

Studienarbeit-Nr. 1624

Ein Kantenbasiertes Stereoverfahren

Christopher Mayer

1997

Universität Stuttgart
Fakultät Informatik

Inhaltsverzeichnis

Abkürzungen und Symbole	4
1 Einleitung	5
2 Das Kameramodell	7
2.1 Das Lochkameramodell.....	7
2.2 Die Abbildung Raum -> Bildebene.....	8
2.2.1 Die Zentralperspektive.....	8
2.2.2 Die räumliche Transformation.....	9
2.2.2.1 Die Translation.....	9
2.2.2.2 Die Rotation.....	9
2.2.3 Direkte Lineare Transformation.....	10
2.3 Diskussion des Kameramodells.....	11
3 Kalibrierung	12
4 Epipole und Epipolarlinien	14
4.1 Der Epipol.....	15
4.2 Die Epipolarlinien.....	16
4.3 Beispiele von Epipolarlinien.....	20
5 Matching	21
5.1 Eigenschaften der Kanten.....	21
5.2 Die Epipolar-Einschränkung.....	23
5.3 Der Matching-Algorithmus.....	24
5.4 Berechnung des Überlappenden Teils zweier Kanten.....	25
5.5 Beispiele.....	27

6 Rekonstruktion	28
6.1 Rekonstruktion durch Weglassen einer Gleichung.....	29
6.2 Rekonstruktion nach der Methode des kleinsten Fehlerquadrates.....	30
6.3 Beispiel.....	31
7 Die Software	32
7.1 Die Kalibrierung.....	33
7.2 Die Bildvorverarbeitung.....	33
7.3 Matching.....	33
7.4 Die Rekonstruktion.....	34
8 Diskussion der Ergebnisse	35
Abbildungsverzeichnis	36
Literaturverzeichnis	37

Verwendete Symbole

$P(x,y,z)$	Szenenpunkt
$P(X,Y,Z,w)$	Szenenpunkt in homogenen Koordinaten
$I(u,v)$	Bildpunkt
$I(U,V,w)$	Bildpunkt in homogenen Koordinaten
c	Brennweite
C	optisches Zentrum
B_1	linke Bildebene
B_2	rechte Bildebene
E_1	Epipol im linken Bild
E_2	Epipol im rechten Bild
M^1	DLT Matrix der linken Kamera
M^2	DLT Matrix der rechten Kamera
*	Mit * gekennzeichnete Bezeichner liegen in homogenen Koordinaten vor.

Kurzfassung

Die Vermessung von beliebigen Raumpunkten unter Verwendung von Vision-Systemen wird durch die Aufnahme von Objekten aus verschiedenen Blickpunkten ermöglicht. In dieser Arbeit wird hierfür eine Stereokamera, wie sie auch für mobile Roboter verwendet wird, eingesetzt. Um einen Raumpunkt ermitteln zu können, müssen die Bildkoordinaten dieses Punktes im linken als auch im rechten Bild bekannt sein. Voraussetzung hierfür ist die Lösung des Korrespondenzproblems, bei dem einem Szenenpunkt die entsprechenden Bildpunkte im linken wie auch im rechten Bild zugeordnet werden müssen. In dieser Arbeit wird das Korrespondenzproblem für Kantenbilder gelöst. D.h. die Kanten des linken Bildes werden denen des rechten Bildes zugeordnet. Dabei wird von einer kalibrierten Kamerageometrie ausgegangen, so daß durch die Epipolarlinien der Suchraum auf eine Gerade reduziert wird.

Das Stereoverfahren wird in eine graphische Oberfläche integriert, mit der es möglich ist das Verfahren zu testen und den Einfluß der verschiedenen Parametereinstellungen zu veranschaulichen.

1 Einleitung

Das räumliche Sehen ist für den Menschen selbstverständlich und fordert keinerlei Anstrengungen. Der Mensch erhält den räumlichen Eindruck seiner Umgebung dadurch, daß das linke und das rechte Auge Bilder sehen, die gewisse Unterschiede aufweisen. Diese Unterschiede kommen von den verschiedenen Blickwinkeln der Augen. Aus der Verschiebung der zueinander gehörenden Punkte im linken und im rechten Auge, der Disparität, entsteht die Tiefenwirkung.

Hier wird nun versucht dieses Verhalten auf den Rechner zu übertragen. Mit Hilfe einer Stereokamera werden Aufnahmen der Umgebung gemacht, aus den Bildern werden die Kanten extrahiert und einander zugeordnet. Aufgrund der Disparität läßt sich nun eine aufgenommene Szene rekonstruieren.

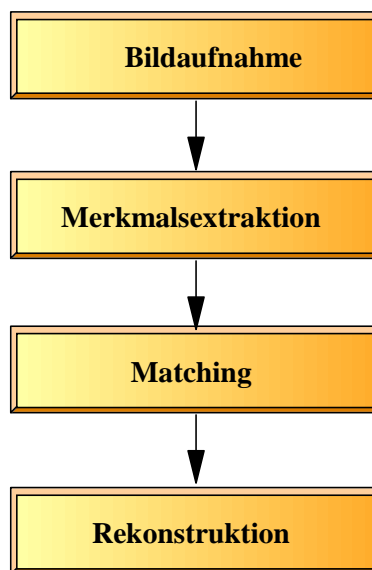


Abbildung 1.1 Stereobildverarbeitung

In der Robotik werden Stereoverfahren für autonome Fahrzeuge eingesetzt. In Kombination mit anderen Sensoren ermöglichen sie es dem Fahrzeug bzw. dem Roboter sich in einer ihm unbekanntem Umgebung zurechtzufinden. Auf andere Sensoren wie z.B. Ultraschall kann allerdings nicht verzichtet werden, da es beispielsweise so gut wie unmöglich ist den Abstand einer glatten Wand mit Hilfe einer Stereoverfahrens zu berechnen. Die Aufnahme einer solchen Wand würde nicht genügend Merkmale liefern. Bei solchen Anwendungen kommen noch die Probleme der Rechenleistung hinzu. Ein autonomes Fahrzeug sollte sehr schnell neue Aufnahmen auswerten können um rechtzeitig zu reagieren. Daher sollte die Bildvorverarbeitung unwesentliche Details herausfiltern können.

Eine andere Anwendung findet sich in der Photogrammetrie. Ziel ist es Höhenmodelle aus sich überlappenden Luftbildaufnahmen zu berechnen. Bei solchen Anwendungen ist die Rechenzeit nicht wesentlich, allerdings sind hier möglichst genaue Ergebnisse wünschenswert. Die Bildvoverarbeitung sollte somit nicht zu viele Details herausfiltern. Voraussichtlich werden bei solchen Verfahren auch die einzelnen Bildpunkte als Merkmale ausgewählt, was die Komplexität erheblich steigert.

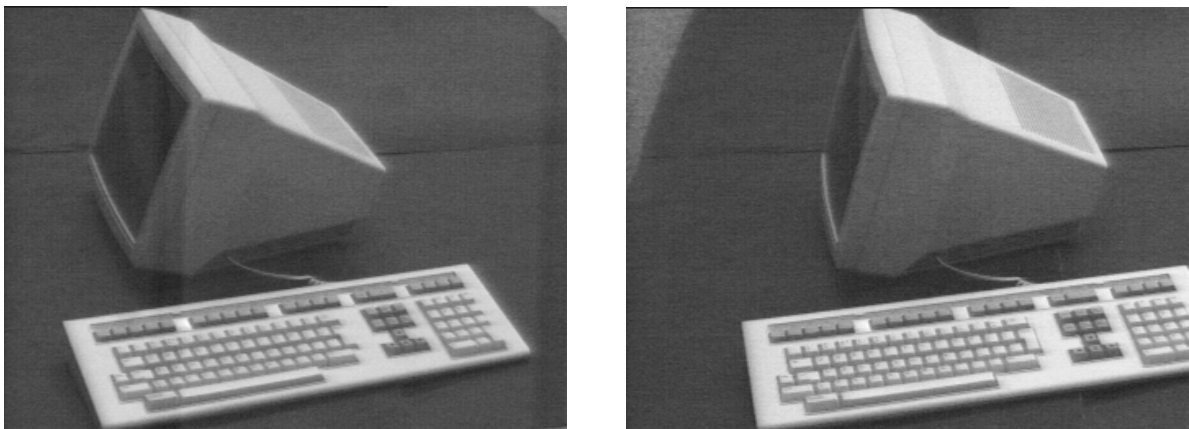


Abbildung 1.2 Stereobildaufnahme

2 Das Kameramodell

Um die Abbildung einer Szene auf die Bildebene der Kamera beschreiben zu können muß ein geeignetes Modell für die Kamera eingeführt werden. Somit läßt sich die Abbildung eines Szenenpunktes auf die Bildebene mathematisch beschreiben.

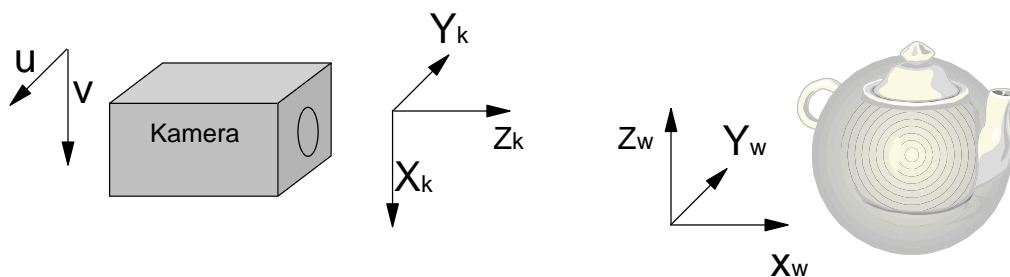


Abbildung 2.1 Bildaufnahme

2.1 Das Lochkameramodell

Bei der Lochkamera fallen durch ein unendlich kleines Loch die Lichtstrahlen aus der Welt ein. Diese projizieren die Szene spiegelverkehrt auf die Rückwand der Kamera, der Bildebene. Dieses ist ein idealisiertes Modell, im Realfall bündelt eine Linse die Lichtstrahlen und bildet diese auf einen Punkt ab. Dies geschieht nach den gleichen Gesetzmäßigkeiten, nur daß noch ein Abbildungsfehler entsteht.

Da es umständlich ist mit umgekehrten Objekten, wie sie das Lochkameramodell liefert, zu arbeiten, wird es durch das Mattscheibenmodell ersetzt. Das Mattscheibenmodell unterscheidet sich vom Lochkameramodell nur dadurch, daß die Bildebene vor dem Projektionszentrum liegt.

2.2 Die Abbildung Szene -> Bildebene

Die Abbildung eines Szenenpunktes kann nun aufgrund des Kameramodells mathematisch durch eine lineare Transformation beschrieben werden. Die Transformationsmatrix wird bei der Kamerakalibrierung bestimmt und beschreibt die geometrische Transformation, die Einflüsse der Zentralperspektive, und Abbildungseigenschaften der Kamera soweit dies mit einem linearen Modell möglich ist.

2.2.1 Die Zentralperspektive

Gegeben sei ein Kamerakoordinatensystem mit dem Ursprung im optischen Zentrum und einer Bildebene parallel zur x,y Ebene an der Stelle $z=c$.

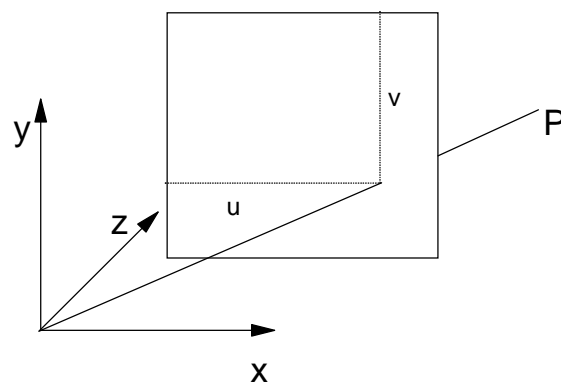


Abbildung 2.2 Das Mattscheibenmodell

Daraus ergeben sich folgende Abbildungsvorschriften für den Szenenpunkt P auf den Bildpunkt $I(u,v)$:

$$u = c \cdot \frac{x}{z} \quad \text{und} \quad v = c \cdot \frac{y}{z} \quad (2.1)$$

Gleichung 2.1 beschreibt die perspektivische Abbildung eines 3D Punktes auf die Bildebene. Wie man sieht werden die Objekte mit zunehmendem Abstand kleiner.

Um die geometrische Transformation vollständig zu beschreiben sind zudem noch die Rotation und die Translation notwendig.

2.2.2 Die Räumliche Transformation

Die bisher verwendeten Koordinaten werden nun zu homogenen Koordinaten erweitert. Dies geschieht durch Ergänzung um eine vierte bzw. eine dritte Komponente. Diese Komponente hat den Wert 1, der so entstandene Vektor kann nun mit einem beliebigen Skalar w multipliziert werden.

$$\text{Homogene Darstellung: } (X, Y, Z, w) = (x \cdot w, y \cdot w, z \cdot w, w) \quad (2.2)$$

2.2.2.1 Die Translation

Die Translation beschreibt die Verschiebung eines Punktes $P(x, y, z)$ um einen Vektor v . Der neu entstandene Punkt $P' = P + \begin{pmatrix} x_0 & y_0 & z_0 \end{pmatrix}$. Diese Verschiebung kann auch durch eine Matrixmultiplikation von homogenen Koordinaten beschrieben werden.

$$T = \begin{pmatrix} 1 & 0 & 0 & x_0 \\ 0 & 1 & 0 & y_0 \\ 0 & 0 & 1 & z_0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X' \\ Y' \\ Z' \\ w \end{pmatrix} = T \cdot \begin{pmatrix} X \\ Y \\ Z \\ w \end{pmatrix} \quad (2.3)$$

2.2.2.2 Die Rotation

Die dreidimensionale Rotation setzt sich aus den Drehungen um die 3 Achsen des kartesischen Koordinatensystems zusammen.

$$\text{Rotation um die x-Achse: } R_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \gamma & -\sin \gamma & 0 \\ 0 & \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\text{Rotation um die y-Achse: } R_y = \begin{bmatrix} \cos \beta & 0 & \sin \beta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \beta & 0 & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\text{Rotation um die z-Achse : } R_z = \begin{bmatrix} \cos a & -\sin a & 0 & 0 \\ \sin a & \cos a & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Die Rotationsmatrix für die gesamte Rotation erhält man durch Multiplikation der einzelnen Matrizen. Da die Matrixmultiplikation nicht kommutativ ist, muß die Reihenfolge der Drehungen festgelegt werden. Eine sinnvolle Reihenfolge ergibt sich aus der Anlehnung an die Kugelkoordinaten.

-Drehung um die z -Achse mit dem Winkel θ , neue Koordinaten (x', y', z')

-Drehung um die y' -Achse mit dem Winkel $(90 - \vartheta)$, neue Koordinaten (x'', y'', z'')

-Drehung um die x'' -Achse mit dem Winkel ψ

Die resultierende Matrix ergibt sich zu :

$$R = R_\theta \cdot R_\vartheta \cdot R_\alpha \tag{2.4}$$

2.2.3 Direkte Lineare Transformation

Die direkte lineare Transformation bildet einen Szenenpunkt $P(x,y,z,1)$ in homogenen Koordinaten auf den Bildpunkt $I(U,V,w)$ ebenfalls in homogenen Koordinaten ab.

$$\begin{bmatrix} U \\ V \\ w \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{2.5}$$

Die DLT-Matrix beinhaltet Informationen über die Orientierung des Kamerakoordinatensystems bezüglich des Weltkoordinatensystems, den Abstand der beiden Koordinatensysteme und über Abbildungseigenschaften der Linse. Der letzte Punkt stellt aber auch einen Schwachpunkt der DLT dar, da diese nur lineare Linsenfehler berücksichtigt.

2.3 Diskussion des Kameramodells

Das hier verwendete Kameramodell ist ein stark idealisiertes Modell, das die physikalische Wirklichkeit nur zum Teil widerspiegelt. Dieses Modell reicht allerdings für viele Anwendungen aus. Um die Abbildungseigenschaften einer Kamera vollständig zu beschreiben, müssen die Ungenauigkeiten die bei der Fertigung von handelsüblichen Kameras entstehen berücksichtigt werden. Zum Beispiel steht die Bildebene bei solchen Kameras in der Regel nicht senkrecht zur optischen Achse. Zur vollständigen Beschreibung muß ein Szenenpunkt P durch mehrere Transformationen auf die Bildebene abgebildet werden.

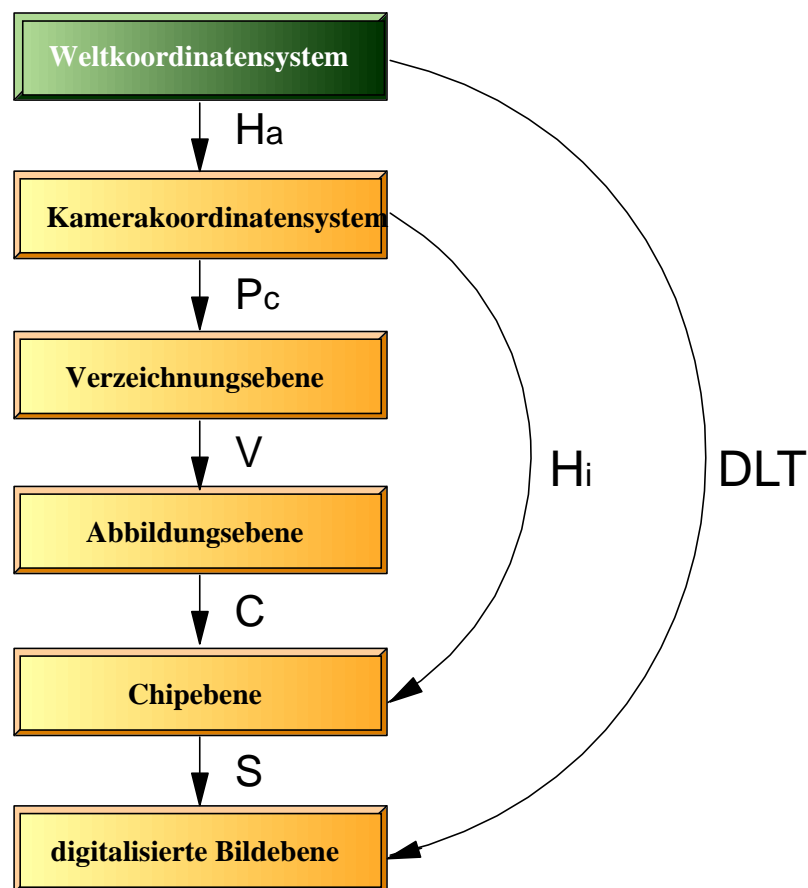


Abbildung 2.3 Abbildungen im vollständigen Kameramodell

Eine genaue Beschreibung des vollständigen Kameramodells wird in [Hetzel 94] vorgestellt.

3 Kalibrierung

Die Kalibrierung einer Kamera dient dazu, die verschiedenen Parameter, die zur Beschreibung der Abbildungsfunktion notwendig sind, zu bestimmen. Da hier mit der direkten linearen Transformation gearbeitet wird beschränkt sich die Kalibrierung auf die Bestimmung der zwölf Elemente der Transformationsmatrix. Um diese Parameter zu bestimmen werden Aufnahmen von einem genau vermessenen Referenzobjekt gemacht. Auf diesem Referenzobjekt befinden sich sechs helle nicht komplanare Punkte deren Koordinaten bekannt sind. Für jede Zuordnung eines Referenzpunktes mit einem Bildpunkt entstehen aus Gleichung 2.5 zwei Gleichungen :

$$u = \frac{m_{11}x+m_{12}y+m_{13}z+m_{14}}{m_{31}x+m_{32}y+m_{33}z+m_{34}} \quad (3.1)$$

$$v = \frac{m_{21}x+m_{22}y+m_{23}z+m_{24}}{m_{31}x+m_{32}y+m_{33}z+m_{34}} \quad (3.2)$$

Es entsteht also ein Gleichungssystem mit 12 Unbekannten und 12 Gleichungen. Um bei der Lösung dieses Gleichungssystems nicht die triviale Lösung zu erhalten, muß das Gleichungssystem eingeschränkt werden. Üblicherweise wird $m_{34} = 1$ gesetzt. Dies ist laut [Krauss 83] kein Problem, da m_{34} dem Abstand vom Kamerazentrum zum Szenenursprung entspricht.

Die exakte Lösung ist theoretisch kein Problem. Aufgrund der numerischen Ungenauigkeit werden allerdings oft wesentlich mehr Referenzpunkte genommen und das so entstehende überdefinierte Gleichungssystem wird z.B. nach der Methode des kleinsten Fehlerquadrates gelöst.

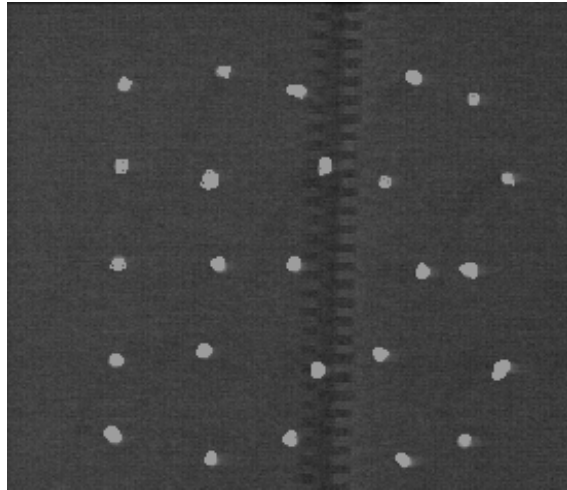


Abbildung 3.1 Aufnahme eines Referenzobjektes

4 Epipole und Epipolarlinien

Ein großes Problem bei der Auswertung und Rekonstruktion von Stereobildern ist es zu einem Punkt bzw. einer Kante im linken Bild den entsprechenden bzw. die entsprechende im rechten Bild zu finden. Dieses sogenannte Matching-Problem wird in einem späteren Kapitel noch ausführlich beschrieben. Eine große Hilfe bei der Lösung des Matching-Problems ist die Epipolar-Einschränkung. Durch sie wird der Suchraum auf eine Gerade im anderen Bild beschränkt.

Ist, wie in Abbildung 4.1 dargestellt, ein Bildpunkt I_1 gegeben kann der korrespondierende Punkt nur auf der Geraden D_{E_2} liegen. Die Gerade D_{E_2} ist der Schnitt der Ebene, die durch C_1 , C_2 und I_1 aufgespannt wird, mit der Bildfläche B_2 .

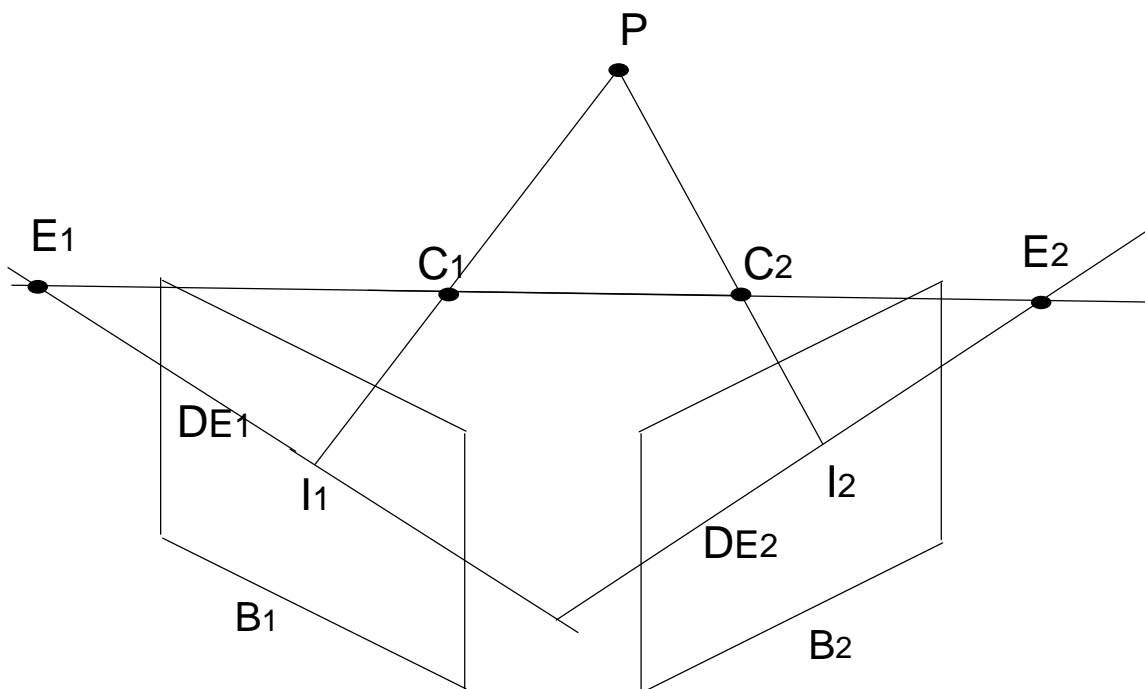


Abbildung 4.1 Geometrie der Epipolarlinien

Die folgenden Berechnungen beziehen sich nur auf das linke Bild. Analog hierzu kann die Berechnung für das rechte Bild durchgeführt werden.

4.1 Der Epipol

Der Epipol ist der Punkt in dem sich alle Epipolarlinien schneiden. Er kann endliche Koordinaten haben oder er liegt im Unendlichen. Im zweiten Fall laufen dann alle Epipolarlinien in dem entsprechenden Bild parallel.

Der Epipol der linken Bildebene E_1 ist das Bild des optischen Zentrums C_2 der rechten Kamera. Die Berechnung von E_1 geschieht also in zwei Schritten :

1) Bestimmung des optischen Zentrums C_2 der rechten Kamera :

$$C_2 = (x_{c2}, y_{c2}, z_{c2})$$

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = M^2 \cdot \begin{bmatrix} x_{c2} \\ y_{c2} \\ z_{c2} \\ 1 \end{bmatrix} \quad (4.1)$$

Aus dem lineare Gleichungssystem 4.1 lassen sich die Koordinaten des optischen Zentrums berechnen.

2) Abbildung des optischen Zentrums C_2 auf die Bildebene B_1 :

Aus Gleichung 2.5 folgt :

$$E_1^* = M^1 \cdot \begin{bmatrix} x_{c2} \\ y_{c2} \\ z_{c2} \\ 1 \end{bmatrix} \quad (4.2)$$

$$E_1^* = \begin{bmatrix} U_{E1} \\ V_{E1} \\ w_{E1} \end{bmatrix}$$

Ist w_{E1} gleich Null so liegt der Epipol im Unendlichen, ansonsten hat er die Bildkoordinaten:

$$E_1 = \begin{bmatrix} U_{E1}/w_{E1} \\ V_{E1}/w_{E1} \end{bmatrix}$$

4.2 Die Epipolarlinien

Die Epipolarlinie im linken Bild berechnet sich aus dem Bildpunkt I_2 im rechten Bild und dem optischen Zentrum C_2 der rechten Kamera. Die Epipolarlinie D_{E1} ist das Bild der Geraden $\overline{I_2 C_2}$ auf der linken Bildebene. Auf dieser Epipolarlinie muß sich der Punkt I_1 befinden.

Aus Gleichung 3.1 und Gleichung 3.2 folgt :

$$u_2 \cdot (M_3^2 \cdot P + m_{34}^2) = M_1^2 \cdot P + m_{14}^2 \quad (4.3)$$

$$v_2 \cdot (M_3^2 \cdot P + m_{34}^2) = M_2^2 \cdot P + m_{24}^2 \quad (4.4)$$

mit

$$M_i^j = (m_{i1}^j, m_{i2}^j, m_{i3}^j)$$

$$P = (x, y, z, 1)^T$$

wobei j der Index der Kamera und i die Zeilennummer der entsprechenden DLT Matrix ist.

Gleichung 4.3 und 4.4 lassen sich Umformen in Ebenengleichungen in Normalform :

$$(M_1^2 - u_2 \cdot M_3^2) \cdot P - u_2 \cdot m_{34}^2 + m_{14}^2 = 0 \quad (4.5)$$

$$(M_2^2 - v_2 \cdot M_3^2) \cdot P - v_2 \cdot m_{34}^2 + m_{24}^2 = 0 \quad (4.6)$$

Dies sind die Gleichungen zweier Ebenen deren Schnitt die Gerade $\overline{I_2 C_2}$ ist. Der Richtungsvektor dieser Geraden ist das Kreuzprodukt der Normalenvektoren der Ebenen.

$$n_2 = (M_1^2 - u_2 \cdot M_3^2) \times (M_2^2 - v_2 \cdot M_3^2) \quad (4.7)$$

Die Geradengleichung für $\overline{I_2 C_2}$ lautet folglich :

$$g_2 = C_2 + \lambda \cdot n_2 \quad (4.8)$$

Die Abbildung dieser Geraden ist die Epipolarlinie D_{E1}^* in homogenen Koordinaten.

$$D_{E1}^* = M^1 \cdot \begin{bmatrix} C_2 + \lambda \cdot n_2 \\ 1 \end{bmatrix} = M^1 \cdot \begin{bmatrix} C_2 \\ 1 \end{bmatrix} + \begin{bmatrix} \lambda \cdot n_2 \\ 0 \end{bmatrix}$$

Mit Gleichung 4.1 folgt :

$$D_{E1}^* = E_1^* + M^1 \cdot \begin{bmatrix} \lambda \cdot n_2 \\ 0 \end{bmatrix} \quad (4.9)$$

Die homogenen Koordinaten von D_{E1}^* lauten :

$$U_1 = M_1^1 \cdot C_2 + m_{14}^1 + M_1^1 \cdot \lambda \cdot n_2$$

$$V_1 = M_2^1 \cdot C_2 + m_{24}^1 + M_2^1 \cdot \lambda \cdot n_2$$

$$w = M_3^1 \cdot C_2 + m_{34}^1 + M_3^1 \cdot \lambda \cdot n_2$$

Hieraus ergeben sich die kartesischen Koordinaten :

$$u_1 = \frac{M_1^1 \cdot C_2 + m_{14}^1 + M_1^1 \cdot \lambda \cdot n_2}{M_3^1 \cdot C_2 + m_{34}^1 + M_3^1 \cdot \lambda \cdot n_2} = \frac{U_{E1} + \lambda \cdot U_{F1}}{S_{E1} + \lambda \cdot S_{F1}} \quad (4.10)$$

$$v_1 = \frac{M_2^1 \cdot C_2 + m_{24}^1 + M_2^1 \cdot \lambda \cdot n_2}{M_3^1 \cdot C_2 + m_{34}^1 + M_3^1 \cdot \lambda \cdot n_2} = \frac{V_{E1} + \lambda \cdot V_{F1}}{S_{E1} + \lambda \cdot S_{F1}} \quad (4.11)$$

Nach der Kalibrierung lassen sich die Epipole berechnen. Nun ist bekannt ob sie im endlichen oder im unendlichen liegen. Liegt der Epipol im endlichen muß nur noch der Richtungsvektor der Epipolarlinie bestimmt werden. Dieses läßt sich durch eine Transformation realisieren. Die Elemente dieser Matrix sind nicht vom Bildpunkt abhängig und können somit nach der Kalibrierung berechnet werden.

mit

$$\begin{bmatrix} \Delta u_1 \\ \Delta v_1 \end{bmatrix} = T_1 \cdot \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} \quad (4.12)$$

$$T_1 = \begin{bmatrix} S_{E1} & 0 & -U_{E1} \\ 0 & S_{E1} & -V_{E1} \end{bmatrix} \cdot \left(\begin{bmatrix} M_1^1 \\ M_2^1 \\ M_3^1 \end{bmatrix} \cdot \begin{bmatrix} M_2^2 \times M_3^2 \\ M_3^2 \times M_1^2 \\ M_1^2 \times M_2^2 \end{bmatrix} \right)^T$$

Liegt der Epipol im unendlichen muß für die Epipolarlinie der Richtungsvektor und der vanishing Point bestimmt werden. Der vanishing Point ist ein Punkt durch den die Epipolarlinie D_{E1} läuft. Der Richtungsvektor läßt sich direkt nach der Kalibrierung bestimmen.

$$\begin{bmatrix} \Delta u_1 \\ \Delta v_1 \end{bmatrix} = \begin{bmatrix} U_{E1} \\ V_{E1} \end{bmatrix} \quad (4.13)$$

Der vanishing Point berechnet sich wie folgt :

$$F_1^* = N_1 \cdot \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} \quad (4.14)$$

mit

$$N_2 = \begin{bmatrix} M_1^1 \\ M_2^1 \\ M_3^1 \end{bmatrix} \cdot \begin{bmatrix} M_2^2 \times M_3^2 & M_3^2 \times M_1^2 & M_1^2 \times M_2^2 \end{bmatrix}$$

In Anwendungen empfiehlt es sich allerdings, aufgrund der Ungenauigkeiten bei der Kalibrierung, mit Epipolarbändern zu arbeiten. Je nach Qualität der Kalibrierung sollte diese Bänder eine Breite von 2 bis 6 Pixel haben.

4.3 Beispiele von Epipolarlinien

Abbildung 4.2 zeigt die Epipolarlinien der Anfangs und Endpunkte der gestrichelten Kanten.

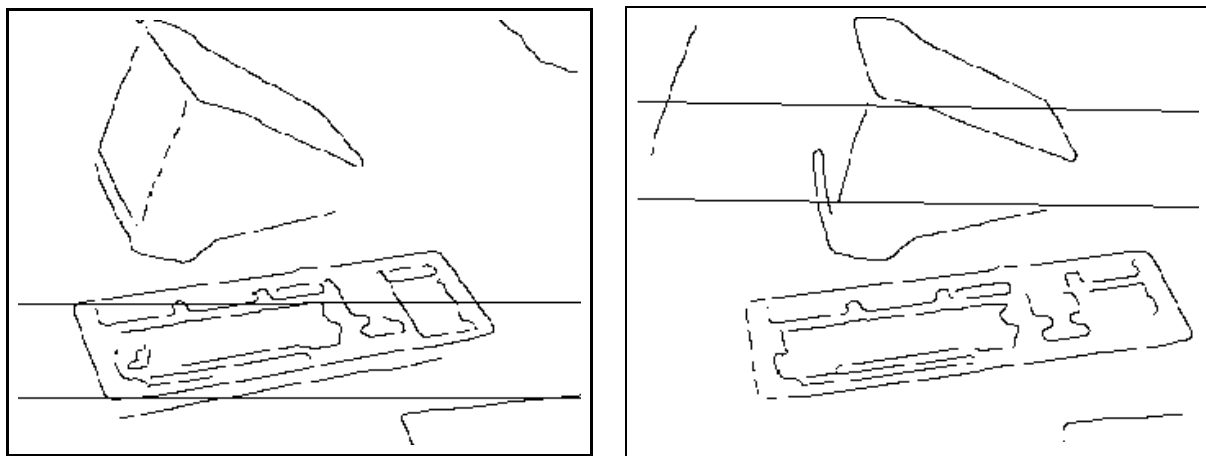


Abbildung 4.2 Epipolarlinien von Kanten

5 Matching

Bevor aus zwei Bildern die Tiefeninformation berechnet werden kann, müssen die Elemente des einen Bildes den Elementen des anderen zugeordnet werden. Diese Zuordnung wird als Matching bezeichnet. Als Bildelemente kommen Pixel, Kanten, Segmente oder Konturen in Frage. Je nach Anwendung und Rechenleistung muß die Wahl der Bildelemente getroffen werden. Es ist offensichtlich, daß die Zuordnung von einzelnen Pixel eine wesentlich höhere Komplexität als die Zuordnung von linearen Kantensegmenten hat. Zudem sollten die Bildelemente folgende Eigenschaften haben :

- Ein Element sollte so viele Informationen wie möglich beinhalten
- Invarianz gegenüber verschiedenen Blickwinkeln
- Kleinere Bildstörungen sollten nicht ins Gewicht fallen
- Alle Objekte einer Szene sollten so repräsentiert werden können

Das im Folgenden beschriebene Verfahren beschränkt sich auf die Zuordnung von linearen Kantensegmenten.

5.1 Eigenschaften von Kanten

Aufgrund der Merkmale, die sich in der Umgebung einer Kante befinden, lassen sich eine Vielzahl von Kanten im anderen Bild ausschließen ohne großen Rechenaufwand zu betreiben. Verwendete Eigenschaften :

- 1) Länge der Kante
- 2) Gemittelter Grauwertgradient entlang der Kante
- 3) Orientierung der Kante (Winkel zur Horizontalen)

Beim Vergleich der Eigenschaften von zwei Kanten in verschiedenen Bildern muß bedacht werden, daß die Eigenschaften nicht exakt gleich sind. Für die Auswertung müssen geeignete Toleranzen eingestellt werden. Die Längen der Kanten können unterschiedlich sein, da z.B. das Pendant einer Kante im anderen Bild durch zwei Kanten gegeben ist (Kantenabriß). Aufgrund einer Längendifferenz kommt es auch zu unterschiedlichen gemittelten Grauwertgradienten. Hier empfiehlt es sich die Toleranz des Grauwertgradienten von der Längendifferenz abhängig zu machen. Die Orientierung zweier zusammengehörenden Kanten wird von den verschiedenen Orientierungen der Kameras beeinflußt. Es ist hier natürlich möglich eine Toleranz aus den DLT-Matrizen der Kameras zu berechnen. Andererseits hat [Arnold et al. 80] gezeigt, daß bei einer Kameraausrichtung, wie sie üblicherweise bei Stereokameras für mobile Roboter verwendet wird, die Wahrscheinlichkeit für eine Winkeldifferenz von mehr als 9° kleiner 0.5 ist. Somit werden mit einer Toleranz von 10° - 20° auch sehr gute Ergebnisse erzielt.

Eine weitere Eigenschaft die öfters verwendet wird, ist die Nachbarschaft einer Kante. D.h. eine Kante wird mit den Kanten die sie umgeben in Relation gesetzt. Wie in Abbildung 5.1 gezeigt wird, liegt Kante 1 links von Kante 2, Kante 2 liegt rechts von Kante 1 und links von Kante 3 usw. Leider ist die Nachbarschaftsrelation von der Orientierung der Kameras abhängig und die Beziehungen wie sie in Bild 1 gelten müssen in Bild 2 nicht die gleichen sein.

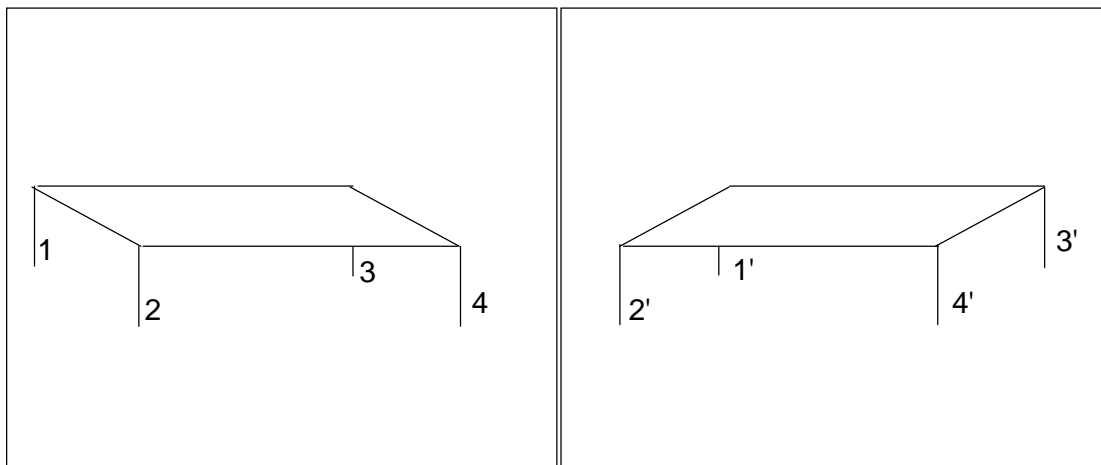


Abbildung 5.1 Widerspruch bei der Nachbarschaftsrelation

5.2 Die Epipolar-Einschränkung

Da die Koordinaten der Anfangs bzw. Endpunkte einer Kante, bei nicht parallelen optischen Achsen, nicht direkt miteinander verglichen werden können, wird die Epipolar-Einschränkung verwendet. Mit ihrer Hilfe können Kante gefunden werden deren Position es Wahrscheinlich macht, daß sie zu der entsprechenden Kante im anderen Bild gehören. Leider läßt sich dieses Kriterium nur auf die y-Position anwenden. Parallel verschobene Kanten können mit dieser Einschränkung nicht ausgeschlossen werden.

Ausgehend von einer Kante wird zu ihrem Mittelpunkt die Epipolarlinie im anderen Bild berechnet. Alle Kanten die diese Epipolarlinie schneiden erfüllen die Epipolar-Einschränkung. An dieser Stelle wird noch nicht überprüft, ob der Schnittpunkt in der Nähe des Mittelpunktes der geschnittenen Kante liegt, da ansonsten Kanten die abgerissen sind nicht gematched werden könnten. Und es ist wünschenswert, daß bei einer abgerissenen Kante zumindest ein Teil gematched wird.

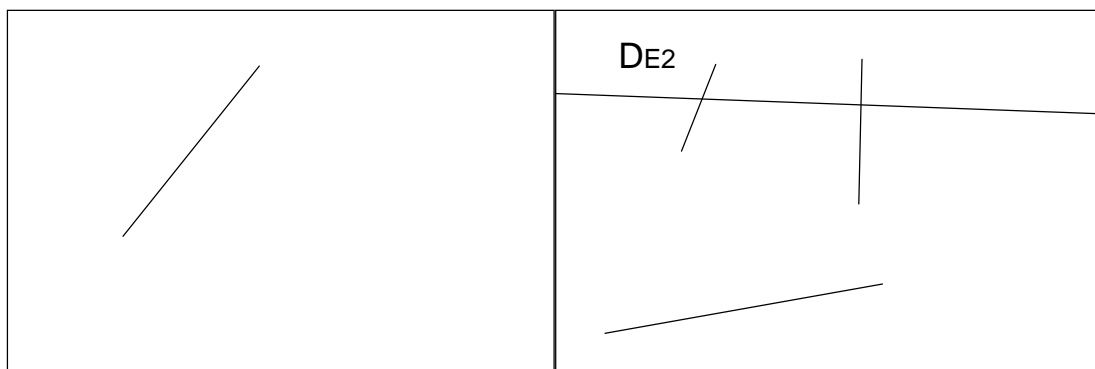


Abbildung 5.2 Epipolar-Einschränkung

Es hat sich gezeigt, daß bei Bildern in denen wenig parallel verschobene Kanten existieren, mit Hilfe der Eigenschaften von Kanten und der Epipolar-Einschränkung, die in beide Richtungen angewandt wird, schon einige Kanten eindeutig zugeordnet werden können. Dies hat den Vorteil, daß in den Folgenden Schritten des Matching, die wesentlich rechenintensiver sind, nicht mehr so viele Kanten zugeordnet werden müssen.

5.3 Der Matching-Algorithmus

Der Matching-Algorithmus ordnet einer gegebenen Kante K die korrespondierende Kante aus der Kantenmenge KM zu.

$$\text{Matching} : \{K\} \times KM \rightarrow \{\text{True}, \text{False}\}$$

Der hier verwendete Algorithmus, der in Abbildung 5.4 dargestellt ist, löst dieses Problem in drei Schritten. Als erstes wird aufgrund der Eigenschaften der Kante K und der Epipolar-Einschränkung die Menge PM der in Frage kommenden Kanten (Possible Matches) gebildet. Nun wird für jede Kante K' aus PM ein Backmatch durchgeführt. D.h. es wird erneut aufgrund der Eigenschaften und der Epipolar-Einschränkung von K' eine Kantenmenge BM gebildet, diesmal allerdings in umgekehrter Richtung. Die Kante K' bleibt in der Menge der möglichen Kanten PM , falls die ursprüngliche Kante K in der Backmatchmenge BM enthalten ist. Jetzt besteht die Möglichkeit, daß beide Mengen nur noch ein Element enthalten. In diesem Fall kann der Algorithmus abgebrochen werden, da eine eindeutige Zuordnung vorliegt. Der Algorithmus wird ebenfalls nicht weiter ausgeführt wenn BM kein Element enthält, da nun keine Zuordnung mehr möglich ist. Stehen nun noch mehrere Kanten zur Auswahl wird im dritten Schritt erneut mit Hilfe der Epipolarlinien eine Auswahl getroffen. Jetzt werden die Epipolarlinien des Anfangs und des Endpunktes der Kante K gebildet. Diese werden mit den Kanten die noch in PM verbleiben geschnitten. Schneidet eine der Epipolarlinien eine Kante K' aus PM genau im entsprechenden Punkt und ist der Schnitt mit den anderen Kanten aus PM qualitativ schlechter so wird K' als Pendant ausgewählt. Aufgrund der möglichen Längendifferenzen und Kantenabrisse reicht hier ein exakter Schnitt mit einer der Epipolarlinie.

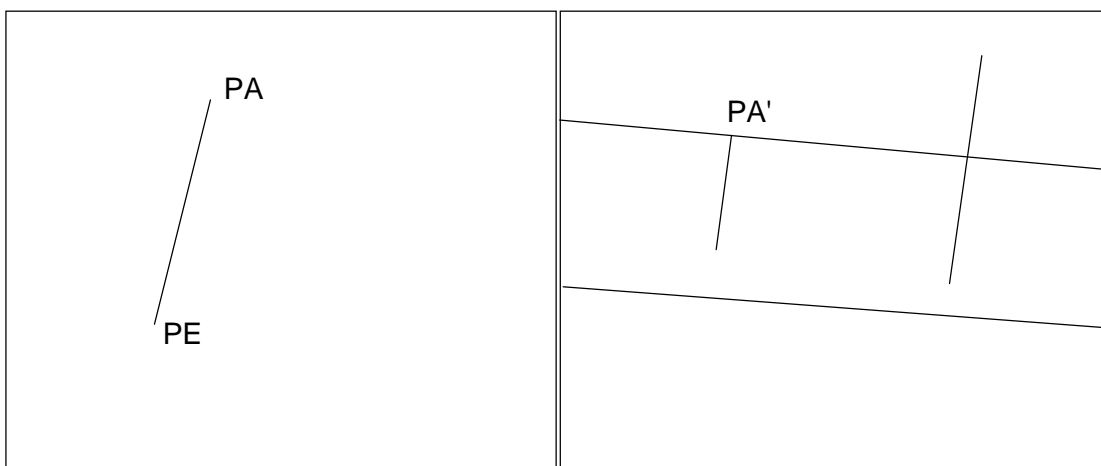


Abbildung 5.3 Problem des Kantenabrisse

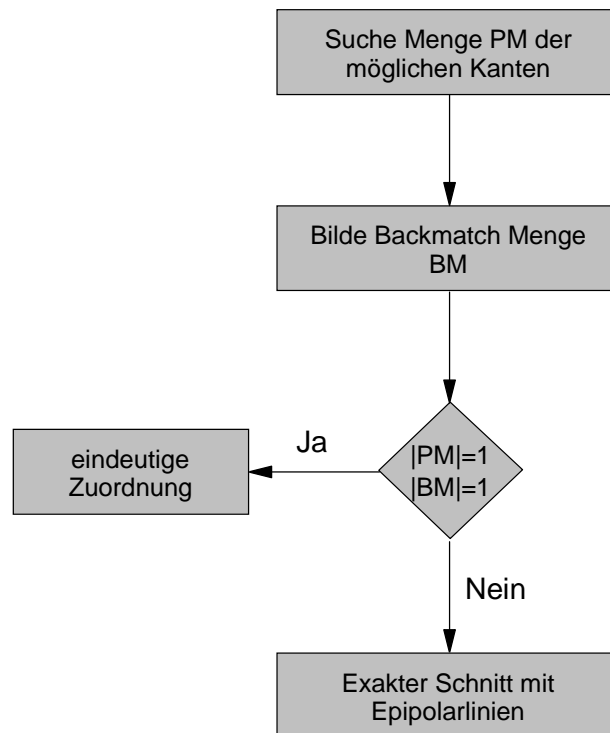


Abbildung 5.4 Der Matching-Algorithmus

5.4 Berechnung des Überlappenden Teils zweier Kanten

Da nicht gewährleistet ist, daß zwei korrespondierende Kanten in verschiedenen Bildern immer die gleiche Länge haben, dürfen die entsprechenden Anfangs und Endpunkte nicht zur Berechnung der Szenenkoordinaten herangezogen werden. Um die korrespondierenden Anfangs und Endpunkte zu erhalten werden erneut die Epipolarlinien zu Hilfe genommen.

Ausgehend von den Punkten I_1 und I_2 im linken Bild werden die Epipolarlinie D_{E2}^1 und D_{E2}^2 berechnet. Sie werden mit der korrespondierenden Kante geschnitten. Die ursprüngliche Kante wird nun durch den Teil von ihr ersetzt, der zwischen den Epipolarlinien D_{E2}^1 und D_{E2}^2 liegt.

Jetzt wird ausgehend von den neu entstandenen Anfangs und Endpunkten I'_1 und I'_2 die Epipolarlinien im linken Bild berechnet. Auch hier wird nun die original Kante durch den Teil von ihr ersetzt der zwischen den Epipolarlinien liegt. Abbildung 5.5 veranschaulicht diese Vorgehensweise.

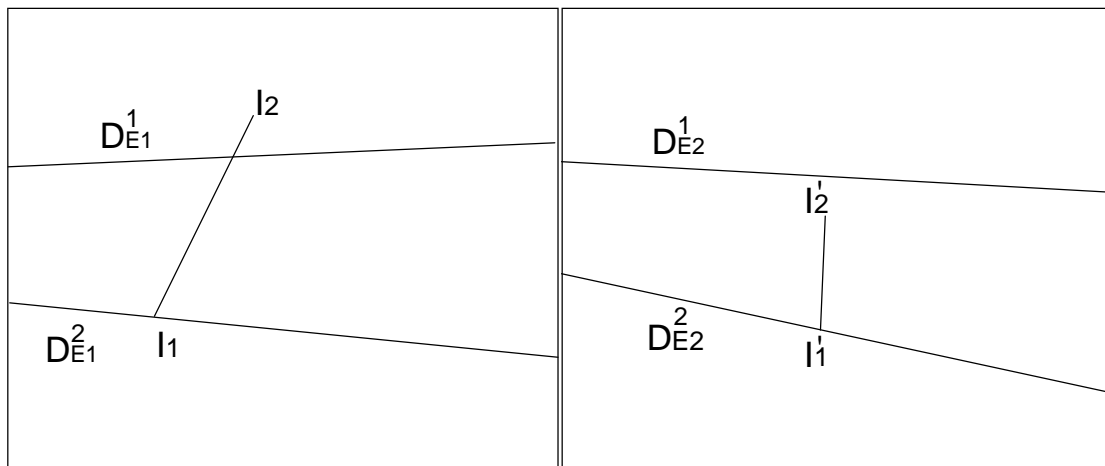


Abbildung 5.5 Der überlappende Teil zweier korrespondierender Kanten

5.5 Beispiele

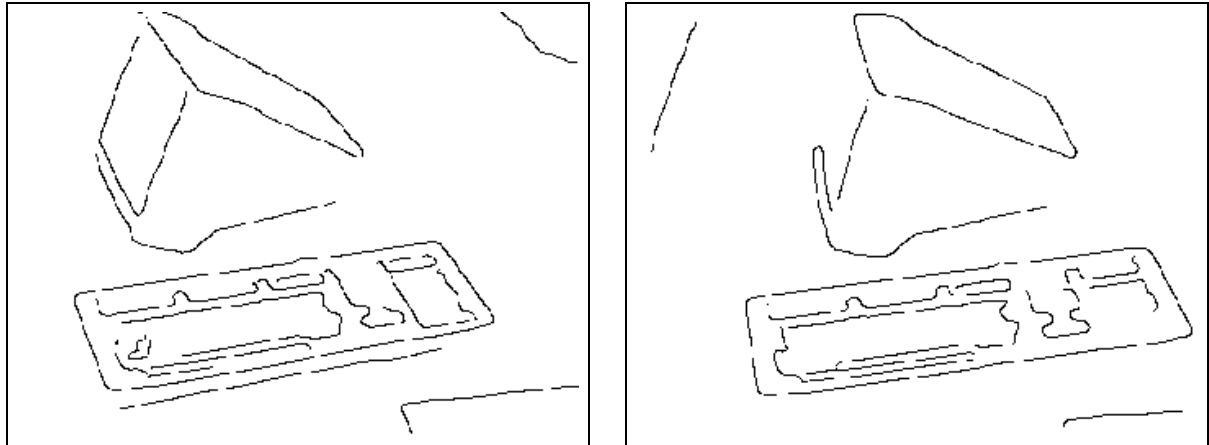


Abbildung 5.6 Kantenbilder

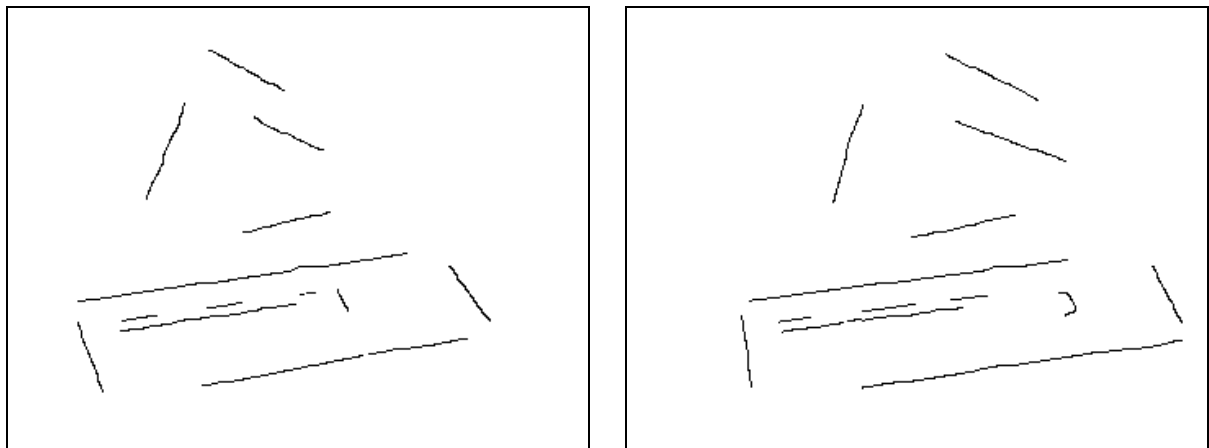


Abbildung 5.7 Überlappender Teil der zugeordnete Kanten

6 Rekonstruktion

Nachdem nun das Korrespondenzproblem gelöst ist, können aus dem überlappenden Teil der zugeordneten Kanten die 3D Szenenkoordinaten berechnet werden. Wurden lineare Kantensegmente als Bildelemente ausgewählt genügt es die Rückrechnungen für den Anfangs und Endpunkt einer Kante durchzuführen. Da die Information für die Berechnungen aus den Kalibriermatrizen M^1 und M^2 bezogen wird, beziehen sich die so gewonnenen Koordinaten auch auf das Koordinatensystem bezüglich dem die Kameras kalibriert wurden.

Aus Gleichung 2.5 folgt für die linke Kamera :

$$u_1 = \frac{M_1^1 \cdot P + m_{14}^1}{M_3^1 \cdot P + m_{34}^1}$$

$$v_1 = \frac{M_2^1 \cdot P + m_{24}^1}{M_3^1 \cdot P + m_{34}^1}$$

und für die rechte :

$$u_2 = \frac{M_1^2 \cdot P + m_{14}^2}{M_3^2 \cdot P + m_{34}^2}$$

$$v_2 = \frac{M_2^2 \cdot P + m_{24}^2}{M_3^2 \cdot P + m_{34}^2}$$

Die Koordinaten der Bildpunkte u_1, v_1, u_2 und v_2 sind bekannt. Die Gleichungen lassen sich zu einem linearen Gleichungssystem umschreiben.

$$u_1(M_3^1 \cdot P + m_{34}^1) - (M_1^1 + m_{14}^1) = 0 \quad (6.1)$$

$$v_1(M_3^1 \cdot P + m_{34}^1) - (M_2^1 + m_{24}^1) = 0 \quad (6.2)$$

$$u_2(M_3^2 \cdot P + m_{34}^2) - (M_1^2 + m_{14}^2) = 0 \quad (6.3)$$

$$v_2(M_3^2 \cdot P + m_{34}^2) - (M_2^2 + m_{24}^2) = 0 \quad (6.4)$$

Es gilt nun die Koordinaten von P zu bestimmen. Da es sich hierbei um ein überbestimmtes Gleichungssystem handelt, kann es entweder durch das Weglassen einer beliebigen Gleichung gelöst werden oder es wird die Lösung mit dem kleinsten Fehlerquadrat gesucht.

6.1 Rekonstruktion durch Weglassen einer Gleichung

Der einfachste Weg zur Lösung dieses Gleichungssystems ist das Weglassen einer beliebigen Gleichung. Dies entspricht dem Schnitt einer Geraden mit einer Ebene. Der Schnittpunkt ist der gesuchte Szenenpunkt. Wird beispielsweise Gleichung 6.4 weggelassen so ergibt sich das eindeutig bestimmte Gleichungssystem :

$$\begin{bmatrix} m_{31}^1 u_1 - m_{11}^1 & m_{32}^1 u_1 - m_{12}^1 & m_{33}^1 u_1 - m_{13}^1 \\ m_{31}^1 v_1 - m_{21}^1 & m_{32}^1 v_1 - m_{22}^1 & m_{33}^1 v_1 - m_{23}^1 \\ m_{31}^2 u_2 - m_{11}^2 & m_{32}^2 u_2 - m_{12}^2 & m_{33}^2 u_2 - m_{13}^2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} m_{14}^1 - m_{34}^1 u_1 \\ m_{24}^1 - m_{34}^1 v_1 \\ m_{14}^2 - m_{34}^2 u_2 \end{bmatrix} \quad (6.5)$$

6.2 Rekonstruktion nach der Methode des kleinsten Fehlerquadrates

Durch die Methode des kleinsten Fehlerquadrates läßt sich eine Lösung für ein überbestimmtes Gleichungssystem finden, die dieses am günstigsten erfüllt.

Es gelten folgende Gleichungen :

$$G_1 : u_1(M_3^1 \cdot P + m_{34}^1) - (M_1^1 + m_{14}^1) = 0 \quad (6.6)$$

$$G_2 : v_1(M_3^1 \cdot P + m_{34}^1) - (M_2^1 + m_{24}^1) = 0 \quad (6.7)$$

$$G_3 : u_2(M_3^2 \cdot P + m_{34}^2) - (M_1^2 + m_{14}^2) = 0 \quad (6.8)$$

$$G_4 : v_2(M_3^2 \cdot P + m_{34}^2) - (M_2^2 + m_{24}^2) = 0 \quad (6.9)$$

Diese vier Gleichungen werden quadriert und addiert :

$$G = G_1^2 + G_2^2 + G_3^2 + G_4^2 \quad (6.10)$$

Es wird nicht möglich sein, eine Lösung für Gleichung 6.10 zu finden, so daß $G=0$ ist. Statt dessen wird das Minimum der Funktion $G(x,y,z)$ gesucht. Das Minimum einer Funktion mit mehreren Veränderlichen wird aus den partiellen Ableitungen bestimmt.

$$G_x = \frac{\delta G}{\delta x} \quad (6.11)$$

$$G_y = \frac{\delta G}{\delta y} \quad (6.12)$$

$$G_z = \frac{\delta G}{\delta z} \quad (6.13)$$

Die Gleichungen 6.11 bis 6.13 bilden ein eindeutig lösbares Gleichungssystem, das z.B. mit dem Gauß-Algorithmus gelöst werden kann. Die Lösung entspricht dem rekonstruierten Szenenpunkt P.

6.3 Beispiel

In dem in Abbildung 6.1 gezeigten Beispiel stehen an den Kanten die berechneten Längen in cm. Die Rekonstruktion dieses Beispiels wurde mit Gleichung 6.5 durchgeführt.

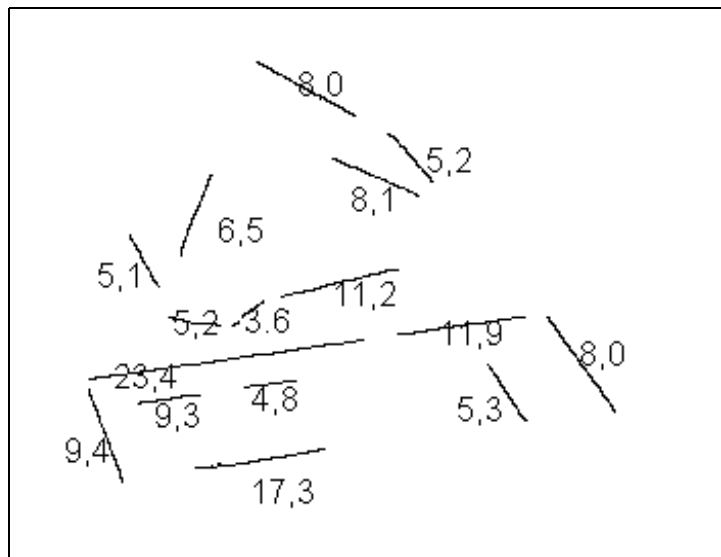


Abbildung 6.1 Absolute Längen der rekonstruierten Kanten

7 Die Software

Um die Funktionsweisen der Stereobildverarbeitung zu veranschaulichen wurde unter OSF Motif eine graphische Benutzeroberfläche entwickelt, in die Funktionen zur Kalibrierung, Bildvorverarbeitung, Matching und Rekonstruktion integriert wurden.



Abbildung 7.1 Die Software Cgui

7.1 Die Kalibrierung

Mit Hilfe von Paßpunktaufnahmen lassen sich die DLT-Matrizen berechnen. Diese können dann abgespeichert werden. Der hier verwendete Kalibrier-Algorithmus ist in [Hetzel 94] beschrieben. Werden später Bilder geladen muß das System nicht neu kalibriert werden, sondern die Matrizen können nachgeladen werden.

7.2 Die Bildvorverarbeitung

Zur Extraktion der Kanten steht eine Kombination aus Gauß und Canny Filter zur Verfügung. Mit Hilfe des Gauß-Filters lassen sich in Abhängigkeit des Parameters Sigma unwesentliche Details aus den Bildern herausfiltern. Der Canny-Filter detektiert dann die Kanten. Hier kann auch nochmals über eine Mindestlänge, die eine Kante haben muß um akzeptiert zu werden, Einfluß auf die Anzahl der entstehenden Kanten genommen werden.

7.3 Matching

Der für das Matching verwendete Algorithmus ist in Kapitel 5 beschrieben. Die Kanten lassen sich einzeln selektieren. Für jede Kante lassen sich die Epipolarlinien und ihre Eigenschaften anzeigen. Es besteht die Möglichkeit sich auch die Mengen der möglichen Matches (PM) und der Backmatches (BM) visualisieren zu lassen. Des weiteren können Kanten entweder einzeln zugeordnet werden oder die Zuordnung erfolgt über alle vorhandenen Kanten. Im zweiten Fall werden nur noch die überlappenden Teile der Kanten eingezeichnet. Sämtliche Toleranzen die verwendet werden lassen sich einstellen um so ihren Einfluß ausprobieren zu können.

7.4 Die Rekonstruktion

Um die Szenenkoordinaten eines Punktes zu berechnen wird das Gleichungssystem 6.5 gelöst. Die so erhaltenen Ergebnisse lassen sich entweder durch einen Grundrißplan oder durch Ausgabe der Koordinaten bezüglich des Objektkoordinatensystems veranschaulichen.

8 Diskussion der Ergebnisse

In dieser Arbeit wurde die Vorgehensweise für ein kantenbasiertes Stereoverfahren beschrieben. Wie leicht zu erkennen ist, hängen die Ergebnisse stark von der Kamerakalibrierung und von der Übereinstimmung des idealisierten Kameramodells mit der realen Kamera ab. Je besser die Übereinstimmungen in diesem Punkt sind, desto besser werden auch die Ergebnisse sein.

Eine weitere Qualitätssteigerung ist durch den Einsatz einer dritten Kamera möglich. Ein solches Verfahren ist in [AYACHE 91] beschrieben. Durch den Einsatz einer dritten Kamera würde sich für die Rekonstruktion eines Szenenpunktes ein Gleichungssystem mit drei Unbekannten und 6 Gleichungen ergeben. Wird ein solches System mit der Methode des kleinsten Fehlerquadrates gelöst, können Ungenauigkeiten in den DLT-Matrizen wesentlich besser ausgeglichen werden. Die hier gezeigten Ergebnisse könnten durch die Rekonstruktion nach der Methode des kleinsten Fehlerquadrates auch noch verbessert werden.

Der hier verwendete Matching-Algorithmus liefert ganz gute Ergebnisse. Probleme gibt es allerdings wenn viele vertikale Kanten gleicher Länge parallel nebeneinander liegen und sich ihre Umgebung kaum unterscheidet. Für einen solchen Fall reicht die Epipolar-Einschränkung nicht aus. Hier wäre eine Konfliktauflösung mit Hilfe der Nachbarschaftsrelation sehr vorteilhaft.

Abbildungsverzeichnis

1.1 Stereobildverarbeitung.....	5
1.2 Stereobildaufnahme.....	6
2.1 Bildaufnahme.....	7
2.2 Das Mattscheibenmodell.....	8
2.3 Abbildungen im vollständigen Kameramodell.....	11
3.1 Aufnahme eines Referenzobjektes.....	13
4.1 Geometrie der Epipolarlinien.....	14
4.2 Epipolarlinien von Kanten.....	20
5.1 Widerspruch bei der Nachbarschaftsrelation.....	22
5.2 Epipolar-Einschränkung.....	23
5.3 Problem des Kantenabrisses.....	24
5.4 Der Matching-Algorithmus.....	25
5.5 Der Überlappende Teil zweier korrespondierender Kanten.....	26
5.6 Kantenbilder.....	27
5.7 Überlappender Teil der zugeordneten Kanten.....	27
6.1 Absolute Länge der rekonstruierten Kanten.....	31
7.1 Die Software CGui.....	32

Literaturverzeichnis

- [Arnold et al. 80] R.D. Arnold, T.O. Binford: Geometric constraints in stereovision; In SPIE 238; Seite 281-292; San Diego 1980.
- [Ayache 91] Nicholas Ayache Artificial Vision for Mobile Robots: Stereo Vision Multisensory Perception; MIT Press; Cambridge, Massachusetts 1991.
- [Bräunl et al. 95] Thomas Bräunl, Stefan Feyrer, Wolfgang Rapf, Michael Reinhardt : Parallele Bildverarbeitung; Addison-Wesley; Deutschland 1995.
- [Foehr 90] R. Foehr: Photogrammetrische Erfassung räumlicher Informationen aus Videobildern; Fortschritt der Robotik Vol. 7; W. Ameling und M. Weck; 1990.
- [Hetzel 94] Günter Hetzel: Kalibrierung eines Stereokamerasystems; Diplomarbeit, Institut für Prozeßrechen-technik und Robotik, Universität Karlsruhe, 1994.
- [Posch 89] Stefan Posch: Automatische Bestimmung von Tiefeninformation aus Grauwert-Stereobildern; Doktorarbeit, Technische Fakultät der Universität Erlangen-Nürnberg, 1989.
- [Weckesser 93] Peter Weckesser: Entwicklung und Aufbau des aktiven Echtzeit Stereo Sichtsystems Kastor zur Navigation eines mobilen Robotersystems; Diplomarbeit, Institut für Prozeßrechen-technik und Robotik, Universität Karlsruhe, 1993.

Erklärung :

Ich versichere, daß ich diese Arbeit selbstständig verfaßt und nur die angegebenen Quellen und Hilfsmittel verwendet habe.

Stuttgart, den 21 Juli 1997