Kalibrierung der internen Parameter einer aktiven Kamera

Diplomarbeit
am Lehrstuhl für Kognitive Systeme
des Instituts für Informatik und Praktische Mathematik
der Christian–Albrechts–Universität zu Kiel

vorgelegt von
Jörg Ernst

Betreuer: Dr. Kostas Daniilidis
Prof. Dr. Gerald Sommer

August 1996

Zusammenfassung

Diese Diplomarbeit beschäftigt sich mit der Kalibrierung der internen Parameter der Kameras einer binokularen Plattform mit vier mechanischen Freiheitsgraden. Diese ist nötig, um dreidimensionale metrische Informationen aus Bilddaten zu extrahieren.

Die Bestimmung der internen Parameter einer Kamera erlaubt die Abbildung von Pixelkoordinaten auf die entsprechenden Sehstrahlen im Raum. Interne Parameter sind die Position des Schnittpunktes der optischen Achse mit der Bildebene und die Skalierungsfaktoren, die die Angabe der Projektion eines Punktes in Pixeleinheiten bestimmen.

Verfahren, die visuelle Aufgaben mit Hilfe einer solchen Plattform lösen können, setzen diese Informationen voraus.

Kern dieser Arbeit ist ein Verfahren, das diese Parameter für eine Kameraplattform dadurch bestimmt, daß deren Freiheitsgrade ausgenutzt werden. Im Gegensatz zu Verfahren, die ein Referenzobjekt benutzen, setzt das hier vorgestellte Verfahren nur die Existenz einer Schar paralleler Geraden in den Bildausschnitten voraus.

Die meisten Menschen sind genau so glücklich, wie sie sein wollen.

A. Lincoln

Danksagung

Für die Unterstüzung durch Diskussionen und Anregungen bedanke ich mich bei I. Ahrns, K. Daniilidis, R. Hoschke, B. Leoniuk und G. Sommer.

Inhaltsverzeichnis

1 Einleitung **4**

1.1 Aktives Sehen 4

1.2 Binokulare Kameraplattformen 5

1.3 Kalibrierung 5

1.4 Einführendes Beispiel 6

1.5 Ziel der Arbeit 7

1.6 Aufbau der Arbeit 8

2 Geometrie und Architektur **10**

2.1 Definition der Koordinatensysteme 11

2.2 Projektive Geometrie 13

2.2.1 Homogene Koordinaten von Punkten in der Ebene 14

2.2.2 Homogene Koordinaten von Geraden in der Ebene 14

2.3 Geometrische Beschreibung 16

2.3.1 Kinematik der Kameraplattform 16

2.3.2 Beschreibung der Zentralprojektion 18

2.3.3 Transformation zwischen Szene und Kamera 23

2.4 Fluchtpunkte 23

2.5 Systemarchitektur 24

3 Kalibrierung **27**

3.1 Kalibrierverfahren 27

3.2 Klassische Kalibrierverfahren 29

3.2.1 Definition der klassischen Kalibrierung 30

3.2.2 Erweiterung der Zentralprojektion 31

3.2.3 Perspektivische Transformationsmatrix 32

3.2.4 Bestimmung der perspektivischen Transformationsmatrix . 33

3.2.5 Zerlegung der perspektivischen Transformationsmatrix . . 35

3.3 Kalibrierung mit Fluchtpunkten 36

3.3.1 Kalibrierung der internen Parameter 37

3.3.2 Kalibrierung der Rotationsmatrix 38

3.4 Aktive Kalibrierverfahren 38

3.4.1 Kalibrierung des Bildhauptpunktes durch Zoomen 38

3.4.2 Aktive Kalibrierung durch Rotationen 40
3.4.3 Kalibrierung mittels bekannter Rotationen 41
3.4.4 Kalibrierung mittels unbekannter Rotation 44
3.4.5 Lösungsansätze anderer Autoren 50
3.5 Klassifizierung der Kalibrierverfahren 58

4 Detektion von Referenzmerkmalen 61
4.1 Kalibrierkörper 62
4.1.1 Kalibrierwürfel 62
4.1.2 Zuordnung der Merkmale eines Kalibrierwürfels 63
4.2 Detektion von Bildmerkmalen 65
4.2.1 Die Hough-Transformation 66
4.2.2 Anwendung der Hough-Transformation zur Geradendetektion 68
4.2.3 Verfahren zur Fluchtpunktschätzung 69
4.2.4 Verfahren zur Hyperbelschätzung 71
4.2.5 Detektion von Bildmerkmalen eines Kalibrierwürfels 74
4.2.6 Zuordnung der Bildmerkmale eines Kalibrierwürfels 77
4.2.7 Detektion von konjugierten Fluchtpunkten 79

5 Ergebnisse 81
5.1 Simulation mit synthetischen Daten 81
5.1.1 Meßfehler in den Punkten und im Rotationswinkel 83
5.1.2 Entfernung des Kalibrierobjekts 86
5.1.3 Gesamtschwenkwinkel 87
5.1.4 Symmetrieindex 89
5.1.5 Relative Orientierung der Parallelenschar 90
5.1.6 Anzahl der Bildmerkmale 91
5.1.7 Anzahl der Zoomschritte 92
5.2 Realdatenexperiment 93

6 Diskussion und Ausblick 98
6.1 Diskussion der vorgestellten Methoden 98
6.2 Ausblick auf zukünftige Erweiterungen 102

A Kalibrieralgorithmen 103
A.1 Die klassische Kalibriermethode 103
A.2 Die Zoom-Methode 104
A.3 Die Methode der konjugierten Fluchtpunkte 104
A.4 Die Rotationsmethde 105
A.5 Die Hyperbelmethode 106

B Merkmalsalgorithmen 108
B.1 Bildpunktdetektion auf einem Kalibrierwürfel 109
B.1.1 Segmentierung 109
B.1.2 Merkmalsextraktion 109
B.2 Geradenscharen mit parallelen Urbildern 110
B.3 Fluchtpunktschätzung 111

C Zusammenspiel der Merkmals- und Kalibrieralgorithmen 112

Kapitel 1

Einleitung

1.1 Aktives Sehen

Das maschinelle Sehen beschäftigt sich mit der Auswertung von Bildern oder Bildfolgen mit dem Ziel, notwendige Informationen für eine bestimmte Aufgabe zu gewinnen. In den letzten Jahren entwickelte sich innerhalb des Gebiets des maschinellen Sehens das Paradigma der *verhaltensbasierten* Informationsverarbeitung. Darunter versteht man die Entwicklung von *allgemeinen* Lösungen für visuelle Aufgaben, die zur Umsetzung eines *speziellen* Verhaltens in einem aktiven System notwendig sind [Som95]. Diesem Paradigma wird auch der Ansatz des aktiven Sehens zugewiesen.

Systeme des aktiven Sehens besitzen Mechanismen, die aktiv die Parameter von Kamerasystemen in Abhängigkeit von den Anforderungen der zu bewältigenden Aufgaben kontrollieren. Diese Parameter sind z.B. Orientierung, Fokussierung, Zoom, Blende und in einem binokularen Kamerasystem Vergenz. Durch Steuerung dieser Parameter ist das System in der Lage, robustere und schnellere Lösungen zu erhalten. Im weiteren Sinne beschäftigt sich das aktive Sehen mit der Aufmerksamkeit eines Systems, wobei diese als ein selektives Begreifen des Raumes, der Auflösung und der Zeit zu verstehen ist.

Somit ist die maschinelle Aufmerksamkeit vergleichbar mit der Aufmerksamkeit des menschlichen okulomotorischen Systems, welches verschiedene Mechanismen zur Steuerung der Augenbewegungen besitzt. Das menschliche visuelle System reagiert auf bestimmte visuelle Stimuli, wie etwa das Schließen der Iris bei großem Lichteinfall auf die Retina. Ein aktives Verhalten des Menschen erzeugt Stimuli, die die visuelle Wahrnehmung optimieren. Dies sind z.B. die Vergenzbewegungen zur Einschränkung der Stereodisparität und der optokinetische Reflex zur Bildstabilisierung [DHKS95].

Durch die maschinelle Aufmerksamkeitssteuerung des aktiven Sehens ist es möglich, einfachere Regelungsalgorithmen und Repräsentationen von Szenen zu implementieren. Dies führt zu einer Verkürzung der Reaktionszeit eines Systems und sichert außerdem einen größeren Erfolg bei der Ausführung geplanter Aktionen zu [SS93].

Die Forschung auf diesem Gebiet wurde erst in den letzten Jahren durch die Schaffung der entsprechenden Hardware (leichtgewichtige Kameras, kameraunabhängige Steuerung der Motoren und Rechner zur Bildverarbeitung in Echtzeit) vorangetrieben. Die weite Verbreitung von experimentellen Anwendungen ist ein deutliches Zeichen für das steigende Interesse, das diesem Gebiet zuteil wird.

1.2 Binokulare Kameraplattformen

Eine besondere Rolle spielen in diesem Zusammenhang die binokularen Kameraplattformen. Die erste ihrer Art, die in der Lage war, Bewegungen eines Objekts durch Fixieren desselben im Bildzentrum beider Kameras zu verfolgen, wurde in den frühen achtziger Jahren an der Universität von Pennsylvania, USA, hergestellt [SS93]. Aufgrund der konstruktiven Nähe zur Anatomie des menschlichen visuellen Apparats bezeichnet man eine solche Plattform auch als Kamerakopf.

Heutzutage verfügen viele Forschungsgruppen über solche Kameraplattformen, um aktives Sehen zu erforschen und anzuwenden. Die Plattform der Forschungsgruppe des Lehrstuhls für Kognitive Systeme der Christian-Albrechts-Universität zu Kiel gehört zu den ersten, welche diese Aufgabe in Video-Echtzeit bewältigt [DHKS95].

Ein Ziel des Einsatzes binokularer Kameraplattformen besteht darin, aus den zwei verschiedenen zweidimensionalen Abbildern des dreidimensionalen Raumes auf die Entfernung von Objekten im Sichtbereich des Systems zu schließen. Ein besonderer Vorteil einer solchen Plattform gegenüber einem monokularen System zur Gewinnung von Tiefeninformationen besteht darin, daß die Bildaufnahmen zeitgleich geschehen. Wird nur eine Kamera (z.B. montiert an einen Robotergreifer) benutzt, um aus zwei Perspektiven dieselbe Szene aufzunehmen, so besteht zwischen den Aufnahmen eine zeitliche Differenz. Diese kann zu großen Fehlern in der Tiefenberechnung führen, falls sich Objekte innerhalb der Szene bewegen. Ein solches System würde weiterhin von Verzögerungen beeinträchtigt.

1.3 Kalibrierung

Für das maschinelle Sehen ist eine Kalibrierung von Kameras notwendig. In diesem Zusammenhang ist die Kalibrierung der Prozeß der Bestimmung der internen Kameraparameter und der externen Parameter des Systems. Die internen Parameter beschreiben die Relation zwischen Bildkoordinaten auf der Sensorfläche einer Kamera und Richtungen von Strahlen im dreidimensionalen Koordinatensystem der Kamera. Die externen Parameter sind die Parameter einer Robotereinheit, die die Transformation von dem Kamerakoordinatensystem auf ein

Koordinatensystem beschreiben, das, bezogen auf die Welt oder den Roboter, fest ist.

In den neunziger Jahren wurden viele Verfahren beschrieben, die visuelle Probleme ohne Wissen über die internen Kameraparameter lösen. Dennoch bleibt eine ausreichende Anzahl an Problemen übrig, die metrische Informationen benötigen, um eine Abbildung zwischen den Pixelkoordinaten der Sensorfläche und den dreidimensionalen Strahlen herzustellen. Diese Arbeit beschäftigt sich hauptsächlich mit der Kalibrierung der internen Parameter.

Die externen Parameter teilt man in zwei Gruppen ein. Die erste Gruppe besteht aus den Parametern, die die Transformation zwischen einem Szenen- und dem Kamerakoordinatensystem beschreiben. Die zweite Gruppe besteht aus den Parametern, die die Transformation zwischen dem Kamerakoordinatensystem und verschiedenen Koordinatensystemen einer Einheit beschreiben, auf welche die Kamera montiert ist. Die Kalibrierung letztgenannter Parameter wird als Sensor-Effektor- oder Hand-Auge-Kalibrierung bezeichnet.

Die Kalibrierung der externen Parameter zwischen Szenen- und Kamerakoordinatensystem ist u.a. von Faugeras [Fau93] gelöst worden. In dieser Arbeit wird ebenfalls auf die Kalibrierung dieser Parameter eingegangen. Die Hand-Auge Kalibrierung wird u.a. von Wang [Wan92], Young et al. [YHHY92], Tsai und Lenz [TL88], Horaud et al. [HDM94], Chen [Che91] und Li et al. [LB95] untersucht. Auf diese Kalibrierung wird hier nicht eingegangen.

Manche Kameras verfügen über den Freiheitsgrad der Brennweite. Die Kalibrierung einer solchen Zoomfunktion wird im allgemeinen gesondert durchgeführt und daher in dieser Arbeit nicht behandelt. Diese Kalibrierung wurde u.a. von Tarabanis et al. [TTG94] und von Li und Lavest [LL95] untersucht.

1.4 Einführendes Beispiel

Zur Einführung wird im folgenden auf die Kalibrierung von zwei Kameras eingegangen, die zur Entfernungsberechnung eines Objekts herangezogen werden. Im besonderen wird gezeigt, wie sich Fehler in der Kalibrierung auf diese Entfernungsberechnung auswirken.

Für dieses Beispiel betrachte man zwei Kameras, die im Raum so aufgestellt sind, daß folgende Annahmen gelten. Die optischen Achsen beider Kameras liegen in einer Ebene und besitzen einen Schnittpunkt im dreidimensionalen Raum. Beide Kameras besitzen je eine Rotationsachse um das optische Zentrum. Diese Rotationsachsen seien parallel zueinander und stehen senkrecht zu den jeweiligen optischen Achsen.

Fixiert man ein Objekt so, daß eine Ecke dieses Objekts in beiden Kameras im Bildmittelpunkt gehalten wird, so kann man die Entfernung der Ecke zur Kamera durch $Z = \frac{-b cos\theta_L cos\theta_R}{sin\theta}$ berechnen, wobei b die Entfernung der optischen Zentren,

θ_L den Rotationswinkel der linken, θ_R den Rotationswinkel der rechten Kamera und θ den Winkel zwischen den optischen Achsen angibt. Die Entfernung Z ist die Entfernung in einem Weltkoordinatensystem, dessen Ursprung zwischen den beiden optischen Zentren liegt.

Man nehme an, daß die Entfernung der optischen Zentren durch $b = 20cm$ gegeben ist. Die Winkel betragen $\theta_L = 3°$, $\theta_R = 3°$ und $\theta = 6°$. Nach obiger Formel wird das Objekt somit in einer Entfernung $Z \approx 191cm$ fixiert.

Erhält man durch Übernahme einer fehlerhaften Angabe des Herstellers oder Durchführung einer ungenauen Kalibrierung beispielsweise die Größen $b = 19cm$, $\theta_L = 3.5°$, $\theta_R = 3.5°$ und $\theta = 7°$, so erhält man aus der Tiefenformel $Z \approx 155cm$. Der Fehler in der Tiefenberechnung beträgt somit $\approx 36cm$ bzw. $\approx 23\%$.

Ist ein solcher Algorithmus in einen autonomen Roboter eingebettet, der nach einer Entfernungsmessung zu einem Gegenstand fahren soll, so wird dieser den Gegenstand nicht erreichen, da er den Gegenstand in einer kleineren Entfernung erwartet hat.

1.5 Ziel der Arbeit

In dieser Arbeit wird auf unterschiedliche Möglichkeiten der Kalibrierung der internen Kameraparameter eingegangen. Das vorrangige Ziel dieser Arbeit besteht darin, eine Kalibriermethode für aktive Kamerasysteme zu entwickeln, die auf eine mobile Robotereinheit montiert sind.

Das Ausnutzen der Freiheitsgrade, die eine solche Einheit besitzt, führt zu zusätzlichen Informationen, die bei der Kalibrierung einer stationären Kamera nicht zur Verfügung stehen. Der Einsatz einer aktiven Kamera auf einer mobilen Einheit führt zu Anforderungen an den Kalibrierprozeß, die durch traditionelle Verfahren unter Ausnutzung von dreidimensionalen Daten bekannter Objekte nicht gelöst werden können.

Da aktive Kameras im allgemeinen ständig ihre Zoom- und Fokuscharakteristik verändern, werden zur Kalibrierung Verfahren benötigt, die eine wiederholte Kalibrierung während des Betriebs der Robotereinheit durchführen können. Somit darf das Verfahren nicht von Merkmalen abhängen, deren Lage als bekannt vorausgesetzt wird.

Das Verfahren sollte in der Lage sein, die Kalibrierung auszuführen, ohne die Robotereinheit stoppen zu müssen, und es sollte Bildmerkmale in unterschiedlichen Bildern einander zuordnen können. Die Kalibrierung der internen Parameter sollte von denen der externen entkoppelt durchgeführt werden, da dieser Teil der Kalibrierung von einem festen Weltkoordinatensystem abhängt.

Das Ziel einer aktiven Kalibrierung ist es nicht, ähnlich exakte Ergebnisse zu berechnen, wie sie bei der Kalibrierung stationärer Kameras unter Zuhilfenahme eines Referenzobjektes erzielt werden, sondern vielmehr eine Kalibrierung zu erhalten, die während des aktiven Betriebs der Kameras eine stabile Rekalibrierung realisiert. Im folgenden wird ein solches Referenzobjekt mit Kalibrierkörper bezeichnet werden.

Im Sinne der Vorgehensweise von verhaltensbasierten Systemen wird hier eine allgemeine Lösung zu einer visuellen Aufgabe angestrebt, bei der das Zusammenspiel zwischen Genauigkeit und Anforderungen an die Umwelt den bestmöglichen Kompromiß bildet.

Es ist nicht zu erwarten, daß das angestrebte Verfahren vergleichbar gute Ergebnisse liefert, wie beispielsweise das Verfahren von Tsai [Tsa86], da Tsai die exakt bekannten dreidimensionalen Daten von mehreren hundert Punkten voraussetzt.

Um diesen Anforderungen gerecht zu werden, wird in dieser Arbeit ein neuer Ansatz zur Kalibrierung der internen Parameter einer Kamera präsentiert, für den lediglich folgendes angenommen wird:

- Die einzige Voraussetzung an die Umgebung ist die Existenz einer parallelen Geradenschar.
- Die Kamera muß in der Lage sein, um zwei Achsen des Kamerakoordinatensystems zu rotieren (Neigungs- und Vergenzachse).

Zusätzlich zu diesem Verfahren sollen weitere vorgestellt und untersucht werden. Anschließend sollen sämtliche Verfahren implementiert und durch einen systematischen Vergleich voneinander abgegrenzt werden. Dazu wird detailiert auf die unterschiedlichen Voraussetzungen der einzelnen Verfahren eingegangen. Es werden sowohl Testreihen mit synthetischen Daten als auch durchgeführte Realexperimente präsentiert.

Die Verfahren, die in dieser Arbeit untersucht werden, besitzen verschiedene Eingabeparameter. Um diese Parameter aus Bilddaten zu extrahieren, wird insbesondere auf die Geradendetektion mit der Hough-Transformation [Hou62] eingegangen. Einige der Verfahren benutzen Kalibrierkörper, andere nur Bild- oder Fluchtpunkte zur Kalibrierung. Läßr man eine Kamera um spezielle Achsen rotieren und nimmt so eine Bildfolge auf, so enstehen Hyperbeln. Diese sind durch mehrere Bild- oder Fluchtpunkte gegeben. Ein weiteres Ziel dieser Arbeit ist es, Verfahren zur Berechnung dieser Hyperbelparameter zu untersuchen.

1.6 Aufbau der Arbeit

Im zweiten Kapitel wird auf die der Kalibrierung zugrundeliegende Geometrie eingegangen und das zu kalibrierende Kamerasystem beschrieben. Es werden grundlegenden Definitionen vorgenommen, die projektive Geometrie eingeführt und die Kinematik der Kameraplattform, die Zentralprojektion und die Systemarchitektur des zu kalibrierenden Systems beschrieben.

Das dritte Kapitel befaßt sich mit der Kalibrierung der internen Kameraparameter. Es werden unterschiedliche Methoden zur Kalibrierung von Kameras besprochen. Es werden auf bekannten Verfahren beruhende Methoden besprochen, und es wird ein Verfahren entwickelt, das den Anforderungen eines autonomen Robotersystems genügt. Die unterschiedlichen Kalibrierverfahren zur Bestimmung der internen Kameraparametern werden klassifiziert.

Da die meisten Verfahren zur Kalibrierung der internen Parameter Referenzmerkmale benutzen, wird im vierten Kapitel auf die Gewinnung solcher Merkmale eingegangen. Dazu wird auf unterschiedliche Kalibrierkörper eingegangen, und es werden Möglichkeiten besprochen, Referenzmerkmale aus Bildern zu gewinnen.

Im fünften Kapitel werden die Ergebnisse von Testreihen mit synthetischen Daten und Ergebnisse von Realexperimenten untersucht und mit anderen Arbeiten verglichen. Dazu wird ein systematischer Vergleich der Methoden durchgeführt und auf die Einflußnahme der Aufnahmegeometrie auf diese Methoden eingegangen.

Im sechsten Kapitel wird eine Diskussion der vorgestellten Methoden und ein Ausblick folgen. Im Anhang werden die implementierten Algorithmen beschrieben.

Kapitel 2

Geometrie und Architektur

Eine binokulare Kameraplattform besteht aus mehreren Bauteilen, die insgesamt eine Einheit bilden (Abb 2.1). Die Basis der Plattform besteht aus einem Bauteil ohne Freiheitsgrade. Die übrigen Bauteile, ein Hals, ein Verbindungsstück und zwei Kameras, sind auf die Basis montiert. Der Hals ist auf der Basis angebracht und besitzt den Freiheitsgrad einer Drehbewegung um eine Achse. Diese Achse wird als Schwenkachse (pan axis) bezeichnet. Das Verbindungsstück ist auf dem Hals angebracht und besitzt ebenfalls den Freiheitsgrad einer Drehbewegung um eine Achse. Diese Achse wird als Neigungsachse (tilt axis) bezeichnet.

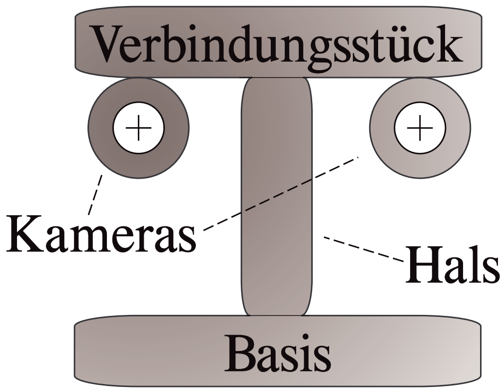


Abbildung 2.1: Bauteile einer binokularen Kameraplattform

Die Schwenk- und die Neigungsachse besitzen einen Schnittpunkt. Beide Kameras sind an das Verbindungsstück montiert und besitzen wiederum den Freiheitsgrad einer Drehbewegung um eine Achse. Die Achsen der Kameras bezeichnet man als Vergenzachsen. Beide Vergenzachsen besitzen jeweils einen Schnittpunkt mit der Neigungsachse.

Binokulare Kameraplattformen sind in ihrer Gestaltung an die Anatomie des menschlichen visuellen System angelehnt:

- Die zwei Kameras entsprechen den Augen.
- Bewegungen um die Vergenzachsen entsprechen horizontalen Bewegungen der Augen. Dadurch ist ein Fixieren von Objekten möglich.
- Bewegungen um die Neigungsachse entsprechen einer vertikalen Nackenbewegung.
- Bewegungen um die Schwenkachse entsprechen einer horizontalen Halsbewegung.

Daher werden binokulare Kameraplattformen auch als Kameraköpfe bezeichnet. Die verwendeten Kameras sind unabhängig von der Definition der Kameraplattform und können über weitere Freiheitsgrade wie Zoom und Fokus verfügen.

In diesem Kapitel werden grundsätzliche Definitionen aufgestellt und auf die projektive Geometrie, die Zentralprojektion und die Kinematik einer Kameraplattform eingegangen.

Die Notationen, die in diesem Kapitel und auch in den folgenden Kapiteln verwendet werden, gehen überwiegend auf Gengenbach [Gen94] und Daniilidis et al. [DHKS95] zurück.

2.1 Definition der Koordinatensysteme

Sämtliche Koordinatensysteme, die in dieser Arbeit verwendet werden, sind rechtshändige Koordinatensysteme. Die x-Achsen aller Koordinatensysteme der binokularen Plattform verlaufen in horizontaler Richtung, die y-Achsen in vertikaler Richtung zu der Fläche, auf der die Kameraplattform angebracht ist. Des weiteren besitzen alle entsprechenden Achsen in einem globalen Weltkoordinatensystem denselben Richtungsvektor.

Koordinatensysteme, die für eine Kamera definiert werden, sind alternativ für die rechte und für die linke Kamera einer binokularen Kameraplattform gültig. Der Übersichtlichkeit halber wird hier auf eine sperate Darstellung verzichtet und die Koordinatensysteme für eine Kamera definiert.

- **Die Rechnerkoordinatensysteme**
  Nomenklatur : $\boldsymbol{P_p} = (x_p, y_p)^T$ in *Pixel*
  Die Koordinaten des Rechnerkoordinatensystems sind durch die im Bildspeicher abgelegte Grauwertmatrix gegeben. Dazu sei das Bildformat durch $x_{p_{\max}} \times y_{p_{\max}}$ Bildpunkte und die Startadresse a_0 im Bildspeicher gegeben. Alle diese Konstanten seien dabei aus dem Bereich der natürlichen Zahlen. Der Wert eines Punktes $(x_p, y_p)^T$ im Bildspeicher berechnet sich wie folgt durch die Funktion Υ, die den Wert einer Speicheradresse liefert:

  $$\Upsilon : (\mathbb{N}_+, \mathbb{N}_+) \rightarrow \mathbb{N}_+,\ (x_p, y_p)^T \mapsto \mathrm{a}_0 + y_p x_{p_{\max}} + x_p$$

  mit $0 \leq y_p < y_{p_{\max}}$ und $0 \leq x_p < x_{p_{\max}}$.

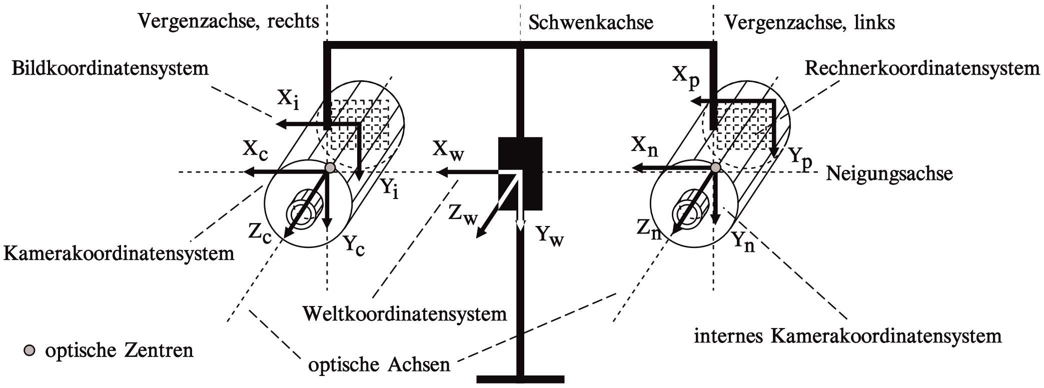


Abbildung 2.2: Die Koordinatensysteme der binokularen Kameraplattform.

- **Die Bildkoordinatensysteme**
  Nomenklatur : $\boldsymbol{P_i} = (x_i, y_i)^T$ in mm
  Die Ursprünge dieser Koordinatensysteme liegen in den Schnittpunkten der optischen Achsen mit den Bildebenen. Die beiden Achsen der Koordinatensysteme beschreiben die Bildebenen, die den sensitiven Oberflächen der CCD-Bildaufnehmer in den Kameras entsprechen. Die x-Achsen der Koordinatensysteme liegen parallel zu den Pixelzeilen, die y-Achsen parallel zu den Pixelspalten.

- **Die internen Kamerakoordinatensysteme**
  Nomenklatur : $\boldsymbol{P_n} = (x_n, y_n, z_n)^T$ in mm
  Die Ursprünge dieser Koordinatensysteme liegen in den optischen Zentren der durch die Kameras realisierten Abbildungen. Die z-Achsen der Koordinatensysteme entsprechen den optischen Achsen der Kameras und zeigen von der Bildebene in Richtung Szene.

- **Die Kamerakoordinatensysteme**
  Nomenklatur : $\boldsymbol{P_c} = (x_c, y_c, z_c)^T$ in mm
  Die Ursprünge dieser Koordinatensysteme liegen in den Schnittpunkten der Neigungsachse mit den Vergenzachsen. Die z-Achsen der Koordinatensysteme zeigen von der Bildebene in Richtung Szene. Diese Koordinatensysteme sind von den Rotationen um jeweils drei der vier Achsen der Plattform abhängig.

- **Die lokalen Referenzkoordinatensysteme der Kameras**
  Nomenklatur : $\boldsymbol{P_r} = (x_r, y_r, z_r)^T$ in mm
  Die Ursprünge dieser Koordinatensysteme liegen ebenfalls in den Schnittpunkten der Schwenkachse mit den Vergenzachsen. Diese Koordinatensysteme sind allerdings nur von einer Rotation um die Schwenkachse abhängig.

Somit entsprechen diese Koordinatensysteme den Kamerakoordinatensystemen genau dann, wenn sich die Neigungsachse und die Vergenzachsen in der Ruheposition befinden.

- **Das Effektorkoordinatensystem**
  Nomenklatur : $\boldsymbol{P_e} = (x_e, y_e, z_e)^T$ in mm
  Der Ursprung dieses Koordinatensystems wird durch den Schnittpunkt der Schwenk- und Neigungsachse der Kameraplattform festgelegt. Die z-Achse des Koordinatensystems zeigt orthogonal zur Neigungs- und Schwenkachse in Richtung Szene. Dieses Koordinatensystem ist von den Rotationen um die Schwenkachse abhängig.

- **Das Weltkoordinatensystem**
  Nomenklatur : $\boldsymbol{P_w} = (x_w, y_w, z_w)^T$ in mm
  Der Ursprung dieses Koordinatensystems wird durch den Schnittpunkt der Schwenk- und Neigungsachse der Kameraplattform festgelegt. Die z-Achse des Koordinatensystems zeigt orthogonal zur Kameraplattform in Richtung Szene. Wenn sich die Schwenkachse in Ruhelage befindet, entspricht dieses Koordinatensystem dem Effektorkoordinatensystem.

- **Das Szenenkoordinatensystem**
  Nomenklatur : $\boldsymbol{P_s} = (x_s, y_s, z_s)^T$ in mm
  Dieses Koordinatensystem wird durch ein Objekt festgelegt, welches sich in dem Sichtbereich der Kameraplattform befindet.

Die Rechner-, Bild-, Kamera- und internen Kamerakoordinatensysteme sind sich bewegende Koordinatensysteme, da sie von sämtlichen Bewegungen der Kameraplattform abhängig sind. Die lokalen Referenzkoordinatensysteme der Kameras und das Effektorkoordinatensystem sind ebenfalls bewegte Koordinatensysteme. Ihre Bewegung ist nur von der Schwenkachse der Plattform abhängig. Das Weltkoordinatensystem und das Szenenkoordinatensystem sind feststehende Koordinatensysteme, die nur durch die Lage der Plattform bzw. eines Objekts bestimmt werden.

2.2 Projektive Geometrie

Im Gegensatz zur affinen Geometrie stehen in der projektiven Geometrie keine metrischen Aspekte, wie Längen und Winkel, sondern Kollinearität von Punkten und Kopunktalität von Geraden im Vordergrund [Kol95], [Sch90], [Fis94]. Eine ausführliche Darstellung der projektiven Geometrie für die Bildauswertung wird durch Faugeras [Fau93] gegeben. Diese Einführung orientiert sich an Kollnig [Kol95]. In dieser Arbeit werden homogene Koordinatenvektoren stets mit eckigen, inhomogene Koordinatenvektoren mit runden Klammern angegeben.

2.2.1 Homogene Koordinaten von Punkten in der Ebene

Jeder Punkt $\boldsymbol{P}$ eines n-dimensionalen projektiven Raumes wird durch einen homogenen $n+1$-dimensionalen Koordinatenvektor $[m_1, \cdots, m_{n+1}]^T$ beschrieben. Mindestens eine der $n+1$ Koordinaten muß von Null verschieden sein. Ein Punkt mit den homogenen Bildkoordinaten $[m_1, \cdots, m_{n+1}]^T$ wird in dem Fall

- $m_{n+1} \neq 0$ mit dem eigentlichen Punkt $(\frac{m_1}{m_{n+1}}, \cdots, \frac{m_n}{m_{n+1}})^T$ identifiziert, wobei $x_1 = \frac{m_1}{m_{n+1}}, \cdots, x_n = \frac{m_n}{m_{n+1}}$ inhomogene Koordinaten genannt werden,
- $m_{n+1} = 0$ mit einem im Unendlichen liegenden uneigentlichen Punkt oder Fernpunkt identifiziert.

Die Gerade durch den Ursprung des projektiven Raumes und einen Punkt in homogenen Koordinaten wird als Sehstrahl des Punktes bezeichnet. Durch Einführung der homogenen Koordinaten ist es möglich, Bildkoordinaten über Sehstrahlen in dreidimensionale interne Kamerakoordinaten abzubilden. Ein Bildpunkt $(x_i, y_i)^T$ wird durch den zugehörigen Sehstrahl auf den Punkt $(x_c, y_c, z_c)^T$ abgebildet, der in dem Kamerakoordinatensystem liegt (Abb. 2.3). Dieser Sachverhalt beschreibt die Zentralprojektion, die der Bildaufnahme zugrunde liegt und auf welche in Kapitel 2.3.2 noch näher eingegangen wird.

Der Zusammenhang zwischen Punkten in homogenen Koordinaten und der Zentralprojektion spiegelt sich hauptsächlich in den zwei folgenden Aussagen wieder:

- Betrachtet man den homogenen Koordinatenvektor eines Punktes im Bildkoordinatensystem, so folgt aus der Eigenschaft der Zentralprojektion, daß sich das Urbild dieses Bildpunktes im Szenenkoordinatensystem auf dem Sehstrahl befindet.
- Aus der Definition der homogenen Koordinaten folgt sofort, daß der zu einem Punkt im Bildkoordinatensystem mit den inhomogenen Koordinaten $(x_i, y_i)^T$ gehörende homogene Koordinatenvektor $[x_i, y_i, z_i]^T$ und der Richtungsvektor des Sehstrahls des Bildpunktes linear abhängig sind.

2.2.2 Homogene Koordinaten von Geraden in der Ebene

Eine Gerade wird durch ein vom Nullvektor verschiedenes homogenes Koordinatentupel $[l_1, \cdots, l_{n+1}]^T$ definiert. Eine Gerade mit den homogenen Koordinaten $[l_1, \cdots, l_{n+1}]^T$ stellt im Fall

- $l_i \neq 0$ für mindestens ein $i \in \mathbb{N}$ mit $0 < i \leq n$ die eigentliche Gerade $l_1 x_1 + \cdots + l_n x_n + l_{n+1} = 0$ dar,
- $l_i = 0$ für alle $i \in \mathbb{N}$ mit $0 < i \leq n$ eine uneigentliche Gerade im Unendlichen, die Ferngerade, dar.

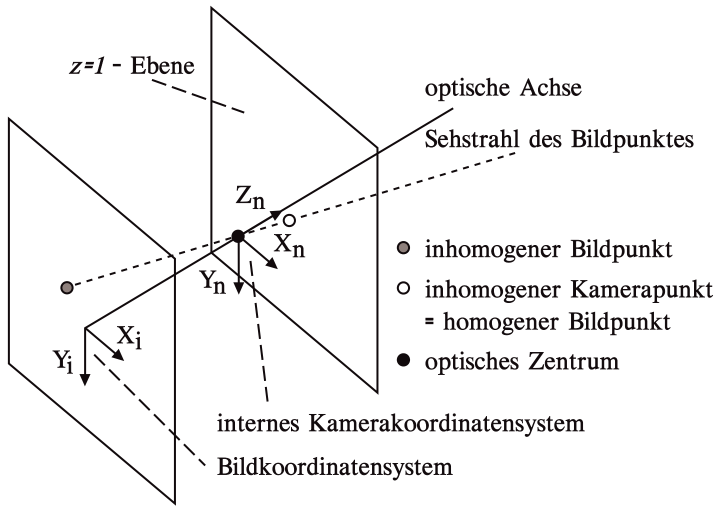


Abbildung 2.3: Veranschaulichung von homogenen Koordinaten

Als die Menge aller Fernpunkte ergänzt die Ferngerade die affine Ebene [Fis94] zur projektiven Ebene. Die Ebene durch den Ursprung und eine Gerade heißt Projektionsebene der Geraden. Der Zusammenhang zwischen Geraden in homogenen Koordinaten und der Zentralprojektion kann wie folgt ausgedrückt werden:

1. Aus der Eigenschaft der Zentralprojektion folgt, daß die Projektionsebene einer Bildgeraden das Urbild der Geraden im Raum enthält.

2. Des weiteren folgt aus der Definition der projektiven Geraden, daß der zu einer Bildgeraden $ax + by + c = 0$ gehörende homogene Koordinatenvektor $(a, b, c)^T$ zugleich ein Normalenvektor der Projektionsebene der Geraden ist.

Kennzeichnend für die projektive Geometrie ist das Dualitätsprinzip, welches besagt, daß man Punkte durch Geraden bzw. Geraden durch Punkte ersetzen kann und nach wie vor eine projektive Ebene erhält. Somit entsprechen die Definitionen und Sätze zu homogenen Punktkoordinaten denen der homogenen Koordinatenvektoren.

Weiterhin gilt, daß der dreidimensionale Richtungsvektor einer Raumgeraden orthogonal zum homogenen Koordinatenvektor der zugehörigen Bildgeraden ist. Dies folgt mit 2. aus der Tatsache, daß nach 1. die dreidimensionale Raumgerade, die das Urbild der Bildgeraden ist, in der Projektionsebene der Bildgeraden liegt. Aus diesem Zusammenhang folgt sofort, daß der Normalenvektor der Projektionsebene einer Bildgeraden orthogonal zur Raumgeraden ist.

Homogene Koordinaten werden weiterhin benutzt, um affine Abbildungen zu linearisieren. Sind beispielsweise $\boldsymbol{x_1}$, $\boldsymbol{x_2}$ und $\boldsymbol{t}$ Vektoren aus dem $\mathbb{R}^n$ mit $n \in \mathbb{N}$

und R eine $(n \times n)$-Rotationsmatrix, so läßt sich die affine Abbildung $\boldsymbol{x_1} = R\boldsymbol{x_2} + \boldsymbol{t}$ in die lineare Abbildung

$$\begin{bmatrix} \boldsymbol{x_1} \\ 1 \end{bmatrix} = \begin{pmatrix} R & \boldsymbol{t} \\ 0^T & 1 \end{pmatrix} \begin{bmatrix} \boldsymbol{x_2} \\ 1 \end{bmatrix}$$

transformieren. Diese Transformation hat den Vorteil, affine Abbildungen durch Matrizenoperationen darstellen zu können. Dazu wird ein n-dimensionaler Vektor um einen Index auf einen $(n+1)$-dimensionalen Vektor erweitert und der letzte Index des erweiterten Vektors mit dem Wert 1 belegt.

2.3 Geometrische Beschreibung

In diesem Abschnitt wird zunächst die Kinematik einer binokularen Kameraplattform beschrieben. Anschließend wird die Zentralprojektion als Abbildung einer dreidimensionalen Szene auf die zweidimensionale Abtastfläche einer Kamera eingeführt. Zuletzt wird auf die Transformation zwischen Szene und Kamera eingegangen.

2.3.1 Kinematik der Kameraplattform

In diesem Unterabschnitt wird die direkte und inverse Abbildung von einem Referenzkoordinatensystem, dem Weltkoordinatensystem der Kameraplattform, zu den Kamerakoordinatensystemen beschrieben. Der Übersichtlichkeit halber wird erneut auf eine separate Darstellung in den Koordinaten beider Kameras verzichtet.

Es wird im folgenden von der Annahme ausgegangen, daß die optischen Zentren identisch mit den Schnittpunkten der Vergenz- und Neigungsachsen sind und daß die Bildebene und die $(z = 0)$-Ebene des Kamerakoordinatensystems parallel sind. Diese Annahme ist eine Vereinfachung. Die Sensor-Effektor- oder Hand-Auge- Kalibrierung bildet insbesondere ein selbständiges Problem, wenn die Position des Projektionszentrums von der Steuerung der Brennweite und der Fokussierung verstellt werden kann. Ein weiteres Problem ist durch eine Drehung des Aufnahmesensors oder der verwendeten Kameras gegeben, da im allgemeinen nicht davon ausgegangen werden kann, daß die Achsen des CCD-Chips parallel zu den Achsen des Gelenks verlaufen. Da in dieser Arbeit die Zoom-Funktion nicht kalibriert wird, kann das optische Zentrum so eingestellt werden, daß keine Translation zwischen den internen Kamerakoordinatensystemen und den Kamerakoordinatensystemen vorhanden ist. Aus diesem Grund werden im folgenden die internen Kamerakoordinatensysteme mit den Kamerakoordinatensystemen gleichgesetzt:

$$\begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} = \begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} . \tag{2.1}$$

Die binokulare Kameraplattform besitzt vier Freiheitsgrade, beschrieben durch den Schwenkwinkel χ, den Neigungswinkel ϕ und die Vergenzwinkel θ für die linke und rechte Kamera (Abb. 2.4). b bezeichnet die Basislänge der Kamera-

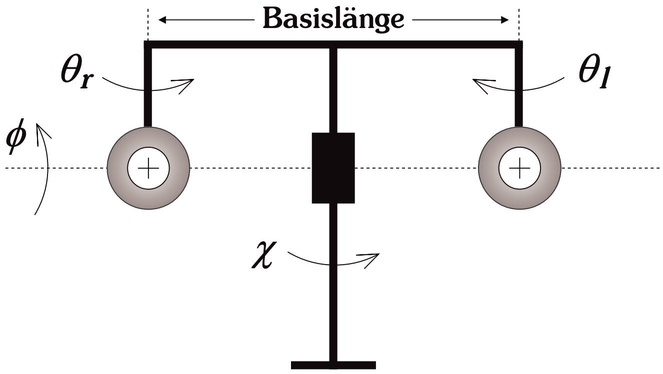


Abbildung 2.4: Die vier Freiheitsgrade einer Kameraplattform (aus [DHKS95])

plattform, also den Abstand der Bildzentren beider Kameras. Es wird angenommen, daß die Vergenzachsen die Entfernung $b/2$ von der Schwenkachse haben. Auch diese Annahme ist nur angenähert. Die Abweichungen sind so gering, daß sie keine Rolle spielen. Somit kann die Transformation der homogenen Kamerakoordinaten in homogene Weltkoordinaten wie folgt beschrieben werden:

$$\begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = T_\chi T_\phi T_\theta \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} , \tag{2.2}$$

wobei

$$T_\chi = \begin{pmatrix} \cos\chi & 0 & \sin\chi & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\chi & 0 & \cos\chi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , T_\phi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi & 0 \\ 0 & \sin\phi & \cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} ,$$

und

$$T_\theta = \begin{pmatrix} \cos\theta & 0 & \sin\theta & -\frac{b}{2} \\ 0 & 1 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} .$$

Im folgenden werden die normierten Bildkoordinaten in den Kamerakoordinatensystemen betrachtet. Ein normiertes Bildkoordinatentupel beschreibt den Schnittpunkt des Sehstrahls eines Bildpunktes mit der $z = 1$-Ebene des Kamerakoordinatensystems. Die Darstellung dieses Sehstrahls in den Kamerakoordinatensystemen ist $\lambda[x_i, y_i, 1]^T$ mit $\lambda \in \mathbb{R}_{\neq 0}$. Bildet man letztere Sehstrahlen auf Sehstrahlen $\lambda(x_r, y_r, 1)^T$ mit $\lambda \in \mathbb{R}_{\neq 0}$ in den lokalen Referenzkoordinatensystemen der Kameras ab, so gilt für diese Abbildung:

$$\lambda \begin{pmatrix} x_r \\ y_r \\ z_r \end{pmatrix} = R_\phi R_\theta \begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} \ , \ \lambda \in \mathbb{R}_{\neq 0} \ , \tag{2.3}$$

wobei

$$R_\phi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix} \text{ und } R_\theta = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} .$$

Diese Abbildung stellt eine projektive Kollineation zwischen Strahlenbündeln dar. Man erkennt daraus die Tatsache, daß sich jede Rotation eines Koordinatensystems im Gegensatz zu einer Translation unter Kenntnis des Winkels und ohne Wissen über die Tiefen der abgebildeten Punkte kompensieren läßt. Aus der Vektordarstellung von Gleichung (2.3) läßt sich λ eliminieren. Man erhält

$$\begin{aligned} x_r &= \frac{\cos\theta x_c + \sin\theta}{-\cos\phi\sin\theta x_c + \sin\phi y_c + \cos\phi\cos\theta} \ , \\ y_r &= \frac{\sin\phi\sin\theta x_c + \cos\phi y_c - \sin\phi\cos\theta}{-\cos\phi\sin\theta x_c + \sin\phi y_c + \cos\phi\cos\theta} \ . \end{aligned} \tag{2.4}$$

Mittels dieser Beziehung läßt sich das direkte kinematische Problem beschreiben. Wird ein Sehstrahl $\boldsymbol{P_r} = (x_r, y_r, 1)^T$ als Stellgröße für die optische Achse $(0,0,1)^T$ der Kamera angegeben, so erhält man die Lösung des inversen kinematischen Problems:

$$\tan\phi = -y_r \text{ und } \tan\theta = \frac{x_r}{\sqrt{1 + {y_r}^2}} \ . \tag{2.5}$$

2.3.2 Beschreibung der Zentralprojektion

Die Abbildung der Kameras wird durch die Zentralprojektion beschrieben. Ein durch mehrere Linsen zusammengesetztes Linsensystem bzw. Objektiv läßt sich durch die Bildbrennweite und zwei Hauptebenen beschreiben, sofern es sich um ein zentriertes System handelt. Unter einem zentrierten System versteht man in diesem Zusammenhang, daß sich die Krümmungsmittelpunkte aller Linsenoberflächen auf der optischen Achse befinden müssen.

Da sich in den hier verwendeten Objektiven vor und hinter den Linsen das gleiche Medium (Luft) befindet, sind Objektbrennweite und Bildbrennweite gleich. Wenn r_1 und r_2 die Krümmungsradien einer Linse sind, gilt für symmetrische

Linsen aus Glas in Luft, daß die Brennweite f gleich dem Krümmungsradius ist ($f = r_1 = r_2$). Alle Parameter werden in mm angegeben.

Für ein zusammengesetztes symmetrisches Linsensystem mit gemeinsamer optischer Achse und den Brennweiten f_1 und f_2 gilt für die Gesamtbrennweite f:

$$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{d}{f_1 f_2} \ , \tag{2.6}$$

wobei d der Abstand der beiden Linsenebenen in mm ist. Verändert man also die Abstände zwischen den Linsen eines Objektivs, so erhält man unterschiedliche Brennweiten. Besitzt eine Kamera den Freiheitsgrad der Brennweite, so bezeichnet man diesen als Zoomen.

Als Hauptebene 1 bezeichnet man die Linsenebene eines zusammengesetzten Systems, die der Szene am nächsten ist. Die Linsenebene, die der Aufnahmefläche am nächsten ist, wird als Hauptebene 2 bezeichnet. Die Entfernung zwischen einem Gegenstand und der Hauptebene 1 wird als Gegenstandsweite bezeichnet. Der Abstand der Bildebene (Sensorfläche) von der Hauptebene 2 wird als Bildweite bezeichnet. Auch diese beiden Parameter werden in mm angegeben. Im allgemeinen wird die Bildweite so eingestellt, daß Abbildungsstrahlen eines Punktes in der Szene sich in einem Punkt auf der Bildebene schneiden (Schärfeebene).

Die Bildgröße bezeichnet den Abstand eines Bildpunktes von der optischen Achse. Die Gegenstandsgröße bezeichnet den Abstand eines Gegenstandspunktes von der optischen Achse.

Die Schnittpunkte der beiden Hauptebenen mit der optischen Achse werden Hauptpunkte genannt. Der in der Photogrammetrie eingeführte Begriff des Bildhauptpunktes bezeichnet den Schnittpunkt der optischen Achse mit der Bildebene. Der Abstand zwischen Bildebene und Gegenstand ergibt sich aus der Summe von Gegenstandsweite, Bildweite und dem Abstand der beiden Hauptebenen (Abb. 2.5).

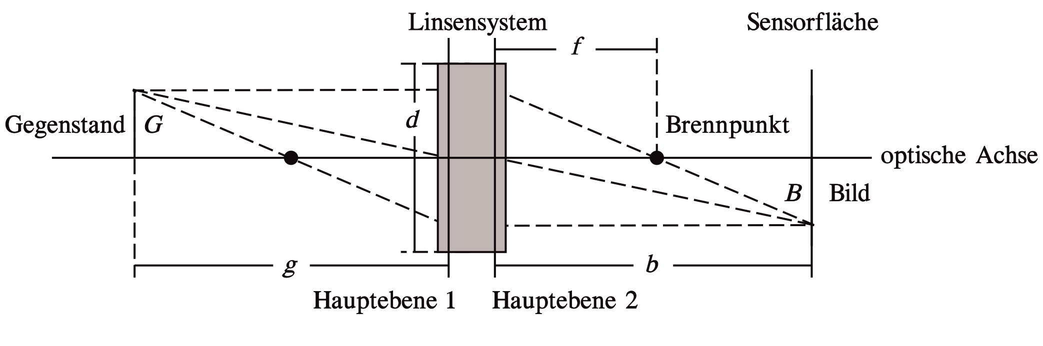


Abbildung 2.5: Modell eines Kameraobjektivs

Wird die durch das Kameraobjektiv durchgeführte Abbildung mit Hilfe der Zentralprojektion modelliert, wird der Abstand der beiden Hauptebenen vernachlässigt, da er im allgemeinen sehr viel kleiner als die Bild- und Gegenstandsweite ist. Somit fallen die beiden Hauptpunkte zusammen und werden als optisches Zentrum bezeichnet. Weiter ergibt sich der Abstand zwischen Bildebene und Gegenstand nur noch aus der Summe von Gegenstandsweite und Bildweite.

Mit Hilfe des Strahlensatzes ergibt sich folgende Beziehung zwischen den Kamera- und Bildkoordinatensystemen:

$$\frac{x_c}{x_i} = \frac{z_c}{b} \Rightarrow x_i = \frac{x_c b}{z_c} \text{ und } \frac{y_c}{y_{il}} = \frac{z_c}{b} \Rightarrow y_{il} = \frac{y_c b}{z_c} \,, \tag{2.7}$$

wobei b die Bildweite der Kamera angibt. Verwendet man homogene Koordinaten, so kann die projektive Transformation von Sehstrahlen $\boldsymbol{P_i}$ in Bildkoordinaten auf Sehstrahlen $\boldsymbol{P_c}$ in Kamerakoordinaten aus Gleichung (2.7) in Matrizenform folgendermaßen dargestellt werden [Fau93] (siehe Abb. 2.7):

$$\lambda \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = P' \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} \text{, wobei } P' = \begin{pmatrix} -b & 0 & 0 & 0 \\ 0 & -b & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \text{ und } \lambda \in \mathbb{R} \,. \tag{2.8}$$

Die Gegenstandsweite g ergibt sich mit der Brennweite f aus dem Linsengesetz und wird ebenfalls in mm angegeben:

$$\frac{1}{f} = \frac{1}{g} + \frac{1}{b} \quad \Rightarrow \quad g = \frac{bf}{b-f} \,. \tag{2.9}$$

Zwischen der Bildgröße B und der Gegenstandsgröße G gilt folgender Abbildungsmaßstab β (ohne Einheit):

$$\beta = \frac{b}{g} = \frac{B}{G} \,. \tag{2.10}$$

Der Durchmesser des Kreises in, in dem Licht bzw. die Sehstrahlen in das Linsensystem einfallen können und dessen Mittelpunkt in der optischen Achse liegt, wird mit d bezeichnet und in mm angegeben. Aus diesem Durchmesser und der Brennweite ergibt sich die Blendenzahl $k = \frac{f}{d}$ (ohne Einheit). Gegenstände, die sich in einer von der Gegenstandsweite abweichenden Entfernung von der Linse befinden, werden auf der Bildebene auf einen Zerstreuungskreis mit dem Durchmesser u abgebildet. Es gibt genau zwei Gegenstandsweiten, die einen Gegenstand auf die Bildebene mit einem Zerstreuungskreis von einem bestimmten Durchmesser abbilden. Diese Gegenstandsweiten werden zu gegebenem u mit Vordertiefe g_v und Hintertiefe g_h bezeichnet. Der Schärfentiefebereich ist der Bereich, in dem ein Gegenstand auf der Bildebene auf einen Durchmesser abgebildet wird, der kleiner als ein bestimmtes u ist. Dieser Bereich befindet sich zwischen der Vordertiefe g_v und Hintertiefe g_h. Die zur Vorder- und Hintertiefe korrespondierenden Bildweiten ergeben sich aus Gleichung (2.9) zu b_v und b_g (Abb. 2.6). Aus dem

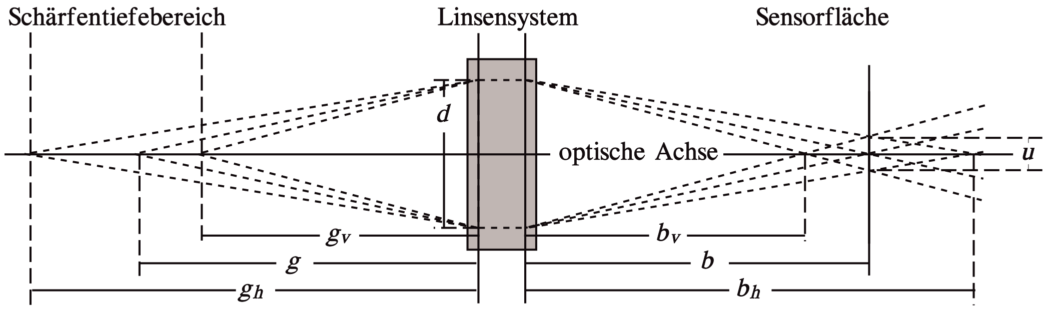


Abbildung 2.6: Modell des Schärfentiefebereichs eines Kameraobjektivs

Strahlensatz ergibt sich folgender Zusammenhang [Kra86]:

$$\frac{d}{b_v} = \frac{u}{b_v - b} \text{ und } \frac{d}{b_h} = \frac{u}{b - b_h} \ . \tag{2.11}$$

Durch Einsetzen von Gleichung (2.9) und der Blendenzahl k und Auflösen nach g_v und g_h erhält man somit:

$$g_v = \frac{bf}{b - f + ku} \text{ und } g_h = \frac{bf}{b - f - ku} \ . \tag{2.12}$$

Das Sichtfeld ($G_{\max}$) ist der maximale Ausschnitt einer Szene, der mit einer Einstellung der Kamera, abhängig vom Abbildungsmaßstab β, aufgenommen werden kann. Die Auflösung ($G_{\min}$) ist die minimale Größe eines Szenenmerkmals, so daß es durch Bildverarbeitung noch als Bildmerkmal detektiert werden kann. Diese Größen berechnen sich jeweils für die x- bzw. y-Koordinatenachse aus der Sensorgröße des CCD-Chips, der Anzahl ausgewerteter Bildpunkte und einem Subpixelfaktor:

$$G_{\max} = \frac{\text{Anzahl der Bildpunkte} \times \text{Sensorzellengröße}}{\beta} \tag{2.13}$$

$$G_{\min} = \frac{\text{Sensorzellengröße} \times \text{Subpixelfaktor}}{\beta} \tag{2.14}$$

Der Subpixelfaktor gibt die Bildverarbeitungsqualität an. Ein Subpixelfaktor von 1 bedeutet, daß die Auflösung des Bildes gleich der Auflösung des Chips ist.

Die Transformation zwischen Bild- und Rechnerkoordinatensystem hat nach Faugeras [Fau93] zunächst vier Freiheitsgrade. Dies sind zum einen Skalierungsfaktoren k_x und k_y, die in *Pixel/mm* angegeben werden, und zum anderen die Koordinaten des Bildhauptpunktes x_0 und y_0 in Rechnerkoordinaten, die in *Pixel* angegeben werden. Die Skalierungsfaktoren in y-Richtung ergeben sich direkt aus dem vertikalen Abstand der Sensorzellen auf dem CCD-Bildaufnehmer, in

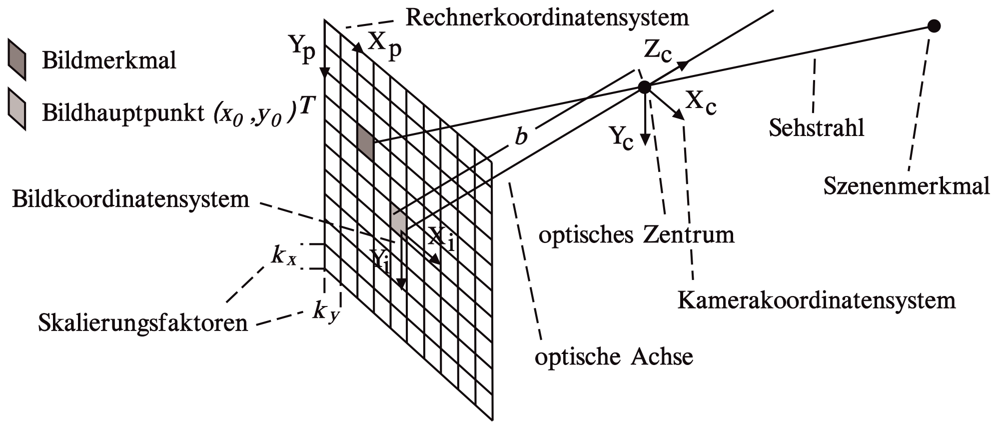


Abbildung 2.7: Abbildung der Sehstrahlen durch die projektive Transformation

x-Richtung muß zusätzlich noch das Verhältnis von Pixeltaktfrequenz zur Abtastfrequenz des Bildauswertungsrechners berücksichtigt werden [Gen94] (siehe Abb. 2.7):

$$k_x = \frac{1}{\text{horizontaler Pixelabstand}} \frac{\text{Abtastfrequenz}}{\text{Kamerapixeltakt}} , \tag{2.15}$$

$$k_y = \frac{1}{\text{vertikaler Pixelabstand}} . \tag{2.16}$$

Diese Transformation wird durch homogene Koordinaten in Matrizenform folgendermaßen dargestellt [Fau93]:

$$\begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = P'' \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} , \text{ wobei } P'' = \begin{pmatrix} k_x & 0 & x_0 \\ 0 & k_y & y_0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{2.17}$$

Das Betrachtungsverhältnis k (aspect ratio) ist definiert durch $k = \frac{k_x}{k_y}$ (ohne Einheit).

Faßt man Gleichung (2.8) und (2.17) durch Einführen neuer Skalierungsvariablen s_x und s_y mit $s_x = bk_x$ und $s_y = bk_y$ zusammen, so erhält man die Transformation zwischen Kamera- und Rechnerkoordinatensystem:

$$\lambda \begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = P''' \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} , \text{ wobei } P''' = \begin{pmatrix} s_x & 0 & x_0 & 0 \\ 0 & s_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \text{ und } \lambda \in \mathbb{R} . \tag{2.18}$$

Die Skalierungsvariablen werden in *Pixel/mm* angegeben. Aus Gleichung (2.18) folgt die projektive Transformation der Sehstrahlen $\boldsymbol{P_p}$ in Rechnerkoordinaten auf die Sehstrahlen $\boldsymbol{P_c}$ in Kamerakoordinaten:

$$x_p = s_x \frac{x_c}{z_c} + x_0 \ , \ y_p = s_y \frac{y_c}{z_c} + y_0 \ . \tag{2.19}$$

2.3.3 Transformation zwischen Szene und Kamera

Die Transformation zwischen Szenen- und Kamerakoordinatensystem besitzt sechs Freiheitsgrade, nämlich drei Winkel für die Rotation und drei Komponenten für die Translation. Diese Abbildung läßt sich als Transformation in Matrizenform folgendermaßen darstellen:

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = R \begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix} , \text{ wobei } R = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{2.20}$$

2.4 Fluchtpunkte

Die Bilder im Raum paralleler, jedoch nicht notwendigerweise koplanarer Geraden schneiden sich bei perspektivischer Projektion in der Bildebene. Dieser gemeinsame, eventuell im Unendlichen liegende Schnittpunkt heißt Fluchtpunkt. Die perspektivische Projektion der dreidimensionalen Welt auf die zweidimensionalen Bildkoordinaten wird durch die Zentralprojektion beschrieben.

Sei $G = (\bar{x_c}, \bar{y_c}, \bar{z_c})^T + \lambda(x_c, y_c, z_c)^T$ mit $\lambda \in \mathbb{R}$ eine Gerade im Kamerakoordinatensystem. Aus Gleichung (2.18) erhält man folgende homogene Koordinatendarstellung des Fluchtpunktes der perspektivischen Transformation in Rechnerkoordinaten:

$$\begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} = \lim_{\lambda \to \infty} \begin{pmatrix} s_x(\bar{x_c} + \lambda x_c) + (\bar{z_c} + \lambda z_c) x_0 \\ s_y(\bar{y_c} + \lambda y_c) + (\bar{z_c} + \lambda z_c) y_0 \\ \bar{z_c} + \lambda z_c \end{pmatrix} = P''' \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} , \tag{2.21}$$

wobei P''' wie in Gleichung(2.18) definiert ist. Für den Fall $z_{pl} \neq 0$ existieren die eigentlichen inhomogenen Bildkoordinaten des Fluchtpunktes mit

$$(x_p, y_p)^T = (s_x \frac{x_c}{z_c} + x_0 \ , \ s_y \frac{y_c}{z_c} + y_0)^T \ . \tag{2.22}$$

Im Fall $z_{pl} = 0$ verlaufen die Raumgeraden parallel zur Bildebene. Daraus ergibt sich der uneigentliche Fluchtpunkt $[x_p, y_p, 0]^T$ in homogenen Koordinaten.

Der Richtungsvektor einer Geraden im Raum ist gleichzeitig der homogene Koordinatenvektor des Fluchtpunktes seiner Projektion, was äquivalent dazu ist, daß der Verbindungsvektor vom optischen Zentrum zum Fluchtpunkt kolinear

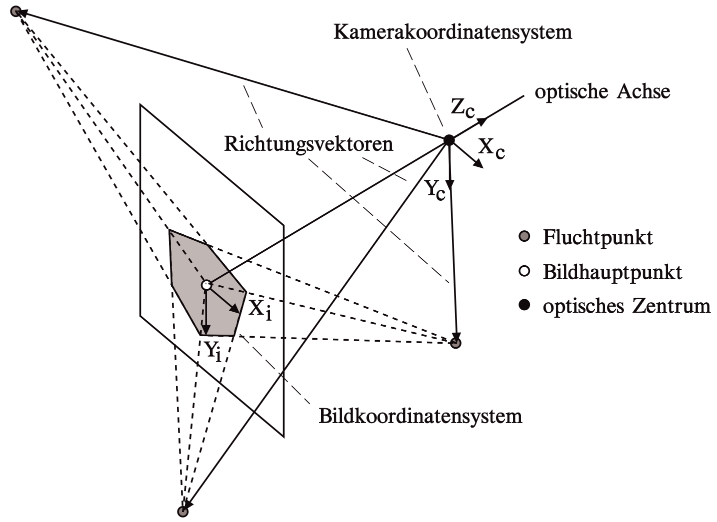


Abbildung 2.8: Geometrie von Fluchtpunkten

zum Richtungsvektor aller Raumgeraden ist, deren Bilder sich im Fluchtpunkt schneiden (Abb. 2.8).

Fluchtpunkte sind translationsinvariant, da beim Grenzübergang in Gleichung (2.21) der Translationsvektor $(\bar{x}_c, \bar{y}_c, \bar{z}_c)^T$ keine Rolle spielt. Allerdings sind Fluchtpunkte von der Orientierung und den inneren Kameraparametern abhängig.

Verlaufen zwei Geraden orthogonal im Raum, so heißen die beiden zugehörigen Fluchtpunkte der perspektivischen Abbildung konjugiert. Ist in den folgenden Kapiteln von im Ursprung parallelen Geraden die Rede, so bezeichnet dies Geraden, die in der Welt parallel sind und durch die Zentralprojektion auf eine projektive Ebene abgebildet werden.

2.5 Systemarchitektur

Zur Anwendung der aus dieser Arbeit gewonnenen Ergebnisse wurde ein System verwendet, das sich aus kommerziellen Hardwarekomponenten zusammensetzt (Abb. 2.9).

Die verwendete binokulare Kameraplattform BiSight der Firma Transition Research Corporation (TRC) besitzt vier Freiheitsgrade. Zum einen zwei unabhängige Vergenzachsen der Kameras, zum anderen eine Schwenk- und eine Neigungsachse. Die beiden Vergenzachsen und die Schwenkachse sind vertikale Achsen, die parallel zueinander im Raum liegen. Die Neigungsachse ist eine horizontale Achse. Die drei vertikalen Achsen haben jeweils einen Schnittpunkt mit der horizontalen Achse.

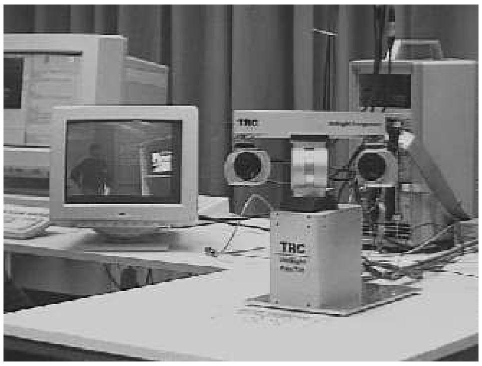


Abbildung 2.9: Abbildung des Kamerasystems

Die beiden Kameras, die an die Kameraplattform angebracht sind, sind Modelle XC-77RR der Firma Sony. Die Aufnahmeelemente beruhen auf dem Prinzip eines Interline-Transfer CCD-Bildwandlers und besitzen 768 $\times$ 493 Bildelemente[1]. Die Zellengröße beträgt $11\mu m \times 13\mu m$.

Die verwendeten Objektive der Kameras sind Modelle HGX11.5B-MPX31 der Firma Fujinon. Die Brennweite der Objektive ist zwischen 11.5mm und 69mm variabel, wodurch man ein Zoom-Ratio von fünf erhält. Somit besitzt das System zusätzlich zu den vier Freiheitsgraden der Plattform noch zwei weitere optische Freiheitsgrade (Zoom und Fokus).

Die untergelagerte Servolenkung der Achsen der Kameraplattform erfolgt auf einer Steuereinheit PMAC der Firma DeltaTau. Diese Steuereinheit liefert auch die Winkelpositionen. Die Kommunikation zwischen den Systemkomponenten ist in Abbildung 2.10 dargestellt.

Das Videoausgangssignal der CCD-Kameras wird in je einem MaxVideo200 Pipeline Rechner der Firma Datacube, Inc. digitalisiert. Das Digitalisierungsver-

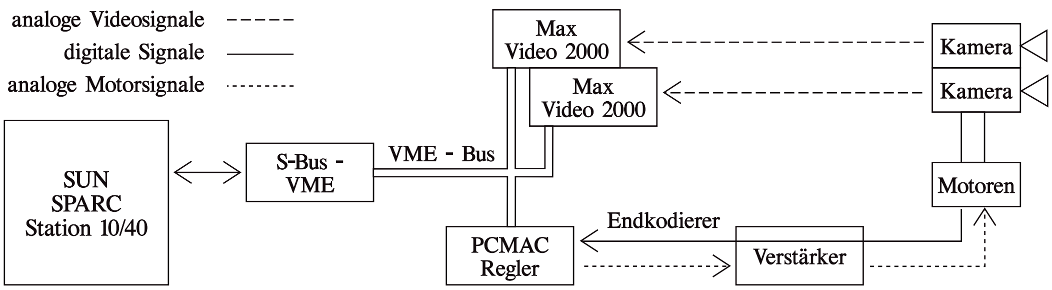


Abbildung 2.10: Architektur des realisierten Regelkreises (aus [DHKS95])

[1] Diese Arbeit wurde vor der Installation mit der CCIR77-Norm durchgeführt.

fahren extrahiert dabei zusammenhängend die ersten 480 der 493 zur Verfügung stehenden Zeilen. Aus den 768 zur Verfügung stehenden Spalten werden 512 extrahiert, so daß das digitalisierte Ausgangssignal eine Auflösung von 512×480 *Pixel* besitzt. Somit gilt: $x_{p_{\max}} = 512$ und $y_{p_{\max}} = 480$ für beide Kameras. In diesem Rechner findet ebenfalls die in der Einleitung angesprochene Video-Echtzeit-Bildverarbeitung zur Bewegungsdetektion statt.

Das digitalisierte Videosignal wird anschließend über einen VME-Bus an eine Sun-Sparcstation 10/40 weitergeleitet. Die Auswertung der Videosignale und die Kalibrierung der Kameraplattform erfolgen auf diesem Rechner.

Kapitel 3

Kalibrierung der internen Kameraparameter

In diesem Kapitel wird die Kalibrierung von internen Kameraparametern besprochen. Die Kalibrierung von kamerainternen Parametern hat im engeren Sinne wenig mit einer Kameraplattform zu tun und bezieht sich nur auf eine Kamera. In dieser Arbeit werden Freiheitsgrade solcher Plattformen ausgenutzt. Die Kalibrierung der internen Parameter beider Kameras einer Kameraplattform wird entkoppelt durchgeführt.

Die internen Parameter einer Kamera bestehen u.a. aus den Skalierungsfaktoren und dem Bildhauptpunkt. Auf die Kalibrierung dieser Parameter wird in diesem Kapitel eingegangen. Neben diesen Parametern besitzen einige Kameras auch variable interne Parameter, wie beispielsweise eine veränderbare Brennweite. Auf die Kalibrierung dieser Parameter wird in dieser Arbeit nicht eingegangen (siehe dazu u.a. Li und Lavest [LL95] oder Tarabanis et al. [TTG94]).

Zunächst werden hier die verschiedenen Verfahren der Kalibrierung der internen Parameter voneinander abgegrenzt. Anschließend werden die verschiedenen Verfahren der Kalibrierung interner Parameter differenziert besprochen. Im Kapitel 3.4.4 wird auf das Hauptresultat dieser Arbeit eingegangen. Dies ist ein Verfahren, das auf einer unbekannten Rotation um eine bekannte Achse beruht.

3.1 Kalibrierverfahren

Die Verfahren zur Kalibrierung der internen Parameter kann man in zwei Gruppen unterteilen:

Die erste Gruppe ist die der klassischen Kalibrierverfahren. In diese Gruppe ordnet man die Verfahren ein, die sich mit der Kalibrierung der internen Kameraparameter einer unbewegten Kamera beschäftigen und welche die Kameraparameter unter Zuhilfenahme eines Kalibrierkörpers ermitteln. Diese Kalibrierverfahren berechnen dabei eine Transformation zwischen der abgebildeten Szene und der Kamera. Die Freiheitsgrade, die beispielsweise eine Kameraplattform besitzt (siehe Kap. 2.3.1), werden dabei nicht ausgenutzt.

Die zweite Gruppe ist die Gruppe der aktiven Verfahren. Aktiv bedeutet in diesem Zusammenhang, daß

- man eine Kamera oder ein Objekt rotieren läßt,
- eine Kamera oder ein Objekt aus dem Sichtbereich verschoben wird,
- variable Parameter einer Kamera verändert werden.

Die Kamera wird also an verschiedenen Orten positioniert, oder es werden die Freiheitsgrade der Montageeinheit, auf der die Kamera installiert ist, ausgenutzt. Variable Parameter der Kamera können Brennweite und Fokus sein.

Im Gegensatz zu den klassischen Kalibrierverfahren werden bei den aktiven Verfahren vermehrt Erkenntnisse aus der projektiven Geometrie einfließen. Im Rahmen dieser Arbeit wird dies hauptsächlich Wissen über Fluchtpunkte von im Urbild parallelen Geraden sein.

Mit den aktiven Verfahren kann ebenfalls unter Zuhilfenahme eines Kalibrierkörpers kalibriert werden, man ist jedoch bestrebt, einen solchen zu vermeiden.

Man unterscheidet weiterhin zwischen linearen und nichtlinearen Verfahren. Die nichtlinearen Verfahren nehmen im allgemeinen eine genauere Kalibrierung vor. Dafür liegen bei diesen Methoden oftmals längere Rechenzeiten vor, und in manchen Fällen konvergieren sie in einem Nebenminimum. Um die Konvergenz in einem Nebenminimum zu vermeiden, werden Startwerte vorgegeben, die schon in etwa den tatsächlichen Werten entsprechen. Dies sind im allgemeinen Ergebnisse, die mit einem linearen Verfahren erzielt wurden.

Eine erstmalige Kalibrierung mit einem nichtlinearen Verfahren stellt somit ein Problem dar, da keine geeigneten Startwerte zur Verfügung stehen. Aus diesem Grund ist man bestrebt, Verfahren so zu vereinfachen, daß die Kalibrieraufgabe linear lösbar wird.

Im allgemeinen ist für die Kalibrierung eine Menge von Zuordnungen nötig. Diese werden bei verschiedenen Verfahren zwischen unterschiedlichen Merkmalen benötigt. Diese unterschiedlichen Merkmale können Szenen- oder Bildmerkmale oder Winkel sein.

Szenenmerkmale können beispielsweise Ecken eines Objektes in der Welt sein und werden in Szenenkoordinaten angegeben. Eine Menge von Szenenmerkmalen ist somit durch

$$\mathcal{S} \subseteq \mathbb{R} \times \mathbb{R} \times \mathbb{R} \tag{3.1}$$

gegeben. Die Szenenmerkmalskoordinaten müssen sehr genau bekannt sein. Bildmerkmale können die Ecken eines Objektes sein, nachdem sie durch die Zentralprojektion auf die Bildfläche einer Kamera abgebildet wurden. Sie werden in

Rechnerkoordinaten angegeben. Da Bildmerkmale mit Subpixelgenauigkeit bestimmt werden können, kann eine Menge von Bildmerkmalen durch

$$\mathcal{P} \subseteq \mathbb{R} \times \mathbb{R} \tag{3.2}$$

beschrieben werden. Die Bildmerkmale werden durch Bildauswertungsverfahren detektiert. Winkel sind in diesem Zusammenhang Rotationswinkel einer Kamera und durch

$$\mathcal{W} \subseteq \mathbb{R} \tag{3.3}$$

gegeben. Dabei gilt für alle $w \in \mathcal{W}$: $0 \leq w < 2\pi$. Eine Zuordnung ist durch n Merkmale gegeben. Somit ist eine Menge von Zuordnungen durch

$$\mathcal{Z} \subseteq (\mathcal{S} \cup \mathcal{P} \cup \mathcal{W})^n \tag{3.4}$$

gegeben.

3.2 Klassische Kalibrierverfahren

Die klassischen Kalibrierverfahren benutzen einen Kalibrierkörper. Dieser wird mit der Kamera aufgenommen und muß über Szenenmerkmale wie Kanten, Ecken, Bohrungen oder ähnliches verfügen, die möglichst zuverlässig durch Bildauswertungsverfahren den Szenenkoordinaten zugeordnet werden können. Die klassischen Kalibrierverfahren benötigen zur Bestimmung der internen Parameter eine Aufnahme des Kalibrierkörpers.

Zieht man mehrere Aufnahmen heran, so ist für jede Aufnahme eine getrennte Kalibrierung notwendig. Aus den Einzelergebnissen kann anschließend ein Gesamtergebnis durch eine Ausgleichsrechnung bestimmt werden, da sich die internen Parameter durch eine Rotation oder Translation der Kamera oder des Kalibrierkörpers nicht verändern.

Der in diesem Abschnitt vorgestellte Ansatz der klassischen Kalibrierung ist ein zweistufiger Prozeß. Im ersten Schritt berechnet man eine Transformationsmatrix. Für manche Aufgaben ist es schon ausreichend, diese Transformationsmatrix zu kennen (siehe z.B. Stereorekonstruktion bei Faugeras [Fau93]). Andere Aufgaben erfordern die Kenntnis der internen Kameraparameter. Diese erhält man im zweiten Schritt aus der Transformationsmatrix. Die freien Parameter der Transformation sind sowohl die internen Parameter der Kamera als auch die externen Parameter der Translation und Rotation zwischen Rechner- und Szenenkoordinatensystem.

Ein anderer Ansatz der klassischen Kalibrierung wird u.a. von Tsai [Tsa86] vorgestellt. In diesem Ansatz wird ein einstufiges Verfahren vorgestellt, welches die internen Kameraparameter und die externen Parameter der Rotation und Translation zwischen Kamera und Kalibrierkörper direkt aus den Zuordnungen bestimmt.

Ein Überblick über verschiedene klassische Kalibriermethoden wird durch Tsai [Tsa89] gegeben. Dort werden die verschiedenen Methoden voneinander abgegrenzt und praktische Hinweise gegeben, welche Kalibriermethode für welche Anforderungen die zweckmäßigste ist.

3.2.1 Definition der klassischen Kalibrierung

Bei dem Einsatz einer Kamera wird eine dreidimensionale Szene durch eine Funktion f auf die Elemente der zweidimensionalen Grauwertmatrix im Bildspeicher abgebildet:

$$\begin{pmatrix} x_p \\ y_p \end{pmatrix} = f\left(\begin{pmatrix} x_s \\ y_s \\ z_s \end{pmatrix}\right) . \tag{3.5}$$

Diese Abbildung wird durch ein physikalisches Modell beschrieben, auf das sowohl der physikalische Aufbau der Kamera, die Abtastrate der Sensorfläche als auch die Digitalisierung durch den Bildwandler gemäß der Zentralprojektion (siehe Kap. 2.3.2) und eventueller Erweiterungen (siehe Kap. 3.2.2) Einfluß nimmt. Des weiteren wird durch das Auswählen eines Bildausschnitts durch eine Bildverarbeitungsmaschine auf die Abbildung Einfluß genommen (Kap. 2.5). Zusätzlich geht eine beliebige Translation und Rotation zwischen Rechner- und Szenenkoordinatensystem gemäß Gleichung (2.20) in diese Transformation ein. Die Aufgabe der klassischen Kalibrierung ist es zunächst, diese Transformation zu bestimmen, um anschließend die freien Parameter zu berechnen.

Für die klassische Kalibrierung ist eine Menge von Zuordnungen nach (3.4) notwendig. Es müssen Szenen- und Bildmerkmale einander zugeordnet werden. Somit ist die Zuordnungsmenge durch

$$\mathcal{Z} \subseteq \mathcal{S} \times \mathcal{P} \tag{3.6}$$

gegeben. Dazu muß ein Szenenmerkmal jeweils genau dem Bildmerkmal zugeordnet werden, auf dessen Sehstrahl es liegt. Auf Methoden, die eine solche Zuordnung ermöglichen, wird in Kapitel 4 eingegangen.

Mit Hilfe dieser Zuordnung kann bei der klassischen Kalibrierung ein optimaler Parametervektor für das Kameramodell bestimmt werden. Optimal heißt in diesem Fall, daß die Summe der quadrierten Abstände zwischen den mit Hilfe des Kameramodells bestimmten Projektionen der Szenenmerkmale und den detektierten Bildmerkmalen minimal wird:

$$\min \sum_{(P_s,P_p)\in\mathcal{Z}} \left\|\boldsymbol{P_s} - f\left(\boldsymbol{P_p}\right)\right\|^2 , \tag{3.7}$$

wobei f wie in (3.5). Diese nichtlineare Problemstellung wird im allgemeinen durch Umformung auf eine lineare Form gebracht, wobei die resultierende Problemstellung nicht äquivalent zur nichtlinearen Problemstellung ist.

3.2.2 Erweiterung der Zentralprojektion

Zusätzlich zu den in Kapitel 2.3.2 eingeführten freien Parametern kann das Kameramodell noch um weitere Abbildungen zur Modellierung von Linsenverzerrungen erweitert werden. Dies sind beispielsweise radiale Linsenverzerrungen, die aufgrund einer fehlerhaften Krümmung der Oberflächen der Linsen entstehen. Außerdem können solche Verzerrungen tangentiale und radiale Linsenverzerrungen sein, die aufgrund einer nicht zentrierten Montage der Linsen im Linsensystem entstehen.

Des weiteren können solche Verzerrungen aufgrund leichter Neigungen der Linsen zueinander im Objektiv entstehen. Gengenbach [Gen94] bietet einen Überblick über verschiedene Linsenverzerrungen und darüber, wie diese von verschiedenen Autoren modelliert wurden. Als Beispiel sei hier die Modellierung einer radialen Linsenverzerrung durch einen Faktor κ nach Tsai [Tsa86] angegeben:

$$x_i' = \frac{x_i}{z} \text{ und } y_i' = \frac{y_i}{z} \text{ , wobei } z = 1 + \kappa({x_i}^2 + {y_i}^2) \ . \tag{3.8}$$

Die gestrichenen Variablen bedeuten in diesem Zusammenhang, daß von einem verzerrten Bildkoordinatensystem auf ein unverzerrtes Bildkoordinatensystem abgebildet wird.

Für die Realisierung von Steuerungsaufgaben für einen Roboter oder auch für eine Kameraplattform ist zu beachten, daß sämtliche aus der Aufnahme extrahierten Bildmerkmale mit Hilfe des Modells der Linsenverzerrungen korrigiert werden müssen. Dies stellt insbesondere dann ein Problem dar, wenn etwa Polynome höherer Ordnung ausgewertet werden müssen, da der dadurch entstehende Rechenaufwand hoch ist. Um die Echtzeitfähigkeit eines Systems nicht zu beeinträchtigen, wird daher im allgemeinen auf eine aufwendige Korrektur der Linsenverzerrung verzichtet [Gen94]. Solche Linsenverzerrungen spielen hauptsächlich dann eine Rolle, wenn die Diagonale der Abtastfläche der verwendeten Kameras größer ist als die Brennweite des zugehörigen Objektives. Die Diagonale des zur Aufnahme relevanten Aufnahmefeldes der verwendeten Kameras beträgt weniger als $9mm$, die Brennweite beträgt mindestens $11.5mm$ (Kap. 2.5). Somit spielen Linsenverzerrungen bei der Kalibrierung der verwendeten Kameraplattform keine nennenswerte Rolle.

Faugeras und Toscani [FT86] modellieren ebenfalls keine Linsenverzerrungen, lassen jedoch eine Drehung der x_p-Achse zu:

$$x_p = k_x \frac{x_c}{z_c} + k_{xy} \frac{y_c}{z_c} + x_0 \ , \ y_p = k_y \frac{y_c}{z_c} + y_0 \ . \tag{3.9}$$

Diese Methode hat den Vorteil, daß nur eine zusätzliche Multiplikation zur Entzerrung der Bildpunkte notwendig ist und somit auch die Echtzeitfähigkeit eines Systems nicht stark beeinträchtigt wird.

Faugeras [Fau93] modelliert einen zusätzlichen Parameter, indem er einen Winkel zwischen den Richtungsvektoren der x_p- und y_p-Achse des Rechnerkoordinatensystems einführt. Er begründet seine Modellierung durch die Feststellung, daß die Basisvektoren des Rechnerkoordinatensystems nicht zwingendermaßen eine Orthogonalbasis bilden müssen. Diese Modellierung führt zu den folgenden Gleichungen:

$$x_p = k_x \frac{x_c}{z_c} - k_x \cot\theta \frac{y_c}{z_c} + x_0 \;,\; y_p = \frac{k_y}{\sin\theta} \frac{y_c}{z_c} + y_0 \;. \tag{3.10}$$

Für $\theta = \pi$ und $k_{xy} = 0$ sind die beiden Modellierungen gleich. In anderen Fällen scheint zunächst keine Äquivalenz vorzuliegen. Berücksichtigt man, daß in der ersten Modellierung verfälschte Werte für k_y aufgrund der Drehung um die x_p-Achse auftreten, so kann man dennoch von einer äquivalenten Modellierung sprechen. Im folgenden wird die erste Modellierung verwendet.

Durch Einführung einer neuen Skalierungsvariable s_{xy} mit $s_{xy} = -bk_{xy}$ erhält man somit die Matrizendarstellung:

$$\lambda \begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = P \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} \text{,wobei } P = \begin{pmatrix} s_x & s_{xy} & x_0 & 0 \\ 0 & s_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \text{ und } \lambda \in \mathrm{I\!R}_{\neq 0} \;. \tag{3.11}$$

Aus dieser Gleichung folgt die zu Gleichung (2.19) äquivalente projektive Transformation der Sehstrahlen $\boldsymbol{P_p}$ in Rechnerkoordinaten auf die Sehstrahlen $\boldsymbol{P_c}$ in Kamerakoordinaten:

$$x_p = s_x \frac{x_l}{z_c} + s_{xy} \frac{y_l}{z_c} + x_0 \;,\; y_p = s_y \frac{y_l}{z_c} + y_0 \;. \tag{3.12}$$

3.2.3 Perspektivische Transformationsmatrix

Durch Verkettung des Kameramodells ohne Linsenverzerrung, aber mit Drehung der x_p - Achse nach Gleichung (3.11) erhält man die perspektivische Transformation von Sehstrahlen $\boldsymbol{P_s}$ in Szenenkoordinaten auf Sehstrahlen $\boldsymbol{P_p}$ in Rechnerkoordinaten. Diese läßt sich mit homogenen Koordinaten in Matrizenform darstellen:

$$\lambda \begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = PR \begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix} \text{,wobei } \lambda \in \mathrm{I\!R}_{\neq 0} \;, \tag{3.13}$$

$$P = \begin{pmatrix} s_x & s_{xy} & x_0 & 0 \\ 0 & s_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \text{ und } R = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \;.$$

Die linke, obere 3×3 Untermatrix der Matrix R ist eine Rotationsmatrix und somit orthogonal und regulär. Daher beträgt der Rang der Matrix R vier. Man definiert $T = PR$ und bezeichnet T als perspektivische Transformationsmatrix. Die perspektivische Transformation vom Szenenkoordinatensystem ins Rechnerkoordinatensystem läßt sich durch Einführen der perspektivischen Transformationsmatrix in homogenen Koordinaten wie folgt darstellen:

$$\lambda \begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = T \begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix} \text{,wobei } \lambda \in \mathbb{R}_{\neq 0} \text{ und T} = \begin{pmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & t_{24} \\ t_{31} & t_{32} & t_{33} & t_{34} \end{pmatrix} \quad (3.14)$$

mit den Koeffizienten

$$t_{11} = s_x r_{11} + s_{xy} r_{21} + x_0 r_{31}\ ,\ t_{12} = s_x r_{12} + s_{xy} r_{22} + x_0 r_{32}\ ,$$
$$t_{13} = s_x r_{13} + s_{xy} r_{23} + x_0 r_{33}\ ,\ t_{14} = s_x t_x + s_{xy} t_y + x_0 t_z\ ,$$
$$t_{21} = s_y r_{21} + y_0 r_{31}\ ,\ t_{22} = s_y r_{22} + y_0 r_{32}\ ,$$
$$t_{23} = s_y r_{23} + y_0 r_{33}\ ,\ t_{24} = s_y t_y + y_0 t_z\ ,$$
$$t_{31} = r_{31}\ ,\ t_{32} = r_{32}\ ,\ t_{33} = r_{33}\ ,\ t_{34} = t_z\ .$$

Diese Transformation stellt eine Funktion in Abhängigkeit von den fünf internen Parametern der Kamera und sechs externen Parametern der Transformation zwischen Szene und Kamera dar. Die internen Parameter sind die Skalierungsvariablen s_x, s_y und s_{xy} und die Koordinaten des Bildhauptpunktes x_0 und y_0. Die externen Parameter sind drei Parameter für die Rotation und drei Parameter für die Translation zwischen Bild- und Szenenkoordinatensystem. Der Rang der Matrix T beträgt genau dann drei, wenn die internen Parameter s_x und s_y ungleich null sind, da dann der Rang der Matrix P aus Gleichung (3.13) drei beträgt. Der Rang der Matrix R beträgt ohnehin vier.

Für Matrizen $A \in \mathbb{R}^{m \times n}$, die einen projektiven Raum der Dimension n auf einen projektiven Unterraum der Dimension m abbilden, gilt, daß die Gleichungen $A\boldsymbol{x} = 0$ und $A(\lambda \boldsymbol{x}) = 0$ für $\lambda \in \mathbb{R}$ mit $\lambda \neq 0$ und $\boldsymbol{x} \in \mathbb{R}^n$ gleichbedeutend sind [Fis94]. Daraus folgt, daß die perspektivische Transformationsmatrix einen Freiheitsgrad besitzt [Fau93]: Die Matrizen T und λT repräsentieren für alle $\lambda \in \mathbb{R}$ mit $\lambda \neq 0$ dieselbe Projektion.

3.2.4 Bestimmung der perspektivischen Transformationsmatrix

Um die perspektivische Transformationsmatrix bestimmen zu können, untersucht man zunächst, wann eine 3×4-Matrix in Form der Gleichung (3.14) geschrieben werden kann, so daß eine Lösung für die internen und externen Parameter existiert. Faugeras [Fau93] beweist, daß genau dann vier Mengen von Lösungen der internen und externen Parameter der Transformationsmatrix existieren, wenn die Nebenbedingung

$$\|(t_{11}, t_{12}, t_{13})^T\| = 1 \quad (3.15)$$

erfüllt wird.

Die durch die erweiterte Zentralprojektion festgelegte Abbildung (3.14) ist nicht linear. Somit ist die Auflösung der perspektivischen Transformationsmatrix nach ihren 11 Parametern unter Minimierung des Abbildungsfehlers nach Gleichung (3.7) mit einem linearen Verfahren nicht möglich. Eine nichtlineare Minimierung stellt ebenfalls ein Problem dar, da im allgemeinen keine Startparameter für die Minimierungsaufgabe gegeben sind.

Um die Transformationsmatrix zu berechnen, wird das Problem daher durch geeignetes Umformen auf eine lineare Problemstellung reduziert, die nicht äquivalent zur Minimierungsaufgabe (3.7) ist. Dazu werden die dreidimensionalen und zweidimensionalen Koordinaten linear in Beziehung gesetzt. Aus Gleichung (3.14) folgt unmittelbar durch Auflösen nach x_p und y_p:

$$\begin{aligned} t_{11}x_s + t_{12}y_s + t_{13}z_s + t_{14} - x_p t_{31}x_s - x_p t_{32}y_s - x_p t_{33}z_s - x_p t_{34} = 0 \ , \\ t_{21}x_s + t_{22}y_s + t_{23}z_s + t_{24} - y_p t_{31}x_s - y_p t_{32}y_s - y_p t_{33}z_s - y_p t_{34} = 0 \ . \end{aligned}$$

Für n Punkte erhält man $2n$ homogene Gleichungen, die linear in den Komponenten der perspektivischen Transformationsmatrix sind. Man erhält folgendes lineares Gleichungssystem:

$$A\boldsymbol{t} = 0 \ , \tag{3.16}$$

wobei

$$A = \begin{pmatrix} x_s & y_s & z_s & 1 & 0 & 0 & 0 & 0 & -x_p x_s & -x_p y_s & -x_p z_s & x_p \\ 0 & 0 & 0 & 0 & x_s & y_s & z_s & 1 & -y_p x_s & -y_p y_s & -y_p z_s & y_p \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \end{pmatrix} ,$$

$$\boldsymbol{t} = (t_{11}, t_{12}, t_{13}, t_{14}, t_{21}, t_{22}, t_{23}, t_{24}, t_{31}, t_{32}, t_{33}, t_{34})^T \ .$$

Die Matrix A ist für n Punkte eine $2n \times 12$-Matrix und $\boldsymbol{t}$ ein 12-dimensionaler Vektor. Da die perspektivische Transformationsmatrix einen Freiheitsgrad besitzt, ist $\boldsymbol{t}$ bis auf einen Skalierungsfaktor eindeutig bestimmt.

Um den Freiheitsgrad der Transformationsmatrix festzulegen, kann in Gleichung (3.16) $t_{34} = 1$ definiert werden. Man erhält dann für n Punkte eine $2n \times 11$-Matrix A und einen 11-dimensionalen Parametervektor $\boldsymbol{t}$. In diesem Fall kann zur Lösung des Gleichungssystems ein lineares Ausgleichsverfahren, wie beispielsweise die Householder-Reduktion [Sto93] oder die Methode der kleinsten Quadrate, angewandt werden.

In diesem Fall wird nicht der reine Abbildungsfehler nach Gleichung (3.7) minimiert, sondern ein Fehlermaß, das von der Position der Kamera abhängig ist [FT86].

Ein weiteres lineares Verfahren zur Berechnung der internen Kameraparameter wird von Faugeras [Fau93] vorgeschlagen. Durch Beachtung der Nebenbedingung (3.15) wird die Abhängigkeit von der Aufnahmeposition vermieden. Dieses

Verfahren minimiert die quadrierte Norm von $\boldsymbol{At}$ unter der Nebenbedingung nach Gleichung (3.15):

$$\min_{\|(t_{31},t_{32},t_{33})^T\|=1} \|\boldsymbol{At}\|^2 . \tag{3.17}$$

Man erhält einen Ausdruck in geschlossener Form. Das Verfahren wird analog zur Bestimmung der Eigenvektoren einer 3×3-Matrix und dem Invertieren einer 9×9 Matrix durchgeführt.

Zuletzt wird in diesem Abschnitt darauf eingegangen, wann eine Lösung für die Matrix A vorliegt. Dazu wird zunächst der Rang dieser Matrix betrachtet. Die Dimensionsformel [Fis91] hat folgende Form:

$$Rang(A) + Dim(Kern(A)) = m . \tag{3.18}$$

In diesem Fall gilt $m = 12$. Somit existieren die folgenden Fälle:

- $Rang(A) = 12$. Dann ist mit (3.18) die Dimension des Kerns von $A = 0$. Weiterhin ist $\boldsymbol{t} = 0$ die einzige Lösung der Gleichung $\boldsymbol{At} = 0$. Diese Lösung ist nicht sinnvoll. Faugeras [Fau93] zeigt, daß eine solche Lösung nicht wahrscheinlich ist.

- $Rang(A) = 11$. Dann hat der Kern von A eine Dimension von 1, und es existiert eine bis auf einen Skalierungsfaktor eindeutige Lösung für Gleichung (3.16).

- $Rang(A) < 11$. Dann hat der Kern von A eine größere Dimension als 1. Somit hat der Lösungsraum ebenfalls eine größere Dimension als 1.

Faugeras [Fau93] beweist, daß der Rang der Matrix A genau dann 11 beträgt, wenn zur Lösung des Gleichungssystems (3.16) mindestens 6 Zuordnungen aus mindestens zwei verschiedenen Ebenen verwendet werden. Werden mindestens 4 Zuordnungen aus einer Ebene betrachtet, so beträgt der Rang der Matrix 8. Für koplanare Punkte ist es somit nicht möglich, eine Kalibrierung aller internen Kameraparameter durchzuführen. Lenz [Len87] gibt ein Verfahren an, das auch für koplanare Punkte interne Kameraparameter bestimmt. In diesem Verfahren werden allerdings einige interne Parameter als bekannt vorausgesetzt.

3.2.5 Zerlegung der perspektivischen Transformationsmatrix

Um aus der perspektivischen Transformationsmatrix die internen Kameraparameter zu berechnen, betrachte man die linke 3×3 - Untermatrix der Transformationsmatrix T aus Gleichung (3.14):

$$\overline{T} = \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix} . \tag{3.19}$$

Mit Gleichung (3.13) läßt sich diese durch $\overline{T} = \overline{P}\,\overline{R}$, wobei $\overline{P} = P$ und $\overline{R}$ die linke 3×3-Untermatrix von R nach Gleichung (3.13) ist, zerlegen. Da $\overline{R}$ eine reine Rotationsmatrix ist, ist diese orthogonal, und es gilt die folgende Umformung:

$$\overline{T}\,\overline{T}^T = (\overline{P}\,\overline{R})(\overline{P}\,\overline{R})^T = \overline{P}\,\overline{R}\,\overline{R}^T\,\overline{P}^T = \overline{P}\,\overline{R}\,\overline{R}^{-1}\,\overline{P}^T = \overline{P}\,\overline{P}^T = K\ ,$$

wobei

$$K = \begin{pmatrix} {s_x}^2 + {s_{xy}}^2 + {x_0}^2 & s_{xy}s_y + x_0 y_0 & x_0 \\ s_{xy}s_y + x_0 y_0 & {s_y}^2 + {y_0}^2 & y_0 \\ x_0 & y_0 & 1 \end{pmatrix} . \tag{3.20}$$

Aus der Matrix K lassen sich die internen Parameter der Kamera direkt berechnen:

$$x_0 = k_{13}\ , \tag{3.21}$$

$$y_0 = k_{23}\ , \tag{3.22}$$

$$s_y = \sqrt{k_{22} - {k_{23}}^2}\ , \tag{3.23}$$

$$s_{xy} = \frac{k_{12} - k_{13}k_{23}}{s_y}\ , \tag{3.24}$$

$$s_x = \sqrt{k_{11} - {k_{13}}^2 - {s_{xy}}^2}\ . \tag{3.25}$$

Mit Hilfe der internen Parameter und der perspektivischen Transformationsmatrix aus Gleichung (3.14) lassen sich die externen Parameter der Rotation und Translation zwischen Rechner- und Szenenkoordinatensystem berechnen.

3.3 Kalibrierung unter Ausnutzung von Fluchtpunkten

In diesem Paragraphen werden Methoden besprochen, die die internen Parameter einer Kamera durch Ausnutzen der projektiven Fluchtpunktgeometrie bestimmen. Solche Verfahren wurden unter anderem von Caprile und Torre [CT90], Echigo [Ech90], Kolnig [Kol95] und Li [Li94] veröffentlicht. Alle diese Ansätze benutzen stationäre Kameras und setzen Wissen über die Winkel zwischen mehreren Mengen von im Raum parallelen Geraden voraus. Keinem dieser Verfahren ist es möglich, sämtliche internen Parameter einer Kamera zu bestimmen.

In [Kol95] wird ein Verfahren vorgeschlagen, die Brennweite der Kameras anhand eines bekannten Schnittwinkels der Raumgeraden zu ermitteln. Die Raumgeraden müssen dabei nicht orthogonal zueinander sein. Umgekehrt kann man auch bei bekannter Brennweite den Schnittwinkel zwischen den Raumgeraden berechnen. Dieses Verfahren ist in unserem Fall nicht von Interesse, da weder die Brennweite noch der Schnittwinkel zwischen den Raumgeraden als bekannt vorausgesetzt werden kann.

Das im folgenden Abschnitt beschriebene Verfahren benutzt konjugierte Fluchtpunkte, also Fluchtpunkte zu im Raum orthogonalen Geradenbüscheln.

Durch dieses Verfahren sind nicht alle Parameter bestimmbar. Somit setzt man einige Parameter als bekannt voraus. Anschließend lassen sich die übrigen Parameter der Kamera und die Rotation zwischen Kamera und Szene bestimmen. Da Fluchtpunkte translationsinvariant sind, läßt sich die Translation zwischen Kamera und Szene nicht bestimmen (vgl. Kap.2.4).

3.3.1 Kalibrierung der internen Parameter

Ausgehend von Gleichung (2.21) erhält man den Richtungsvektor einer zu einem homogenen Fluchtpunktvektor im Kamerakoordinatensystem kolinearen Geraden (vgl. Kap.2.4, Abb. 2.8):

$$\begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = \begin{pmatrix} \frac{x_p - z_p x_0}{s_x} \\ \frac{y_p - z_p y_0}{s_y} \\ z_p \end{pmatrix} . \tag{3.26}$$

Seien $\boldsymbol{g_1}$, $\boldsymbol{g_2}$ und $\boldsymbol{g_3}$ drei paarweise verschiedene und konjugierte Fluchtpunkte gemäß Gleichung (3.26) mit den bekannten Rechnerkoordinaten $[x_{p_1}, y_{p_1}, z_{p_1}]^T$, $[x_{p_2}, y_{p_2}, z_{p_2}]^T$ und $[x_{p_3}, y_{p_3}, z_{p_3}]^T$.

Da die Richtungsvektoren der Raumgeraden orthogonal zu den homogenen Bildkoordinaten des Fluchtpunktes sind (vgl. Kap.2.4), gilt:

$$\boldsymbol{g_1} \cdot \boldsymbol{g_2} = 0 \ , \ \boldsymbol{g_2} \cdot \boldsymbol{g_3} = 0 \text{ und } \boldsymbol{g_3} \cdot \boldsymbol{g_1} = 0 \ . \tag{3.27}$$

Daraus erhält man folgendes Gleichungssystem mit drei Gleichungen und den vier Unbekannten x_0, y_0, s_x und s_y:

$$\begin{aligned} &(x_{p_1} - z_{p_1} x_0)(x_{p_2} - z_{p_2} x_0) s_x^{-2} \\ &\qquad + (y_{p_1} - z_{p_1} y_0)(y_{p_2} - z_{p_2} y_0) s_y^{-2} + z_{p_1} z_{p_2} = 0 \ , \\ &(x_{p_2} - z_{p_2} x_0)(x_{p_3} - z_{p_3} x_0) s_x^{-2} \\ &\qquad + (y_{p_2} - z_{p_2} y_0)(y_{p_3} - z_{p_3} y_0) s_y^{-2} + z_{p_2} z_{p_3} = 0 \ , \\ &(x_{p_3} - z_{p_3} x_0)(x_{p_1} - z_{p_1} x_0) s_x^{-2} \\ &\qquad + (y_{p_3} - z_{p_3} y_0)(y_{p_1} - z_{p_1} y_0) s_y^{-2} + z_{p_3} z_{p_1} = 0 \ . \end{aligned} \tag{3.28}$$

Liegen mindestens zwei Fluchtpunkte im Endlichen, existieren also $i, j \in \{1, 2, 3\}$ mit $i \neq j$ und $z_{p_i} \neq 0 \neq z_{p_j}$, so muß man mindestens eine der vier Unbekannten kennen, um die restlichen zu bestimmen. Liegen zwei oder drei Fluchtpunkte im Unendlichen, so läßt sich das Betrachtungsverhältnis (aspect ratio) berechnen, das sich aus $k = \frac{k_x}{k_y} = \frac{s_x}{s_y}$ ergibt. Die Berechnungen ergeben sich direkt aus Gleichung (3.28).

[Li94] modelliert ein ähnliches Verfahren, das jedoch nur für Fluchtpunkte mit inhomogenen Koordinaten gilt. Man erhält die Modellierung von [Li94] aus (3.28), wenn man $z_{p_i} = 1$ für alle $i \in \{1, 2, 3\}$ setzt.

3.3.2 Kalibrierung der Rotationsmatrix

Seien G_1, G_2 und G_3 drei Geraden, die orthogonal zueinander im Raum stehen und sich in genau einem Punkt schneiden. Diese Geraden werden so gewählt, daß der Ursprung eines Szenenkoordinatensystems durch den Schnittpunkt aller Geraden gegeben ist und die drei Einheitsvektoren $\boldsymbol{N}_{G_1} = \boldsymbol{E}_x$, $\boldsymbol{N}_{G_2} = \boldsymbol{E}_y$ und $\boldsymbol{N}_{G_3} = \boldsymbol{E}_z$ der Orthonormalbasis des Szenenkoordinatensystems durch die Richtungsvektoren der Geraden gegeben sind.

Die zugehörigen Fluchtpunkte der perspektivischen Abbildung seien erneut durch $\boldsymbol{g_1}$, $\boldsymbol{g_2}$ und $\boldsymbol{g_3}$ gegeben. Die zugehörigen normierten Einheitsvektoren sind durch $\boldsymbol{n_{g_1}} = \frac{\boldsymbol{g_1}}{|\boldsymbol{g_1}|}$, $\boldsymbol{n_{g_2}} = \frac{\boldsymbol{g_2}}{|\boldsymbol{g_2}|}$ und $\boldsymbol{n_{g_3}} = \frac{\boldsymbol{g_3}}{|\boldsymbol{g_3}|}$ gegeben. Durch eine Drehung R geht $\boldsymbol{n_{g_i}}$ für $i \in \{1, 2, 3\}$ aus $\boldsymbol{N}_{G_i}$ hervor. Daher gilt folgender Zusammenhang:

$$\boldsymbol{n_{g_1}} = R\boldsymbol{N}_{g_1} = R\boldsymbol{E}_x \ , \ \boldsymbol{n_{g_2}} = R\boldsymbol{N}_{g_2} = R\boldsymbol{E}_y \text{ und } \boldsymbol{n_{g_3}} = R\boldsymbol{N}_{g_3} = R\boldsymbol{E}_z \ . \tag{3.29}$$

Für die Rotationsmatrix R gilt also:

$$R = (\boldsymbol{n_{g_1}}, \boldsymbol{n_{g_2}}, \boldsymbol{n_{g_3}}) \ . \tag{3.30}$$

3.4 Aktive Kalibrierverfahren

Für aktive Kalibrierverfahren werden genauso wie bei den klassischen Verfahren Zuordnungen benötigt. Diese Zuordnungen bestehen in diesem Fall allerdings nicht zwischen Rechner- und Szenenkoordinaten, sondern zwischen Rechnerkoordinaten zu verschiedenen Zeitpunkten. Es kann genauso wie bei der Kalibrierung mit klassischen Verfahren ein Kalibrierkörper eingesetzt werden, bei dem durch Bildauswertungsverfahren Bildmerkmale detektiert werden. Allerdings wird nicht vorausgesetzt, daß die Szenenkoordinaten der detektierten Bildmerkmale bekannt sind. Daher kann auch auf einen Kalibrierkörper verzichtet werden. Man detektiert dazu aus jeder einzelnen Aufnahme Bildmerkmale, die anschließend einander zugeordnet werden müssen. Diese Zuordnungen haben in den unterschiedlichen Verfahren unterschiedliche Ausprägungen, so daß diese in den einzelnen Verfahren definiert werden.

3.4.1 Kalibrierung des Bildhauptpunktes durch Zoomen

In diesem Abschnitt wird eine Kalibriermethode besprochen, die allein der Bestimmung des Bildhauptpunktes dient. Die Zoomfunktion einer Kamera stellt

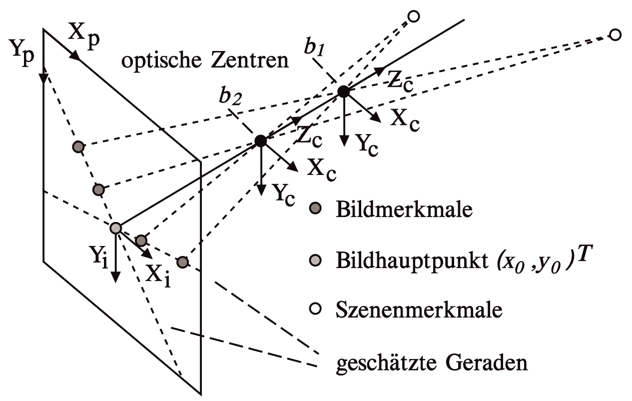


Abbildung 3.1: Bestimmung des Bildhauptpunktes durch Fokus und Zoom. Zwei unterschiedliche Szenenmerkmale werden auf die Sensorfläche bei zwei unterschiedlichen Brennweiten abgebildet. Für ein Szenenmerkmal ergeben sich die Abbildungen jeweils zu einer Gerade. Die Geraden schneiden sich im Bildhauptpunkt.

unterschiedliche Brennweiten der Objektive ein. Da diese in die theoretische Betrachtung nicht eingeht, sei folgendes angemerkt:

Betrachtet man die Zentralprojektion, so geht zunächst nur eine variierende Bildweite, nicht aber eine variierende Brennweite in die perspektivische Transformation ein (vgl. Gl. (2.19)). Der Zusammenhang zwischen Brennweite und Bildweite wird in der Linsengleichung (2.9) ausgedrückt. Daraus ergibt sich die Schärfentiefe in Gleichung (2.12). Wird im folgenden auch nur über die Bildweite gesprochen, so ist diese stets über die Schärfentiefe mit der Brennweite gekoppelt.

Wenn man fokussiert oder zoomt, so wird sich jeder Bild- oder Fluchtpunkt auf einer Geraden durch das Kamerabild bewegen. Alle diese Geraden schneiden sich genau in dem Schnittpunkt der optischen Achse mit der Sensorfläche, also dem Bildhauptpunkt (Abb. 3.1). Eine Zuordnung gemäß (3.4) ist somit ein n-Tupel von Bildmerkmalen, so daß alle Elemente dasselbe Szenenmerkmal repräsentieren. Die gesamte Zuordnungsmenge ist durch

$$\mathcal{Z} \subseteq \mathcal{B}^n \tag{3.31}$$

beschreibbar. Es werden nur detektierte Bildpunkte zugeordnet, die Szenenkoordinaten der Merkmale werden nicht benötigt. Außerdem werden die sich verändernden Werte für die Bildweite und die Brennweite ebenfalls nicht benötigt.

Sei $(x_{p_{ref}}, y_{p_{ref}})^T$ ein beliebiger Referenzpunkt im Rechnerkoordinatensystem. Abhängig von einer variablen Bildweite b erhält man aus Gleichung (2.19) die parametrisierte Geradendarstellung

$$g : \mathrm{I\!R} \to \mathrm{I\!R}^2, \mathrm{b} \mapsto (x_p, y_p)^T \text{ , wobei } \begin{pmatrix} x_p \\ y_p \end{pmatrix} = \begin{pmatrix} b(x_{p_{ref}} - x_0) + x_0 \\ b(y_{p_{ref}} - y_0) + y_0 \end{pmatrix} . \tag{3.32}$$

Durch Auflösen nach b erhält man die nicht-parametrisierte Gerade

$$\rho - x_p \cos\theta + y_p \sin\theta = 0 \ , \tag{3.33}$$

wobei

$$\sin\theta = \frac{\overline{x}}{\sqrt{\overline{x}^2 + \overline{y}^2}} \ , \ \rho = \frac{\overline{y}x_0 - \overline{x}y_0}{\sqrt{\overline{x}^2 + \overline{y}^2}} \ , \ \overline{x} = x_{p_{ref}} - x_0 \text{ und } \overline{y} = y_{p_{ref}} - y_0 \ .$$

Wie leicht einzusehen ist, befindet sich der Bildhauptpunkt $(x_0, y_0)^T$ auf allen Geraden g, bzw. g'.

Bestimmt man durch Variieren der Brennweite n korrespondierende Punkte im Bildkoordinatensystem, so kann man durch ein Ausgleichsverfahren eine Gerade anpassen. Zur Geradenanpassung benötigt man mindestens zwei korrespondierende Punkte. Findet man auf diese Weise m Geraden, so schneiden sich diese alle in dem Bildhauptpunkt. Um den Bildhauptpunkt zu finden, muß eine Ausgleichsrechnung durchgeführt werden, beispielsweise nach der Methode der kleinsten Quadrate. Zur Lösung dieser Aufgabe benötigt man mindestens zwei Geraden.

Durch Zoomen kann man den Bildhauptpunkt finden, wenn mindestens zwei Punkte im Bild detektiert werden, denen man nach mindestens einmaligem Zoomen die korrespondierenden Punkte zuweisen kann.

Die Stabilität dieser Methode ist abhängig von der Stabilität der Zoomfunktion. Da die Linsen durch das Zoomen bewegt werden, treten im allgemeinen Verzerrungen auf. Weiterhin ist die Stabilität von der Güte der durch ein Bildverarbeitungsverfahren detektierten Bildpunkte abhängig. Li [Li94] hat diese Methode der Bestimmung des Bildhauptpunktes benutzt, und verwertbare Ergebnisse erzielt.

3.4.2 Aktive Kalibrierung durch Rotationen

Dieses Kapitel beschäftigt sich mit verschiedenen Verfahren der Kalibrierung der internen Parameter einer Kamera, wobei jeweils Rotationsbewegungen einer zugrundeliegenden Kameraplattform ausgenutzt werden (Abb. 3.2). Dazu wird wie in Kapitel 2.3.1 vorausgesetzt, daß das interne Kamerakoordinatensystem dem Kamerakoordinatensystem entspricht.

Die verwendete Kameraplattform läßt um die optischen Zentren nur Neigungs- und Vergenzrotationen zu. Die hier vorgestellten Verfahren nutzen jeweils nur eine Neigungs- oder eine Vergenzrotation aus. Dies hat mehrere Gründe:

- Eine Rotation um eine beliebige Achse ist nur eingeschränkt möglich.
- Das Gleichungssystem einer allgemeinen Rotation ist nichtlinear und liegt nicht in geschlossener Form vor. Eine Linearisierung ist nicht möglich. Es ist davon auszugehen, daß keine sinnvollen Startwerte zur Verfügung stehen.

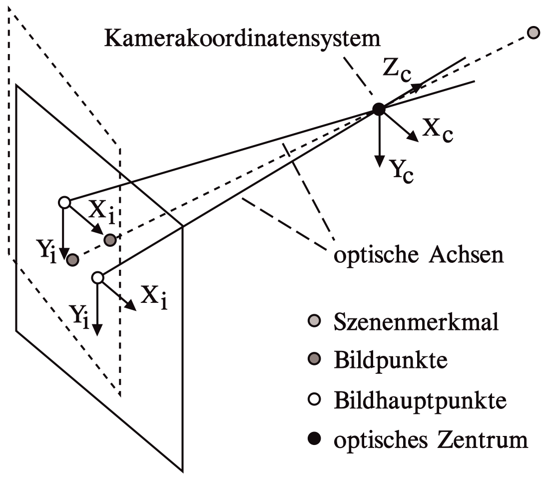


Abbildung 3.2: Punktverschiebung auf der Sensorfläche durch Rotation einer Kamera. Auf der rechten Sensorfläche (durchgezogene Linien) besitzt das abgebildete Szenenmerkmal eine negative y-Komponente. Auf der linken Sensorfläche besitzt das abgebildete Szenemerkmal eine positive y-Komponente.

Zunächst werden zwei eigene Lösungsansätze für die Kalibrierung vorgestellt. Der erste Ansatz beruht auf der Annahme, daß der Rotationswinkel zwischen den einzelnen Aufnahmen der Kamera bekannt ist. Der zweite Ansatz beruht auf der Tatsache, daß sich die Abbildungen der Szenenmerkmale auf einer Hyperbel eines Kreiskegelschnitts befinden. Für diesen Ansatz wird der Rotationswinkel zwischen den Kamerabildern nicht als bekannt vorausgesetzt. Anschließend werden weitere Lösungsansätze von anderen Autoren besprochen.

3.4.3 Kalibrierung mittels bekannter Rotationen

Bei der Kalibrierung der internen Kameraparameter durch Rotation der zugrundeliegenden Kamera werden in unterschiedlichen Aufnahmen Bildpunkte detektiert, die anschließend einander zugeordnet werden müssen. Der Ansatz, der in diesem Abschnitt besprochen wird, geht von der Annahme aus, daß die Rotationswinkel zwischen den einzelnen Aufnahmen mit den Kameras bekannt sind. Eine Zuordnung besteht jeweils aus zwei Bildmerkmalen, die dasselbe Szenenmerkmal repräsentieren, und aus einem Rotationswinkel zwischen den beiden Aufnahmen des Merkmals. Für die Zuordnungsmenge gemäß (3.4) gilt somit:

$$\mathcal{Z} \subseteq \mathcal{B} \times \mathcal{B} \times \mathcal{W} \; . \tag{3.34}$$

Sei $(x_{p_{ref}}, y_{p_{ref}})^T$ ein Referenzpunkt im Rechnerkoordinatensystem. Ausgehend von Gleichung (2.4) erhält man durch Einsetzen von (2.19) folgende allgemeine Gleichungen für Rotationen von Punkten im Rechnerkoordinatensystem:

$$\begin{pmatrix} \frac{x_{p_{ref}}-x_0}{s_x} \\ \frac{y_{p_{ref}}-y_0}{s_y} \end{pmatrix} = \begin{pmatrix} \dfrac{\cos\theta\frac{x_p-x_0}{s_x}+\sin\theta}{-\cos\phi\sin\theta_l\frac{x_p-x_0}{s_x}+\sin\phi\frac{y_p-y_0}{s_y}+\cos\phi\cos\theta_l} \\ \dfrac{\sin\phi\sin\theta_l\frac{x_p-x_0}{s_x}+\cos\phi\frac{y_p-y_0}{s_y}-\sin\phi\cos\theta_l}{-\cos\phi\sin\theta_l\frac{x_p-x_0}{s_x}+\sin\phi\frac{y_p-y_0}{s_y}+\cos\phi\cos\theta_l} \end{pmatrix}$$

$$\Leftrightarrow \quad \begin{pmatrix} x_{p_{ref}} \\ y_{p_{ref}} \end{pmatrix} = \begin{pmatrix} z^{-1}(s_x s_y x_p' \cos\theta + {s_x}^2 s_y \sin\theta) + x_0 \\ z^{-1}({s_y}^2 x_p' \sin\phi\sin\theta - s_x {s_y}^2 \sin\phi\cos\theta + s_x s_y y_p' \cos\phi) + y_0 \end{pmatrix} ,$$

wobei

$$\begin{aligned} x_p' = x_p - x_0 \ , \ y_p' = y_p - y_0 \ \text{und} \\ z = s_x s_y \cos\theta\cos\phi - s_y x_p' \sin\theta\cos\phi + s_x y_p' \sin\phi \ . \end{aligned} \tag{3.35}$$

Man erhält also aus einer beliebigen Rotation um Vergenz- und Neigungsachse des Kamerakopfes zwei Gleichungen mit den vier internen Parametern x_0, y_0, s_x und s_y als Unbekannte. Mit mindestens einem Referenzpunkt und zwei Rotationen oder zwei Referenzpunkten und einer Rotation ist ein daraus resultierendes nichtlineares Gleichungssystem mit iterativen Verfahren lösbar. Hat man sinnvolle Startwerte gewählt, erhält man die internen Parameter des Bildhauptpunktes und der Skalierungsvariablen.

Diese Gleichungen sind nichtlinear und liegen nicht in geschlossener Form vor. Sie sind des weiteren nicht durch eine Variablensubstitution linearisierbar. Somit ist eine zu lösende Ausgleichsrechnung über viele Referenzpunkte und Rotationen sehr zeitintensiv. Weiterhin kann im allgemeinen nicht davon ausgegangen werden, daß sinnvolle Startwerte für die nichtlineare Ausgleichsrechnung vorgegeben werden können.

3.4.3.1 Vergenzrotation der Kamera

Hier wird besprochen, auf welche Weise sich eine reine Vergenzrotation auf das Verfahren nach (3.35) auswirkt. In diesem Fall gilt für den Neigungswinkel $\phi = 0$. Ausgehend von (3.35) erhält man:

$$\begin{pmatrix} x_{p_{ref}} \\ y_{p_{ref}} \end{pmatrix} = \begin{pmatrix} z^{-1}(s_x x_p' \cos\theta + {s_x}^2 \sin\theta) + x_0 \\ z^{-1} s_x y_p' \cos\phi + y_0 \end{pmatrix} , \tag{3.36}$$

wobei $z = s_x \cos\theta - x_p' \sin\theta$ und x_p' und y_p' wie in Gleichung (3.35) gewählt seien. Durch Umformen erhält man aus (3.36) zwei nichtlineare Gleichungen mit den Unbekannten x_0, y_0 und s_x:

$${s_x}^2 + {x_0}^2 + \cot\theta(x_p - x_{p_{ref}})s_x - (x_p + x_{p_{ref}})x_0 + x_p x_{p_{ref}} = 0 \tag{3.37}$$

$$\sin^{-1}\theta(\cos\theta - 1)s_x y_0 + \sin^{-1}\theta(y_p - y_{p_{ref}}\cos\theta)s_x$$
$$+x_0 y_0 - y_{p_{ref}} x_0 - x_p y_0 + x_p y_{p_{ref}} = 0 \qquad (3.38)$$

Mit mindestens einem Referenzpunkt und zwei Rotationen oder zwei Referenzpunkten und einer Rotation ist ein daraus resultierendes nichtlineares Gleichungssystem mit iterativen Verfahren lösbar. Im Gegensatz zur allgemeinen Rotation um eine beliebige Achse ist dieses Gleichungssystem überbestimmt.

Schränkt man also die Rotationsbewegung auf die Vergenzachse ein, läßt sich der Skalierungsfaktor s_y nicht berechnen. Außerdem stellt die Rotation um die Vergenzachse dann ein nichtlineares Problem dar.

Gleichung (3.37) hängt nur von den internen Parametern x_0 und s_x ab und weist keine gemischten Terme auf. Gleichung (3.38) hängt von den internen Parametern x_0, y_0 und s_x ab und weist gemischte Terme auf.

Da in (3.37) keine gemischten Terme vorhanden sind und da vor den quadratischen Termen ${x_0}^2$ und ${s_x}^2$ jeweils der gleiche Koeffizient steht, läßt sich (3.37) durch Einführen einer neuen Variablen $s = {s_x}^2 + {x_0}^2$ wie folgt umformen:

$$s + \cot\theta\left(x_p - x_{p_{ref}}\right)s_x - \left(x_p + x_{p_{ref}}\right)x_0 + x_p x_{p_{ref}} = 0\ . \qquad (3.39)$$

Man erhält eine lineare Gleichung mit den drei Unbekannten x_0, s_x und s. Mit mindestens einem Referenzpunkt und drei Rotationen oder drei Referenzpunkten und einer Rotation ist ein daraus resultierendes lineares Gleichungssystem lösbar. Man erhält die internen Parameter x_0 und s_x. Die Nebenbedingung $s = {s_x}^2 + {x_0}^2$ wird nicht berücksichtigt.

Führt man zunächst eine Ausgleichsrechnung für mehrere Rotationen und Referenzpunkte nach (3.39) durch, erhält man Startwerte für x_0 und s_x, um ein nichtlineares Ausgleichsverfahren anzuwenden, das aus (3.37) resultiert.

Da auch nach einer Ausgleichsrechnung nach (3.39) kein Startwert für y_0 vorliegt, ist die Einbeziehung von (3.38) in ein nichtlineares Verfahren zunächst nicht möglich. Möchte man auch (3.38) einbeziehen, ist ein geeigneter Startwert auf andere Weise zu berechnen.

Die in diesem Paragraphen geschilderte Methode liefert, linear oder nichtlinear, nur Werte für den Skalierungsfaktor s_x und die x-Koordinate des Bildhauptpunktes.

3.4.3.2 Neigungsrotation der Kamera

In diesem Abschnitt wird besprochen, welche Auswirkungen eine reine Neigungsrotation auf das Verfahren nach (3.35) hat. In diesem Fall gilt für den Vergenzwinkel $\theta = 0$. Ausgehend von (3.35) erhält man:

$$\begin{pmatrix} x_{p_{ref}} \\ y_{p_{ref}} \end{pmatrix} = \begin{pmatrix} z^{-1}s_y x_p' + x_0 \\ z^{-1}(s_y y_p' \cos\phi - {s_y}^2 \sin\phi) + y_0 \end{pmatrix}\ , \qquad (3.40)$$

wobei $z = s_y \cos\phi + y_p' \sin\phi$ und x_p' und y_p' wie in (3.35). Durch Umformen erhält man aus (3.40) zwei nichtlineare Gleichungen mit den drei Unbekannten x_0, y_0 und s_y:

$$
\begin{aligned}
&{s_y}^2 + {y_0}^2 - \cot\phi(y_p - y_{p_{ref}})s_y - (y_p + y_{p_{ref}})y_0 + y_p y_{p_{ref}} = 0 && (3.41)\\
&\sin^{-1}\phi(\cos\phi - 1)s_y x_0 + \sin^{-1}\phi(x_p - x_{p_{ref}}\cos\phi)s_y \\
&\qquad\qquad -x_0 y_0 + y_p x_0 + x_{p_{ref}} y_0 - y_p x_{p_{ref}} = 0\ . && (3.42)
\end{aligned}
$$

Wiederum ist ein daraus resultierendes nichtlineares Gleichungssystem durch iterative Verfahren mit mindestens einem Referenzpunkt und zwei Rotationen oder zwei Referenzpunkten und einer Rotation lösbar. In diesem Fall ist neben den internen Parametern des Bildhauptpunktes der Skalierungsfaktor s_y bestimmbar, wogegen sich s_x nicht berechnen läßt.

Weiter läßt sich (3.41) durch Einführen einer neuen Variablen $s = {s_y}^2 + {y_0}^2$ umformen:

$$s - \cot\phi(y_p - y_{p_{ref}})s_y - (y_p + y_{p_{ref}})y_0 + y_p y_{p_{ref}} = 0\ . \qquad (3.43)$$

Äquivalent zur Vergenzrotation erhält man eine lineare Gleichung mit den drei Unbekannten y_0, s_y und s. Durch Lösen eines linearen Gleichungssystems erhält man die beiden internen Parameter y_0 und s_y. Die Nebenbedingung bleibt wiederum unberücksichtigt.

3.4.3.3 Kombination von Neigungs- und Vergenzrotation

Führt man eine Vergenzrotation durch, lassen sich die internen Parameter x_0 und s_x linear berechnen. Führt man eine Neigungsrotation durch, lassen sich die internen Parameter y_0 und s_y linear berechnen. Man erhält somit alle relevanten internen Parameter der Kamera.

Anschließend ist es möglich, auch nichtlineare Verfahren unter Einbeziehung der Gleichungen (3.38) oder (3.42) durchzuführen, da für alle Parameter Startwerte vorliegen. Ein nichtlineares Verfahren unter Einbezug von (3.37) und (3.38) liefert die internen Parameter x_0, y_0 und s_x. Ein nichtlineares Verfahren unter Einbeziehung von (3.41) und (3.42) liefert die internen Parameter x_0, y_0 und s_y. Da bei der getrennten Ausführung die internen Parameter des Bildhauptpunktes zweimal bestimmt werden, ist es sinnvoll, nur eine nichtlineare Ausgleichsrechnung unter Einbeziehung aller vier Gleichungen durchzuführen. Es kann allerdings auch eine anschließende Ausgleichsrechnung über den Schwerpunkt, gewichtet mit den entsprechenden Varianzen, stattfinden.

3.4.4 Kalibrierung mittels unbekannter Rotation

In diesem Abschnitt wird das Hauptresultat dieser Arbeit besprochen. Das hier vorgestellte Kalibrierverfahren nutzt genauso wie das im vorhergehenden Abschnitt besprochene Verfahren kinematische Freiheitsgrade aus. Im Gegensatz zu

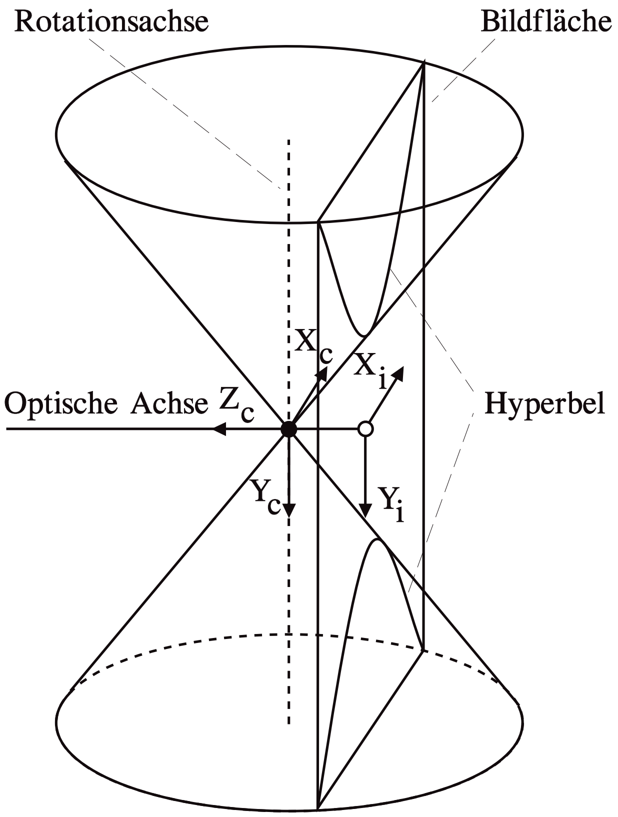


Abbildung 3.3: Durch den Schnitt eines Rotationskegels um die x-Achse des Kamerakoordinatensystems mit einer zur xy-Ebene des Kamerakoordinatensystems parallelen Ebene ensteht ein Kegelschnitt. In diesem Fall ist der Schnitt eine Hyperbel.

diesem Verfahren wird hier keine Kenntnis der Rotationswinkel zwischen mehreren Aufnahmen vorausgesetzt.

Die Idee, die der in diesem Abschnitt beschriebenen Methode zugrunde liegt, geht auf die Anpassung eines Kegelschnitts an eine Punktfolge zurück. Detektiert man ein Szenenmerkmal unter Rotation einer Kamera um eine feste Rotationsachse in mehreren Bildern, so liegen diese Punkte auf einem Kegelschnitt. Der Rotationskegel ist dabei abhängig von einem detektierten Referenzpunkt und der Rotationsachse. Der Kegelschnitt entsteht dann durch den Schnitt des Rotationskegels mit der Sensorfläche der Kamera (Abb. 3.3). Wie im vorhergehenden Verfahren wird hier eine getrennte Neigungs- und Vergenzrotation durchgeführt, um alle internen Kameraparameter zu erhalten.

Man nehme an, daß sich eine Kamera um eine normierte Rotationsachse $\widehat{\boldsymbol{N}} = (n_x, n_y, n_z)^T$ durch den Ursprung des Kamerakoordinatensystems bewegt ($||\widehat{\boldsymbol{N}}|| = 1$). Dann beschreibt der Sehstrahl zu einem Szenenpunkt im Kamerakoordinatensystem einen Rotationskegel. Dieser Kegel ist abhängig vom Öffnungswinkel ω, der sich aus der Lage des betrachteten Szenenpunktes und der Rotationsachse ergibt. Die Kegelspitze befindet sich im optischen Zentrum (Abb. 3.3). Der Rotationskegel hat für beliebige Punkte $\boldsymbol{P_c}$ im Kamerakoordinatensystem die Form:

$$\left(\widehat{\boldsymbol{N}}^T \boldsymbol{P_c}\right)^2 = \cos^2\omega \|\boldsymbol{P_c}\|^2 \ . \tag{3.44}$$

Daraus ergibt sich folgende Umformung:

$$\begin{aligned} (3.44) \quad \Leftrightarrow \quad & \boldsymbol{P_c}^T \widehat{\boldsymbol{N}} \widehat{\boldsymbol{N}}^T \boldsymbol{P_c} - \cos^2\omega \boldsymbol{P_c}^T \boldsymbol{P_c} = 0 \\ \Leftrightarrow \quad & \boldsymbol{P_c}^T \left(\widehat{\boldsymbol{N}}\widehat{\boldsymbol{N}}^T - \cos^2\omega I\right) \boldsymbol{P_c} = 0 \ . \end{aligned} \tag{3.45}$$

Dabei ist I die Einheitsmatrix. Mit Gleichung (2.18) ergibt sich dann

$$\begin{aligned} (3.45) \quad \Leftrightarrow \quad & \left(S^{-1}\boldsymbol{P_p}\right)^T \left(\widehat{\boldsymbol{N}}\widehat{\boldsymbol{N}}^T - \cos^2\omega I\right) S^{-1}\boldsymbol{P_p} = 0 \\ \Leftrightarrow \quad & \boldsymbol{P_p}^T S^{-1^T} \left(\widehat{\boldsymbol{N}}\widehat{\boldsymbol{N}}^T - \cos^2\omega I\right) S^{-1}\boldsymbol{P_p} = 0 \ , \end{aligned} \tag{3.46}$$

wobei

$$S = \begin{pmatrix} s_x & 0 & x_0 \\ 0 & s_y & y_0 \\ 0 & 0 & 1 \end{pmatrix} \ .$$

Diese Gleichung repräsentiert den Schnitt eines Kegels mit einer Ebene. Die fünf Koeffizienten des Kegelschnitts hängen von den beiden Parametern der Rotationsachse, den vier internen Parametern der Kamera und dem Winkel zwischen einem Szenenmerkmal und der Rotationsachse ab.

Beardsley et al. [BMZ92] zeigen, daß mehr als ein Kegelschnitt benötigt wird, um die internen Parameter einer Kamera zu berechnen. Die unterschiedlichen Kegelschnitte haben auf verschiedenen Rotationsachsen zu beruhen. Daher werden im folgenden Spezialfälle betrachtet, für die die Rotationsachsen bekannt sind und mit denen sich sämtliche internen Kameraparameter bestimmen lassen.

Um mit diesem Verfahren die internen Kameraparameter zu bestimmen, benötigt man einen Kegelschnitt. Dieser berechnet sich aus einer Folge von Punkten. Es werden somit Bildmerkmale einander zugeordnet, die alle dasselbe Szenenmerkmal repräsentieren müssen. Durch

$$\mathcal{Z} \subseteq \mathcal{B} \tag{3.47}$$

kann diese Zuordnung gemäß (3.4) beschrieben werden.

3.4.4.1 Vergenzrotation der Kamera

Im Fall einer Vergenzrotation hat die Rotationsachse die Form $\widehat{\boldsymbol{N}} = (0,1,0)^T$. Durch Einsetzen erhält man aus (3.46):

$$\frac{(y_p - y_0)^2}{{s_y}^2}\,\frac{1-\cos^2\omega}{\cos^2\omega} - \frac{(x_p - x_0)^2}{{s_x}^2} = 1\ . \tag{3.48}$$

Allgemein gilt für den Winkel α zwischen zwei Vektoren $\boldsymbol{u}$ und $\boldsymbol{v}$: $\cos\alpha = \frac{\boldsymbol{u}^T\boldsymbol{v}}{||\boldsymbol{u}||\,||\boldsymbol{v}||}$. Somit gilt für einen Referenzpunkt $(x_{ref}, y_{ref})^T$ im Kamerakoordinatensystem $\cos^2\omega = \frac{{y_{ref}}^2}{1+{y_{ref}}^2}$. Setzt man hier (2.19) ein, so erhält man aus (3.48) für einen Referenzpunkt $(x_{p_{ref}}, y_{p_{ref}})^T$ im Rechnerkoordinatensystem:

$$\frac{{s_x}^2 + (x_0 - x_p)^2}{{s_x}^2 + {x_0}^2} = \frac{(y_0 - y_p)^2}{\left(y_0 - y_{p_{ref}}\right)^2}\ . \tag{3.49}$$

Setzt man in Gleichung (2.4) $\phi = 0$, so erhält man zwei Gleichungen. Eliminiert man aus diesen den Rotationswinkel θ und setzt in die resultierende Gleichung (2.19) ein, so erhält man eine Gleichung mit dem Referenzpunkt $(x_{p_{ref}}, y_{p_{ref}})^T$. Dieser Referenzpunkt besitzt einen Freiheitsgrad. Legt man diesen durch $x_{p_{ref}} = 0$ fest, so erhält man ebenfalls (3.49). Durch Umformen von (3.49) erhält man eine Hyperbel:

$$a{x_p}^2 + b{y_p}^2 + cx_p + dy_p + 1 = 0\ , \tag{3.50}$$

wobei

$$\begin{aligned}
a &= -e^{-1}\left(y_0 - y_{p_{ref}}\right)^2\ ,\\
b &= e^{-1}\left({s_x}^2 + {x_0}^2\right)\ ,\\
c &= 2e^{-1}x_0\left(y_0 - y_{p_{ref}}\right)^2\ ,\\
d &= -2e^{-1}y_0\left({s_x}^2 + {x_0}^2\right)\ ,\\
e &= \left({s_x}^2 + {x_0}^2\right)y_{p_{ref}}\left(y_{p_{ref}} - 2y_0\right)\ .
\end{aligned}$$

Bei bekannten Hyperbelkoeffizienten besitzt dieses Gleichungssystem die vier Unbekannten x_0, y_0, s_x und $y_{p_{ref}}$, wobei die y-Koordinate des Referenzpunktes nicht von Interesse ist.

Aus (3.50) erhält man die internen Parameter des Bildhauptpunktes durch:

$$x_0 = -\tfrac{1}{2}ca^{-1} \text{ und } y_0 = -\tfrac{1}{2}db^{-1}\ . \tag{3.51}$$

Man entnimmt daraus die bekannte Tatsache, daß sich der Schnittpunkt der Asymptoten einer Hyperbel aus (3.51) ergibt. Der Skalierungsfaktor s_x ergibt sich aus (3.50) zu:

$${s_x}^2 = \frac{1}{a} - \frac{d^2}{4ab} - \frac{c^2}{4a^2}\ . \tag{3.52}$$

Gleichung (3.50) entnimmt man weiter, daß die durch die perspektivische Transformation abgebildete Hyperbel nicht gedreht ist und somit kein gemischter Term der Form $x_p y_p$ auftaucht.

Stellt man ein Gleichungssystem gemäß (3.50) mit den bekannten Parametern x_p und y_p und den Unbekannten a, b, c und d auf, läßt sich mit mindestens drei Rotationen eine Hyperbel mit linearen Methoden bestimmen (siehe Kap. 4.2.4, [DE96a], [DE96b]).

Auch hier läßt sich der Skalierungsfaktor s_y durch eine alleinige Vergenzrotation nicht bestimmen. Somit muß der interne Parameter s_y mit anderen Methoden berechnet werden.

Möchte man mehr als einen Referenzpunkt in die lineare Berechnung der internen Kameraparameter einbeziehen, sind die internen Kameraparameter für jeden Referenzpunkt durch die korrespondierende Hyperbel zu berechnen. Anschließend hat man eine Ausgleichsrechnung über die verschiedenen Lösungen von internen Parametern durchzuführen. In diesem Fall ist es angebracht, die verschiedenen Lösungen mit ihren Varianzen zu gewichten.

Eine nichtlineare Berechnung der internen Kameraparameter x_0, y_0 und s_x durch iterative Verfahren kann im Gegensatz zu der im vorherigen Kapitel beschriebenen Methode direkt mit (3.50) erfolgen, da durch lineare Verfahren Startwerte für alle Parameter berechenbar sind. Nichtlineare iterative Verfahren lassen es zu, mehrere Hyperbeln direkt in eine Berechnung einfließen zu lassen.

3.4.4.2 Neigungsrotation der Kamera

Im Fall einer Neigungsrotation hat die Rotationsachse die Form $\widehat{\boldsymbol{N}} = (1,0,0)^T$. Durch Einsetzen erhält man aus (3.46):

$$\frac{(x_p - x_0)^2}{{s_x}^2} \, \frac{1 - \cos^2\omega}{\cos^2\omega} + \frac{(y_p - y_0)^2}{{s_y}^2} = 1 \tag{3.53}$$

Für einen Referenzpunkt $(x_{ref}, y_{ref})^T$ im Kamerakoordinatensystem gilt dann $\cos^2\omega = \frac{{x_{ref}}^2}{1+{x_{ref}}^2}$. Durch Einsetzen von (2.19) erhält man aus (3.53):

$$\frac{{s_y}^2 + (y_0 - y_p)^2}{{s_y}^2 + {y_0}^2} = \frac{(x_0 - x_p)^2}{\left(x_0 - x_{p_{ref}}\right)^2} \; . \tag{3.54}$$

Geht man von Gleichung (2.4) aus, setzt in diese $\theta = 0$ ein, eliminiert den Rotationswinkel ϕ und setzt in die resultierende Gleichung (2.19) ein, so erhält man Gleichung (3.54). Dazu lege man den Freiheitsgrad des Referenzpunktes durch $y_{p_{ref}} = 0$ fest. Umformen von (3.54) ergibt eine Hyperbel:

$$a{x_p}^2 + b{y_p}^2 + cx_p + dy_p + 1 = 0 \; , \tag{3.55}$$

wobei

$$\begin{aligned}
a &= e^{-1}\left({s_y}^2 + {y_0}^2\right) \ , \\
b &= -e^{-1}\left(x_0 - x_{p_{ref}}\right)^2 \ , \\
c &= -2e^{-1}x_0\left({s_y}^2 + {y_0}^2\right) \ , \\
d &= 2e^{-1}y_0\left(x_0 - x_{p_{ref}}\right)^2 \ , \\
e &= \left({s_y}^2 + {y_0}^2\right)x_{p_{ref}}\left(x_{p_{ref}} - 2x_0\right) \ .
\end{aligned}$$

Aus (3.50) lassen sich, genauso wie bei der Vergenzrotation, die internen Parameter des Bildhauptpunktes durch (3.51) berechnen. Der Skalierungsfaktor s_y ergibt sich aus (3.55) zu:

$$ {s_y}^2 = \frac{1}{b} - \frac{c^2}{4ab} - \frac{d^2}{4b^2} \ . \tag{3.56}$$

Der Skalierungsfaktor s_x ist nicht bestimmbar. Auch in diesem Fall ist die durch die perspektivische Transformation abgebildete Hyperbel nicht gedreht, und die Hyperbelparameter lassen sich erneut durch lineare Methoden bestimmen. Die nichtlineare Bestimmung der internen Parameter aus den Hyperbelkoeffizienten erfolgt analog zur nichtlinearen Bestimmung der internen Parameter bei einer Vergenzrotation.

3.4.4.3 Kombination von Neigungs- und Vergenzrotation

Führt man eine alleinige Vergenzrotation durch, so lassen sich die internen Parameter x_0, y_0 und s_x linear berechnen. Führt man eine alleinige Neigungsrotation durch, so lassen sich die internen Parameter x_0, y_0 und s_y linear berechnen. Um sämtliche internen Parameter zu berechnen, muß somit sowohl eine Vergenz- als auch eine Neigungsrotation durchgeführt werden. Die Berechnung der internen Parameter kann dann wie folgt geschehen:

1. Ausführung einer Vergenzrotation um die y-Achse des Kamerakoordinatensystems mit anschließender Anpassung einer Hyperbel. Aus den Hyperbelkoeffizienten berechne man das Zentrum und die Achsen der Hyperbel. Anschließend berechne man die internen Parameter x_0 und s_x aus diesen Koeffizienten.

2. Ausführung einer Neigungsrotation um die x-Achse des Kamerakoordinatensystems mit anschließender Anpassung einer Hyperbel und Berechnung des Zentrums und der Achsen. Daraus berechne man die internen Parameter y_0 und s_y.

Als Alternative können die internen Parameter des Bildhauptpunktes durch eine Ausgleichsrechnung über die Werte beider Rotationen berechnet werden.

Nutzt man nichtlineare iterative Verfahren, so lassen sich beide Rotationen zusammenfassen, und sämtliche internen Parameter werden durch einmaliges Anwenden eines iterativen nichtlinearen Verfahrens ermittelt.

3.4.5 Lösungsansätze anderer Autoren

In diesem Abschnitt wird auf Lösungsansätze anderer Autoren eingegangen, die Bewegungen entweder einer Kamera oder von Gegenständen im Blickfeld einer Kamera ausnutzen.

In dieser Übersicht werden die Ansätze von Maybank und Faugeras [MF92], Luong und Viéville [LV94] und Hartley [Har94] nicht beschrieben. Diese Verfahren bestimmen die Kameraparameter als Nebenprodukt einer Bewegungsbestimmung.

Zunächst wird auf die Ansätze von Du und Brady [DB93] und von Basu [Bas95] eingegangen, die jeweils den optischen Fluß kleiner Rotationen ausnutzen. Basu [Bas95] geht dabei von zwei unabhängigen Rotationen aus (Pan und Tilt), benutzt Konturen statt Punkte und initialisiert den Bildhauptpunkt mit einem geschätzten Wert. Du und Brady [DB93] hingegen nutzen die Hyperbel der Trajektorien von Punkten aus, um eine stabilere Berechnung des Bildhauptpunktes zu gewährleisten.

Anschließend werden die Ansätze von Stein [Ste95] und Viéville [Vié94] besprochen. Beide Ansätze lösen nichtlineare Gleichungen, die auf Bewegungen mit bekannten Rotationswinkeln um eine feste Rotationsachse beruhen. Diese Verfahren sind dem in Kapitel 3.4.3 beschriebenen Verfahren am ähnlichsten.

Schließlich wird ein Verfahren von Beardsley et al. [BMZ92] vorgestellt. In diesem Verfahren werden die Trajektorien von Fluchtpunkten benutzt, die durch Bewegungen eines Rotationskörpers um eine feste Achse im Blickfeld der Kamera rotieren. Dieses Verfahren ist dem in Kapitel 3.4.4 beschriebenen Verfahren am ähnlichsten.

3.4.5.1 Das Verfahren von Basu

Wie auch die in Kapitel 3.4.3 und 3.4.4 beschriebenen Verfahren geht auch das Verfahren von Basu [Bas95] von der Gleichung (2.3) aus. Er führt ebenfalls eine getrennte Vergenz- und Neigungsrotation durch. Zunächst werden wir die Neigungsrotation beschreiben. Basu [Bas95] nimmt an, daß nur kleine Rotationen durchgeführt werden. Dadurch vereinfacht sich die Rotationsmatrix der Neigung aus Gleichung (2.3) zu

$$R_\phi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\phi \\ 0 & \phi & 1 \end{pmatrix} . \tag{3.57}$$

Wegen der kleinen Rotationswinkel kann eine Taylor-Entwicklung des Nenners der Gleichung (3.40) durchgeführt werden, was zu den Gleichungen

$$x_p{}' \approx x_p \left(1 + \phi \frac{y_p}{s_y}\right) \text{ und } y_p{}' \approx (y_p + \phi s_y) \left(1 + \phi \frac{y_p}{s_y}\right) \tag{3.58}$$

führt. Dabei ist $(x_p{}', y_p{}')^T)$ ein Bildpunkt nach der Neigungsrotation. Führt man selbiges Verfahren für einen Vergenzschwenk durch, so erhält man die Gleichungen

$$x_p{}' \approx (x_p + \theta s_x) \left(1 + \theta \frac{x_p}{s_x}\right) \text{ und } y_p{}' \approx y_p \left(1 + \theta \frac{x_p}{s_x}\right) \ , \tag{3.59}$$

wobei $(x_p{}', y_p{}')^T)$ ein Bildpunkt nach der Vergenzrotation ist.

Anschließend nimmt Basu [Bas95] an, daß der Bildhauptpunkt bis auf einen Fehler $(\delta_x, \delta_y)^T$ bekannt ist. Des weiteren geht er von der Detektion einer Kontur aus[1], das bedeutet, es werden nicht wie in den bereits bekannten Verfahren Bildpunkte, sondern Konturen einander zugeordnet.

Zunächst beschreibt Basu [Bas95] ein Verfahren, das zuerst den Bildhauptpunkt korrigiert, um anschließend die Skalierungsfaktoren zu berechnen. Auf dieses Verfahren wird nicht näher eingegangen, da es aufgrund der Berechnung der Differenzen zwischen den Punkten vor und nach der Rotation instabil ist. Dies liegt an der Tatsache, daß für einen Neigungsschwenk $(x_p{}' - x_p) \approx 0$ gilt. Dadurch wirken sich Fehler in der Konturdetektion sehr stark aus.

In einem weiteren Verfahren werden zunächst die Skalierungsfaktoren berechnet. Durch die Annahme des bekannten Bildhauptpunktes mit Abweichung $(\delta_x, \delta_y)^T$ ersetzt man zunächst jeweils jede x-, bzw y-Koordinate durch die Abweichung

$$(x_p - \delta_x) \text{ und } (y_p - \delta_y) \ . \tag{3.60}$$

Durch die Annahme, daß die Abweichung $(\delta_x, \delta_y)^T$ relativ klein ist, kann man die erhaltenen Gleichungen weiter vereinfachen. Dabei werden Terme, in denen δ_x oder δ_y vorkommen, eliminiert. Man erhält so aus der zweiten Gleichung von (3.58) für einen Neigungsschwenk

$$s_y{}^2 - s_y \frac{\overline{y_p}' - \overline{y_p}(1 + \phi^2)}{\phi} + \overline{y_p{}^2} = 0 \tag{3.61}$$

und aus der ersten Gleichung von (3.59) für einen Vergenzschwenk

$$s_x{}^2 - s_x \frac{\overline{x_p}' - \overline{x_p}(1 + \theta^2)}{\theta} + \overline{x_p{}^2} = 0 \ . \tag{3.62}$$

[1]Im folgenden werden Variablen, die eine Kontur bezeichnen, mit einem „$^-$" versehen, um sie von den ursprünglichen Variablen abzugrenzen.

Anschließend kann der Bildhauptpunkt $(x_0, y_0)^T$ aus (3.60) abgeleitet werden, indem man eine zweite Kontur bestimmt und die bekannten Skalierungsfaktoren substituiert. Dies geschieht mit den Gleichungen

$$s_y \frac{\overline{x_p}' - \overline{x_p}}{\phi} = \overline{x_p y_p} - \overline{x_p}\delta_y - \overline{y_p}\delta_x \quad \text{und} \tag{3.63}$$

$$s_y \frac{\overline{x_p}' - \overline{x_p}}{\theta} = \overline{x_p y_p} - \overline{x_p}\delta_y - \overline{y_p}\delta_x \tag{3.64}$$

gleichfalls aus einer Vergenz- bzw. Neigungsrotation.

Die Annahme des bekannten Bildhauptpunktes ist eine Einschränkung des Verfahrens. Bevor dieses Verfahren angewandt werden kann, müssen Ergebnisse für den Bildhauptpunkt vorliegen. Dieser wird durch das Verfahren lediglich verbessert. Die Zuordnung der Kontur wird wenig besprochen, so daß ein fehlerspezifischer Vergleich zu der üblichen Methode der Bestimmung von korrespondierenden Punkten oder Fluchtpunkten nicht möglich ist. Die einzigen freien Parameter, die unabhängig bestimmt werden, sind die Skalierungsfaktoren. Des weiteren müssen bekannte Winkel für die Rotation zwischen zwei Konturen vorliegen. Die Voraussetzungen sind somit restriktiver als in dem in Kapitel 3.4.3 vorgestellten Verfahren.

3.4.5.2 Das Verfahren von Du und Brady

Auch bei dem Verfahren von Du und Brady [DB93] wird eine getrennte Vergenz- und Neigungsrotation durchgeführt. Sei $\Omega = (\omega_x, \omega_y, \omega_z)^T$ eine Rotation und $V = \{v_x, v_y, v_z\}$ eine Translation. Wenn das optische Zentrum der Kamera einer allgemeinen Rotation und Translation unterliegt, ergibt sich für einen Bildpunkt im Rechnerkoordinatensystem:

$$\begin{pmatrix} k_x x_p' \\ k_y y_p' \end{pmatrix} = \begin{pmatrix} \frac{k_x x_p v_z - v_x}{Z} + k_x k_y x_p y_p \omega_x - \left(k_x{}^2 x_p{}^2 + 1\right)\omega_y + k_y y_p \omega_z \\ \frac{k_y y_p v_z - v_y}{Z} + \left(k_y{}^2 y_p{}^2 + 1\right)\omega_x - k_x k_y x_p y_p \omega_y - k_x x_p \omega_z \end{pmatrix},$$

wobei k_x und k_y in diesem Fall normalisierte Skalierungsfaktoren in x- und y-Richtung sind, die die Pixelkoordinaten in die Ebene $z = 1$ des Kamerakoordinatensystems transformieren. Der Bildpunkt vor der Bewegung sei durch (x_p, y_p), der nach der Bewegung durch (x_p', y_p') bezeichnet. Im folgenden wird eine alleinige Rotation und keine Translation angenommen. Für eine Neigungsrotation erhält man aus obiger Gleichung:

$$\begin{pmatrix} k_x \Delta x_p \\ k_y \Delta y_p \end{pmatrix} = \begin{pmatrix} k_x k_y (x_p - x_0)(y_p - y_0) \\ 1 + k_y{}^2 (y_p - y_0)^2 \end{pmatrix}, \tag{3.65}$$

wobei $\Delta x_p = \frac{x_p'}{\omega_x}$ und $\Delta y_p = \frac{y_p'}{\omega_x}$ den optischen Fluß in Pixel geteilt durch die relative Rotation um die x-Achse (ω_x) beschreiben. Mit mindestens drei Punkten kann dieses Gleichungssystem gelöst werden. Man erhält y_0 und den Skalierungsfaktor k_y . Die Bestimmung von x_0 und des Skalierungsfaktors k_x wird mit einer Vergenzrotation und zu (3.65) analogen Gleichungen durchgeführt.

Bis zu diesem Punkt ist dieses Verfahren dem in Kapitel 3.4.3 beschriebenen Verfahren sehr ähnlich. Du und Brady [DB93] untersuchen jetzt allerdings den Spezialfall, daß sich die untersuchten Bildpunkte nahe der x-Achse bei einer Neigungsrotation oder nahe der y-Achse bei einer Vergenzrotation befinden. Dies führt zu den wesentlich einfacheren Gleichungen[2]:

$$k_y = \Delta {y_p}^{-1} \text{ und } k_x = \Delta {x_p}^{-1} \tag{3.66}$$

zur Bestimmung der Skalierungsfaktoren mit nur einem Punkt. Anschließend läßt sich der Bildhauptpunkt durch

$$\begin{array}{llll} x_0 = \frac{x_{p_1} - \Gamma_x x_{p_2}}{1-\Gamma_x} & \text{mit} & \Gamma_x = \frac{\Delta x_{p_1}}{\Delta x_{p_2}} & \text{und} \\ y_0 = \frac{y_{p_1} - \Gamma_y y_{p_2}}{1-\Gamma_y} & \text{mit} & \Gamma_x = \frac{\Delta y_{p_1}}{\Delta y_{p_2}} & \text{finden}^2, \end{array} \tag{3.67}$$

wobei zwei Punkte $(x_{p1}, y_{p1})^T$ und $(x_{p2}, y_{p2})^T$ betrachtet werden müssen. Die Skalierungsfaktoren können also mit einem Punkt berechnet werden. Um den Bildhauptpunkt zu ermitteln benötigt man zwei Punkte. Zur Berechnung des optischen Flusses wird vorausgesetzt, daß eine merkmalsbasierte Methode verwendet wird [Nag83].

In einem dritten Schritt stellen Du und Brady [DB93] ein Verfahren vor, das durch Ausnutzen von Trajektorien von Strukturen zunächst nur den Bildhauptpunkt bestimmt. Dazu werden zunächst Bildmerkmale detektiert, die sich nicht in der Nähe des Bildhauptpunktes befinden. Anschließend läßt man die Kamera um die Vergenz- oder Neigungsachse rotieren. Man erhält eine Trajektorie, die die Hälfte einer Hyperbel darstellt. Durch ein Anpassungsverfahren wird eine Hyperbelhälfte bestimmt, und anschließend werden die Symmetrieachsen und somit der Bildhauptpunkt bestimmt. Man erhält x_0 aus einer Vergenz- und y_0 aus einer Neigungsrotation.

Zusammenfassend kann gesagt werden, daß dieses Verfahren in den ersten beiden Schritten jeweils Zuordnungen von Bildpunkten wie auch die bekannten Rotationen benötigt. Dann ist es in beiden Fällen möglich, alle vier unbekannten Parameter zu berechnen. Der zweite Schritt schränkt die Umgebung dadurch weiter ein, daß der optische Fluß in der Umgebung des Bildhauptpunktes vorausgesetzt wird. Im dritten Schritt ist eine Kenntnis der Rotation nicht mehr notwendig, es wird allerdings vorausgesetzt, daß sich die detektierten Trajektorien jenseits des Bildhauptpunktes befinden. Berechnet wird in diesem Fall nur der Bildhauptpunkt.

[2] Die erste Gleichung resultiert aus einer Neigungsrotation, die zweite Gleichung aus einer Vergenzrotation.

3.4.5.3 Das Verfahren von Viéville

Viéville [Vié94] beschreibt zunächst ein polynomiales Modell einer Kamerabewegung. Dieses Modell basiert auf der Rodriguez-Formel für die Rotation von Punkten:

$$R(\boldsymbol{u}, \delta) = I + \sin\delta\tilde{U} + (1 - \cos\delta)\tilde{U}^2 \ . \tag{3.68}$$

Dabei ist I die Einheitsmatrix und $R(\boldsymbol{u}, \delta)$ eine Rotationsmatrix, abhängig von einer normierten Rotationsachse $\boldsymbol{u} = (u_x, u_y, u_z)^T$ und einem Rotationswinkel δ. Des weiteren ergibt sich die antisymmetrische Matrix $\tilde{U}$ aus der Rotationsachse:

$$\tilde{U} = \begin{pmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{pmatrix} \ .$$

Das aus (3.68) resultierende polynomiale Modell besitzt neun Unbekannte. Sie seien durch die Vektoren $\boldsymbol{O} = (O_1, O_2, O_3)^T$, $\boldsymbol{S} = (S_1, S_2, S_3)^T$ und $\boldsymbol{C} = (C_1, C_2, C_3)^T$ gegeben und repräsentieren implizit die vier internen Kameraparameter, die Rotationsachse und den Rotationswinkel. Die Parameter $\boldsymbol{O}$, $\boldsymbol{S}$ und $\boldsymbol{C}$ sind bis auf einen Skalierungsfaktor unter der Nebenbedingung $O_3 + C_3 = 1$ eindeutig bestimmbar und lassen sich linear mit mindestens vier Punkten aus einer unbekannten Rotation um eine feste Rotationsachse bestimmen. Die internen Parameter ergeben sich anschließend aus:

$$\begin{aligned}
x_{p_1}O_3 + 2x_{p_2}S_3 &= 2/3x_{p_1} - 2x_{p_3} \ , \\
y_{p_1}O_3 + 2y_{p_2}S_3 &= 2/3x_{p_1} - 2x_{p_3} \ , \\
C_3 &= 1 - O_3 \ , \\
O_1 &= 2x_{p_2} + x_{p_0}O_3 + 2x_{p_1}S_3 \ , \\
S_1 &= x_{p_1} + x_{p_0}S_3 \ , \\
C_1 &= -2x_{p_2} + x_{p_0}C_3 - 2x_{p_1}S_3 \ , \\
O_2 &= 2y_{p_2} + y_{p_0}O_3 + 2y_{p_1}S_3 \ , \\
S_2 &= y_{p_1} + y_{p_0}S_3 \ , \\
C_2 &= -2y_{p_2} + y_{p_0}C_3 - 2y_{p_1}S_3 \ .
\end{aligned} \tag{3.69}$$

Dabei sind x_{p_i} und y_{p_i} mit $i \in \{1 \cdots 4\}$ die Rechnerkoordinaten eines Szenenmerkmals zu verschiedenen Zeitpunkten. Zur Verbesserung der ermittelten Werte für die internen Parameter schlägt Viéville [Vié94] ein nichtlineares Verfahren vor. Dabei wird die Rotation als bekannt vorausgesetzt. Die Bestimmung der internen Parameter des Bildhauptpunktes $(x_0, y_0)^T$ und der Skalierungsfaktoren s_x und s_y geschieht mit diesem Verfahren durch:

$$\begin{aligned}
0 &= \frac{(C_1 - x_0C_3)(S_1 - x_0S_3)}{{s_x}^2} + \frac{(C_2 - y_0C_3)(S_2 - y_0S_3)}{{s_y}^2} + C_3C_2 \text{ und} \\
0 &= \frac{(C_1 - x_0C_3)^2 - (S_1 - x_0S_3)^2}{{s_x}^2} + \frac{(C_2 - y_0C_3)^2 - (S_2 - y_0S_3)^2}{{s_y}^2} + {C_3}^2 - S_3 - {S_3}^2 \ .
\end{aligned} \tag{3.70}$$

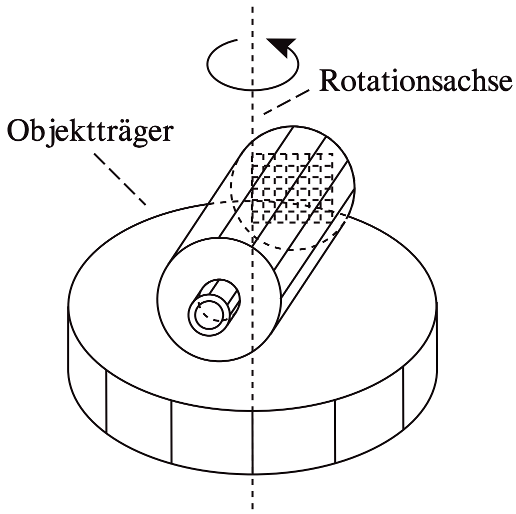


Abbildung 3.4: Objektträger zur Rotation einer Kamera. Stein [Ste95] gibt ein manuelles Verfahren an, um die Kamera so auszurichten, daß das optische Zentrum auf der Rotationsachse liegt.

Um die internen Parameter aus diesen zwei Gleichungen bestimmen zu können, braucht man mindestens zwei unterschiedliche polynomiale Modelle, also mindestens zwei Rotationen.

Dieses Verfahren setzt zur Bestimmung der internen Parameter die Rotation als bekannt voraus. Auf eine bekannte Rotation kann verzichtet werden, falls man nur eine lineare Berechnung der Parameter durchführt. Dieser Fall wird von Viéville [Vié94] nicht untersucht. Detektiert werden jeweils korrespondierende Bildpunkte. Das Verhalten der Methoden unter der Detektion von Fluchtpunkten wird nicht untersucht. Besonders erwähnenswert ist die Tatsache, daß nur eine Rotation durchzuführen ist, um alle internen Parameter zu berechnen. Da allerdings um eine feste Rotationsachse gedreht werden muß, bleibt die Frage zu beantworten, in welcher Weise dies zu geschehen hat. Darauf geht der Artikel ebenfalls nicht ein.

3.4.5.4 Das Verfahren von Stein

Stein [Ste95] beschreibt ein Verfahren, das sich insbesondere dadurch von allen übrigen unterscheidet, daß in diesem keine Kameraplattform gedreht wird, sondern ein Objektträger, auf dem eine Kamera angebracht ist. Da auch hier eine Rotation um das optische Zentrum zu gewährleisten ist, stellt dies ein Problem dar. Stein [Ste95] gibt zur Ermittlung der Position der Kamera auf dem Objektträger ein Verfahren an, das allerdings manuell zu realisieren ist.

Mit dieser Methode sind wiederum zwei Rotationen, eine Neigungs- und eine

Vergenzrotation durchzuführen, um die internen Kameraparameter zu berechnen. Dabei geht Stein [Ste95] von einer allgemeinen Rotation nach der Rodriguez-Formel (3.68) für die Rotation von Punkten aus. Das Verfahren berechnet die internen Parameter der Kamera und die Rotationsachse durch eine nichtlineare Suche nach (3.71). Als interne Kameraparameter werden nicht nur die Skalierungsfaktoren und der Bildhauptpunkt berechnet, sondern auch Faktoren, die die Linsenverzerrungen der Kamera wiedergeben sollen (vgl. (3.8)). Dazu werden in zwei benachbarten Bildern jeweils korrespondierende Merkmale detektiert und insgesamt eine Rotation von 360 Grad durchgeführt. Die zu minimierende Kostenfunktion hat die Form:

$$E = \sum_{i=1}^{N} \sum_{j=1}^{M} \left(\left(x_{p_{ij}}' - x_{p_{ij}} \right)^2 + \left(y_{p_{ij}}' - y_{p_{ij}} \right)^2 \right) \ , \tag{3.71}$$

wobei M die Anzahl der Rotationen und N die Anzahl der detektierten Bildmerkmale ist. Weiterhin bezeichnet $(x_{p_{ij}}, y_{p_{ij}})$ den Bildpunkt vor und $(x_{p_{ij}}', y_{p_{ij}}')$ den zugehörigen Bildpunkt nach der Rotation. Stein [Ste95] verwendet zur Berechnung der Parameter Rotationen mit bekanntem Rotationswinkel.

Das Verfahren, das von Stein [Ste95] vorgestellt wird, ist umständlich durchzuführen, da die zu kalibrierende Kamera jeweils manuell auf einem Objekt so zu installieren ist, daß das optische Zentrum der Kamera auf der Rotationsachse des Objekts liegt. In diesem Verfahren werden Bildpunkte detektiert und einander zugeordnet. Fluchtpunkte werden nicht betrachtet. Es werden zwei Rotationen zur Berechnung der internen Kameraparameter benötigt. Ein Vorteil des Verfahrens liegt in der Möglichkeit, eine Rotation um 360 Grad durchzuführen, was das Verfahren stabilisiert. Im allgemeinen kann eine solche Rotation nicht durchgeführt werden, so daß die Kalibrierung stark von dem zu rotierenden Objekt abhängt. Deswegen ist diese Methode für die Kalibrierung in einem aktiven System nicht geeignet.

3.4.5.5 Das Verfahren von Beardsley et al.

Beardsley et al. [BMZ92] beschreiben ein Verfahren, das eine feststehende Kamera aufgrund eines rotierenden Objekts kalibriert. Dieses Verfahren unterscheidet sich von den bisher beschriebenen insbesondere dadurch, daß man keine Kamera, sondern ein Objekt rotieren läßt. In diesem Verfahren wird die Fluchtpunktgeometrie ausgenutzt (siehe Kap. 2.4). Dadurch, daß die Bildebene die Trajektorie der Punkte nicht schneidet, beschreiben detektierte Fluchtpunkte eine Ellipse (Abb. 3.5).

Um eine solche Ellipse möglichst genau zu beschreiben, wird eine Rotation um 360 Grad vorgenommen. Eine Ellipse läßt sich jeweils aus den detektierten Fluchtpunktkoordinaten durch Anpassung an die Ellipsengleichung

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{3.72}$$

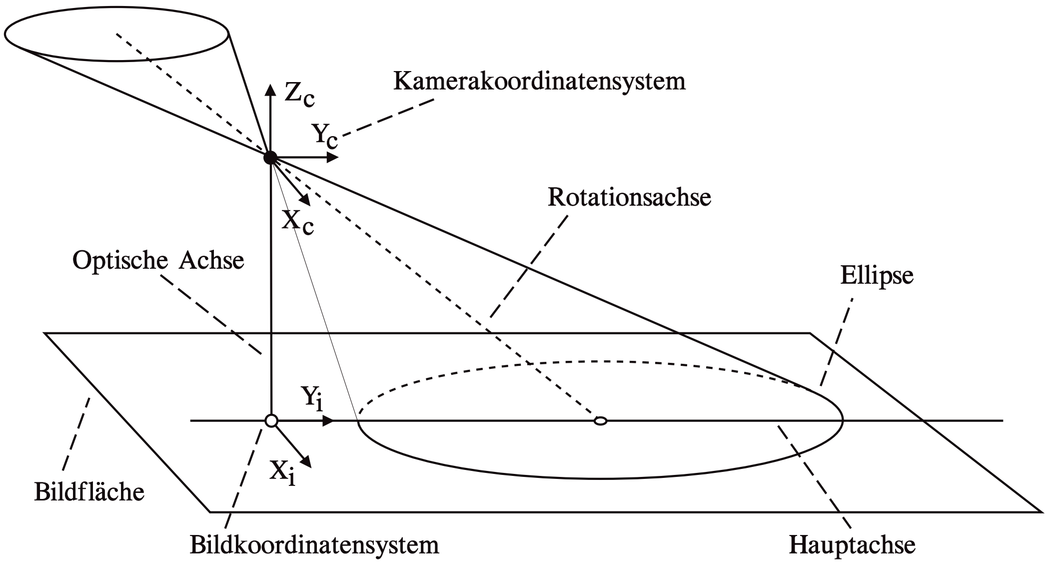


Abbildung 3.5: Durch den Schnitt eines Rotationskegels mit einer zur xy-Ebene des Kamerakoordinatensystems parallelen Bildebene ensteht ein Kegelschnitt. Der Schnitt ist eine Ellipse, wenn die Bildebene nur eine Hälfte des Kegels schneidet.

ermitteln. Die Bestimmung der internen Parameter wird anschließend in drei Schritten durchgeführt:

1. Bestimmung des Betrachtungsverhältnisses (aspect ratio $= k = \frac{k_x}{k_y}$),

2. Bestimmung des Bildhauptpunktes $(x_0, y_0)^T$,

3. Bestimmung der Brennweite f.

Die Hauptachsen[3] mehrerer Ellipsen, die durch Rotation um verschiedene Rotationsachsen auf die Sensorfläche der Kamera abgebildet werden, schneiden sich in einem Punkt [BMZ92].

Zur Berechnung des Betrachtungsverhältnisses werden daher mindestens drei Ellipsen benötigt. Ist das Betrachtungsverhältnis falsch, so sind die Ellipsen verzerrt, und die Hauptachsen der Ellipsen schneiden sich nicht in einem Punkt. Korrigiert man das Betrachtungsverhältnis so, daß sich die Hauptachsen aller drei Ellipsen in einem Punkt schneiden, so erhält man das korrekte Betrachtungsverhältnis.

Ist das Betrachtungsverhältnis bekannt, so ergibt sich der Bildhauptpunkt $(x_0, y_0)^T$ als Schnittpunkt der Hauptachsen mindestens zweier Ellipsen.

[3]Die Hauptachse einer Ellipse ist eine Gerade durch zwei Punkte der Ellipse, so daß der Abstand der Punkte maximal ist.

Die Brennweite f ergibt sich schließlich mit bekanntem Bildhauptpunkt aus einer Ellipse:

$$\frac{d^2}{a^2 - b^2} - \frac{f^2}{b^2} = 1 \ . \tag{3.73}$$

Dabei ist d die Entfernung des Bildhauptpunktes vom Zentrum der Ellipse, und a und b werden direkt aus (3.72) übernommen.

Dieses Verfahren ist in der Lage, alle internen Parameter einer Kamera zu berechnen. Des weiteren ist es dazu nur notwendig, Fluchtpunkte einander zuzuordnen. Die Rotationswinkel werden nicht als bekannt vorausgesetzt. Beardsley et al. [BMZ92] führen eine Rotation um 360 Grad durch. Zur Bestimmung der internen Kameraparameter sind mindestens drei Rotationen um verschiedene Rotationsachsen nötig. Die Durchführung der Methode ist nicht aufwendig, allerdings wird ein Kalibrierkörper benötigt. Zur Kalibrierung in einem aktiven System ist dieses Verfahren daher nicht geeignet.

3.5 Klassifizierung der Kalibrierverfahren

In diesem Abschnitt wird eine Klassifizierung der untersuchten Kalibrierverfahren gegeben (Abb. 3.6). Die Klassifizierung richtet sich nach Art und Umfang des vorausgesetzten Wissens. Die Unterteilung der Kalibriermethoden erfolgt in sechs Gruppen. Je höher die Gruppennummer ist, um so geringer ist das benötigte Vorwissen. Dies führt dazu, daß ein Verfahren einer höheren Gruppe einer natürlichen Umwelt angepaßter ist, als ein Verfahren einer niedrigeren Gruppe.

- **Gruppe 1 - klassische Kalibriermethoden (inaktiv)**
  Für diese Methoden sind sowohl die Szenenkoordinaten wie auch die zugehörigen Pixelkoordinaten wichtig. Für diese Methode benötigt man des weiteren einen Kalibrierkörper. In diese Kategorie fallen das unter Kapitel 3.2 beschriebene Verfahren, sowie die dort zitierten Ansätze. Im allgemeinen werden alle internen Parameter mit diesen Verfahren berechnet.

- **Gruppe 2 - Rotationsmethoden (aktiv)**
  Diese Methoden setzen neben der Zuordnung von Pixelkoordinaten auch bekannte Rotationen voraus. In einigen Verfahren werden zusätzlich noch Kalibrierkörper benötigt. Manche Methoden lassen es zu, Pixelkoordinaten durch Fluchtpunktkoordinaten zu ersetzen. Das in Kapitel 3.4.3 beschriebene Verfahren fällt in diese Gruppe, so wie auch die Verfahren von Du und Brady [DB93], Viéville [Vié94] und Basu [Bas95]. Mit diesen Methoden werden in der Regel alle internen Parameter berechnet. Teilweise werden interne Parameter als bekannt vorausgesetzt.

- **Gruppe 3 - Zoommethoden (aktiv)**
  Die Zoommethoden setzen nur eine Zuordnung von Pixelkoordinaten voraus. Fluchtpunkte können nicht verwendet werden. In diese Gruppe fallen

das unter Kapitel 3.4.1 beschriebene und das von Li [Li94] verwendete Verfahren. Mit diesen Verfahren ist es nur möglich, den Bildhauptpunkt zu bestimmen.

- **Gruppe 4 - Fluchtpunktmethoden (inaktiv)**
  Die inaktiven Fluchtpunktmethoden setzen die Kenntnis über verschiedene Fluchtpunkte voraus. Weiterhin können nicht alle internen Parameter der Kamera berechnet werden, was dazu führt, daß weitere interne Parameter schon als bekannt vorausgesetzt werden müssen. In diese Gruppe fallen das unter Kapitel 2.4 beschriebene Verfahren und die von Li [Li94], Caprille und Torre [CT90], Echigo [Ech90] und Kollnig [Kol95] beschriebenen Verfahren.

- **Gruppe 5 - Ellipsenmethode (aktiv)**
  Die Ellipsenmethode setzt Kenntnisse über zugeordnete Fluchtpunkte oder Fluchtpunktlinien (Horizonte) voraus. Es können auch Bildpunkte benutzt werden. Des weiteren muß ein Kalibrierkörper benutzt werden. In diese Gruppe fällt das von Beardsley et al. [BMZ92] beschriebene Verfahren. Es können alle internen Parameter berechnet werden.

- **Gruppe 6 - Hyperbelmethode (aktiv)**
  Die Hyperbelmethode setzt lediglich Kenntnisse über zugeordnete Fluchtpunkte voraus. Es können auch Bildpunkte benutzt werden. In diese Gruppe fällt das in Kapitel 3.4.4 beschriebene Verfahren. Diese Methode ist das Hauptresultat dieser Arbeit. Es werden sämtliche internen Parameter berechnet.

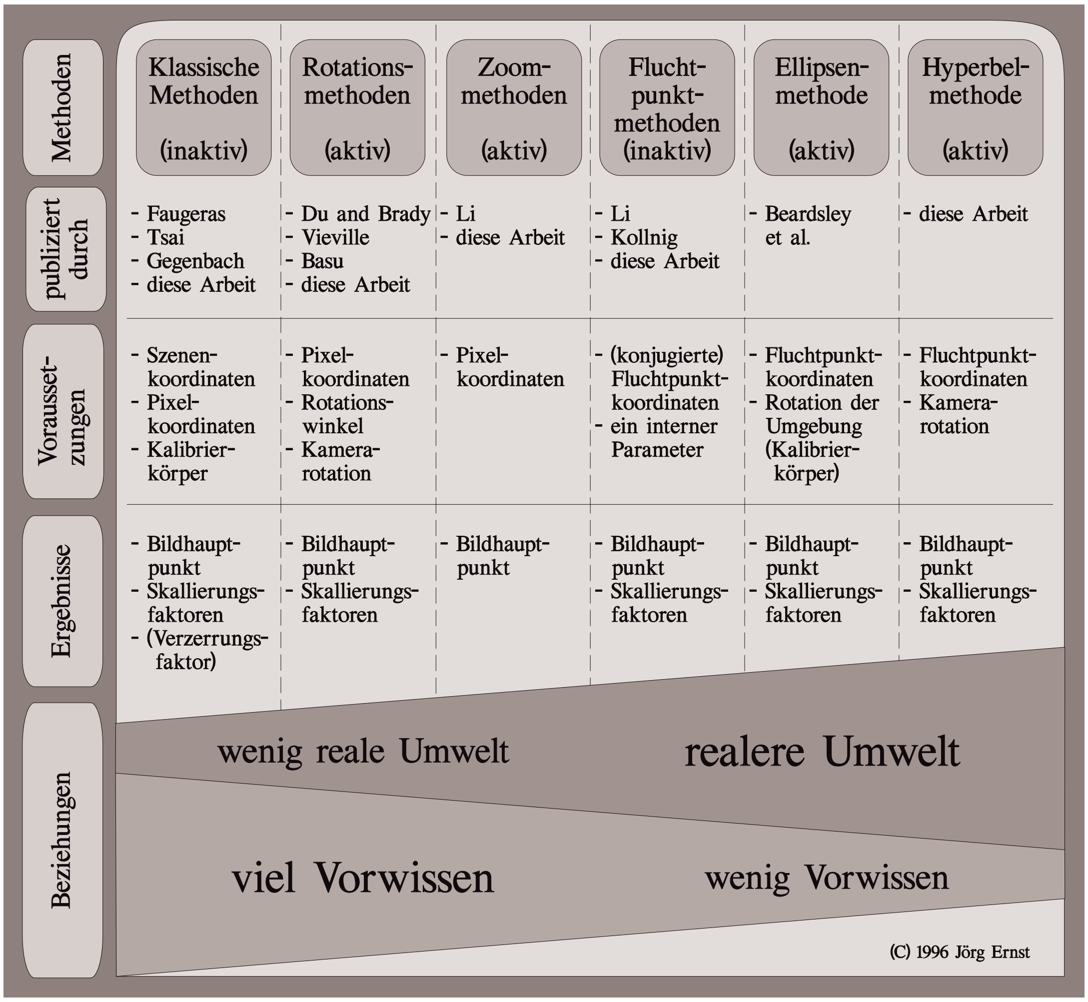


Abbildung 3.6: Klassifizierung der Kalibrierverfahren nach der Umfang des Vorwissens und somit die Einordnung in eine realere oder weniger reale Umwelt.

Kapitel 4

Detektion der Referenzmerkmale zur Kalibrierung der internen Parameter

In diesem Kapitel werden Möglichkeiten der Gewinnung von Referenzmerkmalen aus Kamerabildern und deren Zuordnung zu Szenenmerkmalen beschrieben.

Zunächst wird auf Kalibrierkörper eingegangen, die für klassische Kalibrierverfahren von Bedeutung sind. Es wird ein Kalibrierwürfel erläutert und seine Szenenmerkmale durch eine Abbildung festgelegt.

Anschließend werden Bildverarbeitungsmethoden aufgezeigt, die der Findung von Bildmerkmalen dienen. Weiterhin wird die Detektion von Fluchtpunkten und die Detektion von Bildpunkten eines Kalibrierwürfels besprochen. Sämtlichen Verfahren liegt die Hough-Transformation [Hou62] zugrunde. Fluchtpunkte werden aufgrund der Detektion von Geradenbüscheln, deren Urbilder parallel sind, berechnet. Die Detektion der Bildmerkmale geschieht ebenfalls durch die Detektion von Geraden. Dieses Verfahren wurde neu entwickelt und hat insbesondere die folgenden Eigenschaften:

- Die Bildmerkmale werden als Schnittpunkt mehrerer Geraden definiert. Daher haben sie reellen Charakter. Das bedeutet, daß die Bildmerkmale mit Subpixelgenauigkeit detektiert werden Können.

- Durch eine Zuordnungsvorschrift ist es möglich, den Kalibriervorgang vollständig automatisch ablaufen zu lassen.

Weiterhin wird ein Verfahren vorgestellt, das einer vorgegebenen Punktfolge eine Hyperbel anpaßt. Diese Punktfolge kann aus Bild- oder Fluchtpunkten bestehen.

4.1 Kalibrierkörper

Für die Kalibrierung mit klassischen Verfahren (siehe Kap. 3.2) ist die Verwendung eines Kalibrierkörpers unerläßlich, da bei diesem Verfahren genaue Kenntnisse über Szenenkoordinaten vorliegen müssen.

Im allgemeinen werden zwei verschiedene Kalibrierkörper bevorzugt. Zum einen kann man sich eine Kalibrierplatte erstellen, über der eine Kamera so angebracht wird, daß die optische Achse eine Normale der Oberfläche darstellt. Die Kalibrierplatte muß über ausgezeichnete Merkmale verfügen, die durch Bildverarbeitungsverfahren detektiert werden können. Da es für die klassischen Kalibrierverfahren notwendig ist, daß die ausgezeichneten Punkte auf mindestens zwei unterschiedlichen Ebenen im Raum liegen (siehe Kap. 3.2.4), sind die ausgezeichneten Merkmale oftmals auf mehreren Ebenen angeordnet. Ein solcher Kalibrierkörper wird unter anderem durch Tsai [Tsa86] und Beß [Beß90] verwendet. Die Kamera kann dabei im allgemeinen durch Auf- bzw. Abwärtsbewegungen auf verschiedene Höhen eingestellt werden. Die Entfernung der Kamera von der Kalibrierplatte muß sehr genau bestimmbar sein.

Zum anderen besteht die Möglichkeit, einen Kalibrierwürfel zu benutzen. Auf den Seiten des Würfels müssen Merkmale ausgezeichnet sein, so daß ein Bildverarbeitungssystem diese möglichst zuverlässig detektieren kann. Wiederum ist es für klassische Verfahren notwendig, mindestens zwei Ebenen zu kennen (siehe Kap. 3.2.4).

4.1.1 Kalibrierwürfel

Für die relevanten Verfahren dieser Arbeit wurde ein Kalibrierwürfel benutzt (siehe Abb. 4.1). Die Verwendung eines Würfels hat gegenüber einer Kalibrierplatte den Vorteil, daß er leicht transportabel ist und somit an unterschiedlichen Orten eingesetzt werden kann.

Auf dem verwendeten Würfel sind auf drei benachbarten Seiten jeweils 81 Quadrate abgebildet. Die Ecke des Kalibrierwürfels, die sich zwischen diesen drei Seiten befindet, bildet den Ursprung eines rechtshändigen Szenenkoordinatensystems (siehe Kap. 2.1). Die z-Achse des Koordinatensystems verläuft in vertikaler Richtung vom Ursprung in Richtung gegenüberliegender Ecke des Kalibrierwürfels und bildet eine Kante des Würfels. Die beiden anderen Basisvektoren zeigen ebenfalls vom Ursprung in Richtung der gegenüberliegenden Würfelecke. Die Quadrate auf einer Seite sind jeweils so angeordnet, daß das erste Quadrat, ausgehend vom Ursprung, in einer Entfernung von $10mm$ von zwei Achsen des Szenenkoordinatensystems liegt. Die Quadrate selbst haben die Größe von $40mm \times 40mm$ und sind voneinander $10mm$ entfernt.

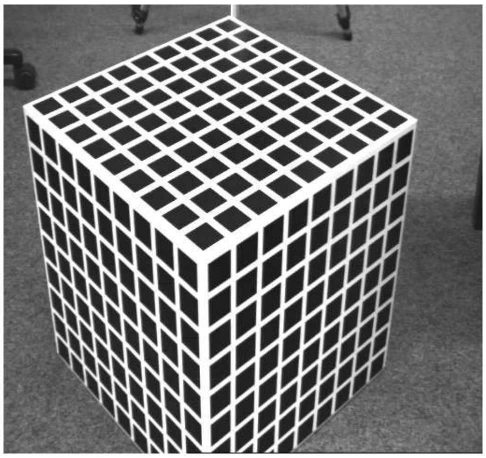


Abbildung 4.1: Der Kalibrierwürfel

4.1.2 Zuordnung der Merkmale eines Kalibrierwürfels

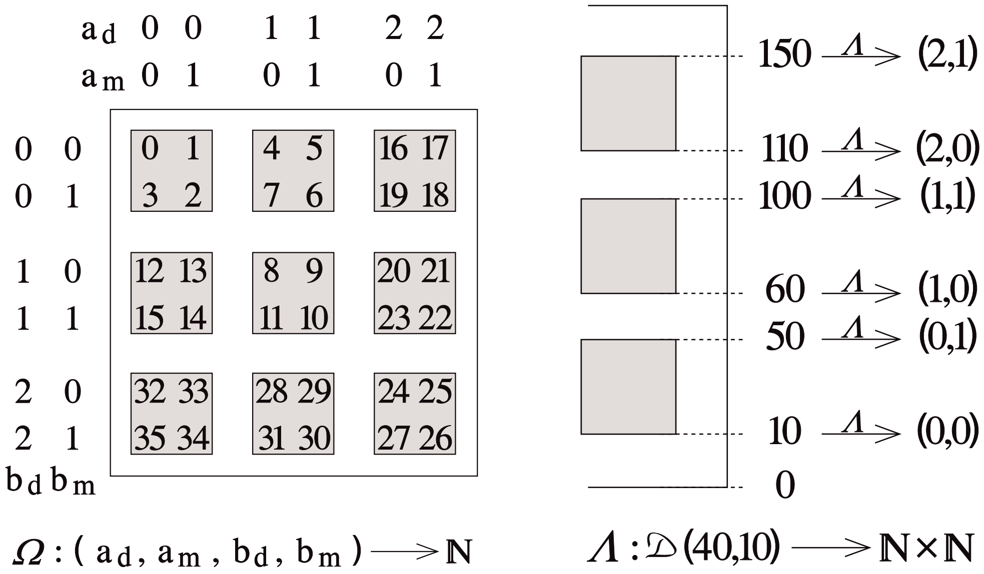
Um eine eindeutige Zuordnung von Szenen- zu den Bildkoordinaten gemäß (3.6) zu ermöglichen, wird in diesem Abschnitt eine Abbildung angegeben, nach der Ecken von Quadraten, die auf einem Kalibrierwürfel abgebildet sind, einem Index zugeordnet werden. Dies entspricht einer Numerierung der Ecken von Quadraten auf einem Kalibrierwürfel, wie er in Kapitel 4.1.1 beschrieben wurde. Die Art dieser Zuordung wurde im Rahmen dieser Arbeit entwickelt. Betrachtet man eine Ebene des Kalibrierwürfels, so hat diese Zuordung den Vorteil, daß die Numerierung unabhägig von der Zahl der Quadrate, die betrachtet werden, dieselbe ist.

Dazu wird zunächst eine Funktion Ω definiert, die ein Quadrupel auf eine natürliche Zahl abbildet:

$$\begin{aligned}\Omega : \mathbb{N} \times \mathbb{B} \times \mathbb{N} \times \mathbb{B} \rightarrow \mathbb{N}\ ,\ (\mathrm{a_d}, \mathrm{a_m}, \mathrm{b_d}, \mathrm{b_m}) \mapsto \\ 4\left((\max\{a_d; b_d\})^2 + \max\{a_d; b_d\} + b_d - a_d\right) + a_m + b_m\,(3 - 2a_m)\ .\end{aligned} \tag{4.1}$$

Um diese Funktion zu erläutern, ist Abbildung 4.2(a) als Beispiel angegeben. Es ist ein Ausschnitt einer Seite des unter Kapitel 4.1.1 beschriebenen Kalibrierwürfels abgebildet. Die Parameter a_d und b_d der Funktion Ω kann man als Numerierung der Quadrate in a-, bzw. b-Richtung betrachten. Die Parameter a_m und b_m beschreiben die Numerierung der Ecken eines Quadrats.

Im folgenden wird davon ausgegangen, daß auf drei Seiten eines Kalibrierwürfels Quadrate angeordnet sind. Das erste Quadrat hat die gleiche Entfernung zum Kalibrierwürfelrand, wie zwei benachbarte Quadrate zueinander, und alle Quadrate auf dem Kalibrierwürfel besitzen dieselbe Größe. Um die Koordinaten festzulegen, wird eine Menge $\mathcal{D}(q_k, q_d)$ von natürlichen Zahlen für q_k, $q_d \in \mathbb{N}$



(a) Beispiel zur Funktion Ω

(b) Beispiel zur Funktion Λ

Abbildung 4.2: Funktionen zur Szenenkoordinatenzuordung

durch

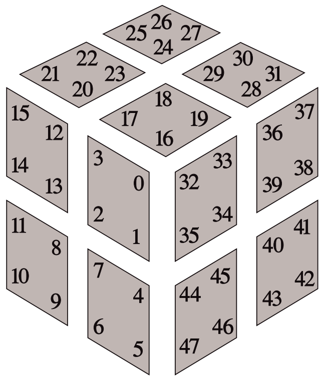
$$\mathcal{D}\left(q_k, q_d\right) = \left\{x \in \mathbb{N}_+ | \mathrm{x} \bmod \left(\mathrm{q_k} + \mathrm{q_d}\right) = 0 \vee \left(\mathrm{x} - \mathrm{q_d}\right) \bmod \left(\mathrm{q_k} + \mathrm{q_d}\right) = 0\right\} \quad (4.2)$$

definiert. Die Kantenlänge eines Quadrats auf einem Kalibrierwürfel werde durch q_k bezeichnet und in mm angegeben. Die Entfernung zweier Quadrate voneinander werde durch q_d bezeichnet und ebenfalls in Millimeter angegeben. Mit dieser Definition sind sämtliche Szenenkoordinaten der Eckpunkte der abgebildeten Quadrate des in Kapitel 4.1.1 beschriebenen Kalibrierwürfels Elemente aus $\mathcal{D}^3(40, 10)$ (siehe Abb. 4.1).

Weiterhin wird eine Funktion Λ definiert, die eine Zahl aus der Menge $\mathcal{D}(q_k, q_d)$ auf ein Tupel abbildet:

$$\begin{aligned} \Lambda : \mathcal{D}\left(q_k, q_d\right) &\rightarrow \mathbb{N} \times \mathbb{B} \ , \qquad (4.3) \\ n &\mapsto \left(\left(n - q_d\right) \ \operatorname{div} \ \left(q_k + q_d\right), \left(n \bmod \left(q_k + q_d\right)\right) \ \operatorname{div} \ q_d\right) \ . \end{aligned}$$

Um die Arbeitsweise dieser Funktion zu demonstrieren, ist Abbildung 4.2(b) als Beispiel angegeben. Wiederum ist ein Ausschnitt einer Seite des unter Kapitel 4.1.1 beschriebenen Kalibrierwürfels abgebildet. Es werden die Koordinaten in einer Richtung aus der Menge $\mathcal{D}(q_k, q_d)$ mit den Parametern $q_k = 40mm$ und $q_d = 10mm$ betrachtet. Den ersten Ausgabeparameter der Funktion Λ kann man als Numerierung der Quadrate in dieser Richtung betrachten, den zweiten Parameter als Beschreibung einer durch den ersten Parameter festgelegten Ecke eines Quadrates in einer Richtung.



x	y	z	Π_s	x	y	z	Π_s	x	y	z	Π_s
10	0	10	0	0	10	10	16	10	10	0	32
10	0	50	1	0	50	10	17	50	10	0	33
50	0	50	2	0	50	50	18	50	50	0	34
50	0	10	3	0	10	50	19	10	50	0	35
10	0	60	4	0	60	10	20	60	10	0	36
10	0	100	5	0	100	10	21	100	10	0	37
50	0	100	6	0	100	50	22	100	50	0	38
50	0	60	7	0	60	50	23	60	50	0	39
60	0	60	8	0	60	60	24	60	60	0	40
60	0	100	9	0	100	60	25	100	60	0	41
100	0	100	10	0	100	100	26	100	100	0	42
100	0	60	11	0	60	100	27	60	100	0	43
60	0	10	12	0	10	60	28	10	60	0	44
60	0	50	13	0	50	60	29	50	60	0	45
100	0	50	14	0	50	100	30	50	100	0	46
100	0	10	15	0	10	100	31	10	100	0	47

Abbildung 4.3, Tabelle 4.1 : Beispielzuordnung der Szenenmerkmale

Mit diesen Definitionen ist eine Zuordnung der Eckpunkte der Quadrate auf einem in Kapitel 4.1.1 beschriebenen Kalibrierwürfel definierbar:

$$\Pi_s : \mathcal{D}^3(q_k, q_d) \times \mathbb{N} \to \mathbb{N}, \tag{4.4}$$

$$(x, y, z, m) \mapsto \begin{cases} 8m^2 + \Omega(\Lambda(y), \Lambda(z)) \text{ , falls } x = 0 \\ \Omega(\Lambda(z), \Lambda(x)) \text{ , falls } y = 0 \\ 4m^2 + \Omega(\Lambda(x), \Lambda(y)) \text{ , falls } z = 0 \end{cases} ,$$

wobei

$$x, y, z \leq (q_k + q_d)m \text{ und } (x = 0)\dot{\vee}(y = 0)\dot{\vee}(z = 0)^1 \text{ .}$$

Der Parameter $m \in \mathbb{N}$ bezeichnet die Anzahl der Quadrate auf dem Würfel, die auf drei unterschiedlichen Ebenen des Würfels betrachtet werden.

Um die Zuordnung gemäß (4.4) zu erläutern, sind Abbildung 4.3 und Tabelle 4.1 abgebildet. Hier wird der Fall $m = 2$, $q_k = 40mm$ und $q_d = 10mm$ demonstriert, welcher der Betrachtung des verwendeten Kalibrierwürfels (Abb. 4.1) entspricht, falls man nur 2×2 Quadrate auf jeder Seite berücksichtigt.

4.2 Detektion von Bildmerkmalen

In diesem Kapitel wird die Detektion von verschiedenen Arten von Bildmerkmalen dargestellt. Es wird auf die Schätzung von Fluchtpunktkoordinaten von in der Welt parallelen Geradenbüscheln (siehe Kap. 2.4), auf die Schätzung von Hyperbeln aufgrund einer Folge von Fluchtpunkten und auf die Detektion von Bildpunkten auf einem Kalibrierwürfel eingegangen.

[1] Durch $\dot{\vee}$ werde das exklusive *ODER* bezeichnet.

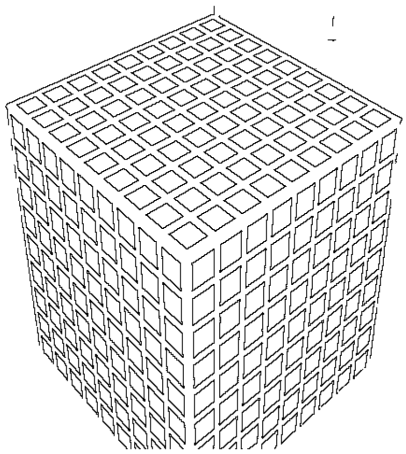


Abbildung 4.4: Kalibrierwürfel nach Kantendetektion

Diese Verfahren setzen im Kamerabild detektierte Geraden voraus. Zur Geradenermittlung wurde in dieser Arbeit die Hough-Transformation [Hou62] verwendet.

Zum Schluß wird auf die Zuordnung der Bildmerkmale eines Kalibrierwürfels zu den korrespondierenden Szenenmerkmalen eingegangen und die Detektion von konjugierten Fluchtpunkten eines Kalibrierwürfels besprochen.

Da der Geradendetektion dieser Arbeit ein Binärbild zugrunde liegt, sei hier als Beispiel das Binärbild eines Kalibrierwürfels angegeben (Abb. 4.4). Dieses Bild ist aufgrund der Anwendung eines Kantendetektors (siehe [Jäh93]) aus dem in Abbildung 4.1 dargestellten Grauwertbild entstanden.

4.2.1 Die Hough-Transformation

Die klassische Hough-Transformation [Hou62] wird angewendet, falls man in einem digitalisierten Bild Objekte sucht, über deren Auftreten wenig bekannt ist. Das Objekt selbst muß allerdings gut als parametrisierte Kurve beschrieben werden können (z.B. eine Gerade oder ein Kreis). Der Bildraum hat als endliche Menge von Bildpunkten vorzuliegen. Das Wesentliche der Hough-Transformation ist die gleichzeitige Gruppierung und Schätzung von Parametern und die Unempfindlichkeit gegenüber Fehlern im Verlauf der Kurven und gegenüber Rauschen in den Bilddaten [BB82].

Der Ausgangspunkt für die Definition der Hough-Transformation ist ein parametrischer Zusammenhang der Form

$$\begin{aligned} & f(X,\Omega) = 0, \qquad (4.5) \\ \text{mit} \quad & X = (x_1, \cdots, x_M)^T \text{ und } \Omega = (\omega_1, \cdots, \omega_N)^T \ , \end{aligned}$$

wobei X ein Punkt im Bildraum und Ω ein Parametervektor im Parameterraum ist, der die spezielle Form eines Objektes festlegt, das es zu erkennen gilt. Diese Definition geht auf Princen et al. [PIK92] zurück.

In dieser Arbeit werden Bildpunkte jeweils detektierte Kantenpunkte eines Grauwertbildes sein, die selbst einen binären Bildraum ergeben. Jeder Punkt Ω_0 im Parameterraum repräsentiert ein parametrisiertes Objekt im Bildraum. Die Menge aller Bildpunkte, die die Parameter Ω_0 erfüllen, wird durch

$$B(\Omega_0) = \{X | f(X, \Omega_0) = 0\} \tag{4.6}$$

dargestellt. Die Menge aller möglichen Objekte des Parameterraumes, auf denen ein vorgegebener Punkt X_0 des Bildraumes liegen kann, wird durch

$$P(X_0) = \{\Omega | f(X_0, \Omega) = 0\} \tag{4.7}$$

angegeben. Die erzeugten Mengen $P(X_0)$ können geometrisch so interpretiert werden, daß jede Menge in dem Parameterraum eine Kurve darstellt. Schneiden sich sehr viele Kurven in dem Punkt Ω_0, so ist die Wahrscheinlichkeit hoch, daß ein Objekt mit den Parametern Ω_0 in dem Bildraum zu finden ist.

Um in einem Bildraum diejenigen Objekte zu ermitteln, die darin am wahrscheinlichsten enthalten sind, definiert man die für jeden Punkt Ω_0 des Parameterraumes die Hough-Transformierte $H(\Omega_0)$ als die Mächtigkeit der Menge aller Bildpunkte, die die Parameter Ω_0 erfüllen :

$$H(\Omega_0) = |B(\Omega_0)| \; . \tag{4.8}$$

Die Hough-Transformierte $H(\Omega)$ gibt ein Maß dafür an, wie wahrscheinlich es ist, daß das Objekt Ω_0 sich in dem Bildraum befindet. Je größer der Wert einer Hough-Transformierten ist, um so wahrscheinlicher ist es, daß man ein Objekt mit diesen Parametern in dem Bildraum findet. Aus den berechneten Hough-Transformierten schließt man durch Schwellwertbildung, Bestimmung von lokalen Maxima oder gewichteten Mittelwertbildungen auf Objekte des Bildraumes.

Die Implementierung der Hough-Transformation für den praktischen Einsatz macht Schwierigkeiten, da der Parameterraum im allgemeinen nicht endlich, also beispielsweise der $\mathbb{R}^n$ mit $n \in \mathbb{N}$ ist. Daher ist es nötig, den Parameterraum zu diskretisieren. Dazu wird ein Feld von Zellen definiert, das einem Teilraum des Parameterraumes entspricht. Für jede dieser Zellen wird ein Zähler definiert. Dieses Feld von Zählern wird als Akkumulatorfeld bezeichnet.

Zu Beginn eines Algorithmus werden alle Zähler mit dem Wert null initialisiert. Anschließend werden für alle Punkte X_0 des Bildraumes diejenigen Zellen des Akkumulatorfeldes um den Wert eins erhöht, die durch die Elemente der Menge $P(X_0)$ berührt werden. Hat man alle Mengen $P(X_0)$ auf das Akkumulatorfeld angewendet, so stellt das Akkumulatorfeld eine diskrete Variante aller Hough-Transformierten dar, und es korrespondieren lokale Maxima zu erkannten Objekten.

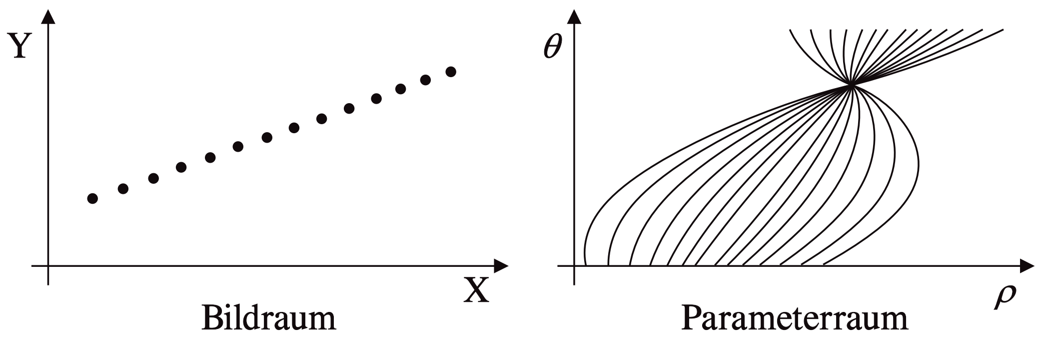


Abbildung 4.5: Bildraum und Parameterraum der Hough-Transformation zur Geradendetektion nach (4.9). In einem zweidimensionalen Bildraum befindet sich eine durch eine Punktreihe repräsentierte Gerade. Der Parameterraum ist ebenfalls zweidimensional und wird durch die Parameter ρ und θ repräsentiert. In diesem Fall gilt: $\rho \in \mathbb{R}$ und $0 \leq \theta < \pi$. Für jeden Punkt der Geraden aus dem Bildraum erhält man eine Kurve im Parameterraum. Sämtliche Kurven im Parameterraum schneiden sich in genau einem Punkt. Dies sind die Parameter für ρ und θ zur Darstellung der Geraden im Bildraum nach (4.9).

Bei dieser Beschreibung ist zu beachten, daß dem Programmierer viele Freiheitsgrade offengelassen wurden. Diese Freiheitsgrade sind hauptsächlich die Wahl der Parametrisierung und die Größe, Gestalt und Anordnung der Zellen des Akkumulatorfeldes.

Die Auswirkungen der Wahl dieser Freiheitsgrade nehmen einen großen Einfluß auf die Güte der Ergebnisse. Diese Wahl hat von Fall zu Fall sehr unterschiedlich zu geschehen. Somit ist vor Implementierung der Hough-Transformation große Aufmerksamkeit auf die Festlegung dieser Freiheitsgrade zu richten.

4.2.2 Anwendung der Hough-Transformation zur Geradendetektion

Aufgrund der oben angesprochenen Freiheitsgrade soll hier zunächst die Wahl der Geradendarstellung begründet werden. Die sicherlich bekannteste Geradendarstellung ist

$$y + mx = c\ .$$

Diese Repräsentation hat den Nachteil, daß zur y-Achse parallele Geraden nicht dargestellt werden können. Aus diesem Grund wird die Geradendarstellung

$$\rho - x\cos\theta + y\sin\theta = 0 \tag{4.9}$$

verwendet. Man erhält den Bildraum $X = (x, y)^T$ und den Parameterraum $\Omega = (\theta, \rho)^T$ (Abb. 4.5).

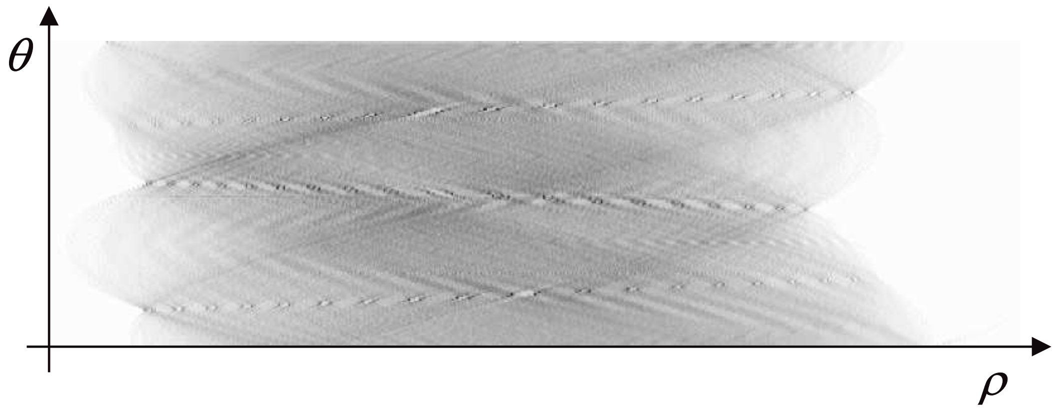


Abbildung 4.6: Akkumulatorfeld der Hough-Transformation, welches sich aus Abbildung 4.4 ergibt, wenn man Geraden nach (4.9) bestimmen möchte. Je dunkler eine Zelle ist, um so größer ist ihr Akkumulatorwert. Geraden des Bildraumes der Abbildung 4.4 werden durch diejenigen Zellen repräsentiert, die in einer lokalen Umgebung den höchsten Akkumulatorwert besitzen.

Um die Eindeutigkeit der Geraden festzulegen, gelte im folgenden $0 \leq \theta < \pi$. Der Parameter ρ kann auf $-\rho_{max} \leq \rho \leq \rho_{max}$ beschränkt werden. Dabei gelte $\rho_{max} = ({x_{p_{\max}}}^2 + {y_{p_{\max}}}^2)^{\frac{1}{2}}/2$ für $x_{p_{\max}}$ als Bildbreite und $y_{p_{\max}}$ als Bildhöhe des Bildraumes. Dazu wird der Ursprung des Bildraumes auf $(x_{p_{\max}}/2, y_{p_{\max}}/2)$ festgelegt.

Da die Zellgröße, deren Anordnung und Gestalt innerhalb des Akkumulatorfeldes variabel gehandhabt wurde, wird auf die Wahl der Freiheitsgrade nicht näher eingegangen. Es sei hier auf die Autoren Eckhardt und Maderlechner [EM94], Davies [Dav94], Princen et al. [PIK92], Leavers [Lea93] und Illingworth und Kittler [IK87] verwiesen, in deren Veröffentlichungen auf die Wahl dieser Freiheitsgrade ausführlich eingegangen wird.

In Abbildung 4.6 ist ein aufgrund der Hough-Transformation entstandenes Akkumulatorfeld abgebildet. Als Bildraum liegt diesem Akkumulatorfeld das Binärbild eines Kalibrierwürfels (Abb. 4.4) zugrunde. Auf die Detektion der Geraden anhand des Akkumulatorfeldes wird in Kapitel 4.2.5 eingegangen.

4.2.3 Verfahren zur Fluchtpunktschätzung

Sei im folgenden $n \in \mathbb{N}$ die Anzahl der Geraden, die zur Fluchtpunktbestimmung gegeben sind, und ein Geradenbüschel durch

$$\overline{\mathcal{L}} = \{l_i = (\theta_i, \rho_i) | 1 \leq i < n\} \tag{4.10}$$

nach Gleichung (4.9) gegeben, wobei die Urbilder aller Geraden parallel verlaufen. Ein solches Büschel von Geraden erhält man u.a. aus dem Akkumulatorfeld der Hough-Transformation unter der Annahme, daß die Winkel θ aller Geraden aus der Menge $\overline{\mathcal{L}}$ in einem Toleranzbereich liegen[2]. Die Wahl des Toleranzbereichs hängt von der perspektivischen Verzerrung der durch die Zentralprojektion realisierten Abbildung ab.

Die perspektivische Transformation zwischen dem Weltkoordinatensystem und dem Rechnerkoordinatensystem wurde durch die projektive Geometrie beschrieben (siehe Kap. 2.2). Diese hat den Vorteil, daß zwischen im Endlichen und im Unendlichen liegenden Fluchtpunkten nicht unterschieden werden muß (siehe Kap. 2.4).

Sei im folgenden $\boldsymbol{P}_\infty = (x, y, z)^T$ ein Fluchtpunktvektor in homogenen Koordinaten. Nach Kapitel 2.4 ist der homogene Koordinatenvektor des Fluchtpunktes orthogonal zu den homogenen Koordinatenvektoren der Bildgeraden der Schar. Dies führt unter Verwendung von Gleichung (4.9) zum folgenden Minimierungsproblem:

$$\min \sum_{i=1}^{n} (\rho_i z - x\cos\theta_i + y\sin\theta_i)^2 \tag{4.11}$$

unter der Nebenbedingung $x^2 + y^2 + z^2 = 1$, wobei $(\theta_i, \rho_i) \in \overline{\mathcal{L}}$. Dieses Problem ist äquivalent zu

$$\min_{\|\boldsymbol{E}\|_2 = 1} \|A\boldsymbol{E}\|_2 \ , \tag{4.12}$$

wobei

$$A = \begin{pmatrix} & \vdots & \\ -\cos\theta_i & \sin\theta_i & \rho_i \\ & \vdots & \end{pmatrix} \text{ und } \boldsymbol{E} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{4.13}$$

sind, was wiederum äquivalent zu der Minimierungsaufgabe

$$\min \|A\boldsymbol{E}\|_2{}^2 + \lambda(1 - \|\boldsymbol{E}\|_2{}^2) \tag{4.14}$$

mit dem Lagrange-Multiplikator λ ist. Um diese Problemstellung zu lösen, ist ein Einheitsvektor $\boldsymbol{E}$ gesucht, der

$$(A^T A - \lambda I)\boldsymbol{E} = 0$$

erfüllt, wobei I die Einheitsmatrix ist. Folglich ist der gesuchte Einheitsvektor ein Eigenvektor von

$$A^T A = \begin{pmatrix} \sum\cos^2\theta_i & -\sum\cos\theta_i\sin\theta_i & -\sum\cos\theta_i\rho_i \\ -\sum\cos\theta_i\sin\theta_i & \sum\sin^2\theta_i & \sum\sin\theta_i\rho_i \\ -\sum\cos\theta_i\rho_i & \sum\sin\theta_i\rho_i & \sum{\rho_i}^2 \end{pmatrix} . \tag{4.15}$$

[2]Diese Annahme ist sehr restriktiv. In einer natürlichen Umgebung gestaltet sich dieser Gruppierungsprozeß schwieriger.

Dieses Minimierungsproblem wird durch den Einheitseigenvektor von A^TA zum kleinsten Eigenwert gelöst. Zur Lösung dieses Problems sind Verfahren aus der numerischen Mathematik bekannt (siehe z.B. Stoer [Sto93]).

Da Fluchtpunkte im folgenden zur Schätzung von Hyperbeln verwendet werden, findet hier noch eine statistische Betrachtung statt. Die Fehler ϵ_i werden mit den Varianzen $\sigma^2_{\epsilon_i}$ normalverteilt angenommen. Nach der Maximum-Likelihood-Schätzung gilt es dann, die Parameter so zu schätzen, daß der folgende Ausdruck maximiert wird:

$$e^{-\sum_{i=1}^{N} \frac{\epsilon_i^2}{2\sigma^2_{\epsilon_i}}} \quad \text{, wobei} \quad \epsilon_i = x_p \cos\theta_i + y_p \sin\theta_i - \rho_i$$

und σ^2_ϵ die Varianz des Fehlers ist, die in ungefähr der Varianz σ^2_ρ der Entfernung der Geraden vom Ursprung entspricht. In diesem Fall ist die Varianz von der Auflösung des zugrunde gelegten Parameterraumes der Hough-Transformation (siehe Kap. 4.2.1) abhängig. Die Kovarianzmatrix des berechneten homogenen Fluchtpunktes hat die Form:

$$\Sigma_{\boldsymbol{P}_\infty} = \begin{pmatrix} \sigma^2_x & \sigma^2_{xy} & \sigma^2_{xz} \\ \sigma^2_{xy} & \sigma^2_y & \sigma^2_{yz} \\ \sigma^2_{xz} & \sigma^2_{yz} & \sigma^2_z \end{pmatrix} = \sigma^2_\rho \left(A^TA\right)^{-1}. \tag{4.16}$$

4.2.4 Verfahren zur Hyperbelschätzung

In diesem Abschnitt wird ein Verfahren vorgestellt, das einer vorgegebenen Punktfolge eine Hyperbel anpaßt. Die Punktfolge ist entweder durch Bildpunkte in Rechnerkoordinaten oder durch Fluchtpunkte nach Kapitel 4.2.3 gegeben. Um eine stabile Berechnung der Hyperbel zu gewährleisten, werden die Bild- bzw. Fluchtpunkte mit ihren Varianzen gewichtet. Dazu wird die Kovarianzmatrix der Eingangsdaten verwendet. Diese Matrix wird außerdem zur Eliminierung der Verzerrung bei der Schätzung der Hyperbelkoeffizienten verwendet.

Die Gleichung des Kegelschnitts (3.46) kann in quadratischer Form als

$$\boldsymbol{P_p}^T A \boldsymbol{P_p} = 0$$

formuliert werden, wobei A eine symmetrische 3×3-Matrix ist. Für A gilt in diesem Fall $A_{12} = A_{21} = 0$, da für die Rotationsachse $\widehat{\boldsymbol{N}}$ aus Gleichung (3.46) entweder $\widehat{\boldsymbol{N}} = (1,0,0)$ oder $\widehat{\boldsymbol{N}} = (0,1,0)$ angenommen wird. Sei im folgenden $n \in \mathbb{N}$ die Anzahl der Punkte der zugrundeliegenden Punktfolge. Entkoppelt man die Fluchtpunktkoordinaten von den Unbekannten, so erhält man

$$\boldsymbol{d}^T \boldsymbol{a} = 0 \ ,$$

wobei

$$\boldsymbol{d}^T = \begin{pmatrix} {x_p}^2 & {y_p}^2 & x_p & y_p & 1 \end{pmatrix} \tag{4.17}$$

und

$$\boldsymbol{a}^T = \begin{pmatrix} A_{11} & A_{22} & 2A_{13} & 2A_{23} & A_{33} \end{pmatrix} . \tag{4.18}$$

Dies führt zur gewichteten Minimierung der kleinsten Quadrate des Terms

$$\min \sum_{i=1}^{n} \frac{(\boldsymbol{d}_i^T \boldsymbol{a})^2}{\mathrm{Var}[\boldsymbol{d}_i^T \boldsymbol{a}]} . \tag{4.19}$$

Werden Bildpunktfolgen in der Minimierungsaufgabe verwendet, so hat die Kovarianzmatrix $\Sigma_{\boldsymbol{P_p}}$ eines Bildpunktes die Form:

$$\Sigma_{\boldsymbol{P_p}} = \begin{pmatrix} \sigma_x^2 & \sigma_{xy}^2 \\ \sigma_{xy}^2 & \sigma_y^2 \end{pmatrix} . \tag{4.20}$$

Die Varianzen der Bildpunkte hängen von dem Verfahren ab, mit denen die Bildpunkte detektiert werden, und unterscheiden sich von den Varianzen der Fluchtpunkte der Kovarianzmatrix (4.16). Im folgenden bezeichnet Σ die Kovarianzmatrix eines Bild- oder Fluchtpunktes entsprechend der Punktfolge, die zur Hyperbelschätzung verwendet wird.

Berechnet man die Varianzen $\mathrm{Var}[\boldsymbol{d}_i^T \boldsymbol{a}]$ exakt, so stellt Gleichung (4.19) ein nichtlineares Minimierungsproblem dar, da der Vektor der Unbekannten $\boldsymbol{a}$ im Nenner jedes Summanden auftritt. Daher wird die Gewichtung approximativ durch die Spur $sp(\Sigma_i)$ der Kovarianzmatrix Σ_i jedes Punktes $\boldsymbol{P}_{\infty i}$ oder $\boldsymbol{P}_{\boldsymbol{p}_i}$ mit $i \in \mathbb{N}_{\leq n}$ vorgenommen [Kan93].

Da die Unbekannten des Vektors $\boldsymbol{a}$ bis auf einen Skalierungsfaktor eindeutig bestimmt sind, ist eine Nebenbedingung einzuführen, wenn nicht nach der exakten Gewichtung der Gleichung (4.19) minimiert wird. Nach Kanatani [Kan93] kann die Nebenbedingung

$$A_{11}^2 + A_{22}^2 = 1 ,$$

die die Invarianz gegenüber der Anpassung von Rotation und Translation ausnutzt, in dem Fall von nichtisotropischen Fehlern in den Punkten und ungleicher Gewichtung der Punkte nicht angewendet werden. Daher schlägt Kanatani [Kan93] vor, die Minimierung unter der gewöhnlichen Nebenbedingung $\|\boldsymbol{a}\| = 1$ durchzuführen. Dies führt zu dem Minimierungsproblem:

$$\min_{\|\boldsymbol{a}\|=1} \boldsymbol{a}^T C \boldsymbol{a} \text{ , wobei } C = \sum_{i=1}^{n} \frac{1}{sp(\Sigma_i)} \boldsymbol{d}_i \boldsymbol{d}_i^T . \tag{4.21}$$

Dieses Minimierungsproblem wird durch den Einheitseigenvektor von C zum kleinsten Eigenwert mit bekannten Verfahren (siehe z.B. Stoer [Sto93]) gelöst.

Die Matrix C ist durch ein Rauschen δC gestört, da das Rauschen der Bild - oder Fluchtpunkte $(x_p, y_p)^T$ in diese einfließt. Nimmt man normalverteiltes Rauschen $(\delta x_p, \delta y_p)$ mit dem Erwartungswert null in den Punkten an, wobei dieses durch die Kovarianzmatrix Σ nach Gleichung (4.16) oder (4.20) gegeben ist, so erhält man als Erwartungswert für $C + \delta C$:

$$E[C + \delta C] = \sum_{i=1}^{n} E[C_i + \delta C_i].$$

Um den Erwartungswert $E[C_i + \delta C_i]$ für einen Punkt i zu berechnen, wird der Index i zunächst sowohl für die Punktkoordinaten als auch für die Varianzen der Punkte weggelassen:

$$E[C_i + \delta C_i] = \frac{1}{sp(\Sigma_i)} E\left[\begin{pmatrix} (x_p + \delta x_p)^4 & (x_p + \delta x_p)^2 (y_p + \delta y_p)^2 & \cdots \\ (x_p + \delta x_p)^2 (y_p + \delta y_p)^2 & (y_p + \delta y_p)^4 & \cdots \\ \vdots & \vdots & \end{pmatrix}\right].$$

Nach der Berechnung der Momente erster bis vierter Ordnung der Variablen x_p und y_p erhält man aus obiger Matrix:

$$E[C_i + \delta C_i] = C_i + \frac{1}{sp(\Sigma_i)} \begin{pmatrix} C_{11} & C_{12} & 3x_p\sigma_x^2 & C_{14} & \sigma_x^2 \\ C_{21} & C_{22} & C_{23} & 3y_p\sigma_y^2 & \sigma_y^2 \\ 3x_p\sigma_x^2 & C_{32} & \sigma_x^2 & \sigma_{xy}^2 & 0 \\ C_{41} & 3y_p\sigma_y^2 & \sigma_{xy}^2 & \sigma_y^2 & 0 \\ \sigma_x^2 & \sigma_y^2 & 0 & 0 & 0 \end{pmatrix}, \tag{4.22}$$

wobei

$$\begin{aligned} C_{11} &= 6{x_p}^2\sigma_x^2 + 3\sigma_x^4 , \\ C_{12} = C_{21} &= {x_p}^2\sigma_y^2 + {y_p}^2\sigma_x^2 + 4x_p y_p \sigma_{xy}^2 + \sigma_x^2\sigma_y^2 + 2\sigma_{xy}^4 , \\ C_{22} &= 6{y_p}^2\sigma_y^2 + 3\sigma_y^4 , \\ C_{23} = C_{32} &= 2y_p\sigma_{xy}^2 + x_p\sigma_y^2 \text{ und} \\ C_{41} = C_{14} &= 2x_p {s_{xy}}^2 + y_p\sigma_x^2 . \end{aligned}$$

Aus dem Störungstheorem für Eigenvektoren symmetrischer Matrizen (siehe Kanatani [Kan93] und Golub und van Loan [GvL89]) kann die Verzerrung der Lösung $\hat{\boldsymbol{a}}$ berechnet werden:

$$E[\delta\hat{\boldsymbol{a}}] = -\Big(\Sigma_{j=1}^{4} \frac{\boldsymbol{u}_j \boldsymbol{u}_j^T}{\lambda_j}\Big) E[\delta C]\hat{\boldsymbol{a}} , \tag{4.23}$$

wobei $(\lambda_j, \boldsymbol{u}_j)$ die vier zugehörigen Eigenwert-Eigenvektoren Paare sind.

Nach (4.22) ist die Verzerrung $E[\delta\hat{\boldsymbol{a}}]$ genau dann gleich null, wenn auch der Erwartungswert $E[\delta C]$ null ist. Unglücklicherweise ist der Erwartungswert $E[\delta C]$ gleich der Summe der Matrizen $E[\delta C_i]$ jedes Punktes des zweiten Terms der

Gleichung (4.22) und ungleich null. Aus diesem Grund wird der Erwartungswert $E[\delta C_i]$ im zweiten Term der Gleichung (4.22) von jedem Punkt subtrahiert. Dies führt schließlich zu dem verzerrungsfreien Minimierungsproblem:

$$\min_{\|\boldsymbol{a}\|=1} \boldsymbol{a}^T \Big(\sum_{i=1}^{M} \frac{1}{sp(\Sigma_i)} \boldsymbol{d}_i \boldsymbol{d}_i^T - E[\delta C_i] \Big) \boldsymbol{a} \ . \tag{4.24}$$

4.2.5 Detektion von Bildmerkmalen eines Kalibrierwürfels

Sei im folgenden das Akkumulatorfeld einer Hough-Transformation (siehe Kap. 4.2.1) mit dem Parameter ρ als horizontaler Achse und dem Parameter θ als vertikaler Achse gegeben (siehe Abb. 4.6). Der Bildraum der Hough-Transformation sei das Binärbild eines Kalibrierwürfels nach einer Kantendetektion[3]. Der Kalibrierwürfel sei so gestaltet wie in Kapitel 4.1.1 und 4.1.2 beschrieben (Abb. 4.4).

Um aus den Geraden, welche die Hough-Transformation liefert, die Eckpunkte der Quader den Szenenkoordinaten zuordnen zu können, erfolgt die Bearbeitung des Bildraumes in drei Schritten. Im ersten Schritt werden die Szenenkoordinatensystemachsen dem Bild zugeordnet und so eine Segmentierung des Bildes erreicht (Abb. 4.7a). Anschließend werden Geraden in dem Akkumulatorfeld detektiert (Abb. 4.7b), um schließlich die durch Schnittpunkte detektierten Bildmerkmale (Abb. 4.9) den Szenenkoordinaten zuzuordnen (vgl. Kap. 4.1.2).

4.2.5.1 Segmentierung

Auf einem Kalibrierwürfel, wie er in Kapitel 4.1.1 beschrieben wurde, existieren drei Geradenscharen, die jeweils zu einer Achse des Szenenkoordinatensystems parallel sind. Daher ist es möglich, in dem Akkumulatorfeld drei vorherrschende Richtungen zu detektieren. Diese entsprechen dann den durch die Zentralprojektion (siehe Kap. 2.3.2) abgebildeten Szenenkoordinatensystemachsen bis auf eine Unsicherheit.

Um die vorherrschenden Richtungen der segmentierenden Geraden zu finden, kann wie folgt vorgegangen werden:

1. Wende auf das Akkumulatorfeld einen Schwellwertoperator an.

2. Summiere die Zeilen des Akkumulatorfeldes in einem eindimensionalen Array.

3. Falls nicht genau drei Zusammenhangskomponenten in dem entstandenen Array zu finden sind, beginne erneut mit 1. und einem anderen Schwellwert.

[3] Zu Verfahren der Kantendetektion sei auf Jähne [Jäh93] verwiesen.

4. Ansonsten bilde ein gewichtetes Mittel über jede Zusammenhangskomponente. Die vorherrschenden Richtungen sind durch die Mittel gegeben.

Hat man diese vorherrschenden Richtungen gefunden, so ist es möglich, die zu den Achsen des Szenenkoordinatensystems äquivalenten Geraden zu finden. Für jede Richtung führt man dazu folgenden Algorithmus aus:

1. Lege ein eindimensionales Array entsprechend der Größe des Parameters ρ in dem Akkumulatorfeld an.

2. Für jedes Feld in dem Array führe folgende Anweisungen durch:

   (a) Lege eine Gerade (θ, ρ) über das binäre Ausgangsbild, wobei θ durch die vorherrschende Richtung und ρ durch den Laufindex gegeben sind.

   (b) Trage in dem Feld die Länge der längsten Strecke dieser Geraden ein, die keine Punkte des Ausgangsbildes berührt.

3. Der Parameter ρ der gesuchten Geraden ist durch das Maximum in dem entstandenen Array gegeben.

Schneiden sich alle so generierten Geraden annähernd in einem Punkt, so hat man die segmentierenden Geraden gefunden. Ihr gemeinsamer Schnittpunkt entspricht dann bis auf einen Unsicherheitsfaktor dem Ursprung des Szenenkoordinatensystems.

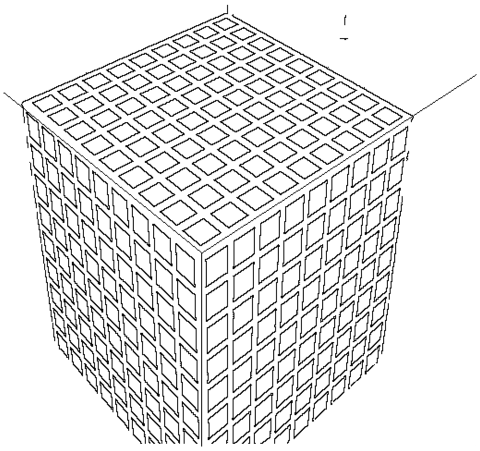
Schneiden sich die drei Geraden nicht annähernd in einem Punkt, so liegt eine fehlerhafte Berechnung vor. Dann kann entweder der Suchbereich für die segmentierenden Geraden auf einen Ausschnitt des Ausgangsbildes beschränkt werden oder die Bestimmung der vorherrschenden Richtungen erneut mit Schwellwerten durchgeführt werden, die noch nicht verwendet wurden.

Durch die drei Geraden wird eine Sechsteilung des Ausgangsbildes erreicht. Da die Geraden auf einer Seite des Ursprungs keine Schnittpunkte mit den Kanten der Quadrate haben, ist so eine Dreiteilung des Ausgangsbildes durch Zusammenfassen jeweils zweier benachbarter Gebiete möglich.

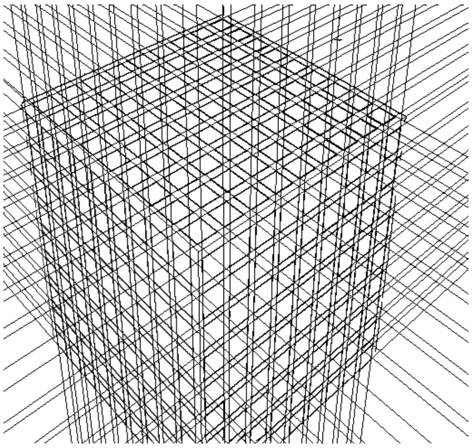
Diese Dreiteilung stellt eine Segmentierung des Ausgangsbildes dar, bei der jeweils genau eine Seite des Kalibrierwürfels in einem Segment abgebildet ist. Eine solche Segmentierung sei durch die Menge $\overline{\mathcal{S}}$ repräsentiert:

$$\overline{\mathcal{S}} = \{s_p = (\theta_p, \rho_p, i_p) | p \in \{0,1,2\}, i_p \in \{0,1\}\} \ , \tag{4.25}$$

wobei θ_p und ρ_p für alle p eine segmentierende Gerade festlegen. Der Parameter i_p legt für alle Geraden s_p fest, welcher Abschnitt einer Geraden, ausgehend vom Schnittpunkt aller Geraden, für die Segmentierung relevant ist. Der relevante Teil einer segmentierenden Geraden ist aus Abbildung 4.7a) zu ersehen. Die Menge $\overline{\mathcal{S}}$ repräsentiert aufgrund des dritten Parameters eine erweiterte Geradendarstellung.



(a) Segmentierung



(b) detektierte Geraden

Abbildung 4.7: Segmentierung des Kalibrierwürfels und detektierte Geraden : (a) Szenenkoordinatensystemachsen zur Segmentierung des Kalibrierwürfels, (b) detektierte Geraden auf dem Kalibrierwürfel.

Zum Verständnis des Parameters i_p betrachte man eine zu s_p orthogonale Gerade $g = (\theta_g, \rho_g)$ durch den Schnittpunkt aller Geraden aus $\overline{\mathcal{S}}$. Der relevante Abschnitt der Geraden s_p befindet sich dann auf einer Seite der orthogonalen Geraden g. Gilt für alle zu g parallelen Geraden $h = (\theta_h, \rho_h)$, die s_p an einer relevanten Stelle schneiden, $\rho_h \leq \rho_g$, so wird $i_p = 0$ gesetzt. Gilt für alle diese Geraden $\rho_h \geq \rho_g$, so setzt man $i_p = 1$.

Das erste Segment sei durch die Geraden s_0 und s_1, das zweite Segment durch die Geraden s_1 und s_2 und das dritte Segment durch die Geraden s_2 und s_0 gegeben.

In Abbildung 4.7a) ist eine Segmentierung des in Abbildung 4.4 abgebildeten Kalibrierwürfels dargestellt, die mit Hilfe des Akkumulatorfeldes aus Abbildung 4.6 erstellt wurde.

4.2.5.2 Geradendetektion

Zur Geradendetektion sei im folgenden eine Segmentierung gemäß (4.25) gegeben. Dann existieren zu jedem Segment zwei Geradenbüschel, deren Urbilder bezüglich der Zentralprojektion parallel sind und deren Richtung sich aus den vorherrschenden Richtungen der Segmentierung ergibt (vgl. Abb. 4.4). Diese werden mit der Hough-Transformation (siehe Kap. 4.2.2) berechnet, indem lokale Maxima in dem Akkumulatorfeld detektiert werden. Diese Geraden-Bestimmung erfolgt jeweils ausgehend von einer segmentierenden Geraden. Für jedes Segment erhält man so zwei Geradenscharen. Sämtliche Geraden werden durch die Geradenmengen $\mathcal{L}(m)_{ij}$ repräsentiert. Die Anzahl der detektierten Geraden ist abhängig von einer Zahl $m \in \mathbb{N}$, welche die Anzahl der Quadrate auf dem abgebildeten Würfel repräsentiert (vgl. Kap. 4.1.2), die in den drei Segmenten betrachtet werden:

$$\mathcal{L}(m)_{ij} = \{l_{ijk} = (\theta_{ijk}, \rho_{ijk}) | i \in \{0,1,2\}, j \in \{0,1\}, 0 \leq k < 2m\} \ . \tag{4.26}$$

Dabei gibt der Index i das Segment an, in dem eine Gerade detektiert wurde, der Index j, zu welcher segmentierenden Geraden diese Gerade in der Welt parallel verläuft. Der Index k sortiert die Geraden und steigt mit dem Abstand, die eine Gerade zu einer segmentierenden Geraden hat. Um die detektierten Geraden eindeutig einer Menge $\mathcal{L}(m)_{ij}$ und einem Index k zuzuordnen, gelte folgende Annahme: Für $(\theta_{ijk}, \rho_{ijk})$, $(\theta_{ijl}, \rho_{ijl}) \in \mathcal{L}(m)_{ij}$ mit $l > k$ und $(\theta_p, \rho_p) \in \overline{\mathcal{S}}$ gilt:

1. $(i + j) \bmod 3 = p \Rightarrow \theta_{ijk} \approx \theta_p$ und
2. $(i + j) \bmod 3 = p \Rightarrow |\rho_{ijk} - \rho_p| < |\rho_{ijl} - \rho_p|$.

Dabei bedeutet die erste Annahme, daß eine Gerade, die sich in der Welt in demselben Parallelenbüschel wie $s_0 \in \overline{\mathcal{S}}$ befindet, die Indizes $i = 0$ und $j = 0$ oder $i = 2$ und $j = 1$ hat. Eine Gerade, die sich in der Welt in demselben Parallelenbüschel wie $s_1 \in \overline{\mathcal{S}}$ befindet, hat die Indizes $i = 0$ und $j = 1$ oder $i = 1$ und $j = 0$ und eine Gerade, die sich in der Welt in demselben Parallelenbüschel wie $s_2 \in \overline{\mathcal{S}}$ befindet, die Indizes $i = 1$ und $j = 1$ oder $i = 2$ und $j = 0$. Diese Definition besitzt einen Toleranzbereich, der sich aus dem Zeichen „$\approx$“ der ersten Annahme ergibt. Durch den Benutzer kann eine kleine perspektivische Verzerrung der Abbildung durch die Zentralprojektion zugesichert werden, so daß die Wahl dieses Toleranzbereiches kein Problem darstellt.

Die zweite Annahme drückt aus, daß die Entfernung in jedem Geradenbüschel eines jeden Segments zwischen der Geraden und der zugehörigen segmentierenden Geraden (siehe Annahme 1.) mit zunehmendem dritten Index steigt.

Geraden aus dem ersten Segment tragen den Index $i = 0$, Geraden aus dem zweiten Segment den Index $i = 1$ und Geraden aus dem dritten Segment den Index $i = 2$.

In Abbildung 4.7b) ist die Detektion von Geraden eines Kalibrierwürfels dargestellt, die aufgrund der Segmentierung in Abbildung 4.7a) vorgenommen wurde. Des weiteren zeigt die Abbildung 4.8 zusammen mit der Tabelle 4.2, wie eine solche Geradenmenge aussehen kann. Die Indizes i, j und k sind dabei wie in (4.26) zu verstehen.

4.2.6 Zuordnung der Bildmerkmale eines Kalibrierwürfels

Der Schnittpunkt zweier Geraden muß bekannt sein, um Bildmerkmale eines Kalibrierwürfels aus Geraden zu bestimmen. Dieser ergibt sich aus (4.9) zu:

$$\Psi : (\mathbb{R} \times \mathbb{R}) \times (\mathbb{R} \times \mathbb{R}) \rightarrow \mathbb{R} \times \mathbb{R} \ , \ ((\theta_1, \rho_1), (\theta_2, \rho_2)) \mapsto (\mathrm{x}, \mathrm{y}),$$

$$\text{wobei} \quad x = \rho_1 \sin\theta_2 - \rho_2 \sin\theta_1 \text{ und } y = \rho_2 \cos\theta_1 - \rho_1 \cos\theta_2. \tag{4.27}$$

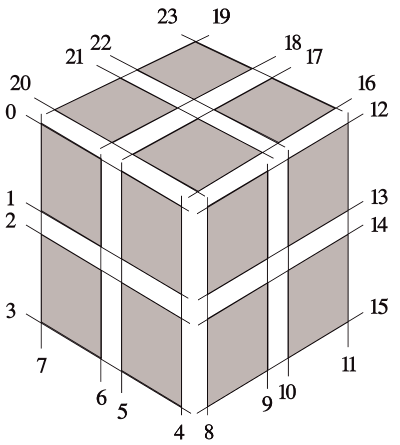


Abbildung 4.8: Beispielgeraden zur Zuordnung der Bildmerkmale eines Kalibrierwürfels. In diesem Fall gilt $m = 2$.

i	j	k	θ_{ijk}	ρ_{ijk}	i	j	k	θ_{ijk}	ρ_{ijk}
0	0	0	1.54	-51.1	1	1	0	2.57	-4.6
0	0	1	1.52	-70.6	1	1	1	2.59	-26.8
0	0	2	1.52	-76.4	1	1	2	2.59	-35.6
0	0	3	1.51	-94.0	1	1	3	2.61	-55.9
0	1	0	0.70	-44.5	2	0	0	2.55	11.1
0	1	1	0.68	-19.7	2	0	1	2.54	34.3
0	1	2	0.67	-10.1	2	0	2	2.53	41.6
0	1	3	0.65	13.4	2	0	3	2.51	63.1
1	0	0	0.72	-61.3	2	1	0	1.55	-37.9
1	0	1	0.74	-82.5	2	1	1	1.57	-13.8
1	0	2	0.74	-89.6	2	1	2	1.57	-6.8
1	0	3	0.76	-108.3	2	1	3	1.59	15.3

Tabelle 4.2: Zahlenwerte zur Illustration der Beispielgeraden aus Abb. 4.8.

Um die Ecken der Quadrate eindeutig den Bildmerkmalen zuzuordnen, müssen auch diese, wie die Szenenmerkmale, durchnumeriert werden. Dazu wird die Geradenmenge $\mathcal{L}(m)_{ij}$ aus (4.26) verwendet, da bestimmte Schnittpunkte zweier Geraden aus der Menge $\mathcal{L}(m)_{ij}$ ein Bildmerkmal auszeichnen. Sei dazu eine Funktion Π_b definiert, die als Argumente zwei Geraden und eine natürliche Zahl $m \in \mathbb{N}$ besitzt, wobei m die Anzahl der Quader angibt, die in x-, y- und z-Richtung betrachtet werden (vgl. Kap. 4.1.2):

$$\Pi_b : (\mathbb{R} \times \mathbb{R}) \times (\mathbb{R} \times \mathbb{R}) \times \mathbb{N} \to \mathbb{N} \;,$$
$$(g_{i0k}, g_{i1k'}, m) \mapsto \Omega(k/2, k \bmod 2, k'/2, k' \bmod 2) + 4i(m^2) \;, \qquad (4.28)$$

wobei

$$g_{i0k} \in \mathcal{L}(m)_{i0} \text{ und } g_{i1k'} \in \mathcal{L}(m)_{i1} \;.$$

Die Funktion Ω wurde in (4.1) definiert, und Π_b ist partiell definiert, da nicht alle Schnittpunkte der Geraden aus $\mathcal{L}(m)_{ij}$ Bildmerkmale darstellen.

Betrachtet man die Zuordnungvorschrift der Gleichung (3.6) der klassischen Kalibriermethoden (siehe Kap. 3.2), so hat diese unter Anwendung der beschriebenen Zuordnungsvorschrift folgende Form:

$$\mathcal{Z} = \{(x_s, y_s, z_s, \Psi(g_{i0k}, g_{i1k'}))|\Pi_s(x_s, y_s, z_s, m, q_k, q_d) = \Pi_b(g_{i0k}, g_{i1k'}, m)\},$$

wobei

$$m, q_k, q_d, i, k, k' \in \mathbb{N} \;, \mathrm{g_{i0k}} \in \mathcal{L}(\mathrm{m})_{\mathrm{i0}} \text{ und } \mathrm{g_{i1k'}} \in \mathcal{L}(\mathrm{m})_{\mathrm{i1}} \;.$$

Anschließend wird den Schnittpunkten zweier Geraden (siehe Kap. 4.2.6), die einem Bildmerkmal auf dem Kalibrierwürfel entsprechen, eine natürliche Zahl über die Funktion Π_b zugeordnet. Diese Abbildung ist abhängig von dem Parameter $m \in \mathbb{N}$. Wählt man für die Zuordnung der Szenenmerkmale (siehe Kap. 4.1.2) und der Bildmerkmale denselben Parameter $m \in \mathbb{N}$, so erhält man eine

i	k	k'	Ψ_x	Ψ_y	Π_b	i	k	k'	Ψ_x	Ψ_y	Π_b	i	k	k'	Ψ_x	Ψ_y	Π_b
0	0	0	205	277	0	1	0	0	218	277	16	2	0	0	212	261	32
0	0	1	206	306	1	1	0	1	242	262	17	2	0	1	194	246	33
0	1	1	187	289	2	1	1	1	242	291	18	2	1	1	216	232	34
0	1	0	186	260	3	1	1	0	219	306	19	2	1	0	234	248	35
0	0	2	206	316	4	1	0	2	249	257	20	2	0	2	187	239	36
0	0	3	207	343	5	1	0	3	271	243	21	2	0	3	169	225	37
0	1	3	189	325	6	1	1	3	270	271	22	2	1	3	191	212	38
0	1	2	187	299	7	1	1	2	248	286	23	2	1	2	208	226	39
0	2	2	181	293	8	1	2	2	248	295	24	2	2	2	217	221	40
0	2	3	183	320	9	1	2	3	270	280	25	2	2	3	199	207	41
0	3	3	165	303	10	1	3	3	269	307	26	2	3	3	220	195	42
0	3	2	163	277	11	1	3	2	248	322	27	2	3	2	237	209	43
0	2	0	180	255	12	1	2	0	219	316	28	2	2	0	243	242	44
0	3	1	181	283	13	1	3	1	242	300	29	2	3	1	224	227	45
0	2	1	163	267	14	1	2	1	242	327	30	2	2	1	245	215	46
0	3	0	161	238	15	1	3	0	219	343	31	2	3	0	264	229	47

Tabelle 4.3: Beispielzuordnung der Bildmerkmale für den Fall $m = 2$

eindeutige und korrekte Zuordnung der Szenenmerkmale zu den Bildmerkmalen, denen durch die Funktionen Π_s und Π_b dieselbe natürliche Zahl zugewiesen wird.

Um diese Zuordnung zu verdeutlichen, ist in Tabelle 4.3 eine Beispielzuordnung für den Fall $m = 2$ abgebildet. Diese Tabelle knüpft an das Beispiel aus Abbildung 4.3 und Tabelle 4.1 an. Die Parameter i, j, k und k' sind dabei die Parameter der Funktion Π_b, durch die zwei Geraden aus der Menge $\mathcal{L}$ eindeutig bestimmt werden. Der Ausgabeparameter der Funktion ist in der Spalte Π_b zu finden. Die Schnittpunkte der jeweiligen Geraden sind in den Spalten Ψ_x und Ψ_y zu finden und wurden aus den Beispielgeraden der Tabelle 4.2 berechnet. Jedes Szenenmerkmal der Tabelle 4.1 mit derselben Zuordnungsnummer wie ein Bildmerkmal aus der Tabelle 4.3 wird dem angegebenen Schnittpunkt (Ψ_x, Ψ_y) zugeordnet.

Abschließend sind in Abbildung 4.9 die detektierten Schnittpunkte der Geraden der Abbildung 4.7b) dargestellt, die durch die beschriebene Vorschrift den Szenenkoordinaten zugeordnet werden.

4.2.7 Detektion von konjugierten Fluchtpunkten eines Kalibrierwürfels

Setzt man die Verwendung eines unter Kapitel 4.1.1 beschriebenen Kalibrierwürfels voraus, so erhält man mit der Hough-Transformation die Geradenmengen $\mathcal{L}(m)_{ij}$ nach Gleichung (4.26) nach dem in Kapitel 4.2.5 beschriebenen Verfahren. Daraus erhält man drei Mengen von Geraden:

$$\begin{aligned}
\overline{\mathcal{L}}_0 &= \mathcal{L}(m)_{01} \cup \mathcal{L}(m)_{10} \ , \\
\overline{\mathcal{L}}_1 &= \mathcal{L}(m)_{11} \cup \mathcal{L}(m)_{20} \ , \\
\overline{\mathcal{L}}_2 &= \mathcal{L}(m)_{21} \cup \mathcal{L}(m)_{00} \ , \\
\text{wobei } m &\in \mathbb{N} \ .
\end{aligned}$$

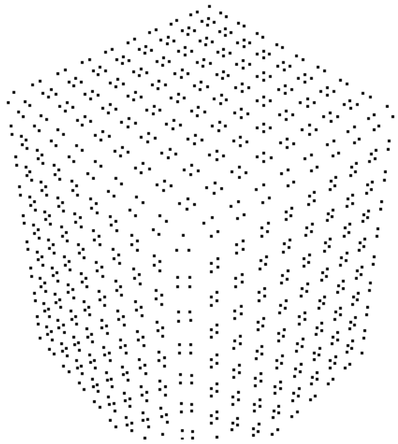


Abbildung 4.9: Detektierte Bildmerkmale auf einem Kalibrierwürfel

Für jede dieser Geradenmengen sind bezüglich der Zentralprojektion die Urbilder aller Geraden parallel. Des weiteren sind die Urbilder von Geraden aus unterschiedlichen Mengen $\overline{\mathcal{L}}_i$ und $\overline{\mathcal{L}}_j$ orthogonal zueinander. Durch Anwendung der in Kapitel 4.2.3 beschriebenen Verfahren zur Fluchtpunktermittlung erhält man aus den drei Geradenmengen $\overline{\mathcal{L}}_0$, $\overline{\mathcal{L}}_1$ und $\overline{\mathcal{L}}_2$ drei konjugierte Fluchtpunkte.

Eine weitere Möglichkeit, konjugierte Fluchtpunkte zu schätzen, wird durch Lutton et al. [LMLK94] beschrieben. Die Schätzung der Fluchtpunkte wird dabei direkt mit der Hough-Transformation (vgl. Kap. 4.2.1) durchgeführt.

Kapitel 5

Ergebnisse

In diesem Kapitel werden durchgeführte Simulationen mit synthetischen Daten und Realexperimente beschrieben und deren Ergebnisse dargestellt. In allen Experimenten werden folgende Kalibriermethoden untersucht:

- Die klassische Methode nach Kapitel 3.2.
- Eine kombinierte Methode aus der Zoommethode nach Kapitel 3.4.1 und der Methode der konjugierten Fluchtpunkte nach Kapitel 3.3.
- Die Rotationsmethode nach Kapitel 3.4.3.
- Die Hyperbelmethode nach Kapitel 3.4.4.

Alle Verfahren berechnen die internen Parameter des Bildhauptpunktes und der Skalierungsfaktoren. Keiner der internen Parameter wird zur Berechnung als bekannt vorausgesetzt. Die kombinierte Methode aus der Zoommethode und der Methode der konjugierten Fluchtpunkte berechnet zunächst die internen Parameter des Bildhauptpunktes mit der Zoommethode. Anschließend werden die Skalierungsfaktoren durch die Methode der konjugierten Fluchtpunkte mit bekanntem Bildhauptpunkt berechnet.

5.1 Simulation mit synthetischen Daten

In diesem Abschnitt werden die Resultate verschiedener Simulationen mit synthetischen Daten vorgestellt. Durch die Simulation mit idealen Daten wurden die implementierten Methoden auf Korrektheit getestet. Es wurde das Verhalten der Verfahren in bezug auf die Variation aller Experimentfaktoren untersucht. Eine solche Untersuchung ist bei einem Realexperiment nicht möglich. Des weiteren ist der Verlauf des Schätzfehlers als Funktion der Aufnahmegeometrie nicht aus einer analytischen Fortpflanzung des Meßfehlers in den zugrundeliegenden Gleichungen ermittelbar.

Es werden zwei verschiedene Kalibrierobjekte simuliert. Für die klassische Methode, die Zoommethode und die Methode der konjugierten Fluchtpunkte wird ein Kalibrierwürfel (Kap. 4.1.1) und für die Rotations- und Hyperbelmethode ein Schachbrettmuster simuliert.

Eine Simulation hat jeweils die folgende Form: Es wird die relative Lage des Kalibrierobjekts bezüglich der virtuellen Kamera festgelegt. Im Fall einer Rotation der Kamera werden die resultierenden relativen Lagen des Kalibrierobjekts, ausgehend von der Ruheposition der Kamera, neu berechnet. Dazu wird eine Rotationsachse und ein Rotationswinkel vorgegeben.

Es werden die internen Parameter der Kamera vorgegeben. Für die Zoommethode werden unterschiedliche Parameter vorgegeben. Die relevanten Punkte des Kalibrierobjekts werden auf die virtuelle Sensorfläche der Kamera projiziert. Auf die Bildpunkte wird ein Gaußsches Rauschen addiert. Da die Rotationsmethode den Rotationswinkel als bekannt voraussetzt, wird auch auf diesen ein Gaußsches Rauschen addiert. Aus diesen simulierten Daten werden die internen Parameter rekonstruiert.

Durch Addition von unterschiedlichem Rauschen erhält man unterschiedliches Datenmaterial. Die Konstruktion dieser Daten und die anschließende Rekonstruktion der internen Parameter wird in allen Simulationen 1000-fach durchgeführt. Durch Mittelung der Werte der rekonstruierten internen Kameraparameter wurde nachgewiesen, daß sämtliche Verfahren unverzerrt sind. Das heißt, daß die Erwartungswerte der internen Parameter den tatsächlichen Werten entsprechen.

Um den relativen Fehler einer Simulation zu ermitteln, wird zu jedem Versuch die Abweichung der rekonstruierten Parameter von den vorgegebenen Parametern berechnet. Anschließend wird der Mittelwert der Beträge über alle durchgeführten Versuche gebildet. Im einzelnen wird die Rekonstruktion der internen Parameter auf folgende Weise durchgeführt:

Bei der klassischen Methode und der Zoommethode werden die konstruierten Daten direkt als Eingabe verwendet. Zur Rekonstruktion mit der Methode der konjugierten Fluchtpunkte werden den konstruierten Daten Geraden angepaßt, die den Geraden entsprechen, die ein Bildverarbeitungsprogramm in der Aufnahme eines Kalibrierwürfels detektieren würde. Anschließend werden die Geraden so gruppiert, daß drei Geradenbüschel im Urbild paralleler Geraden entstehen. Zwei Geraden aus unterschiedlichen Geradenbüscheln stehen im Urbild jeweils orthogonal zueinander (Kap. 4.2.7). Zu jedem Geradenbüschel wird ein Fluchtpunkt geschätzt (Kap. 4.2.3). Diese Fluchtpunkte dienen zusammen mit dem rekonstruierten Bildhauptpunkt eines Versuchs der Zoommethode als Eingabedaten der Methode der konjugierten Fluchtpunkte.

Den aktiven Verfahren (Rotations- und Hyperbelmethode) liegt ein Schachbrettmuster zugrunde. Die Eckpunkte dieses Schachbrettmusters werden durch obiges Verfahren auf die virtuelle Sensorfläche der Kamera projiziert. Anschlies send werden diesen Punkten Geraden so angepaßt, daß ein Geradenbüschel in der Welt paralleler Geraden entsteht. Zu diesem Geradenbüschel wird ein Fluchtpunkt bestimmt (Kap. 4.2.3).

Zur Kalibrierung der internen Parameter x_0 und s_x wird ein Vergenzschwenk um die y-Achse des Kamerakoordinatensystems simuliert. Zur Kalibrierung der internen Parameter y_0 und s_y werden Aufnahmen eines Neigungsschwenks um die x-Achse des Kamerakoordinatensystems simuliert. In jeder dieser Simulationen wird ein Fluchtpunkt nach obigem Verfahren geschätzt.

Bei der Rotationsmethode werden die ermittelten Fluchtpunkte direkt zur Kalibrierung verwendet (Kap. 3.4.3). Die Hyperbelmethode paßt den Fluchtpunkten eine Hyperbel an, um die internen Parameter aus den Hyperbelkoeffizienten zu berechnen (Kap. 3.4.4).

Die Ergebnisse der Testdaten, die in diesem Abschnitt präsentiert werden, beschränken sich auf die internen Parameter x_0 und s_x, da die internen Parameter y_0 und s_y sich exakt symmetrisch verhalten. In Abbildung 5.1 ist die Legende zu sämtlichen Fehlerdiagrammen der Simulationen mit synthetischen Daten abgebildet. Die Ergebnisse aus der Analyse der Simulationen mit synthetischen Daten werden jeweils als relative Fehler zu den vorgegebenen Werten der internen Parameter in der Bildgenerierung angegeben.

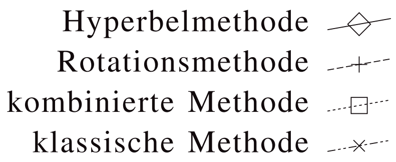


Abbildung 5.1: Legende zu den Fehlerdiagrammen der Simulationen mit synthetischen Daten. Die kombinierte Methode ermittelt den Bildhauptpunkt mit der Zoommethode, die Skalierungsfaktoren werden anschließend mit der Methode der konjugierten Fluchtpunkte ermittelt.

5.1.1 Meßfehler in den Punkten und im Rotationswinkel

Zunächst wird die Abhängigkeit des relativen Fehlers der verschiedenen Methoden von der Standardabweichung des Meßfehlers in den Bilddaten und in der Winkelbestimmung untersucht.

Für alle Verfahren wird eine Entfernung zwischen der Kamera und dem Kalibrierobjekt (Kalibrierwürfel oder Schachbrettmuster) von $1000mm$ angenommen. Bei der klassischen Methode werden jeweils 972 Bildmerkmale produziert. Bei der Zoommethode werden 10 unterschiedliche Aufnahmen mit unterschiedlicher Brennweite simuliert. Für diese Methode werden pro Aufnahme ebenfalls 972 Bildmerkmale produziert.

Für die Methode der konjugierten Fluchtpunkte werden auf einem Kalibrierwürfel drei Geradenbüschel zu je 36 Geraden aus den Bilddaten geschätzt.

Für die Rotations- bzw. Hyperbelmethode werden 36 Geraden eines Schachbrettmusters zur Fluchtpunktbestimmung produziert. Weiterhin wird bei diesen Methoden ein Gesamtschwenk von 120 Grad mit 12 Rotationen von je 10 Grad durchgeführt.

Es zeigt sich, daß alle Verfahren von dem Fehler in den Pixelkoordinaten abhängig sind. Man stellt fest, daß sämtliche Verfahren mit steigender Standardabweichung des Meßfehlers in den Bilddaten einem größeren relativen Fehler ausgesetzt sind (Abb. 5.2). Die Standardabweichung des Meßfehlers in der Winkelbestimmung der Rotationsmethode (Kap. 3.4.3) wird in diesem Fall nicht betrachtet.

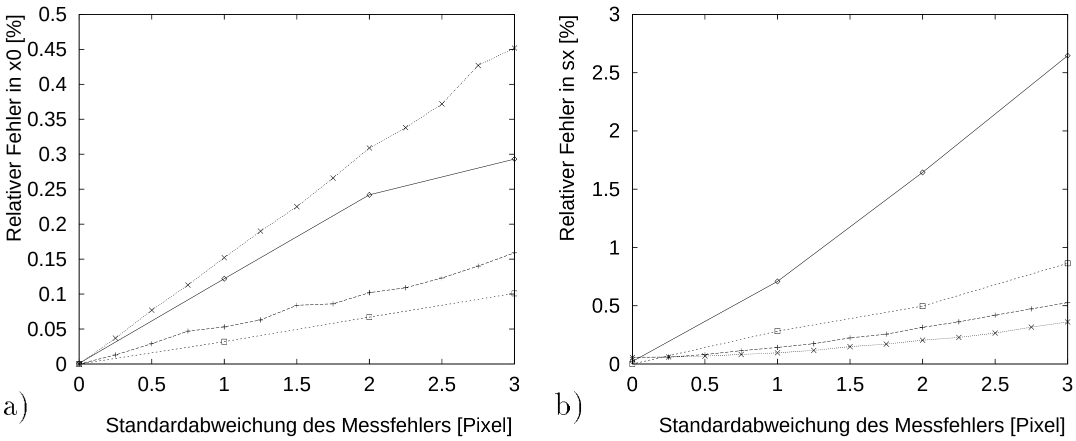


Abbildung 5.2: Der relative Fehler in x_0 ist in Abbildung a) und in s_x in Abbildung b) als Funktion der Standardabweichung des Meßfehlers in den Bilddaten dargestellt.

Jetzt wird die Standardabweichung des Meßfehlers in der Winkelbestimmung der Rotationsmethode (Kap. 3.4.3) untersucht. Abbildung 5.3 ist zu entnehmen, daß der relative Fehler mit zunehmender Standardabweichung steigt. Die Standardabweichung des Meßfehlers in den Bilddaten wird in dieser Betrachtung vernachlässigt.

Um die Rotationsmethode unter Einwirkung beider Standardabweichungen mit den übrigen Methoden vergleichen zu können, ist in Tabelle 5.1 die Kombination der Standardabweichungen ergänzt. Aus Abbildung 5.2 und Tabelle 5.1 ersieht man, daß der relative Fehler in den internen Parametern x_0 und s_x bei der klassischen Methode (Kap. 3.2) etwa derselben Größenordnung entspricht. Bei der kombinierten Methode wirkt sich der Meßfehler in den Bilddaten mehr auf s_x als auf den Parameter x_0 aus. Dies ist dadurch zu begründen, daß bei der Bestimmung von s_x neben dem Meßfehler in den Bilddaten zusätzlich der Meßfehler von x_0 eingeht.

Die Hyperbelmethode (Kap. 3.4.4) liefert wesentlich stabilere Ergebnisse für den internen Parameter x_0 als für s_x. Fluchtpunkte, die sich in einer geringeren Entfernung vom Bildzentrum befinden, besitzen eine geringere Standardabweichung. Somit besitzen Fluchtpunkte, die die Steigung der Hyperbelschenkel

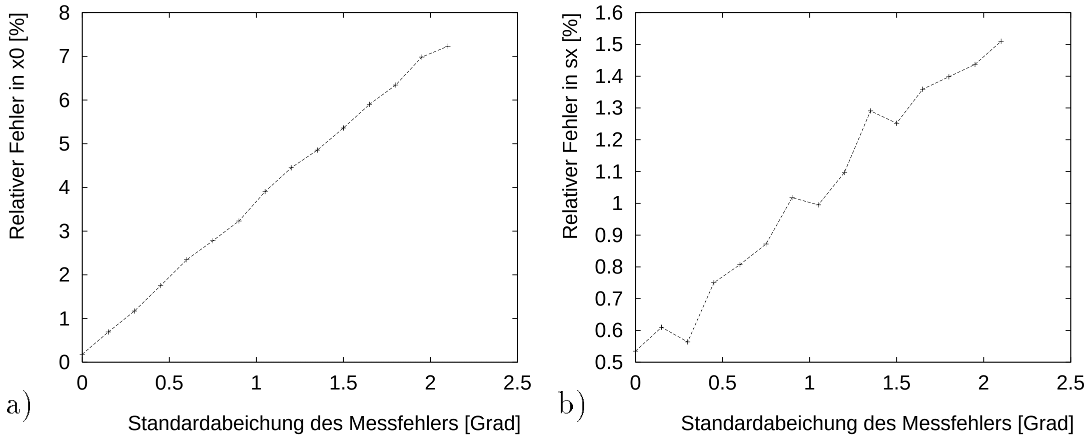


Abbildung 5.3: Der relative Fehler in x_0 ist in Abbildung a) und in s_x in Abbildung b) als Funktion der Standardabweichung des Meßfehlers in der Winkelbestimmung dargestellt.

bestimmen, eine größere Unsicherheit als die Fluchtpunkte, die den Hyperbelscheitel bestimmen. Der Unterschied in dem relativen Fehler der Bestimmung von x_0 und s_x läßt sich somit dadurch begründen, daß der interne Parameter x_0 durch den Hyperbelscheitel bestimmt wird, s_x hingegen durch die Steigung der Schenkel der Hyperbel.

Die Rotationsmethode (Kap. 3.4.3) ist wesentlich stärker von der Standardabweichung in der Bestimmung der Rotationswinkel als von der der Bilddaten abhängig. Insgesamt wirkt sich eine Unsicherheit wesentlich stärker auf x_0 als auf s_x aus.

Aus Abbildung 5.2 ersieht man, daß die klassische Methode (Kap. 3.2) in dem Parameter x_0 die schlechtesten Ergebnisse liefert. In s_x hingegen liefert sie die besten Ergebnisse. Beachtet man bei der Rotationsmethode (Kap. 3.4.3) auch den Meßfehler in der Winkelbestimmung (Abb. 5.3), so ersieht man aus Tabelle 5.1, daß die klassische Methode in beiden Parametern bessere Werte liefert.

Die kombinierte Methode liefert für x_0 die besten Ergebnisse. Für s_x erhält man nur durch die klassische Methode bessere Ergebnisse. Die beiden auf Rotationen basierenden Verfahren liefern insgesamt schlechtere Ergebnisse als die übrigen Methoden. Eine Ausnahme bildet die Hyperbelmethode in der Schätzung von x_0. In diesem Fall liefert eine auf Rotationen basierende Methode stabilere Ergebnisse als die klassische Methode.

Aufgrund des großen relativen Fehlers durch die Standardabweichung des Meßfehlers in der Winkelbestimmung liefert die Rotationsmethode in dem Parameter x_0 schlechtere Ergebnisse als die Hyperbelmethode. Dies ist um so interessanter, da für die Hyperbelmethoden weniger Informationen als bekannt vorausgesetzt werden. Diese Tatsache läßt sich dadurch begründen, daß die Standardabweichung in den Bilddaten in die Berechnung der Hyperbel mit einfließt (Kap. 4.2.4). Eine solche Gewichtung findet bei der Rotationsmethode nicht statt.

$intP$	$Fehler$	$klass$	$ZFpkt$	$Hyper$	$Rot0.5$	$Rot1.0$	$Rot2.0$
x_0	0.5	0.08	0.02	0.06	1.83	3.93	7.13
	1.0	0.15	0.03	0.12	1.85	3.95	7.15
	2.0	0.31	0.07	0.24	1.90	4.00	7.20
	3.0	0.45	0.10	0.29	1.96	4.06	7.26
s_x	0.5	0.07	0.14	0.35	0.84	1.08	1.57
	1.0	0.10	0.28	0.71	0.88	1.12	1.61
	2.0	0.20	0.50	1.64	0.93	1.17	1.66
	3.0	0.35	0.86	2.65	1.03	1.27	1.76

Tabelle 5.1: Vergleich des relativen Fehlers in x_0 und s_x unter Beachtung der Standardabweichung des Meßfehlers in den Bilddaten und in der Winkelbestimmung. Die Bezeichnungen in dieser Tabelle sind $intP$ für den betrachteten internen Kameraparameter, $Fehler$ für die Standardabweichung des Meßfehlers in den Bilddaten, $klass$ für die klassische Methode, $ZFpkt$ für die kombinierte Methode aus der Zoommethode und der Methode der konjugierten Fluchtpunkte, $Hyper$ für die Hyperbelmethode und $Rot0.5$, $Rot1.0$ bzw. $Rot2.0$ für die Rotationsmethode mit einer Standardabweichung des Meßfehlers in der Winkelbestimmung von 0.5, 1.0 bzw. 2.0 Grad.

Bei einem großen Meßfehler in den Punkten liefert die Hyperbelmethode für s_x schlechtere Ergebnisse als die Rotationsmethode. Dies liegt an der schon oben beschriebenen Tatsache, daß die Berechnung der Steigung der Hyperbelschenkel einen größeren Unsicherheitsfaktor birgt. Bei geringem Meßfehler in den Punkten liefert die Hyperbelmethode bessere Ergebnisse für den internen Parameter s_x als die Rotationsmethode.

5.1.2 Entfernung des Kalibrierobjekts

Hier wird die Abhängigkeit des relativen Fehlers der verschiedenen Methoden von der Entfernung des Kalibrierobjekts untersucht.

Für alle Verfahren wird eine Standardabweichung des Meßfehlers in den Bilddaten von 1.0 Pixel angenommen. Für die klassische Methode werden 972 Bildmerkmale produziert. Für die Zoommethode werden in 10 Aufnahmen mit unterschiedlicher Brennweite ebenfalls 972 Bildmerkmale produziert und einander zugeordnet. Für die Methode der konjugierten Fluchtpunkte werden drei Geradenbüschel zu je 36 Geraden auf einem Kalibrierwürfel aus den Bilddaten geschätzt.

Bei der Rotations- bzw. Hyperbelmethode werden 36 Geraden eines Schachbrettmusters zur Fluchtpunktbestimmung aus den Bilddaten geschätzt. Für die Rotationsmethode wird des weiteren eine Standardabweichung des Meßfehlers in der Winkelbestimmung von 0.5 Grad angenommen. Bei der Rotations- und der

Hyperbelmethode wird ein Gesamtschwenk von 120 Grad durch 12 Rotationen zu je 10 Grad simuliert.

Mit zunehmender Entfernung des Kalibrierobjekts verringern sich die Abstände der einzelnen Bildmerkmale und der Geraden in den Bildern zueinander. Abbildung 5.4 ist zu entnehmen, daß der relative Fehler in der Bestimmung der internen Parameter in allen Verfahren mit zunehmender Entfernung des Kalibrierobjekts steigt.

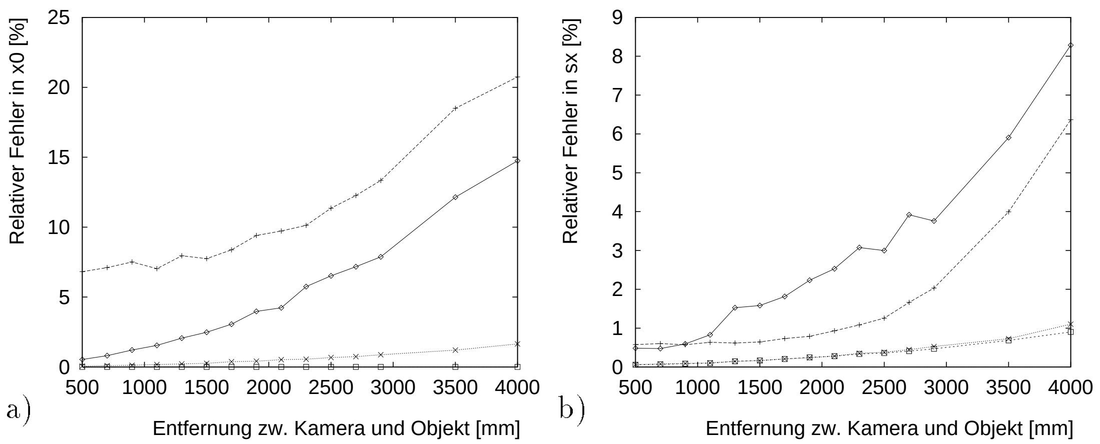


Abbildung 5.4: Der relative Fehler in x_0 ist in Abbildung a) und in s_x in Abbildung b) als Funktion der Entfernung zwischen der Kamera und dem Objekt dargestellt.

In dem Fall, daß die Methode Bildpunkte verwendet, liegt dies an der Tatsache, daß die Diskretisierung der Bildpunktdaten mit steigender Entfernung ebenfalls zunimmt. Der zugrundeliegende Kalibrierwürfel wird auf einen kleineren Ausschnitt auf die Sensorfläche projiziert.

Verwendet das Verfahren Fluchtpunkte, so wird der eben beschriebene Unsicherheitsfaktor noch durch einen zweiten ergänzt. Mit zunehmender Entfernung des Kalibrierobjekts nimmt die Variation der Richtungen der angepaßten Geraden ab. Somit steigt die Unsicherheit in der Bestimmung der Fluchtpunkte, obwohl diese translationsinvariant sind.

Aus diesem Grund wirken sich Variationen in der Entfernung des Kalibrierobjekts bei der Rotations- und Hyperbelmethode stärker aus als bei der klassischen Methode. Die Methode der konjugierten Fluchtpunkte wird ebenfalls nicht so stark durch diesen Faktor beeinflußt, da die Zoommethode kaum durch Entfernungsvariationen beeinflußt wird.

5.1.3 Gesamtschwenkwinkel

Für auf Rotationen basierenden Kalibriermethoden ist der wichtigste Unsicherheitsfaktor die Abhängigkeit des relativen Fehlers von dem Gesamtschwenkwinkel. Diese Abhängigkeit wird hier untersucht.

Für die Rotationsmethode und die Hyperbelmethode wird eine Standardabweichung des Meßfehlers in den Bilddaten von 1.0 Pixel angenommen. Es werden 36 Geraden eines Schachbrettmusters zur Fluchtpunktbestimmung angepaßt. Für die Rotationsmethode wird des weiteren eine Standardabweichung des Meßfehlers in der Winkelbestimmung von 0.5 Grad angenommen. Das Kalibrierobjekt wird in einer Entfernung von $1000mm$ vor der Kamera angenommen. Zu den unterschiedlichen Gesamtschwenkwinkeln werden jeweils 12 Rotationen simuliert.

In Abbildung 5.5 a) und b) sind die Fluchtpunkte für den Fall eines Gesamtschwenks von 120 und 20 Grad auf einer angepaßten Hyperbel dargestellt. Abbildung 5.5 c) und d) illustrieren den relativen Fehler in den Parametern x_0 und s_x als Funktion des Gesamtschwenkwinkels der Rotation.

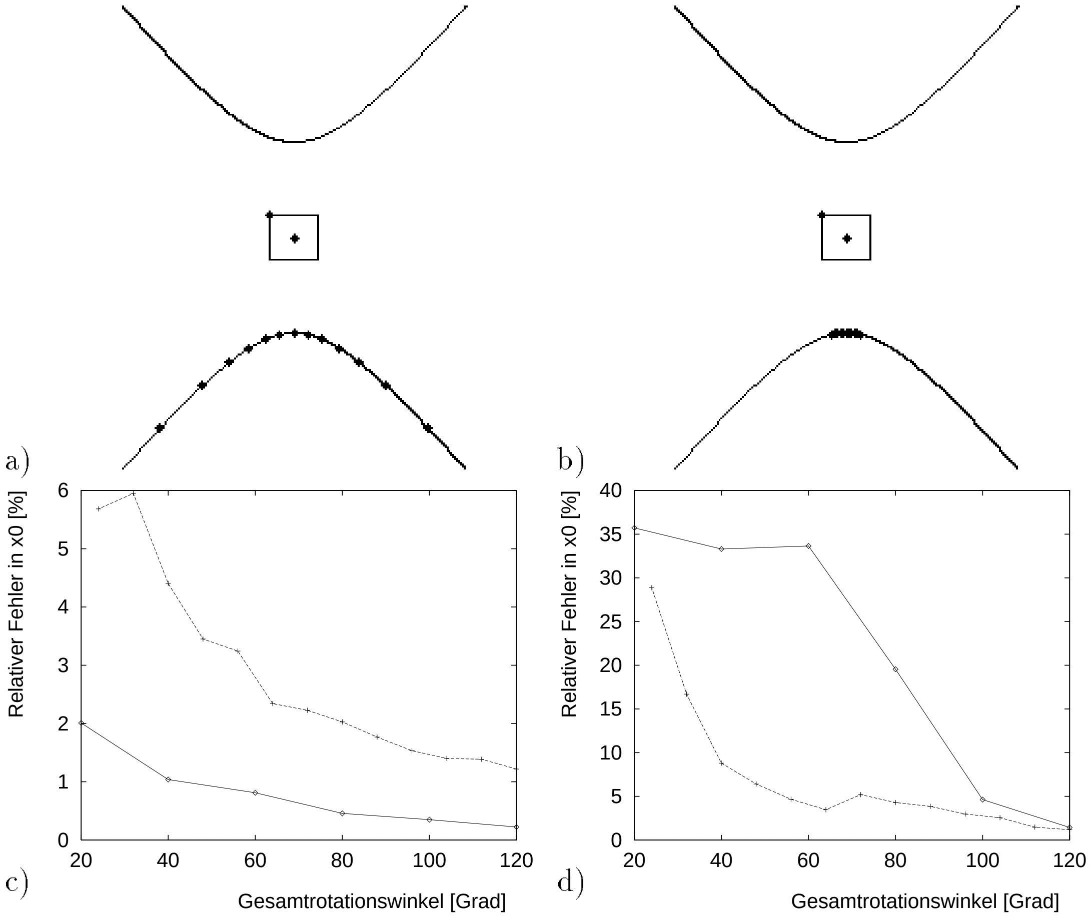


Abbildung 5.5: Fluchtpunkte und angepaßte Hyperbel für einen Gesamtrotationsschwenk von a) 120 und b) 20 Grad. Das Rechteck stellt die Sensorfläche der Kamera mit dem Ursprung in der linken oberen Ecke dar. Der Punkt innerhalb des Rechtecks gibt den geschätzten Bildhauptpunkt an. Der relative Fehler in x_0 ist in Abbildung c) und in s_x in Abbildung d) als Funktion des Gesamtschwenkwinkels der Rotation dargestellt.

Wie zu erwarten, sinkt der relative Fehler mit wachsender Gesamtrotation. Der Grund dafür ist bei der Hyperbelmethode, daß der Bereich der Fluchtpunkte, an denen die Hyperbel angepaßt wird, ein wesentlich größerer ist, wenn ein größerer Kameraschwenk mit gleichbleibender Anzahl an Rotationen durchgeführt wird. Bei der Rotationsmethode wird keine Hyperbel angepaßt, das Argument für den fallenden relativen Fehler ist jedoch dasselbe.

Des weiteren wird erneut festgestellt, daß die Hyperbelmethode für x_0 und die Rotationsmethode für s_x genauere Resultate schätzt. Außerdem entnimmt man, daß die Ergebnisse für s_x stärker durch einen kleineren Gesamtkameraschwenk beeinflußt werden.

5.1.4 Symmetrieindex

In diesem Abschnitt wird die Abhängigkeit des relativen Fehlers der auf Rotationen basierenden Kalibriermethoden von einem Symmetrieindex untersucht. Dieser Symmetrieindex ist durch die Position der Fluchtpunkte bestimmt, die zur Berechnung der Hyperbel dienen.

Für die Rotationsmethode und die Hyperbelmethode wird eine Standardabweichung des Meßfehlers in den Bilddaten von 1.0 Pixel angenommen. Für die Rotationsmethode wird des weiteren eine Standardabweichung des Meßfehlers in der Winkelbestimmung von 0.5 Grad vorausgesetzt. Bei beiden Verfahren werden 36 Geraden eines Schachbrettmusters zur Fluchtpunktbestimmung angepaßt. Es wird jeweils ein Gesamtschwenkwinkel von 80 Grad durch 16 Rotationen von jeweils 5 Grad simuliert. Das Kalibrierobjekt befindet sich in einer Entfernung von $1000mm$ vor der Kamera.

Zur Illustration des Symmetrieindexes dient Abbildung 5.6 a) und b). Bei einem Symmetrieindex von null befinden sich die Fluchtpunkte achsensymmetrisch auf einem Hyperbelzweig, wobei die Symmetrieachse die reelle Achse der Hyperbel ist. Mit steigendem Symmetrieindex werden die Fluchtpunkte auf einen Schenkel eines Hyperbelzweigs verschoben. Für den Fall, daß der Symmetrieindex von null verschieden ist, liegt keine Achsensymmetrie vor. Ein Symmetrieindex von null repräsentiert einen Kameraschwenk von -40 bis 40 Grad, ein Symmetrieindex von 45 bedeutet ein Kameraschwenk von -85 bis -5 Grad. Dabei bedeutet ein Kamerawinkel von null Grad, daß der resultierende endliche Fluchtpunkt auf der reellen Achse der Hyperbel liegt. Abbildung 5.6 a) und b) zeigen die Extremfälle der Symmetrieindizes null bzw. 45. Die Abbildungen 5.6 c) und d) illustrieren den relativen Fehler in den Parametern x_0 und s_x als Funktion des Symmetrieindexes.

Man entnimmt Abbildung 5.5 c), daß bei beiden Methoden der Fehler des internen Parameters x_0 mit abnehmender Symmetrie steigt. Dies liegt bei der Hyperbelmethode erneut daran, daß der Bildhauptpunkt durch den Schnittpunkt der Hyperbelachsen bestimmt wird. Weiter entnimmt man Abbildung 5.5 d), daß bei beiden Methoden der Fehler in dem internen Parameter s_x mit abnehmender Symmetrie fällt. Dies liegt an der Tatsache, daß die Hyperbelschenkel durch eine

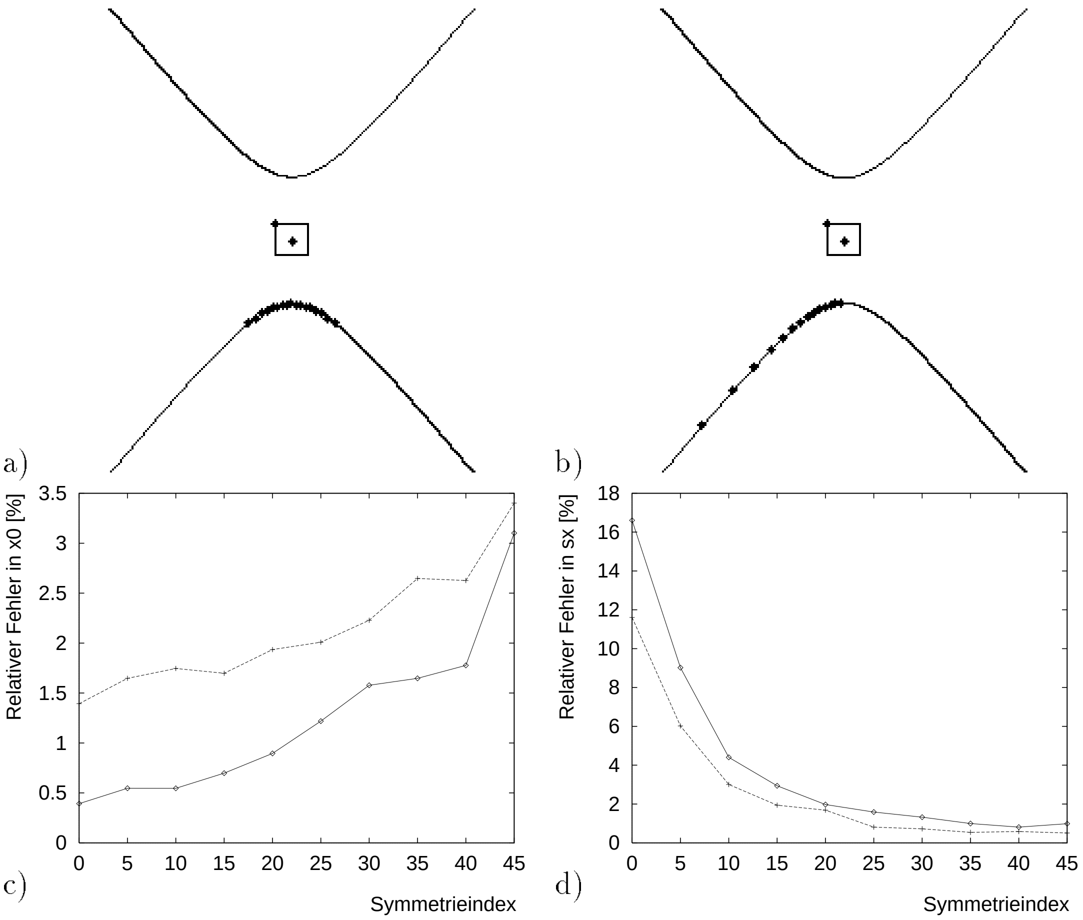


Abbildung 5.6: Fluchtpunkte und die angepaßte Hyperbel für einen Symmetrieindex von 0 a) und 45 b). Der relative Fehler in x_0 ist in Abbildung c) und in s_x in Abbildung d) als Funktion des Symmetrieindexes dargestellt. Der Symmetrieindex variiert zwischen 0 und 45, was zur Folge hat, daß sich die Fluchtpunkte auf den linken Schenkel eines Hyperbelzweiges verlagern.

größere Anzahl von Fluchtpunkten bestimmt werden.

5.1.5 Relative Orientierung der Parallelenschar

Die Abhängigkeit des relativen Fehlers der auf Rotationen basierenden Kalibriermethoden von dem Winkel zwischen dem Richtungsvektor der optischen Achse und denen der parallelen Geraden in einem globalen Weltkoordinatensystem wird in diesem Abschnitt untersucht. Der Winkel wird im folgenden als Slantwinkel bezeichnet.

Für die Hyperbel- und die Rotationsmethode wird eine Standardabweichung des Meßfehlers in den Bilddaten von 1.0 Pixel angenommen. Für die Rotationsmethode wird eine Standardabweichung des Meßfehlers in der Winkelbestimmung von 0.5 Grad vorausgesetzt. Weiterhin werden 36 Geraden eines Schachbrettmusters zur Fluchtpunktbestimmung angepaßt und das Kalibrierobjekt in einer Entfernung von $1000mm$ vor der Kamera angenommen. Insgesamt wird ein Rotationsschwenk von 120 Grad durch 12 Rotationen zu je 10 Grad simuliert.

Die Abbildungen 5.7 a) und b) illustrieren den relativen Fehler in den Parametern x_0 und s_x als Funktion des Slantwinkels.

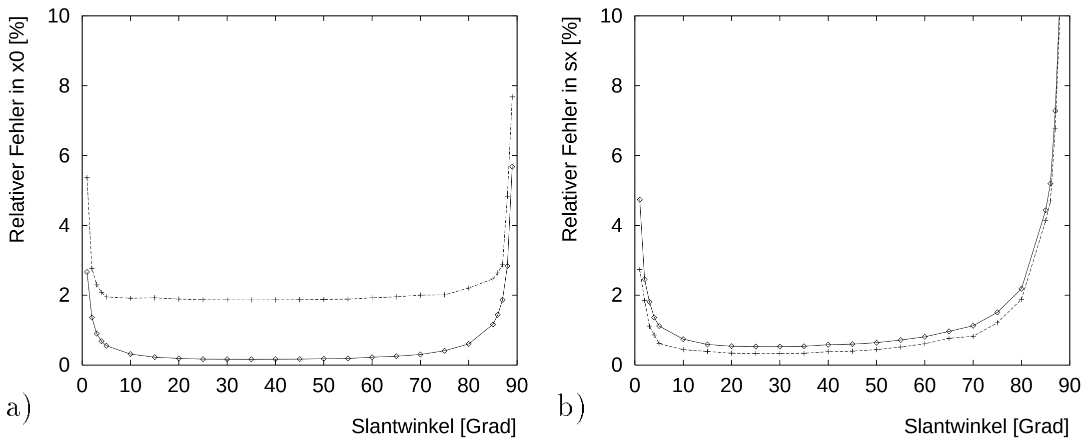


Abbildung 5.7: Der relative Fehler in x_0 ist in Abbildung a) und in s_x in Abbildung b) als Funktion des Slantwinkels dargestellt. In dem Fall, daß der Slantwinkel null beträgt, ist die optische Achse parallel zu den übrigen parallelen Geraden. Es liegt der singuläre Fall vor, daß die Hyperbel zu einer Geraden degeneriert. Beträgt der Slantwinkel 90 Grad, so steht die optische Achse orthogonal zu den parallelen Geraden. Es liegt der singuläre Fall vor, daß sich die Projektionen der Geraden im Bild nicht schneiden. In diesem Fall liegen sämtliche Fluchtpunkte auf der Geraden im Unendlichen.

In den beiden Extremfällen liegen Singularitäten vor. Ist die optischen Achse parallel zu den übrigen parallelen Geraden (null Grad), so liegen alle Fluchtpunkte auf einer Geraden. Steht die optische Achse senkrecht zu den parallelen Geraden (90 Grad), so liegen alle Fluchtpunkte im Unendlichen.

Bei beiden Methoden und bei beiden internen Parametern ergibt sich, daß der relative Fehler in der Schätzung der Parameter ein Minimum bei etwa 45 Grad besitzt. Zu beiden Seiten dieses Minimums steigt der Fehler an. In den Singularitäten können keine Werte berechnet werden.

5.1.6 Anzahl der Bildmerkmale

Hier wird die Abhängigkeit des relativen Fehlers der Kalibriermethoden, die extrahierte Bildmerkmale eines Kalibrierwürfel verwenden, von der Anzahl der produzierten Bildmerkmale untersucht. Die Methoden sind zum einen die klassische Methode und zum anderen die Zoommethode. Letztlich wirken sich die Fehler in der Zoommethode auch auf die Methode der konjugierten Fluchtpunkte aus. Daher werden auch diese Ergebnisse untersucht.

Es wird angenommen, daß sich der Kalibrierkörper in einer Entfernung von $1000mm$ vor der Kamera befindet. Bei der Zoommethode werden 10 unterschiedliche Bilder mit unterschiedlicher Brennweite simuliert. Weiterhin wird für sämtliche Methoden eine Standardabweichung des Meßfehlers in den Bilddaten von 1.0 Pixel angenommen. In Abbildung 5.8 ist der relative Fehler in x_0 in a) und in s_x in b) als Funktion der Anzahl der Bildmerkmale abgebildet.

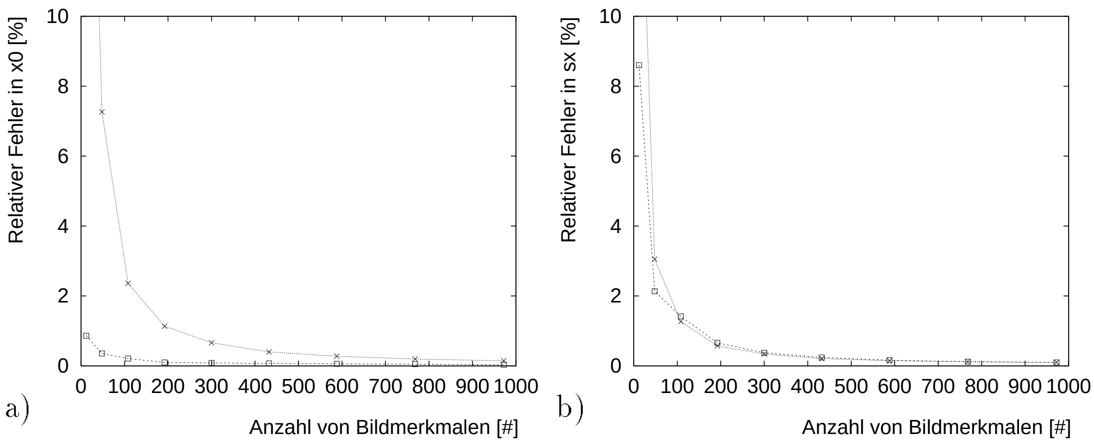


Abbildung 5.8: Der relative Fehler in x_0 ist in Abbildung a) und in s_x in Abbildung b) als Funktion der Anzahl der Bildmerkmale dargestellt.

Beide Methoden schätzen die internen Parameter genauer, wenn die Zahl der produzierten Bildmerkmale steigt. Ein kleiner Fehler in dem Bildhauptpunkt wirkt sich bei der Berechnung von s_x mit der Methode der konjugierten Fluchtpunkte stark aus. Die klassische Methode ist bei der Bestimmung von x_0 wesentlich stärker von der Anzahl der produzierten Bildmerkmale abhängig.

5.1.7 Anzahl der Zoomschritte

Hier wird die Abhängigkeit der Zoommethode von der Anzahl der Zoom-Schritte in der Bestimmung des Bildhauptpunktes untersucht. Weiterhin wird untersucht, wie sich der Fehler in der Bestimmung des Bildhauptpunktes bei der Bestimmung der Skalierungsfaktoren mit der Methode der konjugierten Fluchtpunkte fortpflanzt.

Es wird angenommen, daß sich der Kalibrierkörper in einer Entfernung von $1000mm$ vor der Kamera befindet. Es werden in jedem Bild 972 Bildmerkmale produziert und einander zugeordnet. In Abbildung 5.9 ist der relative Fehler in a) x_0 und b) s_x als Funktion der Anzahl der Zoomschritte dargestellt.

Wie erwartet verringert sich der relative Fehler in der Bestimmung der Parameter x_0 und s_x mit zunehmender Anzahl der Zoomschritte.

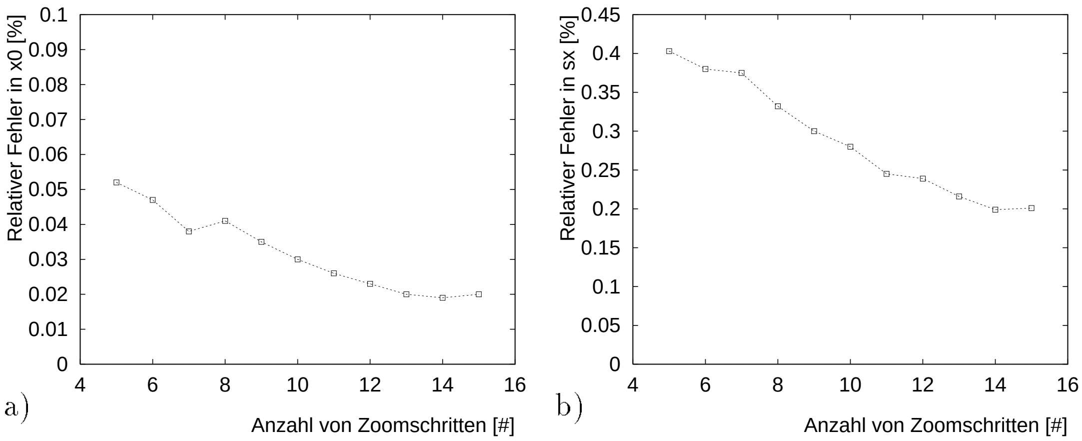


Abbildung 5.9: Der relative Fehler in x_0 ist in Abbildung a) und in s_x in Abbildung b) als Funktion der Anzahl der Zoomschritte.

5.2 Realdatenexperiment

In diesem Abschnitt werden Realexperimente mit den vier in dieser Arbeit besprochenen Verfahren zur Kalibrierung einer aktiven Kamera untersucht. Als Referenzwerte wurden x_0=256.0, y_0=226.0, s_x=820.0 und s_y=998.0 benutzt.

Zunächst wurde eine klassische Kalibrierung durchgeführt. Es wurde ein Kalibrierwürfel benutzt (Abb. 5.10). Auf diesem Würfel wurden 972 Merkmale detektiert, zu denen die Szenenmerkmale bekannt waren. Die Detektion der Bildmerkmale wurde mit der in Kapitel 4 beschriebenen Methode durchgeführt. Die anschließende Kalibrierung wurde mit der klassischen Methode nach Kapitel 3.2 durchgeführt. Als Ergebnis erhielt man für den Bildhauptpunkt x_0=256.3 und y_0=226.6. Die Skalierungsfaktoren ergaben sich zu s_x=816.7 und s_y=998.1. Dies bedeutet eine Abweichung von den Referenzwerten von 0.1% in dem Parameter x_0, 0.2% in dem Parameter y_0, 0.4% in dem Parameter s_x und 0.0% in dem Parameter s_y. Die Ergebnisse der Kalibrierung sind in Tabelle 5.2 zusammengefaßt. Für den Verzerrungsfaktor s_{xy} wurde ein Wert von null angenommen. Daher geht dieser in die Betrachtungen nicht ein.

Als nächstes wurde die kombinierte Methode benutzt, um die internen Parameter zu bestimmen. Zur Kalibrierung des Bildhauptpunktes wurde eine Sequenz von sechs Zoomschritten der Abbildung eines Kalibrierwürfels verwendet (Abb. 5.10). Auf dem Kalibrierwürfel werden in jedem Bild dieselben 81 Bildmerkmale detektiert. Den sechs Zuordnungen eines Bildmerkmals wurde jeweils eine Gerade angepaßt. Daraus ergaben sich 81 Geraden, deren gemeinsamer Schnittpunkt durch eine Ausgleichsrechnung nach Kapitel 3.4.1 ermittelt wurde. Als Ergebnis für den Bildhauptpunkt erhielt man x_0=257.3 und y_0=229.4. Dies ist eine Abweichung von 0.5% bzw. 1.5%.

Zur Bestimmung der Skalierungsfaktoren wurden in der letzten Aufnahme der Zoomschritte (Abb. 5.10b) auf dem Kalibrierwürfel drei Geradenbüschel von

Methode	Bildhauptpunkt		Skalierungsfaktoren	
	x_0	y_0	s_x	s_y
Referenzwerte	256.0	226.0	820.0	998.0
klassische Methode	256.3 (0.1%)	226.6 (0.2%)	816.7 (0.4%)	998.1 (0.0%)
Zoommethode	257.3 (0.5%)	229.4 (1.5%)	-	-
konj. Flpktmethode	-	-	820.6 (0.1%)	988.9 (0.9%)
Rotationsmethode	245.2 (4.2%)	275.6 (7.6%)	660.9 (19.4%)	510.3 (37.8%)
Hyperbelmethode	268.8 (5.0%)	206.1 (9.8%)	842.1 (2.7%)	945.1 (5.3%)

Tabelle 5.2: Ergebnisse von Realexperimenten zur Kalibrierung einer aktiven Kamera und deren Vergleich zu Referenzwerten. Die in Klammern angegebenen Werte sind die relativen Abweichungen der Ergebnisse zu den Referenzwerten.

jeweils 18 Geraden detektiert, deren Urbilder parallel waren und für die sich je zwei Geraden unterschiedlicher Geradenbüschel orthogonal zueinander verhalten. Anschließend wurden die Skalierungsfaktoren mit dem in Kapitel 3.3 beschriebenen Verfahren zu s_x=820.6 und s_y=988.9 berechnet. Dies entspricht einer Abweichung von den Referenzwerten von 0.1% bzw. 0.9%.

Die Bestimmung der internen Parameter mit der Rotationsmethode nach Kapitel 3.4.3 und mit der Hyperbelmethode nach Kapitel 3.4.4 erfolgte jeweils in zwei getrennten Schritten. Zur Berechnung der internen Parameter x_0 und s_x wurde eine Sequenz von Aufnahmen verwendet, die eine Vergenzrotation um die y-Achse des Kamerakoordinatensystems beschreiben. Abbildung 5.11 a) zeigt die erste, Abbildung 5.11 b) die letzte Aufnahme einer solchen Vergenzrotation. In Abbildung 5.11 c) sind sämtliche Fluchtpunkte und die angepaßte Hyperbel abgebildet. Das Rechteck entspricht dem Bildausschnitt der Kamera, der Punkt in der Mitte dieses Rechtecks dem ermittelten Bildhauptpunkt. Die Berechnung der internen Parameter y_0 und s_y erfolgte durch eine Sequenz von Aufnahmen, die eine Neigungsrotation um die x-Achse des Kamerakoordinatensystems beschreiben. Ein solche Neigungsrotation ist in Abbildung 5.12 dargestellt. Dabei entspricht Abbildung 5.12 a) der Aufnahme vor dem ersten, Abbildung 5.12 b) der Aufnahme vor dem letzten Rotationsschritt. In beiden Fällen wurden 25 Aufnahmen durchgeführt. Zwischen jeweils zwei Aufnahmen wurde ein Kameraschwenk von etwa 5 Grad durchgeführt, so daß die gesamte Rotation 120 Grad betrug.

Mit der Rotationsmethode erhielt man aus der Vergenzrotation die Ergebnisse x_0=245.2 und s_x=660.9, was einer Abweichung von 4.2% bzw. 19.4% entspricht. Mit der Hyperbelmethode erhält man x_0=268.8 und s_x=842.1, was einer Abweichung von 5.0% bzw. 2.7% entspricht.

Aus der Neigungsrotation erhielt man mit der Rotationsmethode die Ergebnisse y_0=275.6 und s_y=510.3 mit den Abweichungen 7.6% und 37.8%. Die Hyperbelmethode ergab in diesem Fall y_0=206.1 und s_y=945.1, was einem relativen

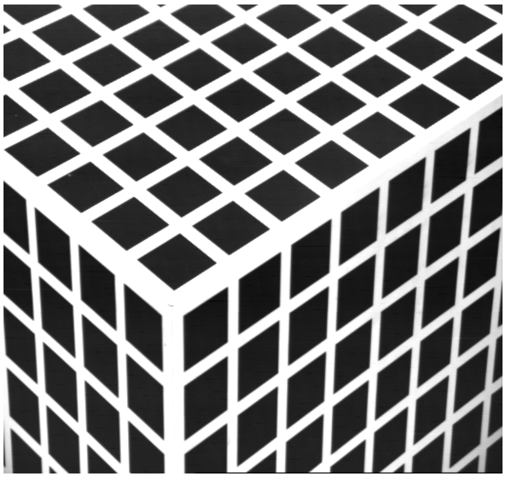
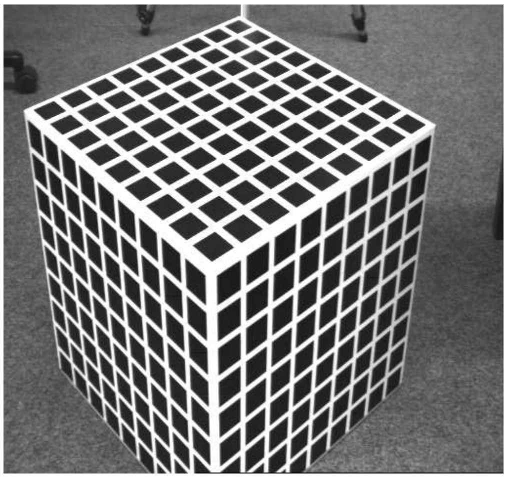
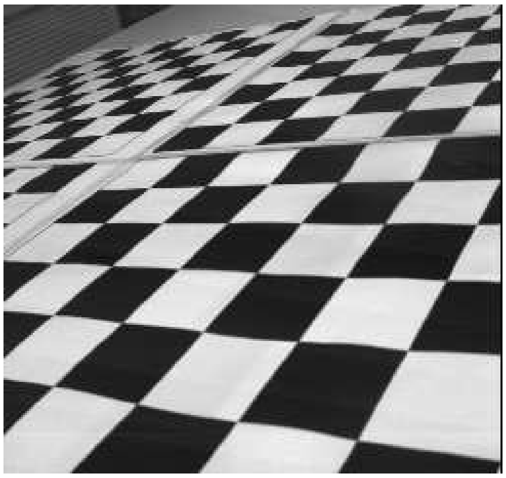
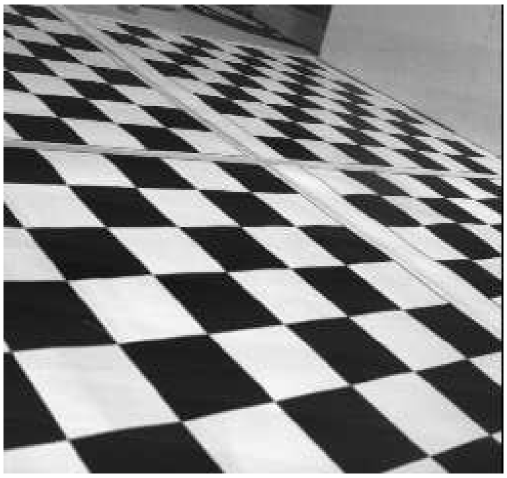
a)  b) 

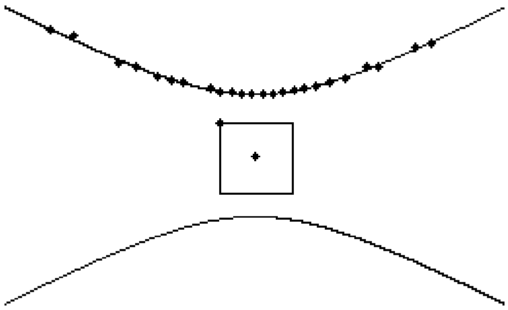
Abbildung 5.10: Beschreibung der Kalibrierung des Bildhauptpunktes durch Zoomen mit einem Kalibrierwürfel. In Abbildung a) ist der erste Zoomschritt und in Abbildung b) der letzte Zoomschritt dargestellt. Die Skalierungsfaktoren werden anschließend durch Bestimmen der konjugierten Fluchtpunkte in der Aufnahme b) kalibriert. Aufnahme b) kann des weiteren herangezogen werden, um die internen Parameter anhand der klassischen Methode zu kalibrieren.

Fehler von 9.8% bzw. 5.3% entspricht.

Aus Tabelle 5.2 lassen sich die Resultate ablesen, die aufgrund der Simulationen zu erwarten waren. Das beste Ergebnis lieferte die klassische Methode, bei der die Voraussetzungen an die Umwelt am größten waren. Da für die kombinierte Methode ein Kalibrierwürfel benutzt wurde, waren ähnlich gute Ergebnisse zu erwarten. Des weiteren zeigt sich, daß die Rotationsmethode schlechtere Ergebnisse liefert als die Hyperbelmethode, was dadurch zu begründen ist, daß bei der Hyperbelmethode eine statistische Betrachtung stattfindet, bei der Rotationsmethode hingegen nicht. Beide Methoden sind aufgrund des geringen Vorwissens, das an die Umgebung gestellt wird, in den Ergebnissen schlechter als die klassische und die kombinierte Methode.

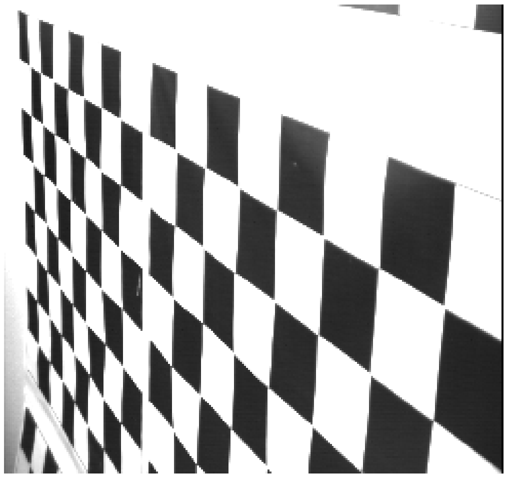
Verglichen mit den Ergebnissen anderer Arbeiten, wie der von Li [Li94] oder der von Beardsley et al. [BMZ92], bewältigt die Hyperbelmethode die Kalibrieraufgabe zufriedenstellend, da weniger Informationen an die Umwelt vorausgesetzt werden. Beardsley et al. [BMZ92] führen beispielsweise eine Rotation von 360 Grad durch und lassen nicht die Kamera sondern ein Kalibrierobjekt rotieren. Die Einordnung der Ergebnisse dieser Arbeit, die durch Daniilidis und den Autor [DE96a] und [DE96b] publiziert wurden, ist unter Beachtung des Vorwissens in Kapitel 3.6 nachzulesen.

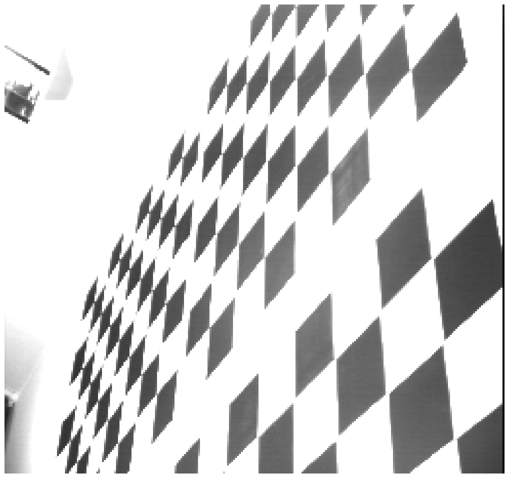
a)  b) 



c)

Abbildung 5.11: Beschreibung einer Vergenzrotation um die y-Achse des Kamerakoordinatensystems. Die Bilder a) und b) beschreiben die Aufnahme vor der ersten und nach der letzten Rotation. In Bild a) ist die verwendete Schar an parallelen Geraden durch den Fluchtpunkt in der rechten oberen Ecke, in Bild b) durch den Fluchtpunkt in der linken oberen Ecke gegeben. In Abbildung c) ist die an die Fluchtpunkte aller Rotationen angepaßte Hyperbel abgebildet. Das Rechteck in Abbildung c) entspricht dem Bildausschnitt der Kamera, mit dem ermittelten Bildhauptpunkt in der Mitte. Die Kalibrierung mit der Rotationsmethode kann ebenfalls mit der zugrundeliegenden Bildfolge geschehen.

a) 

b) 

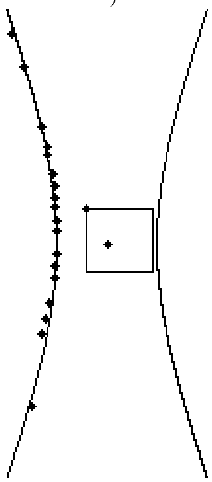

c)

Abbildung 5.12: Beschreibung einer Neigungsrotation um die x-Achse des Kamerakoordinatensystems. Die Bilder a) und b) sind die Abbildungen vor der ersten und nach der letzten Rotation. Die verwendete Schar an parallelen Geraden ist in Bild a) die Schar von Geraden, deren zugehöriger Fluchtpunkt in der linken oberen Ecke zu finden ist. Entsprechend ist die verwendete Geradenschar in Bild b) durch den linken unteren Fluchtpunkt gegeben. Die Abbildung c) zeigt die Fluchtpunkte sämtlicher Rotationen und die daran angepaßte Hyperbel. Wiederum ist der Bildausschnitt der Kamera durch ein Rechteck mit ermitteltem Bildhauptpunkt in der Mitte abgebildet. Die zugrundeliegende Bildfolge wird ebenfalls zur Kalibrierung mit der Rotationsmethode verwendet.

Kapitel 6

Diskussion und Ausblick

Zum Abschluß dieser Arbeit werden die erzielten Ergebnisse zusammengefaßt und bewertet. In einem darauffolgenden Ausblick werden Erweiterungsmöglichkeiten angegeben, die Gegenstand zukünftiger Arbeiten sein können.

6.1 Diskussion der vorgestellten Methoden

In dieser Arbeit wurden verschiedene Methoden zur Kalibrierung der internen Parameter einer Kamera vorgestellt. Der Beitrag dieser Arbeit liegt dabei in

1. der Darstellung und dem systematischen Vergleich von vier existierenden Methoden zur Berechnung der internen Parameter einer Kamera, welche neu implementiert wurden. Dies sind:
   - die klassische Methode nach Kapitel 3.2.
   - die Zoommethode nach Kapitel 3.4.1.
   - die Methode der konjugierten Fluchtpunkte nach Kapitel 3.3.
   - die Rotationsmethode nach Kapitel 3.4.3.
2. in der Entwicklung und Implementierung einer neuen Methode zur Berechnung der internen Kameraparameter, die ein gutes Verhältnis zwischen a-priori-Wissen und Genauigkeit aufweist. Diese Methode ist
   - die Hyperbelmethode nach Kapitel 3.4.4.
3. in der eingehenden Untersuchung des Verhaltens sowohl der neuen als auch der existierenden Methoden in bezug auf den Einfluß der Aufnahmegeometrie.

Es wurden neben den erwähnten Methoden auch Methoden anderer Autoren besprochen (Kap. 3.4.5). Die meisten Methoden anderer Autoren sind mit einer der vorgestellten Methoden in eine Gruppe einzuordnen. Diese Klassifizierung ist in Kapitel 3.5 dargestellt. Es werden hier kurz die Erfahrungen mit den existierenden Methoden geschildert.

Zunächst soll hier auf die klassische Kalibriermethode eingegangen werden. Die klassische Kalibriermethode wurde durch zahlreiche Autoren beschrieben. Sie ist in der Lage, sämtliche interne Parameter auf der Basis einer Aufnahme eines Kalibrierkörpers zu bestimmen. Dazu müssen Bildmerkmale des Kalibrierkörpers in der Aufnahme detektiert werden.

Die internen Parameter sind die Koordinaten des Bildhauptpunktes und die Skalierungsfaktoren, die sich aus der Brennweite und der Größe der Sensorzellen auf der Aufnahmefläche ergeben. Weiterhin kann durch diese Methode ein Verzerrungsfaktor bestimmt werden. Dieser Verzerrungsfaktor ist insbesondere dann von Interesse, wenn Weitwinkel- oder Fischaugekameras verwendet werden. Für die Kameras, die in dieser Arbeit verwendet wurden, spielte der Verzerrungsfaktor keine Rolle.

In der Simulation mit synthetischen Daten und bei Realexperimenten hat sich gezeigt, daß die klassische Methode hauptsächlich von der Anzahl der detektierten Bildmerkmale auf einem Kalibrierkörper und von der Entfernung des Kalibrierkörpers von der Kamera abhängig ist.

Im Hinblick auf die Kalibrierung von aktiven Kameras ist die Verwendung dieser Kalibriermethode nicht von großem Interesse, da ein Kalibrierkörper benutzt werden muß. Ein solcher Kalibrierkörper ist bei einem Einsatz von aktiven Kameras im allgemeinen nicht gegeben. Aufgrund der großen Menge an Vorkenntnissen, die über einen solchen Kalibrierkörper gegeben sein müssen, ist eine genaue Kalibrierung zu erwarten. Durch Testreihen mit synthetischen Daten und Realexperimente wurde nachgewiesen, daß diese Methode vergleichbar genaue Ergebnisse für die internen Parameter liefert.

Als weitere Kalibriermethode wurde die Zoommethode behandelt. Mit der Zoommethode lassen sich nur die internen Parameter des Bildhauptpunktes ermitteln. Es muß mit der verwendeten Kamera möglich sein, die Brennweite zu verändern. In den Aufnahmen mit unterschiedlichen Brennweiten haben Bildmerkmale einander zugeordnet zu werden. In dieser Arbeit wurde zur Zuordung ein Kalibrierwürfel benutzt. Steht ein solcher nicht zur Verfügung, so können beispielsweise Ecken einer realen Umgebung einander zugeordnet werden. Diese Aufgabe stellt an sich schon ein nicht triviales Problem dar, wodurch der Einsatz dieser Methode in einer realen Umgebung problematisch wird.

Eine Kamera, die in einer realen Umgebung eingesetzt werden soll, wird den Freiheitsgrad der Brennweite im allgemeinen nicht sehr häufig nutzen. Denkbar ist in diesem Fall nur, daß ein Programm, das eine Kamera steuert, seine Aufmerksamkeit auf ein bestimmtes Objekt, wie beispielsweise ein Verkehrschild, lenkt, um es mit einer höheren Brennweite zu untersuchen. Konventionelle Kameras, die mit dem Freiheitsgrad zur Veränderung der Brennweite ausgerüstet sind, können die Funktion der Veränderung der Brennweite nur relativ langsam ausführen. Daher wird ein aktives System, welches in Echtzeit arbeiten soll, eher auf das Zoomen verzichten.

Es ist mit dieser Methode nur möglich, die internen Parameter des Bildhauptpunktes zu kalibrieren. Dies spricht gegen den Einsatz dieser Methode in einem aktiven System.

Die Ergebnisse, die mit dieser Methode erzielt wurden, waren im Realexperiment vergleichbar gut. Dies lag hauptsächlich an der großen Menge an Vorabinformationen, die in diese Methode investiert wurde. Andere Autoren haben diese Methode ebenfalls mit Kalibrierkörpern getestet und sind ebenfalls zu guten Ergebnissen gelangt.

In der Simulation mit synthetischen Daten hat sich gezeigt, daß die Zoommethode insbesondere von der Anzahl der detektierten Bildmerkmale und von der Anzahl der Zoomschritte abhängig ist. Des weiteren fließt auch die Entfernung des Kalibrierkörpers in die Güte der Ergebnisse ein.

Eine weitere Methode, die in dieser Arbeit besprochen wurde, ist die Methode der konjugierten Fluchtpunkte. Diese Methode ist in der Lage, aus drei konjugierten Fluchtpunkten drei interne Parameter zu berechnen. Da in jedem Fall vier interne Parameter vorhanden sind, ist somit mindestens ein Parameter als bekannt vorauszusetzen, um die anderen bestimmen zu können. Sind nur zwei konjugierte Fluchtpunkte bekannt, so können höchstens zwei interne Parameter einer Kamera berechnet werden.

In dieser Arbeit wurde zur Bestimmung der konjugierten Fluchtpunkte ein Kalibrierkörper benutzt. Die internen Parameter des Bildhauptpunktes wurden in einem ersten Schritt durch die Zoommethode ermittelt. Die Bestimmung der Skalierungsfaktoren wurde anschließend mit der Methode der konjugierten Fluchtpunkte vollzogen. Es gibt in der vom Menschen geschaffenen Umwelt eine Reihe von Geraden, die orthogonal zueinander stehen. Die Aufgabe, diese Geraden in einer Aufnahme zu ermitteln und zu gruppieren, ist allerdings nicht trivial. Die Annahmen, die dazu gemacht werden müssen, schränken den Einsatz dieser Methode stark ein. Neben der bedingten Einsatzfähigkeit durch restriktive Annahmen stellt auch die partielle Berechnung der internen Parameter eine Einschränkung dar. Somit erscheint auch dieser Ansatz für den Einsatz in einem aktiven System nicht als besonders geeignet.

Der Schwerpunkt dieser Arbeit lag in der Entwicklung einer Kalibriermethode für aktive Kameras. Dazu wurden zwei Kalibriermethoden untersucht, die die Freiheitsgrade der Kamera durch Rotationen um feste Achsen ausnutzen. Beide Verfahren nutzen Fluchtpunkte. Es genügt somit, während der Rotation einer Kamera über die abgebildete Umgebung nur zu wissen, daß im Urbild parallele Geraden vorhanden sind. Dies ist eine wesentlich kleinere Menge an Vorwissen, das zur Kalibrierung mit diesen Methoden vorausgesetzt wird. Im Urbild parallele Geraden sind in einer vom Menschen geschaffenen Umgebung sehr häufig vorzufinden. Als Beispiel lassen sich Hausfassaden, Straßenverläufe und Gänge, Korridore oder Räume in Gebäuden anführen.

Die zwei Methoden, die Rotationen einer aktiven Kamera zur Kalibrierung der internen Parameter ausnutzen, sind zum einen die Rotationsmethode und zum anderen die Hyperbelmethode. Der hauptsächliche Unterschied in den Voraussetzungen zwischen diesen beiden Methoden liegt darin, daß für die Rotationsmethode zusätzlich die Rotationswinkel zwischen den einzelnen Aufnahmen bekannt sein müssen.

Die Rotationsmethode ist in Varianten durch zahlreiche Autoren veröffentlicht worden. Die Hyperbelmethode wurde in dieser Arbeit neu entwickelt. Beide Methoden nutzen Fluchtpunkte zur Kalibrierung der internen Parameter. Da Fluchtpunkte translationsinvariant sind, kann auch während der Translation der Kamera eine Rotation mit integrierter Kalibrierung vorgenommen werden. Des weiteren ist die Detektion von Geraden stabiler als die Detektion von Bildpunkten.

In der Simulation wurde für beide Verfahren nachgewiesen, daß folgende Faktoren die Güte der Ergebnisse beeinflussen:

- die Größe des Gesamtschwenks der Rotation und
- die Symmetrie der Fluchtpunkte entlang eines Hyperbelzweigs
- die relative Orientierung der Parallelenschar und
- die Entfernung des Kalibrierobjekts.

Die ersten beiden Faktoren sind dabei die kritischsten. Weiterhin hat sich durch die Simulation mit synthetischen Daten und bei Realexperimenten gezeigt, daß die Hyperbelmethode trotz weniger Voraussetzungen an die Umwelt (die Rotationswinkel werden nicht als bekannt vorausgesetzt) gleichwertige bis bessere Ergebnisse für die Kalibrierung der internen Parameter liefert. Insbesondere bei den Realexperimenten hat die Hyperbelmethode wesentlich bessere Ergebnisse geliefert. Obwohl beide auf Rotationen basierende Methoden aufgrund der kleinen Menge an Vorwissen geeignet sind, in einem aktiven System zum Einsatz zu kommen, ist die Hyperbelmethode aus Stabilitätsgründen der Rotationsmethode vorzuziehen.

Für den Einsatz in einem aktiven System bietet die Hyperbelmethode folgende Vorteile:

- Die Anforderungen an die Umgebung sind gering. Es wird lediglich eine Schar von parallelen Geraden in der Umgebung vorausgesetzt. Dadurch wird eine ständige Rekalibrierung möglich.
- Die Anforderungen an die Hardware der Kamera sind gering. Es wird lediglich verlangt, daß die Kamera in der Lage ist, eine Vergenz- und eine Neigungsrotation durchzuführen.
- Die Rotationswinkel werden nicht als bekannt vorausgesetzt.

- Es werden sämtliche interne Kameraparameter kalibriert.
- Aufgrund der Verwendung von Fluchtpunkten können die internen Kameraparameter während einer Translation berechnet werden. Daher braucht ein aktives System zur Kalibrierung nicht zu stoppen.
- Die internen Parameter werden entkoppelt von den externen kalibriert.

6.2 Ausblick auf zukünftige Erweiterungen

Ist ein autonomes Robotersystem mit einer aktiven Kamera ausgestattet, so ist diese Bewegungen ausgesetzt. Da durch diese Bewegungen auch die internen Parameter der Kamera verändert werden können, ist eine ständige Rekalibrierung notwendig. Die vorgestellte Hyperbelmethode besitzt in einer vom Menschen geschaffenen Umwelt die Möglichkeit, diese Aufgabe wahrzunehmen.

Um die Hyperbelmethode sinnvoll zu nutzen, muß vor allem eine stabile Detektion der im Urbild parallelen Geraden unter Realweltbedingungen gegeben sein. Die Detektion von Geraden ist mit der Hough-Transformation lösbar. Die Gruppierung der Geraden ist eine nicht triviale Aufgabe, zu der bisher keine robusten Verfahren zur Verfügung stehen.

Im Hinblick auf den aktiven Einsatz der Hyperbelmethode zur ständigen Rekalibrierung der internen Kameraparameter sollte außerdem eine optimale Position der Kamera zu einer Schar von parallelen Geraden bestimmt werden, so daß die zu erwartenden Ergebnisse weiter optimiert werden.

Weiterhin bedarf ein aktives Kamerasystem einer Kalibrierung der externen Parameter. Diese Hand-Auge- oder Sensor-Effektor-Kalibrierung wurde von zahlreichen Autoren beschrieben und implementiert. Von besonderem Interesse ist in diesem Zusammenhang eine Sensor-Effektor-Kalibrierung, die ähnliche Anforderungen an den Kalibriervorgang stellen, wie die Hyperbelmethode.

Schließlich ist es von Interesse, diese beiden Methoden miteinander zu koppeln. Besitzt das aktive Kamerasystem den Freiheitsgrad einer Zoomfunktion, so muß auch diese kalibriert werden. Auch eine Kombination mit einer solchen Kalibrierung sollte untersucht werden.

Anhang A

Beschreibung der implementierten Kalibrieralgorithmen

Es wurden fünf verschiedene Algorithmen zur Berechnung der internen Kameraparameter implementiert:

- **classic** - klassische Kalibriermethode nach Kapitel 3.2
- **zoom** - Zoom-Methode nach Kapitel 3.4.1
- **vp** - die Methode der konjugierten Fluchtpunkte nach Kapitel 3.3
- **rotate** - Rotationsmethode nach Kapitel 3.4.3
- **hyperbel** - Hyperbelmethode nach Kapitel 3.4.4

Somit wurde aus fünf Gruppen der in Kapitel 3.5 klassifizierten Verfahren jeweils ein Verfahren implementiert. Im folgenden werden die einzelnen Algorithmen beschrieben. Dazu werden zunächst die Eingabeparameter, anschließend der implementierte Algorithmus und letztlich die Ausgabeparameter beschrieben.

A.1 Die klassische Kalibriermethode

classic[1] erwartet als Eingabe eine endliche Zahl von Fünftupeln gemäß (3.6). Diese können in Form einer Datei übergeben werden:

```
classic < Datei
```

Die Datei besitzt das Format:

[1] Dieses Programm lag in einer Version von Dr. Daniilidis vor, die nur geringfügig modifiziert wurde.

N
x_{s1} y_{s1} z_{s1} x_{p_1} y_{p_1}
$\vdots$
x_{sN} y_{sN} z_{sN} x_{p_N} y_{p_N}

Dabei ist N die Anzahl der Zuordnungen von Szenenkoordinaten zu Pixelkoordinaten. Jede Zeile repräsentiert jeweils eine Zuordnung. Es werden mindestens sechs Zuordnungen von Koordinatentupeln benötigt.

Der implementierte Algorithmus realisiert das in Kapitel 3.2 beschriebene klassische Kalibrierverfahren. Als Ausgabe erhält man die internen Parameter des Bildhauptpunktes x_0 und y_0, die Skalierungsfaktoren s_x und s_y und den Skalierungsfaktor s_{xy}. Die Ausgabeparameter besitzen das Format:

x_0 y_0 s_x s_y s_{xy}

A.2 Die Zoom-Methode

zoom erwartet als Eingabeparameter zwei oder mehr Dateien:

```
zoom Datei1 Datei2 [Datei3 ...]
```

Jede Datei besitzt das Format:

N
x_{p_1} y_{p_1}
$\vdots$
x_{p_N} y_{p_N}

$N \geq 2$ gibt die Anzahl der detektierten Bild- oder Fluchtpunkte an, die in allen Dateien übereinstimmen muß. Des weiteren müssen dieselben Zeilen in unterschiedlichen Dateien dasselbe Bildmerkmal kennzeichnen. Der Algorithmus realisiert die in Kapitel 3.4.1 beschriebenen Zoommethode zur Bestimmung des Bildhauptpunktes. Als Ausgabe erhält man die internen Parameter des Bildhauptpunktes in der Form:

x_0 y_0

A.3 Die Methode der konjugierten Fluchtpunkte

vp erwartet als Eingabe zwei oder drei homogene Fluchtpunktkoordinaten von konjugierten Fluchtpunkten und mindestens den Wert eines internen Parameters:

```
vp [-x0 Wert][-y0 Wert][-sx Wert][-sy Wert] Datei
```

Die Datei besitzt das Format:

$$\begin{array}{lll} N & & \\ x_{p1} & y_{p1} & z_{p1} \\ \vdots & & \\ x_{pN} & y_{pN} & z_{pN} \end{array}$$

$N \in \{2,3\}$ gibt die Anzahl der detektierten konjugierten Fluchtpunkte in homogenen Rechnerkoordinaten an. Wird nur der Wert eines internen Parameters übergeben, so erwartet das Programm drei homogene Fluchtpunktvektoren. Werden nur zwei konjugierte Fluchtpunkte übergeben, so erwartet das Programm mindestens die Werte zweier interner Parameter. Liegen zwei oder drei Fluchtpunkte im Unendlichen, so werden die Kameraparameter des Bildhauptpunktes sowie ein bekannter Skalierungsfaktor benötigt. Der Algorithmus ist eine Implementierung der in Kapitel 3.3 beschriebenen Methode. Als Ausgabe erhält man sämtliche internen Parameter. Die Ausgabeparameter besitzen die Form:

$$x_0 \quad y_0 \quad s_x \quad s_y$$

A.4 Die Rotationsmethde

rotate erwartet als Eingabe einen Modus und ein oder zwei Dateien:

```
rotate Modus Datei1 [Datei2]
```

Es gibt die folgenden drei Modi:

- `Modus`=1 - alleinige Vergenzrotation,
- `Modus`=2 - alleinige Neigungsrotation,
- `Modus`=3 - Vergenz- und Neigungsrotation.

In den Modi 1 und 2 wird nur eine Datei entsprechend einer Vergenz- oder Neigungsrotation vorausgesetzt. In Modus 3 werden sowohl eine Dateien für eine Vergenz, als auch eine Datei für eine Neigungsrotation erwartet. Die Datei der Vergenzrotation wird dabei als erstes verlangt. Die Dateien besitzen das Format:

$$\begin{array}{lllllll} N & & & & & & \\ \theta & x_{p1} & y_{p1} & z_{p1} & x'_{p1} & y'_{p1} & z'_{p1} \\ \vdots & & & & & & \\ \theta & x_{pN} & y_{pN} & z_{pN} & x'_{pN} & y'_{pN} & z'_{pN} \end{array}$$

$N \geq 2$ gibt die Anzahl der detektierten, zugeordneten und rotierten Bildpunkte oder Fluchtpunkte in Rechnerkoordinaten an. Es können Bild- oder Fluchtpunkte von unterschiedlichen Merkmalen benutzt werden. Der Rotationswinkel zwischen dem ungestrichenen Vektor vor und dem gestrichenen Vektor nach einer Rotation wird mit θ bezeichnet.

In Modus 1 werden die internen Parameter x_0 und s_x nach Kapitel 3.4.3.1 berechnet. Die Ausgabeparameter besitzen das Format:

$x_0 \quad 0 \quad s_x \quad 0$

In Modus 2 werden die internen Parameter y_0 und s_y nach Kapitel 3.4.3.2 berechnet. Die Ausgabeparameter besitzen die Form:

$0 \quad y_0 \quad 0 \quad s_y$

In Modus 3 werden sämtliche internen Kameraparameter nach Kapitel 3.4.3.3 berechnet. Die Ausgabeparameter besitzen das Format:

$x_0 \quad y_0 \quad s_x \quad s_y$

In allen drei Modi findet zunächst eine lineare Berechnung der internen Parameter statt. Anschließend werden die Parameter durch ein iteratives nichtlineares Verfahren verbessert.

A.5 Die Hyperbelmethode

hyperbel erwartet als Eingabe einen Modus und mindestens eine Datei:

```
hyperbel Modus Datei1 [Datei1.1 ...] [Datei2 [Datei2.1 ...]]
```

Es gibt die drei Modi:

- `Modus`=1 - alleinige Vergenzrotation,
- `Modus`=2 - alleinige Neigungsrotation,
- `Modus`=3 - Vergenz- und Neigungsrotation.

In den Modi 1 und 2 wird nur eine Datei, entsprechend einer Vergenz- oder Neigungsrotation, erwartet. In Modus 3 werden sowohl $M \geq 1$ Dateien für Vergenz als auch M Dateien für Neigungsrotationen vorausgesetzt. Die Dateien der Vergenzrotationen werden in diesem Fall als erstes verlangt. In einer Datei werden jeweils die Daten eines rotierten Bild- und Fluchtpunktes benötigt. Jede Datei besitzt das Format:

$$\begin{array}{lll} N & & \\ x_{p_1} & y_{p_1} & z_{p_1} \\ \vdots & & \\ x_{p_N} & y_{p_N} & z_{p_N} \end{array}$$

$N \geq 4$ gibt die Anzahl der detektierten, zugeordneten und rotierten Bildpunkte oder Fluchtpunkte in Rechnerkoordinaten an. Sämtliche Bild- oder Fluchtpunkte einer Datei müssen dasselbe Merkmal repräsentieren. Unterschiedliche Merkmale werden in verschiedenen Dateien übergeben.

In Modus 1 werden die internen Parameter x_0, y_0 und s_x nach Kapitel 3.4.4.1 berechnet. Die Ausgabeparameter besitzen die Form:

$$x_0 \quad y_0 \quad s_x \quad 0$$

In Modus 2 werden die internen Parameter x_0, y_0 und s_y nach Kapitel 3.4.4.2 berechnet. Die Ausgabeparameter besitzen das Format:

$$x_0 \quad y_0 \quad 0 \quad s_y$$

In Modus 3 werden sämtliche internen Kameraparameter nach Kapitel 3.4.4.3 berechnet. Die Ausgabeparameter besitzen die Form:

$$x_0 \quad y_0 \quad s_x \quad s_y$$

In allen drei Modi wird zunächst für jede Datei eine Hyperbel nach Kapitel 4.2.4 berechnet. In den Modi 1 und 2 werden daraus jeweils interne Parameter berechnet. Im Modus 3 werden jeweils aus einem Dateienpaar, bestehend aus einer Vergenz- und einer Neigungsrotation, die internen Parameter berechnet. Weiterhin findet in allen drei Modi eine lineare Berechnung mit anschließender nichtlinearer iterativer Verbesserung statt. Zuletzt wird über eine Menge von internen Parametern eine Ausgleichsrechnung durchgeführt. Die Gewichtung der internen Parameter in dieser Ausgleichsrechnung erfolgt über die Varianzen der einzelnen Parameter, die jeweils parallel zu den internen Parametern berechnet werden (Kap. 4.2.4).

Anhang B

Beschreibung der implementierten Merkmalsalgorithmen

Die gesamte Bildverarbeitung wurde unter der Software-Entwicklungsumgebung *Khoros*[1] implementiert. Die Vorverarbeitung der Kamerabilder wird mit bereits implementierten Verfahren aus der *Khoros*-Umgebung vorgenommen. Diese Vorverarbeitung besteht insbesondere in der Bearbeitung der Grauwertbilder der Kameras mit einem Kantendetektor, so daß als Ausgabeparameter ein Binärbild entsteht. Kantendetektoren sind in der Computer Vision hinreichend bekannt, so daß hier auf Jähne [Jäh93] verwiesen wird. Außerdem wurden folgende Bildverarbeitungsalgorithmen unter der *Khoros*-Umgebung entwickelt und implementiert:

- **segment** - Segmentierung eines Kalibrierwürfels nach Kapitel 4.2.5.1.
- **feature** - Bildmerkmalsextraktion eines Kalibrierwürfels nach Kapitel 4.2.5.2 und deren Zuordnung zu Szenenmerkmalen nach Kapitel 4.2.6.
- **parlines** - Detektion von Geraden mit parallelen Urbildern nach Kapitel 4.2.2.

Zusätzlich wurde ein Algorithmus zur Schätzung von Fluchtpunkten implementiert.

- **comvp** - Fluchtpunktbestimmung nach Kapitel 4.2.3.

Dieser Algorithmus wurde unabhängig von der *Khoros*-Umgebung implementiert. Im folgenden werden die einzelnen Algorithmen der Bildverarbeitung und der Algorithmus zur Schätzung von Fluchtpunkten beschrieben. Dazu werden zunächst die Eingabeparameter, anschließend der implementierte Algorithmus und zuletzt die Ausgabeparameter beschrieben.

[1] *Khoros* ist ein registriertes Produkt der Firma Khoral Research Inc.

B.1 Bildpunktdetektion auf einem Kalibrierwürfel

Um Bildmerkmale auf einem Kalibrierwürfel zu detektieren, wurde ein zweistufiger Algorithmus implementiert. Die erste Stufe nimmt eine Segmentierung des Ausgangsbildes in die drei sichtbaren Seiten des Kalibrierwürfels vor, die zweite Stufe detektiert anschließend Bildmerkmale auf dem Kalibrierwürfel und ordnet sie den korrespondierenden Szenenmerkmalen zu.

B.1.1 Segmentierung

segment erwartet als Eingabeparameter ein Binärbild eines kantendetektierten Kalibrierwürfels und einige optionale Parameter:

```
segment Binärbild Optionen
```

Die Optionen sind manuell einzustellende Parameter, die angeben, mit welcher Schrittweite in den Bildern nach einer Segmentierung zu suchen ist.

Der Algorithmus generiert mit der Hough-Transformation (Kap. 4.2.1) ein Akkumulatorfeld zur Geradendetektion (Kap. 4.2.2). In diesem Feld werden vorherrschende Richtungen detektiert und so eine Segmentierung nach Kapitel 4.2.5.1 vorgenommen. Als Ausgabeparameter erhält man die Segmentierungsdaten des Eingangsbildes.

B.1.2 Merkmalsextraktion

feature erwartet als Eingabeparameter ein Binärbild eines kantendetektierten Kalibrierwürfels, die Segmentierungsdaten aus *segment* und einige optionale Parameter:

```
feature Binärbild Segmentierungsdaten Optionen
```

Die optionalen Parameter sind manuell einzustellende Parameter, die angeben, mit welcher Schrittweite in den Bildern nach Bildmerkmalen gesucht wird, und ein Parameter, der die Anzahl der zu extrahierenden Bildmerkmale bestimmt.

Zunächst wird durch die Hough-Transformation für jedes Segment ein Akkumulatorfeld zur Geradendetektion angelegt. Anschließend werden Geraden nach dem Verfahren aus Kapitel 4.2.5.2 detektiert. Dann werden Bildmerkmale nach Kapitel 4.2.6 detektiert und deren Szenenkoordinaten nach Kapitel 4.1.2 zugeordnet. Zuletzt werden drei Geradenmengen nach Kapitel 4.2.7 generiert. Als Ausgabeparameter erhält man fünf Dateien. Die erste Datei enthält Merkmalszuordnungen in der Form:

N
$x_{s1} \quad y_{s1} \quad z_{s1} \quad x_{p_1} \quad y_{p_1}$
$\vdots$
$x_{sN} \quad y_{sN} \quad z_{sN} \quad x_{p_N} \quad y_{p_N}$

Dabei ist N die Anzahl der Zuordnungen von Szenenkoordinaten zu Pixelkoordinaten. Jede Zeile repräsentiert eine Zuordnung. Weiterhin erhält man eine Datei von Bildmerkmalen in dem Format:

N
$x_{p_1} \quad y_{p_1}$
$\vdots$
$x_{p_N} \quad y_{p_N}$

Dabei ist N die Anzahl der Bildmerkmale. Des weiteren erhält man drei Dateien von Geraden `Ger0`, `Ger1` und `Ger2` in dem Format:

N
$\theta_1 \quad \rho_1$
$\vdots$
$\theta_N \quad \rho_N$

Dabei gibt N die Anzahl der Geraden an. Alle Geraden einer solchen Datei haben parallele Urbilder. Geraden aus unterschiedlichen Dateien stehen senkrecht zueinander.

B.2 Berechnung von Geradenscharen mit parallelen Urbildern

parlines erwartet als Eingabeparameter ein Binärbild eines kantendetektierten Kamerabildes und einige optionale Parameter:

```
parlines Binärbild Optionen
```

Die optionalen Parameter sind manuell einzustellende Parameter, die angeben, mit welcher Schrittweite in den Bildern nach Bildmerkmalen zu suchen ist, und ein Parameter, der die Anzahl der zu extrahierenden Geradenbüschel bestimmt.

Zunächst wird durch die Hough-Transformation ein Akkumulatorfeld zur Geradendetektion nach Kapitel 4.2.2 angelegt und darin mehrere Geradenbüschel detektiert. Als Ausgabeparameter erhält man mehrere Dateien von Geraden in dem Format:

N
$\theta_1 \quad \rho_1$
$\vdots$
$\theta_N \quad \rho_N$

Dabei gibt N die Anzahl der Geraden an. Sämtliche Geraden einer solchen Datei haben parallele Urbilder.

B.3 Fluchtpunktschätzung

comvp erwartet als Eingabeparameter mindestens eine Datei von Geraden, deren Urbilder parallel sind:

```
comvp Datei1 [Datei2 ...]
```

Jede Datei besitzt das Format:

$$\begin{array}{ll} N & \\ \theta_1 & \rho_1 \\ \vdots & \\ \theta_N & \rho_N \end{array}$$

Dabei gibt N die Anzahl der Geraden an. Der Algorithmus ist eine Implementierung der in Kapitel 4.2.3 beschriebenen Methode zur Schätzung von Fluchtpunkten. Als Ausgabeparameter erhält man eine Datei von homogenen Fluchtpunkten in der Form:

$$\begin{array}{lll} M & & \\ x_{p_1} & y_{p_1} & z_{p_1} \\ \vdots & & \\ x_{p_N} & y_{p_N} & z_{p_N} \end{array}$$

Dabei gibt M die Anzahl der generierten Fluchtpunkte an, die der Anzahl der übergebenen Dateien entspricht. Weiterhin entspricht die Reihenfolge der Fluchtpunkte der Reihenfolge der übergebenen Dateien.

Anhang C

Zusammenspiel der Merkmals- und Kalibrieralgorithmen

In diesem Abschnitt wird darauf eingegangen, wie die detektierten Referenzmerkmale der Algorithmen aus Anhang B als Eingabe der Kalibrieralgorithmen aus Anhang A verwendet werden.

- Durch *feature* wird eine Datei von Zuordnungen von Szenen- zu Merkmalskoordinaten generiert. Diese Datei kann direkt als Eingabe der klassischen Kalibrierung mit *classic* verwendet werden.

- Werden mehrere Bilder mit unterschiedlichen Zoom-Schritten aufgenommen, so können die Dateien der Bildmerkmale, die der Algorithmus *feature* zu jeder Aufnahme generiert, als Eingabe des Kalibrierverfahrens *zoom* dienen.

- Durch *feature* werden drei Dateien im Urbild paralleler Geraden erzeugt. Verwendet man diese Dateien als Eingabe des Algorithmus *comvp*, so erhält man eine Datei von drei konjugierten Fluchtpunkten. Diese Datei kann als Eingabeparameter des Kalibrieralgorithmus *vp* dienen. Als bekannte interne Parameter können die Ergebnisse der Zoommethode benutzt werden.

- Werden mehrere Bilder mit der Kamera so aufgenommen, daß die Voraussetzungen für die Kalibrierung mit der Hyperbel- oder Rotationsmethode vorliegen, kann auf jedes Bild der Algorithmus *feature* oder der Algorithmus *parlines* angewendet werden. Dann können die korrespondierenden Geradenbüschel aller Aufnahmen mit dem Algorithmus *comvp* bearbeitet werden. Man erhält daraus mehrere Dateien mit Fluchtpunkten, die als Eingabe des Kalibrieralgorithmus *hyperbel* dienen können. Anschließend können aus den vorliegenden Dateien mit einem Hilfsprogramm Dateien als Eingabeparameter des Algorithmus *rotate* generiert werden.

Möchte man Bildpunkte als Eingabe der Algorithmen *hyperbel* oder *rotate* benutzen, so hat dies manuell zu geschehen.

Abbildungsverzeichnis

2.1 Bauteile einer binokularen Kameraplattform 10
2.2 Koordinatensysteme 12
2.3 Veranschaulichung von homogenen Koordinaten 15
2.4 Freiheitsgrade der Kameraplattform 17
2.5 Modell eines Kameraobjektivs 19
2.6 Modell des Schärfentiefebereichs eines Kameraobjektivs 21
2.7 Abbildung der Sehstrahlen durch die projektive Transformation . 22
2.8 Geometrie von Fluchtpunkten 24
2.9 Abbildung des Kamerasystems 25
2.10 Architektur des realisierten Regelkreises 25

3.1 Bestimmung des Bildhauptpunktes durch Fokus und Zoom 39
3.2 Punktverschiebung auf der Sensorfläche durch Rotation 41
3.3 Schnitt eines Rotationskegels mit einer Ebene 45
3.4 Objektträger zur Rotation einer Kamera 55
3.5 Schnitt eines Rotationskegels mit einer Ebene 57
3.6 Klassifizierung der Kalibrierverfahren 60

4.1 Der Kalibrierwürfel 63
4.2 Funktionen zur Szenenkoordinatenzuordung 64
4.3 Beispielzuordnung der Szenenmerkmale 65
4.4 Kalibrierwürfel nach Kantendetektion 66
4.5 Bildraum und Parameterraum bei der Hough-Transformation . . . 68
4.6 Akkumulatorfeld der Hough-Transformation 69
4.7 Segmentierung des Kalibrierwürfels und detektierte Geraden . . . 76
4.8 Beispielgeraden 78
4.9 Detektierte Bildmerkmale auf einem Kalibrierwürfel 80

5.1 Legende 83
5.2 Standardabweichung des Meßfehlers in den Punkten 84
5.3 Standardabweichung des Meßfehlers in der Winkelbestimmung . . 85
5.4 Entfernung zwischen Kamera und Szene 87
5.5 Gesamtrotationsschwenk 88
5.6 Symmetrieindex 90

5.7 Relative Orientierung der Parallelenschar 91
5.8 Anzahl der Bildmerkmale 92
5.9 Anzahl der Zoomschritte 93
5.10 Kalibrierung durch Zoomen 95
5.11 Vergenzrotation zur Kalibrierung 96
5.12 Neigungsrotation zur Kalibrierung 97

Tabellenverzeichnis

4.1 Beispielzuordnung der Szenenmerkmale 65
4.2 Beispielgeraden 78
4.3 Beispielzuordnung der Bildmerkmale 79

5.1 Standardabweichung des Meßfehlers 86
5.2 Ergebnisse von Realexperimenten 94

Literaturverzeichnis

[Bas95] A. Basu. Active Calibration of Cameras. *IEEE Trans. Systems, Man, and Cybernetics*, 25(2):256–265, 1995.

[BB82] D.H. Ballard und C.M. Brown. *Computer Vision.* Prentice-Hall, 1982.

[Beß90] R. Beß. Kalibrierung einer beweglichen, monokularen Kamera zur Tiefengewinnung aus Bildfolgen. In *Proc. of the 16. DAGM-Symposium, Wien*, Seiten 524–531, 1990.

[BMZ92] P. Beardsley, D. Murray und A. Zisserman. Camera Calibration using Multiple Images. In *Proc. Second European Conference on Computer Vision*, Seiten 312–320. Santa Margerita, Italy, May 23-26, G. Sandini (Ed.), Lecture Notes in Computer Science 588, Springer-Verlag, Berlin et al., 1992.

[Che91] H.H. Chen. A Screw Motion Approach to Uniqueness Analysis of Head-Eye Geometry. In *IEEE Conf. Computer Vision and Pattern Recognition*, Seiten 145–151. Maui, Hawaii, June 3-6, 1991.

[CT90] B. Caprile und V. Torre. Using Vanishing Points for Camera Calibration. *International Journal on Computer Vision*, 4:127–140, 1990.

[Dav94] E.R. Davies. The Performance of the Generalized Hough Transform: Concavities, Ambiguities and Positional Accuray. In *Proc. Alvey Vision Conference*, Seiten 327–333. Reading, UK, Sept. 25-28, 1994.

[DB93] F. Du und M. Brady. Self-Calibration of the Intrinsic Parameters of Cameras for Active Vision Systems. In *IEEE Conf. Computer Vision and Pattern Recognition*, Seiten 477–482. New York, NY, June 15-17, 1993.

[DE96a] K. Daniilidis und J. Ernst. Active Intrinsic Calibration using Vanishing Points. In *IEEE Conf. Computer Vision and Pattern Recognition*, Seiten 708–713. San Francisco, CA, 1996.

[DE96b] K. Daniilidis und J. Ernst. Active Intrinsic Calibration using Vanishing Points. *Pattern Recognition Letters*, 1996. to appear.

[DHKS95] K. Daniilidis, M. Hansen, C. Krauss und G. Sommer. Auf dem Weg zum künstlichen aktiven Sehen: Modellfreie Bewegungsverfolgung durch Kameranachführung. In G. Sagerer, S. Posch und F. Kummert, Hrsg., *Mustererkennung 1995*, Seiten 277–284, Tagungsband 17. DAGM–Symposium, Bielefeld, Springer Verlag, 1995.

[Ech90] T. Echigo. A Camera Calibration Technique using three Sets of Parallel Lines. *Machine Visions and Applications*, 3:159–167, 1990.

[EM94] U. Eckhardt und M. Maderlechner. Parametrization of the Hough Transform. In *Proc. Int. Conf. on Pattern Recognition*, Seiten 52–59. Atlantic City, NJ, June 17-21, 1994.

[Fau93] O. Faugeras. *Three-Dimensional Computer Vision.* MIT-Press, Cambridge, MA, 1993.

[Fis91] G. Fischer. *Lineare Algebra.* Verlag Vierweg, 9. Auflage, 1991.

[Fis94] G. Fischer. *Analytische Geometrie.* Verlag Vierweg, 6. Auflage, 1994.

[FT86] O.D. Faugeras und G. Toscani. The Calibration Problem for Stereo. In *IEEE Conf. Computer Vision and Pattern Recognition*, Seiten 15–20, Miami Beach, FL, June 22–26, 1986.

[Gen94] V. Gengenbach. *Einsatz von Rückkoplungen in der Bildauswertung bei einem Hand-Auge-System zur automatischen Demontage.* Sankt Augustin, Infix, DISKI72, 1994.

[GvL89] G.H. Golub und C.F. van Loan. *Matrix Computations.* The Johns Hopkins University Press, Baltimore, Maryland, 2. Auflage, 1989.

[Har94] R.J. Hartley. Self-Calibration from Multiple Views with a Rotating Camera. In *Proc. Third European Conference on Computer Vision*, Seiten 471–478. Stockholm, Sweden, May 2-6, J.O. Eklundh (Ed.), Springer LNCS 800, 1994.

[HDM94] R. Horaud, F. Dornaika und R. Mohr. Self Calibration of a Stereo Head Mounted onto a Robot Arm. In *Proc. Fourth European Conference on Computer Vision*, Seiten 455–462, 1994.

[Hou62] P.V.C. Hough. Methods and Means for Recognizing Complex Patterns, 1962. U.S. Patent 3,069,654.

[IK87] J. Illingworth und J. Kittler. A Survey of Efficient Hough Transform Methods. In *Proc. Alvey Vision Conference*, Seiten 319–326. Cambridge, UK, Sept. 15-17, 1987.

[Jäh93] B. Jähne. *Digitale Bildverarbeitung.* Springer Verlag, 3. Auflage, 1993.

[Kan93] K. Kanatani. *Geometric Computation for Machine Vision.* Oxford University Press, Oxford, UK, 1993.

[Kol95] H. Kollnig. *Ermittlung von Verkehrsgeschehen durch Bildfolgenauswertung.* Sankt Augustin, Infix, DISKI88, 1995.

[Kra86] K. Kraus. *Photogrammetrie.* F. Dümmler Verlag, Bonn, 1986.

[LB95] M. Li und D. Betsis. Hand-Eye Calibration. In *Proc. Int. Conf. on Computer Vision*, Seiten 40–46. Boston, MA, June 20-23, 1995.

[Lea93] V.L. Leavers. Survey - Which Hough Transform. *CVGIP: Image Understanding*, 58(2):250–264, 1993.

[Len87] R. Lenz. Linsenfehlerkorrigierte Eichung von Halbleiterkameras mit Standartobjektiven für hochgenaue 3D-Messungen in Echtzeit. In *Proc. of the 9. DAGM-Symposium*, Seiten 212–215, 1987.

[Li94] M. Li. Camera Calibration of a Head-Eye System for Active Vision. In *Proc. Third European Conference on Computer Vision*, Seiten 543–554, Stockholm, Sweden, May 2-6, J.O. Eklundh (Ed.), Springer LNCS 800, 1994.

[LL95] M. Li und J.M. Lavest. Some Aspects of Zoom-Lens Camera Calibration. *Report from Computational Vision and Active Perception Laboratory, Stockholm*, 1995.

[LMLK94] E. Lutton, H. Maître und J. Lopez-Krahe. Constribution to the Determination of Vanishing Points using Hough Transform. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 16(4):430–438, 1994.

[LV94] Q. Luong und T. Viéville. Canonic Representations for the Geometries of Multiple Projective Views. In *Proc. Third European Conference on Computer Vision*, Seiten 589–599. Stockholm, Sweden, May 2-6, J.O. Eklundh (Ed.), Springer LNCS 800, 1994.

[MF92] S.J. Maybank und O.D. Faugeras. A Theory of Self-Calibration of a Moving Camera. *International Journal of Computer Vision*, 8:123–151, 1992.

[Nag83] H.-H. Nagel. Overview on Image Sequence Analysis. In T. Huang, Hrsg., *Image Sequence Processing and Dynamic Scene Analysis*, Seiten 2–39. Springer-Verlag, 1983.

[PIK92] J. Princen, J. Illingworth und J. Kittler. A Formal Definition of the Hough Transform: Properties and Relationships. *Journal of Mathematical Imaging and Vision*, 1:153–168, 1992.

[Sch90] R. Schnabel. *Geometrie I*. Christian-Albrechts-Universität Kiel, 1990. Ausarbeitung zu einer Vorlesung gehalten im WS 1989/90 an der Universität Kiel.

[Som95] G. Sommer. Verhaltensbasierter Entwurf technisch visueller Systeme. *Künstliche Intelligenz*, 5(3):42–45, 1995.

[SS93] M.J. Swain und M.A. Stricker. Promising Directions in Active Vision. *International Journal of Computer Vision*, 11(2):109–126, 1993. Written by the Attendees of the NSF Actice Vision Workshop, University of Chicago, August 1991.

[Ste95] G. Stein. Accurate Internal Camera Calibration using Rotation, with Analysis of Sources of Errors. In *International Conference on Computer Vision*, Seiten 230–236. Boston, MA, June 20-23, 1995.

[Sto93] J. Stoer. *Numerische Mathematik 1 und 2*. Springer Lehrbuch, Springer Verlag Berlin, 1.-5.. Auflage, 1972-93.

[TL88] R. Y. Tsai und R.K. Lens. Real Time Versatile Robotics Hand/Eye Calibration using 3D Machine Vision. In *IEEE International Conference on Robotics and Automation*, Seiten 554–561, 1988.

[Tsa86] R.Y. Tsai. An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. In *IEEE Conf. Computer Vision and Pattern Recognition*, Seiten 364–374. Miami Beach, FL, June 22–26, 1986.

[Tsa89] R. Tsai. Synopsis of Recent Progress on Camera Calibration for 3D Machine Vision. In O. Khatib et al., Hrsg., *The Robotics Review*, Seiten 147–159. The MIT Press, Cambridge, MA, 1989.

[TTG94] K. Tarabanis, R.Y. Tsai und D.S. Goodman. Calibration of a Computer Controlled Robotic Vision Sensor with a Zoom Lens. *Computer Vision, Graphics, and Image Processing*, 59(2):226–241, 1994.

[Vié94] T. Viéville. Auto-Calibration of Visual Sensor Parameters on a Robotic Head. *Image and Vision Computing*, 12(4):227–237, 1994.

[Wan92] C. C. Wang. Extrinsic Calibration of a Vision Sensor Mounted Robot. *IEEE Transaction on Robotics and Automation*, 8(2):161–175, 1992.

[YHHY92] G.S. Young, T.H. Hong, M. Herman und J.C.S. Yang. Kinematic Calibration of an Active Camera System. In *IEEE Conf. Computer Vision and Pattern Recognition*, Seiten 748–751, Champaign, IL, June 15-18, 1992.

Erklärung:

Hiermit versichere ich, daß ich die vorliegende Arbeit selbständig verfaßt und ausschließlich die angegebenen Hilfsmittel und Quellen verwendet habe.

Kiel, den