

Untersuchungen zur Bewegungsanalyse
von aus starren Einzelteilen
zusammengesetzten Objekten

Harald Lange

FBI-HH-B-148/90

22. Juni 1990

Universität Hamburg
Fachbereich Informatik
Bodenstedtstraße 16
D-2000 Hamburg 50

Zusammenfassung

Die Bewegung eines starren Körpers läßt sich mathematisch zwischen zwei Zeitpunkten t_1 und t_2 durch eine lineare Abbildung beschreiben, der man Rotations- und Translationsparameter entnehmen kann. Aus diesen Bewegungsparametern ergibt sich die Orientierung der Rotationsachse, nicht aber ihre genaue Lage im Raum. Bei Bewegungsabläufen eines Objektes, das aus beweglich miteinander verbundenen starren Einzelteilen zusammengesetzt ist, stellen die Verbindungsachsen der Einzelteile eine wesentliche Bewegungscharakteristik dar. Bei der Bewegungsbeschreibung eines starren Einzelteils ist daher eine körperfeste Rotationsachse gesucht, die mit einer Verbindungsachse (physikalischen Rotationsachse) des Einzelteils zusammenfällt. Um eine geeignete Wahl treffen zu können, wird die Lage der Rotationsachse durch eine Glattheitsannahme eingeschränkt, welche in Form eines Bewegungsmodells ausgedrückt wird. Es ergibt sich eine einfache und natürliche Bewegungsbeschreibung, die die inhärenten physikalischen Bewegungsabhängigkeiten der Einzelteile zum Ausdruck bringt.

Abstract

The motion of a rigid body can be mathematically described by a linear transformation, which includes both rotational and translational parameters. These parameters determine the orientation of the rotation axis, but they are not sufficient to determine its exact position in space. The motion of an object, which is composed of jointed rigid parts can be characterized by the joint axis. Therefore, to describe the motion of a rigid component, we are interested in a rotation axis fixed to the part, which coincides with the joint axis. To find a proper axis, we use a smoothness constraint expressed by a motion model. This results in a simple and natural description of the object motion. In addition, the motion constraints between the rigid components are adequately modeled.

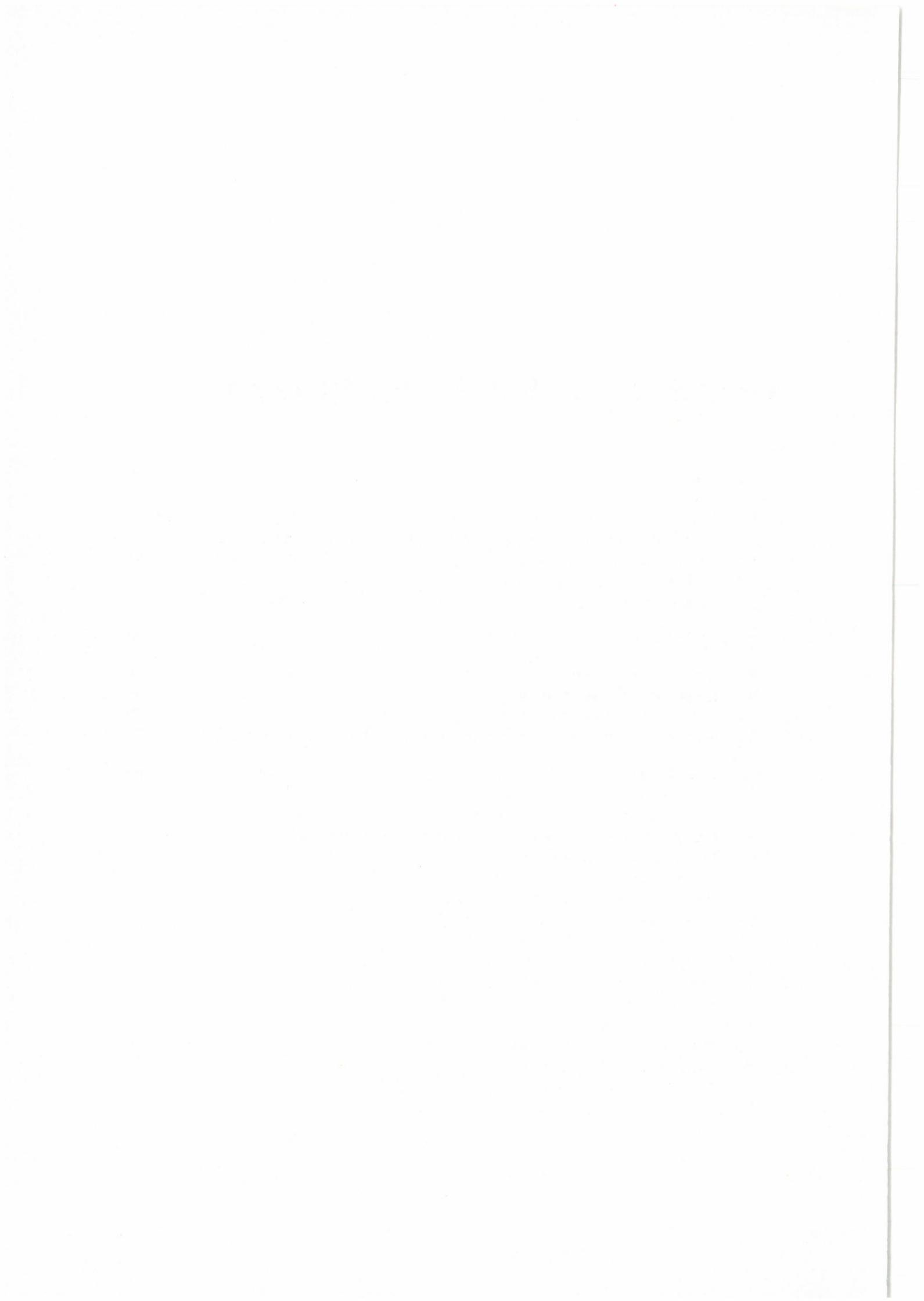
Inhaltsverzeichnis

1	Einleitung	1
2	“Structure from Motion”	4
2.1	Arbeiten von Ullman	5
2.1.1	Direkter perspektivischer Ansatz	8
2.1.2	Polar/parallel Methode	9
2.2	Weitere Arbeiten zu “Structure from Motion”	10
2.3	Nichtstarre Bewegung	14
2.4	Zusammenfassung	17
3	Bewegungsbeschreibung einzelner Punkte	20
3.1	Approximation der Punkttrajektorien	20
3.2	Wahl der Basisfunktionen	23
3.3	Geschwindigkeit und Beschleunigung von Punkten	25
3.4	Probleme bei der Bewegungsbeschreibung einzelner Punkte	27
4	3D-Bewegung starrer Körper	37
4.1	Isolieren von starren Körpern	37
4.2	Mathematik der Bewegung starrer Körper	41
4.2.1	Rotationsbewegung	41
4.2.2	Allgemeine Bewegung im Raum	47
4.3	Beschreibung von allgemeiner Bewegung im Raum	48
4.3.1	Linearer Ansatz	49
5	Bewegungsbeschreibung von zusammengesetzten Objekten	52
5.1	Der Ansatz von Asada und Tsuji	57
5.1.1	Zerlegung der Rotationsbewegung	58
5.1.2	Bestimmung der Rotationsachsenlage	62
5.1.3	Zusammenfassung	65
5.2	Eigene Vorgehensweise	66
5.2.1	Relative Bewegungsbeschreibung	67
5.2.2	Bestimmung der Rotationsachsenlage	68
5.2.3	Zusammenfassung	73

5.3	Implementation und Probleme bei realen Daten	74
6	Ergebnisse	78
6.1	1. Testfolge	78
6.2	2. Testfolge	91
7	Zusammenfassung	97
A	Mathematischer Anhang	100
	Literaturverzeichnis	103

Verzeichnis der Abbildungen

3.1	Punktbewegung	29
3.2	Gekoppelte Bewegung	30
3.3	Punktbewegungsgraphen	31
3.4	Approximation der Punkttrajektorie (algebraisches Polynom)	32
3.5	Approximation der Punkttrajektorie (trigonometrisches Polynom)	33
3.6	Vorhersage der Punkttrajektorie mit einem algebraischen Polynom	34
3.7	Vorhersage der Punkttrajektorie mit einem trigonometrischen Polynom	35
4.1	Eulerwinkel	46
5.1	Rotationsachsenlage	53
5.2	Körperfeste Rotationsachse 1	56
5.3	Körperfeste Rotationsachse 2	56
5.4	Objekt, das aus starren bewegbaren Einzelteilen zusammengesetzt ist	57
6.1	1. Testfolge	80
6.2	Orientierungen der ermittelten Rotationsachsen von Teil 1	81
6.3	Orientierungen der ermittelten Rotationsachsen von Teil 2	81
6.4	Orientierungen der ermittelten Rotationsachsen von Teil 3	82
6.5	Bewegung von Teil 1 relativ zum WKS	83
6.6	Trajektorie der Rotationsachse von Teil 1 im WKS	84
6.7	Bewegung von Teil 2 relativ zu Teil 1	85
6.8	Bewegung von Teil 1 relativ zu Teil 2	86
6.9	Bewegung von Teil 2 relativ zu Teil 3	87
6.10	Bewegung von Teil 3 relativ zu Teil 2	88
6.11	Bewegung von Teil 3 relativ zu Teil 1	89
6.12	Bewegung von Teil 1 relativ zu Teil 3	90
6.13	2. Testfolge	92
6.14	Bewegung von Teil 1 relativ zum WKS	93
6.15	Bewegung von Teil 2 relativ zu Teil 1	94
6.16	Korrekte Bewegung von Teil 2 relativ zu Teil 1	96



Kapitel 1

Einleitung

Die vorliegende Arbeit entstand im Arbeitsbereich Kognitive Systeme des Fachbereichs Informatik der Universität Hamburg. An diesem Arbeitsbereich beschäftigt man sich seit längerem mit der Analyse von Bildfolgen. Ein Ziel dieser Analyse ist es, eine geometrische Szenenbeschreibung [Neumann 82] zu erstellen, welche die Objekte und deren Bewegungen in der Szene räumlich beschreibt. Die Aufgabe besteht also zunächst darin, aus den zweidimensionalen Bildvorlagen eine Zerlegung der Szene in eigenständig sich bewegende Objekte zu gewinnen. Für jedes dieser Objekte ist die dreidimensionale Form und die dreidimensionale Bewegung zu rekonstruieren. Dabei gibt die gesuchte Bewegungsbeschreibung darüber Aufschluß, wie sich die Objekte in der Szene zwischen den Bildaufnahmezeitpunkten bewegt haben.

Um räumliche Informationen zu erhalten, muß die beim Abbildungsprozeß verlorengegangene Tiefeninformation zurückgewonnen werden. Dieses läßt sich unter anderem dadurch erreichen, daß man die verschiedenen Ansichten der Objekte nutzt, die eine Bildfolge bereitstellt. Dazu sucht man zunächst in jedem Bild dieser Bildfolge lokal nach markanten Merkmalen. Diese sollen so beschaffen sein, daß sie die Bilder von markanten Objektmerkmalen darstellen (wie z.B. Eckpunkte, Schnittpunkte von Kanten etc.). Aus diesen Bildmerkmalen lassen sich die räumlichen Tiefen der zugehörigen Objektmerkmale zurückgewinnen, wenn es gelingt, das Abbild jedes Objektmerkmals in mehreren Bildern der Bildfolge (idealerweise in allen Bildern der Folge) zu finden. Zuordnungen von Bildmerkmalen, die aus unterschiedlichen Bildern stammen und Bilder desselben Objektmerkmals sind, nennt man *Korrespondenzen*. Sind die Bilder zu unterschiedlichen Zeitpunkten entstanden, so spricht man von zeitlichen Korrespondenzen. Sind sie zum selben Zeitpunkt aus unterschiedlichen Positionen aufgenommen worden, so nennt man die Zuordnungen der Bildmerkmale räumliche Korrespondenzen.

Verwendet man eine monokulare Bildfolge, so muß man die Form und die Bewegung der Objekte allein aus den zeitlichen Korrespondenzen ermitteln. Das Kapitel 2 beschäftigt sich mit verschiedenen Arbeiten zu diesem Problem. In allen vorgestellten Verfahren müssen einschränkende Annahmen gemacht werden, damit das Problem mathematisch gelöst werden kann. In den meisten Fällen setzt man voraus, daß die Objekte starr sind. Andere Ansätze machen einschränkende Annahmen über die Objektbewegung.

Im SISSY (Stereo Image Sequence SYstem) Projekt, in dem auch diese Arbeit entstand, wird mit zwei sich an unterschiedlichen Positionen befindenden Farbbildkameras eine bewegte Szene aufgenommen. Man erhält eine zeitlich diskrete Folge von farbigen Stereobildern. Da man zu jedem Zeitpunkt zwei unterschiedliche Ansichten derselben Szene erhält, kann man aus den räumlichen Korrespondenzen eines Raumpunktes \mathbf{p} dessen 3D-Koordinaten direkt bestimmen. Der gesuchte Punkt im Raum muß nämlich auf den Projektionsstrahlen der beiden korrespondierenden Bildpunkte liegen. Die Koordinaten von \mathbf{p} erhält man also, indem man den Schnittpunkt dieser beiden Geraden bestimmt. Um aus den räumlichen Korrespondenzen die Tiefe des Raumpunktes bestimmen zu können, benötigt man die genauen Positionen und Orientierungen der Kameras sowie deren Abbildungsparameter (z. B.: Brennweite, Bildgröße etc.). Diese lassen sich aus Bildern von genau vermessenen Objekten bestimmen. Dieser Vorgang wird als Kamerakalibrierung bezeichnet. Einen Einblick in die Probleme der Kamerakalibrierung gibt [Bock 89]. Da die Tiefeninformation nicht mehr aus dem zeitlichen Kontext gewonnen werden muß, kann die Einschränkung auf starre Objekte, wie sie bei der dynamischen Szenenanalyse von monokularen Bildfolgen benötigt wird, entfallen.¹ Damit sind wir in der Lage, nicht nur Form und Bewegung von Objekten, sondern auch zeitliche Formveränderungen der Objekte zu beschreiben.

Die räumlichen Koordinaten von Szenenpunkten können sowohl aus den zeitlichen als auch aus den räumlichen Korrespondenzen der zugehörigen Bildpunkte ermittelt werden. Trotz der genannten Vorteile, die man durch die Verwendung einer Stereokamera erhält, existieren Grenzfälle, wie noch gezeigt werden wird, in denen nur monokulare Methoden weiterhelfen können.

Wir werden in dieser Arbeit voraussetzen, daß das Korrespondenzproblem näherungsweise gelöst wurde und daß die räumlichen Koordinaten von korrespondierenden Bildpunkten zu aufeinanderfolgenden Zeitpunkten bereits ermittelt wurden. Mit Hilfe dieser rekonstruierten Punkte werden wir versuchen, die räumliche Bewegung der Objekte zu beschreiben. Die Beschreibung soll die Bewegung auf eine möglichst einfache Weise wiedergeben und darüberhinaus geeignet sein, die Bewegung eines Objektes vorherzusagen, unter der Annahme, daß es sich unverändert weiterbewegt. Wir werden voraussetzen, daß die Objekte aus starren Einzelteilen zusammengesetzt sind. Dabei sollen die Verbindungen der Einzelteile so beschaffen sein, daß die relative Rotationsbewegung der verbundenen Teile bis auf einen verbleibenden Freiheitsgrad eingeschränkt wird. Damit rotieren die verbundenen Teile um eine Rotationsachse, die ihre Orientierung bezüglich der Teile beibehält. Die relative Translationsbewegung der verbundenen Teile werden wir durch ein Bewegungsmodell einschränken. Ein Beispiel für ein derartiges Objekt wäre ein Roboter. Er besteht aus starren Teilen, die über Gelenke oder Achsen miteinander verbunden sind.

In Kapitel 3 wird zunächst versucht eine Bewegungsbeschreibung für einzelne Punkte zu finden, die die zukünftige Bewegung des betrachteten Punktes vorhersagen kann. Dazu

¹Für ein Stereobildpaar ist die Forderung nach starren Objekten trivialer Weise erfüllt, da die Aufnahmen zum selben Zeitpunkt gemacht wurden.

wird eine Approximationsfunktion gesucht, die die räumliche Bewegungsbahn (Trajektorie) des betrachteten Punktes beschreibt. Die gewonnene Funktion wird dabei nur einen sehr beschränkten Einblick in die physikalische Bewegung und deren Ursachen geben. Zu Bewegungsvorhersagen eignet sie sich kaum.

Wir werden daher im weiteren versuchen, die Bewegung von makroskopischen Körpern zu beschreiben, indem wir für alle ermittelten Punkte des Körpers eine geschlossene Bewegungsbeschreibung suchen. Dazu müssen die Punkte zunächst zu Körpern gruppiert werden. Da im SISSY-Projekt kein szenenspezifisches Wissen verwendet wird, ist nicht von vornherein klar, welche Punkte zum selben Objekt gehören. In diesem Fall lassen sich nur allgemeine Annahmen ausnutzen. So erwartet man von Punkten desselben Objektes ähnliche Eigenschaften, wie z. B. räumliche Nähe oder gleichartige Bewegung. Bewegen sich Punkte zueinander starr, so liefert dies einen Hinweis darauf, daß sie zum selben Objekt gehören. Dieses strenge Bewegungskriterium werden wir in Abschnitt 4.1 zur Gruppierung verwenden. Für jedes dieser Einzelteile wird nun zunächst eine Bewegungsbeschreibung gesucht. Der Abschnitt 4.2 beschäftigt sich mit den mathematischen Grundlagen, die wir für die Beschreibung der Bewegung starrer Körper benötigen. In Abschnitt 4.3 werden wir dann ein Verfahren vorstellen, mit dem sich die Bewegungsparameter für ein starres Einzelteil aus den Punktkoordinaten ermitteln lassen.

Die ermittelten Bewegungsparameter beschreiben die Bewegung eines starren Körpers durch Rotation und Translation. In der Physik stellen Rotation und Translation zwei unterschiedliche und unabhängige Bewegungsphänomene dar. Möchte man die ermittelten Bewegungsparameter physikalisch interpretieren, so muß man bedenken, daß sich die Translation nicht eindeutig ergibt, sondern von der räumlichen Lage der Rotationsachse abhängt. Man muß also versuchen, die Lage der Rotationsachse so zu wählen, daß die resultierenden Translationsparameter geeignet sind, um physikalische Bewegungsgrößen wie Geschwindigkeit und Beschleunigung wiederzugeben. In Kapitel 5 werden wir zeigen, wie eine geeignete Wahl getroffen werden kann. Weiterhin werden wir versuchen, die Kopplungen der Einzelteile zu bestimmen, indem wir die Bewegungen der Einzelteile miteinander in Beziehung setzen. Sind zwei Teile miteinander verbunden, z. B. über eine gemeinsame Achse, so sind ihre Bewegungen korreliert. Um die strukturelle Beziehung der beiden Teile zum Ausdruck zu bringen, empfiehlt es sich, die Bewegung eines Teiles relativ zu seinem Verbindungspartner zu beschreiben. Diese Beschreibung wird sich dadurch auszeichnen, daß sie einfacher ist als andere Beschreibungen. Eine einfache Bewegungsbeschreibung gibt darüber hinaus die physikalischen Ursachen der Bewegung am besten wieder. Diese sind für das Verstehen der Bewegung von großer Bedeutung. So gewinnt man, um beim Beispiel des Roboters zu bleiben, einen Überblick über seine Zusammensetzung und über die Art der herrschenden Verbindungen.

Ich werde in dieser Arbeit versuchen, eine einfache und natürliche Bewegungsbeschreibung für solche Objekte zu erstellen, die aus verbundenen starren Teilen zusammengesetzt sind. Die Bewegungen derartiger Objekte wurden von *Asada + Tsuji* 83 bereits untersucht. Wir werden uns zunächst mit ihrem Verfahren beschäftigen und danach einen eigenen Ansatz vorstellen, der eine flexiblere und allgemeinere Beschreibung derartiger Bewegungen ermöglicht.

Kapitel 2

“Structure from Motion”

Das visuelle System des Menschen ist in der Lage, beim Betrachten einer zeitlichen monokularen Bildfolge, allein durch die Wahrnehmung der Objektbewegung auf die 3D-Form und -Bewegung der Objekte zu schließen. Die Bilder von bewegten Objekten werden richtig interpretiert, selbst wenn die Objekte dem Betrachter unbekannt sind und/oder ein einzelnes Bild keine 3D-Interpretation zuläßt. Diesen Interpretationsprozeß bezeichnet man gewöhnlich als “*Structure from Motion*”.

Ausgehend von diskreten Merkmalen in diskreten Bildern ist aus mathematischer Sicht die folgende Frage interessant: Wieviele Punkte in wievielen Ansichten werden benötigt, um die Form und die Bewegung eines Objektes ermitteln zu können? Theoretisch sind die Umkehrungen der Projektionen immer mehrdeutig, d. h. für eine beliebige Anzahl von Punkten und Ansichten gibt es unendlich viele Bewegungen der Punkte im Raum, die alle dieselben Projektionen erzeugen. Daher müssen zusätzliche Annahmen gemacht werden, welche die Lösungsmenge genügend einschränken. Hierzu wird meistens die Starrheit der Objekte vorausgesetzt (siehe Seite 6), das heißt, es wird angenommen, daß die paarweisen Punktabstände konstant bleiben. Unter dieser Voraussetzung läßt sich nun fragen, wieviele Punkte in wievielen Ansichten für eine eindeutige Lösung nötig sind. Setzt man perspektivische Projektion voraus, so läßt sich die Größe und Verschiebung der Objekte bis auf einen beliebigen Skalierungsfaktor genau bestimmen.¹ Sei M die Anzahl der Bilder und N die Anzahl der Punkte, so ergeben sich

$$MN - 1 \quad \text{Unbekannte.}$$

Dieses sind jeweils die Tiefen der N Punkte in den M Ansichten, wobei die Tiefe eines Punktes aufgrund des unbestimmbaren Skalierungsfaktors frei gewählt werden kann. Aus der Starrheitsannahme des Objektes lassen sich nun einschränkende Gleichungen formulieren. Um die Starrheit einer Punktconfiguration zu gewährleisten, müssen genügend viele paarweise Punktabstände festgelegt werden. Bei drei Punkten müssen drei Abstände konstant bleiben. Für jeden weiteren Punkt sind drei weitere Abstände notwendig. Somit

¹Dies beruht auf der Tatsache, daß die Größe eines Objektes aus den perspektivischen Projektionen nicht bestimmt werden kann, solange nicht mindestens ein absoluter Abstand bekannt ist.

erhält man aus der Starrheitsannahme $3(N - 2)$ Einschränkungen pro Bildintervall, d.h. insgesamt:

$$3(M - 1)(N - 2) \quad \text{Einschränkungen.}$$

Eine eindeutige Lösung des Problems kann nur dann existieren, wenn die Anzahl der Einschränkungen größer oder gleich der Anzahl der Unbekannten ist. Wir erhalten somit die folgende Beziehung [Ballard + Brown 82, Seite 214]:

$$\begin{aligned} 3(M - 1)(N - 2) &\geq MN - 1 \\ \Rightarrow 2MN - 3N - 6M + 7 &\geq 0 \\ \Rightarrow M &\geq \frac{3}{2} + \frac{2}{2N - 6} \end{aligned} \quad (2.1)$$

$$\Rightarrow N \geq 3 + \frac{2}{2M - 3} \quad (2.2)$$

Für eine gegebene Anzahl von Punkten/Ansichten läßt sich aus diesen Gleichungen ermitteln, wieviele Ansichten/Punkte mindestens vorliegen müssen, damit die Bestimmung möglich wird. So ergibt sich beispielsweise aus Gleichung 2.1, daß bei fünf Punkten mindestens zwei Ansichten benötigt werden, um das Problem lösen zu können. Umgekehrt genügen bei drei Ansichten bereits vier Punkte (Gleichung 2.2). Bei dieser Betrachtung wurde perspektivische Projektion vorausgesetzt, welche zu nichtlinearen Gleichungen führt. Bei nichtlinearen Gleichungen ist eine eindeutige Lösung nicht unbedingt gesichert, auch wenn die Anzahl der Unbekannten kleiner oder gleich der Anzahl der gegebenen Einschränkungen ist.

Erste Untersuchungen hierzu wurden von Wallach + O'Connell 53 unternommen. In ihren Versuchen zum – wie sie es nannten – "Kinetic Depth Effect" wurden Drahtrahmenmodelle hinter einem durchscheinenden Schirm bewegt und die Schatten ihrer Parallelprojektion von der anderen Seite des Schirmes beobachtet. Die Beobachter konnten aus der Veränderung der Schattenbilder die Form und Bewegung der ihnen unbekanntem Drahtrahmenmodelle erkennen. In einem einzelnen Schattenbild war die Form nicht erkennbar. Wallach + O'Connell 53 vermuteten, daß die Wahrnehmung der sich ändernden Kantenlängen und Kantenorientierungen den Interpretationsprozeß steuern.

2.1 Arbeiten von Ullman

Versuche, wie sie von Ullman 79 durchgeführt wurden, scheinen die Vermutung von Wallach + O'Connell 53 zu widerlegen. Ullman 79 machte Versuche mit Bildern, die nur einzelne Punkte enthielten. Die Parallelprojektionen von zwei mit Punkten versehenen, koaxial rotierenden Zylindern wurden von den Beobachtern richtig räumlich interpretiert, obwohl die Zylinder selbst unsichtbar waren, man also nur die Punkte sehen konnte. Obwohl die Bewegungen dieser Punkte in den Bildern nur zweidimensional ablaufen, nimmt der Beobachter sie als Elemente eines rotierenden 3D-Zylinders wahr. Da es unendlich viele 3D-Interpretationen gibt, müssen einschränkende Annahmen gemacht werden, wel-

che die "naheliegendste" bzw. "wahrscheinlichste" Lösung charakterisieren. Hierzu lassen sich die Eigenschaften der physikalischen Welt ausnutzen.

Ullman 79 vermutete, daß der Beobachter stets versucht, die Projektionen als Bewegung eines starren Objektes zu interpretieren, und so zu einer eindeutigen 3D-Interpretation gelangt. Seine sogenannte "rigidity assumption" in [Ullman 79, S. 146] lautet:

Any set of elements undergoing a two-dimensional transformation which has a unique interpretation as a rigid body moving in space should be interpreted as such a body in motion.

Unter Ausnutzung dieser Annahme gelang es ihm, eine hinreichende Bedingung für die Gewinnung einer starren 3D-Punktconfiguration zu formulieren. Sein "Structure from Motion" Theorem in [Ullman 79, S. 148] lautet:

Given three distinct orthographic views of four non-coplanar points in a rigid configuration, the structure and motion compatible with the three views are uniquely determined up to a reflection about the image plane.

Ullman verfährt dazu wie folgt: Gegeben seien die Punkte O, A, B, C in drei orthographischen Ansichten. Es wird vorausgesetzt, daß die Korrespondenzen bekannt sind. Ferner wird bei der Analyse von einer stationären Szene und einer bewegten Kamera ausgegangen. Ein äquivalentes Problem wäre der Fall einer stationären Kamera und eines bewegten Objekts. Die relative Bewegung von Objekt zu Kamera ist in beiden Fällen identisch. Jedes Bild Π_i besitze ein eigenes Bildkoordinatensystem mit den orthonormalen Basisvektoren p_i, q_i und dem Bild des Punktes O im Ursprung. Aus den gegebenen Daten lassen sich nun die Schnittgeraden und damit die Lage der Projektionsebenen bestimmen. Ist die Lage jeder Projektionsebene bekannt, so ergeben sich die Raumpunkte durch die Schnittpunkte der Projektionsstrahlen. Aus der relativen Lage der berechneten Projektionsebenen läßt sich die Bewegung der Kamera bzw. der vier Punkte im Raum ablesen. Die Unbekannten werden durch Lösen eines linearen Gleichungssystems bestimmt. Da orthographische Projektion vorausgesetzt wurde (Projektionsstrahlen von Punkten desselben Bildes sind parallel), ist der Translationsanteil in die Tiefe nicht bestimmbar, er wird daher bei der Berechnung auf null gesetzt. Weiterhin gibt es zwei spiegelsymmetrische Lösungen der Objektform und der zugehörigen Rotationsachse (Bildebene ist Spiegelebene). Eine Rotation der gewonnenen 3D-Form um den Winkel α entspricht einer Rotation um den Winkel $-\alpha$ der Spiegelbildlösung. Diese Mehrdeutigkeit der orthographischen Projektion kann sich beim Beobachter derartiger Bildfolgen durch eine plötzliche Tiefenumkehr bemerkbar machen.²

Liegt also ein starres Objekt (mindestens vier Punkte) vor, so ist die zugehörige Form und Bewegung des Objektes aus den drei Ansichten mit Ausnahme der obigen Einschränkungen eindeutig bestimmbar. Da das obige Theorem aber nur eine hinreichende Bedingung darstellt, besteht die Möglichkeit, daß bei einem nichtstarrten Objekt die Berechnung eine starre Konfiguration liefert. Nach Ullman 79 ist die Wahrscheinlichkeit

²Ein bekanntes Beispiel, bei dem dieses Phänomen beobachtet werden kann, ist die orthographische Ansicht eines Drahtrahmenwürfels ("Neckers Würfel").

für einen solchen Fall äußerst gering. Er kann nur bei speziellen Objektverformungen auftreten. Diese können im allgemeinen ausgeschlossen werden (Generalitätsprinzip).

Enthalten die Bilder mehrere Objekte, die sich unabhängig voneinander bewegen, oder sind die Objekte aus bewegten starren Einzelteilen zusammengesetzt, so können zunächst Punktgruppen zu je vier benachbarten Punkten gebildet werden. Für jede Gruppe wird eine starre 3D-Interpretation gesucht. Mehrere benachbarte Vierergruppen, die eine gemeinsame starre Interpretation besitzen, lassen sich danach zu größeren starren Objekten zusammenfassen. Im Randbereich von Objekten werden vielfach nichtstarre Punktconstellationen auftreten, falls die Punkte zu unterschiedlichen Objekten gehören. Hier muß für jeden Punkt entschieden werden zu welchem Objekt er gehört. Liegen die Punkte genügend dicht beieinander (z. B. beim optischen Fluß), so daß das Starrheitskriterium für jede Vierergruppe nur lokal gefordert werden muß, so ließe sich das Verfahren auch bei nichtstarrten Objekten anwenden.³ In einem solchen Fall wirken sich jedoch – als Folge der geringen Punktabstände – Ungenauigkeiten in der Punktlokalisierung stärker aus. Darüberhinaus erschweren die Mehrdeutigkeiten der orthographischen Projektion die Bewegungsbeschreibung eines nichtstarrten Objektes. Ohne vorherige Annahmen über die Bewegungsform scheint daher die Anwendung eines "Structure from Motion"-Verfahrens auf nichtstarre Objekte nicht sinnvoll zu sein.

Um die Mehrdeutigkeiten der orthographischen Projektion zu vermeiden, unternahm Ullman 79 auch Versuche mit perspektivischer Projektion. Bei dieser Projektionsart ist, wie wir bereits erwähnten, die Größe und Verschiebung der Objekte bis auf einen verbleibenden Skalierungsfaktor bestimmbar. Die Frage, welche Projektionsart die geeignetere ist, beurteilt Ullman 79 wie folgt:

Betrachtet man das menschliche Sehsystem, so fällt auf, daß die Bilder auf der Retina durch perspektivische Projektion hervorgerufen werden. Trotzdem ist der Mensch in der Lage, orthographisch projizierte Bildfolgen von bewegten Objekten richtig zu interpretieren (siehe oben). Darüberhinaus sind lokal, d. h. für kleine Bildregionen, beide Projektionsarten gleich. Ist die Entfernung von zwei Punkten zur Bildebene gegenüber ihrem räumlichen Abstand groß, so sind die Projektionsstrahlen der beiden Punkte bei perspektivischer Projektion nahezu parallel. Verlangt man von dem Berechnungsverfahren, daß es fehlertolerant gegenüber Lokalisationsfehlern der Punkte im Bild ist, so darf auch die gewählte Projektionsart keine gravierende Rolle spielen, da, wie oben bereits ausgeführt, die Unterschiede lokal vernachlässigbar sind. Dies rechtfertigt in lokalen Bildbereichen die Annahme von orthographischer Projektion.

Ullman 79 untersuchte zwei verschiedene Ansätze mit perspektivischer Projektion: Der erste Ansatz wurde von Ullman als "direkter perspektivischer" Ansatz bezeichnet, da er die perspektivische Projektion im Gegensatz zum zweiten Ansatz direkt behandelt. Die Nachteile bei dieser Vorgehensweise führten zu seinem zweiten sogenannten "polar/parallel" Ansatz, in dem Ullman 79 lokal orthographische Projektion voraussetzte.

³Jedes nichtstarre Objekt läßt sich in erster Näherung lokal als starr betrachten.

2.1.1 Direkter perspektivischer Ansatz

Mit seinem *direkten perspektivischen* Ansatz untersuchte Ullman 79 die folgenden Bewegungen:

1. Rotation um eine vertikale Achse mit beliebiger Translation.
2. Reine Translation, d. h. keine Rotation.
3. Uneingeschränkte Bewegung.

In den ersten beiden Fällen genügen drei Punkte in zwei Ansichten. Aus den Bewegungsgleichungen dieser drei Punkte leitet Ullman sechs nichtlineare Gleichungen ab, die die sechs unbekanntenen Größen enthalten. Diese sind drei Translationskomponenten, zwei Punkttiefen (die dritte Tiefe ist aufgrund des unbestimmbaren Skalierungsfaktors frei wählbar) und ein Rotationswinkel. Die sechs Gleichungen lassen sich durch Substitution auf eine einzige Gleichung reduzieren, die nur noch den Rotationswinkel als unbekanntene Größe enthält. Diese sogenannte Polargleichung ist eine Gleichung vierten Grades und hat somit im allgemeinen vier Lösungen. Um eine eindeutige Lösung zu erhalten, müssen mehrere Polargleichungen für jeweils Dreierpunktgruppen einer starren Konfiguration gelöst werden. Die Schnittmenge der einzelnen Lösungsmengen liefert die eindeutige Lösung. Mit dem so erhaltenen Rotationwinkel lassen sich die restlichen Unbekannten unter Zuhilfenahme der Bildkoordinaten direkt berechnen.

Im Falle einer reinen Translation (Rotationswinkel $\theta = 0$) vereinfacht sich die Polargleichung zu einer Gleichung, die nur noch meßbare Größen (Bildkoordinaten der Punkte) enthält. Diese Gleichung läßt sich zum Test auf reine Translation heranziehen. Sie ist zwar nur eine notwendige und keine hinreichende Bedingung für eine reine Translationsbewegung, jedoch stellte sich dieser Test bei den Experimenten von Ullman 79 als sehr zuverlässig heraus. Auch hier kann man sich wieder auf das Generalitätsprinzip berufen. Die Form und Bewegung ergeben sich wie oben direkt aus den Bildkoordinaten.

Den Fall uneingeschränkter Bewegung behandelt Ullman 79 als Verallgemeinerung der obigen Methode. Da die Orientierung der Rotationsachse nicht bekannt ist, müssen viele Polargleichungen gelöst werden, um die richtige Rotationsachse zu finden. Diese Vorgehensweise erfordert einen erheblichen Suchaufwand und wurde daher von Ullman 79 als nicht praktikabel und ineffizient bezeichnet. Er benutzte sie vielmehr um herauszufinden, wieviele Punkte in wievielen Ansichten benötigt werden. Er fand heraus, daß fünf Punkte in drei Ansichten genügen, um die Form und Bewegung der Objekte zuverlässig bestimmen zu können.

Da die Starrheitsannahme nur lokal als allgemeingültig angesehen werden kann, sollte man sich bei der Bestimmung von Form und Bewegung auf lokale Bildbereiche beschränken. In ihnen sind die perspektivischen Effekte allerdings gering, so daß das Verfahren auf sehr genaue Eingabedaten und damit auf genaue Lokalisation der Punkte im Bild angewiesen ist. Daraus ergibt sich eine starke Anfälligkeit gegenüber geringen Fehlern in der Lokalisation der Punkte. Daher ist der *direkte perspektivische* Ansatz für die Bestimmung von "Structure from Motion" bei der die Starrheit vorausgesetzt wird,

weniger gut geeignet. Experimente haben weiterhin gezeigt, daß der bisher vorgestellte perspektivische Ansatz im Falle von flüchtender Bewegung, bei der die Translation parallel zur optischen Achse verläuft, dem menschlichen Sehsystem überlegen ist. Das menschliche Sehsystem hat in solchen Fällen Schwierigkeiten, die Form und Bewegung eindeutig wahrzunehmen. Diese Überlegenheit wird allerdings mit einer starken Empfindlichkeit gegenüber Fehlern und einem komplexen Algorithmus erkauft. Diese Eigenschaften haben Ullman 79 veranlaßt, seine Methode zu vereinfachen.

2.1.2 Polar/parallel Methode

Bei der sogenannten *polar/parallel* Methode nahm er innerhalb lokaler Bildregionen orthographische Projektion an. Diese Annahme läßt sich, wie bereits erwähnt, dadurch rechtfertigen, daß die Projektionsstrahlen von dicht beieinanderliegenden Punkten nahezu parallel sind und damit der Unterschied zwischen den Projektionsarten vernachlässigbar wird. Die Richtung der Projektionsstrahlen für eine Bildregion ergibt sich jeweils aus der Verbindung des optischen Zentrums mit dem Mittelpunkt der betreffenden Region, wodurch verschiedene Regionen unterschiedliche Projektionsstrahlen haben. Werden die Ergebnisse der einzelnen Regionen zusammengefaßt, um zu einer globalen Interpretation zu gelangen, so können die Mehrdeutigkeiten der orthographischen Projektion aufgelöst werden. Das heißt, es kann die Spiegelbildlösung von der richtigen Lösung unterschieden werden, und es kann unter Umständen die Bewegung in die Tiefe bestimmt werden.

Um zu verstehen, wie sich die richtige Lösung von der Spiegelbildlösung unterscheiden läßt, müssen wir uns vor Augen halten, daß die zwei Lösungen, die wir innerhalb eines Bildbereiches erhalten, spiegelsymmetrisch sind bezüglich einer Ebene senkrecht zum Projektionsstrahl (Bei orthographischer Projektion ist dies die Bildebene). Unterschiedliche Bildbereiche, die dasselbe starre Objekt zeigen, sollten in einer Lösung übereinstimmen, damit das Objekt eine eindeutige Rotationsachse besitzt. Tun sie es, so müssen aufgrund der unterschiedlichen Spiegelebenen die zweiten Lösungen sich voneinander unterscheiden, so daß zwei verschiedene Bereiche nur eine gemeinsame Lösung haben können. Damit läßt sich durch Vergleich mehrerer Bildbereiche desselben starren Objektes die Orientierung der Rotationsachse eindeutig bestimmen. Je weiter die beiden Bildbereiche auseinander liegen, desto stärker unterscheiden sich die falschen Lösungen, so daß sich die Rotationsachse umso leichter bestimmen läßt, je größer der von dem Objekt eingenommene Teil des Bildes ist.

Flüchtende Bewegung, die parallel zur optischen Achse verläuft, ist zwar prinzipiell bestimmbar, jedoch treten bei dieser Bewegungsart die größten Fehler auf. Damit läßt sich diese Bewegung am schwersten erkennen. Der Grund hierfür liegt darin, daß die Änderung der Objektgröße im Bild (Kontraktion bzw. Expansion), durch die die flüchtende oder nähernde Bewegung erkannt werden kann, bei der *polar/parallel* Methode als Rauschen empfunden wird. In dieser Eigenschaft ähnelt die Methode dem menschlichen Sehsystem, welches ebenfalls Schwierigkeiten hat, derartige Bewegungen zu interpretieren.

2.2 Weitere Arbeiten zu "Structure from Motion"

Roach + Aggarwal 80 bestimmen die Form und die Bewegung der Objekte aus zwei perspektivischen Ansichten. Dazu benötigen sie mindestens fünf korrespondierende Punkte. Die 27 unbekannt Parameter (Kamerastandorte, Kameraorientierungen und 3D-Koordinaten der fünf Punkte) werden durch die perspektivischen Projektionsgleichungen der fünf Punkte in Beziehung gesetzt zu den gegebenen Bildkoordinaten. Da man pro Punkt und Bild jeweils zwei Projektionsgleichungen erhält, ergeben sich 20 nichtlineare Gleichungen mit 27 Unbekannten. Um eine eindeutige Lösung zu ermöglichen, müssen folglich mindestens sieben Unbekannte vorgegeben werden. Roach + Aggarwal 80 erreichen dies, indem sie den Kamerastandort und die Rotationswinkel der ersten Kamera auf null setzen. Weiterhin setzen sie die z -Komponente eines Punktes, den sie als Referenzpunkt bezeichnen, auf einen konstanten Wert. Durch diese Festlegung sind die 3D-Koordinaten dieses Punktes bestimmt, wodurch zwei Gleichungen, die die Projektion des Referenzpunktes in Bildebene 1 beschreiben, nur bekannte Größen enthalten. Gleichzeitig ist der Skalierungsfaktor der perspektivischen Projektion festgelegt. Es verbleiben folglich nur 18 nichtlineare Gleichungen mit 18 Unbekannten.

Für nichtlineare Gleichungssysteme gibt es leider keine vergleichbare geschlossene Lösung, wie es für lineare Gleichungssysteme der Fall ist. Daher kommen hier meistens iterative Verfahren zur Anwendung, mit deren Hilfe man eine approximative Lösung sucht. Weiterhin ist bei nichtlinearen Gleichungen die Existenz einer eindeutigen Lösung nicht gewährleistet.

Roach + Aggarwal 80 benutzen ein modifiziertes "Least-Squared-Error"-Verfahren, welches die Unbekannten so bestimmt, daß der quadratische Fehler der 18 Gleichungen minimiert wird. Die Methode benutzt eine Glättungsoperation, durch die Nebenminima vermieden werden sollen. Damit wird das Auffinden der korrekten Lösung erleichtert. Die Methode ist iterativ und erfordert eine Anfangsschätzung der unbekannt Parameter. Eine gute Anfangsschätzung garantiert schnelle Konvergenz und reduziert das Risiko, auf ein Nebenminima zu laufen. Die Schätzung erhalten Roach + Aggarwal 80, indem sie annehmen, daß die Bildebenen koplanar sind. Der Translationsvektor der Kameras kann dann mit Hilfe des gewählten Referenzpunktes bestimmt werden. Ist der Translationsvektor bekannt, so sind die Kameraparameter eindeutig festgelegt, so daß sich die 3D-Koordinaten der verbleibenden vier Punkte ebenfalls bestimmen lassen. Für jeden möglichen Referenzpunkt erhalten sie eine Schätzung der festzulegenden Parameter. Aus diesen Parametern wählen sie diejenigen aus, deren Summe der quadratischen Fehler zwischen den wahren Bildkoordinaten und den projizierten Koordinaten der geschätzten 3D-Punkte minimal ist.

Roach + Aggarwal 80 testen ihre Methode zunächst mit künstlich erzeugten Bildern. Da die 3D-Koordinaten der Punkte genau bekannt sind, können sie ihre Ergebnisse direkt bewerten. Bei rauschfreien Daten ermittelt ihr Verfahren stets die richtigen Ergebnisse. Schwierigkeiten treten allerdings bei speziellen Ansichten auf. So schlägt ihr Verfahren fehl, wenn die Objekte um eine Rotationsachse rotieren, die parallel zur Projektionsachse verläuft, und sie weiterhin keiner Translation unterliegen.

Da ihr Verfahren eine quadratische Fehleranpassung durchführt, ist es auch für vertauschte Daten gut geeignet. Die Unbekannten müssen dann allerdings genügend überbestimmt sein. So zeigte sich in ihren Versuchen, daß bei zwei Ansichten mindestens 12 bis 15, bei drei Ansichten mindestens 7 bis 8 Punkte benötigt wurden.

Webb + Aggarwal 82 bestimmen die Form und Bewegung von starren, miteinander verbundenen Objekten aus parallel projizierten Bildern. Ihr Ansatz unterscheidet sich von den bisherigen darin, daß bereits zwei Punkte in einer starren Konfiguration genügen, um deren Struktur und Bewegung zu bestimmen. Zu dieser Fähigkeit soll auch das menschliche Sehsystem in der Lage sein. Da aus mathematischer Sicht die Starrheit von zwei Punkten für eine eindeutige Bestimmung eine zu geringe Einschränkung darstellt, machen Webb + Aggarwal 82 die Annahme, daß die Rotationsachse für einen kürzeren Zeitraum ihre Lage nicht verändert. Diese sogenannte "fixed-axis-assumption" trifft für viele Bewegungen von Objekten aus unserer Umwelt zu. Darüberhinaus nutzen Webb und Aggarwal die Verbindung von starren Teilen aus, um die eindeutige räumliche Lage der Einzelteile zu bestimmen. Ist ihre Annahme einer festen Rotationsachse gerechtfertigt, so läßt sich die Bewegung der Objektpunkte relativ zu einem beliebig gewählten Referenzpunkt auf der Rotationsachse durch einen Kreis beschreiben. Diese räumlich kreisförmige Bewegung ergibt im Bildkoordinatensystem eine Bewegung, die auf einer Ellipse verläuft. Aus der Form dieser Ellipse, d. h. aus der Orientierung der Halbachsen sowie dem Verhältnis der Achsenlängen, läßt sich die Lage der Rotationsachse sowie die Struktur der Objekte bestimmen. Da Parallelprojektion verwendet wird, kann der Translationsanteil in die Tiefe nicht bestimmt werden. Ferner müssen bestimmte Anforderungen an die Bewegung gestellt werden, damit die Methode anwendbar ist. So liefern reine Translationsbewegungen und Bewegungen, bei denen das Objekt um eine senkrecht zur Bildebene verlaufende Achse rotiert, keine räumliche Information über die Objekte, da die Bilder keine verschiedenen Ansichten des Objektes darstellen. Die Bilder können durch Verschiebung innerhalb der Bildebene ineinander überführt werden. Weiterhin darf die Rotationsachse nicht parallel zur Bildebene verlaufen, da in einem solchen Fall die kleinen Halbachsen der Ellipsen gleich null werden und die Ellipsen damit im Bild degenerieren. Die einzelnen Stufen ihres Algorithmus sind:

1. Finden von starren Einzelteilen durch Ellipsenanpassung an die zeitlichen Punkt-korrespondenzen im Bildkoordinatensystem.
2. Suchen von sichtbaren Verbindungspunkten. Ein Verbindungspunkt ist ein Punkt, der zu mehreren starren Teilen gehört.
3. Vorhersage von unsichtbaren Verbindungspunkten.
4. Ausnutzung der Verbindungspunkte, um die Mehrdeutigkeiten der räumlichen Lage und Bewegung der Einzelteile aufzulösen.

Ihr Algorithmus liefert befriedigende Ergebnisse, wenn die Eingangsdaten geringe Fehler aufweisen und die Bewegungen der starren Teile die Annahme von festen Rotationsachsen

erfüllen. Bei fehlerbehafteten Daten und Bewegungen, die der Annahme einer festen Rotationsachse nicht genügen, schlägt ihr Algorithmus häufig fehl.

Tsai und Huang bestimmen in [Tsai + Huang 81, Tsai + Huang 82] die Bewegungsparameter von starren ebenen Flächen und von Objekten mit gekrümmten Oberflächen aus acht korrespondierenden Punkten in zwei perspektivischen Ansichten. Ihr Verfahren unterscheidet sich von den bisherigen in folgenden Punkten:

1. Das Verfahren erfordert nur das Lösen von linearen Gleichungen. Dadurch können Tsai und Huang einen Beweis für eine eindeutige Lösung erbringen. Die bisherigen Ansätze für perspektivische Projektion führen alle auf nichtlineare Gleichungen, für die keine eindeutige Lösung sichergestellt ist.
2. Bisherige Methoden erfordern spezielle Annahmen, um die Berechnung zu vereinfachen, wie z. B.:
 - Parallelprojektion [Ullman 79, Webb + Aggarwal 82]
 - Bewegungseinschränkung:
 - "fixed-axis-assumption" [Webb + Aggarwal 82]
 - Rotation um eine vertikale Achse mit beliebiger Translation oder reine Translation [Ullman 79]

Aus acht korrespondierenden Punkten in zwei perspektivischen Ansichten bestimmen Tsai und Huang die Koeffizienten der Matrix

$$\mathbf{E} = \begin{pmatrix} e_1 & e_2 & e_3 \\ e_4 & e_5 & e_6 \\ e_7 & e_8 & e_9 \end{pmatrix}$$

für die gilt:

$$\begin{aligned} \begin{pmatrix} x'_i & y'_i & 1 \end{pmatrix} \mathbf{E} \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} &= 0 & (2.3) \\ \Rightarrow \mathbf{n}'_i{}^T \mathbf{E} \mathbf{n}_i &= 0 & \forall i = 1, \dots, 8 \end{aligned}$$

Die Größen x_i , y_i und x'_i , y'_i sind die Bildkoordinaten von zeitlich korrespondierenden Bildpunkten. Die Vektoren \mathbf{n}_i und \mathbf{n}'_i stellen die Abbildungsstrahlen dieser Bildpunkte dar. Die Matrix \mathbf{E} drückt die folgende geometrische Beziehung zwischen diesen Strahlen und der Rotationsmatrix \mathbf{R} sowie dem Translationsvektor \mathbf{t} aus.

$$\begin{aligned} \mathbf{n}'_i{}^T (\mathbf{t} \times \mathbf{R} \mathbf{n}_i) &= 0 \\ \Rightarrow \mathbf{n}'_i{}^T \mathbf{E} \mathbf{n}_i &= 0 \end{aligned}$$

Diese Gleichung bringt zum Ausdruck, daß die drei Vektoren \mathbf{t} , $\mathbf{R} \mathbf{n}_i$ und \mathbf{n}'_i koplanar sind. Die Matrix \mathbf{E} enthält alle Informationen, die man aus den gegebenen zeitlichen Punktkorrespondenzen

$$(x_i \ y_i) \rightarrow (x'_i \ y'_i)$$

erhalten kann. Ihre Koeffizienten werden daher von Tsai und Huang als "essential parameters" bezeichnet. Da durch Multiplikation der Matrix \mathbf{E} mit einem Skalar die Gleichung nicht beeinflußt wird, sind die Koeffizienten von \mathbf{E} bis auf einen verbleibenden Skalierungsfaktor bestimmbar. Dieser Skalierungsfaktor beeinflußt lediglich die Translationskomponenten und ist, wie bereits mehrfach erwähnt wurde, eine unvermeidliche Mehrdeutigkeit der perspektivischen Projektion. Ein Koeffizient kann also ohne Beschränkung der Allgemeinheit frei gewählt werden. Nimmt man acht korrespondierende Punkte, so erhält man nach Gleichung 2.3 entsprechend acht lineare Gleichungen mit den Unbekannten e_1, \dots, e_8 . Aus diesen sogenannten "essential parameters" lassen sich die Bewegungsparameter durch Lösen einer polynomiellen Gleichung sechster Ordnung eindeutig bestimmen. Zur eindeutigen Berechenbarkeit der Matrix \mathbf{E} , müssen die Objektpunkte zu beiden Zeitpunkten zusammengenommen die folgenden Bedingungen erfüllen:

1. Die Punkte liegen nicht auf zwei Ebenen, von denen eine den Ursprung enthält.
2. Die Punkte liegen nicht auf einem Kegel, der den Ursprung enthält.

Sind beide Bedingungen erfüllt, so ist die Bestimmung der "essential parameters" eindeutig und damit auch die Berechnung der Bewegungsparameter. Aus den berechneten Bewegungsparametern läßt sich dann direkt auf die 3D-Koordinaten der Punkte schließen.

Williams 80 analysiert Bilder, die von einer geradlinig in Richtung der optischen Achse bewegten Kamera geliefert werden. Die Szene, durch die sich die Kamera bewegt, ist stationär. Die Bilder werden zunächst in Regionen segmentiert, wobei jede Region als Bild einer ebenen Fläche interpretiert wird. Für die Orientierung einer ebenen Fläche im Raum erlaubt Williams 80 nur die beiden folgenden Möglichkeiten:

1. Fläche ist parallel zur Bildebene (XY -Ebene), d. h. vertikal.
2. Fläche ist parallel zur XZ -Ebene, d. h. horizontal (Z -Achse weist in Richtung der optischen Achse).

Derartige Flächen lassen sich durch die Angabe der z -Koordinate (Tiefe) bzw. durch Angabe der y -Koordinate (Höhe) genügend charakterisieren. Zu Anfang werden alle Flächen als gleich tief bzw. gleich hoch angenommen. Durch Messung der Punktverschiebung im Bild werden die Angaben der Flächenhöhen und -tiefen verfeinert (inkrementieren, beibehalten oder dekrementieren der Werte). Dabei werden die Werte derart verändert, daß sie mit der durch die Translationsbewegung der Kamera hervorgerufenen Verschiebung im Bild im Einklang sind. Für die Verfeinerung des Szenenmodells setzt Williams eine "hypothesize and test" Strategie ein.

Nagel + Neumann 81 betrachten fünf Punkte eines starren Körpers in zwei perspektivischen Ansichten. Wie Tsai + Huang 81, Tsai + Huang 82 (siehe Seite 12) nutzen sie in ihren aufgestellten Gleichungen die geometrische Beziehung zwischen dem Translationsvektor \mathbf{t} , der Rotationsmatrix \mathbf{R} und den Abbildungsstrahlen \mathbf{n}_i und \mathbf{n}'_i ($i = 1, \dots, 5$) der korrespondierenden Bildpunkte aus. Diese Beziehung bringen sie in einer alternativen Form zum Ausdruck. Ihre Gleichungen drücken aus, daß die von den Vektorpaaren

$(\mathbf{Rn}_i, \mathbf{n}'_i)$ aufgespannten Ebenen sich alle in dem Translationsvektor \mathbf{t} schneiden. Die Ebenennormalen sind damit koplanar. Ihre Gleichungen lauten:

$$((\mathbf{Rn}_1 \times \mathbf{n}'_1) \times (\mathbf{Rn}_2 \times \mathbf{n}'_2))(\mathbf{Rn}_k \times \mathbf{n}'_k) = 0 \quad k = 3, \dots, 5 \quad (2.4)$$

Diese drei nichtlinearen Gleichungen enthalten die Polargleichung von Ullman als Spezialfall. Sie lassen sich in Polynome vierten Grades transformieren, die mit einem iterativen Verfahren gelöst werden können. Als Lösung ergeben sich die drei unbekannt Rotationsparameter. Durch das Lösen der Gleichungen wird ferner die Richtung der Translation eindeutig festgelegt.

Die Rekonstruktion schlägt fehl, wenn die Bewegung der Punkte durch eine reine Rotation um das Kamerazentrum beschrieben werden kann, so daß keine Translation vorliegt, oder wenn die Translation entlang des Abbildungsstrahls eines Punktes verläuft.

Bonde + Nagel 79 rekonstruieren die Form und Bewegung eines starren Objektes aus einer monokularen Bildfolge. Die aufgenommene Straßenszene, in der ein Taxi in eine Straße einbiegt, erlaubt es, die räumliche Bewegung auf eine Ebene zu beschränken. Die Rotationsachse muß damit senkrecht zu dieser Ebene sein. Das Weltkoordinatensystem wird so gelegt, daß die XY-Ebene koplanar zur Straße ist. Damit erfolgt die Rotation um die Z-Achse. Die Neigung der Kamera, d. h. der Winkel zwischen optischer Achse und Bewegungsebene wurde mittels Kamerakalibrierung bestimmt. Unter Ausnutzung der obigen Bewegungseinschränkung ergibt sich ein nichtlineares Gleichungssystem, welches die unbekannt Tiefen der Punkte bezüglich des Objektkoordinatensystems und die unbekannt Rotationswinkel enthält. Mit einem numerischen Ansatz von Fletcher-Powell werden die Unbekannt derart bestimmt, daß die 3D-Koordinaten bezüglich des gewählten Objektkoordinatensystems konstant bleiben und damit ein starres Objekt bilden. Dazu wird die Summe der quadratischen Abweichungen minimiert. Die Einschränkungen ergeben weniger Unbekannte, was zuverlässigere Ergebnisse bei der Lösungssuche gewährleistet.

2.3 Nichtstarre Bewegung

Um auch nichtstarre Bewegung zu berücksichtigen, schlägt Ullman in [Ullman 83] einen inkrementellen Ansatz zur Lösung von "Structure from Motion" vor. Er verwendet ein sogenanntes internes Modell, welches das aktuelle Wissen über die 3D-Form des Objektes repräsentiert. Jede orthographische Ansicht liefert neue Information über die Form des betrachteten Objektes. Diese Information wird zur Anpassung des Modells an die wahre Objektform verwendet. Die Anpassung kann dabei aus zwei unterschiedlichen Gründen erfolgen:

1. Das Modell stimmt nicht mit der wahren Objektstruktur überein. Die zusätzliche Ansicht hat neue Informationen über das Objekt geliefert (starre Bewegung).
2. Das Objekt hat sich nicht starr bewegt. Das Modell muß der neuen Objektform angepaßt werden.

Nach jeder neuen Ansicht versucht Ullman, die Bewegung des Objektes so starr wie möglich zu beschreiben. Dazu verfährt er wie folgt:

Gegeben sei die orthographische Projektion I der n Punkte p_1, \dots, p_n . Die XZ -Ebene des Weltkoordinatensystems sei die Bildebene, die Y -Achse sei die optische Achse und zeige vom Betrachter weg (rechtshändiges Koordinatensystem). $M(t)$ stelle das interne Modell des Objektes zum Zeitpunkt t dar. Es enthält die angenommenen Weltkoordinaten der n Punkte p_1, \dots, p_n . Die x - und z -Koordinaten der Punkte sind durch die Bildkoordinaten gegeben. Die y -Koordinaten und damit die Tiefen der Punkte werden am Anfang geschätzt oder stammen aus anderen Informationsquellen. Damit liegen die räumlichen Abstände der Punkte untereinander fest. Der räumliche Abstand zwischen den Punkten p_i und p_j nach dem bisherigen Modell zur Zeit t sei mit $l_{ij}(t)$ bezeichnet. Für eine neue orthographische Ansicht I' zum Zeitpunkt t' sind nun die Tiefen der Punkte y'_i so zu wählen, daß die Punktabstände $l_{ij}(t)$ möglichst erhalten bleiben. Hierdurch wird eine starre Transformation bevorzugt. Seien die sich ergebenden räumlichen Punktabstände für gewählte Tiefen der neuen Ansicht mit $l_{ij}(t')$ bezeichnet, und sei die Funktion $D(l_{ij}(t), l_{ij}(t'))$ ein Maß für die Abweichung zwischen $l_{ij}(t)$ und $l_{ij}(t')$, so ist die folgende Funktion zu minimieren:

$$\begin{aligned} \mathcal{D}(M(t), M(t')) &= \sum_{i,j} D(l_{ij}(t), l_{ij}(t')) \\ &= \sum_{i,j} \frac{(l_{ij}(t) - l_{ij}(t'))^2}{l_{ij}(t)^3} \\ &= \sum_{i,j} \frac{1}{l_{ij}(t)} \frac{(l_{ij}(t) - l_{ij}(t'))^2}{l_{ij}(t)^2} \end{aligned} \quad (2.5)$$

Für die gegebenen Bildkoordinaten (x'_i, z'_i) der Ansicht I' , sind die Punkttiefen y'_i so zu wählen, daß die Summe der quadrierten relativen Differenzen der Punktabstände minimal ist. Zusätzlich wird durch den Abstand $l_{ij}(t)$ dividiert, wodurch dicht zusammenliegende Punkte stärker gewichtet werden. Der Grund hierfür liegt in der Annahme, daß weit voneinander entfernte Punkte eher zu verschiedenen Objekten gehören, die sich unabhängig bewegen, so daß eine starre Konstellation nicht gerechtfertigt ist. Die Funktion nimmt im Falle einer starren Bewegung den Wert null an. Die sich ergebenden Tiefen y'_i bilden zusammen mit den Bildkoordinaten (x'_i, z'_i) das neue Modell $M(t')$. Mit der nächsten orthographischen Ansicht und dem neuen Modell wird der beschriebene Vorgang wiederholt.

Wie wir bei *Webb + Aggarwal 82* bereits gesehen haben (siehe Seite 11), liefern eine reine Translation sowie eine Rotation um die optische Achse keine neuen Informationen über die Form des Objektes, da die Bildpunkte bei derartigen Bewegungen ihre relative Konstellation nicht ändern. Daher kann nur eine Rotation des Objektes bzw. der Kamera um eine nicht zur optischen Achse parallel verlaufenden Achse einen Beitrag zur Verbesserung des Objektmodells liefern.

Um dem Fall, daß das Modell nicht mit der wahren Objektform übereinstimmt (siehe Fall 1 auf Seite 14) besser Rechnung zu tragen, schlägt Ullman eine geänderte Version

vor, in der er bei jeder Minimierung zunächst Korrekturen des Modells $M(t)$ vornimmt und für das resultierende Modell $M'(t)$ eine möglichst starre Bewegung sucht. Die zu minimierende Funktion \mathcal{E} sieht wie folgt aus:

$$\mathcal{E} = \mathcal{D}(M(t), M'(t)) + \mathcal{D}(M'(t), M(t'))$$

Mit dieser Methode lassen sich zusätzliche 3D-Informationen über $M'(t)$ in die Modellmodifikation einbeziehen. Ullman stellte bei seinen Tests fest, daß die Objektform bei starrer Bewegung mit der neuen Fehlerfunktion \mathcal{E} zwar schneller gefunden werden kann, das Modell aber stark um die korrekte Objektform oszilliert. Das Konvergenzverhalten ist damit nicht monoton.

Interessant für parallele Berechnungen ist das von ihm beschriebene, analoge mechanische Modell. Man stelle sich die räumlichen Punkte paarweise durch Hooke'sche Federn⁴ in Gleichgewichtslage verbunden vor. Entsprechend der oben verwendeten Distanzfunktion $D(l_{ij}(t), l_{ij}(t'))$ kann die Federkonstante mit zunehmendem Punktabstand verkleinert werden. Gesucht ist nun eine Bewegung, durch die die Deformationsenergie E des Federmodells minimiert wird. Bei starrer Bewegung ist die Energie gleich null. Die Energie des Modells ergibt sich aus der Summe der Federenergien, wobei die Feder zwischen den Punkten p_i und p_j die Federkonstante K_{ij} habe:

$$E = \frac{1}{2} \sum_{i,j} K_{ij} (l_{ij}(t) - l_{ij}(t'))^2$$

Diese Gleichung ist äquivalent zu Gleichung 2.5.

Kommen wir nun zu den Versuchsergebnissen, die Ullman mit seinen Verfahren erzielte: Ullman untersuchte seinen Ansatz an künstlich erzeugten Bildern von sich starr verhaltenden Objekten. Für seine Untersuchungen standen ihm damit exakte Bildpunktkoordinaten zur Verfügung. Er fand heraus, daß sich die Bilder genügend stark voneinander unterscheiden müssen, damit das Verfahren konvergiert. Genauer gesagt, ist eine genügend große Winklerschiebung notwendig, um stabiles Konvergenzverhalten zu gewährleisten. Ullmans Versuche bringen dies in sehr krasser Form zum Ausdruck. Eine künstliche Bildfolge eines rotierenden starren Körpers wurde zum Test des Verfahrens herangezogen. Wurde nun die Bildfolge stärker aufgelöst, d. h. dieselbe Bewegungsfolge mit mehr Bildern pro Zeitintervall verwendet, so ergaben sich zunehmend schlechtere Ergebnisse. Diese überraschende Eigenschaft zeigt, daß die Modellkorrekturen bei kleineren Punktverschiebungen überproportional schwächer werden. Ein Grund für diese Eigenschaft mag darin liegen, daß die Suche nach einer starren Bewegung die Anpassungsfähigkeit des Modells stark einschränkt. Damit können mit dem Verfahren nur dann gute Ergebnisse erzielt werden, wenn das Modell nur geringfügig von der Objektform abweicht oder aber die Bilder sich derart stark unterscheiden, daß eine Modellanpassung trotz Suche

⁴Derartige Federn gehorchen dem Hooke'schen Gesetz. Es besagt daß die rücktreibende Kraft proportional zur Auslenkung und damit die Deformationsenergie proportional zum Quadrat der Auslenkung ist.

nach starrer Bewegung genügend groß ausfällt. Weiterhin waren selbst bei starrer Bewegung deutlich mehr Ansichten notwendig, um die Objektform zu bestimmen, als bei den direkten Ansätzen.

Für die Anwendung des inkrementellen Ansatzes auf realen Bildern bleibt zu berücksichtigen, daß zur Lösung des Korrespondenzproblems eine hohe zeitliche Auflösung der Bildfolge unumgänglich ist. Gerade bei solchen Bildfolgen ergeben sich jedoch beim inkrementellen Ansatz die größten Schwierigkeiten. Weiterhin müssen bei realen Bildern zusätzliche Rauscheffekte berücksichtigt werden, die zu ungenauen Punktlokalisationen im Bild führen. Nicht zuletzt diese Eigenschaften lassen den inkrementellen Ansatz für die Formbestimmung aus realen Bildfolgen als ungeeignet erscheinen.

Hildreth + Grzywacz 86 verallgemeinern Ullmans Verfahren auf den kontinuierlichen Fall von infinitesimal kleinen Zeitintervallen. Sie untersuchen, in wieweit sich das Modell der korrekten Objektform anpaßt. Als Maß für die Abweichung des Modells von der korrekten Objektform berechnen sie nach jeder neuen orthographischen Projektion die Summe der quadratischen Abweichungen zwischen den wahren räumlichen Punktabständen d_{ij} der Punkte p_i und p_j und den Punktabständen $l_{ij}(t)$ innerhalb des aktuellen Modells zur Zeit t . Für jede neue Ansicht wird wie bei Ullman eine Modellanpassung vorgenommen. Mit einer mathematischen Betrachtung beweisen sie, daß für einen Großteil von Bewegungen die Fehlerfunktion \mathcal{E} nicht gegen null konvergiert.

$$\mathcal{E} = \sum_{i,j} (d_{ij} - l_{ij})^2$$

Weiterhin beweisen sie für den diskreten Fall, daß mit zunehmender Winkeldifferenz aufeinanderfolgender Ansichten die Geschwindigkeit der Konvergenz zunimmt.

2.4 Zusammenfassung

"Structure from Motion"-Methoden basieren auf einer monokularen Bildfolge. Die räumliche Form und die Bewegung der Objekte kann nur dann aus den korrespondierenden Bildpunkten zurückgewonnen werden, wenn einschränkende Annahmen gemacht werden. In den meisten Fällen setzt man die Starrheit der Objekte voraus. Im letzten Abschnitt wurde der Versuch unternommen, auch nichtstarre Objekte zu berücksichtigen. Es ergaben sich jedoch schon bei den künstlich erzeugten Bildern von sich starr verhaltenden Objekten derartige Schwierigkeiten bei der Modellanpassung, daß eine allgemeine Anwendung auf nichtstarre Objekte nicht sinnvoll erscheint.

Im Gegensatz zu den bisher vorgestellten Verfahren werden im Sissy-Projekt Stereobildfolgen analysiert. Bestimmt man in einem Stereobildpaar die räumlichen Korrespondenzen von markanten Bildpunkten, so lassen sich mit Hilfe der Kameraparameter die räumlichen Koordinaten der realen Objektpunkte berechnen. Damit stehen uns zu jedem Zeitpunkt 3D-Koordinaten zur Verfügung. Haben wir für die Bildpunkte auch die zeitlichen Korrespondenzen bestimmt, so können wir die Bewegung der Objekte beschreiben, indem wir die räumlichen Objektpunktkonstellationen zu verschiedenen Zeitpunkten ver-

gleichen. Im Gegensatz zu den vorgestellten "Structure from Motion"-Ansätzen versetzt uns diese Vorgehensweise in die Lage, auch nichtstarre Körper behandeln zu können.

Trotz der Möglichkeit, zu jedem Zeitpunkt räumliche Koordinaten von Objektpunkten bestimmen zu können, wird man auch weiterhin "Structure from Motion"-Verfahren große Beachtung schenken müssen. Diese Notwendigkeit ergibt sich aus den Unzulänglichkeiten der Stereoskopie. Diese lassen sich durch zwei Extremfälle veranschaulichen: Objekte die eine geringe oder eine große Entfernung zu den beiden Kameras haben,⁵ bereiten Probleme bei der räumlichen Rekonstruktion.

Betrachten wir sehr nahe Objekte, so ergeben sich für korrespondierende Bildmerkmale große Disparitäten, was für eine genaue Berechnung der räumlichen Tiefe wünschenswert ist. Gleichzeitig aber unterscheiden sich die Objektansichten stark, so daß eine Zuordnung von korrespondierenden Bildmerkmalen durch das unterschiedliche Aussehen erschwert wird. Außerdem enthalten die Bilder unterschiedliche Objektteile, für die eine Zuordnung ohnehin nicht möglich ist (Verdeckung).

Der andere Extremfall sind weit entfernte Objekte. Hier unterscheiden sich die Ansichten nur geringfügig. Dies erleichtert zwar die Ermittlung der Korrespondenzen, da Objektteile in beiden Bildern nahezu gleich erscheinen. Gleichzeitig verringert sich jedoch die Disparität, so daß die Berechnung der Tiefe unsicher wird. Die Projektionsstrahlen der korrespondierenden Bildpunkte sind nahezu parallel. Der Schnittpunkt solcher Geraden ändert sich stark, wenn der Winkel zwischen den Geraden geringfügig schwankt. Damit ist die Bestimmung des Schnittpunktes bei derartigen Geraden stark fehleranfällig. In beiden Fällen versagt die stereoskopische Tiefenberechnung. Gelingt es, die zeitlichen Korrespondenzen zu ermitteln, so können "Structure from Motion"-Methoden zur Berechnung von Struktur und/oder Bewegung der Objekte eingesetzt werden. Die sich ergebenden Rekonstruktionsergebnisse können dann dazu verwendet werden, um die räumlichen Korrespondenzen ausfindig zu machen.

Umgekehrt können sich jedoch auch bei der räumlichen Rekonstruktion mittels zeitlicher Korrespondenzen Schwierigkeiten ergeben. Durch Abschattungen oder auch durch Rauschen können sich zeitlich korrespondierende Bildbereiche stark unterscheiden, was die Zuordnung erschwert. Ferner lassen sich auch bei zeitlichen Bildfolgen zwei Extremfälle aufzeigen, die den möglichen Anwendungsbereich begrenzen. In diesem Fall betrifft dies die Bewegung der Objekte. Nahe Objekte bewirken durch ihre Bewegung starke Änderungen im Bild, welche die Korrespondenzanalyse erschweren. Objekte, die sich nicht bewegen, rufen dagegen keine Änderungen in zeitlich aufeinanderfolgenden Bildern hervor, was die Suche der Korrespondenzen zwar enorm erleichtert, dafür aber keine neuen Informationen liefert. Gelingt es zu den einzelnen Zeitpunkten, die räumlichen Korrespondenzen zu ermitteln, so können die resultierenden stereoskopischen Rekonstruktionen dazu verwendet werden, das zeitliche Korrespondenzproblem im 3D-Raum zu lösen. *Jenkin + Tsotsos 86* stellen Glattheitsanforderungen an die Bewegung der einzelnen Punkte im Raum, um die Zuordnungsvielfalt einzuschränken.

⁵Ein Objekt befindet sich in einer geringen Entfernung zu den Kameras, wenn sein Abstand in der Größenordnung der Basislinie der beiden Kameras liegt. Ist der Abstand hingegen groß gegenüber der Basislinie, so besitzt das Objekt eine große Entfernung.

Beide Rekonstruktionsmethoden können also dazu dienen, um zum einen räumliche Informationen zu gewinnen und um zum anderen die Korrespondenzanalyse zu unterstützen. Beide Vorgehensweisen haben allerdings Grenzen. Während bei der Stereoskopie das Verhältnis zwischen Basislinie der Kameras zur Entfernung der Objekte für eine erfolgreiche Rekonstruktion maßgebend ist, betrifft dies bei der Betrachtung von monokularen Bildfolgen das Verhältnis zwischen Objektbewegung und Entfernung zur Kamera. Mißlingt die Stereoberechnung, weil einer der beiden genannten Extremfälle vorliegt, so kann eine Analyse der zeitlichen Korrespondenzen genau dann weiterhelfen, wenn die zurückgelegten Entfernungen der Objekte nicht in derselben Größenordnung liegen wie die Basislinie der Kameras.

Wir werden im weiteren Verlauf davon ausgehen, daß die räumlichen und die zeitlichen Korrespondenzen bereits ermittelt wurden, und uns mit der Frage beschäftigen, wie sich aus den zu unterschiedlichen Zeitpunkten berechneten räumlichen Koordinaten die Bewegungen der Punkte beschreiben lassen.

Eine gute Bewegungsbeschreibung sollte dazu geeignet sein, die Bewegung der Punkte oder der Objekte vorherzusagen. Diese Vorhersagen können der Korrespondenzanalyse dazu dienen, weitere Zuordnungen vorzunehmen. Bewegungsbeschreibung und Korrespondenzanalyse hängen damit voneinander ab. Die Bewegung läßt sich nur dann beschreiben, wenn Punktkorrespondenzen zugeordnet wurden, während die Korrespondenzanalyse ihrerseits die Bewegungsbeschreibung benötigt, um leichter Zuordnungen treffen zu können.⁶ Korrespondenzanalyse und Bewegungsbeschreibung sollten also nicht nacheinander, sondern nebeneinander verlaufen. Die Bewegungsvorhersagen beruhen dabei stets auf der Annahme, daß sich der betrachtete Punkt bzw. das betrachtete Objekt unverändert weiterbewegt. Gilt diese Annahme nicht, weil das Objekt seinen Bewegungszustand ändert, so sind die Bewegungsvorhersagen falsch. In diesem Fall könnte die Korrespondenzanalyse in die Irre geleitet werden. Verläßt sie sich nur auf diese Information, so kann sie falsche Zuordnungen vornehmen, die wiederum zu fehlerhaften Bewegungsbeschreibungen führen usw.

Hieraus ergeben sich zwei Folgerungen: Erstens muß man bei der Korrespondenzsuche mit Unstetigkeiten im Bewegungsverlauf rechnen, so daß man sich nicht vollständig auf die Bewegungsvorhersagen verlassen sollte, und zweitens benötigt die Korrespondenzanalyse weiteres Wissen, um sichere Zuordnungen vornehmen zu können. Mit diesem zusätzlichen Wissen ist sie in der Lage, eventuelle Ungereimtheiten in der Bewegungsvorhersage zu entdecken.

⁶Es wird deutlich, daß ein geeigneter Anfang gefunden werden muß (Initialisierung), in dem die Korrespondenzanalyse, ohne Unterstützung der Bewegungsbeschreibung, Kandidaten zuordnet. Auf diesen Korrespondenzen kann eine Bewegungsanalyse aufsetzen. Die gelieferten Bewegungsvorhersagen können die weitere Korrespondenzanalyse unterstützen und damit dazu beitragen, daß mehr Zuordnungen gemacht werden können usw. Das Problem einen geeigneten Anfang zu finden, wurde von *Dreschler-Fischer* 87 als "Bootstrap-Problem" bezeichnet.

Kapitel 3

Bewegungsbeschreibung einzelner Punkte

In dem nun folgenden Kapitel beschränken wir uns auf die Bewegungsbeschreibung einzelner Punkte. Diese Beschränkung kann immer dann nötig sein, wenn keine Kenntnisse über die Art des zugrundeliegenden Objekts existieren. Haben wir die 3D-Koordinaten eines Punktes zu aufeinanderfolgenden Zeitpunkten bestimmt, so versuchen wir nun, die räumliche Bewegung des Punktes zu beschreiben. Dazu suchen wir nach einer geeigneten Funktion, welche die Bewegungsbahn (Trajektorie) des Punktes im Raum wiedergibt. Eine derartige Funktion liefert uns eine kontinuierliche Verlaufsbeschreibung und kann darüberhinaus eine Vorhersage machen, wohin sich ein Punkt bewegen wird. Diese Vorhersage würde die Suche nach dem Punkt in späteren Bildern einschränken, was sowohl die Merkmalsextraktion als auch die Korrespondenzanalyse erleichtert.

3.1 Approximation der Punkttrajektorien

In dem nun folgenden Abschnitt werden wir uns damit beschäftigen, wie sich die Trajektorie eines Punktes funktional beschreiben läßt. Die dabei verwendeten Approximationsmethoden finden sich in [Engeln-Müllges + Reutter 88]; Ableitungsregeln für Matrizen finden sich in [Horn 86] (siehe auch den Anhang dieser Arbeit).

Um die Bewegung eines Raumpunktes \mathbf{p} kontinuierlich beschreiben zu können, müssen wir folgendes Problem lösen: Gegeben seien die Koordinaten $\mathbf{p}(t_0), \dots, \mathbf{p}(t_n)$ eines Raumpunktes \mathbf{p} zu den verschiedenen Zeitpunkten $t = t_0, \dots, t_n$. Gesucht ist eine Funktion $\Phi(t)$, die zu jedem Zeitpunkt $t \in [t_0, \dots, t_n]$ die 3D-Koordinaten des Punktes \mathbf{p} liefert. Die Funktion $\Phi(t)$ ist somit an endlich vielen diskreten Stellen bekannt, denn es gilt:

$$\Phi(t_i) = \begin{pmatrix} x(t_i) \\ y(t_i) \\ z(t_i) \end{pmatrix} = \mathbf{p}(t_i) \quad \forall i = 0, \dots, n \quad (3.1)$$

Im allgemeinen versucht man die Komponenten der unbekanntes Vektorfunktion $\Phi(t)$ durch eine Linearkombination von geeigneten Basisfunktionen φ_k darzustellen:

$$\Phi(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = \begin{pmatrix} c_{10} & c_{11} & \cdots & c_{1m} \\ c_{20} & c_{21} & \cdots & c_{2m} \\ c_{30} & c_{31} & \cdots & c_{3m} \end{pmatrix} \begin{pmatrix} \varphi_0(t) \\ \varphi_1(t) \\ \vdots \\ \varphi_m(t) \end{pmatrix} \quad \begin{array}{l} t \in [t_0, \dots, t_n] \\ c_{ij} \in \mathcal{R} \end{array}$$

Die unbekanntes Koeffizienten c_{ij} müssen folgendes Gleichungssystem erfüllen:

$$\begin{pmatrix} x(t_0) & x(t_1) & \cdots & x(t_n) \\ y(t_0) & y(t_1) & \cdots & y(t_n) \\ z(t_0) & z(t_1) & \cdots & z(t_n) \end{pmatrix} = \begin{pmatrix} c_{10} & \cdots & c_{1m} \\ c_{20} & \cdots & c_{2m} \\ c_{30} & \cdots & c_{3m} \end{pmatrix} \begin{pmatrix} \varphi_0(t_0) & \varphi_0(t_1) & \cdots & \varphi_0(t_n) \\ \varphi_1(t_0) & \varphi_1(t_1) & \cdots & \varphi_1(t_n) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_m(t_0) & \varphi_m(t_1) & \cdots & \varphi_m(t_n) \end{pmatrix}$$

$$\Rightarrow \mathbf{P} = \mathbf{C} \mathbf{B} \quad (3.2)$$

Für $n = m$ ist dieses Gleichungssystem eindeutig lösbar, da die Matrix \mathbf{B} quadratisch und außerdem, durch die Verwendung von linear unabhängigen Basisfunktionen, auch regulär ist. Die Gleichung 3.1 ist erfüllt, d. h. die resultierende Funktion $\Phi(t)$ geht genau durch die gegebenen Punkte $\mathbf{p}(t_0), \dots, \mathbf{p}(t_n)$. In diesem Fall spricht man von *Interpolation*.

Ist $n > m$, so ist das Gleichungssystem überbestimmt, da es mehr Gleichungen als Unbekannte enthält. Es braucht damit keine genaue Lösung zu existieren. Die Koeffizienten werden so bestimmt, daß ein gewähltes Fehlermaß minimiert wird. In diesem Fall spricht man von *Approximation*. Wir wollen annehmen, daß sich der betreffende Punkt "einfach" bewegt, so daß sich die Bewegung durch eine Approximation niedrigen Grades ($n \gg m$) gut beschreiben läßt. Eine Approximation wirkt darüberhinaus glättend auf zufällige Fehler der Eingabepunkte. In unserem Fall sind dies die ermittelten 3D-Koordinaten. Diese Fehler können z. B. durch ungenaue Korrespondenzen hervorgerufen worden sein. Diese Fehler würden bei einer Interpolation vollständig eingehen und damit zu einer Verfälschung der Bewegungsbeschreibung führen.

Der Fehler, den wir bei der Approximation machen, kann als Indiz dafür dienen, wie gut sich die Bewegung durch die gewählten Funktionen und den gewählten Grad m beschreiben läßt. Als Fehlermaß verwendet man in den meisten Fällen das gaußsche Fehlerquadratkriterium, bei dem der Fehler \mathcal{E} an der Summe der quadratischen Abweichungen gemessen wird:¹

$$\begin{aligned} \mathcal{E} &= \sum_{i=0}^n w_i (\mathbf{p}(t_i) - \Phi(t_i))^2 \\ &= \sum_{i=0}^n w_i \left(\mathbf{p}(t_i) - \mathbf{C} \begin{pmatrix} \varphi_0(t_i) \\ \varphi_1(t_i) \\ \vdots \\ \varphi_m(t_i) \end{pmatrix} \right)^2 \rightarrow \text{Minimum} \end{aligned}$$

¹Im weiteren verwenden wir auch die Bezeichnung "least-square-fit".

Die Koeffizienten w_i sind reelle Gewichtungsfaktoren, mit denen sich die einzelnen quadratischen Abweichungen unterschiedlich bewerten lassen. Konnten die 3D-Koordinaten eines Punktes genau bestimmt werden, so wird man einen großen Gewichtungsfaktor wählen. Umgekehrt wird man einen Punkt, bei dem sich die Raumkoordinaten nicht so genau bestimmen ließen, niedriger gewichten, so daß ihm eine geringere Bedeutung zukommt. Wie wir auf Seite 18 gesehen haben, lassen sich die 3D-Koordinaten eines Punktes mit zunehmendem Abstand zu den Stereokameras immer ungenauer rekonstruieren. Daher empfiehlt es sich, für die Gewichtungsfaktoren den Kehrwert des quadrierten Punktabstandes zur Kamera zu verwenden. Sind die Punkte bezüglich des Kamerakoordinatensystems angegeben, so setzt man $w_i = 1/\mathbf{p}(t_i)^T \mathbf{p}(t_i)$. Damit minimieren wir die Quadratsumme

$$\sum_{i=0}^n \frac{1}{\mathbf{p}(t_i)^T \mathbf{p}(t_i)} (\mathbf{p}(t_i) - \Phi(t_i))^2$$

der relativen Fehler. Sind alle 3D-Koordinaten mit denselben Fehlern behaftet, so setzt man die Koeffizienten w_i der Einfachheit halber alle auf eins. In Matrixschreibweise ergibt sich:

$$\begin{aligned} \mathbf{E} &= \mathbf{P} - \mathbf{C}\mathbf{B} \\ \mathbf{W} &= \begin{pmatrix} w_0 & & & 0 \\ & w_1 & & \\ & & \ddots & \\ 0 & & & w_n \end{pmatrix} \\ \mathcal{E} &= \text{Spur}(\mathbf{E}^T \mathbf{E} \mathbf{W}) \\ &= \text{Spur}((\mathbf{P} - \mathbf{C}\mathbf{B})^T (\mathbf{P} - \mathbf{C}\mathbf{B}) \mathbf{W}) \\ &= \text{Spur}(\mathbf{P}^T \mathbf{P} \mathbf{W} - \mathbf{P}^T \mathbf{C} \mathbf{B} \mathbf{W} - (\mathbf{C}\mathbf{B})^T \mathbf{P} \mathbf{W} + (\mathbf{C}\mathbf{B})^T \mathbf{C} \mathbf{B} \mathbf{W}) \\ &= \text{Spur}(\mathbf{P}^T \mathbf{P} \mathbf{W} - \mathbf{P}^T \mathbf{C} \mathbf{B} \mathbf{W} - ((\mathbf{P}\mathbf{W})^T \mathbf{C}\mathbf{B})^T + \mathbf{B}^T \mathbf{C}^T \mathbf{C} \mathbf{B} \mathbf{W}) \end{aligned}$$

Um das Minimum der Fehlerfunktion \mathcal{E} zu bestimmen, müssen wir nach der Matrix \mathbf{C} ableiten. Wir erhalten:

$$\begin{aligned} \frac{d\mathcal{E}}{d\mathbf{C}} &= \frac{d\text{Spur}(\mathbf{E}^T \mathbf{E} \mathbf{W})}{d\mathbf{C}} & (3.3) \\ &= -\mathbf{P}(\mathbf{B}\mathbf{W})^T - \mathbf{P}\mathbf{W}\mathbf{B}^T + \mathbf{C}\mathbf{B}\mathbf{W}\mathbf{B}^T + \mathbf{C}\mathbf{B}\mathbf{W}^T \mathbf{B}^T \\ &= -\mathbf{P}\mathbf{W}^T \mathbf{B}^T - \mathbf{P}\mathbf{W}\mathbf{B}^T + 2\mathbf{C}\mathbf{B}\mathbf{W}\mathbf{B}^T \\ &= -2\mathbf{P}\mathbf{W}\mathbf{B}^T + 2\mathbf{C}\mathbf{B}\mathbf{W}\mathbf{B}^T \end{aligned}$$

Hierbei wurde ausgenutzt, daß \mathbf{W} eine Diagonalmatrix ist, so daß $\mathbf{W} = \mathbf{W}^T$ gilt. Um die notwendige Bedingung für ein Minimum zu erfüllen, setzen wir die Ableitung der

Fehlerfunktion \mathcal{E} auf Null. Wir erhalten:

$$\begin{aligned} \frac{d\mathcal{E}}{d\mathbf{C}} &= \mathbf{0} \\ \Rightarrow \mathbf{CBWB}^T &= \mathbf{PWB}^T \\ \Rightarrow \mathbf{C} &= \mathbf{PWB}^T(\mathbf{BWB}^T)^{-1} \end{aligned}$$

Werden die Gewichtungsfaktoren w_i alle auf eins gesetzt, so ist die Matrix \mathbf{W} gleich der Einheitsmatrix und die Gleichung vereinfacht sich zu

$$\mathbf{C} = \mathbf{PB}^T(\mathbf{BB}^T)^{-1}$$

Die Matrix $\mathbf{B}^T(\mathbf{BB}^T)^{-1}$ nennt man *Pseudoinverse* der nicht quadratischen Matrix \mathbf{B} . Für $n = m$ wäre die Matrix \mathbf{B} quadratisch. Sie wäre weiterhin regulär, da die verwendeten Basisfunktionen (wie man dem Namen unschwer entnehmen kann) linear unabhängig sind. Damit ist sie invertierbar und die *Pseudoinverse* von \mathbf{B} ist gleich der Inversen von \mathbf{B} .

3.2 Wahl der Basisfunktionen

Bislang wurden die Basisfunktionen φ_i noch nicht näher spezifiziert. Der nun folgende Abschnitt gibt eine Übersicht über mögliche Funktionensystemen und deren Eigenschaften. Jedes Funktionensystem repräsentiert einen bestimmten Funktionstyp. Es hängt also von der Art der Bewegung des betreffenden Punktes bzw. von den zugrundeliegenden physikalischen Ursachen der Bewegung ab, ob das eine oder das andere Funktionensystem besser geeignet ist. Ist uns nichts über die Bewegung des Punktes bekannt, so ist eine angemessene Wahl nicht möglich. In diesem Fall kann die Approximation mit verschiedenen Basisfunktionen durchgeführt werden, wobei bei gleichem Approximationsgrad m dasjenige Funktionensystem die Bewegung am besten wiedergibt, bei dem der geringste Fehler aufgetreten ist.

Algebraische Polynome

Die Approximationsfunktion Φ ist mit $\varphi_k(t) = t^k$ ein algebraisches Polynom vom Höchstgrad m :

$$\Phi(t) = \sum_{k=0}^m \mathbf{c}_k t^k$$

Die Koeffizienten \mathbf{c}_k sind die Spaltenvektoren der Matrix \mathbf{C} aus Gleichung 3.2. Dieses häufig verwendete Funktionensystem hat den Nachteil, daß die Matrix \mathbf{BB}^T oft schlecht konditioniert ist,² so daß die Invertierung und damit die Berechnung der *Pseudoinversen* erschwert wird.

Beispiel: Wirkt auf den Punkt eine konstante Kraft, so läßt sich die Bewegungsbahn

² $\det(\mathbf{BB}^T) \approx 0$

durch ein algebraisches Polynom zweiten Grades beschreiben. Das Weg-Zeit-Gesetz einer gleichmäßig beschleunigten Bewegung lautet nämlich:

$$\mathbf{s} = \frac{1}{2} \mathbf{a} t^2 + \mathbf{v}_0 t + \mathbf{s}_0, \quad (3.4)$$

wobei \mathbf{a} die konstante Beschleunigung, \mathbf{v}_0 die Anfangsgeschwindigkeit zum Zeitpunkt $t = 0$, \mathbf{s}_0 der zum Zeitpunkt $t = 0$ bereits zurückgelegte Weg und t die Zeit ist. Eine Approximation mit einem algebraischen Polynom würde also eine angemessene Beschreibung für diese Punktbeziehung liefern.

Trigonometrische Polynome

Keht ein Punkt während seiner Bewegung zu seinem Ausgangspunkt zurück, so kann angenommen werden, daß er sich periodisch bewegt. In einem solchen Fall empfiehlt es sich, periodische Basisfunktionen zu wählen. Verwendet man

$$1, \sin(kx), \cos(kx) \quad \begin{array}{l} k \in [1, \dots, m] \\ x \in [0, \dots, 2\pi] \end{array}$$

als Basisfunktionen, so ist die Funktion Φ ein trigonometrisches Polynom vom Grad m :

$$\Phi(t) = \frac{1}{2} \mathbf{a}_0 + \sum_{k=1}^m (\mathbf{a}_k \cos(2\pi kt) + \mathbf{b}_k \sin(2\pi kt)) \quad t \in [0, \dots, 1]$$

Für die gegebenen Punkte erhalten wir das folgende Gleichungssystem:

$$\begin{pmatrix} x(t_0) & x(t_1) & \cdots & x(t_n) \\ y(t_0) & y(t_1) & \cdots & y(t_n) \\ z(t_0) & z(t_1) & \cdots & z(t_n) \end{pmatrix} = \begin{pmatrix} a_{0x} & a_{1x} & \cdots & a_{mx} & b_{1x} & \cdots & b_{mx} \\ a_{0y} & a_{1y} & \cdots & a_{my} & b_{1y} & \cdots & b_{my} \\ a_{0z} & a_{1z} & \cdots & a_{mz} & b_{1z} & \cdots & b_{mz} \end{pmatrix} \\ \times \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \cdots & \frac{1}{2} \\ \cos(2\pi t_0) & \cos(2\pi t_1) & \cdots & \cos(2\pi t_n) \\ \vdots & \vdots & & \vdots \\ \cos(2\pi m t_0) & \cos(2\pi m t_1) & \cdots & \cos(2\pi m t_n) \\ \sin(2\pi t_0) & \sin(2\pi t_1) & \cdots & \sin(2\pi t_n) \\ \vdots & \vdots & & \vdots \\ \sin(2\pi m t_0) & \sin(2\pi m t_1) & \cdots & \sin(2\pi m t_n) \end{pmatrix} \\ \Rightarrow \mathbf{P} = \mathbf{C}\mathbf{B}$$

Auch hier könnten wir aus der Pseudoinversen der Matrix \mathbf{B} die Koeffizienten der Basisfunktionen berechnen. Da die obigen Basisfunktionen jedoch ein Orthonormalsystem bilden, können wir die Koeffizienten direkt angeben. Betrachtet man nämlich den Funktionenraum, der von den orthonormalen Basisfunktionen aufgespannt wird, und möchte man nun eine gegebene Funktion f durch eine Linearkombination dieser Basisfunktionen

darstellen, so erhält man die zu den Basisfunktionen zugehörigen Koeffizienten, indem man das Skalarprodukt der Funktion f mit der jeweiligen Basisfunktion bildet. Damit erhalten wir:

$$\begin{aligned} \mathbf{a}_l &= \frac{2}{n+1} \sum_{k=0}^n \mathbf{p}(t_k) \cos(2\pi l t_k) \\ \mathbf{b}_l &= \frac{2}{n+1} \sum_{k=0}^n \mathbf{p}(t_k) \sin(2\pi l t_k) \end{aligned}$$

Dabei sind die t_k so zu skalieren, daß t_0, \dots, t_n den Bereich $[0, \dots, 1]$ aufspannen. Es muß sichergestellt sein, daß der Anfangs- und der Endpunkt zusammenfallen, denn erst in diesem Fall kann eine periodische Bewegung vorliegen, die durch ein trigonometrisches Polynom beschrieben werden kann.

Beispiel: Unterliegt ein Punkt einer rücktreibenden Kraft, die proportional zu seiner Auslenkung ist, so führt der Punkt eine harmonische Schwingung aus. Das Weg-Zeit-Gesetz einer harmonischen Schwingung lautet

$$\mathbf{s} = \mathbf{a} \sin(\omega t) + \mathbf{s}_0 \quad (3.5)$$

Wobei \mathbf{a} die Amplitude der Schwingung, $\omega = 2\pi/T$ die Winkelgeschwindigkeit³ und \mathbf{s}_0 der zum Zeitpunkt $t = 0$ bereits zurückgelegte Weg ist. Die Trajektorie des Punktes läßt sich folglich durch ein trigonometrisches Polynom ersten Grades beschreiben.

3.3 Geschwindigkeit und Beschleunigung von Punkten

Die Bewegung läßt sich gegebenenfalls einfacher beschreiben, wenn man statt des zurückgelegten Weges die Geschwindigkeit oder aber die Beschleunigung des gegebenen Punktes betrachtet. Die Geschwindigkeit \mathbf{v} eines Punktes \mathbf{p} ist definiert als Ableitung des zurückgelegten Weges \mathbf{s} nach der Zeit t :

$$\mathbf{v} = \frac{d\mathbf{s}}{dt}$$

Ferner ergibt sich die Beschleunigung \mathbf{a} aus der Ableitung der Geschwindigkeit nach der Zeit:

$$\mathbf{a} = \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{s}}{(dt)^2}$$

Eine Schätzung für die Geschwindigkeit $\mathbf{v}(t_k)$ des Punktes \mathbf{p} zum Zeitpunkt t_k erhält man aus der Differenz der Punktkoordinaten $\mathbf{p}(t_{k-1})$ und $\mathbf{p}(t_k)$:

$$\mathbf{v}(t_k) = \frac{\mathbf{p}(t_k) - \mathbf{p}(t_{k-1})}{t_k - t_{k-1}} \quad k \in [1, \dots, n]$$

³ T ist die Schwingungsdauer oder auch Periode

Entsprechend erhält man eine Schätzung für die Beschleunigung, indem man die Differenz zeitlich aufeinanderfolgender Geschwindigkeiten bildet:

$$\begin{aligned} \mathbf{a}(t_k) &= \frac{\mathbf{v}(t_k) - \mathbf{v}(t_{k-1})}{t_k - t_{k-1}} \quad k \in [2, \dots, n] \\ &= \frac{\frac{\mathbf{p}(t_k) - \mathbf{p}(t_{k-1})}{t_k - t_{k-1}} - \frac{\mathbf{p}(t_{k-1}) - \mathbf{p}(t_{k-2})}{t_{k-1} - t_{k-2}}}{t_k - t_{k-1}} \end{aligned}$$

Sind die Zeitintervalle alle gleich lang, so erhält man:

$$\mathbf{a}(t_k) = \frac{\mathbf{p}(t_k) - 2\mathbf{p}(t_{k-1}) + \mathbf{p}(t_{k-2}))}{(t_k - t_{k-1})^2}$$

Mit den erhaltenen Geschwindigkeiten $\mathbf{v}(t_1), \dots, \mathbf{v}(t_n)$ bzw. mit den Beschleunigungen $\mathbf{a}(t_2), \dots, \mathbf{a}(t_n)$ läßt sich nun das obige Approximationsverfahren durchführen.

Beispiel: Im folgenden wird eine Bewegung beschrieben, bei der es sinnvoll ist, die Geschwindigkeit zu approximieren.

Gegeben sei ein mit konstanter Winkelgeschwindigkeit ω rotierender starrer Körper, dessen Rotationszentrum sich mit konstanter Geschwindigkeit \mathbf{v}_R senkrecht zur Rotationsachse fortbewegt (siehe Abbildung 3.1). Die Translationsbewegung des Körpers sei auf die XY-Ebene beschränkt, so daß die Rotationsachse senkrecht zu dieser Ebene orientiert ist. Die Bewegung läßt sich damit in der XY-Ebene beschreiben. Das Weg-Zeit-Gesetz für einen beliebigen Punkt dieses starren Körpers kann nun folgendermaßen geschrieben werden:

$$\mathbf{s}(t) = r \begin{pmatrix} \sin(\omega t) \\ \cos(\omega t) - 1 \end{pmatrix} + \mathbf{v}_R t + \mathbf{s}_0 \quad (3.6)$$

Dabei sei r der radiale Abstand des beobachteten Punktes zum Rotationszentrum und \mathbf{s}_0 der zum Zeitpunkt $t = 0$ zurückgelegte Weg (Koordinaten des Punktes \mathbf{p} zum Zeitpunkt $t = 0$). Für diese einfache Bewegung erhalten wir also eine Bewegungsgleichung, die sich als Summe eines trigonometrischen und eines algebraischen Polynoms darstellen läßt. Damit wären beide oben vorgestellten Basissysteme für sich allein genommen zur Approximation ungeeignet. Statt dessen kann man eine neue Basis aus einer Kombination von trigonometrischen und algebraischen Basisfunktionen bilden. Diese neue Funktionenbasis trägt nunmehr verschiedenen physikalischen Bewegungserscheinungen gleichzeitig Rechnung. Mit ihr kann wiederum das vorgestellte Pseudoinversenverfahren durchgeführt werden.⁴ Die ermittelte Approximationsfunktion bringt aber nur dann die Gesetzmäßigkeit in Gleichung 3.6 zum Ausdruck, wenn der Beobachtungszeitraum ein ganzzahliges Vielfaches der Rotationsperiode ist. Denn nur dann kann der Rotationsanteil in dem beobachteten Zeitraum eindeutig als Summe von trigonometrischen Funktionen dargestellt werden. Dieses ist jedoch schwierig sicherzustellen, da der betrachtete Punkt aufgrund

⁴Die direkte Angabe der Funktionskoeffizienten, wie sie bei den trigonometrischen Polynomen möglich ist, kann hier nicht erfolgen, da das neue Basissystem nicht mehr orthogonal ist.

des Translationsanteils in der Bewegung nicht wieder zu seinem Ausgangspunkt zurückkehrt. Man hat damit keinen direkten Hinweis mehr, wo eine Rotationsperiode beendet sein könnte.

Anstatt durch Wahl eines neuen Basissystems kann man sich mit der Ableitung der Gleichung 3.6 weiterhelfen. Man erhält dann:

$$\mathbf{v}(t) = \frac{d\mathbf{s}(t)}{dt} = r\omega \begin{pmatrix} \cos(\omega t) \\ -\sin(\omega t) \end{pmatrix} + \mathbf{v}_R$$

Diese Gleichung stellt das Geschwindigkeit-Zeit-Gesetz der obigen Bewegung dar und läßt sich durch ein reines trigonometrisches Polynom beschreiben. Damit kann die Approximation der Punktgeschwindigkeiten mit der trigonometrischen Funktionenbasis erfolgen.

Da wir unsere Approximationsfunktion zur Vorhersage der Punktlokalisierung nutzen wollen, müssen wir die erhaltene Geschwindigkeitsfunktion hinterher jedoch noch nach der Zeit integrieren. Wir erhalten:

$$\begin{aligned} \mathbf{s}'(t) &= \int_0^t \mathbf{v}(t) dt \\ &= r\omega \int_0^t \begin{pmatrix} \cos(\omega t) \\ -\sin(\omega t) \end{pmatrix} dt + \mathbf{v}_R \int_0^t dt \\ &= r \begin{pmatrix} \sin(\omega t) \\ \cos(\omega t) - 1 \end{pmatrix} + \mathbf{v}_R t \end{aligned}$$

Dies entspricht bis auf den fehlenden konstanten Summanden \mathbf{s}_0 der Gleichung 3.6. Die Koordinaten des Punktes \mathbf{p} zum Zeitpunkt $t = 0$ müssen noch zu $\mathbf{s}'(t)$ addiert werden, um die gewünschte Gleichung zu erhalten.

Bei der Approximation der Punktgeschwindigkeiten bleibt zu bedenken, daß die Geschwindigkeiten durch Differenzbildung entstanden sind. Sie geben damit die mittlere Geschwindigkeit für das vorliegende Zeitintervall an. Inwieweit diese Geschwindigkeiten eine gute Näherung für die Geschwindigkeiten zu den gewünschten Zeitpunkten darstellen, hängt zum einen von der Größe der Zeitintervalle und zum anderen von den in der Szene auftretenden Beschleunigungen ab. Je größer die Beschleunigungen sind, desto kleiner müssen die Zeitintervalle gewählt werden.

Die Betrachtungen lassen sich in derselben Form auch für die Punktbeschleunigungen anstellen. Hier kann die zweite Ableitung des Weg-Zeit-Gesetzes zu einer weiteren Vereinfachung der zu approximierenden Funktion führen. Dafür wirken sich jedoch hier die Probleme der Differenzbildung gleich doppelt aus.

3.4 Probleme bei der Bewegungsbeschreibung einzelner Punkte

Die vorgestellte Approximationsmethode verlangt die Wahl von geeigneten Basisfunktionen. Diese Wahl kann nur dann vernünftig erfolgen, wenn eine Modellvorstellung der

anzunähernden Kurve existiert. Dies setzt jedoch voraus, daß wir die physikalischen Ursachen der Bewegung kennen. So kann eine Approximation mittels eines trigonometrischen Polynoms nur dann erfolgreiche Vorhersagen machen, wenn eine Periodizität in der Punktbeziehung auch wirklich vorliegt. Dieses müßte also geprüft werden, bevor die Approximation erfolgt. Nun läßt sich zwar eine Trajektorie innerhalb des gegebenen Zeitintervalls, unabhängig vom gewählten Basissystem, beliebig genau approximieren, indem man den Grad m der Approximationsfunktion erhöht; zu Vorhersagen eignet sich die erhaltene Funktion jedoch nur dann, wenn die gewählten Basisfunktionen die physikalische Bewegungsursache wiedergeben können. Ein algebraisches Polynom eignet sich beispielsweise nicht zur Beschreibung von periodischen Bewegungen, wie sie sich z. B. bei Schwingungsphänomenen ergeben. Möchte man mit einer derartigen Approximationsfunktion die weitere Bewegung des Punktes vorhersagen, so wird man kläglich scheitern.

Ein weiteres Problem ist, daß einzelne Punkte häufig keinen einfachen physikalischen Verhältnissen unterliegen. Der Grund hierfür liegt darin, daß ein Punkt nur einen sehr kleinen Teil eines makroskopischen Körpers repräsentiert.⁵ Dieser Körper unterliegt in seiner Gesamtheit den herrschenden physikalischen Gesetzen. Da alle Punkte des Körpers miteinander in starrer oder nichtstarrer Verbindung stehen, läßt sich die Bewegung eines Punktes nicht losgelöst von den anderen Punkten betrachten. Nehmen wir z. B. an, daß ein starrer Körper einer konstanten Kraft ausgesetzt ist. In einem solchen Fall läßt sich die Gleichung 3.4 lediglich auf den Schwerpunkt dieses Körpers anwenden. Die Bewegung eines beliebigen Punktes kann selbst für diese einfache physikalische Situation eine kompliziertere Bewegung vollführen, wie die Abbildung 3.1 verdeutlicht.

Im Fall von gekoppelter Bewegung kann die Bewegungsbahn eines Punktes noch komplizierter werden, nämlich dann, wenn der Punkt eine Rotationsbewegung relativ zu einem anderen rotierenden Objektteil ausführt. Die Bildfolge in Abbildung 3.2 zeigt ein derartiges Objekt, welches aus drei verbundenen starren Einzelteilen besteht. Diese sind über Achsen verbunden und führen jeweils eine eigene Rotation mit konstanter Winkelgeschwindigkeit aus. Der in den Bildern durch \square markierte Punkt führt eine periodische Bewegung mit einer Schwingungsdauer von $T = 360$ Zeiteinheiten aus. Für diesen Punkt wurde in Abbildung 3.3 der zurückgelegte Weg, die Geschwindigkeit und die Beschleunigung in x - und y -Richtung gegen die Zeit aufgetragen.⁶ Berücksichtigt wurden nur die Punktkoordinaten zu den 10 gegebenen Bildaufnahmezeitpunkten $t = 0, 10, 20, \dots, 90$. In Abbildung 3.4 und Abbildung 3.5 wurden die erzielten Approximationsergebnisse dem zurückgelegten Weg in x - und y -Richtung überlagert und die aufgetretenen Fehler gegen die Zeit aufgetragen. Dabei wurde im ersten Fall ein algebraisches Polynom vierten Grades und im zweiten Fall ein trigonometrisches Polynom zweiten Grades verwendet. Inwieweit die berechneten Approximationsfunktionen zu Vorhersagen geeignet sind, wird durch die Abbildungen 3.6 und 3.7 veranschaulicht. Hier wurde der zurückgelegte Weg

⁵Mathematisch hat ein Punkt keine Ausdehnung, so daß diese Feststellung eigentlich trivial ist. Die in unserem Fall rekonstruierten Punkte jedoch repräsentieren räumlich ausgedehnte Flächen, die genau in einen Bildpunkt abgebildet werden.

⁶Aus Skalierungsgründen wurde in den Abbildungen 3.3, 3.4 und 3.5 auf die Darstellung der z -Komponente des Weges, der Geschwindigkeit und der Beschleunigung verzichtet.

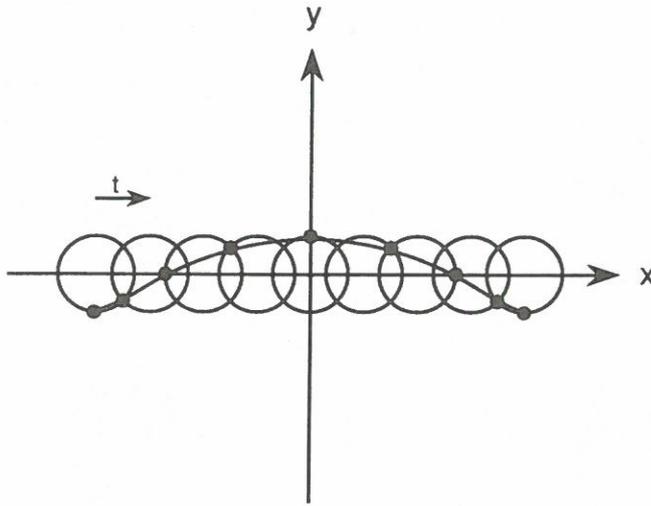


Abbildung 3.1: Bewegung eines Punktes auf einem sich kräftefrei bewegenden, rotierenden starren Körper

des gewählten Punktes über den betrachteten Zeitraum hinaus aufgetragen, unter der Annahme, daß sich das Objekt unverändert weiterbewegt. Überlagert wurden die Approximationsfunktionen. Man sieht im Falle von algebraischen Polynomen (Abbildung 3.6), daß die Trajektorie des Punktes für den Aufnahmezeitraum zwar sehr genau wiedergegeben wird, die Funktion zu Vorhersagen aber kaum geeignet ist. Ein algebraisches Polynom besitzt nur eine endliche Zahl von lokalen Extremstellen und eignet sich damit nicht zur Darstellung von periodischen Funktionen. Allerdings hilft uns auch kein trigonometrisches Polynom weiter, da der Beobachtungszeitraum keine vollständige Periode umfaßt. Diese Annahme liegt der Approximation mit trigonometrischen Polynomen zugrunde (Abbildung 3.7).

Betrachten wir die Geschwindigkeit und die Beschleunigung des Punktes (siehe Abbildung 3.3), so erhalten wir wieder periodische Funktionen. Eine Approximation der Geschwindigkeit bzw. der Beschleunigung des Punktes, wie wir sie in Abschnitt 3.3 kennengelernt haben, befreit uns damit nicht von den genannten Problemen.

Als Fazit läßt sich sagen, daß die Approximation von Punkttrajektorien kein geeignetes Mittel darstellt, um Vorhersagen über die zukünftige Bewegung eines Punktes abzuleiten. Erst wenn wir die zugrundeliegenden physikalischen Ursachen der Bewegung (Bewegungsgesetz) kennen, können geeignete Basisfunktionen gewählt werden, die in der Lage sind, die Bewegung wiederzugeben. Dieses Wissen würde uns aber bereits in die Lage versetzen, die Bewegung eines Punktes vorherzusagen und damit eine Approximation überflüssig machen. Eine Approximation ist damit lediglich geeignet eine kontinuierliche Verlaufsbeschreibung eines Punktes innerhalb des betrachteten Zeitintervalls zu liefern. Diese Information könnte für verschiedene Situationen genutzt werden. Für den Fall, daß der Punkt für kurze Zeit nicht in den Bildern aufgefunden werden konnte, sei es

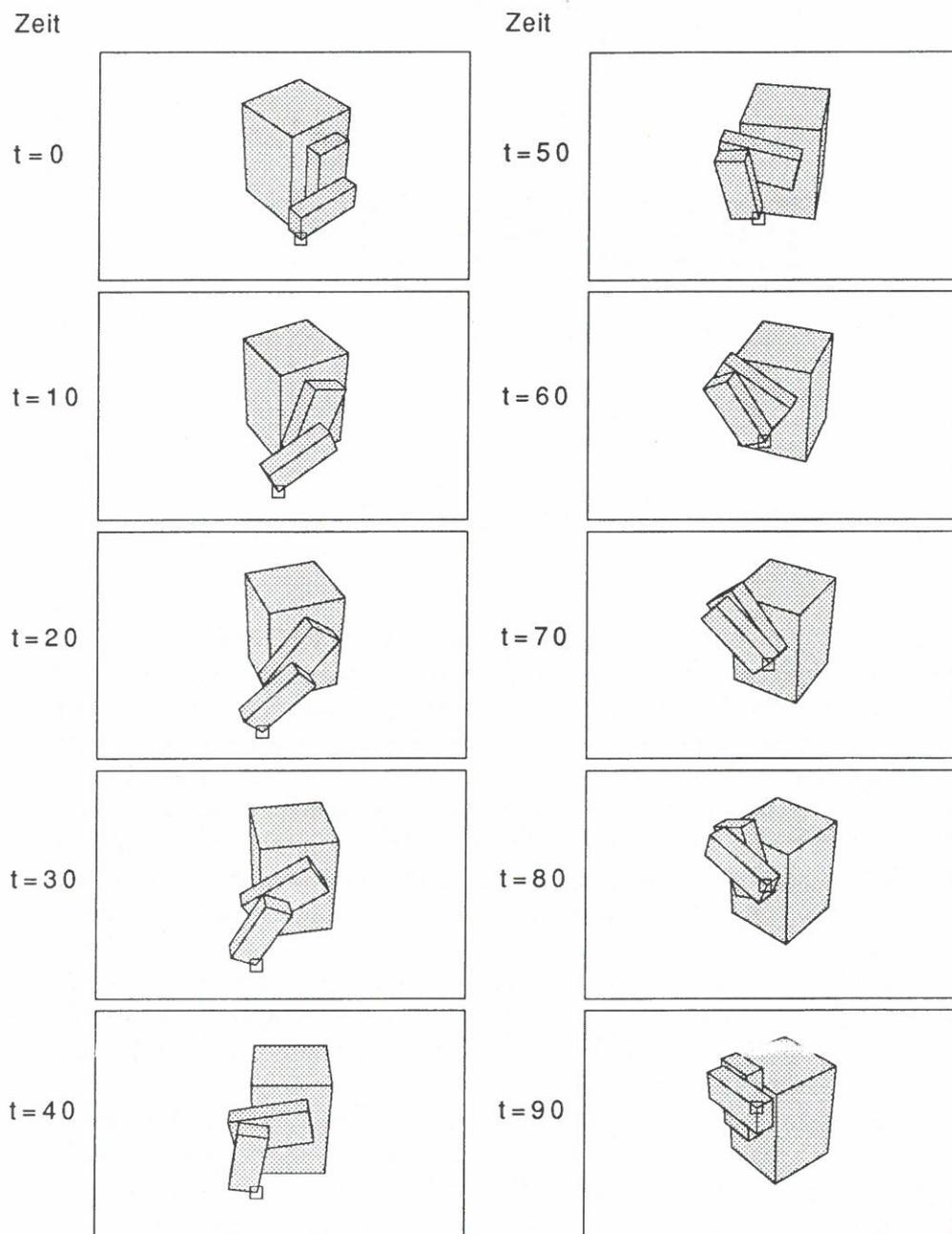


Abbildung 3.2: Gekoppelte Bewegung

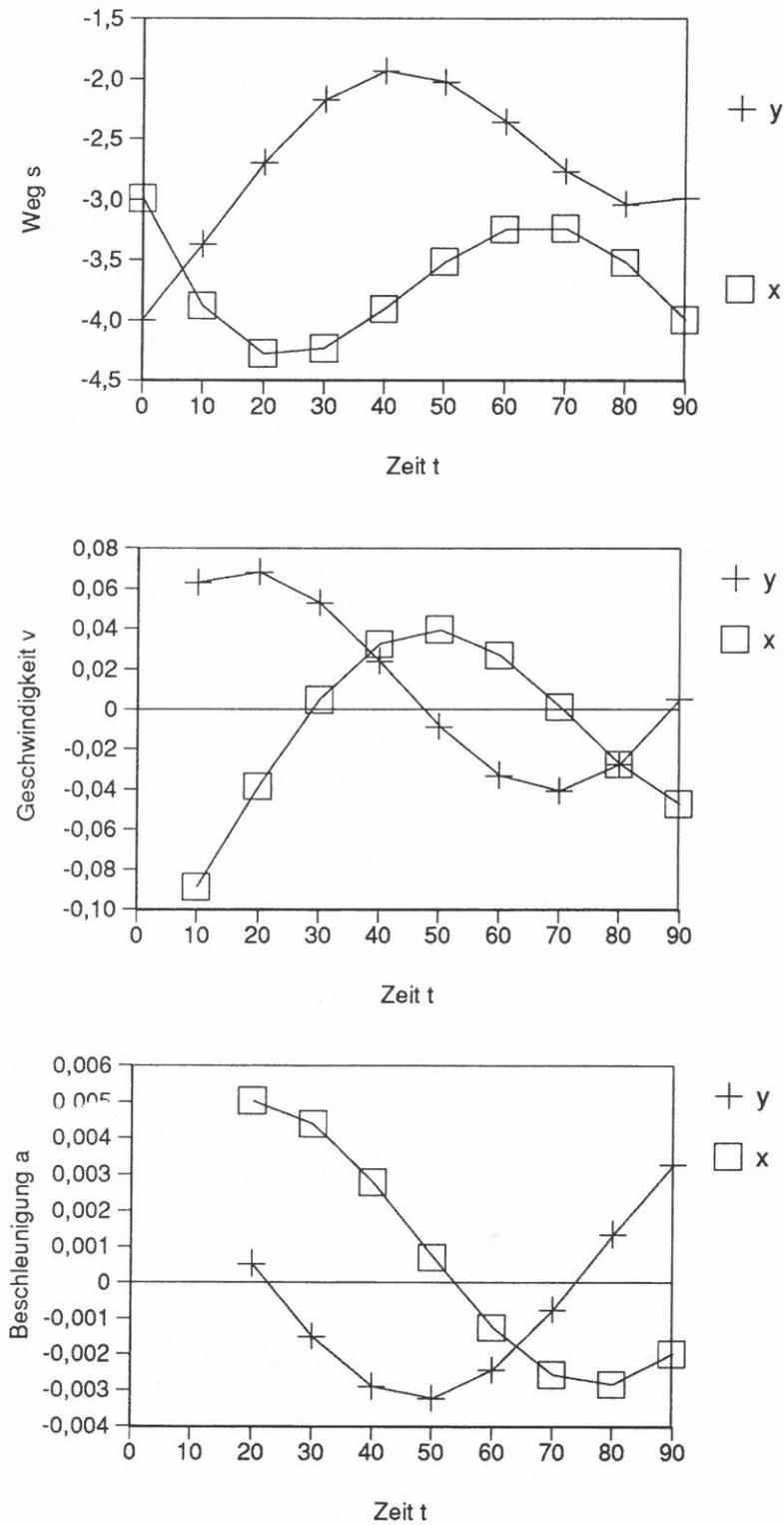


Abbildung 3.3: Weg, Geschwindigkeit und Beschleunigung des markierten Punktes aus Abbildung 3.2

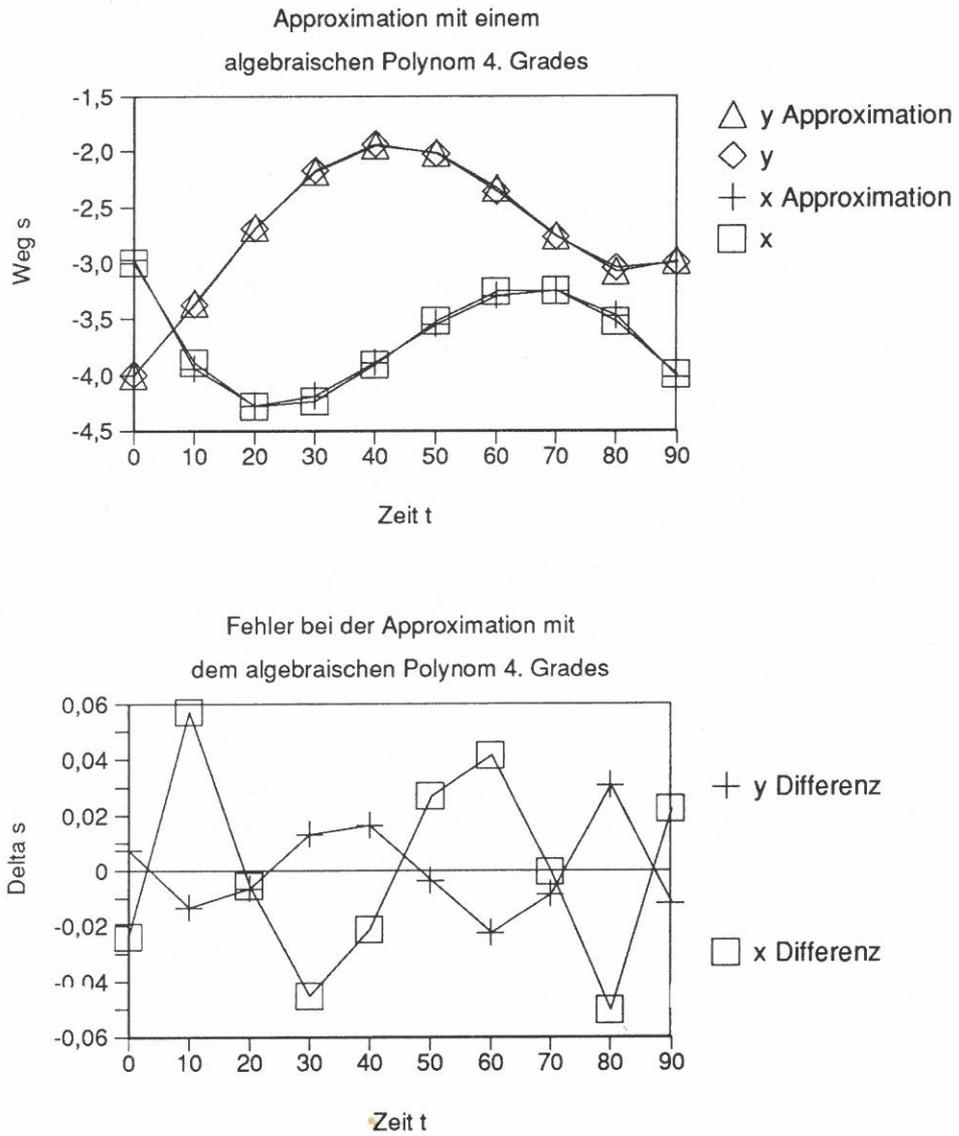


Abbildung 3.4: Approximation der Punkttrajektorie mit einem algebraischen Polynom und die erhaltenen Fehler

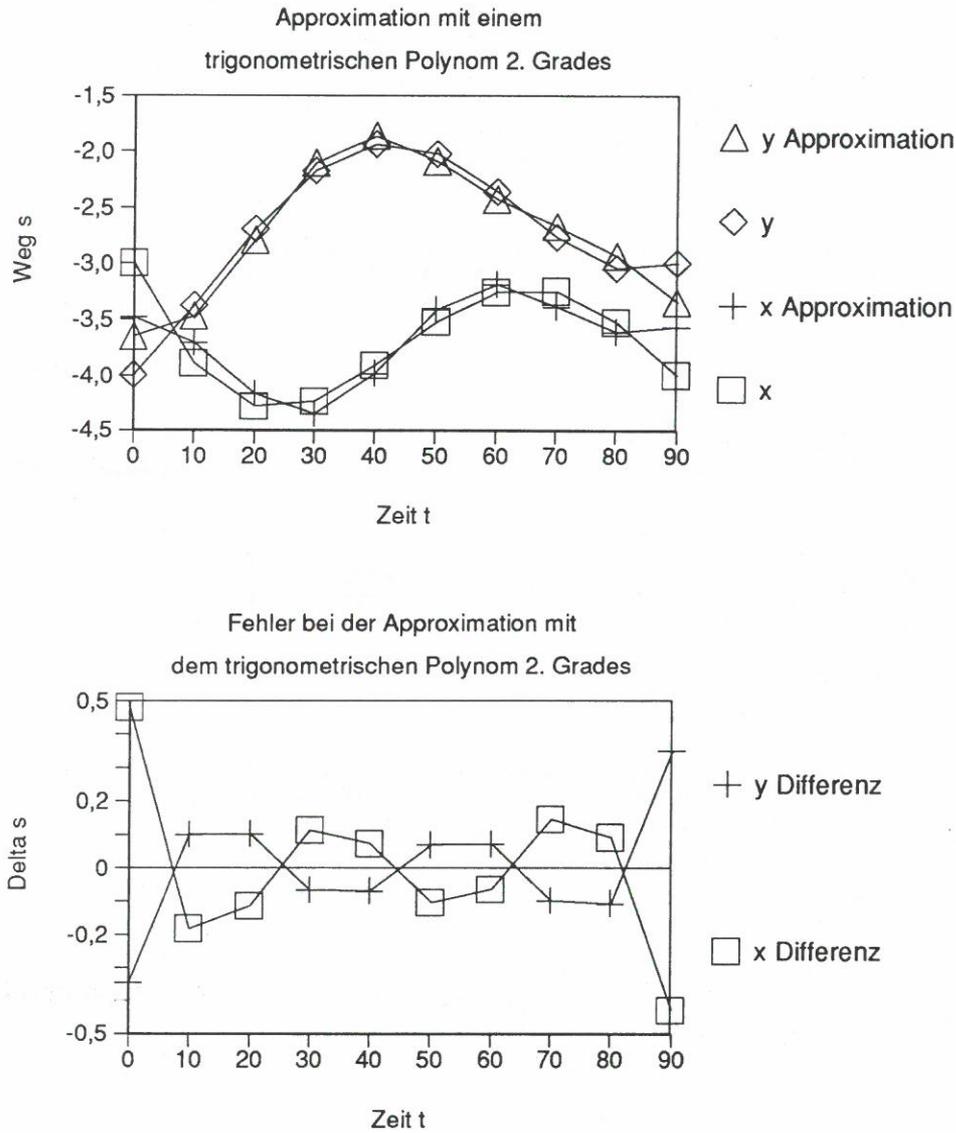


Abbildung 3.5: Approximation der Punkttrajektorie mit einem trigonometrischen Polynom und die erhaltenen Fehler

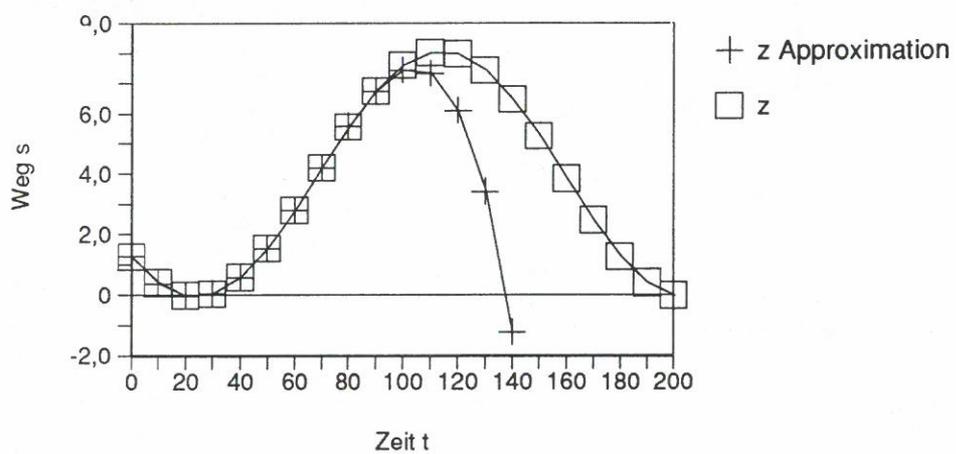
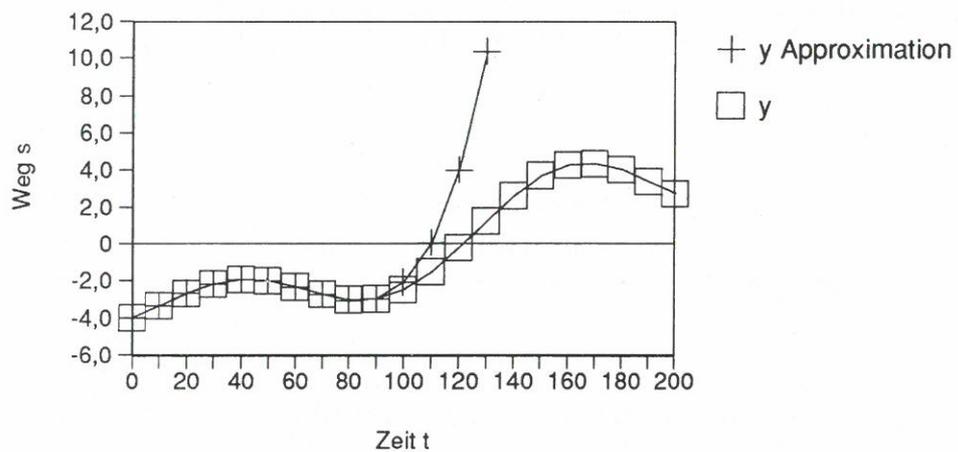
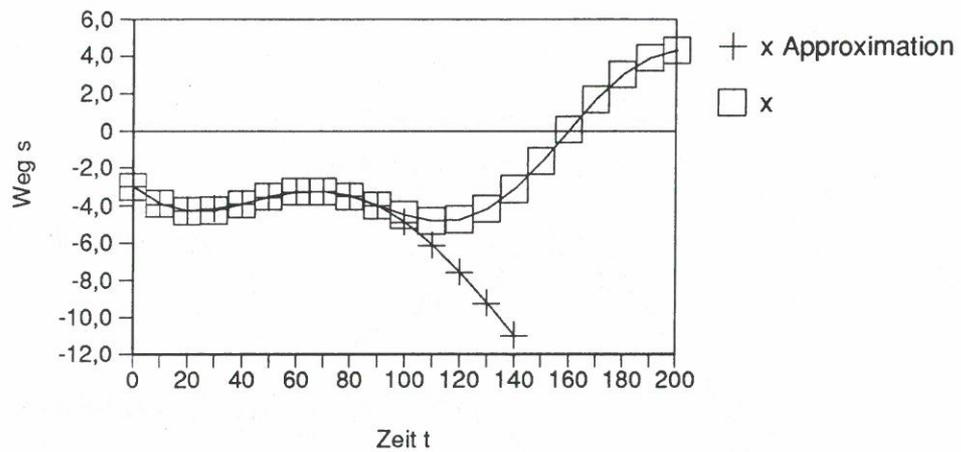


Abbildung 3.6: Vorhersage der Punkttrajektorie mit einem algebraischen Polynom 4. Grades, Beobachtungsintervall: $t = 0, \dots, 90$

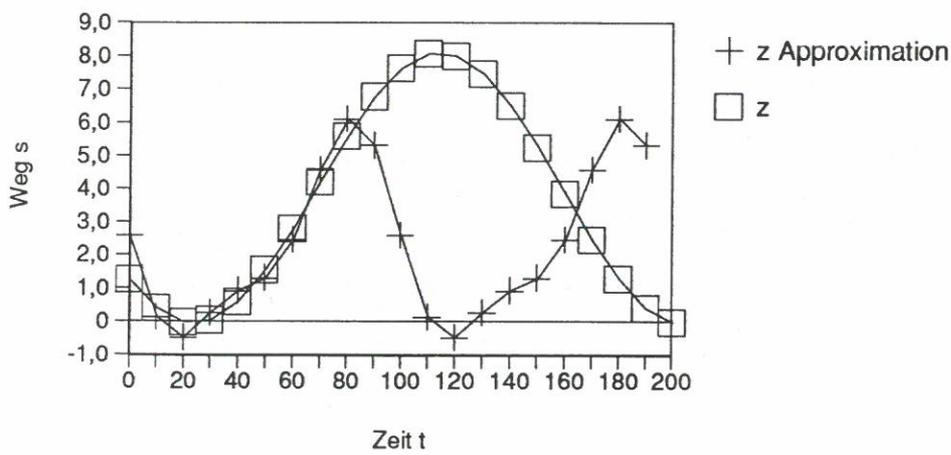
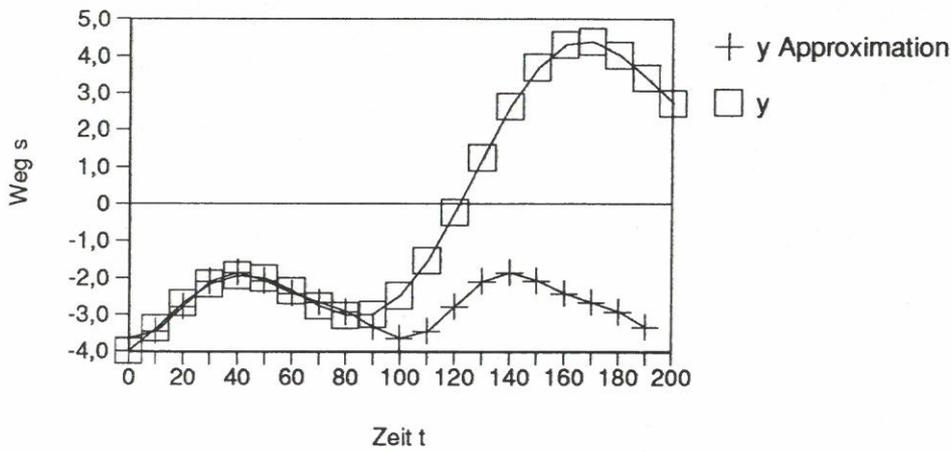
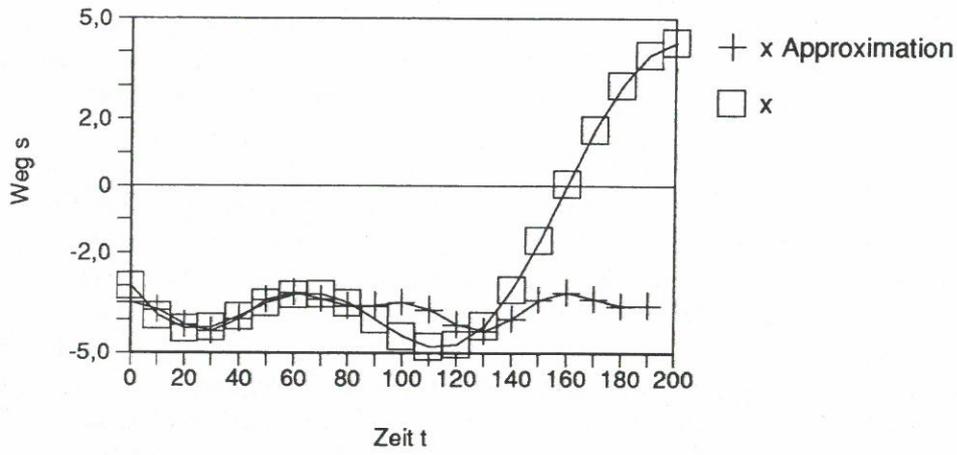


Abbildung 3.7: Vorhersage der Punkttrajektorie mit einem trigonometrischen Polynom 2. Grades, Beobachtungsintervall: $t = 0, \dots, 90$

durch räumliche Verdeckung oder aber durch unvollständige Segmentierung hervorgerufen, läßt sich mit Hilfe der Approximationsfunktion die Trajektorie des Punktes ergänzen und damit seine Lage in den betroffenen Bildern bestimmen. Weiterhin ließen sich bei vereinzelt starken Abweichungen von der Approximationsfunktion die Korrespondenzen korrigieren.

Unabhängig von den Schwierigkeiten bei der Approximation wird man es im allgemeinen immer schwer haben, die Bewegung von Punkten vorherzusagen. Der Grund dafür liegt darin, daß die Bewegung eines Punktes, wenn man sie isoliert betrachtet, nicht erkennen läßt, auf welche Art und Weise sie zustande gekommen ist. Wir werden daher im folgenden versuchen, die Bewegung von makroskopischen Körpern zu beschreiben, indem wir alle ermittelten Punkte des Körpers berücksichtigen. Die Bewegung läßt sich dann durch Rotation und Translation erklären. Diese beiden Bewegungsformen genügen, um die Bewegung von starren Körpern mit physikalischen Mitteln zu beschreiben. Gelingt es darüberhinaus, die Kopplungen der beteiligten starren Körpern zu ermitteln, so gewinnt man ein größeres Verständnis für die Bewegung. Dieses sollte eine leichtere Vorhersage der Bewegung ermöglichen.

Kapitel 4

3D-Bewegung starrer Körper

In dem nun folgenden Kapitel werden wir uns mit der Bewegungsbeschreibung von starren Körpern beschäftigen. Kennen wir die räumliche Position eines starren Körpers zu zwei aufeinanderfolgenden Zeitpunkten, so läßt sich die Bewegung in dem Zeitintervall durch Rotation und Translation beschreiben. Diese unterschiedlichen Bewegungsphänomene werden auch in der Physik verwendet, um die Bewegung von starren Körpern darzustellen.

Da uns nur einzelne Raumpunkte und deren räumliche Bewegungen vorliegen, müssen wir zunächst feststellen, welche Punkte einen starren Körper bilden. Der folgende Abschnitt 4.1 beschreibt eine Möglichkeit, um die vorliegende Punktmenge zu starren Körpern zu gruppieren. Für jeden dieser starren Körper lassen sich nun für jedes Zeitintervall Rotations- und Translationsparameter ermitteln. Der Abschnitt 4.2 beschreibt zunächst die mathematische Darstellung und die Eigenschaften der Bewegungsparameter von starren Körpern. Im letzten Abschnitt 4.3 werden wir dann ein Verfahren angeben, welches aus den räumlichen Punktkoordinaten eines starren Körpers die Bewegungsparameter bestimmt. Es wird ein “least-square-fit”-Verfahren zur Anwendung kommen.

4.1 Isolieren von starren Körpern

Ein starrer Körper zeichnet sich dadurch aus, daß der Abstand zwischen beliebigen Punkten des Körpers konstant bleibt. Diese Eigenschaft werden wir durch den Aufbau eines Starrheitsgraphen für alle Punktpaare überprüfen. Der Starrheitsgraph enthält für jeden Objektpunkt \mathbf{p} einen Knoten v . Dabei existiert zwischen zwei Knoten v_i und v_j eine ungerichtete Kante $e(v_i, v_j)$, falls die zugehörigen Punkte \mathbf{p}_i und \mathbf{p}_j ihren räumlichen Abstand zu den gemessenen Zeitpunkten beibehalten haben. Da die Berechnung der 3D-Punktkoordinaten einer Vielzahl von Fehlerquellen ausgesetzt ist, werden wir geringe Änderungen im Punktabstand zulassen. Wir nehmen an, daß diese Fehler normalverteilt sind und ermitteln daher für einen gegebenen Zeitraum t_0, \dots, t_{n-1} für jedes Punktpaar $(\mathbf{p}_i, \mathbf{p}_j)$ den Mittelwert $\overline{d_{ij}}$ des Punktabstandes und die Standardabweichung σ_{ij} von

diesem Mittelwert:

$$\overline{d_{ij}} = \frac{1}{n} \sum_{k=0}^{n-1} |\mathbf{p}_i(t_k) - \mathbf{p}_j(t_k)|$$

$$\sigma_{ij} = \sqrt{\frac{1}{n} \sum_{k=0}^{n-1} (|\mathbf{p}_i(t_k) - \mathbf{p}_j(t_k)| - \overline{d_{ij}})^2}$$

Ist die Standardabweichung σ_{ij} klein gegenüber dem mittleren Punktabstand $\overline{d_{ij}}$ der Punkte \mathbf{p}_i und \mathbf{p}_j , so rechnen wir die beiden Punkte dem gleichen starren Körper zu und verbinden somit die Knoten v_i und v_j im Starrheitsgraphen durch eine Kante. An das Verhältnis zwischen Standardabweichung und Abstandsmittelwert wird also ein Schwellwert s angelegt. Liegt das Verhältnis unter der gewählten Schwelle s , so betrachten wir die Punkte als starr verbunden, andernfalls betrachten wir sie als nicht zu einem gemeinsamen starren Körper gehörend.

Für eine Punktmenge, die sich starr verhält, müssen alle Knotenpaare der korrespondierenden Knotenmenge im Starrheitsgraphen durch eine Kante verbunden sein. Diese Knoten bilden also im Starrheitsgraphen einen vollständig verbundenen Teilgraphen, den man auch als *Clique* bezeichnet.

Ein Algorithmus zum Auffinden von maximalen *Cliquen* im Starrheitsgraphen kann somit dazu dienen, die starren Körper zu ermitteln. Es ist allerdings zu berücksichtigen, daß eine maximale *Clique* im Starrheitsgraphen nur eine notwendige Bedingung für einen starren Körper darstellt. Wird ein Objekt in sein Spiegelbild überführt, so bleiben die paarweisen Punktabstände zwar erhalten, die Bewegung wäre aber im physikalischen Sinne nicht starr. Wir werden in Abschnitt 4.2 dieses Phänomen aufgreifen und eine hinreichende Bedingung für eine starre Bewegung formulieren. Zum Auffinden von starren Körpern beschränken wir uns hier auf die Cliquensuche im Starrheitsgraphen, da der Spezialfall der Spiegelung ein kaum anzutreffendes Bewegungsphänomen darstellt. Außerdem würden die in Abschnitt 4.2 beschriebenen Bewegungsparameter die Spiegelung erkennen lassen, so daß wir uns um dieses Problem hier noch nicht zu kümmern brauchen.

Im weiteren werden wir uns damit beschäftigen, wieviele Zeitpunkte betrachtet werden sollen, um die starren Punktconstellationen zu ermitteln. Berücksichtigen wir nur zwei Zeitpunkte, so kann der Abstand zweier Punkte konstant bleiben, obwohl die Punkte nicht zum selben starren Körper gehören. Dadurch würden wir zusätzliche *Cliquen* erhalten, die sich durch folgende Eigenschaften auszeichnen:

- Diese *Cliquen* beinhalten Knoten, die auch in anderen *Cliquen* enthalten sind. Es werden folglich bestimmte Punkte mehreren starren Körpern zugeordnet.
- Diese *Cliquen* beinhalten nur wenige Punkte, da die Punkte nur aufgrund einer speziellen Bewegung ihren Abstand beibehalten.

Damit wir derartige *Cliquen* entfernen können, fordern wir die Disjunktheit aller *Cliquen*.¹ Dieses läßt sich am einfachsten durch Entfernen von *Cliquen* erreichen. Sind zwei *Cliquen*

¹Es sei angemerkt, daß ein Punkt sehr wohl mehreren starren Objekten zugesprochen werden kann.

nicht disjunkt, so könnten wir z. B. die kleinere Clique entfernen. Das Problem von zufällig sich starr verhaltenden Punkten läßt sich schnell dadurch umgehen, indem mehr als zwei Zeitpunkte berücksichtigt werden. Dies setzt allerdings voraus, daß uns die 3D-Punktkoordinaten zu genügend vielen Zeitpunkten zur Verfügung stehen. Dies stellt hohe Anforderungen an die Ermittlung der räumlichen und zeitlichen Korrespondenzen. Aus diesem Grund kann es sinnvoll sein, sich zunächst auf eine geringe Anzahl von Zeitpunkten zu beschränken.

Folgender Algorithmus zum Auffinden von starren Körpern aus den gegebenen 3D-Punkten wurde verwendet. Die betreffenden Raumpunkte müssen zu mindestens zwei Zeitpunkten zur Verfügung stehen.

1. **Aufbau des Starrheitsgraphen:** Für alle Punktpaare p_i und p_j berechne den Mittelwert \bar{d}_{ij} ihres Abstandes und die Standardabweichung σ_{ij} . Liegt das Verhältnis zwischen Standardabweichung und Mittelwert unter einem gewählten Schwellwert s , so betrachten wir die Punkte p_i und p_j als starr und verbinden die zugehörigen Knoten im Starrheitsgraphen durch eine Kante.

2. **Bestimmung aller Cliques:** (siehe [Ballard + Brown 82, S. 376])

Die Funktion $\text{Cliques}(X, Y)$ erhält als Argumente eine Knotenmenge X , die im Graph eine Clique darstellt und eine Knotenmenge Y , mit $X \subseteq Y$. Sie liefert alle Cliques, die X enthalten und die in Y enthalten sind. Damit liefert der Ausdruck $\text{Cliques}(\emptyset, N)$ alle Cliques des gegebenen Graphen mit den Knoten der Menge N .

$\text{Cliques}(X, Y) :=$

```

IF    kein Knoten der Menge  $Y - X$  mit allen Elementen
      der Knotenmenge  $X$  verbunden ist
THEN  $X$ 
ELSE  $\text{Cliques}(X \cup \{y\}, Y) \cup \text{Cliques}(X, Y - \{y\})$ 

```

wobei y ein Knoten aus der Menge Y ist, der mit allen Knoten der Menge X verbunden ist.

3. **Finden von maximalen Cliques:** Dazu entfernt man alle Cliques, die in einer anderen Clique enthalten sind.

4. Für alle gefundenen Cliques C_i und C_j , für deren Knotenmengen N_i und N_j gilt:

$$N_i \cap N_j \neq \emptyset \text{ für } i \neq j$$

entferne C_i , falls $\text{size}(N_i) < s' \cdot \text{size}(N_j)$. Dabei ist s' ein festzulegender Schwellwert mit $s' \ll 1$.

Besitzen zwei starre Körper eine gemeinsame Drehachse (Verbindung), so verhalten sich alle Punkte auf der Achse zu beiden Körpern starr. Obwohl wir in Kapitel 5 derartige gekoppelte Bewegungen noch genauer untersuchen werden, wollen wir hier diesen Fall zunächst ausschließen. Es genügt anzunehmen, daß Achsenpunkte nicht detektiert wurden.

5. Für jede noch existierende Clique: Bestimme für ihre Knotenmenge die Anzahl der nichtleeren Schnittmengen mit den Knotenmengen der anderen Cliques. Für alle Cliques C_i und C_j , für deren Knotenmengen N_i und N_j gilt:

$$N_i \cap N_j \neq \emptyset \text{ für } i \neq j$$

entferne C_i , falls $\text{size}(N_i) < \text{size}(N_j)$ und N_i mehr nichtleere Schnittmengen mit den Knotenmengen der anderen Cliques aufweist als N_j .

Die letzten beiden Schritte 4 und 5 wurden lediglich dazu verwendet, um bei der Berücksichtigung von nur zwei oder wenigen Zeitpunkten, zufällig sich starr verhaltende Punktgruppen ausfindig zu machen und zu entfernen. Es wird angenommen, daß diese Cliques verhältnismäßig klein sind (siehe Schritt 4) und/oder daß sie mehr nichtleere Schnittmengen mit den anderen Cliques aufweisen (siehe Schritt 5). Je mehr Zeitpunkte herangezogen werden, desto geringer ist die Wahrscheinlichkeit von Punktgruppen, die sich nur zufällig starr verhalten, so daß die Schritte 4 und 5 immer stärker an Bedeutung verlieren.

Das Problem der Cliquensuche gehört zu der Klasse der NP-vollständigen Probleme und läßt sich daher im allgemeinen nur mit exponentiellem Aufwand lösen. Dieser Aufwand ist nur bei kleineren Punktmengen vertretbar.

Um zu zeigen, daß ein Punkt zu einem gegebenen starren Körper gehört, genügt es festzustellen, daß die Abstände dieses Punktes zu drei nicht kollinearen Punkten des starren Körpers konstant bleiben.² Man braucht also nicht alle Abstände des Punktes zu allen Punkten des starren Körpers zu überprüfen, um die Starrheit zu zeigen. Der Aufwand beim Aufbau des Starrheitsgraphen und die daraus resultierende Cliquensuche ist also gar nicht erforderlich, um die Starrheit einer Punktmenge sicherzustellen. Da man geringe Abstandsänderungen zulassen will, kann es bei der Überprüfung von nur drei Abständen allerdings von der Wahl dieser drei Punkte aus der bereits als starr angesehenen Punktmenge abhängen, ob der untersuchte Punkt auch dieser Menge zugesprochen wird oder nicht. Dieser Nachteil ist mit dem erhöhten Aufwand der Cliquensuche abzuwägen. Ein Kompromiß läge wahrscheinlich zwischen diesen Extremen.

Es sei angemerkt, daß die Starrheit von Punktmenge für die Korrespondenzanalyse eine wichtige Informationsquelle darstellt. Setzt man voraus, daß sich eine bisher starr verhaltende Punktmenge auch weiterhin starr verhält, so benötigt man zu einem späteren Zeitpunkt nur die räumlichen Positionen von drei nicht kollinearen Punkten dieser Menge, um die Positionen aller anderen Punkte der Punktmenge vorhersagen zu können.³ Werden Punkte aufgrund ihrer zeitlichen Bewegung in späteren Bildern verdeckt, so lassen sie sich mit Hilfe der Starrheitsannahme im Raum lokalisieren.

²Diese Tatsache wurde auf Seite 4 ausgenutzt, um die Einschränkungen für die Starrheit von N Punkten zu bestimmen.

³Siehe dazu Theorem 4.4 auf Seite 49

4.2 Mathematik der Bewegung starrer Körper

Ein starrer Körper ist die Idealisierung eines physikalischen Objektes, welches sich unter jeder physikalisch möglichen Bewegung derart verhält, daß die Abstände zwischen beliebigen Punkten des Objektes unverändert bleiben. Der nun folgende Abschnitt beschäftigt sich mit den mathematischen Grundlagen, die zur Beschreibung räumlicher Bewegung starrer Körper benötigt werden. Eine Übersicht findet sich z. B. in [Angeles 82].

Die Bewegung eines starren Körpers kann durch eine bijektive Abbildung M beschrieben werden, welche jeden 3D-Punkt \mathbf{p} des Körpers auf einen Punkt $M(\mathbf{p})$ im Raum eindeutig abbildet (M ist ein Endomorphismus des \mathcal{R}^3). Dabei stellt $M(\mathbf{p})$ den Ort im Raum dar, zu dem sich der Punkt \mathbf{p} bewegt hat. Damit die Abbildung eine starre Bewegung beschreibt, müssen die Abstände zwischen beliebigen Punkten \mathbf{p}_i und \mathbf{p}_j erhalten bleiben, das heißt:

$$|\mathbf{p}_i - \mathbf{p}_j| = |M(\mathbf{p}_i) - M(\mathbf{p}_j)|$$

Aus dieser Bedingung folgt, daß der Winkel zwischen beliebigen Geraden des starren Körpers konstant bleibt. Damit muß auch das Skalarprodukt zwischen den Punktabstandsvektoren durch die Abbildung erhalten bleiben, da es das Produkt aus Vektorbeträgen und dem Cosinus des eingeschlossenen Winkels ist:

$$(\mathbf{p}_k - \mathbf{p}_i)^T(\mathbf{p}_j - \mathbf{p}_i) = (M(\mathbf{p}_k) - M(\mathbf{p}_i))^T(M(\mathbf{p}_j) - M(\mathbf{p}_i)) \quad (4.1)$$

4.2.1 Rotationsbewegung

Im allgemeinen ist die Abbildung M nicht linear. Es gibt jedoch eine Klasse von Bewegungen, die durch eine lineare Transformation beschrieben werden kann.

Theorem 4.1 *Bleibt bei der Bewegung eines starren Körpers ein Punkt fixiert, dann (und nur dann) läßt sich die Bewegung durch eine lineare Transformation beschreiben.*

Beweis (siehe auch [Angeles 82]): Es sei o. B. d. A. angenommen, daß sich der fixierte Punkt im Ursprung des gewählten Koordinatensystems befindet. Unter dieser Bedingung muß für jeden Punkt \mathbf{p}_i des Körpers gelten:

$$|\mathbf{p}_i| = |M(\mathbf{p}_i)|$$

Diese Beziehung wird im folgenden ausgenutzt. Nehmen wir zunächst an, daß die Abbildung M nicht linear sei. Sei \mathbf{e} die Abweichung von einer linearen Abbildung, so erhalten wir:

$$\begin{aligned} \mathbf{e} &= M(\mathbf{p}_i + \mathbf{p}_j) - (M(\mathbf{p}_i) + M(\mathbf{p}_j)) \\ \Rightarrow |\mathbf{e}|^2 &= |M(\mathbf{p}_i + \mathbf{p}_j)|^2 + |M(\mathbf{p}_i) + M(\mathbf{p}_j)|^2 \\ &\quad - 2M(\mathbf{p}_i + \mathbf{p}_j)^T(M(\mathbf{p}_i) + M(\mathbf{p}_j)) \end{aligned}$$

$$\begin{aligned}
&= |\mathbf{p}_i + \mathbf{p}_j|^2 + |M(\mathbf{p}_i)|^2 + |M(\mathbf{p}_j)|^2 + 2M(\mathbf{p}_i)^T M(\mathbf{p}_j) \\
&\quad - 2M(\mathbf{p}_i + \mathbf{p}_j)^T (M(\mathbf{p}_i) + M(\mathbf{p}_j)) \\
&= |\mathbf{p}_i|^2 + |\mathbf{p}_j|^2 + 2\mathbf{p}_i^T \mathbf{p}_j + |\mathbf{p}_i|^2 + |\mathbf{p}_j|^2 + 2\mathbf{p}_i^T \mathbf{p}_j \\
&\quad - 2(\mathbf{p}_i + \mathbf{p}_j)^T \mathbf{p}_i - 2(\mathbf{p}_i + \mathbf{p}_j)^T \mathbf{p}_j \\
&= 2|\mathbf{p}_i|^2 + 2|\mathbf{p}_j|^2 + 4\mathbf{p}_i^T \mathbf{p}_j - 2|\mathbf{p}_i|^2 - 2|\mathbf{p}_j|^2 - 4\mathbf{p}_i^T \mathbf{p}_j \\
&= 0
\end{aligned}$$

Aus der Eigenschaft der Norm

$$|\mathbf{v}| \geq 0 \quad \text{und} \quad \forall \mathbf{v} : |\mathbf{v}| = 0 \Leftrightarrow \mathbf{v} = \mathbf{0}$$

folgt:

$$M(\mathbf{p}_i + \mathbf{p}_j) = M(\mathbf{p}_i) + M(\mathbf{p}_j)$$

Damit ist gezeigt, daß ein fixierter Punkt bei der Bewegung eines starren Körpers eine hinreichende Bedingung für eine lineare Abbildung darstellt. Umgekehrt bleibt bei jeder starren Bewegung, die durch eine lineare Abbildung beschrieben werden kann, ein Punkt fixiert, da der Nullvektor durch eine lineare Abbildung stets auf sich selbst abgebildet wird. Dieser stellt in dem Koordinatensystem, in dem die Bewegung beschrieben wird, den fixierten Punkt dar. Damit ist das Theorem 4.1 bewiesen. Der fixierte Punkt bei der Bewegung eines starren Körpers ist sowohl notwendige als auch hinreichende Bedingung für eine lineare Abbildung. Eine lineare Abbildung M läßt sich in der Mathematik durch eine Matrix M repräsentieren.⁴

Eine lineare Abbildung M , bei der der Betrag und damit auch das Skalarprodukt zwischen Vektoren (siehe Gleichung 4.1) erhalten bleibt, nennt man *Isometrie*. In unserem Fall läßt sie sich, bezüglich einer vorgegebenen Basis \mathcal{B} , durch eine 3×3 Matrix darstellen. Handelt es sich bei der linearen Abbildung um eine *Isometrie*, so ist die zugehörige Matrix M orthogonal, daß heißt:

$$M^{-1} = M^T$$

Beweis: $\mathbf{p}_t, \mathbf{p}_{t+1}$ seien die Koordinatenvektoren des Punktes \mathbf{p} zu den Zeitpunkten t und $t + 1$, dann gilt:

$$\begin{aligned}
&\mathbf{p}_{t+1} = M\mathbf{p}_t \\
\Rightarrow \mathbf{p}_{t+1}^T \mathbf{p}_{t+1} &= (M\mathbf{p}_t)^T (M\mathbf{p}_t) \\
&= \mathbf{p}_t^T M^T M \mathbf{p}_t \\
&= \mathbf{p}_t^T \mathbf{p}_t && (|\mathbf{p}_{t+1}| = |\mathbf{p}_t|) \\
\Rightarrow M^T M &= \mathbf{I} \\
\Rightarrow M^T &= M^{-1}
\end{aligned}$$

⁴Im folgenden benutze ich synonym für eine lineare Abbildung und deren darstellende Matrix dieselbe Bezeichnung M

Die Matrix \mathbf{I} stelle die Einheitsmatrix dar. Die Spaltenvektoren der Matrix \mathbf{M} bilden ein Orthonormalsystem⁵, da $\mathbf{M}^T\mathbf{M} = \mathbf{I}$. Die Orthogonalität von \mathbf{M} läßt sich auch anhand der kanonischen Basisvektoren $(1, 0, 0)^T$, $(0, 1, 0)^T$ und $(0, 0, 1)^T$ verdeutlichen. Da diese Vektoren normalisiert und orthogonal (=orthonormal) sind, müssen diese Eigenschaften auch für ihre Bilder gelten, damit der Betrag der Vektoren und der Winkel zwischen ihnen erhalten bleibt. Da die Bilder aber genau die Spaltenvektoren der Matrix \mathbf{M} sind, müssen diese folglich ein Orthonormalsystem bilden, so daß \mathbf{M} orthogonal ist.

Die Orthogonalität von \mathbf{M} ist jedoch noch keine hinreichende Bedingung für eine starre Bewegung, da die Eigenschaften der Isometrie auch für Spiegelungen gelten. Betrachten wir die kanonischen Basisvektoren, so gilt die Beziehung:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Vektoren mit dieser Eigenschaft bilden ein Rechtshandsystem. Die Bilder der kanonischen Basisvektoren ($\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3$) müssen genau eine der folgenden Bedingungen erfüllen, wenn die Matrix \mathbf{M} orthogonal ist:

$$\mathbf{m}_1 \times \mathbf{m}_2 = \mathbf{m}_3 \quad \text{oder} \quad \mathbf{m}_1 \times \mathbf{m}_2 = -\mathbf{m}_3$$

Gilt die erste Gleichung, so bilden auch die Bilder der kanonischen Basisvektoren und damit die Spaltenvektoren der Matrix \mathbf{M} ein Rechtshandsystem. Damit beschreibt \mathbf{M} eine mögliche starre Bewegung. Im zweiten Fall beschreiben die Spaltenvektoren ein Linkshandsystem, so daß die Abbildung geometrisch eine Spiegelung verursacht. Die Unterscheidung zwischen Spiegelung und Nichtspiegelung läßt sich auch noch anders ausdrücken. Dazu zunächst folgender Satz:

Satz 4.1 Die Determinante $\det(\mathbf{M})$ einer orthogonalen Matrix \mathbf{M} ist entweder +1 oder -1.

Besitzt die Determinante den Wert -1, so beschreibt die Matrix \mathbf{M} eine Spiegelung. Andernfalls beschreibt sie eine starre Bewegung (Rotation). Eine orthogonale Matrix mit einer Determinanten von +1 wird in der Mathematik als *eigentlich orthogonal*, andernfalls als *uneigentlich orthogonal* bezeichnet.

Als nächstes werden wir die Eigenwerte und Eigenvektoren von orthogonalen Matrizen näher betrachten. Sie führen uns auf den Begriff der Rotationsachse. Die Eigenwerte von quadratischen Matrizen können sowohl reel als auch komplex sein.

Satz 4.2 Die Eigenwerte einer orthogonalen Matrix liegen auf dem Einheitskreis der komplexen Ebene.

⁵Leider bezeichnet man in der Mathematik die zugehörige Matrix nur als orthogonal, obwohl die Spaltenvektoren sogar orthonormal sind!

Beweis: Sei M eine orthogonale $n \times n$ Matrix, λ ein Eigenwert mit dem zugehörigen Eigenvektor e .

$$Me = \lambda e \quad (4.2)$$

Das Transponieren und Konjugieren beider Seiten liefert:

$$e^* M^T = \lambda^* e^* \quad (4.3)$$

Multipliziert man die entsprechenden Seiten der Gleichungen 4.2 und 4.3 so erhält man:

$$e^* M^T M e = \lambda^* \lambda e^* e$$

Da die Matrix M nach Voraussetzung orthogonal ist, bekommen wir:

$$\begin{aligned} e^* e &= |\lambda|^2 e^* e \\ \Rightarrow |\lambda|^2 &= 1 \\ &\text{q. e. d.} \end{aligned}$$

Der Betrag der Eigenwerte, seien sie nun reel oder komplex, ist somit gleich 1. Man kann zeigen, daß eine orthogonale $n \times n$ Matrix genau einen reellen Eigenwert besitzt, vorausgesetzt n ist ungerade. Dieser muß nach dem obigen Satz gleich $+1$ oder -1 sein. Ist die Matrix *eigentlich orthogonal*, so ist der reelle Eigenwert gleich $+1$, ist sie *uneigentlich orthogonal*, so ist -1 ihr reeller Eigenwert. Für eine *eigentlich orthogonale* Matrix M erhalten wir:

$$\begin{aligned} Me &= +1e \\ \Rightarrow M\alpha e &= \alpha e \end{aligned}$$

Damit bleiben alle Punkte längs des Vektors e bei der durch M beschriebenen Bewegung fixiert.

Theorem 4.2 (Theorem von Euler) *Unterliegt ein starrer Körper einer Bewegung, bei der ein Punkt fixiert bleibt, so existiert eine Gerade durch diesen Punkt, so daß alle Punkte längs dieser Geraden ihre Lage im Raum nicht ändern. Die Gerade bezeichnet man als Rotationsachse, wobei der Rotationswinkel auf einer Ebene senkrecht zu dieser Achse gemessen wird.*

Eine Rotationsmatrix läßt sich damit eindeutig durch Rotationsachse und Rotationswinkel beschreiben. Diese Größen bezeichnet man daher auch als *Invarianten* einer Rotationsmatrix.⁶ Im weiteren Verlauf dieser Arbeit werden wir die Rotationsbewegung eines starrer Körpers entweder durch eine Rotationsmatrix oder durch die dazu äquivalente Darstellungsform von Rotationsachse und -winkel repräsentieren. Aus diesem Grund

⁶Die Orientierung einer Geraden im Raum läßt sich durch zwei Größen darstellen. Damit genügen insgesamt drei Größen zur Darstellung einer Rotation.

betrachten wir nun das Problem, wie sich diese beiden unterschiedlichen Repräsentationsformen einer Rotationsbewegung ineinander überführen lassen. Wie lassen sich Rotationsachse und Rotationswinkel aus einer gegebenen Rotationsmatrix bestimmen und umgekehrt, wie ergibt sich die Rotationsmatrix, falls Rotationsachse und Rotationswinkel bekannt sind?

Zur Bestimmung der Rotationsachse aus einer *eigentlich orthogonalen* Matrix M erinnern wir uns, daß der zum Eigenwert $+1$ gehörige Eigenvektor \mathbf{e} in Richtung der Rotationsachse weist:

$$\begin{aligned} M\mathbf{e} &= +1\mathbf{e} \\ \Rightarrow (M - I)\mathbf{e} &= \mathbf{0} \end{aligned}$$

Die Matrix I sei wieder die Einheitsmatrix. Die nichttrivialen Lösungen dieses linearen Gleichungssystems ergeben die Rotationsachse. Um den Winkel zu bestimmen, führen wir zunächst die kanonische Form einer Rotationsmatrix ein. Für jede Rotationsmatrix M bezüglich einer Basis \mathcal{B} läßt sich eine neue orthonormale Basis $\mathcal{B}' = [\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3]$ bestimmen, so daß die Rotationsachse parallel zu – sagen wir – \mathbf{b}'_3 ist. Dieses wird durch folgende Ähnlichkeitstransformation ausgedrückt:

$$M_{\mathcal{B}} = A_{\mathcal{B}'}^{-1} M_{\mathcal{B}'} A_{\mathcal{B}'} = A_{\mathcal{B}'}^T M_{\mathcal{B}'} A_{\mathcal{B}'} \quad (4.4)$$

$M_{\mathcal{B}}$ beschreibt die Rotation bezüglich der Ausgangsbasis $\mathcal{B} = [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3]$, $M_{\mathcal{B}'}$ beschreibt die Rotation bezüglich der Basis $\mathcal{B}' = [\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3]$. $A_{\mathcal{B}'}$ ist eine orthogonale Transformationsmatrix, welche die Komponenten eines Vektors bezüglich der Basis \mathcal{B} transformiert in die Komponenten bezüglich der Basis \mathcal{B}' . Damit besitzt die Matrix $M_{\mathcal{B}'}$ die folgende Gestalt:

$$M_{\mathcal{B}'} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$M_{\mathcal{B}'}$ bezeichnen wir als eine kanonische Form von $M_{\mathcal{B}}$. Folgende Beziehung läßt sich leicht entnehmen:

$$\begin{aligned} \text{Spur}(M_{\mathcal{B}'}) &= 2 \cos(\theta) + 1 \\ \Rightarrow \theta &= \arccos\left(\frac{1}{2}(\text{Spur}(M_{\mathcal{B}'}) - 1)\right) \end{aligned}$$

Da die Spur einer Matrix invariant ist gegenüber Ähnlichkeitstransformationen, gilt diese Beziehung auch für die Matrix $M_{\mathcal{B}}$. Wir haben damit eine einfache Möglichkeit zur Bestimmung des Rotationswinkels gefunden. Das Vorzeichen des Winkels läßt sich jedoch nicht aus dieser Gleichung entnehmen. Das folgende Theorem gibt uns eine Möglichkeit, dieses zu berechnen.

Theorem 4.3 *Gegeben sei die Rotationsbewegung eines starren Körpers um eine Achse, die durch den Ursprung des Koordinatensystems verläuft. Sei \mathbf{e} der Einheitsvektor in*

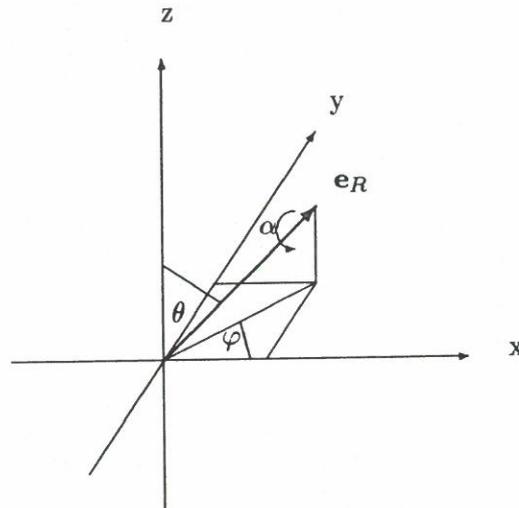


Abbildung 4.1: Darstellung der Rotationsachse \mathbf{e}_R in Eulerwinkeln; φ und θ bestimmen die Orientierung; es wird um den Winkel α rotiert.

Richtung der Rotationsachse, θ der Rotationswinkel. Ferner seien \mathbf{p} und \mathbf{p}' die Ortsvektoren eines Punktes vor und nach der Bewegung des starren Körpers, die nicht auf der Rotationsachse liegen ($\mathbf{p} \neq \mathbf{p}'$). Dann gilt:

$$\operatorname{sgn}((\mathbf{p} \times \mathbf{p}')^T \mathbf{e}) = \operatorname{sgn}(\theta).$$

Bei einer Rechts-/Linksdrehung um den Vektor \mathbf{e} ergibt sich damit ein positives/negatives Vorzeichen.

Im folgenden wird das umgekehrte Problem betrachtet, bei dem die Rotationsmatrix bezüglich der Basis \mathcal{B} aus Rotationsachse und Rotationswinkel bestimmt wird. Auch hier erweist sich die kanonische Form einer Rotationsmatrix als hilfreich. Legen wir unser Koordinatensystem so, daß die z -Achse mit der Rotationsachse zusammenfällt, so läßt sich die Rotationsmatrix entsprechend der obigen kanonischen Form sofort angeben. Gesucht ist somit eine Ähnlichkeitstransformation, durch die die Rotation bezüglich einer neuen Basis $\mathcal{B}' = (\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{e})$ beschrieben werden kann. Dabei sei \mathbf{e} der Einheitsvektor der Rotationsachse. Die Basisvektoren \mathbf{b}'_1 und \mathbf{b}'_2 sind so zu wählen, daß \mathcal{B}' ein Orthornormalsystem bildet. Die Rotationsmatrix $\mathbf{M}_{\mathcal{B}}$ erhält man aus Gleichung 4.4. Wenn die Rotationsachse und der Rotationswinkel in Form von Eulerwinkeln (φ, θ, α) (wie in Abbildung 4.1 gezeigt) gegeben sind, so ergibt sich $\mathbf{A}_{\mathcal{B}'}$ durch eine Rotation um den Winkel φ um die z -Achse, gefolgt von einer Rotation um den Winkel θ um die neue y -Achse. Dadurch weist die neue z -Achse in Richtung der Rotationsachse. Zur Ausführung der eigentlichen Rotation wird nun um den Winkel α um die neue z -Achse rotiert. Danach

werden die durch die Matrix $\mathbf{A}_{B'}$ erfolgten Rotationen wieder rückgängig gemacht. Die Matrizen $\mathbf{A}_{B'}$ und $\mathbf{M}_{B'}$ haben somit die folgende Gestalt:

$$\mathbf{A}_{B'} = \begin{pmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{pmatrix} \begin{pmatrix} \cos(\varphi) & \sin(\varphi) & 0 \\ -\sin(\varphi) & \cos(\varphi) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{M}_{B'} = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

4.2.2 Allgemeine Bewegung im Raum

Bislang haben wir nur Rotationsbewegungen von starren Körpern betrachtet. Diese Bewegungen lassen sich mathematisch durch lineare Abbildungen beschreiben. Um jede beliebige Bewegung eines starren Körpers mathematisch ausdrücken zu können, müssen wir noch einen zusätzlichen Translationsanteil in die Bewegungsgleichung hinzufügen. Eine allgemeine räumliche Bewegung eines starren Körpers läßt sich nun durch folgende Transformation beschreiben:

$$\mathbf{p}' = \mathbf{R}\mathbf{p} + \mathbf{t}$$

$$\Rightarrow \begin{pmatrix} p'_x \\ p'_y \\ p'_z \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}$$

\mathbf{p} und \mathbf{p}' stellen einen Punkt des starren Körpers zu zwei aufeinanderfolgenden Zeitpunkten dar. Wie wir bereits gesehen haben, ist die Matrix \mathbf{R} eine *eigentlich orthogonale* Matrix, die die Rotation des starren Körpers um eine Achse durch den Ursprung beschreibt. Der Vektor \mathbf{t} gibt die zusätzliche geradlinige Verschiebung (Translation) des starren Körpers an.

Der Translationsanteil in der Bewegung bewirkt, daß kein Punkt bei der Bewegung des starren Körpers fixiert bleibt. Dadurch läßt sich nach Theorem 4.1 eine allgemeine Bewegung im Raum nicht mehr durch eine lineare Abbildung darstellen. Diesen Nachteil können wir umgehen, indem wir homogene Koordinaten verwenden (siehe [Ballard + Brown 82] oder auch [Duda + Hart 73]).

Ein Vektor im \mathcal{R}^3 wird in homogenen Koordinaten (vierdimensionaler Raum) durch eine Gerade dargestellt, die durch den Ursprung geht. Für \mathbf{p} schreiben wir jetzt:

$$\begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} \rightarrow \begin{pmatrix} wp_x \\ wp_y \\ wp_z \\ w \end{pmatrix} \quad w \neq 0$$

Die ersten drei Komponenten dividiert durch die vierte Komponente des Vektors ergeben die Koordinaten des Punktes \mathbf{p} im \mathcal{R}^3 . Eine allgemeine Bewegung läßt sich nun in

homogenen Koordinaten durch die folgende lineare Abbildung darstellen:

$$\begin{aligned} \mathbf{p}' &= \mathbf{TRp} & (4.5) \\ \Rightarrow \begin{pmatrix} p'_x \\ p'_y \\ p'_z \\ 1 \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_x \\ p_y \\ p_z \\ 1 \end{pmatrix} \\ \Rightarrow \begin{pmatrix} p'_x \\ p'_y \\ p'_z \\ 1 \end{pmatrix} &= \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_x \\ p_y \\ p_z \\ 1 \end{pmatrix} = \mathbf{Mp} \end{aligned}$$

Die Matrix \mathbf{M} beinhaltet sowohl die Rotations- als auch die Translationsparameter der Bewegung. Derartige Matrizen werden wir im folgenden als Bewegungsmatrizen bezeichnen.

Anmerkung zu homogenen Koordinaten: Wie bei jeder linearen Abbildung bleibt auch bei dieser der Nullvektor $\mathbf{v} = (0, 0, 0, 0)^T$ durch die Abbildung erhalten. Dieser Punkt besitzt allerdings kein Pendant im \mathcal{R}^3 , da der Skalierungsfaktor w nicht Null werden darf. Damit haben wir uns von dem Problem des fixierten Punktes befreit.

In diesem Abschnitt haben wir kennengelernt, wie sich die Bewegung starrer Körper mathematisch beschreiben läßt. Wir müssen uns nun fragen, wie wir diese Bewegungsparameter ermitteln können. Bekannt sind uns die 3D-Koordinaten von Punkten des starren Körpers zu den diskreten Aufnahmezeitpunkten. Für zwei aufeinanderfolgende Zeitpunkte werden wir die Bewegungsparameter so bestimmen, daß die zeitlich korrespondierenden 3D-Punkte zu den beiden Aufnahmezeitpunkten ineinander überführt werden. Wir wissen nicht, ob diese ermittelte Bewegung mit der tatsächlichen Bewegung des starren Körper übereinstimmt, da wir nicht feststellen können, auf welchem Weg sich der Körper zwischen den beiden Zeitpunkten bewegt hat. Die Zeitintervalle müssen also klein genug sein, damit die ermittelten Bewegungsparameter alle wesentlichen Bewegungscharakteristiken des starren Körpers erfassen. Auf der anderen Seite dürfen die Zeitpunkte nicht zu dicht gewählt werden, da sonst die Strecken, die von den Punkten in den betreffenden Zeitintervallen zurückgelegt werden, und die Rekonstruktionsfehler in derselben Größenordnung liegen.

4.3 Beschreibung von allgemeiner Bewegung im Raum

Im weiteren gehen wir davon aus, daß Szenenpunkte zu verschiedenen Bildaufnahmezeitpunkten aus korrespondierenden Bildpunkten rekonstruiert wurden. Diese Szenenpunkte werden wir nun zunächst zu starren Körpern gruppieren. In Abschnitt 4.1 hatten wir ein Verfahren dazu angegeben. Für jeden so erhaltenen starren Körper versuchen wir die Bewegung in jedem Zeitintervall durch Rotation und Translation zu beschreiben. Dazu stehen uns die 3D-Koordinaten der rekonstruierten Szenenpunkte zu jeweils zwei aufeinanderfolgenden Zeitpunkten zur Verfügung. Die Bewegungsparameter sind nun so zu

bestimmen, daß die Szenenpunkte zum ersten Zeitpunkt in die entsprechenden Szenenpunkte zum zweiten Zeitpunkt überführt werden.

In [Martin + Aggarwal 88, Kapitel 10] werden von *Blostein + Huang* verschiedene Verfahren vorgestellt, mit denen sich die Bewegungsparameter von starren Körpern, basierend auf 3D-Punktkorrespondenzen, ermitteln lassen. Blostein und Huang unterscheiden die Verfahren grob nach linearen und nichtlinearen. Sie versuchen die einzelnen Verfahren zu bewerten, indem sie deren Fehleranfälligkeit bei verrauschten Eingangsdaten messen. Wir werden nicht im einzelnen auf die von ihnen erzielten Ergebnisse eingehen, sondern nur die wesentlichen Schlußfolgerungen wiedergeben.

Blostein und Huang stellen fest, daß die linearen Verfahren sich zwar im allgemeinen durch einen größeren Berechnungsaufwand auszeichnen, ihre numerische Stabilität dafür allerdings, im Gegensatz zu den nichtlinearen Verfahren, unabhängig von der Bewegung ist. Weiterhin ergibt sich aus ihren Untersuchungen, daß die Fehleranfälligkeit abnimmt, je mehr 3D-Korrespondenzen bei der Berechnung berücksichtigt werden. Es sind also bessere Ergebnisse zu erwarten, wenn man das Problem genügend überbestimmt. Wir werden aus diesen Gründen ein lineares Verfahren verwenden und die Überbestimmung durch ein "least-square-fit"-Verfahren ermöglichen.

Es stellt sich nun zunächst die Frage, durch wieviele Punkte die Lage eines starren Körpers im Raum eindeutig festgelegt wird.

Theorem 4.4 *Drei nicht kollineare Punkte auf einem starren Körper sind notwendig und hinreichend, um seine Position und Orientierung im dreidimensionalen Raum eindeutig bestimmen zu können.*

Drei verschiedene, nicht kollineare Punkte bestimmen eine Ebene. Diese Ebene legt die Position und Orientierung des gesamten starren Körpers fest. Die Anzahl der notwendigen Punkte stellt eine untere Schranke für die benötigten zeitlichen Korrespondenzen dar.

Im folgenden wird beschrieben, wie die gewünschten Bewegungsparameter bestimmt werden können.

4.3.1 Linearer Ansatz

Kennen wir die räumlichen Koordinaten von vier Punkten eines starren Körpers zu jeweils zwei aufeinanderfolgenden Zeitpunkten, so ergeben sich aus Gleichung 4.5 vier lineare Gleichungssysteme, aus denen die 12 unbekannt Parameter r_{11} , r_{12} , r_{13} , r_{21} , r_{22} , r_{23} , r_{31} , r_{32} , r_{33} , t_x , t_y und t_z bestimmt werden können:

$$\begin{pmatrix} p'_{1x} & p'_{2x} & p'_{3x} & p'_{4x} \\ p'_{1y} & p'_{2y} & p'_{3y} & p'_{4y} \\ p'_{1z} & p'_{2z} & p'_{3z} & p'_{4z} \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_{1x} & p_{2x} & p_{3x} & p_{4x} \\ p_{1y} & p_{2y} & p_{3y} & p_{4y} \\ p_{1z} & p_{2z} & p_{3z} & p_{4z} \\ 1 & 1 & 1 & 1 \end{pmatrix} \quad (4.6)$$

$$\Rightarrow \mathbf{P}' = \mathbf{TRP}$$

$$\Rightarrow \mathbf{P}'\mathbf{P}^{-1} = \mathbf{TR}$$

Damit dieses lineare Gleichungssystem lösbar ist, muß die Matrix \mathbf{P} invertierbar sein. Daraus folgt, daß die Punkte nicht koplanar sein dürfen. Obwohl mathematisch eine Bewegung bereits durch drei nicht kollineare Punkte eindeutig festgelegt ist, benötigt man zur eindeutigen Lösung von Gleichung 4.6 genau vier nicht koplanare Punkte. Der Grund hierfür liegt darin, daß die Abhängigkeiten der neun Rotationsparameter (siehe Abschnitt 4.2) nicht in der Gleichung zum Ausdruck kommen. Damit enthält die Gleichung 4.6 zusätzliche Freiheitsgrade, durch die auch nichtstarre Bewegungen beschrieben werden können. Diese nichtstarrten Bewegungen lassen sich durch Scherungen und Skalierungen⁷ charakterisieren. Es muß daher vorher oder hinterher geprüft werden, ob die Punkte sich starr verhalten haben. Zum einen läßt sich dies an sich ändernden Punkt-Abständen ablesen (siehe Abschnitt 4.1), zum anderen muß die Rotationsmatrix *eigentlich orthogonal* sein (siehe Abschnitt 4.2).

Da man im allgemeinen mit fehlerbehafteten Daten rechnen muß, ist es sinnvoll, möglichst viele Punkte zur Bewegungsberechnung heranzuziehen. Für n Punkte sieht das zu lösende Gleichungssystem dann wie folgt aus:

$$\begin{pmatrix} p'_{1x} & p'_{2x} & \cdots & p'_{nx} \\ p'_{1y} & p'_{2y} & \cdots & p'_{ny} \\ p'_{1z} & p'_{2z} & \cdots & p'_{nz} \\ 1 & 1 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_{1x} & p_{2x} & \cdots & p_{nx} \\ p_{1y} & p_{2y} & \cdots & p_{ny} \\ p_{1z} & p_{2z} & \cdots & p_{nz} \\ 1 & 1 & \cdots & 1 \end{pmatrix} \quad (4.7)$$

$$\Rightarrow \mathbf{P}' = \mathbf{TRP}$$

$$\Rightarrow \mathbf{P}' = \mathbf{MP} \quad (4.8)$$

Für $n > 3$ ist die Matrix \mathbf{P} nicht mehr invertierbar, da sie nicht mehr quadratisch ist. Das Gleichungssystem ist damit überbestimmt, wodurch eine eindeutige Lösung nicht mehr zu existieren braucht. Stattdessen sucht man diejenigen Transformationsparameter, die die Bewegung der gegebenen Punkte am besten wiedergeben. Eine übliche Methode ist, die Summe der quadratischen Fehlerabweichungen zu minimieren.⁸

Fehler:

$$\mathbf{e}_i = \mathbf{p}'_i - \mathbf{M}\mathbf{p}_i$$

$$\text{oder} \quad \mathbf{E} = \mathbf{P}' - \mathbf{MP}$$

Summe der Fehlerquadrate:

$$\begin{aligned} \sum_{i=0}^n \mathbf{e}_i^2 &= \sum_{i=0}^n (\mathbf{p}'_i - \mathbf{M}\mathbf{p}_i)^2 \\ &= \sum_{i=0}^n (\mathbf{p}'_i - \mathbf{M}\mathbf{p}_i)^T (\mathbf{p}'_i - \mathbf{M}\mathbf{p}_i) \end{aligned}$$

⁷Die bereits erwähnte Spiegelung ist ein Spezialfall einer Skalierung.

⁸Der hier verwendete Ansatz, bei dem die Summe der quadratischen Fehler minimiert wird, ist bis auf die fehlenden Gewichtungsfaktoren identisch mit dem Ansatz, den wir zur Approximation einer Punkttrajektorie auf Seite 22 verwendet haben.

$$\begin{aligned}
&= \text{Spur}(\mathbf{E}^T \mathbf{E}) \\
&= \text{Spur}((\mathbf{P}' - \mathbf{M}\mathbf{P})^T (\mathbf{P}' - \mathbf{M}\mathbf{P})) \\
&= \text{Spur}(\mathbf{P}'^T \mathbf{P}' - \mathbf{P}'^T \mathbf{M}\mathbf{P} - (\mathbf{M}\mathbf{P})^T \mathbf{P}' + (\mathbf{M}\mathbf{P})^T \mathbf{M}\mathbf{P}) \\
&= \text{Spur}(\mathbf{P}'^T \mathbf{P}' - \mathbf{P}'^T \mathbf{M}\mathbf{P} - (\mathbf{P}'^T \mathbf{M}\mathbf{P})^T + \mathbf{P}^T \mathbf{M}^T \mathbf{M}\mathbf{P}) \rightarrow \text{Minimum}
\end{aligned}$$

Ableitung nach \mathbf{M} :

$$\begin{aligned}
\frac{d\text{Spur}(\mathbf{E}^T \mathbf{E})}{d\mathbf{M}} &= -2\mathbf{P}'\mathbf{P}^T + 2(\mathbf{P}^T \mathbf{M}^T)^T \mathbf{P}^T = \mathbf{0} \\
\Rightarrow \mathbf{P}'\mathbf{P}^T &= \mathbf{M}\mathbf{P}\mathbf{P}^T \\
\Rightarrow \mathbf{M} &= \mathbf{P}'\mathbf{P}^T (\mathbf{P}\mathbf{P}^T)^{-1}
\end{aligned}$$

Die verwendeten Ableitungsregeln für Matrizen stammen wiederum aus [Horn 86, S. 458] und finden sich auch im Anhang dieser Arbeit. Die Matrix $\mathbf{P}^T (\mathbf{P}\mathbf{P}^T)^{-1}$ nennt man *Pseudoinverse* der Matrix \mathbf{P} .

Wie wir bereits erwähnten, kann die Matrix $\mathbf{M} = \mathbf{T}\mathbf{R}$ auch nichtstarre Bewegungen beschreiben. Damit stellt die Summe der Fehlerquadrate, für sich allein genommen, kein eindeutiges Maß für die Starrheit der Bewegung dar. Zusätzlich ist die Orthogonalität der Matrix \mathbf{R} zu berücksichtigen.

Mit dem hier vorgestellten Verfahren läßt sich die Bewegung eines starren Körpers zwischen zwei aufeinanderfolgenden Zeitpunkten ermitteln. Wir erhalten somit für jedes Zeitintervall Rotations- und Translationsparameter. Im nächsten Kapitel werden wir versuchen, die Bewegung von starren Körpern über einen längeren Zeitraum zu beschreiben. Dazu müssen wir die Bewegungsparameter, die wir für die einzelnen Zeitintervalle erhalten haben, geeignet zusammenfassen. Dabei werden wir für jeden starren Körper eine sogenannte physikalische Rotationsachse bestimmen. Weiterhin werden wir Kopplungen zwischen starren Körpern erkennen und beschreiben. Wir erhalten also einen größeren Einblick in die Bewegungen der starren Körper und in die Kopplungen die zwischen ihnen existieren können. Dieses wird es uns ermöglichen, die Bewegung jedes Körpers besser vorherzusagen.

Kapitel 5

Bewegungsbeschreibung von zusammengesetzten Objekten

Mit den im vorigen Kapitel vorgestellten Verfahren können wir die Bewegung eines starren Körpers zwischen zwei aufeinanderfolgenden Zeitpunkten durch Rotation und Translation beschreiben. Dabei wird die Orientierung der Rotationsachse sowie der Rotationswinkel durch die Bewegung eindeutig festgelegt.¹ Bestimmt man die Bewegungsparameter bezüglich eines anderen Koordinatensystems, so bleibt die räumliche Orientierung der Rotationsachse erhalten. Der Rotationswinkel ist ebenfalls konstant. Fallen die Ursprünge der beiden Koordinatensysteme nicht zusammen, so hat sich die Lage der Rotationsachse im Raum verändert.² Daraus resultiert eine unterschiedliche Translationsbewegung in den beiden Koordinatensystemen. Veranschaulicht wird diese Abhängigkeit in Abbildung 5.1. Die Bewegung des Quaders kann in der XY-Ebene beschrieben werden. Es wird zunächst um eine Rotationsachse rotiert, die senkrecht aus der Bildebene heraustritt. Der Rotationswinkel beträgt 90° . Nach der Rotation wird der Quader in der XY-Ebene verschoben. Dabei hängt der Translationsvektor von der Lage des Koordinatensystemursprungs ab, um den rotiert wurde.

Die hier auf eine Ebene beschränkte Betrachtung läßt sich schnell auf den dreidimensionalen Fall verallgemeinern, indem wir das Koordinatensystem durch eine z-Achse erweitern. In unserem Beispiel verläuft die Rotationsachse dann parallel zur z-Achse. Für die Wahl des Koordinatensystemursprungs erhalten wir nun mit der hinzugekommenen z-Achse einen zusätzlichen Freiheitsgrad. Die Wahl des Ursprungs in Richtung der z-Achse, und damit in Richtung der Rotationsachse, verändert allerdings nicht die räumliche Lage der Achse und damit auch nicht die erforderliche Translation. Als weitere Folgerung ergibt sich, daß lediglich der Translationsanteil senkrecht zur Rotationsachse von der Wahl des Koordinatensystems betroffen ist. Der erforderliche Translationsanteil in Richtung der Rotationsachse wird nicht von der Rotation beeinflusst und ist daher

¹Diese Größen hängen nur von der räumlichen Orientierung des Objektes vor und nach der Bewegung ab. Wir hatten sie bereits in Abschnitt 4.2.1 als Invarianten einer Rotationsmatrix kennengelernt.

²Erinnern wir uns, daß stets um den Ursprung des Koordinatensystems rotiert wird.

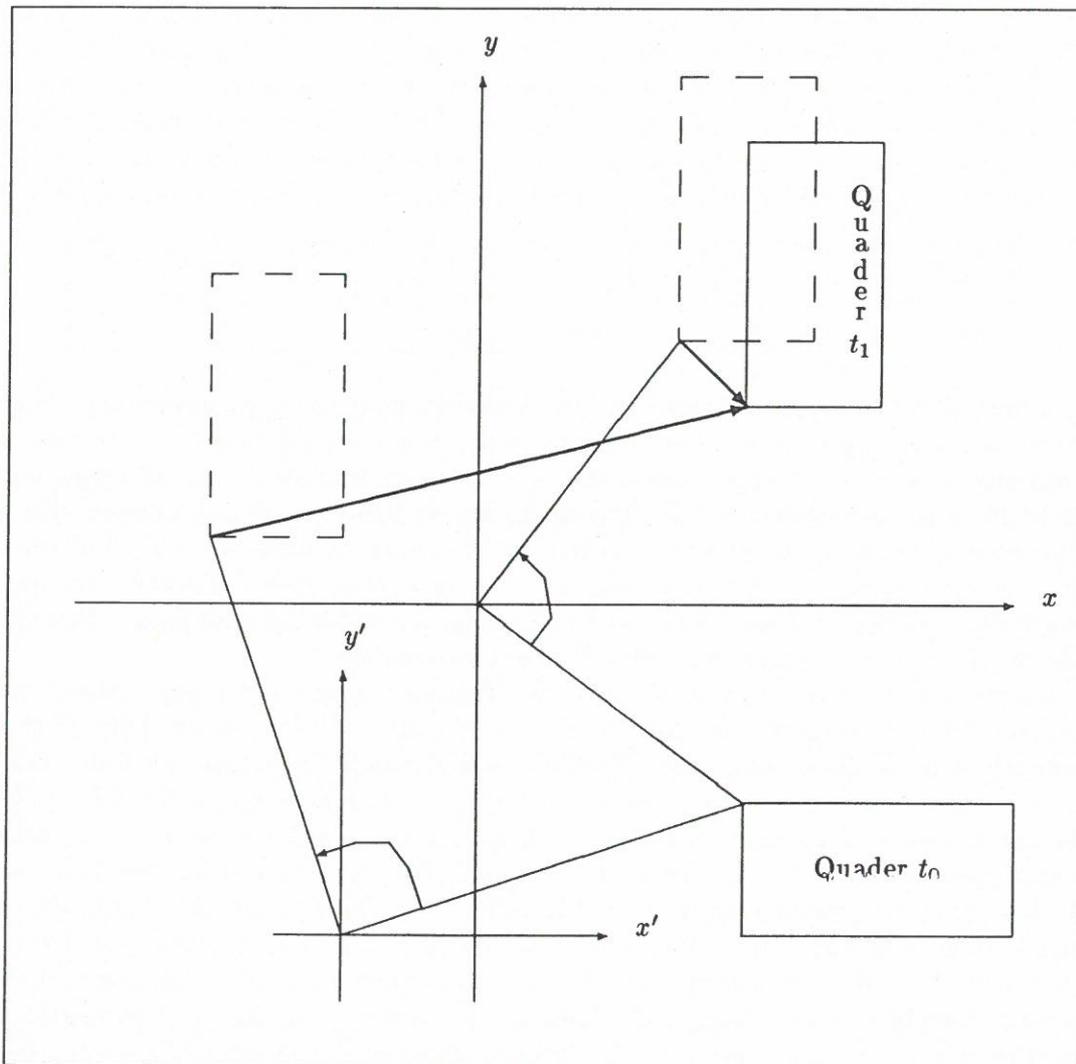


Abbildung 5.1: Bewegung eines Quaders bezüglich zweier verschiedener Koordinatensysteme

ebenfalls eine Invariante der Bewegung.

Wir können damit unsere Betrachtungen auf eine Ebene senkrecht zur Rotationsachse beschränken. Damit verbleiben zwei Freiheitsgrade³ für die Wahl eines geeigneten Koordinatensystems, um die räumliche Lage der Rotationsachse festzulegen.

Die Zerlegung der Bewegung in Rotation und Translation findet sich auch in der Physik wieder. Die Rotationsbewegung eines starren Körpers wird durch den Drehimpuls und die Translationsbewegung durch den Impuls beschrieben. Beide Größen sind sogenannte Erhaltungsgrößen, d. h. sie bleiben konstant, solange von außen keine verändernde Wirkung eintritt. Mögliche Wirkungen wären ein Drehmoment, welches die Winkelgeschwindigkeit und damit den Rotationszustand des Körpers ändert, bzw. eine Kraft, die die Geschwindigkeit und damit den Translationszustand des Körpers beeinflusst. Um

Bewegung	Ursache	Wirkung	Erhaltungsgröße
Rotation	Drehmoment	Winkelbeschleunigung	Drehimpuls
Translation	Kraft	Beschleunigung	Impuls

den ermittelten Bewegungsparametern physikalische Bedeutung beimessen zu können, müssen wir uns fragen, in welchem Koordinatensystem die physikalischen Ursachen der Bewegung am leichtesten erkennbar sind. Anders ausgedrückt stellt sich die Frage, durch welche Rotationsachsenlage wir die Bewegung am einfachsten erfassen können. Da wir nicht wissen, wie sich das Objekt zwischen den Zeitpunkten verhalten hat, sind wir bei nur zwei Zeitpunkten nicht in der Lage, aus den möglichen Beschreibungen, die physikalisch sinnvollste zu wählen. Mehr Aufschluß über die wahre (physikalische) Bewegung können wir uns erst aus dem zeitlichen Kontext verschaffen.

Damit wir die ermittelten Rotations- und Translationsparameter physikalisch interpretieren können, fordern wir, daß die zu bestimmende Rotationsachse körperfest ist. Ermitteln wir die Bewegung eines Objektes in aufeinanderfolgenden Zeitintervallen in einem beliebigen Koordinatensystem, so verlaufen die sich ergebenden Rotationsachsen alle durch dessen Ursprung. Durch diese Einschränkung müssen die Rotationsachsen nicht körperfest zum beschriebenen Objekt sein. Dies ist genau dann der Fall, wenn die Bewegungsbestimmung einen von null verschiedenen Translationsanteil senkrecht zur Rotationsachse ergeben hat. Beschreiben wir die Rotation des Objektes um derartige nicht körperfeste Achsen, so ergeben sich Translationsparameter, die nicht geeignet sind, korrekte Aussagen über physikalische Größen wie Geschwindigkeit und Beschleunigung des Objektes zu machen. Damit wir die Bewegungsparameter physikalisch interpretieren können, suchen wir also zunächst eine zum jeweiligen Objekt körperfeste Rotationsachse. Es genügt, die Lage der Rotationsachse im Raum zu Beginn der Bewegungsmessung festzulegen. Daraus ergeben sich für uns zwei verbleibende Freiheitsgrade. Die sich ergebende Translationsbewegung beschreibt dann die räumliche Bewegung der Rotationsachse.

Um die Bewegung über einen längeren Zeitraum beschreiben zu können, bestimmen wir zunächst die Bewegungsparameter (Rotation und Translation) in dem gegebenen Koordi-

³In unserem Fall der Durchstoßpunkt der Rotationsachse in der XY-Ebene.

natensystem für alle Zeitintervalle (t_k, t_{k+1}) $k = 0, \dots, n - 1$. Aus den Rotationsmatrizen lassen sich, wie in Abschnitt 4.2.1 gezeigt wurde, die Orientierungen der Rotationsachsen bestimmen.

Weisen die Rotationsachsen alle in dieselbe Richtung, so erscheint es physikalisch sinnvoll, die Lage der Rotationsachse so zu legen, daß eine "einfache" Translationsbewegung resultiert. Eine einfache Bewegung gibt die physikalischen Ursachen am besten wieder. Ein bekanntes Beispiel aus der Geschichte der Physik ist die Beschreibung unseres Sonnensystems. Vor Nikolaus Kopernikus wurde die Erde als Zentrum der Welt betrachtet und damit versucht, die Bewegung der übrigen Planeten relativ zu ihr zu beschreiben. Es ergaben sich komplizierte Planetenbahnen, die nur durch komplizierte Gleichungen ausgedrückt werden konnten. Erst die Idee, die Sonne als Zentrum der Planetenbahnen anzusehen und damit der Erde eine Bewegung um die Sonne zuzuschreiben, ergab einfache Verhältnisse, aus denen sich die Bewegungsursachen (Gravitationswechselwirkung) leichter entnehmen ließen. Dies Beispiel sei als Motivation für die Annahme einer "einfachen" Translationsbewegung gedacht. Für den Begriff "einfach" läßt sich hier keine exakte Definition geben. Als Beispiele für "einfache" Translationsbewegungen lassen sich z. B. geradlinige Bewegung mit konstanter Geschwindigkeit oder mit konstanter Beschleunigung anführen.

Wie wir oben bereits ausgeführt haben, stehen uns für die Wahl einer körperfesten Rotationsachse zwei Freiheitsgrade zur Verfügung. Diese werden wir in Abschnitt 5.1.2 nutzen, um eine "einfache" Translationsbewegung zu erhalten. Die resultierende Rotationsachse ist dann eher geeignet, als physikalische Rotationsachse des Objektes interpretiert zu werden. Unter der physikalischen Rotationsachse eines Objektes verstehen wir eine körperfeste Rotationsachse, die die physikalischen Bewegungsgrößen des Objektes wie Geschwindigkeit und Beschleunigung (Translationsgrößen) am besten wiedergibt.

Die Abbildungen 5.2 und 5.3 verdeutlichen die Wahl einer geeigneten Rotationsachsenlage. Der abgebildete Quader (Aufsicht) rotiert mit konstanter Winkelgeschwindigkeit um eine zur Bildebene senkrechten Achse und bewegt sich mit konstanter Geschwindigkeit (5.2) bzw. mit konstanter Beschleunigung (5.3). Als Beispiel für eine körperfeste Rotationsachsenlage sind in beiden Bildern jeweils zwei Punkte des Objektes mit den Symbolen * und \square markiert. In beiden Fällen erscheint die Wahl des Punktes * für die Lage der Rotationsachse geeigneter, da sich eine einfachere, weil geradlinige Bewegung der Rotationsachse bzw. des Objektes ergibt.

Besitzen die ermittelten Rotationsachsen hingegen alle unterschiedliche Orientierungen, so ist dies ein Indiz dafür, daß der betroffene starre Körper K_2 mehreren gleichzeitigen Rotationen unterliegt. Der Körper K_2 rotiert damit relativ zu einem Körper K_1 , der seinerseits eine Rotationsbewegung ausführt. Wir wollen K_1 als den zu K_2 übergeordneten Körper oder auch als Objektvater bezeichnen. Abbildung 5.4 zeigt ein derartiges Objekt, mit dem auch die Bildfolge aus Abbildung 3.2 erzeugt wurde. Die dicken schwarzen Pfeile markieren die Rotationsachsen der Teile. Teil 1 ist dem Teil 2 übergeordnet, da Teil 2 um eine konstante Rotationsachse relativ zu Teil 1 rotiert und Teil 1 ebenfalls eine Rotation um eine konstante Rotationsachse ausführt. Da die beiden Rotationsachsen unterschiedliche Orientierungen im Raum haben, ändert sich die Orientierung der gemes-

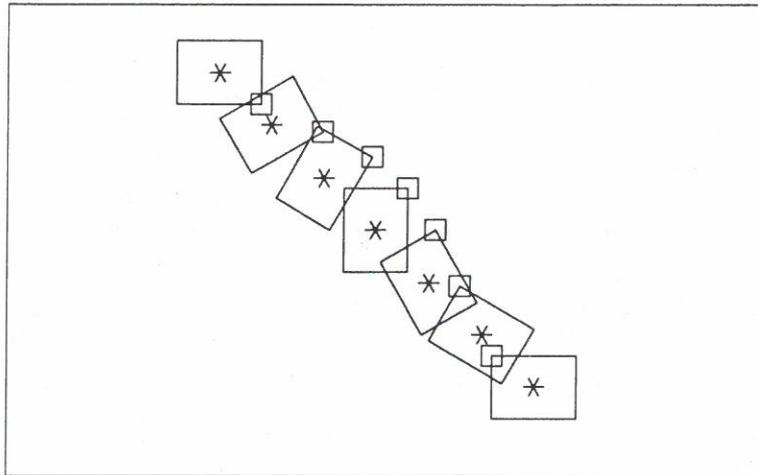


Abbildung 5.2: Beispiele für die Wahl einer körperfesten Rotationsachsenlage; Objekt bewegt sich mit konstanter Geschwindigkeit

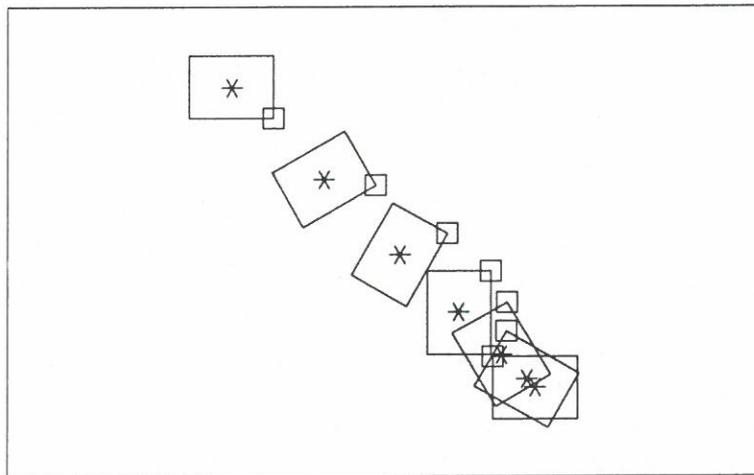


Abbildung 5.3: Beispiele für die Wahl einer körperfesten Rotationsachsenlage; Objekt bewegt sich mit konstanter Beschleunigung

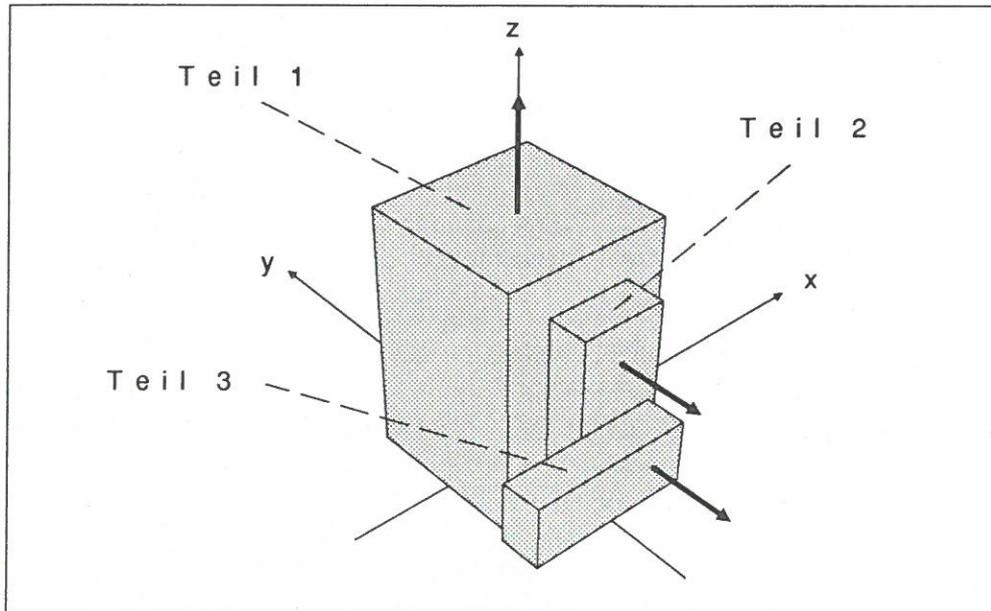


Abbildung 5.4: Objekt, das aus starren bewegbaren Einzelteilen zusammengesetzt ist

senen Rotationsachse von Teil 2 von einem Zeitpunkt zum nächsten. In einem solchen Fall ist es zunächst sinnvoll, die Bewegung von Teil 2 relativ zu Teil 1 zu beschreiben (siehe Abschnitt 5.1.1). Wir erhalten damit eine Bewegungsbeschreibung, bei der die Orientierung der Rotationsachse konstant bleibt. Für diese Achse läßt sich nun wieder die Lage im Raum derart ermitteln, daß eine einfache Translationsbewegung resultiert (siehe Abschnitt 5.1.2).

Die Idee zu dem hier kurz skizzierten Ansatz stammt von *Asada + Tsuji 83*. Wir werden uns zunächst intensiver mit deren Ansatz beschäftigen und danach eine eigene Vorgehensweise beschreiben, die auf dem Ansatz von *Asada + Tsuji 83* aufbaut.

5.1 Der Ansatz von Asada und Tsuji

Asada + Tsuji 83, *Asada + Tsuji 81* versuchen in ihrem Ansatz, die Bewegung von zusammengesetzten Objekten zu beschreiben. Sie setzen voraus, daß ihnen genügend viele räumliche Objektpunkte zu unterschiedlichen Zeitpunkten zur Verfügung stehen.⁴ Weiterhin kennen sie die starren Einzelteile der untersuchten Objekte. Ziel ihrer Untersuchungen ist es, eine einfache und natürliche Bewegungsbeschreibung zu gewinnen. Was sie darunter verstehen, läßt sich dem folgenden Zitat aus [*Asada + Tsuji 83*, Seite 118] entnehmen:

⁴Ihre Untersuchungen wurden an künstlich erzeugten Bildfolgen durchgeführt, vergleichbar mit der in Abbildung 3.2 dargestellten. Dadurch waren die räumlichen Koordinaten der Objektpunkte vorgegeben.

We consider that motion understanding is not to record numerical values of locations and orientations of objects but to interpret the changes in the three-dimensional geometry in terms of a simple and natural representation of the motion through the image sequence ...

A simple and natural representation of motion means that the number of terms describing the motion is as small as possible and that the representation is capable of giving us useful information such as prediction of positions of and relations between objects in the near future.

Obwohl ein starres Teil, wie z. B. Teil 2 in Abbildung 5.4 mehreren gleichzeitigen Rotationen ausgesetzt sein kann, erhalten wir nur eine Rotationsachse und einen Rotationswinkel, wenn wir die Bewegung des Teiles für ein Zeitintervall berechnen. Zur Unterscheidung bezeichnen Asada und Tsuji die berechnete Rotation daher als scheinbar (apparent), da sie die Überlagerung von physikalischen Rotationen darstellt. Für das Verständnis der Bewegung ist eine Zerlegung der scheinbaren Rotation in die physikalischen Rotationen wünschenswert, da so die Verbindungen der Objektteile und damit die Abhängigkeiten zwischen den Bewegungen dieser Objektteile sichtbar werden.

5.1.1 Zerlegung der Rotationsbewegung

Ob eine gemessene Rotation durch eine Überlagerung von verschiedenen physikalischen Rotationen zustande gekommen ist, wird erst deutlich, wenn man die Rotationsparameter über einen längeren Zeitraum zwischen jeweils aufeinanderfolgenden Zeitpunkten ermittelt. Ändert sich die Orientierung der Rotationsachse, so läßt dies auf eine Überlagerung mehrerer Rotationen schließen. Ändert sie sich nicht, so unterliegt das zugehörige Objektteil genau einer Rotation.⁵ Ein derartiges Objektteil bezeichnen Asada und Tsuji als Hauptkörper.

Asada und Tsuji schlagen unterschiedliche Methoden vor, um für ein starres Teil die physikalischen Rotationsachsen aus den gemessenen scheinbaren Rotationen zu bestimmen. Mit Hilfe dieser Methoden läßt sich der Aufbau bzw. die Struktur von Objekten entnehmen, die aus gekoppelten starren Einzelteilen zusammengesetzt sind.

Heuristische Methode

1. Bestimme die scheinbaren Rotationsachsen aller in der Szene befindlichen starren Objektteile.
2. Suche alle Teile, die um Achsen mit konstanter Orientierung rotieren und betrachte sie als Kandidaten für den Hauptkörper einer gekoppelten Bewegung.

⁵Asada und Tsuji setzen voraus, daß die physikalischen Rotationsachsen sich in ihrer Orientierung unterscheiden. Eine Überlagerung führt damit automatisch zu einer Art Kreisbewegung (Präzessionsbewegung).

3. Für jedes Objektteil mit sich ändernder scheinbarer Rotationsachse beschreibe die Bewegung relativ zu einem bereits gefundenen Kandidaten. Bleibt bei dieser Bewegung die Orientierung der Rotationsachse zeitlich konstant, so läßt sich die Bewegung des Objektteiles interpretieren als Rotation um eine zum gewählten Kandidaten feste Rotationsachse.

Damit ist der Kandidat als Hauptkörper einer gekoppelten Bewegung bestätigt.

Asada und Tsuji verwenden diese Methode, um einen oder mehrere Hauptkörper ausfindig zu machen. Dabei betrachten sie nur Objekte, deren Teile um höchstens zwei physikalische Rotationsachsen rotieren. Ein Objektteil ist damit entweder ein Hauptkörper oder aber ein Teil, welches um eine konstante Rotationsachse relativ zum Hauptkörper rotiert. Die Methode liefert die Zusammensetzung derartiger Objekte. Sie funktioniert jedoch nur dann erfolgreich, wenn der zugehörige Hauptkörper auch wirklich detektiert wurde.

Asada und Tsuji schlagen daher eine weitere Methode vor, bei der diese Forderung nicht gestellt zu werden braucht. Weiterhin berücksichtigen sie mehr als zwei physikalische Rotationsachsen.

Geometrisches Rotationsmodell

Asada + Tsuji 83 repräsentieren die Bewegung eines starren Körpers vom Zeitpunkt t_k zum Zeitpunkt t_{k+1} durch

$$\mathbf{p}(t_{k+1}) = \mathbf{R}\mathbf{p}(t_k) + \mathbf{t}$$

wobei $\mathbf{p}(t_k)$, $\mathbf{p}(t_{k+1})$ ein beliebiger Punkt \mathbf{p} des starren Körpers zum Zeitpunkt t_k , t_{k+1} ist, \mathbf{R} eine *eigentlich orthogonale* Rotationsmatrix ist und \mathbf{t} den Translationsvektor darstellt. Bei der Zerlegung der scheinbaren (gemessenen) Rotationen in physikalische Rotationen braucht man die Translationsbewegung nicht zu berücksichtigen, da sie keinen Einfluß auf die Orientierung der Rotationsachsen hat.

Die Überlagerung von physikalischen Rotationen läßt sich durch das Produkt der entsprechenden Rotationsmatrizen beschreiben. Die resultierende Rotationsmatrix stellt die gemessene scheinbare Rotation dar. Legen wir die Anzahl der physikalischen Rotationsachsen fest, so können wir für jede gemessene scheinbare Rotation ein derartiges Gleichungssystem aufstellen. Ob sich aus diesen Gleichungssystemen die unbekanntes Größen der physikalischen Rotationen bestimmen lassen, hängt davon ab, ob die Anzahl der bekannten unabhängigen Größen mindestens so groß ist wie die Anzahl der gesuchten Unbekannten. Asada und Tsuji stellen die bekannten und unbekanntes Größen folgendermaßen gegenüber:

Eine Rotation kann durch drei Größen beschrieben werden (siehe Abbildung 4.1 auf Seite 46). Für n Zeitintervalle erhält man n scheinbare Rotationen, mit insgesamt $3n$ Größen. Vorausgesetzt, der betrachtete starre Körper unterliegt zwei physikalischen Rotationen, so suchen wir $2(2 + n)$ Unbekannte. Dieses sind die Orientierungen der beiden Rotationsachsen und die Rotationswinkel der beiden Achsen in den n Zeiträumen. Damit die Zahl der Unbekannten kleiner gleich der Zahl der bekannten Größen ist, benötigen wir mindestens vier Zeitintervalle, um die beiden Achsen und Winkel bestimmen zu können.

Unterliegt der starre Körper hingegen drei physikalischen Rotationen, so erhalten wir $3(2 + n)$ Unbekannte. Diese lassen sich nicht mehr den $3n$ gemessenen Größen in eindeutiger Weise entnehmen, da die Zahl der Unbekannten stets größer ist, als die Zahl der bekannten Größen. Unterliegt ein starrer Körper also mehr als zwei physikalischen Rotationen, so lassen sich diese nicht mehr eindeutig aus den gemessenen scheinbaren Rotationen bestimmen. Asada + Tsuji 83 machen deshalb die zusätzliche Einschränkung, daß sich das Objekt mit einer konstanten Winkelgeschwindigkeit um die physikalischen Rotationsachsen bewegt. Damit ist der Rotationswinkel um jede Achse pro Zeitintervall konstant, und es verbleiben bei m physikalischen Rotationsachsen genau $3m$ Unbekannte. Es genügen damit m Zeitintervalle bzw. $m + 1$ Zeitpunkte, um die physikalischen Rotationsachsen bestimmen zu können.

Analyse auf der gaußschen Einheitskugel

Szenenobjekte, die aus realen Bildern rekonstruiert wurden, sind erfahrungsgemäß fehlerbehaftet, so daß Szenenpunkte im Raum falsch lokalisiert sind. Bestimmen wir aus den Szenenpunkten die zugehörigen Bewegungsparameter, so sind diese ebenfalls fehlerbehaftet. Die scheinbaren Rotationen können damit nur ungenügend genau gemessen werden. Aus diesem Grund schlagen Asada und Tsuji eine weitere Zerlegungsmethode vor, bei der sie die scheinbaren Rotationen über einen längeren Zeitraum beobachten. Dieses Vorgehen bewirkt eine Mittelung, die dazu führt, daß der Einfluß der störenden Rauscheffekte reduziert wird.

Die normalisierten Orientierungsvektoren der scheinbaren Rotationen lassen sich auf der gaußschen Einheitskugel darstellen. Asada und Tsuji analysieren die sich ergebenden Muster auf der Kugeloberfläche, um die physikalischen Rotationsachsen zu bestimmen. Unterliegt ein starrer Körper zwei physikalischen Rotationen, die mit konstanter Winkelgeschwindigkeit erfolgen, so beschreiben die sich ergebenden scheinbaren Rotationsachsen auf der gaußschen Einheitskugel einen Kreis. Der Normalenvektor der Kreisebene gibt dann die Orientierung der ersten physikalischen Rotationsachse wieder. Findet man einen Hauptkörper, dessen Rotationsachse mit dieser physikalischen Rotationsachse übereinstimmt, so ist der betrachtete starre Körper mit diesem Hauptkörper verbunden. Asada und Tsuji gehen bei ihrer Analyse daher wie folgt vor:

Für jeden starren Körper betrachten sie die scheinbaren Rotationsachsen auf der gaußschen Einheitskugel. Sind die Rotationsachsen alle identisch, so ergibt sich ein Punkt auf der gaußschen Kugel, und das betreffende Objekt wird von ihnen als Hauptkörper betrachtet. Andernfalls passen sie eine Ebene an die gemessenen Rotationsachsen auf der gaußschen Kugel an. Liegt der Fehler der Ebenenanpassung unterhalb eines von ihnen gewählten Schwellwertes, so vermuten sie, daß das Objekt um zwei physikalische Achsen rotiert. Der Normalenvektor der angepaßten Ebene gibt, wie beschrieben, die erste physikalische Rotationsachse an. Liegt der Fehler der Ebenenanpassung oberhalb der Schwelle, so gehen sie davon aus, daß an der Bewegung mehr als zwei physikalische Achsen beteiligt sind. In diesem Fall beschreiben sie die Rotationsbewegung relativ zu einem anderen Objekt, dessen physikalische Rotationsachsen sie bereits kennen, in der

Hoffnung, daß es ein übergeordnetes Objekt darstellt. Ist dies der Fall, so vereinfacht sich die Bewegungsbeschreibung, und die Anzahl der an der relativen Bewegung beteiligten physikalischen Rotationsachsen nimmt ab. Mit den relativen Rotationsbewegungen wird wieder eine Ebenenanpassung durchgeführt usw.

Als Ergebnis bekommen sie die Rotationshierarchie des gesamten Objektes. Läßt sich an die Rotationsachsen eines starren Körpers keine Ebene anpassen, egal zu welchem anderen Objekt man die Bewegung beschreibt, so interpretieren sie dies Objekt als zufällig rotierend.

Anmerkungen

In beiden zuletzt vorgestellten Zerlegungsmethoden machen Asada und Tsuji die Annahme, daß die Winkelgeschwindigkeiten um die physikalischen Rotationsachsen im gesamten Zeitraum konstant bleiben. Bei ihrem geometrischen Rotationsmodell benötigen sie diese Annahme, weil sie die Bewegung der übergeordneten Objektteile nicht berücksichtigen. Es gelingt ihnen dadurch selbst dann eine Zerlegung, wenn keins der übergeordneten Objektteile detektiert wurde.

Bei der Analyse auf der gaußschen Einheitskugel liegt der Grund für diese Annahme dagegen darin, daß Asada und Tsuji sich auf Muster beschränken, bei denen die scheinbaren Rotationsvektoren auf der gaußschen Einheitskugel auf einer Ebene liegen. Ein derartiges Muster erhält man nur dann, wenn das betreffende Objekt um genau zwei physikalische Rotationsachsen rotiert und die Rotationen mit konstanter Winkelgeschwindigkeit erfolgen. Wäre die Rotation um eine der Achsen beschleunigt, so lägen die scheinbaren Rotationsvektoren auf einer spiralförmigen Bahn. Eine Analyse durch Ebenenanpassung wäre hier vollkommen ungeeignet.

Für ein Objekt, welches um mehr als zwei physikalische Rotationsachsen rotiert, ist eine Zerlegung mit Hilfe der gaußschen Einheitskugel nur dann möglich, wenn ein übergeordnetes Objekt existiert, zu welchem die Bewegung als Rotation um höchstens zwei physikalische Rotationsachsen beschrieben werden kann. Im Gegensatz zum geometrischen Rotationsmodells benötigt man für die Zerlegung also wiederum die Bewegung von übergeordneten Objekten.

Wenn es gelingt, das direkt übergeordnete Objekt auszumachen, so erübrigt sich die Analyse auf der gaußschen Einheitskugel. Beschreibt man nämlich die Bewegung eines Objektes O_1 relativ zu einem Objekt O_2 , so läßt sich unabhängig von den Winkelgeschwindigkeiten um die physikalischen Rotationsachsen sofort erkennen, ob sich O_1 mit einer konstanten Rotationsachse bezüglich O_2 bewegt.

Der Vorteil der Analyse auf der gaußschen Einheitskugel gegenüber einer relativen Bewegungsbeschreibung liegt darin, daß für ein betrachtetes Objekt der direkte Vorgänger in der Rotationshierarchie nicht benötigt wird, um eine Zerlegung der scheinbaren Rotationen vornehmen zu können. Das bedeutet, daß eine gesamte Zerlegung auch dann noch gelingt, wenn jedes zweite Objektteil in der Hierarchie fehlt.

Meiner Ansicht nach rechtfertigt dieser Vorteil allerdings nicht die von Asada und Tsuji gemachte Bewegungseinschränkung, die ich als sehr streng empfinde. Dies wird

noch deutlicher, wenn man bedenkt, daß Asada und Tsuji für ihre Analyse – wie anfangs erwähnt – einen längeren Zeitraum betrachten. In einem längeren Zeitraum sollte man damit rechnen, daß sich die Bewegungsgrößen der Objekte in der Szene ändern, so daß die Annahme einer konstanten Winkelgeschwindigkeit damit nicht mehr gerechtfertigt ist.

Asada und Tsuji räumen ein [Asada + Tsuji 83, Seite 142]:

Our future research will investigate what patterns are shown on the sphere under other assumptions and how the system can analyse these patterns.

Dies ist meines Erachtens nicht der richtige Weg, da die Muster auf der Kugel nicht nur von den Orientierungen der physikalischen Rotationsachsen abhängen, sondern auch davon, wie sich die Winkelgeschwindigkeiten verhalten. Damit sind Zerlegungen nur möglich, wenn einschränkende Annahmen über die Bewegung gemacht werden. Diese Einschränkungen sind allerdings nicht erforderlich, wenn man die Bewegungen relativ zueinander beschreibt, was ja bei der Methode ohnehin schon gemacht wird.

Man sollte daher versuchen, durch relative Bewegungsbeschreibung die direkten Vorgänger in der Rotationshierarchie ausfindig zu machen, um damit die physikalischen Rotationsachsen ermitteln zu können. Kann der direkte Vorgänger für ein gegebenes Objekt nicht ermittelt werden, so könnte man den Ansatz von Asada und Tsuji immer noch zur Unterstützung heranziehen.

5.1.2 Bestimmung der Rotationsachsenlage

Mit Hilfe der Rotationsachsenzerlegung ist es möglich, die Bewegung jedes Objektes so zu beschreiben, daß die Orientierung der Rotationsachse konstant ist. Gegebenenfalls muß dazu die Bewegung relativ zu dem übergeordneten Objekt beschrieben werden (falls es ein solches denn gibt).

In diesem Abschnitt wollen wir nun die Lage dieser Rotationsachse im Raum bestimmen. Wir hatten bereits erwähnt, daß die Lage im Raum nicht eindeutig ist. Die verbleibenden zwei Freiheitsgrade lassen sich ausnutzen, um eine einfache Translationsbewegung zu finden. Je einfacher wir die Bewegung des Objektes beschreiben können, um so allgemeiner ist unser Verständnis von der Bewegung und um so leichter sollten wir die Bewegung des Objektes vorhersagen können. Wie im vorangegangenen Abschnitt werden wir zunächst die Vorgehensweise von Asada und Tsuji beschreiben [Asada + Tsuji 83].

Asada und Tsuji bestimmen die Lage der Rotationsachse für einen kurzen Bewegungszeitraum, indem sie an die Bewegung der Rotationsachse und damit an die Translation zwei zusätzliche Einschränkungen stellen. Dadurch werden die zwei verbleibenden Freiheitsgrade der Rotationsachsenlage aufgehoben und die Rotationsachsenlage läßt sich *eindeutig* bestimmen. Als Einschränkung fordern sie entweder eine konstante Geschwindigkeit oder eine konstante Beschleunigung der Rotationsachse.

Die Lage der Rotationsachse beeinflusst lediglich die Translationskomponente senkrecht zur Rotationsachse. Einschränkungen können folglich nur an diesen Translationsanteil gestellt werden. Wir können daher zunächst eine Koordinatentransformation durchführen, durch die die Rotationsachse mit der neuen z-Achse zusammenfällt. Damit

brauchen wir uns nur noch auf die neue XY-Ebene zu beschränken. Die Translationskomponente in Richtung der z-Achse, und damit in Richtung der Rotationsachse, ist bekanntlich unabhängig von der gewählten räumlichen Lage der Rotationsachse, und braucht daher nicht berücksichtigt zu werden.

Asada und Tsuji beschreiben die Bewegung eines starren Körpers in der neuen XY-Ebene zwischen den Zeitpunkten t_0, t_1, \dots, t_n durch die folgende Transformation:

$$\mathbf{p}(t_{i+1}) = \mathbf{R}(t_i) (\mathbf{p}(t_i) - \mathbf{o}(t_i)) + \mathbf{o}(t_i) + \mathbf{t}(t_i) \quad (5.1)$$

$$\mathbf{o}(t_{i+1}) = \mathbf{o}(t_i) + \mathbf{t}(t_i) \quad i = 0, \dots, n-1 \quad (5.2)$$

wobei

$$\mathbf{R}(t_i) = \begin{pmatrix} \cos(\theta_i) & -\sin(\theta_i) \\ \sin(\theta_i) & \cos(\theta_i) \end{pmatrix}$$

Die Rotationsmatrix $\mathbf{R}(t_i)$ beschreibt die Rotationsbewegung vom Zeitpunkt t_i zum Zeitpunkt t_{i+1} . Entsprechend stellt $\mathbf{t}(t_i)$ den Translationsvektor in diesem Zeitintervall dar, $\mathbf{p}(t_i)$ und $\mathbf{p}(t_{i+1})$ sind die XY-Koordinatenvektoren des Raumpunktes \mathbf{p} zu den angegebenen Zeitpunkten, und $\mathbf{o}(t_i)$ gibt die Lage der Rotationsachse in der XY-Ebene zum Zeitpunkt t_i an.

Wenn man Gleichung 5.2 umschreibt zu

$$\mathbf{o}(t_{i+1}) = \sum_{k=0}^i \mathbf{t}(t_k) + \mathbf{o}(t_0) \quad i = 0, \dots, n-1,$$

so kann man anstelle der Gleichungen 5.1 und 5.2 folgende Gleichung schreiben.

$$\mathbf{p}(t_1) = \mathbf{R}(t_0)\mathbf{p}(t_0) + (\mathbf{I} - \mathbf{R}(t_0))\mathbf{o}(t_0) + \mathbf{t}(t_0)$$

$$\mathbf{p}(t_{i+1}) = \mathbf{R}(t_i)\mathbf{p}(t_i) + (\mathbf{I} - \mathbf{R}(t_i)) \left(\sum_{k=0}^{i-1} \mathbf{t}(t_k) + \mathbf{o}(t_0) \right) + \mathbf{t}(t_i) \quad i = 1, \dots, n-1$$

Man erhält für n Zeitintervalle genau $2n$ Gleichungen mit $2(n+1)$ unbekanntten Größen ($\mathbf{o}(t_0), \mathbf{t}(t_0), \dots, \mathbf{t}(t_{n-1})$).

Damit benötigt man noch zwei weitere Beziehungen, damit sich die Gleichungen eindeutig lösen lassen. Folgende Bewegungseinschränkungen werden von Asada und Tsuji verwendet:

1. Konstante Geschwindigkeit in zwei aufeinanderfolgenden Zeitintervallen

Betrachtet man drei aufeinanderfolgende Aufnahmezeitpunkte t_i, t_{i+1}, t_{i+2} (2 Zeitintervalle), so läßt sich die Rotationsachse so legen, daß die Geschwindigkeit in der XY-Ebene konstant ist. Die Translation pro Zeitintervall ist gleich der Geschwindigkeit. Sind die Zeitpunkte äquidistant und damit die Zeitintervalle gleich groß, so ergeben sich die zusätzlichen Einschränkungen aus:

$$\mathbf{t}(t_i) = \mathbf{t}(t_{i+1})$$

2. Konstante Beschleunigung in drei aufeinanderfolgenden Zeitintervallen

Betrachtet man vier aufeinanderfolgende Aufnahmezeitpunkte $t_i, t_{i+1}, t_{i+2}, t_{i+3}$ (3 Zeitintervalle), so läßt sich die Rotationsachse so legen, daß die Beschleunigung in der XY-Ebene konstant ist. Für gleichgroße Zeitintervalle läßt sich die Beschleunigung durch die Differenz von zwei aufeinanderfolgenden Translationsvektoren beschreiben. Man erhält:

$$\mathbf{t}(t_{i+1}) - \mathbf{t}(t_i) = \mathbf{t}(t_{i+2}) - \mathbf{t}(t_{i+1})$$

3. Konstante Bewegungsrichtung in drei aufeinanderfolgenden Zeitintervallen

Wie im vorangegangenen Beispiel läßt sich für die vier aufeinanderfolgenden Aufnahmezeitpunkte die Bewegung derart einschränken, daß die Richtung der Translationsvektoren in den drei Zeitintervallen konstant bleibt. Das Skalarprodukt der Translationsvektoren ist dann gleich dem Produkt ihrer Beträge.

$$\mathbf{t}(t_i)\mathbf{t}(t_{i+1}) = |\mathbf{t}(t_i)||\mathbf{t}(t_{i+1})| \qquad \mathbf{t}(t_{i+1})\mathbf{t}(t_{i+2}) = |\mathbf{t}(t_{i+1})||\mathbf{t}(t_{i+2})|$$

In allen Fällen werden nur so viele Zeitintervalle für die Berechnung der unbekannt Parameter berücksichtigt, wie für zwei hinzugenommene bewegungseinschränkende Gleichungen erforderlich sind. Damit ergibt sich ein eindeutig lösbares Gleichungssystem.

Der Nachteil dieser Methode ist, daß man nicht weiß, ob die Bewegungsannahme, die in die einschränkenden Gleichungen eingebracht wird, überhaupt gerechtfertigt war. Es ist nämlich vollkommen egal, um was für eine Translationsbewegung es sich gehandelt hat. In jedem Fall ergeben sich mit den obigen Einschränkungen eindeutige Lösungen für die Lage der Rotationsachse und die Translationsvektoren. Wenn nur drei aufeinanderfolgende Zeitpunkte berücksichtigt werden, läßt sich also die Bewegung der Rotationsachse für jede beliebige Bewegung so darstellen, als sei sie mit konstanter Geschwindigkeit auf einer Ebene senkrecht zur Rotationsachse erfolgt. Bei vier aufeinanderfolgenden Zeitpunkten läßt sie sich entsprechend so darstellen, als sei sie mit konstanter Beschleunigung erfolgt usw. Damit weiß man allerdings noch nicht, wie sie nun tatsächlich erfolgt ist.

Asada und Tsuji verfahren daher wie folgt: Zunächst betrachten sie alle Möglichkeiten von drei aufeinanderfolgenden Zeitpunkten und beschreiben in diesen Intervallen die Bewegung der Rotationsachse durch eine konstante Geschwindigkeit. Danach tragen sie die x- und y-Komponenten der ermittelten Geschwindigkeiten in einen zweidimensionalen Merkmalsraum ein.⁶ In diesem Merkmalsraum läßt sich nun prüfen, ob die Annahme der konstanten Geschwindigkeit gerechtfertigt war. Eine Ballung von zeitlich aufeinanderfolgenden Geschwindigkeitsvektoren korrespondiert mit einem Zeitintervall, in dem die Geschwindigkeit konstant bleibt. Isolierte Punkte zwischen Ballungen stellen Zeitpunkte dar, in denen die Geschwindigkeit sich abrupt geändert hat. Können keine Ballungen

⁶Um den zeitlichen Zusammenhang zu dokumentieren, verbinden sie in der graphischen Darstellung ihres Merkmalsraumes zeitlich aufeinanderfolgende Einträge durch eine Kante. Man kann annehmen, daß sie bei ihrer Ballungsanalyse den zeitlichen Zusammenhang zusätzlich berücksichtigt haben.

ausgemacht werden, weil die Punkte im Merkmalsraum verstreut sind, so versuchen sie, die Bewegung durch eine konstante Beschleunigung zu beschreiben. Dazu bilden sie alle Möglichkeiten von vier aufeinanderfolgenden Zeitpunkten und bestimmen in jedem dieser Zeiträume die Lage der Rotationsachse unter der Annahme, daß diese sich mit konstanter Beschleunigung durch den Raum bewegt. Die x - und y -Komponenten der sich ergebenden Beschleunigungen tragen sie in ähnlicher Form in einen Merkmalsraum ein. Dieser wird nach derselben Ballungsanalyse untersucht, um die Zeiträume ausfindig zu machen, in denen sich das betreffende Objekt mit konstanter Beschleunigung bewegt hat.

Die Translationskomponente in Richtung der Rotationsachse, die sich ja eindeutig bestimmen läßt, könnte einen Hinweis auf die Art der Translationsbewegung geben. Bleibt sie über einen längeren Zeitraum konstant, so ist die Annahme einer konstanten Geschwindigkeit naheliegend. Ist sie dagegen gleichmäßig beschleunigt, so sollte die Translationsbewegung durch eine konstante Beschleunigung beschrieben werden. Asada und Tsuji nutzen diese Information nicht aus. Die Bewegung in Rotationsachsenrichtung läßt allerdings auch keine eindeutigen Schlußfolgerungen auf die Bewegung senkrecht zur Rotationsachse zu. Die Annahme ist also mehr als Heuristik zu verstehen und ist immer dann gültig, wenn die Ursache für die Translationsbewegung sich weder ausschließlich in Rotationsachsenrichtung noch ausschließlich senkrecht dazu auswirkt. Ist die Ursache für die Translationsbewegung z. B. eine konstante Kraft, deren Richtung nicht mit der Rotationsachse zusammenfällt und nicht senkrecht dazu verläuft, so ist die Bewegung sowohl in Achsenrichtung als auch in einer Ebene senkrecht dazu gleichmäßig beschleunigt. In der Realität treffen wir allerdings häufig den Fall an, in dem die gemachte Annahme nicht gültig ist. So bewegt sich ein um die Ecke biegendes Auto ausschließlich in einer Ebene senkrecht zur Rotationsachse.

5.1.3 Zusammenfassung

Fassen wir noch einmal die wesentlichen Eigenschaften des Verfahrens von Asada und Tsuji zusammen:

Für die Zerlegung der scheinbaren Rotationen halten Asada und Tsuji sowohl ihr geometrisches Rotationsmodell als auch ihre Analyse auf der gaußschen Einheitskugel für sehr nützlich. Beide Verfahren setzen voraus, daß die Winkelgeschwindigkeiten um die physikalischen Rotationsachsen konstant sind. Die Zerlegung in physikalische Achsen gelingt dafür auch dann, wenn das Objekt, zu dem die Rotationsachse fixiert ist, unsichtbar bzw. unbekannt ist.

Bei der Analyse der gaußschen Einheitskugel wurde nur die Richtung der scheinbaren Rotationsachsen verwendet, die gemessenen Rotationswinkel und damit die Winkelgeschwindigkeiten um die scheinbaren Achsen wurden nicht genutzt. Asada und Tsuji stellen fest, daß sich die physikalischen Achsen durch diese zusätzliche Information genauer bestimmen lassen sollten.

Die Autoren unterteilen die Bildsequenz in kurze Teilfolgen und betrachten in jeder dieser Teilfolgen die Translationsbewegung des Hauptkörpers so, als würde sie mit konstanter Geschwindigkeit oder mit konstanter Beschleunigung erfolgen. Durch diese

Bewegungseinschränkung ergibt sich die räumliche Lage der Rotationsachse. In einer darauffolgenden Ballungsanalyse bestimmen sie die Zeitintervalle, in denen die Beschreibung der Translation mit einer konstanten Geschwindigkeit bzw. einer konstanten Beschleunigung angemessen erscheint.

Asada und Tsuji bestimmen die Lage der Rotationsachse nur für den Hauptkörper. Für Teile, die um mehrere physikalische Rotationsachsen rotieren, ließe sich die Lage der (physikalischen) Rotationsachse bezüglich des übergeordneten Objektes in gleicher Weise ermitteln. Dazu müßte man zunächst die Bewegung des Teiles relativ zu seinem übergeordneten Objekt beschreiben. Asada und Tsuji betrachten diese Möglichkeit nicht, weil ihnen ihre ermittelten Bewegungsparameter des Hauptkörpers zu fehlerbehaftet erscheinen. Diese Fehler würden sich fortpflanzen, wenn man die Bewegung des Objektes relativ zum Hauptkörper beschreibt.

Weiterhin setzen Asada und Tsuji voraus, daß sich die physikalischen Rotationsachsen jeweils in ihrer Orientierung unterscheiden. Die Abbildung 5.4 zeigt dagegen ein aus drei verbundenen starren Einzelteilen zusammengesetztes Objekt, bei der dies nicht erfüllt ist. Die physikalischen Rotationsachsen der Teile 2 und 3 (Verbindungsachsen) haben dieselbe Orientierung. Damit läßt sich Teil 3 sowohl relativ zu Teil 2 als auch relativ zu Teil 1 (Hauptkörper) durch eine Bewegung beschreiben, bei der die Orientierung der Rotationsachse konstant bleibt. In diesem Fall muß die Analyse der Translationsbewegung Aufschluß darüber geben, mit welchem Teil das Teil 3 verbunden ist.

5.2 Eigene Vorgehensweise

Die Idee der Bewegungsbeschreibung von zusammengesetzten Objekten, wie sie im vorhergehenden Abschnitt vorgestellt wurde, werden wir in diesem Abschnitt aufgreifen und dabei eine eigene Vorgehensweise wählen, die sich von der in [Asada + Tsuji 83] vorgestellten Methode in mehreren Belangen unterscheidet.

Ausgehen werden wir von einer Menge von 3D-Szenenpunkten, die uns zu $n + 1$ Zeitpunkten t_0, \dots, t_n bekannt sind. Die räumlichen und zeitlichen Korrespondenzen setzen wir also für das obige Zeitintervall als bekannt voraus. Es bleibt abzuwarten, ob die angestrebte Bewegungsbeschreibung zu Vorhersagen geeignet ist, damit man die Korrespondenzanalyse für die folgenden Zeitpunkte unterstützen kann.

Wir setzen voraus, daß sich die Objekte in der Szene in Teilen starr verhalten. Die Menge der gegebenen Raumpunkte ist nun zunächst in Punktgruppen zu zerlegen, die als starre Objektteile interpretiert werden können (Verfahren aus Abschnitt 4.1). Für jedes der so gefundenen Objektteile bestimmen wir in jedem Zeitintervall die Bewegungsparameter (siehe Abschnitt 4.3). Sie bestehen aus einer Rotation um den gewählten Koordinatensystemursprung und einer anschließenden Translation. Bleibt die Orientierung der Rotationsachse eines Objektes O_1 in den aufeinanderfolgenden Zeitintervallen konstant, so betrachten wir es als einen Hauptkörper. Andernfalls versuchen wir festzustellen, ob O_1 relativ zu einem Objekt O_2 mit konstanter Rotationsachsenorientierung rotiert. Wenn ja, betrachten wir die Objekte O_1 und O_2 als gekoppelt. Alle Objekte,

die entweder um eine Rotationsachse mit zeitlich konstanter Orientierung rotieren oder deren Bewegung relativ zu einem anderen Objekt derart dargestellt werden kann, werden im weiteren einer Translationsanalyse unterzogen. Diese wird die Rotationsachse so legen, daß man die Translationsbewegung durch ein gewähltes Bewegungsmodell beschreiben kann. Die beiden Freiheitsgrade bei der Wahl der Rotationsachsenlage werden ausgenutzt, um dem gewählten Bewegungsmodell möglichst genau zu entsprechen. Da wir auch die relative Bewegung berücksichtigen, die ein Objekt bezüglich eines anderen Objektes ausführt, bestimmen wir im Gegensatz zu Asada und Tsuji die Lagen aller physikalischen Rotationsachsen.

5.2.1 Relative Bewegungsbeschreibung

Um festzustellen, ob zwei Objekte O_1 und O_2 gekoppelt sind, werden wir zunächst die Bewegung des einen relativ zum anderen beschreiben:

Gegeben seien die Objekte O_1, O_2, \dots, O_m . Die Bewegung des Objektes O_i im Zeitintervall (t_j, t_{j+1}) sei in homogenen Koordinaten durch die Matrix $M_{O_i}(t_j)$ beschrieben, die wir durch Lösen der Gleichung 4.8 erhalten. In der folgenden Tabelle sind die Bewegungsmatrizen zusammengefaßt. Sie beziehen sich alle auf dasselbe Koordinatensystem, welches wir im weiteren als Weltkoordinatensystem (WKS) bezeichnen werden. Nehmen

Objekte	Bewegungsmatrizen im WKS			
O_1	$M_{O_1}(t_0)$	$M_{O_1}(t_1)$	\dots	$M_{O_1}(t_{n-1})$
O_2	$M_{O_2}(t_0)$	$M_{O_2}(t_1)$	\dots	$M_{O_2}(t_{n-1})$
\vdots	\vdots	\vdots		\vdots
O_m	$M_{O_m}(t_0)$	$M_{O_m}(t_1)$	\dots	$M_{O_m}(t_{n-1})$

wir an, daß das Objekt O_2 relativ zu Objekt O_1 um eine zu O_1 körperfeste Achse rotiert. Die Bewegung von O_2 im Zeitintervall (t_j, t_{j+1}) bezüglich der räumlichen Position von O_1 zum Zeitpunkt t_0 sei durch die Matrix $M_{O_2, O_1}(t_j)$ beschrieben. Die gemessene Bewegungsmatrix $M_{O_2}(t_j)$ läßt sich mit Hilfe dieser Matrix darstellen, indem man zunächst alle erfolgten Bewegungen des Objektes O_1 rückgängig macht, dann die Bewegung von O_2 bezüglich dieser Ausgangsposition von O_1 durchführt ($M_{O_2, O_1}(t_j)$) und danach O_1 wieder in die neue Position zurückbewegt. Damit lassen sich die Bewegungsmatrizen $M_{O_2}(t_j)$ durch folgende Matrixgleichungen darstellen:

$$\begin{aligned}
 M_{O_2}(t_0) &= M_{O_1}(t_0)M_{O_2, O_1}(t_0) \\
 M_{O_2}(t_1) &= M_{O_1}(t_1)M_{O_1}(t_0)M_{O_2, O_1}(t_1)M_{O_1}^{-1}(t_0) \\
 M_{O_2}(t_2) &= M_{O_1}(t_2)M_{O_1}(t_1)M_{O_1}(t_0)M_{O_2, O_1}(t_2)M_{O_1}^{-1}(t_0)M_{O_1}^{-1}(t_1) \\
 &\vdots
 \end{aligned}$$

Bezeichnet man die Bewegung des Objektes O_1 zwischen den Zeitpunkten t_0 und t_j mit $A_{O_1}(t_j)$, so läßt sich verkürzt schreiben:

$$M_{O_2}(t_j) = M_{O_1}(t_j)A_{O_1}(t_j)M_{O_2, O_1}(t_j)A_{O_1}^{-1}(t_j)$$

Der Bewegungsanteil des Objektes O_2 relativ zu O_1 wird also durch eine Ähnlichkeits-
transformation dargestellt, durch die wir die Bewegung $M_{O_2, O_1}(t_j)$ von O_2 , bezüglich der
räumlichen Lage von O_1 zum Zeitpunkt t_0 , auf die räumliche Lage von O_1 zum Zeitpunkt
 t_j beziehen. Aus diesen Matrixgleichungen lassen sich die gesuchten Bewegungsmatrizen
wie folgt bestimmen:

$$\begin{aligned} M_{O_2, O_1}(t_0) &= M_{O_1}^{-1}(t_0)M_{O_2}(t_0) \\ M_{O_2, O_1}(t_1) &= M_{O_1}^{-1}(t_0)M_{O_1}^{-1}(t_1)M_{O_2}(t_1)M_{O_1}(t_0) \\ M_{O_2, O_1}(t_2) &= M_{O_1}^{-1}(t_0)M_{O_1}^{-1}(t_1)M_{O_1}^{-1}(t_2)M_{O_2}(t_2)M_{O_1}(t_1)M_{O_1}(t_0) \\ &\vdots \end{aligned}$$

oder allgemeiner

$$\begin{aligned} M_{O_k, O_l}(t_j) &= \mathbf{A}_{O_l}^{-1}(t_j)M_{O_l}^{-1}(t_j)M_{O_k}(t_j)\mathbf{A}_{O_l}(t_j) \\ &= \mathbf{A}_{O_l}^{-1}(t_{j+1})M_{O_k}(t_j)\mathbf{A}_{O_l}(t_j) \end{aligned} \quad (5.3)$$

Diese Gleichung gibt an, wie sich die Bewegung des Objektes O_k im Zeitintervall (t_j, t_{j+1})
bezüglich des WKS ($M_{O_k}(t_j)$) relativ zum Objekt O_l zum Zeitpunkt t_0 ($M_{O_k, O_l}(t_j)$)
beschreiben läßt.

Eine interessante Beziehung, die sich aus der Gleichung 5.3 ergibt, ist die folgende:

$$\begin{aligned} \mathbf{A}_{O_l}(t_j)M_{O_k, O_l}^{-1}(t_j)\mathbf{A}_{O_l}^{-1}(t_j) &= \mathbf{A}_{O_k}(t_j)M_{O_l, O_k}(t_j)\mathbf{A}_{O_k}^{-1}(t_j) \\ &\Rightarrow M_{O_k, O_l}^{-1}(t_j) \neq M_{O_l, O_k}(t_j) \end{aligned}$$

Möchte man also die Bewegung des Objektes O_l relativ zu O_k bestimmen, so genügt es
nicht, das Inverse der Matrix $M_{O_k, O_l}(t_j)$ zu verwenden!

Die beschriebenen Matrixgleichungen sollen im folgenden dazu verwendet werden, die
Bewegung der Objektteile auf eine einfache Art zu beschreiben. Diese Beschreibung wird
Hinweise über die Kopplungen der Objektteile liefern.

Wenn das Objekt O_k relativ zu O_l um eine konstante Rotationsachse rotiert, dann
müssen die Rotationsachsen, die sich aus den Bewegungsmatrizen $M_{O_k, O_l}(t_j)$ bestimmen
lassen, alle dieselbe Orientierung aufweisen. Nur in diesem Fall scheint eine relative
Bewegungsbeschreibung sinnvoll.

5.2.2 Bestimmung der Rotationsachsenlage

Haben wir für ein Objekt O_2 eine Bewegungsbeschreibung gefunden, bei der die Rotati-
onsachse ihre Orientierung beibehält, so versuchen wir nun, die Lage der Rotationsachse
bezüglich des gegebenen Koordinatensystems zu ermitteln. Ist O_2 ein Hauptkörper, so
verwenden wir das WKS als Koordinatensystem. Andernfalls haben wir die Bewegung
relativ zu einem Objekt O_1 beschrieben, welches O_2 direkt übergeordnet ist. Als Koor-
dinatensystem wird dann das Objektkoordinatensystem von O_1 verwendet.

Um eine physikalisch geeignete Rotationsachse zu finden, soll ein Bewegungsmodell für die Translation zugrundegelegt werden, durch das die gegebenen zwei Freiheitsgrade bei der Wahl der räumlichen Lage der Rotationsachse eingeschränkt werden. Im Gegensatz zu Asada und Tsuji werden wir allerdings ein Gleichungssystem aufstellen, welches stark überbestimmt ist. Durch ein "least-square-fit"-Verfahren ergeben sich dann sowohl die Lage der Rotationsachse als auch die Translationsparameter für das vorausgesetzte Bewegungsmodell. An dem Fehler können wir hinterher ablesen, wie gut sich die Bewegung durch das vorausgesetzte Bewegungsmodell beschreiben läßt.

Als Bewegungsmodell verwenden wir zunächst, ähnlich wie Asada und Tsuji, eine gleichmäßig beschleunigte Bewegung. Um die Bewegung der Rotationsachse zu beschreiben, betrachten wir einen beliebigen Punkt auf der Rotationsachse. Für die Bewegung dieses Punktes erhält man das folgende Weg-Zeit-Gesetz (vergleiche auch Gleichung 3.4):

$$\mathbf{s}(t_j) = \frac{1}{2} \mathbf{a}(t_j - t_0)^2 + \mathbf{v}(t_j - t_0) + \mathbf{r}_0 \quad (5.4)$$

$\mathbf{s}(t_j)$ repräsentiert einen Punkt auf der Rotationsachse zum Zeitpunkt t_j , \mathbf{a} stellt wieder die Beschleunigung und \mathbf{v} die Geschwindigkeit dar, mit der sich der Punkt und damit die Rotationsachse durch den Raum bewegt.⁷ Wenn wir im weiteren von der Lage der Rotationsachse sprechen, so meinen wir die räumliche Lage dieses Punktes.

Im Gegensatz zu Asada und Tsuji berücksichtigt diese Gleichung nicht nur konstante Beschleunigung sondern auch den Fall der konstanten Geschwindigkeit ($\mathbf{a} = 0$). Asada und Tsuji hatten beides getrennt behandelt. Darüberhinaus ist die Bewegung nicht auf eine Ebene senkrecht zur Rotationsachse beschränkt.

Von den gegebenen Punkten unseres starren Körpers betrachten wir zunächst einen beliebig gewählten Punkt \mathbf{p} . Unter Verwendung der folgenden Schreibweise

$$\prod_{k=j-1}^0 \mathbf{R}(t_k) = \mathbf{R}(t_{j-1})\mathbf{R}(t_{j-2}) \cdots \mathbf{R}(t_0)$$

läßt sich die Bewegung dieses Punktes nun durch folgende Gleichung darstellen:

$$\begin{aligned} \mathbf{p}(t_j) &= \left(\prod_{k=j-1}^0 \mathbf{R}(t_k) \right) (\mathbf{p}(t_0) - \mathbf{s}(t_0)) + \mathbf{s}(t_j) \\ &= \left(\prod_{k=j-1}^0 \mathbf{R}(t_k) \right) (\mathbf{p}(t_0) - \mathbf{r}_0) + \frac{1}{2} \mathbf{a}(t_j - t_0)^2 + \mathbf{v}(t_j - t_0) + \mathbf{r}_0 \end{aligned} \quad (5.5)$$

Ausgehend vom Zeitpunkt t_0 wird der Ursprung des Koordinatensystems zunächst in den Rotationsachsenpunkt $\mathbf{s}(t_0) = \mathbf{r}_0$ gelegt. In dem neuen Koordinatensystem erfolgen dann die Rotationen des Zeitintervalls (t_0, t_j) . Danach wird die Lage der Rotationsachse zum

⁷Der Vektor \mathbf{r}_0 ist der betrachtete Punkt auf der Rotationsachse zum Zeitpunkt t_0 .

Zeitpunkt t_j addiert.⁸ Beschreiben wir die Bewegung des Punktes zwischen zwei aufeinanderfolgenden Zeitpunkten t_{j-1} und t_j , so entfällt das Produkt der Rotationsmatrizen:

$$\begin{aligned} \mathbf{p}(t_j) &= \mathbf{R}(t_{j-1})(\mathbf{p}(t_{j-1}) - \mathbf{s}(t_{j-1})) + \mathbf{s}(t_j) \\ &= \mathbf{R}(t_{j-1}) \left(\mathbf{p}(t_{j-1}) - \left(\frac{1}{2} \mathbf{a}(t_{j-1} - t_0)^2 + \mathbf{v}(t_{j-1} - t_0) + \mathbf{r}_0 \right) \right) \\ &\quad + \frac{1}{2} \mathbf{a}(t_j - t_0)^2 + \mathbf{v}(t_j - t_0) + \mathbf{r}_0 \end{aligned} \quad (5.6)$$

Weiterhin berücksichtigt diese Gleichung nur zeitlich aufeinanderfolgende räumliche Punkte. Damit enthält die Gleichung nur solche Punkte, die auch bei der Berechnung der Bewegungsparameter, in diesem Fall der Rotationsmatrix $\mathbf{R}(t_{j-1})$ berücksichtigt wurden. Wie in Gleichung 5.5 wird auch in dieser Gleichung als erstes die Lage der Rotationsachse zum Zeitpunkt t_{j-1} subtrahiert. Als nächstes erfolgt die Rotation im Zeitintervall (t_{j-1}, t_j) . Zum Abschluß wird die Rotationsachsenlage zum Zeitpunkt t_j addiert. Damit wird zum einen der Ursprung des Koordinatensystems wieder in die Ausgangslage gebracht und zum anderen die Translation im Zeitintervall (t_{j-1}, t_j) vorgenommen.

Die Gleichung 5.5 enthält die Vektoren \mathbf{r}_0 , \mathbf{v} und \mathbf{a} als unbekannte Größen. Für $n + 1$ Zeitpunkte t_0, \dots, t_n ergeben sich für den betrachteten Punkt \mathbf{p} genau n lineare Gleichungssysteme. Wir erhalten somit pro Punkt $3n$ Gleichungen mit 3×3 Unbekannten. Wir hatten bereits erwähnt, daß wir im Gegensatz zu Asada und Tsuji das Problem überbestimmen wollen ($n \gg 3$). Die unbekanntenen Größen können damit wiederum durch ein "least-square-fit"-Verfahren ermittelt werden. Bei diesem Verfahren werden wir, wie schon bei der Berechnung der Bewegungsparameter alle ermittelten Punkte unseres starren Körpers berücksichtigen. Würden wir nur einen Punkt unseres starren Körpers verwenden, so hingen die Ergebnisse der Translationsanalyse von den Fehlern des gewählten Punktes ab. Dieses betrifft sowohl die Rekonstruktionsfehler des Punktes als auch die Fehler, die die ermittelten Bewegungsparameter speziell für diesen Punkt ergeben haben. Besteht unser starrer Körper aus den Punkten $\mathbf{p}_1, \dots, \mathbf{p}_m$, so lautet die Fehlerfunktion $\mathbf{e}_i(t_j)$, die die Abweichung der Bewegung des Punktes \mathbf{p}_i im Zeitintervall (t_{j-1}, t_j) vom Bewegungsmodell darstellt:

$$\begin{aligned} \mathbf{e}_i(t_j) &= \mathbf{p}_i(t_j) - (\mathbf{R}(t_{j-1})(\mathbf{p}_i(t_{j-1}) - \mathbf{s}(t_{j-1})) + \mathbf{s}(t_j)) \\ &= \mathbf{p}_i(t_j) - \left(\mathbf{R}(t_{j-1}) \left(\mathbf{p}_i(t_{j-1}) - \left(\frac{1}{2} \mathbf{a}(t_{j-1} - t_0)^2 + \mathbf{v}(t_{j-1} - t_0) + \mathbf{r}_0 \right) \right) \right. \\ &\quad \left. + \frac{1}{2} \mathbf{a}(t_j - t_0)^2 + \mathbf{v}(t_j - t_0) + \mathbf{r}_0 \right) \\ &= \mathbf{p}_i(t_j) - \mathbf{R}(t_{j-1})\mathbf{p}_i(t_{j-1}) + \frac{1}{2}(\mathbf{R}(t_{j-1})(t_{j-1} - t_0)^2 - \mathbf{I}(t_j - t_0)^2)\mathbf{a} \\ &\quad + (\mathbf{R}(t_{j-1})(t_{j-1} - t_0) - \mathbf{I}(t_j - t_0))\mathbf{v} \\ &\quad + (\mathbf{R}(t_{j-1}) \quad \quad \quad - \mathbf{I} \quad \quad \quad)\mathbf{r}_0 \end{aligned} \quad (5.7)$$

⁸Ein Punkt auf der Rotationsachse (z. B. \mathbf{r}_0) wird durch die Rotation nicht verändert und unterliegt damit einer reinen Translationsbewegung!

$$= -\mathbf{d}_i(j) + \frac{1}{2}\mathbf{C}_1(j)\mathbf{a} + \mathbf{C}_2(j)\mathbf{v} + \mathbf{C}_3(j)\mathbf{r}_0$$

Zur kürzeren Schreibweise wurden die folgenden Substitutionen vorgenommen:

$$\begin{aligned}\mathbf{d}_i(j) &= \mathbf{R}(t_{j-1})\mathbf{p}_i(t_{j-1}) - \mathbf{p}_i(t_j) \\ \mathbf{C}_1(j) &= \frac{1}{2}(\mathbf{R}(t_{j-1})(t_{j-1} - t_0)^2 - \mathbf{I}(t_j - t_0)^2)\mathbf{a} \\ \mathbf{C}_2(j) &= (\mathbf{R}(t_{j-1})(t_{j-1} - t_0) - \mathbf{I}(t_j - t_0))\mathbf{v} \\ \mathbf{C}_3(j) &= (\mathbf{R}(t_{j-1}) - \mathbf{I})\mathbf{r}_0\end{aligned}$$

Die unbekannt Vektoren sind nun so zu bestimmen, daß die Summe der Fehlerquadrate E minimiert wird.

$$E = \sum_{i=1}^m \sum_{j=1}^n \mathbf{e}_i(t_j)^2 = \sum_{i=1}^m \sum_{j=1}^n \mathbf{e}_i(t_j)^T \mathbf{e}_i(t_j) \quad (5.8)$$

Um das Minimum dieser Fehlerfunktion zu finden, leiten wir nach den unbekannt Vektoren ab und setzen die Ableitungen gleich null.

$$\begin{aligned}\frac{dE}{d\mathbf{a}} &= \sum_{i=1}^m \sum_{j=1}^n \mathbf{C}_1^T(j)\mathbf{e}_i(t_j) = \mathbf{0} \\ \frac{dE}{d\mathbf{v}} &= 2 \sum_{i=1}^m \sum_{j=1}^n \mathbf{C}_2^T(j)\mathbf{e}_i(t_j) = \mathbf{0} \\ \frac{dE}{d\mathbf{r}_0} &= 2 \sum_{i=1}^m \sum_{j=1}^n \mathbf{C}_3^T(j)\mathbf{e}_i(t_j) = \mathbf{0}\end{aligned}$$

Die verwendete Ableitungsregel findet sich im Anhang A in Gleichung A.1 wieder. Aus diesen 3 x 3 Gleichungen erhält man ein lineares Gleichungssystem mit neun Gleichungen und neun Unbekannt:

$$\begin{aligned}\frac{1}{2} \sum_{j=1}^n \mathbf{C}_1^T(j)\mathbf{C}_1(j)\mathbf{a} + \sum_{j=1}^n \mathbf{C}_1^T(j)\mathbf{C}_2(j)\mathbf{v} + \sum_{j=1}^n \mathbf{C}_1^T(j)\mathbf{C}_3(j)\mathbf{r}_0 &= \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n \mathbf{C}_1^T(j)\mathbf{d}_i(j) \\ \frac{1}{2} \sum_{j=1}^n \mathbf{C}_2^T(j)\mathbf{C}_1(j)\mathbf{a} + \sum_{j=1}^n \mathbf{C}_2^T(j)\mathbf{C}_2(j)\mathbf{v} + \sum_{j=1}^n \mathbf{C}_2^T(j)\mathbf{C}_3(j)\mathbf{r}_0 &= \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n \mathbf{C}_2^T(j)\mathbf{d}_i(j) \quad (5.9) \\ \frac{1}{2} \sum_{j=1}^n \mathbf{C}_3^T(j)\mathbf{C}_1(j)\mathbf{a} + \sum_{j=1}^n \mathbf{C}_3^T(j)\mathbf{C}_2(j)\mathbf{v} + \sum_{j=1}^n \mathbf{C}_3^T(j)\mathbf{C}_3(j)\mathbf{r}_0 &= \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n \mathbf{C}_3^T(j)\mathbf{d}_i(j)\end{aligned}$$

Dieses Gleichungssystem läßt sich nach den unbekannt Vektoren auflösen.

Bei der Berechnung der Summen können einige Vereinfachungen vorgenommen werden, wenn die Winkelgeschwindigkeit konstant bleibt. Dann nämlich sind die Rotationsmatrizen alle gleich,⁹ und anstelle der Summen läßt sich ein arithmetischer Ausdruck angeben.

⁹Die Orientierung der Rotationsachse ist nach Voraussetzung konstant!

Bevor wir dieses Gleichungssystem unter bestimmten Bewegungen näher betrachten werden, sei zunächst etwas über die Eindeutigkeit einer Lösung gesagt. Wir hatten bereits gesehen, daß die Lage der Rotationsachse im Raum lediglich zwei Freiheitsgraden unterliegt. Der Vektor \mathbf{r}_0 ist allerdings als dreidimensionaler Vektor in der Bewegungsgleichung berücksichtigt. Angenommen die Rotationsachse weist in Richtung der z -Achse,¹⁰ dann müßte die z -Komponente von \mathbf{r}_0 durch das Gleichungssystem unbestimmt sein. Dies ist auch der Fall, wie man folgendermaßen sieht:

Die Rotationsmatrizen $\mathbf{R}(t_j)$ haben in unserem Fall die Form:

$$\mathbf{R}(t_j) = \begin{pmatrix} \cos(\theta_j) & \sin(\theta_j) & 0 \\ -\sin(\theta_j) & \cos(\theta_j) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Die dritte Zeile und dritte Spalte der Matrix \mathbf{C}_3 enthalten damit nur Nullen. Dadurch enthalten die neunte Zeile und die neunte Spalte der Koeffizientenmatrix unseres Gleichungssystems 5.9 ebenfalls nur Nullen. Das Gleichungssystem ist also unterbestimmt, denn die z -Komponente von \mathbf{r}_0 kann nunmehr beliebig gewählt werden. Die Koeffizientenmatrix läßt sich folglich nicht mehr invertieren. In dem hier konstruierten Fall ist allerdings auch sofort klar, wie man diesem Problem begegnen muß. Es genügt, die neunte Zeile und die neunte Spalte der Koeffizientenmatrix zu streichen und das verbleibende Gleichungssystem bestehend aus acht Gleichungen und acht Unbekannten zu lösen. Das Gleichungssystem 5.9 bleibt also unverändert bestehen, wenn man die dritte Spalte der Matrix \mathbf{C}_3 bzw. die dritte Zeile der Matrix \mathbf{C}_3^T streicht und \mathbf{r}_0 ein Vektor mit zwei statt mit drei Komponenten ist. Für die Rotationsachsenlage werden jetzt nur noch die beiden Koordinaten auf einer Ebene senkrecht zur Rotationsachse berücksichtigt. Die zu Beginn vorgenommene Koordinatentransformation, durch die wir die Rotationsachse in Richtung der z -Achse unseres Koordinatensystems gelegt haben, muß mit den Lösungsvektoren des Gleichungssystems rückgängig gemacht werden. Dazu setzt man zuvor die z -Komponente des Vektors \mathbf{r}_0 auf einen beliebigen Wert, zum Beispiel auf null.

Um zu sehen, wie sich das Gleichungssystem 5.9 verhält, betrachten wir es unter einigen Spezialfällen der Bewegung:

Die Bewegung ist eine reine Translation:

In diesem Fall sind die Rotationsmatrizen alle gleich der Einheitsmatrix:

$$\begin{aligned} \mathbf{R}(t_{j-1}) &= \mathbf{I} \\ \Rightarrow \mathbf{C}_3(j) &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad j = 1, \dots, n \end{aligned}$$

¹⁰Dieses kann durch eine Koordinatentransformation erreicht werden.

Das Gleichungssystem enthält nunmehr drei Nullzeilen und drei Nullspalten. Der Vektor \mathbf{r}_0 ist dadurch völlig unbestimmt.¹¹ Nach streichen dieser Zeilen und Spalten ergibt sich:

$$\begin{aligned} \frac{1}{2} \sum_{j=1}^n \mathbf{C}_1^T(j) \mathbf{C}_1(j) \mathbf{a} + \sum_{j=1}^n \mathbf{C}_1^T(j) \mathbf{C}_2(j) \mathbf{v} &= \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n \mathbf{C}_1^T(j) (\mathbf{p}_i(t_{j-1}) - \mathbf{p}_i(t_j)) \\ \frac{1}{2} \sum_{j=1}^n \mathbf{C}_2^T(j) \mathbf{C}_1(j) \mathbf{a} + \sum_{j=1}^n \mathbf{C}_2^T(j) \mathbf{C}_2(j) \mathbf{v} &= \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n \mathbf{C}_2^T(j) (\mathbf{p}_i(t_{j-1}) - \mathbf{p}_i(t_j)) \end{aligned}$$

Aus diesem Gleichungssystem ergibt sich die Geschwindigkeit und die Beschleunigung, die jetzt nicht mehr die Bewegung der Rotationsachse sondern die Bewegung des starren Körpers beschreibt. Das zugehörige Weg-Zeit-Gesetz für einen beliebigen Punkt des starren Körpers lautet:

$$\mathbf{p}(t) = \frac{1}{2} \mathbf{a}(t - t_0)^2 + \mathbf{v}(t - t_0) + \mathbf{p}(t_0)$$

Probleme können extrem kleine Rotationswinkel bereiten. Die letzten beiden Zeilen und Spalten unseres 8×8 Gleichungssystems enthalten in diesem Fall nicht null, aber sehr kleine Werte. Der Grenzfall von extrem kleinen Rotationswinkeln bedeutet, daß wir eine Translation durch eine Rotation um eine unendlich weit entfernte Achse ($\mathbf{r}_0 \rightarrow \infty$) um einen verschwindend kleinen Winkel darstellen können. Die numerische Stabilität des Lösungsverfahrens ist bei derart konditionierten Gleichungssystemen nicht mehr gewährleistet, so daß man den Lösungsvektoren keine große Bedeutung beimessen darf ($\mathbf{v} \approx 0$, $\mathbf{a} \approx 0$). Da sich der Rotationswinkel jedoch den ermittelten Bewegungsparametern eindeutig entnehmen läßt, kann dieser Fall vorher erkannt und abgefangen werden.

Die Bewegung ist eine reine Rotation:

In diesem Fall degeneriert die Gleichung 5.6 für einen Punkt $\mathbf{p}_i(t_j)$ zu:

$$\begin{aligned} \mathbf{p}_i(t_j) &= \mathbf{R}(t_{j-1}) (\mathbf{p}_i(t_{j-1}) - \mathbf{r}_0) + \mathbf{r}_0 \\ \Rightarrow \mathbf{e}_i(t_j) &= \mathbf{p}_i(t_j) - \mathbf{R}(t_{j-1}) \mathbf{p}_i(t_{j-1}) + (\mathbf{R}(t_{j-1}) - \mathbf{I}) \mathbf{r}_0 \\ \Rightarrow \mathbf{e}_i(t_j) &= \mathbf{p}_i(t_j) - \mathbf{R}(t_{j-1}) \mathbf{p}_i(t_{j-1}) + \mathbf{C}_3(j) \mathbf{r}_0 \end{aligned}$$

Vergleichen wir diesen Fehler mit der Fehlerfunktion in Gleichung 5.7, so ergibt sich die Geschwindigkeit und die Beschleunigung eindeutig zu null:

$$\mathbf{v} = \mathbf{0} \quad \text{und} \quad \mathbf{a} = \mathbf{0}$$

5.2.3 Zusammenfassung

Der von mir vorgeschlagene mathematische Ansatz versucht, die Lage der Rotationsachse so zu bestimmen, daß die resultierende Objektbewegung einem Bewegungsmodell genügt.

¹¹verständlich, wenn keine Rotation stattgefunden hat!

Die Variation der Winkelgeschwindigkeit ω des betrachteten Objektes ist für den Ansatz unbedeutend, wenn man von dem Spezialfall $\omega = 0$ absieht. Die Winkelgeschwindigkeit muß also, im Gegensatz zu dem Verfahren von Asada und Tsuji, keinesfalls konstant sein.

Das zu lösende Problem ist stark überbestimmt, da wir deutlich mehr Zeitintervalle berücksichtigen als dies für die Einschränkung der zwei verbleibenden Freiheitsgrade notwendig wäre. Zur Lösung wird ein "least-square-fit"-Verfahren verwendet. Der resultierende Fehler gibt uns einen Hinweis darauf, wie gut die gemessene Bewegung durch das angenommene Bewegungsmodell erklärt werden kann. Als Bewegungsmodell wurde lediglich das Weg-Zeit-Gesetz einer gleichmäßig beschleunigten Bewegung berücksichtigt. Es läßt sich als algebraisches Polynom zweiten Grades der Zeit darstellen und enthält in unserem Fall genau acht unabhängige Parameter.¹² Das verwendete Bewegungsmodell schließt folgende Fälle ein:

- Keine Translation, das heißt eine fixierte Achse ($\mathbf{v}=\mathbf{0}$, $\mathbf{a}=\mathbf{0}$).
- geradlinige Bewegung mit
 - konstanter Geschwindigkeit ($\mathbf{a}=\mathbf{0}$)
 - konstanter Beschleunigung ($\mathbf{v}=\mathbf{0}$ oder \mathbf{v} und \mathbf{a} haben dieselbe Richtung, so daß $\mathbf{v} \cdot \mathbf{a} = |\mathbf{v}||\mathbf{a}|$)
- allgemeine gleichmäßig beschleunigte Bewegung (Trajektorie beschreibt eine Parabel im Raum)

Eine andere Bewegungsgleichung wäre zum Beispiel die einer harmonischen Schwingung (Federpendel). In diesem Fall erhielten wir eine Bewegungsgleichung der Form:

$$\mathbf{s}(t) = \boldsymbol{\alpha} \sin(2\pi\nu t + \varphi) + \mathbf{r}_0$$

Der Vektor $\boldsymbol{\alpha}$ stellt die Amplitude, ν die Frequenz und φ die Phasenverschiebung der Schwingung dar. Verwendet man dieses Bewegungsmodell anstelle der Gleichung 5.4, so müssen wir die sich ergebende Fehlerfunktion nach $\boldsymbol{\alpha}$, ν , φ und \mathbf{r}_0 ableiten. Es ergibt sich ein Gleichungssystem mit sieben Unbekannten.

5.3 Implementation und Probleme bei realen Daten

In diesem Abschnitt werden nun alle Schritte beschrieben, die durchgeführt wurden, um aus gegebenen 3D-Szenenpunkten eine Bewegungsbeschreibung von Objekten abzuleiten, die aus beweglich zusammengesetzten starren Einzelteilen bestehen. Der im weiteren beschriebene Algorithmus wurde von mir auf einem ATARI 1040 ST in der Programmiersprache PASCAL implementiert. Zum Testen des Algorithmus wurden künstliche 3D-Objekte einer vorgegebenen Bewegung ausgesetzt und die 3D-Koordinaten von fest gewählten Objektpunkten zu aufeinanderfolgenden Zeitpunkten erzeugt. Diese exakten

¹²Je drei Parameter für die Vektoren \mathbf{a} und \mathbf{v} sowie zwei Parameter für \mathbf{r}_0

3D-Koordinaten dienen dem im weiteren beschriebenen Algorithmus als Eingabe (siehe Kapitel 6).

Verwendet man Szenenpunkte, die aus realen Bildern rekonstruiert wurden, so wird man mit fehlerbehafteten Punktkoordinaten rechnen müssen. Wir werden derartige Fehler berücksichtigen und eine Abschätzung über die Art der tolerierbaren Fehler geben.

Die Bewegungsbeschreibung kann grob in zwei Phasen eingeteilt werden. In der ersten Phase wird versucht, für jedes Zeitintervall Bewegungsparameter für starre Teile zu berechnen. Während in der zweiten Phase versucht wird, für jedes starre Teil aus den Bewegungsparametern der Zeitintervalle eine geschlossene Bewegungsbeschreibung für den gesamten Zeitraum abzuleiten. Diese Beschreibung bringt die Abhängigkeiten in den Bewegungen der verbundenen starren Teile zum Ausdruck.

1. Phase

Da wir Bewegungsparameter von starren Teilen bestimmen wollen, müssen wir zunächst unsere gegebenen 3D-Punkte in Punktgruppen zerlegen, die sich starr verhalten. Da wir bei der Rekonstruktion von Szenenpunkten aus realen Bildern mit fehlerbehafteten Punktlokalisationen im Raum rechnen müssen, betrachten wir bei dem in Abschnitt 4.1 vorgestellten Verfahren eine Punktgruppe auch dann noch als einen starren Körper, wenn die Abstände zwischen Punktpaaren geringen Schwankungen unterliegen. Dazu wird für alle Punktpaare der mittlere Abstand und die Standardabweichung von diesem Mittelwert berechnet. Das Verhältnis aus Standardabweichung und mittlerem Abstand muß unter einem gewählten Schwellwert s liegen, damit die Punkte als zueinander starr betrachtet werden. Die starren Körper ergeben sich aus den maximalen Cliques im Starrheitsgraphen.

Für diese als starr betrachtete Punktgruppe bestimmen wir als nächstes für jedes Zeitintervall die Bewegungsparameter, bestehend aus Rotation und Translation nach dem im Abschnitt 4.3 beschriebenen Verfahren. Dabei werden bei der Berechnung Fehler auftreten, durch die die Bewegungsparameter die exakte Überführung der 3D-Punkte nicht mehr beschreiben können und/oder die Bewegungsparameter nicht mehr die Bewegung eines starren Körpers darstellen.

Die Anwendung der relativen Bewegungsbeschreibung und der Translationsanalyse ist nur in solchen Fällen sinnvoll, in denen die Bewegungsparameter zumindestens annähernd die Bewegung eines starren Körpers wiedergeben. Daher müssen wir uns zunächst fragen, wie groß die Abweichungen von einer starren Konfiguration sein dürfen, damit die ermittelten Bewegungsparameter diese Bedingung erfüllen. Es ergeben sich zwei Bedingungen für die Abstandsänderungen von Punkten:

1. Damit die Punkte in erster Näherung eine starre Konfiguration bilden, müssen die Änderungen der Punktabstände klein sein gegenüber den Abständen der betreffenden Punkte. Diese Bedingung wurde bei dem im Abschnitt 4.1 vorgestellten Verfahren berücksichtigt, indem die Standardabweichung als Maß für die Änderung des Punktabstandes mit dem Mittelwert des Punktabstandes in Relation gesetzt wurde.

Aus dieser Bedingung folgt, daß die Punkte im Raum möglichst weit auseinander liegen sollten, damit sich fehlerhafte Punktlokalisationen nicht so stark auswirken.

2. Damit die Bewegungsmatrix die reale Bewegung des starren Körpers erfassen kann, müssen die Abstandsänderungen der Punkte klein sein gegenüber den zurückgelegten Entfernungen der Punkte.

Daraus folgt, daß die Translation bzw. der Rotationswinkel möglichst groß ausfallen sollten. Eventuell ist dazu ein größeres Zeitintervall zu wählen.

Diese beiden Bedingungen sollten genügen, damit die ermittelte Bewegungsmatrix die Bewegung eines starren Körpers wiedergibt. Sie können beide sichergestellt werden. Die erste Bedingung ist direkt über den Schwellwert s bei der Gruppierung der Punkte zu starren Körpern beeinflussbar, während die zweite Bedingung noch zusätzlich geprüft werden muß.

2. Phase

Aus den ermittelten Rotationsmatrizen werden als nächstes die Orientierungen der Rotationsachsen ermittelt. Für eigentlich orthogonale Matrizen erhält man die Rotationsachse, indem man den zum Eigenwert $+1$ gehörenden Eigenvektor bestimmt. Weicht die ermittelte Rotationsmatrix hingegen von einer eigentlich orthogonalen Matrix ab, so muß kein Eigenwert $+1$ existieren. Wir berechnen daher zunächst die Eigenwerte der ermittelten Rotationsmatrix und betrachten den zum größten Eigenwert gehörenden Eigenvektor als Rotationsachse.

Für jedes starre Teil müssen wir im weiteren entscheiden, ob die Orientierungen der ermittelten Rotationsachsen konstant sind. Auch hier werden wir leichte Schwankungen in den Orientierungen tolerieren. Wir bilden daher zunächst den Betrag \bar{r} des Mittelwertvektors der normierten Achsenvektoren $\mathbf{r}(t_i)$:

$$\bar{r} = \left| \frac{1}{n} \sum_{i=0}^{n-1} \mathbf{r}(t_i) \right|$$

Weisen die Rotationsachsen alle in dieselbe Richtung, so nimmt \bar{r} den Wert eins an. Existiert keine Vorzugsrichtung bei den Orientierungen, so wird \bar{r} gleich null. Mit einem erneuten Schwellwert legen wir die Grenze zwischen konstanter Orientierung und variierender Orientierung. Der Schwellwert sollte dicht bei eins liegen.

Sind die Orientierungen der Rotationsachsen für die gegebenen Punkte eines starren Körpers konstant, so werden alle Punkte und die aus ihnen berechneten Bewegungsparameter der Translationsanalyse unterzogen. Zur Lösung des Gleichungssystems ist eine Koordinatentransformation nötig, durch die die Rotationsachse in Richtung der Z-Achse weist. Als Richtung der Rotationsachse kann der Mittelwertvektor der Rotationsvektoren¹³ dienen. Die Translationsanalyse liefert entsprechend dem zugrundegelegten Bewegungsmodell die körperfeste Achse sowie die Geschwindigkeit und die Beschleunigung,

¹³Die Richtung eines Rotationsvektors ist gleich der Rotationsachsenrichtung, wobei die Rotation stets rechtsdrehend um diesen Vektor erfolgt, und der Betrag ist gleich dem Rotationswinkel.

mit der sich diese Achse durch den Raum bewegt. Der Fehler bei der Translationsanalyse kann als Hinweis dienen, inwieweit das vorausgesetzte Bewegungsmodell erfüllt ist. Dabei sollten allerdings die Fehler, die bei der Berechnung der Bewegungsparametern aufgetreten sind, mit berücksichtigt werden, da diese bei der Translationsanalyse wieder mit eingehen.

Für den Fall, daß die Orientierungen der Rotationsachsen des betrachteten starren Körpers nicht konstant bleiben, wird seine Bewegung relativ zu allen anderen gegebenen starren Körpern beschrieben (siehe Abschnitt 5.2.1).¹⁴ Bei jeder dieser relativen Bewegungen wird erneut geprüft, ob die Orientierungen der ermittelten Rotationsachsen bei der Bewegung konstant bleiben. Sind die Orientierungen konstant, so werden die beiden beteiligten starren Körper als verbunden betrachtet, und die relativen Bewegungsmatrizen sowie die entsprechend transformierten Punkte des starren Körpers einer Translationsanalyse unterzogen. Bei der relativen Bewegungsbeschreibung werden neue Bewegungsmatrizen aus dem Produkt von ermittelten Bewegungsmatrizen berechnet. Sind die ermittelten Bewegungsmatrizen fehlerbehaftet, so pflanzt sich der Fehler durch die Produktbildung fort. Es muß daher sichergestellt sein, daß die von den Bewegungsmatrizen beschriebenen Bewegungen nur geringfügig von starren Bewegungen abweichen.

Bleibt bei keiner der relativen Bewegungen die Orientierung der Rotationsachse konstant, so schlägt die Bewegungsbeschreibung fehl. Dies ist genau dann der Fall, wenn der Vorgänger in der Bewegungshierarchie nicht bekannt ist.

¹⁴Die durchgeführten Tests beschränkten sich auf eine geringere Anzahl starrer Körper, so daß der kombinatorische Aufwand vertretbar war.

Kapitel 6

Ergebnisse

Zum Testen des Algorithmuses wurden zunächst künstliche Objekte konstruiert, die aus zueinander beweglichen starren polyedrischen Teilen bestehen. Die Teile dieser Objekte wurden nach vorgegebenen Bewegungsparametern von einem Zeitpunkt zum nächsten bewegt und die 3D-Koordinaten der Eckpunkte der Polyeder zu jedem Zeitpunkt gespeichert. Die Raumkoordinaten der Eckpunkte dienen dem Algorithmus als Eingabe. Es wurden auf diese Weise zwei unterschiedliche Testfolgen generiert, die sich sowohl in dem betrachteten Objekt als auch in den Bewegungen unterscheiden. Beide Testfolgen sind so beschaffen, daß sie dem Objekt- und Bewegungsmodell entsprechen, die dem vorgeschlagenen Algorithmus zugrundeliegen. Dabei bringen sie unterschiedliche Aspekte der zusammengesetzten Bewegung zum Ausdruck.

6.1 1. Testfolge

Abbildung 5.4 zeigt das Objekt, mit dem die erste Testfolge generiert wurde. Es besteht aus drei starren Einzelteilen, deren Ausmaße wie folgt gewählt wurden:¹ Sowohl die

	ΔX	ΔY	ΔZ
Teil 1	4.00 m	4.00 m	6.00 m
Teil 2	2.00 m	0.95 m	4.00 m
Teil 3	3.75 m	0.95 m	1.50 m

Teile 1 und 2 als auch die Teile 2 und 3 sind über eine Achse miteinander verbunden, die körperfest zu beiden Verbindungspartnern ist. Die Teile rotieren mit jeweils konstanter Winkelgeschwindigkeit um die in Abbildung 5.4 gezeigten Achsen. Die Rotationswinkel pro Zeitintervall betragen:

$$\alpha_1 = -10^\circ \quad \alpha_2 = -20^\circ \quad \alpha_3 = 50^\circ$$

¹Als Einheit wählen wir willkürlich 1 m

Darüberhinaus bewegt sich das Teil 1 bzw. seine Rotationsachse mit gleichmäßig beschleunigter Bewegung im WKS. Die Geschwindigkeit und die Beschleunigung lauten:

$$\mathbf{v} = \begin{pmatrix} 1.0 \\ 0.5 \\ 0.5 \end{pmatrix} \quad [\text{m/Zeitintervall}] \quad \mathbf{a} = \begin{pmatrix} -0.25 \\ 0.0 \\ -0.15 \end{pmatrix} \quad [\text{m/Zeitintervall}^2]$$

Die sich ergebende Bewegungsfolge ist in Abbildung 6.1 zu sehen. Die zeitliche Reihenfolge der Objektbilder erfolgt, wie auch in den folgenden Abbildungen von links oben nach links unten und von rechts oben nach rechts unten. Das Verfahren gruppiert zunächst die gegebenen Punkte zu den drei starren Teilen und errechnet für jedes Teil und jedes Zeitintervall die Bewegungsmatrizen. Aus diesen werden die Rotationsachsen der drei ermittelten Einzelteile bestimmt. Ihre Orientierungen sind in den Abbildungen 6.2, 6.3 und 6.4 bezüglich des WKS aufgetragen. Nur Teil 1 besitzt im gesamten Zeitraum eine konstante Rotationsachse.

Da Teil 1 mit invarianter Rotationsachse relativ zum WKS rotiert, haben wir die Bewegung seiner Rotationsachse im WKS bestimmt. Die durchgeführte Translationsanalyse liefert die Translationsparameter entsprechend dem verwendeten Bewegungsmodell (\mathbf{v} , \mathbf{a}) und die räumliche Lage der Rotationsachse (\mathbf{r}_0). In Abbildung 6.5 sind die Ergebnisse der Translationsanalyse wiedergegeben. Die räumlichen Positionen der Objektpunkte wurden mit Hilfe der Bewegungsgleichung 5.6 auf Seite 70 berechnet. Dazu wurden die Punktkoordinaten zum Zeitpunkt t_0 und die ermittelten Größen der Translationsanalyse verwendet. Die sich ergebenden Objektpositionen zu den Aufnahmezeitpunkten wurden ins Bildkoordinatensystem zurückprojiziert. Ferner ist die Lage der ermittelten Rotationsachse in den Bildern erkennbar. Die ermittelte Bewegung stimmt mit der Eingangsbewegung exakt überein. Die Geschwindigkeit, die Beschleunigung und die Lage der Rotationsachse im Raum wurden eindeutig und exakt ermittelt. In Abbildung 6.6 wird dieselbe Bewegung aus einer anderen Sicht betrachtet. Die Bildebene ist parallel zur XY-Ebene. Die Positionen von Teil 1 zu den 10 Zeitpunkten wurden alle überlagert. Die ermittelte Rotationsachse ist durch ein * markiert. Man erkennt deutlich die parabelförmige Trajektorie der Rotationsachse. Die Bewegung enthält auch einen Beschleunigungsanteil in Richtung der Z-Achse und damit in Richtung der Rotationsachse. Auch dieser wurde durch die Translationsanalyse eindeutig ermittelt.²

Die Orientierungen der Rotationsachsen von Teil 2 und Teil 3 sind nicht invariant bezüglich des WKS. Wir betrachten daher ihre relative Bewegung zu Teil 1 sowie zueinander. Bei jeder dieser relativen Bewegung bleibt die Orientierung der Rotationsachse konstant.

²Er läßt sich allerdings schwer aus Abbildung 6.6 entnehmen. Die Kantenlänge der abgebildeten Quadrate wird mit zunehmender Zeit zunächst größer, da sich das Objekt auf den Betrachterstandpunkt zubewegt und danach wieder kleiner, da die zunehmende Geschwindigkeit das Objekt von dem Betrachterstandpunkt wieder entfernt. Zur Kontrolle sei ein Lineal empfohlen!

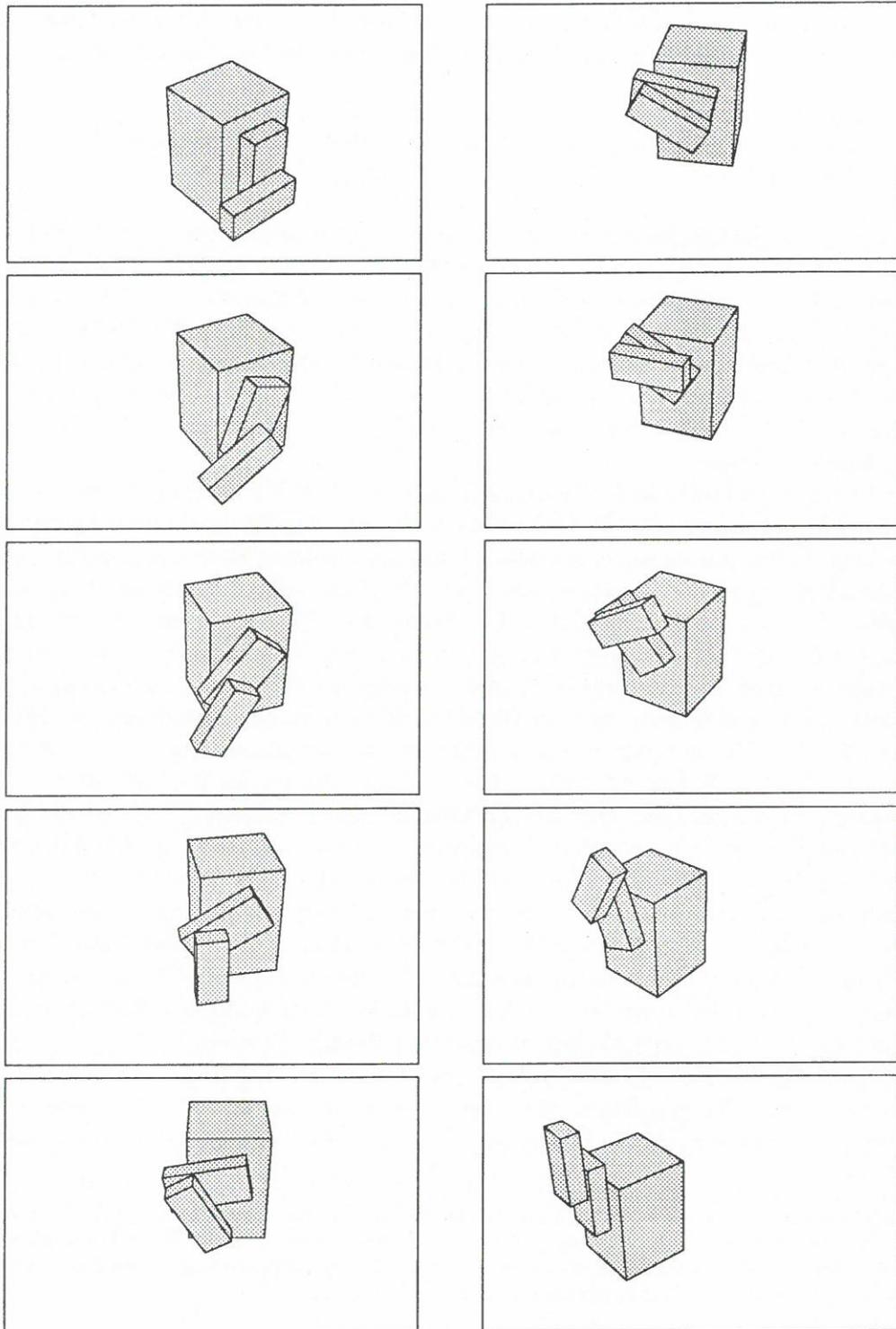


Abbildung 6.1: 1. Testfolge

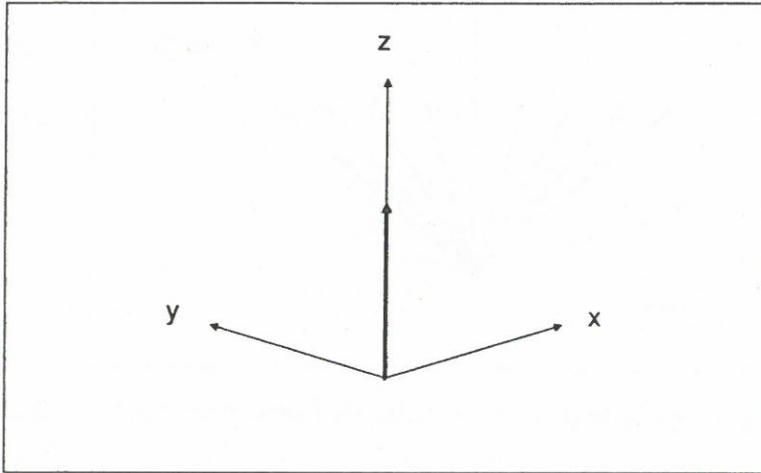


Abbildung 6.2: Orientierungen der ermittelten Rotationsachsen von Teil 1

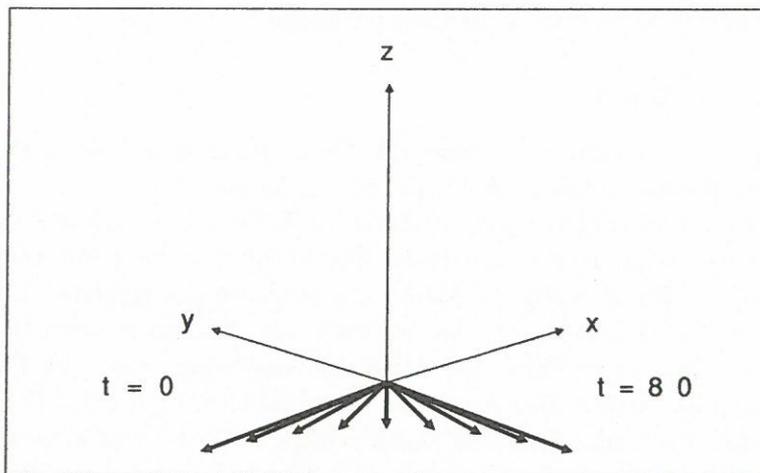


Abbildung 6.3: Orientierungen der ermittelten Rotationsachsen von Teil 2

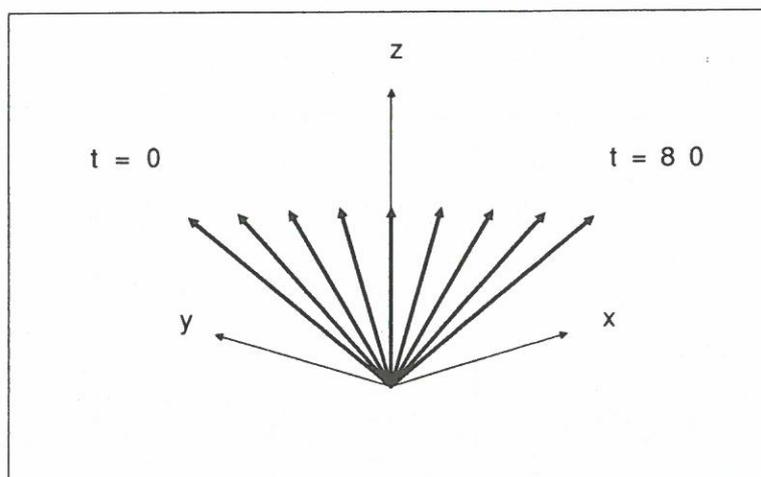


Abbildung 6.4: Orientierungen der ermittelten Rotationsachsen von Teil 3

Relative Bewegung von Teil 2

Die Ergebnisse der Translationsanalyse der relativen Bewegung von Teil 2 zu Teil 1 sind in Abbildung 6.7 dargestellt. Die ermittelte Rotationsachse bewegt sich nach den Ergebnissen der Translationsanalyse mit $\mathbf{v} = \mathbf{0}$ und $\mathbf{a} = \mathbf{0}$ relativ zu Teil 1. Die Rotationsachse ist damit zu beiden Teilen körperfest. Es sollte sich daher dieselbe Rotationsachse ergeben, wenn wir die umgekehrte Bewegung, also die Bewegung von Teil 1 relativ zu Teil 2 betrachten. Die Ergebnisse dieser Bewegung sind in Abbildung 6.8 gezeigt.

Auch die relative Bewegung von Teil 2 zu Teil 3 liefert eine zu beiden Teilen körperfeste Rotationsachse. Die Ergebnisse sind in Abbildung 6.9 gezeigt.

Relative Bewegung von Teil 3

Die relative Bewegung von Teil 3 zu Teil 2 bestätigt, die in der umgekehrten Bewegungsbeschreibung gefundene Rotationsachse (siehe Abbildung 6.10).

Bei der Bewegungsbeschreibung von Teil 3 relativ zu Teil 1 ergibt sich das erste Mal ein deutlicher Fehler bei der Translationsanalyse. Abbildung 6.11 zeigt die ermittelten Ergebnisse dieser Analyse. Dabei wurde zu jedem Zeitpunkt in gestrichelten Linien die räumliche Position von Teil 3 überlagert, wie sie sich aus den exakt ermittelten Bewegungsmatrizen ergibt. Die umgekehrte Bewegungsbeschreibung bringt die Fehler der Translationsanalyse noch deutlicher zum Ausdruck (siehe Abbildung 6.12). Es läßt sich also keine Rotationsachse für Teil 3 finden, die sich relativ zu Teil 1 mit konstanter Beschleunigung bewegt. Die Verbindungsachse von Teil 2 und Teil 3 bewegt sich relativ zu Teil 1 auf einer Kreisbahn. Eine Beschreibung durch eine Bewegung mit konstanter

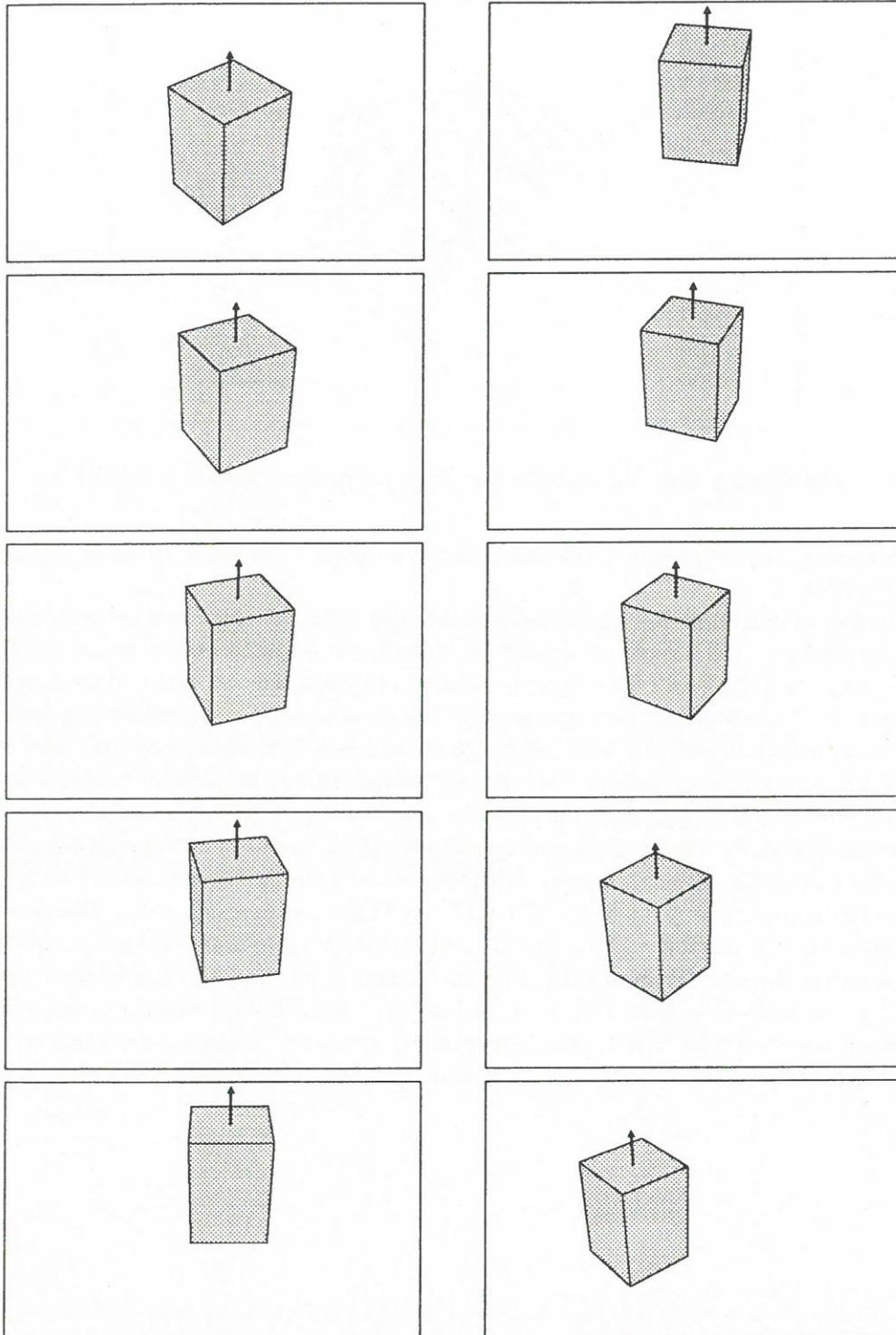


Abbildung 6.5: Bewegung von Teil 1 relativ zum WKS

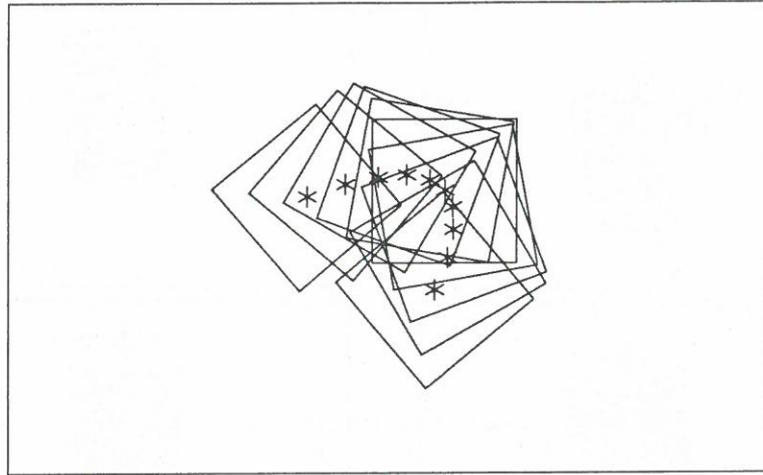


Abbildung 6.6: Trajektorie der Rotationsachse von Teil 1 im WKS

Beschleunigung ist in diesem Falle daher nicht möglich.³ Daher wird diese Achse auch nicht ermittelt.

Aus den ermittelten Bewegungsbeschreibungen läßt sich die Zusammensetzung des Objektes ersehen. Teil 1 ist durch seine invariante Rotationsachsenorientierung ein Hauptkörper. Für die Teile 2 und 3 werden die relativen Bewegungen zu Teil 1 sowie zueinander ermittelt. Da bei all diesen Bewegungen die Rotationsachse ihre Orientierung beibehält, ist nicht sofort klar, welches Teil mit welchem anderen Teil verbunden ist, oder anders ausgedrückt, zu welchem anderen Teil sich die Bewegung am einfachsten beschreiben läßt. Dieses Problem ergibt sich dadurch, daß die physikalischen Rotationsachsen der Teile 2 und 3 parallel sind. Dieser Fall mußte von *Asada + Tsuji 83* ausgeschlossen werden, damit die von ihnen vorgeschlagenen Verfahren funktionieren.⁴ In unserem Fall zeigt uns die Translationsanalyse jedoch deutlich auf, wo Verbindungen bestehen. Die Bewegung von Teil 3 relativ zu Teil 1, fällt durch ihren erheblich größeren Fehler bei der Translationsanalyse heraus. Eine exakte Übereinstimmung des Bewegungsmodells wurde bei der relativen Bewegung von Teil 3 zu Teil 2 (bzw. umgekehrt) sowie bei der relativen Bewegung von Teil 2 zu Teil 1 (bzw. umgekehrt) ermittelt. Dazu ergaben sich in beiden (bzw. allen vier) Fällen Achsen, die zu beiden Verbindungspartnern körperfest sind.

³Da die Beschleunigung ein Vektor ist, bedeutet konstante Beschleunigung, daß sowohl Betrag als auch Richtung dieses Vektors konstant bleiben. Bei einer Bewegung auf einer Kreisbahn bleibt nur der Betrag, nicht aber die Richtung der Beschleunigung konstant.

⁴Eine nicht unerhebliche Einschränkung, wie man an dem hier gezeigten Beispiel sieht.

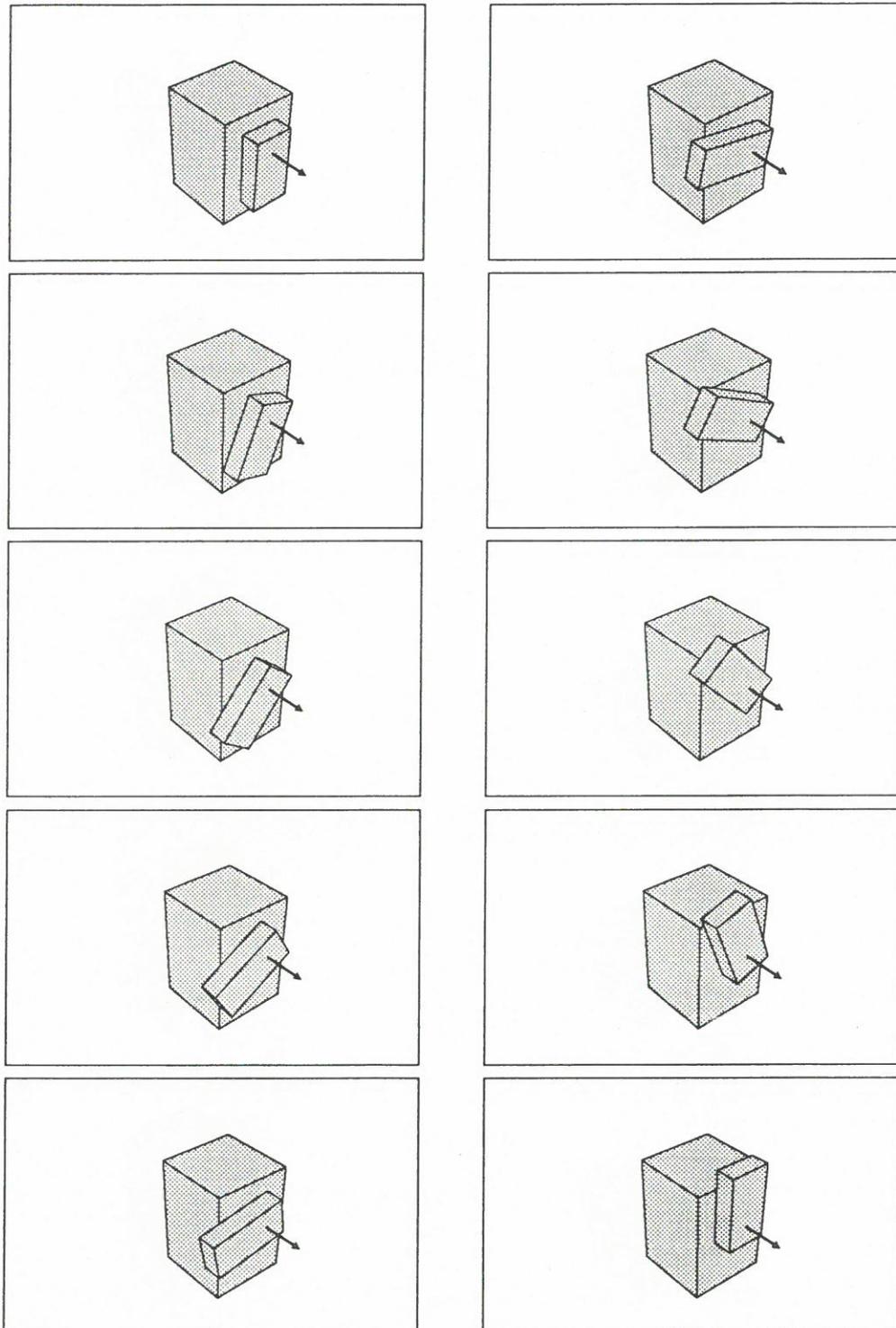


Abbildung 6.7: Bewegung von Teil 2 relativ zu Teil 1

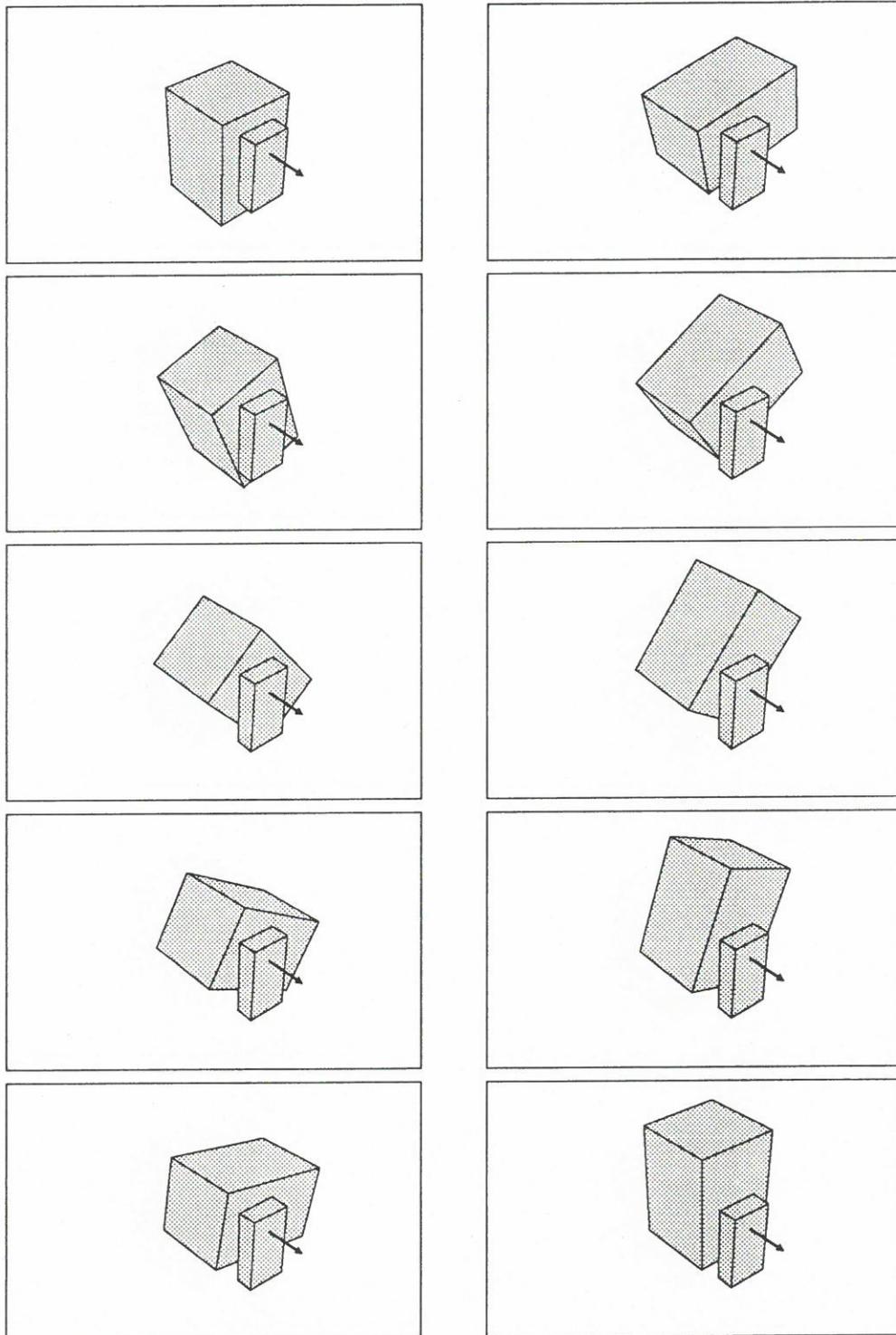


Abbildung 6.8: Bewegung von Teil 1 relativ zu Teil 2

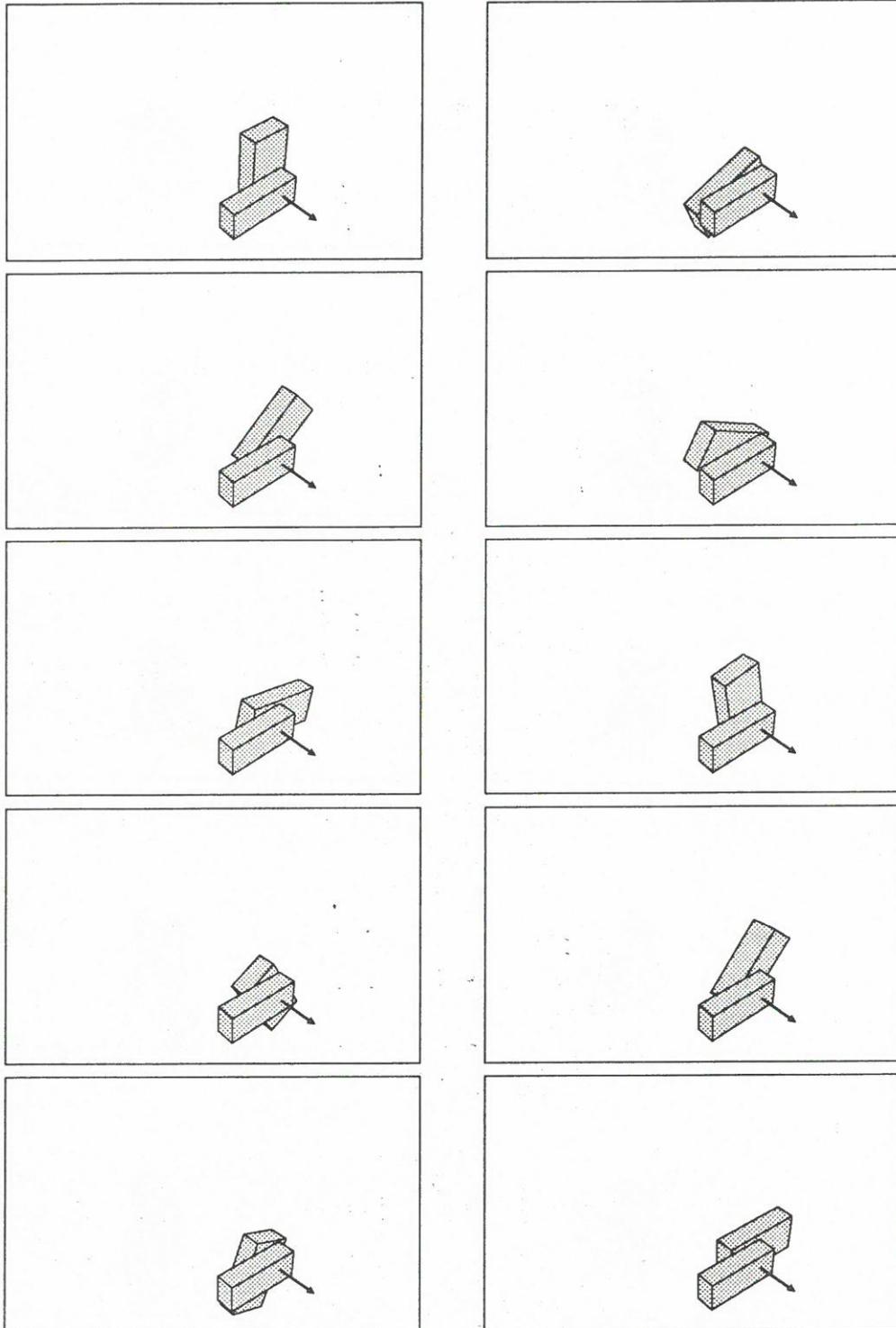


Abbildung 6.9: Bewegung von Teil 2 relativ zu Teil 3

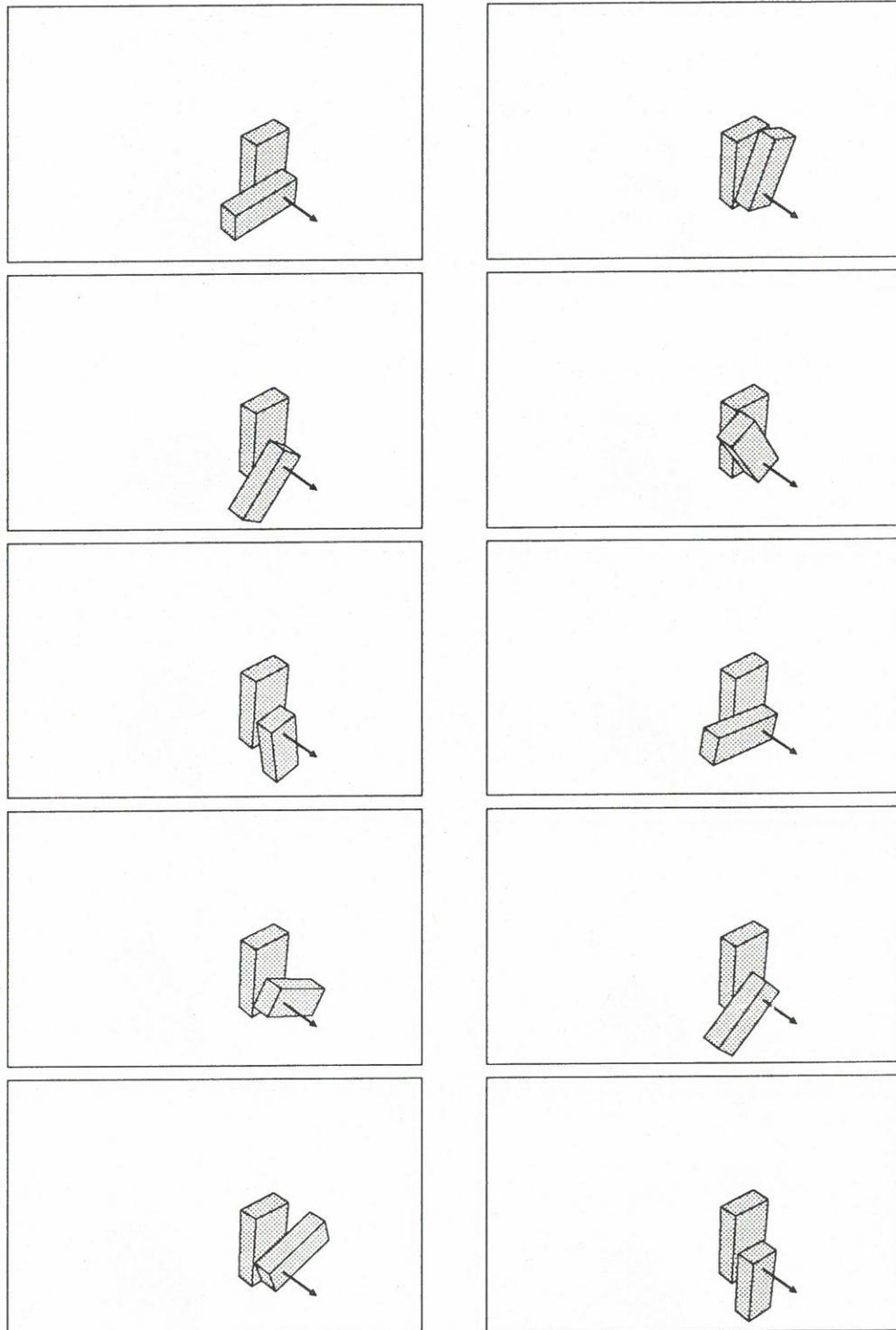


Abbildung 6.10: Bewegung von Teil 3 relativ zu Teil 2

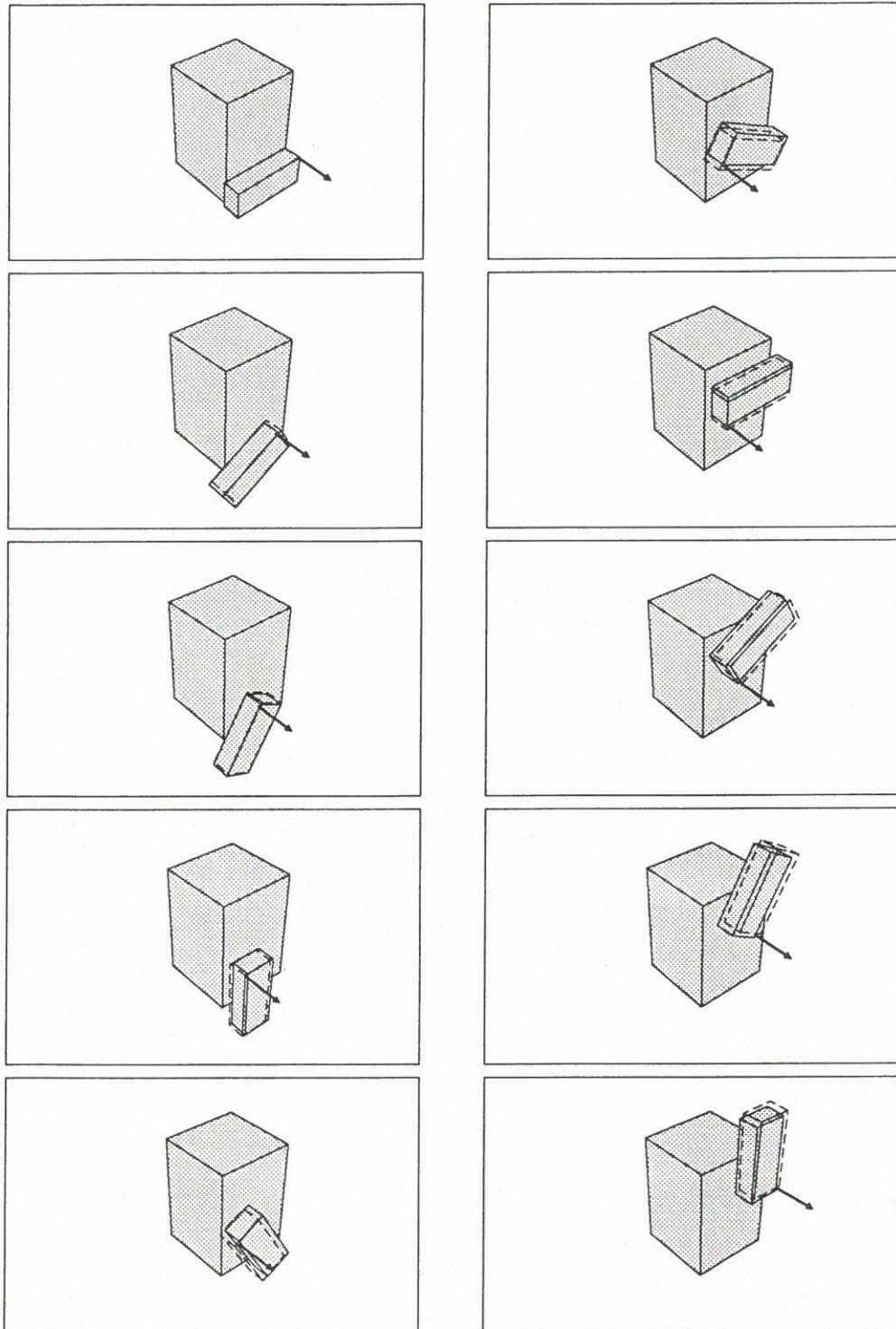


Abbildung 6.11: Bewegung von Teil 3 relativ zu Teil 1

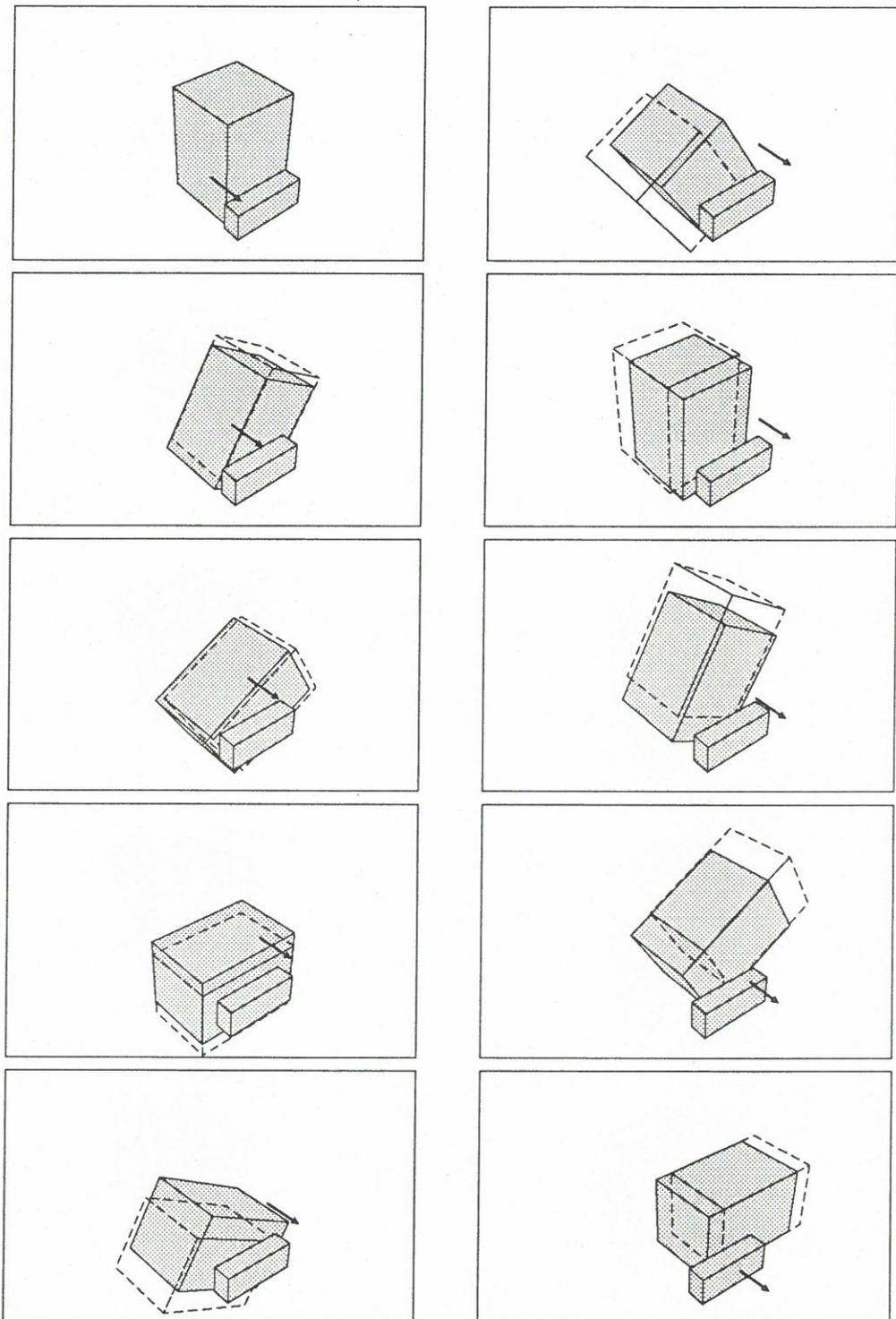


Abbildung 6.12: Bewegung von Teil 1 relativ zu Teil 3

6.2 2. Testfolge

Anhand der Bewegungsfolge aus Abbildung 6.13 wurde mit einem anderen Objekt ein weiterer Test durchgeführt. Die kleinsten umschreibenden Quader der beiden Teile haben die folgenden Ausmaße:

	ΔX	ΔY	ΔZ
Teil 1	4.0 m	1.0 m	4.0 m
Teil 2	3.5 m	1.0 m	1.0 m

Teil 1 führt eine reine Rotation um die Z-Achse mit dem Rotationswinkel $\alpha_1 = -10^\circ$ pro Zeitintervall aus. Die Rotation von Teil 2 relativ zu Teil 1 erfolgt in den ersten fünf Zeitintervallen mit einem Rotationswinkel von $\alpha_2 = -18^\circ$ pro Zeitintervall um die Y-Achse und danach um den Winkel $\alpha_2' = -22.5^\circ$. Darüberhinaus bewegt sich die Rotationsachse von Teil 2 in den ersten fünf Zeitintervallen mit der Geschwindigkeit

$$\mathbf{v} = \begin{pmatrix} -0.5 \\ 0.0 \\ 0.5 \end{pmatrix} \quad [\text{m/Zeitintervall}]$$

relativ zu Teil 1. Danach bricht die Translationsbewegung allerdings ab, und Teil 2 beschreibt nur noch eine reine Rotationsbewegung relativ zu Teil 1.

Im Gegensatz zur ersten Testfolge sind die beiden Teile in den ersten fünf Zeitintervallen also nicht über eine feste Achse miteinander verbunden. Darüberhinaus enthält die relative Bewegung von Teil 2 zu Teil 1 eine abrupte Bewegungsänderung, durch die die Bewegung im gesamten Zeitraum nicht mehr als gleichmäßig beschleunigt beschrieben werden kann.

Für Teil 1 ergibt die Translationsanalyse eine zum WKS fixierte Rotationsachse. Die Ergebnisse sind in Abbildung 6.14 in gewohnter Weise verdeutlicht.

Teil 2 besitzt bezüglich des WKS keine konstante Rotationsachsenorientierung und wird daher relativ zu Teil 1 beschrieben. An diese Bewegungsbeschreibung knüpft sich die Translationsanalyse an. Abbildung 6.15 läßt die Fehler bei dieser Analyse deutlich erkennen. Die ermittelte Rotationsachse liegt deutlich neben der Spitze von Teil 2. Die ermittelte Bewegung dieser Rotationsachse, die mit konstanter Beschleunigung erfolgt, gibt die korrekt ermittelte Objektbewegung, die in gestrichelten Linien dargestellt ist, nur ungenügend wieder.

Die Fehler bei der Translationsanalyse liegen in diesem Fall an der plötzlichen Änderung der Translationsbewegung. Diese Änderung bewirkt, daß die Bewegung, wenn man den gesamten Zeitraum betrachtet, nicht mehr dem benutzten Bewegungsmodell genügt und daher durch dieses schlecht beschrieben werden kann. Dies Beispiel macht deutlich, wie wichtig es ist, Unstetigkeiten im Bewegungsablauf zu entdecken. Zerlegen wir die obige Bewegungsfolge in zwei Teilfolgen, die wir unabhängig voneinander analysieren, so kann die Bewegung von Teil 2 relativ zu Teil 1 exakt beschrieben werden. Beide Teilfolgen genügen nämlich dem zugrundegelegten Bewegungsmodell. Die Ergebnisse dieser

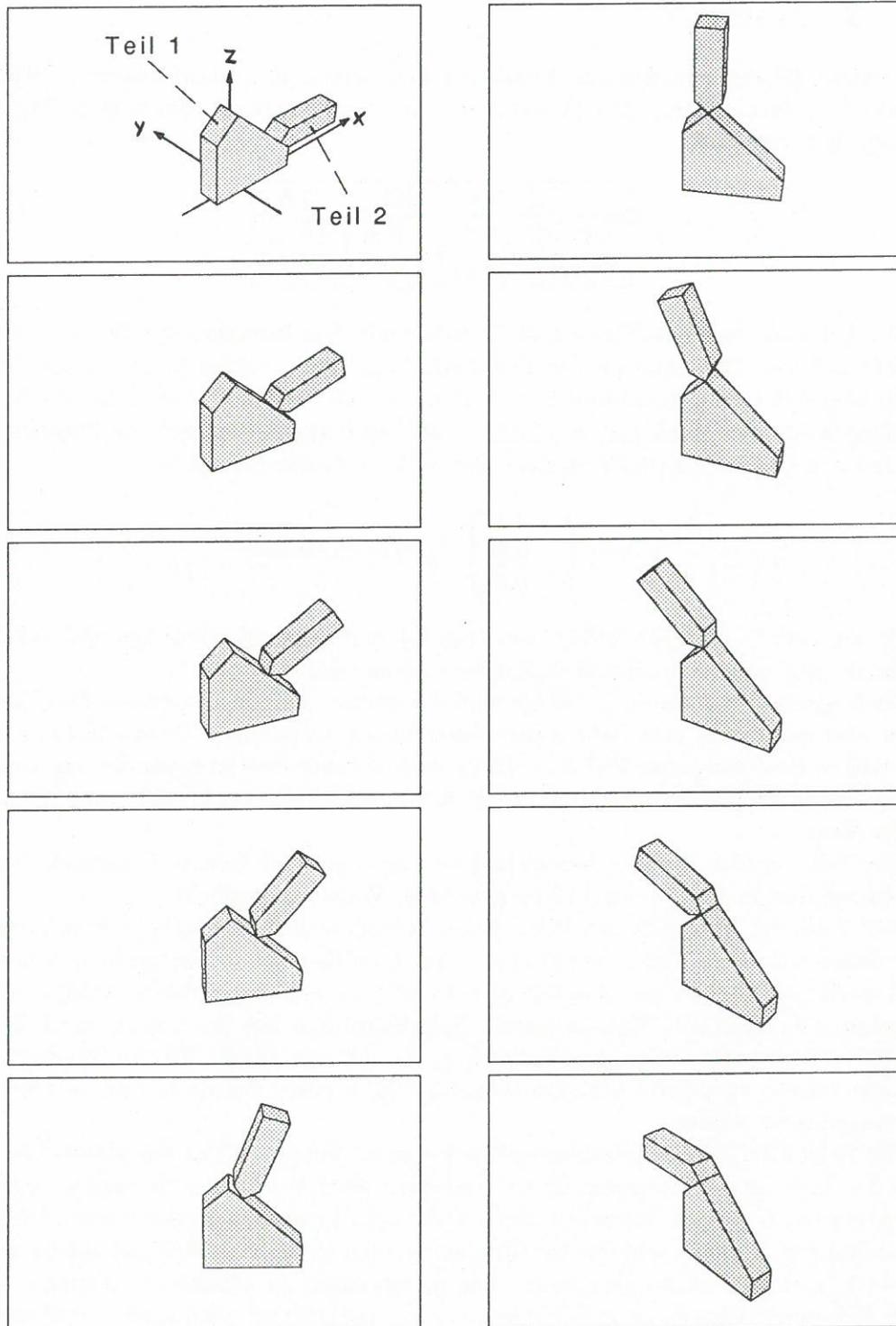


Abbildung 6.13: 2. Testfolge

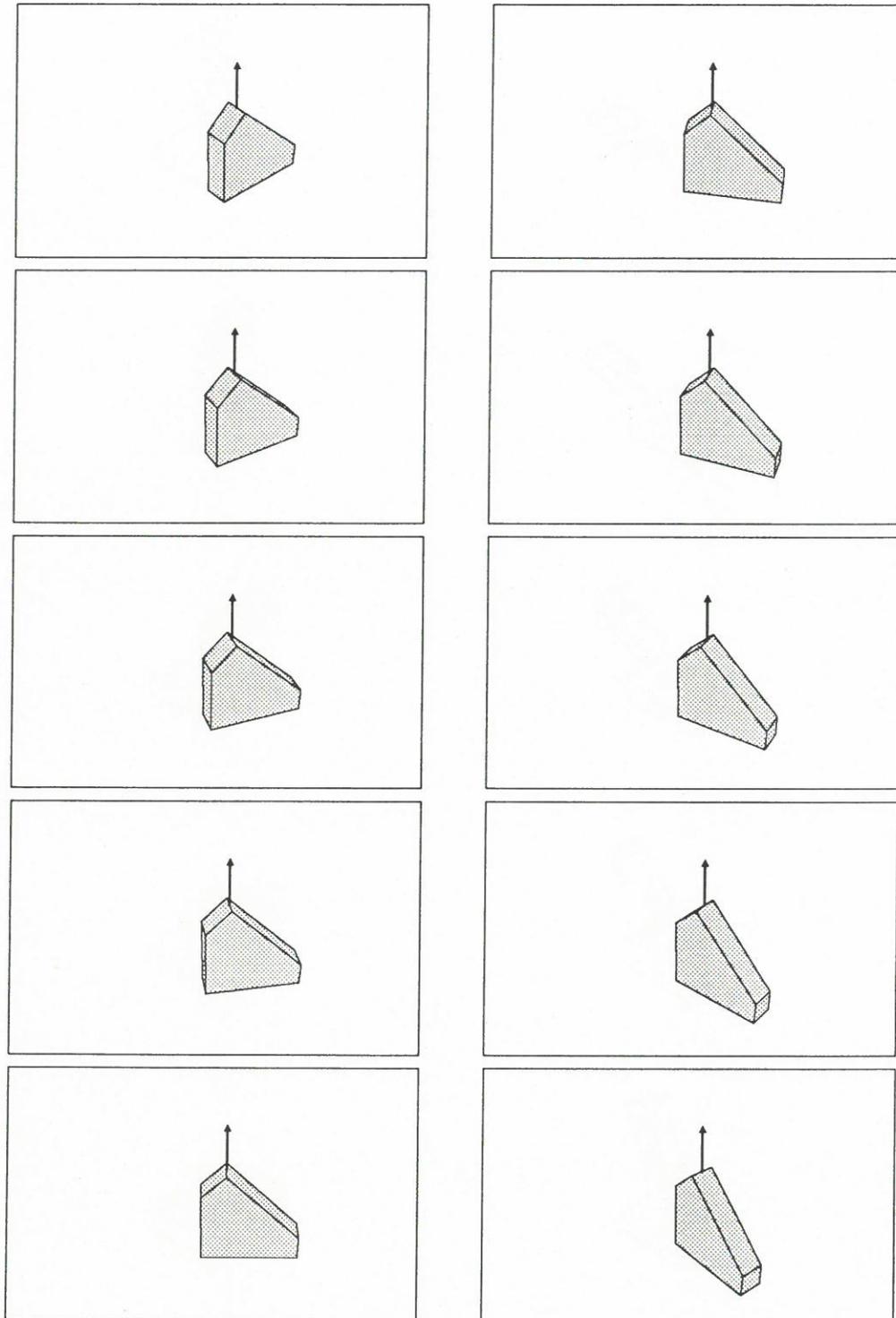


Abbildung 6.14: Bewegung von Teil 1 relativ zum WKS

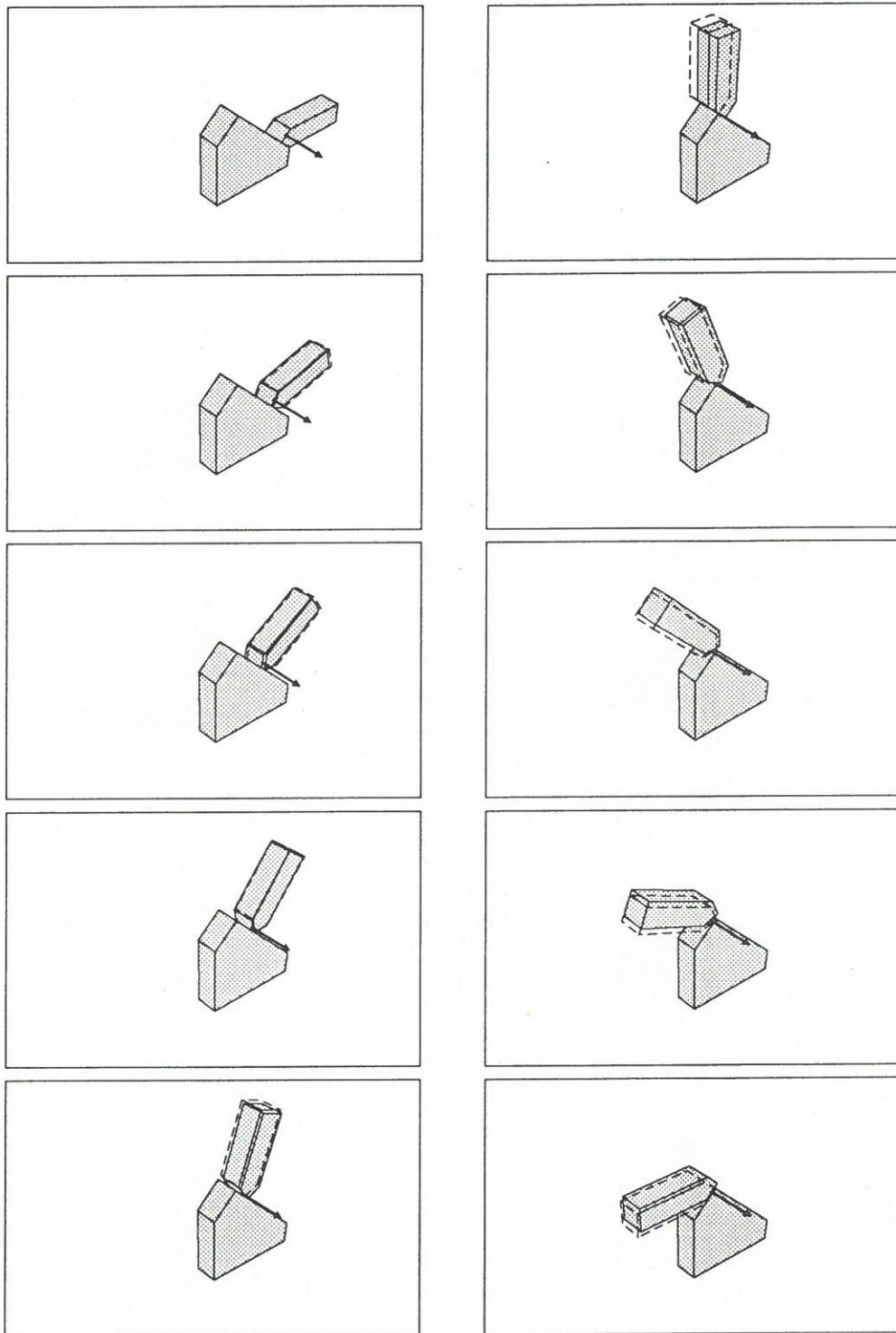


Abbildung 6.15: Bewegung von Teil 2 relativ zu Teil 1

getrennten Analyse sind der Vollständigkeit halber in Abbildung 6.16 gezeigt.

Die Ergebnisse zeigen deutlich, daß der vorgeschlagene Algorithmus geeignet ist, um die Bewegung von Objekten zu beschreiben, die aus starren Teilen zusammengesetzt sind, wobei zwischen den Teilen Verbindungen bestehen, die die Rotationsbewegung bis auf einen verbleibenden Rotationsfreiheitsgrad einschränken. Dadurch behält die Rotationsachse ihre Orientierung bezüglich der miteinander verbundenen Teile bei. Die Translationsbewegung dieser Rotationsachse kann mit konstanter Beschleunigung im Raum erfolgen. Folgende Bewegungsaspekte wurden durch die Tests untersucht:

- Teile sind über eine feste Achse miteinander verbunden, d. h. die Rotationsachse bleibt bei der relativen Bewegung der beiden Teile fixiert.
- Rotationsachse bewegt sich mit konstanter Geschwindigkeit (Teil 2 in der zweiten Folge, erste Hälfte).
- Rotationsachse bewegt sich mit konstanter Beschleunigung im Raum (Teil 1 in der ersten Folge).

Darüberhinaus wurden zwei weitere Phänomene betrachtet. In der ersten Folge betrifft dies die beiden Teile 2 und 3, deren Rotationsachsen dieselbe Orientierung aufweisen. Dadurch geht die Zusammensetzung des Objektes nicht mehr eindeutig aus der relativen Bewegungsbeschreibung hervor. Ein zweites Phänomen betrifft das Teil 2 in der zweiten Folge. In dem betrachteten Zeitraum wurde der Bewegungsverlauf von Teil 2 unterbrochen. Dadurch kann die Bewegung von Teil 2 relativ zu Teil 1 nicht mehr durch eine konstante Beschleunigung beschrieben werden, wenn der gesamten Zeitraum betrachtet wird.

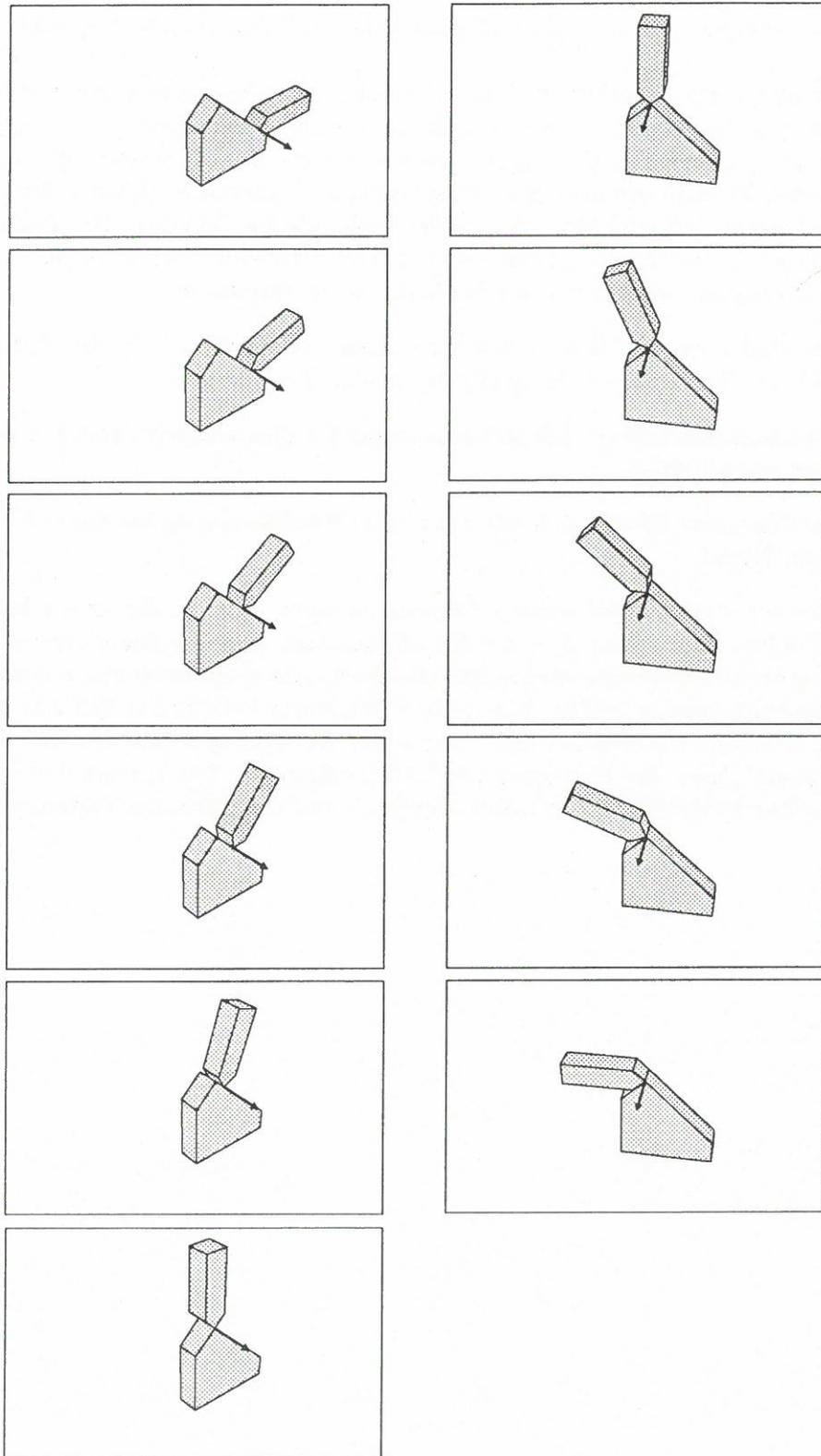


Abbildung 6.16: Bewegung von Teil 2 relativ zu Teil 1 in den beiden Teilfolgen. 1. Teilfolge in der linken, 2. Teilfolge in der rechten Spalte

Kapitel 7

Zusammenfassung

In dieser Arbeit wurde versucht, die räumliche Bewegung von Szenenobjekten zu beschreiben. Diese Beschreibung sollte die Bewegung auf eine einfache und natürliche Weise wiedergeben und darüberhinaus geeignet sein, die Bewegung des Objektes vorherzusagen, unter der Annahme, daß sich das Objekt unverändert weiterbewegt. Die Eigenschaft einer Bewegungsbeschreibung, Vorhersagen über ein Objekt machen zu können, zeigt wie gut die tatsächliche Bewegung verstanden wurde.

Um die Bewegung der Objekte im Raum beschreiben zu können, müssen wir zunächst die 3D-Koordinaten von korrespondierenden Bildpunkten bestimmen. In Kapitel 2 wurden daher zunächst Verfahren vorgestellt, die aus einer monokularen Bildfolge die räumlichen Koordinaten von zeitlich korrespondierenden Bildpunkten ermitteln. Die Verfahren zeichneten sich ausnahmslos durch einschränkende und vereinfachende Annahmen, wie z. B. Starrheit der Objekte und Parallelprojektion, aus. Diese Einschränkungen können wir umgehen, wenn wir eine Stereokamera verwenden. Diese ermöglicht es uns, Szenenpunkte aus räumlich korrespondierenden Bildpunkten zu rekonstruieren. Trotz dieses Vorteils ergeben sich aber auch bei der Stereoanordnung Grenzen, die einen Rückgriff auf "*Structure from Motion*"-Methoden erforderlich machen können. Man wird daher beide Vorgehensweisen bei der räumlichen Rekonstruktion berücksichtigen müssen. Im weiteren wurde vorausgesetzt, daß die 3D-Koordinaten von Objektpunkten in aufeinanderfolgenden Zeitpunkten bereits ermittelt wurden.

Das 3. Kapitel beschäftigte sich mit der Bewegung einzelner Punkte. Hier wurde versucht, die Trajektorie eines Punktes zu approximieren und die gewonnene Approximationsfunktion zu Vorhersagen zu verwenden. Schwierigkeiten ergaben sich bei der Wahl der Basisfunktionen, die geeignet sein mußten, die physikalischen Ursachen der Bewegung wiederzugeben. Eine geeignete Wahl konnte nur dann erfolgen, wenn die physikalischen Ursachen für die Bewegung des betrachteten Punktes bekannt waren. Diese Ursachen lassen sich der Bewegung einzelner Punkte nur schwer entnehmen, wenn man sie losgelöst von dem starren Körper betrachtet, zu dem sie gehören. Dies wird noch schwieriger, wenn wir es mit gekoppelten Bewegungen zu tun haben.

Es wurde daher im weiteren versucht, die Bewegung von starren Körpern zu beschreiben. Mathematisch läßt sich die Bewegung eines starren Körpers zwischen zwei Zeit-

punkten in homogenen Koordinaten durch eine lineare Abbildung darstellen. Wie auch in der Physik wird die Bewegung dabei in Rotation und Translation zerlegt. Die mathematische Beschreibung ist allerdings nicht eindeutig, da die Lage der Rotationsachse frei wählbar ist und die Translationsbewegung senkrecht zur Rotationsachse von dieser Wahl abhängt. Damit bringen die auf mathematischem Wege gewonnenen Rotations- und Translationsparameter die unterschiedlichen physikalischen Bewegungsphänomene der Rotation und Translation nicht eindeutig zum Ausdruck. Erst wenn die Lage der Rotationsachse korrekt gewählt wird, ist eine physikalische Interpretation der ermittelten Bewegungsparameter möglich.

Die erste Bedingung für die zu wählende Rotationsachse ist, daß sie bezüglich des Objektes körperfest ist. Weiterhin sollte die Rotationsachse so gewählt sein, daß die resultierende Translationsbewegung so einfach wie möglich wird. Eine einfache Bewegung gibt die physikalischen Ursachen am besten wieder. Für den Begriff "einfach" läßt sich keine exakte Definition angeben. Von daher ist auch nicht klar, nach was für Translationsbewegungen gesucht werden soll, um die "einfachste" Bewegungsbeschreibung zu erhalten. Dies hat uns zu der Wahl eines Bewegungsmodells veranlaßt, welches geeignet ist, verschiedenen Bewegungserscheinungen Rechnung zu tragen.

In Kapitel 5 wurden aus verbundenen starren Einzelteilen zusammengesetzte Objekte untersucht. Es wurden dabei nur solche Verbindungen (Kopplungen) zwischen den Einzelteilen berücksichtigt, die die relative Rotationsbewegung der verbundenen Einzelteile zueinander bis auf einen verbleibenden Rotationsfreiheitsgrad einschränken. Die relative Rotationsbewegung der verbundenen Teile vollzieht sich damit um eine Achse, die ihre Orientierung zu den beiden Verbindungspartnern nicht ändert. Nur die Winkelgeschwindigkeit unterliegt keinen Einschränkungen. Die Translationsbewegung dieser Achse wurde durch das Bewegungsmodell eingeschränkt. Der Fehler, der sich bei der Translationsanalyse ergab, gibt einen Hinweis darauf, wie gut sich die Bewegung durch das Modell beschreiben läßt.

Die Bewegung jedes Einzelteils wurde in einem Koordinatensystem beschrieben, in dem die Orientierung der Rotationsachse konstant ist. Als Koordinatensystem kam dabei das WKS oder das Objektkoordinatensystem eines anderen starren Teiles in Frage. In dem gewählten Koordinatensystem wurde die Lage der Rotationsachse so gelegt, daß dem Bewegungsmodell am besten entsprochen werden konnte. Das von mir gewählte Bewegungsmodell konnte verschiedene Bewegungen zwischen den Einzelteilen ausdrücken.

Unklar bleibt, was getan werden soll, wenn die Fehler bei der Anpassung des Bewegungsmodells zu groß werden. Ein großer Fehler zeigt, daß das Bewegungsmodell die Bewegung nur ungenügend beschreibt. An den gezeigten Beispielen hatten wir zwei verschiedene Gründe dafür kennengelernt. An der Beispielfolge in Abbildung 6.1 ergab sich ein großer Fehler bei der relativen Bewegung von Teil 3 zu Teil 1, weil zwischen diesen Teilen keine direkte Verbindung bestand. Diese relative Bewegungsbeschreibung wurde daher als nicht angemessen verworfen. In der zweiten Beispielfolge wurde der große Fehler bei der relativen Bewegung von Teil 2 zu Teil 1 allerdings durch eine plötzliche Bewegungsänderung von Teil 2 im betrachteten Zeitabschnitt hervorgerufen. Trotz des großen Fehlers ist das Bewegungsmodell bei diesem Beispiel angemessen. Dies wird deutlich,

wenn die Zeitabschnitte vor und nach der Bewegungsänderung getrennt betrachtet werden. Dies stellt uns vor das Problem, Unstetigkeiten im Bewegungsverlauf zu entdecken. Dieses Problem muß noch gelöst werden, damit man bei der Translationsanalyse nur Zeitintervalle berücksichtigt, in denen eine Konstanz der Bewegungsparameter (z. B. \mathbf{v} , \mathbf{a}) zu erwarten ist.

Als Ergebnis der Translationsanalyse ergibt sich eine Bewegungsbeschreibung, aus der sich Vorhersagen über die Lokalisation der starren Einzelteile und damit des Objektes ableiten lassen. Bisher unberücksichtigt blieb die Winkelgeschwindigkeit, mit denen die Einzelteile rotieren. Sie ist eine eindeutige Bewegungsgröße und brauchte bislang nicht berücksichtigt zu werden. Für eine Bewegungsvorhersage muß ihr Verlauf jedoch bekannt sein. Man könnte also wieder versuchen, die Winkelgeschwindigkeit eines Objektteiles durch eine Approximationsfunktion zu beschreiben. Auch hier ergeben sich die in Kapitel 3 genannten Probleme der Approximation. So könnte das betreffende Teil in dem gesamten betrachteten Zeitraum eine Rotation mit konstanter Winkelgeschwindigkeit ausführen oder aber auch eine Rotationsschwingung, vergleichbar mit einer Unruhe in einer Uhr. Um die geeigneten Basisfunktionen auszuwählen, müssen wir also wieder Annahmen machen über die zugrundeliegenden Bewegungsursachen. Dies ist ein grundlegendes Problem, von dem wir uns nicht befreien können. Auch in dem Bewegungsmodell steckt ja eine Annahme über die Art der Bewegung.

Ungeachtet dieses Problems bringen die ermittelten Bewegungsparameter die Bewegungsursachen klarer zum Ausdruck, als dies bei der Betrachtung eines einzelnen Punktes möglich ist. Die Zusammensetzung des Objektes und die Art der Verbindungen der Einzelteile konnte der gewonnenen Bewegungsbeschreibung entnommen werden.

Die erarbeiteten Verfahren wurden bislang nur auf künstliche Bewegungsfolgen angewendet. Es standen damit exakte räumliche Punktkoordinaten zur Verfügung. Damit der vorgeschlagene Algorithmus auch unter realen Bedingungen zur Anwendung kommen kann, wurden in Abschnitt 5.3 die Auswirkungen von Fehlern auf die einzelnen Schritte des Verfahrens berücksichtigt und entsprechende Maßnahmen vorgeschlagen.

Anhang A

Mathematischer Anhang

Eine kurze Übersicht über Ableitungen von Skalaren, Vektoren und Matrizen findet sich in [Horn 86, Seite 458], aus dem ein Teil der nun folgenden Ableitungsregeln stammt.

Ableitung eines Skalars nach einem Vektor:

Die Ableitung eines Skalars f nach einem Vektor \mathbf{v} ist ein Vektor, dessen Komponenten die Ableitungen des Skalars nach den einzelnen Vektorkomponenten darstellen.

$$\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \frac{df}{d\mathbf{v}} = \begin{pmatrix} \frac{df}{dx} \\ \frac{df}{dy} \\ \frac{df}{dz} \end{pmatrix}$$

Hieraus ergeben sich unter anderem die folgenden Ableitungen:

$$\frac{d(\mathbf{a}^T \mathbf{M} \mathbf{b})}{d\mathbf{a}} = \mathbf{M} \mathbf{b} \quad \text{und} \quad \frac{d(\mathbf{a}^T \mathbf{M} \mathbf{b})}{d\mathbf{b}} = \mathbf{M}^T \mathbf{a}$$

insbesondere gilt:

$$\frac{d(\mathbf{x}^T \mathbf{M} \mathbf{x})}{d\mathbf{x}} = (\mathbf{M} + \mathbf{M}^T) \mathbf{x}$$

Unter Verwendung dieser Ableitungen ergibt sich weiterhin:

$$\frac{d(\mathbf{M}_1 \mathbf{a} + \mathbf{M}_2 \mathbf{b})^T (\mathbf{M}_1 \mathbf{a} + \mathbf{M}_2 \mathbf{b})}{d\mathbf{b}} = 2\mathbf{M}_2^T (\mathbf{M}_1 \mathbf{a} + \mathbf{M}_2 \mathbf{b}) \quad (\text{A.1})$$

Diese Ableitungsregel wurde zur Ableitung der Fehlerfunktion in Gleichung 5.8 verwendet.

Ableitung eines Skalars nach einer Matrix:

Die Ableitung eines Skalars f nach einer Matrix \mathbf{M} ist eine Matrix, deren Komponenten die Ableitungen des Skalars nach den einzelnen Matrixkomponenten darstellen.

$$\mathbf{M} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \quad \frac{df}{d\mathbf{M}} = \begin{pmatrix} \frac{df}{dm_{11}} & \frac{df}{dm_{12}} & \frac{df}{dm_{13}} \\ \frac{df}{dm_{21}} & \frac{df}{dm_{22}} & \frac{df}{dm_{23}} \\ \frac{df}{dm_{31}} & \frac{df}{dm_{32}} & \frac{df}{dm_{33}} \end{pmatrix}$$

Hieraus ergeben sich unter anderem die folgenden Ableitungen:

$$\frac{d(\mathbf{a}^T \mathbf{M} \mathbf{b})}{d\mathbf{M}} = \mathbf{a} \mathbf{b}^T \quad (\text{dyadisches Produkt})$$

$$\frac{d(\mathbf{x}^T \mathbf{M} \mathbf{x})}{d\mathbf{M}} = \mathbf{x} \mathbf{x}^T$$

$$\frac{d\text{Spur}(\mathbf{M})}{d\mathbf{M}} = \mathbf{I} \quad (\text{Einheitsmatrix})$$

$$\frac{d\text{Spur}(\mathbf{A} \mathbf{M} \mathbf{B})}{d\mathbf{M}} = \mathbf{A}^T \mathbf{B}^T$$

$$\Rightarrow \frac{d\text{Spur}(\mathbf{A} \mathbf{B})}{d\mathbf{A}} = \mathbf{B}^T \quad \text{und} \quad \frac{d\text{Spur}(\mathbf{A} \mathbf{B})}{d\mathbf{B}} = \mathbf{A}^T$$

Die folgende Ableitung ergibt sich nach langwieriger Rechnerei:

$$\frac{d\text{Spur}(\mathbf{B}^T \mathbf{M}^T \mathbf{M} \mathbf{B} \mathbf{W})}{d\mathbf{M}} = \mathbf{M} \mathbf{B} \mathbf{W} \mathbf{B}^T + \mathbf{M} \mathbf{B} \mathbf{W}^T \mathbf{B}^T$$

Diese Ableitungsregel wurde auf Seite 22 benötigt, um die Fehlerfunktion in Gleichung 3.3 abzuleiten.

Ableitung eines Vektors nach einem Vektor:

Die Ableitung eines Vektors \mathbf{w} nach einem Vektor \mathbf{v} ist eine Matrix, deren Komponenten die Ableitungen der Vektorkomponenten von \mathbf{w} nach den Vektorkomponenten von \mathbf{v} darstellen.

$$\mathbf{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} \quad \mathbf{w} = \begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix}$$

$$\frac{d\mathbf{w}}{d\mathbf{v}} = \begin{pmatrix} \frac{dw_x}{dv_x} & \frac{dw_x}{dv_y} & \frac{dw_x}{dv_z} \\ \frac{dw_y}{dv_x} & \frac{dw_y}{dv_y} & \frac{dw_y}{dv_z} \\ \frac{dw_z}{dv_x} & \frac{dw_z}{dv_y} & \frac{dw_z}{dv_z} \end{pmatrix}$$

Sei $\mathbf{w} = \mathbf{M}\mathbf{v}$, so erhalten wir:

$$\frac{d(\mathbf{w})}{d\mathbf{v}} = \frac{d(\mathbf{M}\mathbf{v})}{d\mathbf{v}} = \mathbf{M}$$

Die Matrix $d\mathbf{w}/d\mathbf{v}$ bezeichnet man daher auch als *Jakobi*-Matrix der Koordinatentransformation von \mathbf{v} nach \mathbf{w} .

Zusätzliche Hilfsgleichungen:

$$\begin{aligned} \text{Spur}(\mathbf{A} + \mathbf{B}) &= \text{Spur}(\mathbf{A}) + \text{Spur}(\mathbf{B}) \\ \text{Spur}(\mathbf{M}^T) &= \text{Spur}(\mathbf{M}) \end{aligned}$$

Literaturverzeichnis

[Angeles 82]

Spatial Kinematic Chains — Analysis-Synthesis-Optimization.
Jorge Angeles.
Springer-Verlag, Berlin – Heidelberg – New York 1982.

[Asada + Tsuji 81]

Reconstruction of Three-Dimensional Motions from Image Sequences.
Minoru Asada, Saburo Tsuji.
Proc. Int. Conf. Patt. Recogn. Image Proc. PRIP-81, 88-90.

[Asada + Tsuji 83]

Representation of Three-Dimensional Motion in Dynamic Scenes.
Minoru Asada, Saburo Tsuji.
Computer Vision, Graphics, and Image Processing 21 (1983) 118-144.

[Ballard + Brown 82]

Computer Vision.
Dana H. Ballard, Christopher M. Brown.
Prentice-Hall, Englewood Cliffs, New Jersey 1982.

[Blostein + Huang]

Algorithms for Estimation Based on Three-Dimensional Correspondences.
S.D. Blostein, T.S. Huang.
Motion Understanding – Robot and Human Vision, W.N. Martin, J.K. Aggarwal
(eds.), Kluwer Academic Publ., Boston – Dordrecht – Lancaster 1988, 329-352.

[Bock 89]

Bestimmung der relativen Orientierung zweier planarer Objekte zur videogestützten geometrischen Vermessung von Unterkieferbewegungen.
Michael Bock.
Diplomarbeit, Fachbereich Informatik, Universität Hamburg, 1989.

[Bonde + Nagel 79]

Deriving a 3-D Description of a Moving Rigid Object from Monocular TV-Frame Sequences.

Thomas Bonde, Hans-Hellmut Nagel.

Workshop on Computer Analysis of Time-Varying Imagery WCATVI-79, Philadelphia, PA, April 1979, J.K. Aggarwal, N.I. Badler (eds.), 44–45.

[Dreschler-Fischer 87]

Das "Bootstrap-Problem" bei der geometrischen Szenenrekonstruktion — Eine Übersicht.

Leonie S. Dreschler-Fischer.

11th German Workshop on Artificial Intelligence "GWAI-87", Geseke, September 28 – October 2, 1987, Katharina Morik (Hrsg.), Informatik Fachberichte Bd. 152, Springer-Verlag, Berlin – Heidelberg – New York 1987, 1–15.

[Duda + Hart 73]

Pattern Classification and Scene Analysis.

Richard O. Duda, Peter E. Hart.

John Wiley & Sons, New York 1973.

[Engeln-Müllges + Reutter 88]

Formelsammlung zur Numerischen Mathematik mit MODULA 2-Programmen.

Gisela Engeln-Müllges, Fritz Reutter.

BI-Wissenschaftsverlag, Mannheim–Wien–Zürich 1988.

[Hildreth + Grzywacz 86]

The Incremental Recovery of Structure from Motion: Position vs. Velocity Based Formulations.

Ellen C. Hildreth, Norberto M. Grzywacz.

Proc. Workshop on Motion: Representation and Analysis, Charleston, South Carolina, May 1986, 137–143.

[Horn 86]

Robot Vision.

Berthold Klaus Paul Horn.

The MIT Press, Cambridge, Mass. 1986.

[Jenkin + Tsotsos 86]

Applying Temporal Constraints to the Dynamic Stereo Problem.

Michael Jenkin, John K. Tsotsos.

Computer Vision, Graphics, and Image Processing **33** (1986) 16–32.

[Martin + Aggarwal 88]

Motion Understanding — Robot and Human Vision.

W.N. Martin, J.K. Aggarwal (eds.).

Kluwer Academic Publ., Boston – Dordrecht – Lancaster 1988.

[Nagel + Neumann 81]

On 3D Reconstruction from Two Perspective Views.

Hans-Hellmut Nagel, Bernd Neumann.

Proc. 7. Int. Joint Conf. on Art. Intell. IJCAI-81, Vancouver, BC, Canada, August 1981, 661-663.

[Neumann 82]

Knowledge Sources for Understanding and Describing Image Sequences.

Bernd Neumann.

6th German Workshop on Artificial Intelligence "GWAI-82", Bad Honnef, September 27 - October 1, 1982, Wolfgang Wahlster (Hrsg.), Informatik Fachberichte Bd. 58, Springer-Verlag, Berlin - Heidelberg - New York 1982, 1-21.

[Roach + Aggarwal 80]

Determining the Movement of Objects from a Sequence of Images.

John W. Roach, J. K. Aggarwal.

IEEE Trans. on Pattern Analysis and Machine Intelligence PAMI-2 (1980) 554-562.

[Tsai + Huang 81]

Estimating Three-Dimensional Motion Parameters of a Rigid Planar Patch.

R. Y. Tsai, T. S. Huang.

Proc. Int. Conf. Patt. Recogn. Image Proc. PRIP-81, 94-97.

[Tsai + Huang 82]

Uniqueness and Estimation of Three-Dimensional Motion Parameters of Rigid Objects with Curved Surfaces.

R. Y. Tsai, T. S. Huang.

Proc. Int. Conf. Patt. Recogn. Image Proc. PRIP-82, 112-118.

[Ullman 79]

The Interpretation of Visual Motion.

Shimon Ullman.

The MIT Press, Cambridge, Mass. 1979.

[Ullman 83]

Maximizing Rigidity: The Incremental Recovery of 3-D Structure from Rigid and Rubbery Motion.

Shimon Ullman.

MIT AI Laboratory, A.I. Memo No. 721, 1983.

[Wallach + O'Connell 53]

The Kinetic Depth Effect.

H. Wallach, D. N. O'Connell.

Journal of Experimental Psychology 45 (1953) 205-217.

[Webb + Aggarwal 82]

Structure from Motion of Rigid and Jointed Objects.

Jon A. Webb, J. K. Aggarwal.

Artificial Intelligence **19** (1982) 107-130.

[Williams 80]

Depth from Camera Motion in a Real World Scene.

Thomas D. Williams.

IEEE Trans. on Pattern Analysis and Machine Intelligence **PAMI-2** (1980) 511-516.

Danksagung

An dieser Stelle möchte ich mich bei meiner Betreuerin Prof. Dr. Leonie Dreschler-Fischer für die Unterstützung bei der Anfertigung dieser Arbeit bedanken. Sie hatte stets ein offenes Ohr für mich.

Weiterhin richtet sich mein Dank an alle Mitglieder des SISSY-Projektes, für die schon Jahre andauernde freundschaftliche Zusammenarbeit. Ganz besonderer Dank gebührt Carsten Schröder, der in seiner gewohnten Akribie den ersten Entwurf dieser Arbeit begutachtete. Seine kritischen Anmerkungen waren eine große Hilfe.

Ingeborg Heer-Mück und Hartmuth Krüger danke ich für ihre Hilfsbereitschaft bei Problemen mit der Hardware.

Zum Schluß möchte ich ganz besonders meinen Eltern danken, die mir meine langwierige Ausbildung ermöglicht haben.

