes von der Idealkurve zu rechnen. Es kann deshalb insgesamt angenommen werden, dass die gelernten Graphen bei Vorliegen von genügend Trainingsdaten glatt werden.

Als nächstes wurde untersucht, inwieweit sich die Kurven für zwei unterschiedliche Laserscanner voneinander unterscheiden. Dazu wurden auch für den anderen Laserscanner des verwendeten Roboters entsprechende Messungen durchgeführt, wieder am Beispiel einer blauen Jeans und einer orangen Stoffhose. Die Graphiken 9.15 und

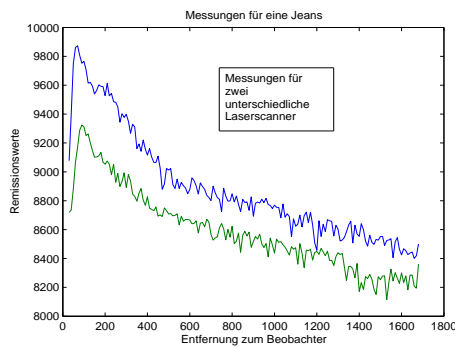


Abbildung 9.15: Messverhalten für eine blaue Jeans.

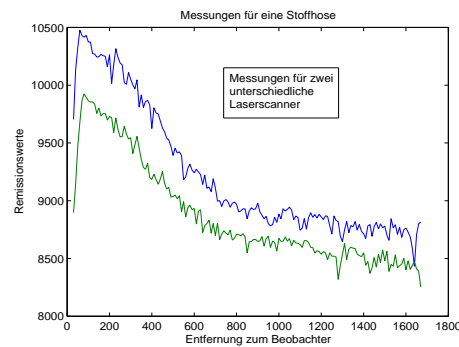


Abbildung 9.16: Messverhalten für eine orange Stoffhose.

Messverhalten für zwei unterschiedliche Laserscanner am Beispiel einer Jeans und einer Stoffhose.

9.16 veranschaulichen, dass unterschiedliche Laserscanner ein deutlich voneinander abweichendes Messverhalten zeigen können. Deshalb müssen die Eichkurven für jeden verwendeten Laserscanner separat gelernt werden.

Als letzte Vorarbeit soll noch untersucht werden, wie groß die Streubreite der Messungen ist. Ein geeignetes Maß hierfür stellt die Standardabweichung dar. Deshalb werden für die schon eingeführte blaue Jeans und orange Stoffhose jeweils in einem Diagramm der Mittelwert, die Standardabweichung addiert zum Mittelwert und die Standardabweichung subtrahiert vom Mittelwert aufgetragen. Die entsprechenden Graphiken 9.17 und 9.18 belegen, dass die Standardabweichung jeweils ungefähr fünf Prozent des Mittelwertes beträgt. Es kann deshalb gehofft werden, dass es Materialien gibt, deren Unterschiede im Mittelwert im Vergleich zu den Standardabweichungen

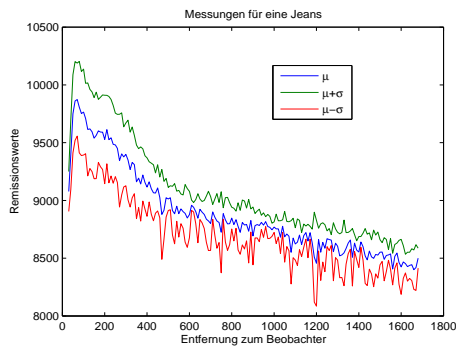


Abbildung 9.17: Darstellung der Standardabweichungen für eine blaue Jeans.

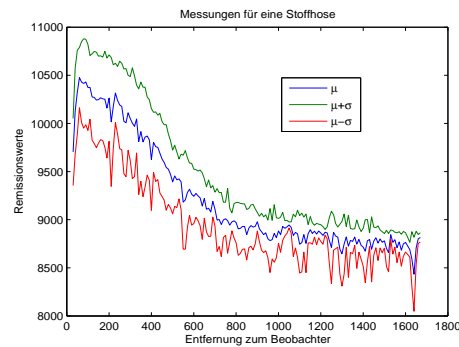


Abbildung 9.18: Darstellung der Standardabweichungen für eine orange Stoffhose.

Darstellung der Standardabweichungen.

groß genug sind, um diese Stoffe anhand ihres Reflexionsverhaltens unterscheiden zu können. Diese Fragestellung soll im Folgenden detailliert untersucht werden.

Zuvor soll aber nicht unerwähnt bleiben, dass die Remission oder Reflektivität von Objekten in den letzten Jahren auch das Interesse anderer Autoren auf sich gezogen hat. In (Andreasson et al., 2005) beispielsweise wurde ein System aus einem Laserscanner und einer Kamera verwendet, um ebene Objekte zu erkennen. Die Kamera und der Laser wurden dabei dadurch aufeinander abgestimmt, dass die von dem Laser gemessene Reflektivität mit der mit der Kamera gewonnenen Helligkeit verglichen wurde. Auch in (Liu und Stamos, 2005) war eine Kalibrierung einer Kamera und eines Lasers nötig. Dort wurden zu diesem Zweck Linien aus Remissions- und Helligkeitsmessungen miteinander verglichen. In (Frintrop et al., 2003) schließlich werden Remissionswerte dazu benutzt, um aus 'Remissionsbildern' wichtige Objekte zu identifizieren. Systematische Untersuchungen zur Abhängigkeit der Reflektivität von der Entfernung oder den Objekteigenschaften finden sich in der angegebenen Literatur allerdings nicht. Es wird lediglich darauf hingewiesen, dass oft eine gewisse Korrelation zwischen den von einer Kamera gemessenen Helligkeitswerten und der Reflektivität besteht (Andreasson et al., 2005; Lamond und Watson, 2004), die Remission andererseits aber auch von den Eigenschaften der reflektierenden Materialien selbst abhängt (Andreasson et al., 2005).

Auch der Hersteller der verwendeten Laserscanner, nämlich die Firma Sick AG, hat natürlich bereits Untersuchungen zum Messverhalten dieser Sensoren durchgeführt. In (Sick AG, 2006) beispielsweise ist eine Abbildung zu finden, die die Abhängigkeit der gemessenen Reflektivität von der Entfernung zum beobachteten Objekt indirekt beschreibt. Diese ist hier mit freundlicher Genehmigung der Firma Sick AG wiedergegeben (Abbildung 9.19). Allerdings zeigt die Graphik zunächst lediglich die Abhängig-

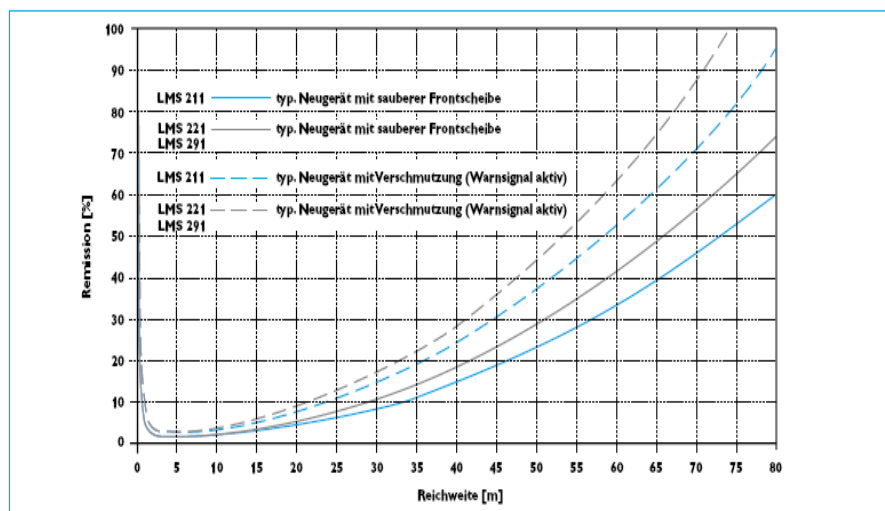


Abbildung 9.19: Erforderliche minimale Remission in Abhängigkeit von der Entfernung. Mit freundlicher Genehmigung der Firma Sick AG.

keit der minimalen Reflektivität eines Objektes, die erforderlich ist, damit vom Laserscanner gerade noch ein Signal gemessen wird, von der Entfernung. Bezeichnet man aber diese Reflektivität mit α_{\min} und mit I_{\min} die minimale Signalstärke, die erforderlich ist, damit der Laserscanner ein Signal noch registriert, so erhält man aus Gleichung 9.18 und dem zugehörigen Modell die Beziehung

$$\alpha_{\min} = \frac{I_{\min}}{I_0} r^2, \quad (9.19)$$

die die Abhängigkeit der minimalen Reflektivität von der Entfernung im Rahmen des oben eingeführten Modells beschreibt. Diese Relation entspricht einer parabelförmigen Abhängigkeit der minimalen Reflektivität von der Entfernung, was in guter Übereinstimmung mit Abbildung 9.19 ist. Das hier vorgestellte Modell und die gemessenen Kurven, wie sie beispielsweise in Abbildung 9.13 oder E.11 dargestellt sind, sind also in guter Kongruenz mit den Angaben des Herstellers der Laserscanner. Auch die Anomalie für sehr kleine Entfernungen, wie sie beispielsweise in den Abbildungen 9.15 oder 9.16 erkennbar ist, findet sich in der Graphik 9.19 der Firma Sick wieder.

In einem ersten Versuch zur Unterscheidbarkeit von Stoffen wurde ein Material verwendet, das ein sehr hohes Reflexionsvermögen aufweist, und zwar eine Folie der Firma 3M mit dem Namen Diamondgrade. Diese Folie zeigt folgendes Reflexionsverhalten: trifft der Laserstrahl senkrecht auf, so wird er nahezu vollständig reflektiert. Paradoxerweise misst ein Laserscanner der Firma Sick in diesem Fall aber lediglich Werte im Bereich von 300, während ein typischer Remissionswert für normale Materialien im Bereich von 9000 liegt. Dieser Umstand ist darin begründet, dass der Messbereich eines solchen Laserscanners nach oben durch den Wert $2^{15} - 1 = 32767$ begrenzt ist und es somit zu einem Überlauf kommt. Der tatsächliche Wert ergibt sich also nach

Addition von 32767 zu dem gemessenen Wert. Im Randbereich, wenn der Laserstrahl nicht mehr senkrecht auf die Folie auftrifft, ist deren Reflexionsvermögen wesentlich kleiner und es werden typischerweise Werte im Bereich von 10000 gemessen, wie sie auch für andere Stoffe üblich sind.

Abbildung 9.20 zeigt die geglätteten Werte einer exemplarischen Messung, wie sie auch in Anhang E.1 ausgeführt wurde. Dabei wurden nur Werte kleiner als 2000 verwendet, d.h. Messungen aus dem Randbereich wurden ausgeschlossen. Das Remissionsverhalten der Folie unterscheidet sich also deutlich vom Remissionsverhalten gewöhnlicher Stoffe.

Es wird nun ein Experiment mit zwei Probanden und einer Interaktion beschrieben, wobei eine Person eine Jeanshose trägt, während die zweite Versuchsperson die Folie vor die Beine hält, so dass der Laserscanner nur Messungen von dieser Folie registriert, nicht aber von der Hose dieser Person. Falls ein Remissionswert unter 6500 lag, wurde zu diesem Wert 32767 addiert, da die Messung dann zweifelsfrei von der Folie stammte. Der minimale Wert, der im Laufe der Untersuchungen für gewöhnliche Materialien gemessen wurde, betrug nämlich 7304. Lag nun ein Messwert, der einer Person zugeordnet worden war, über 11500, so wurde die Wahrscheinlichkeit, dass diese Person die Person mit der Folie ist, auf Eins gesetzt. Der maximale Wert, der für gewöhnliche Materialien gemessen wurde, betrug nämlich 10774. Lag dagegen kein Wert aller derjenigen Messungen, die einer Person zugeordnet worden waren, über 11500, wurde diese Wahrscheinlichkeit auf 0.1 gesetzt. Andererseits wurden die Wahrscheinlichkeiten dafür, dass ein Proband die Person mit der Jeans ist, ähnlich wie bei der Erkennung über Farbhistogramme über eine Gaußverteilung berechnet.

Die zugehörigen Mittelwerte und Standardabweichungen wurden dabei hier und im Folgenden den Trainingsdaten entnommen, wobei jeweils für jede Einzelmessung das Paar aus Entfernung zum Roboter und Remissionswert herangezogen wurde. Die Distanz zum Roboter wurde dabei aus der Positionsschätzung für die jeweilige Person, bestimmt durch den ECSA, und der bekannten Position des Roboters berechnet. Dann wurde diese Entfernung der entsprechenden Zelle aus den Trainingsdaten zugeordnet und der der Zelle entsprechende gelernte Mittelwert zusammen mit der Standardabweichung ausgelesen. Das weitere Vorgehen war dann wie im Falle der Erkennung mit Hilfe von Farbhistogrammen.

Die Abbildungen 9.21 und 9.22 zeigen die Ergebnisse. Die Person mit der Folie befand sich dabei jeweils vor und nach der Interaktion im oberen Bereich des Bildes. Diesem Pro-

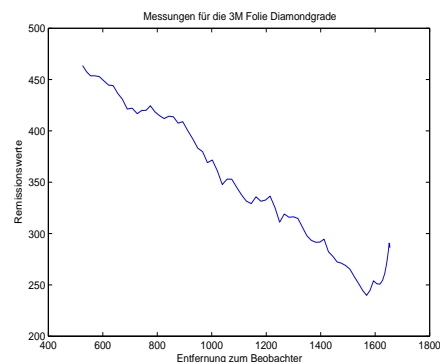


Abbildung 9.20: Registrierte Remissionswerte für eine 3M-Folie vom Typ Diamondgrade.

banden sollte jeweils eine 2 zugeordnet werden. Die Graphiken zeigen, dass die berechneten Identitäten sowohl vor als auch nach der Interaktion richtig sind.

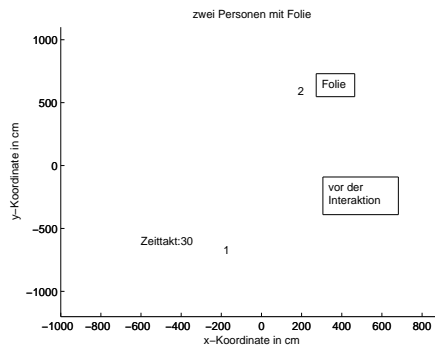


Abbildung 9.21: Begegnung einer Person mit einer normalen Hose und eines Probanden mit einer 3M-Folie, vor der Interaktion.

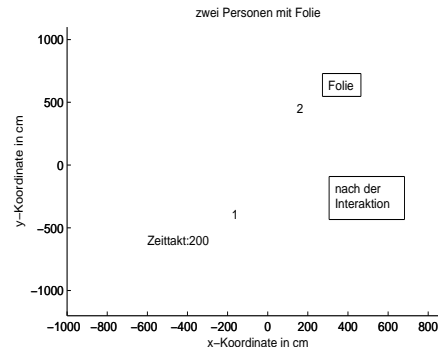


Abbildung 9.22: Begegnung einer Person mit einer normalen Hose und eines Probanden mit einer 3M-Folie, nach der Interaktion.

Interaktion einer Person mit einer normalen Hose und eines Probanden mit einer 3M-Folie vom Typ Diamondgrade.

Das letzte Experiment hat gezeigt, dass es im Prinzip möglich ist, Personen aufgrund der Remissionseigenschaften ihrer Beinbekleidung zu unterscheiden. Es wird nun die Frage untersucht, ob dies auch für gewöhnliche Kleidungsstücke gilt. Dazu werden wieder wie eben eine blaue Jeans und zusätzlich eine orange Stoffhose betrachtet. Die in der Lernphase aufgenommenen Kurven, die die Abhängigkeit der Remissionswerte von der Entfernung zum Laser zeigen, sind in den Abbildungen 9.23 und 9.24 wiedergegeben. Die Kurven der beiden Materialien unterscheiden sich für jeden Laser deutlich voneinander. Während des eigentlichen Experimentes wurden dann wieder die gemessenen Reflektivitätswerte und die Gelernten unter Zugrundelegung einer Normalverteilung miteinander verglichen. Dabei wurden die Messwerte wie auch bei den restlichen Versuchen dieses Kapitels als ungewichtetes Mittel aller derjenigen Messungen bestimmt, die zuvor durch den ECSA mit dem jeweiligen Ziel assoziiert worden waren. Außerdem muss darauf geachtet werden, daß stets auf die gelernten Daten desjenigen Lasers, in dessen Bereich sich die jeweilige Person gerade aufhält, zurückgegriffen wird. Die wahrscheinlichste Zuordnung wird dann schließlich wieder mit dem Algorithmus von Munkres bestimmt.

Es wurden für diese beiden Stoffe insgesamt vier Experimente durchgeführt. Dabei befand sich vor und nach der Interaktion jeweils eine Person im unteren und ein Proband im oberen Bereich der Abbildungen. Die Permutation $12 \mapsto 21$ bezeichnet dabei beispielsweise den Fall, dass sich vor der Interaktion die Person 1 und nach der Interaktion die Person 2 im unteren Bereich aufhielt. Die Abbildungen 9.25 bis

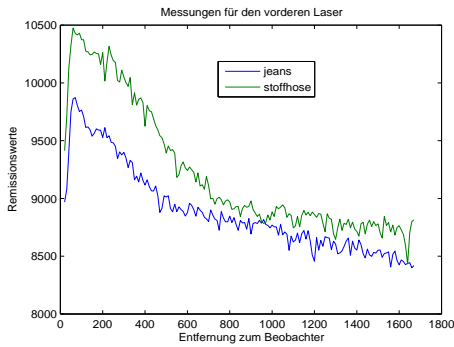


Abbildung 9.23: Remissionswerte für den vorderen Laser.

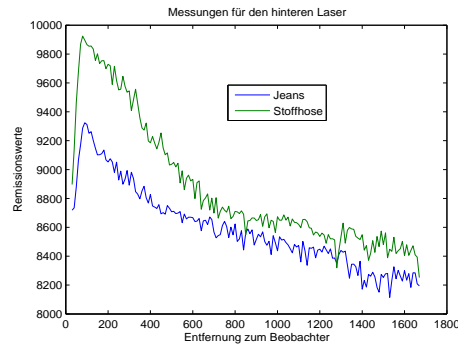


Abbildung 9.24: Remissionswerte für den hinteren Laser.

Remissionswerte für eine blaue Jeans und eine orange Stoffhose.

9.27 und E.18 bis E.23 aus Anhang E.2 zeigen die Versuchsergebnisse. Für die erste Permutation $12 \mapsto 12$ ist in den Abbildungen 9.25 bis 9.27 auch der genaue Versuchverlauf wiedergegeben. Zunächst (Abbildung 9.25) befinden sich die beiden Personen getrennt voneinander im unteren bzw. oberen Bildbereich und werden durch eine Zahl wiedergegeben, die die zugeordnete Identität der Personen anzeigt. Danach findet eine Interaktion statt. Der zugehörige Cluster wird wie in Kapitel 7 als blaues Rechteck abgebildet (Abbildung 9.26). Nach der Interaktion (Abbildung 9.27) bezeichnen wieder die entsprechenden Ziffern die Positionen der beiden Probanden. Die Graphiken für

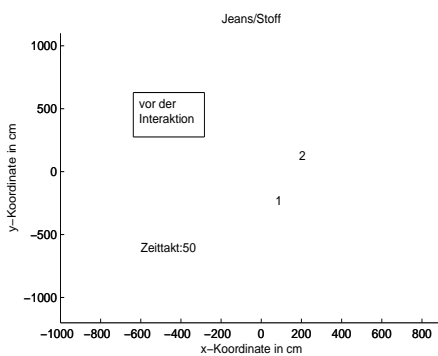


Abbildung 9.25: Das Paar Jeans/Stoff vor der Interaktion.

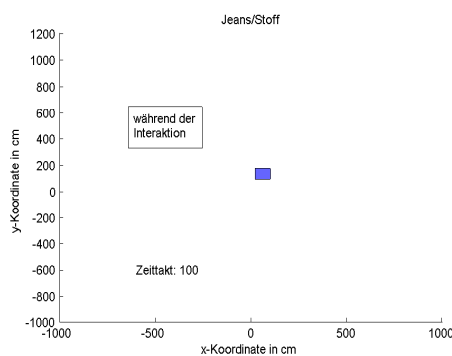


Abbildung 9.26: Das Paar Jeans/Stoff während der Interaktion.

Interaktion einer blauen Jeans und einer orangen Stoffhose, Permutation $12 \mapsto 12$.

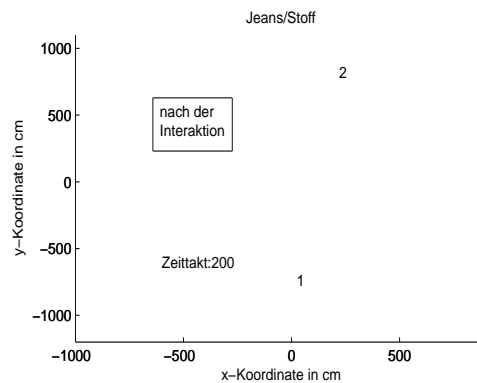


Abbildung 9.27: Das Paar Jeans/Stoff nach der Interaktion, Permutation $12 \mapsto 12$.

die restlichen drei Permutationen finden sich in Anhang E.2 (Abbildungen E.18 bis E.23). Dabei wurden jeweils nur die Ausgangs- und Endpositionen wiedergegeben.

Die Versuche zeigen, dass für alle vier Permutationen die beiden Personen sowohl vor als auch nach der Interaktion durch die Messung des Reflexionsverhaltens ihrer Hosen korrekt identifiziert werden können.

Als nächstes wurde zu der Jeans und der Stoffhose noch ein drittes Material hinzugefügt, nämlich blanke Haut. Die entsprechende Versuchsperson trug dazu eine kurze Hose. Da der Laser etwa auf Kniehöhe misst, rührten die reflektierten Laserstrahlen dann von der Interaktion mit der Haut her. Die Abbildungen 9.28 und 9.29 zeigen das entsprechende Reflexionsverhalten für beide Laser jeweils im Vergleich. Es ist ersichtlich, dass sich Jeans und Haut kaum voneinander unterscheiden. Dennoch ist das Verfahren dazu in der Lage, auch diese drei Materialien sicher zu erkennen. Zum

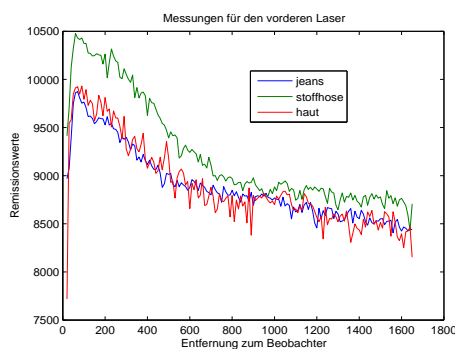


Abbildung 9.28: Remissionswerte für den vorderen Laser.

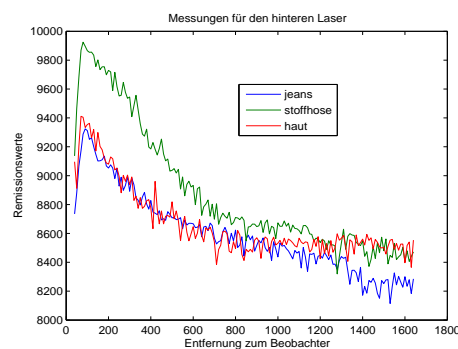


Abbildung 9.29: Remissionswerte für den hinteren Laser.

Remissionswerte für eine blaue Jeans, eine orange Stoffhose und blanke Haut.

Nachweis dieses Sachverhaltes wurden insgesamt sechs Versuche durchgeführt, wobei

die Ausgangsposition der drei Personen immer identisch war, während die Endposition alle sechs möglichen Permutationen durchlief. Position 1 befand sich dabei in den Abbildungen links unten, Position 2 rechts unten und Position 3 oben in der Mitte. Der erste Versuch, in diesem Falle die identische Permutation, wird wieder etwas ausführlicher beschrieben. Zunächst (Abbildung 9.30) bewegen sich alle drei Personen getrennt voneinander. Die Identitäten werden dabei korrekt zugeordnet und in der Graphik als Ziffern dargestellt. In Abbildung 9.31 haben sich zwei Probanden zu einem Cluster vereinigt, der als blaues Rechteck wiedergegeben wird, während die dritte Person als Einzelobjekt wie in Kapitel 7 durch eine grüne Ellipse dargestellt ist. Schließlich laufen alle drei Personen zusammen, d.h. sie bilden einen kombinierten Cluster, der ebenfalls durch ein blaues Rechteck wiedergegeben wird (Abbildung 9.32). Dieser kombinierte Cluster zerfällt zunächst in zwei und dann auch in drei Subcluster, die jeweils durch eine entsprechende Anzahl von blauen Rechtecken dargestellt werden (Abbildung 9.33 und 9.34). Am Ende des Experimentes zerfällt der Cluster wieder in die drei einzelnen Individuen, denen abermals die richtigen Identitäten zugeordnet werden (Abbildung 9.35). Auch für die restlichen fünf Permutationen ist die Zuord-

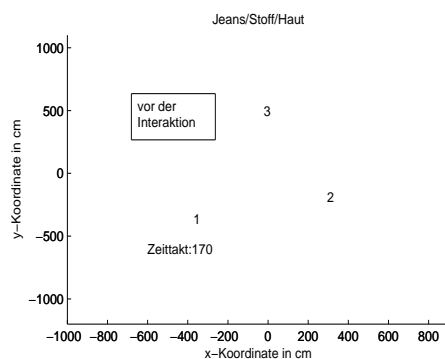


Abbildung 9.30: Die Kombination Jeans/Stoff/Haut vor der Interaktion.

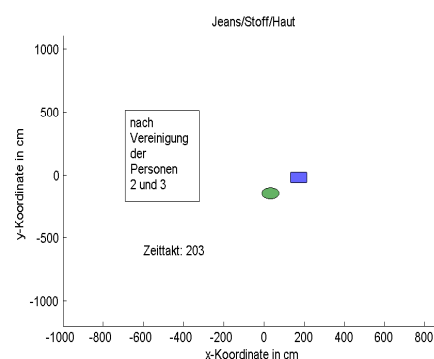


Abbildung 9.31: Nach der Vereinigung von Person 2 (Stoff) und 3 (Haut).

Interaktion einer Jeans, einer Stoffhose und blanker Haut, Permutation

123 \mapsto 123.

nung der Identitäten sowohl vor als auch nach der Interaktion immer korrekt, wie aus Anhang E.2 hervorgeht (Abbildungen E.24 bis E.33). Zusammenfassend kann deshalb festgestellt werden, dass für die betrachtete Kombination dreier Materialien (Jeans, Stoff und Haut) die Identitäten der Probanden sowohl vor als auch nach der Interaktion immer richtig bestimmt wurden.

Im nächsten Schritt soll nun der Fall betrachtet werden, dass sich zwei Materialien sehr ähnlich sind. Deshalb wurde zusätzlich noch eine hellbraune Stoffhose untersucht. Wie die Abbildungen 9.36 und 9.37 zeigen, unterscheidet sich diese in ihrem Remissionsverhalten kaum von der orangen Stoffhose. Es wurden nun zwei Versuche

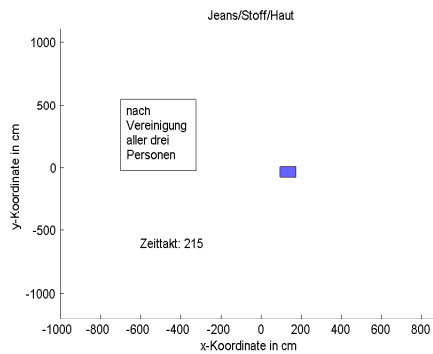


Abbildung 9.32: Nach der Vereinigung aller drei Personen.

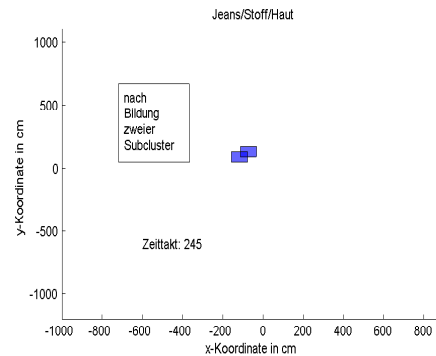


Abbildung 9.33: Nach der Aufteilung in zwei Subcluster.

Interaktion einer Jeans, einer Stoffhose und blanker Haut, Permutation

123 \mapsto 123.

durchgeführt. Die erste Position befand sich dabei jeweils links unten, die zweite Position rechts unten, die dritte Position links oben und die vierte Position rechts oben. Eine 1 entspricht der Jeans, eine 2 der orangen Stoffhose, eine 3 der Haut und eine 4 der hellbraunen Stoffhose. Im ersten Versuch sind die geschätzten Identitäten vor der Interaktion zeitlich nicht konstant. Während die Jeans und die Haut immer korrekt identifiziert werden, werden die beiden Stoffhosen zeitweise richtig erkannt (Abbildung E.35), zeitweise aber auch miteinander vertauscht (Abbildung E.34). Nach der Interaktion werden alle Identitäten zeitlich konstant richtig zugewiesen (Abbildung E.36).

Im zweiten Experiment ist die Zuordnung vor der Interaktion stets korrekt (Abbildung E.37), während sie danach wieder zeitweise richtig (Abbildung E.39) und zeitweise falsch (Abbildung E.38) ist.

Als Ergebnis dieser beiden Versuche kann festgehalten werden, dass die Jeans und die Haut jeweils immer korrekt identifiziert wurden, obwohl diese beiden Materialien sich in ihrem Reflexionsverhalten nur wenig unterscheiden. Die beiden Stoffhosen dagegen werden zu 75 Prozent richtig erkannt, obwohl rein statistisch nur eine Trefferquote von 50 Prozent zu erwarten wäre. Sie werden also noch überzufällig häufig korrekt zugeordnet, obwohl Unterschiede in ihrem Reflexionsverhalten mit bloßem Auge nicht mehr zu erkennen sind (Abbildungen 9.36 und 9.37).

Als nächstes wurden folgende weitere Stoffe auf ihr Reflexionsverhalten hin untersucht:

- schwarze Damenstrümpfe.
- da die Materialien, aus denen gängige Hosen hergestellt werden, begrenzt sind, wurde desweiteren auf Jacken und Mäntel zurückgegriffen. Es handelt sich da-

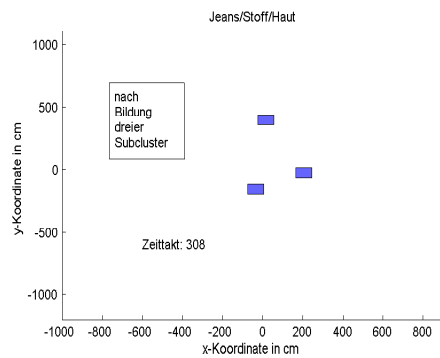


Abbildung 9.34: Nach der Aufteilung in drei Subcluster.

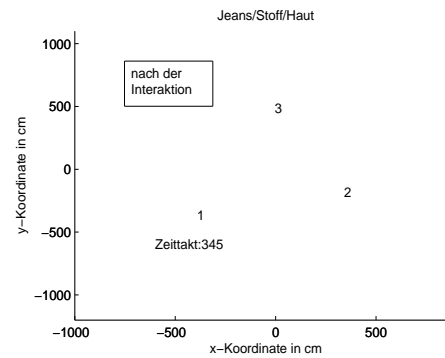


Abbildung 9.35: Die Kombination Jeans/Stoff/Haut nach der Interaktion.

Interaktion einer Jeans, einer Stoffhose und blanker Haut, Permutation

123 \mapsto 123.

bei um Stoffe, aus denen im Prinzip auch Hosen hergestellt werden können bzw. tatsächlich Hosen im Handel erhältlich sind, die aus diesen Materialien hergestellt wurden. Die Oberbekleidungen wurden jeweils von den Versuchspersonen vor ihre Beine in das Sichtfeld der Laserscanner gehalten, so dass die Messungen nur von den Jacken, nicht aber von den jeweiligen Hosen generiert wurden. Es handelte sich dabei im Einzelnen um folgende Materialien:

- ein rotes Leinensakko.
- das Innenfutter dieses Sakkos, das aus Viskose hergestellt worden war.
- eine schwarze Lederjacke.
- eine braune Ziegenvelourslederjacke.
- einen schwarzen Wollmantel.

Die Resultate sind, getrennt für beide Laser, in den Abbildungen 9.38 und 9.39 dargestellt. Dabei wurden auch wieder die Materialien blaue Jeans, orange Stoffhose und blanke Haut wiedergegeben. Es fällt auf, dass zwischen dem Stoff mit dem höchsten Reflexionsvermögen, nämlich der braunen Ziegenvelourslederjacke, und dem Material mit dem niedrigsten Reflexionsvermögen, nämlich der schwarzen Lederjacke, eine absolute Variationsbreite von ca. 1700 Einheiten und eine relative Variationsbreite von ca. 20 Prozent besteht. Es muss auch noch hinzugefügt werden, dass die schwarze Lederjacke auf größere Entfernungen von über zwölf Metern keine Messungen mehr generiert hat. Offensichtlich ist die Intensität des reflektierten Laserstrahls dann so gering, dass sie den Messbereich des verwendeten Laserscanners unterschreitet. Vergleiche hierzu auch Abbildung 9.19.

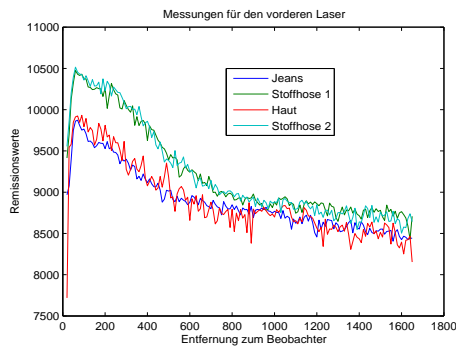


Abbildung 9.36: Remissionswerte für den vorderen Laser.

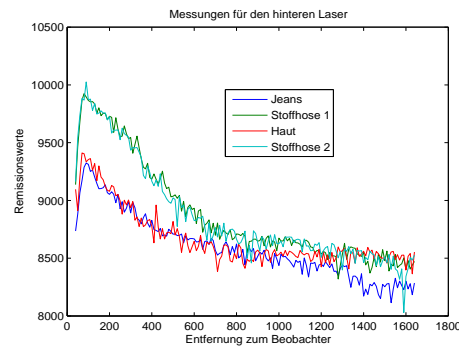


Abbildung 9.37: Remissionswerte für den hinteren Laser.

Remissionswerte für eine blaue Jeans, eine orange Stoffhose, eine hellbraune Stoffhose und blanke Haut.

Als nächstes wurden für die mittlere Gruppe von Stoffen aus Abbildung 9.38 bzw. 9.39 weitergehende Betrachtungen durchgeführt. Es handelte sich dabei um die Materialien blaue Jeans, orange Stoffhose, blanke Haut und schwarze Strümpfe. Das zugehörige Reflexionsverhalten wird in den Abbildungen 9.40 und 9.41 noch einmal zusammengefasst. Es handelt sich also um Materialien, deren Reflexionsverhalten relativ ähnlich ist. Es fällt desweiteren auf, dass der vordere Laser für die Stoffhose ein höheres Reflexionsvermögen misst als für die Strümpfe, während beim hinteren Laser die Verhältnisse gerade umgekehrt sind.

Das angestrebte Ziel der Untersuchungen bestand darin, statistisch fundierte Aussagen treffen zu können über die Wahrscheinlichkeit, mit der diese Materialien voneinander unterschieden werden können. Hierzu wurden insgesamt zwölf Experimente durchgeführt. Bei jedem Versuch bewegten sich die vier Probanden zunächst getrennt voneinander. Danach kamen sie zusammen, so dass eine Interaktion stattfand. Schließlich liefen sie wieder separat. Sowohl vor als auch nach der Interaktion wurde jeweils überprüft, inwieweit die aufgrund des gemessenen Reflexionsverhaltens zugeordneten Identitäten mit den tatsächlichen Identitäten übereinstimmen. Da es sich um zwölf Versuche handelte, die Identitäten jeweils vor und nach der Interaktion überprüft wurden und jeweils vier Personen beteiligt waren, wurde insgesamt $12 \cdot 2 \cdot 4 = 96$ -mal getestet, ob das Verfahren die Identitäten richtig bestimmt.

Die Versuche zeigten zunächst, dass die berechneten Identitäten oft zeitlich nicht konstant waren, sondern vielmehr von den zuerst zugewiesenen Identitäten nach einer kurzen Zeit auf andere Identitäten gewechselt wurde. Es wurden deshalb zur Versuchsauswertung diejenigen Identitäten herangezogen, die sich nach einer gewissen Zeit einstellten und dann auch über einen hinreichend langen Zeitraum beibehalten wurden. Ergebnis der Versuchsreihe war dabei, dass insgesamt 79 Zuordnungen richtig

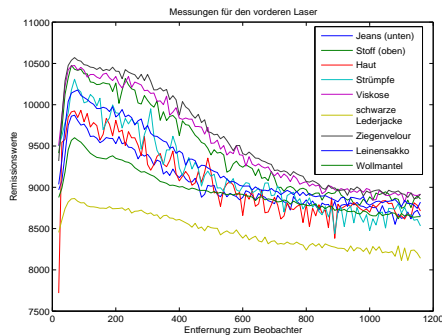


Abbildung 9.38: Remissionswerte für den vorderen Laser.

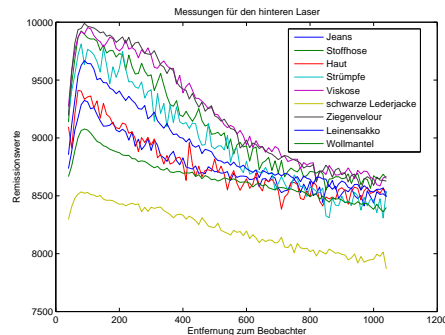


Abbildung 9.39: Remissionswerte für den hinteren Laser.

Remissionswerte für neun verschiedene Materialien.

waren.

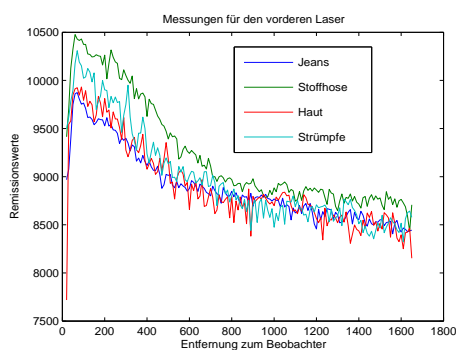


Abbildung 9.40: Remissionswerte für den vorderen Laser.

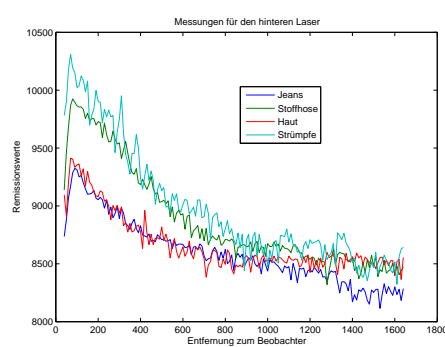


Abbildung 9.41: Remissionswerte für den hinteren Laser.

Remissionswerte für eine blaue Jeans, eine orange Stoffhose, bloße Haut und schwarze Strümpfe.

Dieses Experiment wurde nun mit einem einseitigen Signifikanztest ausgewertet. Dazu soll zunächst kurz beschrieben werden, worum es sich bei einem solchen Test handelt. Als Signifikanztest bezeichnet man einen Hypothesentest, bei dem überprüft wird, ob die Wahrscheinlichkeit dafür, dass eine zu prüfende Hypothese auf Grund einer willkürlichen Entscheidungsregel abgelehnt wird, obwohl sie richtig ist, kleiner oder gleich einer bestimmten Grenzwahrscheinlichkeit, dem so genannten Signifikanzniveau, ist. Man betrachtet dabei ein Zufallsexperiment und interessiert sich für das Eintreten eines bestimmten Ereignisses ω . In unserem Fall besteht das Ereignis ω

darin, dass die zugeteilte Identität richtig ist. Unter der Nullhypothese versteht man nun die Vermutung, dass die Wahrscheinlichkeit $p(\omega)$ für das Eintreten von ω kleiner oder gleich einer bestimmten Schwelle p_0 ist. Es wird dann untersucht, ob diese Nullhypothese tatsächlich zutrifft. Dazu wird das Experiment n -mal durchgeführt und bestimmt, wie oft dabei das Ereignis ω eintritt. Diese Häufigkeit wird dabei mit z bezeichnet. Man unterscheidet nun den Annahme- A und den Ablehnungsbereich \bar{A} . Die Nullhypothese wird angenommen, falls gilt $z \in A$, und abgelehnt, falls $z \in \bar{A}$. Da lediglich Stichproben die Grundlage über Annahme oder Ablehnung der Nullhypothese bilden, besteht natürlich die Möglichkeit einer Fehlentscheidung. Von einem Fehler 1. Art spricht man dabei dann, wenn die Nullhypothese zutreffend ist, aber aufgrund des Versuchsergebnisses abgelehnt werden muss. Die Wahrscheinlichkeit für einen Fehler 1. Art kann berechnet werden. Sie ist abhängig von der Anzahl n der Stichproben, dem gewählten Annahmebereich und der Wahrscheinlichkeit $p(\omega)$ für das Eintreffen des Ereignisses ω . Man beschränkt jetzt die Wahrscheinlichkeit für das Eintreten eines Fehlers 1. Art nach oben und bestimmt den Annahmebereich entsprechend dieser Vorgabe. Diese obere Schranke bezeichnet man auch als Signifikanzniveau. Ist das Signifikanzniveau fünf Prozent und wird die Nullhypothese aufgrund des Versuchsergebnisses verworfen, so spricht man von einem signifikanten Ergebnis, bei einem Signifikanzniveau von einem Prozent analog von einem hoch signifikanten Resultat.

Führt man nun n -mal das gleiche Experiment durch, so spricht man auch von einem Bernoulli-Experiment. Die Wahrscheinlichkeit $B(n, p_0, z)$ dafür, dass das Ereignis ω bei n hintereinander ausgeführten Versuchen genau z -mal auftritt, wenn $p(\omega) = p_0$ gilt, lässt sich dann mit Hilfe der Binomialverteilung berechnen, und es ist

$$B(n, p_0, z) = \binom{n}{z} p_0^z (1 - p_0)^{n-z}. \quad (9.20)$$

In unserem Falle ist $n = 96$ und $z = 79$ und somit $\bar{A} = \{79, \dots, 96\}$. Die Nullhypothese soll sein, dass die Wahrscheinlichkeit für die richtige Zuordnung der Identität kleiner oder gleich $p_0 = 0.7$ ist. Die Wahrscheinlichkeit P für einen Fehler 1. Art ist dann

$$P = \sum_{k \in \bar{A}} B(n, p_0, k) = \sum_{k=79}^{96} B(96, 0.7, k) = 0.0044. \quad (9.21)$$

Die Wahrscheinlichkeit für einen Fehler 1. Art beträgt also 0.44 Prozent. Damit kann die Nullhypothese mit einer Irrtumswahrscheinlichkeit von nur 0.44 Prozent verworfen werden. Es ist also statistisch hoch signifikant, dass die Wahrscheinlichkeit dafür, dass in dem gegebenen experimentellen Setting mit dem vorgestellten Verfahren einer zufällig ausgewählten Person die richtige Identität zugeordnet wird, größer ist als 0.7. Zum Vergleich: da es sich um einen Versuch mit insgesamt vier Probanden handelt, ist die Wahrscheinlichkeit für eine richtige Zuweisung bei einem rein zufälligen Verfahren gleich $\frac{1}{4} = 0.25$. Durch die gewählte Vorgehensweise kann die Wahrscheinlichkeit für die korrekte Zuteilung also nahezu verdreifacht werden, obwohl das Reflexionsverhalten der vier Materialien sich nicht wesentlich voneinander unterscheidet.

In einem zweiten Versuch sollte auf vier Stoffe zurückgegriffen werden, deren Reflexionsverhalten möglichst stark voneinander differiert. Es wurden deshalb die schwarze Lederjacke, die braune Ziegenvelourslederjacke, das rote Leinensakko und der schwarze Wollmantel ausgewählt. Die Abbildungen 9.42 und 9.43 fassen das entsprechende Reflexionsverhalten zusammen. Es fällt schon auf den ersten Blick auf, dass die Unterschiede relativ groß sind, insbesondere für kleinere Entfernungen. Es wurde nun wieder dasselbe Experiment wie oben durchgeführt. Da somit insgesamt zwölf Versuche ausgeführt wurden und die Zuordnung vor und nach jeder Interaktion bestimmt wurde, konnten dabei alle $24 = 4!$ Permutationen der vier Personen ausprobiert werden. Das Ergebnis der Versuchsreihe ist, dass diesmal sogar alle 96 Zuwei-

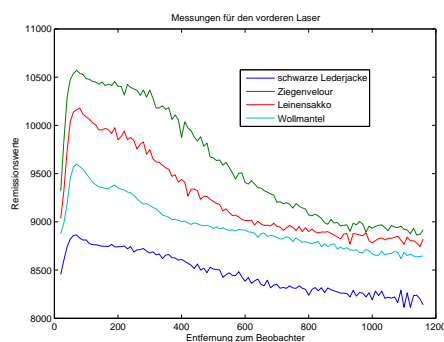


Abbildung 9.42: Remissionswerte für den vorderen Laser.

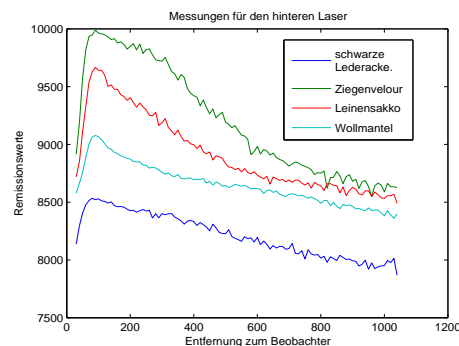


Abbildung 9.43: Remissionswerte für den hinteren Laser.

Remissionswerte für eine schwarze Lederjacke, eine braune Ziegenvelourslederjacke, ein rotes Leinensakko und einen schwarzen Wollmantel.

sungen korrekt waren, und zwar jeweils auch von Anfang an und zeitlich konstant.

Legt man in diesem Falle für den Signifikanztest $p_0 = 0.95$ zugrunde, so kann die Nullhypothese, dass die Identitäten mit einer Wahrscheinlichkeit kleiner oder gleich p_0 richtig zugeteilt werden, mit einer Irrtumswahrscheinlichkeit von 0.73 Prozent abgelehnt werden. In diesem Falle gilt nämlich für den Ablehnungsbereich $\bar{A} = \{96\}$ und die Wahrscheinlichkeit P für den Fehler 1. Art ist somit

$$P = \sum_{k=96}^{96} B(96, 0.95, k) = 0.0073. \quad (9.22)$$

Es ist also in diesem Falle statistisch hoch signifikant, dass eine zufällig ausgewählte Person mit einer Wahrscheinlichkeit von über 95 Prozent richtig identifiziert wird.

Abschließend kann festgestellt werden, dass ein relativer Unterschied im Reflexionsverhalten von nur fünf Prozent bereits ausreicht, um zwei Personen sicher unterscheiden zu können. Da die Remissionswerte üblicher Stoffe einen Bereich zwischen

circa 8500 und 10500 abdecken, ist es deshalb möglich, Hosen aus vier oder fünf verschiedenen Materialien herzustellen, die sicher erkannt werden können. Außerdem gibt es Materialien mit derart charakteristischen Reflexionseigenschaften, wie z.B. die Folie von 3M, dass eine Person, die eine Hose aus einem solchen Stoff trägt, aus einer großen Menge von Personen heraus identifiziert werden kann. Auf diese Weise könnten z.B. in einem Überwachungsszenario Polizisten für Roboter kenntlich gemacht werden.

Kapitel 10

Erweiterungen

10.1 Das Problem der Spuraufnahme und der Präselektion der Messdaten

Die Probleme der Spuraufnahme und der Präselektion der Messdaten sind sehr verwandte Fragestellungen und werden deshalb sinnvollerweise gemeinsam abgehandelt. Das Problem der Spuraufnahme bedeutet dabei, dass sich bewegende Objekte, die neu in das Überwachungsgebiet eines mobilen Roboters geraten, als solche erkannt werden und jeweils ein entsprechender neuer Track initialisiert wird. Die Spuraufnahme (Track Detection) gehört neben der weiteren kontinuierlichen Positionsschätzung (Track Maintenance) bekannter Ziele, d.h. dem eigentlichen Tracking, und der Identifikation der Objekte zu den drei Grundproblemen bei der Überwachung von Gebieten, in denen sich bewegliche Ziele aufhalten. Versucht man dagegen lediglich, Objekte bestimmten Klassen zuzuordnen, so spricht man von Klassifizierung. Mit der Detektierung, dem Tracking und der Klassifizierung von Zielen befasste sich auch der '2nd Workshop on Planning, Perception and Navigation for Intelligent Vehicles' auf der Konferenz 'IEEE/RSJ IROS 2008'. Dies unterstreicht noch einmal die Bedeutung dieser Thematik im Bereich der mobilen Robotik.

Unter einer Präselektion der Messdaten versteht man, dass aus allen Messungen des Laserscanners diejenigen ausgewählt werden sollen, die von sich bewegenden Objekten stammen können. Dieser Ansatz ist aus folgendem Grunde sinnvoll und wichtig. Im Falle einer Interaktion benutzt der ECSA ja den langsameren VBA. Im Rahmen des VBA wird aber, wie auch bei vielen anderen Verfahren, die meiste Rechenzeit auf den Gatingprozess verwendet. In diesem Falle besteht dabei sogar eine lineare Abhängigkeit von der Anzahl der zu verarbeitenden Messwerte. Liegen nämlich zum Beispiel für einen Cluster dreißig Positionsschätzungen vor, was ein üblicher Wert ist, und daneben 360 aktuelle Messungen, wenn die Laser mit einer Winkelauflösung von einem Grad betrieben werden, so muss der Gating-Prozess inklusive der Berechnung der Mahalanobisdistanz ca. zehntausendmal durchlaufen werden. Dies liegt daran, dass für jeden dieser 360 Messwerte und für jedes der ca. dreißig Erwartungsgebiete, die

zu den Positionsschätzungen gehören, überprüft werden muss, ob die Messung in dem jeweiligen Erwartungsgebiet enthalten ist. Vergleiche hierzu auch Kapitel 4.

Durch eine geeignete Vorauswahl der Messdaten lässt sich aber der Anteil derjenigen Messungen, die dafür in Frage kommen, dass sie von beweglichen Objekten und damit auch von dem gerade betrachteten Cluster stammen können, oft auf einige wenige Dutzend einschränken. Der Rechenaufwand, der für den Gating-Prozess erforderlich ist, verringert sich durch dieses Screening deshalb in der Regel um ungefähr eine Zehnerpotenz. Allgemein beruht ein solches Screening auf a-priori-Wissen über die geometrischen Charakteristika des Problems oder auf statistischen Verfahren. Man spricht im Falle einer Hintereinanderschaltung von verschiedenen Screening- oder Gating-Prozeduren häufig auch von Multiple-Gating-Tests (Hall und Llinas, 2001). Im Rahmen dieser Arbeit werden zur Vorauswahl geometrische Tests angewandt.

Im Falle eines stationären Beobachters lassen sich die beiden Probleme sehr leicht lösen. Man speichert einfach für alle Winkel von einem bis 360 Grad die gemessene Entfernung des reflektierenden Objektes beim ersten Scan. Bei jeder neuen Messung vergleicht man dann für jeden Winkel die neu gemessene Entfernung mit der Gespeicherten. Ist die neu gemessene Entfernung kleiner, so bedeutet dies, dass ein sich bewegendes Objekt den Blick auf das ursprünglich beobachtete Ziel versperrt. Die zugehörige Messung muss also dann von einem sich bewegendem Objekt stammen. Allerdings muss hierbei aufgrund der Messungenauigkeiten mit einer geeigneten Sicherheitsschwelle gearbeitet werden, d.h. Messwerte werden nur dann einem beweglichen Objekt zugeschrieben, wenn die neu gemessene Entfernung kleiner ist als die Ursprüngliche abzüglich dieser Sicherheitsschwelle. Eine Sicherheitsschwelle von dreißig Zentimetern hat sich dabei als brauchbar erwiesen. Ist die neu gemessene Entfernung dagegen größer, bedeutet dies, dass das im ersten Scan beobachtete Objekt sich von seinem Platz entfernt hat und dem Roboter den Blick auf weiter entfernte Ziele freigibt. In diesem Falle wird die bisher gespeicherte Entfernung durch die neu Gemessene ersetzt. Dadurch wird der Bereich vergrößert, bezüglich dessen Aussagen über sich bewegende Objekte gemacht werden können. Vergleiche hierzu auch die Ausführungen in Abschnitt 4.3.

Im Falle eines sich bewegenden Beobachters ist die Lösung dieser Probleme ungleich schwieriger, zum Beispiel deshalb, weil sich der Bereich, der von dem Roboter überwacht wird, ständig ändern kann, wenn der Roboter in neue Räume vordringt. Da beispielsweise Menschen in den Lasermessungen bestimmte Muster hervorrufen, können aber folgende Charakteristika ausgenutzt werden, die sich bewegende Objekte in geschlossenen Gebäuden in der mobilen Robotik aufweisen:

1. die Ziele haben meistens einen gewissen Mindestabstand zu den Hindernissen im Hintergrund. In der Regel handelt es sich bei diesen Hindernissen um die Wände der Räume, die überwacht werden sollen. Dies führt dazu, dass zunächst ein starker Abfall und einige Messwerte weiter nach links oder rechts ein starker Anstieg der gemessenen Entfernungen zum Beobachter auftritt (Abbildung 10.1). Dabei muss für diesen Unterschied wieder ein Schwellenwert festge-

legt werden. Ein geeigneter Wert hierfür beträgt zwanzig Zentimeter. Außerdem müssen benachbarte Messungen an diesen Abbruchkanten einen gewissen Mindestabstand voneinander haben. Hierfür wurden als Schwellenwert vierzig Zentimeter zugrunde gelegt. Im ersten Schritt wird also ähnlich wie von zahlreichen anderen Autoren, wie zum Beispiel (Bellotto und Hu, 2006; Schulz et al., 2001b), nach lokalen Minima in den Abstandsmessungen gesucht.

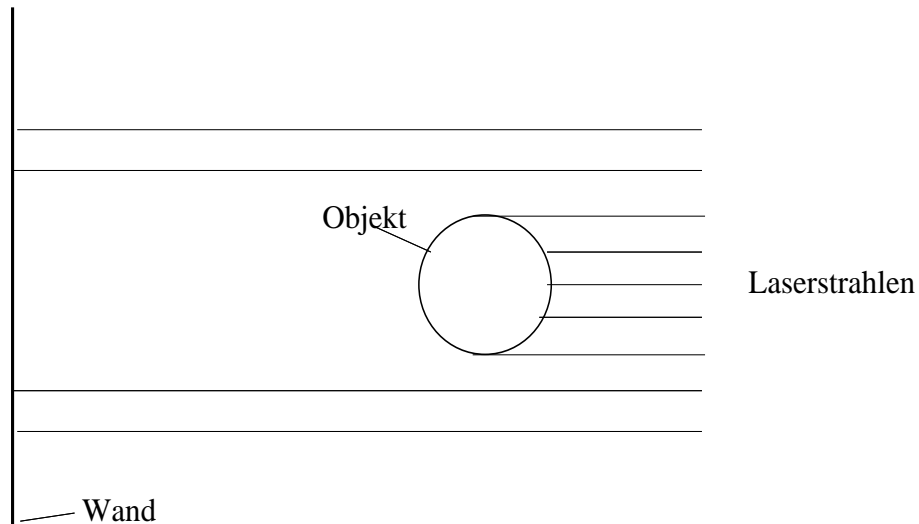


Abbildung 10.1: Abbruchkanten eines Objektes vor einer Wand.

- die räumliche Ausdehnung der beweglichen Objekte (andere mobile Roboter oder Menschen) ist beschränkt. Dadurch wird auch die Anzahl der Messwerte, die von einem Objekt stammen und zwischen dem Ernennungsabfall und dem anschließenden –anstieg liegen können, nach oben begrenzt. Falls man dabei ein kreisförmiges Ziel mit Radius r annimmt und die Entfernung des Beobachters zum Mittelpunkt dieses Objektes gleich d ist, so beträgt der Winkel ϕ , der das beobachtete Objekt einschließt, nach Abbildung 10.2 gerade

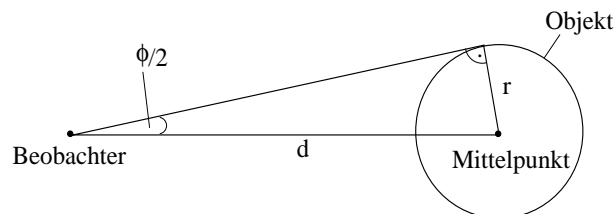


Abbildung 10.2: Berechnung des Winkels ϕ .

$$\phi = 2 \arcsin \frac{r}{d}. \quad (10.1)$$

Aus der Angabe von ϕ in Winkelgraden erhält man dann bei Kenntnis der Winkelauflösung des Sensors sofort eine obere Schranke für die Anzahl der Messungen zwischen dem Abfall und dem Anstieg, die einem beweglichen Ziel zugeordnet werden sollen. Für Menschen ist dabei $r = 50$ Zentimeter ein brauchbarer Wert.

Ein wesentlicher Unterschied zwischen den beiden Problemen muss an dieser Stelle allerdings beachtet werden. Bei der Spuraufnahme genügt es, sich auf Einzelobjekte zu beschränken. Betritt nämlich tatsächlich einmal ein Cluster aus mehreren Personen das Überwachungsgebiet, so werden die Probanden die zweite Bedingung jeweils für sich erfüllen, sobald der Cluster in diese Einzelobjekte zerfallen ist und die Spur der Ziele wird dann aufgenommen. Bei der Präselektion der Daten kann es aber, falls ein Cluster vorliegt, passieren, dass der Cluster so groß ist, dass er die zweite Bedingung nicht mehr erfüllt. Deshalb muss man in diesem Falle stets die maximale Anzahl n der Einzelobjekte bestimmen, die in einem Cluster enthalten sind, und in Gleichung 10.1 r durch $n \cdot r$ ersetzen.

3. die Bedingungen drei und vier betreffen nur die Spuraufnahme. Bei der Spuraufnahme muss nämlich zusätzlich noch der Fall ausgeschlossen werden, dass Messungen, die die ersten beiden Bedingungen erfüllen, von bereits bekannten Zielen wie Einzelobjekten oder Clustern stammen. Zu diesem Zweck wird überprüft, ob zu diesen Objekten jeweils ein gewisser Mindestabstand besteht. Ein geeigneter Schwellenwert ist hierbei 150 Zentimeter. Diese Schwelle ist deshalb so groß gewählt, da hier ja mit der Positionsschätzung des betreffenden Objektes verglichen wird. Da diese Ziele aber insbesondere im Falle eines Clusters eine gewisse Ausdehnung aufweisen, liegen die Messwerte, die zu diesen Objekten gehören, über einen gewissen Bereich verstreut.
4. die ersten drei Bedingungen werden auch von vielen statischen Objekten wie Papierkörben erfüllt. Deshalb wird, sobald diese Bedingungen erfüllt sind, lediglich ein neuer Track angelegt. Das betreffende Ziel wird aber von höheren Funktionen, wie zum Beispiel der Überprüfung auf eine mögliche Interaktion, zunächst noch ausgeschlossen. Erst sobald dieses Objekt einen gewissen Mindestabstand von seiner allerersten Position einnimmt, wird es auch für diese höheren Funktionen zugelassen, da es sich dann um ein bewegliches Ziel handeln muss. Eine geeignete Wahl für den zugehörigen Schwellenwert beträgt wieder 150 Zentimeter.

Mit diesen vier Bedingungen werden in den allermeisten Fällen tatsächlich nur die gewünschten Objekte initialisiert. Ausnahmen bestehen nur dann, wenn der Roboter sich selbst bewegt und das Verfahren zur Eigenlokalisierung nicht immer genau genug ist. Vergleiche hierzu auch die Ausführungen in Kapitel 11.

Neben der Track-Initialisierung gibt es wie schon erwähnt auch noch die Track-Maintenance, zu der neben der Weiterführung bestehender Bahnen auch die Terminie-

rung oder Löschung von Tracks gehört. Im Rahmen dieser Arbeit wird ein Track dann gelöscht, wenn diesem durch den (zweistufigen) Gating-Prozess längere Zeit keine Messungen mehr zugeordnet wurden.

Was die Präselektion oder Vorfilterung der Messdaten anbelangt, so werden die ersten beiden Bedingungen wie folgt ausgenutzt. Zunächst durchlaufen die 360 Messwerte des neuen Scans einen ersten Gating-Prozess. Nur die Messungen, die die ersten beiden Bedingungen erfüllen, werden danach noch weiter betrachtet. Genauer wird nur für die Messwerte, die diese beiden Bedingungen erfüllen, auch überprüft, ob sie in einem geeigneten Erwartungsgebiet liegen. Auf diese Weise lassen sich fast alle Messungen, die von stationären Objekten stammen, in diesem ersten Schritt eliminieren. Diese Messwerte stellen in der Regel den Hauptanteil der 360 Entfernungsmessungen eines Scans dar.

Die ersten beiden Bedingungen entsprechen letzten Endes der Suche nach bestimmten Mustern, wie sie üblicherweise von Personen in Lasermessungen hervorgehoben werden. In abgewandelter Form werden diese auch von anderen Autoren zur Extraktion der Features aus Laserabstandsmessungen verwendet (Bellotto und Hu, 2006; Schulz et al., 2003a).

Es wird nun untersucht, in welchem Umfang der ECSA beschleunigt werden kann, wenn man die Präselektion der Messdaten anwendet. Dazu wurden die schon bekannten fünf Szenarien herangezogen (Kapitel 6.3). In jedem dieser fünf Szenarien liefen zwei Personen in unserer Experimentalhalle umher. Am Anfang gingen sie dabei separat. Nach einer gewissen Zeit trafen sie sich dann, so dass eine Interaktion stattfand. Schließlich trennten sie sich wieder voneinander. Während der Interaktion wurden die beiden Probanden definitionsgemäß mit dem VBA verfolgt. Es wurde nun die Zeit betrachtet, die für die Berechnungen des Algorithmus während der Begegnung mit und ohne eine Vorauswahl der Messdaten benötigt wurde. Tabelle 10.1 fasst die Ergebnisse zusammen. Sie gibt die durchschnittliche Zeit in Millisekunden an, die für die Berechnungen eines Zeittaktes benötigt wurde. Die Algorithmen wurden in MATLAB implementiert und auf einem Pentium IV mit 2.8 GHz ausgeführt.

Tabelle 10.1: Bedarf an Rechenzeit in Millisekunden.

Experiment	1	2	3	4	5
keine Vorauswahl	115.3	111.0	136.5	168.5	120.4
Vorauswahl	42.7	40.3	57.0	69.5	49.0

Die Tabelle zeigt, dass der Bedarf an Rechenzeit um ungefähr siebzig Prozent reduziert werden kann, wenn man die Vorauswahl der Messdaten benutzt. Da für einen Zeittakt dann nicht mehr als hundert Millisekunden benötigt und die Laser mit Frequenzen von ungefähr sechs Hertz betrieben werden, bedeutet dies, dass die Verfahren bei Anwendung der Präselektion der Messdaten in der Regel echtzeitfähig sind. Die Vorauswahl der Messdaten wurde auch noch auf die Experimente angewandt, die in Abschnitt 7.3 diskutiert wurden. Bei dem Versuch mit drei Probanden konnte der

Rechenzeitbedarf pro Zeittakt von 0.2781 Sekunden auf 0.1135 Sekunden reduziert werden. Im Falle des Experimentes mit vier Versuchspersonen dagegen benötigte der Algorithmus ohne Vorauswahl 0.3614 Sekunden und mit Vorauswahl 0.1420 Sekunden pro Zeittakt. Es fällt dabei auf, dass mit steigender Anzahl der Personen auch die benötigte Rechenzeit zunimmt. Dies liegt daran, dass die Anzahl der Messungen, die mit einem Cluster assoziiert werden, in der Regel von der Anzahl der Personen, aus denen der Cluster besteht, abhängt. Natürlich wurden in allen Experimenten auch dann keine Tracks verloren, wenn die Vorauswahl angewandt wurde.

Auch auf die in Abschnitt 6.2 und 6.3 betrachteten Verfahren wurde die Präselektion der Messdaten angewandt. Die Ergebnisse für den Equal Weights Algorithmus sind in Abbildung 10.3 dargestellt. Dabei wird der Rechenzeitbedarf bei Vorfilterung im Verhältnis zum Rechenzeitbedarf ohne Vorfilterung für den gesamten Versuch in Prozent angegeben. Da die Schnelligkeit des EWA kaum von der Anzahl der zu verarbeitenden Messungen abhängt, profitiert dieser Algorithmus kaum von der Präselektion. Der Viterbi-basierte Algorithmus (Abbildung 10.4) und der Cluster-Sorting Algorithmus (Abbildung 10.5), die stets die sehr rechenzeitintensive Form des Gatings anwenden, profitieren dagegen sehr stark von einer Vorfilterung. Der Switching Algorithmus, der sowohl ein einfaches als auch das komplizierte Gating benutzt, profitiert noch relativ gut von einer Vorfilterung, wobei der Gewinn für Szenario Eins, wo die Interaktion den größten Anteil am gesamten Versuch einnimmt, dementsprechend auch am grössten ist (Abbildung 10.6). Der verbesserte Switching Algorithmus, der gegenüber dem Switching Algorithmus während der Interaktion nur noch ungefähr den halben Rechenaufwand betreiben muss (Abbildung 6.29), profitiert dagegen nur noch relativ mäßig (Abbildung 10.7). Allerdings ist der verbesserte Switching Algorithmus mit der Vorfilterung schon schneller als der Equal Weights Algorithmus (Abbildung 10.8), was auch daran liegt, dass der verbesserte Switching Algorithmus während des Crossings nur noch einen Track weiterführt.

Eine geeignete Präselektion der Messdaten kann aber auch die qualitativen Eigenschaften bzw. die Stabilität eines Tracking-Algorithmus verbessern. Eine Schwierigkeit praktisch aller Tracking-Verfahren für People-Tracking in geschlossenen Gebäuden besteht nämlich darin, dass ein Objekt, das zu nahe an eine Wand gerät, schließlich in dieser lokalisiert wird und der zugehörige Track verlorengelht, wenn sich die Person anschließend wieder von der Wand entfernt. Dies liegt daran, dass bei zu großer Annäherung an eine Wand die überwiegende Anzahl der Messungen im Erwartungsgebiet von der Wand stammt und diese Messwerte dann beispielsweise bei der Mittelwertbildung im Rahmen des Equal Weights Algorithmus überwiegen. Im Falle eines statischen Beobachters lässt sich dieses Problem noch dadurch lösen, dass vor dem eigentlichen Versuch die Wände mit einigen Scans vermessen werden und Messungen von diesen Wänden dann während des eigentlichen Versuchs ausgeschlossen werden. Im Falle einer mobilen Roboters ist diese Vorgehensweise dagegen nicht möglich, da sich die Umgebung fortlaufend ändern kann. In diesem Fall kann aber versucht werden, durch eine Vorfilterung die Messwerte von den Wänden auszuschließen, so dass dieses Problem unter Umständen auch hier beherrscht werden kann.

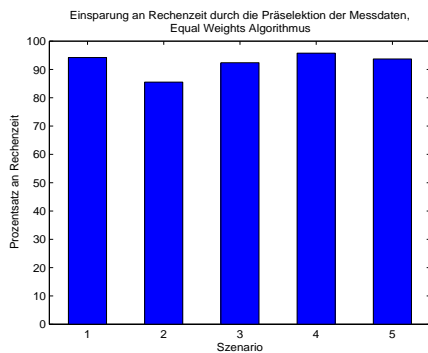


Abbildung 10.3: Prozentsatz an Rechenzeitbedarf bei Verwendung der Präselektion für den Equal Weights Algorithmus.

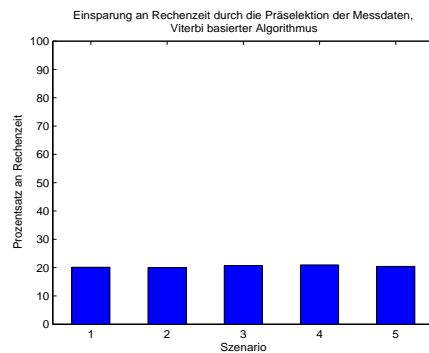


Abbildung 10.4: Prozentsatz an Rechenzeitbedarf bei Verwendung der Präselektion für den Viterbi-basierten Algorithmus.

Zwei interagierende Ziele, Reduktion des Rechenzeitbedarfs des Equal Weights Algorithmus und des Viterbi-basierten Algorithmus bei Präselektion der Messdaten anhand der fünf Szenarien.

Im Falle derjenigen Verfahren, die das komplizierte und aufwendige Gating des Viterbi-Algorithmus benutzen, wie zum Beispiel beim ECSA während eines Crossings, kann eine zusätzliche dramatische Verbesserung erzielt werden, wenn man dieses Gating den Verhältnissen beim People-Tracking anpasst und auf diese Weise optimiert. Dieser Gating-Prozess, wie er in Abschnitt 4.4.2 erläutert wurde, ist im Wesentlichen der Quelle (Pulford und La Scala, 1995) entnommen und ursprünglich für punktförmige Ziele in Clutter entwickelt worden. Er ist deshalb noch nicht an die speziellen Verhältnisse beim People-Tracking adaptiert. Es ist nämlich bei diesem nicht nötig, Messwerte, die bereits als zu der gerade betrachteten Gruppe oder dem Cluster zugehörig eingestuft worden sind, nochmals zu prüfen. Deshalb braucht für eine Messung, die bereits in einem Erwartungsgebiet zu einer alten Positionsschätzung lag, für alle weiteren alten Positionsschätzungen zum Zeitpunkt k nicht mehr unbedingt festgestellt werden, ob sie auch in deren Erwartungsgebieten enthalten ist. Dadurch müssen deutlich weniger Vorhersagen mit Messwerten verglichen werden, so dass gleichfalls deutlich weniger Mahalonisdistanzen berechnet und mit dem Schwellenwert verglichen werden müssen. Außerdem entfällt die Vereinigung der einzelnen Mengen aus den neuen Messungen, die in den jeweiligen Erwartungsgebieten liegen (Schritt 4 auf Seite 50), da ein bestimmter Messwert ja nicht mehr in mehreren dieser Mengen liegen kann. Allerdings müssen dann bei der Bestimmung des Vorgängers gemäß Schritt 5 auf Seite 50 alle alten Positionsschätzungen betrachtet werden, da ja diejenigen Positionsschätzungen, in deren zugehörigem Erwartungsgebiet eine neue Messung liegt, nicht mehr bekannt sind. Dieser geringe zusätzliche Aufwand fällt allerdings im Vergleich zu den Einsparungen kaum ins Gewicht.

Durch diese Vorgehensweise hängt der Rechenaufwand für diese Stufe des Ga-

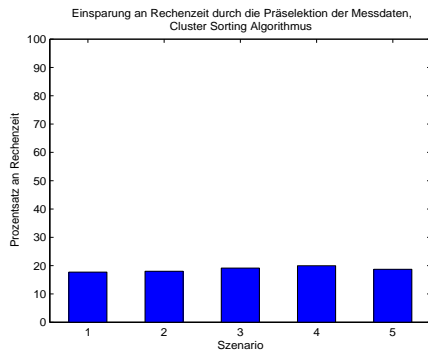


Abbildung 10.5: Prozentsatz an Rechenzeitbedarf bei Verwendung der Präselektion für den Cluster-Sorting Algorithmus.

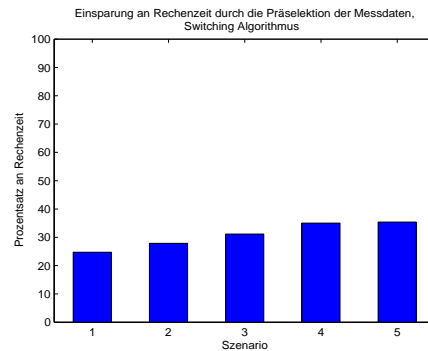


Abbildung 10.6: Prozentsatz an Rechenzeitbedarf bei Verwendung der Präselektion für den Switching Algorithmus.

Zwei interagierende Ziele, Reduktion des Rechenzeitbedarfs des Cluster-Sorting Algorithmus und des Switching Algorithmus bei Präselektion der Messdaten anhand der fünf Szenarien.

tings im Wesentlichen nicht mehr quadratisch (beispielsweise musste bisher für jedes Paar von Vorhersage und neuer Messung die Mahalanobisdistanz berechnet werden), sondern nur noch linear von der Anzahl der Messwerte ab, die durch die Vorauswahl nicht schon verworfen worden sind. Allerdings ist es so, dass für Messungen von statischen Objekten wie Papiereimern, die durch die Präselektion nicht ausgeschlossen werden können, für jede alte Messung immer wieder überprüft werden muss, ob sie im zugehörigen Erwartungsgebiet liegen. Dies rührt daher, dass diese Messwerte in der Regel in keinem der Erwartungsgebiete liegen. Somit hängt der Rechenaufwand noch deutlich vom aktuellen experimentellen Setting ab. Aus diesem Grunde wurde auch auf systematische Untersuchungen zur möglichen Rechenzeiteinsparung durch diese Modifikation verzichtet. Es soll an dieser Stelle lediglich darauf hingewiesen werden, dass sich durch dieses effizientere Gating der Rechenzeitbedarf im Falle der Datensätze, die am Ende von Kapitel 9.3 für die Signifikanztests verwendet wurden, noch einmal um circa neunzig Prozent verringern lässt, so dass das Verfahren damit in jedem Falle echtzeitfähig ist.

Abschließend soll noch einmal ausdrücklich betont werden, dass eine effiziente Methode für das Gating deshalb so wichtig ist, weil das Gating bzw. die Datenassoziiierung während einer Interaktion für einen Großteil des Rechenzeitbedarfs des vorgestellten ECSA verantwortlich ist. Ähnliches gilt auch für viele andere Tracking-Verfahren (Cox, 1993).

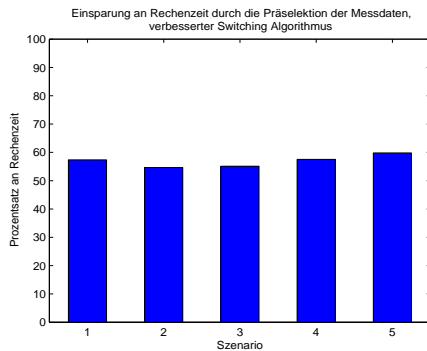


Abbildung 10.7: Prozentsatz an Rechenzeitbedarf bei Verwendung der Präselektion für den verbesserten Switching Algorithmus.

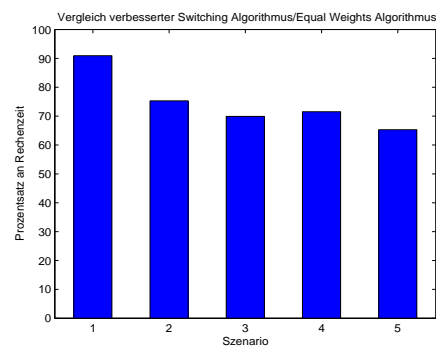


Abbildung 10.8: Verhältnis des Rechenzeitbedarfs des verbesserten Switching Algorithmus im Vergleich zum Equal Weights Algorithmus bei Verwendung der Präselektion der Messdaten.

Zwei interagierende Ziele, Rechenzeitbedarf bei Präselektion der Messdaten anhand der fünf Szenarien.

10.2 Das Problem der Datenfusion bei mehreren beobachtenden Robotern

In diesem Abschnitt soll das Multi-Sensor-Problem, d.h. die Verwendung mehrerer Sensoren, und zwar insbesondere von Laser-Scannern, betrachtet werden. Es wird untersucht, inwieweit die Information, die von den verschiedenen Sensoren bereitgestellt wird, verwendet werden kann, um die Resultate der Tracking-Verfahren zu verbessern. Dabei gibt es qualitative Unterschiede bezüglich der zwei Klassen von Objekten, die in der mobilen Robotik vorkommen können, nämlich andere Roboter oder Menschen. Dies liegt hauptsächlich daran, dass viele mobile Roboter kreisförmig sind und somit ihre Oberfläche eine exakt definierte Form besitzt. Wie in Kapitel 5 gezeigt wurde, kann für solche Ziele nämlich aus den Daten jedes einzelnen Sensors ein wohldefinierter Schwerpunkt mit sehr hoher Genauigkeit geschätzt werden und daraus ebenfalls mit sehr hoher Genauigkeit auf die Position des Mittelpunktes des Roboters geschlossen werden. Demgegenüber liegt bei der Verfolgung von Personen nur wenig Wissen über die aktuelle Form der Objekte, die zudem von Scan zu Scan stark variieren kann, vor. Diese Umstände bewirken, dass jeweils unterschiedliche Ansätze zur Anwendung kommen sollten, wenn Daten von verschiedenen Sensoren fusioniert werden sollen.

Es wird deshalb zunächst der Fall betrachtet, dass ein kreisförmiger Roboter mit bekanntem Radius beobachtet wird. Es kann dann mit den aus Kapitel 5 bekannten Methoden, insbesondere unter Verwendung des EWA, aus den Daten jedes einzelnen Sensors i eine sehr gute Schätzung c_i für den Mittelpunkt dieses Roboters erhalten

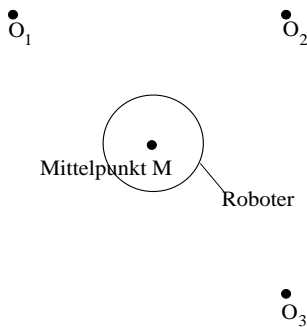


Abbildung 10.9: Die Anordnung der drei Sensoren O_i .

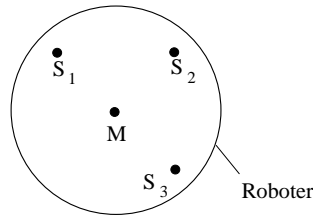


Abbildung 10.10: Darstellung der drei verschiedenen Erwartungswerte S_i .

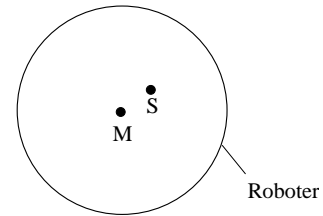


Abbildung 10.11: Die resultierende Schätzung S für die Position des Roboters.

Die Verfolgung eines mobilen Roboters mit drei Sensoren.

werden. Es sollte dann ein ungewichtetes Mittel dieser Schätzungen c_i als Schätzung für den Mittelpunkt des Roboters verwendet werden. Zur Begründung dieser Vorgehensweise wird kurz erläutert, was passieren würde, wenn man die Schätzungen, die man mit dem EWA erhält, direkt fusionieren würde. Dabei wird exemplarisch angenommen, dass der Roboter mittels dreier Sensoren, wie in Abbildung 10.9 dargestellt, beobachtet wird. Dann korrespondieren die drei verschiedenen Schätzungen, die jeweils mit dem EWA berechnet werden, zu drei verschiedenen Erwartungswerten S_1 , S_2 und S_3 , wie sie in Kapitel 5 definiert und in Abbildung 10.10 dargestellt sind. Die Position des Mittels S , das aus diesen drei Schwerpunkten berechnet wird, weicht dann in der in Abbildung 10.11 dargestellten Weise vom wahren Mittelpunkt des Roboters ab. Es macht in diesem Fall auch keinen Sinn, zuerst die Daten aus allen Sensoren zu fusionieren und danach auf dieses Mittel den EWA anzuwenden. Da dieses Mittel dann nicht mehr zu einem einzigen Sensor korrespondiert, würde der EWA nämlich keinen wohldefinierten Erwartungswert mehr berechnen.

Nun wird der Fall der Verfolgung eines Menschen diskutiert. Dabei kann sich die Form des Objektes wie schon erwähnt mit der Zeit sehr stark verändern. Deshalb kann kein Erwartungswert analog zum Fall eines kreisförmigen Zieles berechnet werden. Darüberhinaus weiß man auch nicht, zu welchem Punkt des Objektes die Schätzung der Position, die mit einem Tracking-Algorithmus berechnet wurde, gehört. Da ein Mensch zwei Beine besitzt, kann diese Positionsschätzung sogar zwischen den Beinen, also außerhalb des eigentlichen Objektes liegen. Es erscheint deshalb in diesem Falle am sinnvollsten, ein ungewichtetes Mittel aller Messungen von den verschiedenen Sensoren zu berechnen und dieses Mittel im Rahmen des EWA in der Update-Gleichung des Kalman-Filters zu verwenden. Dabei muss allerdings folgender Faktor berücksichtigt werden. Es kann bei dieser Vorgehensweise nämlich der Fall auftreten, dass derselbe Teil der Oberfläche einer Person von mehreren Sensoren erfasst wird.

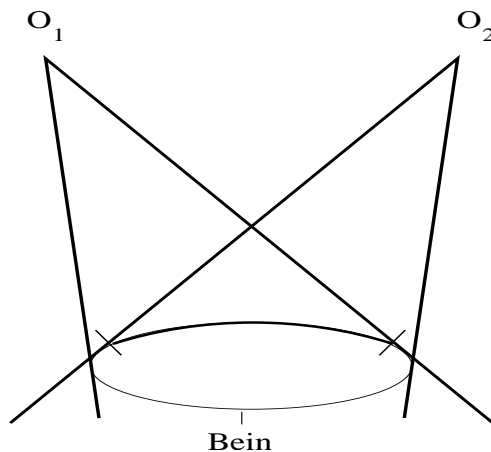


Abbildung 10.12: Ein Teil des Beines wird doppelt erfasst.

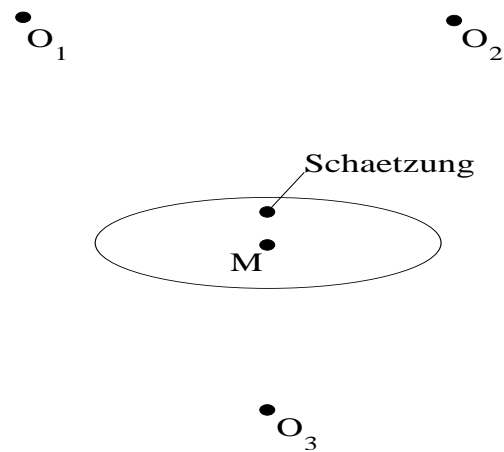


Abbildung 10.13: Abweichung der Positionsschätzung vom wirklichen Mittelpunkt.

Die Beobachtung eines Beines mit mehreren Sensoren.

Dann sollten die zugehörigen Messungen abhängig von der Anzahl dieser Sensoren bei der Berechnung des Mittels weniger stark gewichtet werden. Ansonsten würde die Positionsschätzung, die der EWA liefert, in Richtung dieser Region abdriften. Dieser Sachverhalt wird in den Abbildungen 10.12 und 10.13 erläutert. Dabei wird zunächst der Fall betrachtet, dass ein Bein von zwei Sensoren gleichzeitig beobachtet wird. Für ein solches Bein wird dabei vereinfachend die Form einer Ellipse angenommen (Abbildung 10.12). Ein bestimmter Teil des Beines liegt dabei im Gesichtsfeld beider Sensoren. In Abbildung 10.13 wird dann noch ein dritter Sensor hinzugefügt. Im Endresultat ist die Schätzung für die Position des Beines im Vergleich zum wirklichen Mittelpunkt M in Richtung zu den Beobachtern O_1 und O_2 verschoben, da ein Teil des Beines von diesen beiden Sensoren doppelt erfasst wird.

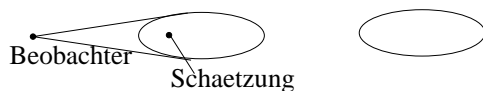


Abbildung 10.14: Beobachtung eines Beines mit einem Sensor.

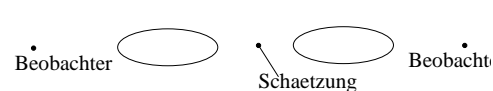


Abbildung 10.15: Beobachtung beider Beine mit zwei Sensoren.

Die Beobachtung der beiden Beine eines Menschen mit einem oder zwei Sensoren.

In den Graphiken 10.14 und 10.15 schließlich wird veranschaulicht, dass es ein

Gewinn sein kann, wenn man einen Menschen mit mehreren Sensoren beobachtet. Die Beine des Menschen werden dabei wieder mit Ellipsen angenähert. In Abbildung 10.14 wird nur ein Sensor benutzt, der nur einen Teil eines Beines der Person abdeckt. Deshalb ist die Positionsschätzung im Innern dieses Beines und weit entfernt von dem anderen Bein. In Abbildung 10.15 dagegen werden zwei Beobachter eingesetzt. In diesem Fall liegt die Positionsschätzung zwischen den beiden Beinen und damit in der Nähe des Schwerpunktes, was wesentlich sinnvoller ist.

Die bisherigen Ausführungen in diesem Abschnitt werden allerdings dadurch wesentlich eingeschränkt, dass die einzelnen Koordinatensysteme der verschiedenen Sensoren nahezu zusammenfallen müssten. Ansonsten würde man durch die Fusion der Daten keinen wesentlichen Vorteil erzielen, da ja der EWA beispielsweise die Position eines kreisförmigen Objektes mit einer Genauigkeit von ungefähr einem Zentimeter schätzen kann. Im Moment lässt sich aber ein Unterschied von ca. zehn Zentimetern in den Koordinatensystemen in der Regel nicht vermeiden.

Auch im Falle der Beobachtung einer Gruppe interagierender Personen mit dem VBA kann es von Vorteil sein, mehrere Sensoren zu verwenden. Bei Benutzung nur eines Sensors kann es nämlich beispielsweise vorkommen, dass sich eine Person im Schatten der anderen Probanden aufhält. Diese Person könnte sich dann in diesem Schatten unbemerkt von dem Sensor von der Gruppe entfernen. Wenn diese Person dann schließlich wieder aus dem Schatten heraustritt, hat sie sich möglicherweise schon soweit von der Gruppe entfernt, dass die zugehörigen Messungen nicht mehr in den Erwartungsgebieten dieser Gruppe liegen. Die Person würde dann nicht mehr der Gruppe zugeordnet werden. Der ECSA könnte dann, wenn sich der Cluster wieder auflöst, niemals diejenige Anzahl an Einzelpersonen detektieren, die die Gruppe ursprünglich gebildet haben. Es wäre somit niemals die erste Bedingung des Endes einer Interaktion erfüllt, wie sie für den ECSA auf Seite 103 definiert worden ist. Deshalb könnte das Ende der Interaktion in diesem Falle niemals festgestellt werden. Andererseits könnte man mit nur drei Sensoren die Gruppe schon von allen Seiten vollständig erfassen, wenn man diese beispielsweise in den Ecken eines gleichseitigen Dreiecks anordnet. Der geschilderte Fall könnte dann nicht mehr auftreten. Man erreicht hier also durch die Verwendung zusätzlicher Sensoren eine bessere räumliche Überdeckung des Überwachungsgebietes. In diesem Falle spielen die geringfügigen Unterschiede in den Koordinatensystemen der einzelnen Roboter auch keine wesentliche Rolle.

Es ist im Übrigen auch möglich, durch Verwendung zusätzlicher Sensoren die Positionsdaten um zusätzliche Dimensionen zu erweitern. In einer Serie von Veröffentlichungen (Bellotto und Hu, 2006; Bellotto und Hu, 2007a; Bellotto und Hu, 2007b; Bellotto und Hu, 2007c) wurde zum Beispiel eine Methode eingeführt, wie mit einer Kamera die Position der Gesichter von Personen bestimmt und so die Positionsmessungen von Laserscannern ergänzt werden können. Weitere Verbesserungen könnten durch ein aktives Sensor-Management erzielt werden, wie zum Beispiel durch Koordination der Roboter, die die Sensoren tragen (Gonzalez-Banos et al., 2002; Jung und Sukhatme, 2004; Lavalle et al., 1997; Murrieta-Cid et al., 2002; Parker, 1997; Pirjanian und Mataric, 2000). Unter Sensor-Management versteht man allgemein die

Optimierung der Ausnutzung der Sensor-Ressourcen. So gab es zum Beispiel auf der Konferenz 'Robotics: Science and Systems 2008' einen Workshop zum Thema 'Topology and Minimalism in Robotics and Sensor Networks', wo auch Anwendungen im Bereich Tracking untersucht wurden. Im weitesten Sinne kann man in einem solchen Falle auch von 'Spatial Computing' sprechen. Davon redet man, wenn eine Ansammlung von Rechnern vorliegt, die über einen physikalischen Raum verteilt sind.

Kapitel 11

Experimente in Echtzeit und die Interaktion eines mobilen Roboters mit einer ausgewählten Person

Um die Leistungsfähigkeit der in dieser Arbeit vorgestellten Verfahren, insbesondere was die Frage ihrer praktischen Anwendbarkeit betrifft, unter Beweis zu stellen, wurden schließlich noch Experimente in Echtzeit durchgeführt. Ein Roboter hatte dabei die Aufgabe, eine Gruppe von Personen zu verfolgen und dabei eine vorher festgelegte Person aus dieser Gruppe zu identifizieren und diesen Probanden zu begleiten. Im Einzelnen wurden dabei folgende, teilweise in dieser Arbeit eingeführte Softwarekomponenten, eingesetzt:

- die Spuraufnahme der Personen erfolgte mit dem in Abschnitt 10.1 vorgestellten Ansatz.
- zur Verfolgung der Probanden wurde der erweiterte Cluster-Sorting Algorithmus (ECSA) aus Abschnitt 7 verwendet. Dieser benutzt die Abstandsmessungen der Laserscanner. Damit wird garantiert, dass die Versuchspersonen ohne Trackverlust verfolgt werden können. Außerdem ist dieser Algorithmus onlinefähig.
- um den Algorithmus insbesondere während einer Interaktion so schnell zu machen, dass er echtzeitfähig ist, wurde in den ECSA zusätzlich das Verfahren zur Präselektion der Messdaten aus Abschnitt 10.1 integriert.
- die Identifizierung der Personen geschah durch Auswertung der Reflexionseigenschaften ihrer Beinbekleidung mit dem in Abschnitt 9.3 eingeführten Verfahren und unter Zuhilfenahme der dort gelernten Eichkurven. Die Reflexionsmessungen wurden dabei ebenfalls mit den Laserscannern ausgeführt.
- die Bestimmung der Eigenbewegung des Roboters erfolgte mit Hilfe der Odometrie, d.h. der Messung der Rotationsbewegung der Räder mit einem entsprechenden Sensor wie in (Bellotto und Hu, 2007b). Damit kann sich der Roboter

selbst lokalisieren, um die Position der Personen stets in einem absoluten Koordinatensystem angeben zu können. Bei dem verwendeten B21 beträgt der akkumulative Fehler in der Lokalisierung lediglich wenige Zentimeter, falls sich der Roboter einige hundert Meter bewegt. Theoretische Überlegungen zur Bestimmung der Eigenbewegung eines mobilen Roboters aus Odometriemessungen finden sich in (Thrun et al., 2005).

- die Steuerung des Roboters zum jeweiligen Zielpunkt mit einem an der Universität Bonn entwickelten Programm (Höller, 2006; Hoeller et al., 2007), das sogenannte Probabilistic Roadmaps (Choset et al., 2005) zur Pfadplanung benutzt.
- die Middleware RoSe (Roboter Service) zur Steuerung des Robotersystems (Tiderko und Bachran, 2007; Tiderko et al., 2007).
- das Framework WNet für das Übertragungsprotokoll zur drahtlosen Kommunikation (wireless LAN) und Datenübertragung zwischen den einzelnen Komponenten (Bachran et al., 2005; Tiderko und Bachran, 2007).

Außerdem wurden folgende Hardwarekomponenten eingesetzt:

- der Versuchsleitstand bestand aus einem Rechner mit vier Opteron AMD-Prozessoren und als Speicher einer RAM mit 4 GB.
- als Roboter wurde ein B21 benutzt. Er war mit einem Rechner mit einem Dual-Prozessor mit 2000 MHz und einer Speicher-Ram mit 2 GB ausgestattet. Die Sensorik bestand aus zwei SICK Laser-Scannern des Typs LMS 200 (Sick AG, 2006) mit jeweils einem Sichtfeld von 180 Grad. Um die gesamte Versuchsumgebung ausmessen zu können, waren diese deshalb Rücken an Rücken auf Kniehöhe an den Roboter montiert. Auf diese Weise konnten die Probleme des Trackings und der Identifikation der Versuchspersonen mit Messdaten von Sensoren nur einer Bauart gleichzeitig gelöst werden. Dies ist möglich, da diese Laserscanner zwei Typen von Informationen liefern, nämlich Abstandsmessungen für das Tracking und Remissionswerte für die Identifizierung.

Jede Anwendung beruht auf mit geeigneten Sensoren beobachtbaren Phänomenen. Diese Sensoren müssen zur weiteren Datenverarbeitung mit Computern kommunizieren. Die Datenverarbeitung ist in diesem Fall hybrid, d.h. es wird sowohl eine autonome Komponente auf dem B21 verwendet, um die Übertragungsrate zu verringern, als auch eine zentrale Komponente auf dem Leitstand, um die Rechenkapazität der onboard-Einheit nicht zu übersteigen. Desweiteren wurden die Algorithmen möglichst optimiert, wie in Kapitel 10.1 dargestellt.

Bezüglich des Punktes, den der Roboter ansteuern soll, müssen zwei Fälle unterschieden werden, nämlich zum einen der Fall, dass Einzelobjekte im Sinne des ECSA vorliegen, und zum anderen der Fall, dass eine Interaktion stattfindet, d.h. sich eine

Personengruppe oder ein Cluster gebildet hat. Falls Einzelobjekte vorliegen und das Verfahren zur Identifizierung die richtige Person erkannt hat, wird als Zielpunkt der Punkt \vec{P} definiert, der auf der Verbindungsgeraden zwischen der Position \vec{p}_r des Roboters und der Position \vec{p}_z der Zielperson liegt und von dieser einen vorher definierten Abstand d hat. d sollte dabei nicht zu klein sein, damit es nicht zu Kollisionen zwischen der Zielperson und dem Roboter kommen kann. In der Praxis hat sich hierfür ein Wert von $d = 50$ Zentimetern bewährt. Dieser Wert wird auch in den folgenden Versuchsreihen zugrunde gelegt. Der Punkt \vec{P} lässt sich mit der Formel

$$\vec{P} = \vec{p}_r + \frac{l-d}{l} (\vec{p}_z - \vec{p}_r) \quad (11.1)$$

berechnen. Dabei ist $l = \|\vec{p}_z - \vec{p}_r\|$ die Länge der Verbindungsstrecke zwischen der Position des Zieles und der Position des Roboters.

Falls eine Interaktion stattfindet und das System somit keine Einzelpersonen mehr verfolgen und auch keine Personenidentifizierung durchführen kann, ist es lediglich möglich, den Roboter auf eine gewisse Entfernung an die Personengruppe heranzusteuern. Dazu wird diejenige Positionsschätzung P_{\min} aus allen durch den ECSA mit der Gruppe oder dem Cluster assoziierten Positionsschätzungen ausgewählt, die zu dem Roboter den kleinsten Abstand hat. Der Zielpunkt ist dann analog zu oben der Punkt auf der Verbindungsgeraden zwischen der Positionsschätzung P_{\min} und der Position des Roboters, der zur Positionsschätzung P_{\min} den Abstand d hat. Durch die Auswahl des Punktes mit dem kleinsten Abstand zum Roboter wird dabei gewährleistet, dass der Roboter mit keiner der Personen aus der Gruppe zusammenstoßen kann. Die Formel zur Berechnung dieses Zielpunktes ist analog zu Gleichung 11.1. Die Ansteuerung des Zielpunktes durch den Roboter erfolgte wie schon erwähnt mit einem Programm, das im Rahmen einer Diplomarbeit an der Universität Bonn entwickelt wurde (Höller, 2006; Hoeller et al., 2007).

Vor jedem Versuch müssen dem System folgende Daten mitgeteilt werden:

- die Anzahl der Probanden. Sobald das System im Rahmen des Spuraufgriffs alle Versuchspersonen detektiert hat, d.h. die Anzahl der Tracks gleich der vorgegebenen Anzahl der Probanden ist, beginnt das System mit der Identifizierung der Personen und der Berechnung des Zielpunktes und lässt den Roboter den Zielpunkt ansteuern.
- die Kennziffern der Eichkurven, mit denen die gemessenen Reflexionswerte verglichen werden sollen. Die Eichkurven sind dabei in einer kleinen Datenbank gespeichert. Der Vergleich von Messdaten mit Daten aus einer Datenbank zur Personenidentifizierung ist ein wohlbekannter Ansatz in der mobilen Robotik (Bellotto und Hu, 2007b).
- die Kennziffer der Person, die der Roboter begleiten soll, aus der Datenbank.
- der oben definierte Abstand d , den der Roboter von der zu begleitenden Person oder einer Gruppe halten soll.

Im Folgenden werden nun exemplarisch drei Experimente mit jeweils zwei Versuchspersonen beschrieben. Die beiden Probanden betraten dabei jeweils zu Beginn der Versuche nacheinander das Überwachungsgebiet des Roboters, so dass es zum Spuraufgriff kam. Sobald das System beide Tracks initialisiert hatte, war es auch möglich, die beiden Personen zu identifizieren und somit zu unterscheiden, so dass der Zielpunkt des Roboters berechnet und der Roboter zu diesem Zielpunkt gesteuert werden konnte. Die Personenidentifizierung und die Robotersteuerung erfolgten dabei jeweils problemlos. Der Rest der Versuche war dann in drei Abschnitte unterteilt. Zunächst bewegten sich die beiden Personen getrennt voneinander und der Roboter folgte der Zielperson. Im mittleren Teil kam es zu einer Begegnung zwischen den beiden Versuchspersonen, so dass der ECSA die Probanden in einem Cluster abbildete und keine Identifizierung mehr möglich war. Der Roboter wurde während dieser Zeit in der oben beschriebenen Art und Weise in die Nähe des Clusters gesteuert. Im dritten Teil des Versuches liefen die beiden Versuchspersonen dann wieder getrennt voneinander, so dass das System die Zielperson erneut identifizieren und begleiten konnte.

Die Idee hinter dieser Versuchsanordnung bestand darin, dass ein Roboter eine vorher definierte Person identifizieren und diese begleiten soll, so dass er diese bei der Durchführung einer Handlung unterstützen kann. Dies kann zum Beispiel durch den Transport einer Last geschehen, die für einen Menschen zu schwer wäre. Deshalb führen Roboter und Zielperson in den Versuchen jeweils gemeinsam eine symbolische Handlung aus. Diese besteht darin, dass die Zielperson zu Beginn des Versuches auf dem Roboter, sobald sich dieser ihr angenähert hat, ein Buch ablegt, das der Roboter dann im mittleren Teil anstelle des Probanden transportiert. So hat die Zielperson in dieser Phase des Versuches beide Hände frei, um beispielsweise mit dem zweiten Probanden gemeinsam weitere Handlungen auszuführen. Am Ende des Versuches, wenn sich die Zielperson wieder alleine bewegt, nimmt diese das Buch wieder von dem Roboter entgegen. Die ersten beiden Versuche wurden mit einer Kamera aufgezeichnet, die an der Decke der Experimentalhalle befestigt war, während der dritte Versuch mit einer handgehaltenen Kamera dokumentiert wurde.

In allen drei Versuchen trug einer der beiden Probanden die orange Stoffhose, die auch schon zur Aufzeichnung der in Abschnitt 9.3 dokumentierten Eichkurven verwendet worden war, während die andere Versuchsperson jeweils eine Hose trug, die dort zwar nicht gelernt worden war, aber einer solchen Hose, was das Material betrifft, ausreichend ähnlich war. In den ersten beiden Versuchen handelte es sich dabei um eine schwarze Lederhose, so dass dem System als Referenzkurven die beiden Eichkurven der schwarzen Lederjacke (für den vorderen und den hinteren Laser) vorgegeben wurden. Im dritten Versuch wurde eine blaue Jeanshose benutzt, so dass als Referenzkurven die beiden Eichkurven ebenfalls einer blauen Jeanshose verwendet wurden. Da der Roboter an Beispielen (eine spezielle blaue Jeans und eine spezielle schwarze Lederjacke) gelernt hatte und daraus auf Eigenschaften unbekannter Objekte (eine andere blaue Jeans und eine schwarze Lederhose) schloss, handelt es sich hierbei auch um Beispiele für maschinelles Lernen, wie schon in der Einleitung erwähnt.

Im ersten Versuch (Abbildungen 11.1 bis 11.12) hatte der Roboter die Aufgabe, der Person mit der schwarzen Lederhose zu folgen. Die beiden Probanden nehmen vor der Begegnung jeweils zwei unterschiedliche Positionen ein (Abbildung 11.1 und 11.3), so dass sich der Roboter der Zielperson zweimal annähert (Abbildung 11.2 und 11.4). Da der Roboter sich wesentlich langsamer als ein Mensch bewegt und außerdem zu Beginn einer neuen Bewegung in der Regel zunächst die Orientierung wechseln muss, dauert es jeweils einige Zeit, bis der Roboter seine Zielposition erreicht hat, so dass die Zielperson immer etwas warten muss, bevor sie eine neue Position einnehmen kann. Dann findet die Begegnung statt (Abbildungen 11.5 bis 11.8), wobei zunächst die Person mit der orangen Hose (Abbildung 11.5) und dann die Person mit der schwarzen Lederhose (Abbildung 11.7) näher zum Roboter steht. Der Roboter nähert sich jeweils der Person, die die kürzere Distanz zu ihm aufweist (Abbildung 11.6 und 11.8). Dann trennen sich die beiden Probanden wieder voneinander (Abbildung 11.9) und der Roboter folgt der Zielperson (Abbildung 11.10). Abschließend ändern die beiden Probanden noch einmal ihre Positionen (Abbildung 11.11) und der Roboter folgt wieder der Zielperson (Abbildung 11.12). Die Ablage des Buches auf dem Roboter erfolgte dabei in Abbildung 11.2 und die Wiederaufnahme in Abbildung 11.12.



Abbildung 11.1: Die beiden Personen haben die Versuchsumgebung betreten.



Abbildung 11.2: Der Roboter hat die Person mit der schwarzen Hose erkannt und sich zu ihr begeben. Die Zielperson legt als Bestandteil einer symbolischen Interaktion mit dem Roboter ein Buch auf dem Roboter ab.

Erstes Experiment.

Im zweiten Versuch (Graphiken 11.13 bis 11.34) hat der Roboter die Aufgabe, der Person mit der orangen Hose zu folgen. Vor der Begegnung nehmen die beiden Probanden insgesamt drei unterschiedliche Positionen ein (Abbildungen 11.13, 11.15 und 11.17), so dass sich der Roboter der Zielperson dreimal annähern muss (Abbildungen



Abbildung 11.3: Die beiden Probanden haben zwei neue Positionen eingenommen.



Abbildung 11.4: Der Roboter hat sich der neuen Position der Person mit der schwarzen Hose angenähert.

Erstes Experiment.

11.14, 11.16 und 11.18). Anschließend kommt es zu einer längeren Begegnung zwischen den beiden Versuchspersonen (Abbildungen 11.19 bis 11.28). Zunächst hat dabei die Person mit der orangefarbenen Hose die kürzere Entfernung zum Roboter (Abbildung 11.19), so dass sich der Roboter dieser annähert (Abbildung 11.20). Danach nimmt der Proband mit der schwarzen Hose die kürzere Entfernung zum Roboter ein (Abbildungen 11.21 und 11.22), so dass der Roboter sich diesem Probanden annähert (Abbildung 11.23). Anschließend nehmen die beiden Probanden zweimal gemeinsam eine neue Position ein (Abbildung 11.24 und 11.27). Der Roboter verfolgt jeweils wieder den Cluster aus den beiden Personen (Abbildung 11.25 und 11.26 bzw. 11.28). Danach trennen sich die beiden Probanden wieder, wobei die Zielperson schließlich noch drei verschiedene Positionen einnimmt (Abbildung 11.29, 11.31 und 11.33). Der Roboter folgt jeweils erfolgreich dieser Zielperson (Abbildung 11.30, 11.32 und 11.34). Die Ablage des Buches auf dem Roboter erfolgte hier in Abbildung 11.18 und die Wiederaufnahme in Abbildung 11.34. In diesem Versuch behielt die Person mit der schwarzen Lederhose, die nicht die Zielperson war, ihre Position in der ersten und dritten Phase stets bei, um zu demonstrieren, dass das System auch dann zurecht kommt, wenn sich einzelne Versuchspersonen zeitweise nicht bewegen.

Im dritten Experiment (Abbildungen 11.35 bis 11.49) hat der Roboter die Aufgabe, der Person mit der blauen Jeans zu folgen. Da die handgehaltene Kamera ein wesentlich kleineres Sichtfeld hat als die Deckenkamera, ist hier, außer während der Begegnung, jeweils nur die Person zu sehen, der der Roboter folgen soll. In Abbildung 11.35 sieht man, wie diese Person die Versuchsumgebung betritt. Vor der Begegnung nimmt die Zielperson zwei verschiedene Positionen ein (Abbildungen 11.36 und 11.38) und der Roboter folgt ihr jeweils (Abbildungen 11.37 und 11.39). In Abbildung 11.39 er-



Abbildung 11.5: Die beiden Personen haben sich einander angenähert, so dass eine Interaktion stattfindet.



Abbildung 11.6: Der Roboter hat die Person mit der orangen Hose aufgesucht, da diese die kürzere Distanz zu ihm aufweist.

Erstes Experiment.

folgte auch die Ablage des Buches. Die Begegnung ist in den Abbildungen 11.40 bis 11.46 dargestellt. Der Proband mit der orangen Hose nimmt dabei zweimal die kürzere Distanz zum Roboter ein (Abbildung 11.41 und 11.43) und der Roboter nähert sich ihr deshalb jeweils (Abbildung 11.42 und 11.44). In Abbildung 11.45 ist auch zu erkennen, wie die beiden Probanden als gemeinsame symbolische Handlung miteinander sprechen. Abschließend wird die Begegnung beendet, indem sich die Zielperson wieder von dem zweiten Probanden entfernt und ihre Endposition einnimmt (Abbildung 11.47 und 11.48). Der Roboter erreicht schließlich erneut diese Zielperson, welche das Buch, das der Roboter während der Begegnung transportiert hat, wieder an sich nimmt (Abbildung 11.49).



Abbildung 11.7: Die Person mit der orangenen Hose hat sich in den Rücken der Person mit der schwarzen Hose begeben, so dass nun Letztere näher zum Roboter steht.



Abbildung 11.8: Der Roboter hat sich nunmehr zu der Person mit der schwarzen Hose begeben.

Erstes Experiment.



Abbildung 11.9: Die beiden Probanden haben sich wieder voneinander getrennt, so dass die Interaktion beendet ist.



Abbildung 11.10: Der Roboter hat sich nun wieder der Person mit der schwarzen Hose angenähert.

Erstes Experiment.



Abbildung 11.11: Die beiden Personen haben noch einmal neue Positionen eingenommen.



Abbildung 11.12: Der Roboter hat sich erneut zu der Person mit der schwarzen Hose bewegt. Diese nimmt das Buch, das der Roboter zwischenzeitlich für sie transportiert hat, wieder an sich.

Erstes Experiment.



Abbildung 11.13: Die beiden Versuchspersonen haben ihre Startpositionen eingenommen.



Abbildung 11.14: Der Roboter hat die Person mit der orangen Hose erkannt und sich zu ihr hinbewegt.

Zweites Experiment.



Abbildung 11.15: Der Proband mit der orangen Hose hat eine neue Position eingenommen.



Abbildung 11.16: Der Roboter hat sich erneut der Person mit der orangen Hose angenähert.

Zweites Experiment.



Abbildung 11.17: Die Versuchsperson mit der orangen Hose hat eine dritte Position aufgesucht.



Abbildung 11.18: Der Roboter hat die Person mit der orangen Hose erneut erreicht. Die Zielperson legt als Teil einer symbolischen interaktiven Handlung mit dem Roboter ein Buch auf dem Roboter ab.

Zweites Experiment.



Abbildung 11.19: Die beiden Versuchspersonen haben sich einander angenähert, so dass eine Interaktion stattfindet.



Abbildung 11.20: Der Roboter hat sich zu der Person mit der orangenen Hose begeben, da diese sich zuvor näher zu ihm befunden hatte.

Zweites Experiment.



Abbildung 11.21: Die Person mit der orangenen Hose hat sich in den Rücken der zweiten Person begeben, so dass nun die Person mit der schwarzen Hose die kürzere Distanz zum Roboter aufweist.



Abbildung 11.22: Beide Personen haben sich noch etwas weiter vom Roboter entfernt.

Zweites Experiment.



Abbildung 11.23: Der Roboter hat sich der Person mit der schwarzen Hose angenähert, da diese nun näher zu ihm benachbart gewesen war.



Abbildung 11.24: Die beiden Versuchspersonen nehmen gemeinsam eine neue Position ein.

Zweites Experiment.



Abbildung 11.25: Die beiden Probanden haben gemeinsam ihre neue Endposition erreicht und der Roboter bewegt sich auf sie zu.



Abbildung 11.26: Der Roboter hat sich der Person mit der schwarzen Hose angenähert, da diese die kürzere Distanz zu ihm aufgewiesen hatte.

Zweites Experiment.



Abbildung 11.27: Die beiden Versuchspersonen haben erneut gemeinsam neue Positionen eingenommen.



Abbildung 11.28: Der Roboter hat sich der Person mit der schwarzen Hose angenähert, da diese sich näher zu ihm befunden hatte.

Zweites Experiment.



Abbildung 11.29: Die beiden Personen haben sich wieder voneinander getrennt, so dass die Interaktion beendet ist.



Abbildung 11.30: Der Roboter hat sich jetzt wieder der Person mit der orangen Hose genähert.

Zweites Experiment.



Abbildung 11.31: Die Person mit der orangenen Hose hat eine neue Position eingenommen.



Abbildung 11.32: Der Roboter hat wieder die Person mit der orangenen Hose erreicht.

Zweites Experiment.



Abbildung 11.33: Die Person mit der orangenen Hose hat letztmals ihren Standort gewechselt.



Abbildung 11.34: Der Roboter hat sich schließlich wieder der Person mit der orangenen Hose genähert. Zum Abschluss der symbolischen interaktiven Handlung mit dem Roboter nimmt die Zielperson das Buch, das der Roboter in der Zwischenzeit für sie transportiert hat, wieder an sich.

Zweites Experiment.



Abbildung 11.35: Die Person mit der Jeans hat die Versuchsumgebung betreten.

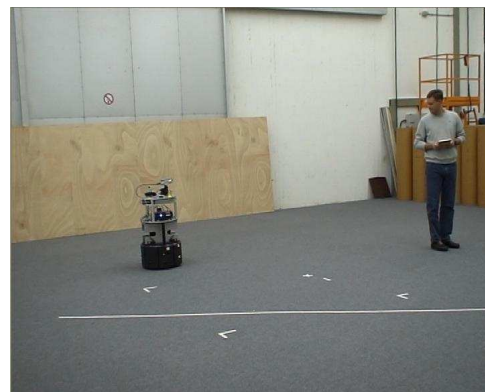


Abbildung 11.36: Die Zielperson hat ihre Startposition eingenommen und wurde vom Roboter bereits erkannt. Der Roboter hat begonnen, sich dieser Person anzunähern.

Drittes Experiment.

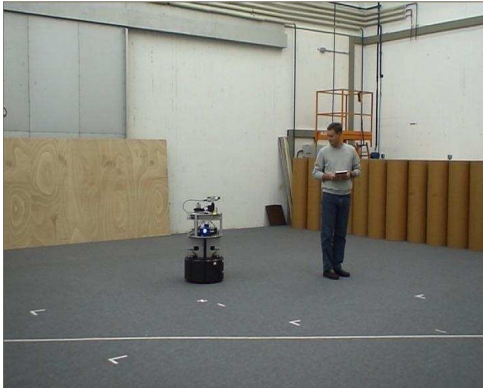


Abbildung 11.37: Der Roboter hat die Person mit der Jeans erreicht.

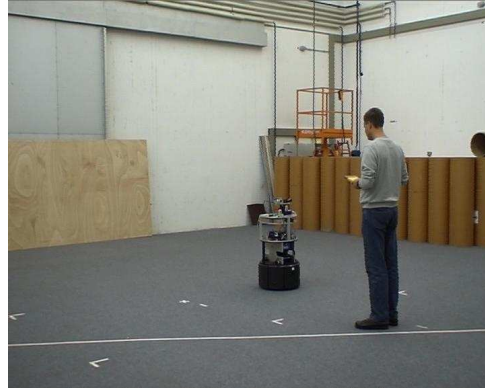


Abbildung 11.38: Die zu begleitende Person hat eine neue Position eingenommen und der Roboter hat begonnen, sich auf diese Position zuzubewegen.

Drittes Experiment.



Abbildung 11.39: Der Roboter hat die neue Zielposition erreicht. Die Versuchsperson legt als Einleitung der symbolischen Interaktion mit dem Roboter wieder ein Buch, das der Roboter transportieren soll, auf dem Roboter ab.



Abbildung 11.40: Die zu begleitende Person bewegt sich in Richtung des zweiten Probanden mit der orangen Hose, der am linken Bildrand gerade noch zu erkennen ist. Der Roboter folgt nach und transportiert dabei das Buch.

Drittes Experiment.

Es wurden auch noch zahlreiche weitere Versuche durchgeführt, die bewiesen, dass das doch sehr komplexe System insgesamt ein recht robustes Verhalten zeigt. Fehler bei den im Rahmen dieser Arbeit vorgestellten Softwarekomponenten traten lediglich



Abbildung 11.41: Die beiden Personen haben sich einander angenähert, so dass eine Interaktion begonnen hat.

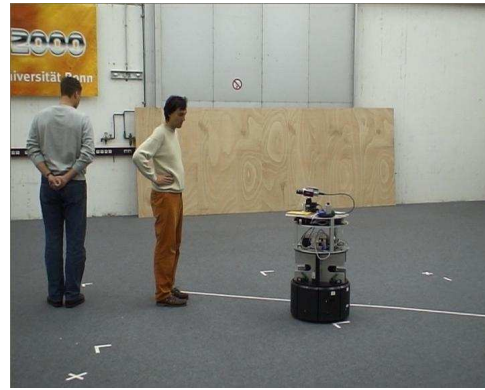


Abbildung 11.42: Da die Person mit der orangen Hose näher zum Roboter steht, sucht der Roboter diese auf.

Drittes Experiment.

manchmal bei der Trackinitialisierung auf. Aufgrund der Eigenbewegung des Roboters und der Fehler in seiner Odometrie kam es nämlich gelegentlich vor, dass der Roboter statische Objekte mit geeigneter Geometrie als sich bewegende Personen identifizierte. Aufgrund der vierten Bedingung der Trackinitialisierung (Kapitel 10.1) sind statische Objekte ja unter normalen Bedingungen nicht schädlich. Dieser Fehler trat aber nur mit einer Häufigkeit von deutlich unter zehn Prozent auf. Ansonsten kam es auch zu Schwierigkeiten bei der drahtlosen Datenübertragung zwischen Leitstand und Roboter oder nach längerer Dauer der Experimente zu Engpässen bei der Onboard-Energieversorgung des Roboters.

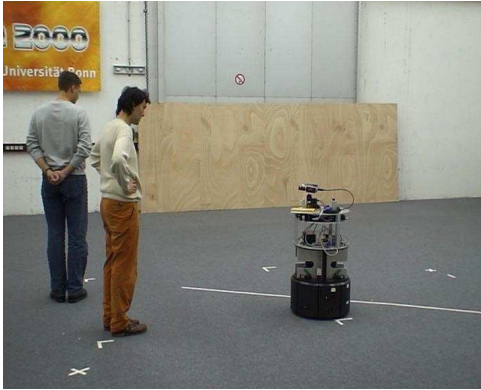


Abbildung 11.43: Die Person mit der orangenen Hose hat ihre Position gewechselt, während die Person mit der Jeans ihre Position beibehalten hat.

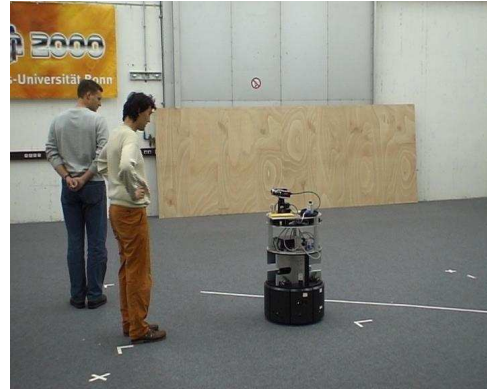


Abbildung 11.44: Da die Person mit der orangenen Hose weiterhin näher zum Roboter steht, begibt sich der Roboter zu ihrer neuen Position.

Drittes Experiment.



Abbildung 11.45: Die beiden Versuchspersonen unterhalten sich kurz miteinander.



Abbildung 11.46: Der Roboter hat sich der neuen Position der beiden Probanden angenähert.

Drittes Experiment.

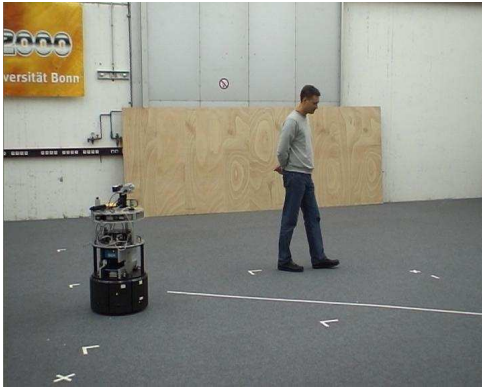


Abbildung 11.47: Der zu begleitende Proband hat sich von der anderen Person gelöst und bewegt sich zu einer neuen Position hin. Der Roboter hat diese Person wieder erkannt und folgt ihr.

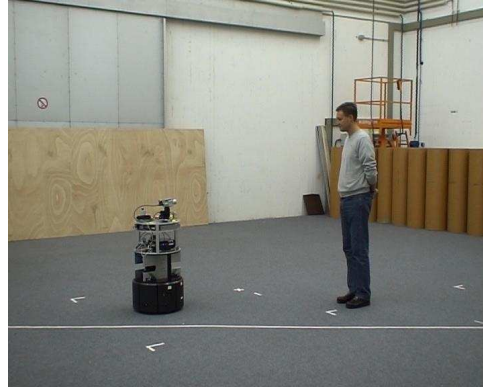


Abbildung 11.48: Die zu begleitende Person hat ihre Zielposition erreicht. Der Roboter bewegt sich weiter auf diese Person zu.

Drittes Experiment.

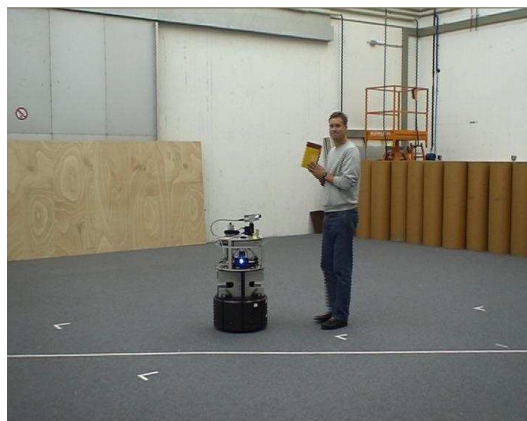


Abbildung 11.49: **Drittes Experiment:** Der Roboter hat die Person mit der Jeans erreicht. Der Proband hat zur Beendigung der symbolischen Interaktion mit dem Roboter das Buch wieder an sich genommen.

Kapitel 12

Zusammenfassung und Schlussfolgerungen

Thema dieser Arbeit war im Wesentlichen die Detektion und die gleichzeitige Verfolgung und Identifizierung von Menschen (People Tracking and Identification) mit mobilen Robotern. Unter Verfolgung oder Tracking versteht man dabei die kontinuierliche Positionsschätzung sich bewogender Objekte, z.B. wie in diesem Falle mit Hilfe von Entfernungsmessungen. Ziel der Untersuchungen war es, einem solchen Vehikel die Begleitung einer ausgewählten Person aus einer Gruppe von Menschen zu ermöglichen. Als Sensoren benutzt der Roboter dabei Laser-Scanner, die die Entfernung zu den reflektierenden Objekten in der Umgebung und zusätzlich noch die Intensität des zurückkommenden Laserstrahles messen. Daneben wurden zur Identifikation auch noch Versuche mit einem Kamerasystem durchgeführt. Als Methoden werden das mathematische Kalkül der Bayesschen Statistik, Sensor-Daten-Fusion, Bildverarbeitung, geometrische Analysen unter Anwendung der ebenen Euklidischen Geometrie und Werkzeuge des maschinellen Lernens verwendet.

In der Einleitung werden das Problem charakterisiert und die Untersuchungen motiviert. Danach folgt eine Darstellung des Stands der Forschung. Hier sind in erster Linie Arbeiten von Dr. Dirk Schulz von der Universität Bonn und einer Arbeitsgruppe aus Essex in Großbritannien, bestehend aus dem Doktoranden Bellotto und dem Betreuer, Professor Hu, hervorzuheben. Insbesondere zeigt sich dabei, dass das Problem des People-Tracking ein vielbeachtetes und grundlegendes Forschungsthema in der mobilen Robotik darstellt. In Kapitel 3 werden die benötigten mathematischen Werkzeuge bereitgestellt. Es handelt sich hierbei vor allem um versteckte lineare Gauß-Markov-Modelle und den Kalman-Filter.

Mit Kapitel 4 beginnen die eigentlichen Arbeiten. Zunächst wird dabei das Modell, das zur Beschreibung der Dynamik der Objekte und der Sensoreigenschaften dient, beschrieben. Die Konvergenzeigenschaften dieses Modells, d.h. das Verhalten nach einer längeren Beobachtungszeit, werden dabei in einem eigenen Anhang A untersucht. Es zeigt sich dabei, dass das System eine sehr schnelle Konvergenz aufweist und es deshalb möglich ist, eine sogenannte Steady-State Lösung anstelle des Kalman-Filters

zu verwenden. Danach werden, basierend auf dem eingeführten Modell, die Grundstruktur und die Kennzeichen eines Algorithmus zur Verfolgung ausgedehnter Objekte wie Menschen dargestellt. Dieses Verfahren ist eine geeignete Modifikation gängiger Ansätze, wie sie zur Verfolgung punktförmiger Ziele verwendet werden. Das wesentliche Kennzeichen der Verfolgung in der mobilen Robotik besteht gerade darin, dass aufgrund des hohen räumlichen Auflösungsvermögens der verwendeten Sensoren mehr als eine Messung zum selben Zeitpunkt von einem Ziel registriert wird. Abschließend werden dann noch einige Möglichkeiten aufgezeigt, wie diese Cluster von Messungen im Rahmen einfacher Tracking-Verfahren verarbeitet und das Problem der sogenannten Datenassoziiierung gelöst werden können. Im Wesentlichen kann man diese Methoden danach unterscheiden, ob sie einen Mittelwert aller Messungen verwenden oder ob für jeden Messwert eine eigene Positionsschätzung berechnet wird.

In Kapitel 5 wird das Problem der Verfolgung eines kreisförmigen Objektes betrachtet. Anhand dieser Fragestellung und aufgrund simulierter Daten werden die im vorhergehenden Kapitel eingeführten Verfahren hinsichtlich der bereitgestellten Information über die tatsächliche Position eines solchen Zieles und ihrer Komplexität evaluiert, da dieses Problem sich dadurch auszeichnet, dass mit Methoden der Integralrechnung ein bestimmter Punkt im Inneren eines solchen Objektes berechnet werden kann, der von den Verfahren mit einer Mittelwertbildung teilweise sehr genau geschätzt wird. Außerdem gibt es eine wichtige Klasse von Zielen in der mobilen Robotik, die eine solche Oberfläche aufweisen, nämlich zahlreiche Service-Roboter. Dieser ausgezeichnete Punkt wird in Anhang B berechnet. Das Verfahren zur Gewinnung der Simulationsdaten wird in Anhang C beschrieben. Anhand dieses Problems werden auch optimale Werte für die Parameter des verwendeten Modells gesucht. Im Weiteren wird diskutiert, wie genau die Positionsschätzungen der einzelnen Ansätze bezogen auf dieses Problem tatsächlich sind. Dabei zeigt sich, dass Verfahren mit Mittelwertbildung generell vorzuziehen sind. Über einen Laufzeitvergleich wird dann die Komplexität der einzelnen Verfahren diskutiert. Die Verhältnisse beim People-Tracking, wo man es mit sehr variablen, sich schnell ändernden Oberflächen zu tun hat, werden kurz angeschnitten. Es stellt sich dabei heraus, dass dieses Problem einer streng analytischen genauen Betrachtung kaum zugänglich ist. Abschließend wird erörtert, wie mit Hilfe der Ergebnisse dieses Kapitels das Problem der Lokalisierung eines mobilen Roboters unter sehr restriktiven Bedingungen, wie sie beispielsweise in einem Kunstmuseum erfüllt sein können, gelöst werden kann. Dabei wird eine Methode präsentiert, wie ein mobiler Roboter in einer solchen Umgebung mit einer Genauigkeit von ungefähr einem Zentimeter zuverlässig lokalisiert werden kann.

In Kapitel 6 wird die Fragestellung zweier interagierender Personen untersucht. Darunter versteht man das häufige Phänomen, dass zwei Menschen sich für eine gewisse Zeit so nahe beieinander bewegen, dass es im Rahmen der Datenassoziiierung zu Auflösungs- und Zuordnungskonflikten kommt. Zunächst wird dabei aufgezeigt, dass die in Kapitel 4 vorgestellten elementaren Ansätze dieses Problem nicht zufriedenstellend lösen können, d.h. eine der beiden Personen verlieren, sobald diese sich wieder trennen. Deshalb werden anschließend drei Verfahren vorgeschlagen, bei denen diese

Schwierigkeiten nicht auftreten. Diese unterscheiden sich lediglich im Rechenzeitbedarf und benutzen zwei der in Kapitel 4 vorgestellten einfacheren Verfahren, insbesondere einen Viterbi-basierten Ansatz. Außerdem verwenden sie Vorwissen über die geometrischen Charakteristika des Problems. Mit diesen Methoden wurden zum ersten Male Verfahren vorgestellt, die das Problem der Begegnung zweier Menschen im Kontext der mobilen Robotik zuverlässig und robust, d.h. ohne Trackverlust, lösen.

In Kapitel 7 wird eine Verallgemeinerung der in Kapitel 6 entwickelten Methoden vorgestellt, die das Problem von Gruppen, die aus einer beliebigen Anzahl von Einzelpersonen bestehen können, so behandelt, dass ebenfalls keine Spuren verloren gehen, wenn sich eine solche Gruppe auflöst. Das Verfahren unterscheidet dabei Einzelpersonen und Personengruppen und lässt auch Übergänge zwischen diesen beiden Objektklassen zu. Unter Anwendung von Vorwissen werden mit Hilfe von Clustering-Methoden Gruppen gegebenenfalls wieder in ihre Einzelobjekte aufgelöst. Anhand zweier Experimente wird nachgewiesen, dass der Ansatz mit allen Situationen und Übergängen, die bei der Verfolgung von Einzelpersonen und Gruppen auftreten können, zurecht kommt. Die Methode ist damit bisher bekannten Verfahren an Robustheit überlegen.

In Kapitel 8 wird das in Kapitel 7 vorgestellte Verfahren, das auch als erweiterter Cluster-Sorting Algorithmus (ECSA) bezeichnet wird, mit einem etablierten Ansatz, und zwar dem SJPDAF von Schulz, bezüglich der Genauigkeit in der Positionsschätzung, dem Rechenzeitbedarf und der Robustheit gegenüber Trackverlust im Falle zweier interagierender Personen verglichen. Die ersten beiden Kriterien werden dabei wieder anhand des exemplarischen Problems der Verfolgung eines kreisförmigen Objektes untersucht. Das Ergebnis dieser Betrachtungen lautet dahingehend, dass der ECSA bezüglich aller dreier Kriterien Vorteile aufweist. Es zeigt sich aber auch, dass der SJPDAF ebenfalls ein sehr gutes Verfahren ist und allen einfachen Verfahren, wie sie in Kapitel 4 vorgestellt worden sind, deutlich überlegen ist.

Nachdem in den vorhergehenden Kapiteln das Problem der zuverlässigen Verfolgung von Einzelpersonen und Personengruppen betrachtet wurde, wird in Kapitel 9 dem zweiten wichtigen Aspekt dieser Arbeit, nämlich der Identifikation von Personen, nachgegangen. Dabei werden zwei Ansätze verfolgt, nämlich die Verwendung von Farbinformation über die Bekleidung der Personen, die mit Hilfe eines Kamerasystems gewonnen wird, und die Analyse der Reflexionswerte der Beinbekleidung. Diese werden direkt durch den Laserscanner zur Verfügung gestellt. Im letzteren Fall werden in einer Lernphase Trainingsdaten gewonnen, die dann in Form von Eichkurven, die die Abhängigkeit von der Entfernung Laser – Person beschreiben, in einer Datenbank gespeichert werden. Während des eigentlichen Versuchs können dann die Versuchspersonen den in der Datenbank gespeicherten Individuen zugeordnet werden. Mit Hilfe von Experimenten und Signifikanztests wurde untersucht, wie zuverlässig diese Vorgehensweise ist. Es zeigte sich dabei, dass ein relativer Unterschied von fünf Prozent in den Remissionswerten bereits ausreicht, um zwei Menschen prompt und zuverlässig unterscheiden zu können. Die Bandbreite der Reflektivitätswerte gängiger Bekleidungen ist dabei außerdem so groß, dass ohne Probleme vier oder fünf Stof-

fe gefunden werden können, die sicher unterschieden werden können. Darüberhinaus gibt es Materialien mit besonders hoher Reflektivität, wie eine Folie der Firma 3M vom Typ Diamondgrade, mit der Personen eindeutig gekennzeichnet werden können.

Im Falle des Kamerasystems können die nötigen Charakteristika auch sehr schnell während des eigentlichen Versuchs gelernt und dann zur richtigen Wiedererkennung der Personen nach einer Interaktion verwendet werden. Hierbei werden die Bilder der Personen im Kamerabild mit einer thermischen Kamera aufgespürt, die Menschen als sogenannte Hotspots wahrnimmt. Auch diese Methode erkennt die Identitäten von Personen zuverlässig, wenn die Bekleidung der Probanden sich in ausreichendem Maße unterscheidet. Allerdings ist hier das beschränkte Sichtfeld eines Kamerasystems zu beachten. Außerdem müssen die Koordinatensysteme der Laserscanner und des Kamerasystems aufeinander abgebildet werden. Einige Graphiken zu Versuchen sind dabei in Anhang D bzw. E.2 wiedergegeben. In Anhang E.1 sind darüber hinaus gewisse Voruntersuchungen zur Personenerkennung anhand des Reflexionsverhaltens ihrer Hosen zusammengefasst. Dabei stellte sich heraus, dass die Reflektivitätswerte sowohl von der Entfernung einer Person vom Sensor als auch von ihrer Beinbekleidung einschließlich Farbe und Oberflächenbeschaffenheit abhängen.

In Kapitel 10 wird das Problem der Spuraufnahme oder Detektion der Ziele behandelt. Dabei werden Methoden vorgestellt, mit denen ein Erkennen dynamischer Objekte und insbesondere eine Unterscheidung von statischen Objekten wie Papierkörben in den meisten Fällen zuverlässig gelingt. Die selben geometrischen Überlegungen, die zur Lösung dieses Problems benutzt werden, werden auch noch dazu verwendet, die Verfahren durch eine geeignete Vorauswahl der Messdaten deutlich zu beschleunigen. Damit werden die Methoden insbesondere echtzeitfähig, d.h. die notwendigen Berechnungen können mit den begrenzten Rechenkapazitäten einer Onboard-Einheit während eines realen Experimentes durchgeführt werden. Daneben werden in diesem Kapitel auch noch kurz geeignete Herangehensweisen zur Datenfusion bei Verwendung von mehreren beobachtenden Robotern vorgeschlagen. Dabei zeigt sich, dass wieder zwischen zwei Klassen von Objekten, nämlich Menschen und mobilen Robotern, unterschieden werden muss.

In Kapitel 11 werden schließlich Experimente in Echtzeit beschrieben. Ein mobiler Roboter hatte dabei die Aufgabe, zwei Versuchspersonen zu detektieren, kontinuierlich deren Positionen zu schätzen, auch wenn sie eine Gruppe bildeten, und die Probanden zu identifizieren. Dazu standen ihm Entfernungs- und Reflexionsmessungen der Laserscanner und die Eichkurven aus der Datenbank zur Verfügung. Zusätzlich hatte die mobile Plattform die Aufgabe, eine vorher ausgewählte Person kontinuierlich zu begleiten und für diese als symbolische interaktive Handlung während des Versuchs ein Buch zu transportieren. Zur Durchführung dieser Aufgaben musste der Roboter mittels Odometrie auch noch kontinuierlich seine eigene Position schätzen, d.h. sich selbst lokalisieren. Die Versuchsreihe zeigte, dass die mobile Einheit mit Hilfe der in dieser Arbeit vorgestellten Verfahren in den meisten Fällen zuverlässig dazu in der Lage ist, die gestellte Aufgabe zu lösen. Probleme traten lediglich dann auf, wenn zum Beispiel die drahtlose Übertragung zwischen dem Leitstand und der mobilen Einheit Fehler

aufweist oder die Energieversorgung des Roboters Engpässe zeigt.

Die autonome Identifikation und Begleitung einer Person aus einer Personengruppe stellt eine Schlüsselfähigkeit einer mobilen Einheit dar, um es dieser zu ermöglichen, mit einer Zielperson zuverlässig und stabil zu kooperieren und gemeinsame, arbeitsteilige Handlungen auszuführen (Bellotto und Hu, 2007a; Tapus et al., 2007). Mit der vorliegenden Arbeit konnte dieses Problem, das als eines der Schlüsselprobleme in der mobilen Robotik angesehen wird, nach bestem Wissen des Autors zum ersten Mal zuverlässig gelöst werden.

Als wesentliche Neuerungen dieser Arbeit sollen noch einmal die folgenden Punkte zusammengefasst werden:

1. die genaue kontinuierliche Positionsschätzung eines kreisförmigen Objektes. Mit Hilfe eines stationären Sensornetzwerkes im Sinne eines sogenannten intelligenten Gebäudes kann dadurch die Position eines geeigneten mobilen Roboters mit einer Genauigkeit im Bereich von ungefähr einem Zentimeter oder besser bestimmt werden.
2. die Verfolgung von Personengruppen, ohne dass Spuren von einzelnen Personen verloren gehen, wenn sich eine solche Gruppe wieder auflöst.
3. die Identifizierung von Menschen mit Hilfe von Laserintensitätsmessungen.

Als Resultat der letzten beiden Punkte ist es insbesondere möglich geworden, eine vorher ausgewählte Person von einem mobilen Roboter begleiten zu lassen. Insgesamt lässt sich deshalb feststellen, dass die in dieser Arbeit vorgestellten Verfahren und Methoden den Stand der Forschung im Bereich des People-Tracking mit Lasern in der mobilen Robotik signifikant verbessern.

Mit der vorliegenden Arbeit scheint das Gebiet der Verfolgung von Personen mit zweidimensionalen Laserbildern in geschlossenen Gebäuden weitgehend erschlossen zu sein. Mögliche Erweiterungen, die in der Zukunft untersucht werden könnten, betreffen zum einen Anwendungen im Outdoor-Bereich, wie zum Beispiel im Straßenverkehr (Broggi, 2008; Ngako Pangop et al., 2008), oder die Verwendung von dreidimensionalen Laserscans. Die Outdoor-Robotik und die 3D-Scans sind zur Zeit Gegenstand intensiver weltweiter Forschungsbemühungen (Hoeller et al., 2008; Nüchter et al., 2005; Nüchter et al., 2007; Röhling, 2008; Röhling und Schulz, 2008). Hier könnte beispielsweise geprüft werden, inwieweit bzw. gegebenenfalls unter welchen Modifikationen die in dieser Arbeit vorgestellten Methoden auf diese Problemfelder übertragen werden können. People-Tracking im Outdoor-Bereich wird beispielsweise dadurch gekennzeichnet, dass hier vor allem in spärlich bewachsenem Gelände die Mehrzahl der Laserstrahlen nicht reflektiert wird. In dichten Wäldern dagegen könnte es, insbesondere wenn der Roboter sich bewegt, schwierig sein, Personen von Bäumen zu unterscheiden. In offenem Gelände könnte auch versucht werden, mit sogenannten UAV's (Unmanned Aerial Vehicle, fliegende Roboter) zu operieren, die dann mit 3D-Scans betrieben werden müssten. Im Falle von 3D-Scans liegen ungleich mehr Daten

vor im Vergleich zu zweidimensionalen Laserscans. Dies macht eine effiziente Lösung des Assoziationsproblems noch viel wichtiger.

Anhang A

Konvergenzeigenschaften des Modells

Die Folge $\{K(k)\}_{k=1}^{\infty}$ der Kalman-Gains konvergiert für das verwendete Modell für ausgedehnte Ziele sehr schnell. Um genauer zu sein, ist dies nicht allgemeingültig. Dies beruht auf der Tatsache, dass die Einträge der Kovarianz R des Messrauschens von den Polarkoordinaten der Vorhersage für die Position des Objektes abhängen. Da sich diese Koordinaten aber während der Bewegung des Zieles ändern, ist die Kovarianz R während des Tracking-Vorganges nicht zeitlich konstant und verhindert damit eigentlich die Konvergenz der Folge der Kalman-Gains. Tatsächlich ist es aber so, dass dieser Nachteil des Modells dadurch geheilt wird, dass die Einträge der Matrix E , die die Ausdehnung der Objekte beschreibt, viel größer sind als die Einträge der Matrix R , so dass die Matrix R in Gleichung 4.13 wie schon im Hauptteil erläutert vernachlässigt werden kann. Genau genommen ist es so, dass die Berücksichtigung der Matrix R lediglich zu kleinen Fluktuationen in den Einträgen der Kalman-Gains während des Tracking-Vorganges führen würde.

Die Konvergenz der Kalman-Gains $K(k)$ und auch der Kovarianzen ist für viele Modelle, die im Bereich des Trackings benutzt werden, bekannt (Ekstrand, 1985; Kalata, 1984; Painter et al., 1990). Da die Einträge der Kalman-Gain in diesen Fällen durch die Konstanten α und β bzw. in dem Falle, dass auch die Beschleunigung modelliert wird, zusätzlich durch eine Konstante γ repräsentiert werden, spricht man dabei oft auch von α - β - bzw. α - β - γ -Trackern. Diese Ansätze waren sogar schon in den 50-er Jahren des vergangenen Jahrhunderts bekannt, also noch vor Einführung des Kalman-Filters, der ja aus dem Jahre 1960 stammt (Kalman, 1960). Deshalb fehlte diesen Ansätzen damals auch noch der theoretische Unterbau.

Die Konstanten werden oft a-priori ausgewählt und treffen eine Gewichtung zwischen der Vorhersage und der aktuellen Messung. Sie bilden die Dynamik des Zieles und die Eigenschaften des Sensors ab und bedeuten einen Kompromiss zwischen der Berücksichtigung des Messrauschens und des Prozessrauschens und der Beschleunigungen. Da keine Matrizen berechnet werden müssen, sind α - β -Filter einfach zu implementieren und haben einen geringen Rechenbedarf.

Im Folgenden wird die Konvergenz mit Methoden der Kontrolltheorie mathema-

tisch schlüssig bewiesen. Dabei wird folgendes Modell zugrundegelegt:

$$x_k = Ax_{k-1} + Gw \quad (\text{A.1})$$

und

$$z_k = Bx_k + v \quad (\text{A.2})$$

mit der Kovarianz Q des Prozessrauschens w und der Kovarianz E des Messprozesses v . Dabei ist E eine positiv definite Diagonalmatrix, die die Ausdehnung der Objekte modelliert. Die Kovarianz R , die das Messrauschen beschreibt, wird also, wie schon besprochen, wieder vernachlässigt.

Eine Eigenschaft des verwendeten Modells besteht darin, dass die Bewegungen in den x - und y -Koordinaten unabhängig voneinander sind. Deshalb genügt es, die Verhältnisse in einer Dimension zu betrachten. Dabei sind die Matrizen in den eben eingeführten Gleichungen

$$A = \begin{pmatrix} 1 & \Delta T \\ 0 & \rho \end{pmatrix} \quad (\text{A.3})$$

mit

$$\rho = \exp(-\Delta T/\Theta), \quad (\text{A.4})$$

$$G = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (\text{A.5})$$

$$Q = \begin{pmatrix} 0 & 0 \\ 0 & \mu \end{pmatrix} \quad (\text{A.6})$$

mit

$$\mu = \Sigma^2 (1 - \exp(-2\Delta T/\Theta)), \quad (\text{A.7})$$

$$B = \begin{pmatrix} 1 & 0 \end{pmatrix} \quad (\text{A.8})$$

und

$$E = \delta. \quad (\text{A.9})$$

Für die Kovarianz der Vorhersage $P(k+1|k)$ ist nun die Rekursionsgleichung

$$P(k+1|k) = AP(k|k)A^\top + GQG^\top \quad (\text{A.10})$$

und für die Kovarianz $P(k+1|k+1)$ die Rekursionsgleichung

$$P(k+1|k+1) = P(k+1|k) - P(k+1|k)B^\top [S(k+1)]^{-1}BP(k+1|k) \quad (\text{A.11})$$

wohlbekannt.¹ Dabei kann die Innovationskovarianz $S(k+1)$ berechnet werden nach der Formel

$$S(k+1) = BP(k+1|k)B^\top + E. \quad (\text{A.12})$$

¹Für eine Herleitung der hier verwendeten Matrixgleichungen des Kalman-Filters vergleiche (Kräußling, 2007a).

Die Kombination dieser drei Gleichungen ergibt die folgende Rekursionsformel für die Kovarianz der Vorhersage:

$$P(k+1|k) = A \left[P(k|k-1) - P(k|k-1)B^\top [BP(k|k-1)B^\top + E]^{-1} \cdot \right. \\ \left. \cdot BP(k|k-1) \right] A^\top + GQG^\top. \quad (\text{A.13})$$

Die letzte Gleichung wird auch eine diskrete algebraische Matrix–Ricatti–Gleichung genannt. Durch diese Gleichung wird eine Folge $\{P(k+1|k)\}_{k=0}^\infty$ von Matrizen definiert. Falls diese Folge konvergiert, konvergieren auch die Folgen der Kalman–Gains, der Kovarianzen und der Innovationskovarianzen.

Das Problem der Konvergenz der Folge $\{P(k+1|k)\}_{k=0}^\infty$ kann unter Zuhilfenahme von Theorem 17.5.3 in (Lancaster und Rodman, 1995) untersucht werden. Dieses Theorem besagt, dass die Folge konvergiert, falls das Paar (B, A) d –beobachtbar ist und das Paar (A, G) d –stabilisierbar. Zusätzlich muss die Matrix E positiv definit sein und die Matrix Q positiv semidefinit. Die letzten beiden Bedingungen sind offensichtlich erfüllt. Darüberhinaus ist, falls das Paar (A, G) steuerbar ist, der Grenzwert der Folge eine positiv definite Matrix.

Die gerade eingeführten Bezeichnungen bedürfen zunächst einiger erläuternder Worte. Das Paar (C, D) heißt d –stabilisierbar, falls eine Matrix K existiert, so dass die Matrix $A + BK$ stabil ist. Eine Matrix heißt stabil, falls alle ihre Eigenwerte im offenen Einheitskreis liegen. Ein Paar (D, C) heißt d –beobachtbar genau dann, wenn das Paar (C^*, D^*) d –stabilisierbar ist. Dabei ist C^* das komplex Konjugierte des Transponierten C^\top von C . Für detailliertere Informationen vergleiche (Lancaster und Rodman, 1995), Seite 90 bis 91. Es gibt nun einen Satz (Theorem 4.4.2 in (Lancaster und Rodman, 1995)), der besagt, dass das Paar (C, D) d –stabilisierbar ist, falls es steuerbar ist. Ein Paar (C, D) , wobei C eine $n \times n$ – und D eine $n \times m$ Matrix ist, heißt dabei steuerbar genau dann, wenn gilt

$$\text{rang} \begin{bmatrix} D & CD & C^2D & \dots & C^{n-1}D \end{bmatrix} = n \quad (\text{A.14})$$

((Lancaster und Rodman, 1995), Seite 84).

Nach diesen vorbereitenden Ausführungen kann die Frage der Konvergenz nun untersucht werden. Zunächst wird dazu gezeigt, dass das Paar (B, A) d –beobachtbar ist. Dies ist genau dann der Fall, wenn das Paar (A^*, B^*) d –stabilisierbar ist. Dies gilt, wenn dieses Paar steuerbar ist. Somit muss mit $n = 2$ überprüft werden, ob

$$\text{rang} \begin{bmatrix} B^* & A^*B^* \end{bmatrix} = \text{rang} \begin{bmatrix} 1 & 1 \\ 0 & \Delta T \end{bmatrix} = 2 \quad (\text{A.15})$$

ist. Da aber

$$\begin{vmatrix} 1 & 1 \\ 0 & \Delta T \end{vmatrix} = \Delta T > 0 \quad (\text{A.16})$$

gilt, ist dies der Fall.

Als nächstes wird gezeigt, dass das Paar (A, G) ebenfalls steuerbar und somit d -stabilisierbar ist. Dazu muss gezeigt werden, dass gilt

$$\text{rang} \begin{bmatrix} G & AG \end{bmatrix} = \text{rang} \begin{bmatrix} 1 & 0 & 1 & \Delta T \\ 0 & 1 & 0 & \rho \end{bmatrix} = 2,$$

was offensichtlich der Fall ist. Somit ist die Konvergenz gezeigt. Außerdem ist der Grenzwert $\lim_{k \rightarrow \infty} P(k+1|k)$ positiv definit. Theorem 17.5.3 in (Lancaster und Rodman, 1995) besagt darüberhinaus, dass dies für jede positiv semidefinite Initialisierung $P(1|0)$ gilt und der Grenzwert der maximalen Lösung der Matrixgleichung

$$X = A \left[X - XB^T [BXB^T + E]^{-1} BX \right] A^T + GQG^T \quad (\text{A.17})$$

entspricht.

Im Hauptteil werden lediglich die Einträge der Kalman-Gain $K(k)$ in der Update-Gleichung 4.27 und der Innovationskovarianz $S(k)$ für die Berechnung der Mahalanobisdistanz in Gleichung 4.21 benötigt. Dabei führt man wie schon angedeutet die Variablen α und β gemäß der Gleichung

$$\lim_{k \rightarrow \infty} K(k) = \begin{pmatrix} \alpha \\ \beta/\Delta T \end{pmatrix} \quad (\text{A.18})$$

und den Grenzwert $S = \lim_{k \rightarrow \infty} S(k)$ ein. Diese Variablen α und β und der Grenzwert S können im Prinzip mit Methoden wie sie in (Bar-Shalom und Fortmann, 1988) vorgestellt werden, berechnet werden. Diese Herleitungen sind aber umständlich und langwierig und würden deshalb den Rahmen dieser Arbeit sprengen. Es soll darum hier ein alternativer Weg eingeschlagen werden. Dieser beruht auf der Tatsache, dass die Folge der Kalman-Gains äußerst schnell konvergiert. Präziser ist eine sehr genaue Konvergenz schon nach weniger als zehn Iterationen erreicht. Die Grenzwerte können somit durch iterative Anwendung der Matrixgleichungen des Kalman-Filters sehr einfach erhalten werden. Die Abbildungen A.1 bis A.5 zeigen die Ergebnisse dieser numerischen Untersuchungen. Der dabei verwendete Parameter λ ist definiert zu

$$\lambda = \frac{\mu[\Delta T]^2}{\delta}. \quad (\text{A.19})$$

Durch Verwendung der hier eingeführten sogenannten Steady-State-Lösungen lässt sich beispielsweise die fortlaufende rechenintensive Matrixinversion bei der Berechnung der Mahalanobis-Distanz zur Lösung des Korrespondenzproblems (vergleiche Gleichung 4.21 in Kapitel 4.3) vermeiden. Allerdings muss hierbei beachtet werden, dass der Eintrag der Matrix S , der für die Berechnung der Mahalanobisdistanz benötigt wird, und die Werte α und β , die in der Update-Gleichung des Kalman-Filters benutzt werden, von der Länge des Zeitintervalls ΔT abhängen, wie in Abbildung A.6 dargestellt. Empirische Untersuchungen, die in Abbildung A.7 dargelegt sind, zeigen aber, dass die Werte von ΔT sehr stabil nur in einem sehr kleinem Bereich liegen, so

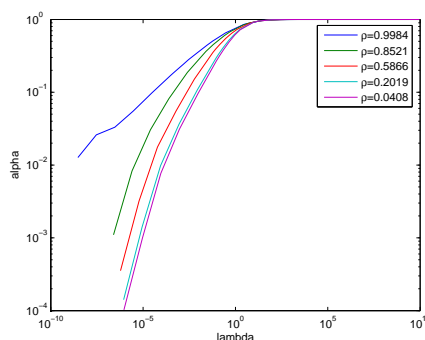
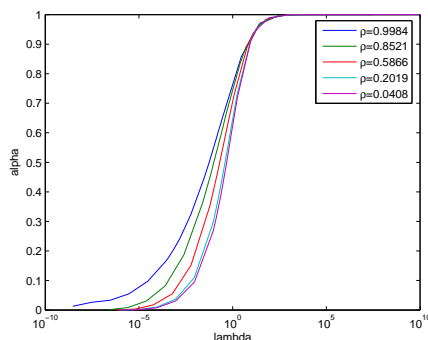


Abbildung A.1: Einfach logarithmische Darstellung. Abbildung A.2: Doppelt logarithmische Darstellung.

α als Funktion von λ und ρ .

dass die Verwendung der Steady–State–Lösungen für einen geeignet gewählten Wert von ΔT eine zulässige Näherung darstellt. Für den Wert $\Delta T = 0.1975$ und die im Hauptteil verwendeten Parameter $\Theta = 20$, $\Sigma = 60$ und $\delta = 780$ erhält man beispielsweise $\alpha = 0.2857$, $\beta = 0.0473$ und $S = 1092.0$. Dieser Wert für S bedeutet dabei für die Gating–Prozedur, dass beispielsweise für eine Schranke von $\lambda = 4$ Messungen, die sechzig Zentimeter von der Vorhersage entfernt liegen, noch dem Ziel zugeordnet werden. Bei einer Schranke von $\lambda = 9$ trifft dies sogar noch für Messungen mit einem Abstand von neunzig Zentimetern zu.

Aufgrund von Gleichung 4.30 existiert auch eine Steady–State–Lösung für die Matrizen $J(k)$ des Kalman–Glätters. Für die eben aufgeführten Werte lautet diese

$$\lim_{k \rightarrow \infty} J(k) = \begin{pmatrix} 0.9499 & -0.1397 \\ 0.2539 & 0.7073 \end{pmatrix}. \quad (\text{A.20})$$

Das negative Vorzeichen des Terms -0.1397 rührt dabei daher, dass hier gemäß Gleichung 4.29 'mit Werten aus der Zukunft' auf die Position in der Gegenwart geschlossen werden soll und deshalb die Geschwindigkeit mit einem negativen Vorfaktor versehen werden muss. Bei den Algorithmen, die nur für den Offline–Betrieb geeignet sind und deshalb auch den Kalman–Glätter benutzen können, wie zum Beispiel die Verfahren aus Kapitel 6, kann auf diese Weise die Komplexität der Methoden weiter reduziert werden.

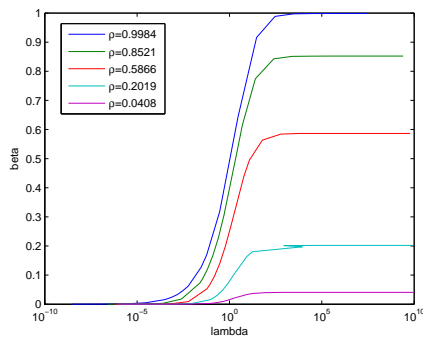


Abbildung A.3: Einfach logarithmische Darstellung.

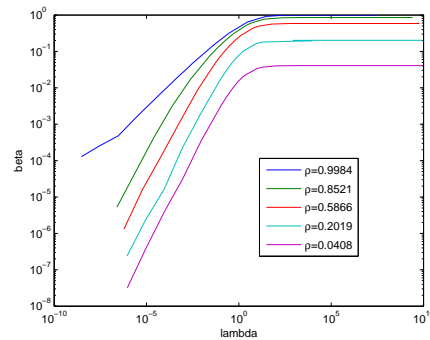
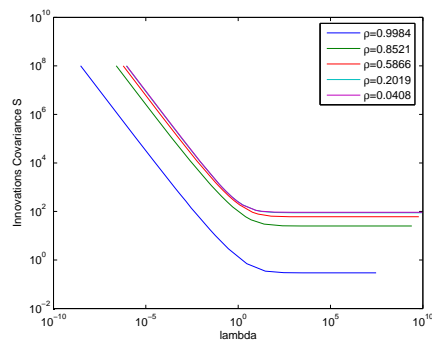
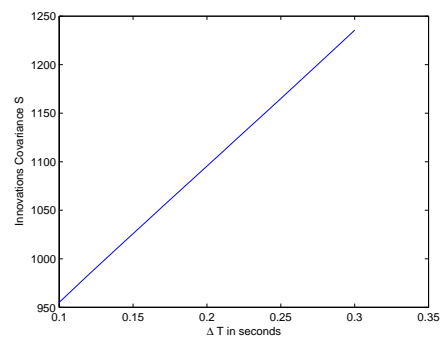
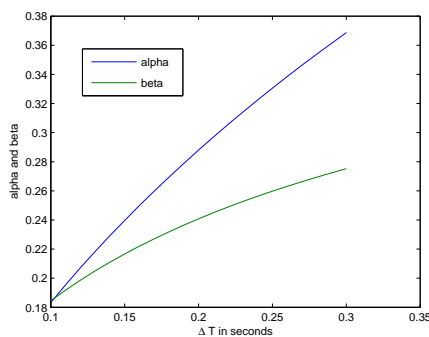


Abbildung A.4: Doppelt logarithmische Darstellung.

β als Funktion von λ und ρ .

Abbildung A.5: S als Funktion von λ und ρ , doppelt logarithmische Darstellung.Abbildung A.6: α , β und S als Funktionen von ΔT .

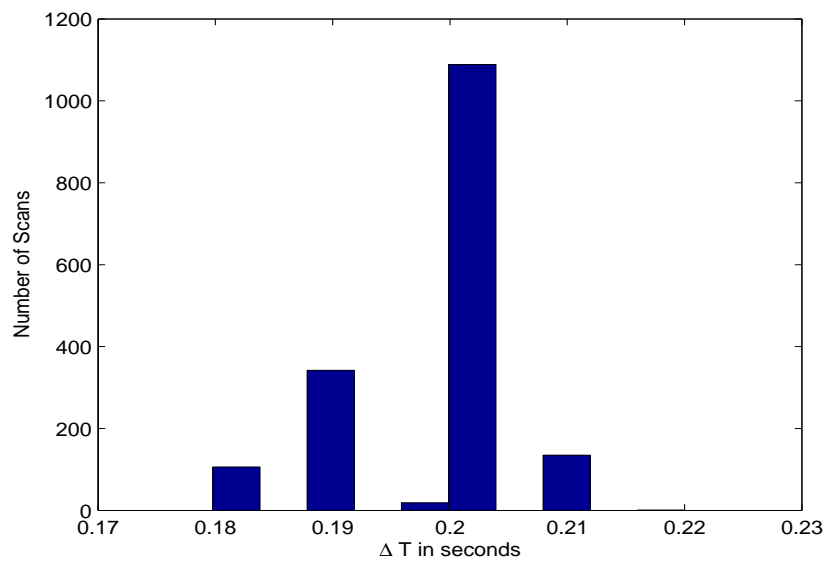


Abbildung A.7: Ein Histogramm für einen Datensatz empirischer Werte des Zeitintervalls ΔT .

Anhang B

Die Berechnung des Schwerpunktes

S_O

Am Anfang der hier dargelegten Betrachtungen steht die Vermutung, dass Algorithmen, die einen Mittelwert aller Messungen von der Oberfläche eines kreisförmigen Objektes, z.B. in der Update-Gleichung des Kalman-Filters, benutzen, einen Punkt im Inneren dieses Objektes schätzen, der der Mittelwert derjenigen Punkte auf der Oberfläche des Objektes ist, die im Sichtfeld des Sensors liegen. Dieses Mittel, das Schwerpunkt S_O (O für Oberfläche) genannt werden soll, wird nun berechnet.

Zur Vereinfachung der Untersuchungen wird zunächst angenommen, dass der Mittelpunkt des beobachteten kreisförmigen Objektes im Ursprung des planaren Koordinatensystems liegt, während sich das als punktförmig angenommene Messgerät auf der x -Achse mit Koordinaten $(d, 0)$ befindet. Die Entfernung des Mittelpunktes des Objektes zum Sensor beträgt somit d . Außerdem sei der Radius des beobachteten Objektes gleich r . Der Schwerpunkt $S_O = (x, y)$ im Innern des Objektes berechnet sich dann wie folgt: aus Symmetriegründen ergibt sich zunächst sofort $y = 0$ und weiter ist

$$x = \frac{1}{\phi} \int_0^\phi x(\theta) d\theta. \quad (\text{B.1})$$

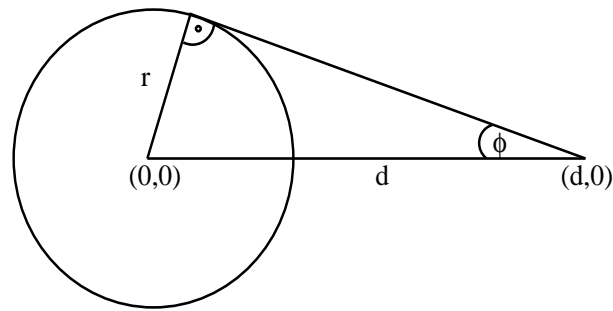
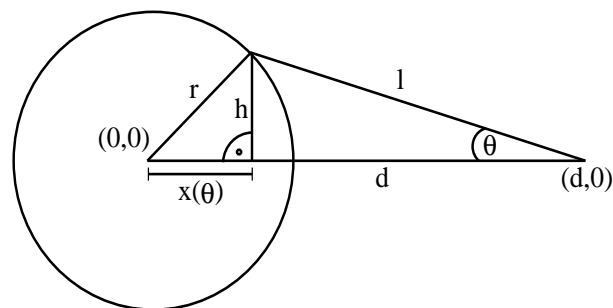
Für die Definition des Grenzwinkels ϕ vergleiche Abbildung B.1 und für die Definition der Abstände $x(\theta)$ vergleiche Abbildung B.2.

Diese Vorgehensweise trägt auch der Tatsache Rechnung, dass die Laserstrahlen das Objekt in der Mitte dichter treffen. Der Grenzwinkel ϕ folgt dabei nach Abbildung B.1 aus der Beziehung

$$\sin \phi = \frac{r}{d}. \quad (\text{B.2})$$

Die Entfernung $x(\theta)$ erhält man aus den bekannten Werten d und r und dem Winkel θ wie folgt (Abbildung B.2): zunächst ist mit dem Satz von Pythagoras

$$r^2 = h^2 + x^2(\theta) \quad (\text{B.3})$$

Abbildung B.1: Bestimmung des Grenzwinkels ϕ .Abbildung B.2: Berechnung der Entfernung $x(\theta)$.

und

$$(d - x(\theta))^2 + h^2 = l^2 \quad (\text{B.4})$$

bzw. nach Ausmultiplizieren der Klammer

$$d^2 - 2dx(\theta) + x^2(\theta) + h^2 = l^2. \quad (\text{B.5})$$

Außerdem gilt die Beziehung

$$\sin \theta = \frac{h}{l}. \quad (\text{B.6})$$

Aus der letzten Gleichung folgt

$$l = \frac{h}{\sin \theta}. \quad (\text{B.7})$$

Eingesetzt in Gleichung B.5 ergibt dies

$$d^2 - 2dx(\theta) + x^2(\theta) + h^2 = \frac{h^2}{\sin^2 \theta}. \quad (\text{B.8})$$

Löst man die letzte Beziehung nach h^2 auf, so erhält man

$$h^2 = \frac{d^2 - 2dx(\theta) + x^2(\theta)}{\frac{1}{\sin^2 \theta} - 1} = \frac{\sin^2 \theta (d^2 - 2dx(\theta) + x^2(\theta))}{1 - \sin^2 \theta} =$$

$$= \frac{\sin^2 \theta (d^2 - 2dx(\theta) + x^2(\theta))}{\cos^2 \theta} = \tan^2 \theta (d^2 - 2dx(\theta) + x^2(\theta)). \quad (\text{B.9})$$

Löst man Gleichung B.3 nach h^2 auf und setzt das Ergebnis in Gleichung B.9 ein, so liefert dies

$$r^2 = d^2 \tan^2 \theta - 2dx(\theta) \tan^2 \theta + x^2(\theta) \tan^2 \theta + x^2(\theta). \quad (\text{B.10})$$

Multiplikation mit $\cos^2 \theta$ ergibt

$$\begin{aligned} r^2 \cos^2 \theta - d^2 \sin^2 \theta &= x^2(\theta) (\sin^2 \theta + \cos^2 \theta) - 2dx(\theta) \sin^2 \theta = \\ &= x^2(\theta) - 2dx(\theta) \sin^2 \theta. \end{aligned} \quad (\text{B.11})$$

Wendet man nun die quadratische Ergänzung auf die letzte Gleichung an, so erhält man

$$r^2 \cos^2 \theta - d^2 \sin^2 \theta + d^2 \sin^4 \theta = (x(\theta) - d \sin^2 \theta)^2. \quad (\text{B.12})$$

Somit gilt

$$x(\theta) - d \sin^2 \theta = \pm \sqrt{r^2 \cos^2 \theta - d^2 (1 - \sin^2 \theta) \sin^2 \theta}. \quad (\text{B.13})$$

Die Lösung mit dem Minuszeichen, die der Lösung mit dem kleineren Wert von $x(\theta)$ entspricht, beschreibt dabei den zweiten Schnittpunkt des Strahls von $(d, 0)$ zum Kreis mit dem Kreis, der dichter am Mittelpunkt liegt und weiter vom Sensor entfernt ist, so dass zur richtigen Lösung das Pluszeichen gehört.¹ Es gilt also

$$\begin{aligned} x(\theta) - d \sin^2 \theta &= \sqrt{r^2 \cos^2 \theta - d^2 (1 - \sin^2 \theta) \sin^2 \theta} = \\ &= \sqrt{r^2 \cos^2 \theta - d^2 \sin^2 \theta \cos^2 \theta} = \pm \cos \theta \sqrt{r^2 - d^2 \sin^2 \theta}. \end{aligned} \quad (\text{B.14})$$

Auch in diesem Fall liefert das Pluszeichen den richtigen Schnittpunkt, es ist folglich

$$x(\theta) - d \sin^2 \theta = \cos \theta \sqrt{r^2 - d^2 \sin^2 \theta}. \quad (\text{B.15})$$

Als Endergebnis für $x(\theta)$ erhält man schließlich

$$x(\theta) = d \sin^2 \theta + \cos \theta \sqrt{r^2 - d^2 \sin^2 \theta}, \quad (\text{B.16})$$

so dass mit Gleichung B.1 insgesamt gilt

$$x = \frac{1}{\phi} \int_0^\phi x(\theta) d\theta = \frac{1}{\phi} \int_0^\phi \left[d \sin^2 \theta + \cos \theta \sqrt{r^2 - d^2 \sin^2 \theta} \right] d\theta. \quad (\text{B.17})$$

Bezüglich der Stammfunktion des ersten Terms im Integral gilt (Bronstein und Semendjajew, 1987)

$$\int \sin^2 \theta d\theta = \frac{1}{2} \theta - \frac{1}{2} \sin \theta \cos \theta = \frac{1}{2} \theta - \frac{1}{2} \sin \theta \sqrt{1 - \sin^2 \theta}. \quad (\text{B.18})$$

¹Vergleiche hierzu auch Abbildung C.2.

Unter Berücksichtigung von Gleichung B.2 folgt daraus

$$\begin{aligned} \frac{1}{\phi} \int_0^\phi d \sin^2 \theta d\theta &= \frac{d}{\phi} \left[\frac{1}{2} \theta - \frac{1}{2} \sin \theta \sqrt{1 - \sin^2 \theta} \right]_0^\phi = \\ &= \frac{d}{\phi} \left(\frac{1}{2} \phi - \frac{1}{2} \sin \phi \sqrt{1 - \sin^2 \phi} \right) = \frac{d}{2} - \frac{d}{2\phi} \frac{r}{d} \sqrt{1 - \frac{r^2}{d^2}} = \\ &= \frac{d}{2} - \frac{r}{2d \arcsin \frac{r}{d}} \sqrt{d^2 - r^2}. \end{aligned} \quad (\text{B.19})$$

Die Stammfunktion des zweiten Terms erhält man mit Hilfe der Integration durch Substitution. Dazu wird $u = d \sin \theta$ gesetzt. Damit ist

$$du = d \cos \theta d\theta \quad (\text{B.20})$$

und folglich

$$\int \cos \theta \sqrt{r^2 - d^2 \sin^2 \theta} d\theta = \frac{1}{d} \int \sqrt{r^2 - u^2} du. \quad (\text{B.21})$$

Die Stammfunktion von $\int \sqrt{r^2 - u^2} du$ ist (Bronstein und Semendjajew, 1987)

$$\frac{1}{2} \left(u \sqrt{r^2 - u^2} + r^2 \arcsin \frac{u}{r} \right). \quad (\text{B.22})$$

Die untere Integrationsgrenze u_1 errechnet sich zu

$$u_1 = d \sin(0) = 0 \quad (\text{B.23})$$

und die obere Integrationsgrenze u_2 zu

$$u_2 = d \sin \arcsin r/d = r. \quad (\text{B.24})$$

Somit gilt

$$\begin{aligned} \frac{1}{\phi} \int_0^\phi \cos \theta \sqrt{r^2 - d^2 \sin^2 \theta} d\theta &= \frac{1}{d\phi} \int_0^r \sqrt{r^2 - u^2} du = \\ &= \frac{1}{d\phi} \left[\frac{1}{2} \left(u \sqrt{r^2 - u^2} + r^2 \arcsin \frac{u}{r} \right) \right]_0^r = \frac{1}{2d\phi} (r^2 \arcsin 1 - r^2 \arcsin 0) = \\ &= \frac{r^2}{2d\phi} \frac{\pi}{2} = \frac{r^2 \pi}{4d \arcsin \frac{r}{d}}. \end{aligned} \quad (\text{B.25})$$

Die Kombination der Gleichungen B.17, B.19 und B.25 liefert schließlich als Endergebnis

$$x = \frac{d}{2} - \frac{r}{2d \arcsin \frac{r}{d}} \sqrt{d^2 - r^2} + \frac{r^2 \pi}{4d \arcsin \frac{r}{d}} = \frac{d}{2} + \frac{r}{2d \arcsin \frac{r}{d}} \left(\frac{\pi r}{2} - \sqrt{d^2 - r^2} \right). \quad (\text{B.26})$$

Zur Herleitung des im Haupttext eingeführten Grenzwertes $\lim_{d \rightarrow \infty} x$ wird die Reihenentwicklung der Arcussinus-Funktion benutzt. Es ist

$$\arcsin(x) = x + o(x^3). \quad (\text{B.27})$$

Somit ist

$$\begin{aligned} \lim_{d \rightarrow \infty} x &= \lim_{d \rightarrow \infty} \left[\frac{d}{2} + \frac{r}{2d \left(\frac{r}{d} + o\left(\left(\frac{1}{d}\right)^3\right)\right)} \left(\frac{\pi r}{2} - \sqrt{d^2 - r^2} \right) \right] = \\ &= \lim_{d \rightarrow \infty} \left[\frac{d}{2} + \frac{1}{2} \frac{1}{1 + o\left((d)^{-2}\right)} \left(\frac{\pi r}{2} - d \sqrt{1 - \frac{r^2}{d^2}} \right) \right]. \end{aligned} \quad (\text{B.28})$$

Mit den wohlbekanntenen Reihenentwicklungen

$$\frac{1}{1+x} = 1 - o(x) \quad (\text{B.29})$$

und

$$\sqrt{1-x} = 1 - \frac{1}{2}o(x) \quad (\text{B.30})$$

folgt daraus

$$\begin{aligned} x_\infty &= \lim_{d \rightarrow \infty} x = \lim_{d \rightarrow \infty} \left[\frac{d}{2} + \frac{1}{2} \left(1 - o\left((d)^{-2}\right) \right) \left(\frac{\pi r}{2} - d \left(1 - \frac{1}{2}o\left((d)^{-2}\right) \right) \right) \right] = \\ &= \lim_{d \rightarrow \infty} \left[\frac{d}{2} + \frac{1}{2} \left(\frac{\pi r}{2} - d + o\left((d)^{-2}\right) \right) \right] = \frac{1}{2} \left(d + \frac{\pi r}{2} - d \right) = \frac{\pi r}{4}. \end{aligned} \quad (\text{B.31})$$

Das letzte Resultat kann auch auf eine sehr verschiedene Art und Weise hergeleitet werden. Im Falle $d = \infty$ treffen nämlich die Laserstrahlen parallel auf das Ziel, ähnlich wie im Falle der Beobachtung eines Objektes aus unendlicher Entfernung in der geometrischen Optik (Abbildung B.3). Somit ist

$$x_\infty = \frac{1}{r} \int_0^r x(s) ds, \quad (\text{B.32})$$

wobei s und $x(s)$ definiert sind wie in Abbildung B.4. Für die Ableitung von $x(s)$ kann dabei wieder der Satz des Phytagoras verwendet werden. Nach Abbildung B.4 ist

$$x(s)^2 + s^2 = r^2 \quad (\text{B.33})$$

und somit

$$x(s) = \sqrt{r^2 - s^2}. \quad (\text{B.34})$$

Das auftretende Integral kann nach (Bronstein und Semendjajew, 1987) gelöst werden zu

$$x_\infty = \frac{1}{r} \int_0^r \sqrt{r^2 - s^2} ds = \frac{1}{2r} \left[s\sqrt{r^2 - s^2} + r^2 \arcsin \frac{s}{r} \right]_0^r = \frac{1}{2r} r^2 \arcsin 1 = \frac{\pi r}{4}. \quad (\text{B.35})$$

Im Falle von $r = 27 \text{ cm}$ ist beispielsweise

$$x_\infty = 21.2058 \text{ cm}. \quad (\text{B.36})$$

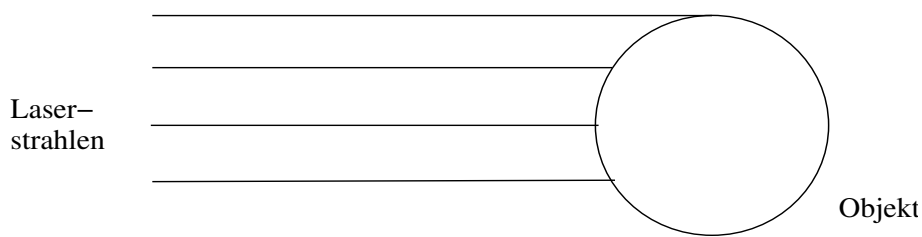


Abbildung B.3: Beobachtung aus unendlicher Entfernung.

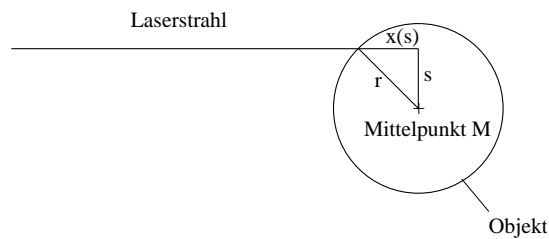


Abbildung B.4: Beobachtung aus unendlicher Entfernung, Ableitung von $x(s)$.

Anhang C

Das Simulationsprogramm

Zur Erzeugung der zum quantitativen Vergleich der Verfahren in Kapitel 5 benötigten Daten wurde ein Simulationsprogramm entwickelt, d.h. es wurden Monte–Carlo–Simulationen durchgeführt. Von der Verwendung von Daten aus realen Experimenten wurde deshalb abgesehen, weil die tatsächliche Position des Objektes jeweils sehr genau bekannt sein musste, da die Unterschiede zwischen den einzelnen Verfahren teilweise im Millimeterbereich liegen. Dies ist aber, wie in einem anderen Zusammenhang schon von anderen Autoren festgestellt wurde (Zhao und Shibasaki, 2005), mit realen Daten nicht möglich, auch nicht mit Methoden wie sie in (Bellotto und Hu, 2006; Bellotto und Hu, 2007a; Bellotto und Hu, 2007b; Bellotto und Hu, 2007c) vorgestellt wurden. Dort wurden mit einem an der Decke befestigten Kamerasystem die Positionen von Personen vermessen, die sich zur Positionsmarkierung jeweils eine Farbplatte über den Kopf hielten.

Bei der Simulation wurde angenommen, dass der Lasersensor seine Umgebung mit einer Winkelauflösung von einem Grad abtastet, wie es in der Regel auch in unserer Experimentalumgebung der Fall ist. Das zu beobachtende kreisförmige Objekt mit bekanntem Radius $r = 27$ cm bewegt sich auf einer Kreisbahn mit Radius d um den Beobachter, der sich im Ursprung des Koordinatensystems befinden soll. Es wird weiter angenommen, dass es bei dieser Bewegung keine Abweichungen von der Kreisbahn gibt. Das Prozessrauschen rührt damit lediglich von der Zentripetalbeschleunigung her, die das Ziel auf der Kreisbahn hält.

Zur Erzeugung der Simulationsdaten ist es zunächst nötig, zu ermitteln, unter welchen Winkeln ϕ der Laserstrahl von dem Objekt reflektiert wird. Dazu werden der maximale und der minimale Winkel berechnet, die von dem Objekt besetzt werden. Diese Winkel werden mit ϕ_{\min} und ϕ_{\max} bezeichnet. Mit dem bekannten Winkel ϕ_P der Polarkoordinaten des Mittelpunktes des Objektes ist nach Abbildung C.1

$$\phi_{\min} = \phi_P - \arcsin \frac{r}{d} \quad (\text{C.1})$$

und

$$\phi_{\max} = \phi_P + \arcsin \frac{r}{d}. \quad (\text{C.2})$$

Damit sind aber auch die Winkel, für die der Laserstrahl das Objekt trifft, bekannt.

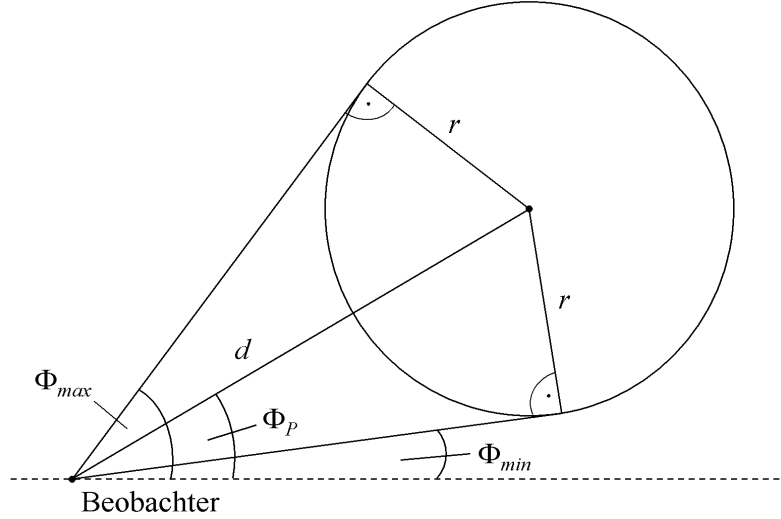


Abbildung C.1: Bestimmung der Winkel ϕ_{min} und ϕ_{max} .

Für jeden dieser Winkel ϕ wird nun der Einheitsvektor

$$e = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \cos \phi \\ \sin \phi \end{pmatrix} \quad (C.3)$$

in Richtung dieses Winkels berechnet. Durch diesen Einheitsvektor wird eine Halbgerade definiert, die von dem Beobachter zum Ziel verläuft. Die Punkte auf dieser Halbgeraden sind alle Punkte mit Koordinaten $\alpha \cdot e$ mit $\alpha \geq 0$. Dem Schnittpunkt einer solchen Halbgeraden mit dem Objekt entspricht dabei nach Abbildung C.2 der Wert α , für den gilt

$$\left\{ l \left[\alpha \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right] \right\}^2 = r^2. \quad (C.4)$$

Dabei sind x_1 und x_2 die kartesischen Koordinaten des Mittelpunktes des Objektes und $l(v)$ ist die euklidische Länge des Vektors $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$. Aus Gleichung C.4 erhält man zunächst

$$\left[l \begin{pmatrix} \alpha y_1 - x_1 \\ \alpha y_2 - x_2 \end{pmatrix} \right]^2 = r^2. \quad (C.5)$$

Mit $l^2(v) = v_1^2 + v_2^2$ ergibt sich daraus

$$\alpha^2 y_1^2 - 2\alpha y_1 x_1 + x_1^2 + \alpha^2 y_2^2 - 2\alpha y_2 x_2 + x_2^2 = r^2 \quad (C.6)$$

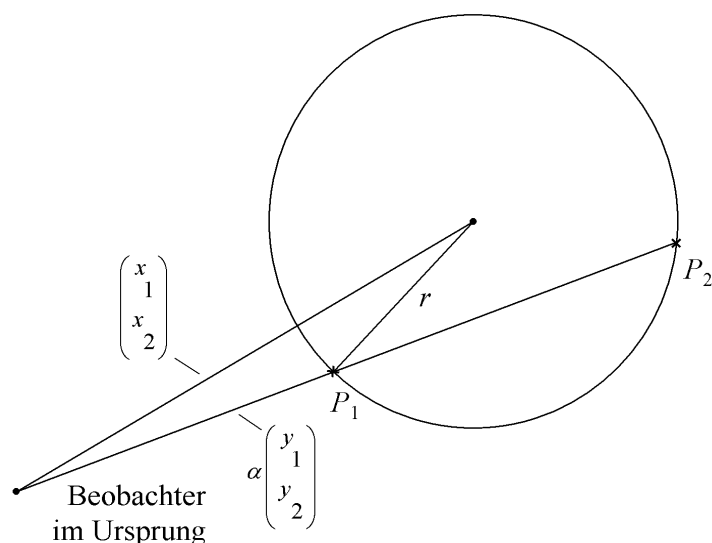


Abbildung C.2: Bestimmung des Faktors α und geometrische Bedeutung der beiden Lösungen $\alpha_{1/2}$.

und damit

$$\alpha^2 (y_1^2 + y_2^2) - 2\alpha(y_1x_1 + y_2x_2) + x_1^2 + x_2^2 - r^2 = 0. \quad (\text{C.7})$$

Da e ein Einheitsvektor war, gilt aber $y_1^2 + y_2^2 = 1$ und somit ist

$$\alpha^2 - 2\alpha(y_1x_1 + y_2x_2) + x_1^2 + x_2^2 - r^2 = 0. \quad (\text{C.8})$$

Mit der Lösungsformel für quadratische Gleichungen liefert dies

$$\begin{aligned} \alpha_{1/2} &= \frac{2(y_1x_1 + y_2x_2) \pm \sqrt{4(y_1x_1 + y_2x_2)^2 - 4(x_1^2 + x_2^2 - r^2)}}{2} = \\ &= y_1x_1 + y_2x_2 \pm \sqrt{y_1^2x_1^2 + 2y_1x_1y_2x_2 + y_2^2x_2^2 - x_1^2 - x_2^2 + r^2} = \\ &= y_1x_1 + y_2x_2 \pm \sqrt{x_1^2(y_1^2 - 1) + x_2^2(y_2^2 - 1) + 2y_1x_1y_2x_2 + r^2} = \\ &= y_1x_1 + y_2x_2 \pm \sqrt{-x_1^2y_2^2 - x_2^2y_1^2 + 2y_1x_1y_2x_2 + r^2}. \end{aligned} \quad (\text{C.9})$$

Das Pluszeichen in obiger Gleichung gehört zu dem größeren Wert von α und damit zu dem Schnittpunkt P_2 der Halbgeraden mit dem Kreis, der weiter vom Beobachter entfernt liegt. Das Minuszeichen entspricht somit dem Schnittpunkt P_1 und daher der richtigen Lösung (Abbildung C.2). Es gilt also

$$\alpha = y_1x_1 + y_2x_2 - \sqrt{2y_1x_1y_2x_2 - x_1^2y_2^2 - x_2^2y_1^2 + r^2}. \quad (\text{C.10})$$

Den Fehler in der Messung führt man nun dadurch ein, dass man zu dem errechneten theoretischen Wert für α ein weißes Rauschen aufaddiert. Es wird also gesetzt

$$\alpha_{real} = \alpha + \sigma\epsilon. \quad (\text{C.11})$$

Dabei ist ϵ eine Realisierung einer Normalverteilung mit Mittelwert Null und Standardabweichung Eins, also einer sogenannten Standardnormalverteilung. Durch die Multiplikation mit dem Faktor σ erhält man daraus eine Realisierung eines Gaußschen Rauschens mit Mittelwert Null und Standardabweichung σ (Renyi, 1970). Die kartesischen Koordinaten der simulierten Messung lauten dann $\alpha_{real} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$. Diese Daten werden schließlich durch die Tracking-Algorithmus ausgewertet.

Anhang D

Graphiken einiger Experimente zur Identifikation mit einem Kamerasystem

In diesem Anhang sind die Graphiken einiger Versuche zur Identifikation von Personen mit Hilfe eines Kamerasystems (Kapitel 9.2) aufgenommen, auf deren Darstellung im Hauptteil verzichtet wurde. Es handelt sich dabei um Experimente mit drei Versuchspersonen. Für Details der graphischen Darstellung und der Versuchsanordnung wird auf den Hauptteil verwiesen.

Zunächst werden drei Versuche dargestellt, bei denen genau zwei der drei Versuchspersonen ihre Positionen während der Interaktion vertauschen (Abbildung D.1 bis D.6).

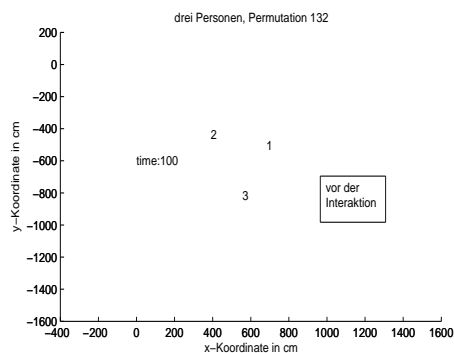


Abbildung D.1: Drei Einzelpersonen vor der Interaktion.

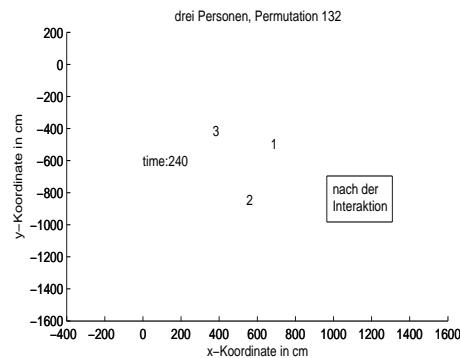


Abbildung D.2: Drei Einzelpersonen nach der Interaktion.

Drei Personen, Permutation 123 \mapsto 132.

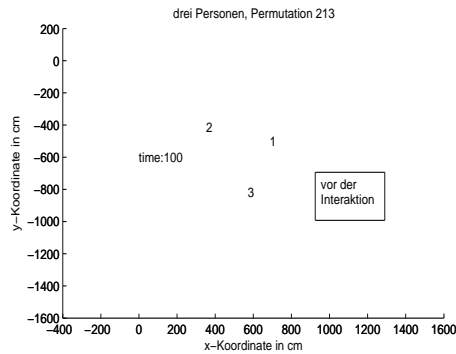


Abbildung D.3: Drei Einzelpersonen vor der Interaktion.

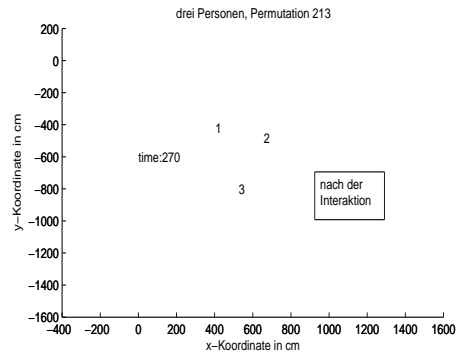


Abbildung D.4: Drei Einzelpersonen nach der Interaktion.

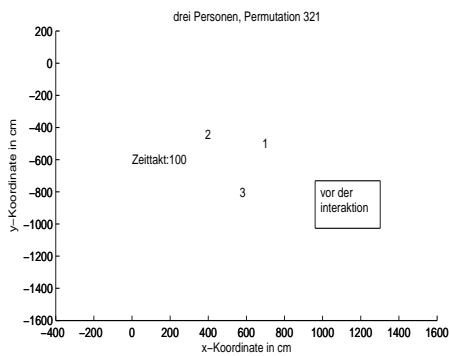
Drei Personen, Permutation 123 \mapsto 213.

Abbildung D.5: Drei Einzelpersonen vor der Interaktion.

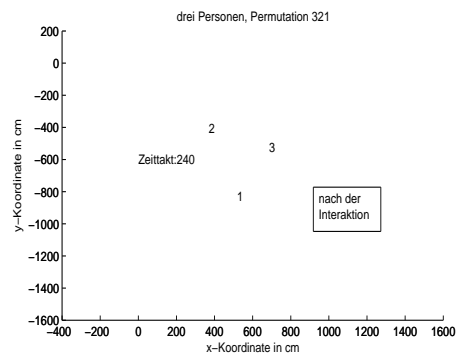


Abbildung D.6: Drei Einzelpersonen nach der Interaktion.

Drei Personen, Permutation 123 \mapsto 321.

Als nächstes werden zwei Versuche wiedergegeben, bei denen alle drei Versuchspersonen nach der Interaktion neue Positionen einnehmen (Abbildung D.7 bis D.10).

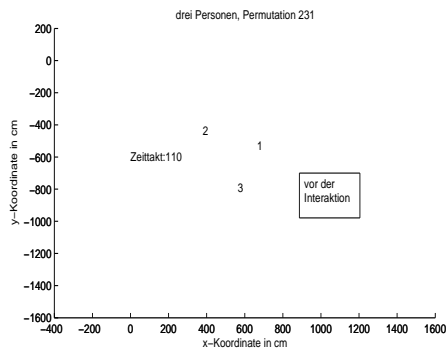


Abbildung D.7: Drei Einzelpersonen vor der Interaktion.

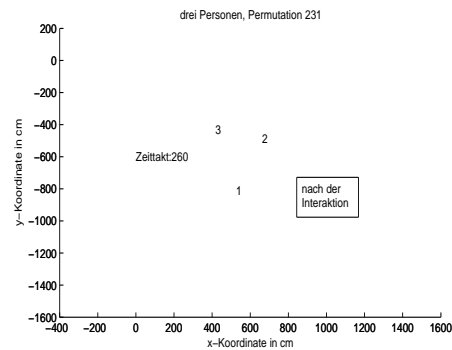


Abbildung D.8: Drei Einzelpersonen nach der Interaktion.

Drei Personen, Permutation $123 \mapsto 231$.

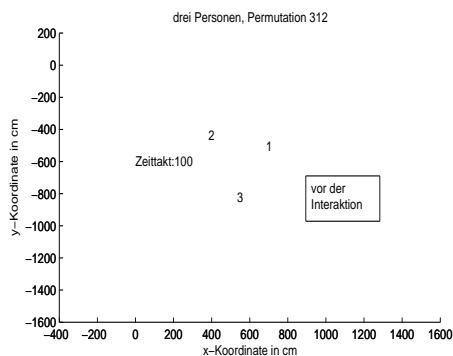


Abbildung D.9: Drei Einzelpersonen vor der Interaktion.

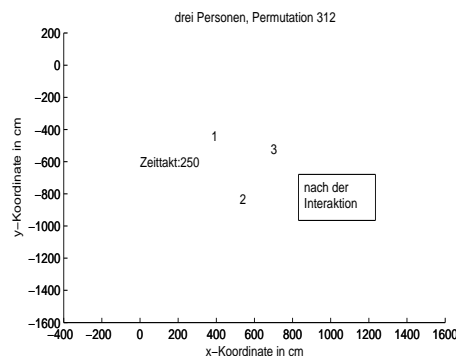


Abbildung D.10: Drei Einzelpersonen nach der Interaktion.

Drei Personen, Permutation $123 \mapsto 312$.

Nun ist ein Versuch wiedergegeben, bei dem die Versuchspersonen vor der Interaktion andere Startpositionen einnehmen (Abbildung D.11 und D.12). Dazu wird die Permutation $213 \mapsto 123$ herangezogen. Dies bedeutet beispielsweise, dass die Anfangsposition der Person 2 Position 1, also der rechte obere Bereich, ist.

Abschließend ist ein Versuch dargestellt, bei dem zwei Interaktionen hintereinander stattfinden (Abbildungen D.13 bis D.15).

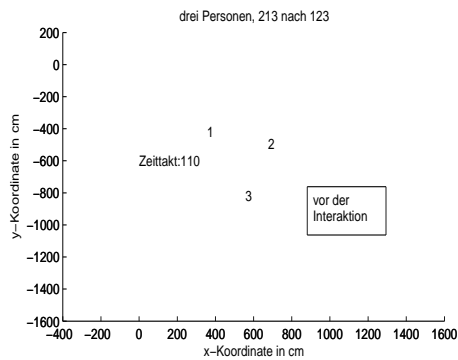


Abbildung D.11: Drei Einzelpersonen vor der Interaktion.

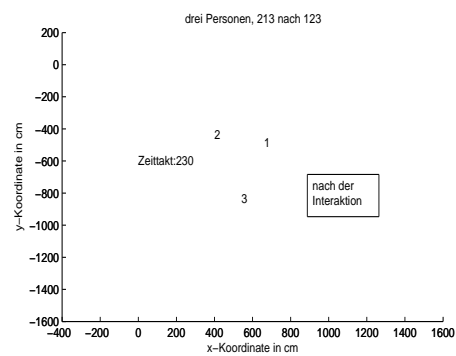


Abbildung D.12: Drei Einzelpersonen nach der Interaktion.

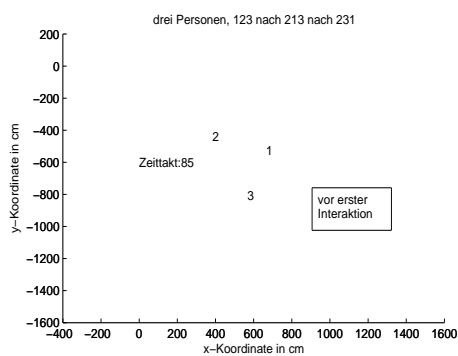
Drei Personen, Permutation 213 \mapsto 123.

Abbildung D.13: Drei Einzelpersonen vor der ersten Interaktion.

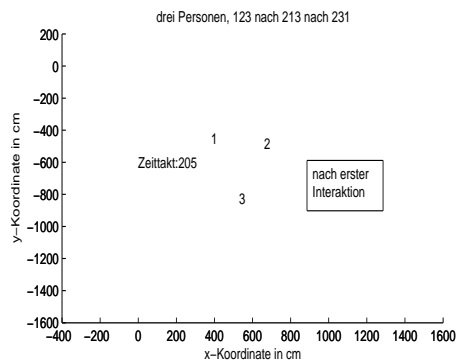


Abbildung D.14: Drei Einzelpersonen nach der ersten Interaktion.

Drei Personen, Permutation 123 \mapsto 213 \mapsto 231.

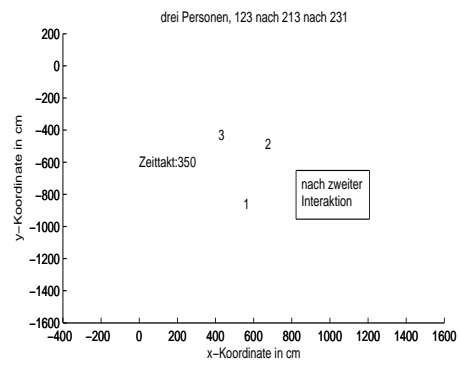


Abbildung D.15: Drei Einzelpersonen nach der zweiten Interaktion, Permutation $123 \mapsto 213 \mapsto 231$.

Anhang E

Voruntersuchungen zur Messung von Remissionswerten mit Laserscannern und Graphiken einiger Experimente

E.1 Voruntersuchungen

Im Rahmen der vorliegenden Arbeit wurde, soweit es dem Autor bekannt ist, zum ersten Mal untersucht, inwieweit die Messungen von Laserscannern zur Identifikation von Personen geeignet sind. Hierfür kommen in erster Linie die Remissionswerte in Frage, die kommerzielle Laserscanner neben den Entfernungswerten noch zusätzlich liefern. Bei diesen Remissionswerten handelt es sich um künstliche Werte, die der Laserscanner als Ausgabe bereitstellt und die ein Maß für die Intensität des registrierten Laserstrahls darstellen. Die Hoffnung, die Ausgangspunkt der Untersuchungen war, bestand dabei darin, dass sich diese Remissions- oder auch Reflexionswerte für unterschiedliche Hosenmaterialien in ausreichendem Maße unterscheiden, so dass aus den Messungen der Remissionswerte Informationen zur Unterscheidung von Personen gewonnen werden können.

Zunächst wurden einige einfache Experimente durchgeführt. Dabei liefen eine oder zwei Personen in unserer Experimentalhalle umher und wurden mit einem Laserscanner beobachtet. Der in Kapitel 7 beschriebene Algorithmus wurde dann angewandt, um die Messungen mit den Personen oder Tracks zu assoziieren. Wegen der Ausdehnung der Objekte lag dabei in den meisten Fällen mehr als eine Messung von derselben Person vor. Deshalb wurde für den Remissionswert jeweils ein ungewichtetes Mittel aller Messungen herangezogen, die zuvor mit derselben Person assoziiert worden waren.

Im einem ersten Experiment wurde eine einzelne Person beobachtet, die ein Paar Blue Jeans trug. Abbildung E.1 zeigt die Ergebnisse. Man erkennt, dass für dieselbe Person eine große Bandbreite in der Variation der Remissionswerte bestehen kann. Der relative Unterschied der Remissionswerte beträgt für diese Hose nämlich mehr als zehn Prozent. Ebenso sind die Unterschiede zwischen aufeinanderfolgenden Scans

oft sehr groß. Somit muss es noch andere Faktoren als die Farbe der Hosen geben, die einen Einfluss auf die Remissionswerte haben. Als mutmaßliche weitere Ursachen können die Entfernung des Laserscanners von der Person oder der Winkel, unter dem der Laserstrahl die Hose trifft, angeführt werden.

Da die Remissionswerte auch für direkt aufeinanderfolgende Scans große Unterschiede zeigen und sich hierbei die Randbedingungen wie die Entfernung nur unwesentlich geändert haben können, kann man vermuten, dass das statistische Rauschen einen weiteren erheblichen Einflussfaktor auf die Messungen darstellt. Deshalb könnte es zunächst mehr Sinn machen, geglättete Remissionswerte zu betrachten, um den Einfluss dieses Rauschens zu reduzieren. Dazu wurde der Durchschnitt für die Entfernung und die Intensität über zwanzig hintereinanderfolgende Messungen herangezogen. Die Ergebnisse werden in Abbildung E.2 gezeigt. Es ist offensichtlich, dass die geglätteten Daten ein wesentlich regelmäßigeres Verhalten zeigen als die ursprünglichen Werte.

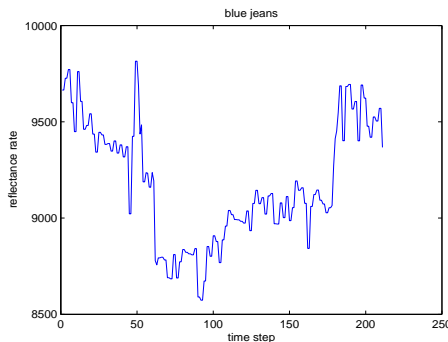


Abbildung E.1: Remissionswerte für eine Blue Jeans.

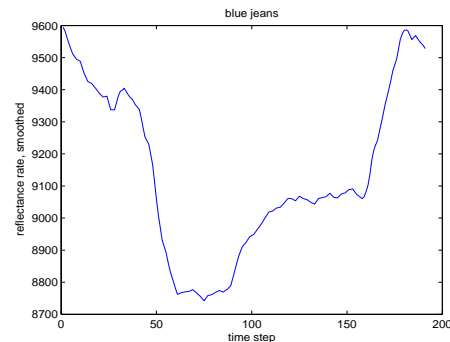


Abbildung E.2: Geglättete Remissionswerte für eine Blue Jeans.

Remissionswerte für eine blaue Jeans.

Als nächstes werden zwei verschiedene Hosen miteinander verglichen. Abbildung E.3 zeigt dazu die Verhältnisse für ein Paar dunkelbeiger Hosen. Es fallen dabei einige bemerkenswerte Unterschiede zu dem Paar blauer Jeans auf. Zunächst besteht eine größere Variabilität von Scan zu Scan. Zweitens besteht eine größere absolute und relative Variabilität über den gesamten Datensatz (mehr als 15 Prozent). Drittens ist das Maximum mit einem Wert von ca. 10400 größer. Abbildung E.4 zeigt auch wieder die geglätteten Werte.

Für dieses Paar Hosen wurde auch die Abhängigkeit der Remissionswerte von der Entfernung der Person zum Beobachter genauer untersucht. Dieser Abstand wurde dabei aus der Positionsschätzung für die Versuchsperson, berechnet durch den Tracking-Algorithmus, und der bekannten Position des Roboters bestimmt. Abbildung E.5 zeigt diese Entfernungen. Der Vergleich von Abbildung E.4 und E.5 zeigt ein sehr beeindruckendes Resultat. Es besteht nämlich eine starke Korrelation zwischen der Remis-

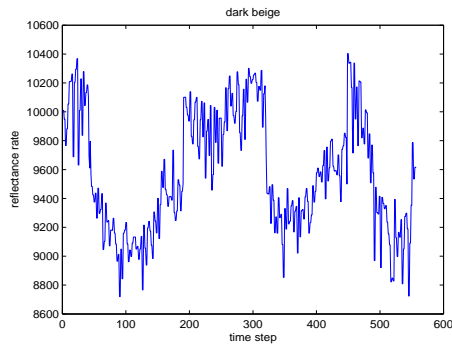


Abbildung E.3: Remissionswerte für eine dunkelbeige Hose.

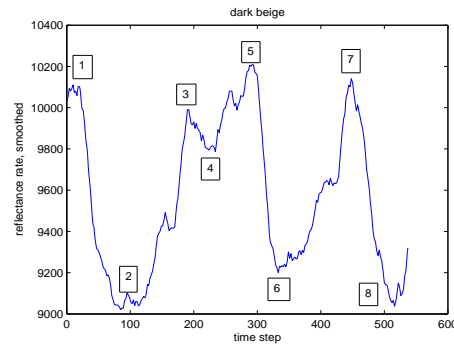


Abbildung E.4: Geglättete Remissionswerte für eine dunkelbeige Hose.

Remissionswerte für eine dunkelbeige Hose.

sion und der Entfernung der Person zum Beobachter. Die Maxima 2, 4, 6 und 8 der Entfernung korrespondieren zu den entsprechenden Minima der Remission, während die Minima 1, 3, 5 und 7 der Entfernung den entsprechenden Maxima der Remission entsprechen.

Als Zwischenresultat kann folglich festgehalten werden, dass es im Prinzip möglich sein sollte, Personen aufgrund der Remissionswerte ihrer Hosen zu identifizieren. Es muss aber dabei berücksichtigt werden, dass diese Remissionswerte nicht nur durch das Material und die Farbe der Hosen, die die Personen tragen, beeinflusst werden, sondern auch von der Entfernung der Probanden zum Beobachter.

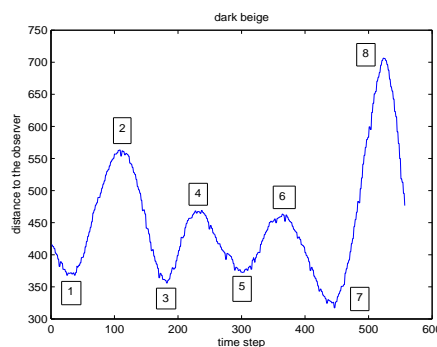


Abbildung E.5: Entfernung zum Beobachter für die dunkelbeige Hose.

Nun wird der Fall einer Begegnung zweier Personen untersucht. Dabei sollte der Fall betrachtet werden, dass eine Person Hosen trägt, die aus einem bestimmten Material mit einem charakteristischen Reflektionsverhalten bestehen. Deshalb wurden die Hosen einer Person mit Aluminiumfolie überklebt. Diese Vorgehensweise ist auch

nicht zu realitätsfern. Als Anwendung kann man sich nämlich einen mobilen Roboter vorstellen, der zusammen mit einigen Wachmännern ein Museum bewachen soll. Diese könnten Hosen mit speziellem Reflektionsverhalten tragen, um dem Roboter zu helfen, diese von Eindringlingen zu unterscheiden.

Abbildung E.6 zeigt die Resultate vor einer Begegnung. Dabei sind die Messungen von der Person mit der Aluminiumfolie mit grüner Farbe und diejenigen von dem anderen Probanden mit blauer Farbe wiedergegeben. Es ist sehr einfach, die Person mit der Aluminiumfolie von der anderen Person zu unterscheiden. Die Remissionswerte sind im ersten Fall im Durchschnitt viel höher und auch die Maxima sind ungefähr 2000 Einheiten größer. Darüberhinaus ist die Variabilität viel höher. Abbildung E.8 zeigt die Verhältnisse nach der Begegnung. Wiederum ist es sehr einfach, zu erkennen, dass der grüne Graph von der Person mit der Aluminiumfolie herrührt. Die Abbildungen E.7 und E.9 zeigen jeweils wieder die geglätteten Werte.

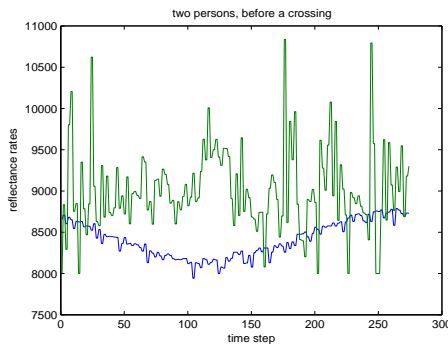


Abbildung E.6: Remissionswerte für ein Paar normaler Hosen und ein Paar mit Aluminiumfolie bedeckter Hosen vor einer Begegnung.

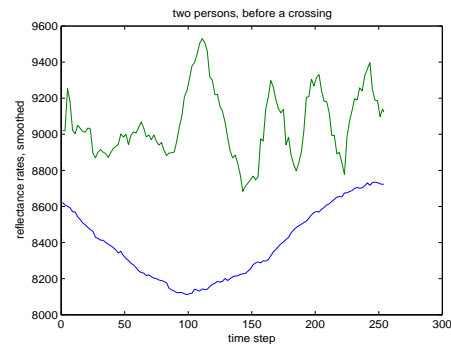


Abbildung E.7: Geglättete Remissionswerte für ein Paar normaler Hosen und ein Paar mit Aluminiumfolie überklebter Hosen vor einer Begegnung.

Remissionswerte für zwei unterschiedliche Paare Hosen vor einer Begegnung.

Es wird nun die Abhängigkeit der Remission von der Entfernung noch genauer untersucht. Zu diesem Zweck lief eine Person in unserem Labor auf einer vorgegebenen Bahn und wurde dabei mit einem Laserscanner beobachtet. Genauer bewegte sich der Proband in einer geraden Linie von dem Roboter weg und trug dabei jeweils unterschiedliche Hosen. Abbildung E.10 zeigt die Abhängigkeit der Remission von der Entfernung für eine braune Stoffhose. Die Graphik zeigt, dass die Remission oder die Intensität des registrierten Laserstrahles mit wachsender Entfernung abfällt. Außerdem sind die Daten wieder durch ein Rauschen gestört. Um den Einfluss dieses Rauschens zu mildern, wurden auch hier wieder zusätzlich die geglätteten Werte aufgetragen (Abbildung E.11). Die Abbildungen E.12 und E.13 zeigen die ursprünglichen bzw. die geglätteten Daten für eine blaue Stoffhose. Auch in diesem Falle nehmen die

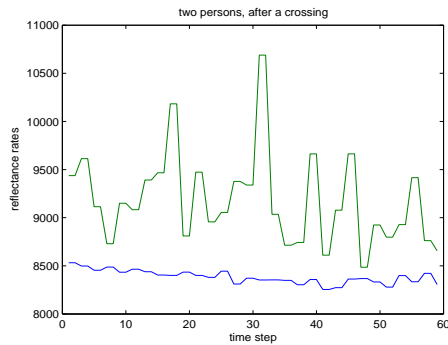


Abbildung E.8: Remissionswerte für ein Paar normaler Hosen und ein Paar mit Aluminiumfolie bedeckter Hosen nach einer Begegnung.

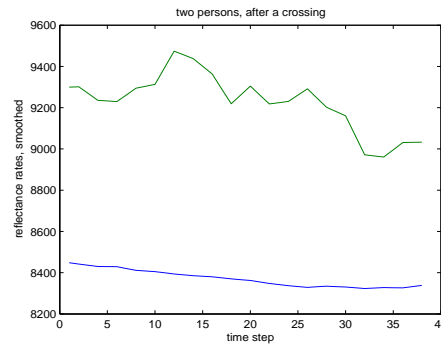


Abbildung E.9: Geglättete Remissionswerte für ein Paar normaler Hosen und ein Paar mit Aluminiumfolie abgedeckter Hosen nach einer Begegnung.

Remissionswerte für zwei unterschiedliche Paare Hosen nach einer Begegnung.

Intensitäten mit wachsender Entfernung ab. Die Abbildungen E.14 und E.15 schließlich zeigen die Verhältnisse für eine gelbe Kordhose.

Als vorläufiges Resultat dieser Versuche kann festgehalten werden, dass zwischen den drei Hosen, die zwar von verschiedener Farbe aber aus einem ähnlichen Material sind, nur geringe Unterschiede im Reflektionsverhalten bestehen, während jeweils eine deutliche Abhängigkeit von der Entfernung vorliegt. Deshalb wurde auch noch eine Hose betrachtet, die aus einem unterschiedlichen Material besteht. Dazu wurde wie oben eine blaue Jeans herangezogen. Es zeigte sich dabei, dass die Reflexionswerte für die blaue Jeans signifikant niedriger liegen als für die anderen drei Hosen (Abbildung E.16 und E.17).

Als Ergebnis kann festgehalten werden, dass die registrierten Intensitäten in erster Linie von der Entfernung der Person zum Beobachter abhängen, aber in manchen Fällen die Unterschiede, die durch die unterschiedlichen Hosenmaterialien hervorgerufen werden, groß genug sein sollten, um am Ende einer Interaktion eine fundierte Entscheidung darüber treffen zu können, welche Person zu welchem Subcluster gehört. Aufbauend auf diesen Voruntersuchungen konnte deshalb eine Methode entwickelt werden, die es innerhalb gewisser Grenzen erlaubt, Menschen anhand des Reflexionsverhaltens ihrer Beinbekleidung zu identifizieren. Diese Methode ist im Detail im Hauptteil dargestellt (Kapitel 9.3).



Abbildung E.10: Remissionswerte für ein Paar brauner Stoffhosen.

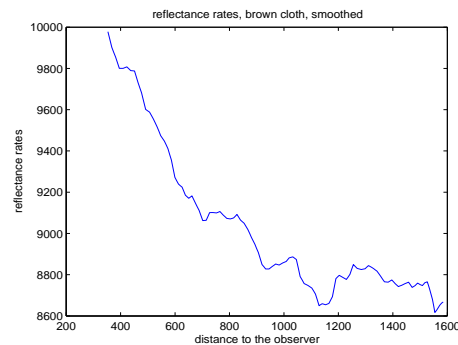


Abbildung E.11: Geglättete Remissionswerte für ein Paar brauner Stoffhosen.

Remissionswerte in Abhängigkeit von der Entfernung für ein Paar brauner Stoffhosen.

E.2 Graphiken zu einigen Experimenten

In diesem Anhang sind die Graphiken einiger Versuche zur Identifikation von Personen mit Hilfe der Laserremissionswerte (Kapitel 9.3) aufgenommen, auf deren Darstellung im Hauptteil verzichtet wurde. Zur genaueren Beschreibung dieser Experimente vergleiche auch wieder den Text dort.

Zunächst werden drei Versuche mit zwei Versuchspersonen dargestellt, bei denen die eine Person eine blaue Jeanshose und der andere Proband eine orange Stoffhose trug. Die drei Versuche unterscheiden sich dabei in der relativen Lage der jeweiligen Anfangs- und Endpositionen zueinander (Abbildung E.18 bis E.23).

Desweiteren wurden Versuche mit drei Personen durchgeführt. Dabei trug die eine Person wieder eine blaue Jeanshose und die zweite Versuchsperson auch wieder eine orange Stoffhose, während die dritte Person dagegen gar keine Hose trug, so dass die Laserstrahlen in diesem Fall direkt von der Haut des betreffenden Probanden reflektiert wurden (Abbildung E.24 bis E.33).

Schließlich wurden zwei Versuche durchgeführt, bei denen vier Personen mitwirkten. Dabei trugen diese Personen eine blaue Jeans, eine orange Stoffhose, eine hellbraune Stoffhose bzw. wieder gar keine Hose. Die beiden Stoffhosen waren sich dabei in ihrem Reflexionsverhalten sehr ähnlich. Die Zuordnungen waren deshalb für diese beiden Hosen teilweise nicht zeitlich konstant, so dass einmal vor und einmal nach der Interaktion in verschiedenen Zeitabschnitten beide mögliche Zuordnungen angezeigt wurden (Abbildung E.34 bis E.39).

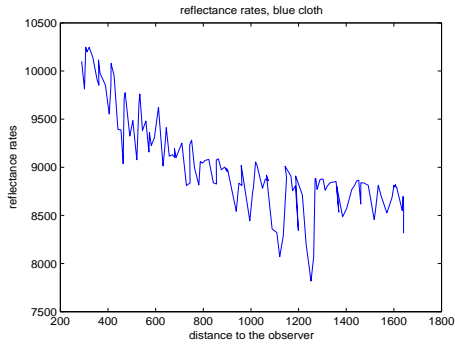


Abbildung E.12: Remissionswerte für ein Paar blauer Stoffhosen.

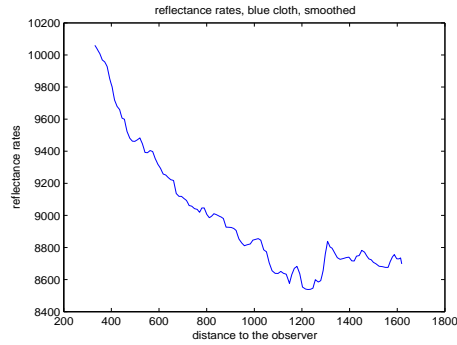


Abbildung E.13: Geglättete Remissionswerte für ein Paar blauer Stoffhosen.

Remissionswerte in Abhängigkeit von der Entfernung für ein Paar blauer Stoffhosen.

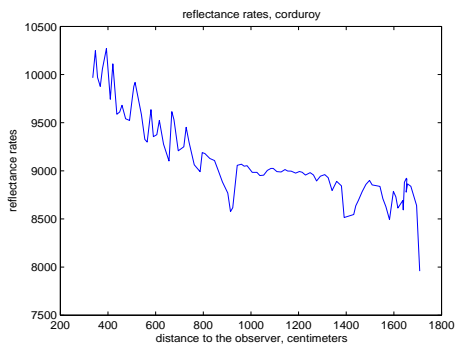


Abbildung E.14: Remissionswerte für ein Paar Kordhosen.

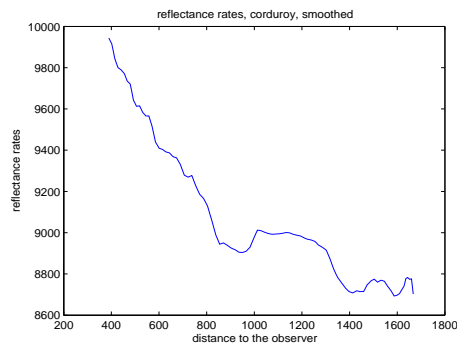


Abbildung E.15: Geglättete Remissionswerte für ein Paar Kordhosen.

Remissionswerte in Abhängigkeit von der Entfernung für ein Paar gelbe Kordhosen.

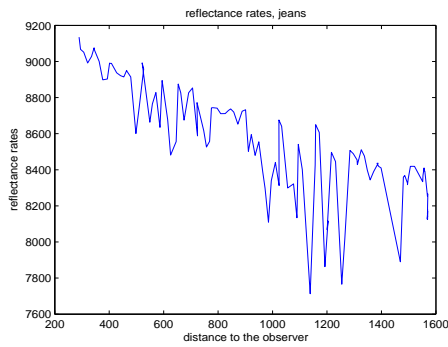


Abbildung E.16: Remissionswerte für ein Paar Blue Jeans.

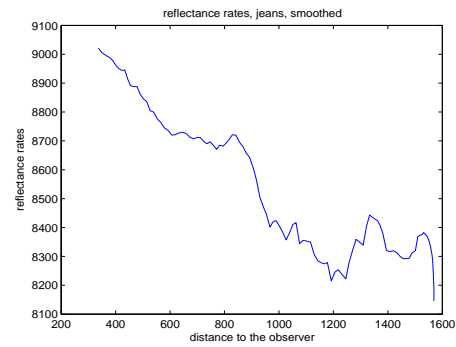


Abbildung E.17: Geglättete Remissionswerte für ein Paar Blue Jeans.

Remissionswerte in Abhängigkeit von der Entfernung für ein Paar Blue Jeans.

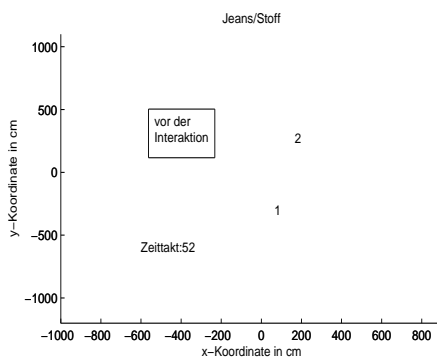


Abbildung E.18: Das Paar Jeans/Stoff vor der Interaktion.

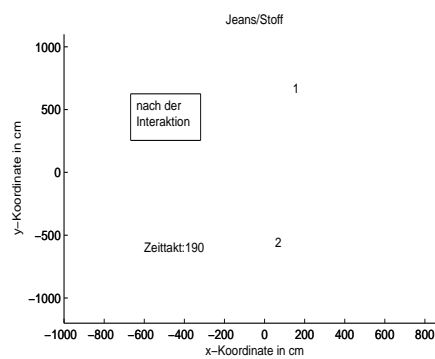


Abbildung E.19: Das Paar Jeans/Stoff nach der Interaktion.

Interaktion einer Jeans und einer Stoffhose, Permutation $12 \mapsto 21$.

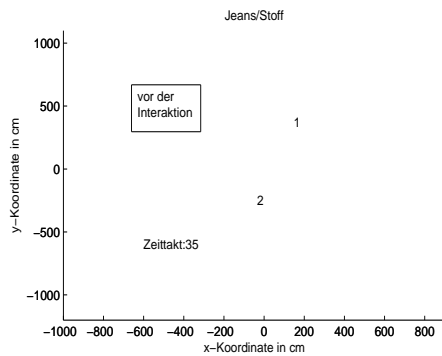


Abbildung E.20: Das Paar Jeans/Stoff vor der Interaktion.

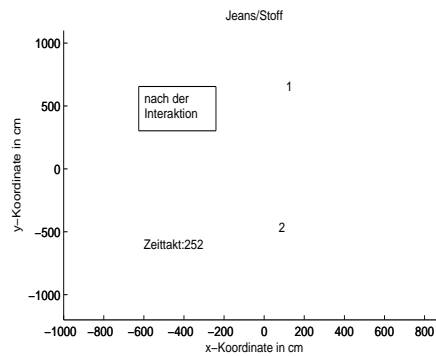


Abbildung E.21: Das Paar Jeans/Stoff nach der Interaktion.

Interaktion einer Jeans und einer Stoffhose, Permutation $21 \mapsto 21$.

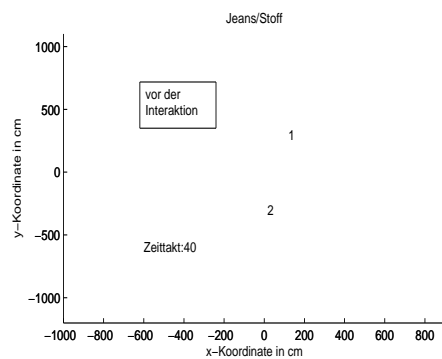


Abbildung E.22: Das Paar Jeans/Stoff vor der Interaktion.

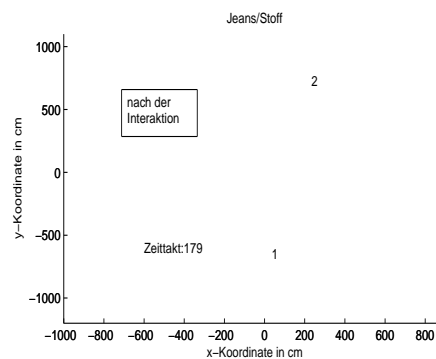


Abbildung E.23: Das Paar Jeans/Stoff nach der Interaktion.

Interaktion einer Jeans und einer Stoffhose, Permutation $21 \mapsto 12$.

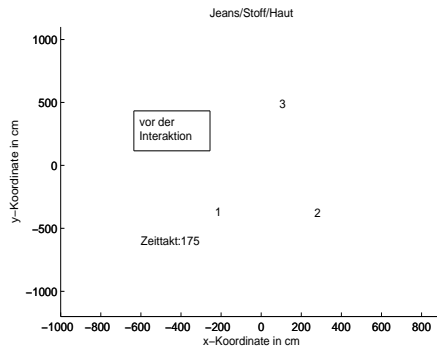


Abbildung E.24: Die Kombination Jeans/Stoff/Haut vor der Interaktion.

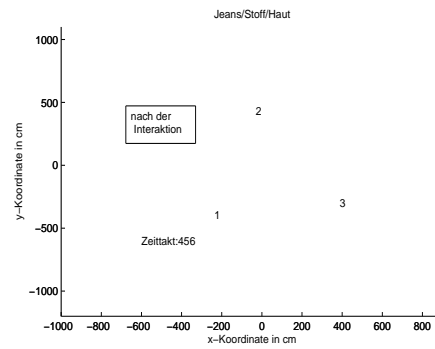


Abbildung E.25: Die Kombination Jeans/Stoff/Haut nach der Interaktion.

Interaktion einer Jeans, einer Stoffhose und einfacher Haut, Permutation
 $123 \mapsto 132$.

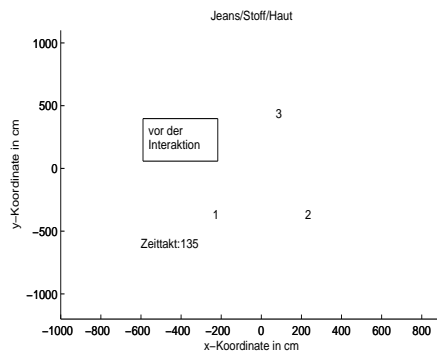


Abbildung E.26: Die Kombination Jeans/Stoff/Haut vor der Interaktion.

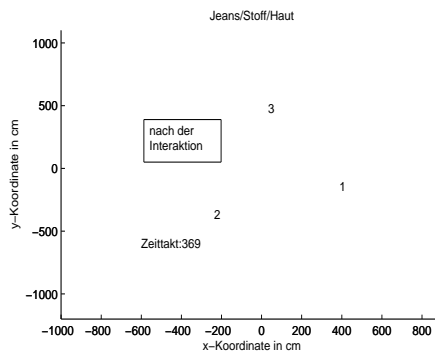


Abbildung E.27: Die Kombination Jeans/Stoff/Haut nach der Interaktion.

Interaktion einer Jeans, einer Stoffhose und einfacher Haut, Permutation
 $123 \mapsto 213$.

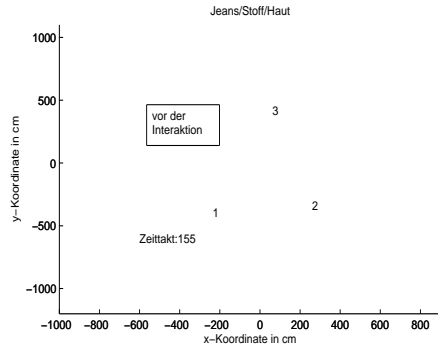


Abbildung E.28: Die Kombination Jeans/Stoff/Haut vor der Interaktion.

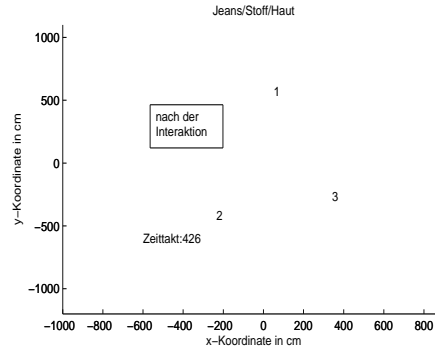


Abbildung E.29: Die Kombination Jeans/Stoff/Haut nach der Interaktion.

Interaktion einer Jeans, einer Stoffhose und einfacher Haut, Permutation
 123 \mapsto 231.

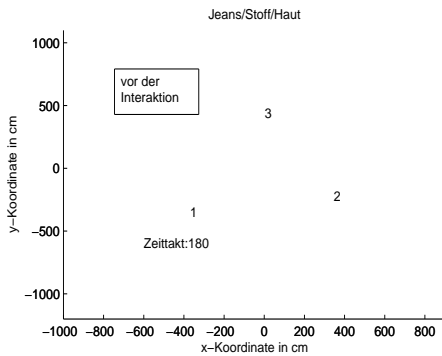


Abbildung E.30: Die Kombination Jeans/Stoff/Haut vor der Interaktion.

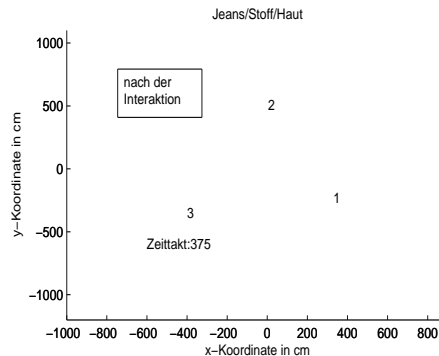


Abbildung E.31: Die Kombination Jeans/Stoff/Haut nach der Interaktion.

Interaktion einer Jeans, einer Stoffhose und einfacher Haut, Permutation
 123 \mapsto 312.

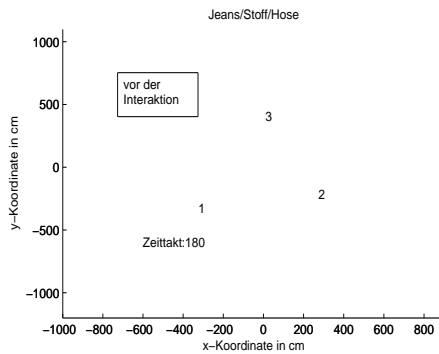


Abbildung E.32: Die Kombination Jeans/Stoff/Haut vor der Interaktion.

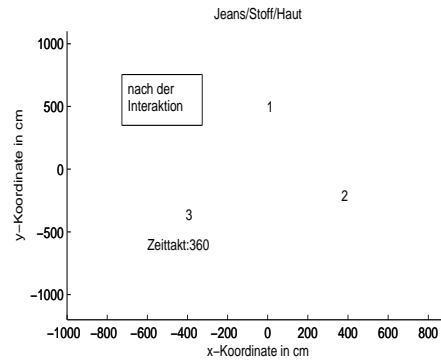


Abbildung E.33: Die Kombination Jeans/Stoff/Haut nach der Interaktion.

**Interaktion einer Jeans, einer Stoffhose und einfacher Haut, Permutation
123 \mapsto 321.**

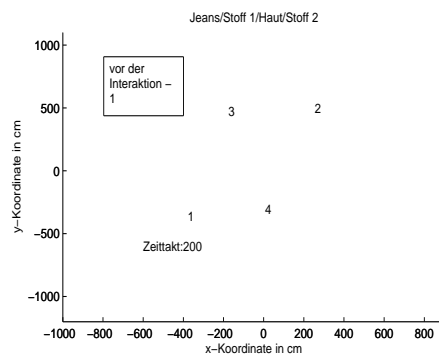


Abbildung E.34: Die Kombination Jeans / Stoff 1 / Haut / Stoff 2 vor der Interaktion — 1: falsche Zuordnung.

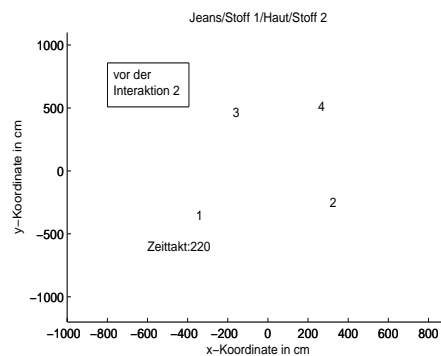


Abbildung E.35: Die Kombination Jeans / Stoff 1 / Haut / Stoff 2 vor der Interaktion — 2: richtige Zuordnung.

**Interaktion einer Jeans, zweier Stoffhosen und einfacher Haut, Permutation
1234 \mapsto 2143.**

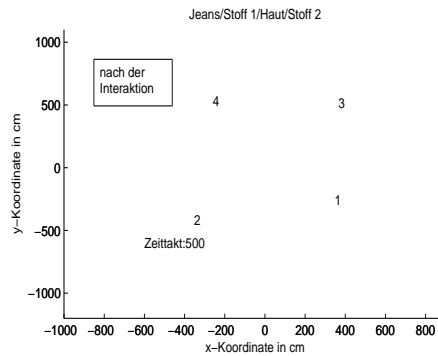


Abbildung E.36: Interaktion einer Jeans, zweier Stoffhosen und einfacher Haut, Permutation 1234 \mapsto 2143, nach der Interaktion.

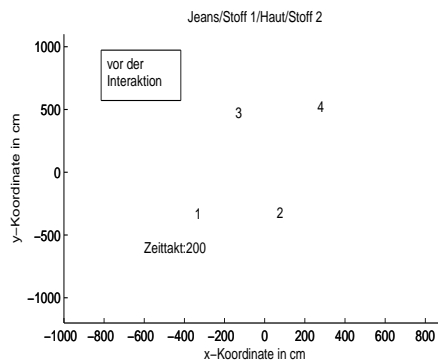


Abbildung E.37: Interaktion einer Jeans, zweier Stoffhosen und einfacher Haut, Permutation 1234 \mapsto 3124, vor der Interaktion.

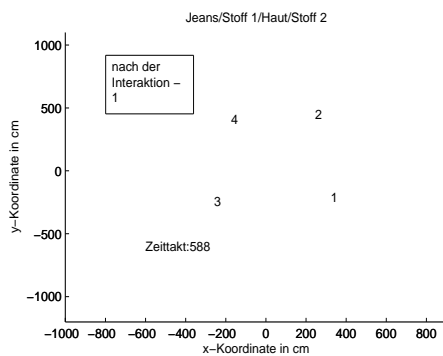


Abbildung E.38: Die Kombination Jeans / Stoff 1 / Haut / Stoff 2 nach der Interaktion — 1: falsche Zuordnung.

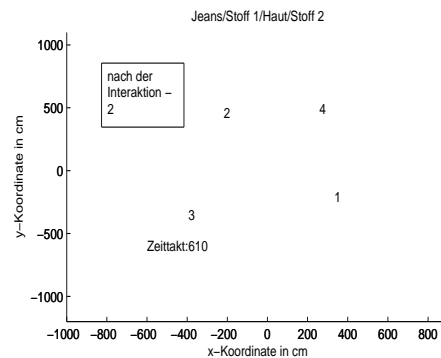


Abbildung E.39: Die Kombination Jeans / Stoff 1 / Haut / Stoff 2 nach der Interaktion — 2: richtige Zuordnung.

Interaktion einer Jeans, zweier Stoffhosen und einfacher Haut, Permutation 1234 \mapsto 3124.

Literaturverzeichnis

- Aggarwal, J. K. und Cai, Q. (1999). Human motion analysis: A review. *Computer Vision and Image Understanding: CVIU*, 73(3):428–440.
- Anderberg, M. R. (1973). *Cluster Analysis for Applications*. Monographs and Textbooks on Probability and Mathematical Statistics. Academic Press, Inc., New York, USA.
- Andreasson, H., Triebel, R., und Burgard, W. (2005). Improving plane extraction from 3D data by fusing laser data and vision. In *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005)*, Edmonton, Alberta, Kanada, Seite 2656–2661.
- Araneda, A. (2003). Statistical inference in mapping and localization for a mobile robot. In Bernardo, J. M., Bayarri, M. J., Berger, J. O., Dawid, A. P., Heckerman, D., Smith, A. F. M., und West, M., Herausgeber, *Bayesian Statistics 7*. Oxford University Press, Oxford, UK.
- Arras, K. O., Grzonka, S., Luber, M., und Burgard, W. (2008). Efficient people tracking in laser range data using a multi-hypothesis leg-tracker with adaptive occlusion probabilities. In *Proceedings of the 2008 IEEE International Conference on Robotics and Automation (ICRA '08)*, Pasadena, Kalifornien, USA, Seite 1710–1715.
- Arras, K. O. und Vestli, S. J. (1998). Hybrid, high-precision localization for the mail distributing mobile robot system MOPS. In *Proceedings of the 1998 IEEE International Conference on Robotics and Automation (ICRA '98)*, Leuven, Belgien, Seite 3129–3134. IEEE Computer Society.
- Asoh, H., Hayamizu, S., Hara, I., Motomura, Y., Akaho, S., und Matsui, T. (1997). Socially embedded learning of the office-conversant robot Jijo-2. In *Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence (IJCAI '97)*, Nagoya, Aichi, Japan, Seite 880–887.
- Bachran, T., Bongartz, H. H.-J., und Tiderko, A. (2005). A framework for multicast and quality based forwarding in MANETS. In *Proceedings of the 3rd IASTED International Conference on Communications and Computer Networks 2005 (CCN*

- 2005), *Marina del Rey, Kalifornien, USA*, Seite 120–125, Calgary, Kanada. ACTA Press.
- Backhaus, K., Erichson, B., Plinke, W., und Weiber, R. (2000). *Multivariate Analysemethoden*. Springer, Berlin Heidelberg New York.
- Ballard, D. H. und Brown, C. M. (1982). *Computer Vision*. Prentice–Hall, Inc., Englewood Cliffs, New Jersey, USA.
- Bar-Shalom, Y. (1974). Extension of the probabilistic data association filter to multi-target environments. In *Proceedings of the Fifth Symposium on Nonlinear Estimation, San Diego, Kalifornien, USA*.
- Bar-Shalom, Y. und Birimiwal, K. (1983). Consistency and robustness evaluation of the PDAF for target tracking in a cluttered environment. *Automatica*, 19:431–437.
- Bar-Shalom, Y. und Chen, H. (2007). Track-to-track association using attributes. *Journal of Advances in Information Fusion*, 2(1):49–59.
- Bar-Shalom, Y. und Fortmann, T. E. (1988). *Tracking and Data Association*, Band 179. Academic Press, Orlando, Florida, USA, Mathematics in Science and Engineering Edition.
- Bar-Shalom, Y., Fortmann, T. E., und Scheffe, M. (1980). Joint probabilistic data association for multiple targets in clutter. In *Proceedings of the Conference on Information Sciences and Systems, Princeton, New Jersey, USA*.
- Bar-Shalom, Y. und Jaffer, A. G. (1972). Adaptive nonlinear filtering for tracking with measurements of uncertain origin. In *Proceedings of the 11th IEEE Conference on Decision and Control, New Orleans, Louisiana, USA*, Seite 243–247.
- Bar-Shalom, Y. und Tse, E. (1975). Tracking in a cluttered environment with probabilistic data association. *Automatica, Pergamon Press*, 11:451–460.
- Becher, C., Foresti, G. L., Kaul, P., Koch, W., Lorenz, F. P., Lubczyk, D., Micheloni, C., Piciarelli, C., Safenreiter, K., Siering, C., Varela, M., Waldvogel, S. R., und Wienecke, M. (2008). A security assistance system combining person tracking with chemical attributes and video event analysis. In *Proceedings of the IEEE/ISIF — 11th International Conference on Information Fusion (FUSION 2008), Köln, Deutschland, CD-ROM*.
- Bellotto, N. (2008). Multisensor data fusion for simultaneous people tracking and recognition with a service robot. Vortrag am 12.9.2008, Universität Bonn.
- Bellotto, N. und Hu, H. (2006). Vision and laser data fusion for tracking people with a mobile robot. In *Proceedings of the 2006 IEEE International Conference on Robotics and Biomimetics (ROBIO 2006), Kunming, China*, Seite 7–12.

- Bellotto, N. und Hu, H. (2007a). Multisensor data fusion for joint people tracking and identification with a service robot. In *Proceedings of the 2007 IEEE International Conference on Robotics and Biomimetics (ROBIO 2007)*, Yalong Bay, Sanya, China, Seite 1494–1499.
- Bellotto, N. und Hu, H. (2007b). People tracking and identification with a mobile robot. In *Proceedings of the 2007 IEEE International Conference on Mechatronics and Automation (ICMA 2007)*, Harbin, Heilongjiang, China, Seite 3565–3570.
- Bellotto, N. und Hu, H. (2007c). People tracking with a mobile robot: a comparison of Kalman and particle filters. In Schilling, K., Herausgeber, *Proceedings of the 13th IASTED International Conference on Robotics and Applications, Würzburg, Deutschland*, Seite 388–393, Anaheim, Calgary, Zürich. The International Association of Science and Technology for Development, ACTA Press.
- Bennewitz, M., Burgard, W., Cielniak, G., und Thrun, S. (2005). Learning motion patterns of people for compliant robot motion. *The International Journal of Robotics Research*, 24(1):31–48.
- Bennewitz, M., Burgard, W., und Thrun, S. (2002a). Learning motion patterns of persons for mobile service robots. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '02)*, Washington, DC, USA, Seite 3601–3606.
- Bennewitz, M., Burgard, W., und Thrun, S. (2002b). Using EM to learn motion behaviours of persons with mobile robots. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '02)*, Lausanne, Schweiz.
- Bennewitz, M., Cielniak, G., und Burgard, W. (2003). Utilizing learned motion patterns to robustly track persons. In *Proceedings of the First Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS '03)*, Nizza, Frankreich, Seite 102–109.
- Betke, M. und Gurvits, L. (1997). Mobile robot localization using landmarks. *IEEE Transactions on Robotics and Automation*, 13(2):251–263.
- Beymer, D. und Konolige, K. (1999). Real-time tracking of multiple people using continuous detection. In *Proceedings of the 7th IEEE International Conference on Computer Vision (ICCV '99)*, Kerkyra, Korfu, Griechenland.
- Beymer, D. und Konolige, K. (2001). Tracking people from a mobile platform. In *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence (IJCAI '01)*, Workshop on Reasoning with Uncertainty in Robotics, Seattle, Washington, USA,
Online: <http://www.aass.oru.se//Agora/RUR01/proceedings.html>.

- Bishop, C. (1994). Novelty detection and neural network validation. *IEEE Proceedings Vision, Image and Signal Processing*, 141(4):217–222.
- Blackman, S. (1991). Persönliche Mitteilung. In Hall, D. L., Herausgeber, *Mathematical Techniques in Multisensor Data Fusion*, Seite 82. Artech House, Boston, London.
- Borenstein, J., Everett, B., und Feng, L. (1996). *Navigating mobile robots: Systems and Techniques*. A. K. Peters, Ltd.
- Bratko, I. (2001). *Prolog Programming for Artificial Intelligence*. Pearson Education, Addison–Wesley, Harlow, England.
- Broggi, A. (2008). Towards the perfect pedestrian detection system. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'08), Workshop on Planning, Perception and Navigation for Intelligent Vehicles, Nizza, Frankreich, Eingeladener Vortrag.
- Bronstein, I. N. und Semendjajew, K. A. (1987). *Taschenbuch der Mathematik*. Verlag Harri Deutsch, Thun und Frankfurt/Main.
- Brooks, A. und Williams, S. (2003). Tracking people with networks of heterogeneous sensors. In *Proceedings of the Australasian Conference on Robotics and Automation (ACRA '03), Brisbane, Australien*, Seite 1–7.
- Brcsic, D. und Hashimoto, H. (2007). Map building and object tracking inside intelligent spaces using static and mobile sensors. In *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007), San Diego, Kalifornien, USA*, Seite 3751–3756.
- Bui, H., Venkatesh, S., und West, G. (2001). Tracking and surveillance in wide–area spatial environments using the Abstract Hidden Markov Model. *International Journal of Pattern Recognition and Artificial Intelligence*, 15(1):177–195.
- Burgard, W., Cremers, A. B., Fox, D., Hähnel, D., Lakemeyer, G., Schulz, D., Steiner, W., und Thrun, S. (1998). The interactive museum tour–guide robot. In *Proceedings of the 15th National Conference on Artificial Intelligence (AAAI–98), Madison, Wisconsin, USA*, Seite 11–18.
- Burgard, W., Cremers, A. B., Fox, D., Hähnel, D., Lakemeyer, G., Schulz, D., Steiner, W., und Thrun, S. (2000). Experiences with an interactive museum tour–guide robot. *Artificial Intelligence (AI)*, 114(1–2):3–55.
- Burgard, W., Fox, D., Hennig, D., und Schmidt, T. (1996). Estimating the absolute position of a mobile robot using position probability grids. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence (AAAI–96), Portland, Oregon, USA*, Seite 896–901. AAAI Press/MIT Press.

- Burgard, W., Moors, M., Stachniss, C., und Schneider, F. (2005). Coordinated multi-robot exploration. *IEEE Transactions on Robotics*, 21(3):376–386.
- Burgard, W., Trahanias, P., Hähnel, D., Moors, M., Schulz, D., Baltzakis, H., und Argyros, A. A. (2002). TOURBOT and WebFAIR: Web-operated mobile robots for Tele-Presence in populated exhibitions. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '02), Workshop on 'Robots in Exhibitions'*, Lausanne, Schweiz.
- Burger, W. und Burge, M. J. (2005). *Digitale Bildverarbeitung*. Springer, Berlin Heidelberg.
- Calisi, D., Iocchi, L., und Leone, G. R. (2007). Person following through appearance models and stereo vision using a mobile robot. In *2nd International Conference on Computer Vision Theory and Applications (VISAPP 2007), The First International Workshop on Robot Vision, Barcelona, Spanien*, Seite 46–56.
- Chakravarty, P. und Jarvis, R. (2006). Panoramic vision and laser range finder fusion for multiple person tracking. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2006), Peking, China*, Seite 2949–2954.
- Cho, J. U., Jin, S. H., Pham, X. D., und Jeon, J. W. (2007). Multiple objects tracking circuit using particle filters with multiple features. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2007), Rom, Italien*, Seite 4639–4644.
- Choo, K. und Fleet, D. J. (2001). People tracking using hybrid Monte-Carlo filtering. In *Proceedings of the Eighth IEEE International Conference on Computer Vision (ICCV 2001), Vancouver, British Columbia, Kanada*, Band 2, Seite 321–328.
- Choset, H., Lynch, K. M., Hutchinson, S., Kantor, G., Burgard, W., Kavraki, L. E., und Thrun, S. (2005). *Principles of Robot Motion: Theory, Algorithms and Implementation*. The MIT Press, Cambridge, Massachusetts und London, England, Intelligent Robotics and Autonomous Agents Edition.
- Cielniak, G. und Duckett, T. (2003). Person identification by mobile robots in indoor environments. In *Proceedings of the First IEEE International Workshop on Robotic Sensing 2003 (ROSE '03), Örebro, Schweden*, Seite 1–5.
- Cielniak, G., Duckett, T., und Lilienthal, A. J. (2007). Improved data association and occlusion handling for vision-based people tracking by mobile robots. In *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007), San Diego, Kalifornien, USA*, Seite 3436–3441.

- Cox, I. (1991). Blanche — an experiment in guidance and navigation of an autonomous robot vehicle. *IEEE Transactions on Robotics and Automation*, 7(2).
- Cox, I. J. (1993). A review of statistical data association techniques for motion correspondence. *International Journal of Computer Vision*, 10(1):53–66.
- Cox, I. J. und Hingorani, S. L. (1996). An efficient implementation of Reids multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 18(2):138–150.
- Crowley, J. L. (1989). Control of translation and rotation in a robot vehicle. In *Proceedings of the 1989 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '89)*, Tsukuba, Japan.
- Darrell, T., Moghaddam, B., und Pentland, A. P. (1996). Active face tracking and pose estimation in an interactive room. In *Proceedings of the IEEE 6th International Conference on Computer Vision and Pattern Recognition (CVPR '96)*, San Francisco, Kalifornien, USA, Seite 67–72, Washington, DC, USA. IEEE Computer Society.
- Dellaert, F., Fox, D., Burgard, W., und Thrun, S. (1999). Monte Carlo localization for mobile robots. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '99)*, Detroit, Michigan, USA, Seite 1322–1328.
- Derenick, J. C., Spletzer, J. R., und Hsieh, M. A. (2007). A graph theoretic approach to optimal target tracking for mobile robot teams. In *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007)*, San Diego, Kalifornien, USA, Seite 4222–4228.
- Dezert, J. (1998). Tracking maneuvering and bending extended targets in a cluttered environment. In *SPIE Proceedings Series, Conference on Signal and Data Processing of Small Targets*, Orlando, Florida, USA, Band 3373, Seite 283–294, Bellingham, Washington, USA. Society of Photo–Optical Instrumentation Engineers.
- Dissanayake, G. M. W. M., Newman, P., Clark, S., Durrant-Whyte, H., und Csorba, M. (2001). A solution to the simultaneous localization and map building (SLAM) problem. *IEEE Transactions on Robotics and Automation*, 17(3):229–241.
- Douillard, B., Fox, D., und Ramos, F. (2008). Laser and vision based outdoor object mapping. In *Proceedings of Robotics: Science and Systems IV*, Zürich, Schweiz, Online: <http://www.roboticsproceedings.org/rss04/p2.html>.
- Durrant-Whyte, H., Majumder, S., Thrun, S., de Battista, M., und Scheduling, S. (2003). A Bayesian algorithm for simultaneous localization and map building. In Jarvis,

- R. A. und Zelinsky, A., Herausgeber, *Robotics Research, The Tenth International Symposium*, Band 6, Seite 49–60. Springer, Berlin Heidelberg, Springer Tracts in Advanced Robotics Edition.
- Durrant-Whyte, H. und Stevens, M. (2001). Data fusion in decentralized sensing networks. In *Proceedings of the IEEE/ISIF — 4th International Conference on Information Fusion (Fusion 2001), Montreal, Quebec, Kanada*, Seite 302–307.
- Durrant-Whyte, H. F. (1996). An autonomous guided vehicle for cargo handling applications. *International Journal of Robotics Research*, 15(5):407–440.
- Ekstrand, B. (1985). Analytical steady state solution for a continuous time Kalman filter. *IEEE Transactions on Aerospace and Electronic Systems*, 21(6):746–750.
- Elfes, A. (1987). Sonar-based real-world mapping and navigation. *IEEE Transactions on Robotics and Automation*, 3(3):249–265.
- Endres, H., Feiten, W., und Lawitzky, G. (1998). Field test of a navigation system: Autonomous cleaning in supermarkets. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '98), Leuven, Belgien*, Seite 1779–1781. IEEE Computer Society.
- Engelberger, J. F. (1993). Health-care robotics goes commercial: The HelpMate experience. *Robotica*, 11:517–523.
- Erding, O., Willett, P., und Coraluppi, S. (2008). Multistatic sensor placement: a tracking approach. *Journal of Advances in Information Fusion*, 2(1):22–34.
- Fischler, M. A. und Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the Association for Computing Machinery*, 24(6):381–395.
- Fod, A., Howard, A., und Mataric, M. J. (2002). Laser-based people tracking. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '02), Washington, DC, USA*, Seite 3024–3029.
- Forney Jr., G.-D. (1973). The Viterbi algorithm. *Proceedings of the IEEE*, 61(3):268–278.
- Fortmann, T. E., Bar-Shalom, Y., und Scheffe, M. (1980). Multi-target tracking using joint probabilistic data association. In *Proceedings of the 19th IEEE Conference on Decision and Control, Albuquerque, New Mexico, USA*, Seite 807–812.
- Fortmann, T. E., Bar-Shalom, Y., und Scheffe, M. (1983). Sonar tracking of multiple targets using joint probabilistic data association. *IEEE Journal of Oceanic Engineering*, OE-8(3):173–184.

- Fox, D., Burgard, W., Dellaert, F., und Thrun, S. (1999a). Monte Carlo localization: efficient position estimation for mobile robots. In *Proceedings of the Sixteenth National Conference on Artificial Intelligence (AAAI-99), Orlando, Florida, USA*, Seite 343–349.
- Fox, D., Burgard, W., und Thrun, S. (1999b). Markov localization for mobile robots in dynamic environments. *Journal of Artificial Intelligence Research (JAIR)*, 11:391–427.
- Fox, D., Burgard, W., Thrun, S., und Cremers, A. B. (1998). Position estimation for mobile robots in dynamic environments. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI-98), Madison, Wisconsin, USA*, Seite 983–988.
- Frank, O., Nieto, J., Guivant, J., und Scheduling, S. (2003). Multiple target tracking using sequential Monte Carlo methods and statistical data association. In *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003), Las Vegas, Nevada, USA*, Band 3, Seite 2718–2723.
- Frintrop, S., Rome, E., Nüchter, A., und Surmann, H. (2003). An attentive, multi-modal laser 'eye'. In Crowley, J., Piater, J. H., Vincze, M., und Paletta, L., Herausgeber, *Proceedings of the 3rd International Conference on Computer Vision Systems (ICVS 2003), Graz, Österreich*, Lecture Notes in Computer Science (LNCS), Seite 202–211, Berlin. Springer.
- Fuerstenberg, K. C., Dietmayer, K. C. J., und Willhoeft, V. (2002). Pedestrian recognition in urban traffic using a vehicle based multilayer laserscanner. In *Proceedings of the IEEE Intelligent Vehicles Symposium (IV 2002), Versailles, Frankreich*, Band 1, Seite 31–35.
- Furukawa, T., Durrant-Whyte, H. F., und Lavis, B. (2007). The element-based method — theory and its application to Bayesian search and tracking —. In *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007), San Diego, Kalifornien, USA*, Seite 2807–2812.
- Garcia, J. G., Robertsson, A., Ortega, J. G., und Johansson, R. (2008). Sensor fusion for compliant robot motion control. *IEEE Transactions on Robotics*, 24(2):430–441.
- Gelb, A. (1974). *Applied Optimal Estimation*. MIT Press, Cambridge, Massachusetts, USA.
- Gelman, A., Carlin, J. B., und Stern, H. S. (2003). *Bayesian Data Analysis*. Chapman and Hall, Boca Raton, Florida, USA.

- Ghahramani, Z. (2001). An introduction to Hidden Markov Models and Bayesian networks. *International Journal of Pattern Recognition and Artificial Intelligence*, 15(1):9–42.
- Gonzalez, R. C. und Woods, R. E. (1992). *Digital Image Processing*. Addison–Wesley.
- Gonzalez-Banos, H. H., Lee, C. Y., und Latombe, J. C. (2002). Real–time combinatorial tracking of a target moving unpredictably among obstacles. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2002)*, Washington, DC, USA, Seite 1683–1690.
- Gordon, N., Salmond, D., und Smith, A. (1993). A novel approach to nonlinear/non–Gaussian Bayesian state estimation. *IEE Proceedings F: Radar and Signal Processing*, 140(2):107–113.
- Gordon, N. J., Maskell, S., und Kirubarajan, T. (2002). Efficient particle filters for joint tracking and classification. In Drummond, O. E., Herausgeber, *SPIE Proceedings Series, Conference on Signal and Data Processing of Small Targets, Orlando, Florida, USA*, Band 4728, Seite 439–449, Bellingham, Washington, USA. Society of Photo–Optical Instrumentation Engineers.
- Gruben, C. (1993). *Teilchendetektoren*. BI Wissenschaftsverlag, Mannheim.
- Guivant, J. und Nebot, E. (2003). Implementation of simultaneous navigation and mapping in large outdoor environments. In Jarvis, R. A. und Zelinsky, A., Herausgeber, *Robotics Research, The Tenth International Symposium*, Band 6, Seite 37–48. Springer, Berlin Heidelberg, Springer Tracts in Advanced Robotics Edition.
- Guivant, J., Nebot, E., und Baiker, S. (2000a). Autonomous navigation and map building using laser range sensors in outdoor applications. *Journal of Robotic Systems*, 17(10):265–283.
- Guivant, J., Nebot, E., und Durrant-Whyte, H. (2000b). Simultaneous localization and map building using natural features in outdoor environments. In *Proceedings of the Sixth IEEE International Conference on Intelligent Autonomous Systems (IAS-6)*, Venedig, Italien, Seite 581–588.
- Guivant, J., Nieto, J., Masson, F., und Nebot, E. (2004). Navigation and mapping in large unstructured environments. *International Journal of Robotics Research*, 23(4–5):449–472.
- Guivant, L. J. (2008). The identity management problem — a short survey. In *Proceedings of the IEEE/ISIF — 11th International Conference on Information Fusion (FUSION 2008)*, Köln, Deutschland, CD–ROM.

- Gutmann, J.-S., Burgard, W., Fox, D., und Konolige, K. (1998). An experimental comparison of localization methods. In *Proceedings of the 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-98)*, Victoria, British Columbia, Kanada.
- Gutmann, J.-S. und Fox, D. (2002). An experimental comparison of localization methods continued. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-02)*, Lausanne, Schweiz, Band 1, Seite 454–459.
- Hähnel, D., Schulz, D., und Burgard, W. (2002). Map building with mobile robots in populated environments. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-02)*, Lausanne, Schweiz, Band 1, Seite 496–501.
- Hall, D. L. (1992). *Mathematical Techniques in Multisensor Data Fusion*. Artech House, Boston London.
- Hall, D. L. und Llinas, J., Herausgeber (2001). *Handbook of Multisensor Data Fusion*. CRC Press, Boca Raton London New York Washington D.C.
- Han, J. und Kamber, M. (2001). *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers, San Francisco San Diego New York Boston London Sydney Tokio.
- Hand, D., Mannila, H., und Smyth, P. (2001). *Principles of Data Mining*. The MIT Press, Cambridge, Massachusetts und London, UK.
- Handschin, J. (1970). Monte Carlo techniques for prediction and filtering of non-linear stochastic processes. *Automatica*, 6:555–563.
- Hastie, T., Tibshirani, R., und Friedman, J. (2001). *The elements of statistical learning: Data mining, Inference, and Prediction*. Springer Series in Statistics. Springer, New York, USA.
- Hightower, J. und Borriello, G. (2001). Location systems for ubiquitous computing. *IEEE Computer*, 34(8):57–66.
- Hoeller, F., Schulz, D., Moors, M., und Schneider, F. E. (2007). Accompanying persons with a mobile robot using motion prediction and probabilistic roadmaps. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007)*, San Diego, Kalifornien, USA, Seite 1260–1265.
- Hoeller, F., Röhling, T., und Königs, A. (2008). Combining coordinated navigation and reactive collision avoidance for GPS-based convoying. In Ramamoorthy, S. und Hayes, G. M., Herausgeber, *Proceedings of Towards Autonomous Robotic Systems 2008 (TAROS 2008)*, Edinburgh, UK, Seite 93–100.

- Hofmann, T. und Buhmann, J. (1997). Pairwise data clustering by deterministic annealing. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 19(1):1–14.
- Hofstadter, D. (1979). *Godel, Escher, Bach: An external golden Braid*. Basic Books, New York, USA.
- Höller, F. (2006). Personenbegleitung mit mobilen Robotern: Lokale Navigation mit probabilistischen Roadmapverfahren. Diplomarbeit, Institut für Informatik III, Rheinische Friedrich–Wilhelms–Universität Bonn, Römerstraße 164, 53117 Bonn.
- Horswill, I. (1993). Polly: A vision–based artificial agent. In *Proceedings of the Eleventh National Conference on Artificial Intelligence (AAAI 1993), Washington, DC, USA*, Seite 824–829.
- Huang, C.-M., Lai, C.-W., und Fu, L.-C. (2007). Real–time multitarget visual tracking with an active camera. In *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007), San Diego, Kalifornien, USA*, Seite 2741–2746.
- Ikedo, T., Ishiguro, H., und Nishimura, T. (2007). People tracking by cross modal association of vision sensors and acceleration sensors. In *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007), San Diego, Kalifornien, USA*, Seite 4147–4151.
- Illmann, J., Kluge, B., und Prassler, E. (2002). Statistical recognition of motion patterns. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2002), Lausanne, Schweiz*, Seite 1259–1264.
- Isard, M. und Blake, A. (1996). Contour tracking by stochastic propagation of conditional density. In Buxton, B. und Cipolla, R., Herausgeber, *Proceedings of the Fourth European Conference of Computer Vision (ECCV '96), Band 1, Cambridge, UK*, Band 1064 der *Lecture Notes in Computer Science*, Seite 343–356, Berlin Heidelberg. Springer.
- Isard, M. und Blake, A. (1998). Condensation–conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29(1):5–28.
- Jung, B. und Sukhatme, G. S. (2004). A generalized region–based approach for multi–target tracking in outdoor environments. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2004), New Orleans, Louisiana, USA*, Seite 2189–2195.
- Kaelbling, L., Cassandra, A., und Kurien, J. (1996). Acting under uncertainty: discrete Bayesian models for mobile–robot navigation. In *Proceedings of the IEEE/RSJ*

- International Conference on Intelligent Robots and Systems (IROS '96), Osaka, Japan, Seite 963–972.*
- Kahn, R. E., Swain, M., Prokopowicz, P. N., und Firby, R. J. (1996). Gesture recognition using the Perseus architecture. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '96), San Francisco, Kalifornien, USA, Seite 734–741.*
- Kalata, P. R. (1984). The tracking index: A generalized parameter for α - β and α - β - γ target trackers. *IEEE Transactions on Aerospace and Electronic Systems*, AES-20(2):174–182.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the American Society of Mechanical Engineers, Series D, Journal of Basic Engineering*, 82:34–45.
- Kanazawa, K., Koller, D., und Russell, S. (1995). Stochastic simulation algorithms for dynamic probabilistic networks. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence (UAI-95), Montreal, Quebec, Kanada, Seite 346–351.*
- Khan, Z., Balch, T., und Dellaert, F. (2003). Efficient particle filter-based tracking of multiple interacting targets using an MRF-based motion model. In *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '03), Las Vegas, Nevada, USA.*
- Kitagawa, G. (1996). Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics*, 5:1–25.
- Klein, R. (2005). *Algorithmische Geometrie*. Springer, Heidelberg.
- Kleinknecht, K. (2005). *Detektoren für Teilchenstrahlung*. Teubner, München.
- Klix, F., Herausgeber (1979). *Human and Artificial Intelligence*. North-Holland Publishing Company, Amsterdam, Niederlande.
- Kluge, B., Koehler, C., und Prassler, E. (2001). Fast and robust tracking of multiple moving objects with a laser range finder. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2001), Seoul, Korea, Seite 1683–1688.*
- Koch, K. R. (2000). *Einführung in die Bayes-Statistik*. Springer, Berlin Heidelberg New York.
- Koch, W. (2000). Target tracking. In Stergiopoulos, S., Herausgeber, *Advanced Signal Processing Handbook: Theory and Implementation for Radar, Sonar and Medical Imaging Real Time Systems*. CRC Press Inc., Boca Raton, Florida, USA, Electrical Engineering & Signal Processing Series Edition.

- Koch, W. (2006). On Bayesian tracking of extended objects. In *Proceedings of the 2006 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2006), Heidelberg, Deutschland*, Seite 209–216.
- Koch, W. und Saul, R. (2005). A Bayesian approach to extended object tracking and tracking of loosely structured target groups. In *Proceedings of the IEEE/ISIF — Eighth International Conference on Information Fusion (FUSION 2005), Philadelphia, Pennsylvania, USA, CD-ROM*.
- Koch, W. und Stannus, W. (2003). A new application of the EM algorithm: Robot-borne extended object tracking. Wissenschaftlicher Bericht 66, SDF/FKIE/FGAN, Neuenahrer Straße 20, 53343 Wachtberg, Deutschland.
- Koch, W. und van Keuk, G. (1997). Multiple hypothesis track maintenance with possibly unresolved measurements. *IEEE Transactions on Aerospace and Electronic Systems*, 33(3):883–892.
- Koller, D. und Fratkina, R. (1998). Using learning for approximation in stochastic processes. In *Proceedings of the Fifteenth International Conference on Machine Learning (ICML '98), Madison, Wisconsin, USA*, Seite 287–295.
- Koller-Meier, E. und Ade, F. (2001). Tracking multiple objects using the condensation algorithm. *Journal of Robotics and Autonomous Systems*, 34(2–3):93–105.
- Kortenkamp, D., Bonasso, R., und Murphy, R. (1997). *AI-based mobile robots: case studies of successful robot systems*. The MIT Press, Cambridge, Massachusetts, USA.
- Kortenkamp, D., Huber, E., und Bonasso, R. P. (1996). Recognizing and interpreting gestures on a mobile robot. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence (AAAI-98), Portland, Oregon, USA*, Band 2, Seite 915–921. AAAI Press/The MIT Press.
- Kräußling, A. (2006a). Tracking extended moving objects with a mobile robot (full paper). In *Proceedings of the 3rd IEEE Conference on Intelligent Systems (IEEE-IS 06), London, UK, CD-ROM*.
- Kräußling, A. (2006b). Tracking multiple objects using the Viterbi algorithm (full paper). In *Proceedings of the 3rd International Conference on Informatics in Control, Automation and Robotics (ICINCO 06), Conference Area Robotics and Automation, Setubal, Portugal*, Seite 18–25.
- Kräußling, A. (2007a). Die Bayes-Statistik — Grundlagen und Anwendungen in der mobilen Robotik. Wissenschaftlicher Bericht 140, EFS/FKIE/FGAN, Neuenahrer Straße 20, 53343 Wachtberg, Deutschland.

- Kräußling, A. (2007b). A novel approach to the mobile robot localization problem using tracking methods. In Schilling, K., Herausgeber, *Proceedings of the 13th IASTED International Conference on Robotics and Applications, Würzburg, Deutschland*, Seite 107–112, Anaheim, Calgary, Zürich. The International Association of Science and Technology for Development, ACTA Press.
- Kräußling, A. (2008a). A switching algorithm for people tracking with mobile robot systems. In Schmidt, L., Schlick, C. M., und Grosche, J., Herausgeber, *Ergonomie und Mensch–Maschine–Systeme*, Seite 331–352. Springer, Heidelberg.
- Kräußling, A. (2008b). Tracking extended moving objects with a mobile robot. In Chountas, P., Petrounias, I., und Kacprzyk, J., Herausgeber, *Intelligent Techniques and Tools for Novel System Architectures*, Band 109, Seite 513–530. Springer, Heidelberg, Studies in Computational Intelligence Edition.
- Kräußling, A. (2008c). Tracking multiple objects using the Viterbi algorithm. In Cetto, J. A., Ferrier, J.-L., Pereira, J. M. C. D., und Filipe, J., Herausgeber, *Informatics in Control Automation and Robotics, Selected Papers from the International Conference on Informatics in Control Automation and Robotics 2006*, Band 15, Seite 233–247. Springer, Heidelberg, Lecture Notes in Electrical Engineering Edition.
- Kräußling, A., Brüggemann, B., und Schulz, D. (2010). People tracking and identification using laser features and colour distributions. *Zur Publikation angenommen im Springer–Buch ausgewählter Veröffentlichungen der Konferenz 'International Conference on Informatics in Control Automation and Robotics 2008 (ICINCO 2008)'*. Springer, Heidelberg.
- Kräußling, A., Brüggemann, B., Schulz, D., und Cremers, A. B. (2008). People tracking using laser range scanners and vision (full paper). In *Proceedings of the 5th International Conference on Informatics in Control, Automation and Robotics (ICINCO 08), Conference Area Robotics and Automation, Funchal, Madeira, Portugal*, Band 1, Seite 29–36.
- Kräußling, A., Schneider, F. E., und Wildermuth, D. (2004a). Tracking expanded objects using the Viterbi algorithm. In *Proceedings of the IEEE Conference on Intelligent Systems (IEEE–IS 04), Varna, Bulgarien*.
- Kräußling, A., Schneider, F. E., und Wildermuth, D. (2004b). Tracking of extended crossing objects using the Viterbi algorithm (full paper). In *Proceedings of the 1st International Conference on Informatics in Control, Automation and Robotics (ICINCO 04), Conference Area Robotics and Automation, Setubal, Portugal*, Seite 142–149.
- Kräußling, A., Schneider, F. E., und Wildermuth, D. (2005a). A switching algorithm for tracking extended targets (full paper). In *Proceedings of the 2nd International*

Conference on Informatics in Control, Automation and Robotics (ICINCO 05), Conference Area Robotics and Automation, Barcelona, Spanien, Seite 126–133.

- Kräußling, A., Schneider, F. E., und Wildermuth, D. (2005b). Zur Verfolgung ausgedehnter Ziele — eine Übersicht über ausgewählte Algorithmen und ein Vergleich deren Güte. Wissenschaftlicher Bericht 94, EFS/FKIE/FGAN, Neuenahrer Straße 20, 53343 Wachtberg, Deutschland.
- Kräußling, A., Schneider, F. E., Wildermuth, D., und Sehestedt, S. (2007). A switching algorithm for tracking extended targets. In Filipe, J., Ferrier, J.-L., Cetto, J. A., und Carvalho, M., Herausgeber, *Informatics in Control, Automation and Robotics*, Band 2, Seite 117–128. Springer, Heidelberg.
- Kräußling, A. und Schulz, D. (2006). Tracking extended targets — a switching algorithm versus the SJPDAF. In *Proceedings of the IEEE/ISIF — 9th International Conference on Information Fusion, Florenz, Italien, CD-ROM*.
- Kräußling, A. und Schulz, D. (2008). Data fusion for person identification. In *Proceedings of the IEEE/ISIF — 11th International Conference on Information Fusion (FUSION 2008), Köln, Deutschland, CD-ROM*.
- Kruse, E. und Wahl, F. M. (1998). Camera-based monitoring system for mobile robot guidance. In *Proceedings of the 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '98), Victoria, British Columbia, Kanada*, Band 2, Seite 1248–1253.
- Lamond, B. und Watson, G. (2004). Hybrid rendering — a new integration of photogrammetry and laser scanning for image based rendering. In *Proceedings of the Second Conference on Theory and Practice of Computer Graphics (TPCG 2004), Bournemouth, UK*, Seite 179–186.
- Lancaster, P. und Rodman, L. (1995). *Algebraic Riccati Equations*. Clarendon Press, Oxford, UK.
- Lavalle, S. M., Gonzalez-Banos, H. H., Becker, G., und Latombe, J.-C. (1997). Motion strategies for maintaining visibility of a moving target. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 1997), Albuquerque, New Mexico, USA*, Seite 731–736.
- Lebeltel, O., Bessiere, P., Diard, J., und Mazer, E. (2004). Bayesian robot programming. *Autonomous Robots*, 16(1):49–79.
- Leonard, J., Newman, P., Rikoski, R., Neira, J., und Tardos, J. (2003). Towards robust data association and feature modeling for concurrent mapping and localization. In Jarvis, R. A. und Zelinsky, A., Herausgeber, *Robotics Research, The Tenth International Symposium*, Band 6, Seite 7–20. Springer, Berlin Heidelberg, Springer Tracts in Advanced Robotics Edition.

- Leonard, J. J. und Durrant-Whyte, H. F. (1992). *Directed Sonar Sensing for Mobile Robot Navigation*. Kluwer Academic Publishers, Dordrecht, Niederlande.
- Lindström, M. und Eklundh, J.-O. (2001). Detecting and tracking moving objects from a mobile platform using a laser range scanner. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2001)*, Maui, Hawaii, USA, Band 3, Seite 1364–1369.
- Liu, H., Sun, F., Yu, L., und He, K. (2007). Vehicle tracking using stochastic fusion-based particle filter. In *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007)*, San Diego, Kalifornien, USA, Seite 2735–2740.
- Liu, L. und Stamos, I. (2005). Automatic 3D to 2D registration for the photorealistic rendering of urban scenes. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2005)*, San Diego, Kalifornien, USA, Band 2, Seite 137–143.
- Luerweg, F. (2007). Punktlandung mit dem Skalpell: Mathematik sorgt im OP für den treffsicheren Schnitt. In Archut, A., Herausgeber, *forsch: Bonner Universitäts-Nachrichten*, Band 4, Seite 20. Universität Bonn, Regina Pacis-Weg 3, 53313 Bonn.
- Luo, R. C. und Lin, M.-H. (1988). Robot multi-sensor fusion and integration: Optimum estimation of fused sensor data. In *Proceedings of the 1988 IEEE International Conference on Robotics and Automation (ICRA '88)*, Philadelphia, Pennsylvania, USA, Band 2, Seite 1076–1081. IEEE Computer Society Press.
- MacCormick, J. und Blake, A. (1999). A probabilistic exclusion principle for tracking multiple objects. In *Proceedings of the 7th International Conference on Computer Vision (ICCV '99)*, Kerkyra, Korfu, Griechenland, Seite 572–587.
- Mahalanobis, P. C. (1936). On the generalized distance in statistics. *Proceedings of the National Institute of Science of India*, 12:49–55.
- Makarenko, A. und Durrant-Whyte, H. (2004). Decentralized data fusion and control in active sensor networks. In *Proceedings of the IEEE/ISIF — Seventh International Conference on Information Fusion (FUSION 2004)*, Stockholm, Schweden, CD-ROM.
- Marsland, S., Nehmzow, U., und Shapiro, J. (2005). On-line novelty detection for autonomous mobile robots. *Robotics and Autonomous Systems*, 51(2–3):191–206.
- Maybeck, P. S. (1979). *Stochastic Models, Estimation and Control*. Academic Press, New York San Francisco London.

- Meier, E. und Ade, F. (1999). Using the condensation algorithm to implement tracking for mobile robots. In Schweitzer, G., Burgard, W., Nehmzow, U., und Vestli, S. J., Herausgeber, *Proceedings of the Third European Workshop on Advanced Mobile Robots (Eurobot '99)*, Zürich, Schweiz, Seite 73–80.
- Michalski, R. S., Bratko, I., und Kubat, M. (1999). *Machine Learning and Data Mining: Methods and Applications*. John Wiley and Sons, Inc., Hoboken, New Jersey, USA.
- Minvielle, P., Marrs, A., Maskell, S., und Doucet, A. (2005). Joint target tracking and identification — part II: Shape video computing. In *Proceedings of the IEEE/ISIF — Eighth International Conference on Information Fusion (FUSION 2005)*, Philadelphia, Pennsylvania, USA, CD-ROM.
- Mitchell, T. M. (1997). *Machine Learning*. McGraw–Hill Series in Computer Science. McGraw–Hill, The City of New York, New York, USA.
- Montemerlo, M. und Thrun, S. (2007). *FastSLAM, a Scalable Method for the Simultaneous Localization and Mapping Problem in Robotics*, Band 27 der Springer Tracts in Advanced Robotics. Springer, Berlin Heidelberg.
- Montemerlo, M., Thrun, S., Koller, D., und Wegbreit, B. (2002a). FastSLAM: A factored solution to the simultaneous localization and mapping problem. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence (AAAI-02)*, Edmonton, Alberta, Kanada, Seite 593–598.
- Montemerlo, M., Thrun, S., und Whittaker, W. (2002b). Conditional particle filters for simultaneous mobile robot localization and people tracking. In *Proceedings of the 2002 IEEE International Conference on Robotics and Automation (ICRA 2002)*, Washington, DC, USA, Seite 695–701.
- Montgomery, D. C. (2001). *Design and Analysis of Experiments*. John Wiley and Sons, Inc., Hoboken, New Jersey, USA.
- Moors, M. (2008). *Multi Robot Intruder Search*. Dissertation, Institut für Informatik III, Rheinische Friedrich–Wilhelms–Universität Bonn, Römerstraße 164, 53117 Bonn.
- Moors, M., Röhling, T., und Schulz, D. (2005). A probabilistic approach to coordinated multi–robot indoor surveillance. In *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005)*, Edmonton, Alberta, Kanada, Seite 3447–3452.
- Moors, M. und Schulz, D. (2006). Improved Markov models for indoor surveillance. In *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2006)*, Peking, China, Seite 4072–4077.

- Moravec, H. P. und Elfes, A. E. (1985). High resolution maps from wide angle sonar. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 1985)*, Seite 116–121.
- Mucientes, M. und Burgard, W. (2006). Multiple hypothesis tracking of clusters of people. In *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '06), Peking, China*, Seite 692–697.
- Munkres, J. (1957). Algorithms for the assignment and transportation problems. *Journal of the Society of Industrial and Applied Mathematics*, 5(1):32–38.
- Munoz-Salinas, R., Aguirre, E., und Garcia-Silvente, M. (2007). People detection and tracking using stereo vision and color. *Image and Vision Computing*, 25(6):995–1007.
- Murillo, A. C., Guerrero, J. J., und Sagues, C. (2006). Robot and landmark localization using scene planes and the 1D trifocal tensor. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '06), Peking, China*, Seite 2070–2075.
- Murrieta-Cid, R., Gonzalez-Banos, H. H., und Tovar, B. (2002). A reactive motion planner to maintain visibility of unpredictable targets. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '02), Washington, DC, USA*, Seite 4242–4248.
- Neira, J., Tardos, J. D., und Castellanos, J. A. (2003). Linear time vehicle relocation in SLAM. In *Proceedings of the 2003 IEEE International Conference on Robotics and Automation (ICRA 2003), Taipeh, Taiwan*, Seite 427–433.
- Neto, H. V. und Nehmzow, U. (2007). Visual novelty detection with automatic scale detection. *Robotics and Autonomous Systems*, 55(9):693–701.
- Ngako Pangop, L., Chausse, F., Chapuis, R., und Cornou, S. (2008). Asynchronous Bayesian algorithm for object classification: Application to pedestrian detection in urban areas. In *Proceedings of the IEEE/ISIF — 11th International Conference on Information Fusion (FUSION 2008), Köln, Deutschland, CD-ROM*.
- Nieto, J., Guivant, J., Nebot, E., und Thrun, S. (2003). Real time data association for FastSLAM. In *Proceedings of the IEEE 2003 International Conference on Robotics and Automation (ICRA 2003), Taipeh, Taiwan*, Seite 412–418.
- Nourbakhsh, I., Powers, R., und Birchfield, S. (1995). DERVISH: an office-navigating robot. *AI Magazine*, 16(2):53–60.

- Nüchter, A., Lingemann, K., Hertzberg, J., und Surmann, H. (2005). 6D SLAM with approximate data association. In *Proceedings of the 11th International Conference on Advanced Robotics (ICAR '05), Seattle, Washington, USA*, Seite 242–249.
- Nüchter, A., Lingemann, K., Hertzberg, J., und Surmann, H. (2007). 6D SLAM – 3D mapping outdoor environments. *Journal of Field Robotics (JFR), special issue on quantitative performance evaluation of robotic and intelligent systems*, 24(8–9):699–722.
- Painter, J. H., Kerstetter, D., und Jowers, S. (1990). Reconciling steady–state Kalman and alpha–beta filter design. *IEEE Transactions on Aerospace and Electronic Systems*, 26(6):986–991.
- Parker, L. E. (1997). Cooperative motion control for multi–target observation. In *Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '97), Grenoble, Frankreich*, Seite 1591–1598.
- Pineau, J., Montemerlo, M., Roy, N., Thrun, S., und Pollack, M. (2003). Towards robotic assistants in nursing homes: Challenges and results. *Robotics and Autonomous Systems, Special Issue on Socially Interactive Robots*, 42(3–4):271–281.
- Pirjanian, P. und Mataric, M. (2000). Multi–robot target acquisition using multiple objective behaviour coordination. In *Proceedings of the 2000 IEEE International Conference on Robotics and Automation (ICRA 2000), San Francisco, Kalifornien, USA*, Seite 2696–2702.
- Pitt, M. und Shephard, N. (1999). Filtering via simulation: auxiliary particle filters. *Journal of the American Statistical Association*, 94(446):590–599.
- Prassler, E., Scholz, J., und Elfes, E. (1999). Tracking people in a railway station during rush–hour. In Christensen, H. I., Herausgeber, *Computer Vision Systems*, Band 1542, Seite 162–179. Springer, Lecture Notes in Computer Science Edition.
- Preve, C. und Marchette, D. (1987). An application of neural networks to a data fusion problem. Tri–service Data Fusion Technical Proceedings, John Hopkins University, Maryland, Naval Air Development Center, Warminster, Pennsylvania, USA, Seite 226–236.
- Priyantha, N. B., Chakraborty, A., und Balakrishnan, H. (2000). The cricket location–support system. In *Proceedings of the Sixth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom 2000), Boston, Massachusetts, USA*, Seite 32–43.

- Pulford, G. W. und La Scala, B. (1995). Over-the-horizon radar tracking algorithm using the Viterbi algorithm — third report to high frequency radar division. Technischer Bericht 27/95, Cooperative Research Centre for Sensor Signal and Information Processing, University of Melbourne.
- Quach, T. und Farooq, M. (1994). Maximum likelihood track formation with the Viterbi algorithm. In *Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, USA*, Seite 271–276.
- Rasmussen, C. und Hager, G. D. (1998). Joint probabilistic techniques for tracking multi-part objects. In *Proceedings of the 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '98), Santa Barbara, Kalifornien, USA*, Seite 16–21. IEEE Computer Society.
- Rauch, H. E., Tung, F., und Striebel, C. T. (1965). Maximum likelihood estimates of linear dynamic systems. *American Institute of Aeronautics and Astronautics Journal*, 3(8):1445–1450.
- Reid, D. B. (1979). An algorithm for tracking multiple targets. *IEEE Transactions on Automatic Control*, 24(6):843–854.
- Renyi, A. (1970). *Probability Theory*. North Holland Publishing Company, Amsterdam London.
- Röhling, T. (2008). Scanclusteringverfahren für den dreidimensionalen Kartenbau. Diplomarbeit, Institut für Informatik III, Rheinische Friedrich-Wilhelms-Universität Bonn, Römerstraße 164, 53117 Bonn.
- Röhling, T. und Schulz, D. (2008). Improving 3D scan registration for SLAM with clustering and deterministic annealing. In Ramamoorthy, S. und Hayes, G. M., Herausgeber, *Proceedings of Towards Autonomous Robotic Systems 2008 (TA-ROS 2008), Edinburgh, UK*, Seite 50–56.
- Romera, M. M., Vazquez, M. A. S., und Garcia, J. C. G. (2004). Tracking multiple and dynamic objects with an extended particle filter and an adapted k-means clustering algorithm. In *Proceedings of the 5th IFAC/EURON Symposium on Intelligent Autonomous Vehicles (IAV 2004), Lissabon, Portugal, CD-ROM*.
- Rosales, R. und Sclaroff, S. (1998). Improved tracking of multiple humans with trajectory prediction and occlusion modeling. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '98), Proceedings of the Workshop on the Interpretation of Visual Motion, Santa Barbara, Kalifornien, USA*.
- Rosencrantz, M., Gordon, G., und Thrun, S. (2003). Locating moving entities in dynamic indoor environments with teams of mobile robots. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2003), Melbourne, Victoria, Australien*, Seite 233–240.

- Roth, V., Laub, J., Kawanabe, M., und Buhmann, J. (2003). Optimal cluster preserving embedding of nonmetric proximity data. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 25(12):1540–1550.
- Rubin, D. (1988). Using the SIR algorithm to simulate posterior distributions. In Bernardo, J. M., de Groot, M. H., Lindley, D. V., und Smith, A. F. M., Herausgeber, *Bayesian Statistics*, Band 3, Seite 395–402. Oxford University Press, Oxford, UK.
- Ryu, H. R. und Huber, M. (2007). A particle filter approach for multi-target tracking. In *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007)*, San Diego, Kalifornien, USA, Seite 2753–2760.
- Salmond, D. J. und Gordon, N. J. (1999). Group and extended object tracking. In *Proceedings of Signal and Data Processing of Small Targets, Denver, Colorado, USA*, Band 3809 der *SPIE Proceedings Series*, Seite 284–296. Society of Photo-Optical Instrumentation Engineers.
- Schmitt, T., Beetz, M., Hanek, R., und Buck, S. (2002). Watch their moves: Applying probabilistic multiple object tracking to autonomous robot soccer. In *Proceedings of the Eigtheenth National Conference on Artificial Intelligence (AAAI-02)*, Edmonton, Alberta, Kanada, Seite 599–604.
- Schmundt, H. (2004). Roboter: Thermoskanne auf Patrouille. In *Der Spiegel*, Nummer 3, Seite 120. Spiegel-Verlag Rudolf Augstein GmbH & Co. KG.
- Schneider, F. E. (2006). EFS/FKIE/FGAN, Neuenahrer Straße 20, 53343 Wachtberg, Persönliche Mitteilung.
- Schulz, D. (2002). *Internet-Based Robotic Tele-Presence*. Dissertation, Institut für Informatik III, Rheinische Friedrich-Wilhelms-Universität Bonn, Römerstraße 164, 53117 Bonn.
- Schulz, D. (2006a). A probabilistic exemplar approach to combine laser and vision for person tracking. In *Proceedings of Robotics: Science and Systems II, Philadelphia, Pennsylvania, USA*,
Online: <http://www.roboticsproceedings.org/rss02/p19.html>.
- Schulz, D. (2006b). Institut für Informatik III, Rheinische Friedrich-Wilhelms-Universität Bonn, Römerstraße 164, 53117 Bonn, Persönliche Mitteilung.
- Schulz, D. (2008). EMS/FKIE/FGAN, Neuenahrer Straße 20, 53343 Wachtberg, Persönliche Mitteilung.
- Schulz, D., Burgard, W., Fox, D., und Cremers, A. B. (2001a). Tracking moving targets with a mobile robot using particle filters and statistical data association. In

- Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '01), Seoul, Südkorea, Seite 1665–1670.*
- Schulz, D., Burgard, W., Fox, D., und Cremers, A. B. (2001b). Tracking multiple moving objects with a mobile robot. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001), Kauai, Hawaii, Band 1, Seite 371–377.*
- Schulz, D., Burgard, W., Fox, D., und Cremers, A. B. (2003a). People tracking with mobile robots using sample-based joint probabilistic data association filters. *The International Journal of Robotics Research*, 22(2):99–116.
- Schulz, D., Fox, D., und Hightower, J. (2003b). People tracking with anonymous and ID-sensors using Rao-Blackwellised particle filters. In *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI 2003), Acapulco, Mexiko, Seite 921–926.*
- Schumitsch, B., Thrun, S., Bradski, G., und Olukotun, K. (2006a). The information-form data association filter. In Weiss, Y., Schölkopf, B., und Platt, J., Herausgeber, *Advances in Neural Information Processing Systems 18*, Seite 1193–1200. The MIT Press, Cambridge, Massachusetts, USA.
- Schumitsch, B., Thrun, S., Guibas, L., und Olukotun, K. (2006b). The identity management Kalman filter (IMKF). In *Proceedings of Robotics: Science and Systems II, Philadelphia, Pennsylvania, USA*,
Online: <http://www.roboticsproceedings.org/rss02/p29.html>.
- Sea, R. G. (1971). An efficient suboptimal decision procedure for associating sensor data with stored tracks in real-time surveillance systems. In *Proceedings of the 10th IEEE Conference on Decision and Control, Miami Beach, Florida, USA*, Seite 33–37.
- Shao, X., Zhao, H., Nakamura, K., Katabira, K., Shibasaki, R., und Nakagawa, Y. (2007). Detection and tracking of multiple pedestrians by using laser range scanners. In *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007), San Diego, Kalifornien, USA*, Seite 2174–2179.
- Shapiro, S. C., Herausgeber (1987). *Encyclopedia of Artificial Intelligence*, Band 1–2. John Wiley and Sons, Inc., New York, USA.
- Shumway, R. H. und Stoffer, D. S. (2000). *Time Series Analysis and Its Applications*. Springer, Heidelberg.
- Sick AG (2006). *Technische Beschreibung Lasermesssysteme LMS 200/LMS 211/LMS 220/LMS 221/LMS 291*. Sick AG, Sick Vertriebs-GmbH, Schiess-Straße 56, 40549 Düsseldorf, Deutschland.

- Sigal, L., Bhatia, S., Roth, S., Black, M. J., und Isard, M. (2004). Tracking loose-limbed people. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04), Washington, DC, USA*, Band 1, Seite 421–428.
- Simmons, R., Goodwin, R., Haigh, K., Koenig, S., und O'Sullivan, J. (1997). A layered architecture for office delivery robots. In *Proceedings of the First International Conference on Autonomous Agents, Marina del Rey, Kalifornien, USA*, Seite 245–252.
- Simmons, R. und Koenig, S. (1995). Probabilistic robot navigation in partially observable environments. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI '95), Montreal, Quebec, Kanada*, Band 2, Seite 1080–1087, San Francisco, Kalifornien. Morgan Kaufmann Publishers Inc.
- Singer, R. A. und Kanyuck, A. J. (1971). Computer control of multiple site correlation. *Automatica*, 7(4):455–463.
- Singer, R. A. und Stein, J. J. (1971). An optimal tracking filter for processing sensor data of imprecisely determined origin in surveillance systems. In *Proceedings of the Tenth IEEE Conference on Decision and Control, Miami Beach, Florida, USA*, Seite 171–175.
- Sittler, R. W. (1964). An optimal data association problem in surveillance theory. *IEEE Transactions on Military Electronics*, 8(2):125–139.
- Smets, P. und Ristic, B. (2004). Kalman filter and joint tracking and classification in the TBM framework. In Svensson, P. und Schubert, J., Herausgeber, *Proceedings of the IEEE/ISIF — Seventh International Conference on Information Fusion (FUSION 2004), Mountain View, Kalifornien, USA*, Band 1, Seite 46–53. International Society of Information Fusion.
- Smith, R., Self, M., und Cheeseman, P. (1990). Estimating uncertain spatial relationships in robotics. In Cox, I. und Wilfong, G., Herausgeber, *Autonomous robot vehicles*, Band 15 der *Lecture Notes in Electrical Engineering*. Springer, Heidelberg.
- Smith, R. C. und Cheeseman, P. (1986). On the representation and estimation of spatial uncertainty. *The International Journal of Robotics Research*, 5(4):56–68.
- Song, X., Cui, J., Wang, X., Zhao, H., und Zha, H. (2008). Tracking interacting targets with laser scanner via on-line supervised learning. In *Proceedings of the 2008 IEEE International Conference on Robotics and Automation (ICRA 2008), Pasadena, Kalifornien, USA*, Seite 2271–2276.

- Stannus, W., Koch, W., und Kräußling, A. (2004). On robot-borne extended object tracking using the EM algorithm. In *Proceedings of the 5th IFAC/EURON Symposium on Intelligent Autonomous Vehicles (IAV 2004), Lissabon, Portugal, CD-ROM*.
- Tadokoro, S., Hayashi, M., Manabe, Y., Nakami, Y., und Takamori, T. (1995). On motion planning of mobile robots which coexist and cooperate with humans. In *Proceedings of the 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '95), Pittsburgh, Pennsylvania, USA, Band 2, Seite 518–523*.
- Tajana, M., Gaspar, J., Nascimento, J., Bernardino, A., und Lima, P. (2007). On the use of perspective catadioptric sensors for 3D model-based tracking with particle filters. In *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007), San Diego, Kalifornien, USA, Seite 2747–2752*.
- Tapus, A., Mataric, M. J., und Scasselati, B. (2007). Socially assistive robotics. *IEEE Robotics and Automation Magazine*, 14(1):35–42.
- Tarassenko, L., Hayton, P., Cerneaz, N., und Brady, M. (1995). Novelty detection for the identification of masses in mammograms. In *Proceedings of the Fourth International Conference on Artificial Neural Networks (ICANN 95), Paris, Frankreich, Seite 442–447*. Cambridge University Press.
- Tardos, J. D., Neira, J., Newman, P., und Leonard, J. (2002). Robust mapping and localization in indoor environments using sonar data. *The International Journal on Robotics Research*, 21(4):311–330.
- Taylor, G. und Kleeman, L. (2004). A multiple hypothesis walking person tracker with switched dynamic model. In *Proceedings of the Australasian Conference on Robotics and Automation (ACRA 2004), Canberra, Australien*.
- Taylor, O. und MacIntyre, J. (1998). Adaptive local fusion systems for novelty detection and diagnostics in condition monitoring. In *Proceedings of Sensor Fusion: Architectures, Algorithms and Applications II, Orlando, Florida, USA, Band 3376 der SPIE Proceedings Series, Seite 210–218*, Bellingham, Washington, USA. Society of Photo-Optical Instrumentation Engineers.
- Thrun, S. (1998a). Bayesian landmark learning for mobile robot localization. *Machine Learning*, 33(1):41–76.
- Thrun, S. (1998b). Learning metric-topological maps for indoor mobile robot navigation. *Artificial Intelligence*, 99(1):21–71.
- Thrun, S. (2003). Robotic mapping: A survey. In Lakemeyer, G. und Nebel, B., Herausgeber, *Exploring Artificial Intelligence in the new Millennium*, Seite 1–35. Morgan Kaufmann Publishers Inc., San Francisco, Kalifornien, USA.

- Thrun, S., Bennewitz, M., Burgard, W., Cremers, A. B., Dellaert, F., Fox, D., Hähnel, D., Rosenberg, C. R., Roy, N., Schulte, J., und Schulz, D. (1999). MINERVA: A second generation mobile tour–guide robot. In *Proceedings of the 1999 IEEE International Conference on Robotics and Automation (ICRA 1999)*, Detroit, Michigan, USA, Seite 1999–2005. IEEE Robotics and Automation Society.
- Thrun, S., Burgard, W., und Fox, D. (2000). A real–time algorithm for mobile robot mapping with applications to multi–robot and 3D mapping. In *Proceedings of the 2000 IEEE International Conference on Robotics and Automation (ICRA 2000)*, San Francisco, Kalifornien, USA, Seite 321–328.
- Thrun, S., Burgard, W., und Fox, D. (2005). *Probabilistic Robotics*. Intelligent Robotics and Autonomous Agents. The MIT Press, Cambridge, Massachusetts und London, UK.
- Thrun, S., Fox, D., und Burgard, W. (1998). A probabilistic approach to concurrent mapping and localization for mobile robots. *Machine Learning, Special Issue on Learning in Autonomous Robots*, 31(1–3):29–53.
- Tiderko, A. und Bachran, T. (2007). A service oriented framework for wireless communication in mobile multi robot systems. In *Proceedings of the First International Conference on Robot Communication and Coordination (ROBOCOMM '07)*, Athen, Griechenland, Band 318 der *ACM International Conference Proceedings Series*.
- Tiderko, A., Hoeller, F., und Bachran, T. (2007). A framework for multicast communication over unreliable networks in multirobot systems. In Wilson, M. S., Labrosse, F., Nehmzow, U., Melhuish, C., und Witkowski, M., Herausgeber, *Proceedings of Towards Autonomous Robotic Systems 2007 (TAROS 2007)*, Aberystwyth, UK, Seite 48–55.
- Treptow, A., Cielniak, G., und Duckett, T. (2005). Active people recognition using thermal and grey images on a mobile security robot. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '05)*, Edmonton, Alberta, Kanada, Seite 2103–2108.
- Ulmke, M. (2003). SDF/FKIE/FGAN, Neuenahrer Straße 20, 53343 Wachtberg, Persönliche Mitteilung.
- van Keuk, G. (1971). Zielverfolgung nach Kalman–Anwendung auf elektronisches Radar. Technischer Bericht 173, Forschungsinstitut für Funk und Mathematik, Neuenahrer Straße 20, 53343 Wachtberg, Deutschland.
- Viterbi, A. J. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory*, 13(2):260–269.

- Waldherr, S., Thrun, S., Romero, R., und Margaritis, D. (1998). Template-based recognition of pose and motion gestures on a mobile robot. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI '98)*, Madison, Wisconsin, USA, Seite 977–982. The MIT Press, Cambridge, Massachusetts und London, UK.
- Wang, C.-C., Lo, T.-C., und Yang, A.-W. (2007). Interacting object tracking in crowded urban areas. In *Proceedings of the 2007 IEEE International Conference on Robotics and Automation (ICRA 2007)*, Rom, Italien, Seite 4626–4632.
- Wang, C.-C., Thorpe, C., und Thrun, S. (2003). Online simultaneous localization and mapping with detection and tracking of moving objects: theory and results from a ground vehicle in crowded urban areas. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '03)*, Taipeh, Taiwan, Seite 842–849.
- Want, R., Hopper, A., Falcao, V., und Gibbons, J. (1992). The active badge location system. *Association for Computing Machinery Transactions on Information Systems*, 10(1):91–102.
- West, D. B. (2000). *Introduction to Graph Theory*. Prentice Hall, Inc., Upper Saddle River, New Jersey, USA.
- Wienecke, M., Safenreiter, K., und Koch, W. (2008a). Combined person tracking and classification in a network of chemical sensors. In *Proceedings of the IEEE/ISIF — 11th International Conference on Information Fusion (FUSION 2008)*, Köln, Deutschland, CD-ROM.
- Wienecke, M., Safenreiter, K., und Koch, W. (2008b). Hazardous material localization and person tracking. In *Proceedings of SPIE — Conferences of Optics and Photonics (Signal and data processing of small targets)*, Orlando, Florida, USA, Band 6969, CD-ROM. Society of Photo-Optical Instrumentation Engineers.
- Witten, I. H. und Eibe, F. (2005). *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann Publishers Inc., San Francisco, Kalifornien, USA.
- Wren, C. R., Azarbayejani, A., Darrell, T., und Pentland, A. P. (1997). Pfunder: Real-time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):780–785.
- Wu, S.-X. und Yang, M. (2007). Landmark pair based localization for intelligent vehicles using laser radar. In *Proceedings of the 2007 IEEE Intelligent Vehicles Symposium (IV 2007)*, Istanbul, Türkei, Seite 292–297.

- Ypma, A. und Duin, R. P. (1997). Novelty detection using self-organizing maps. In *Proceedings of the International Conference on Neural Information Processing and Intelligent Information Systems (ICONIP 1997), Dunedin, Neuseeland*, Seite 1322–1325.
- Zajdel, W., Zivkovic, Z., und Krose, B. J. A. (2005). Keeping track of humans: Have i seen this person before? In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2005), Barcelona, Spanien*, Seite 2093–2098.
- Zhang, X., Liu, Z., und Qiao, H. (2007). A robust multiple cues fusion based Bayesian tracker. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2007), Rom, Italien*, Seite 4614–4619.
- Zhao, H. und Shibasaki, R. (2005). A novel system for tracking pedestrians using multiple single-row laser-range scanners. *IEEE Transactions on Systems, Man and Cybernetics — Part A: Systems and Humans*, 35(2):283–291.
- Zhou, S. und Chellappa, R. (2002). Probabilistic human recognition from video. In *Proceedings of the Seventh European Conference on Computer Vision (ECCV 2002), Kopenhagen, Dänemark*, Seite 681–697.
- Zhu, Q. (1991). Hidden Markov model for dynamic obstacle avoidance of mobile robot navigation. *IEEE Transactions on Robotics and Automation*, 7(3):390–397.

