

1 Dornier-Scanner

Auf den folgenden Seiten ist die Funktionsweise und mathematische Modellierung der Dornier Entfernungsbildkamera beschrieben. Als Ergebnis liegen für jeden Scan 4032 Tripel aus Entfernungswert r_i , Spiegelachsenablenkung α_i und Vertikalachsenablenkung β_i (r_i, α_i, β_i) vor.

Diese können nun genutzt werden, um die kartesische Darstellung der Messdaten zu erzeugen. Sei das kartesische Scannerkoordinatensystem (x, y, z) so orientiert, dass die z -Achse in Messrichtung, die y -Achse parallel zur Rotationsachse des Ablenkspiegels nach unten und die x -Achse orthogonal dazu nach rechts orientiert sei. Dann gilt für die Koordinaten eines Messtripels (x_i, y_i, z_i) (Kugelkoordinaten):

$$\begin{aligned}x_i &= r_i \cdot \cos(\beta_i) \cdot \sin(\alpha_i) \\y_i &= -r_i \cdot \sin(\beta_i) \\z_i &= r_i \cdot \cos(\beta_i) \cdot \cos(\alpha_i)\end{aligned}$$

Hierbei sind die Richtungen für $\alpha = \beta = 0$ genau in z -Richtung orientiert.

2 RoSi-Scanner

2.1 Aufbau des RoSi-Scanners

Der RotatingSick-Scanner besteht aus einem SICK LMS 200 Lasermesssystem, welches über eine Drehdurchführung an eine Dunker-Motor-Getriebe-Kombination gekoppelt ist. Die Rotationsachse entspricht hierbei der mittleren Blickrichtung des Scanners. Als Rotationssensor dient ein optisch inkrementeller Encoder, welcher seine Signale quadraturkodiert an die Auswerteelektronik weitergibt. Zusammengeführt werden die Entfernung- und Winkelmessungen des LMS-Scanners und des Winklenkoders durch einen Standard-PC, welcher über eine Digital-I/O-Karte bzw. eine schnelle RS-422-Verbindung die Kommunikation abwickelt.

Abbildung 1 zeigt den RoSi-Scanner ohne Verarbeitungshardware.

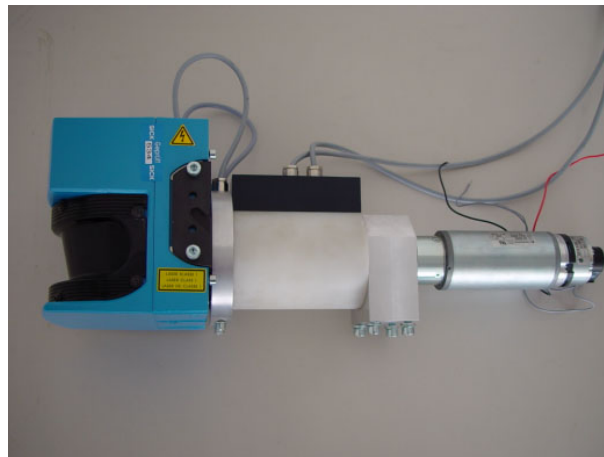


Abbildung 1: Der RoSi-Scanner

2.2 Modellierung des RoSi-Scanners

Der optisch inkrementelle Encoder erzeugt 2000 Impulse pro Achsumdrehung. Er ist direkt an der Motorachse angebracht. Das Planetengetriebe besitzt ein Umsetzverhältnis von 36:1. Daher liefert der inkrementelle Geber pro Scannerumdrehung 72000 Impulse, d.h. die theoretisch erreichbare Auflösung der Roationsachse des Scanners liegt bei

$$\frac{360^\circ}{36 \cdot 2000} = \frac{1}{200}^\circ$$

Das LMS von Sick besitzt folgende Eigenschaften:

- Scanwinkel: 100°
- Winkelauflösung: $0,25^\circ$, d.h. 400 Messwerte pro Scanzeile
- Entfernungsauflösung: 1mm
- Genauigkeit: $\pm 2\text{cm}$
- Reichweite: ca. 50 m
- Scan-Frequenz: 20 Hz
- Einzelscandauer: 40 ms (inkl. Übertragung)

Durch die Rotation des Scanners erhält das Gesamtsystem einen Öffnungswinkel der Scan-Keule von 100° . Bei einer fest eingestellten Rotationsgeschwindigkeit von 30° pro Sekunde misst der Scanner 6 Sekunden für eine halbe Umdrehung und erhält damit ein vollständiges Tiefenbild.

Bei dieser Rotationsgeschwindigkeit beträgt die rotatorische Winkelauflösung ca. $1,5^\circ$ pro Scan. Ein Tiefenbild besteht dann aus $6 \times 20 \times 400 = 48000$ Messwerten.

Sei das Scannerkoordinatensystem derart orientiert, dass die z-Achse der Rotationsachse des Scanners entspricht, die x-Achse in der Scanebene bei Rotationswinkel 0° liegt und die y-Achse dementsprechend orthogonal nach unten zeigt.

Dann gilt als einfache Approximation des Scans ohne Rotation für alle Tupel (r_i, α_i) mit r_i als gemessener Entfernung und α_i als Messwinkel in der Scanebene (0° -Richtung auf der x-Achse):

$$\begin{aligned}x_i &= \cos(\alpha_i) \cdot r_i \\y_i &= 0 \\z_i &= \sin(\alpha_i) \cdot r_i\end{aligned}$$

Betrachtet man nun zusätzlich die Rotation des Scanners um die Rotationsachse mit einem nach dem Scan aufgenommenen Winkelmaß β , so gilt:

$$\begin{aligned}
x_i &= \cos(\beta) \cdot \cos(\alpha_i) \cdot r_i \\
y_i &= \sin(\beta) \cdot \cos(\alpha_i) \cdot r_i \\
z_i &= \sin(\alpha_i) \cdot r_i
\end{aligned}$$

Berücksichtigt man nun noch die tatsächliche Scandauer von $t_S = 20ms$ als der Zeit, die der Laser benötigt, um über die 180° gefahren zu werden, so muss man die Rotation des Scanners innerhalb dieser Zeit auch einrechnen:

$$\begin{aligned}
x_i &= \cos(\beta - (400 - i)RG/(400t_S)) \cdot \cos(\alpha_i) \cdot r_i \\
y_i &= \sin(\beta - (400 - i)RG/(400t_S)) \cdot \cos(\alpha_i) \cdot r_i \\
z_i &= \sin(\alpha_i) \cdot r_i
\end{aligned}$$

mit RG als aktueller Rotationsgeschwindigkeit (rad/s).

2.3 Kalibrierung des RoSi-Scanners

Zwei grundsätzliche Probleme bleiben bei der Scannermodellierung:

- Da ein inkrementeller optischer Encoder verwendet wurde, ist die Anfangsposition des Scanners nicht bekannt, d.h. der Winkel β kann nicht global angegeben werden.
- Zwischen der Aufnahme des Scans im LMS und der Auswertung und Zuordnung eines Rotationswinkels β liegt eine gewisse Zeitspanne t_d (delay). Diese Zeitspanne ist durch die Dauer der Kommunikation sowie sonstige Betriebssystemaktivitäten gegeben. Der Effekt zeigt sich darin, dass einem aktuellen Scan ein falscher Winkel β zugewiesen wird. Es handelt sich hierbei um einen systematischen Fehler.

Beide Probleme können durch Angabe eines Korrekturwinkels $\delta\beta$, welcher als Offset zum aktuellen β addiert wird, behoben werden.

Wie lässt sich der Korrekturwinkel bestimmen ?

In der Kalibrierphase wird ein spezielles, bekanntes Kalibrierobjekt vermessen (in unserem Fall eine einfache, senkrecht stehende Stange, ca. 1m lang mit einer Kantenlänge von 4 cm). Diese Stange befindet sich zu Beginn der Messung in einem vorher festgelegten Winkelausschnitt und Entfernungsbereich des Scanners und kann daher leicht aus der generierten Punktwolke extrahiert werden.

Im Idealfall liefert das System bereits einen korrekt orientierten Scan, d.h. der Korrekturwinkel $\Delta\beta = 0$. In diesem Fall, sollte man durch die Punktwolke des Kalibrierobjektes eine Regressionsgerade legen können, welche im globalen Koordinatensystem senkrecht liegt, d.h. in negativer y-Richtung. Um das nachfolgende Verfahren numerisch stabiler zu halten, sollten die x- und y-Koordinaten vertauscht werden, so dass die zu berechnende Steigung nahe bei 0 und nicht bei ∞ liegt.

Seien nun die Punkte $(x_i, y_i, z_i), i = 1, \dots, n$ als zum Kalibrierobjekt gehörig erkannt, so kann man aus der Geradengleichung in 2D

$$y_i = mx_i + b$$

folgendes, überbestimmtes LGS ableiten:

$$\underbrace{\begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}}_X \cdot \underbrace{\begin{pmatrix} m \\ b \end{pmatrix}}_{\vec{M}} = \underbrace{\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}}_{\vec{Y}}$$

Dieses lässt sich durch Anwendung der kleinsten Fehlerquadrate-Methode (least square) lösen zu:

$$\begin{aligned} X \cdot \vec{M} &= \vec{Y} \\ \Leftrightarrow X^T X \vec{M} &= X^T \vec{Y} \\ \Leftrightarrow (X^T X)^{-1} X^T X \vec{M} &= (X^T X)^{-1} X^T \vec{Y} \\ \Leftrightarrow \vec{M} &= (X^T X)^{-1} X^T \vec{Y} \end{aligned}$$

Damit ist nun die tatsächliche Steigung m bestimmt und die Kalibrierstange wird im Winkel $\beta = \text{atan}(m)$ angemessen. Um dies nun zu korrigieren, kann die Winkeldifferenz $\Delta\beta = 90^\circ - \beta$ als Offset in die Berechnung der kartesischen Koordinaten aufgenommen werden:

$$\begin{aligned}
x_i &= \cos(\beta + \Delta\beta - (400 - i)RG/(400t_S)) \cdot \cos(\alpha_i) \cdot r_i \\
y_i &= \sin(\beta + \Delta\beta - (400 - i)RG/(400t_S)) \cdot \cos(\alpha_i) \cdot r_i \\
z_i &= \sin(\alpha_i) \cdot r_i
\end{aligned}$$

mit RG als aktueller Rotationsgeschwindigkeit (rad/s).

3 Modellierung des Kraft-Momenten-Sensors

Kräfte und Momente lassen sich generell nicht direkt messen. Sie werden beispielsweise durch Messen von Deformationen, die mit Hilfe von Dehnmessstreifen oder Piezokristallen in eine elektrische Spannung umgewandelt werden können, indirekt ermittelt.

Solange die auf den Sensor einwirkenden Kräfte nicht zu groß sind, darf in erster Näherung angenommen werden, dass die von den Dehnmessstreifen erzeugten Spannungen linear von den wirkenden Kräften und Momenten abhängig sind. Bei der Verwendung von n Sensorelementen kann die Transformation zwischen den gemessenen Spannungen W_i und den Kräften und Drehmomenten als lineare Überlagerung der einzelnen orthogonalen Komponenten des angreifenden Kraft-Momenten-Vektors beschrieben werden.

Hieraus ergibt sich folgender Ansatz für die einzelne Messspannung W_i

$$W_i = c_{i,1}F_x + c_{i,2}F_y + c_{i,3}F_z + c_{i,4}M_x + c_{i,5}M_y + c_{i,6}M_z$$

mit den Kraftkomponenten F_x, F_y und F_z sowie den Momenten um die Achsen M_x, M_y und M_z .

Betrachtet man einen typischen Kraft-Momenten-Sensor auf der Basis von Dehnmessstreifen, so sind zumeist über 16 DMS-Signale zu digitalisieren und danach zu verarbeiten. Hieraus ergibt sich ein linear unabhängiges, überbestimmtes lineares Gleichungssystem bestehend aus $n \gg 6$ Gleichungen für die W_i mit

$$\vec{W} = C\vec{F}$$

und

$$\vec{W} = \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \end{pmatrix}, \quad \vec{F} = \begin{pmatrix} F_x \\ F_y \\ F_z \\ M_x \\ M_y \\ M_z \end{pmatrix}, \quad C = \begin{pmatrix} c_{1,1} & \cdots & c_{1,6} \\ c_{2,1} & & c_{2,6} \\ \vdots & \ddots & \vdots \\ c_{n,1} & \cdots & c_{n,6} \end{pmatrix}$$

Die Matrix C wird üblicherweise als Kopplungsmatrix bezeichnet. Die Auflösung dieses überbestimmten LGS nach dem Kraft-Momenten-Vektor F

kann über die Methode der Kleinsten-Fehlerquadrate mittels der sogenannten pseudoinversen Matrix R_F der Kopplungsmatrix C durchgeführt werden:

$$\vec{F} = R_F \vec{W}$$

mit R_F als

$$R_F = \begin{pmatrix} r_{1,1} & \cdots & r_{1,n} \\ \vdots & \ddots & \vdots \\ r_{6,1} & \cdots & r_{1,n} \end{pmatrix}.$$

Die pseudoinverse Matrix R_F ergibt sich durch Umformung aus der Kopplungsbeziehung $\vec{W} = C\vec{F}$

$$\begin{aligned} \vec{W} &= C\vec{F} \\ \Leftrightarrow C^T \vec{W} &= C^T C \vec{F} \\ \Leftrightarrow (C^T C)^{-1} C^T \vec{W} &= \vec{F} \\ \Leftrightarrow R_F \vec{W} &= \vec{F} \end{aligned}$$

mit

$$R_F = (C^T C)^{-1} C^T.$$

3.1 Kalibrierung des KMS

Die Kopplungsmatrix C bzw. die Entkopplungsmatrix R_F sind zunächst nicht bekannt und werden bei der Kalibrierung des Kraft-Momenten-Sensors (KMS) bestimmt. Die Parameter der Kopplungsmatrix analytisch aus den geometrischen und technologischen Eigenschaften des Kraft-Momenten-Sensors bzw. des Einzeldehnmessstreifens abzuleiten, führt nicht zu brauchbaren Ergebnissen. Daher wird ein KMS üblicherweise empirisch über einen Versuchsaufbau beim Hersteller kalibriert.

Ziel der Kalibrierung ist es also, die $6n$ Parameter der Matrizen C bzw. R_F zu bestimmen. Für den statischen Fall wird hierbei ein Versuchsaufbau eingesetzt, mit dem es möglich ist, definierte Kräfte und Momente auf den KMS wirken zu lassen. Hierbei werden dann die auftretenden Einzelspannungen W_i protokolliert. Bei m Versuchen erhält man m Paare aus Kraftvektoren und Messspannungsvektoren $(\vec{F}^i, \vec{W}^i)$ mit $m > n$.

Durch Umstellung von $\vec{W} = C\vec{F}$ zu $\vec{W}^T = \vec{F}^T C^T$ erhält man wiederum ein überbestimmtes LGS für die $6n$ Unbekannten $C_{i,j}$

$$\begin{pmatrix} \vec{W}_1^T \\ \vec{W}_2^T \\ \vdots \\ \vec{W}_m^T \end{pmatrix} = \begin{pmatrix} \vec{F}_1^T \\ \vec{F}_2^T \\ \vdots \\ \vec{F}_m^T \end{pmatrix} \cdot C^T$$

oder

$$W = F \cdot C^T.$$

Hieraus ergibt sich durch Auflösung mit der Pseudoinversen $(F^T F)^{-1} F^T$ die transponierte Koppelmatrix C^T zu

$$C^T = (F^T F)^{-1} F^T W.$$

Aus C^T lässt sich wie bereits gezeigt die Entkopplungsmatrix R_F berechnen.

Durch eine große Anzahl an Messungen m mit einem breiten Spektrum verschiedener Kraft- und Momentvorgaben ist die mathematische Stabilität des Verfahrens gewährleistet.

4 Datenhandschuh

In diesem Fall wird der Cyberglove Datenhandschuh der Firma Virtex Technologies betrachtet. Dieser Datenhandschuh verfügt über 22 Aufnehmer, zur Messung der Fingergelenkstellungen der rechten Hand. Die Finger- bzw. Daumenstellung wird nicht direkt gemessen (beispielsweise über absolute optische Encoder), da ein Zugriff auf die Rotationsachsen der menschlichen Hand verwehrt ist und sich ein Exoskelett für den Datenhandschuh aus Gründen der Beweglichkeit und intuitiven Bedienbarkeit des selben verbietet. Daher wurde hier der Weg gewählt, die Gelenkstellungen durch die auf den Handschuhstoff bei einer Fingerkrümmung einwirkenden Kräfte zu bestimmen. Kräfte und Momente lassen sich generell nicht direkt messen. Sie werden beispielsweise durch Messen von Deformationen, die mit Hilfe von Dehnungsmessstreifen oder Piezokristallen in eine elektrische Spannung umgewandelt werden können, indirekt ermittelt. In diesem Fall wurden 22 Dehnungsmessstreifen als Kraftaufnehmer genutzt. Diese wurden jeweils Mäanderförmig auf dem entsprechenden Flächenelement des Handschuhs verlegt, um die Messbarkeit der Widerstands- und damit Spannungsänderung zu verbessern.

Eine experimentelle Untersuchung ergab für den Einzeldehnungsmessstreifen einen affinen Zusammenhang zwischen Messspannung und Fingergelenkwinkel (in gewissen Grenzen), d.h. die Messspannung W_i lässt sich darstellen als

$$W_i = c_i \alpha_i + t_i$$

mit α_i als Gelenkwinkel, t_i als konstantem Offset und c_i als konstantem Proportionalitätsfaktor. Dies bedeutet, dass für jeden DMS eine lineare Beziehung mit zwei Unbekannten (c_i, t_i) gültig ist, aus der dann der Zusammenhang zwischen der gemessenen Spannung W_i und dem gesuchten Fingergelenkwinkel α_i hervorgeht.

4.1 Kalibrierung des Datenhandschuhs

Die Kalibrierung des Datenhandschuhs bezeichnet das Problem, zu einer gegebenen Benutzerhand die Parameter (c_i, t_i) für alle $i = 1, \dots, 22$ zu bestimmen.

Das normale Vorgehen hierbei ist nun, ein lineares Gleichungssystem für die gesuchten 44 Unbekannten aufzustellen. Hierzu müssen jedoch in ausreichen-

der Menge m bekannte Tupel (W_i^m, α_i^m) aus Messspannungen und korrespondierenden Fingergelenkwinkeln für jeden DMS i vorliegen.

Es wird also ein Versuchsaufbau benötigt, welcher die Messung des ohmschen Widerstandes der einzelnen Dehnmessstreifen i unter der Nebenbedingung vorgegebener Fingergelenkwinkel α_i ermöglicht. Die Problematik der Fingergelenkwinkelvorgabe lässt sich in erster Näherung durch den Einsatz spezialisierter Kalibrierobjekte lösen. Die Benutzerhand muss diese Kalibrierobjekte unter Verwendung vorgegebener Griffmuster ergreifen, so dass alle Fingergelenkwinkelstellungen erzwungen und daher bekannt sind. Die hierbei auftretenden Effekte bzw. Spannungen W_i an den DMS werden protokolliert und dienen nun zur Lösung des linearen Gleichungssystems.

Seien W_i^j die im Versuch j am Fingergelenk i aufgetretene Messspannung und α_i^j der zugehörige Fingergelenkwinkel (für $j = 1, \dots, m, i = 1, \dots, 22$), so lässt sich für jedes $i = 1, \dots, 22$ folgendes LGS über die Unbekannten c_i und t_i aufstellen:

$$\begin{aligned} W_i^1 &= c_i \alpha_i^1 + t_i \\ W_i^2 &= c_i \alpha_i^2 + t_i \\ &\vdots \\ W_i^m &= c_i \alpha_i^m + t_i \end{aligned}$$

Dies lässt sich schreiben als:

$$\begin{pmatrix} W_i^1 \\ W_i^2 \\ \vdots \\ W_i^m \end{pmatrix} = \begin{pmatrix} \alpha_i^1 & 1 \\ \alpha_i^2 & 1 \\ \vdots & \vdots \\ \alpha_i^m & 1 \end{pmatrix} \cdot \begin{pmatrix} c_i \\ t_i \end{pmatrix}$$

oder mit geeignet definierten Matrizen und Vektoren als

$$\vec{W}_i = A_i \cdot \vec{CT}_i.$$

Für alle $m > 2$ handelt es sich hierbei um ein überbestimmtes lineares Gleichungssystem, welches nach der Methode der kleinsten Fehlerquadrate (least square) mit der Pseudoinversen gelöst wird

$$\begin{aligned}
\vec{W}_i &= A_i \vec{C}T_i \\
\Leftrightarrow A_i^T \vec{W} &= A_i^T A_i \vec{C}T_i \\
\Leftrightarrow (A_i^T A_i)^{-1} A_i^T \vec{W}_i &= \vec{C}T_i \\
\Leftrightarrow P_{A_i} \vec{W}_i &= \vec{C}T_i
\end{aligned}$$

mit der Pseudoinversen P_{A_i}

$$P_{A_i} = (A_i^T A_i)^{-1} A_i^T.$$

4.2 Modellierung des Datenhandschuhs

Nachdem nun für alle i die Pseudoinversen P_{A_i} berechnet wurden und daher die (c_i, t_i) bekannt sind, kann der Datenhandschuh über seine Fingergelenke modelliert werden.

Hierbei wird jeder Finger bzw. Daumen als offene kinematische Kette mit entsprechend vielen und geeignet positionierten Gelenken modelliert. Abbildung 2 zeigt ein Beispiel eines aus den Sensordaten berechneten Modells.

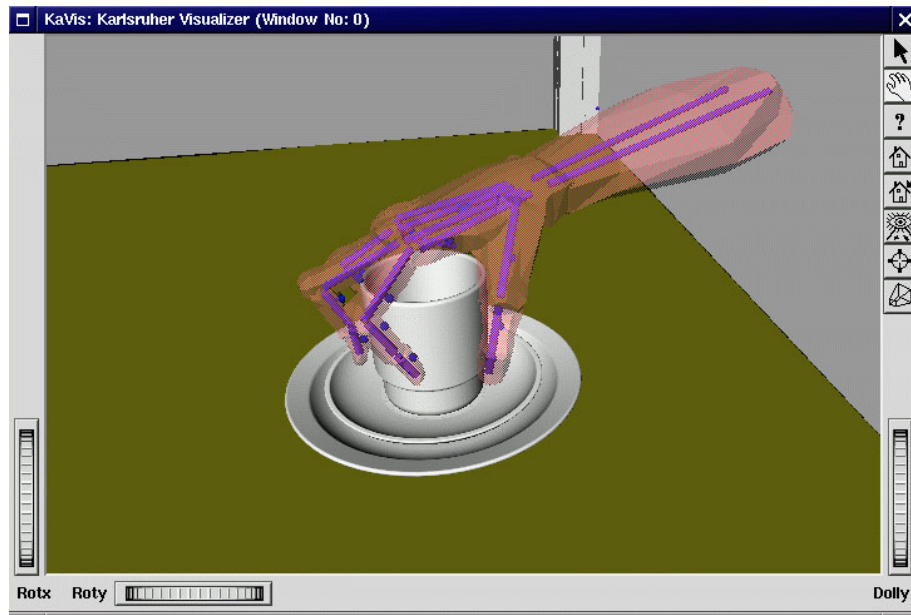


Abbildung 2: Visualisierung des Datenhandschuhmodells

5 Modellierung des Kamerasensors - Lochkameramodell

Bei der Modellierung einer Kamera gibt es grundsätzlich zwei unterschiedliche Aspekte, die modelliert werden können. Zum einen handelt es sich um die Modellierung der einfallenden Lichtintensität und deren Abbildung auf das resultierende, digitalisierte Signal, zum anderen um die Modellierung der geometrischen Eigenschaften der perspektivischen Abbildung durch die Kameralinse auf die Bildebene.

Während die an einer bestimmten Position auf der Bildebene einfallende Lichtintensität abhängig ist von

- dem Umgebungslicht
- sowie den Lagen und Reflexionseigenschaften der vorhandenen Objekte,

muss bei der geometrischen Modellierung die Abbildung des dreidimensionalen Raumes auf die zweidimensionale Bildebene konstruiert werden (siehe Abbildung 3).

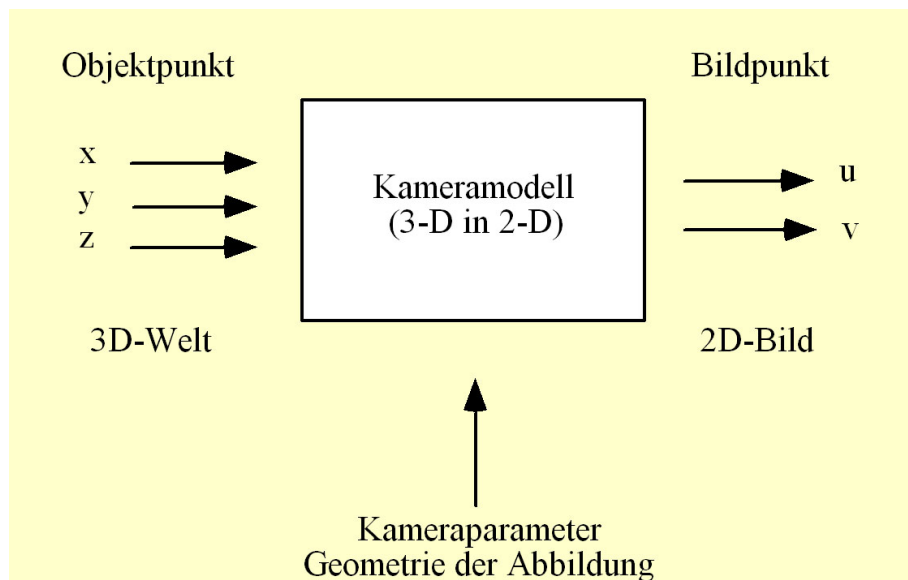


Abbildung 3: Schema des geometrischen Kameramodells

Ein relativ einfaches, aber für viele Anwendungen durchaus geeignetes geometrisches Modell einer Kamera stellt das sogenannte Lochkameramodell dar.

Bei diesem mathematischen Modell schrumpft das Objektiv/das Linsensystem des Sensors zu einem geometrischen Punkt zusammen, dem sogenannten Projektionszentrum. Hinter dem Projektionszentrum befindet sich im Abstand c , der sogenannten Brennweite, die Bildebene (aus der Perspektive von Objekten betrachtet, die sich vor der Kamera befinden). Der orthogonale Schnitt eines Lichtstrahls mit der Bildebene durch das Projektionszentrum (die sogenannte optische Achse) bildet den sogenannten Bildhauptpunkt, also die Mitte der Bildebene. Abbildung 4 verdeutlicht diese Zusammenhänge, wobei hierbei das Projektionszentrum mit dem Nullpunkt des globalen Weltkoordinatensystems zusammenfällt und die optische Achse entlang der z -Achse verläuft. Dementsprechend befindet sich der Bildhauptpunkt an der Stelle $(0, 0, -c)$.

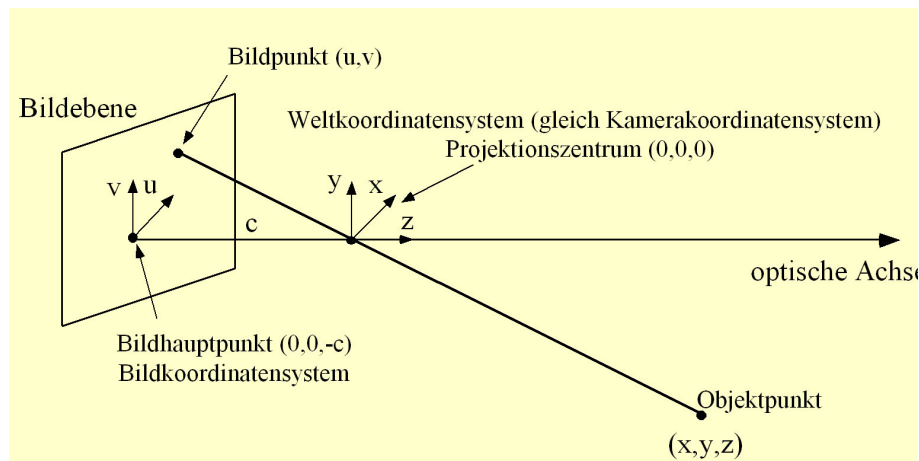


Abbildung 4: Das Lochkameramodell

Ein beliebiger Punkt (x, y, z) wird - folgt man dem Strahlensatz - abgebildet auf den Bildpunkt (u, v) mit

$$u = -c \frac{x}{z}, \quad v = -c \frac{y}{z},$$

d.h. die Abbildung

$$f\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) \rightarrow \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -c \frac{x}{z} \\ -c \frac{y}{z} \end{pmatrix}$$

ist leider keine lineare Abbildung. Für viele geometrische Operationen wäre eine lineare perspektivische Abbildung jedoch äußerst praktisch.

Daher beschreibt man das Kameramodell gerne mit Mitteln der projektiven Geometrie, d.h. man transformiert das Modellierungsproblem vom euklidischen Vektorraum zum projektiven Vektorraum und arbeitet mit sogenannten homogenen Koordinaten.

5.1 Grundlegendes zu homogenen Koordinaten

Transformation euklidischer Koordinaten in homogene Koordinaten:

$$\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \rightarrow \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

Für alle homogenen Vektoren gilt:

$$\vec{x} \equiv \vec{y} : \exists a \in R : \vec{x} = a \cdot \vec{y},$$

d.h. zwei Vektoren sind dann äquivalent (gleich), wenn sie reelle Vielfache voneinander sind.

Die Rücktransformation homogener Koordinaten in euklidische Koordinaten:

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} \rightarrow \begin{pmatrix} x/w \\ y/w \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \rightarrow \begin{pmatrix} x/w \\ y/w \\ z/w \end{pmatrix}$$

Bei homogenen Vektoren (x, y, z, w) mit $w \neq 0$ spricht man von Punkten im dreidimensionalen Raum (zweidimensional analog), bei $w = 0$ spricht man von Richtungen.

Nun wozu das ganze ?

Betrachten wir die Translation eines Vektors (X, Y, Z, W) um einen Vektor (x_0, y_0, z_0) (analog in 2D):

$$\begin{pmatrix} X' \\ Y' \\ Z' \\ W' \end{pmatrix} = T(x_0, y_0, z_0) \cdot \begin{pmatrix} X \\ Y \\ Z \\ W \end{pmatrix}, \text{ mit } T(x_0, y_0, z_0) = \begin{pmatrix} 1 & 0 & 0 & x_0 \\ 0 & 1 & 0 & y_0 \\ 0 & 0 & 1 & z_0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

d.h. aus einer im euklidischen affinen Transformation - der Translation - ist durch Verwendung homogener Koordinaten eine lineare Transformation geworden.

Ausgehend von der euklidischen Rotationsmatrix R_0 für die drei Achsen $(\varphi, \vartheta, \psi)$ mit

$$R_0(\varphi, \vartheta, \psi) = ROT(z, \varphi) \cdot ROT(y, \vartheta) \cdot ROT(x, \psi) = \begin{pmatrix} \cos(\varphi) \cos(\vartheta) & \cos(\varphi) \sin(\vartheta) \sin(\psi) - \sin(\varphi) \cos(\psi) & \cos(\varphi) \sin(\vartheta) \cos(\psi) + \sin(\varphi) \sin(\psi) \\ \sin(\varphi) \cos(\vartheta) & \sin(\varphi) \sin(\vartheta) \sin(\psi) + \cos(\varphi) \cos(\psi) & \sin(\varphi) \sin(\vartheta) \cos(\psi) - \cos(\varphi) \sin(\psi) \\ -\sin(\vartheta) & \cos(\vartheta) \sin(\psi) & \cos(\vartheta) \cos(\psi) \end{pmatrix}$$

ergibt sich die Rotation eines Vektors (X, Y, Z, W) zu

$$\begin{pmatrix} X' \\ Y' \\ Z' \\ W' \end{pmatrix} = R(\varphi, \vartheta, \psi) \cdot \begin{pmatrix} X \\ Y \\ Z \\ W \end{pmatrix}, \text{ mit } R(\varphi, \vartheta, \psi) = \begin{pmatrix} R_0(\varphi, \vartheta, \psi) & 0 \\ 0 & 1 \end{pmatrix},$$

d.h. auch hier liegt eine lineare Abbildung vor. Zuletzt betrachten wir die im euklidischen Raum nichtlineare perspektivische Projektion $P(c)$ mit

$$\begin{pmatrix} X' \\ Y' \\ W' \end{pmatrix} = P(c) \cdot \begin{pmatrix} X \\ Y \\ Z \\ W \end{pmatrix} = \begin{pmatrix} -cX \\ -cY \\ Z \end{pmatrix} = \begin{pmatrix} uZ \\ vZ \\ Z \end{pmatrix}, \text{ mit } P(c) = \begin{pmatrix} -c & 0 & 0 & 0 \\ 0 & -c & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Die Rücktransformation von (uZ, vZ, Z) in den euklidischen Raum zeigt, dass hier die perspektivische Projektion als lineare Transformation durchgeführt werden kann.

5.2 Das Lochkameramodell

Sei nun eine Kamera mit Brennweite c gegeben, deren Projektionszentrum an der Stelle (x_0, y_0, z_0) in der Welt liegt und deren optische Achse parallel zur durch die drei Raumwinkel $(\varphi, \vartheta, \psi)$ gegebene Gerade verläuft. Ein beliebiger Punkt im euklidischen Raum $r_o = (x, y, z)$ kann nun abgebildet werden auf seine Koordinaten auf der Bildebene $r_b = (u, v)$ durch Anwendung der homogenen Transformationen:

$$\begin{pmatrix} ut \\ vt \\ t \end{pmatrix} = P(c) \cdot R(\varphi, \vartheta, \psi) \cdot T(x_0, y_0, z_0) \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

bzw. mit der Kalibriermatrix $A(c, \varphi, \vartheta, \psi, x_0, y_0, z_0) = P(c) \cdot R(\varphi, \vartheta, \psi) \cdot T(x_0, y_0, z_0)$ als

$$\begin{pmatrix} ut \\ vt \\ t \end{pmatrix} = A(c, \varphi, \vartheta, \psi, x_0, y_0, z_0) \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

Somit ist es durch Anwendung projektiver Geometrie und homogenen Koordinaten gelungen, die geometrische Abbildung eines Raumpunktes auf die Bildebene einer Lochkamera durch eine lineare Transformation zu beschreiben.

Das Kameramodell ist dann durch die Matrix A gegeben. Wie zu sehen, besteht das einfache Modell aus dem einzigen internen Parameter c , der Brennweite, sowie aus den sechs externen Parametern $\varphi, \vartheta, \psi, x_0, y_0, z_0$, der Lage der Kamera im Raum.

Leider erfasst das Modell mit nur einem internen Parameter die Abbildungseigenschaften realer Kamerasysteme nur unzureichend.

Gründe hierfür sind beispielsweise:

- der Ursprung des Bildkoordinatensystems liegt nicht im Bildhauptpunkt (zusätzliche Translation in der Bildebene notwendig, zwei zusätzliche Parameter)
- die u - und v -Achse liegen fertigungsbedingt nicht orthogonal (es tritt ein zusätzlicher Scherungsparameter auf)

- die Skalierungen entlang der u - und v -Achsen sind nicht identisch (zwei zusätzliche Skalierungsfaktoren)

Hierdurch ergeben sich fünf weitere interne Faktoren, welche die Matrix P und damit die Matrix A beeinflussen. Da diese Faktoren analytisch nur ungenügend hergeleitet werden können, geht man üblicherweise den Weg, die Matrix A durch ein Kalibrierverfahren empirisch zu bestimmen.

6 Kalibrierung des Kamerasensors - Lochkameramodell

Ein relativ einfaches, aber für viele Anwendungen durchaus geeignetes geometrisches Modell einer Kamera stellt das sogenannte Lochkameramodell dar. Bei diesem mathematischen Modell schrumpft das Objektiv/das Linsensystem des Sensors zu einem geometrischen Punkt zusammen, dem sogenannten Projektionszentrum. Hinter dem Projektionszentrum befindet sich im Abstand c , der sogenannten Brennweite, die Bildebene (aus der Perspektive von Objekten betrachtet, die sich vor der Kamera befinden). Abbildung 5 verdeutlicht diese Zusammenhänge.

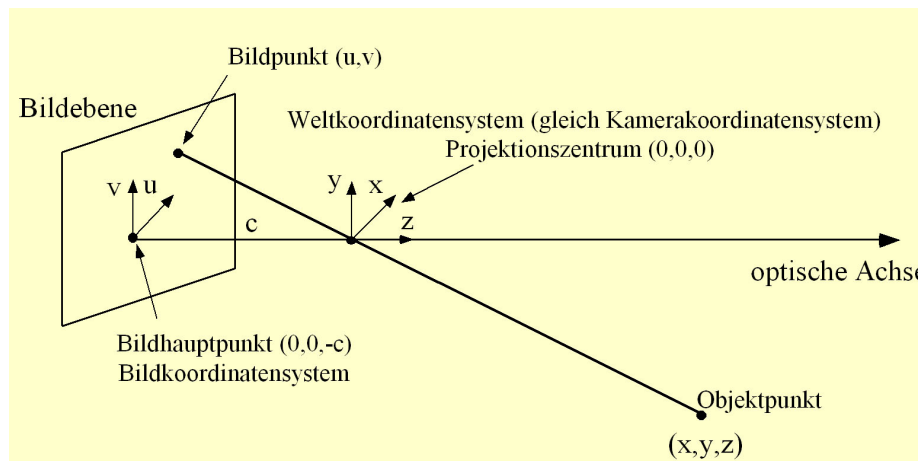


Abbildung 5: Das Lochkameramodell

Ein beliebiger Punkt (x, y, z) wird - folgt man dem Strahlensatz - abgebildet auf den Bildpunkt (u, v) mit

$$u = -c \frac{x}{z}, \quad v = -c \frac{y}{z},$$

d.h. die Abbildung

$$f\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) \rightarrow \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -c \frac{x}{z} \\ -c \frac{y}{z} \end{pmatrix}$$

ist leider keine lineare Abbildung. Für viele geometrische Operationen wäre eine lineare perspektivische Abbildung jedoch äußerst praktisch.

Daher beschreibt man das Kameramodell gerne mit Mitteln der projektiven Geometrie, d.h. man transformiert das Modellierungsproblem vom euklidischen Vektorraum zum projektiven Vektorraum und arbeitet mit sogenannten homogenen Koordinaten.

6.1 Grundlegendes zu homogenen Koordinaten

Transformation euklidischer Koordinaten in homogene Koordinaten:

$$\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \rightarrow \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

Für alle homogenen Vektoren gilt:

$$\vec{x} \equiv \vec{y} : \exists a \in R : \vec{x} = a \cdot \vec{y},$$

d.h. zwei Vektoren sind dann äquivalent (gleich), sie reelle Vielfache voneinander sind.

Die Rücktransformation homogener Koordinaten in euklidische Koordinaten:

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} \rightarrow \begin{pmatrix} x/w \\ y/w \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \rightarrow \begin{pmatrix} x/w \\ y/w \\ z/w \end{pmatrix}$$

Bei homogenen Vektoren (x, y, z, w) mit $w \neq 0$ spricht man von Punkten im dreidimensionalen Raum (zweidimensional analog), bei $w = 0$ spricht man von Richtungen.

Nun wozu das ganze ?

6.2 Das Kalibrierverfahren

Die im euklidischen Vektorraum nichtlineare perspektivische Projektion kann in homogenen Koordinaten als lineare Transformation dargestellt werden.

Sei eine Kamera mit Brennweite c gegeben, deren Projektionszentrum an der Stelle (x_0, y_0, z_0) in der Welt liegt und deren optische Achse parallel zur durch die drei Raumwinkel $(\varphi, \vartheta, \psi)$ gegebene Gerade verläuft. Ein beliebiger Punkt im euklidischen Raum $r_o = (x, y, z)$ kann nun abgebildet werden auf seine Koordinaten auf der Bildebene $r_b = (u, v)$ durch Anwendung der homogenen Transformationen:

$$\begin{pmatrix} ut \\ vt \\ t \end{pmatrix} = A(c, \varphi, \vartheta, \psi, x_0, y_0, z_0) \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}, \text{ mit } A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \end{pmatrix}$$

Aufgabe der Kalibrierung ist es nun, die 12 Unbekannten $a_{i,j}$ empirisch zu bestimmen. Hierzu wird folgender Ansatz gewählt.

Ausgehend aus obiger Gleichung werden die Parameter u, v und t bestimmt zu:

$$\begin{aligned} ut &= a_{1,1}x + a_{1,2}y + a_{1,3}z + a_{1,4} \\ vt &= a_{2,1}x + a_{2,2}y + a_{2,3}z + a_{2,4} \\ t &= a_{3,1}x + a_{3,2}y + a_{3,3}z + a_{3,4} \end{aligned}$$

Hieraus erhält man u und v zu:

$$\begin{aligned} u &= \frac{a_{1,1}x + a_{1,2}y + a_{1,3}z + a_{1,4}}{a_{3,1}x + a_{3,2}y + a_{3,3}z + a_{3,4}} \\ v &= \frac{a_{2,1}x + a_{2,2}y + a_{2,3}z + a_{2,4}}{a_{3,1}x + a_{3,2}y + a_{3,3}z + a_{3,4}} \end{aligned}$$

Da ja alle reellen Vielfachen der Matrix A dieser äquivalent sind, kann o.B.d.A. ein Parameter normiert werden. Hier sei $a_{3,4} := 1$ gewählt. Man erhält durch Umformung (mit dem Nenner erweitern)

$$\begin{aligned} a_{1,1}x + a_{1,2}y + a_{1,3}z + a_{1,4} - a_{3,1}xu - a_{3,2}yu - a_{3,3}zu &= u \\ a_{2,1}x + a_{2,2}y + a_{2,3}z + a_{2,4} - a_{3,2}xv - a_{3,2}yv - a_{3,3}zv &= v \end{aligned}$$

Hierbei handelt es sich nun um zwei linear unabhängige Gleichungen für die noch 11 Unbekannten $a_{i,j}$, falls eine Korrespondenz aus gültigen Bildkoordinaten (u, v) und dem entsprechenden Raumpunkt (x, y, z) bekannt ist.

Das Kalibrierverfahren muss nun sicherstellen, dass eine ausreichend große Menge derartiger Korrespondenzen $((x_i, y_i, z_i), (u_i, v_i))$ vorliegt. Bei 11 Unbekannten und 2 Gleichungen pro Korrespondenz müssen mindestens 6 Korrespondenzen bekannt sein.

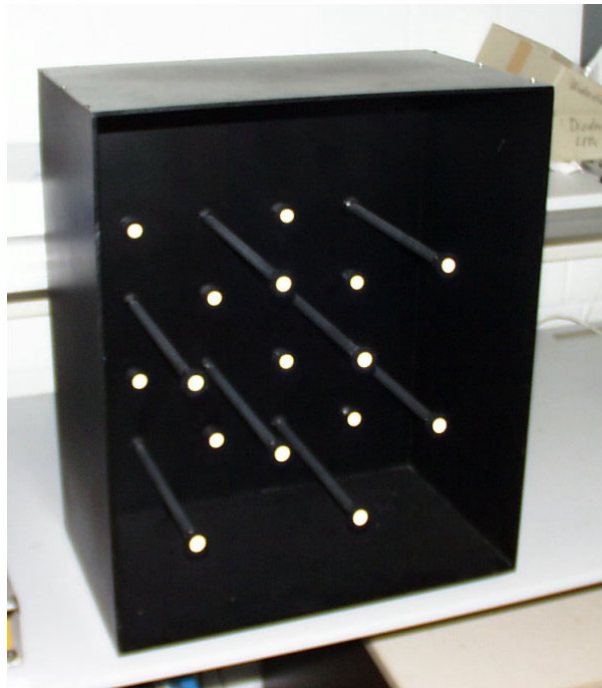


Abbildung 6: Das Kalibrierobjekt

Am IAIM wird das in Abbildung 6 gezeigte Kalibrierobjekt genutzt, welches über eine Koordinatenmessmaschine hochgenau vermessen wurde. Die 16 unterschiedlich langen Stäbe innerhalb der Box sind mit Retrotargets bestückt, um eine möglichst gute Reflexion einstrahlenden Lichtes zu ermöglichen. Der Ablauf der Kalibrierung ist nun wie folgt:

1. Das Kalibrierobjekt wird so vor die Kamera gestellt, dass die 16 Retrotargets alle zu sehen sind und möglichst das gesamte Blickfeld abdecken.
2. Das Kalibrierobjekt wird aus der Kameraperspektive angeleuchtet, um die Retrotargets im Kamerabild als äußerst hell leuchtende Punkte abzubilden.

3. Durch eine Schwellwertfilterung (Binarisierung) auf dem Kamerabild kann der dunkle Hintergrund der Box ausgeblendet werden und nur die hellen Flecken der Retrotargets bleiben im Bild erhalten.
4. Anhand der relativen Positionen der Flecken im Bild, können die Schwerpunkte der Flecken den global bekannten Targetpositionen zugewiesen werden. Hierdurch ist die Menge an Korrespondenzen $((x_i, y_i, z_i), (u_i, v_i))$ bestimmt.

Aus obigen Gleichungen ergibt bei $i = 1, \dots, n$ Korrespondenzen folgendes, überbestimmtes LGS für die 11 Unbekannten:

$$\underbrace{\begin{pmatrix} x_1 & y_1 & z_1 & 1 & 0 & 0 & 0 & 0 & -u_1x_1 & -u_1y_1 & -u_1z_1 \\ 0 & 0 & 0 & 0 & x_1 & y_1 & z_1 & 1 & -v_1x_1 & -v_1y_1 & -v_1z_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & z_n & 1 & 0 & 0 & 0 & 0 & -u_nx_n & -u_ny_n & -u_nz_n \\ 0 & 0 & 0 & 0 & x_n & y_n & z_n & 1 & -v_nx_n & -v_ny_n & -v_nz_n \end{pmatrix}}_X \cdot \underbrace{\begin{pmatrix} a_{1,1} \\ a_{1,2} \\ a_{1,3} \\ a_{1,4} \\ a_{2,1} \\ \vdots \\ a_{3,3} \end{pmatrix}}_{\vec{A}} = \underbrace{\begin{pmatrix} u_1 \\ v_1 \\ \vdots \\ u_n \\ v_n \end{pmatrix}}_{\vec{UV}}$$

Das LGS lässt sich mit der Methode der kleinsten Fehlerquadrate lösen. Hierzu formt man um

$$\begin{aligned} X \cdot \vec{A} &= \vec{UV} \\ \Leftrightarrow X^T X \cdot \vec{A} &= X^T \vec{UV} \\ \Leftrightarrow (X^T X)^{-1} X^T X \cdot \vec{A} &= (X^T X)^{-1} X^T \vec{UV} \\ \Leftrightarrow \vec{A} &= (X^T X)^{-1} X^T \vec{UV} \end{aligned}$$

und erhält somit die noch fehlenden 11 Parameter der Kalibriermatrix A .

Die so gewonnene Kamerakalibrierung ist natürlich abhängig von der aktuellen Kameraposition und -orientierung in der Welt (den externen Kameraparametern) sowie den internen Parametern (fertigungsspezifische Parameter, aktuelle Brennweite etc.).

Wird die Kamera definiert in der Welt bewegt, so kann die Kalibriermatrix durch Transformation mit den entsprechenden homogenen Translations- und Rotationsmatrizen an die neue Lage angepasst werden. Interne Kameraparameter sollten bzw. können sich nicht ändern.

Genutzt werden kann die Kalibrierung, um 3D-Punkte auf die Bildebene abzubilden. Der Hauptanwendungsfall liegt jedoch darin, mit mehreren kalibrierten Kameras 3D-Rekonstruktionen korrespondierender Pixel in den Bildebenen durchzuführen.

7 Stereokamerasysteme - 3D Sensorik

Ein relativ einfaches, aber für viele Anwendungen durchaus geeignetes geometrisches Modell einer einzelnen Kamera stellt das sogenannte Lochkameramodell dar. Bei diesem mathematischen Modell schrumpft das Objektiv/das Linsensystem des Sensors zu einem geometrischen Punkt zusammen, dem sogenannten Projektionszentrum. Hinter dem Projektionszentrum befindet sich im Abstand c , der sogenannten Brennweite, die Bildebene (aus der Perspektive von Objekten betrachtet, die sich vor der Kamera befinden). Abbildung 7 verdeutlicht diese Zusammenhänge.

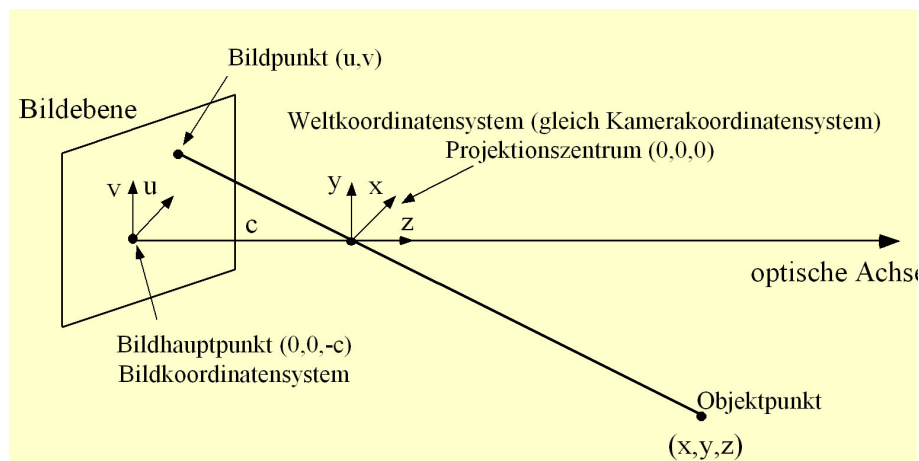


Abbildung 7: Das Lochkameramodell

Ein beliebiger Punkt (x, y, z) wird - folgt man dem Strahlensatz - abgebildet auf den Bildpunkt (u, v) mit

$$u = -c \frac{x}{z}, \quad v = -c \frac{y}{z},$$

d.h. die Abbildung

$$f\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) \rightarrow \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -c \frac{x}{z} \\ -c \frac{y}{z} \end{pmatrix}$$

ist leider keine lineare Abbildung. Für viele geometrische Operationen wäre eine lineare perspektivische Abbildung jedoch äußerst praktisch.

Daher beschreibt man das Kameramodell gerne mit Mitteln der projektiven Geometrie, d.h. man transformiert das Modellierungsproblem vom euklidischen Vektorraum zum projektiven Vektorraum und arbeitet mit sogenannten homogenen Koordinaten.

Die im euklidischen Vektorraum nichtlineare perspektivische Projektion einer Kamera k kann in homogenen Koordinaten als lineare Transformation über eine Transformationsmatrix A_k dargestellt werden.

7.1 Die Stereorekonstruktion

Ein beliebiger Punkt im euklidischen Raum $r_o = (x, y, z)$ kann nun abgebildet werden auf seine Koordinaten auf der Bildebene $r_b = (u, v)$ durch Anwendung der homogenen Transformationen:

$$\begin{pmatrix} ut \\ vt \\ t \end{pmatrix} = A_k \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}, \text{ mit } A_k = \begin{pmatrix} a_{1,1}^k & a_{1,2}^k & a_{1,3}^k & a_{1,4}^k \\ a_{2,1}^k & a_{2,2}^k & a_{2,3}^k & a_{2,4}^k \\ a_{3,1}^k & a_{3,2}^k & a_{3,3}^k & a_{3,4}^k \end{pmatrix}$$

(im Folgenden gilt ein hochgestelltes k nicht als Potenz, sondern als Kameraindex).

Nach der Kamerakalibrierung sind die 12 Parameter $a_{i,j}^k$ jeder Kameramatrix A_k bekannt. Ausgehend von obiger Gleichung werden die Parameter u, v und t bestimmt zu:

$$\begin{aligned} u^k t &= a_{1,1}^k x + a_{1,2}^k y + a_{1,3}^k z + a_{1,4}^k \\ v^k t &= a_{2,1}^k x + a_{2,2}^k y + a_{2,3}^k z + a_{2,4}^k \\ t &= a_{3,1}^k x + a_{3,2}^k y + a_{3,3}^k z + a_{3,4}^k \end{aligned}$$

Man erhält durch Umformung (t einsetzen, ausmultiplizieren und umsortieren)

$$\begin{aligned} a_{1,1}^k x + a_{1,2}^k y + a_{1,3}^k z - a_{3,1}^k x u^k - a_{3,2}^k y u^k - a_{3,3}^k z u^k &= a_{3,4}^k u^k - a_{1,4}^k \\ a_{2,1}^k x + a_{2,2}^k y + a_{2,3}^k z - a_{3,1}^k x v^k - a_{3,2}^k y v^k - a_{3,3}^k z v^k &= a_{3,4}^k v^k - a_{2,4}^k \end{aligned}$$

und nun durch Ausklammern der Unbekannten x, y und z

$$\begin{aligned} (a_{1,1}^k - a_{3,1}^k u^k) x + (a_{1,2}^k - a_{3,2}^k u^k) y + (a_{1,3}^k - a_{3,3}^k u^k) z &= a_{3,4}^k u^k - a_{1,4}^k \\ (a_{2,1}^k - a_{3,1}^k v^k) x + (a_{2,2}^k - a_{3,2}^k v^k) y + (a_{2,3}^k - a_{3,3}^k v^k) z &= a_{3,4}^k v^k - a_{2,4}^k \end{aligned}$$

Um die unbekannte Raumposition (x, y, z) eines ausgewählten Bildpunktes mit den Koordinaten (u, v) auf der Bildebene zu berechnen, liegen bei der Verwendung einer einzelnen Kamera, nicht genügend Informationen vor. Dies zeigt sich daran, dass nur zwei unabhängige Gleichungen für die Information (u, v) gegeben sind.

Für jede weitere Kamera, welche denselben Raumpunkt (x, y, z) betrachtet, können nun jedoch zwei weitere Gleichungen zur Suche der unbekannten Position des Raumpunktes eingesetzt werden. Schon bei zwei Kameras $k = 1, 2$ ist es also möglich, die gesuchte Position zu rekonstruieren, da hier bereits vier Gleichungen für drei Unbekannte vorliegen.

Grundlage für diese Rekonstruktion ist jedoch die Sicherheit, dass die Koordinaten (u^1, v^1) und (u^2, v^2) auf den jeweiligen Bildebenen der Kameras 1 und 2 tatsächlich Abbilder des selben Raumpunktes (x, y, z) darstellen. Das Problem, derartig korrespondierende Bildkoordinaten in unterschiedlichen Kamerabildern aufzufinden, nennt man das Stereokorrespondenzproblem.

Sei eine Menge von $k = 1, \dots, n, n \geq 2$ korrespondierenden Bildkoordinaten (u^k, v^k) in den jeweiligen Kamerabildern k gegeben, so lässt sich zur Rekonstruktion des korrespondierenden Raumpunktes (x, y, z) folgendes, überbestimmtes LGS aufstellen:

$$\underbrace{\begin{pmatrix} a_{1,1}^1 - a_{3,1}^1 u^1 & a_{1,2}^1 - a_{3,2}^1 u^1 & a_{1,3}^1 - a_{3,3}^1 u^1 \\ a_{2,1}^1 - a_{3,1}^1 v^1 & a_{2,2}^1 - a_{3,2}^1 v^1 & a_{2,3}^1 - a_{3,3}^1 v^1 \\ \vdots & \vdots & \vdots \\ a_{1,1}^n - a_{3,1}^n u^n & a_{1,2}^n - a_{3,2}^n u^n & a_{1,3}^n - a_{3,3}^n u^n \\ a_{2,1}^n - a_{3,1}^n v^n & a_{2,2}^n - a_{3,2}^n v^n & a_{2,3}^n - a_{3,3}^n v^n \end{pmatrix}}_X \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \underbrace{\begin{pmatrix} a_{3,4}^1 u^1 - a_{1,4}^1 \\ a_{3,4}^1 v^1 - a_{2,4}^1 \\ \vdots \\ a_{3,4}^n u^n - a_{1,4}^n \\ a_{3,4}^n v^n - a_{2,4}^n \end{pmatrix}}_{\vec{R}}$$

Das LGS lässt sich durch Anwendung der Methode der kleinsten Fehlerquadrate (least square) lösen, indem die Pseudoinverse der Matrix X gebildet wird:

$$\begin{aligned} X \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \vec{R} \\ \Leftrightarrow X^T X \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= X^T \vec{R} \end{aligned}$$

$$\Leftrightarrow (X^T X)^{-1} X^T X \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (X^T X)^{-1} X^T \vec{R}$$

$$\Leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (X^T X)^{-1} X^T \vec{R}$$

Somit kann die gesuchte Raumposition (x, y, z) unter Verwendung der Kamerakalibrierungen sowie der gefundenen korrespondierenden Pixelkoordinaten der Abbilder des Raumpunktes rekonstruiert werden.

7.2 Das Stereokorrespondenzproblem

Die Lösung des Stereokorrespondenzproblems ist trotz aller Fortschritte auf diesem Gebiet immer noch ein aktuelles Forschungsthema. Die Strategien der Problemlösung hängen sehr von den benutzten Merkmalen wie auch den zur Verfügung stehenden Daten ab.

Bei der Merkmalsauswahl stellt sich die Frage, zu welchen Elementen in einem gegebenen Kamerabild die entsprechenden Korrespondenten in dem oder den anderen Kamerabild(ern) gesucht werden sollen. Hierbei gibt es unterschiedliche Strategien, wie beispielsweise die Suche nach Einzelpixeln, die Suche nach Kantenpixeln oder die Suche nach Flächenpixelelementen.

Die zur Verfügung stehenden Daten entscheiden dann bei gegebener Merkmalsauswahl über die anwendbaren Strategien. Als Grundlage können reine Schwarz/Weiß-Bilder, Grauwertbilder mit unterschiedlichen Quantisierungsaufösungen sowie Farbbilder mit verschiedenen Farbtiefen vorhanden sein.

Um beispielsweise in Grauwertbildern korrespondierende Einzelpixel zu suchen, werden zumeist Korrelationsverfahren eingesetzt. Zu einem gegebenen Pixel mit den Koordinaten (u^1, v^1) in Bild 1 wird dessen näher Nachbarschaft betrachtet und mit der Nachbarschaft eines hypothetischen Korrespondenten (u^2, v^2) in Bild 2 verglichen.

Seien $g_1(u, v)$ und $g_2(u, v)$ die Funktionen (Tabellen, Arrays), welche zu einer gegebenen Pixelposition den Grauwert an dieser Stelle enthalten, so lässt sich beispielsweise die SSD-Korrelation (Sum of squared differences) in einer

$(2n + 1) \times (2n + 1)$ -Nachbarschaft derart formulieren:

$$\sum_{i=-n}^n \sum_{j=-n}^n (g_1(u^1 + i, v^1 + j) - g_2(u^2 + i, v^2 + j))^2$$

Die Hypothese ist hierbei, dass der Korrespondent dieses Differenzmaß minimiert. Die Durchführung einer solchen Korrelation ist sehr rechenaufwendig (besonders bei größeren n). Eine Aussage, ob der gesuchte Pixel überhaupt auf der Chipfläche der zweiten Kamera enthalten ist oder vielleicht außerhalb zu liegen kam, ist nicht möglich.

Die epipolare Geometrie hilft glücklicherweise, die Suche nach dem Korrespondenten von der Gesamtfläche auf eine Gerade, Halbgerade oder gar Strecke einzuschränken, je nach relativer Lage der Kameras.

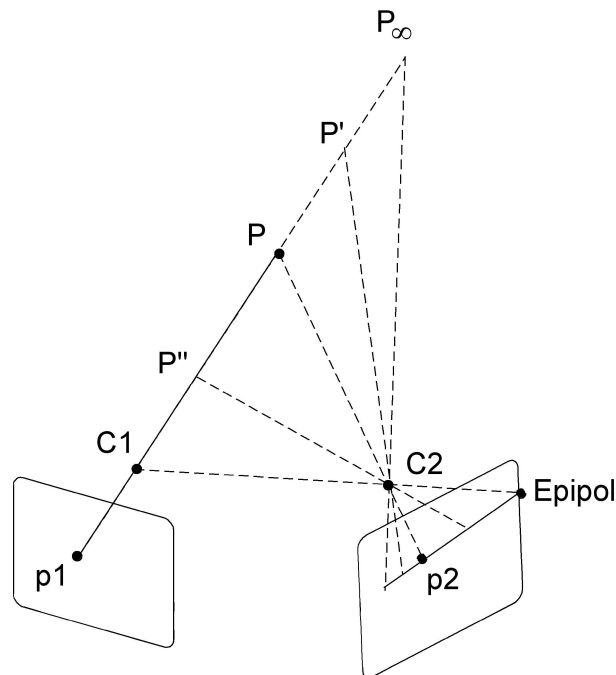


Abbildung 8: Epipolare Geometrie

Ein gegebener Raumpunkt P und seine Abbilder p_1 und p_2 in den jeweiligen Bildebenen spannen eine Ebene auf, die sogenannte epipolare Ebene. Der Schnitt dieser Ebene mit den Bildebenen ergibt die Epipolarlinien, also Geraden. Der zu einem Pixel p_1 gehörende Korrespondent kann also wegen der Abbildungsgeometrie der Kameras nur auf der zugehörigen Epipolarlinie im anderen Kamerabild liegen. Wie berechnet man nun aber diese Epipolarlinie?

Anschaulich betrachtet, verfolgt man einen Lichtstrahl von $p1$ durch das optische Zentrum $C1$ zurück bis ins Unendliche und bildet jeden möglichen Punkt auf dieser Halbgeraden durch das optische Zentrum $C2$ auf die zweite Bildebene ab. Es gibt hierbei zwei Extrempunkte: das eine Ende der Halbgeraden wird durch das optische Zentrum $C1$ gegeben, das andere durch den Punkt im Unendlichen, der in der projektiven Geometrie eine einfache und definierte Darstellungsform hat. Bildet man also die Abbildung dieser beiden Punkte auf die zweite Bildebene, hat man die Endpunkte der Epipolarlinie zu $p1$ gefunden.

Die Abbildung von $C1$ durch $C2$ ist bei beliebig gewähltem Punkt $p1$ immer identisch. Dieses Abbild nennt man den Epipol und es ist klar, dass sich alle Epipolarlinien immer in diesem Epipol schneiden bzw. ihren Anfang nehmen (bei Bildebenen, die in einer Ebene liegen, liegt der Epipol im Unendlichen). Der Epipol ist daher nur einmal zu berechnen. Nur das andere Ende der Epipolarlinie ist von $p1$ abhängig.

Es ist zwar nicht möglich, über die Kalibrierung A aus dem Abbild (u, v) auf die Raumposition (x, y, z) zurückzuschliessen, aber die Richtung, in der (x, y, z) liegt, ist ermittelbar. Als Richtung bezeichnet man in homogenen Koordinaten Punkte, die im Unendlichen liegen. Deren letzte Komponente ist 0.

Aus dem Ansatz

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = A_1 \cdot \begin{pmatrix} X \\ Y \\ Z \\ 0 \end{pmatrix}, \text{ mit } A_1 = \begin{pmatrix} a_{1,1}^1 & a_{1,2}^1 & a_{1,3}^1 & a_{1,4}^1 \\ a_{2,1}^1 & a_{2,2}^1 & a_{2,3}^1 & a_{2,4}^1 \\ a_{3,1}^1 & a_{3,2}^1 & a_{3,3}^1 & a_{3,4}^1 \end{pmatrix}$$

ergibt sich schnell

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} a_{1,1}^1 & a_{1,2}^1 & a_{1,3}^1 \\ a_{2,1}^1 & a_{2,2}^1 & a_{2,3}^1 \\ a_{3,1}^1 & a_{3,2}^1 & a_{3,3}^1 \end{pmatrix} \cdot \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

bzw.

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} a_{1,1}^1 & a_{1,2}^1 & a_{1,3}^1 \\ a_{2,1}^1 & a_{2,2}^1 & a_{2,3}^1 \\ a_{3,1}^1 & a_{3,2}^1 & a_{3,3}^1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

und damit die gesuchte Richtung $(X, Y, Z, 0)$. Durch Abbildung in die zweite

Kamera

$$\begin{pmatrix} X \\ Y \\ Z \\ 0 \end{pmatrix} \cdot A_2 = \begin{pmatrix} U \\ V \\ W \end{pmatrix}$$

ist dann der zweite Endpunkt der Epipolarlinie als $(U/W, V/W)$ gefunden.

Nun kann also auf dem zweiten Kamerabild die Korrelationssuche auf die Verbindungslinie zwischen den Epipolkoordinaten und $(U/W, V/W)$ eingeschränkt werden.

8 Zusatzaufgabe: Relokalisation mobiler Roboter in bekannten 2D-Umgebungen

8.1 Aufgabe

Entwerfen Sie in ihrer Gruppe ein oder mehrere Verfahren zur Lösung des unten beschriebenen Relokalisationsproblems. Bereiten Sie eine Präsentation auf Overhead-Folien vor.

Länge der Präsentation: max. 3 Minuten

8.2 Das Relokalisationsproblem

Teilaufgaben der Navigation mobiler Roboter sind die Lokalisation, die Bahnplanung und die Plandurchführung. Bei der Plandurchführung wird eine vorher geplante geometrische Bahn durch das Robotersystem abgefahren. Hierbei kommt permanent ein Soll-Ist-Vergleich der Roboterlage, also Roboterposition und -orientierung, zum Einsatz.

Die aktuelle Ist-Lage des Robotersystems gewinnt dieses aus einer initial bekannten Anfangslage und sich darauf auswirkender Fahrtooperationen wie beispielsweise Geradeausfahrt, Kreisbogenfahrt, Drehung auf der Stelle durch Anwendung von Odometriesensoren (zumeist optisch inkrementelle Encoder).

Leider entstehen auf die Dauer über diese Odometrie Ist-Positions-Fehler, da die gemessenen Radumdrehungen nicht immer mit der tatsächlichen Bewegung des Robotersystems korrelieren. Gerade auf Teppichböden kann es häufiger passieren, dass sich das Robotersystem durch Schlupf der Räder nicht so bewegt, wie die Radumdrehungen es eigentlich implizieren. Schon kleine Winkelfehler in der Ist-Orientierung wirken sich nach wenigen Metern Geradeausfahrt dramatisch auf den Ist-Positionsfehler aus, so dass ein rein odometriegeführter Dauerbetrieb wenig sinnvoll erscheint.

Abhilfe hierfür schafft man normalerweise durch Anwendung von Relokalisationsalgorithmen, welche zumeist auf der Perzeption ausgewählter Umweltmerkmale und deren Abgleich mit einer a priori bekannten Umweltkarte beruhen.

In dieser Aufgabe sei das Robotersystem mit einem planaren Laserscanner ausgestattet, welcher starr in Fahrtrichtung des Roboters zentral montiert sei

(es gibt also eine feste, affine Abbildung SF zwischen Scannermittelpunkt und Fahrzeugmittelpunkt). Der Scanner liefere etwa 20 mal pro Sekunde einen Datensatz

$$SCAN_i = \{(r_j, \alpha_j) | j = 0, \dots, 360\}$$

bestehend aus einer Entfernungsmessung r_j und dem zugehörigen Messwinkel α_j in Scannerkoordinaten.

Die aktuell vom Robotersystem angenommene Ist-Position sei (x_r, y_r, ϑ_r) und stellt eine brauchbare Approximation der tatsächlichen Ist-Position (x, y, ϑ) dar. Die Aufgabe des Relokalisationsproblems ist es also, den Differenzvektor $(\Delta x_r, \Delta y_r, \Delta \vartheta_r)$ zu bestimmen, mit dessen Hilfe man das Robotersystem wieder über seine tatsächliche Ist-Position belehren kann.

A priori vorgegeben ist unserem Robotersystem eine 2D-Landkarte der Umgebung in Weltkoordinaten in Form eines Grundrißplans bestehend aus signifikanten 2D-Kanten in der Scanebene des Laserscanners. Diese Kantenkarte ist einfach als Liste von Einzelkanten

$$KL = \{(s_i, e_i) | i = 0, \dots, n, s_i = \begin{pmatrix} x_{s_i} \\ y_{s_i} \end{pmatrix}, e_i = \begin{pmatrix} x_{e_i} \\ y_{e_i} \end{pmatrix}\}$$

mit ihren jeweiligen Start- und Endpunkten s_i und e_i gegeben.

Relokalisationsverfahren werden normalerweise relativ häufig ausgeführt, damit der Differenzvektor $(\Delta x_r, \Delta y_r, \Delta \vartheta_r)$ nicht zu gross wird und die Suche nach ihm schnell durchgeführt werden kann. Dementsprechend sollten diese Verfahren relativ schnell sein und müssen dafür nicht von zu großen Abweichungen zwischen interner Ist-Position und tatsächlicher Ist-Position ausgehen.