

Vergleich von 2d- und 3d-Bewegungsmodellen für die Vorhersage der Bewegung eines Sportlers beim Hochgeschwindigkeitstracking

Diplomarbeit

zur Erlangung des Grades eines/r Diplom-Informatikers / Diplom-Informatikerin im Studiengang Computervisualistik

vorgelegt von

Jochen Handels

Betreuer: Prof. Dr.-Ing. Dietrich Paulus, Institut für Computervisualistik,
Fachbereich Informatik
Erstgutachter: Prof. Dr.-Ing. Dietrich Paulus, Institut für Computervisualistik,
Fachbereich Informatik
Zweitgutachter: Dipl. Inf. Peter Decker, Institut für Computervisualistik, Fachbe-
reich Informatik

Koblenz, im Juni 2008

Erklärung

Ich versichere, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen wurde. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

Die Richtlinien der Arbeitsgruppe für Studien- und Diplomarbeiten habe ich gelesen und anerkannt, insbesondere die Regelung des Nutzungsrechts.

Mit der Einstellung dieser Arbeit in die Bibliothek bin ich einverstanden. ja nein

Der Veröffentlichung dieser Arbeit im Internet stimme ich zu. ja nein

Koblenz, den

Unterschrift

Überblick

Diese Arbeit behandelt einen Vergleich verschiedener Algorithmen zur Vorhersage der Bewegung einer Person bei der Ausführung einer sportlichen Aktivität. Als Grundlage für die Vorhersage dienen Bildströme, welche mittels zweier Hochgeschwindigkeitskameras aufgezeichnet wurden. Im Laufe der Arbeit werden Vor- und Nachteile der umgesetzten Ansätze theoretisch erläutert und anschliessend an einer Reihe von Messergebnissen nachgewiesen. Für die Messungen wurde eine Anwendung eingesetzt, welche ebenfalls im Rahmen der Arbeit entwickelt wurde. Neben realen Aufnahmen, wurden zusätzlich synthetische Bildfolgen betrachtet, um Erkenntnisse über das Verhalten der betrachteten Algorithmen unter optimalen Bedingungen zu erlangen.

Inhaltsverzeichnis

1	Einleitung	9
2	Stand der Wissenschaft	11
2.1	Überblick	12
2.2	Dynamische Systeme	12
2.2.1	Freie mechanische Systeme	13
2.2.2	Gebundene mechanische Systeme	15
2.2.3	Stochastische Verfahren	17
2.3	Eigenschaften der verfügbaren Koordinatensysteme	19
2.4	Menschmodelle	20
2.4.1	Skelettmodelle	22
2.4.2	Hautmodelle	25
2.4.3	2-D Menschmodelle	27
2.5	Muskelmodelle	27
2.6	Motorische Kontrolle	29
2.7	Methoden zur Ermittlung des Systemzustands	32
2.7.1	Indirekte Verwendung des Modellzustands	32
2.7.2	Direkte Verwendung des Modellzustands	33

2.8	Konkrete Methoden zur Bewegungsvorhersage	34
2.8.1	Beispiele expliziter Verfahren	34
2.8.2	Beispiele impliziter Verfahren	36
3	Vergleich von 2d- und 3d-Verfahren	37
3.1	Überblick	38
3.2	Verallgemeinerung des dynamischen Modells	38
3.3	2d vs. 3d	39
3.4	Vorhersagegenauigkeit	40
3.5	Die entwickelte Anwendung	41
3.5.1	Die initiale Ermittlung der Kameraparameter	41
3.5.2	Merkmalsdetektion	44
3.5.3	Korrespondenzsuche	44
3.5.4	Triangulierung	45
3.5.5	Vorhersage mittels polynomieller Extrapolation	46
3.5.6	Unterstützte Ein- und Ausgangsformate	50
3.5.7	Abhängigkeiten	50
3.5.8	Verifikation des Trackers	51
4	Evaluierung	53
4.1	Überblick	53
4.1.1	Synthetische Bildfolgen	53
4.1.2	Reale Bildfolgen	55
4.2	Validierung der Ergebnisse	58
4.2.1	Validierung der synthetischen Bildfolgen	58
4.2.2	Validierung der realen Aufnahmen	68

<i>INHALTSVERZEICHNIS</i>	7
4.3 Verifikation der Messergebnisse	73
5 Zusammenfassung und Ausblick	77
5.1 Zusammenfassung	77
5.2 Ausblick	78
A Aufgabenstellung	79
B Mathematische Verfahren	81
B.1 Lösung eines homogenen Gleichungssystems	81
B.2 Regressionsanalyse	82
C Weitere Berechnungsverfahren	85
C.1 Detektion von Reflektionsmarkern	85
D Grundlagen der klassischen Mechanik	87
D.1 Inertiale Bezugssysteme	87
D.2 Kinematik von Massepunkten	88
D.2.1 Kinematik von Starrkörpern	88
D.3 Dynamik	89
D.3.1 Dynamik von Massepunkten	90
D.3.2 Dynamik von Starrkörpern	91
D.3.3 Kraftmodelle	92
D.3.4 Das gemeinsame Wirken mehrere Kräfte	93
E Aufbau der DVDs	95
E.1 Verzeichnisstruktur	96

E.2	Dokumentation erstellen	97
E.3	Den Quelltext kompilieren	97
E.4	Die Ausarbeitung und den Seminarvortrag erstellen	97
E.5	Testprogramme zur Verifikation der Tracking-Module	98
E.6	Videos und Konfigurationsdateien	99
Literaturverzeichnis		102

Kapitel 1

Einleitung

Für viele Anwendungen kann die Fähigkeit zur Vorhersage von Objektbewegungen in Kamerabildfolgen eine wertvolle Bereicherung darstellen. Bei einigen Trackingverfahren werden bereits Bewegungsvorhersagen verwendet, um den negativen Effekt von kurzzeitigen Verdeckungen zu kompensieren. Auch kann eine Bewegungsvorhersage der Beschleunigung interner Optimierungsprozesse dienen, was eine höhere Performanz und eine verbesserte Stabilität der verwendeten Algorithmen mit sich bringt [KM96] [GD96]. Eine weitere Anwendung findet sich im Bereich der autonomen mobilen Systeme. Dort werden Bewegungsvorhersagen bereits eingesetzt, um Kollisionen mit bewegten Objekten in der Umgebung zu vermeiden [MS03].

Ein spezieller Anwendungsbereich für Bewegungsvorhersagen stellt das Tracking von Personenbewegungen dar. Hierbei ist es möglich, zusätzliche Semantik in Form diverser Modelle, welche die anthropometrischen und physiologischen Eigenschaften der beobachteten Person abstrahieren, in den Vorhersageprozess einzubringen. Die Vorhersage selbst kann dabei in verschiedenen Koordinatensystemen stattfinden. Üblich ist es, bei der Verwendung eines Modells, diese Vorhersage in den verallgemeinerten Koordinaten des Modells durchzuführen. Oft steht ein solches personenbezogenes Modell jedoch nicht zur Verfügung. Liegt ein solcher Fall vor, kann eine Vorhersage der Bewegung von natürlichen Merkmalen oder Markern in den Kamerabildern durchgeführt werden. Stehen zudem mehrere Kameras zur Verfügung, bietet sich eine Vorhersage im Weltkoordinatensystem

als eine weitere mögliche Alternative an.

Im Rahmen dieser Arbeit wurde ein Verfahren implementiert, welches einerseits eine Vorhersage der Bewegung in den Kamerabildern und andererseits eine Vorhersage der räumlichen Bewegung einer Person ermöglicht. Da beide Ansätze in der Theorie sowohl Vor- als auch Nachteile mit sich bringen, wurden experimentelle Versuche durchgeführt, um die Vorhersagequalität der beiden Ansätze zu vergleichen. Im Folgenden werden die Ergebnisse dieser Arbeit vorgestellt.

Der weitere Aufbau dieser Arbeit gliedert sich wie folgt:

Kapitel 2 bietet eine kurze Einführung in die theoretischen Grundlagen dynamischer Systeme und darauf aufbauende Modelle und Verfahren, welche zur Vorhersage von Bewegungen eingesetzt werden können. In Kapitel 3 werden die verwendeten Methoden und Modelle des Verfahrens, welches im Rahmen dieser Arbeit implementiert wurde, erläutert. Ferner wird der eigene Beitrag von den bereits bestehenden Ansätzen abgegrenzt. Im Anschluss werden die durchgeführten Experimente beschrieben und die zugehörigen Ergebnisse präsentiert. Hier wird im Speziellen die Güte der realisierten Vorhersagemethoden verglichen. Kapitel 5 schließt die Arbeit mit einer Zusammenfassung der Ergebnisse ab.

Kapitel 2

Stand der Wissenschaft

Optisches Tracking kann als iterativer Prozess der Detektion von Merkmalen und der subsequenten Verfolgung ihrer Bewegungen in einem Strom von Bilddaten verstanden werden. Eine zeitliche Änderung der Koordinaten, welche eingesetzt werden, um die Lage der betrachteten Merkmale zu repräsentieren, wird in der technischen Mechanik i. A. als Bewegung aufgefasst [Gre98, Begr.: Bewegung].

Werden mehrere Kameras eingesetzt, so ist es möglich, die Struktur der betrachteten Szene in Teilen zu rekonstruieren. Über die Zeit kann somit auch, neben der zweidimensionalen Bewegung in den Bildern, eine entsprechende dreidimensionale Bewegung im Raum ermittelt werden.

Ein Modul zur Bewegungsvorhersage, welches ein System zur reinen Bewegungsmessung erweitert, muss den folgenden Kernanforderungen genügen:

- Eine möglichst kompakte Repräsentation der jeweiligen Zustände zu jedem Zeitpunkt der Gesamtbewegung,
- eine möglichst genaue Repräsentation der Dynamik, welche der Gesamtbewegung zugrunde liegt *und*
- eine möglichst hohe Robustheit gegenüber Messfehler.

Im weiteren Verlauf des Kapitels werden Verfahren vorgestellt, welche den oben genannten Anforderungen genügen.

2.1 Überblick

Abschnitt 2.2 dieses Kapitels bietet eine Einführung in die grundlegenden Begriffe der *dynamischen Systeme*, welche die Grundlage für die physikalisch-basierte Berechnung von Bewegungsabläufen bilden. Um die anthropometrischen und physiologischen Eigenschaften des Menschen auf ein dynamisches System abzubilden, sind adäquate Modelle erforderlich. Die Anthropometrie des Körpers wird dabei anhand von Menschmodellen (Abschn. 2.4) abstrahiert, die Kraftentwicklung durch das Zusammenspiel von Muskelmodellen (Abschn. 2.5) und motorischer Kontrolle (Abschn. 2.6). Gegen Ende des Kapitels wird erläutert, wie die bis dahin beschriebenen Methoden integriert werden können. Abschnitt 2.7 beschäftigt sich u. a. mit einer Überführung der gemessenen Koordinaten der Merkmale in Modellkoordinaten. Konkrete Beispiele für integrierte Systeme werden anschließend in Abschn. 2.8 erläutert.

2.2 Dynamische Systeme

Dynamische Systeme ermöglichen die Simulation der Zeitentwicklung von physikalischen, biologischen und anderen realen System [BSMM01, S. 819]. Im Allgemeinen bestehen ein solches System aus einem Satz aus endlich vielen dynamischen Variablen, welche über die Zeit verändert werden. Der *Zustand* eines dynamischen Systems wird als Spaltenvektor $\boldsymbol{x} = (x_1, x_2, \dots, x_N)^T \in \mathbb{R}^N$, über dem als *Phasenraum* $M \subseteq \mathbb{R}^n$ bezeichneten Vektorraum, aufgefasst. Jeder Zustandsvektor setzt sich dabei i. d. R aus den genannten dynamischen Variablen, sowie aus deren zeitlichen Ableitungen zusammen. Betrachtet man die dynamischen Variablen ohne ihre Ableitungen, so spricht man auch von der sogenannten *Konfiguration* des Systems. Dieser ist ein Vektor über dem als *Konfigurationsraum* bezeichneten Vektorraum. [Gre03, S. 376]

Die Dynamik des Systems $\boldsymbol{x}(t)$ wird im Allgemeinen durch eine nichtlineare Funktion \boldsymbol{F}

in Form einer Differentialgleichung beschrieben:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{F}(\mathbf{x}(t), t; \lambda)$$

Diese Funktion \mathbf{F} wird auch als *Geschwindigkeitsfeld* oder als *Vektorfeld* bezeichnet. Ihre Abhängigkeit von der Zeit t repräsentiert einen möglichen Einfluss von zeitlich variablen Kräften auf das System. Ist eine solche Abhängigkeit nicht gegeben, so nennt man das System *autonom*. Das 3. Argument λ der Differentialgleichung repräsentiert eine mögliche Abhängigkeit von vorgegebenen Konstanten, welche als *Kontrollparameter* dienen sollen. [Gre03, S. 377-378]

Während das Geschwindigkeitsfeld des Systems die allgemeine Dynamik der wirkenden Kräfte repräsentiert, beschreibt der *Fluß* des Systems die Entwicklung der Dynamik von einem gegebenen Startzustand aus. Der Fluß entspricht somit einer Abbildung $\Phi_t : \mathbb{R}^N \mapsto \mathbb{R}^N$ mit $\Phi_t(\mathbf{x}_0) = \mathbf{x}(t)$ und beschreibt eine *Trajektorie* in dem gegebenen Phasenraum. Ob die Trajektorie für jeden Startvektor eindeutig ist, hängt von der verwendeten Differentialgleichung ab. Ein Kriterium für den Nachweis der Eindeutigkeit bietet beispielsweise der Satz von Picard-Lindelöf [NEBES98, S. 209]. Ist der Fluß nur für diskrete Zeitschritte definiert, so bezeichnet man das System i. A. als *diskret*, ansonsten als *kontinuierlich* [BSMM01, S. 819].

Lässt sich der Fluß mithilfe einer linearen Abbildung beschreiben, wird das System als *Lineares dynamisches System* (LDS) bezeichnet. Durch eine entsprechende Linearisierung, beispielsweise mithilfe der Taylor-Formel, lässt sich für viele nichtlineare Systeme eine Näherung mit einem angemessenen Fehler finden. Ein wichtiges Kriterium ist hierbei die Komplexität der Bewegung. Während sich kleine Bewegungen mit einem geringen Fehler linearisieren lassen, gilt dies nicht für umfangreiche Bewegungen [SE04, S. 219].

2.2.1 Freie mechanische Systeme

Besteht in der Anwendung des Systems ein Bezug zu Mechanik, so bezeichnet man ein dynamisches System im Allgemeinen auch als *mechanisches System*. Ein mechanisches System heißt *diskret*, wenn kein kontinuierlicher Einzelkörper betrachtet wird, sondern

Art des Elements	eben	räumlich
Massepunkt	2	3
Starrer Körper	3	6

Tabelle 2.1: Anzahl der Freiheitsgrade für ausgewählte Elemente eines diskreten mechanischen Systems [SE04, S. 11].

ausschließlich Mengen von diskreten Körpern oder Massepunkten, welche im Folgenden als *Elemente* des Systems bezeichnet werden.

Sind die einzelnen Bewegungen der betrachteten Massepunkte nicht eingeschränkt, so wird ein mechanisches System als *frei* bezeichnet. Die *Anzahl der Freiheitsgrade* e eines freien Systems, also die Anzahl der unabhängigen Koordinaten, welche notwendig sind um die Lage aller Systemelemente eindeutig zu beschreiben, genügt der Relation $e = \sum_{i=1}^p e_i$, wobei e_i die Anzahl Freiheitsgrade des i -ten Elements des Systems repräsentiert (siehe auch Tabelle 2.2.1). [SE04, S. 11]

Kräfte bestimmen die Dynamik des Systems. Gelingt es für die wirkenden Kräfte ausreichend präzise Modelle in Form von von Differenzialgleichungen zu finden, so lassen sich zukünftige Systemzustände, und dadurch auch die damit verbundenen Bewegungen, vorausberechnen [SE04, S. 103]. In der Regel existiert für jeden Freiheitsgrad eines diskreten Systems eine gewöhnliche Differenzialgleichung erster Ordnung, welche die Bewegung des Freiheitsgrades in Vergangenheit und Zukunft bestimmt. Eine Bewegungsvorhersage ist folglich durch das Lösen eines *Anfangswertproblems*, welches durch das Gleichungssystem und den zuletzt bekannten Zustand gegeben ist, möglich. Da eine analytische Lösung des Problems in vielen Fällen nicht ermittelt werden kann, oder der zeitliche Aufwand ihrer Berechnung nicht vertretbar ist, werden in der Praxis im Regelfall Näherungsverfahren eingesetzt. Eine umfangreiche Auswahl solcher Verfahren werden in [Ebe04, S. 457f.] beschrieben.

2.2.2 Gebundene mechanische Systeme

Der bisher beschriebene Ansatz führt unter Umständen zu sehr komplexen Systemen. Unter Umständen lässt sich jedoch diese Komplexität, und der damit verbundene Aufwand, bei der Berechnung stark reduzieren. Betrachtet man beispielsweise zwei Punkte, welche sich jeweils an einem Ende eines Metallstabs befinden, so lässt sich leicht feststellen, dass der Abstand dieser beiden Punkte in einem gewissen Intervall liegen muss, sofern nicht allzu große Kräfte wirken. Vielmehr ist er durch materialspezifische Eigenschaften wie z. B. Flexibilität und Dehnung eingeschränkt. Ein System, dessen Elemente sich nicht uneingeschränkt bewegen können, wird als *gebundenes System* bezeichnet.

Es gibt unterschiedliche Arten von *Bindungen* in einem solchen System. Grundlegend unterscheidet man zwischen *kinematischen Bindungen* und *physikalischen Bindungen*. Bei kinematischen Bindungen wird die Einschränkung der Lage über eine *Zwangsbedingung*, welche zu jedem Zeitpunkt erfüllt sein muss, erzwungen. Ein konstanter Abstand zwischen zwei Massepunkten lässt sich beispielsweise durch eine kinematische Bindung verwirklichen. Kinematische Bindungen werden oft verwendet, um ideale *Gelenke*, *Lager*, *Führungen* oder andersartige ideale Verbindungen zwischen den Elementen eines Systems zu repräsentieren. Dagegen stellt eine physikalische Bindung eine Beziehung zwischen den gegebenen Abständen der Systemelemente und in dem System wirkenden Kräfte her. Ein Beispiel für eine physikalische Bindung ist ein Feder-Masse-System. [GHSW06, S. 81]

Zwangsbedingungen können diverse Eigenschaften zugeschrieben werden. Sind sie zeitlich invariant, bezeichnet man sie als *skleronom*, andernfalls als *rheonom*. Weiterhin bezeichnet man Zwangsbedingungen, welche als Gleichung formulierbar sind, als *holonom*, ansonsten als *nichtholonom* [Gre03, S. 237]. Im Folgenden sind vorerst ausschließlich skleronome und holonome Zwangsbedingungen von Interesse. Auf nichtholonome Zwangsbedingungen wird am Ende des Abschnitts kurz eingegangen.

Das Vorhandensein von Zwangsbedingungen erfordert i. A. eine Integration von geeigneten *Zwangskräften* in die Bewegungsgleichungen. Der Beitrag von Zwangskräften zur Gesamtbewegung besteht darin, denjenigen Kraftanteil der ursprünglichen Gleichung aufzuheben, welcher zu einer Verletzung der entsprechenden Zwangsbedingung führen würde. Oft sind sie jedoch nicht von essentiellen Interesse für die Lösung einer gegebenen Pro-

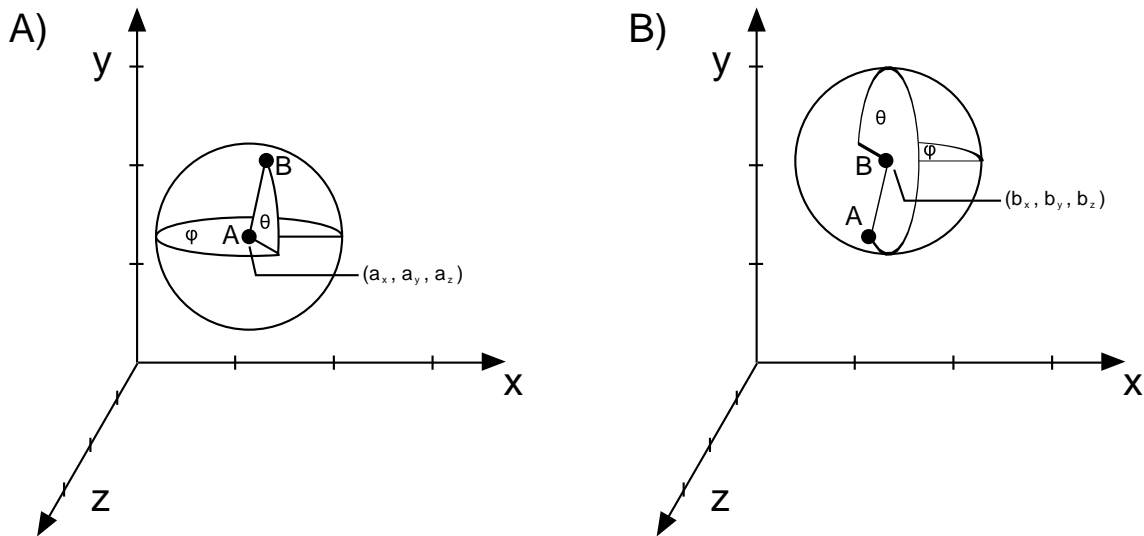


Bild 2.1: Illustration zu dem im Text erläuterten Beispiel.

blemstellung, da sie den Bewegungsverlauf nicht aktiv beeinflussen. Wie im Folgenden erläutert wird, ist es jedoch möglich, eine alternative Koordinatendarstellung des Systems zu finden, in der die Zwangsbedingungen, und somit auch die Zwangskräfte, nicht mehr in den Gleichungen auftreten.

Ein gebundenes System, in dem Bewegungen durch holonome Zwangsbedingungen beschränkt sind, benötigt stets weniger Freiheitsgrade als ein freies System, welches eine identische Menge an Körpern besitzt. Die tatsächliche Anzahl der sogenannten *Lagefreiheitsgrade* f genügt dabei der Relation $f = e - p$, wobei p die Anzahl der unabhängigen Zwangsbedingungen repräsentiert [SE04, S. 10f.].

Wie diese Freiheitsgrade gewählt werden, hängt vom gegebenen Kontext ab. Wird beispielsweise ein konstanter Abstand zwischen zwei Punkten \mathbf{a} und \mathbf{b} durch die Bedingung $|\mathbf{a} - \mathbf{b}|^2 = r^2$ erzwungen, beschreibt eine Kugeloberfläche, mit einem der Punkte als Mittelpunkt, alle möglichen Positionen des anderen. Die benötigten Freiheitsgrade des Systems könnten hier durch die Angabe des Mittelpunktes und die zwei Winkel der Kugelkoordinaten besetzt werden. Durch Winkel wird in diesem Fall die relative Lage eines Punktes zum anderen in Kugelkoordinaten angegeben, wobei eine redundante Koordinate, der Radius, wegfällt (Bild 2.2.2.B). Letzteres bewirkt, dass die zuvor benötigte Zwangs-

gleichung obsolet wird, denn es sind genau alle 5 Lagefreiheitsgrade des Systems besetzt. An dem Beispiel lässt sich auch erkennen, dass es nicht allein eine Möglichkeit gibt, die Freiheitsgrade eines mechanischen Systems zu wählen: Der Mittelpunkt kann wahlweise anhand von \mathbf{a} aber auch anhand von \mathbf{b} festgelegt werden (Bild 2.2.2.B). Eine derartige Koordinatendarstellung wird mit dem Begriff der *generalisierten* oder auch der *verallgemeinerten Koordinaten* bezeichnet.

Neben der Wahl einer geeigneten verallgemeinerten Koordinatendarstellung, muss zusätzlich eine Möglichkeit gegeben werden, diese in die kartesische Darstellung zu transformieren. Nach der Wahl geeigneter generalisierter Koordinaten und den zugehörigen Transformationen, ist es möglich, auch für die gegebenen Bewegungsgleichungen eine alternative Darstellung zu finden, welche das Existenz von Zwangskräften erübrigt. Diese sogenannten *verallgemeinerten Kräfte*, lassen sich mit Hilfe der *Langrang'schen Mechanik* gewinnen. Für weitere Details wird an dieser Stelle jedoch auf [Ebe04] oder [SE04, S. 103] verwiesen.

Nichtholonome Zwangsbedingungen erfordern i. d. R. eine gesonderte Behandlung bei der Dynamiksimulation. Da die mit ihnen assoziierten Zwangskräfte nicht zu jedem Zeitpunkt wirken, ist eine entsprechende *Detektion* mit einer subsequenten *Behandlung* erforderlich. Dies kann unter Umständen zu einem hohen Mehraufwand bei den Berechnungen führen [Ebe04].

2.2.3 Stochastische Verfahren

Der bisherige Ansatz der Bewegungsvorhersage, wie er in Abschnitt 2.2.1 beschrieben wurde, setzt u. a. voraus, dass der Systemzustand bekannt ist. Messungen, sind jedoch stets von einem Rauschen überlagert. Der tatsächliche Systemzustand gilt somit als unsicher. Je nach *Stabilität*¹ der betrachteten Trajektorie können jedoch selbst kleine Änderungen des Startzustands zu stark abweichenden Ergebnissen bei der Vorhersage führen. Eine möglichst genaue Bestimmung des Systemzustandes ist somit ein ausschlaggebendes Kriterium für eine qualitativ hochwertige Vorhersage.

¹Der hier verwendete Stabilitätsbegriff bezieht sich auf die *Lyapunov-Stabilität* [Gre03, S. 398].

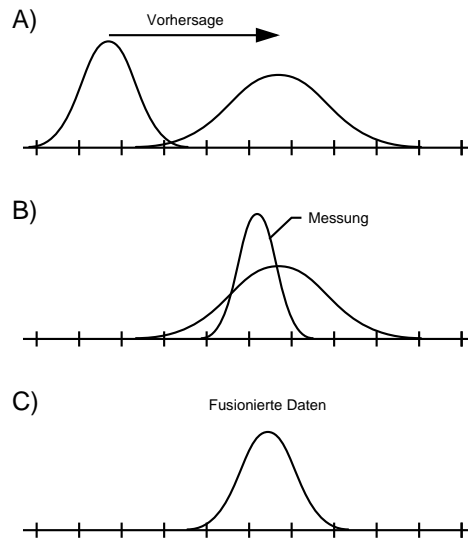


Bild 2.2: Vorhersage eines normalverteilten Zustandes (A) und eine anschließende Fusion (C) mit normalverteilten Messungen (B).

Anstatt eines einzelnen, vermeintlich wahren Systemzustands, kann der Ansatz verfolgt werden, eine Menge von Hypothesen zu betrachten, welche durchgehend mit einer bestimmten Wahrscheinlichkeit behaftet sind. Hierbei kommen, je nach Anwendung und Voraussetzungen, diskrete *Partikelsysteme* oder kontinuierliche Verteilungsmodelle zum Einsatz. Eine wahrscheinlichkeitsbehaftete Vorhersage des Systemzustands wird in diesem Kontext auch als *a-priori Schätzung* bezeichnet. Da die Streuung der Verteilung im Regelfall durch die Vorhersage verbreitert wird, ist es notwendig, diesem Prozess durch eine geeignete *Fusion* der Vorhersage mit den erfolgten Messungen entgegenzuwirken (Bild 2.2.3). Die korrigierte Verteilung wird auch als *a-posteriori Schätzung* bezeichnet.

Die Wahl des Verfahrens, welches zur Fusion der Daten eingesetzt werden kann, hängt sowohl von der Art des dynamischen Systems, als auch von der verwendeten Wahrscheinlichkeitsverteilung ab. Oft genügt eine Normalverteilung als Modell für ein zufälliges Rauschen. Ist das zugrundeliegende dynamische System zudem zeitdiskret und linear, kann der *Kalman-Filter* [Kal] für die Fusion verwendet werden. Dieser gilt dabei als *optimaler Schätzer*, sofern die erforderlichen Voraussetzungen erfüllt sind. Liegt ein zeitkontinu-

ierliches lineares System vor, kann stattdessen der Kalman-Bucy-Filter [KB61] verwendet werden. Im Falle eines nichtlinearen Systemes bietet sich der *erweiterte Kalmanfilter* (EKF) als Näherung, sowie *Partikelfilter*, wie beispielsweise der *Condensation Algorithmus* [Isa98], für eine angenäherte Berechnung an.

2.3 Eigenschaften der verfügbaren Koordinatensysteme

Im vorangegangenen Abschnitt wurde in erster Linie eine theoretische Basis für die physikalisch-basierte Vorhersage vorgegeben, welche sowohl für eine Vorhersage in den Bildern als auch für ihr räumliches Äquivalent gültig ist. Bevor jedoch eine genauere Betrachtung konkreter Modelle erfolgen kann, ist es erforderlich, Eigenschaften der Räume hervorzuheben, welche die Modellbildung entscheidend beeinflussen.

Unter der Annahme, dass die Bildentstehung durch das *Lochkameramodell* erklärt werden kann, lassen sich Bilder als eine projektive Abbildung der Szenenstruktur auf die Bildebene der Kamera begreifen. Dieser Abbildungsschritt geschieht jedoch nicht ohne Informationsverlust. Zusammen mit den Tiefeninformationen der Weltpunkte, gehen weitere höherwertige Informationen, wie beispielsweise Form, Winkel, Abstände und Abstandsverhältnisse verloren, welche im Ursprungsraum zur Verfügung stehen. Die einzige Eigenschaft welche erhalten bleibt, ist die der Kollinearität [HZ03, S. 1]. Um die Bindungen eines mechanischen Systems festlegen zu können, sind jedoch gerade diese Informationen von großem Wert.

Gelingt es, die Struktur der Szene mit geeigneten Mitteln metrisch zu rekonstruieren, so stehen die verlorengangenen Informationen wieder zur Verfügung und können somit in die Modellbildung einbezogen werden. Eine Voraussetzung für eine metrische Rekonstruktion ist jedoch die Kenntnis der *intrinsischen Kameraparameter*. Fehlt dieses Wissen, so kann maximal eine projektive Rekonstruktion erzielt werden [HZ03, S. 264].

Ein Nachteil bei der Verwendung der rekonstruierten räumlichen Struktur, besteht jedoch in einer verringerten Genauigkeit der Daten. Aufgrund der diskreten Natur des Bildes, eventuellen Linearisierungen und Rundungsfehlern, muss mit einem größeren Fehler gerechnet werden (siehe Bild 2.3).

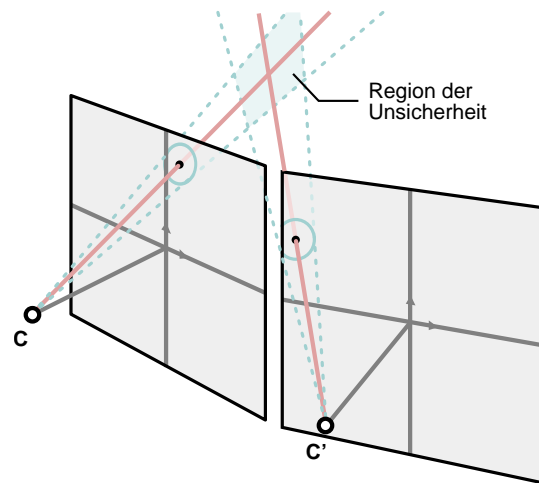


Bild 2.3: Idealisierte Darstellung der Region der Unsicherheit, bei der Rekonstruktion eines Weltpunktes.

2.4 Menschmodelle

Im Allgemeinen ist ein mechanisches System ein repräsentatives Modell für reale mechanische Abläufe. Man spricht daher auch von einem *mechanischem Ersatzsystem*. Durch die Anzahl der Systemfreiheitsgrade wird implizit die Realitätstreue des Modells bestimmt. Je nach Anwendung, muss daher zwischen der Genauigkeit des Modells und dem damit verbundenen Berechnungsaufwand abgewägt werden [SE04, S. 10].

Ein *Menschmodell* spiegelt das Ergebnis eines Bestrebens wieder, die anthropometrischen Eigenschaften des menschlichen Körpers auf ein abstraktes mechanisches System abzubilden [Sei03, S. 21]. Der Grad der Abstraktion reicht hierbei von *Strichmännchenmodellen* bis hin zu *deformierbaren Mehrkörpersystemen* mit dutzenden Möglichkeiten zur individuellen Parametrisierung, wobei die eigentliche Anwendung den entscheidenden Faktor bei der Modellwahl darstellt [MG01].

Von manchen Autoren wird eine liberalere Sicht auf Menschmodelle vertreten, wobei neben anthropometrischen Aspekten auch physiologische mit in die Modellbildung einbezogen werden:

Bezeichnung	Weitere Informationen ...
RAMSIS	http://www.human-solutions.com
Boeing HMS	http://www.boeing.com
SIEMENS Jack	http://www.plm.automation.siemens.com
SAFEWORK	http://www.safework.com

Tabelle 2.2: Auflistung kommerziell erwerblicher Menschmodelle.

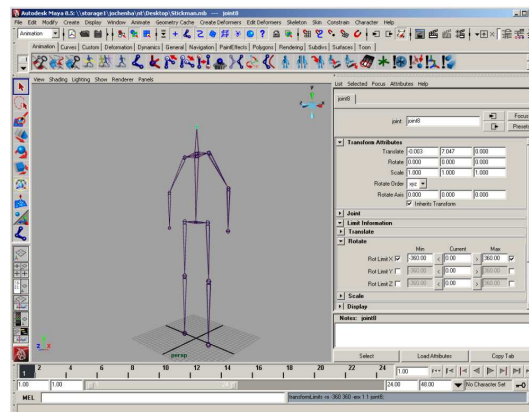


Bild 2.4: Die Erstellung eines Menschmodells unter Verwendung von Autodesk Maya.

Usually a geometric model of the human body is applied but other models, such as motion models, may also be applied. [...] The general concept of using a human model is to exploit the fact that the system is dedicated to analyse a human, and therefore may incorporate knowledge about humans into its processing. [MG01, Abschn. 5]

Die hier erwähnten Bewegungsmodelle werden im Rahmen dieser Arbeit jedoch getrennt von Menschmodellen behandelt.

Wegen ihres breiten Anwendungsspektrums, sind z. Z. verschiedene kommerziell erwerbliche Menschmodelle verfügbar. Eine Auswahl der verfügbaren Modelle findet sich in Ta-

belle 2.4. Zudem existieren mehrere Anwendungen aus dem Bereich der Computergrafik, welche das manuelle Erstellen von Menschmodellen unterstützen (siehe Bild 2.4).

Menschmodelle bestehen im Regelfall aus zwei Teilmodellen. Das *innere Modell*, auch als *Skelettmodell* bezeichnet, spiegelt die Bewegungsfreiheit des menschlichen Körpers, und damit die Freiheitsgrade des mechanischen Systems wieder. Die geometrische Repräsentation hingegen wird hingegen durch das äußere Modell realisiert [Gav99, Aschn. 5].

Im Folgenden wird zuerst auf Aspekte der Modellbildung im dreidimensionalen Fall eingegangen. Im Anschluss daran folgt ein kurzer Einblick in die Ansätze Menschmodellbildung im zweidimensionalen Fall.

2.4.1 Skelettmodelle

Das menschliche Skelett erfüllt diverse Funktionen im menschlichen Organismus. Es setzt sich aus durchschnittlich 206 bis 214 Knochen zusammen, welche mittels Gelenke in Verbindung stehen [HN99, S. 65]. Dadurch erfüllt es u. a. zwei wichtige Aufgaben im Rahmen der allgemeinen Körperbewegung [HN99, S. 64]:

- Es bietet Unterstützung gegenüber äußeren Kräften (wie z. B. Gravitation) und
- es repräsentiert ein Hebelsystem, über welches Muskelkräfte auf die Gelenke übertragen werden können.

Von einem makroskopischen Standpunkt aus, lassen sich die gesamten Bewegungsfreiheiten des menschlichen Körpers auf das Skelett zurückführen. Es gilt also, diese Bewegungsfreiheiten mit angemessenen Vereinfachungen auf die Freiheitsgrade des mechanischen Systems zu übertragen, welches dem Menschmodell zu Grunde liegt. Obwohl reale Knochen eine gewisse Elastizität aufweisen [HN99, S. 76] und Gelenke i. d. R. über alle 6 Freiheitsgrade verfügen [HN99, S. 207] lässt sich hier in beiden Fällen eine bedenkenlose Idealisierung vornehmen. Die Auswirkung dieser Idealisierung fällt bei einer makroskopischer Betrachtung des Körpers verschwindend gering aus.

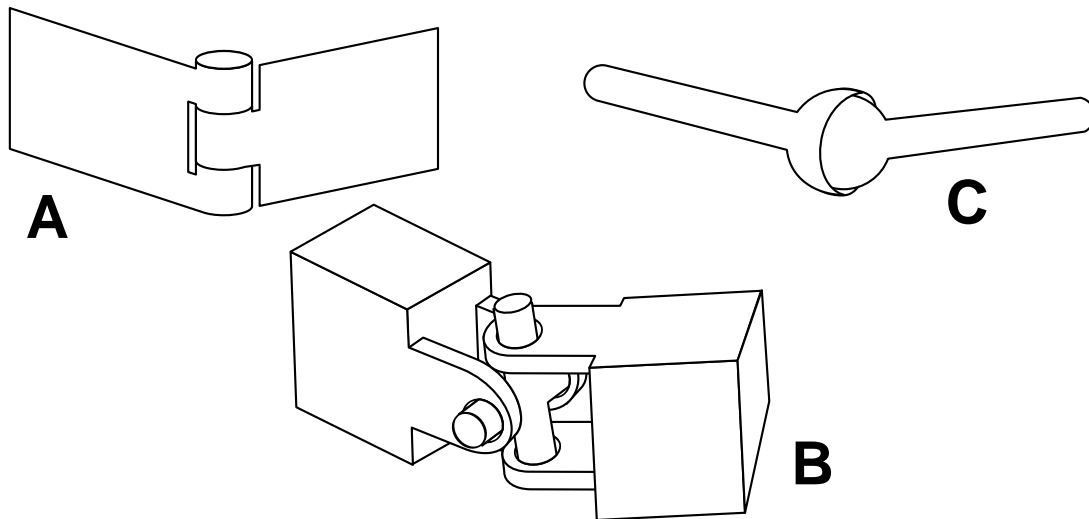


Bild 2.5: Idealisierte Gelenke: A) Ein Scharniergelenk, B) Ein Kreuzgelenk, C) Ein Kugelgelenk

Ideale Gelenke verfügen alleine über ihre Rotationsfreiheitsgrade, wobei auch diese reduziert sein können. Der Dreh- und Angelpunkt des Gelenks wird als *Gelenkpunkt* bezeichnet. Im Rahmen der unterschiedlichen Anwendungsbereiche werden ideale Gelenke oft durch mechanische Gelenke, wie sie im Maschinenbau verwendet werden, charakterisiert:

- Kugelgelenke (3 Freiheitsgrade),
- Kreuzgelenke (2 Freiheitsgrade, keine Rotation um die eigene Achse) und
- Scharniergelenke (1 Freiheitsgrad)

Eine bildliche Illustration der vorgestellten Gelenke befindet sich in Bild 2.5.

Unter der Annahme, Knochen seien ideale Starrkörper, gilt für zwei Gelenkpunkte, dass ihr Abstand zueinander stets konstant bleibt. Jedes Gelenk zwischen 2 Knochen des Skelettmodells erweitert das System somit um 2 bis 4 Freiheitsgrade, abhängig vom verwendeten Gelenktyp. Zudem ist die räumliche Position eines Referenzknochens erforderlich, um die Konfiguration, welche in diesem Kontext auch als *Pose* bezeichnet wird [MG01, Abschn. 5][Par02, S. 185], eindeutig festzulegen.

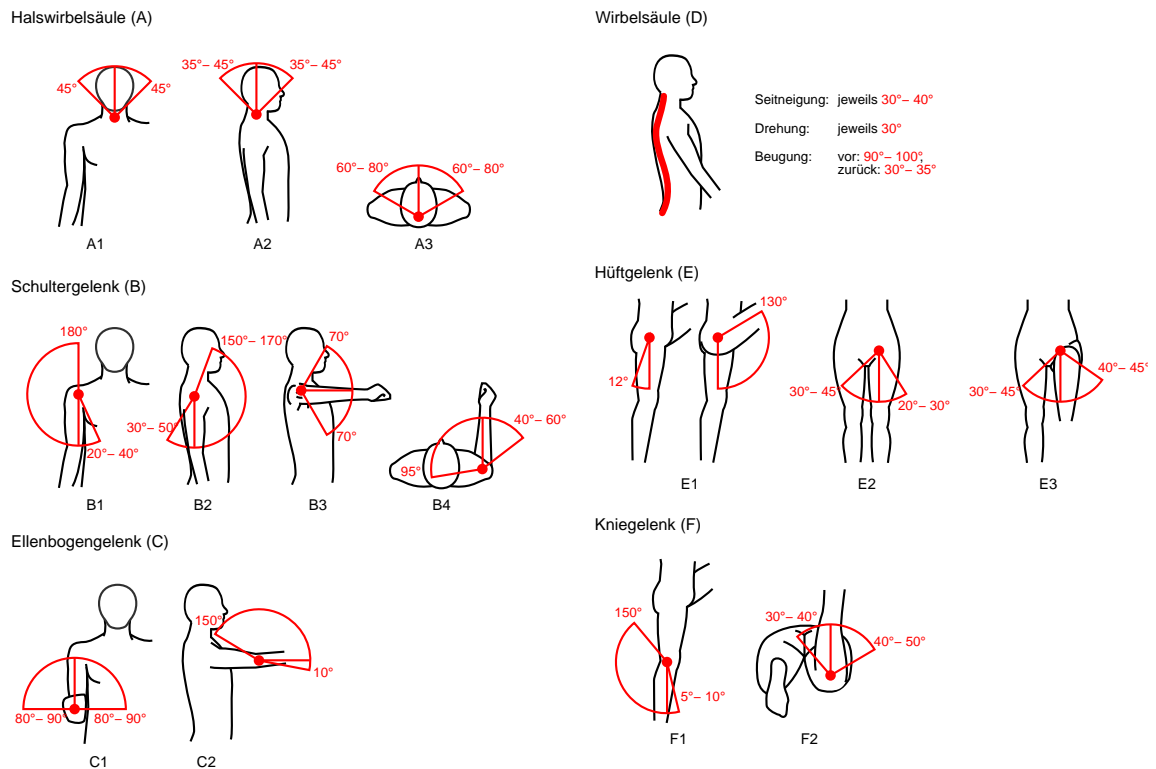


Bild 2.6: Der durchschnittliche Bewegungsspielraum einzelner Gelenke

Eine vollständige Modellierung aller Knochen würde in einer Komplexität jenseits jeder vertretbaren Grenzen resultieren. Daher ist es notwendig, diese Komplexität in einem angemessenen Rahmen zu verringern. Das zu Beginn des Abschnitts erwähnte Menschmodell RAMSIS umfasst beispielsweise über 26 Knochen, was eine erhebliche Vereinfachung gegenüber der Realität darstellt.

Die Bewegungsfreiheit des Skeletts wird jedoch nicht ausschließlich durch die Freiheitsgrade bestimmt. Reale menschliche Gelenke verfügen über einen stark eingeschränkten Spielraum bezüglich der möglichen Winkelstellungen. Dies erfordert die Einführung von nichtholonomen Zwangsbedingungen, welche eine gesonderte Behandlung erfordern, wie in Abschnitt 2.2.2 erläutert wurde. In Bild 2.6 sind durchschnittlich ermittelte Parameterbereiche der einzelnen Gelenke einer erwachsenen Person dargestellt [NP05].

Betrachtet man die Knochen als Knoten und die Gelenke als Kanten, lässt sich die Topologie eines Skelettmodells i. A. als Baum darstellen, wobei der zuvor erwähnte Referenzknochen als Wurzel dient [Par02, S. 177]. Eine Verallgemeinerung der Topologie auf einen Graphen ist dabei nur möglich, falls entsprechende Zwangskräfte eingeführt werden, da Schleifen im Graphen zusätzliche kinematische Bindungen erfordern würden. Eine gleichwertige Alternative wäre die Wahl einer neuen Koordinatenrepräsentation.

2.4.2 Hautmodelle

Während das Skelettmodell alleine die Bewegungsfreiheiten des menschlichen Körpers abstrahiert, legt das Hautmodell die räumliche Struktur des Körpers fest. Neben der Möglichkeit zur Visualisierung des Modells, erfüllt das Hautmodell weitere wichtige Funktionen im Rahmen der Simulation:

- Wichtige kinematische Eigenschaften wie Trägheit und Schwerpunkte werden erst durch die Teilkörper des Hautmodells festgelegt und
- die räumliche Ausdehnung der einzelnen Teilkörper bietet die Grundlage für eine Kollisionserkennung, welche zusätzlichen Realismus in die Simulation einfließen lässt.

Der letzte Punkt dient vor allem zur Einhaltung des sogenannten *non-penetration constraints*, welches das Eindringen eines Körperteils in ein anderes per se ausschließt.

Entspricht die Topologie des Modells einer Baumstruktur, so lässt sich für jeden Knoten eine Transformation in Form einer Transformmatrix bestimmen. Die Transformation der Wurzel dient der Bestimmung des ersten lokale Koordinaten im Rahmen des globalen Koordinatensystems. Jede weitere Transformation dient der Überführung des LKS des Elternknotens in das eigene. Mithilfe dieser Transformationshierarchie lässt sich das Hautmodell mit Hilfe der Gelenkkordinaten des Skelettmodells ausrichten. [Par02, S. 177][Gav99, Abschn. 5.1]. Zur Überführung von Gelenkkordinaten in die genannten Transformationsmatrizen hat sich die *Denavid-Hartenberg-Konvention* bewährt [BM98][Par02, S. 186].

Segment	Anteil [%]
Kopf	7,3
Rumpf	50,7
Oberarm	2,6
Unterarm	1,6
Hand	0,7
Oberes Bein	10,3
Unteres Bein	4,4
Fuß	1,5

Tabelle 2.3: Beispielhafte relative Masseverteilung von Körpersegmenten [HN99, S. 382].

Die Repräsentation des Hautmodells kann dabei unterschiedliche Formen annehmen. Üblicherweise werden für jeden Knoten des Skelettmodells ein oder mehrere Teilkörper definiert, welche im LKS des jeweiligen Knotens modelliert wurden. Bei der Teilkörperrepräsentation lässt sich zudem zwischen Ansätzen unterscheiden, welche Starrkörper verwenden, und Ansätzen, welche verformbare Körper einsetzen [Gav99, Abschn. 5.1]. Letztere werden bevorzugt, wenn Deformationen und die damit verbundene Änderung in der Dynamik für die Simulation von Belang sind. Andernfalls lässt sich durch die Abstraktion des menschlichen Körpers durch Starrkörper eine erhebliche Vereinfachung in der Berechnung der Dynamik erzielen.

Einfache Starrkörpermodelle werden im Regelfall aus geometrischen Primitiven, wie beispielsweise Boxen, Kegelstümpfe oder Ellipsoide zusammengestellt. Drahtgittermodelle oder Körper aus Freiformflächen stellen eine weitere Möglichkeit zur Repräsentation von Starrkörpern dar, erfordern jedoch einen vergleichsweise höheren Aufwand bei der Detektion von Kollisionen. Für verformbare Körpermodelle eignen sich beispielsweise verformbare Superquadricken [BM98] oder auch Feder-Masse-Systeme [HN99, S. 423][Par02, S. 347]. Sehr einfache Menschmodelle verzichten zudem gänzlich auf ein Hautmodell mit räumlicher Ausprägung. Der Schwerpunkt und die Gesamtmasse der einzelnen Knoten wird in einem solchen Fall in Form von Massepunkten angegeben.

Die Bestimmung der Massen einzelner Teilsegmente ist kein triviales Problem, da diese Werte zwischen einzelnen Individuen stark variieren können. Es existieren jedoch auch Methoden um eine angenäherte Masseverteilung einer Person zumindest halb-automatisiert zu bestimmen [HN99, S. 376]. Alternativ kann auf Tabellenwerke mit ermittelten Durchschnittswerten zurückgegriffen werden (siehe Tabelle 2.3).

2.4.3 2-D Menschmodelle

Aufgrund der fehlenden Winkel- und Längeninformatoren besteht bei zweidimensionalen Menschmodellen nur eine eingeschränkte Möglichkeit existierendes Wissen, über den menschlichen Körper zu verwenden. Die in [Gav99, Abschn. 4] und [MG01, Abschn. 5.3] beschriebenen Ansätze nutzen fast ausschließlich die Silhouette der Person um anschließend ein Strichmännchenmodell in diese einzupassen. Zudem existieren Modellierungsansätze, welche gänzlich ohne eine explizite Repräsentation der Körperstrukturen auskommen. Auf genauere Details dieser Ansätze wird im weiteren Verlauf dieser Arbeit jedoch nicht eingegangen. Stattdessen wird auf [Gav99, Abschn. 3] verwiesen.

2.5 Muskelmodelle

Menschmodelle, wie sie im vorangegangenen Abschnitt beschrieben wurden, dienen zum einen der Belegung der Freiheitsgrade und der weiteren Beschneidung des Konfigurationsraums durch weitere Einschränkungen (Parameterbereiche der Winkel, non-penetration constraint, etc.). Zu einer Dynamiksimulation gehören jedoch auch Kraftmodelle und die dazugehörigen Bewegungsgleichungen.

Reale Bewegungen werden durch eine Vielzahl von Kräften beeinflusst. Bei diesen Kräften unterscheidet man, je nach Sicht auf das Modell, zwischen *äußeren Kräften* und *inneren Kräften*. Die wichtigsten äußeren Kräfte sind vor allem Gravitations- und Reibungskräfte, für welche jedoch vereinfachte Standardmodelle existieren (siehe [Ebe04]). Bei inneren Kräften dominieren, bei aktiver Bewegung, vor allem die Kräfte welche von der Skelettmuskulatur erzeugt werden.

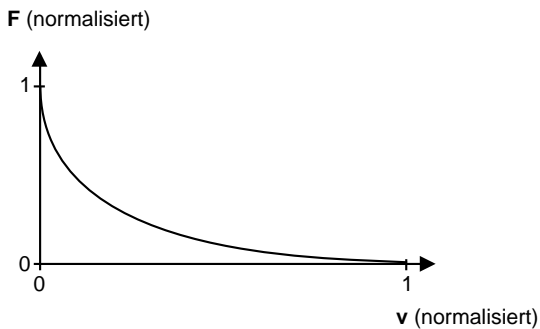


Bild 2.7: Die Kraftentwicklung einer Muskelfaser in Abhängigkeit von ihrer Länge [HN99, S. 606]

Die Skelettmuskulatur ist in der Lage, gespeicherte chemische Energie zu einem Teil in Zugkräfte umzusetzen, indem sie die Muskelfasern anregt zu kontrahieren. Diese Kräfte werden wiederum über das Bindegewebe, zu denen beispielsweise auch die Sehnen zählen, auf das Skelett übertragen. Dadurch, dass das Skelett ein System aus Hebeln darstellt, wird die Zugkraft in Form von Drehmomenten auf die Gelenke übertragen, was schließlich zur Bewegung des Körpers führt.

Ein naiver Ansatz für ein Modell, welches zur Muskelkraftsimulation eingesetzt werden kann, wäre ein Federkraftmodell. Ein solch Lineares Modell entspricht jedoch nur bedingt der Realität. Bild 2.7 zeigt die typische Kraftentwicklung einer Muskelfaser in Abhängigkeit von ihrer Länge [HN99, S. 606]. Die Kraftentwicklung eines realen Muskels verläuft also keineswegs linear. Ein weiterer Faktor bei der Kraftentwicklung ist zudem durch einen ebenfalls nicht linearen Zusammenhang zwischen Kraft und der Kontraktionsgeschwindigkeit gegeben. Zusammen mit einem dritten Parameter, der *Aktivierung*, welche den Grad der Kontraktion in Prozent widerspiegelt, lässt sich nach [HN99, S. 606] ein plausibles Muskelmodell entwickeln.

Ein Muskelmodell welches die genannten Kriterien erfüllt, ist das Modell von Hill [Hil38]. Hierbei wird neben dem Einfluß der Muskelfasern auf die Kraftentwicklung, die ebenfalls relevanten Auswirkungen von Sehnen und anderen elastischen Geweben mit einbezogen. Dieses relativ einfache Modell hat sich weitgehend bei der Simulation durchgesetzt und

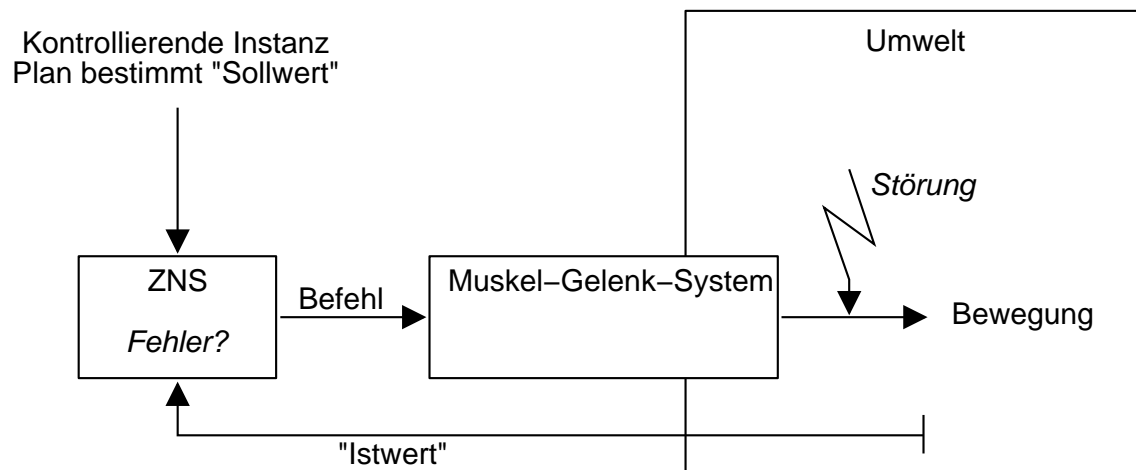


Bild 2.8: Bewegungsregelung: Durch Feedback-Schleife wird es dem ZNS ermöglicht, korrigierende Maßnahmen zu ergreifen, falls der tatsächliche Zustand vom geplanten aufgrund von Störungen abweicht [Ric01].

wird fast ausschließlich für Muskelsimulationen verwendet [HN99, S. 603]. Zwar existieren weitere Muskelmodelle, jedoch erfordern sie einen vergleichsweise hohen Berechnungsaufwand [HN99, S. 607].

2.6 Motorische Kontrolle

Alleine die Verfügbarkeit adäquater Mensch- und Muskelmodelle reicht nicht für eine realistische Simulation komplexer Bewegungen aus. Ohne eine kontrollierende Instanz, welche die Kontraktion der Muskeln in Abhängigkeit von bestimmten Eingangsvariablen steuert, sind komplexe Bewegungsabläufe nicht realisierbar.

Im menschlichen Organismus übernimmt das Zentralnervensystem (ZNS) die vollständige Koordination aller motorischen Abläufe. Es stellt dabei eine Verknüpfung zwischen der Gesamtheit aller zutragenden Nervenfasern, welche *Afferenzen* bezeichnet werden, und der Gesamtheit der wegführenden Nervenfasern, den *Efferenzen*, dar. Dabei haben diverse Eingangssignale einen Einfluß auf den reaktiven Ausgang der Verarbeitung im ZNS (siehe Bild 2.6). Im Rahmen der motorischen Kontrolle sind vor allem die folgenden Informatio-

nen von Belang:

- Sensorische Informationen über die aktuelle Muskellänge, welche von den Muskelspindeln übermittelt werden,
- sensorische Informationen der Wahrnehmungsorgane (vor allem haptische und optische) *sowie*
- eine gezielte Einflußnahme des Gehirns und speziell des motorischen Gedächtnisses.

Für den letztgenannten Punkt wurden im Rahmen der Experimentalpsychologie verschiedene Modelle entwickelt, welche jeweils verschiedene Aspekte der motorischen Kontrolle abstrahieren. In der Theorie der *Generalisierten Motorischen Programme* wird motorische Kontrolle auf der Basis eines festen Befehlssatzes, wobei die einzelnen Befehlswörter jedoch über eine sehr hohe Anzahl an Eingangsparametern verfügen, erklärt. Die Eingangsparameter werden mithilfe erlernter Schemata belegt, die je nach gegebener Situation aktiviert werden. Die *Äquilibrium-Punkt-Theorie* versucht dagegen Bewegungsabläufe anhand von Zielzuständen in einem internen Zustandsraum, welcher über die Muskelfaserlängen aufgespannt wird, zu erklären. Um diese Zielzustände zu erreichen, muss wiederum ein Gleichgewicht der internen und externen Kräfte und Momente erreicht werden, was der Theorie ihren Namen verleiht. Des Weiteren existieren die sogenannten *Vorwärts- und Rückwärtsmodelle*, welche nach dem Vorbild der Natur eine Menge Eingangsparameter (Sensorinformationen) über ein *künstliches neuronales Netz* (NN) mit den Ausgangsparametern (Gelenkwinkel) verknüpft. Letztere werden wahlweise mittels Vorwärts- oder Rückwärtskinematischer Ansätze bestimmt. Durch eine angemessene Lernmethode können bei diesem Ansatz die richtigen Bewegungsabläufe erlernt werden. [Ric01]

Die Ansätze, welche diesen Theorien zugrunde liegen, können zum Teil im Bereich der *dynamischen Bewegungserkennung* (vergleiche hierzu [MG01, Abschn. 6.2]) wiedergefunden werden. Die referenzierten Quellen bieten einen guten Überblick über diverse Ansätze mittels künstlicher neuronaler Netze, *hidden Markov Modellen* (HMM) etc.. Durch die Erkennung einer Bewegung und der subsequenten *Regelung* der Dynamik des Systems, wird die Grundlage für die Simulation der motorischen Kontrolle mittels Bilddaten gelegt.

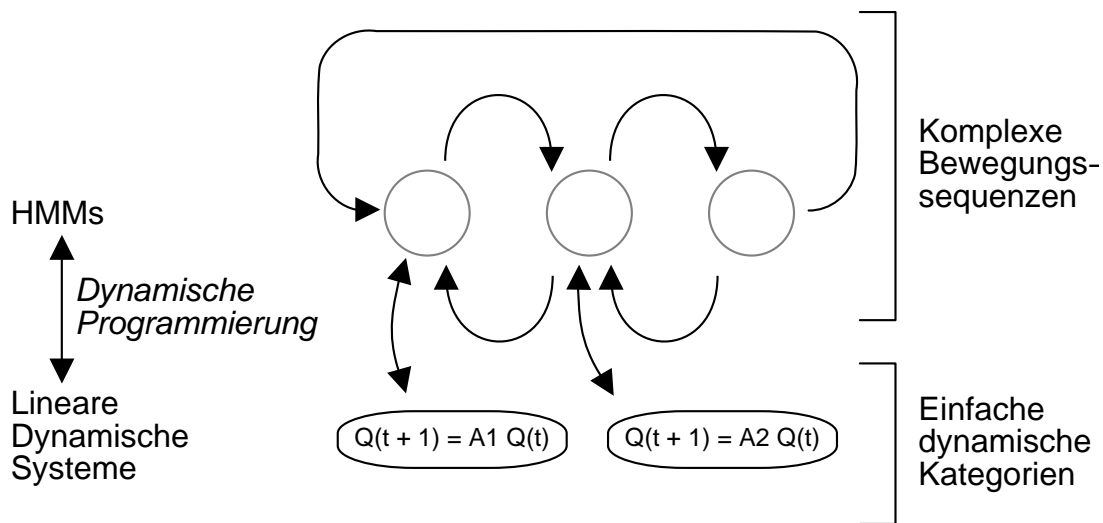


Bild 2.9: Aufbau eines geschalteten linearen dynamischen Systems [Bre97].

Dabei ist es wünschenswert, dass die Bewegungsabläufe, wie beim Einsatz von NNs und HMMs, erlernt werden können.

Ein bisher unerwähnter Ansatz, welcher aber an dieser Stelle hervorgehoben werden soll, ist der Einsatz von *geschalteten linearen dynamischen Systemen* (engl.: switched linear dynamic system; SLDS). Ein geschaltetes lineares dynamisches System wird auf höchster Ebene durch ein HMM repräsentiert. Die Zustände des HMMs repräsentieren dabei jeweils eine Phase in einer Folge komplexer Bewegungen. Die einzelnen Phasen werden dabei so gewählt, dass sich der Bewegungsablauf innerhalb einer Phase gut durch ein lineares Modell beschreiben lässt. Während der Verarbeitung der Sensordaten wechselt das HMM jeweils zum wahrscheinlichsten Zustand, welcher dann die passende Dynamik für eine Bewegungsvorhersage in das System integriert (siehe Bild 2.6). Auch dieser Ansatz hat zudem den Vorteil, dass individuelle Bewegungsmuster, wie auch bei Neuralen Netzen, erlernt werden können. [PRC00]

2.7 Methoden zur Ermittlung des Systemzustands

Für die Bewegungsvorhersage ist die Kenntnis des Systemzustands oder einer entsprechenden Wahrscheinlichkeitsverteilung unabdingbar. Ohne dieses Wissen ist kein Anfangswertproblem gegeben. Dieses bildet jedoch die Grundlage für eine Bewegungsvorhersage.

Werden ausschließlich Merkmale betrachtet, so reicht es aus, diese über einen angemessenen Zeitraum zu beobachten. Durch die Bildung von Differenzenquotienten aus den gemessenen Daten kann dann eine Näherung der Geschwindigkeiten und der Beschleunigungen zu einem gegebenen Zeitpunkt berechnet werden.

Ist jedoch ein Menschmodell gegeben, so muss der wahrscheinlichste Modellzustand anhand der Messungen ermittelt werden. Dabei wird zwischen Methoden unterschieden, welche den aktuellen Zustand des mechanischen Modells direkt oder nur indirekt verwenden [MG01, Abschn. 5].

2.7.1 Indirekte Verwendung des Modellzustands

Werden punktartige Merkmale verwendet, so lässt sich die Einpassung über eine Minimierung der Abstände von Datenpunkten und Modellpunkten erreichen. Am einfachsten gelingt dies, wenn Korrespondenzen zwischen den Datenpunkten und Modellpunkten bekannt sind. In der Regel besteht jedoch keine direkte Möglichkeit, einen solchen Zusammenhang zwischen natürlichen Merkmalen und den realen Objekten der Realität per se zu erschließen. Anders verhält sich dies jedoch bei der Verwendung von *kalibrierten Markern*. Bei diesen besteht stets ein zuvor festgelegter Zusammenhang zwischen Marker-id und einem bekannten Punkt in der Modellstruktur. Eine Abstandsminimierung kann in diesem Fall direkt zwischen den korrespondierenden Punkten vorgenommen werden.

Bei Bewegungen kommt es jedoch, aufgrund von Weichteilverformungen, zu Verschiebungen der Marker, welche im Regelfall auf der Hautoberfläche angebracht werden. Durch eine ausreichend hohe Anzahl an Markern kann ein Teil des entstandenen Fehlers jedoch geglättet werden. Zudem existieren spezialisierte Methoden um den Fehler weiter zu reduzieren. Ein kompakter Überblick über diese Methoden befindet sich in [HN99, S. 316f.].

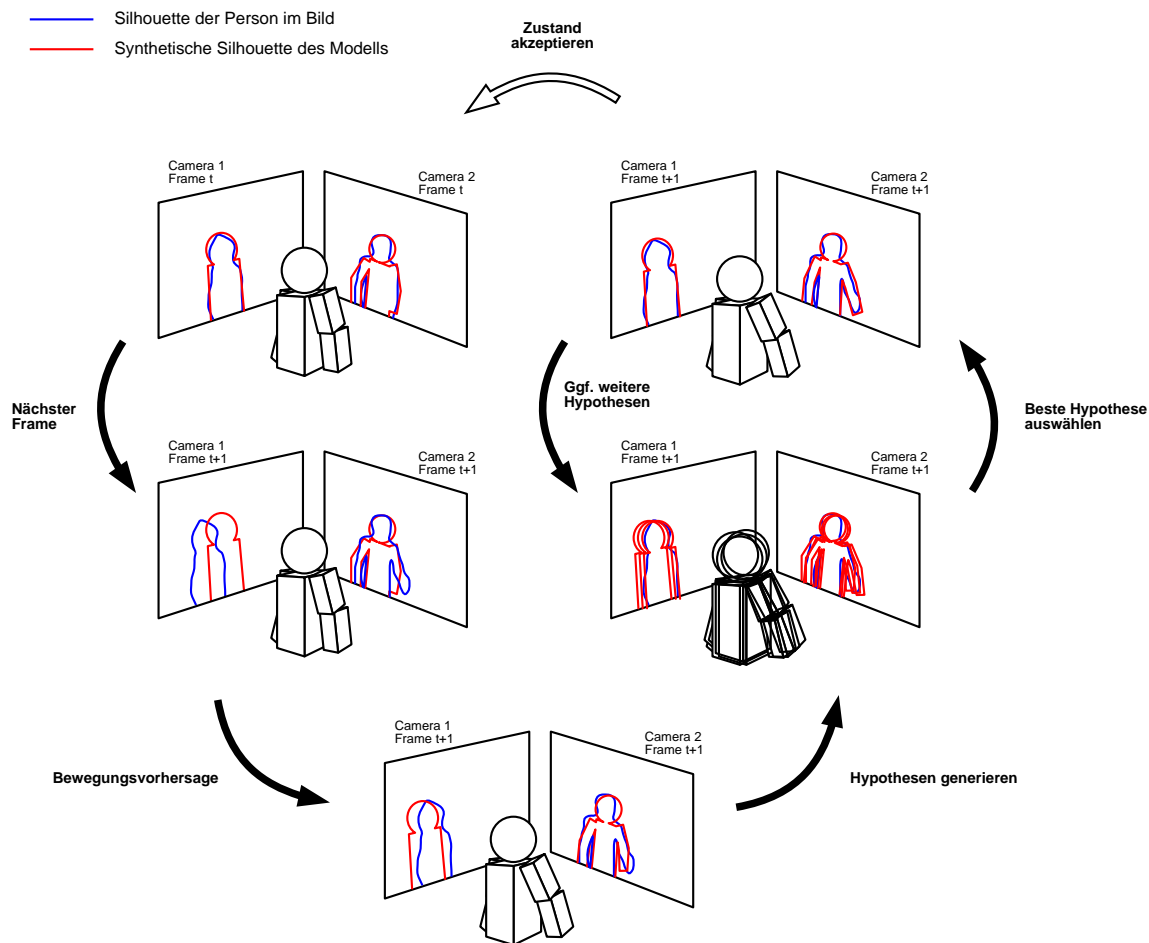


Bild 2.10: Beispiel eines Evolutionsprozesses beim modellbasierten Tracking.

2.7.2 Direkte Verwendung des Modellzustands

Trackingverfahren, mit direktem Einsatz der Modelle während des Trackingprozesses, werden unter dem Oberbegriff der *Analyse durch Synthese* zusammengefasst [MG01, Abschn. 5.3]. Anstatt ein Modell nachträglich in bereits erfasste Daten einzupassen, wird das Modell aktiv dazu verwendet, synthetische Merkmale, wie beispielsweise eine Silhouette des Modells zu erzeugen, welche anschließend mit den realen Merkmalen im Bild verglichen werden. Solange die Ähnlichkeit noch nicht ausreichend hoch ist, wird ein stochastisches Optimierungsverfahren dazu verwendet, Hypothesen mit leicht veränderten

Modellparametern zu erzeugen. Dies geschieht mit dem Ziel, per Zufall eine Pose zu generieren, welche mit der Pose der beobachteten Person übereinstimmt (siehe Bild 2.7.2). Abhängig vom verwendeten Optimierungsverfahren, kann der Optimierungsprozess iterativ verfeinert werden, bis alle notwendigen Kriterien erfüllt sind.

Im Rahmen dieser Verfahren hat sich eine Bewegungsvorhersage bereits als essentieller Bestandteil etabliert. Mithilfe einer guten Vorhersage ist es möglich, den Startzustand für die anschließende heuristische Optimierung signifikant zu verbessern, was u. a. in einer gesteigerten Performanz des Verfahrens resultiert. Zudem ist es möglich, die Bewegung von kurzzeitig verdeckten Körperteilen vorherzusagen, damit ein anschließender Wiederfindungsprozess erleichtert wird. [GD96]

2.8 Konkrete Methoden zur Bewegungsvorhersage

Im Folgenden sollen mögliche Integrationsvarianten der zuvor beschriebenen Methoden und Modelle, anhand konkreter Beispiele demonstriert werden. In der Literatur, welche im Rahmen dieser Arbeit gesichtet wurde, ließen sich zwei verschiedene Grundansätze zur Bewegungsvorhersage erkennen, welche sich in der Verwendung der dynamischen Modelle unterscheiden. Die dazugehörigen Ansätze werden im Rahmen dieser Arbeit als explizit und implizit bezeichnet, und im folgenden vorgestellt.

2.8.1 Beispiele expliziter Verfahren

Bei den expliziten Verfahren ist das zugrundeliegende dynamische Modell explizit gegeben. Alle gesichteten Ansätze nutzten dabei Zustände mit einer Wahrscheinlichkeitsverteilung, wie es in Abschn. 2.2.3 beschrieben wurde. Zudem wurden stets möglichst präzise dynamische Modelle verwendet.

[PRC00] beschreibt einen Ansatz, in dem die Dynamik einer Gangbewegung mithilfe eines SLDS modelliert wird. Hierbei dient der Kalman-Filter zur Berechnung der Modellzustände eines 2d-Modells, wie es in [MR98] vorgeschlagen wird. Aufgrund der Zweidimensionalität werden hier nur Bewegungen betrachtet, welche parallel zur Kameraebene

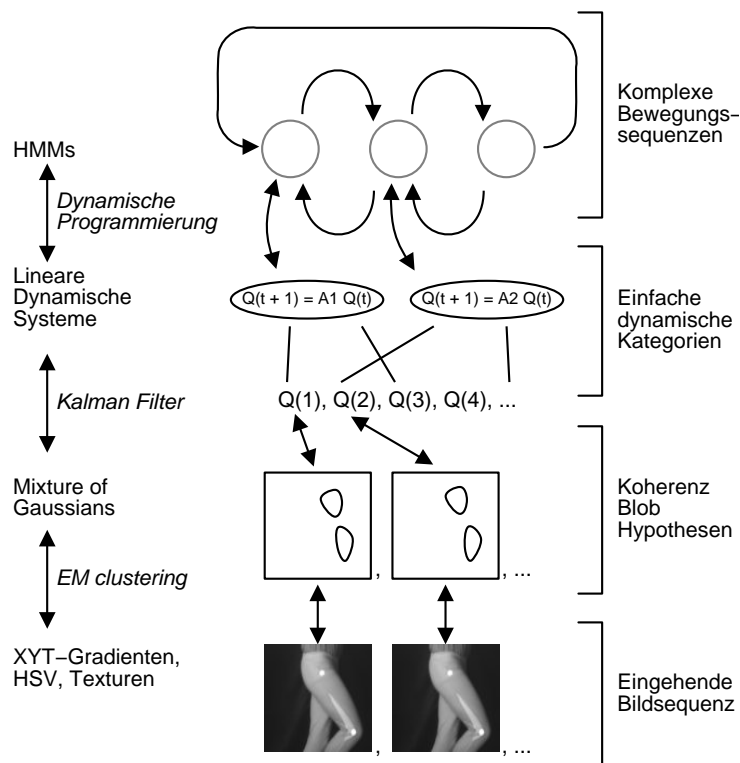


Bild 2.11: Verschiedene Ebenen der Abstraktion beim Tracking unter Verwendung eines SLDS [Bre97].

stattfinden.

Weitere Ansätze, welche ebenfalls die Dynamik von komplexen Bewegungen mithilfe eines SLDS simulieren, finden sich in [WP99] und [Bre97]. Beide detektieren Blobs in Bildern aus 2 Kameras, welche die Positionen von Körperteilen wie Hände und Füße repräsentieren, um den probabilistischen Zustand eines 3d-Modells zu ermitteln (Bild 2.8.1). Auch hier wird der Kalman-Filter für die notwendigen Vorhersagen und Korrekturen der Modellzustände verwendet.

Die Verwendung eines dynamischen Frameworks unter der Verwendung des EKF wird beispielsweise in [KM96] oder auch in [PW98] beschrieben. Ein Verfahren, welches die Dynamik von verformbaren Körpern simuliert, ist in [BM98] beschrieben. Dieser Ansatz basiert wiederum auf dem allgemeinen Rahmenwerk von [BAHH92]. Weitere Verfahren,

in denen u. a. auch Partikelfilter zum Einsatz kommen, lassen sich aus [SBF00] entnehmen.

2.8.2 Beispiele impliziter Verfahren

Die impliziten Verfahren nutzen ein dynamisches Modell nicht direkt, sondern leiten eine Funktion für den zu erwartenden Bewegungsverlauf in einem kurzen Zeitintervall ab. Im Allgemeinen stellt diese also die allgemeine Lösung der zugrundeliegenden Differentialgleichung des dynamischen Modells dar. Durch die Aufzeichnung des zeitlichen Verlaufs der einzelnen Koordinaten ist es möglich, die Bewegungsfunktion dort einzupassen, was eine Extrapolation des vermeintlichen Zustandsvektors zu einem zukünftigen Zeitpunkt ermöglicht. Die betrachteten Zustände werden somit nicht als probabilistisch aufgefasst. Damit eine robuste Vorhersage der Bewegung erfolgen kann, müssen geeignete Methoden zum Einpassen der Funktion verwendet werden.

In [KM96] wird ein solcher Ansatz zur Bewegungsvorhersage im Rahmen einer Analyse durch Synthese Verfahren verwendet. Die Ermittlung des Ähnlichkeitsmaßes erfolgt dabei durch das Abgleichen der Zustandshypothesen mit der Silhouette der beobachteten Person. Bei der Vorhersage wurde angenommen, dass die in dem System wirkenden Gesamtkräfte konstant sind. Da auch die Masse der Teilkörper als konstant angenommen wird, muss der zeitliche Verlauf der Koordinaten durch ein Polynom 2. Grades beschreibbar sein, sofern die Modellannahme zutrifft. Dieses Polynom wird in eine ausreichend große Reihe von Messungen eingepasst, was eine anschließende Extrapolation eines vermeintlichen Folgezustands ermöglicht. Genauere Details zu den verwendeten Methoden wurden jedoch nicht angegeben.

Kapitel 3

Vergleich von 2d- und 3d-Verfahren

Eine physikalisch-basierte Bewegungsvorhersage nutzt Modellinformationen, um die Wahrscheinlichkeit, dass der vorhergesagte Zustand eintritt, zu maximieren. Das Bereitstellen dieser Informationen ist jedoch nicht in allen Fällen möglich. Zum einen benötigen präzise Modelle i. d. R. eine Initialisierungsphase vor dem eigentlichen Tracking, in der eine ausreichende Parametrisierung stattfinden kann. Ist die Zielperson oder der zu erwartende Bewegungsablauf jedoch unbekannt, so können wichtige Parameter nicht belegt werden. Zwar gibt es Bemühungen, eine Initialisierung der Modelle zu automatisieren [Sei03], jedoch scheint dies bisher nur im Fall von sehr einfachen Modellen, wie z. B. in [WAPD97], zu funktionieren.

Eine interessante Frage ist deshalb, inwiefern parameterfreie Modelle im Rahmen einer physik-basierten Bewegungsvorhersage eingesetzt werden können. Ein parameterfreies dynamisches Modell wurde bereits in Abschn. 2.8.2 erwähnt. Aufgrund der Annahme, dass die Dynamik des Systems durch eine konstante Krafteinwirkung bestimmt wird, begründet [KM96] die Wahl eines Polynoms 2. Grades, welches in die aufgezeichnete Datenmenge eingepasst wird. Basierend auf diesem Ansatz lassen sich die folgenden Überlegungen anstellen:

- Bewährt sich das Modell ebenfalls im Kontext eines freien mechanischen Systems?
- Inwiefern wirkt sich die Wahl des Koordinatensystems auf die Qualität der Vorher-

sage bei der Verwendung dieses Modells aus?

- Kann die Qualität der Vorhersage durch den Einsatz von Polynomen eines höheren Grades verbessert werden?

Alle drei Fragen sind dabei stets im Kontext der Vorhersage von Personenbewegungen zu betrachten.

Meinem Kenntnisstand nach zu urteilen, sind diese Fragen bisher in keiner wissenschaftlichen Arbeit behandelt worden. Daher wurde eine Anwendung entwickelt, welche die Bewegungen in den Bildströmen von einem kalibrierten Stereokamerasystem trackt und auf dieser Grundlage die zukünftige Bewegung für einen gewählten Zeitpunkt vorhersagt. Die Vorhersage erfolgt dabei sowohl in den Bildern, als auch räumlich, damit ein anschließender Vergleich der Ergebnisse durchgeführt werden kann.

3.1 Überblick

Im Folgenden werden zuerst theoretische Überlegungen zur Problemlösung vorgestellt. Im Anschluss daran werden ausgewählte Verfahren erläutert, welche im Rahmen der entwickelten Anwendung zum Einsatz kommen.

3.2 Verallgemeinerung des dynamischen Modells

Eine Möglichkeit, die Wahl einer konstanten Kraft zu motivieren, ist die Approximation der tatsächlich wirkenden Gesamtkraft durch eine *Taylor-Approximation*. Nach dem *Satz von Taylor* kann eine stetig differenzierbare Funktion durch ein Polynom in einem gegebenen Intervall beliebig genau approximiert werden. Die Genauigkeit der Approximation hängt dabei u. a. vom Grad des Polynoms und von der Breite des gewählten Intervalls ab. Verkleinert man das Intervall, so sind i. A. weniger Polynomglieder für eine ausreichende Annäherung erforderlich. Dies ist ein wichtiges Prinzip bei allen Linearisierungsansätzen. Im Kontext der Bewegungsvorhersage über eine polynomielle Extrapolation bedeutet dies,

dass die Länge des Zeitintervalls, welches als Grundlage für die Extrapolation verwendet wird, einen bedeutenden Faktor für die Vorhersagegenauigkeit darstellt.

Eine Erhöhung des Polynomgrades muss jedoch nicht zwangsweise zu einem besseren Ergebnis führen. Polynome neigen dazu zu oszillieren, wenn sie in eine ungeeignete Datenmenge eingepasst werden. Die Eignung von Messdaten kann in diesem Sinne beispielsweise durch eine hohe Streuung oder eine zu geringe Anzahl an Stichproben beeinträchtigt werden. Somit sind sowohl das Messrauschen, als auch die Bildrate der verwendeten Kameras weitere ausschlaggebende Faktoren im Kontext der Vorhersagegenauigkeit.

Ein letzter Faktor, welcher an dieser Stelle nicht unerwähnt bleiben soll, ist die Größe des Zeitschritts für die Vorhersage. Da durch das Polynom nur eine kleine Umgebung der eigentlichen Bewegungsfunktion angenähert wird, ist eine Langzeitvorhersage der Bewegung mit dem hier verwendeten Ansatz i. d. R. mit einem großen Fehler behaftet. Bessere Ergebnisse sollten hier Methoden, welche explizite dynamische Modelle verwenden, erzielen.

Im Rahmen dieser Arbeit wurden Hochgeschwindigkeitskameras zur Aufzeichnung der zu analysierenden Bildfolgen verwendet. Unter diesen Bedingungen kann angenommen werden, dass in einem kurzen gewählten Zeitintervall die wirkende Kraft näherungsweise konstant ist und zudem genügend Datenpunkte vorhanden sind. Der Fehler, welcher bei der Vorhersage entsteht, sollte somit bei einer Vorhersage nicht allzu groß werden, sofern der Zeitschritt der Vorhersage klein im Vergleich zur Länge des Zeitintervalls gewählt wird.

3.3 2d vs. 3d

Unter der Annahme, dass die Bewegungen im euklidischen Raum durch ein Polynom niedrigen Grades beschreibbar ist, lassen sich weitere Überlegungen anstellen, inwiefern dieses Modell auch für die Bewegungen im Bild zutrifft. Auf Basis des Lochkameramodells hän-

gen die Koordinaten der Bildebene (x, y) und des Raums (X, Y, Z) wie folgt zusammen:

$$\begin{aligned}x(t) &= \frac{X(t)}{Z(t)} \\y(t) &= \frac{Y(t)}{Z(t)}\end{aligned}$$

Verläuft die räumliche Bewegung parallel zur Kameraebene, so ist die Funktion $Z(t)$ konstant. In diesem Fall lässt sich die Bewegung mit dem zuvor erwähnten Polynom ebenfalls exakt beschreiben. Theoretisch sollte die Vorhersage in der Bildebene in einem solchen Fall besser sein als die räumliche Vorhersage, da ihre Grunddatenmenge im Rahmen von realen Messungen eine geringere Streuung aufweist.

Im Falle einer zur Kameraebene orthogonalen Bewegung sind die Funktion $X(t)$ und $Y(t)$ konstant. Der Bewegungsverlauf im Bild wird diesmal durch eine echt gebrochen rationale Funktion beschrieben. Unter Umständen reicht der Grad des Modellpolynoms in diesem Fall nicht aus, um den lokalen Bewegungsverlauf im Bild ausreichend anzunähern. Daher kann es sein, dass die Vorhersage im Bild schlechtere Ergebnisse liefert als die räumliche Vorhersage.

Die hypothetischen Vermutungen, welche in den vorangegangenen Absätzen erläutert wurden, bilden die Grundlage für die in Kapitel 4 beschriebenen Experimente.

3.4 Vorhersagegenauigkeit

Damit die Genauigkeit einer Vorhersage quantifiziert werden kann, ist ein geeignetes Fehlermaß erforderlich. Dieses ist jedoch stets an ein spezielles Koordinatensystem geknüpft. Um einen Vergleich zwischen den Vorhersagen verschiedener Koordinatensysteme zu ermöglichen, muss eine der Vorhersagen, in einer geeigneten Art und Weise in das Koordinatensystem der anderen überführt werden. Eine Möglichkeit besteht in der Abbildung der 3d-Vorhersagen in die beiden Pixelkoordinatensysteme, was mithilfe der vorhandenen Kameramatrizen erfolgen kann. Alternativ könnten korrespondierende 2d-Vorhersagen trianguliert werden, um einen Vergleich im Weltkoordinatensystem zu ermöglichen. Da dies

jedoch einen zusätzlichen Fehler in die Messungen einbringen würde, ist die erste Variante vorzuziehen.

Für die Ermittlung der Vorhersagegenauigkeit bietet es sich an, die Vorhersagen für einen bestimmten Zeitpunkt mit der tatsächlichen Position des zugehörigen Merkmals desselben Zeitpunkts zu vergleichen. Ein ähnlicher Ansatz wird bei der Ermittlung des *Reprojektionsfehlers* verwendet. Dieser definiert sich wie folgt:

$$\sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2,$$

wobei d die euklidische Distanz darstellt. Die Punkte \mathbf{x} und \mathbf{x}' bezeichnen des weiteren die tatsächlichen Bildpunkte und $\hat{\mathbf{x}}$ und $\hat{\mathbf{x}}'$ die angenäherten.

Dieses Fehlermaß eignet sich dabei vor allem im Rahmen von Minimierungsansätzen. Bei einer individuellen Betrachtung des Fehlers für jede einzelne Vorhersage, kann jedoch auf die Summe verzichtet werden. Auch die Quadrierung der Distanzen ist nicht notwendig, da eine stärkere Gewichtung größerer Abstände in dem hier gegebenen Kontext eher Nachteilig wirkt. Das hier verwendete Fehlermaß für die Vorhersagegenauigkeit nimmt daher die folgende Form an:

$$d(\mathbf{x}_i, \hat{\mathbf{x}}_i) + d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i).$$

3.5 Die entwickelte Anwendung

Der entwickelte Tracker basiert in seinen Grundzügen auf den in [HZ03] beschriebenen Verfahren. Bild 3.5 illustriert die einzelnen Phasen, welche der Gesamtprozess im Idealfall durchläuft. Im Folgenden wird auf einige dieser Teilprozesse genauer eingegangen.

3.5.1 Die initiale Ermittlung der Kameraparameter

Bei der Entwicklung des Trackers wurde davon ausgegangen, dass sich die *extrinsischen Kameraparameter* nach der anfänglichen Initialisierung des Trackers nicht mehr ändern. Dies hat zur Folge, dass die *Kameramatrizen* beider Kameras einmalig zu Beginn ermittelt

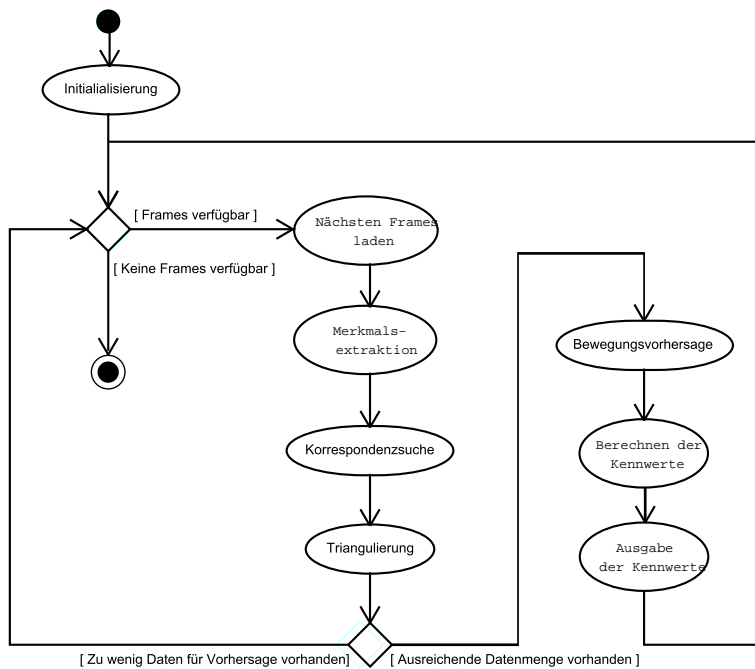


Bild 3.1: Vereinfachte Dartsellung Gesamtprozesses.

werden müssen. Für die Ermittlung der erforderlichen Parameter wurden verschiedene Verfahren implementiert, welche im Folgenden kurz vorgestellt werden sollen.

Zum einen wird eine direkte Bestimmung der Matrizen aus *ground-truth* Daten unterstützt. Hierbei kann beispielsweise ein vermessenes 3d-Kalibrierungsmuster eingesetzt werden, welches in beiden Kameras gleichzeitig sichtbar sein muss. Das hierzu implementierte Verfahren bestimmt in einem ersten Schritt eine Näherungslösung mit Hilfe des in [HZ03, S. 179] vorgeschlagenen linearen Verfahrens. Diese Näherungslösung wird mittels der in der OpenCV-Bibliothek enthaltenen Kalibrierungsfunktion weiter optimiert. Letztere basiert auf einer Implementierung des Verfahrens von [Zha00], welches neben den intrinsischen und extrinsischen Kameraparametern zudem die Parameter eines von [HZ03, S. 189] angeleiteten Modells für die radiale Linsenverzerrung berechnet. Diese werden zur Entzerrung der detektierten Merkmalspositionen und zur Entzerrung der ausgegebenen Bilder verwendet.

Lässt sich das 3d-Kalibrierungsmuster mit beiden Kameras nicht gleichzeitig erfassen,

besteht die Möglichkeit, die intrinsischen und extrinsischen Kameraparameter getrennt zu berechnen. Hierzu wird für jede Kamera getrennt eine Kameramatrix nach dem oben genannten Verfahren berechnet. Mittels einer RQ-Zerlegung lassen sich anschließend die intrinsischen Kameraparameter, in Form der *Kalibrierungsmatrizen* \mathbf{K} und \mathbf{K}' , aus den ermittelten Kameramatrizen faktorisieren.

Um die extrinsischen Kameraparameter zu ermitteln, ist die *Fundamentalmatrix* \mathbf{F} erforderlich. Sie lässt sich beispielsweise anhand des *8-Punkte Algorithmus* aus einer Menge von mindestens 8 Punktkorrespondenzen ermitteln, welche zwischen zwei zeitgleich aufgenommenen Bildern der Kameras existieren. Für die Berechnung der *essentiellen Matrix* \mathbf{E} gilt nun $\mathbf{E} = \mathbf{K}'^T \cdot \mathbf{F} \cdot \mathbf{K}$. Unter der Annahme, dass eine Kamera der kanonischen Kameramatrix $\mathbf{P} = \mathbf{K} \cdot [\mathbf{I}|\mathbf{0}]$ entspricht, lassen sich nun 4 mögliche Kandidaten für die zweite Kamera wie folgt bestimmen:

$$\begin{aligned} \mathbf{P}'_1 &= [\mathbf{U} \cdot \mathbf{W} \cdot \mathbf{V}^T | + \mathbf{u}_3] \\ \mathbf{P}'_2 &= [\mathbf{U} \cdot \mathbf{W} \cdot \mathbf{V}^T | - \mathbf{u}_3] \\ \mathbf{P}'_3 &= [\mathbf{U} \cdot \mathbf{W}^T \cdot \mathbf{V}^T | + \mathbf{u}_3] \\ \mathbf{P}'_4 &= [\mathbf{U} \cdot \mathbf{W}^T \cdot \mathbf{V}^T | - \mathbf{u}_3] \\ \mathbf{W} &= \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ \text{svd}(\mathbf{E}) &= \mathbf{U} \cdot \text{diag}(1, 1, 0) \cdot \mathbf{V}^T, \end{aligned}$$

wobei \mathbf{u}_3 für den letzten Spaltenvektor der Matrix \mathbf{U} steht.

Die Mehrdeutigkeit lässt sich durch die Rekonstruktion eines Weltpunktes anhand einer der Korrespondenzen auflösen, welche schon zur Ermittlung der Fundamentalmatrix verwendet wurden. Nur bei der korrekten Kandidatenmatrix liegt der rekonstruierte Punkt vor beiden Kameras, was als entscheidendes Kriterium für die Wahl der richtigen Kameramatrix dienen kann. Das Verfahren orientiert sich an den in [HZ03, S. 257f.] beschriebenen Vorgaben. Die Rekonstruktion ist zwar metrisch, jedoch existiert eine Mehrdeutigkeit in Form eines unbekanntem Skalierungsfaktors.

Die dritte Alternative besteht in der Verwendung des zuvor erwähnten Kalibrierungsverfahrens von [Zha00]. Hierbei wird ein ebenes Kalibrierungsmuster verwendet, welches

zeitgleich und aus mindestens 2 verschiedenen Perspektiven von den Kameras aufgenommen werden muss.

Eine letzte Möglichkeit besteht in der direkten Angabe der Kameraparameter, was bei der Verwendung synthetischer Bildfolgen von Nutzen sein kann.

3.5.2 Merkmalsdetektion

Die Detektion der Merkmale in den jeweiligen Kamerabildfolgen kann wahlweise anhand eines Harris-Eckendetektors [HS88] oder eines Verfahrens zur Detektion von Reflektionsmarkern erfolgen. Letzteres basiert auf Eigenüberlegungen, welche in Anh. C.1 näher erläutert werden.

Beide Detektoren liefern punktartige Merkmale. Um sie vergleichen zu können, was eine Voraussetzung für die Suche nach korrespondierenden Merkmalen darstellt, sind daher zusätzliche Informationen über die Umgebung notwendig. Diese werden nach dem in [NBN06] beschriebenen Ansatz intern gespeichert. Die Ähnlichkeit zweier Merkmale kann dann unter der Verwendung der *normierten Kreuzkorrelation* bestimmt werden.

3.5.3 Korrespondenzsuche

Ziel der Korrespondenzsuche ist es, bereits detektierte Merkmale sowohl in einem Folgebild als auch im zeitgleich erfassten Bild der anderen Kamera zu finden (Bild 3.5.3). Gelingt dies, so bezeichnet man die beteiligten Merkmale als *Korrespondenzen*. Ein naiver Ansatz wäre, alle detektierten Merkmale paarweise miteinander zu vergleichen. Dies führt im Regelfall zu einem inakzeptablen Aufwand, weshalb an dieser Stelle die Verwendung von geeigneten Strategien zur Einschränkung des Suchraums unerlässlich sind.

Bei der Korrespondenzsuche in zwei aufeinanderfolgenden Bildern einer Kamera, kann i. d. R. eine maximale Verschiebung zwischen den Positionen zweier korrespondierender Merkmale angenommen werden. Anhand dieser lassen sich Suchfenster festlegen, welche die Menge der möglichen Kandidaten für einen Vergleich festlegen. Dabei ist es wichtig, dass die Suche sowohl vorwärts als auch rückwärts durchgeführt wird, wobei nur dieje-

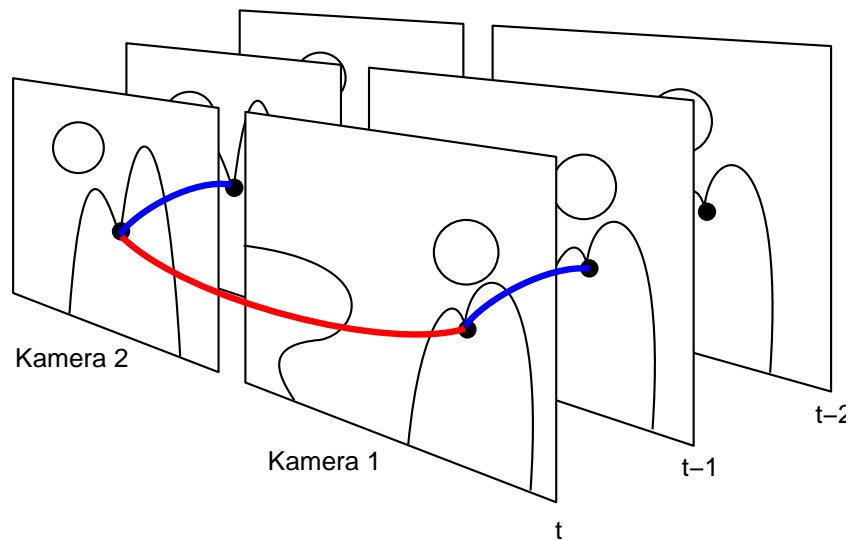


Bild 3.2: Korrespondierende Merkmale in einem Bild.

nigen Merkmale als Korrespondenzen akzeptiert werden, welche in beiden Richtungen einander zugeordnet wurden (Bild 3.5.3). Diese Absicherung wird als *mutual consistency check* bezeichnet [NBN06].

Bei der Korrespondenzsuche zwischen zwei zeitgleich aufgenommenen Bildern der jeweiligen Kameras, kann die Verwendung der Fundamentalmatrix in Betracht gezogen werden. Sie bietet die Möglichkeit, ein Merkmal in einem Bild auf seine *Epipolarlinie* im anderen Bild abzubilden. Die Kandidaten für einen Vergleich können dann anhand ihres Abstands zu der gegebenen Linie ausgewählt werden. Da die *Epipolarbedingung* aufgrund von Rauschen auch bei korrespondierenden Merkmalen im Regelfall nicht erfüllt ist, muss ein angemessener Schwellwert festgelegt werden, welcher den Suchraum ausreichend einschränkt (Bild 3.5.3).

3.5.4 Triangulierung

Damit eine räumliche Vorhersage möglich wird, muss für jeden der ermittelten Korrespondenzen der entsprechende Weltpunkt der Szene rekonstruiert werden. Aufgrund des Rauschens in den Merkmalspositionen, ist dies jedoch nur näherungsweise möglich. Für

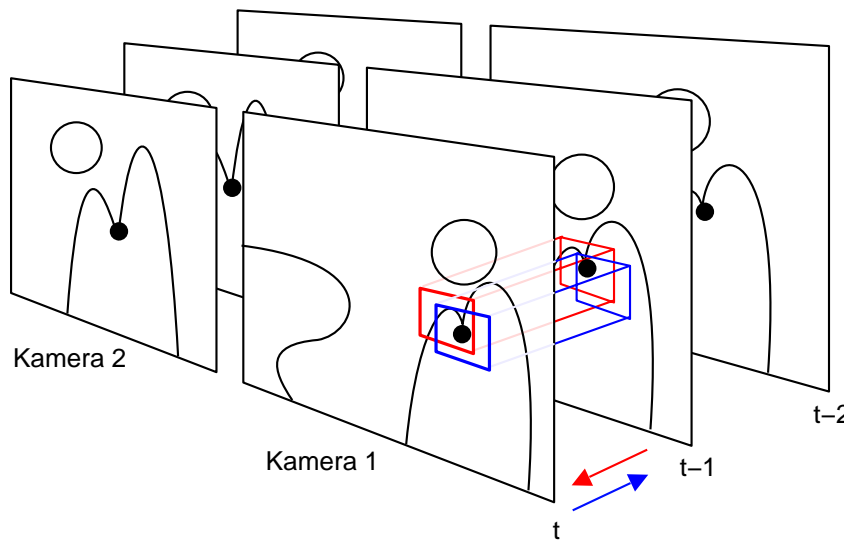


Bild 3.3: Vorwärts- und rückwärtsgerichtete Korrespondenzsuche im Bildstrom einer Kamera.

diesen als *Triangulierung* bezeichneten Vorgang existieren verschiedene Ansätze. Im Rahmen der entwickelten Anwendung wurde ein linearer Ansatz nach [HZ03, S. 312] realisiert.

3.5.5 Vorhersage mittels polynomieller Extrapolation

Wurde eine ausreichend große Menge an Frames verarbeitet, so stehen genug Daten zur Verfügung, um eine Bewegungsvorhersage durchzuführen. Die Menge der Datenpunkte, welche für die Vorhersage verwendet werden soll, kann dabei zu Beginn des Trackingprozesses festgelegt werden. Auch die Schrittweite für die Vorhersage ist variabel.

Nach [PTVF07, S 110f.] lässt sich eine Extrapolation bekannter Daten i. A. anhand der folgenden Teilschritte durchführen:

1. Das Einpassen einer interpolierenden Funktion in die bekannten Daten.
2. Das Ermitteln des gesuchten Wertes durch eine Auswertung der Funktion an der entsprechenden Stelle.

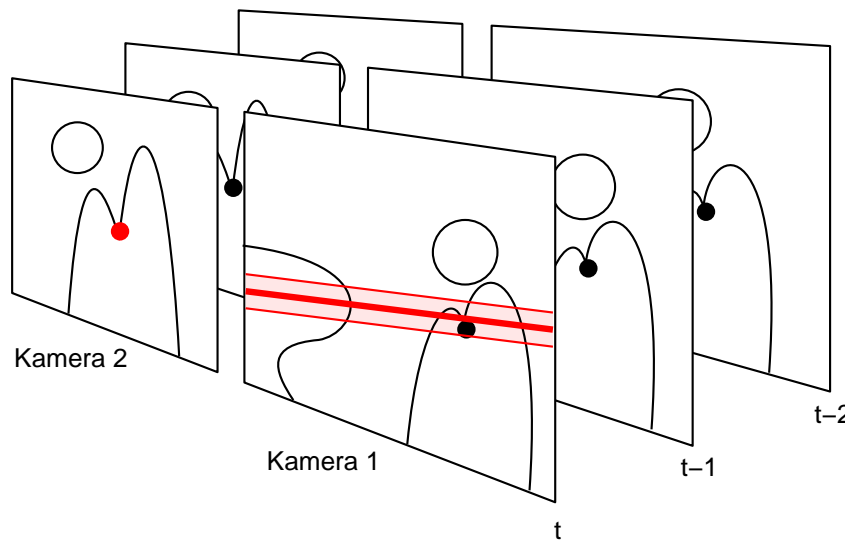


Bild 3.4: Korrespondenzsuche zwischen den Kamerabildern eines Zeitpunktes unter Verwendung der Epipolarlinie.

In dem hier vorliegenden Fall bietet es sich somit an, ein Polynom für die Einpassung in die gegebenen Daten zu verwenden. Dies wird auch als *polynomielle Extrapolation* bezeichnet.

Dieser Ansatz erweist sich jedoch als problematisch, wenn es sich bei den vorhandenen Daten um verrauschte Messungen handelt. Ist die Verteilung des Rauschen jedoch bekannt, lassen sich stochastische Methoden zur Einpassung des Polynoms verwenden. Unter der Annahme, dass ein normalverteiltes Rauschen vorliegt, bietet beispielsweise die Regressionsanalyse [LW04, S. 186] eine solche Möglichkeit. Sie wird in der Regel dazu verwendet, Zusammenhänge zwischen Variablen und somit Trends in einer Menge von gestreuten Daten zu ermitteln. Um dies zu erreichen, wird für eine zu analysierende Trendfunktion ein festgelegtes Fehlermaß minimiert. Üblicherweise wird hierzu die Methode der *Summe der Fehlerquadrate* verwendet.

Die Regressionsanalyse liefert gute Resultate, sofern die gegebenen Voraussetzungen erfüllt sind. Existieren jedoch Ausreißer in den Messdaten, so haben diese, in Abhängigkeit von ihrer relativen Häufigkeit, einen stark verfälschenden Einfluß auf die Ermittlung der Modellparameter (siehe Bild 3.5.5). Eine gesteigerte Robustheit gegenüber Ausreißern ist

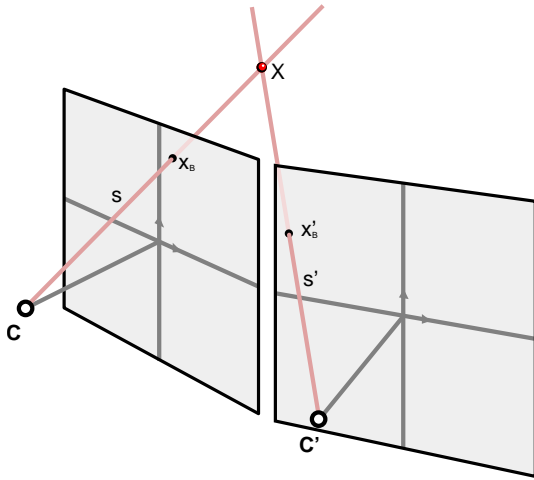


Bild 3.5: Bei Abwesenheit von Rauschen schneiden sich die Strahlen durch die Kamerazentren und Bildpunkte in einem Weltpunkt.

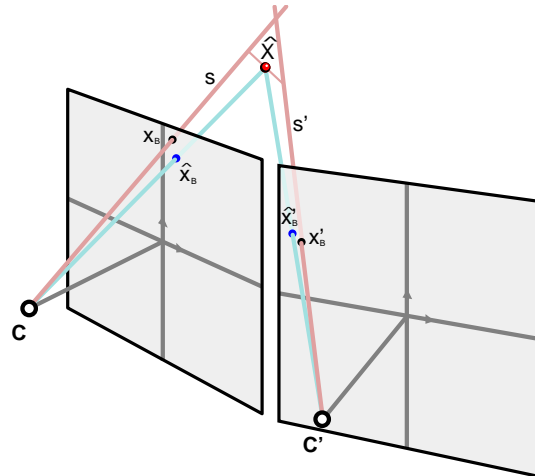


Bild 3.6: Aufgrund von Rauschen muss die Ideallösung angenähert werden.

somit wünschenswert.

Um diese zu erlangen, kann das RANSAC-Verfahren [FB81] eingesetzt werden. Es handelt sich dabei um ein Verfahren, welches zur Bestimmung einer Untermenge der gegebenen Messdaten herangezogen wird. Dabei wird versucht, diese Untermenge so zu bestimmen, dass sie die Ausreißer der Grunddatenmenge nicht enthält und doch eine ausreichend große Menge verwertbarer Messwerte, das sogenannte *Consensus set*, liefert. Das RANSAC-Verfahren kann also als Vorstufe zur Regressionsanalyse eingesetzt werden, welches dann die eigentliche Modelleinpassung vornimmt.

Das Verfahren bestimmt dabei in n Durchläufen jeweils eine zufällige Menge von Datenpunkten aus der Grunddatenmenge. Es werden dabei nur die kleinstmögliche Anzahl an Punkten gewählt, welche zur Parametrisierung des verwendeten Modells gebraucht werden. Für diese als *Inlier* bezeichneten Punktmenge wird angenommen, dass sie Ausreißerfrei ist. Die übrigen Punkte werden subsequent an dem festgelegten Modell getestet. Liegt dabei der Fehler innerhalb einer zuvor festgelegten Toleranzgrenze, so wird der entsprechende Punkt ebenfalls als *Inlier* betrachtet, ansonsten als *Outlier*. Beim ersten Durchlauf

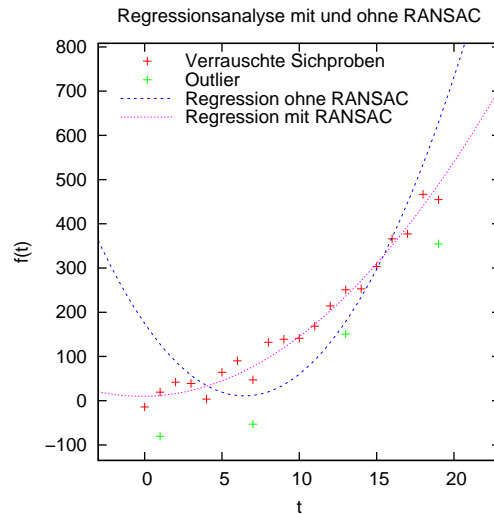


Bild 3.7: Regression mit und ohne RANSAC.

wird die Menge der ermittelten Inlier dann als hypothetisches Consensus set gespeichert. Für jeden weiteren Durchlauf wird, nachdem alle Punkte klassifiziert wurden, geprüft, ob die Anzahl der gefundenen Inlier die Anzahl des bisherigen Consensus sets übersteigt. Ist dies der Fall, so wird das vorherige Consensus set durch die aktuelle Inliermenge ersetzt. Andernfalls wird das vorherige Consensus set beibehalten.

Die Anzahl der Durchläufe n , welche benötigt werden um mit einer Wahrscheinlichkeit p ein ausreißerfreies Consensus set zu finden, lässt sich anhand der Relation

$$n = \frac{\log(1 - p)}{\log(1 - (1 - \epsilon)^s)},$$

festlegen, wobei ϵ für die relative Häufigkeit der Ausreißer und s die Anzahl der Datenpunkte darstellt.

Die Toleranzgrenze dagegen lässt sich nur dann bestimmen, wenn genauere Informationen über die Verteilung bekannt sind. Im Falle einer Normalverteilung kann Basis der *3-Sigma-Regel* behauptet werden, dass $2 \cdot \sigma \leq l \leq 3 \cdot \sigma$ eine gute Wahl darstellt. Ist die Verteilung jedoch unbestimmt, muss l in der Regel empirisch festgelegt werden.

3.5.6 Unterstützte Ein- und Ausgangsformate

Alle erforderlichen Einstellungen müssen dem Tracker zu Beginn in Form einer XML-Datei zur Verfügung gestellt werden. Dadurch wird u. a. eine einfache Reproduktion der Ergebnisse ermöglicht. Eine detaillierte Beschreibung aller Konfigurationsmöglichkeiten wird im Rahmen der Dokumentation zur Software auf der beiliegenden DVD geboten (siehe hierzu Anh. E).

Zum Einlesen der Videodateien wird die OpenCV-Bibliothek als Backend verwendet. Diese unterstützt zum gegebenen Zeitpunkt jedoch ausschließlich das AVI-Format. Des Weiteren können jedoch auch stetig nummerierte Folgen von Bilddateien eines Formats als Quelle verwendet werden.

Die Anwendung unterstützt die direkte Ermittlung verschiedener Messwerte:

- Die Positionen der detektierten Merkmale in den Kamerabildern,
- die zugehörigen Positionen der rekonstruierten Weltpunkte,
- der Reprojektionsfehler dieser Rekonstruktion,
- die Positionen der Vorhergesagten Punkte im Bild sowie
- die entsprechenden Fehler der Vorhersagen (siehe Abschn. 3.4).

Die Messungen können wahlweise als Datei im CSV-, Gnuplot- oder SPSS-Format ausgegeben werden. Zudem ist es möglich, in festgelegten Abständen Bilder der erzeugten Trajektorien zu erhalten. Letzteres kann vor allem als Feedback für die Justierung der Eingangsparameter genutzt werden.

3.5.7 Abhängigkeiten

Bei der Implementierung wurden verschiedene externe Programmbibliotheken verwendet. Diese sind in Tabelle 3.5.7 aufgelistet.

Bibliothek	Version	Wozu verwendet ...
Lapack++	2.5.1	Berechnungen im Rahmen der Linearen Algebra
OpenCV	1.0	Aspekte der Kalibrierung sowie Aspekte der Merkmalsdetektion
TinyXPath	1.0	Einlesen der Konfigurationsdatei

Tabelle 3.1: Abhängigkeiten zu externen Programmbibliotheken.

3.5.8 Verifikation des Trackers

Um die Korrektheit der essentiellen Kernmodule des Trackers zu stützen, wurde eine Reihe von Unit-Tests entwickelt. Diverse Nebenmodule wurden im Rahmen dieser Tests ebenfalls für korrekt befunden. Weitere Informationen hierzu sind in Anh. E.5 gegeben.

Kapitel 4

Evaluierung

Im weiteren Verlauf dieses Kapitels werden die Experimente beschrieben, welche im Rahmen dieser Arbeit durchgeführt wurden, um die in Abschn. 3.3 aufgestellten Thesen zu stützen oder gegebenenfalls zu widerlegen.

4.1 Überblick

Zu Beginn wird der Aufbau der Experimente beschrieben. Betrachtet werden dabei einerseits Experimente mit synthetischen Bildfolgen (Abschn. 4.1.1) und andererseits Experimente mit realen Bildfolgen aus zwei Hochgeschwindigkeitskameras (Abschn. 4.1.2). Abschließend werden die Ergebnisse der Experimente vorgestellt (Abschn. 4.2) und verifiziert (Abschn. 4.3).

4.1.1 Synthetische Bildfolgen

Mittels synthetischer Bildfolgen lässt sich die Vorhersagegenauigkeit unter idealen Bedingungen messen. Ideal bedeutet in diesem Zusammenhang:

- Die erfassten Messungen minimal verrauscht,

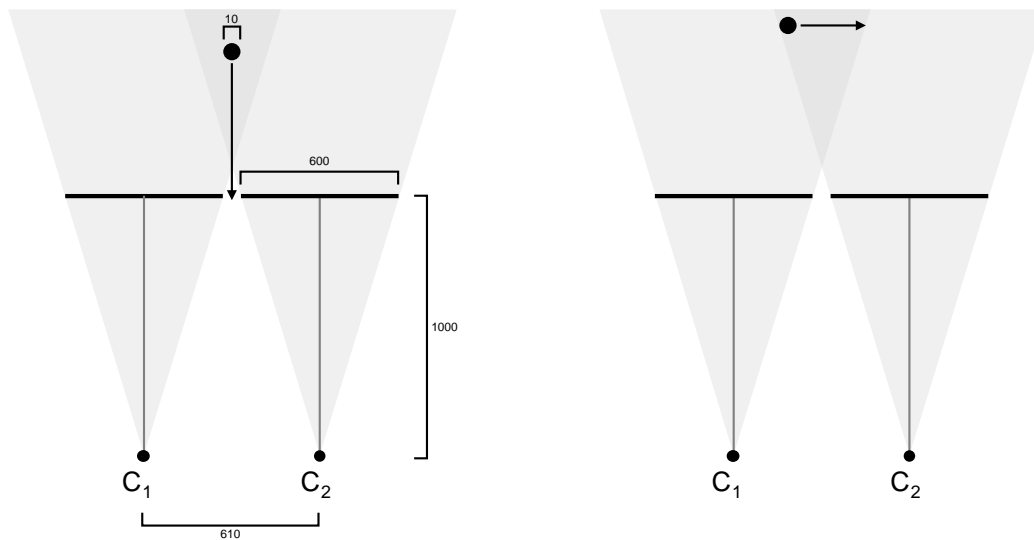


Bild 4.1: Allgemeiner Versuchsaufbau der synthetischen Bildfolgen. Links eine zur Basislinie parallele Bewegung (Versuche 1 bis 4), rechts eine orthogonale Bewegung (Versuche 5 und 6).

- alle Parameter der beiden Kameras sind bekannt,
- die zu messende Trajektorie ist bekannt.

Betrachtet werden sollen hier ausschließlich parallele und orthogonale Bewegungen in Bezug auf die Kameraebenen. In insgesamt 8 Situationen wurde, jeweils für beide Kameraansichten, eine weiße Kugel vor einem schwarzem Hintergrund gerendert, welche über 100 Frames eine gradlinige Bewegung vollführt. Die einzelnen Situationen unterscheiden sich, neben der Bewegungsrichtung, durch die Art der Beschleunigung. Abbildung 4.1.1 illustriert den Versuchsaufbau und Tabelle 4.1.1 gibt die kinematischen Daten an, mit denen die Simulation durchgeführt wurde. Die Parameter der beiden Kameras unterscheiden sich lediglich um eine Translation auf der x-Achse. Damit wird u. a. sichergestellt, dass eine orthogonale Bewegung für beide Bildebenen der Kameras gleichzeitig gilt.

#	Position	Geschwindigkeit	Beschleunigung
1	$\mathbf{X}(t) = (305; 0; 3000)$	$\mathbf{v}(t) = (0; 0; -20)$	$\mathbf{a}(t) = (0; 0; 0)$
2	$\mathbf{X}(t) = (305; 0; 3000)$	$\mathbf{v}(t) = (0; 0; 0)$	$\mathbf{a}(t) = (0; 0; -4 \cdot 10^{-1})$
3	$\mathbf{X}(t) = (305; 0; 3000)$	$\mathbf{v}(t) = (0; 0; 0)$	$\mathbf{a}(t) = (0; 0; -1,3 \cdot 10^{-2} \cdot t)$
4	$\mathbf{X}(t) = (305; 0; 3000)$	$\mathbf{v}(t) = (0; 0; 0)$	$\mathbf{a}(t) = (0; 0; -2,5 \cdot 10^{-4} \cdot t^2)$
5	$\mathbf{X}(t) = (20; 0; 2000)$	$\mathbf{v}(t) = (5,9; 0; 0)$	$\mathbf{a}(t) = (0; 0; 0)$
6	$\mathbf{X}(t) = (20; 0; 2000)$	$\mathbf{v}(t) = (0; 0; 0)$	$\mathbf{a}(t) = (1,2 \cdot 10^{-1}; 0; 0)$
7	$\mathbf{X}(t) = (20; 0; 2000)$	$\mathbf{v}(t) = (0; 0; 0)$	$\mathbf{a}(t) = (3,5 \cot 10^{-3} \cdot t; 0; 0)$
8	$\mathbf{X}(t) = (20; 0; 2000)$	$\mathbf{v}(t) = (0; 0; 0)$	$\mathbf{a}(t) = (7 \cdot 10^{-5} \cdot t^2; 0; 0)$

Tabelle 4.1: Die kinematischen Daten zu den 8 synthetischen Bildfolgen.

4.1.2 Reale Bildfolgen

Um die Modelle an realen Bildfolgen zu testen, wurden mehrere Aufnahmen mit zwei Hochgeschwindigkeitskameras (siehe Bild 4.1.2) durchgeführt. Im Gegensatz zu den synthetischen Aufnahmen, muss hierbei mit einem verstärkten Rauschen in den erfassten Merkmalskoordinaten gerechnet werden. Weitere Fehler lassen sich zudem bei der Bestimmung der Kameraparameter erwarten.

Auch bei den realen Aufnahmen wurde versucht, Situationen mit parallelen und orthogonalen Bewegungen zu erfassen. Da eine ideale Aufstellung der Kameras ohne Hilfestellung nicht garantiert werden kann und zudem der geringe *Bildwinkel* der verwendeten Objektivs eine sehr geringe Distanz der Kameras erfordert hätte, wurden die Kameras stattdessen leicht gedreht aufgestellt. Anstatt der Kameraebenen wurden die orthogonalen Bewegungen in Bezug auf die Basislinie der Kameras durchgeführt. Des Weiteren kann auch eine ideale Beschleunigung, wie im Falle der synthetischen Aufnahmen, nicht garantiert werden. Stattdessen wurden prototypische Bewegungsabläufe aufgenommen. Diese lassen sich grob in fließende Bewegungen, wie z. B. Laufen oder seitliches Ausweichen, und gestoppte Bewegungen, wie z. B. ein Tritt oder Abbremsen, unterteilen. Dabei war es notwendig, auf Ganzkörperaufnahmen zu verzichten, da bei einer Aufnahme mit hoher Bildfrequenz eine ausreichende Beleuchtung erforderlich ist. Durch eine erhöhte Distanz der Kameras zur Szene, wäre die Beleuchtung ebenfalls im Bild zu sehen gewesen. Dies



Bild 4.2: Eine der verwendeten Hochgeschwindigkeitskameras.

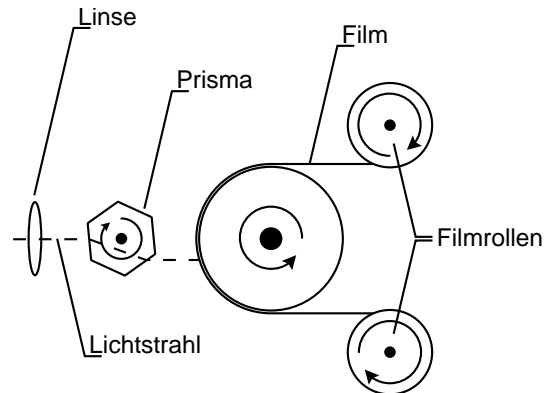


Bild 4.3: Schematischer Aufbau einer Hochgeschwindigkeitskamera mit Hochgeschwindigkeitsfilm [HN99, S. 303].

wurde vermieden, um eventuelle Störungen bei den nachfolgenden Messungen zu vermeiden. Oberkörper und Unterkörper wurden stattdessen separat betrachtet.

Um ausgeprägten Schatten entgegenzuwirken, wurde der Hintergrund der Szene vor den Aufnahmen mithilfe zweier schwarzer Kulissen und einer schwarzen Decke für den Boden präpariert. Zur Beleuchtung der Szene wurden verschiedene Lampentypen eingesetzt. Eine ausreichend hohe Leistung ist wegen der sehr kurzen Belichtungszeit ebenso eine Voraussetzung, wie eine hohe *Flimmerfrequenz* der Lampen. Die Leistung der eingesetzten Lampen beträgt 500 W bzw. 1.500 W.

Es wurden Vorversuche und zwei Hauptversuchsreihen durchgeführt. Zur Ermittlung der Kameraparameter wurden in beiden Versuchsreihen verschiedene Methoden angewandt. Vor der ersten Hauptversuchsreihe wurden die intrinsischen Kameraparameter, unter Zuhilfenahme eines 3d-Kalibrierungsmusters (siehe Bild 4.6), getrennt von den extrinsischen Kameraparametern ermittelt. Für die Ermittlung der intrinsischen Kameraparameter wurden jeweils 31 Punktkorrespondenzen verwendet, für die Ermittlung der Fundamentalmatrix 50. Alle Korrespondenzen wurden dabei händisch ausgelesen.

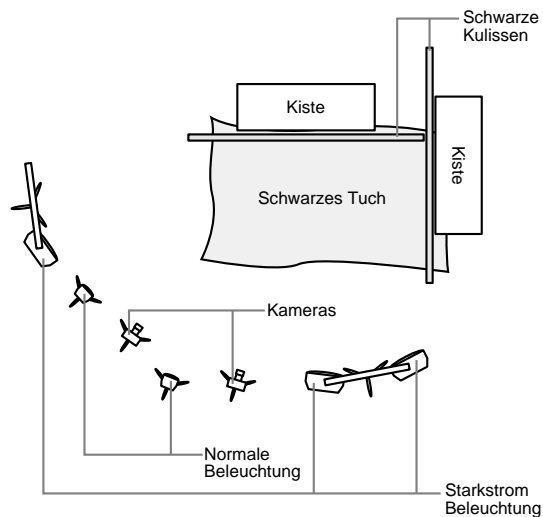


Bild 4.4: Skizze der Aufbauten für die Aufnahme realer Bewegungen.



Bild 4.5: Foto der Aufbauten während der 1. Versuchsreihe.

Die Ermittlung der Kameraparameter der zweiten Hauptversuchsreihe erfolgte vorab anhand mehrere Aufnahmen eines Schachbrett-Kalibrierungsmusters (siehe Bild 4.7). Anhand dieser Aufnahmen konnten insgesamt 576 Korrespondenzen für die linke Kamera und 720 Korrespondenzen für die rechte Kamera automatisiert ermittelt werden.

Die Kameras selbst wurden für Aufnahmen mit einer Bilderfassungsfrequenz von 500 Hz konfiguriert. Diese relativ hohe Frequenz, im Vergleich mit gewöhnlichen Kameras, resultiert aus dem Einsatz von Prismen. Diese brechen das einfallende Licht auf einen laufenden Film oder einen Sensorchip, wodurch sich das sonst erforderliche Stoppen des Films, vor dem öffnen des Verschlusses, vermeiden lässt (siehe Bild 4.3). Weitere Daten zu den verwendeten Kameras und den Objektiven, lassen sich den Tabellen Tabelle 4.1.2 und Tabelle 4.1.2 entnehmen.

Für einen großen Teil der Aufnahmen wurde die erfasste Person mit flachen, kreisförmigen Markern aus gebleichtem Papier an den Gelenken prepariert. Die übrigen Aufnahmen sind ausschließlich für markerfreies Tracking geeignet.

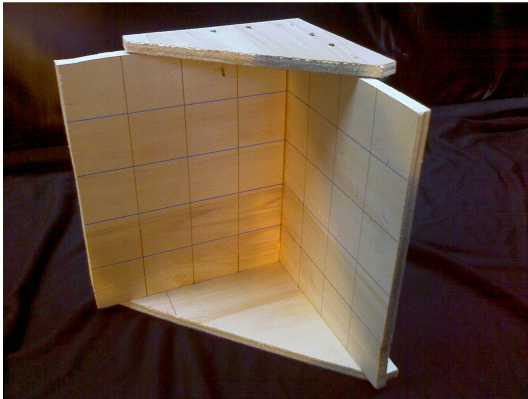


Bild 4.6: Ein 3-D Kalibrierungsmuster.

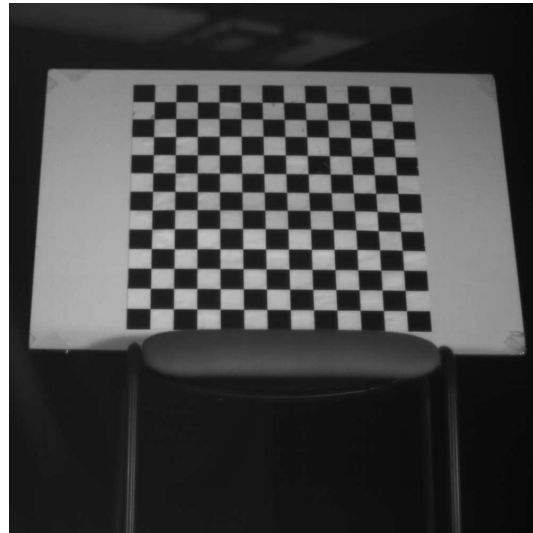


Bild 4.7: Ein planares Kalibrierungsmuster.

4.2 Validierung der Ergebnisse

Im Folgenden werden die Ergebnisse der zuvor beschriebenen Experimente dargestellt und bewertet.

4.2.1 Validierung der synthetischen Bildfolgen

Die Bildsequenzen wurden mit dem entwickelten Tracker unter der Verwendung verschiedener Konfigurationen ausgewertet. Da keine Ausreißer zu erwarten sind, wurden die Vorhersagen zuerst mit dem regressionsbasierten Verfahren mit einer Historie von 20 Frames durchgeführt. Weiterhin wurden die Vorhersagen unter der Annahme einer konstanten Beschleunigung und für einen einzelnen Zeitschritt in die Zukunft durchgeführt.

Bild 4.2.1 zeigt die Entwicklung des Fehlers bei einer Auswahl dieser Aufnahmen. Zum einen lässt sich erkennen, dass ein erwartungsgerechter Zusammenhang zwischen dem Grad der Beschleunigung und den resultierenden Fehlern besteht. Die Auswirkung auf den Vorhersagefehler ist jedoch bei parallelen Bewegungen sehr gering, auch wenn ein Be-

Kamerahersteller	Basler
Kameramodell	A504k
Sensorhersteller	Micron
Sensormodell	MI-MV13 CMOS
Auflösung	1280 × 1024 Pixel
Sensorzellendim.	12 μm × 12 μm
Sensordim.e	15,36 mm × 12,29 mm
Verschuß	0,1 ms bis 33 ms
Bildrate bei voller Auflösung	500 Hz
Bildrate maximal	16000 Hz

Tabelle 4.2: Spezifikationen der verwendeten Kameras.

Hersteller	Nikon
Brennweite	35 mm
Bildwinkel	62°
Lichtstärkeverhältnis	1:1,4

Tabelle 4.3: Spezifikationen der verwendeten Objektive.

schleunigungsgrad > 2 vorliegt. Zum anderen lassen sich Beobachtungen machen, welche die Thesen aus Abschn. 3.3 stützen. Im Vergleich zwischen 2d- und 3d-Vorhersage bei den orthogonalen Bewegungen, ist die 3d-Vorhersage zeitweise um ein Vielfaches besser als die Vorhersage im Bild, wobei die Vorhersagen im parallelen Fall ungefähr gleich gute Ergebnisse erzielen.

Als Kennzahl für den mittleren Fehler, kann der Erwartungswert E der gegebenen Werteverteilung dienen. Betrachtet man die einzelnen Verteilungen in den Diagrammen, so scheint eine Exponentialverteilung als Modell für die tatsächliche Verteilung geeignet zu sein. Der Erwartungswert einer Exponentialverteilung genügt der Relation $E(X) = \frac{\tilde{x}}{\ln 2}$, wobei \tilde{x} den Median der Verteilung repräsentiert. Für die Standardabweichung σ gilt des weiteren $\sigma(X) = E(X)$.

Die Erwartungswerte der in Bild 4.2.1 dargestellten Messungen sind in Tabelle 4.2.1 auf-

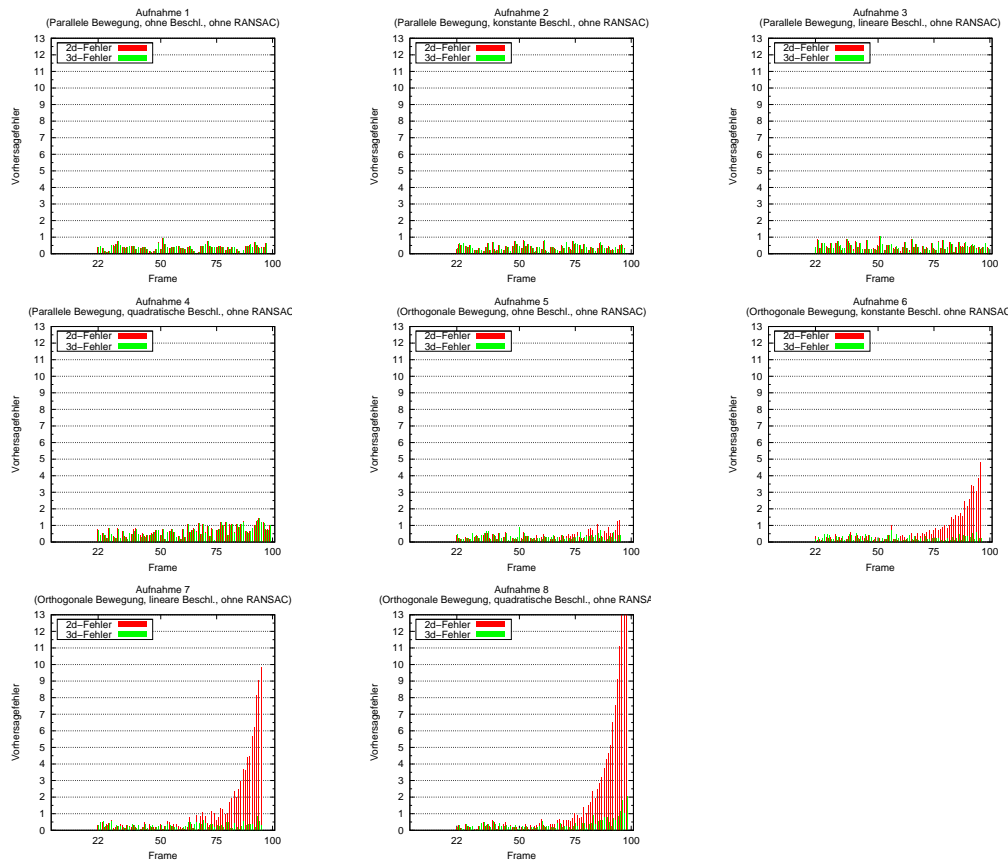


Bild 4.8: Die Messungen der Vorhersagefehler bei den synthetischen Bildfolgen.

gelistet. Sie stützen die Feststellung, dass die 3d-Vorhersage für die erfolgten Messungen genauer ist. Jedoch zeigen sie auch, dass die zu erwartenden Fehler beider Verfahren im Vergleich nur geringe Abweichungen aufzeigen.

Neben den Messungen mit der oben genannten Standardkonfiguration wurden weitere Konfigurationen getestet. Diese zusätzlichen Messungen wurden ausschließlich für die beiden Aufnahmen mit quadratischer Beschleunigung durchgeführt, da hier die Entwicklung der Vorhersagefehler am deutlichsten zu erkennen ist. Es wurden verschiedene Variationen der Bildrate, der Menge der Stützpunkte, des verwendeten Polynomgrades und der Zeitschritte für die Vorhersage getestet. Zudem wurde der RANSAC-Algorithmus hinzugenommen, um festzustellen, inwiefern sich dieser auf die Genauigkeit auswirkt. Da-

Messung	$E(X)$ (2d)	$E(X)$ (3d)
4.2.1.1	0,56	0,56
4.2.1.2	0,53	0,53
4.2.1.3	0,58	0,59
4.2.1.4	0,89	0,89
4.2.1.5	0,48	0,39
4.2.1.6	0,51	0,30
4.2.1.7	0,56	0,41
4.2.1.8	0,56	0,48

Tabelle 4.4: Erwartungswerte der Messungen aus Bild 4.2.1.

Abbildung	Beschreibung
4.2.1	Die Anzahl der Stützpunkte wurde variiert
4.2.1	Die Framerate wurde halbiert
4.2.1	Der Grad des Modellpolynoms und die Anzahl der Stützpunkte wurden variiert
4.2.1	Die Schrittweite bei der Vorhersage wurde auf 3 erhöht
4.2.1	Der RANSAC-Algorithmus wurde hinzugenommen

Tabelle 4.5: Eine Beschreibung der Parametervariationen bei den vom Standard abweichenden Messungen der synthetischen Bildfolgen.

mit einer eindeutigen Rekonstruierbarkeit der Ergebnisse gewährleistet ist, wurde bei dem RANSAC-basierten Vorhersageverfahren auf eine zufällige Auswahl des initialen Consensus sets verzichtet. Stattdessen werden stets alle möglichen Kombinationen für das initiale Consensus set verwendet. Eine Auflistung der hier relevanten Konfigurationen befindet sich in Tabelle 4.2.1.

Das Ergebnis der Messungen mit der variierten Anzahl an Stützpunkten wird in Bild 4.2.1 dargestellt. Da durch eine Verringerung der Stützpunktmenge auch die betrachtete Zeitspanne verkleinert wird, gelingt es sichtlich besser, die Modellfunktion in die Messungen einzupassen. Ein entgegengesetzter Effekt setzt erwartungsgemäß bei einer Erhöhung der Stützpunkte ein. Anhand dieser Beobachtungen darf jedoch nicht geschlossen werden,

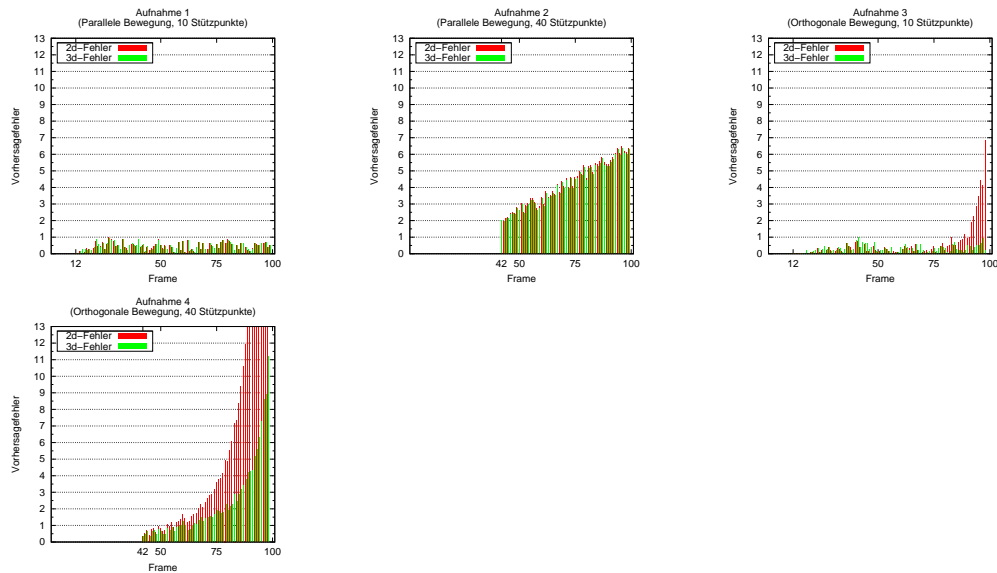


Bild 4.9: Die Messungen der Vorhersagefehler bei den synthetischen Bildfolgen unter der Verwendung von unterschiedlich vielen Stützpunkten.

dass eine Verringerung der Stützpunkte bis auf ein Minimum die Vorhersage stetig verbessert. In realen Aufnahmen ist es wahrscheinlich, dass die Messdaten stärker verrauscht sind, als es bei den betrachteten Bildfolgen der Fall ist. Eine ausreichend hohe Menge an Stützpunkten ist dann notwendig, um die Vorhersage zu stabilisieren.

Bei der Betrachtung der zugehörigen Erwartungswerte (siehe Tabelle 4.2.1) lässt sich auch hier feststellen, dass diese sich zum Großteil kaum unterschieden. Lediglich der Erwartungswert der Messung 4.2.1.4 liegt bei einer räumlichen Vorhersage deutlich unter dem

Messung	$E(X)$ (2d)	$E(X)$ (3d)
4.2.1.1	0,68	0,67
4.2.1.2	5,97	5,84
4.2.1.3	0,40	0,35
4.2.1.4	3,45	2,10

Tabelle 4.6: Erwartungswerte der Messungen aus Bild 4.2.1.

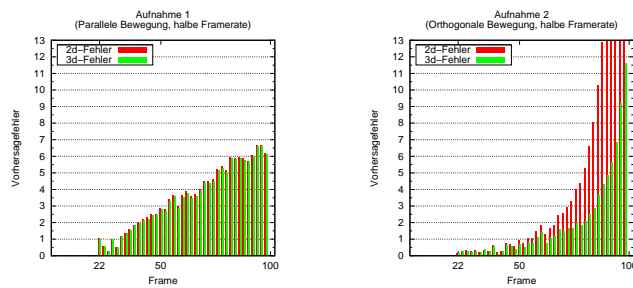


Bild 4.10: Die Messungen der Vorhersagefehler bei den synthetischen Bildfolgen mit halbierte Framerate.

Messung	$E(X)$ (2d)	$E(X)$ (3d)
4.2.1.1	5,28	5,13
4.2.1.2	2,14	1,55

Tabelle 4.7: Erwartungswerte der Messungen aus Bild 4.2.1.

Erwartungswert einer 2d-Vorhersage.

In Bild 4.2.1 sind Messungen mit einer halbierten Framerate dargestellt. Vergleicht man diese mit den Messungen aus Bild 4.2.1, so kann man eine deutliche Verschlechterung der Vorhersagegenauigkeit erkennen. Im Umkehrschluss kann wiederum gefolgert werden, dass eine höhere Dichte an Bildern in einem festen Zeitrahmen, sich positiv auf die Vorhersage auswirken wird.

Auch hier liefert ein Blick auf die Erwartungswerte ähnliche Erkenntnisse, wie bei den vorangegangenen Messungen. Die in Tabelle 4.2.1 aufgelisteten Werte unterschieden sich bei einer parallelen Bewegung kaum, weichen jedoch im orthogonalen Fall sichtlich von einander ab.

Die Ergebnisse der Messung mit einem Modellpolynoms eines höheren Grades, sind in Bild 4.2.1.3 und Bild 4.2.1.4 dargestellt. Diese sind insofern interessant, als das sie der intuitiven Annahme, ein höherer Polynomgrad impliziere auch eine bessere Vorhersagegenauigkeit, widersprechen. Die Annahme bestätigt sich zwar teilweise im Falle der 2d-Vorhersagen (vergleiche Bild 4.2.1.3 bis Bild 4.2.1.6 mit den entsprechenden Unterabbildungen aus Bild 4.2.1), der hohe Durchschnittsfehler bei der 3d-Vorhersagen spricht

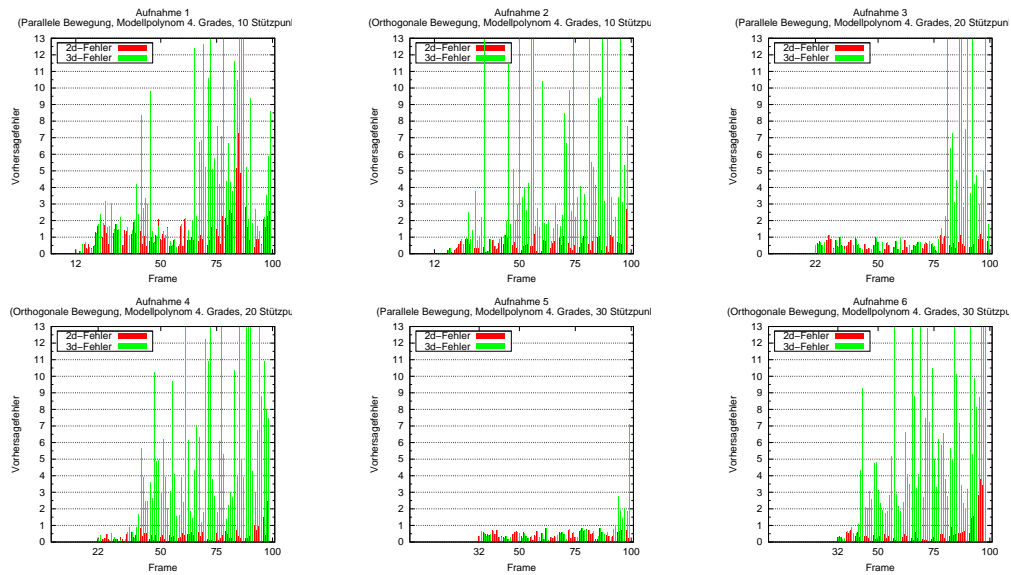


Bild 4.11: Die Messungen der Vorhersagefehler bei den synthetischen Bildfolgen unter der Verwendung eines Modellpolynoms 4. Grades und unterschiedlich vielen Stützpunkten.

jedoch klar gegen sie.

Die Tatsache, dass dieses Phänomen in erster Linie die 3d-Vorhersagen betrifft, legt nahe, dass ein zusätzliches Rauschen in den Rekonstruierten Daten ein stärkeres Oszillieren des Modellpolynoms bewirkt. Aus diesem Grund wurden, neben der Standardanzahl an Stützpunkten (20 Punkte), weitere Variationen der Stützpunktzahl gemessen. Die Ergebnisse dieser Messungen zeigen einen positiven Zusammenhang zwischen der Erhöhung der Stützpunktmenge und der Genauigkeit im Fall der Parallelbewegung (siehe Bild 4.2.1). Bei der orthogonalen Bewegung lässt sich dagegen kaum eine Veränderung beobachten, eher eine leichte Verschlechterung der Ergebnisse.

Bei einem Blick auf die Genauigkeit der Rekonstruktion zieht eine starke Ungenauigkeit bez. der z-Koordinate des rekonstruierten Weltpunktes (siehe Bild 4.2.1). Die beiden anderen Koordinaten weisen dagegen eine Genauigkeit von unter einer Raumeinheit auf.

Auf einen Vergleich der Erwartungswerte der in Bild 4.2.1 dargestellten Messergebnisse, wird an dieser Stelle verzichtet. Stattdessen soll an dieser Stelle ein Vergleich der 2d-Vorhersagen Bild 4.2.1.3 und Bild 4.2.1.4 mit den den räumlichen Vorhersagen Bild 4.2.1.4

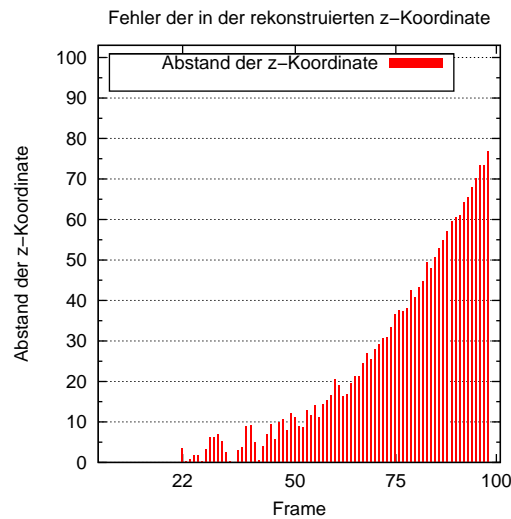


Bild 4.12: Darstellung des Rauschens in der z-Koordinate des rekonstruierten Weltpunkts zu Bild 4.2.1.

und Bild 4.2.1.8 durchzuführen werden. In Bild 4.2.1 sind die Messungen einander gegenübergestellt. An dieser Stelle fällt eine auffällig hohe Vergleichbarkeit auf. Diese lässt sich jedoch erklären. Betrachtet man Bild 4.2.1, so erkennt man die Möglichkeiten, welche ein Polynom 3. Grades für eine Approximation der echt gebrochen-rationalen Bewegungsfunktion bietet. Alleine ein Polynom 2. Grades würde für eine angemessene Näherung nicht ausreichen.

Bei der Betrachtung der Erwartungswerte der Verteilungen (Tabelle 4.2.1) lässt sich erkennen, dass die zu erwartenden Fehler in beiden Fällen für eine Verwendung einer 2d-

Messung	$E(X)$ (2d)	$E(X)$ (3d)
4.2.1.3	0,77	
4.2.1.4		0,89
4.2.1.4	0,51	
4.2.1.8		0,48

Tabelle 4.8: Erwartungswerte ausgewählter Messungen aus Bild 4.2.1. und Bild 4.2.1.

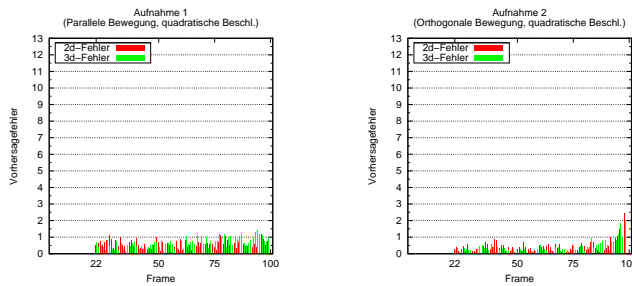


Bild 4.13: Vergleich ausgewählter Messungen aus Bild 4.2.1. und Bild 4.2.1.

Messung	$E(X)$ (2d)	$E(X)$ (3d)
4.2.1.1	1,61	1,57
4.2.1.2	1,08	0,57

Tabelle 4.9: Erwartungswerte der Messungen aus Bild 4.2.1.

Vorhersage sprechen. Es lässt sich vermuten, dass das geringfügig schlechtere Ergebnis der räumlichen Vorhersage auf das erhöhte Rauschen (aufgrund der zusätzlichen Rekonstruktion) zurückgeführt werden kann.

Das Verhalten des Fehlers, im Falle einer erhöhten Vorhersageschrittweite, wird in Bild 4.2.1 dargestellt. Erwartungsgemäß fällt die Genauigkeit der Vorhersage schlechter aus, als bei der Vorhersage für 1 Frame. Jedoch liegt sie noch in einem akzeptablen Rahmen.

Ein Blick auf die Erwartungswerte (Tabelle 4.2.1) ermöglicht auch hier vergleichbare Erkenntnisse in Bezug auf die vorangegangenen Messungen. Die zu erwartenden Fehler bei der parallelen Bewegung sind beinahe gleich, wobei eine überlegene Genauigkeit der räumlichen Vorhersage bei der orthogonalen Bewegung vorherrscht.

Die letzten Messungen an synthetischen Bildfolgen, wurden unter der Zuhilfenahme des RANSAC-Verfahrens durchgeführt. Motiviert wurde der Versuch durch die Überlegung, ob sich das Wegfallen von Datenpunkten positiv oder negativ auf die Einpassung eines Modellpolynoms in eine eher ungeeignete Datenmenge auswirkt. Dabei wurde der Fehlerparameter so gewählt, dass gegen Ende der Bildfolge nur noch die halbe Stützpunktmenge als Inlier gilt. Die Ergebnisse sind in Bild 4.2.1 dargestellt. Man kann erkennen, dass der Einfluß des Algorithmus einen negativen Effekt auf die Vorhersage hat.

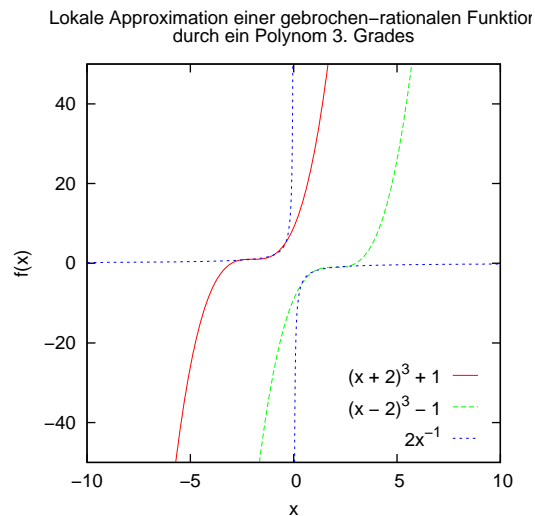


Bild 4.14: Lokale Approximation einer echt gebrochen rationalen Funktion durch ein Polynom 3. Grades.

Messung	$E(X)$ (2d)	$E(X)$ (3d)
4.2.1.1	0,56	0,48
4.2.1.2	0,89	0,96

Tabelle 4.10: Erwartungswerte der Messungen aus Bild 4.2.1.

Die Betrachtung der Erwartungswerte (Bild 4.2.1) zeigt auch hier durchaus vergleichbare Werte bei den zu erwartenden Fehlern. Allerdings ist in diesem Spezialfall die 2d-Vorhersage im orthogonalen Fall geringfügig besser.

Als vorläufiges Fazit, basierend auf den bisher durchgeführten Vergleichen, kann an dieser Stelle festgehalten werden, dass die räumliche Vorhersage, der Vorhersage in den Bildern im direkten Vergleich fast durchgängig überlegen war. Der Vergleich Bild 4.2.1 hat gezeigt, dass im Falle unterschiedlicher Modellpolynome auch gleichwertige Ergebnisse bei einer Bewegungsvorhersage im Bild erzielt werden können. Die Thesen aus Abschn. 3.3 wurden weitestgehend von den Messungen gestützt.

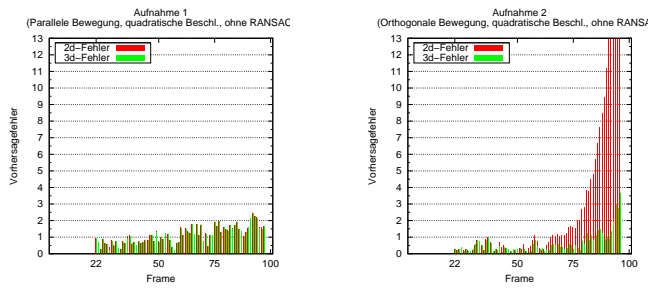


Bild 4.15: Die Messungen der Vorhersagefehler bei den synthetischen Bildfolgen mit einer Vorhersageschrittweite von 3 Frames.

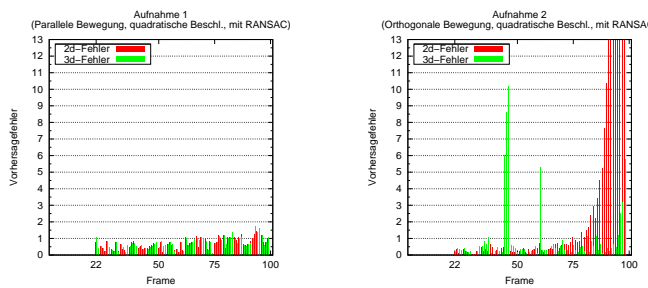


Bild 4.16: Die Messungen der Vorhersagefehler bei den synthetischen Bildfolgen mit RANSAC.

4.2.2 Validierung der realen Aufnahmen

Zuerst sollen im Folgenden die Messergebnisse der Aufnahmen mit Markern betrachtet werden. Da hierbei nicht mit Ausreißern gerechnet werden muß, wurde auch hier auf die Verwendung des RANSAC-Verfahrens bei den Messungen verzichtet.

In Bild 4.2.2 werden die Trajektorien der aufgenommenen Trittbewegung dargestellt. Diese ist im Rahmen der 1. Aufnahmeserie entstanden. Es fällt auf, dass sich die Trajektorie der 3d-Vorhersage von der Trajektorie der erfassten Markerpunkte um einen Versatz von ca. 80 Pixel unterscheidet. Die Ursache für diesen hohen Fehler lässt sich in durch eine ungenügende Genauigkeit der ermittelten Kameraparameter finden. In der 2. Aufnahmeserie wurden bei der Kalibrierung erheblich bessere Ergebnisse erzielt, jedoch ist auch der hier der Fehler nicht unerheblich (siehe hierzu Bild 4.2.2).



Bild 4.17: Die Trajektorien einer Trittbewegung.

Ein absoluter Vergleich der Vorhersagefehler ist, aufgrund der gegebenen Umstände, nicht sinnvoll. Auch eine Betrachtung des methodenspezifischen Verhaltens, bei diversen Variationen der Eingangsparameter, wird durch den zusätzlich einwirkenden Fehler zu sehr verfälscht. Man vergleiche hierzu die Abbildungen Bild 4.2.2 und Bild 4.2.2. Während bei der 2d-Vorhersage eine stetige Verschlechterung der Vorhersage zu erkennen ist, nimmt die Genauigkeit bei der 3d-Vorhersage dagegen sogar etwas zu. Betrachtet man jedoch zusätzlich den Reprojektionsfehler des triangulierten Weltpunktes in Bild 4.2.2, so erkennt man eine geringe, jedoch stetige Steigerung der Rekonstruktionsgenauigkeit über die Zeit. Die dadurch entstandene zusätzliche Beschleunigung in Richtung der tatsächlichen Markerpositionen wirkt sich in diesem Fall positiv auf die Vorhersagegenauigkeit aus. Für weitergehende Erkenntnisse, bezüglich des Verhaltens bei einer guten Kalibrierung, reicht die Datengrundlage somit nicht aus.

Aus dem Verhalten des Fehlers in Bild 4.2.2 kann jedoch ein verwertbares Erkenntnis gewonnen werden. War bei den Messungen der synthetischen Bildfolgen vor allem die orthogonale Bewegung ein Schwachpunkt der 2d-Vorhersage, so erkennt man bei den realen Messungen kaum eine signifikante Verschlechterung. Gründe hierfür lassen sich einerseits in der vergleichsweise sehr hohen Bildrate, andererseits aber auch in dem notwendigen Mindestabstand zu den Bildebenen erkennen. Wurde die Kugel in den synthetischen Bildfolgen bis an die Bildebenen herangeführt, so muss bei realen Aufnahmen ein bestimmter Abstand eingehalten werden, damit eine ausreichende Tiefenschärfe gewährleistet ist.

Ein weiteres Erkenntnis, welches anhand der Messdaten gewonnen werden kann, ist das

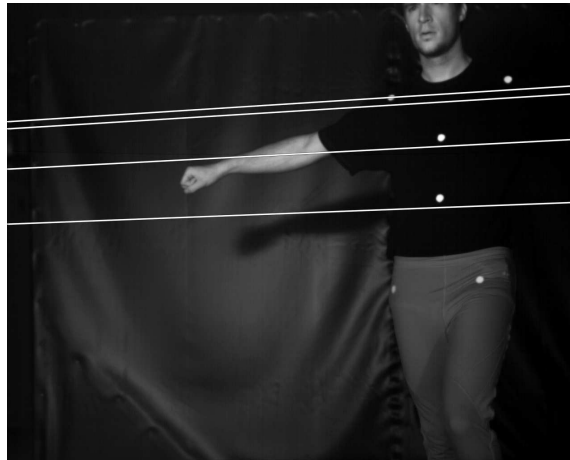


Bild 4.18: Bild einer zur Basislinie orthogonalen Gangbewegung. Zusätzlich sind die Epipolarlinien zu den Markerpunkten im korrespondierenden Kamerabild zu sehen.

Polynomgrad	2d-Vorhersage	3d-Vorhersage
2	0,7	12,7
3	2,0	12,9
4	7,7	14,6

Tabelle 4.11: Mittelwerte zu den Messungen in Tabelle 4.2.2.

Verhalten der Vorhersagegenauigkeit bei Einsatz eines höhergradigen Modellpolynoms. Die Abbildungen Bild 4.2.2.1 bis Bild 4.2.2.3 zeigen einen direkten Vergleich der Vorhersagegenauigkeit unter dem Einsatz verschiedener Polynomgrade. Anders als bei den synthetischen Bildfolgen verschlechtert sich die Genauigkeit der 2d-Vorhersage auffällig, während die Genauigkeit der 3d-Vorhersage nur leicht sinkt. Ein weiterer Blick auf die zugehörigen Mittelwerte (Tabelle 4.2.2) bestätigt die Beobachtung. Der Einsatz höhergradiger Polynome wirkt sich, in dem betrachteten Fall, aber grundsätzlich negativ aus.

Auch für diese Messungen der markerlosen Aufnahmen besteht das Problem der Ungenauigkeiten in den ermittelten Kameraparametern weiter fort. Da eine robuste Korrespondenzsuche, wie sie in Abschn. 3.5.3 illustriert wurde, auf eine möglichst genaue Bestimmung

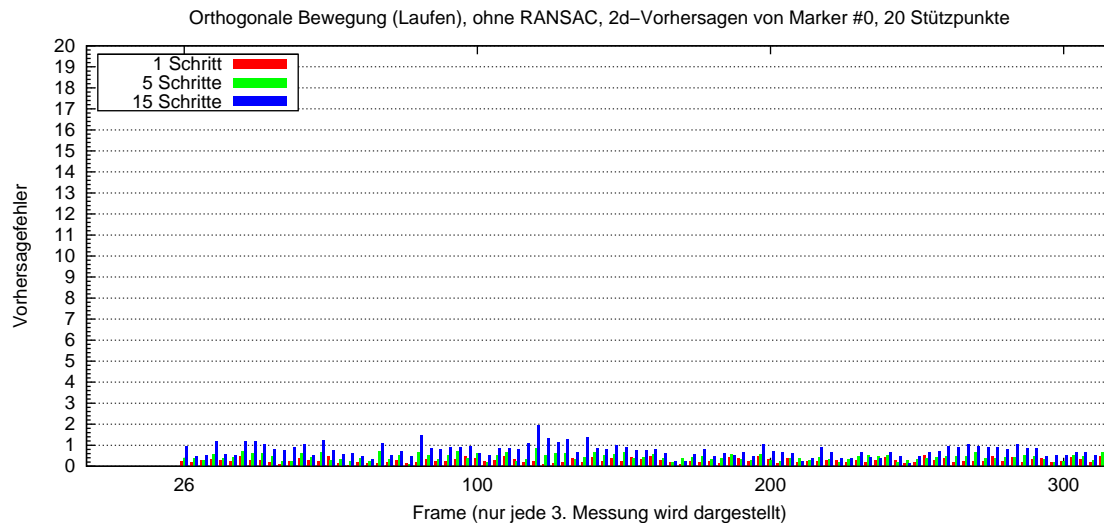


Bild 4.19: Fehlerwerte bei der 2d-Vorhersage einer zur Basislinie orthogonalen Gangbewegung.

dieser angewiesen ist, muss bei den Messungen der Aufnahmen ohne Marker mit einem erhöhten Vorkommen von Fehlzuordnungen, gerechnet werden. Auch hier erscheint ein direkter Vergleich der Fehlerwerte nicht sinnvoll. Als Erkenntnis kann lediglich festgehalten werden, dass ein Einsatz von Vorhersagealgorithmen bei der Verwendung von natürlichen Merkmalen wohlüberlegt sein sollte. Bei einem 2d-Vorhersage von einem Zeitschritt wurde bei der in Bild 4.2.2 dargestellten Messung ein mittlerer Fehlerwert von ca. 13, bei der gemeinsamen Betrachtung aller Merkmale, ermittelt. Für einzelne Merkmale kann jedoch auch eine vergleichsweise hohe Genauigkeit erzielt werden. Die Bewegungsvorhersage von Merkmal 1 aus der selben Aufnahme, erzielte beispielsweise einen mittleren Fehlerwert von 11 (Frames 41 bis 265) bei der Vorhersage von 10 Zeitschritten. Dieser entsprechende Abweichung von durchschnittlich 5,5 Pixel in jedem Bild kann gegebenenfalls als vertretbar angesehen werden.

Im Rahmen dieses Kapitels wurde nur ein Teil der erstellten Hochgeschwindigkeitsaufnahmen vorgestellt. Da die verwertbaren Messungen der übrigen Aufnahmen jedoch keine weiteren Erkenntnisse bezüglich des bisher durchgeführten Vergleichs bieten, wurde auf

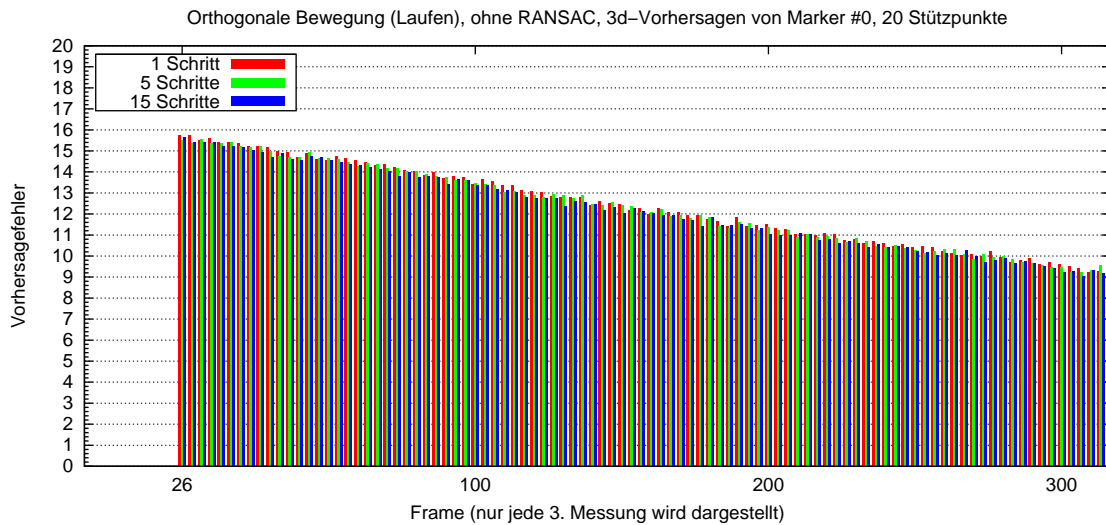


Bild 4.20: Fehlerwerte bei der 3d-Vorhersage einer zur Basislinie orthogonalen Gangbewegung.

eine Darstellung dieser Messergebnisse verzichtet. Zudem musste die Verwendung vieler Messergebnisse, aufgrund der unzureichenden Kalibrierung im Rahmen der ersten Aufnahmeserie, per se ausgeschlossen werden. Die Aufnahmen befinden sich jedoch, zusammen mit den zuvor vorgestellten Aufnahmen, auf den beiliegenden Datenträgern (siehe Anh. E).

Obwohl die vorgestellten Messungen an Realaufnahmen schwierig zu bewerten sind, so deutet sich jedoch ein wichtiger Vorteil des hier betrachteten zweidimensionalen Bewegungsvorhersageverfahrens an: Durch die Unempfindlichkeit gegenüber Fehlern bei der Ermittlung der Kameraparameter, konnte durchweg ein besseres Ergebnis für zweidimensionale Vorhersagen erzielt werden. Auch konnte gezeigt werden, dass gewisse Eigenschaften der Vorhersagemethoden, welche bei synthetischen Bildfolgen beobachtet werden konnten, bei Realaufnahmen nicht im gleichen Maße zur Geltung kommen. Die Thesen aus Abschn. 3.3 wurden i. A. nicht bestätigt.

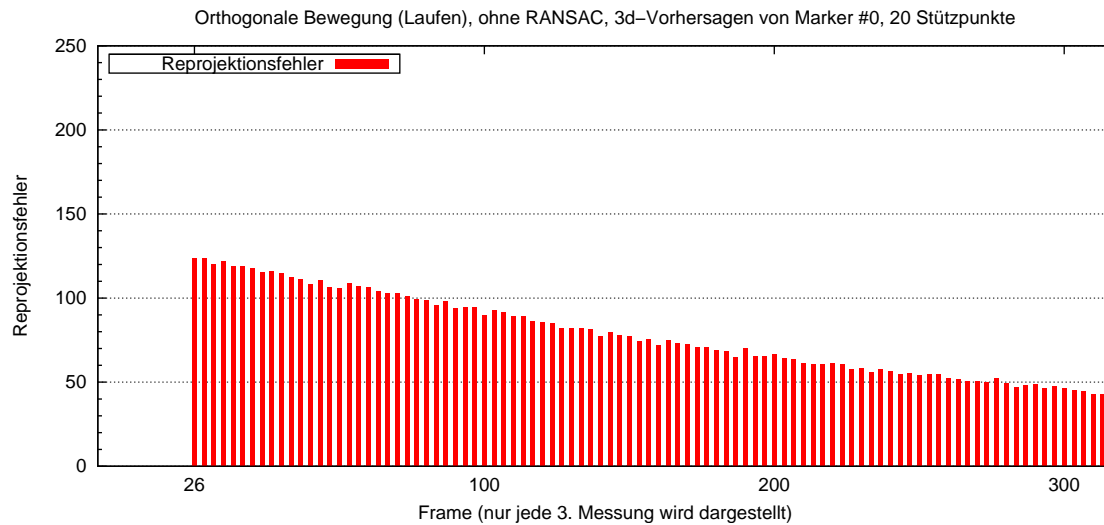


Bild 4.21: Reprojektionsfehler der triangulierten Punkte für die 3d-Vorhersage in Bild 4.2.2.

4.3 Verifikation der Messergebnisse

Im Falle der synthetischen Bilder ist eine direkte Berechnung der Kugelpositionen im Raum und im Bild anhand der in 4.1.1 gegebenen Daten möglich. Bei einem Vergleich mit den ermittelten Merkmalspositionen wurde eine durchschnittliche Abweichung von unter 1 Pixeln anhand von Stichproben festgestellt. Bei der Verifikation der rekonstruierten Welt-punkte dominierte ein hoher Fehler in der z-Koordinate, welcher jedoch erwartet werden konnte (siehe hierzu [HZ03, S. 321]). Die beiden anderen Koordiaten konnten jedoch mit einer hohen Genauigkeit bestimmt werden. Eine vergleichbare Verifikation der Messergebnisse ist bei den realen Aufnahmen nicht möglich, da keine Ground-truth in Form von bekannten Raumpunkten und bekannten Kameraparametern vorhanden ist.

Auch bei der Betrachtung der Reprojektionsfehler im Rahmen der synthetischen Aufnahmen, lag die durchschnittliche Abweichung bei ca. $1 \cdot 10^2$ Pixeln. Die maximale Abweichung kann mit ca. $4 \cdot 10^2$ Pixeln beziffert werden.

Für die Bestimmung der Korrespondenzen bei der Verwendung von 3d-Kalibrierungsmustern

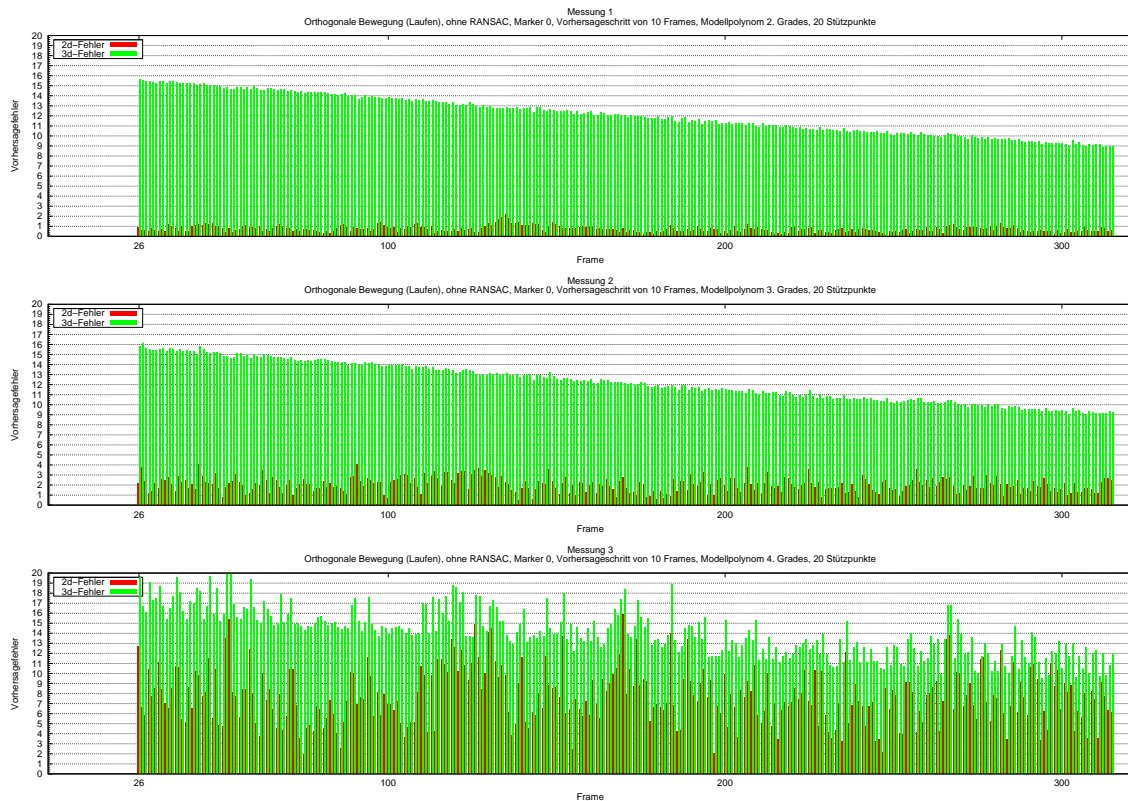


Bild 4.22: Bewegungsvorhersagen einer zur Basislinie orthogonalen Gangbewegung unter der Verwendung unterschiedlicher Modellpolynome.

und initiale Korrespondenzen für eine Berechnung der Fundamentalmatrix aus Bildinformationen, wurde kein automatisiertes Verfahren implementiert. Um die per Hand eingegebenen Korrespondenzen zu verifizieren, wurden zwei Programme geschrieben, welche die Punkte der jeweiligen Korrespondenzen in die entsprechenden Bilder einzeichnen und kennzeichnen. Einzelheiten zur Verwendung dieser Programme können dem Anhang E.5 entnommen werden.



Bild 4.23: Trajektorien der Harris-Corners und den korrespondierenden Vorhersagen bei einer parallelen Bewegung.

Kapitel 5

Zusammenfassung und Ausblick

5.1 Zusammenfassung

Zu Beginn der vorliegenden Arbeit wurden die theoretischen Grundlagen für physikalisch basierte Vorhersagen von Merkmalsbewegungen vorgestellt. Aufbauend auf dem Ansatz von [KM96] wurden zwei verallgemeinerte Methoden zur Bewegungsvorhersage umgesetzt, welche einerseits eine Vorhersage im Bild- und andererseits eine Vorhersage im Weltkoordinatensystem ermöglichen. Die Fehler bei der Vorhersage wurden anschließend für beide Verfahren, unter Verwendung von synthetischen Bildfolgen und realen Aufnahmen, gemessen. Die Ergebnisse der Messungen wurden vorgestellt, verglichen und, soweit dies möglich war, verifiziert.

Als abschließendes Fazit dieser Arbeit lässt sich festhalten, dass eine räumliche Vorhersage in der Theorie durchaus Vorzüge mit sich bringt. Diese verlieren in der Praxis, soweit dies festgestellt werden konnte, jedoch stark an Bedeutung. Vor allem eine hohe Empfindlichkeit gegenüber Ungenauigkeiten bei der Ermittlung der Kameraparameter, welche die Güte der Vorhersage maßgeblich beeinflussen, muss als größter Schwachpunkt des räumlichen Verfahrens genannt werden. Die Vorhersage im Bild ist dagegen unempfindlich gegenüber dieser Art Fehler. Im Falle einer sehr präzisen Ermittlung der Kameraparameter, kann der Einsatz einer räumlichen Vorhersage sich jedoch unter Umständen lohnen. Kann dies nicht garantiert werden, so stellt die direkte Vorhersage im Bild in der Praxis

die geeignetere Wahl dar.

5.2 Ausblick

Diese Arbeit kann in vielerlei Hinsicht fortgeführt werden. Die Kameraparameter zu Aufnahmen, welche im Rahmen dieser Arbeit erstellt wurden, konnten durchgehend nur sehr ungenau ermittelt werden. Daher erscheint es sinnvoll, die vorgestellten Ergebnisse durch weitere Aufnahmen, mit möglichst präzise ermittelten Parametern, zu festigen oder gegebenenfalls zu widerlegen.

Eine weitere Möglichkeit für eine Fortführung der Arbeit wäre die Integration eines Menschmodells in den entwickelten Tracker. Es lässt sich jedoch annehmen, dass ein Vergleich auf der Ebene der Merkmale im Bild ebenfalls zugunsten der Vorhersage im Bild ausfällt. Dies lässt sich durch die starke Abstraktion durch das Menschmodell erklären. Wird beispielsweise ein System aus Starrkörpern verwendet, so muss dieses in die Punktmasse eingepasst werden. Bei der Rücktransformation der Markerpunkte in das Weltkoordinatensystem, muss dann eine Methode für die Korrektur des so entstandenen Fehlers eingesetzt werden. Ein Vergleich mit der vorgestellten räumlichen Vorhersage im Raum könnte jedoch interessant sein, besonders für Vorhersagen über große Zeitschritte hinweg.

Anhang A

Aufgabenstellung

Ziel der Arbeit ist es, aus Bildströmen von einem kalibrierten Mehrkameranystem modellbasiert die Bewegung eines Sportlers vorherzusagen. Hierzu werden Bildströme aus Hochgeschwindigkeitskameras verwendet. Beginnend mit markerbasierten Verfahren soll hierbei durch eine Evaluation die Qualität der Vorhersage von 2d- und 3d-Bewegungsmodellen beim Tracking verglichen werden. Zunächst kann dabei von einer anwendungsbezogenen Einschränkung der Bewegung des Sportlers ausgegangen werden. Erwartet wird eine ausführliche Literaturrecherche im Bereich modell- und markerbasierter Trackingmethoden, um das entwickelte Verfahren vom Stand der Technik abzugrenzen. Die Implementierung erfolgt nach den Prinzipien der Softwaretechnik in C++. Das Ergebnis soll zum Schluss ansprechend visualisiert werden.

Anhang B

Mathematische Verfahren

Im Folgenden werden verschiedene mathematische Verfahren erläutert, welche in der Arbeit verwendet wurden.

B.1 Lösung eines homogenen Gleichungssystems

Die Lösungsmenge eines homogenen Gleichungssystems der Form $A \cdot x = 0$ mit bekannter Matrix A und einem unbekanntem Lösungsvektor x , wird durch den Nullraum von A aufgespannt. Stellen Messungen die Grundlage von A dar, kann es aufgrund von Rauschen vorkommen, dass keine Lösung für das gegebene Gleichungssystem existiert.

Im Allgemeinen gilt, dass eine Matrix mit vollem Rang keinen Nullraum besitzt. Ist nur eine Lösung gesucht, so kann eine Verringerung des Rangs einer solchen Matrix mithilfe der Singulärwertzerlegung erzwungen werden. Alternativ kann die optimale Approximation \hat{x}

zur gesuchten Lösung \boldsymbol{x} jedoch auch direkt abgelesen werden:

$$\text{svd}(\boldsymbol{A}) = \boldsymbol{U} \cdot \boldsymbol{\Sigma} \boldsymbol{V}^T \quad (\text{B.1})$$

$$\boldsymbol{V}^T = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & & v_{2n} \\ \vdots & & \ddots & v_{1n} \\ v_{n1} & v_{n2} & & v_{nn} \end{pmatrix} \quad (\text{B.2})$$

$$\hat{\boldsymbol{x}} = (v_{11}, v_{12}, \dots, v_{1n})^T \quad (\text{B.3})$$

Die Approximation ist optimal in der Hinsicht, dass die *Summe der Fehlerquadrate* $\sum_{i=1}^n (y_i - p(x_i))^2$ minimiert wird.

B.2 Regressionsanalyse

Ziel der *Regressionsanalyse* ist es, eine Modellfunktion möglichst genau in eine gegebene Datenmenge einzupassen um einen Trend in dem Verlauf der Daten aufzuzeigen. Dabei soll der dabei entstehende Fehler minimal sein. In der Praxis wird dabei i. d. R. die Summe der Fehlerquadrate als Abstandsmaß verwendet.

Ohne die Allgemeinheit einzuschränken, wird im weiteren Verlauf ein Polynom 2. Grades $p: \mathbb{R} \mapsto \mathbb{R}$, $p(x) = \hat{y} = a \cdot x^2 + b \cdot x + c$ betrachtet, welches als Modellfunktion dienen soll. Es sind mindestens 3 Punkte notwendig, um ein quadratisches Polynom eindeutig festzulegen. In diesem Spezialfall liegen sogar alle Punkte auf dem Graphen. Werden mehr als 3 Punkte in Betracht gezogen, so gilt es einen Fehlerterm zu minimieren, um eine bestmögliche Approximation der Punktmenge zu erreichen.

Um an eine Lösung zu gelangen, setzt man alle Werte der Punktmenge in die Gleichung des Polynoms ein, wodurch man ein Gleichungssystem erhält, welches linear in den Komponenten \hat{a} , \hat{b} und \hat{c} ist. Dieses lässt sich durch eine Homogenisierung in die folgende Form bringen

$$\boldsymbol{A} \cdot \boldsymbol{z} = \mathbf{0}, \quad \boldsymbol{A} = \begin{pmatrix} x_1^2 & x_1 & 1 - y_1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 - y_n \end{pmatrix}, \quad \boldsymbol{z} = (\hat{a}, \hat{b}, \hat{c})$$

Eine gewünschte Lösung lässt sich nun leicht mit der Methode aus Abschn. B.1 berechnen. Für Modellpolynome eines höheren Grades besteht eine analoge Vorhergehensweise.

Die Regressionsanalyse liefert in der Regel eine gute Näherung, wenn das verwendete Modell (hier also das Polynom) das zugrunde liegende Problem ausreichend gut umschreibt und die Daten einem vergleichsweise schwachen Rauschen ausgesetzt sind. Gegenüber stark verrauschten Daten gilt die Regressionsanalyse jedoch als nicht robust.

Damit das Verfahren numerisch stabil ist, sollte vor der Berechnung eine Normalisierung der Datenmenge vorgenommen werden. In dieser Arbeit wurde der Vorschlag aus [HZ03, S. 180] auf das gegebene Problem übertragen. Eine deutliche Verbesserung der Genauigkeit konnte empirisch festgestellt werden.

Anhang C

Weitere Berechnungsverfahren

C.1 Detektion von Reflektionsmarkern

Die Detektion von Markern ist stets an den Typ der verwendeten Marker geknüpft. Reflektionsmarker weisen in einem bestimmten Frequenzspektrum einen besonders hohen Reflektionsgrad auf, weshalb oft die Anwendung eines einfachen Schwellwerts genügt, um die Marker vom Rest des Bildes zu separieren. Im sichtbaren Lichtspektrum kann jedoch nicht davon ausgegangen werden, dass alleine die verwendeten Marker das Licht ausreichend stark reflektieren. Dies hat besonders fatale Konsequenzen, wenn der Marker auf Materialien wie z. B. der Haut, welche ebenfalls vergleichsweise gute Reflektionseigenschaften besitzt, angebracht werden. Ein globaler Schwellwert führt hierbei zu schlechten Ergebnissen, da dieser so gewählt werden muss, dass selbst bei ungünstigen Beleuchtungssituationen die Marker separiert werden können. Dabei ist es möglich, dass die Region der Marker und der Hautpartien durch den Schwellwert zusammenfällt. Ein bessere Ansatz wäre hier ein adaptiver Schwellwert, welcher sich beispielsweise aus dem Mittelwert s_a einer Nachbarschaft R plus einer Konstanten k ergibt:

$$s_a = \mu(R) + k$$

Der im Rahmen der Arbeit verwendete Algorithmus glättet das Bild vor der Anwendung des adaptiven Schwellwerts mit einem Gauskern der Größe 9×9 . Im Anschluss wird

sowohl ein adaptiver Schwellwert als auch ein normaler Schwellwert auf das Bild angewendet. Der normale Schwellwert soll dabei sehr niedrig gewählt werden. Er dient dazu, den Hintergrund auszugrenzen, wobei angenommen wird, dass der Hintergrund der Szene besonders dunkel gehalten wurde. Die Konstante des adaptiven Schwellwerts soll dagegen einen möglichst hohen Wert erhalten. Dieser Schwellwert dient, wie zuvor erwähnt, zur Separierung der unterschiedlichen Materialien. Die Ergebnisse des vorangegangenen Schritts sind zwei Binärbilder, welche mittels einer Durchschnittsoperation zusammengeführt werden.

Der Folgeschritt wird nun durch die Form der Marker bestimmt. Handelt es sich bei den Markern um flache Aufkleber, wie sie in dieser Arbeit verwendet wurden, so bietet es sich an, den Schwerpunkt der Konturen im Binärbild als Markerpunkt zu betrachten. Dabei kann es vorkommen, dass auch Regionen ohne Marker verarbeitet werden, wenn sie über eine ausreichend hohe Intensität und zudem über einen ausreichenden Abstand zu den Markern verfügen. Solche fehlerhaften Detektionen können eingeschränkt werden, indem man eine Region angibt, in der initial nach Markern gesucht wird. Die übrigen fehlerhaften Detektionen stellen keine allzu große Beeinträchtigung dar, da sie sich bei etwas schlechteren Beleuchtungsverhältnissen in der Regel verlieren.

Anhang D

Grundlagen der klassischen Mechanik

Im diesem Kapitel werden ausgewählte Grundlagen der klassischen Mechanik erläutert.

D.1 Inertiale Bezugssysteme

Ob sich ein Punkt mit konstanter Geschwindigkeit bewegt oder ob er stattdessen ruht, hängt in der klassischen Mechanik vom Standpunkt des Beobachters ab. Dieser wird durch ein *Bezugssystem* repräsentiert, welches wiederum die Grundlage für ein Koordinatensystem darstellt. Bezugssysteme und Koordinatensysteme stellen rein formal unterschiedliche Konzepte dar, da ein Bezugssystem im Gegensatz zum Koordinatensystem, koordinatenfrei charakterisiert werden sollte. Jedoch wird diese strikte Trennung oft vernachlässigt und die Begriffe werden, wie auch im weiteren Verlauf dieses Kapitels, synonym verwendet [Gre98, Begr.: Bezugssystem].

Nimmt man zwei beliebig gewählte Bezugssysteme, so ist es zulässig, dass sich diese mit einer konstanten Geschwindigkeit relativ zueinander bewegen dürfen. Nicht zulässig ist dagegen eine Änderung der relativen Orientierung. Ein solches Bezugssystem wird als *inertiales Bezugssystem* bezeichnet.

Ein Beispiel hierfür ist ein Beobachter in einem Zug und ein Beobachter ausserhalb des Zuges. Der erwähnte Zug soll sich zudem gradlinig und mit einer konstanten Geschwin-

digkeit fortbewegen. Für den Beobachter im Zug scheint sich die Position eines am Zug befestigten Gegenstandes nicht zu ändern. Von dem anderen Standpunkt aus gesehen bewegt sich der Gegenstand jedoch mit einer konstanten Geschwindigkeit.

D.2 Kinematik von Massepunkten

Ein *Massepunkt* ist ein Punkt in einem Koordinatensystem, welcher zusätzlich zu seiner Position durch eine Masse charakterisiert wird. Im Folgenden soll zudem die Annahme gelten, diese Masse sei über die Zeit konstant. Betrachtet man den Verlauf seiner Positionen $\mathbf{X}(t)$ über einen kontinuierlichen Zeitraum, so beschreiben diese eine Trajektorie in dem gewählten Koordinatensystem. Neben der Position lassen sich weitere Charakteristika wie Geschwindigkeit und Beschleunigung festlegen. Dabei wird im Folgenden von einem kartesischen Koordinatensystem ausgegangen.

Unter Geschwindigkeit versteht man eine vektorielle Größe, welche die relative Änderung der Position \mathbf{X} in dem gewählten Bezugssystem über die Zeit beschreibt. Ist ihr Betrag gleich 0, so verharrt der Punkt zu diesem Zeitpunkt an seiner gegebenen Position. Ist die Geschwindigkeit dagegen konstant, bewegt sich der Punkt gleichmäßig in die gegebene Richtung. Die Geschwindigkeit \mathbf{v} zu einem Zeitpunkt t entspricht also der Relation

$$\mathbf{v}(t) = \frac{d\mathbf{X}(t)}{dt}.$$

Eine konstante Geschwindigkeit repräsentiert die einfachste Bewegung in einem inertialen Bezugssystem. I. d. R. beschreibt eine Bewegung jedoch eine gekrümmte Trajektorie, in der sich auch die Geschwindigkeit stetig ändert. Der Punkt erfährt also eine gerichtete Beschleunigung \mathbf{a} zu einem gegebenen Zeitpunkt t , ausgedrückt durch

$$\mathbf{a}(t) = \frac{d\mathbf{v}(t)}{dt}.$$

D.2.1 Kinematik von Starrkörpern

Im Gegensatz zu Massepunkten, welche keine räumliche Ausdehnung beschreiben, muss bei Körpermodellen neben der Translationsbewegung auch die Rotation betrachtet werden.

Zu den einfachsten Körpermodellen gehören die Starrkörper. In einem Starrkörper ist der Abstand von 2 beliebig gewählten Punkten stets konstant. Es kann also keine Deformation des Körpers stattfinden.

In Analogie zu den translationsbezogenen Größen, lassen sich hier Winkelgeschwindigkeit ω und Winkelbeschleunigung α definieren, um die Änderung der Orientierung φ über die Zeit zu beschreiben:

$$\omega(t) = \frac{d\varphi(t)}{dt}, \quad (\text{D.1})$$

$$\alpha(t) = \frac{d\omega(t)}{dt}. \quad (\text{D.2})$$

Diese winkelbezogenen Größen sind ebenfalls vektorieller Natur. Die Winkelgeschwindigkeit zeigt dabei entlang der Drehachse und bestimmt zudem die Richtung der Drehbewegung (Rechte-Hand-Regel).

D.3 Dynamik

Mithilfe der erwähnten kinematischen Größen lassen sich die spatialen und temporalen Eigenschaften einer gegebenen Bewegung beschreiben. Die *Dynamik* liefert dagegen entsprechende Größen, mit denen die Ursachen für einen bestimmten Bewegungsverlauf beschrieben werden können. Im Allgemeinen wird die Ursache für eine Änderung der Geschwindigkeit mithilfe von *Kräften* quantifiziert. Ist eine bestimmte Trajektorie gegeben, so lässt sich, mittels der zuvor beschriebenen kinematischen Größen, die Summe aller Kräfte, welche zu einem gegebenen Zeitpunkt t auf den betrachteten Punkt einwirken, ermitteln. Sind wiederum alleine die Kräfte gegeben, welche über die Zeit auf einen Punkt einwirken, so lässt sich die zugehörige Trajektorie mitsamt der zugehörigen kinematischen Eigenschaften bestimmen. Dieses Prinzip stellt die wichtigste Grundlage für die physikalisch-basierte Vorhersage von Bewegungen dar.

Die grundlegenden Gesetze, welche die Wirkungsweise von Kräften bestimmen, werden durch die Axiome von Newton definiert:

1. Gesetz: Ein Körper verharrt im Zustand der Ruhe oder der gleichförmigen Transla-

tion, solange die Summe aller auf ihn einwirkenden Kräfte und die Summe aller einwirkenden Momente jeweils Null ist. Man bezeichnet diesen Zustand auch als *Äquilibrium*.

2. **Gesetz:** Die Änderung der Bewegung einer Masse ist der Einwirkung der bewegenden Kraft proportional und geschieht nach der Richtung derjenigen geraden Linie, nach welcher jene Kraft wirkt.
3. **Gesetz:** Kräfte treten immer paarweise auf. Übt ein Körper A auf einen anderen Körper B eine Kraft aus (*actio*), so wirkt eine gleich große, aber entgegengerichtete Kraft von Körper B auf Körper A (*reactio*).

Da es oft sehr aufwändig ist, ein komplexes System zu betrachten, unterscheidet man im Regelfall zwischen *inneren Kräften* und *äußeren Kräften*. Innere Kräfte treten stets paarweise in dem System auf. Sie erfüllen dadurch das *Energieerhaltungsprinzip*. Alle äußeren Kräfte werden ohne die Betrachtung einer Gegenkraft in das System eingebracht. Welche Kräfte nun als innere und welche Kräfte als äußere gelten, hängt von den Ansprüchen an das Modell ab, welche mithilfe der dynamischen Größen beschrieben werden sollen. Diese freie Wahl des Abstraktionsgrades wird als *Schnittprinzip* bezeichnet.

D.3.1 Dynamik von Massepunkten

Damit eine Änderung eines Bewegungsablaufes erfolgt, muss eine Kraft einen Massepunkt beschleunigen. Die Beschleunigung alleine reicht für die Charakterisierung von Kräften jedoch nicht aus, da schwerere Körper sich von Natur aus weniger von einwirkenden Kräften in ihrer Bewegung beeinträchtigen lassen als vergleichsweise leichtere. Diese *Trägheit* wird durch die Masse m eines Punktes oder eines Körpers charakterisiert, wobei eine antiproportionale Beziehung zwischen Masse und der effektiven Beschleunigung besteht. Eine Kraft \mathbf{F} lässt sich also wie folgt definieren:

$$\mathbf{F}(t) = m \cdot \mathbf{a}(t).$$

Der *Impuls* $\mathbf{p} = m \cdot \mathbf{v}(t)$ repräsentiert die aufgesparten Kräfte, welche einem Massepunkt zu einem bestimmten Zeitpunkt zugeschrieben werden können. Impulse werden bei Kon-

takten mit anderen Punkten oder Körpern übertragen. Der Impulsaustausch erfolgt dabei mittels Kontaktkräften.

D.3.2 Dynamik von Starrkörpern

Wirkt eine Kraft auf einen Starrkörper statt auf einen Punkt, so verändert die Kraft nicht nur die translationsbezogene Geschwindigkeit des Körpers, sondern gegebenenfalls auch seine Winkelgeschwindigkeit. Um die Anteile der Kraft zu bestimmen, welche jeweils auf die verschiedenen Größen übertragen werden, sind sowohl der Schwerpunkt als auch die Masseverteilung des Körpers von Belang. Beide sind im Falle von Starrkörpern konstant, was eine erhebliche Vereinfachung bei der Simulation von mechanischen Abläufen darstellt.

Wirkt eine Kraft \mathbf{F} auf einen Punkt P eines Starrkörpers mit einem Schwerpunkt C ein, so wird der Anteil der Kraft, welche auf die translationsbezogene Beschleunigung übergeht, anhand des Kosinus des Winkels zwischen \mathbf{F} und CP bestimmt. Sind diese Vektoren kollinear, so findet keinerlei Winkelbeschleunigung statt. Sind sie dagegen orthogonal so findet keinerlei translationsbezogene Beschleunigung statt. Der erste Spezialfall wird im Rahmen der Kollisionsbehandlung von Körpern auch als *zentrisch* bezeichnet, alle übrigen Fälle als *azentrisch*.

Derjenige Anteil der Kraft, welcher auf die Winkelgeschwindigkeit einwirkt, wird als *Drehmoment* $\tau(t)$ bezeichnet. Analog zu den Kräften werden Drehmomente von der Trägheit des Körpers beeinträchtigt, welche wiederum durch ihre Masseverteilung charakterisiert werden kann. Im Falle von Starrkörpern ist diese konstant und lässt sich mittels der *Hauptträgheitsachsen* eindeutig in Form eines Tensors, dem sogenannten *Trägheitstensor* \mathbf{J} , beschreiben. Die Verwendung eines Tensors ist notwendig, da anders als bei einem Punkt, die Masseverteilung eines Körpers anisotropisch ist. Daher muss die effektive Trägheit des Körpers für jeden Angriffswinkel einer Kraft unterschieden werden.

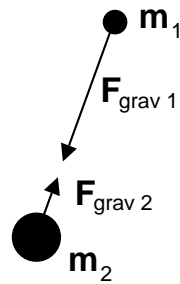


Bild D.1: Gravitationskräfte zwischen zwei Massepunkten mit ungleicher Masse

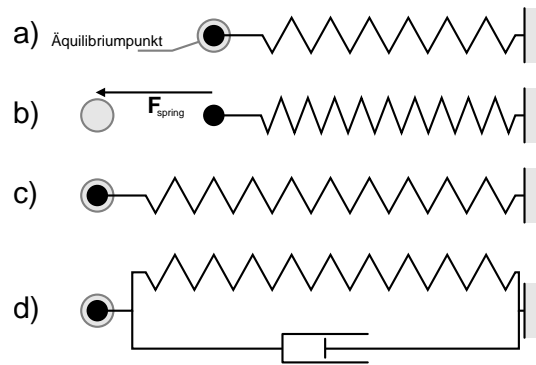


Bild D.2: Federkräfte: a) Der Federpunkt im Äquilibriumpunkt, b) Eine Federkraft wirkt mit $k_s = 1$, c) Situation b ist wieder ausgeglichen, d) Eine gedämpfte Feder

D.3.3 Kraftmodelle

Im Allgemeinen lassen sich Kräfte mithilfe eines Differentialgleichungssystems beschreiben. Dies kann, je nach gewünschter Genauigkeit, jedoch sehr komplex werden, was eine geschlossene Lösung unter Umständen ausschließt. Daher sind oft vereinfachende Modelle erforderlich, welche die effektive Wirkung einer komplexen Kraft unter gegebenen Voraussetzungen gut annähern.

Ein Beispiel stellen Gravitationskräfte dar. Im Allgemeinen kann angenommen werden, diese Kraft zwischen den Schwerpunkten der entsprechenden Körper wirkt, wodurch sich eine Richtung ergibt. Der Betrag dieser Kraft ergibt sich dabei aus dem Produkt der Gravitationskonstante G , welche eine Naturkonstante darstellt, und den beiden Gesamtmassen m_1 und m_2 der betrachteten Körper. Zudem sinkt der Einfluss der Kraft umgekehrt proportional zum Abstand r der beiden Körper, womit sich folgendes Gesamtbild ergibt:

$$|\mathbf{F}_{\text{grav}}| = \frac{G \cdot m_1 \cdot m_2}{r^2}$$

In einer kleinen Umgebung bietet es sich jedoch alternativ an, einen konstanten Gravitationsvektor zu wählen, da sich eine Positionsänderung nur geringfügig auf die Richtung

und den Betrag des Vektors auswirken würde.

D.3.4 Das gemeinsame Wirken mehrere Kräfte

Komplexere Kräfte lassen sich oft durch das Zusammenwirken von mehreren Kräften beschreiben. Daher stellen Kraftmodelle auch ein Baukastensystem dar, wobei es auch verschiedene Abhängigkeiten zwischen den Kräften geben kann. Sind Kräfte jedoch von einander unabhängig, wie beispielsweise Gravitation und Muskelkraft, so wirken sie stets als Summe (*Superpositionsprinzip*). Wichtige Bausteine, welche auch im Rahmen dieser Arbeit erwähnt werden, sind beispielsweise Modelle für (*lineare*) *Federkräfte* und *Dämpfungskräfte*.

Der Betrag der Kraft, welcher eine ideale Feder auf einen Punkt ausübt, verhält sich proportional zur Auslenkung der Feder. Unter der Auslenkung einer Feder versteht man dabei den Betrag des Vektors $\mathbf{u}(t)$ vom Massepunkt zum Äquilibriumspunkt. Unter Angabe einer Proportionalitätskonstanten k_s , der sogenannten Federkonstanten, ergibt sich das folgende Modell für lineare Federkräfte:

$$\mathbf{F}_{\text{spring}} = -k_s \cdot \mathbf{u}(t), \quad k_s > 0.$$

In der Regel neigen weiche Federn dazu, um den Äquilibriumspunkt zu oszillieren. Für Oszillationen mit einer geringen Amplitude kann, auf der Basis von *Stokes Gesetz*, eine geeignete Dämpfungskraft formuliert werden. Stokes Gesetz besagt im übertragenen Sinn, dass ein linearer und proportionaler Zusammenhang zwischen dem mechanischen Widerstand und der Geschwindigkeit des Massepunktes besteht [Par02, S. 486]. Eine einfache Dämpfungskraft genügt also der Relation

$$\mathbf{F}_{\text{damping}} = -k_d \cdot \mathbf{v}(t), \quad k_d > 0,$$

wobei k_d die sogenannte *Dämpfungskonstante* darstellt.

Eine detailliertere Betrachtung von diesen und weiteren Kraftmodellen befindet sich beispielsweise in [Ebe04, S. 457f.].

Anhang E

Aufbau der DVDs

Zu dieser Arbeit gehören 2 Datenträger. Diese enthalten zum einen den Quellcode der Anwendung, welche in Abschn. 3.5 dieser Arbeit vorgestellt wurde, sowie die vorliegende Ausarbeitung, die Präsentation der Zwischenergebnisse und die zur Arbeit gehörigen Hochgeschwindigkeitsaufnahmen.

E.1 Verzeichnisstruktur

```

ROOT
+- Ausarbeitung
| +- Arbeit
| | +- ...
| +- VortragOberseminar
|   +- ...
+- Programmcode
| +- DoxygenDoku
| +- ...
+- Sonstiges
  +- Aufnahmen
  +- WeitereBerechnungen

```

Die Verzeichnisstruktur der 1. DVD entspricht dem vorangegangenen Schema. Das Verzeichnis `Arbeit` enthält dabei den Quelltext und die Abbildungen der vorliegenden Ausarbeitung. Im Verzeichnis `VortragOberseminar` befinden sich die Präsentation der Arbeit, welche im Oberseminar der AGAS am 24. Januar gehalten wurde.

Das Verzeichnis `Programmcode` enthält den Quelltext des Trackers sowie weitere Informationen, welche zur Generierung der zugehörigen Dokumentation benötigt werden. Letztere sind in dem Unterverzeichnis `DoxygenDoku` untergebracht.

Ein Teil der Videos, mit den zugehörigen Konfigurationsdateien, befindet sich im Verzeichnis `Aufnahmen1`. Weitere Aufnahmen befinden sich auf den anderen DVDs, welche der Arbeit beiliegen. Die Verzeichnisstruktur der übrigen DVDs ist selbsterklärend und wird daher an dieser Stelle nicht weiter erläutert.

Im Verzeichnis `WeitereBerechnungen` befinden sich die berechneten Erwartungswerte und Mittelwerte, welche in Kapitel 4.2 vorgestellt wurden.

E.2 Dokumentation erstellen

Die entwickelte Anwendung wurde durchgehend unter Verwendung der Doxygen-Syntax kommentiert. Um aus den Kommentaren eine HTML- oder Latex-basierte Gesamtdokumentation zu generieren, ist ein entsprechender Kompiler notwendig, um die Kommentare zu übersetzen. Das Doxygen-Projekt (<http://www.doxygen.org>) stellt einen solchen Kompiler frei zur Verfügung. Nach der Installation des Softwarepakets, ist der Kompiler über die Kommandozeile verfügbar.

Ist zudem das Entwicklungswerkzeug `make` installiert, lässt sich die Dokumentation leicht über einen Aufruf von `make docs` im Verzeichnis `Programcode` auf der ersten CD erstellen. Ist `make` nicht installiert, so kann die Dokumentation stattdessen durch einen Aufruf von `doxygen Doxyfile` im Verzeichnis `Programcode/DoxygenDoku` direkt erstellt werden.

E.3 Den Quelltext kompilieren

Wie der Quelltext kompiliert werden kann, wird in der zugehörigen Dokumentation beschrieben. Um die Dokumentation zu erstellen, kann die Anleitung des vorangegangenen Abschnittes zu Rate gezogen werden.

E.4 Die Ausarbeitung und den Seminarvortrag erstellen

Sowohl Arbeit als auch Seminarvortrag sind in \LaTeX gesetzt. Um die Arbeit in ein ansprechendes Format zu übersetzen, müssen jedoch Bilder konvertiert und Plots gerendert werden. Diese Arbeit wird automatisch von einem jeweils beiliegenden `Makefile` übernommen. Ist das Entwicklungswerkzeug `make` auf dem System vorhanden, so reicht ein Aufruf von `make pdf` in dem entsprechenden Verzeichnis, um die Ausarbeitung oder den Vortrag in das PDF-Format zu übersetzen.

Test #	Name	Verwendung
1	prediction_test	Verifikation der Vorhersagealgorithmen
2	comparison_and_matching_test	Verifikation der Tracking-komponente (Mekmalsextraktion und Korrespondenzsuche)
3	matrices_test	Verifikation diverser Matrizen und Funktionen (Kameramatrizen, Trinagulierung, etc.)
4	calibrationpattern_verifikation	Verifikation manuell erstellter Korrespondenzen eines 3d-Patterns
5	correspondences_verifikation	Verifikation manuell erstellter Korrespondenzen für eine Fundamentalmatrix
6	planar_calibration_verifikation	Verifikation der Kalibrierung mittels eines Schachbrettmusters
7	3d_calibration_verifikation	Verifikation der Kalibrierung mittels eines 2d-Kalibrierungsmusters

Tabelle E.1: Auflistung der Testprogramme und deren Verwendung.

E.5 Testprogramme zur Verifikation der Tracking-Module

Die in Tabelle 4.3 erwähnten Programme, lassen sich durch einen Aufruf von `make tests` im Verzeichnis `Programcode` der 1. DVD erstellen. Tabelle E.5 listet die einzelnen Programme auf und erläutert deren Verwendung. Alternativ können einzelne Programme separat durch einen Aufruf von `make testN` erstellt werden, wobei N für die jeweilige Testnummer auf der erwähnten Tabelle steht. Die Verwendung der Programme ersliesst sich aus dem Kommentaren in den jeweiligen Quelltexten.

Zwei weitere Zusatzprogramme befinden sich in der Ordner `synthetic_image_generator` und `calibration_image_generator` im `Programcode`-Verzeichnis. Ersteres kann zur verwendung der in Kapitel 4 beschriebenen synthetischen Bildfolgen verwendet werden. Das zweite Zusatzprogramm erstellt die für Test 6 notwendigen synthetischen

Schachbrettkalibrierungsmuster.

E.6 Videos und Konfigurationsdateien

Die Videos wurden, aufgrund ihrer Größe, auf die beiliegenden Datenträger verteilt. Bei den Videos befinden sich die Konfigurationsdateien für die Messungen, auf denen die Plots in Kapitel 4 basieren. In Tabellen Tabelle E.6 und Tabelle E.6 wird eine direkte Zuordnung der Abbildungen zu den entsprechenden Konfigurationsdateien des Trackers angegeben.

Bild	DVD	Dateiname
4.2.1.1	1	Synthetic_X_noacc_regression.xml
4.2.1.2	1	Synthetic_X_constacc_regression.xml
4.2.1.3	1	Synthetic_X_linacc_regression.xml
4.2.1.4	1	Synthetic_X_quadacc_regression.xml
4.2.1.5	1	Synthetic_Z_noacc_regression.xml
4.2.1.6	1	Synthetic_Z_constacc_regression.xml
4.2.1.7	1	Synthetic_Z_linacc_regression.xml
4.2.1.8	1	Synthetic_Z_quadacc_regression.xml
4.2.1.1	1	Synthetic_X_quadacc_smallhistory.xml
4.2.1.2	1	Synthetic_X_quadacc_largehistory.xml
4.2.1.3	1	Synthetic_Z_quadacc_smallhistory.xml
4.2.1.4	1	Synthetic_X_quadacc_largehistory.xml
4.2.1.1	1	Synthetic_X_quadacc_reducedframerate.xml
4.2.1.2	1	Synthetic_Z_quadacc_reducedframerate.xml
4.2.1.1	1	Synthetic_X_quadacc_degree4_smallhistory.xml
4.2.1.2	1	Synthetic_Z_quadacc_degree4_smallhistory.xml
4.2.1.3	1	Synthetic_X_quadacc_degree4.xml
4.2.1.3	1	Synthetic_Z_quadacc_degree4.xml
4.2.1.5	1	Synthetic_X_quadacc_degree4_largehistory.xml
4.2.1.6	1	Synthetic_Z_quadacc_degree4_largehistory.xml
4.2.1.1	1	Synthetic_X_quadacc_largestep.xml
4.2.1.2	1	Synthetic_Z_quadacc_largestep.xml
4.2.1.1	1	Synthetic_X_quadacc_ransac.xml
4.2.1.2	1	Synthetic_Z_quadacc_ransac.xml

Tabelle E.2: Eine Zuordnung der Abbildungen des Abschnitts 4.2.1 zu den beiliegenden Konfigurationsdateien des Trackers.

Bild	DVD	Dateiname
4.2.2	1	Aufnahme1_Beine_Tritt1_regression_snapshots.xml
4.2.2	2	Aufnahme2_Koerper_marker_Z_regression.xml
4.2.2	2	Aufnahme2_Koerper_marker_Z_regression.xml
4.2.2	2	Aufnahme2_Koerper_marker_Z_regression.xml
4.2.2.1	2	Aufnahme2_Koerper_marker_Z_largestep10.xml
4.2.2.2	2	Aufnahme2_Koerper_marker_Z_degree3_largestep10.xml
4.2.2.3	2	Aufnahme2_Koerper_marker_Z_degree4_largestep10.xml
4.2.2	1	Aufnahme1_Koerper_markerlos1_ransac.xml

Tabelle E.3: Eine Zuordnung der Abbildungen des Abschnitts 4.2.2 zu den beiliegenden Konfigurationsdateien des Trackers.

Literaturverzeichnis

- [BAHH92] BERGEN, James R. ; ANANDAN, P. ; HANNA, Keith J. ; HINGORANI, Rajesh: Hierarchical Model-Based Motion Estimation. In: *ECCV '92: Proceedings of the Second European Conference on Computer Vision*. London, UK : Springer-Verlag, <http://www.springer.com>, 1992. – ISBN 3–540–55426–2, S. 237–252
- [BM98] BREGLER, Christoph ; MALIK, Jitendra: *Tracking People with Twists and Exponential Maps*. 1998
- [Bre97] BREGLER, Christoph: Learning and Recognizing Human Dynamics in Video Sequences. In: *Proc. IEEE Conf. Comp. Vision and Pattern Recognition*, 1997
- [BSMM01] BRONSTEIN, I. N. ; SEMENDJAJEW, K. A. ; MUSIOL, G. ; MÜHLIG, H.: *Taschenbuch der Mathematik*. Verlag Harri Deutsch, 2001
- [Ebe04] EBERLY, David: *Game Physics*. Morgan Kaufmann Publishers, 2004
- [FB81] FISCHLER, Martin A. ; BOLLES, Robert C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. In: *Communications of the ACM* 24 (1981), Nr. 6, S. 381–395. ISBN 0001–0782
- [Gav99] GAVRILA, D. M.: The Visual Analysis of Human Movement: A Survey. In: *Computer Vision and Image Understanding* 73 (1999), 1, Nr. 1, S. 82–98

- [GD96] GAVRILA, D. M. ; DAVIS, L. S.: Tracking of humans in action: a 3-D model-based approach. In: *ARPA Image Understanding Workshop*, 1996
- [GHSW06] GROSS, Dietmar ; HAUGER, Werner ; SCHRÖDER, Jörg ; WALL, Wolfgang A.: *Technische Mechanik 3: Kinetik*. Springer-Verlag Berlin-Heidelberg,, 2006
- [Gre98] GREULICH, Walter (Hrsg.): *Lexikon der Physik*. Spektrum, 1998
- [Gre03] GREINER, Walter: *Klassische Mechanik II: Teilchensysteme – Lagrange-Hamiltonische Dynamik – Nichtlineare Phänomene*. Wissenschaftlicher Verlag Harri Deutsch GmbH, 2003
- [Hil38] HILL, A. V.: The Heat of Shortening and the Dynamic Constants of Muscle. In: *Proceedings of the Royal Society 126 (B)*, 1938, S. 136–195
- [HN99] HERZOG, Walter (Hrsg.) ; NIGG, Benno M. (Hrsg.): *Biomechanics of the Musculo-skeletal System*. JOHN WILEY & SONS, 1999
- [HS88] HARRIS, C. ; STEPHENS, M.: A combined corner and edge detector. In: *Fourth Alvey Vision Conference*. Manchester, UK, 1988, S. 147–151
- [HZ03] HARTLEY, Richard I. ; ZISSERMAN, Andrew: *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003
- [Isa98] ISARD, Michael: *Visual Motion Analysis by Probabilistic Propagation of Conditional Density*, Oxford University, Diss., 1998
- [Kal] KALMAN, Rudolph E.: A new approach to linear filtering and prediction problems, 35–45
- [KB61] KALMAN, R.E. ; BUCY, R.S.: *New results in linear filtering and prediction theory*. 1961
- [KM96] KAKADIARIS, Ioannis A. ; METAXAS, Dimitris: Model-Based Estimation of 3D Human Motion with Occlusion Based on Active Multi-Viewpoint Selection. In: *IEEE Computer Vision and Pattern Recognition Conference*, 1996, S. 81–87

- [LW04] LEHN, Jürgen ; WEGMANN, Helmut: *Einführung in die Statistik*. 4. Teubner Verlag Stuttgart, 2004
- [MG01] MOSELUND, Thomas B. ; GRANUM, Erik: *A Survey of Computer Vision-Based Human Motion Capture*. 2001
- [MR98] MORRIS, Daniel D. ; REHG, James M.: Singularity Analysis for Articulated Object Tracking. In: *Computer Vision and Pattern Recognition*, 1998, S. 289–296
- [MS03] MADHAVAN, R. ; SCHLENOFF, C.: Moving Object Prediction for Off-road Autonomous Navigation / National Institute of Standards and Technology, <http://www.nist.gov>. 2003. – Forschungsbericht
- [NBN06] NISTÉR, David ; BERGEN, James R. ; NARODITSKY, Oleg: Visual Odometry for Ground Vehicle Applications. In: *Journal of Field Robotics* 23 (2006), Nr. 1. – inaugural issue
- [NEBES98] NEUNZERT, Helmut ; ESCHMANN, Winfried G. ; BLICKENSDÖRFER-EHLERS, Arndt ; SCHELKES, Klaus: *Analysis 2*. Springer Verlag, 1998
- [NP05] NIETHARD, Fritz U. ; PFEIL, Joachim: *Orthopädie*. Thieme, Stuttgart, 2005
- [Par02] PARENT, Rick: *Computer Animation: Algorithms and Techniques*. Morgan Kaufmann Publishers, 2002
- [PRC00] PAVLOVIC, Vladimir ; REHG, James M. ; CHAM, Tat-Jen: *A Dynamic Bayesian Network Approach to Tracking Using Learned Switching Dynamic Models*. 2000
- [PTVF07] PRESS, William H. ; TEUKOLSKY, Saul A. ; VETTERLING, William T. ; FLANNERY, Brian P.: *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, 2007
- [PW98] PENTLAND, Alex P. ; WREN, Christopher R.: Dynamic Models of Human Motion. In: *IEEE Proceedings of FG 98*, 1998

- [Ric01] RICHTER, Stefanie: *Vorwärtsmodelle und die Vorhersage des Bewegungsverlaufs: Ein Beitrag zur Erforschung interner motorischer Modelle*, Heinrich-Heine Universität Düsseldorf, Diss., 2001
- [SBF00] SIDENBLADH, Hedvig ; BLACK, Michael J. ; FLEET, David J.: *Stochastic tracking of 3D human figures using 2D image motion*. 2000
- [SE04] SCHIEHLEN, Werner ; EBERHARD, Peter: *Technische Dynamik: Modelle für Regelung und Simulation*. Teubner, 2004
- [Sei03] SEITZ, Thomas: *Videobasierte Messung menschlicher Bewegungen konform zum Menschmodell RAMSIS*, Technische Universität München, Diss., 2003
- [TV98] TRUCCO, E. ; VERRI, A.: *Introductory Techniques for 3-D Computer Vision*. New York : Prentice Hall, 1998
- [WAPD97] WREN, Christopher ; AZARBAYEJANI, Ali ; PENTLAND, Alex ; DARRELL, Trevor: Pfänder: Real-Time Tracking of the Human Body. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* Bd. 19, 1997, S. 780–785
- [WP99] WREN, Christopher R. ; PENTLAND, Alex P.: *Understanding Purposeful Human Motion* / Massachusetts Institute of Technology. 1999. – Forschungsbericht
- [Zha00] ZHANG, Zhengyou: A flexible new technique for camera calibration. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 (2000), Nr. 11, S. 1330–1334