Technische Universität Ilmenau
Fakultät für Informatik und Automatisierung
Fachgebiet Neuroinformatik und Kognitive Robotik

Laserbasiertes Verfahren zur vollautomatischen Erfassung und Protokollierung von Personenbewegungstrajektorien

Masterarbeit zur Erlangung des akademischen Grades Master of Science

Konrad Schenk

Betreuer: Dipl.-Inf. Erik Schaffernicht, FG Neuroinformatik und Kognitive Robotik
Dipl.-Inf. Markus Eisenbach, FG Neuroinformatik und Kognitive Robotik
Verantwortlicher Hochschullehrer:
Prof. Dr. H.-M. Groß, FG Neuroinformatik und Kognitive Robotik

Die Masterarbeit wurde am 15.02.2011 bei der Fakultät für Informatik und Automatisierung der Technischen Universität Ilmenau eingereicht.

An dieser Stelle möchte ich allen danken, die mich in meiner Masterarbeit unterstützt haben. Besondere Erwähnung verdienen meine Betreuer Prof. Dr. H.-M. Groß, Dipl.-Inf. Erik Schaffernicht und Dipl.-Inf. Markus Eisenbach, welche mir mit konstruktiver Kritik geholfen haben und immer ein offenes Ohr für mich und meine Fragen hatten. Ohne sie wäre diese Arbeit nicht möglich gewesen. Des Weiteren möchte ich auch allen Korrekturlesern danken. Diese sind im einzelnen mein Bruder Sebastian und mein Vater, welche aus Sicht einer anderen ingenieurswissenschaftlichen Disziplin inhaltliche Anregungen gegeben haben. Meiner Oma Trautchen möchte ich ganz speziell dafür danken, dass diese Arbeit nun auch weitestgehend den Regeln der deutschen Sprache entspricht. Meinem guten Freund und akademischen Leidensgenossen Johannes Rühle gebührt Dank dafür, dass er mich die ersten Schritte auf dem Gebiet des Personentrackings begleitet hat und wir zusammen im Projektseminar einiges an Vorarbeit für den Fußtracker leisten konnten. Allen anderen Freunden sei dafür gedankt, dass sie mir immer mal die nötige Ablenkung gegeben haben, damit ich mich wieder mit frischer Energie auf die Personenbewegungstrajektorien stürzen konnte.
Erwähnung verdient auch noch mein treuer Computer, der wider Erwarten die sechs Monate ohne Absturz, ohne Defekt und ohne Datenverlust überlebt hat, trotzdem ich ihn mit stundenlangen Berechnungen traktiert habe.

Erklärung: „Hiermit versichere ich, dass ich diese Masterarbeit selbständig verfasst und nur die angegebenen Quellen und Hilfsmittel verwendet habe. Alle von mir aus anderen Veröffentlichungen übernommenen Passagen sind als solche gekennzeichnet.“

Ilmenau, 15.02.2011 .....

Konrad Schenk

Inhaltsverzeichnis

1 Einleitung **1**

1.1 Motivation 1

1.2 Aufgabenstellung 4

2 State of the Art **5**

2.1 Einleitung 5

2.2 Personentracking auf Fußhöhe 5

2.3 Personentracking auf Oberkörperhöhe 10

2.4 Repräsentation der Personenspuren 12

2.5 Zusammenfassung 13

3 Theoretische Grundlagen **15**

3.1 Einleitung 15

3.2 Funktionsweise von Laser-Range-Scannern 15

3.3 Kombination mehrerer Laserscanner 17

3.4 Hintergrundsegmentierung 20

3.5 Mean-Shift Clustering 23

3.6 Hauptkomponentenanalyse 25

3.7 Partikelfilter 27

3.7.1 Fußmodell 34

3.7.2 Oberkörpermodell 38

3.7.3 Löschen von Partikelfiltern 38

3.8 Trajektorienrepräsentation 41

3.8.1 Polynomapproximation 41
3.8.2 Chebychev-Approximation 43
3.8.3 Glättung der Trajektorien 45
3.8.4 Zusammenfassung 46

4 Realisierung eines Personentrackers 49
4.1 Einleitung 49
4.2 Messequipment 50
4.3 Vorverarbeitung 51
4.3.1 Segmentierung von Laserdaten 52
4.3.2 Anpassung des Mean-Shifts auf Segmente 53
4.4 Personendetektion 53
4.4.1 Personendetektion auf Fußhöhe 54
4.4.2 Personendetektion auf Oberkörperhöhe 57
4.5 Personentracking 58
4.5.1 Personentracking auf Fußhöhe 59
4.5.2 Personentracking auf Oberkörperhöhe 64
4.6 Nachverarbeitung 66
4.7 Zusammenfassung 68

5 Softwaretechnische Umsetzung 71

6 Experimentelle Untersuchungen 77
6.1 Einleitung 77
6.2 Einsatzumgebungen 77
6.3 Ergebnisse des Fußtrackings 79
6.4 Ergebnisse des Oberkörpertrackings 84
6.5 Vergleich beider Trackingverfahren 91
6.6 Auswertung der Remittanzdaten 93
6.7 Zusammenfassung 94

7 Zusammenfassung und Ausblick **97**

7.1 Zusammenfassung der Arbeit 97

7.2 Weiterführende Arbeiten 98

Literaturverzeichnis **99**

Kapitel 1

Einleitung

1.1 Motivation

In der kognitiven Robotik und auch im Bereich der Mensch-Maschine-Interaktion ist die Erkennung von Personen ein wichtiger Bestandteil. Zum Beispiel muss im Supermarkt ein Serviceroboter wissen, wo sich mögliche Kundschaft aufhält, um gezielt seine Dienste anbieten zu können. Allerdings ermöglicht eine einfache Detektion von Personen in den Sensordaten (z.B. Kamerabilder) noch keine Aussage über dynamische Eigenschaften der Person, wie die Geschwindigkeit, mit der sie sich fortbewegt, oder in welche Richtung sie geht. Des Weiteren gestaltet sich die Wiedererkennung, je nach

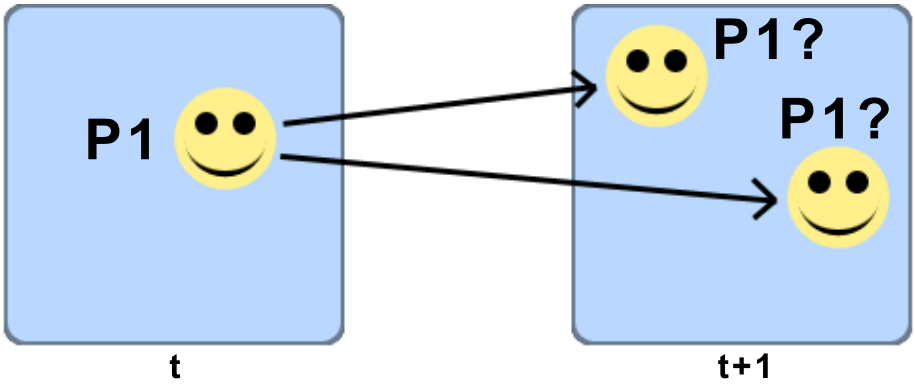


Abbildung 1.1: Welche Person im Zeitschritt $t + 1$ ist die Person $P1$ aus dem Zeitschritt t? Es wird versucht, dieses Problem mit Hilfe von Trackingalgorithmen zu lösen.

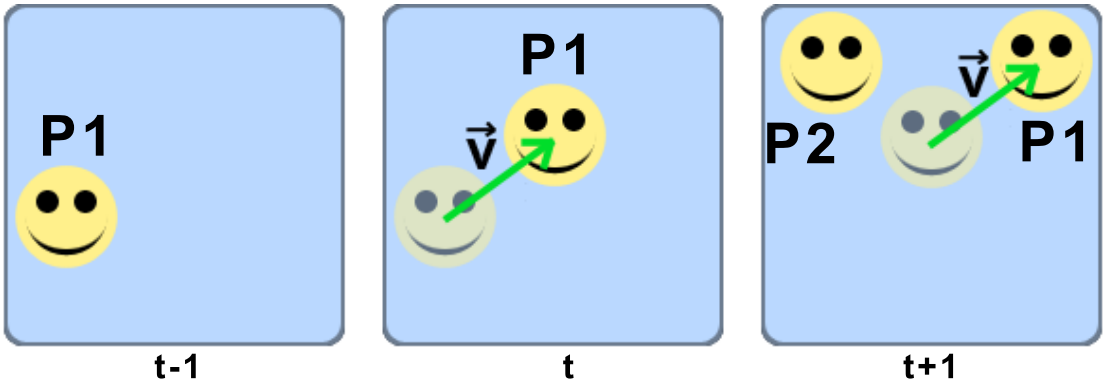


Abbildung 1.2: Anhand vorangegangener Beobachtungen der Person $P1$ aus dem Zeitschritt $t-1$ konnte der Bewegungsvektor v ermittelt werden und eine sichere Zuordnung im neuen Zeitschritt $t+1$ erfolgen.

Sensorik, relativ schwierig. So lässt sich nicht immer eindeutig sagen, ob die Person, welche vor kurzem entdeckt wurde, jene ist, die wenige Zeit später einen halben Meter weiter links erkannt wurde (Abbildung 1.1). Erst durch Berücksichtigung der vergangenen Beobachtungen lässt sich, mit Hilfe von Trackingalgorithmen, eine Zuordnung der Personendetektionen aus den einzelnen Zeitschritten untereinander durchführen und die jeweiligen Personen über die Zeit verfolgen (siehe Abbildung 1.2). Somit ist dann auch die Bestimmung von dynamischen Eigenschaften der Personen möglich.

In einem der aktuellen Projektvorhaben des Fachgebiets Neuroinformatik und kognitive Robotik der TU Ilmenau namens APFel (**A**nalyse von **P**ersonenbewegungen an **F**lughäfen mitt**e**ls zeit**l**ich rückwärts- und vorwärtsgerichteter Videodatenströme) ist ebenfalls die Detektion, das Tracken und die Wiedererkennung von Personen von zentraler Bedeutung, allerdings nicht auf mobilen Robotern, wie im vorangegangenen Beispiel, sondern anhand fest installierter Videokameras. Auf Flughäfen werden sämtliche Bilder der Überwachungskameras in einer Zentrale von Sicherheitspersonal gesichtet und nach auffälligen Personen untersucht. Da es den Mitarbeitern kaum möglich ist, alle Kamerabilder gleichzeitig im Auge zu behalten und es mit viel Aufwand verbunden ist, im Nachhinein zu überprüfen, wo sich eine bestimmte Person entlang bewegt hat, wäre eine automatisierte Unterstützung wünschenswert. Diese Unterstützung soll das Projekt APFel in Form einer Auswertungssoftware liefern. Hierbei wird auf Anfra-

ge eines Operators die bereits zurückgelegte Bewegungsspur einer markierten Person anhand der Videodaten ermittelt und darstellt. Des Weiteren sollen auch mögliche weitere Bewegungsspuren der Person vorhergesagt werden.
Um die nötigen Algorithmen zu entwickeln, die in Videodaten Personen detektieren, wiedererkennen und tracken können, ist für Versuchsaufnahmen eine Wissensbasis notwendig, die schon während der Aufnahmen die genaue Position der einzelnen Personen zu den einzelnen Zeitpunkten beschreibt. Es ist offensichtlich, dass diese Wissensbasis nicht aus den Videodaten selbst gewonnen werden kann, da die dafür notwendigen Algorithmen jene sind, die erst entwickelt werden sollen. Somit müssen die Bewegungsspuren, die diese Wissensbasis bilden, auf andere Art und Weise ermittelt werden. Eine dafür gut geeignete Möglichkeit bieten Entfernungsmessungen mit Laser-Range-Scanner, welche im Gegensatz zu Kameras wesentlich genauere und zuverlässigere Personendetektionen ermöglichen. Allerdings wurden bisher am Fachgebiet Neuroinformatik und kognitive Robotik noch keine geeigneten Algorithmen für das Projektvorhaben entwickelt, welche die Bewegungsspuren mehrerer Personen in einem großen Areal anhand von Laserscandaten erfassen. Im Rahmen dieser Arbeit soll ein Personentracker entwickelt werden, welcher die Bewegungsspuren erfasst.

1.2 Aufgabenstellung

Ziel dieser Masterarbeit ist es, ein Verfahren zu entwickeln, welches aus den Daten mehrerer Laser-Range-Scanner die Bewegungsspuren von mehreren Personen im Erfassungsgebiet der Scanner extrahiert. Dazu sollen im Einzelnen zwei Ansätze umgesetzt und verglichen werden. Der eine Ansatz basiert auf Entfernungsscans der Füße, bzw. Fußknöchel der Personen, wohingegen der andere Ansatz Entfernungsscans der Oberkörper der Personen nutzt. Diese Ansätze sollen dann für das Detektieren und Tracken mehrerer Personen angepasst werden. Zudem müssen auch mehrere Laser-Range-Scanner eingesetzt werden, um bei größerem Personenaufkommen Verdeckungen in den Entfernungsscans zu vermeiden. Dazu müssen die einzelnen Scanner sowohl räumlich, als auch zeitlich synchronisiert werden. Die in den Experimenten verwendeten Laser-Range-Scanner bieten auch die Möglichkeit, Remittanzdaten der erfassten Objekte zu ermitteln. Inwiefern sich diese Daten zur Verbesserung der Trackinggüte verwenden lassen, soll ebenfalls untersucht werden. Schlussendlich ist es auch Bestandteil dieser Arbeit, die ermittelten Bewegungsspuren zu glätten und in geeigneter Form weiterer Auswertungssoftware zur Verfügung zu stellen.
Diese Arbeit ist folgendermaßen gegliedert: Kapitel 2 gibt eine kurze Übersicht über den Stand der Arbeiten anderer Personen und Forschungsgruppen, welche sich mit dem Erkennen und Verfolgen von Personen, anhand von Laserscandaten, beschäftigen. Des Weiteren wird auch kurz auf mögliche Repräsentationsformen der Bewegungsspuren eingegangen. Im Kapitel 3 werden die Verfahren aus der theoretischen Sicht im Einzelnen näher erklärt, mit denen im darauffolgenden Kapitel 4 die zwei Varianten des Personentrackers realisiert wurden. Die dafür notwendige Software und deren Architektur wird im Kapitel 5 beschrieben. Die während der Evaluation erzielten Ergebnisse der Personentracker werden in Kapitel 6 vorgestellt und verglichen. Schlussendlich werden im Kapitel 7 noch einmal die Ergebnisse der Arbeit zusammengefasst und mögliche Erweiterungen vorgeschlagen.

Kapitel 2

State of the Art

2.1 Einleitung

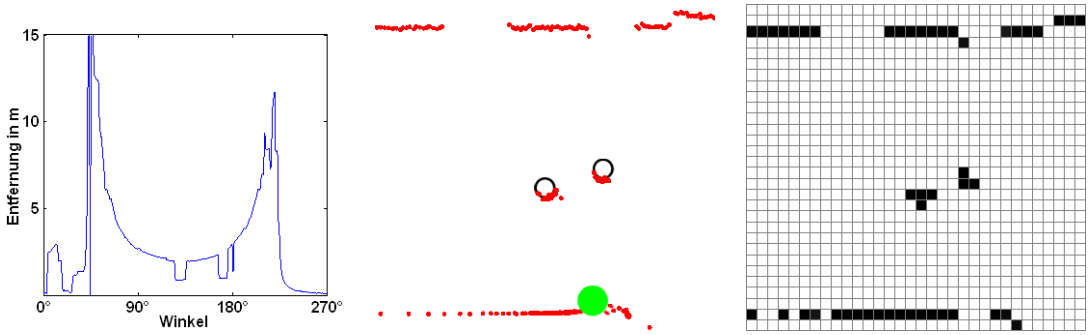
Nachdem im vorherigen Kapitel dargelegt wurde, welche Probleme mit dieser Arbeit gelöst werden sollen, werden in diesem Kapitel ausgewählte Arbeiten anderer Personen und Forschungsgruppen vorgestellt, welche ein ähnliches Problem adressieren. Diese Arbeiten werden zudem daraufhin geprüft, ob sie in dieser Masterarbeit verwendet werden können.

2.2 Personentracking auf Fußhöhe

Mit dem laserbasiertem Mehrpersonentracking auf Fußhöhe haben sich bereits viele Forschungsgruppen befasst. So werden in [Schulz et al. 2001] und [Schulz et al. 2003] Personen, oder bewegte Objekte, mit einem mobilen Roboter getrackt. Dazu werden die lokalen Minima eines einzelnen Laserscanners (deren Funktionsweise wird in Abschnitt 3.2 näher erläutert) ermittelt und mit einer Gridkarte der Umwelt abgeglichen, um die dynamischen Objekte zu extrahieren. Dabei dient die Gridkarte als Repräsentation der statischen Objekte, welche als Hintergrund angesehen werden. Die Karte ist aus kleinen Zellen aufgebaut, mit denen der Raum diskretisiert wird (Abbildung 2.1(c) und 2.2(c) verdeutlichen dies). In den Zellen wird gespeichert, ob sich in den jeweiligen Bereichen ein Hindernis befindet, oder ob er frei ist. Als Tracker

wird ein Partikelfilter (vergleiche dazu Abschnitt 3.7) eingesetzt, der Position, Richtung und Geschwindigkeit der Person schätzt. Mit der Hilfe eines Joint Probabilistic Data Association Filters, welcher die Zuordnung von Elementen zweier Mengen unter Berücksichtigung von deren statistischen Verteilungen realisiert, wird die Zuordnung der geschätzten Personenhypothesen zu den aktuellen Messungen umgesetzt. Somit ließen sich mehrere Personen in der näheren Umgebung des Roboters, trotz Verdeckungen, zuverlässig verfolgen. Da die statischen Objekte ermittelt werden, indem die Differenz der letzten zwei zeitlich aufeinander folgenden Gridkarten betrachtet wird, werden still stehende Personen als Hintergrund angesehen und somit auch nicht weiter getrackt. Trotzdem die Ergebnisse vielversprechend sind, wurde dieser Ansatz hier nicht verwendet, da einerseits nur drei Scans pro Sekunde verarbeitet werden konnten und andererseits auch still stehende Personen kein Problem bereiten dürfen. Zudem wurden einige der verwendeten Verfahren in Hinblick auf einen mobilen Roboter als Einsatzplattform gewählt, was in dieser Masterarbeit jedoch nicht zielführend ist.

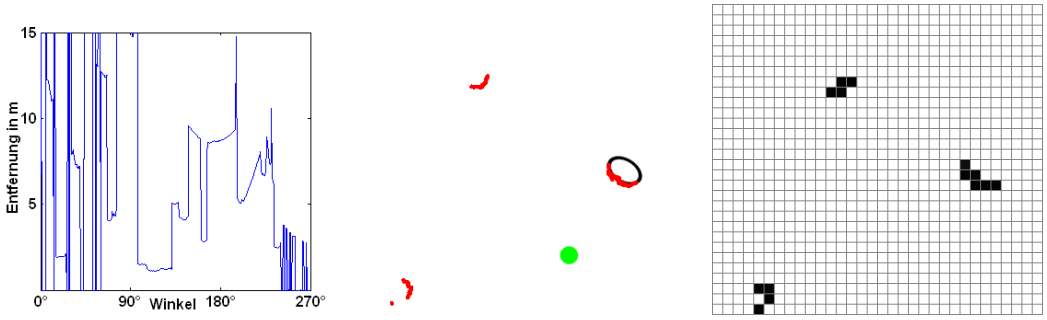
Eine Erweiterung der Arbeiten von [Schulz et al. 2001] stellt [Topp und Christensen 2005] dar. In deren Arbeit wurde ebenfalls ein mobiler Roboter eingesetzt, welcher mit einem Laser-Range-Scanner versucht, Personen in seiner Umgebung zu tracken. Die Detektion von Personen erfolgt mit einer Suche von drei verschiedenen Mustern in den Daten, nachdem der Hintergrund mit Hilfe einer Gridkarte, in der gleichen Weise wie in [Schulz et al. 2001], entfernt wurde: „single leg“, also ein einzelnes sichtbares Bein, „two legs“, was zwei nahe beieinander stehende Beine entspricht (in Abbildung 2.1(b) ist diese Situation auch zu sehen) und „person-wide blob“ für den Fall, dass eine Person mit Rock oder weiten Hosen in der Szene vorhanden ist. Die Erweiterung bezüglich [Schulz et al. 2001] stellt das Behandeln von still stehenden Personen dar, allerdings wird dies in der Arbeit nicht näher erläutert. In aufgeräumten Umgebungen mit wenigen Personen wurden gute Ergebnisse erzielt, jedoch bestand noch Verbesserungsbedarf bei der Personendetektion in vollgestellten und komplizierten Szenen. Da der Personentracker dieser Masterarbeit auf einem Flughafen eingesetzt werden soll, wo die Bewegungsspuren von vielen Personen in einer komplizierten Umgebung ermittelt werden sollen, wurde der Ansatz aus [Topp und Christensen 2005] nicht weiter verfolgt.



(a) Scandaten (b) Punktemenge (c) Gridkarte

Abbildung 2.1: Ein typisches Beispiel für zwei Füße im Laserscan. In (a) ist die gemessene Entfernung in Abhängigkeit zum Scanwinkel abgebildet. Die beiden Minima zwischen 90° und 180° resultieren aus den Füßen. In (b) sind die Daten aus (a) in ein globales Koordinatensystem als Punktemenge in der Draufsicht eingetragen. Die Position des Laserscanners ist grün eingezeichnet, die Scanpunkte sind rot eingefärbt und die ungefähren Umrisse der Füße sind mit schwarzen Kreisen angedeutet. In (c) ist die daraus gebildete Gridkarte, welche auch als Belegtheitskarte bezeichnet wird, exemplarisch dargestellt. Dabei repräsentieren weiße Kästchen Freiraum und schwarze Kästchen Hindernisse.

Die Arbeit von [Lee et al. 2006] befasst sich ebenfalls mit dem Mehrpersonentracking mit der Hilfe eines Laser-Range-Scanners auf einem mobilen Roboter. Dabei werden Personen, bzw. deren Füße, anhand von charakteristischen Geometrien von Punktegruppen detektiert (z. B. ein Kreis, wie in Abbildung 2.1(b)). Diese Punktegruppen werden durch einfache Abstandsmaße (so, wie es auch in Abschnitt 4.3.1 beschrieben wird) in den Laserdaten gebildet. Für das Tracking kommt ein Extended Kalman Filter zum Einsatz. Dieser Filter ist in der Lage, nichtlineare veränderliche Werte auf Basis von verrauschten Sensordaten und unter Berücksichtigung von vorherigen Messungen möglichst genau zu schätzen. Dazu ist ein Modell nötig, welches das zugrundeliegende System beschreibt. Im konkreten Fall kam ein Bewegungsmodell zum Einsatz, welches die pendelartige Bewegung der Füße modelliert. In den Experimenten konnte gezeigt werden, dass mit diesem Verfahren erfolgreich mehrere Personen getrackt werden konnten. Da das Verfahren unter Anderem auf Vorarbeiten von [Shao et al.



(a) Scandaten (b) Punktemenge (c) Gridkarte

Abbildung 2.2: Ein typisches Beispiel für ein Oberkörper im Laserscan. In (a) ist die gemessene Entfernung in Abhängigkeit zum Scanwinkel abgebildet. Das Minimum zwischen 90° und 140° resultiert aus dem Oberkörper. Die oszillierenden Werte von 0° bis etwa 80° und ab etwa 230° sind auf eine verdreckte Schutzblende des Laserscanners zurückzuführen. In (b) sind die Daten aus (a) in ein globales Koordinatensystem als Punktemenge in der Draufsicht eingetragen. Die Position des Laserscanners ist grün eingezeichnet, die Scanpunkte sind rot eingefärbt und der ungefähre Umriss der Person ist als schwarze Ellipse angedeutet. Die anderen beiden halbkreisförmigen Punktemengen wurden durch Säulen hervorgerufen. In (c) ist die daraus gebildete Gridkarte exemplarisch dargestellt.

2007] beruht (dieses Verfahren wird am Ende dieses Abschnittes erläutert) und für eine mobile Plattform mit einem Laser-Range-Scanner angepasst ist, findet es in dieser Masterarbeit keine direkte Verwendung.

Ein Trackingalgorithmus, welcher mit einem statischen Laser-Range-Scanner und einer Panoramakamera mehrere Personen verfolgen kann, wurde in [Chakravarty und Jarvis 2006] umgesetzt. Die Personen werden in den Laserdaten detektiert, indem Mechanismen verwendet werden, die aus der Bildverarbeitung bekannt sind. Die Laserscans werden in eine Gridkarte eingetragen, welche in den nachfolgenden Berechnungsschritten als Bild behandelt werden. Solch eine Karte ist in Abbildung 2.1(c) und 2.2(c) zu sehen. Nachdem ein Differenzbild mit der Gridkarte aus dem vorherigen Zeitschritt berechnet wurde, werden zuerst Störungen entfernt und dann zusammenhängende Pixel gesucht. Anhäufungen von Pixeln stellen dabei Personendetektionen dar, wobei nicht näher auf den genauen Mechanismus der Detektion eingegangen wurde. Die De-

tektionen werden daraufhin in den Kameradaten validiert. Als Tracker kommen mehrere unabhängige Partikelfilter zum Einsatz (vergleiche dazu Abschnitt 3.7), welche die Positionen der Personen schätzen. Indem die Farbinformationen der Personen aus den Kameradaten mit verwendet werden, kann sichergestellt werden, dass jeder Partikelfilter nur die ihm zugehörige Person trackt und nicht auf andere Personen wechselt. Die Experimente wurden mit bis zu drei Personen durchgeführt, jedoch lässt sich keine Aussage finden, wie sich das Vefahren bei wesentlich mehr Personen verhält. Da in der Masterarbeit kein Einsatz von Kameras zur Unterstützung des Trackings vorgesehen sind, welche jedoch in diesem Verfahren essentiell sind, und die Trackingperformance bei einer hohen Personenanzahl unklar ist, fand es keine weitere Anwendung.

In [Shao et al. 2007] wird ein Verfahren beschrieben, welches unter Einsatz mehrerer stationärer Laserscanner in der Lage ist, mehrere Personen in einem größeren Areal zu tracken. Dazu werden zuerst sämtliche Scandaten in einem globalen Koordinatensystem in Form von Punkten eingetragen (siehe dazu Abschnitt 3.3). Nachdem die Scanpunkte mittels eines Mean-Shift-Clusterers (dieser wird in Abschnitt 3.5 genauer erläutert) zu Punktegruppen zusammengefasst wurden, werden diese Gruppen, welche auch als Cluster bezeichnet werden, geprüft, ob sie Personen darstellen könnten. Die Detektion erfolgt durch Auswertung vorangegangener Zeitschritte und basiert im Wesentlichen auf dem Erkennen eines Fußes, welcher an einem anderen vorbeischwingt (die Detektion wird in Abschnitt 4.4.1 näher beschrieben). Nachdem Personen erfolgreich detektiert wurden, werden deren Positionen und Bewegungsparameter jeweils mit einem Partikelfilter (vergleiche dazu Abschnitt 3.7 und 4.5.1) in den nachfolgenden Zeitschritten verfolgt. Dabei kommt ein Bewegungsmodell zum Einsatz, welches in der Lage ist, die periodischen Bewegungen von zwei Füßen einer Person zu beschreiben (Abschnitt 3.7.1 geht darauf näher ein), um somit möglichst genau die Position der Füße in den neuen Zeitschritten vorherzusagen. Damit ein Partikelfilter bei Personen, welche nahe beieinander gehen, nicht die ursprüngliche Person verliert und eine andere weiterverfolgt, wurde ein Interaktionsterm eingeführt (im Abschnitt 4.5.1 wird dieser unter Anderem behandelt), der eine abstoßende Wirkung unter den Partikelfiltern realisiert. Dieses Verfahren hat auch bei einem hohen Personenaufkommen vielversprechende Ergebnisse geliefert und adressiert die Problematik, wie sie in Kapitel 1

beschrieben wird, weshalb es auch weitestgehend in dieser Masterarbeit Verwendung fand.

2.3 Personentracking auf Oberkörperhöhe

Im Bereich des laserbasierten Personentrackings auf Oberkörperhöhe sind vergleichsweise wenige Arbeiten vorhanden. In [Luo et al. 2007] wird ein Laser-Range-Scanner zusammen mit einer Farbkamera in einem mobilen Roboter verwendet, um Personen detektieren und tracken zu können. Die Detektion der Personen erfolgt, indem in den Laserscandaten bestimmte Muster gesucht werden. Von Interesse sind dabei drei nebeneinander liegende, konkave Bereiche, deren jeweilige Breite ausgewertet wird. Sind die beiden äußeren Bereiche zwischen sieben und 20cm breit und hat das mittlere Segment eine Breite zwischen 25 und 60cm, so kann davon ausgegangen werden, dass eine Person vor dem Roboter steht und dessen Arme mit Oberkörper erkannt wurden (vergleiche dazu Abbildung 2.3). In den Kameradaten werden parallel dazu die Gesichter der Personen detektiert und getrackt. Die Personenhypothesen beider Sensoren werden dann mit einem nicht näher benannten Mechanismus fusioniert. Dieses Verfahren

Abbildung 2.3: Personentracking nach [Luo et al. 2007]. Links wurde im Kamerabild das Gesicht der Person detektiert. Rechts ist die Person in den Entfernungsdaten des Laserscanners durch drei nebeneinander liegende, konkave Bereiche mittig im unteren Drittel zu sehen. Die beiden äußeren Bereiche werden durch die Arme hervorgerufen, wohingegen der breitere Bereich in der Mitte vom Torso stammt. Bildquelle: [Luo et al. 2007]

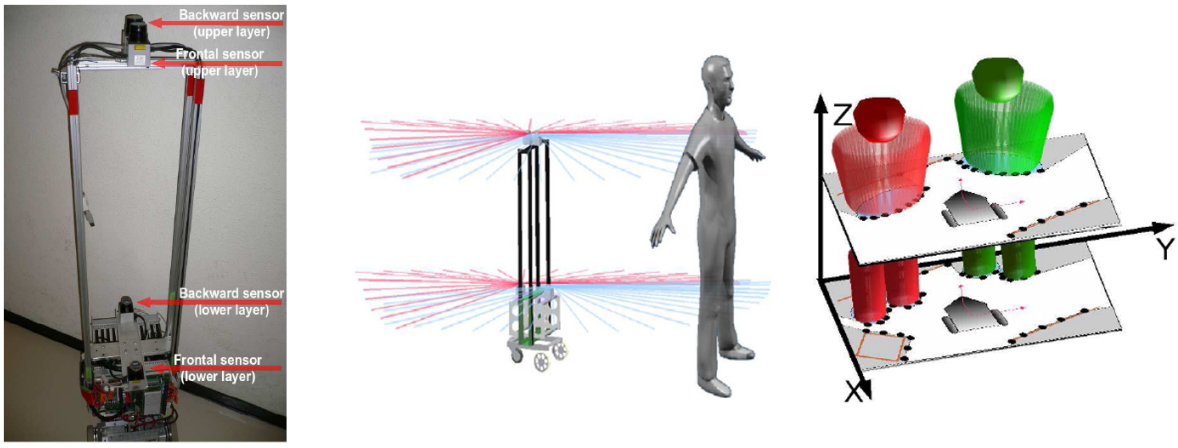


Abbildung 2.4: Scanneranordnung, wie sie in [Carballo et al. 2008] verwendet wurde. Es befinden sich jeweils zwei Laser-Range-Scanner in Höhe der Oberkörper und zwei auf Fußhöhe in entgegengesetzter Richtung, um eine Rundumüberwachung zu gewährleisten. Bildquelle: [Carballo et al. 2008]

scheint nur auf kurzen Distanzen bei frontaler Ausrichtung zum Laserscanner zu funktionieren, da bei größeren Abständen die Auflösung des Scanners nicht hoch genug ist, um die Arme von dem Oberkörper separieren zu können. Um beispielsweise bei einer Scanauflösung von 0,5° (vergleiche dazu Abschnitt 3.2) noch Arme vom Oberkörper separieren zu können, müssten die Arme auf einer Entfernung von 10m zum Scanner mindestens 17cm Zwischenraum zum Oberkörper aufweisen, was nicht unbedingt einer natürlichen Haltung entspricht. Aufgrund dieser Defizite wurde davon abgesehen, das Verfahren in der Masterarbeit weiter zu verfolgen. Zudem wird in der Arbeit das Tracken von nur einer Person beschrieben und nicht auf das Mehrpersonentracking eingegangen.

Eine weitere Publikation, welche ein Verfahren zur Detektion von Personen beschreibt, ist [Carballo et al. 2008]. Dabei werden vier Laser-Range-Scanner eingesetzt, wobei sich zwei auf Fußhöhe und zwei auf Oberkörperhöhe befinden (vergleiche dazu Abbildung 2.4). Die Anordnung der Scanner ist dabei so gewählt, dass eine lückenlose Beobachtung der Umgebung gegeben ist. Nachdem die Scandaten der jeweils auf der gleichen Ebene liegenden Scanner in ein gemeinsames Koordinatensystem eingetragen wurden, werden die Scanpunkte segmentiert. Dabei werden mit einer Abstandsme-

trik, welche die Messunsicherheit der Laserscanner berücksichtigt, zusammenhängende Punkte gesucht. Diese Punktemengen werden daraufhin nach elliptischen Mustern durchsucht. Wurden in den Daten der unteren Laserscanner zwei kleine, kreisförmige Punktemengen und an gleicher Stelle in den Daten der oberen Laserscanner eine große elliptische Form gefunden, so wird dies als eine Personenhypothese angesehen. Indem die Ausrichtung der großen Ellipse des Oberkörper betrachtet wird, ist eine Schätzung der Richtung möglich, in der die Person schaut. Ein Tracking der Personen wird in dieser Publikation allerdings nicht behandelt. In den Ergebnissen wird diese Art der Personendetektion als besonders robust beschrieben. Dieser Detektionsmechanismus scheint vielversprechend und robust zu sein, benötigt allerdings eine Vielzahl von Laserscannern. Da in dieser Masterarbeit ein großes Areal beobachtet werden soll und somit schon viele Laser-Range-Scanner vonnöten sind, um eine Scanebene abzudecken, wurde davon abgesehen dieses Verfahren für zwei Scanebenen umzusetzen, zumal auch kein passender Trackingmechanismus beschrieben wurde.

2.4 Repräsentation der Personenspuren

Die Bewegungstrajektorien von Personen können auf unterschiedliche Art und Weise repräsentiert werden. Die einfachste Variante geht davon aus, dass die Position einer Person zu bestimmten Zeitpunkten ermittelt wird und zusammen mit diesem Zeitpunkt in einer Liste fortlaufend gespeichert wird. Somit ist die Bewegungsspur in Form von kontinuierlichen Positionen, oder Raumkoordinaten zu diskreten Zeitpunkten beschrieben. Bei fast allen Verfahren zur Positionsbestimmung von Personen werden die Positionen im bestimmten Takt ermittelt, wodurch diese Form der Bewegungsrepräsentation als grundlegend angesehen werden kann. In der Masterarbeit wurde diese Repräsentationsform auch verwendet. Sobald man jedoch Bewegungsspuren miteinander vergleichen will, oder eine Klassifikation vornehmen möchte, sind andere Formen von Vorteil. Zwei Möglichkeiten, welche in der praktischen Umsetzung der Masterarbeit jedoch nicht zur Verwendung kamen, da die einfache Repräsentation durch Raumkoordinaten zu diskreten Zeitpunkten ausreichend ist, werden im Folgenden vorgestellt. Eine Repräsentationsform von Trajektorien auf Basis von Polynomapproximationen

wurde in [Melo et al. 2006] und in [Naftel und Khalid 2006] vorgestellt. Dabei werden die Koordinaten der Person als Funktionen der Zeit aufgefasst und versucht, diese Funktionen durch eine Summe von Polynomen anzunähern (siehe dazu Abschnitt 3.8.1). Durch Minimierung einer Fehlerfunktion werden die Polynomkoeffizienten ermittelt, welche dann die Trajektorie repräsentieren. Da diese Trajektorie durch eine lineare Kombination von Polynomen angenähert wird, ist es auch möglich, zwischen zwei Trajektorien zu interpolieren, indem man zwischen deren Polynomkoeffizienten interpoliert. Allerdings tritt bei der Polynomapproximation eine als Runges Phänomen bekannte Schwierigkeit auf. Diese äußert sich durch einen hohen Approximationsfehler an den Intervallgrenzen der Funktionen.
In [Naftel und Khalid 2006] wird ebenfalls eine Approximation der Trajektorien mit der Hilfe von Chebychev-Polynomen thematisiert. Diese, auch in [Mason und Handscomb 2003] ausführlich behandelten, Polynome minimieren Runges Phänomen, welches sich durch eine starke Oszillation am Rand der approximierten Funktion bei hohen Polynomgraden bemerkbar macht, durch eine geeignete Wahl der Stützstellen. Die Approximation anhand einer Summe von Chebychev-Polynomen wird auch in dieser Arbeit im Abschnitt 3.8.2 näher erläutert. Die Trajektorie wird dabei durch die Summenkoeffizienten beschrieben, welche es ermöglichen, zwei Trajektorien miteinander zu vergleichen und auch zwischen zwei Trajektorien zu interpolieren, da die Approximation eine lineare Operation darstellt.

2.5 Zusammenfassung

Im Bereich des Fußtrackings sind die Arbeiten von [Shao et al. 2007] gut für das Szenario eines Flughafens geeignet, da deren Verfahren bereits in öffentlichen Bereichen erfolgreich eingesetzt wurden. Die Arbeiten von [Schulz et al. 2001], [Topp und Christensen 2005] und [Lee et al. 2006] wurden für einen mobilen Roboter angepasst und verwenden nur einen Laser-Range-Scanner, wodurch keine größeren Areale beobachtet werden können. Das Verfahren von [Chakravarty und Jarvis 2006] benötigt zusätzlich noch eine Videokamera und gibt keine Aussage, wie es sich bei einem hohen Personenaufkommen verhält. Somit floss nur der Mechanismus von [Shao et al.

2007] in diese Arbeit mit ein, wohingegen von einer Verwendung der anderen Verfahren abgesehen wurde.
Im Bereich des Oberkörpertrackings liefert das Verfahren von [Luo et al. 2007] nur gute Ergebnisse, wenn eine einzelne Person nahe vor dem Laserscanner steht und in dessen Richtung schaut. Zudem wird zusätzlich eine Kamera benötigt. Da das Detektionsverfahren für einen Einsatz auf einem Flughafen, wo Personen in weiter Entfernung und aus jeder Richtung detektiert werden müssen, nicht praktikabel ist, hat die Arbeit von [Luo et al. 2007] keine Verwendung in dieser Arbeit gefunden. Das Detektionsverfahren aus [Carballo et al. 2008] liefert gute Ergebnisse, allerdings benötigt es viele Laser-Range-Scanner und behandelt nur die Personendetektion, ohne auf ein mögliches Tracking einzugehen. Somit wurde auch dieses Verfahren nicht in der Masterarbeit verwendet, womit ein Tracker auf Oberkörperhöhe in dieser Arbeit neu entwickelt werden muss.
Die theoretischen Grundlagen, welche für die Umsetzung des Trackingverfahrens nach [Shao et al. 2007] und für die Entwicklung des Oberkörpertrackers nötig sind, werden im folgenden Kapitel beschrieben.

Kapitel 3

Theoretische Grundlagen

3.1 Einleitung

Nachdem im vorherigen Kapitel der aktuelle Stand der Forschung umrissen und die Ziele dieser Arbeit dargelegt wurden, sollen im folgenden Kapitel die theoretischen Grundlagen für die Umsetzung dieser Ziele dargelegt werden. Dabei wird zuerst auf die Aufbereitung der Scandaten eingegangen (Abschnitt 3.2 bis 3.6), damit auf Basis dieser Daten Personen detektiert und getrackt werden können. Dem folgt die theoretische Aufarbeitung des Partikelfilters, dem essentiellen Bestandteil des Personentrackers (Abschnitt 3.7). Daraufhin werden die dafür notwendigen Bewegungsmodelle erläutert (ebenfalls im Abschnitt 3.7), gefolgt von Mechanismen, um die ermittelten Bewegungsspuren der Personen in geeigneter Form repräsentieren zu können (Abschnitt 3.8).

3.2 Funktionsweise von Laser-Range-Scannern

Um Personen tracken zu können, müssen diese zuerst durch Sensoren erfasst werden. Einen geeigneten Sensor stellt der Laser-Range-Scanner dar, welcher im Wesentlichen ein Messgerät ist, das die Laufzeit eines Laserimpulses vom Gerät bis zu einem Hindernis und wieder zurück misst und anhand dessen die Entfernung bestimmt. Wenn t_l die gemessene Laufzeit des Laserimpulses ist, so lässt sich die Entfernung d zum Hindernis nach Gleichung 3.1 berechnen, wobei c gleich der Lichtgeschwindigkeit ist

Abbildung 3.1: In den Experimenten verwendeter Laserscanner LMS151 der Firma SICK.

(Abbildung 3.2 verdeutlicht die Messung).

$$d = \frac{t_l \cdot c}{2} \tag{3.1}$$

Diese Messungen werden allerdings nicht nur in eine bestimmte Richtung vorgenommen. Am konkreten Beispiel des LMS151 der Firma SICK (Abbildung 3.1), welcher auch in den Experimenten (siehe Kapitel 6) verwendet wurde, wird der Laserstrahl zwischen den einzelnen Messungen durch einen rotierenden Spiegel in 0,5° Schritten horizontal abgelenkt. Bei insgesamt 541 einzelnen Schritten ergibt sich ein Scanbereich

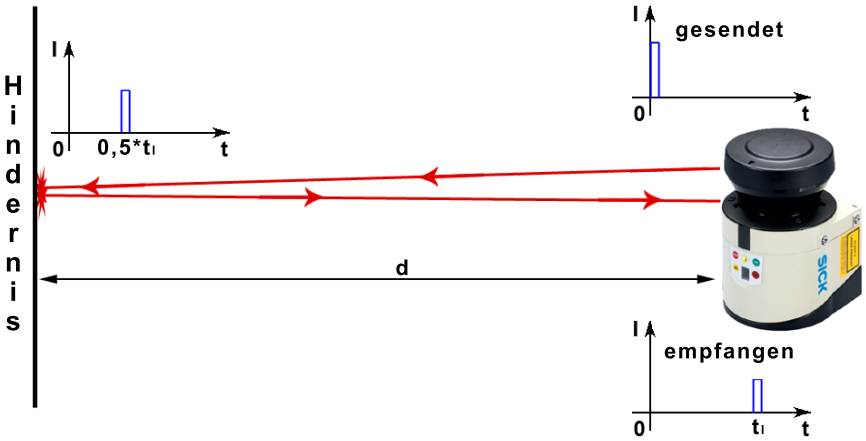


Abbildung 3.2: Entfernungsmessung durch Laufzeitmessung eines Laserimpulses. Das Messgerät auf der rechten Seite sendet einen Laserimpuls aus. Zum Zeitpunkt $0,5 \cdot t_l$ trifft dieser auf das Hindernis in der Entfernung d und wird reflektiert. Nachdem der reflektierte Impuls die Strecke d nochmals durchlaufen hat, wird er zum Zeitpunkt t_l vom Messgerät registriert.

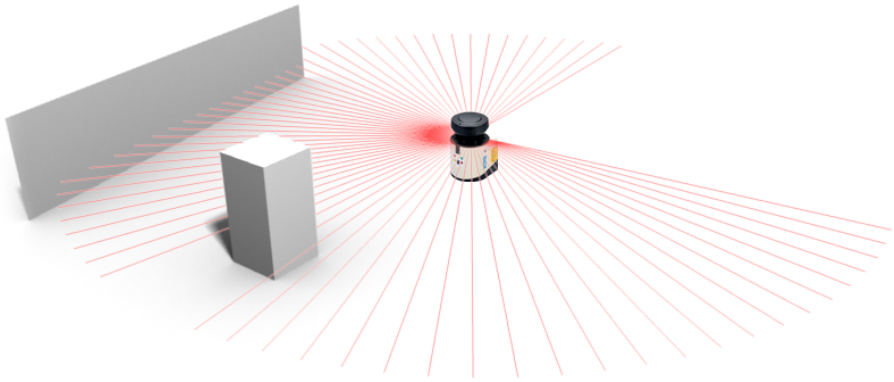
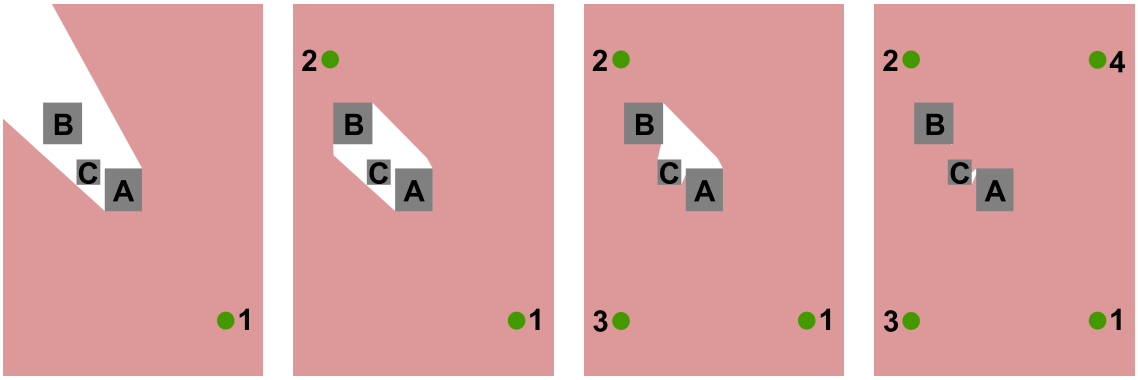


Abbildung 3.3: 541 Messschritte mit 0,5° Schrittwinkel ermöglichen es eine Scanebene mit einem Öffnungswinkel von 270° zu erfassen.

von 270° in der Ebene um den Laser-Range-Scanner (Abbildung 3.3). Ein kompletter Scan wird innerhalb von 20ms durchgeführt, wodurch es möglich ist, 50 Scans pro Sekunde zu erhalten. Die Daten der einzelnen Scans beinhalten die Entfernungen der einzelnen Messrichtungen, mitsamt der Remittanzen der Hindernisse, welche vermessen wurden. Die Remittanz beschreibt hierbei, wieviel des ausgesendeten Laserlichts wieder empfangen wurde. Des Weiteren sind auch Statusinformationen des Messgerätes und die Zeitstempel der einzelnen Scans in den Daten enthalten. Diese werden dann per Ethernet an einen Rechner übertragen, wo sie gespeichert und weiter verarbeitet werden.

3.3 Kombination mehrerer Laserscanner

Da es in Szenarien mit erhöhtem Personenaufkommen, wie z.B. auf Flughäfen, vorkommen kann, dass eine Person eine oder mehrere andere Personen aus Sicht eines Laser-Range-Scanners verdeckt, ist es sinnvoll die Szene mit mehreren Laser-Range-Scannern von unterschiedlichen Positionen aus zu beobachten (siehe Bild 3.4). Dies erfordert allerdings, dass die Scanner untereinander synchronisiert werden, sowohl räumlich, als auch zeitlich. Die zeitliche Synchronisierung lässt sich, unter Vernachlässigung der benötigten Zeit der Datenübertragung vom Laser-Range-Scanner zum Rechner, anhand



(a) ein Scanner (b) zwei Scanner (c) drei Scanner (d) vier Scanner

Abbildung 3.4: Erfassung der Szene mit mehreren Laserscannern. In den Draufsichten sind die Scanner grün, Objekte grau und der erfasste Bereich blassrot eingetragen. In (a) kann Scanner 1 nur Objekt A erfassen. B und C werden von A verdeckt. In (b) ermöglicht Scanner 2 die Erfassung von Objekt B. Objekt C ist jedoch noch immer verdeckt und wird somit auch nicht erfasst. Erst durch Hinzunahme des Scanners 3 in (c) wird auch Objekt C registriert. Mit insgesamt 4 Laser-Range-Scannern lässt sich die konkrete Szene in (d) fast vollständig erfassen, wobei es jedoch noch immer kleine Bereiche der Verdeckung gibt.

der Zeitpunkte des Eintreffens der Datenpakete am Zielrechner realisieren. Die räumliche Synchronisierung hingegen ist mit mehr Aufwand verbunden.

Um aus den Entfernungsdaten eines einzelnen Scanners auf Punkte in einem globalen Koordinatensystem zu schließen, muss einerseits die Position des Laser-Range-Scanners im Raum, als auch dessen Ausrichtung bekannt sein. Unter der Voraussetzung, dass der Laser-Range-Scanner auf einer Höhe in der horizontalen Ebene scannt, lässt sich dessen Pose durch $\mathbf{s} = (x, y)^T$ und ω beschreiben, wobei x und y die Koordinaten im Raum und ω die Drehung um die z-Achse ist. Angenommen, dass in jeder Richtung eine Entfernung bestimmt wurde, ergeben sich für einen einzelnen Scan insgesamt 541 Punkte $\mathbf{p}_i$, $i = 0, \ldots, 540$, die im Raum gemessen wurden. Diese Punkte errechnen sich nach

$$\mathbf{p}_i = \begin{pmatrix} \cos\left(i \cdot 0.5^\circ - \omega\right) \\ \sin\left(i \cdot 0.5^\circ - \omega\right) \end{pmatrix} \cdot d_i + \mathbf{s}, \tag{3.2}$$

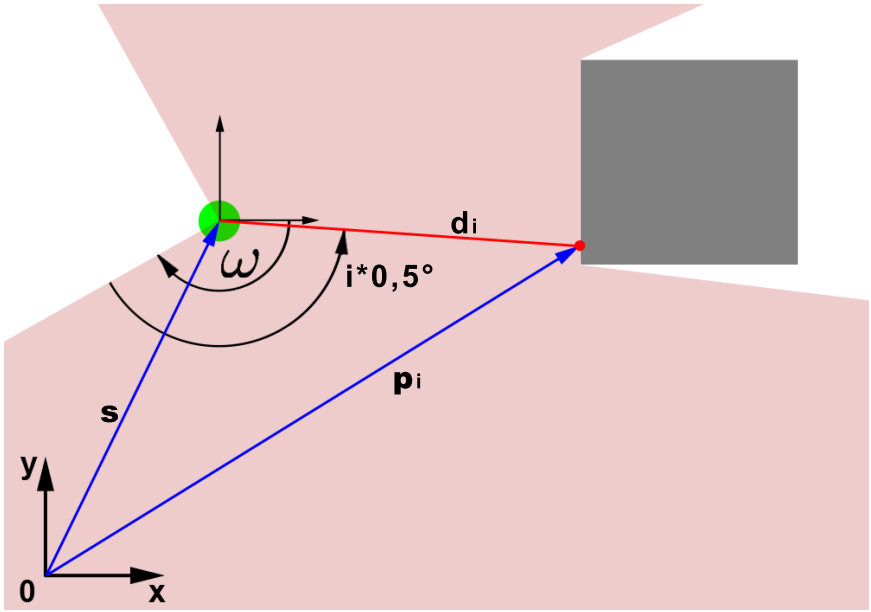


Abbildung 3.5: Berechnung der globalen Punktkoordinaten. In dieser Draufsicht sind der Laserscanner grün, das Objekt grau und der vom Scanner erfasste Bereich blassrot eingetragen. Der eingetragene Scanner hat einen Öffnungswinkel von 270° und scannt von rechts nach links. Der rechte Rand des Erfassungsbereiches ist um den Winkel ω gegenüber der X-Achse gedreht. Der Laserscanner befindet sich an der Position $\mathbf{s}$. Die i-te Messung ergab die Entfernung d_i, welche mit dem Punkt $\mathbf{p}_i$ korrespondiert. Dabei ist der Messstrahl um $i \cdot 0,5°$ Grad vom rechten Rand des Erfassungsbereiches gedreht. Durch Gleichung 3.2 lässt sich der gemessene Punkt $\mathbf{p}_i$ berechnen.

wobei d_i der Entfernungswert der jeweiligen Richtung ist (Abbildung 3.5 verdeutlicht diese Berechnung).

Sofern die Positionen und Ausrichtungen aller Laser-Range-Scanner bekannt sind, können alle Entfernungsmessungen anhand der Gleichung 3.2 in Punkte im globalen Koordinatensystem überführt werden, wodurch dann eine räumliche Synchronisierung gegeben ist. In dieser Punktemenge können nun Personen detektiert und getrackt werden.

3.4 Hintergrundsegmentierung

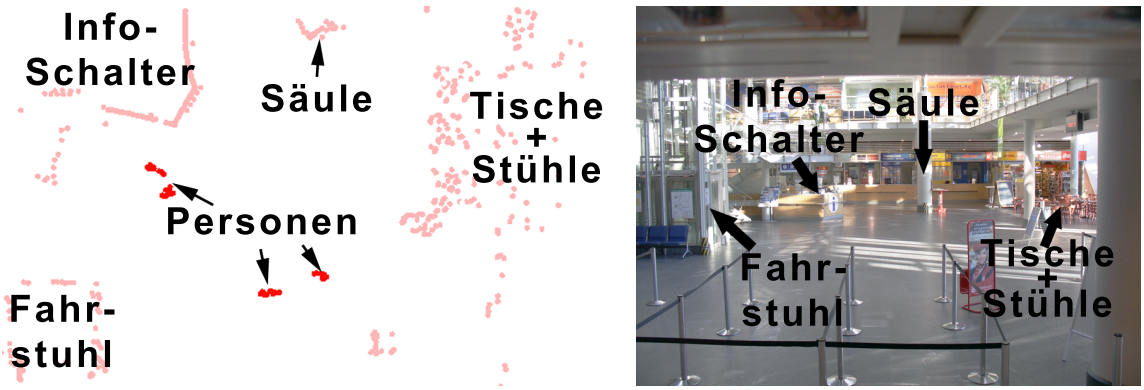
Vor der eigentlichen Detektion von Personen ist es sinnvoll die Punktewolke von den statischen Punkten, welche z. B. durch Wände und andere unbewegliche Objekte hervorgerufen werden und als Hintergrund der beobachteten Szene zu betrachten sind, zu befreien (siehe Bild 3.6). Somit reduziert sich einerseits die zu verarbeitende Punktemenge, was unnötige Berechnungen erspart, und die Detektion wird vereinfacht, da Personen, auch wenn sie zeitweise stillstehen, über einen längeren Zeitraum betrachtet als dynamisch angenommen werden können.

Der gewählte Ansatz zur Hintergrundsegmentierung findet in der Datenverarbeitungskette schon recht früh bei den Entfernungsdaten d_i, $i = 0, \ldots, 540$ der Laserscanner statt. Unter der Voraussetzung, dass die einzelnen Laser-Range-Scanner sich während des Einsatzes nicht bewegen, lässt sich der Hintergrund dadurch definieren, dass er die maximal gemessene Entfernung $d_{i,max}$ in den einzelnen Messrichtungen, bei genügend langer Beobachtung der Szene, hervorruft. Da die Laser-Range-Scanner eine Messunsicherheit besitzen und somit die Messergebnisse verrauscht sind, kann nicht einfach beim Auftreten der maximalen Entfernung der dazugehörige Punkt im globalen Koordinatensystem als Hintergrund klassifiziert werden. Stattdessen ist es sinnvoller für jede Messrichtung einen Schwellwert τ_i zu definieren, der um eine bestimmte Distanz β geringer ist, als die maximal gemessene Entfernung (Gleichung 3.3). Alle Messungen, welche größer als die Schwelle sind, werden als Hintergrund klassifiziert und alle anderen Messungen als Vordergrund (Gleichung 3.4).

$$\tau_i = d_{i,max} - \beta \tag{3.3}$$

$$d_i \in \begin{cases} \text{Vordergrund}, & \text{wenn } d_i \leq \tau_i \\ \text{Hintergrund}, & \text{wenn } d_i > \tau_i \end{cases} \tag{3.4}$$

Diese Sicherheitsdistanz sollte nicht allzu groß gewählt werden, weil sonst fälschlicherweise Teile von Personen als Hintergrund erkannt werden können. Zudem kann es in vereinzelten Fällen auch zu Messungen kommen, die weit größer als die tatsächliche Entfernung zum Hintergrund sind. Solche Messungen treten zum Beispiel bei spiegelnden Oberflächen auf, oder wenn ein Objekt nicht genug Laserlicht reflektiert, um



(a) (b)

Abbildung 3.6: Ziel der Hintergrundsegmentierung ist das Trennen von Scanpunkten statischer Objekte, z. B. Wände und Möbel, von den Scanpunkten dynamischer Objekte. In der Draufsicht (a) sind die Hintergrundpunkte blassrot und die Vordergrundpunkte rot eingetragen. In (b) ist die selbe Szene als Foto ohne Personen zu sehen.

vom Laser-Range-Scanner erfasst zu werden. Aus diesem Grund können die maximal gemessenen Distanzen nicht als Kriterium für die Schwellwerte genommen werden. Eine Möglichkeit, um diesem Problem zu begegnen, liefern Histogramme. Ein Histogramm ist eine diskretisierte Häufigkeitsverteilung einer Zufallsgröße. Im konkreten Fall wird für jede Messrichtung i ein Histogramm angelegt, welches die Entfernung d in Blöcke, oder auch Bins, $h_{i,j}$, $j = 1, \ldots, H$ gleicher Länge unterteilt, in denen gezählt wird, wie oft ein Messwert in dem jeweiligen Bereich des Bins lag (in Abbildung 3.7 ist ein Histogramm, welches während der Experimente entstand, zu sehen). Eine Blocklänge von 10cm hat sich als brauchbar herausgestellt, was bei der maximalen Reichweite d_{max} der verwendeten Laser-Range-Scanner von 50m (wenn ein Hindernis weiter vom Scanner entfernt ist, werden trotzdem 50m gemessen) in $H = 500$ Bins resultiert.

$$h_{i,j}^{neu} = h_{i,j}^{alt} + 1, \text{wenn } \frac{j \cdot d_{max}}{H} \leq d_i < \frac{(j+1) \cdot d_{max}}{H} \tag{3.5}$$

Nachdem genügend Messungen in die Histogramme eingetragen wurden (siehe Gleichung 3.5), kann die Schwelle τ_i errechnet werden, indem der unterste Rand des Bins bestimmt wird (Gleichung 3.6 und 3.7), ab welchem mehr als eine bestimmte relative

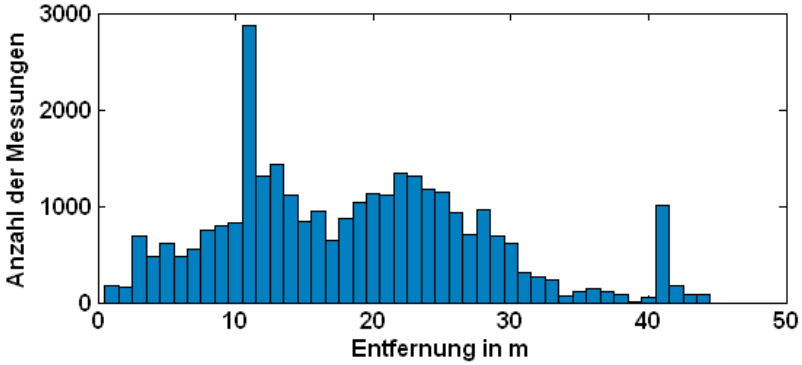


Abbildung 3.7: Histogramm der Entfernungsmessung in eine bestimmte Richtung mit einer Blocklänge von 1m. Das Maximum bei etwa 40m deutet auf ein statisches Objekt, im konkreten Fall eine Wand, hin. Die Messungen aus den weiter entfernten Bins werden durch Messrauschen verursacht. Die vielen Messungen mit einer Entfernung geringer als 40m stammen von Personen, welche sich zwischen der Wand und dem Laserscanner bewegt haben.

Menge μ der Messungen kleiner oder gleich diesem Rand sind und einen Sicherheitsabstand β, wie auch im vorherigen Absatz, abgezogen wird (Gleichung 3.8).

$$h_i = \sum_{j=1}^{H} (h_{i,j}) \tag{3.6}$$

$$j_{min} := \sum_{k=1}^{j_{min}} \left(\frac{h_{i,k}}{h_i} \right) < \mu \leq \sum_{k=1}^{j_{min}+1} \left(\frac{h_{i,k}}{h_i} \right) \tag{3.7}$$

$$\tau_i = \frac{j_{min} \cdot d_{max}}{H} - \beta. \tag{3.8}$$

In den Experimenten aus Kapitel 6 haben sich die empirisch ermittelten Werte $\mu = 0.95$ und $\beta = 40$cm als sinnvoll erwiesen. Der Wert von $\mu = 0.95$ beschreibt, dass das Bin, ab welchem 95% aller Punkte im Histogramm aufsummiert wurden, als Hintergrundschwelle gewählt wird. Der Sicherheitsabstand von $\beta = 40$cm legt die Schwelle 40cm vor dieses Bin, damit möglichst viele der Hintergrundmessungen auch als solche erkannt werden, aber Personen, welche dicht an dem Hindernis entlanggehen, nicht schon als Hintergrund klassifiziert werden. Sobald für alle Messrichtungen aller Laser-Range-Scanner die jeweilige Schwelle τ_i bestimmt und damit die Messung nach Hintergrund

und Vordergrund gemäß Gleichung 3.4 klassifiziert wurde, können die Vordergrundmessungen, wie in Abschnitt 3.3 beschrieben, in eine Punktewolke im globalen Koordinatensystem überführt werden.

3.5 Mean-Shift Clustering

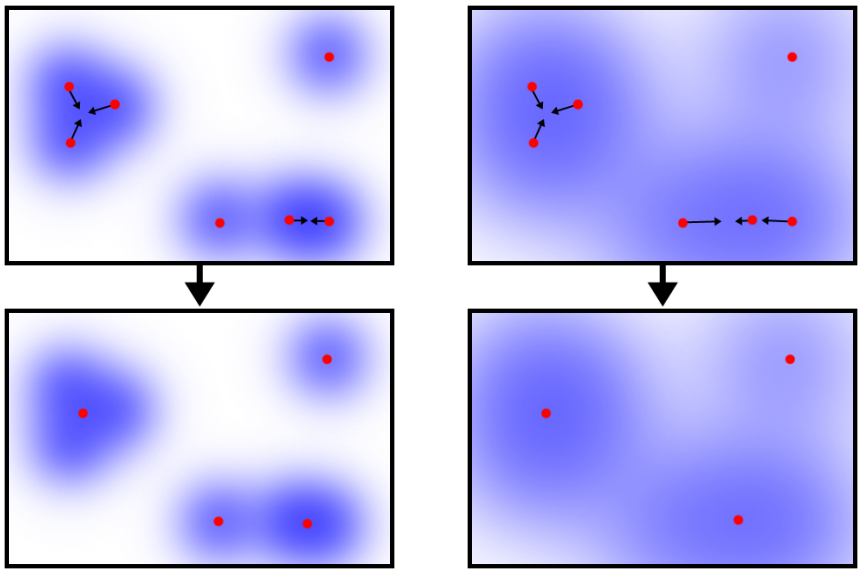
Um Personen detektieren zu können, ist es nutzbringend, in der Punktemenge nach lokalen Ansammlungen zu suchen, welche einzelne Objekte darstellen können. Der Mean-Shift-Algorithmus findet in Punktewolken Anhäufungen von Punkten, auch Cluster genannt, indem er die Punkte iterativ in Richtung deren Masse- oder auch Dichteschwerpunkte zusammenzieht [Fukunaga und Hostetler 1975]. Diese Dichtefunktion wird mit Hilfe einer Kernelfunktion gebildet, die über jeden einzelnen Punkt gelegt wird und sich additiv überlagert (In Abbildung 3.8 blau eingetragen). Diese Kernelfunktion bestimmt dabei, wie die jeweiligen Punkte zur Dichte beitragen. In den Iterationen werden dann sämtliche Punkte in Richtung des Gradienten der Dichtefunktion an der jeweiligen Punktposition verschoben. Als Kernel wird üblicherweise die Gaußfunktion $G(x)$ (siehe Gleichung 3.9) verwendet. Für die Punkte $\mathbf{p}_i$, $i = 1, \ldots, N$ berechnet sich eine Iteration nach Gleichung 3.11, wobei $g\,(x)$ die Ableitung der Kernelfunktion ist (siehe Gleichung 3.10).

$$G\,(x) = \exp\left(-\frac{1}{2}x^2\right) \tag{3.9}$$

$$g\,(x) = -G'\,(x) \tag{3.10}$$

$$\mathbf{p}_i' = \frac{\sum\limits_{j=1}^{N} \mathbf{p}_j \cdot g\left(\left\|\frac{\mathbf{p}_i - \mathbf{p}_j}{\gamma}\right\|\right)}{\sum\limits_{j=1}^{N} g\left(\left\|\frac{\mathbf{p}_i - \mathbf{p}_j}{\gamma}\right\|\right)} \tag{3.11}$$

Der Parameter γ steuert hierbei die Breite des Gaußkernels. Ein hoher Wert von γ bewirkt, dass die einzelnen Punkte auch noch in weiter Entfernung einen Beitrag zur Dichtefunktion liefern, wohingegen ein kleiner Wert von γ den Einfluss auf die Dichte auf eine geringe Entfernung beschränkt (vergleiche Abbildung 3.8). Nach jeder Iteration wird $\mathbf{p}_i$ gleich $\mathbf{p}_i'$ gesetzt und so lange wiederholt, bis sich kein Punkt mehr bewegt,



(a) kleiner Kernel (b) großer Kernel

Abbildung 3.8: Auswirkung der Kernelfunktion. In den Bildern sind die Scanpunkte rot und die Dichtefunktion blau eingezeichnet. Jeder Punkt trägt zur Dichtefunktion mittels der Kernelfunktion bei. Die Punkte in den oberen Bildern werden zu den Clusterzentren in den unteren Bildern zusammengezogen. In (a) wurde ein Gaußkernel mit kleiner Kernelbreite γ (siehe Gleichung 3.11) gewählt. Der eine, etwas weiter weg stehende Punkt rechts unten wurde aufgrund der geringen Kernelbreite nicht zu den anderen beiden Punkten gezogen und bildet sein eigenes Cluster. In (b) wurde eine große Kernelbreite gewählt, wodurch alle drei Punkte rechts unten zu einem Cluster zusammengefasst wurden.

oder eine maximale Anzahl an Iterationen durchgeführt wurde. Sobald das Abbruchkriterium erreicht wurde, bilden die jeweils übereinanderliegenden Punkte die Zentren der somit gefundenen Cluster.

Der Algorithmus kann auch so verstanden werden, dass sich alle Punkte gegenseitig anziehen. Dabei bestimmt der Kernel $G(x)$ aus Gleichung 3.9, wie sich die Anziehungskraft zwischen zwei Punkten in Abhängigkeit zur Entfernung untereinander verhält. Indem die Kraft nur auf eine lokale Umgebung beschränkt ist, ziehen sich nahe beieinander liegende Punkte an und weit auseinander liegende Punkte haben so gut wie

keine Wechselwirkung aufeinander. Dadurch bilden sich nach einigen Iterationen Punktemengen, deren Punkte aufeinander liegen. Diese Punkte gehören dann zum gleichen Cluster und der Ort, an dem sich die Punkte zusammengezogen haben, bildet den Mittelpunkt des Clusters.

3.6 Hauptkomponentenanalyse

Um die Ausdehnung von Punktemengen zu bestimmen, wie es bei der Personendetektion im Abschnitt 4.4 benötigt wird, kann die Hauptkomponentenanalyse verwendet werden. Die Hauptkomponentenanalyse ist eine Koordinatentransformation, welche eine Verteilung von Punkten $\mathbf{v}_i$, $i = 1, \ldots, Z$ der Dimension W in ein neues Koordinatensystem projiziert, in welchem es die geringste Korrelation zwischen den einzelnen Dimensionen besitzt (siehe Abbildung 3.9). Sie wurde erstmals in [Pearson 1901] vorgestellt.

Für die Hauptkomponentenanalyse ist es notwendig, die Verteilung der Punkte $\mathbf{v}_i$

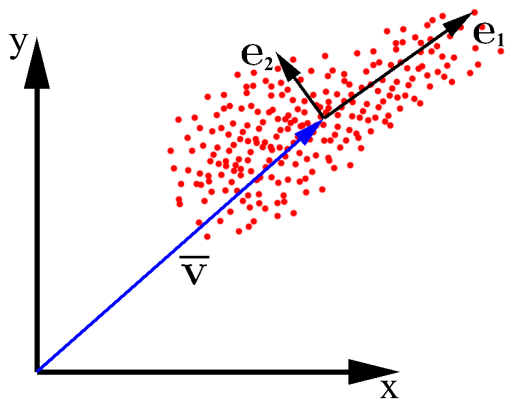


Abbildung 3.9: Veranschaulichung der Hauptkomponentenanalyse. Die rote Punktemenge stellt die zu untersuchende Verteilung mit dem Mittelwert $\overline{\mathbf{v}}$ dar. Durch die Hauptkomponentenanalyse werden die Raumachsen $\mathbf{e}_1$ und $\mathbf{e}_2$, welche auch als Eigenvektoren bezeichnet werden, ermittelt. Diese sind derart ausgerichtet, dass die Verteilung in deren Raum eine möglichst geringe Kovarianz aufweist. Die erste Raumachse liegt dabei entlang der größten Varianz der Verteilung.

durch ihre Kovarianzmatrix $\mathbf{C}$ nach Gleichung 3.12 und 3.13 zu beschreiben.

$$\overline{\mathbf{v}} = \frac{1}{Z} \sum_{i=1}^{Z} \mathbf{v}_i \tag{3.12}$$

$$\mathbf{C} = \frac{1}{Z-1} \sum_{i=1}^{Z} (\mathbf{v}_i - \overline{\mathbf{v}}) (\mathbf{v}_i - \overline{\mathbf{v}})^T \tag{3.13}$$

Diese Matrix beschreibt, inwiefern die einzelnen Dimensionen der Verteilung voneinander abhängen. Dabei beschreibt der Wert an der Stelle (j, k) der Matrix $\mathbf{C}$, wie die j-te Komponente des Vektors $\mathbf{v_i}$ von der k-ten Komponente abhängt. Ein Wert von 1 signalisiert, dass eine vollständige lineare Abhängigkeit vorhanden ist, wohingegen ein Wert von 0 auf Unabhängigkeit deutet. Ein Wert von -1 tritt auf, wenn eine antiproportionale Abhängigkeit vorhanden ist. Es ist ersichtlich, dass die Kovarianzmatrix $\mathbf{C}$ auf der Hauptdiagonalen nur Einsen enthält, da jede Komponente des Vektors $\mathbf{v_i}$ von sich selbst abhängig ist.

Nachdem die Kovarianzmatrix bestimmt wurde, müssen ihre Eigenvektoren und Eigenwerte bestimmt werden. Dies geschieht durch Lösen des Eigenwertproblems von $\mathbf{C}$, welches durch Gleichung 3.14 definiert wird. Das Eigenwertproblem ist die Suche nach allen von Null verschiedenen Vektoren $\mathbf{e}$, welche durch eine Multiplikation mit $\mathbf{C}$ auf sich selbst abgebildet werden und sich dabei höchstens um ein Skalar η unterscheiden.

$$\mathbf{C}\mathbf{e} = \eta \mathbf{e}, \qquad \mathbf{e} \neq 0 \tag{3.14}$$

Die Gleichung 3.14 lässt sich in Gleichung 3.15 umformen, so dass $\eta_1, \ldots, \eta_W$ durch Berechnen der Nullstellen der linken Seite der Gleichung ermittelt werden kann. Dabei ist $\mathbf{E}$ die Einheitsmatrix.

$$\det (\mathbf{C} - \eta \mathbf{E}) \, \mathbf{e} = 0 \tag{3.15}$$

Diese Nullstellen $\eta_1, \ldots, \eta_W$ werden auch als Eigenwerte der Matrix $\mathbf{C}$ bezeichnet. Setzt man diese in Gleichung 3.14, so lassen sich W orthogonal zueinander liegende Lösungen für $\mathbf{e}$ finden. Diese Vektoren $\mathbf{e}_1, \ldots, \mathbf{e}_W$ werden auch als Eigenvektoren, oder Hauptkomponenten bezeichnet. Sie bilden die Basis des Koordinatensystems, in welchem die Verteilung der Punkte die geringste Korrelation zwischen den Dimensionen besitzt.

3.7 Partikelfilter

Zum Ermitteln von Bewegungsspuren ist es notwendig, die Detektionen zu den einzelnen Zeitschritten den jeweiligen Personen zuzuordnen und fehlende Detektionen zu kompensieren. Dies wird von einem Trackingalgorithmus geleistet. Die Methodik des Trackings lässt sich als Schätzung eines Zustandes x_t zum Zeitpunkt t anhand vorangegangener Beobachtungen $z_0, \ldots, z_t$ und Aktionen $u_0, \ldots, u_{t-1}$ beschreiben. Da sämtliche Beobachtungen mit Unsicherheiten behaftet sind, bietet sich ein probabilistischer Zustandsschätzer nach dem Bayes-Filter, wie er in [Thrun et al. 2005] beschrieben wird, zum Taxieren des Zustandes an. Dieser schätzt den aktuellen Zustand x_t in Form einer Wahrscheinlichkeitsdichte $\mathrm{bel}\,(x_t)$, welche sich aus den vorangegangenen Beobachtungen und Aktionen nach Gleichung 3.16 bildet.

$$\mathrm{bel}(x_t) = p\,(x_t | z_t, u_{t-1}, z_{t-1}, u_{t-2}, \ldots, u_0, z_0) \tag{3.16}$$

Die Wahrscheinlichkeitsdichte $\mathrm{bel}\,(x_t)$ ist also für jeden Zustand x_t die Wahrscheinlichkeit, dass er vorliegt, wenn vorher die Beobachtungen $z_0, \ldots, z_t$ gemacht wurden und die Aktionen $u_0, \ldots, u_{t-1}$ getätigt wurden. Unter der Annahme, dass das System ein Markov-Prozess ist, also dass der Zustand des Systems nur vom vorherigen Zustand x_{t-1} und der ausgeführten Aktion u_{t-1} abhängig ist (siehe Gleichung 3.17), lässt sich die Wahrscheinlichkeitsdichte durch eine rekursive Vorschrift nach Gleichung 3.18 und 3.19 bilden.

$$p\,(x_t | z_t, u_{t-1}, z_{t-1}, u_{t-2}, \ldots, u_0, z_0) = p\,(x_t | z_t, u_{t-1}) \tag{3.17}$$

$$\overline{\mathrm{bel}}\,(x_t) = \int p\,(x_t | x_{t-1}, u_{t-1}) \cdot \mathrm{bel}\,(x_{t-1})\, dx_{t-1} \tag{3.18}$$

$$\mathrm{bel}\,(x_t) = \alpha \cdot p\,(z_t | x_t) \cdot \overline{\mathrm{bel}}\,(x_t) \tag{3.19}$$

Dabei ist $\overline{\mathrm{bel}}\,(x_t)$ die prädizierte Wahrscheinlichkeitsdichte, welche den erwarteten Zustand anhand des letzten Zustandes x_{t-1}, der letzten Aktion u_{t-1} und vorangegangenen Wahrscheinlichkeitsdichte $\mathrm{bel}\,(x_{t-1})$ beschreibt. $p\,(x_t | x_{t-1}, u_{t-1})$ ist das Bewegungsmodell, welches die Wahrscheinlichkeit beschreibt, mit der sich der vorherige Zustand x_{t-1} bei einer gegebenen Aktion u_{t-1} zu x_t verändert. Sobald eine neue Messung z_t vorliegt,

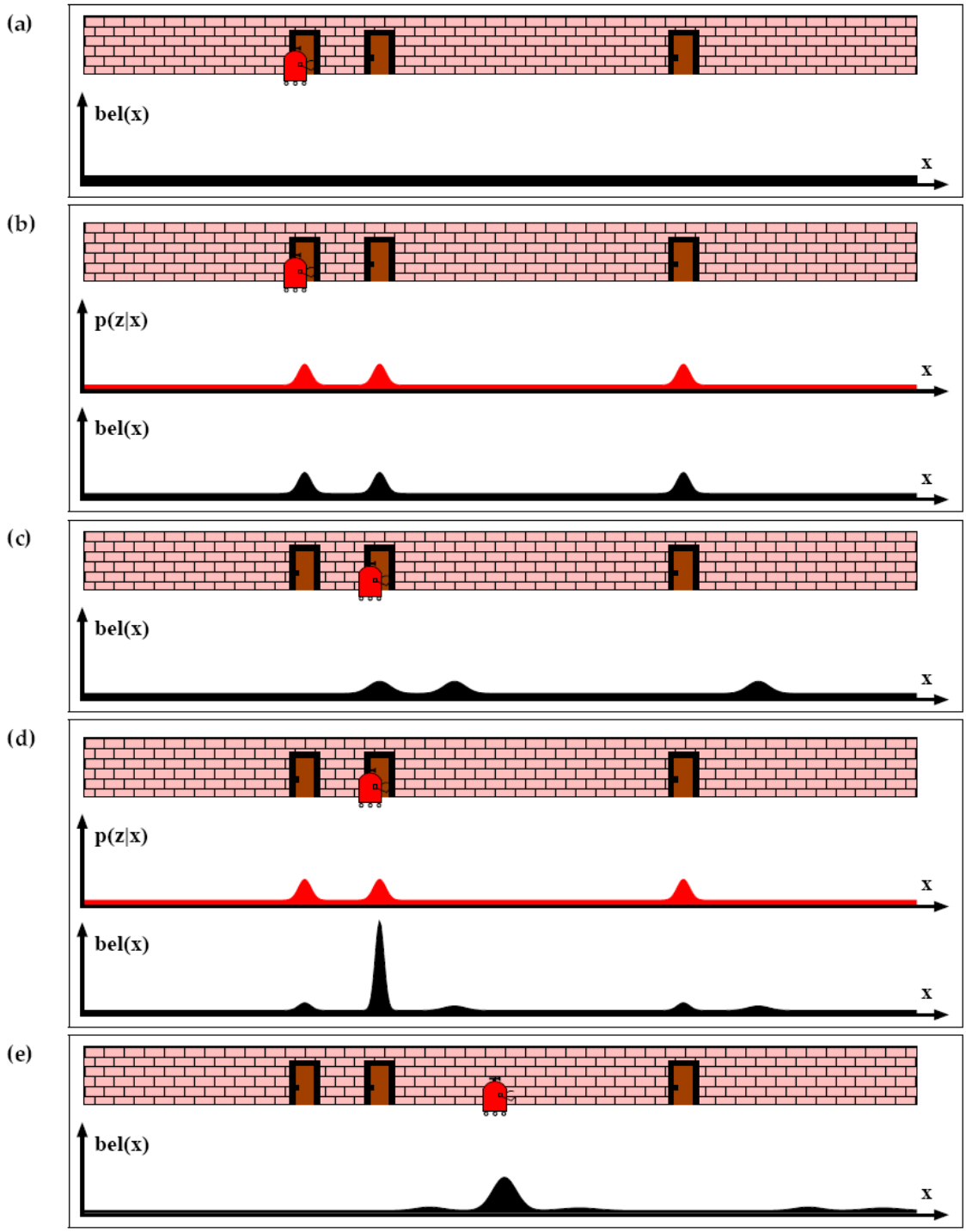


Abbildung 3.10: Veranschaulichung des Bayes-Filter am Beispiel eines Roboters, welcher seine Position anhand von Türen schätzt. Bildquelle: [Thrun et al. 2005]

kann die Prädiktion $\overline{\text{bel}}\,(x_t)$ aus Gleichung 3.18 mit Hilfe der Gleichung 3.19 korrigiert werden, wobei $p\,(z_t|x_t)$ das Beobachtungsmodell ist, welches die Wahrscheinlichkeit beschreibt, dass die Beobachtung z_t im Zustand x_t gemacht wurde. Dadurch können sensorspezifische Messunsicherheiten und Störungen berücksichtigt werden. Die Normalisierungskonstante α wurde in die Formel 3.19 eingebracht um sicherzustellen, dass $\int \text{bel}\,(x_t)\,dx_t = 1$ gilt.

Abbildung 3.10 veranschaulicht den Bayes'schen Zustandsschätzer am Beispiel eines Roboters (rot), welcher seine Position entlang der x-Achse anhand von Türen (braun) schätzt: Anfangs ist in (a) die Position des Roboters unbekannt, da die Wahrscheinlichkeit $\text{bel}\,(x)$ für jede Position gleich ist. Nachdem er in (b) eine Tür detektiert hat, ergeben sich anhand des Beobachtungsmodells $p\,(z|x)$ und dem Wissen über die Positionen aller Türen drei mögliche Orte, wo sich der Roboter aufhalten könnte. Mit dem Beobachtungsmodell können dann die Wahrscheinlichkeiten $\text{bel}\,(x)$ für die Positionen nach Gleichung 3.19 aktualisiert werden. Nun bewegt sich der Roboter in (c) vorwärts, führt also eine Aktion u aus, wodurch Wahrscheinlichkeitsdichtefunktion der Position mit dem Bewegungsmodell nach Gleichung 3.18 mitgeführt wird. Nun hat der Roboter in (d) wieder eine Tür detektiert, wodurch das Beobachtungsmodell wieder drei mögliche Orte liefert und die Positionshypothesen mit Gleichung 3.19 aktualisiert werden. Dabei ist nun ein deutliches Maximum in der Wahrscheinlichkeitsverteilung zu erkennen, welches die wahrscheinlichste Position des Roboters — vor der mittleren Tür — darstellt. Nachdem sich der Roboter in (e) wieder vorwärts bewegt hat, werden nach Gleichung 3.18 die Positionshypothesen anhand des Bewegungsmodells aktualisiert. Dabei bleibt das Maximum in den Hypothesen erhalten und der Roboter ist zuverlässig in der Lage, seine Position zu schätzen.

Eine Anwendung des Bayes-Filter stellt der Partikelfilter dar. Dieser repräsentiert die Wahrscheinlichkeitsverteilung $\text{bel}\,(x_t)$ mit einer endlichen Anzahl von M Partikeln P_t, wobei jeder Partikel für sich einen konkreten Zustand $x_t^{(i)}$ und ein Gewicht $w^{(i)}$ zur Zeit t darstellt (siehe Gleichung 3.20), wobei das Gewicht $w^{(i)}$ der Wahrscheinlichkeit des Zustandes $x_t^{(i)}$ entspricht.

$$\text{bel}\,(x_t) \approx \left\{ \left\langle x_t^{(i)}, w_t^{(i)} \right\rangle | i = 1, \ldots, M \right\} = P_t \tag{3.20}$$

Die Einbringung der Aktion u_t wird durch entsprechende Modifikation des Zustandes $x_t^{(i)}$ der einzelnen Partikel realisiert, was in Rahmen dieser Arbeit jedoch keine Anwendung findet, da die Laser-Range-Scanner statisch sind und keinerlei Aktuatorik auf Seiten des Messequipments involviert ist. Stattdessen werden die Zustände der Partikel nur anhand eines Bewegungsmodells aus ihrem vorherigen Zustand berechnet (siehe Gleichung 3.21). Die Messung z_t fließt ein, indem anhand dessen die Gewichte der Partikel neu berechnet werden (siehe Gleichung 3.22).

$$x_t^{(i)} = f\left(x_{t-1}^{(i)}\right) \tag{3.21}$$

$$w_t^{(i)} \propto p(z_t|x_t^{(i)}) \tag{3.22}$$

Dies entspricht bisher im Weitesten den Ansätzen des Bayes-Filter aus Gleichung 3.18 und 3.19. Der eigentliche Unterschied besteht darin, dass nach Abarbeitung der beiden Schritte ein Resampling durchgeführt wird. Das bedeutet, dass aus der Menge der neu berechneten Partikel $\tilde{P}_t$ eine neue Menge P_{t+1} durch Ziehen mit Zurücklegen erzeugt wird. Dabei werden die Gewichte $w^{(i)}$ als Wahrscheinlichkeit dafür genommen, dass das jeweilige Partikel gezogen wird. Dadurch werden viele der Partikel, die eine relativ unsichere Hypothese repräsentieren, gelöscht, jedoch nicht alle, damit eine gewisse Plastizität gewahrt bleibt. Analog werden viele der Hypothesen mit relativ hoher Wahrscheinlichkeit in die neue Menge übernommen. Um der Nichtlinearität der Zustände Rechnung zu tragen, werden die Zustände $x_t^{(i)}$ der Partikel nach dem Resampling verrauscht.

Im Rahmen dieser Arbeit ist der Zustand x_t ein mehrdimensionaler Vektor, welcher die Bewegungsparameter der zu trackenden Person beschreibt. Je nach dem, ob die Person auf Fußhöhe, oder auf Oberkörperhöhe erfasst wurde, ergeben sich unterschiedliche Parameter, welche in den Abschnitten 4.5.1 und 4.5.2 beschrieben werden. Ebenso variiert das Update (siehe Gleichung 3.21) der Partikel. Die Beobachtungen z_t sind die Menge der im aktuellen Zeitschritt erfassten Vordergrundpunkte (siehe dazu Abschnitt 3.4). Das Resampling der Partikel erfolgt mit einem Low-Variance-Sampler-Algorithmus ([Thrun et al. 2005]), welcher folgendermaßen funktioniert: Wenn P_{t-1} eine Menge von M Partikeln mit den Gewichten $w_{t-1}^{(i)}$, $i = 1, \ldots, M$, ist, dann wird die selektierte Partikelmenge P_t durch Ziehen mit Zurücklegen aus der Menge P_{t-1} gebildet, indem

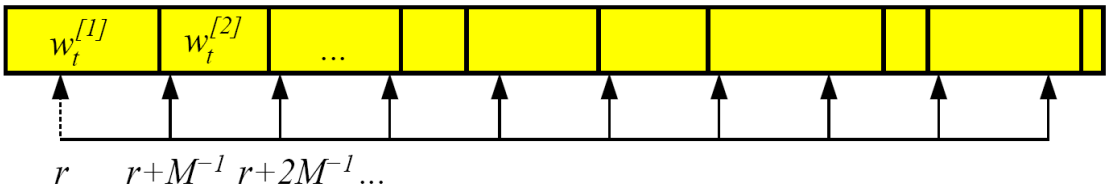


Abbildung 3.11: Veranschaulichung des Low-Variance-Samplers zum Bilden einer neuen Partikelmenge durch Ziehen mit Zurücklegen. Nachdem eine Zufallszahl r aus $\left[0, M^{-1}\right]$ als Schwelle gewählt wurde, werden die normierten Partikelgewichte der Reihe nach aufsummiert. Das Partikel, dessen Gewicht die Summe über die Schwelle brachte, wird gezogen (in diesem Fall das erste Partikel). Danach wird die Schwelle um M^{-1} erhöht und weiter aufsummiert, bis auch diese Schwelle überschritten wurde (hier bei dem zweiten Partikel), und das entsprechende Partikel genommen. Dies wird so lange durchgeführt, bis die Schwelle den Wert 1 überschritten hat. Es ist durchaus möglich, dass ein Partikel mehrfach gezogen wird, wenn sein Gewicht ausreichend groß ist. Bildquelle: [Thrun et al. 2005]

zuerst eine Zufallszahl $r \in [0, M^{-1}]$ gewählt wird. Daraufhin werden die Gewichte $w_{t-1}^{(i)}$ so lange aufsummiert, bis deren Summe größer r ist und der Index m ermittelt (siehe Gleichung 3.23).

$$m = \underset{j}{\operatorname{argmin}} \sum_{i=1}^{j} w_{t-1}^{(i)} \geq r \tag{3.23}$$

Das Partikel an der Stelle m in P_{t-1} wird dann der Menge P_t hinzugefügt. Daraufhin wird r um M^{-1} erhöht und die Gewichte weiter aufsummiert. Dieses Erhöhen und Aufsummieren wird solange wiederholt, bis die Menge P_t insgesamt M Partikel enthält (Abbildung 3.11 veranschaulicht dies). Abschließend werden die Gewichte $w_t^{(i)}$ der Partikel der neuen Menge nach Gleichung 3.24 normiert.

$$w_{t,norm}^{(i)} = \frac{w_t^{(i)}}{\sum\limits_{j=1}^{M} w_t^{(j)}} \tag{3.24}$$

Um aus der Menge der Partikel eine einzelne Personenhypothese x_t zu erhalten, wird nach dem Resampling mit Gleichung 3.25 der Mittelwert von sämtlichen Hypothesen

der Partikel unter Annahme einer unimodalen Verteilung gebildet.

$$x_t = \frac{1}{M} \sum_{i=1}^{M} x_t^{(i)} \tag{3.25}$$

In Abbildung 3.12 wird der Partikelfilter am gleichen Beispiel von Abbildung 3.10 verdeutlicht. Anfangs sind in (a) alle Partikel (hier als senkrechte Striche dargestellt, dessen Position den geschätzten Ort des Partikels repräsentieren und deren Höhe das Gewicht des Partikels angibt) entlang der x-Achse mit dem gleichen Gewicht verstreut. Nachdem die Tür in (b) detektiert wurde, sind mit dem Wissen über die Positionen aller Türen und dem Beobachtungsmodell $p(z|x)$ Aussagen über mögliche Aufenthaltsorte des Roboters möglich. Entsprechend Gleichung 3.22 werden die Gewichte der Partikel, welche die Position nahe der Türen repräsentiert haben, erhöht, wohingegen die Gewichte der Partikel, welche nicht an einer Tür lagen, verringert werden. Im Resamplingschritt, welcher in der Abbildung nicht dargestellt wird, wird die neue Partikelmenge mit dem vorhin beschrieben Low-Variance-Samplers aus der alten Menge anhand der Partikelgewichte gezogen. Es häufen sich die Partikel in der neuen Menge an den Positionen der Türen, da die dortigen Partikel aus der vorherigen Menge wesentlich höhere Gewichte besaßen und somit auch mehrfach gezogen wurden. In (c) hat sich der Roboter vorwärts bewegt und diese Bewegung wurde mit der Hilfe des Bewegungsmodells aus Gleichung 3.21 auf die Partikel übertragen. Die Aufenthaltswahrscheinlichkeit des Roboters entlang der x-Achse wird durch die Dichte der Partikel repräsentiert. Nachdem in (d) eine erneute Beobachtung einer Tür erfolgte, werden die Gewichte der Partikel, welche die Position des Roboters an einer Tür geschätzt haben, wie auch in (b) nach Gleichung 3.22 erhöht. Die Partikel, welche die Position in (b) an der mittleren und rechten Tür geschätzt haben, befinden sich in (d) nicht an einer Tür und erhalten somit ein sehr geringes Gewicht. Jene Partikel, welche zuvor die Position an der linken Tür schätzten, erhalten nun jedoch ein hohes Gewicht, da sie in (d) ebenfalls an einer Tür liegen. Da nun viele Partikel mit hohem Gewicht an der mittleren Tür liegen und nur sehr wenige mit hohem Gewicht die Position an den anderen Türen angaben, werden nach dem Resamplingschritt in der neuen Partikelmenge sehr viele Partikel die Position an der mittleren Tür schätzen. Nachdem sich der Roboter in (e) wieder vorwärts bewegt hat und die Bewegung auch mit dem Bewegungsmodell aus

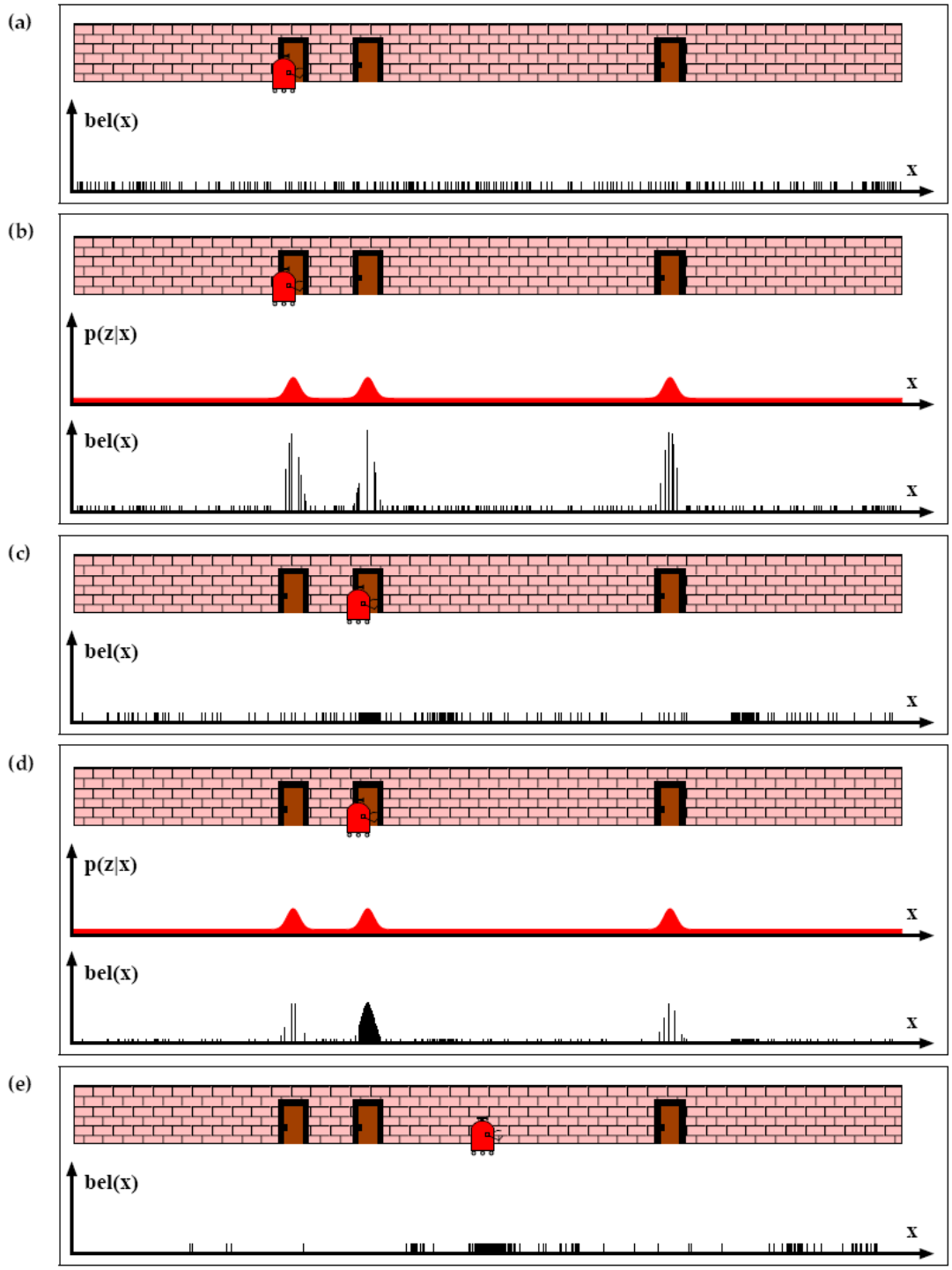


Abbildung 3.12: Veranschaulichung des Partikelfilter am Beispiel eines Roboters, welcher seine Position anhand von Türen schätzt. Bildquelle: [Thrun et al. 2005]

Gleichung 3.21 auf die Partikel übertragen wurde, schätzt eine Vielzahl der Partikel die Position des Roboters annähernd korrekt, wodurch der Roboter eine gute Hypothese zu seiner Position erzeugen kann.

3.7.1 Fußmodell

Zum Tracken von Personen ist es dringend notwendig, deren Bewegungen in einem mathematischen Modell zu beschreiben. Erst damit kann der aktuelle Zustand aus dem des vorherigen Zeitschrittes (siehe Gleichung 3.21) berechnet werden. Dadurch ist es unter Anderem auch erst möglich, im Falle von Verdeckungen die Position der Person zu schätzen. Für Scandaten von Fußknöcheln muss die Bewegung der Füße beschrieben werden. In [Shao et al. 2007] wurde das Modell vorgeschlagen, welches hier unverändert vorgestellt wird — lediglich Gleichung 3.33 bis 3.36 sind aus eigenen Überlegungen entstanden. Bei genauerer Betrachtung der Füße einer gehenden Person lässt sich feststellen, dass sich jeder Fuß periodisch bewegt. Während der linke Fuß fest auf dem Boden aufsetzt, wird der Rechte zuerst angehoben und positiv beschleunigt. Sobald er neben dem linken Fuß vorbeischwingt, hat die Beschleunigung des rechten Fußes aufgehört. Von dort an wird dieser abgebremst, bis er stillsteht und abgesetzt wird. Dann wiederholt der linke Fuß diesen Ablauf, während der Rechte stillsteht (vergleiche dazu Abbildung 3.13). Unter der Annahme, dass die Muskulatur des Menschen die wirkende Kraft auf die Beine linear verändert, lässt sich die Beschleunigung $a_R(t)$ des rechten Fußes als eine Funktion mit der Gleichung 3.26 beschreiben.

$$a_R(t) = \begin{cases} A(1 - 4\frac{t}{T}), & 0 \leq t < \frac{T}{2} \\ 0, & \frac{T}{2} \leq t < T \end{cases} \tag{3.26}$$

Dabei ist A gleich der maximalen Beschleunigung des Fußes und T die Periodendauer von zwei Schritten. Durch zweifaches Integrieren lässt sich daraus der zurückgelegte Weg $S_R(t)$ des rechten Fußes nach Gleichung 3.27 bestimmen.

$$S_R(t) = \begin{cases} \frac{1}{6}At^2\left(3 - 4\frac{t}{T}\right), & 0 \leq t < \frac{T}{2} \\ \frac{1}{24}AT^2, & \frac{T}{2} \leq t < T \end{cases} \tag{3.27}$$

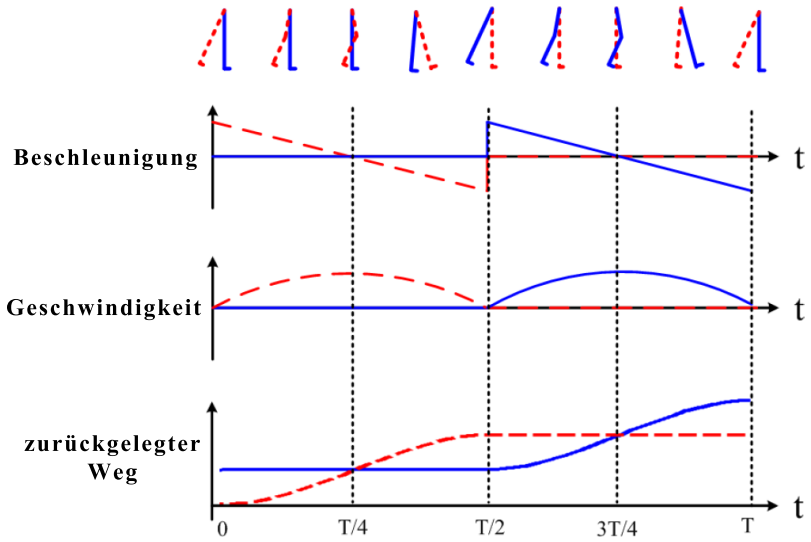


Abbildung 3.13: Position, Geschwindigkeit und Beschleunigung der beiden Füße in einer Periode T. Der rechte Fuß ist rot gestrichelt und der linke Fuß blau eingetragen. Bildquelle: [Shao et al. 2007]

Auf gleiche Art lässt sich der zurückgelegte Weg $S_L(t)$ des linken Fußes ermitteln, welcher in Gleichung 3.28 beschrieben wird.

$$S_L(t) = \begin{cases} \frac{1}{48}AT^2, & 0 \leq t < \frac{T}{2} \\ \frac{1}{6}A\left(t - \frac{T}{2}\right)^2\left(5 - 4\frac{t}{T}\right) + \frac{1}{48}AT^2, & \frac{T}{2} \leq t < T \end{cases} \tag{3.28}$$

Die Distanz $S_d(t)$ zwischen den beiden Füßen lässt sich demnach mit Gleichung 3.29 beschreiben.

$$S_d(t) = S_R(t) - S_L(t) = \begin{cases} -\frac{1}{48}AT^2\left[32\left(\frac{t}{T}\right)^3 - 24\left(\frac{t}{T}\right)^2 + 1\right], & 0 \leq t < \frac{T}{2} \\ -\frac{1}{48}AT^2\left[-32\left(\frac{t}{T}\right)^3 + 72\left(\frac{t}{T}\right)^2 - 48\left(\frac{t}{T}\right) + 9\right], & \frac{T}{2} \leq t < T \end{cases} \tag{3.29}$$

Im Plot 3.14 scheint diese Funktion einem Kosinus zu ähneln und wie in [Shao et al. 2007] festgestellt wurde, ist die Gleichung 3.29 in der Tat die exakte kubische Spline-Interpolation einer Kosinusfunktion. Somit kann die Distanz der beiden Füße auch durch die Gleichung 3.30 beschrieben werden.

$$\widehat{S}_d(t) = -\frac{1}{48}AT^2\cos\left(2\pi\frac{t}{T}\right) = -S\cos(\varphi) \tag{3.30}$$

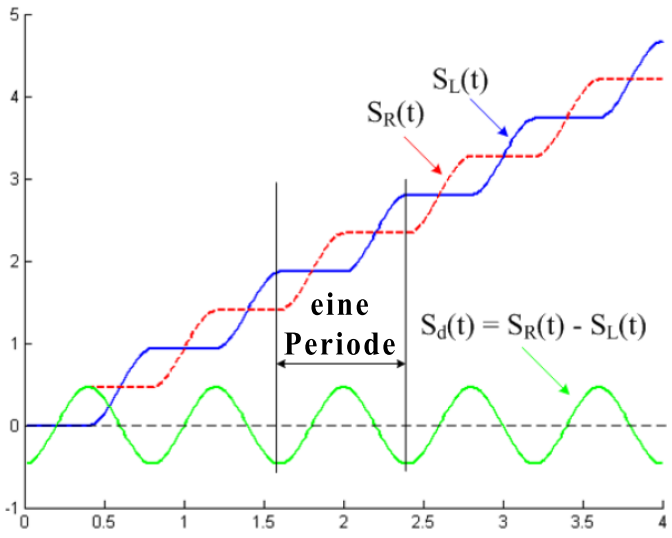


Abbildung 3.14: Position des rechten Fußes $S_R(t)$ (rot gestrichelt) und des linken Fußes $S_L(t)$ (blau), sowie die Distanz beider Füße zueinander (grün). In der Distanz $S_d(t)$ ist eine Funktion zu erkennen, die dem Kosinus ähnelt. Bildquelle: [Shao et al. 2007]

Dabei ist S die Schrittweite der Person und φ die Phase der Bewegung.
Der Bewegungszustand $\mathbf{x}$ der Füße einer Person lässt sich mit einem Vektor, wie er in Gleichung 3.31 dargestellt ist, beschreiben.

$$\mathbf{x}_t = (\mathbf{l}_t, \mathbf{r}_t, S_t, \varphi_t, \vartheta_t, T_t)^T \tag{3.31}$$

Dabei ist $\mathbf{l}_t$ die Position des linken Fußes, $\mathbf{r}_t$ die Position des rechten Fußes, S die Schrittweite, φ die Phase der Bewegung, ϑ die Richtung, in der sich die Person bewegt und T_t die Schrittdauer. Mit dem Wissen über die kosinusartige Bewegung der Füße lässt sich somit ein Bewegungsmodell für die Gleichung 3.21 bilden, welches den Zustandsvektor $\mathbf{x}_t$ für jeden neuen Zeitschritt bestimmt:
Wenn δ die Zeitdauer zwischen dem alten Zustand $\mathbf{x}_{t-1}$ und dem neuen zu berechnenden Zustand $\mathbf{x}_t$ ist, dann kann zuerst die Phase φ_t des neuen Zeitschrittes mit Gleichung 3.32 bestimmt werden, wobei ω_φ ein Rauschterm ist. Ein gleichverteiltes Rauschen von $\frac{2{,}5rad \cdot \delta}{1s}$ hat sich in den Experimenten als geeignete Größe herausgestellt. Bei $\delta = 20$ms wird also die Phase um $\pm 0{,}05$rad verrauscht.

$$\varphi_t = \varphi_{t-1} + 2\pi \cdot \frac{\delta}{T_{t-1}} + \omega_\varphi \tag{3.32}$$

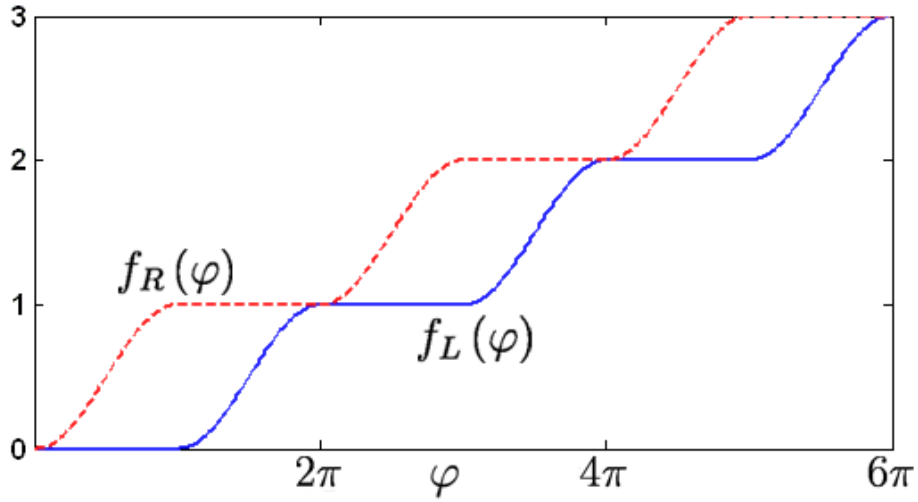


Abbildung 3.15: Hilfsfunktion, welche für den rechten (rot gestrichelt) und den linken Fuß (blau) in Abhängigkeit von der Phase beschreibt, wie viel Schritte sie sich vorwärts bewegt haben.

Die zwei Hilfsfunktionen aus Gleichung 3.33 und 3.34 beschreiben, wieviel Schritte, bzw. Bruchteile von Schritten, sich die beiden Füße zu einer gegebenen Phase vorwärts bewegt haben (siehe Abbildung 3.15). Mit diesen Funktionen ist es möglich, auch über die Grenze von einzelnen Perioden hinweg zu bestimmen, wie weit sich der linke und rechte Fuß im Vergleich zum vorherigen Zeitschritt bewegt haben. So lassen sich dann mit Gleichung 3.35 und 3.36 unter Berücksichtigung der Schrittweite S_{t-1} und der Bewegungsrichtung ϑ_{t-1} die neuen Positionen der Füße berechnen.

$$f_L(\varphi) = \begin{cases} \left\lfloor \frac{\varphi}{2\pi} \right\rfloor, & 0 \leq \varphi \bmod (2\pi) \leq \pi \\ \left\lfloor \frac{\varphi}{2\pi} \right\rfloor + \frac{1-\cos(\varphi+\pi)}{2}, & \pi < \varphi \bmod (2\pi) < 2\pi \end{cases} \tag{3.33}$$

$$f_R(\varphi) = \begin{cases} \left\lfloor \frac{\varphi}{2\pi} \right\rfloor + \frac{1+\cos(\varphi+\pi)}{2}, & 0 \leq \varphi \bmod (2\pi) \leq \pi \\ \left\lfloor \frac{\varphi}{2\pi} \right\rfloor + 1, & \pi < \varphi \bmod (2\pi) < 2\pi \end{cases} \tag{3.34}$$

$$\mathbf{l}_t = S_{t-1} \cdot (f_L(\varphi_t) - f_L(\varphi_{t-1})) \cdot \begin{pmatrix} \cos\vartheta_{t-1} \\ \sin\vartheta_{t-1} \end{pmatrix} + \mathbf{l}_{t-1} \tag{3.35}$$

$$\mathbf{r}_t = S_{t-1} \cdot (f_L(\varphi_t) - f_R(\varphi_{t-1})) \cdot \begin{pmatrix} \cos\vartheta_{t-1} \\ \sin\vartheta_{t-1} \end{pmatrix} + \mathbf{r}_{t-1} \tag{3.36}$$

3.7.2 Oberkörpermodell

Ebenso wie im Fall von Fußscans im Abschnitt 3.7.1, ist es bei Oberkörperscans notwendig, ein Bewegungsmodell zur Verfügung zu haben, womit ein Zustand vom vorherigen Zeitschritt in den aktuellen Zeitschritt überführt werden kann (siehe Gleichung 3.21). Allerdings ist die Bewegung der Oberkörper im Vergleich zu Füßen wesentlich einfacher beschreibbar.
Unter der Annahme, dass sich der Oberkörper einer sich bewegenden Person annähernd linear in eine Richtung bewegt, kann der Bewegungszustand $\mathbf{x}_t$ durch den Vektor, wie er in Gleichung 3.37 dargestellt ist, beschrieben werden.

$$\mathbf{x}_t = (X_t, Y_t, \vartheta_t, v_t)^T \tag{3.37}$$

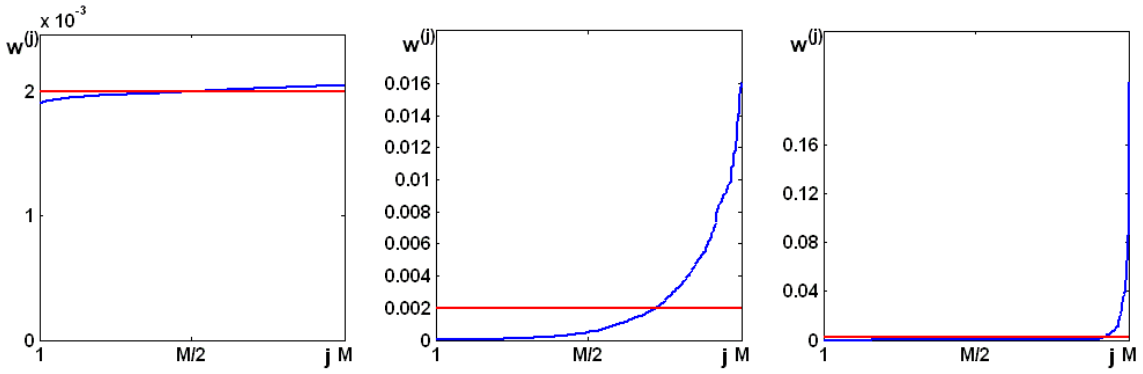
Die Position des Oberkörpers wird durch $\mathbf{m}_t = (X_t, Y_t)^T$ repräsentiert. ϑ_t gibt die Bewegungsrichtung und v_t die Geschwindigkeit an, mit der sich die Person geradeaus bewegt. Wenn δ die Zeit zwischen dem vorangegangenen und dem aktuellen Zeitschritt ist, kann das Bewegungsmodell mit Gleichung 3.38 beschrieben werden.

$$\mathbf{x}_t = \begin{pmatrix} X_t \\ Y_t \\ \vartheta_t \\ v_t \end{pmatrix} = \mathbf{x}_{t-1} + \begin{pmatrix} \cos(\vartheta_{t-1}) \\ \sin(\vartheta_{t-1}) \\ 0 \\ 0 \end{pmatrix} \cdot \delta \cdot v_{t-1} \tag{3.38}$$

Dieses Modell beschreibt eine lineare Bewegung in eine Richtung, wohingegen Personen durchaus ihre Bewegungsrichtung und Geschwindigkeit ändern können. Dem wird allerdings im Resamplingschritt (siehe Abschnitt 3.7) Rechnung getragen, indem dort alle vier Parameter anschließend verrauscht werden und somit deren Variabilität gewahrt bleibt. Dies wird im Abschnitt 4.5.2 genauer erläutert.

3.7.3 Löschen von Partikelfiltern

Da Personen die Szene betreten und wieder verlassen können, müssen Partikelfilter nicht nur bei neuen Detektionen erzeugt werden, sondern auch gelöscht werden, sobald die vom Filter getrackte Person die Szene verlassen hat. Dies geschieht auf zwei verschiedene Arten.



(a) 0m (b) 2m (c) 5m

Abbildung 3.16: Verteilung der Partikelgewichte eines Partikelfilters für unterschiedliche Entfernungen zur Person. In diesen Plots sind die Partikelgewichte der M Partikel, der Größe nach geordnet, blau dargestellt. Der Mittelwert 1/M ist rot eingetragen. In (a) befinden sich die Partikel direkt bei einer Person, weswegen die Partikelgewichte annähernd gleichverteilt sind und etwa die Hälfte der Partikel ein höheres Gewicht als 1/M besitzen. In (b) sind die Partikel im Mittel etwa 2 Meter von der nächsten Person entfernt, wodurch einige Partikel, die näher an der Person liegen, ein höheres Gewicht besitzen, als alle anderen. Nun sind nur noch etwa ein Drittel der Partikelgewichte größer als 1/M. In (c) sind die Partikel im Mittel etwa 5 Meter von der nächsten Person entfernt. Es ist deutlich zu sehen, dass nur noch sehr wenige Partikel ein Gewicht über 1/M besitzen. Dieser Partikelfilter hat seine Person offensichtlich verloren.

Einerseits müssen Durchgangsgebiete im globalen Koordinatensystem definiert werden, in denen Personen die Szene betreten und verlassen. Sobald sich die Personenhypothese eines Partikelfilter ununterbrochen mehr als eine bestimmte Anzahl von Zeitschritten in einem dieser Gebiete befindet, so kann davon ausgegangen werden, dass die Person die Szenerie verlässt und der Partikelfilter kann daraufhin gelöscht werden. Indem Partikelfilter nur auf Personendetektionen erzeugt werden, welche sich nicht mehr in einem Durchgangsgebiet befinden, kann verhindert werden, dass neu erzeugte Partikelfilter nach kurzer Zeit unerwünschterweise wieder gelöscht werden, falls z.B. die zu trackende Person im Durchgangsgebiet verweilt.

Ein zweiter Mechanismus zum Löschen von Partikelfiltern verhindert, dass Filter, wel-

che die zu trackende Person verloren haben, andere Filter nachteilig beeinflussen. Dazu ist es notwendig zu erkennen, wann ein Partikelfilter seine Zielperson verloren hat. Eine Möglichkeit bietet eine Metrik, welche anhand der Partikelgewichte $w^{(j)}$, $j = 1, \ldots, M$ des Filters vor dem Resamplingschritt dessen Nützlichkeitswert berechnet. Wird eine Person korrekt vom Partikelfilter getrackt, dann befinden sich die Partikel nahe an den Scanpunkten und deren Gewichte sind annähernd gleichverteilt. Sind die Partikel hingegen weiter von den Scanpunkten entfernt, wird die Verteilung eher unimodal, da es viele Partikel mit geringem Gewicht gibt und nur wenige eine etwas bessere Personenhypothese liefern als die anderen und damit ein geringfügig höheres Gewicht besitzen (vergleiche Abbildung 3.16). Dies bewirkt, dass die Anzahl der Partikel, deren Gewicht über dem Mittelwert $\frac{1}{M}$ liegt, mit steigendem Abstand zur Person abnimmt. Dieser Zusammenhang kann genutzt werden, um den Nützlichkeitswert ν für den Partikelfilter nach Gleichung 3.39 und 3.40 zu berechnen.

$$n\left(w^{(j)}\right) = \begin{cases} 1, & \text{wenn } w^{(j)} > \frac{1}{M} \\ 0, & \text{wenn } w^{(j)} \leq \frac{1}{M} \end{cases} \tag{3.39}$$

$$\nu = \frac{\sum_{j=1}^{M} n\left(w^{(j)}\right)}{M} \tag{3.40}$$

Ein Partikelfilter, der die Person mit hoher Güte trackt, hat somit einen Wert von $\nu \approx 0.5$, wohingegen ein Partikelfilter, der die zu trackende Person verloren hat, einen Wert von $\nu \approx 0$ aufweist. Da es prinzipbedingt bei dem Erfassen von Personen mit Laser-Range-Scannern zu zeitweisen Verdeckungen kommen kann, sollte ein Partikelfilter nicht gelöscht werden, sobald sein Nützlichkeitswert eine gewisse Schwelle unterschreitet. Stattdessen sollten dem Filter einige Zeitschritte gelassen werden, um die Person wiederzufinden, sobald sie aus der Verdeckung wieder heraustritt. Dies kann durch einfaches Zählen der Zeitschritte, in dem der Nützlichkeitswert ununterbrochen eine Schwelle unterschreitet, realisiert werden. Sobald eine Anzahl von Zeitschritten überschritten wird, kann der Partikelfilter in der Vermutung, dass er seine Person verloren hat, gelöscht werden.

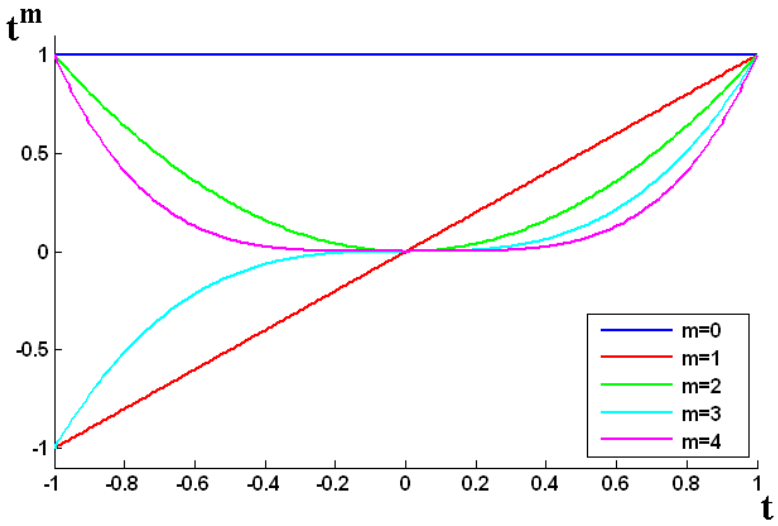


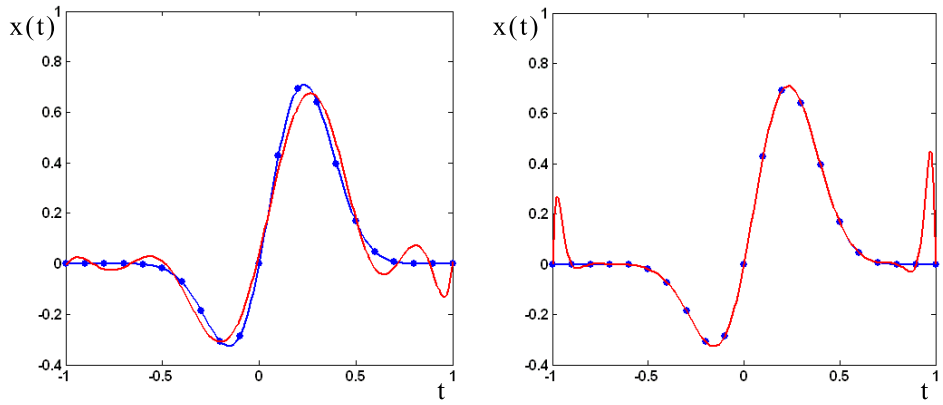
Abbildung 3.17: Die ersten fünf Basisfunktionen der Polynomapproximation.

3.8 Trajektorienrepräsentation

Es gibt mehrere Möglichkeiten Trajektorien zu beschreiben. Der einfachste Ansatz, welcher auch der zeitlichen Diskretisierung des Partikelfilters aus Abschnitt 3.7 Rechnung trägt, ist, die Bewegungszustände der Personen zu diskreten Zeitpunkten abzuspeichern. Wenn der Bewegungszustand als $\mathbf{x}$ bezeichnet wird, dann ist die Trajektorie B eine Folge von Tupeln $B = \{\langle \mathbf{x}_i, t_i \rangle \,|i = 1, \ldots, I\}$. Diese Art der Repräsentation ist in der Praxis die Genaueste, da die Zustände nur zu diskreten Zeitschritten gemessen werden können. Allerdings kann bei hohen Abtastraten mit der Zeit die Datenmenge enorm werden. Je nach dem, was mit den Trajektorien berechnet werden soll, ist diese Repräsentation auch nicht unbedingt nutzbringend. Eine Parametrisierung der Trajektorie, also eine Beschreibung dieser mit einer endlichen Menge von Parametern, kann den Speicheraufwand der Trajektorie minimieren und Vorteile für die weitere Verarbeitung der Trajektorie bringen.

3.8.1 Polynomapproximation

Eine parametrisierte Repräsentationsform auf Basis von Polynomapproximation wurde in [Melo et al. 2006] und in [Naftel und Khalid 2006] vorgestellt. Dabei wird der



(a) m = 10 (b) m = 20

Abbildung 3.18: Polynomapproximation einer Funktion an 21 äquidistanten Stützstellen auf der t-Achse. Die zu approximierende Funktion ist blau eingezeichnet und die für die Berechnung verwendeten Stützstellen als blaue Kreise. Die Approximation dieser Funktion ist rot eingetragen. In (a) wurde ein Polynom des Grades $m = 10$ verwendet und in (b) ein Polynom vom Grad $m = 20$. In (b) ist Runges Phänomen erkennbar, welches sich durch eine starke Oszillation am Rand der Funktion bei hohen Polynomgraden bemerkbar macht.

Zustand $\mathbf{x}$ der Person als Funktionen der Zeit aufgefasst und versucht, diese Funktionen durch ein Polynom $P_{\mathbf{x}}(t)$ des Grades m anzunähern (siehe Gleichung 3.41).

$$\mathbf{x}(t) \approx P_{\mathbf{x}}(t) = \sum_{i=0}^{m} \mathbf{a}_i \cdot t^i \tag{3.41}$$

Die Koeffizientenvektoren $\mathbf{a}_i$, $i = 0, \ldots, m$ können durch Minimierung einer Fehlerfunktion (siehe Gleichung 3.42) ermittelt werden.

$$E(\mathbf{a}_0, \ldots, \mathbf{a}_m) = \sum_{i=1}^{n} \mathbf{x}_i - P_{\mathbf{x}}(t_i) \tag{3.42}$$

Dabei sind $\mathbf{x}_i$ die insgesamt n gemessenen Zustände zu den Zeitpunkten t_i. Die parametrisierte Repräsentation der Trajektorie ist dann durch $A = \{\mathbf{a}_0, \ldots, \mathbf{a}_m\}$ gegeben. Da die Polynomapproximation eine lineare Operation darstellt, kann auch zwischen zwei Trajektorien interpoliert werden, indem deren Parameter interpoliert werden.

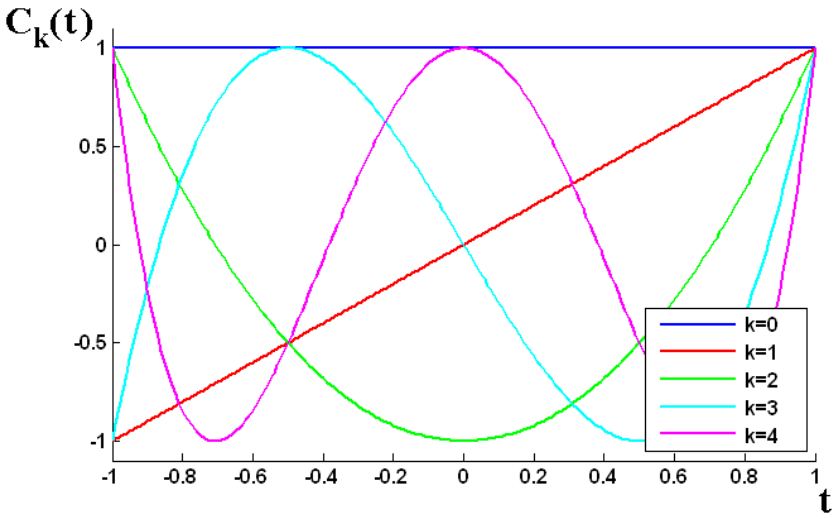


Abbildung 3.19: Die ersten fünf Basisfunktionen der Chebychev-Approximation.

3.8.2 Chebychev-Approximation

In [Naftel und Khalid 2006] wurde ebenfalls die Repräsentation von Trajektorien mit Hilfe von Chebychev-Polynomen $C_k(t)$ beschrieben, welche auch in [Mason und Handscomb 2003] erklärt werden. Diese Polynome bilden sich nach Gleichung 3.43, wobei t auf $[-1, 1]$ normiert werden muss und $k = 0, \ldots, K$ den Grad des Polynoms angibt.

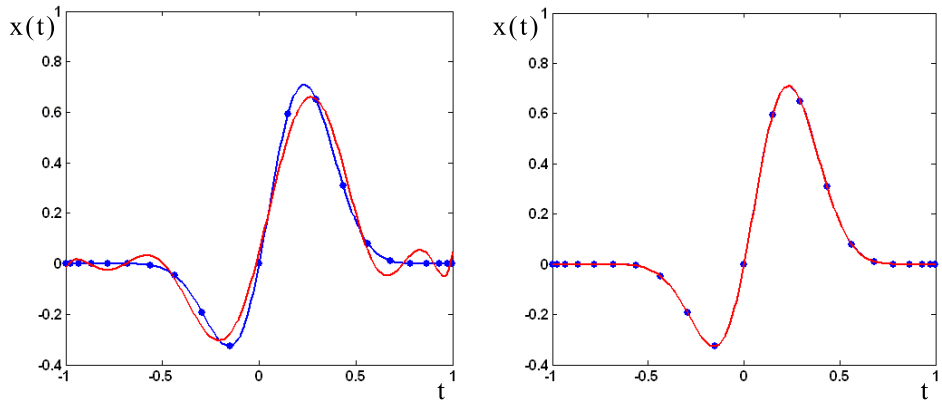
$$C_k(t) = \cos(k \cdot \arccos(t)) \tag{3.43}$$

Die auf $t = [-1, 1]$ normierte Funktion $\mathbf{x}(t)$ des Bewegungszustandes wird mit diesen Polynomen durch Gleichung 3.44 angenähert, wobei $\mathbf{b}_k$ die Chebychev-Koeffizienten sind, die diese Funktion parametrisieren.

$$\mathbf{x}(t) \approx \sum_{k=0}^{K} \mathbf{b}_k C_k(t) - \frac{\mathbf{b}_0}{2} \tag{3.44}$$

Die Koeffizienten lassen sich durch Gleichung 3.45 anhand der zu approximierenden, auf $t \in [-1, 1]$ normierten Funktion $\mathbf{x}(t)$ an den Stützstellen t_q, $q = 1, \ldots, Q$ errechnen.

$$\mathbf{b}_k = \frac{2}{Q} \sum_{q=1}^{Q} \mathbf{x}(t_q) C_k(t_q) \tag{3.45}$$



(a) K = 10 (b) K = 20

Abbildung 3.20: Chebychev-Approximation einer Funktion an 21 Stützstellen. Die zu approximierende Funktion ist blau eingezeichnet und die für die Berechnung verwendeten Chebychev-Knoten als blaue Kreise. Die Approximation dieser Funktion ist rot eingetragen. In (a) wurden Chebychev-Polynome bis zum Grad $K = 10$ verwendet und in (b) bis zum Grad $K = 20$. Es ist erkennbar, dass im Gegensatz zur Polynomapproximation keine Überschwinger am Funktionsrand auftreten.

Diese Stützstellen könnte man äquidistant auf $[-1, 1]$ verteilt wählen, allerdings erhält man den geringsten Approximationsfehler, wenn man die so genannten Chebychev-Knoten wählt, welche sich nach Gleichung 3.46 berechnen lassen.

$$t_q = \cos\left(\frac{2q-1}{2Q} \cdot \pi\right) \tag{3.46}$$

Auf diese Art lässt sich eine Trajektorie durch insgesamt $Q + 2$ Koeffizienten repräsentieren, wobei Q Koeffizienten die Chebychev-Koeffizienten darstellen und noch zwei Koeffizienten notwendig sind, um den zeitlichen Anfang und das Ende der Trajektorie zu speichern, da diese ja im Prozess der Approximation auf $t \in [-1, 1]$ normalisiert wurde. Da die Chebychev-Approximation eine lineare Operation darstellt, kann auch zwischen zwei Trajektorien interpoliert werden, indem deren Koeffizienten interpoliert werden. Es ist ebenfalls möglich ein Distanzmaß für Trajektorien auf Basis der Chebychev-Koeffizienten zu realisieren, indem die euklidische Distanz zwischen den Koeffizienten selbst verwendet wird. Zudem wird Runges Phänomen minimiert, welches

sich durch eine starke Oszillation an den Intervallgrenzen der Funktionen manifestiert.

3.8.3 Glättung der Trajektorien

Da die Personenhypothesen der Partikelfilter durchaus ein Rauschen enthalten können, ist es sinnvoll, die ermittelten Trajektorien zu glätten. Eine einfache Möglichkeit bietet dabei ein Tiefpassfilter, welcher hochfrequente Anteile aus der diskreten Funktion $\mathbf{x}(j), j = 1, \ldots, J$ der J Trajektorienpunkte herausfiltert. Dies geschieht, indem die Funktion mit einer Filterfunktion $F(i), i = -D, \ldots, D$ gefaltet wird. Eine Faltung

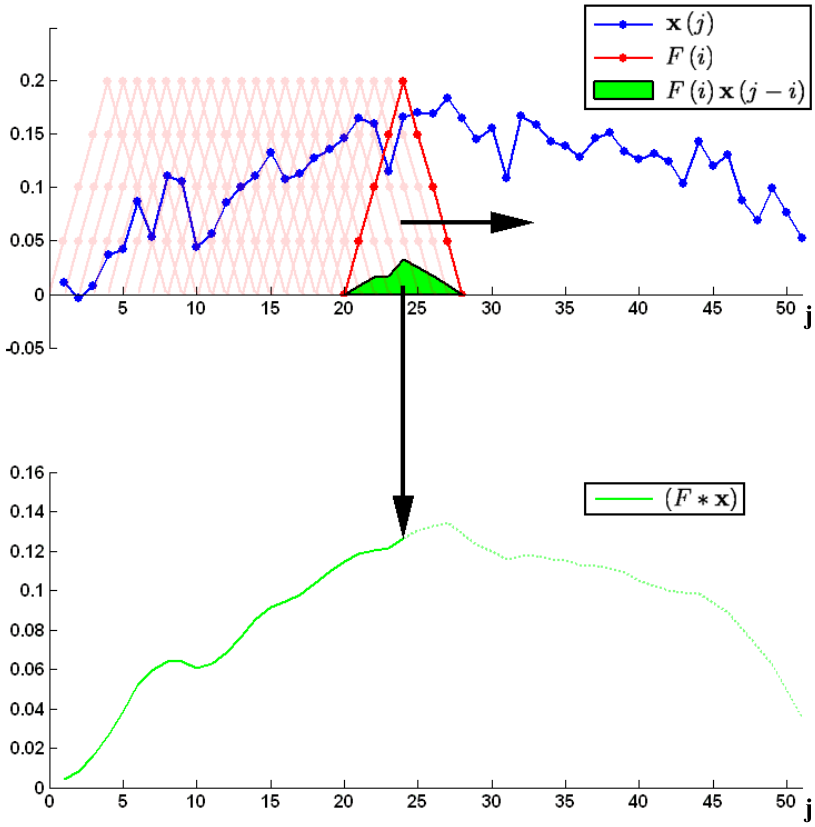


Abbildung 3.21: Veranschaulichung der Faltung. Die rot eingefärbte Filterfunktion wird über die blaue Trajektorie entlang von j geschoben und an jedem Stützpunkt mit ihr multipliziert (schwarze Kurve). Die Fläche unter dem Produkt, welche hier Grün markiert ist, ist der Wert für die geglättete, darunter ebenfalls grün eingezeichnete, Trajektorie an dieser Stelle.

$(F * \mathbf{x})(j)$ ist eine mathematische Operation, die durch 3.47 definiert ist.

$$(F * \mathbf{x})(j) := \sum_{i=-D}^{D} F(i)\,\mathbf{x}(j-i) \tag{3.47}$$

Da $\mathbf{x}(j)$ für $j < 1$ und $j > J$ nicht definiert ist und diese Grenzen während der Faltung überschritten werden können, wird dort für gewöhnlich der Funktionswert 0 verwendet. Die Faltung kann so interpretiert werden, dass jedes Element aus den diskreten Funktionen neu berechnet wird, indem die lokale Nachbarschaft gewichtet gemittelt wird, wobei nahe beim Element liegende Nachbarn sich stark auf den neuen Wert auswirken und weiter weg liegende Nachbarn nur gering. In der Abbildung 3.21 wird die Faltung im Eindimensionalen verdeutlicht. Eine einfache Filterfunktion, welche auch in den Experimenten verwendet wurde, ist eine Dreiecksfunktion (siehe Gleichung 3.48). Auch andere Funktionen, wie die Gaußfunktion, sind denkbar.

$$F(i) = \frac{D - \|i\|}{D^2}, \qquad i = -D, \ldots, D \tag{3.48}$$

3.8.4 Zusammenfassung

Nachdem in diesem Kapitel beschrieben wurde, wie man mit mehreren Laser-Range-Scannern die Umgebung in Form einer Punktemenge im globalen Koordinatensystem erfassen kann, folgte die Hintergrundsegmentierung, um diese Punktemenge von statischen Objekten, wie Wände und Säulen, zu befreien. Daraufhin wurde der Mean-Shift-Clusterer vorgestellt, mit dem die von den statischen Objekten befreite Punktemenge in lokale Gruppen zusammenfasst werden kann. Dem folgte die Beschreibung der Hauptkomponentenanalyse, mit der die Ausdehnungen der Punktegruppen ermittelt werden kann. Anschließend wurde der Partikelfilter mit den beiden Bewegungsmodellen für die Füße und den Oberkörper vorgestellt, womit die Personen getrackt und dessen Bewegungsspuren ermittelt werden können. Zum Schluss folgte eine Beschreibung unterschiedlicher Methoden, mit denen die Bewegungstrajektorien repräsentiert werden können.

All diese Grundlagen sind notwendig, um mit Partikelfiltern einen Personentracker auf Basis von Laser-Range-Scannern realisieren zu können. Wie diese Verfahren miteinander kombiniert werden können, um den Tracker umzusetzen, wird im folgenden Kapitel behandelt.

Kapitel 4

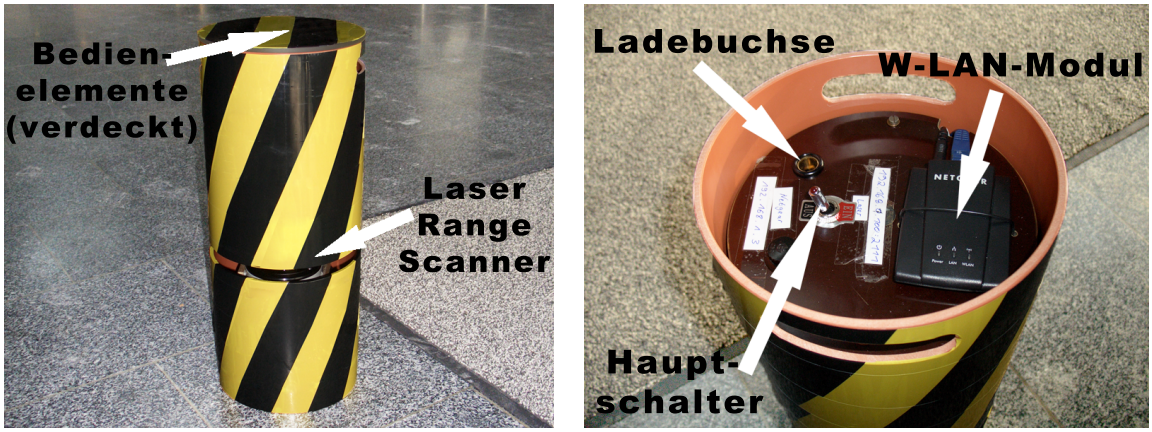
Realisierung eines Personentrackers

4.1 Einleitung

Dieses Kapitel befasst sich mit der praktischen Umsetzung der im vorherigen Kapitel vorgestellten Mechanismen in einen funktionstüchtigen Personentracker. Zuerst wird im Abschnitt 4.2 der Aufbau der Scannerhardware erläutert. Die Daten welche von diesen Laser-Range-Scannern ausgegeben werden und in jedem neuen Zeitschritt vorliegen, werden im Abschnitt 4.3 mit einem Vorverarbeitungsschritt in ein brauchbares Format gebracht. In diesen Daten wird dann nach neuen Personen gesucht (Abschnitt 4.4). Nachdem eventuell neu erkannte Personen dem Tracking-Modul mitgeteilt wurden, kann der Tracker die Hypothesen der ihm bekannten Personen anhand der Daten des neuen Zeitschritts aktualisieren (Abschnitt 4.5). Sobald der Tracker seine Hypothesen erneuert hat, können mit diesen Hypothesen in der Nachverarbeitung Personenspuren gebildet werden (Abschnitt 4.6). Sobald dies abgeschlossen ist, wartet der Personentracker auf neue Scandaten der Laser, damit die gerade beschriebene Iteration wieder der Reihe nach abgearbeitet werden kann. Nachdem in diesem Kapitel auf die konkreten Umsetzungen der einzelnen Elemente eingegangen wurde, wird der Gesamtablauf des Personentrackings noch einmal in der Zusammenfassung des Kapitels dargestellt (Abschnitt 4.7).

4.2 Messequipment

Für die Experimente wurden die Laser-Range-Scanner per Bleiakku mit Strom versorgt und die Ethernetverbindung zum Zielrechner wurde mit Hilfe einer W-LAN-Brücke hergestellt. Der Scanner, der Akku, das W-LAN-Modul und sämtliche Verbindungskabel wurden in ein Gehäuse verbaut (Abbildung 4.1). Damit entfällt das Verlegen von Kabeln, wodurch ein sicherer Einsatz in stark frequentierten Umgebungen, ohne Gefährdung der Passanten durch Stolpergefahren, möglich wird. Der im Scanner verwendete Laser arbeitet mit einer Wellenlänge von 905nm und entspricht der Laserklasse 1. Somit ist er selbst bei kontinuierlicher Einwirkung ungefährlich für die Augen und die menschliche Haut und kann ohne Bedenken im öffentlichen Bereich eingesetzt werden. Dadurch, dass keine Verkabelung nötig ist, um die Laserscanner einzusetzen, können sie auch innerhalb kürzester Zeit aufgebaut und in Betrieb genommen werden. Das Design der Gehäuse ermöglicht zudem auch einen einfachen Transport zur Einsatzumgebung. Bei den Experimenten in Ilmenau und Erfurt aus Kapitel 6 waren somit weniger als zehn Minuten nötig, um alle Laser-Range-Scanner aufzustellen und mit der Datenaufnahme zu beginnen. Der Akku ermöglicht Betriebszeiten von mehreren Stunden. Die maximale Einsatzzeit konnte noch nicht bestimmt werden, allerdings



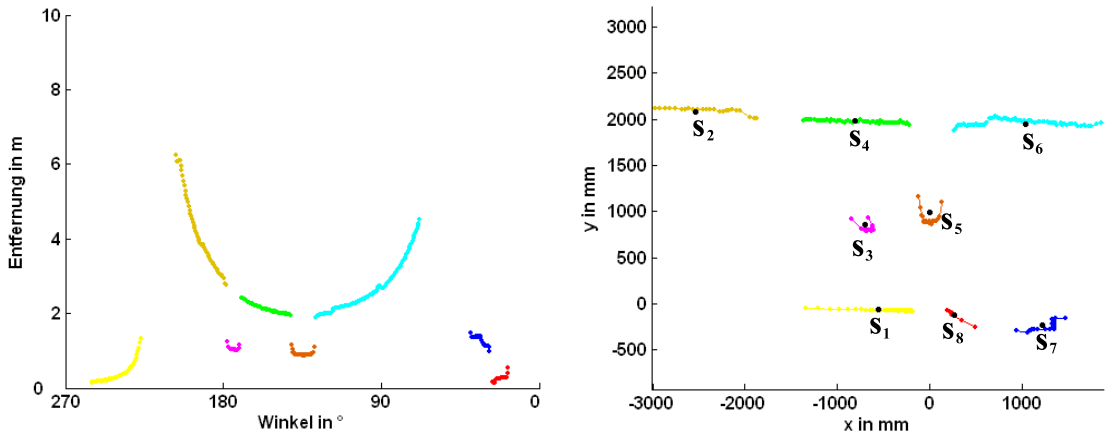
(a) Laser-Range-Scanner im Gehäuse (b) Bedienelemente

Abbildung 4.1: Das Gehäuse mit sämtlichen Messequipment. Der Akku ist nicht sichtbar und befindet sich über dem eigentlichen Laser-Range-Scanner.

waren die Laserscanner in Erfurt mehr als sieben Stunden kontinuierlich in Betrieb, ohne dass sie zwischendurch aufgeladen wurden.

4.3 Vorverarbeitung

Bevor der eigentliche Tracker zum Einsatz kommen kann, müssen einige wichtige Vorverarbeitungsschritte durchgeführt werden, damit die Rohdaten der Laserscanner in geeigneter Form vorliegen. Diese Schritte sind im Einzelnen sowohl das räumliche, als auch das zeitliche Synchronisieren der Rohdaten aller eingesetzter Laser-Range-Scanner, wie es im Abschnitt 3.3 beschrieben wurde, um ein komplettes Abbild der Szene in Form einer Punktemenge zu erhalten. Diese Menge muss noch von statischen Objekten durch eine Hintergrundsegmentierung bereinigt werden, welche in Abschnitt 3.4 bereits näher erläutert wurde. Dadurch erhält man die Scandaten, welche zu den dynamischen Objekten, also hauptsächlich zu den Personen, gehören. Damit Personen in diesen Daten auch erkannt werden können, ist es notwendig, dass die Punktemenge in einzelne, jeweils zu einer Person gehörenden, Gruppen zusammengefasst wird. Diese Zusammenfassung, oder auch Clusterung genannt, wird in zwei Schritten durchgeführt. Zuerst werden in den Rohdaten der Laserscanner mit einer einfachen Abstandsmetrik zusammenhängende Segmente gesucht (Abschnitt 4.3.1). Nach der räumlichen Synchronisierung werden die Punkte dieser Segmente mit dem Mean-Shift-Verfahren zu Punktegruppen gebündelt (Abschnitt 4.3.2). Der Zwischenschritt der Segmentierung ist rein praktischen Überlegungen entsprungen. Es kann auch die Punktemenge der Laserscanner direkt mit dem Mean-Shift-Verfahren geclustert werden. Da aber der Rechenaufwand des Algorithmus quadratisch zur Anzahl der Punkte wächst, wird dafür erheblich Rechenzeit benötigt. Der Rechenaufwand für die Segmentierung aus Abschnitt 4.3.1 wächst hingegen nur linear mit der Anzahl der betrachteten Punkte und ist je nach aktueller Szenerie in der Lage große Mengen von Punkten in wenige Segmente zusammenzufassen. Somit müssen mit dem Mean-Shift-Verfahren, anstatt großer Punktemengen, nur noch wenige Segmente geclustert werden, wodurch die benötigte Zeit für das Zusammenfassen der Punktemenge in Cluster erheblich verringert werden kann.



(a) Entfernungsscan der Szene

(b) Draufsicht auf die Szene

Abbildung 4.2: Segmentierung eines Laserscans. Die Segmentierung findet direkt im Entfernungsscan von (a) statt. Die einzelnen gefunden Segmente sind mit unterschiedlicher Farbe in den Bildern eingetragen. Es ist in (b) zu sehen, dass diese einfache Art, zusammenhängende Punkte zu finden, gute Ergebnisse liefert. Die schwarzen Punkte $\mathbf{s}_1$ bis $\mathbf{s}_8$ geben die Mittelpunkte der Segmente an.

4.3.1 Segmentierung von Laserdaten

Die Segmentierung in den Entfernungsdaten der einzelnen Laserscanner lässt sich mit relativ geringem Aufwand realisieren. Ziel der Segmentierung ist es, schon in den Rohdaten der Laser-Range-Scanner zusammenhängende Messungen zu finden, welche zu einem Objekt gehören könnten, und diese durch einen gemeinsamen Mittelpunkt $\mathbf{s}_i$ im globalen Koordinatensystem zu repräsentieren (siehe Abbildung 4.2). Da die Entfernungsdaten schon prinzipbedingt dem Ablenkwinkel nach geordnet von 0° bis 270° in einem Vektor $\mathbf{d} = (d_0, \ldots, d_{540})$ vorliegen, lassen sich die Daten durch Betrachtung der benachbarten Entfernungswerte segmentieren. Wenn die Differenz zweier benachbarter Entfernungswerte d_i und d_{i+1} aus dem Scanvektor $\mathbf{d}$ geringer als ein bestimmter Wert ist, so gehören diese Werte zum gleichen Segment, ansonsten gehören sie zu unterschiedlichen Segmenten. Es ist nicht notwendig die euklidische Distanz der zu den Messungen gehörenden Punkte zu verwenden, da bei einem Winkel von 0,5° die euklidische Distanz in etwa gleich der Differenz der beiden Messwerte selbst ist. Eine maximale Differenz von 20cm hat sich als brauchbarer Schwellwert für zusammenhängende Seg-

mente herausgestellt. Somit wird auch noch auf einer Entfernung von 20m eine Wand im Winkel von 45° zum Messstrahl als zusammenhängendes Objekt erkannt. Nachdem für jeden einzelnen Laser-Range-Scanner die Segmente ermittelt wurden, können diese dann gemeinsam im globalen Koordinatensystem mit dem Mean-Shift-Clusterer, welcher im folgenden Abschnitt beschrieben wird, zusammengefasst werden.

4.3.2 Anpassung des Mean-Shifts auf Segmente

Da der Mean-Shift-Algorithmus allerdings so, wie er in Abschnitt 3.5 beschrieben wird, nur auf Punkte und nicht auf Segmente, also Mengen von Punkte, anwendbar ist, muss er angepasst werden.
Damit die Segmente, bzw. deren Mittelpunkte $\mathbf{s}_i$, geclustert werden können, muss die Gleichung 3.11 zu Gleichung 4.1 angepasst werden, um der unterschiedlichen Anzahl an Punkten in den einzelnen Segmenten Rechnung zu tragen.

$$\mathbf{s}_i' = \frac{\sum_{j=1}^{N} \mathbf{s}_j \cdot n\left(\mathbf{s}_j\right) \cdot g\left(\left\|\frac{\mathbf{s}_i - \mathbf{s}_j}{\gamma}\right\|\right)}{\sum_{j=1}^{N} g\left(\left\|\frac{\mathbf{s}_i - \mathbf{s}_j}{\gamma}\right\|\right)} \tag{4.1}$$

Dabei ist $n(\mathbf{s}_j)$ gleich der Anzahl der Punkte, welche im Segment $\mathbf{s}_j$ zusammengefasst wurden. Mit dieser Modifikation werden alle Punkte eines Segments quasi gleichzeitig verschoben, indem angenommen wird, dass die Punkte nahe an deren Mittelpunkt liegen und durch diesen auch repräsentiert werden können.
Die gefundenen Cluster können nun dahingehend geprüft werden, ob sie eine Person darstellen könnten, oder nicht. Dazu müssen jedoch noch die Ausdehnungen der Cluster bestimmt werden. Unter der Annahme, dass die Punkte der Cluster elliptisch angeordnet sind, kann deren Ausdehnung mit der Hauptkomponentenanalyse aus Abschnitt 3.6 für die Personendetektion im folgenden Abschnitt bestimmt werden.

4.4 Personendetektion

Die Personendetektion und das Tracking der Personen wird in dieser Arbeit auf zwei verschiedene Arten betrachtet. Einerseits anhand von Laserscans in Höhe der Fußknö-

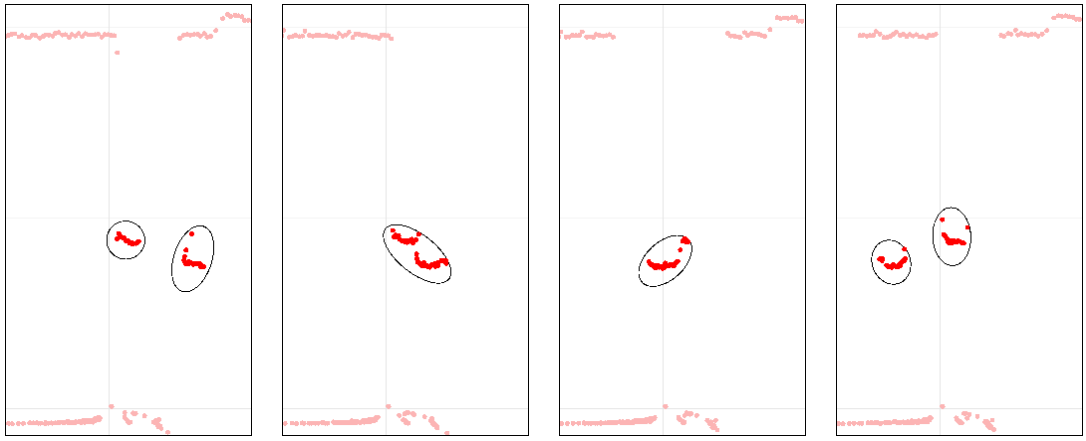
chel und andererseits auf Höhe der Oberkörper. Dadurch ergeben sich zwei verschiedene typische Muster einer Person in der Punkterepräsentation der Daten. Auf Knöchelhöhe erscheint eine Person für gewöhnlich als zwei kreisförmig angeordnete Punktemengen mit einem Durchmesser von ungefähr 15cm (siehe Abbildung 2.1(b)). Auf Höhe des Oberkörpers tritt eine Person meist als elliptisch angeordnete Punktewolke auf, die etwa 40cm mal 60cm misst — abhängig davon, wie die Person proportioniert ist (siehe Abbildung 2.2(b)).

Da die Punktemengen bereits durch die Clusterung ermittelt und deren Ausdehnung mit Hilfe der Hauptkomponentenanalyse bestimmt wurden, können diese Muster mit wenig Aufwand in den Clustern gesucht werden. Wie diese Suche im einzelnen funktioniert, wird in den folgenden zwei Abschnitten erläutert.

4.4.1 Personendetektion auf Fußhöhe

Um Personen, bzw. deren Fußknöchel erfassen zu können, müssen die Laser-Range-Scanner in einer Ebene scannen, die sich etwa 20cm über dem Boden befindet. Bei Laserscans in dieser Höhe lassen sich in einer Momentaufnahme nur schwer einzelne Personen identifizieren, da zu jeder Person zwei Knöchel gehören, welche einander zugeordnet werden müssen. Steht eine Person alleine im freien Raum, so lassen sich sicherlich die beiden Fußknöchel einander recht einfach zuordnen. Wenn allerdings mehrere Personen dicht beieinander stehen, so ist diese Zuordnung nicht mehr so einfach möglich und erfordert die Betrachtung vorangegangener Zeitschritte.

Werden die Punktedaten und Cluster einer gehenden Person auf Knöchelhöhe betrachtet, ergibt sich bei geeigneter Wahl der Clusterbreite γ des Mean-Shift-Clusterers (siehe Abschnitt 3.5 und 4.3.2) ein typischer Ablauf (zu sehen in Abbildung 4.3). Sind die Beine weit auseinander, so bilden die einzelnen Knöchel jeweils ein eigenes Cluster. Schwingt das eine Bein am anderen, feststehenden Bein vorbei, so werden jedoch beide Knöchel zu einem Cluster zusammengefasst. Bewegt sich das eine Bein vom feststehenden Bein weg, so trennt sich das große Cluster wieder in zwei einzelne Cluster auf. Wird dieser Ablauf beobachtet, so kann davon ausgegangen werden, dass die beiden beteiligten Cluster zu den Füßen einer einzelnen Person gehören und somit erfolgreich

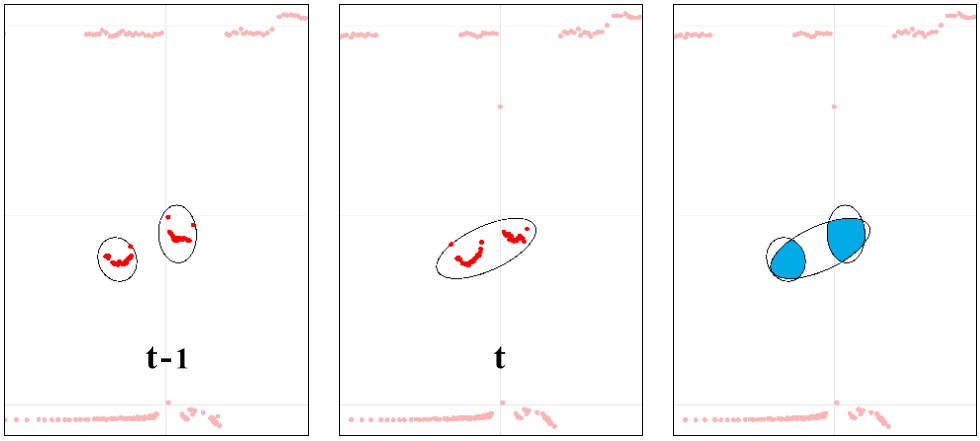


(a) (b) (c) (d)

Abbildung 4.3: Typischer Bewegungsablauf in Laserscans auf Fußhöhe in der Draufsicht. Die Scanpunkte der Füße sind rot markiert. Die Cluster, welche sich aus diesen Punkten bilden, sind durch schwarze Ellipsen angedeutet. In (a) steht der rechte Fuß fest auf und der Linke ist kurz davor, an diesem vorbeizuschwingen (im Bild von Rechts nach Links). Für jeden Fuß existiert hier noch ein einzelnes Cluster. In (b) hat sich der linke Fuß dem Rechten schon soweit genähert, dass die Scanpunkte beider Füße in einem Cluster zusammengefasst werden. In (c) hat der linke Fuß den rechten Fuß passiert und in (d) werden beide Füße wieder in je einem Cluster zusammengefasst. Die Beobachtung dieser Abfolge löst eine Personendetektion aus.

eine Person detektiert wurde.

Dafür ist es allerdings notwendig, die jeweiligen Cluster aus dem Zeitschritt $t-1$ auch im nachfolgenden Zeitschritt t wiederzufinden. Da die verwendeten Laser-Range-Scanner alle 20ms einen Scan liefern und ein sich bewegender Fuß langsam genug ist, so dass die dem Fuß zugehörigen Clustermittelpunkte in zwei aufeinanderfolgenden Zeitschritten $\mathbf{c}_{i,t-1}$ und $\mathbf{c}_{j,t}$ nahe beieinander liegen, kann durch einen Nearest-Neighbor-Ansatz die Zuordnung realisiert werden. Dieser Ansatz ordnet den Clustern aus dem aktuellen Zeitschritt t jeweils das Cluster aus dem vorhergehenden Zeitschritt $t-1$ zu, welches die geringste euklidische Distanz zum aktuellen Cluster aufweist — also den



(a) (b) (c)

Abbildung 4.4: Erkennung einer Vereinigung in Laserscans auf Fußhöhe. In dieser Draufsicht sind die Scanpunkte der Füße rot markiert und die Cluster mit einer schwarzen Ellipse angedeutet. Die Scanpunkte der beiden Füße in (a) bilden jeweils ein Cluster und werden einen Zeitschritt weiter in (b) zu einem Cluster zusammengefasst. Um diese Vereinigung algorithmisch zu erkennen, wird geprüft, ob sich die Cluster aus (a) mit dem aus (b) überlappen. Zur Veranschaulichung wurden die Cluster beider Zeitpunkte in (c) übereinandergelegt. Die überlappende Fläche ist blau eingetragen. Beide Cluster aus (a) besitzen mit dem Cluster aus (b) eine Schnittfläche, womit eine Vereinigung vorliegt.

nächsten Nachbarn darstellt (siehe Gleichung 4.2).

$$\mathbf{c}_{i,t-1} \hat{=} \mathbf{c}_{j,t}, \text{ wenn } \|\mathbf{c}_{i,t-1} - \mathbf{c}_{j,t}\| < \epsilon \ \wedge \ j = \underset{k}{\operatorname{argmin}} \|\mathbf{c}_{i,t-1} - \mathbf{c}_{k,t}\| \tag{4.2}$$

Ist die geringste Distanz größer als eine Schwelle ϵ, dann wird keine Zuordnung vorgenommen. In den Experimenten hat sich ein Wert von $\epsilon = 40$cm als brauchbar erwiesen, da selbst die Beine einer rennenden Person diese Strecke innerhalb von zwei Scans nicht überbrücken können.

Ob zwei Cluster im nachfolgenden Zeitschritt zu einem Cluster vereinigt werden, lässt sich durch die Betrachtung der Überschneidungen der beteiligten Cluster erkennen. Schneidet sich das Cluster aus dem Zeitschritt t mit zwei Clustern aus dem Zeitschritt $t-1$, so liegt eine Vereinigung vor (siehe Abbildung 4.4). Schneiden sich zwei Cluster aus dem Zeitschritt t mit nur einem Cluster aus dem Zeitschritt $t-1$, so liegt eine

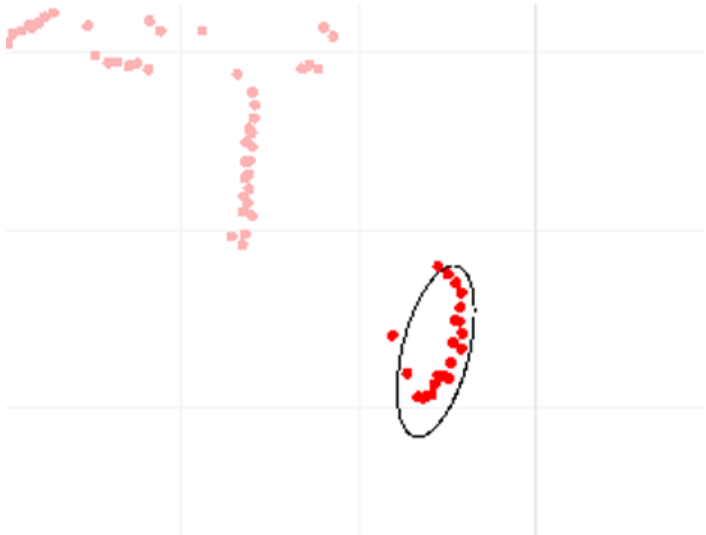


Abbildung 4.5: Detektion eines Oberkörpers in einem Laserscan. In dieser Draufsicht sind die Scanpunkte des Oberkörpers rot und dessen Clusterumfang als schwarze Ellipse eingetragen. Der größte Eigenwert, welcher der Ausdehnung der Clusterellipse entlang der großen Halbachse entspricht, beträgt in diesem Fall etwa 40cm. Somit kann angenommen werden, dass diese Menge an Scanpunkten aus einem Oberkörper resultiert und dadurch erfolgreich eine Person detektiert wurde.

Auftrennung vor.

Wurde der Ablauf des Vereinigen und Auftrennen zweier Cluster beobachtet, also eine Person erfolgreich erkannt, so kann geprüft werden, ob in einem bestimmten Abstand ψ zur Detektion bereits ein Partikelfilter existiert und diese Person trackt. Ist dies nicht der Fall, so wird dort, wo die Person detektiert wurde, ein neuer Partikelfilter (siehe Abschnitt 3.7 und 4.5.1) initialisiert, der dann das Tracking dieser Person übernimmt.

4.4.2 Personendetektion auf Oberkörperhöhe

Die Detektion von Personen auf Höhe der Oberkörper gestaltet sich im Vergleich zur Detektion auf Knöchelhöhe recht einfach. Da in der Ebene, welche sich etwa 1,2m über dem Boden befindet, in den Scandaten hauptsächlich Menschen die dynamischen Objekte in öffentlichen Bereichen darstellen, reicht eine einfache Prüfung der Clusterausdehnung aus. Wird ein Cluster, dessen größter Eigenwert (siehe Abschnitt 3.6)

größer als 15cm ist, gefunden, so wird an dieser Stelle eine Person detektiert (siehe Abbildung 4.5). Der größte Eigenwert ist gleich der größten Halbachse des Clusters, was bei einem Oberkörper die halbe seitliche Ausdehnung darstellt. Eine Untergrenze von 15cm ermöglicht es auch schmale Personen, oder Kinder erfolgreich zu detektieren.

Sobald eine Person erkannt wurde, muss wie im Abschnitt 4.4.1 geprüft werden, ob in einem bestimmten Abstand ψ zur Detektion bereits ein Partikelfilter existiert und diese Person trackt. Ist dem nicht so, kann ein neuer Partikelfilter (siehe Abschnitt 3.7 und 4.5.2) für diese Person initialisiert werden.

4.5 Personentracking

Falls eine neue Person erkannt wurde und für diese Person ein neuer Partikelfilter (siehe dazu auch Abschnitt 3.7) erstellt wurde, muss dieser zuerst in geeigneter Weise initialisiert werden. Dies geschieht, indem die Partikel des Filters um den Ort der Detektion gestreut werden. In den Experimenten hat sich eine Streuung der Position um ±50cm als nutzbringend herausgestellt. Allerdings wird nicht nur die Position der Partikel, sondern auch alle anderen Bewegungsparameter der Partikel in einem bestimmten Bereich gestreut. Näheres dazu wird in den Abschnitten 4.5.1 und 4.5.2 erläutert.

Das eigentliche Tracking nach der Initialisierung erfolgt mit dem im Abschnitt 3.7 beschriebenen Ablauf. In der praktischen Umsetzung des Partikelfilters muss dazu jedoch problemspezifisch das Bewegungsmodell aus Gleichung 3.21 und das Beobachtungsmodell aus Gleichung 3.22 konkretisiert werden, wie es in den nächsten zwei Abschnitten erfolgt. Dabei behandelt Abschnitt 4.5.1 das Personentracking in Laserscandaten, welche auf einer Höhe von ca. 20cm über dem Boden aufgenommen wurden und Abschnitt 4.5.2 beschäftigt sich mit dem Tracking in Laserscandaten, welche auf Höhe des Oberkörpers bei etwa 120cm gemessen wurden. In den Abschnitten wird zuerst auf die eben genannte Streuung der Partikel während der Initialisierung eingegangen. Danach wird das jeweilige Bewegungsmodell beschrieben und gegen Ende der Abschnitte wird das Beobachtungsmodell auf die spezifischen Gegebenheiten der Sensordaten angepasst.

4.5.1 Personentracking auf Fußhöhe

Der Bewegungszustand der Personen, bzw. deren Füße, kann, wie in Abschnitt 3.7.1 schon behandelt, mit einem Vektor $\mathbf{x}_t$ (siehe Gleichung 3.31) beschrieben werden. In diesem Zustandsvektor werden die aktuellen Positionen der beiden Füße $\mathbf{l}_t$ und $\mathbf{r}_t$, die Schrittweite S_t, die Phase φ_t, in welcher sich die periodische Fußbewegung befindet, die Bewegungsrichtung ϑ_t und die Dauer T_t von zwei Schritten gespeichert.

Während der Initialisierung ist es sinnvoll, die Bewegungsparameter der Partikel in geeigneter Weise zu streuen, damit der tatsächliche Bewegungszustand schnell geschätzt werden kann. Die Positionen der Füße $\mathbf{l}_t$ und $\mathbf{r}_t$ werden, wie im vorherigen Abschnitt schon beschrieben, im Umkreis von 50cm um die Personendetektion gestreut. Die Partikelparameter für die Schrittweite S_t wird zwischen 50cm und 150cm angenommen, um sowohl kurze Schritte, als auch lange Schritte, wie sie bei rennenden Personen vorkommen können, erfassen zu können. Da neue Personen detektiert werden, sobald das in Abschnitt 4.4.1 beschriebene Schwingen eines Fußes vorbei am anderen Fuß erkannt wurde, muss die Phase φ_t nicht den ganzen Bereich von 0 bis 2π abdecken, sondern wird um $\frac{\pi}{2}$ und $\frac{3\pi}{2}$ mit $\pm$0,9rad gestreut. Die Bewegungsrichtung hingegen wird im ganzen Bereich von 0 bis 2π angenommen, da Versuche, die Bewegungsrichtung ϑ_t bei der Personendetektion zu schätzen, keine nutzbringenden Ergebnisse geliefert haben. Die Schrittdauer T_t wird zwischen 0,5s, um rennende Personen auch noch tracken zu können, und 1,5s, um langsam gehende Personen nicht zu verlieren, gewählt.

Die Umsetzung des Bewegungsmodells aus Gleichung 3.21 erfolgt durch Berechnen des neuen Zustandes $\mathbf{x}_t$ aus dem vorhergehenden Zustand $\mathbf{x}_{t-1}$. Im ersten Schritt werden dazu die Bewegungsrichtung ϑ_{t-1}, die Dauer zweier Schritte T_{t-1} und die Schrittweite S_{t-1} aus dem letzten Zustand unverändert in den neuen Zustand übernommen. Lediglich die Fußpositionen $\mathbf{l}_t$ und $\mathbf{r}_t$ und die Phase φ_t werden nach Gleichung 3.32 bis 3.36 neu berechnet. Diese Berechnung lässt sich durch Gleichung 4.3 zusammenfassen, wobei allerdings φ_t schon vorher nach Gleichung 3.32 ermittelt werden muss, damit die Fußpositionen korrekt berechnet werden können. Die Definitionen für $f_L(\varphi)$ und

$f_R(\varphi)$ sind in Gleichung 3.33 und 3.34 gegeben.

$$\tilde{\mathbf{x}}_t = \begin{pmatrix} l_{t,x} \\ l_{t,y} \\ r_{t,x} \\ r_{t,y} \\ S_t \\ \varphi_t \\ \vartheta_t \\ T_t \end{pmatrix} = \begin{pmatrix} l_{t-1,x} \\ l_{t-1,y} \\ r_{t-1,x} \\ r_{t-1,y} \\ S_{t-1} \\ \varphi_{t-1} \\ \vartheta_{t-1} \\ T_{t-1} \end{pmatrix} + \begin{pmatrix} S_{t-1} \cdot (f_L(\varphi_t) - f_L(\varphi_{t-1})) \cdot \cos\vartheta_{t-1} \\ S_{t-1} \cdot (f_L(\varphi_t) - f_L(\varphi_{t-1})) \cdot \sin\vartheta_{t-1} \\ S_{t-1} \cdot (f_R(\varphi_t) - f_R(\varphi_{t-1})) \cdot \cos\vartheta_{t-1} \\ S_{t-1} \cdot (f_R(\varphi_t) - f_R(\varphi_{t-1})) \cdot \sin\vartheta_{t-1} \\ 0 \\ 2\pi \cdot \frac{\delta}{T_{t-1}} \\ 0 \\ 0 \end{pmatrix} \tag{4.3}$$

Die Umsetzung des Bewegungsmodells aus Gleichung 3.21 kann somit im zweiten und letzten Schritt durch Gleichung 4.4 erfolgen, in welcher dem neuen Bewegungszustand $\tilde{\mathbf{x}}_t$ ein Rauschterm ω hinzugefügt wird.

$$\mathbf{x}_t = f(\mathbf{x}_{t-1}) = \tilde{\mathbf{x}}_t + (\omega_{l,x}, \omega_{l,y}, \omega_{r,x}, \omega_{r,y}, \omega_S, \omega_\varphi, \omega_\vartheta, \omega_T)^T \tag{4.4}$$

Die Rauschterme für die Fußpositionen $\omega_{l,x}$, $\omega_{l,y}$, $\omega_{r,x}$ und $\omega_{r,y}$ wurden in den Versuchen gleichverteilt mit $\pm$10cm angesetzt. Die Schrittweite wird durch ω_S mit $\pm$15cm verrauscht. Die Terme ω_φ für die Phase, ω_ϑ für die Bewegungsrichtung und ω_T für die Schrittdauer werden mit einem gleichverteilten Rauschen im Bereich, $\pm$0,05rad, $\pm$0,25rad und $\pm$0,05s versehen.

Nachdem das Bewegungsmodell umgesetzt wurde, kann im Folgenden das Beobachtungsmodell spezifiziert werden.

Für die Neuberechnung der Partikelgewichte, anhand der aktuellen Beobachtung, lassen sich die euklidischen Distanzen zwischen den prädizierten Fußpositionen und den gemessenen Scanpunkten im globalen Koordinatensystem derart nutzen, wie es in [Shao et al. 2007] vorgeschlagen wurde. Wenn $\mathbf{p}_i, i = 1, \ldots, N$ die N Scanpunkte des momentanen Zeitschritts t sind, dann lassen sich die Gewichte $w_t^{(j)}$, $j = 1, \ldots, M$ der Partikel nach Gleichung 4.5 und 4.6 bestimmen.

$$\hat{w}_t^{(j)} = o_l \cdot o_r \cdot o_b \cdot o_d \cdot o_f \tag{4.5}$$

$$w_t^{(j)} = \frac{\hat{w}_t^{(j)}}{\sum_{i=1}^{M} \hat{w}_t^{(i)}} \tag{4.6}$$

Dabei sind o_l, o_r, o_b, o_d und o_f Terme, welche verschiedene Gegebenheiten in die Gewichtsberechnung einfließen lassen. Diese Terme werden im Folgenden erklärt:

- o_l und o_r sorgen dafür, dass Partikel, deren Fußpositionen nahe an Scanpunkten liegen, ein hohes Gewicht erhalten, wohingegen die Partikel, deren Fußpositionen weiter von Scanpunkten entfernt liegen, ein geringes Gewicht erhalten. Die Terme werden nach Gleichung 4.7 und 4.8 errechnet, wobei ς bestimmt, auf welche Entfernung die Scanpunkte einen Einfluss auf das Partikelgewicht üben. Ein hoher Wert für ς bewirkt einen Einfluss der Scanpunkte auf großer Distanz und ein geringer Wert beschränkt die Einflussnahme auf nahe liegende Punkte. In den Experimenten wurde $\varsigma = 20$cm verwendet. τ, was in den Experimenten mit $\tau = 0{,}03$ verwendet wurde, ist in den Arbeiten von [Shao et al. 2007] nicht mit enthalten und dient dazu, dass die Gewichtsfunktion in Abhängigkeit zur Entfernung der Partikel zu den Scanpunkten ihre Monotonie beibehalten. Auf diese Problematik wird am Ende des Abschnitts genauer eingegangen.

$$o_l = \sum_{i=1}^{N} \exp\left(\frac{-\left\|\mathbf{p}_i - \mathbf{l}_t^{(j)}\right\|^2}{\varsigma^2}\right) + N \cdot \tau \tag{4.7}$$

$$o_r = \sum_{i=1}^{N} \exp\left(\frac{-\left\|\mathbf{p}_i - \mathbf{r}_t^{(j)}\right\|^2}{\varsigma^2}\right) + N \cdot \tau \tag{4.8}$$

- Der Term o_b sichert die Konsistenz des Fußabstandes mit dem Bewegungsmodell und wird nach Gleichung 4.9 und 4.10 berechnet. Dabei ist s_d der gemessene Fußabstand entlang der Bewegungsrichtung und ξ steuert, bei welcher Differenz von s_d zum erwarteten Fußabstand das Partikel nur noch das 0,6065 fache Gewicht erhält. In den Versuchen wurde $\xi = 25$cm gewählt.

$$s_d = \left(\mathbf{r}_t^{(j)} - \mathbf{l}_t^{(j)}\right)^T \cdot \begin{pmatrix} \cos \vartheta_t^{(j)} \\ \sin \vartheta_t^{(j)} \end{pmatrix} \tag{4.9}$$

$$o_b = \exp\left(\frac{-\left(s_d + S_t^{(j)} \cdot \cos \varphi_t^{(j)}\right)^2}{2\xi^2}\right) \tag{4.10}$$

- Da sich ursprünglich in den Experimenten die Fußpositionen der Partikelfilter oft zusammengezogen haben und auf einem Fuß, anstatt auf zwei Füßen konzentriert haben, wurde der Term o_d eingeführt, welcher in den Arbeiten von [SHAO et al. 2007] nicht vorhanden ist. Dieser erhöht das Gewicht von Partikeln, deren Füße aufgrund der Phase weiter auseinander sind und berechnet sich nach Gleichung 4.11.

$$o_d = \frac{1}{2}\cos^2\left(\varphi_t^{(j)}\right) + 1 \tag{4.11}$$

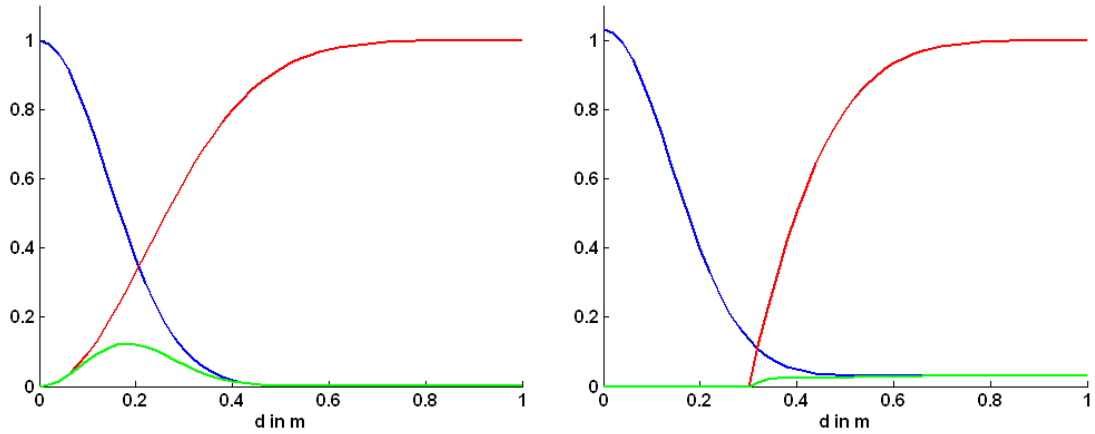
- Damit bei parallel gehenden Personen der Partikelfilter der einen Person nicht von den Scanpunkten der anderen Person angezogen wird, ist es notwendig, dass sich Partikelfilter untereinander abstoßen, was mit dem Term o_f (beschrieben in Gleichung 4.12 und 4.13) realisiert wird, welcher allerdings in [SHAO et al. 2007] nicht näher erläutert wird und in eigener Arbeit umgesetzt werden musste. Dabei ist $d_{j,k}$ die geringste Distanz der Fußpositionen des Partikels j zu den Fußhypothesen des Partikelfilters k.

$$o_p\left(d\right) = \max\left(0, 1 - \exp\left(\frac{-d^2 + \sigma^2}{\left(\rho - \sigma\right)^2}\right)\right) \tag{4.12}$$

$$o_f = \prod_k o_p\left(d_{j,k}\right) \tag{4.13}$$

  σ ist ein Parameter, welcher die Monotonie der Gewichtsfunktion in Abhängigkeit zur Entfernung der Partikel zu den Scanpunkten sichert (vergleiche Abbildung 4.6) und mit $\sigma = 0{,}3$m bestimmt wurde. Das bedeutet, dass ein Partikel bei einer Entfernung von weniger als 30cm zur Hypothese eines andern Partikelfilters ein Gewicht von 0 erhält. ρ steuert, bis zu welcher Entfernung sich Partikelfilter untereinander noch maßgeblich abstoßen und wurde mit $\rho = 0{,}63$m bestimmt. Das bedeutet, dass der Einfluss eines anderen Partikelfilters bei einer Distanz von 0,63m bereits vernachlässigbar gering ist.

  Ohne die Parameter $\tau = 0{,}03$ und $\sigma = 0{,}3$m kann es passieren, dass die Partikel eines Filters, welcher eine Person trackt, in Richtung einer anderen, parallel nebenher laufenden Person wandern. Wird diese Person ebenfalls von einem Partikelfilter getrackt, dann halten die Partikel aufgrund des Terms o_f eine gewisse



(a) Gewichtsfunktion mit $\sigma = 0$m und $\tau = 0$.

(b) Gewichtsfunktion mit $\sigma = 0{,}3$m und $\tau = 0{,}03$.

Abbildung 4.6: Gewichtsfunktion eines Partikels in Abhängigkeit der Entfernung zu Scanpunkten, auf welchen schon die Hypothese eines anderen Partikelfilters liegen. Der anziehende Faktor der Scanpunkte ist blau, der abstoßende Faktor des anderen Partikelfilters ist rot und das Produkt beider Faktoren, was die eigentliche Gewichtsfunktion darstellt, ist grün dargestellt. Ohne die Korrekturparameter σ und τ bildet sich in (a) ein Maximum bei etwa 18cm, was bewirkt, dass sich dort in kurzer Zeit die Partikel des Filters sammeln. Mit $\sigma = 0{,}3$m und $\tau = 0{,}03$ wird das Maximum in (b) vermieden, wodurch erfolgreich verhindert werden kann, dass sich ein anderer Partikelfilter auf die bereits getrackten Scanpunkte festsetzt.

Distanz zur anderen Person, bei der sich der Einfluss der Anziehung und der Abstoßung aufheben. Nach einigen Zeitschritten hat dadurch der Partikelfilter die ursprüngliche Person verloren und dessen Personenhypothese verweilt in einigem Abstand zur anderen Person. Da der Partikelfilter, welcher die andere Person trackt, von diesen Hypothesen aufgrund des Terms o_f ebenfalls abgestoßen wird, befinden sich nach wenigen Zeitschritten beide Partikelfilter neben der Person auf gegenüberliegenden Seiten. Dieses unerwünschte Verhalten resultiert aus einem lokalen Maximum in der Gewichtsfunktion (siehe Abbildung 4.6). Diese Funktion beschreibt das Gewicht eines Partikels in Abhängigkeit zur Distanz zu einem Scanpunkt, auf dem bereits eine Personenhypothese eines anderen Partikelfilters liegt. Da prinzipbedingt die Partikel mit hohem Gewicht mit höherer

Wahrscheinlichkeit in den nächsten Zeitschritt übernommen werden, sammeln sich nach einer gewissen Zeit die Partikel in diesem Maximum. Durch die Parameter τ und σ kann dieses lokale Maximum allerdings vermieden werden.

4.5.2 Personentracking auf Oberkörperhöhe

Der Bewegungszustand der Personen, bzw. deren Oberkörper, kann unter der Annahme, dass sich der Oberkörper einer Person annähernd linear in eine Richtung bewegt, wie in Abschnitt 3.7.2 schon behandelt, mit einem Vektor $\mathbf{x}_t$ (siehe Gleichung 3.37) beschrieben werden. In diesem Zustandsvektor wird die aktuelle Position des Oberkörpers X_t und Y_t, die Bewegungsrichtung ϑ_t und die Geschwindigkeit v_t der Person gespeichert.

Während der Initialisierung ist es sinnvoll, die Bewegungsparameter der Partikel in geeigneter Weise zu streuen, damit der tatsächliche Bewegungszustand schnell geschätzt werden kann. Die Positionen des Oberkörpers werden, wie im Abschnitt 4.5 beschrieben, im Umkreis von 50cm um die Personendetektion gestreut. Die Partikelparameter für die Bewegungsrichtung ϑ_t werden im ganzen Bereich von 0 bis 2π angenommen, da bei ungünstigem Zusammenspiel aus Verdeckungen die Bewegungsrichtung nicht zuverlässig aus einer einzigen Beobachtung geschätzt werden kann. Um sowohl still stehende Personen, als auch rennende Personen von Anfang an erfolgreich tracken zu können, wird die Geschwindigkeit v_t in den Partikeln von $0\frac{\mathrm{m}}{\mathrm{s}}$ bis $4\frac{\mathrm{m}}{\mathrm{s}}$ gestreut.

Die Umsetzung des Bewegungsmodells aus Gleichung 3.21 erfolgt durch Berechnen des neuen Zustandes $\mathbf{x}_t$ aus dem vorhergehenden Zustand $\mathbf{x}_{t-1}$, anhand von Gleichung 3.38. Dabei werden die Bewegungsrichtung und die Geschwindigkeit vorerst unverändert aus dem vorhergehenden Zustand übernommen. Die Position wird anhand der Geschwindigkeit und Bewegungsrichtung für den aktuellen Zeitschritt neu errechnet. Daraufhin werden sämtliche Parameter mit einem Rauschterm ω behaftet. Gleichung 4.14 fasst nochmals beide Schritte in Einem zusammen. Dabei ist δ die Zeit, welche

zwischen dem vorherigen und dem aktuellen Zeitschritt vergangen ist.

$$\mathbf{x}_t = \begin{pmatrix} X_t \\ Y_t \\ \vartheta_t \\ v_t \end{pmatrix} = \begin{pmatrix} X_{t-1} \\ Y_{t-1} \\ \vartheta_{t-1} \\ v_{t-1} \end{pmatrix} + \begin{pmatrix} \cos(\vartheta_{t-1}) \\ \sin(\vartheta_{t-1}) \\ 0 \\ 0 \end{pmatrix} \cdot \delta \cdot v_{t-1} + \begin{pmatrix} \omega_X \\ \omega_Y \\ \omega_\vartheta \\ \omega_v \end{pmatrix} \tag{4.14}$$

Die Rauschterme für die Position ω_X und ω_Y wurden in den Versuchen gleichverteilt mit $\pm$5cm angesetzt. Die Bewegungsrichtung wird durch ω_ϑ mit $\pm$0,2rad verrauscht. Der Term ω_v für die Geschwindigkeit wird mit einem gleichverteilten Rauschen im Bereich von $\pm 0{,}1\frac{\mathrm{m}}{\mathrm{s}}$ versehen.

Das Beobachtungsmodell berücksichtigt, ähnlich dem Fußtracking aus Abschnitt 4.5.1, die euklidischen Distanzen zwischen den prädizierten Oberkörperpositionen $\mathbf{m}_t^{(j)} = \left(X_t^{(j)}, Y_t^{(j)}\right)^T$ und den gemessenen Scanpunkten im globalen Koordinatensystem: Wenn $\mathbf{p}_i$, $i = 1, \ldots, N$ die Scanpunkte des momentanen Zeitschritts sind, dann lassen sich die Gewichte $w_t^{(j)}$, $j = 1, \ldots, M$ der Partikel nach Gleichung 4.15 bestimmen.

$$w_t^{(j)} = \sum_{i=1}^{N} \exp\left(\frac{-\left\|\mathbf{p}_i - \mathbf{m}_t^{(j)}\right\|^2}{\zeta^2}\right) \tag{4.15}$$

Der Parameter ζ steuert dabei, auf welcher Entfernung sich noch Scanpunkte auf das Gewicht auswirken. Ein hoher Wert ermöglicht es, dass weit entfernt liegende Punkte das Partikelgewicht beeinflussen, wohingegen ein kleiner Wert für ζ den Einfluss auf eine geringe Distanz beschränkt. In den Versuchen wurde $\zeta = 50$cm gewählt.

Um zu vermeiden, dass Partikelfilter von einer Person auf eine Andere wechseln und somit mehrere Partikelfilter ein und die selbe Person tracken, wurde ein anderer Ansatz verwendet, als im Abschnitt 4.5.1. Die Partikelfilter können dazu gebracht werden, nur ein und die selbe Person zu tracken, indem zum Updateschritt nur Punkte verwendet werden, die auch zur getrackten Person gehören. Um zu beurteilen, welcher Punkt zu welcher Person gehört, können die Hypothesen der Partikelfilter selbst verwendet werden. Dies mag paradox erscheinen, da sich die Partikelfilter im Prinzip selbst mitteilen, welche Punkte zur Person gehören, die erkannt und getrackt werden soll. Dies ist jedoch nur ein vermeintliches Paradoxon, da bereits bei der Initialisierung des Partikelfilters wie in Abschnitt 4.4.2 beschrieben, eine zuverlässige Personendetektion vorliegt.

Somit kann schon von Beginn an davon ausgegangen werden, dass die Hypothese des Partikelfilters zumindest in nächster Nähe zur korrekten Position der Person liegt und im initialen Updateschritt nur die Scanpunkte verwendet werden, welche zur Person selbst gehören. Dadurch ist es dem Partikelfilter in den ersten Zeitschritten möglich, die Bewegungsparameter $\mathbf{x}_t$ mit ausreichender Genauigkeit zu schätzen und die Person auch zukünftig zuverlässig zu tracken, sofern sie nicht sofort nach der Detektion verdeckt wird.

Die Zuordnung der Punkte $\mathbf{p}_i$ zu den K Partikelfiltern lässt sich durch eine Parzellierung der Scanpunkte vornehmen. Dabei wird für jeden Punkt die minimale Entfernung $d_{i,k,\min}$ zu den aktuellen Personenhypothesen $\hat{\mathbf{m}}^{(k)}$, $k = 1, \ldots, K$ der Partikelfilter ermittelt (siehe Gleichung 4.16 und 4.17).

$$k_{i,\min} = \underset{k}{\operatorname{argmin}} \left\| \hat{\mathbf{m}}^{(k)} - \mathbf{p}_i \right\| \tag{4.16}$$

$$d_{i,k,\min} = \left\| \hat{\mathbf{m}}^{\left(k_{i,\min}\right)} - \mathbf{p}_i \right\| \tag{4.17}$$

$$\mathbf{p}_i \mapsto \begin{cases} \text{Partikelfilter } k_{i,\min}, & \text{wenn } d_{i,k,\min} < \eta \\ \text{Partikelfilter } k = 1, \ldots, K & \text{sonst} \end{cases} \tag{4.18}$$

Ist diese Distanz geringer als eine gewisse Schwelle η, welche in den Versuchen mit $\eta = 1{,}5\text{m}$ festgelegt wurde, so wird der Punkt dem Partikelfilter $k_{i,\min}$ zugeordnet. Ist die Distanz größer, so wird der Punkt allen Partikelfiltern zugeordnet (siehe Gleichung 4.18).

4.6 Nachverarbeitung

Nachdem in den vorherigen Abschnitten geklärt wurde, wie auf Basis von Laserscannern Personen über die Zeit mit Hilfe von Partikelfiltern verfolgt werden können, müssen noch die Personenhypothesen der einzelnen Filter ausgewertet werden, um die eigentlichen Personenspuren zu erhalten. Dazu wird jedem Partikelfilter bei der Erzeugung eine eineindeutige ID zugewiesen. Diese ID wird ermittelt, indem ganz zu Beginn des Trackings, also noch bevor irgendein Partikelfilter angelegt wurde, ein

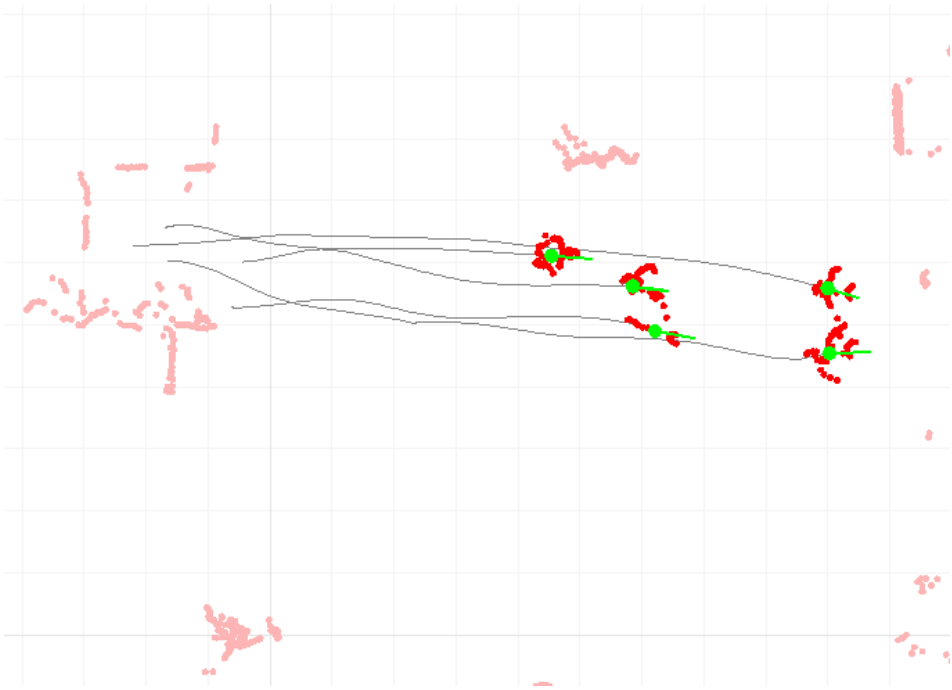


Abbildung 4.7: Ermittelte Bewegungsspuren von fünf Personen in der Draufsicht. Die zu den Personen gehörenden Scanpunkte sind rot und die Hintergrundpunkte sind blassrot eingetragen. Die Personenhypothesen der Partikelfilter zum aktuellen Zeitschritt sind grün eingetragen, wobei der Kreis die Position und der vom Kreis ausgehende Strich die Bewegungsrichtung angibt. Die Spuren sind grau eingetragen. Die Personen haben diese Szene über eine Treppe betreten, die im Bild links als eine Menge von Hintergrundpunkten zu sehen ist.

Zähler $U = 0$ gesetzt wird. Wurde eine neue Person in der Szene entdeckt und daraufhin ein Partikelfilter an dieser Stelle initialisiert, so wird diesem Partikelfilter die ID U zugewiesen und U selbst um eins inkrementiert. Somit kann sichergestellt werden, dass jede ID nur einmal verwendet wird. In den Versuchen wurde die einfache Art der Trajektorienrepräsentation gewählt, wie sie in Abschnitt 3.8 beschrieben wird, da zum Zeitpunkt der Experimente keine weitere Auswertung der Trajektorien, bis auf eine visuelle Überprüfung, vorgesehen war. Somit werden also für jeden Zeitschritt t die Bewegungsparameter (oder auch Personenhypothesen) der O Partikelfilter $\mathbf{x}_t^{(i)}$, $i = 1, \ldots, O$ nach Gleichung 3.25 ermittelt und zusammen mit dem Zeitstempel des aktuellen Zeitschrittes und der ID des Partikelfilters gespeichert. Um die Trajektori-

en $B^{(i)}$ zu bilden, muss für jede ID eine Liste angelegt werden, in der hintereinander weg die Zeitstempel und die dazugehörigen Bewegungsparameter gespeichert werden. Danach können die Trajektorien für die weitere Verarbeitung geglättet werden, wie es in Abschnitt 3.8.3 beschrieben wird. Diese Trajektorien können dann beispielsweise in einer Grafik ausgegeben werden, so wie es auch in Abbildung 4.7 zu sehen ist.

4.7 Zusammenfassung

In diesem Kapitel wurden neben dem Messequipment (Abschnitt 4.2) die Teile des Personentrackers näher erläutert, welche iterativ zu jedem Zeitschritt abgearbeitet werden. Zuerst werden die Daten der Laser-Range-Scanner im Vorverarbeitungsschritt für die Personendetektion und für das Personentracking aufbereitet (Abschnitt 4.3). Dabei werden die Scandaten von Hintergrundobjekten befreit, in eine globale Punktemenge überführt und zu Punktecluster zusammengefasst. Diese Cluster werden bei der Personendetektion (Abschnitt 4.4) dahingehend geprüft, ob sie eine Person darstellen, oder nicht. Im Falle einer Detektion wird ein neuer Partikelfilter initialisiert, der diese Person dann trackt, sofern dies nicht schon ein bereits existierender Partikelfilter tut. Nachdem die Personendetektion abgeschlossen ist, werden beim Personentracking (Abschnitt 4.5) zuerst alle einzelnen Partikelfilter daraufhin geprüft, ob sie noch eine Person tracken, oder nicht. Im letzteren Fall werden die entsprechenden Partikelfilter gelöscht. Alle verbleibenden Partikelfilter werden dann mit den Scanpunkten aktualisiert. Nachdem daraufhin die Personenhypothesen der Filter ausgelesen und die Bewegungsspuren aktualisiert wurden (Abschnitt 4.6), kann mit der nächsten Iteration begonnen werden. In Abbildung 4.8 wurde der Ablauf einer Iteration noch einmal in Form von Pseudocode dargestellt. Als Eingaben werden die Posen der Laser-Range-Scanner $\left(x_t^{(s)}, y_t^{(s)}, \omega_t^{(s)}\right)^T$, die aktuellen Entfernungsmessungen $\mathbf{d}_t^{(s)}$, die aus den Histogrammen gebildeten Schwellwerte $\tau^{(s)}$ und die Partikelfilter $\left\langle P_{t-1}^{(f)}, \mathbf{x}_{t-1}^{(f)}, ID^{(f)} \right\rangle$ des letzten Zeitschrittes übergeben. Daraus werden mit dem gerade beschriebenen Ablauf die aktualisierten Partikelfilter $\left\langle P_t^{(f)}, \mathbf{x}_t^{(f)}, ID^{(f)} \right\rangle$ berechnet.
Die konkrete Umsetzung dieses Ablaufes in einer Software wird im folgenden Kapitel genauer beschrieben.

Eingaben

1 $\left(x_t^{(s)}, y_t^{(s)}, \omega_t^{(s)}\right)^T, s = 1, \ldots, \Sigma$ //Posen der Σ Scanner

2 $\mathbf{d}_t^{(s)}, s = 1, \ldots, \Sigma$ //aktuelle Entfernungsvektoren der Laserscanner (Abschnitt 3.2)

3 $\tau^{(s)}, s = 1, \ldots, \Sigma$ //Hintergrundschwellwerte der Laserscanner (Abschnitt 3.4)

4 $\left\langle P_{t-1}^{(f)}, \mathbf{x}_{t-1}^{(f)}, ID^{(f)} \right\rangle, f = 1, \ldots, O_{t-1}$ //Partikel, Hypothesen und IDs der O_{t-1} Filter

Algorithmus

Vorverarbeitung:

5 FOREACH Scanner s;

6 Bilde $Punkte^{(s)} = \left\{\mathbf{p}_i^{(s)} | i = 0, \ldots, 540\right\}$ aus $\mathbf{d}_t^{(s)}$ mit $\left(x_t^{(s)}, y_t^{(s)}, \omega_t^{(s)}\right)^T$; //Gl. 3.2

7 Bilde $Segmente^{(s)} = \left\{\mathbf{s}_i^{(s)} | i = 1, \ldots, N^{(s)}\right\}$ aus $Punkte^{(s)}$ mit $\mathbf{d}_t^{(s)}$; //Abs. 4.3.1

8 Entferne Hintergrund aus $Punkte^{(s)}$ u. $Segmente^{(s)}$ mit $\tau^{(s)}$ und $\mathbf{d}_t^{(s)}$; //Abs.3.4

9 ENDFOREACH;

10 Bilde $Punkte = \left\{Punkte^{(s)} | s = 1, \ldots, \Sigma\right\}$; //Punktemenge aller Laserscanner

11 Bilde $Segmente = \left\{Segmente^{(s)} | s = 1, \ldots, \Sigma\right\}$; //Segmente aller Laserscanner

12 Berechne Cluster $\mathbf{c}_j$ aus $Segmente$; //Abschnitt 3.5, 3.6 und 4.3.2

Personendetektion:

13 FOREACH $\mathbf{c}_j$;

14 IF $\mathbf{c}_j$ ist Person AND $dist\left(\mathbf{c}_j, P_{t-1}^{(f)}\right) > \psi \ \forall \ f$; //Neue Person detektiert (Abs. 4.4)

15 Initialisiere neuen Partikelfilter mit $ID = U$ bei $\mathbf{c}_j$;

16 Inkrementiere U;

17 ENDIF;

18 ENDFOREACH;

Personentracking:

19 FOREACH Partikelfilter f;

20 IF f hat Person verloren; //Abschnitt 3.7.3

21 Lösche f;

22 ELSE;

23 Berechne $\tilde{P}_t^{(f)}$ aus $P_{t-1}^{(f)}$; //Gleichung 3.21 und Abschnitt 4.5

24 Berechne $\mathbf{w}_t^{(f)}$ für $\tilde{P}_t^{(f)}$ mit $Punkte$; //Gleichung 3.22 und Abschnitt 4.5

25 Bilde $P_t^{(f)}$ aus $\tilde{P}_t^{(f)}$; //siehe Resampling im Abschnitt 3.7

26 Berechne Personenhypothese $\mathbf{x}_t^{(f)}$ aus $P_t^{(f)}$; //Gleichung 3.25

27 ENDIF;

28 ENDFOREACH;

Rückgabe

29 $\left\langle P_t^{(f)}, \mathbf{x}_t^{(f)}, ID^{(f)} \right\rangle, f = 1, \ldots, O_t$ //aktualisierte Partikelfilter

Abbildung 4.8: Pseudocodedarstellung für den Trackingalgorithmus.

Kapitel 5

Softwaretechnische Umsetzung

Damit die Mechanismen aus den vorherigen Kapiteln praktisch verwendet werden können, müssen sie in einer Software umgesetzt werden. In diesem Kapitel wird beschrieben, wie die Umsetzung in dieser Masterarbeit realisiert wurde.

Die Trackingsoftware, welche im Rahmen dieser Arbeit entstand, ist in mehrere, aneinandergekettete Module gegliedert (siehe Abbildung 5.1). Der modulare Aufbau ermöglicht es, verschiedene Ansätze zur algorithmischen Umsetzung des Trackers verhältnismäßig einfach auszuprobieren, da so nur einzelne Module ersetzt werden müssen, und nicht der komplette Programmcode neu geschrieben werden muss. Es kann z. B. das Trackingmodul ausgetauscht werden, um zwischen dem Oberkörpertracker und dem Fußtracker zu wechseln. Es wäre durchaus auch möglich, dies während der Ausführung des Programms zu tun, ohne dass der Code neu compiliert werden muss.

Die Wahl der Modularisierung trägt dem Datenfluss Rechnung. So steht am Beginn der Kette die Datenbeschaffung von den Laser-Range-Scannern per Ethernet mittels mehrerer LRS-Clienten. Dazu wird für jeden Scanner eine eigene Modulinstanz in einem Thread erzeugt. Diese Instanz kann sich zum Laserscanner verbinden, ermöglicht deren Steuerung und verarbeitet die empfangenen Datenpakete. Die Einbettung der Clienten in Threads wurde zum einen gewählt, um ein paralleles Auslesen der Daten zu ermöglichen und zum anderen kann es durch die Auslagerung in Threads ermöglicht werden, dass die dem Programm verfügbaren Scandaten immer auf dem neuesten Stand sind. Des Weiteren ist auch ein kontinuierliches Mitschreiben der Scandaten auf die Fest-

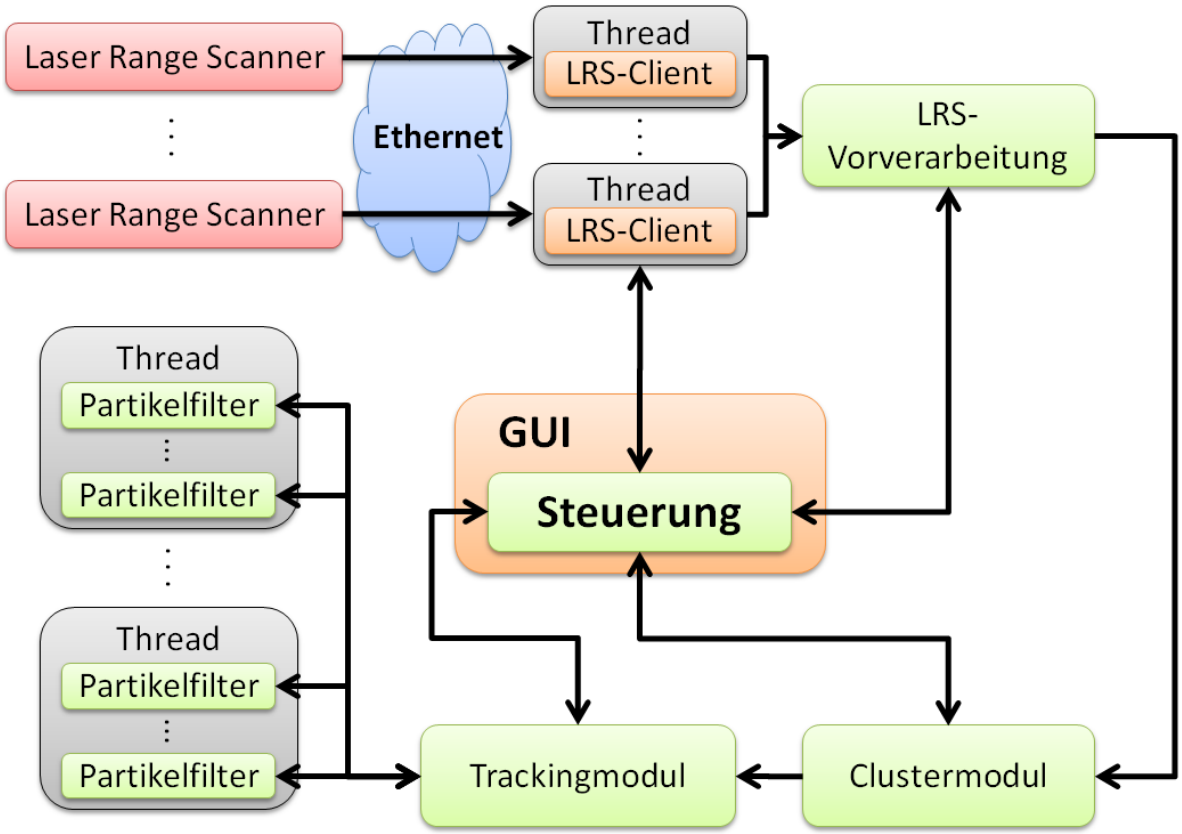


Abbildung 5.1: Modularer Aufbau der Trackingsoftware. Im Zentrum steht die Steuerung und das Graphische Nutzerinterface (GUI), über welche sämtliche Komponenten gesteuert werden können. Die Laser-Range-Scanner sind periphere Geräte und kein eigentlicher Bestandteil der Software. Die LRS-Clienten und die GUI bilden die Schnittstellen der Software zur Außenwelt.

platte möglich, ohne dass der Trackingalgorithmus die CPU für die Datenbeschaffung blockiert und so Scans verloren gehen. Nachdem die eingangsseitigen Datenpakete der jeweiligen Laserscanner ausgelesen wurden, werden die letzten empfangenen Daten auf der Ausgangsseite der Module dem nachfolgendem Modul zur Verfügung gestellt. Die Scandaten beinhalten im einzelnen den Entfernungsvektor (siehe Abschnitt 3.3), den Remittanzwertvektor, den Gerätestatus des Scanners und den Zeitstempel, wann das Datenpaket eintraf.

Die letzten empfangenen Scandaten aller Client-Module werden daraufhin von einem Vorverarbeitungsmodul entgegengenommen, welches die Entfernungsscans in ein glo-

bales Koordinatensystem als Punktemenge überführt (siehe dazu Abschnitt 3.3). Für jede Instanz der Clienten wird auch eine Konfiguration mitgeführt. Sie besteht aus der Pose des jeweiligen Laser-Range-Scanners im Raum und aus einem Histogramm, anhand dessen die jeweiligen Punkte nach Vordergrund und Hintergrund klassifiziert werden (siehe dazu Abschnitt 3.4). Diese Konfiguration kann sowohl auf Festplatte gespeichert, als auch gelesen werden, damit es möglich ist, zuerst während eines geringen Personenaufkommens die Histogramme erstellen zu lassen und die Posen der Laser-Range-Scanner zu bestimmen um sie dann später während des eigentlichen Einsatzes sofort zur Verfügung zu haben. Das Vorverarbeitungsmodul dient neben der Umrechnung und Klassifikation der Scandaten auch als Schnittstelle des Steuerungsmoduls zu den Laserscannern, um dem Nutzer der Software die Möglichkeit zu geben, deren Betriebszustand zu kontrollieren. Als Ausgabe des Vorverarbeitungsmoduls stehen einerseits die klassifizierten Punkte im globalen Koordinatensystem zu Verfügung und andererseits auch die schon zu Segmenten zusammengefassten Punkte der einzelnen Scanner (siehe Abschnitt 4.3.1).

Diese Segmente werden dann dem Clustermodul zur Verfügung gestellt, welches, wie in Abschnitt 3.5 beschrieben, die Segmente und deren Vordergrundpunkte zu Punktecluster zusammenfasst. Zu den gefundenen Clustern werden neben dessen Mittelpunkt und Kovarianzmatrix auch noch die Eigenvektoren und Eigenwerte, sowie die im Cluster enthaltenen Punkte gespeichert.

Die gefundenen Cluster nimmt das Trackingmodul entgegen. Dieses Modul hält eine variable Anzahl von Partikelfiltern, welche am Anfang von jedem neuen Zeitschritt daraufhin geprüft werden, ob sie noch eine Person tracken (vergleiche dazu Abschnitt 3.7.3). Danach werden deren Personenhypothesen ausgelesen. Mit der Hilfe dieser Hypothesen wird versucht, in den gefundenen Clustern des Clustermoduls Personen zu detektieren. Wenn neue Personen gefunden wurden, werden an entsprechender Stelle neue Partikelfilter angelegt. Sobald dieser Schritt abgeschlossen ist, werden die Punkte der Cluster den Partikelfiltern übergeben, deren Updateschritte dann in mehreren Threads parallel berechnet werden. Die Berechnung in Threads ermöglicht es, von Mehrkernprozessoren zu profitieren, da somit mehrere Partikelfilter zeitlich parallel ihre rechenaufwändigen Kalkulationen durchführen können. Dabei ist die Anzahl der

Abbildung 5.2: Benutzeroberfläche der Trackingsoftware. Links sind die Bedienelemente angeordnet, während der restliche Platz in einer Draufsicht der Visualisierung der Daten dient.

Threads von Anfang an festgelegt, um die Menge der zur Verfügung stehenden CPU-Kerne optimal auszunutzen und Threadwechsel nach Möglichkeit vermeiden zu können. Die variable Anzahl von Partikelfiltern wird gleichmäßig auf die feste Anzahl von Threads verteilt. Somit werden, je nach Anzahl der Partikelfilter, mehrere Filter in einem Thread sequentiell und die Threads an sich parallel abgearbeitet. Nachdem die Partikelfilter ihre Berechnungen abgeschlossen haben, stehen die Bewegungsparameter der getrackten Personen des aktuellen Zeitschritts für die nachfolgenden Module zur Verfügung.

Die Partikelfilter als eigenständiges Modul nehmen zwecks Initialisierung Parameter wie die anfängliche Position der Person, Anzahl der zu berechnenden Partikel und die Streubreiten der Bewegungsparameter als Eingabe an. Nach der Initialisierung werden den Partikelfiltern zu jedem neuen Zeitschritt die zu berücksichtigenden Punkte übergeben, damit anhand derer die Partikelgewichte, so wie es in Abschnitt 3.7 beschrieben

wurde, neu berechnet werden können. Dabei werden erst alle Partikelhypothesen für den aktuellen Zeitschritt berechnet. Dem folgt die eigentliche Berechnung der Partikelgewichte und des Nützlichkeitswertes, wie er in Abschnitt 3.7.3 erklärt wurde. Daraufhin wird der Resamplingschritt ausgeführt, der aus der aktuellen Partikelmenge anhand der Gewichte eine neue Menge bildet. Schlussendlich können die aktuellen Personenhypothesen vom Partikelfilter abgefragt werden.
Die Spuren der Personenhypothesen werden vom Steuerungsmodul gebildet, verwaltet und auf Festplatte mitgeschrieben. Zudem wird von diesem Modul auch der gesamte Datenfluss sichergestellt und überwacht. Über die GUI (gezeigt in Abbildung 5.2) ist es dem Nutzer der Software möglich, die Steuerung zu beeinflussen und wahlweise Daten aus verschieden Stellen des Datenflusses anzeigen zu lassen.
Mit Hilfe dieser Software wurden auch die Experimente durchgeführt, welche im folgenden Kapitel erläutert werden.

Kapitel 6

Experimentelle Untersuchungen

6.1 Einleitung

Nachdem dargelegt wurde, wie mit den Mechanismen aus Kapitel 3 und 4 die Bewegungsspuren von Personen ermittelt werden können und im Kapitel 5 diese Mechanismen in einer Software umgesetzt wurden, werden in diesem Kapitel die damit erzielten Ergebnisse des Oberkörpertrackers und des Fußtrackers erläutert, um eine Aussage über deren Leistungsfähigkeit zu erhalten und die Ursachen für eventuelle Schwachstellen herauszufinden.

6.2 Einsatzumgebungen

Die Experimente wurden in drei verschiedenen Einsatzumgebungen durchgeführt. Die Daten für die ersten Versuche wurden in einem Durchgangsbereich des Flughafen Hannover aufgenommen. Später wurden Aufnahmen im Foyer des Humboldt-Gebäudes der TU-Ilmenau durchgeführt und die letzten Aufnahmen stammen vom Eingangsbereich des Flughafen Erfurt.
In Hannover kamen insgesamt drei Laser-Range-Scanner auf Fußhöhe zum Einsatz. Der Durchgangsbereich verbindet zwei Gates und ist daher stark frequentiert. In dem von den Scannern erfassten Areal befindet sich zudem der Zugang zu den Toiletten und der Eingang zu einer Sparkassenfiliale (siehe Abbildung 6.1). Wie auf Flughäfen

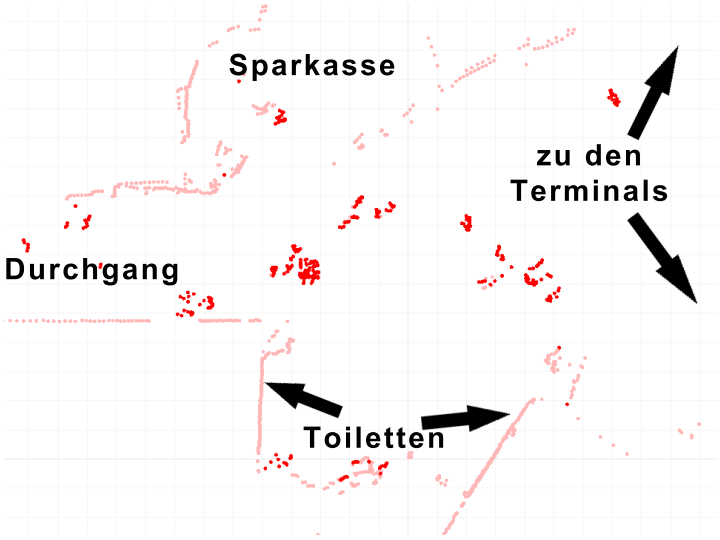


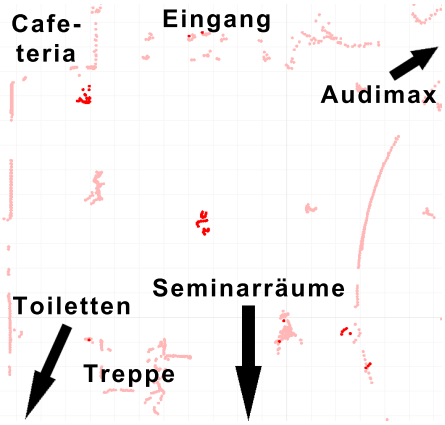
Abbildung 6.1: Laserscan der Szene auf dem Flughafen Hannover in der Draufsicht. Der Hintergrund, welcher hauptsächlich aus Wänden besteht, ist blassrot eingetragen und alle dynamischen Objekte, hier Personen und Trolleys, sind rot markiert. Der Hauptstrom der Personen floss zwischen dem Durchgang und den Terminals. Einige haben zwischendrin die Toilette besucht und wenige sind zur Sparkasse abgebogen.

üblich, sind die Personen, welche die Szene durchquert haben, oftmals mit Gepäck und Kofferwagen unterwegs gewesen, was das Tracking erschwert. Insgesamt wurden etwa 15 Minuten lang Daten aufgenommen.

Im Foyer des Humboldt-Gebäudes (Abbildung 6.2) kamen vier Laser-Range-Scanner zum Einsatz: jeweils zwei auf Fußhöhe und zwei auf Oberkörperhöhe. Dabei befanden sich die Oberkörper-Scanner in etwa einen Meter über den Fuß-Scannern, um die beiden entsprechenden Verfahren möglichst gut miteinander vergleichen zu können. Neben einer gestellten Szene mit zwei Personen, die aufeinander zugehen, wurden auch etwa neun Minuten lang Daten aufgenommen, während Studenten das Gebäude betraten, oder verließen.

Auf dem Flughafen Erfurt wurde mit fünf Laser-Range-Scannern die Haupthalle (Abbildung 6.3) auf Fußhöhe beobachtet, kurz nachdem eine Maschine gelandet ist. Demzufolge haben viele Personen den Weg von der Gepäckausgabe bis zum Ausgang durchquert. Der Großteil der Personen war mit Gepäck unterwegs, welches in den Scandaten

(a) Übersicht über die Szene an der TU-Ilmenau, während keine Personen in der Szene vorhanden waren. Die Stellwände rechts standen während der Aufnahmen dort nicht.



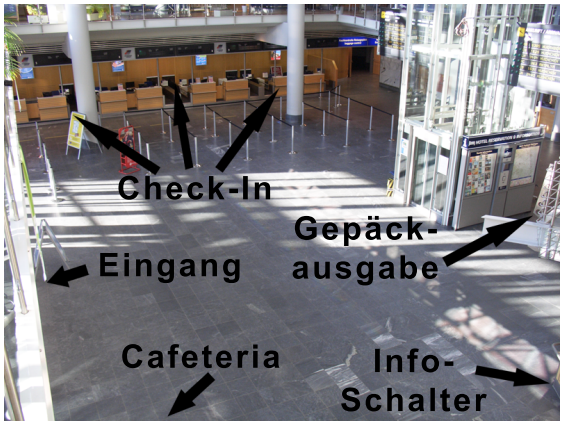
(b) Laserscan der Szene in der Draufsicht. Fest stehende Objekte sind blassrot eingetragen und alle Personen sind rot markiert.

Abbildung 6.2: Szene im Foyer des Humboldt-Gebäudes der TU-Ilmenau. Viele Personen bewegten sich zwischen den Seminarräumen, der Treppe und dem Eingang. Einige haben das Gebäude nur betreten, um direkt zur Cafeteria zu gelangen. Das Audimax war zur Zeit der Aufnahmen nicht belegt, weswegen in dieser Richtung sich auch keine Personen bewegt haben.

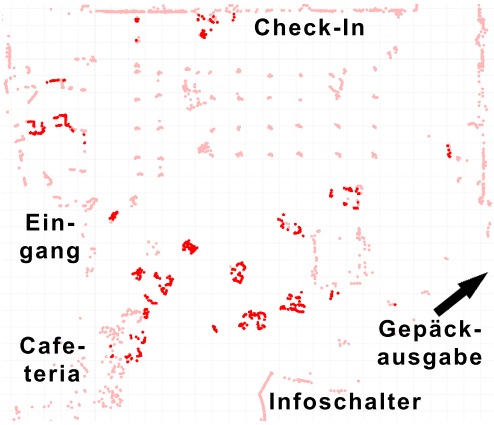
auch zu erkennen ist. Die Aufnahme dauerte etwa 14 Minuten.
Die Aufnahmen in Ilmenau sind die Einzigen, bei denen die Szene sowohl auf Fuß- als auch auf Oberkörperhöhe erfasst wurde. Da somit ein direkter Vergleich zwischen beiden Verfahren möglich ist, wurden die Ilmenauer Aufnahmen hauptsächlich zur Evaluierung verwendet, wohingegen die anderen Aufnahmen zum Entwickeln und Testen der Verfahren genutzt wurden.
Die Hardware, auf die sich die Angaben zur Rechenperformanz in den nachfolgenden Abschnitten beziehen, ist ein Intel Core2Quad, welcher vier CPU-Kerne mit jeweils 2,83GHz Taktfrequenz besitzt. Die Software wurde mit SSE2-Optimierung compiliert.

6.3 Ergebnisse des Fußtrackings

Da keine Ground-Truth zu den aufgenommenen Daten vorhanden ist, ist eine quantitative Bewertung der Algorithmen kaum möglich. Die Experimente haben gezeigt, dass



(a) Übersicht über die Szene auf dem Flughafen Erfurt, während keine Personen in der Szene vorhanden waren.



(b) Laserscan der Szene in der Draufsicht. Fest stehende Objekte sind blassrot eingetragen und alle Personen sind rot markiert.

Abbildung 6.3: Szene in der Eingangshalle des Flughafen Erfurt. Anfangs bewegten sich viele Personen vom Eingang zum Check-In. Nach dem Check-In sind viele gleich über den Fahrstuhl, oder die Treppe (welche sich neben dem Infoschalter befindet) zum Terminal gegangen. Kurze Zeit später haben sich viele Personen von der Gepäckausgabe zum Eingang bewegt. Nur sehr wenige sind zum Infoschalter, oder zur Cafeteria gegangen.

es mit fünf Laser-Range-Scannern möglich ist, mehr als 20 Personen in Echtzeit zu tracken. In den Aufnahmen vom Humboldt-Foyer wurden insgesamt 195 Personen erfasst, wovon sich 145 durch die Szene bewegt haben und 40 nur am Rand entlang gingen (von der Cafeteria zum Eingang). Davon wurden 187 Personen erfolgreich erkannt und getrackt, wobei 49 von denen mehr als zwei Meter (was in etwa zwei Schritten entspricht) gehen mussten, bis die Person detektiert wurde. 14 Personen wurden zwischendrin verloren und neun Personenhypothesen haben die Person gewechselt. Insgesamt kam es fünf mal vor, dass eine Person von mehreren Partikelfiltern getrackt wurde.

Wie es in Abschnitt 4.4.1 bereits erläutert wurde, wird eine Person detektiert, wenn die beiden Cluster der Füße zu einem Cluster zusammengefasst und kurz darauf wieder in zwei einzelne Cluster aufgetrennt werden. Dieser Ablauf ist zu beobachten, wenn das eine Bein am anderen Bein vorbeischwingt (vergleiche dazu Abbildung 4.3). Treten währenddessen Verdeckungen auf, kann es vorkommen, dass der Ablauf vom Tracker

nicht registriert wird. Dies ist auch der Grund für die 49 späten und die acht fehlenden Detektionen. Von den acht Personen, welche nicht detektiert wurden, bewegten sich sieben zwischen der Cafeteria und dem Eingang, traten also nur für kurze Zeit im Scan auf. Die wenigen Möglichkeiten, bei denen die Personen detektiert werden konnten, wurden dabei von Verdeckungen verhindert. Eine Person bewegte sich mit zwei weiteren sogar quer durch die Szene, ohne getrackt zu werden. Hierbei waren auch ungünstigen Verdeckungen zu den Zeitpunkten des Vorbeischwingens die Ursache für die fehlende Detektion dieser Person. Von den 49 verspätet detektierten Personen wurden die ersten Vorbeischwinger der Beine vom Tracker ebenfalls aufgrund von Verdeckungen nicht beobachtet, wodurch sie erst nach mehr als zwei Metern erkannt wurden.
Von den 14 Partikelfiltern, welche ihre Person verloren haben, wechselten neun die Person. Dieses Verhalten wird jedoch erst im nächsten Absatz genauer erklärt. Der Verlust der fünf anderen Personen ist einerseits auf Verdeckungen und andererseits auf schlecht geschätzte Bewegungsparameter zurückzuführen. Wird eine Person verdeckt, welche bereits von einem Partikelfilter getrackt wird, dann wird deren Position anhand der Bewegungsparameter geschätzt. Ist die Person wieder in den Scandaten zu sehen, so befindet sich die Positionshypothese bei gut geschätzten Bewegungsparametern nahe an der realen Personenposition, wodurch der Partikelfilter die Person wiederfinden kann und weiterverfolgt. Wenn die Bewegungsparameter allerdings nicht ausreichend gut geschätzt wurden, dann kann auch die Position der verdeckten Person nicht zuverlässig prognostiziert werden, wodurch die Positionshypothese zu weit von der realen Position der Person entfernt ist, wenn sie wieder in den Scandaten erscheint. Die Scanpunkte der wieder erschienenden Person haben somit, aufgrund der hohen Entfernung, keine Wirkung auf die Partikel des ursprünglichen Partikelfilters, wodurch der Filter die Person auch nicht wiederfindet. Sobald die Person dann eine neue Detektion auslöst, wird ein neuer Partikelfilter an ihrer Position initialisiert und trackt sie an Stelle des ursprünglichen Partikelfilters weiter.
Gehen mehrere Personen in einer Gruppe durch die Szene, so kann es vorkommen, dass ein Partikelfilter die Person wechselt, so wie es neun mal im Experiment vorkam. Die Ursache für dieses Verhalten ist der Interaktionsterm (vergleiche Abschnitt 4.5.1 und Gleichung 4.13), welcher eine abstoßende Wirkung zwischen zwei Partikelfiltern

verursacht. Gehen z. B. drei Personen nebeneinander durch die Szene und wurden nur die linke und die mittlere Person erkannt und von einem Partikelfilter getrackt, wohingegen die rechte Person weder detektiert wurde, noch getrackt wird, so wechselt der Partikelfilter der mittleren Person auf die rechte Person. Die Ursache dafür ist, dass der Partikelfilter der mittleren Person vom linken Partikelfilter abgestoßen wird und bei der rechten Person ebenfalls Scanpunkte findet, die getrackt werden können. Sobald die mittlere Person eine Detektion auslöst, wird sie von einem neuen Partikelfilter weitergetrackt.

Dass eine Person von mehreren Partikelfiltern getrackt wird, hat zwei verschiedene Ursachen, welche aber beide einen schlecht geschätzten Bewegungszustand gemeinsam haben. Dieser hat zur Folge, dass die beiden Fußhypothesen des Partikelfilters auf dem hinteren Bein der Person liegen. Eine Situation, welche es ermöglicht, dass ein zweiter Partikelfilter diese Person trackt, tritt auf, wenn sich der freie Fuß bei großen Schritten dicht an einem anderen Objekt vorbeibewegt. Dadurch kann fälschlicher Weise eine Personendetektion ausgelöst werden, welche in einem neuen Partikelfilter resultiert, der dann das zweite Bein trackt. Die andere Situation tritt auf, wenn sich die Person in einer Gruppe bewegt. Dabei kann sich der Partikelfilter einer anderen Person auf den freien Fuß setzen, wenn dieser aufgrund des Interaktionstermes (in Abschnitt 4.5.1 und Gleichung 4.13 wird dieser erklärt) von umliegenden Filtern abgestoßen wird.

In den Experimenten wurden auffallend oft die Bewegungsparameter der Personen falsch geschätzt. Bei genauerer Untersuchung dieses Umstandes hat sich herausgestellt, dass viele Partikelfilter noch kurz nach ihrer Initialisierung die Bewegungen der Füße gut getrackt haben und erst mit Auftreten einer zeitweisen Verdeckung eines Beines die Bewegungsparameter nicht mehr korrekt schätzen. Dies äußert sich dadurch, dass beide Fußhypothesen des Partikelfilters am hinteren Bein festhaften und bei jedem Vorbeischwingen des Fußes auf das andere, nun hinten stehende Bein, wechselt (siehe Abbildung 6.4). Die Ursache dafür ist, dass während der Verdeckung die Partikel, bei denen beide Fußhypothesen in der Nähe des noch sichtbaren Fußes liegen, ein höheres Gewicht erhalten, als alle anderen Partikel. Dadurch ziehen sich die Fußhypothesen auf den noch sichtbaren Fuß zusammen. Die Hypothesen verweilen auch noch auf dem einen Fuß, sobald das andere Bein nicht mehr verdeckt wird, da die Bewegungspara-

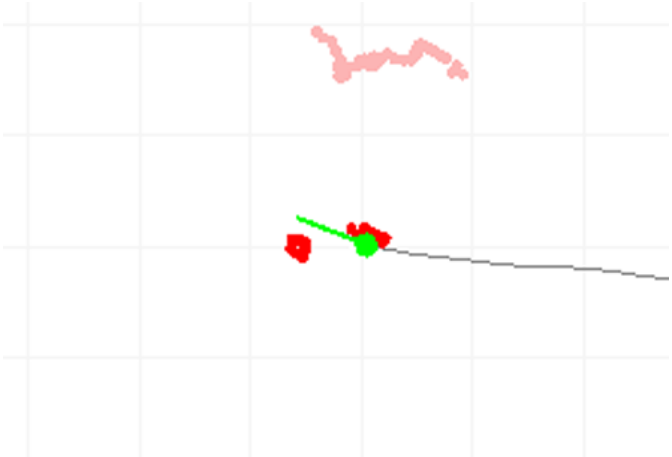


Abbildung 6.4: Fehlerhaftes Fußmodell. Die zu den Füßen gehörenden Scanpunkte sind rot und die Hintergrundpunkte sind blassrot eingetragen. Die Personenhypothese des Partikelfilters ist grün eingetragen, wobei der Kreis die Position und der vom Kreis ausgehende Strich die Bewegungsrichtung angibt. Die Bewegungsspur ist grau markiert. Es ist zu sehen, dass die Personenhypothese auf dem hinteren Bein liegt. Diese fehlerhafte Bewegungsverfolgung lag in diesem Fall die ganze Zeit vor, bis die Person die Szene verlassen hat.

meter mittlerweile gänzlich falsch geschätzt werden und der Partikelfilter auch nicht mehr durch das Verrauschen im Resamplingschritt in der Lage ist, den korrekten Bewegungszustand zu schätzen. Somit wird die Person dann allein aufgrund der Streuung der Position im Resamplingschritt notdürftig weitergetrackt. Allerdings wird dabei nur die Position und die Bewegungsrichtung der Person nutzbringend geschätzt. Es könnte ein Mechanismus umgesetzt werden, der solche Situationen detektiert und dann den Partikelfilter neu initialisiert, allerdings hat der Zeitrahmen der Masterarbeit nicht ausgereicht, um diesen Mechanismus noch zu implementieren.

In den Aufnahmen von Hannover und Erfurt hat sich zudem gezeigt, dass Gepäckwagen die Partikelfilter nachteilig beeinflussen und gelegentlich auch einer Person den Partikelfilter „wegnehmen“, woraufhin dann der Gepäckwagen vom ursprünglichen Partikelfilter getrackt wird. Da ein Gepäckwagen in den Scandaten meist mehr Scanpunkte aufweist, als die Füße der Personen, erhalten Partikel in der Nähe des Wagens alleine durch die hohe Anzahl der Scanpunkte ein höheres Gewicht, als die Partikel in Nä-

he zu den Füßen (vergleiche dazu Gleichung 4.7 und 4.8). Dadurch sammeln sich die Partikel nach kurzer Zeit beim Gepäckwagen und der beteiligte Partikelfilter trackt nur noch diesen. Sobald die Person jedoch eine neue Detektion auslöst, wird sie durch einen neuen Partikelfilter getrackt.
In den Erfurter Aufnahmen war ebenso zu erkennen, dass der Fußtracker Probleme mit still stehenden Personen hat. So war in den Fußhypothesen für die stehende Person immer eine langsame, gehende Bewegung zu erkennen. Der Grund dafür ist, dass es nicht möglich ist, mit dem Bewegungsmodell aus Gleichung 4.3 eine still stehende Person zu modellieren. Dazu müsste die Schrittdauer T unendlich groß sein, was mit einem Partikelfilter nicht praktikabel umsetzbar ist. Da bei einem Partikelfilter die Variabilität des geschätzten Bewegungszustandes hauptsächlich durch das Verrauschen des Zustandes während des Resamplingschritts gegeben ist, bräuchte es unendlich viele Zeitschritte, bis $T = \infty$ geschätzt wird. Eine mögliche Lösung wäre, das Bewegungsmodell aus Gleichung 4.3 derart umzustellen, dass der Bewegungszustand nicht mit Hilfe der Schrittdauer T dargestellt wird, sondern das Reziprok $\theta = \frac{1}{T}$ verwendet wird, was der Anzahl der Schritte pro Zeit entsprechen würde. Diser Wert würde bei Stillstand den Wert $\theta = 0$ annehmen, was mit einem Partikelfilter durchaus geschätzt werden kann. Diese Lösung konnte allerdings nicht mehr in dieser Masterarbeit umgesetzt werden, da die Ursache für das Problem und deren Lösung erst recht spät festgestellt wurde. Somit ließ sich die Umsetzung nicht mehr mit dem Zeitrahmen der Arbeit vereinbaren.

6.4 Ergebnisse des Oberkörpertrackings

Beim Oberkörpertracking ist, wie auch beim Fußtracker, in Ermangelung der Ground-Truth nur wenig Möglichkeit gegeben, eine quantitative Auswertung vorzunehmen. Die Experimente haben gezeigt, dass es mit fünf Laser-Range-Scannern möglich ist, mehr als hundert Personen in Echtzeit zu tracken. In den Aufnahmen vom Humboldt-Foyer wurden insgesamt 195 Personen erfasst, wovon sich 145 durch die Szene bewegt haben und 40 nur am Rand entlang gingen (von der Cafeteria zum Eingang). Davon wurden 192 Personen erfolgreich erkannt und getrackt, wobei zwölf von denen mehr als zwei Meter (was in etwa zwei Schritten entspricht) gehen mussten, bis die Person detektiert

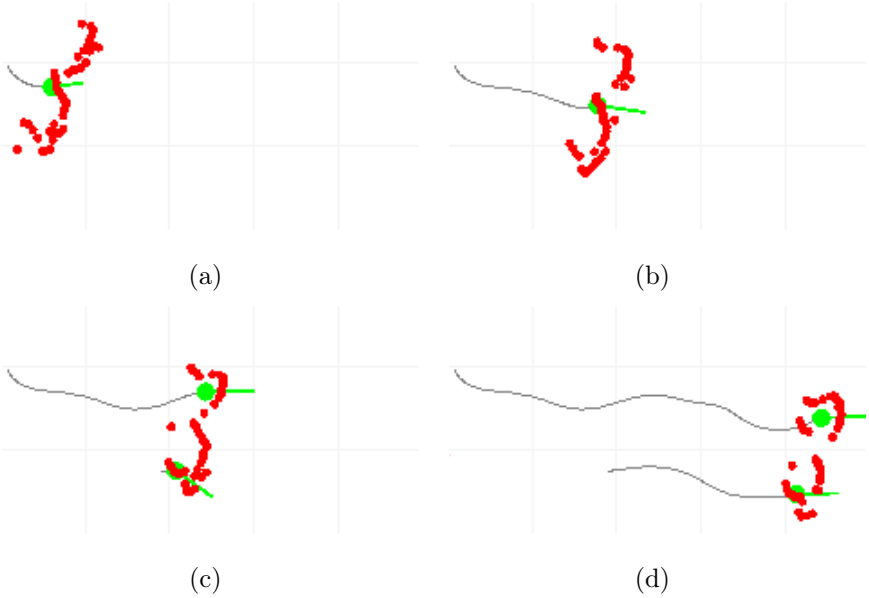


Abbildung 6.5: Verspätete Personendetektion. Die zu den Personen gehörenden Scanpunkte sind rot eingetragen. Die Personenhypothesen der Partikelfilter sind grün eingetragen, wobei der Kreis die Position und der vom Kreis ausgehende Strich die Bewegungsrichtung angibt. Die Bewegungsspuren sind grau markiert. In (a) betreten zwei Personen die Szene, wobei sie nur als eine Person erkannt wurden und dementsprechend auch nur ein Partikelfilter angelegt wurde. Dieser Partikelfilter trackt beide Personen bis (b), wo er sich in Richtung der oberen Punktemenge bewegt. In diesem Zeitraum wurde die untere Person schon mehrfach detektiert, allerdings wurde kein zweiter Partikelfilter angelegt, da die Personenhypothese des ersten Partikelfilters zu nahe an der Persondetektion liegt. In (c) ist der erste Partikelfilter weit genug von der unteren Person entfernt, so dass an ihrer Position ein neuer Partikelfilter angelegt werden kann, der diese Person dann trackt, wie es in (d) zu sehen ist.

wurde. Zehn Personen wurden zwischendrin verloren und neun Personenhypothesen haben die Person gewechselt. Insgesamt kam es drei mal vor, dass eine Person von mehreren Partikelfiltern getrackt wurde.

Die drei nicht detektierten Personen haben sich alle zwischen der Cafeteria und dem Eingang am Rand der beobachteten Szene bewegt und wurden jeweils von einer anderen Person teilweise bis vollständig verdeckt. Sie wurden zusätzlich auch nur von einer

Seite erfasst, wodurch die jeweiligen Punktecluster der drei Personen nicht die zur Detektion nötige Größe erreicht haben (vergleiche dazu Abschnitt 4.4.2). Ebenfalls der Verdeckung geschuldet sind ungefähr die Hälfte der zwolf verspäteten Detektionen. Die beteiligten Personen bewegten sich in Gruppen durch die Szene und wurden anfangs durch die Oberkörper der anderen Personen fast vollständig verdeckt. Erst nach einiger Zeit wurde ein ausreichend großer Teil der Personen von den Laser-Range-Scannern erfasst, um eine Detektion auszulösen. Die andere Hälfte der verspäteten Detektionen trat bei eng zusammen gehenden Personen auf. Dabei wurde zuerst eine der Personen detektiert und ein Partikelfilter an der entsprechenden Stelle initialisiert. Dieser schätzte allerdings innerhalb weniger Zeitschritte die Position zwischen den beiden Personen, da dort die Partikel des Filters aufgrund vieler eng beieinander liegender Scanpunkte ein hohes Gewicht besaßen. Bei einer Detektion der zweiten Person wurde die Initialisierung eines zweiten Partikelfilters durch die Nähe der Personenhypothese des Partikelfilters verhindert. Erst als die Scanpunkte einer der beiden Personen den ersten Partikelfilter mehr anzogen, als die Scanpunkte der anderen Person (z. B. aufgrund einer höheren Anzahl von Scanpunkten, oder weil sich der Abstand zwischen den Personen vergrößerte), wurde ein zweiter Partikelfilter initialisiert, wodurch dann beide Personen getrackt wurden. In Abbildung 6.5 wird dieser Ablauf noch einmal verdeutlicht.

In neun der zehn Fälle, in denen der Partikelfilter seine Person verloren hat, wurde die getrackte Person vom Partikelfilter gewechselt. Dieser Fall wird allerdings erst im nächsten Absatz behandelt. Der Verlust der zehnten Person kam durch ein ungünstiges Zusammenspiel von Verdeckung und Fehlklassifikation der Scanpunkte zustande. Gerade als die Person für zwei Sekunden durch eine Verdeckung hinter einer Säule ging, traten am Säulenrand fehlerhafte Messungen auf, welche als Vordergrundpunkte klassifiziert wurden. Solche Messfehler treten auf, wenn Objektkanten während des Messvorgangs nur zum Teil von einem Laserstrahl erfasst werden. Befindet sich hinter der Objektkante in nächster Nähe (bis etwa zwei Meter) ein anderes Objekt, so wird vom Laser-Range-Scanner ein Entfernungswert gemessen, der zwischen den Entfernungen der beiden Objekten liegt. Dies resultiert in Messpunkten, die im globalen Koordinatensystem dann zwischen beiden Objekten liegen (siehe Abbildung 6.6). Solche

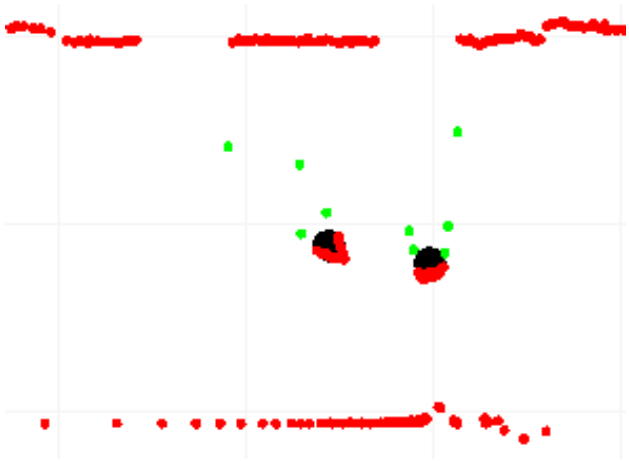
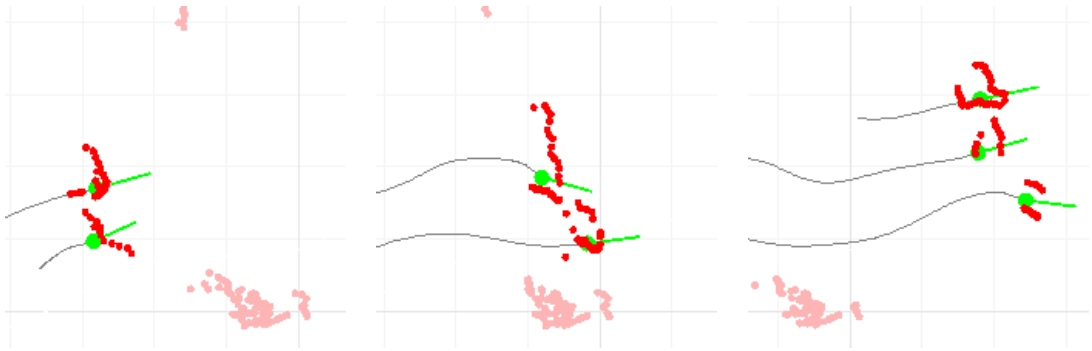


Abbildung 6.6: Fehlerhafte Scandaten an Objektkanten in der Draufsicht. In dieser Szene befindet sich oben und unten jeweils eine Wand und mittig zwei schwarz eingezeichnete Beine. Die korrekten Scanpunkte sind rot eingefärbt. Die an den Beinrändern entlangstreifenden Messstrahlen liefern fehlerhafte Scanpunkte (grün markiert), welche zwischen den Beinen und der oberen Wand liegen. An diesen Stellen befinden sich allerdings keine Objekte, sondern Freiraum.

Punkte haben den Partikelfilter der Person derart abgelenkt, dass die Personenhypothese des Filters zu weit von der wieder erscheinenden Person weg war, um sie dann weitertracken zu können.

Ursache für die neun Wechsel der getrackten Personen waren Verdeckungen in Personengruppen. Bei sieben der Wechsel wurde eine Person in der Gruppe noch nicht detektiert und somit auch noch nicht von einem Partikelfilter getrackt. Wird in solch einer Situation eine der anderen Personen, welche schon von einem Partikelfilter getrackt wird, verdeckt, wechselt dessen Filter auf die noch nicht detektierte, aber schon in den Scanpunkten sichtbare Person und trackt diese weiter, da die Partikel in der Nähe dieser Scanpunkte ein höheres Gewicht erhalten, als die Partikel, welche nun keine Scanpunkte in ihrer Nähe haben. In Abbildung 6.7 ist eine ähnliche Situation dargestellt. In einem einzelnen Fall haben zwei Partikelfilter die getrackte Person getauscht. Dies geschah, als in einer Dreiergruppe, in welcher alle Personen von einem Partikelfilter getrackt wurden, die mittlere Person verdeckt wurde (zum besseren Verständnis wird der Partikelfilter der rechten Person mit R bezeichnet, der Filter der



(a) (b) (c)

Abbildung 6.7: Wechsel der getrackten Person. Die zu den Personen gehörenden Scanpunkte sind rot und die Hintergrundpunkte sind blassrot eingetragen. Die Personenhypothesen der Partikelfilter sind grün eingetragen, wobei der Kreis die Position und der vom Kreis ausgehende Strich die Bewegungsrichtung angibt. Die Bewegungsspuren sind grau markiert. In (a) haben drei Personen die Szene betreten, wovon zwei erkannt und bereits getrackt werden. In (b) erscheint die mittlere der drei Personen in den Scandaten und wird vom Partikelfilter, welcher zuvor die linke Person getrackt hat, übernommen, da die mittlere Person durch mehr Scanpunkte repräsentiert wird, als die linke Person. In (c) hat ein neuer Partikelfilter die in (b) verlorene linke Person weiter getrackt.

mittleren Person mit M und der Filter der linken Person mit L). Der Partikelfilter M der mittleren Person hat daraufhin den Rand der rechten Person weitergetrackt. Da bei dem Wechsel die geschätzte Geschwindigkeit etwas angestiegen ist und der Partikelfilter M somit gewissermaßen ein Moment besaß, wanderte die Personenhypothese zum vorderen rechten Rand der rechten Person und hat den ursprünglichen Filter R nach hinten verdrängt. Nachdem die mittlere Person wieder in den Scandaten erschien, lag die Personenhypothese des Partikelfilters R näher an der mittleren Person, als die Hypothese von Filter M, wodurch der Filter R auf die mittlere Person wechselte und diese weitergetrackt hat. Abbildung 6.8 zeigt diesen Ablauf noch einmal bildlich.

Ist eine Person besonders dick, oder trägt sie eine große Tasche über einer Schulter, kann es vorkommen, dass sie zwei Personenhypothesen auslöst, da deren Scanpunkte aufgrund der Kernelbreite in zwei Cluster gefasst werden (siehe Abbildung 3.8). Die

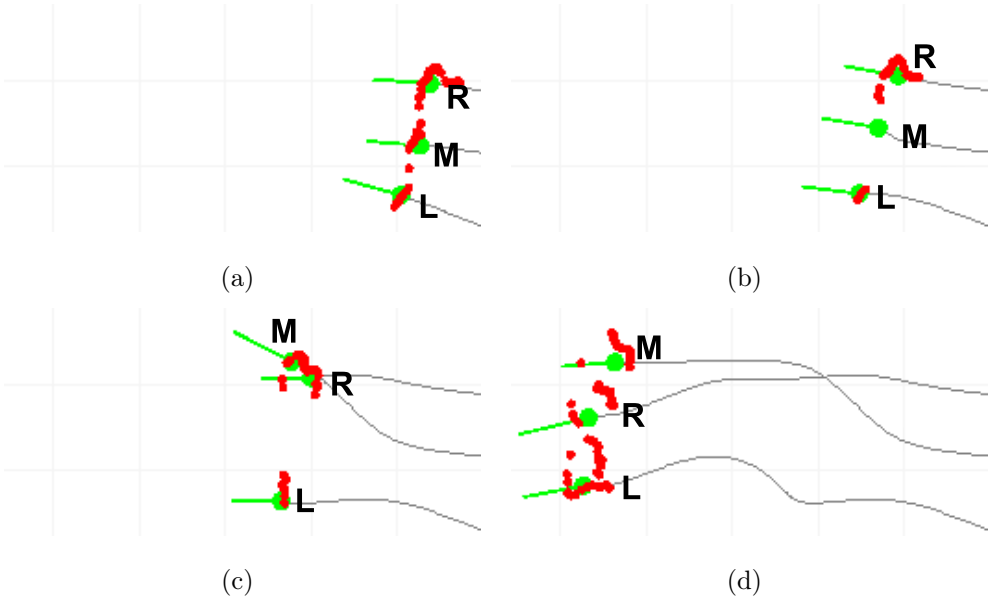
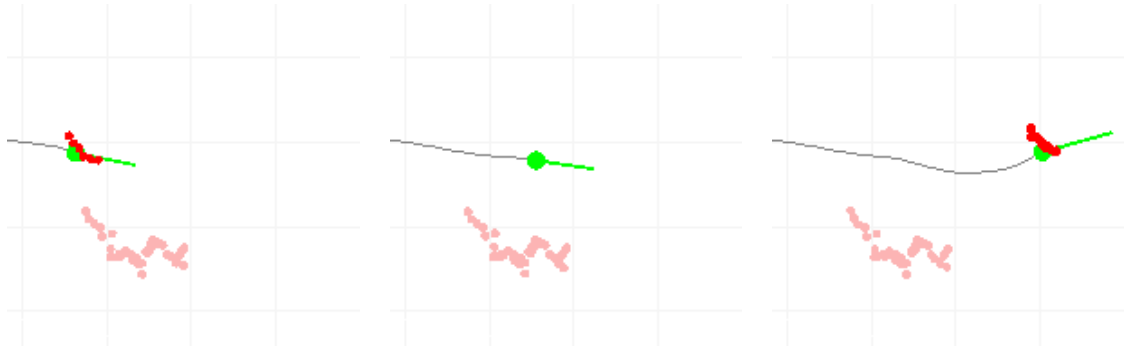


Abbildung 6.8: Wechsel der getrackten Personen. Die zu den Personen gehörenden Scanpunkte sind rot eingetragen. Die Personenhypothesen der Partikelfilter sind grün eingetragen, wobei der Kreis die Position und der vom Kreis ausgehende Strich die Bewegungsrichtung angibt. Die Bewegungsspuren sind grau markiert. In (a) haben drei Personen die Szene betreten und werden von den Partikelfiltern L, M und R getrackt. In (b) wird die mittlere Person von den beiden äußeren Personen in den Scandaten verdeckt. Dadurch, dass der Partikelfilter R am rechten Rand der rechten Person liegt, werden einige Scanpunkte dieser Person dem Filter M zugeordnet, der daraufhin seine Personenhypothese in diese Richtung verschiebt. Da Partikelfilter prinzipbedingt eine gewisse Trägheit besitzen, bewegt sich der Filter M bis zum vorderen rechten Rand der rechten Person, so wie es in (c) zu sehen ist. Dadurch wird der Partikelfilter R an den hinteren linken Rand der Person verdrängt. In (d) ist die mittlere Person nicht mehr verdeckt und wird vom Filter M weitergetrackt, da er näher an den Scanpunkten liegt, als der Partikelfilter M. Damit haben R und M die getrackte Person getauscht.

zwei daraufhin initialisierten Partikelfilter tracken dann diese eine Person, wie es auch in einer Situation der Fall war. In den anderen beiden Fällen, wo eine Person von zwei Partikelfiltern getrackt wurde, war jeweils ein anderer Partikelfilter beteiligt, welcher kurz zuvor seine Person verloren hat. Dessen Personenhypothese kam, aufgrund des



(a) (b) (c)

Abbildung 6.9: Bewegungsverfolgung trotz Verdeckung. Die zu der Person gehörenden Scanpunkte sind rot und die Hintergrundpunkte sind blassrot eingetragen. Die Personenhypothese des Partikelfilters ist grün eingetragen, wobei der Kreis die Position und der vom Kreis ausgehende Strich die Bewegungsrichtung angibt. Die Bewegungsspur ist grau markiert. In (a) bewegt sich eine bereits getrackte Person auf eine Region zu, welche nicht von Laserscannern erfasst werden kann. In (b) ist die Person vollständig vom Objekt, was unten im Bild als Hintergrund markiert zu sehen ist, verdeckt. Trotzdem liefert der Partikelfilter brauchbare Personenhypothesen. In (c) verlässt die Person die Verdeckung und wird vom Partikelfilter wiedergefunden.

geschätzten Bewegungsparameters, so nahe an die Scanpunkte der anderen Person heran, dass diese Scanpunkte dem fremden Partikelfilter zugeordnet wurden. Somit konnte dieser Filter dann auch die andere Person mittracken.

Bei der qualitativen Betrachtung hat sich herausgestellt, dass das vergleichsweise einfache Bewegungsmodell durchaus in der Lage ist, zuverlässig Bewegungen von Personen zu modellieren. Dabei sollte jedoch darauf geachtet werden, dass die Laser-Range-Scanner derart angeordnet sind, dass Verdeckungen möglichst vermieden werden. Wenn eine Person, welche bereits von einem Partikelfilter verfolgt wird, trotzdem einen verdeckten Bereich passiert (beispielsweise hinter einer breiten Säule), so wird sie beim Verlassen der Verdeckung, bis auf wenige Ausnahmen, welche in den vorherigen Absätzen erläutert wurden, jedoch zuverlässig wiedererkannt und weiter verfolgt, sofern sie während der Verdeckung nicht gänzlich die Bewegung ändert (siehe Abbildung 6.9). Da oftmals die Oberkörper nicht von allen Seiten erfasst werden, liegen deren

Tabelle 6.1: Übersicht über die Trackingergebnisse des Fußtrackers (FT) und des Oberkörpertrackers (OT) im Humboldt-Foyer der TU-Ilmenau.

	Personen	Erkannt	nach 2m	Verluste	Wechsel	mehrfach
FT	195	187	49	14	9	5
OT	195	192	12	10	9	3

Personenhypothesen nicht immer in der Mitte, sondern verschieben sich an den vom Scanner sichtbaren Personenrand. Einzeln gehende Personen werden, sofern sie nicht verdeckt werden, von dem ersten Zeitschritt an erfolgreich detektiert und zuverlässig getrackt. Auch schnelle Änderungen in der Bewegung können vom Oberkörpertracker erkannt und verfolgt werden, ohne dabei die Person zu verlieren.

6.5 Vergleich beider Trackingverfahren

Nachdem in den vorherigen Abschnitten beide Trackingverfahren im Einzelnen ausgewertet wurden, folgt in diesem Abschnitt ein Vergleich des Fußtrackers mit dem Oberkörpertracker. Die quantitativen Ergebnisse beider Verfahren sind noch einmal in Tabelle 6.1 zusammengefasst.

Bezüglich der Detektionsleistung ist der Oberkörpertracker dem Fußtracker überlegen. Beim Fußtracker können Personen nur im Moment des Vorbeischwingens des einen Beins am Anderen detektiert werden, wohingegen auf Oberkörperhöhe zu jeder Zeit eine Detektion möglich ist. Somit reicht es beim Oberkörpertracker aus, dass in nur einem einzelnen Scan ein ausreichend großer Teil der Person erfasst wird, um einen Partikelfilter zu erzeugen, der dann diese Person trackt (vergleiche dazu Abschnitt 4.4.2). Beim Fußtracker muss hingegen eine bestimmte Sequenz ohne Unterbrechung gut sichtbar sein, damit die Person getrackt wird (diese Sequenz wird in Abschnitt 4.4.1 genauer erklärt), was bei ungünstigen Verdeckungen erst sehr spät der Fall sein kann. In Tabelle 6.1 ist dies auch erkennbar, da der Fußtracker gegenüber dem Oberkörpertracker etwa vier mal mehr Personen erst nach zwei Metern oder mehr in den Scandaten detektiert und verfolgt hat.

Bei der Trackinggüte erzielt der Oberkörpertracker ebenfalls bessere Ergebnisse, als der Fußtracker. Auch wenn die Personenhypothese des Oberkörpertrackers meist nicht genau mittig vom Oberkörper liegt, wird die Position der Person fast immer innerhalb der Punktemenge des Oberkörpers vermutet. Auch werden die anderen Bewegungsparameter, wie die Geschwindigkeit und Bewegungsrichtung, gut geschätzt. Der Fußtracker hingegen prognostiziert die Bewegungsparameter nur richtig, wenn das Fußmodell nicht versagt, wie es in Abschnitt 6.3 und in Abbildung 6.4 erläutert wird. Läuft eine einzelne Person unverdeckt durch die Szene, so ermittelt der Oberkörpertracker die Bewegungsparameter dem entsprechend sehr gut. Beim Fußtracker hingegen werden, bis auf die Bewegungsrichtung und Position, die Bewegungsparameter nur selten richtig geschätzt. Es ist zu vermuten, dass das Bewegungsmodell für die Füße aus Abschnitt 4.5.1 aufgrund seiner Komplexität anfälliger für Störungen ist, als das einfach gehaltene Bewegungsmodell der Oberkörper aus Abschnitt 4.5.2 und die Bewegungsparameter somit öfter falsch ermittelt werden. Der Fußtracker muss insgesamt 4 Parameter (Bewegungsrichtung, Schrittweite, Phase und Schrittdauer) korrekt schätzen, damit die Bewegungen der Person verfolgt werden können. Wenn schon ein Parameter falsch geschätzt wird, wird die Fußbewegung nicht mehr zuverlässig modelliert und die Person wird einzig durch das Streuen der Position im Resamplingschritt (siehe Abschnitt 3.7, Gleichung 4.4 und 3.21) getrackt. Der Oberkörpertracker muss nur zwei Parameter (Geschwindigkeit und Bewegungsrichtung) zuverlässig schätzen, wodurch er robuster gegen Störungen ist, als der Fußtracker. Wenn eine Person ihre Geschwindigkeit verändert, so tut sie dies, indem sie die Schrittweite und die benötigte Zeit für einen Schritt gleichzeitig verändert. Somit muss der Fußtracker also zwei Parameter anpassen, wenn eine Person z. B. langsamer wird, wobei der Oberkörpertracker lediglich einen Parameter anpassen muss und dadurch zuverlässiger und schneller auf die Veränderung reagieren kann.

Bei einem hohen Personenaufkommen hat der Fußtracker gegenüber dem Oberkörpertracker den Vorteil, dass in den Scandaten ein geringerer Grad an Verdeckungen auftritt, da eine Person auf Oberkörperhöhe mehr Platz einnimmt, als auf Fußhöhe. Dadurch erscheinen auch bei dicht gedrängten und großen Personengruppen die Füße der Personen im Inneren der Gruppe zumindest zeitweise in den Scandaten, womit eine

gewisse Chance besteht, dass diese Personen auch detektiert werden. Auf Oberkörperhöhe ist es allerdings wahrscheinlicher, dass diese Personen nur zu kleinen Teilen in den Scandaten vorkommen, wodurch deren Detektion verhindert wird. Dieser Fall trat allerdings nicht in der Aufnahmen auf und ist nur in Extremsituationen zu erwarten.
Im Vergleich der Rechenperformanz beider Verfahren hat sich herausgestellt, dass der Oberkörpertracker wesentlich mehr Bewegungsspuren in Echtzeit ermitteln kann, als der Fußtracker. Der Großteil der Berechnungen wird durch das Gewichtsupdate in Gleichung 4.5 bis 4.13 für den Fußtracker, bzw. Gleichung 4.15 für den Oberkörpertracker eingenommen, da diese Gleichungen für jedes Partikel von jedem Partikelfiter berechnet werden müssen. Es ist ersichtlich, dass bei dem Fußtracker wesentlich mehr Operationen pro Gewichtsupdate durchgeführt werden müssen, als beim Oberkörpertracker. Hinzu kommt, dass beim Fußtracker alle Scanpunkte in allen Gewichtsupdates berücksichtigt werden, also jeder Scanpunkt Einfluss auf jedes Partikelgewicht hat. Bei dem Oberkörpertracker hat allerdings fast jeder Scanpunkt nur auf die Partikel eines einzelnen Partikelfiters Einfluss (siehe dazu den letzten Absatz aus Abschnitt 4.5.2 und Gleichung 4.16 bis 4.18) und nur wenige, vereinzelt liegende Scanpunkte haben einen Einfluss auf die Partikelgewichte aller Partikelfilter. Somit können in gleicher Zeit die Partikelgewichte von mehr Partikelfiltern auf Oberkörperhöhe berechnet werden, als auf Fußhöhe.
Zusammenfassend lässt sich sagen, dass die Defizite des Oberkörpertrackers hauptsächlich durch Verdeckungen zustande kommen, wohingegen die Defizite des Fußtrackers sowohl durch Verdeckungen, als auch durch falsch geschätzte Bewegungszustände verursacht werden.

6.6 Auswertung der Remittanzdaten

Da die verwendeten Laser-Range-Scanner in der Lage sind, neben den Entfernungen auch Remittanzdaten zu erfassen, liegt die Idee nahe, diese Daten hinsichtlich einer möglichen Unterscheidung von Personen und Objekten zu untersuchen. So ist zu vermuten, dass ein dunkles Hosenbein eine andere Remittanz aufweist, als ein Helles. In kurzen Distanzen (bis ungefähr 3m) ist dies auch möglich, allerdings fangen bei größe-

ren Distanzen die Remittanzdaten an, stark zu streuen. Um dieses Verhalten erklären zu können, bedarf es einer genaueren Betrachtung des Laserscanners. Dieser sendet keinen Laserstrahl aus, wie es von Laserpointern bekannt ist, sondern eine Laserkeule, die sich mit steigender Entfernung aufweitet. Der Keulendurchmesser d kann für eine bestimmte Entfernung e mit $d = e \cdot 0{,}015 + 8\text{mm}$ berechnet werden. So hat die Keule auf eine Entfernung von 10m schon einen Durchmesser von 15,8cm. Sofern das Bein nicht mittig im Strahl liegt, wird es nicht komplett von der Keule erfasst und wirft dadurch weniger Licht zurück. Erschwerend kommt noch hinzu, dass Hosenbeine nicht zylinderförmig und glatt in Erscheinung treten, sondern in fast allen Fällen Falten werfen. Da die Remittanz eines Objektes auch vom Einfallswinkel des Lichts abhängt, kann ein und das selbe Hosenbein unterschiedliche Remittanzwerte bei gleicher Entfernung liefern. Somit lässt sich eine Verbesserung der Trackinggüte anhand von Vergleichen der Remittanzdaten nicht umsetzen. Allerdings hat sich gezeigt, dass Reflektionsstreifen, welche z. B. auf Warnwesten angebracht sind, in den Remittanzwerten erkennbar sind. Durch Anbringung solcher Streifen an die Gehäuse der Laserscanner ist es möglich, diese in den Scandaten anderer Laser-Range-Scanner zu erkennen und somit die räumliche Kalibrierung zu vereinfachen.

6.7 Zusammenfassung

Sowohl das Trackingverfahren auf Fußhöhe, als auch das Verfahren auf Oberkörperhöhe ist in der Lage, zuverlässig die Bewegungsspuren mehrerer Personen auf Basis von Laserscans zu ermitteln. Dabei ist der Oberkörpertracker dem Fußtracker in Hinblick auf die Detektionsleistung überlegen. Dadurch, dass nicht erst eine Sequenz von bestimmten Ereignissen beobachtet werden muss, um eine Detektion auszulösen, ist der Oberkörpertracker in der Lage, wesentlich früher, als der Fußtracker Personen zu verfolgen. Durch das einfachere Bewegungsmodell verliert der Oberkörpertracker auch weniger Personen, da er schneller auf Änderungen in der Bewegung reagieren kann und die Bewegungsparameter auch zuverlässiger, als der Fußtracker schätzt, wodurch ein besseres Wiederfinden von Personen nach Verdeckungen ermöglicht wird. Auch ist der Oberkörpertracker in Hinblick auf die Trackinggüte besser, als der Fußtracker, da er die

Bewegungsparameter wesentlich besser schätzt. Ebenso ist der Berechnungsaufwand bei dem Oberkörpertracker geringer, wodurch die Bewegungsspuren von wesentlich mehr Personen in Echtzeit ermittelt werden können. Allerdings ist die Installation der Laserscanner auf Oberkörperhöhe umständlicher. Trotzdem die Scanner in den Experimenten kein Verlegen von Kabeln (vergleiche Abbildung 4.1) erforderten, müssen sie für den Oberkörpertracker in entsprechender Höhe aufgestellt werden, wohingegen es für den Fußtracker reicht, sie ohne Weiteres auf den Fußboden zu stellen. Die Remittanzdaten haben keine Verbesserung der Trackinggüte ermöglicht, da sie in ihren Werten zu sehr streuen. Allerdings könnte mit der Hilfe der Remittanzen die Detektion von anderen Laser-Range-Scannern in den Scandaten umgesetzt werden, um so die räumliche Kalibrierung zu automatisieren.

Kapitel 7

Zusammenfassung und Ausblick

7.1 Zusammenfassung der Arbeit

Diese Arbeit hat gezeigt, dass es mit heutiger Rechentechnik möglich ist, in einem großen Areal eine Vielzahl von Personen mit Laser-Range-Scannern zu erfassen und in Echtzeit deren Bewegungsspuren zu extrahieren. Es wurde das Tracking auf Fußhöhe nach [Shao et al. 2007] erfolgreich umgesetzt und bestätigt, dass dieses Verfahren zuverlässig Personen tracken kann. Allerdings versagt das Fußmodell dieses Trackers zeitweise, wodurch die Bewegungsparameter, mit Ausnahme der Position, falsch geschätzt werden. Zudem wurde ein selbst entwickeltes Trackingverfahren auf Oberkörperhöhe umgesetzt, welches dem Fußtracker in Trackinggüte und Rechenperformanz überlegen ist und zuverlässig die Position, Geschwindigkeit und Bewegungsrichtung von Personen ermitteln kann. Bei einer sehr hohen Anzahl von Personen treten in den Scandaten auf Oberkörperhöhe jedoch mehr Verdeckungen auf, als auf Fußhöhe.

Es kann festgehalten werden, dass das Tracking auf Oberkörperhöhe zu bevorzugen ist, sofern nicht mit extremen Menschenmengen in der Szene gerechnet wird und die Aufstellung der Laserscanner in entsprechender Höhe kein Hindernis darstellt. Andernfalls sollte das Tracking auf Fußhöhe verwendet werden, wenn die Schätzung der Geschwindigkeit und Bewegungsrichtung gegenüber der Erkennung möglichst aller Personen eine untergeordnete Rolle spielt und es auch toleriert werden kann, wenn Personen zwischendurch verloren gehen.

Die Remittanzdaten der Laserscanner lassen sich nicht zur Verbesserung der Trackinggüte verwenden. Allerdings ist es mit diesen Daten möglich, andere Laser-Range-Scanner, welche mit speziellen Reflektorstreifen versehen sind, in den Scans zu erkennen, wodurch eine räumliche Kalibrierung der Scanner vereinfacht werden kann.
Es war sogar schon vor Abschluss der Arbeit möglich, die Bewegungsspuren in Testaufnahmen auf dem Flughafen Erfurt zu ermitteln und somit eine nutzbringende Wissensgrundlage für das Personentracking in Kameradaten, welche parallel zu den Laseraufnahmen erstellt wurden, zu bilden. Zum Zeitpunkt der Abgabe dieser Masterarbeit stellt sowohl das umgesetzte Trackingverfahren auf Fußhöhe, als auch das neu entwickelte Trackingverfahren auf Oberkörperhöhe vermutlich den besten Mehrpersonentracker auf Basis stationärer Laser-Range-Scanner am Fachgebiet Neuroinformatik und Kognitive Robotik der TU-Ilmenau dar. Allerdings muss diese Vermutung noch durch vergleichende Experimente belegt werden.

7.2 Weiterführende Arbeiten

Durch ein Versehen wurde während den Aufnahmen in Erfurt zeitweise der Oberkörpertracker anstatt des Fußtrackers auf den Fußscans angewendet. Unerwarteter Weise lieferte der Oberkörpertracker jedoch gute Trackingergebnisse. Dieser Umstand stellt eine Motivation dar, den Oberkörpertracker auf die Fußscans anzupassen, um somit den Vorteil der besseren Erfassung aller Personen mit der guten Trackingperformance des Oberkörpertrackers zu kombinieren.
Bisher wurden die Positionen und Ausrichtungen der Laser-Range-Scanner manuell bestimmt. Da es möglich ist, auf Grundlage der Remittanzdaten andere Laser-Range-Scanner zu erkennen, wäre es nutzbringend, eine automatische räumliche Kalibrierung der Laserscanner umzusetzen.
Ein logischer Schritt wäre auch, die während der Masterarbeit schon ermittelten Bewegungsspuren von Personen nicht nur zur Entwicklung von Trackingalgorithmen auf Kameradaten zu verwenden, sondern diese Spuren schon zu analysieren und für das APFel-Projekt des Fachgebiets Neuroinformatik und Kognitive Robotik der TU-Ilmenau Verfahren zur Trajektorienauswertung zu entwickeln.

Literaturverzeichnis

[Carballo et al. 2008] Carballo, A., A. Ohya und S. Yuta (2008). *Fusion of Double Layered Multiple Laser Range Finders for People Detection from a Mobile Robot*. In: *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*.

[Chakravarty und Jarvis 2006] Chakravarty, Punarjay und R. Jarvis (2006). *Panoramic Vision and Laser Range Finder Fusion for Multiple Person Tracking.*. In: *IROS*, S. 2949–2954. IEEE.

[Fukunaga und Hostetler 1975] Fukunaga, K. und L. Hostetler (1975). *The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition*. IEEE Transactions on Information Theory, 21(1):32–40.

[Lee et al. 2006] Lee, Jae Hoon, T. Tsubouchi, K. Yamamoto und S. Egawa (2006). *People Tracking Using a Robot in Motion with Laser Range Finder*. In: *IROS*, S. 2936–2942.

[Luo et al. 2007] Luo, R. C., Y. J. Chen, C. T. Liao und A. C. Tsai (2007). *Mobile Robot Based Human Detection andTracking Using Range and Intensity Data Fusion*. In: *IEEE Workshop on Advanced Robotics and its Social Impacts*.

[Mason und Handscomb 2003] Mason, John C. und D. Handscomb (2003). *Chebyshev Polynomials*. Chapman & Hall/CRC.

[Melo et al. 2006] Melo, J., A. Naftel, A. Bernardino und J. Santos-Victor (2006). *Detection and Classification of Highway Lanes Using Vehicle Motion Trajectories*. IEEE Trans. Intelligent Transportation Systems, 7(2):188–200.

[Naftel und Khalid 2006] Naftel, Andrew und S. Khalid (2006). *Classifying spatiotemporal object trajectories using unsupervised learning in the coefficient feature space*. Multimedia Systems, 12(3):227–238.

[Pearson 1901] Pearson, Karl (1901). *On lines and planes of closest fit to systems of points in space*. Philosophical Magazine, 2(6):559–572.

[Schulz et al. 2001] Schulz, Dirk, W. Burgard, D. Fox und A. B. Cremers (2001). *Tracking Multiple Moving Targets with a Mobile Robot using Particle Filters and Statistical Data Association*. In: *ICRA*, S. 1665–1670.

[Schulz et al. 2003] Schulz, Dirk, W. Burgard, D. Fox und A. B. Cremers (2003). *People Tracking with Mobile Robots Using Sample-based Joint Probabilistic Data Association Filters*. I. J. Robotic Res., 22(2):99–116.

[Shao et al. 2007] Shao, Xiaowei, H. Zhao, K. Nakamura, K. Katabira, R. Shibasaki und Y. Nakagawa (2007). *Detection and tracking of multiple pedestrians by using laser range scanners*. In: *IROS*, S. 2174–2179.

[Thrun et al. 2005] Thrun, Sebastian, W. Burgard und D. Fox (2005). *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press.

[Topp und Christensen 2005] Topp, E. A. und H. I. Christensen (2005). *Tracking for Following and Passing Persons*. In: *Intl Conf. on Intelligent Robotics and Systems (IROS)*, S. 70–77, Edmundton, Canada.