

Liste der noch zu erledigenden Punkte

abstract schreiben	V
einleitung schreiben	1
ist noch nicht fertig	8
etwas mehr ausführen	17
analyse einleitung schreiben	33
umbenennen	36
Literatur überprüfen, insbesondere Seitenangaben	41
muss das hier überhaupt rein?	43



ENTZERRUNG VON KEGELOBERFLÄCHEN AUS EINER EINKAMERAANSICHT BASIEREND AUF PROJEKTIVER GEOMETRIE

BACHELORARBEIT
zur Erlangung des akademischen Grades
BACHELOR OF SCIENCE

Westfälische Wilhelms-Universität Münster
Fachbereich Mathematik und Informatik
Institut für Informatik

Betreuung:
Dimitri Berh

Erstgutachten:
Prof. Dr. Xiaoyi Jiang

Zweitgutachten:
Prof. Dr. Klaus Hinrichs

Eingereicht von:
Lars Haalck

Münster, September 2016

Zusammenfassung

abstract schrei-
ben

Inhaltsverzeichnis

1. Einleitung	1
2. Theoretische Grundlagen	3
2.1. Kegel	3
2.2. Lineares Ausgleichsproblem	7
2.3. Kamerakalibrierung und projektive Geometrie	8
2.4. Blob-Detektor	10
2.5. Canny	10
2.6. Hough-Transformation	11
2.7. RANSAC	11
2.8. Ellipsen	12
2.8.1. Definition	12
2.8.2. Abstand: Punkt zu Ellipse	15
2.9. Analytical Deformable Templates	17
3. Methodik	19
3.1. Kalibrierungsmuster	19
3.1.1. Aufbau des Kalibrierungsmusters	19
3.1.2. Anzahl der Samples	19
3.2. Intrinsische Kamerakalibrierung	21
3.3. Detektion der charakteristischen Punkte	21
3.4. Ellipsen-Detektion	21
3.4.1. RANSAC	21
3.4.2. Analytical Deformable Templates	24
3.5. Zuordnung der Punkte	26
3.6. Weltkoordinaten bestimmen	26
3.7. Entfaltung	27
3.7.1. Vorwärtsentfaltung	27
3.7.2. Rückwärtsentfaltung	29
4. Implementierung	31
5. Analyse	33
5.1. Vergleich Vorwärtsentfaltung und Rückwärtsentfaltung	33
5.2. Einfluss der intrinsischen Kalibrierung	35
5.3. Einfluss der Rotation der Kamera	35
5.4. Laufzeit der Entfaltung	35
5.5. Evaluierung Deformable Templates	36
5.6. Evaluierung des RANSAC-Verfahrens zur Ellipsendetektion	38
6. Fazit und Ausblick	41
6.1. Parallelisierung	41
6.2. Verbesserung der Linien-Detektion	41
6.3. Verbesserung der Vorwärtsentfaltung	41
6.4. Verbesserung der Rückwärtsentfaltung	41

INHALTSVERZEICHNIS

A. Ergebnisse	43
Abbildungsverzeichnis	45

1. Einleitung

hier setup mit bild?

einleitung
schreiben

2. Theoretische Grundlagen

2.1. Kegel

2.1.1 Definition (Kegel)

Ein Kegel ist ein geometrischer Körper, der durch eine beliebige Grundfläche, sowie einen Punkt definiert wird. Ein Kegel mit Kreis als Grundfläche wird als Kreiskegel bezeichnet. Liegt die Kegelspitze auf einer Geraden durch die Normale der Grundebene, bezeichnen wir den Kegel als gerade.

In der weiteren Arbeit betrachten wir nur gerade Kreiskegel. Ein Kegel mit Spitze $T(0, 0, 0)$, Radius R und Höhe H kann parametrisch beschrieben werden als:

$$\begin{aligned}x &= \frac{u}{H}R \cos\theta \\y &= u \\z &= \frac{u}{H}R \sin\theta\end{aligned}\tag{2.1}$$

mit $u \in [0, H]$ und $\theta \in [0, 2\pi]$

S bezeichne hierbei die Seitenhöhe und sei definiert durch das rechtwinklige Dreieck mit den Seitenlängen H , R und S (siehe Abbildung 2.1). Es gilt $S = \sqrt{H^2 + R^2}$.

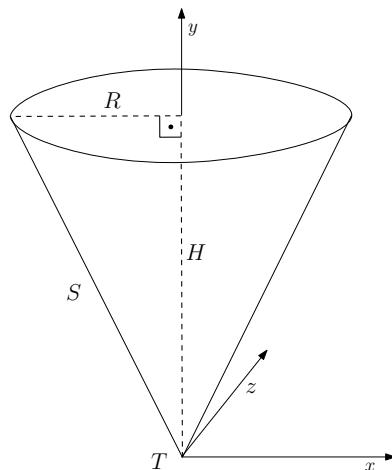


Abbildung 2.1.: Gerader Kreiskegel

2.1.2 Definition (Kegelstumpf und Ergänzungskegel)

Ein Kegelstumpf entsteht als Schnitt eines geraden Kreiskegels mit einer zur Grundfläche parallelen Ebene (siehe Abbildung 2.2). Der Kegel, der definiert ist durch die Schnittfläche und die Spitze des ursprünglichen Kegels bezeichnen wir als Ergänzungskegel. Die Differenz des eigentlichen Kegels und des Ergänzungskegels bezeichnen wir als Kegelstumpf.

H, R, S sind dabei weiterhin die Höhe, Radius und Seitenhöhe des gesamten Kegels. Hinzu kommen h, r, s als Höhe, Radius und Seitenhöhe des Ergänzungskegels. Die Höhe, sowie die Seitenhöhe des Kegelstumpfs werden durch die Differenzen $\Delta S = S - s$, $\Delta H = H - h$ charakterisiert (siehe Abbildung 2.3).

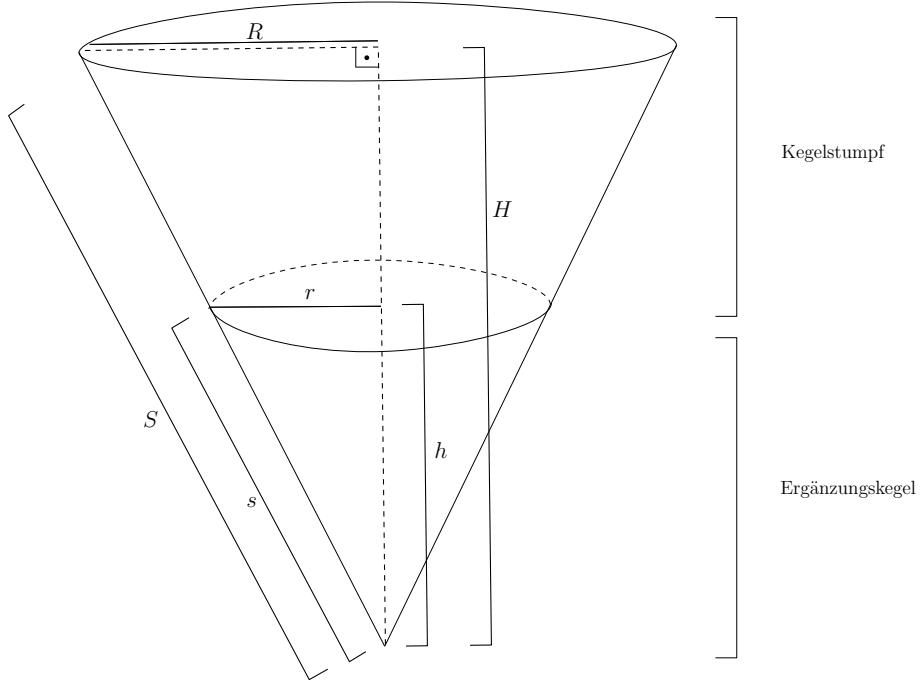


Abbildung 2.2.: Kegelstumpf und Ergänzungskegel

Wir definieren einen Kegelstumpf mit Zentrum $(0, 0, 0)$ des kleineren Kreises mit Radius r durch folgende Parametrisierung:

$$\begin{aligned} x &= \left(r + \frac{u}{\Delta H}(R - r)\right) \cos\theta \\ y &= u \\ z &= \left(r + \frac{u}{\Delta H}(R - r)\right) \sin\theta \end{aligned} \tag{2.2}$$

mit $u \in [0, \Delta H]$ und $\theta \in [0, 2\pi]$. Die parametrische Form eines Kegelstumpfs ist somit ein Verallgemeinerung der Parametrisierung von klassischen Kegeln (siehe Gleichung 2.1), wobei beim klassischen Kegel $r = 0$ gilt. Wie in Abbildung 2.3 zu sehen, wollen wir bei Höhe $u = 0$ einen Radius von r erreichen. Wir führen also eine Skalierung des Intervalls $[0, R]$ auf das Intervall $[r, R]$ durch. Dies erreichen wir durch den Term $\left(r + \frac{u}{\Delta H}(R - r)\right)$.

Die Mantelfläche des Kegelstumpfs aus Abbildung 2.4 kann dann in Polarkoordinaten parametrisch beschrieben werden als

$$\begin{aligned} x &= -(s + \frac{u}{\Delta H}(S - s)) \sin\phi \\ y &= (s + \frac{u}{\Delta H}(S - s)) \cos\phi \end{aligned} \tag{2.3}$$

mit $u \in [0, \Delta H]$ und $\phi \in [0, \alpha] \subseteq [0, 2\pi]$.

Die Mantelfläche ergibt sich durch Entrollen des Kegelstumpfs. Da der Umfang des Kreises mit Radius R , $2\pi R$ beträgt, muss, wie in Abbildung 2.4 zu sehen, der äußere Kreisbogen die Bogenlänge $2\pi R$ haben. Analog

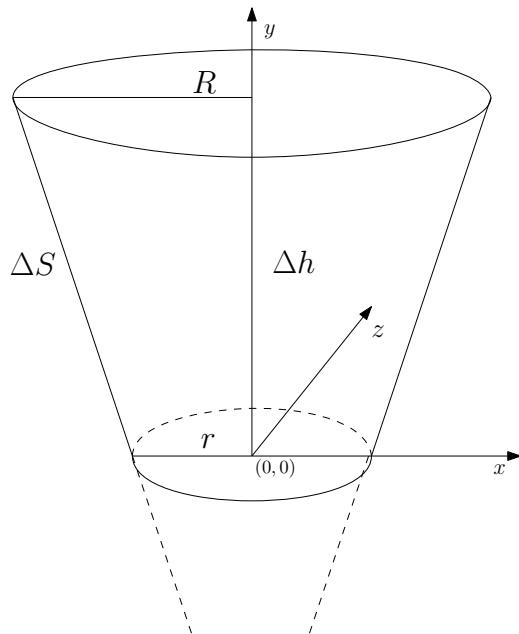


Abbildung 2.3.: Kegelstumpf

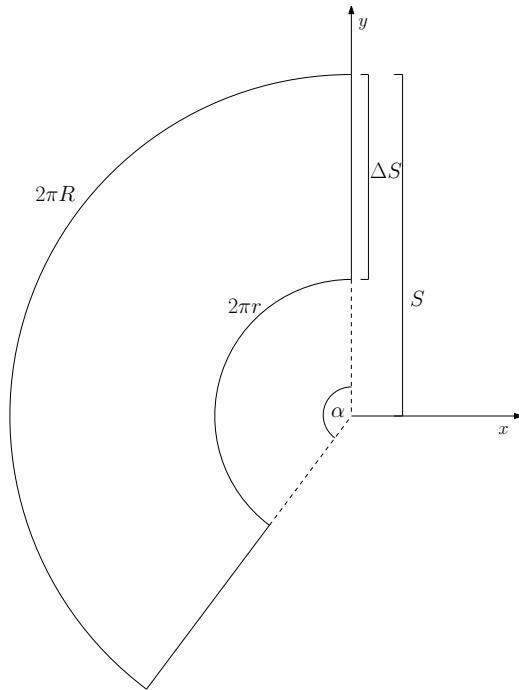


Abbildung 2.4.: Kegelmantelfläche

beträgt die Bogenlänge des inneren Kreisbogens $2\pi r$. Für den Winkel α muss demnach gelten $\alpha S = 2\pi R$, also $\alpha = 2\pi \frac{R}{S}$. Da sich die Seitenhöhe, wie in Abbildung 2.5 zu sehen, linear zur Höhe des Kegelstumpfs verhält, kann die Seitenhöhe durch die Höhe ausgedrückt werden mit $(s + \frac{u}{\Delta H}(S - s))$. Bei Höhe $u = 0$ ergibt sich somit, wie erwartet die Seitenhöhe s , bei Höhe $u = \Delta H$ die Seitenhöhe S .

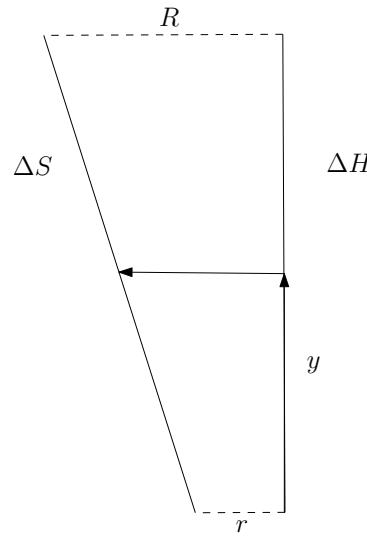


Abbildung 2.5.: Abbildung der Kegelstumpfhöhe auf die Seitenhöhe

Ein Punkt auf der Oberfläche des Kegelstumpfs kann eindeutig einem Punkt auf der Mantelfläche (und umgekehrt) zugeordnet werden. Dazu konstruieren wir folgende Abbildung und ihr Inverses:

Sei dazu ein Punkt $C(x, y, z)$ auf der Oberfläche des Kegelstumpfs gegeben. Gegeben durch die parametrischen Gleichung 2.2, hat C die Form

$$C(x, y, z) = \left(r + \frac{u}{\Delta H}(R - r) \right) \cos\theta, u, \left(r + \frac{u}{\Delta H}(R - r) \right) \sin\theta$$

für ein $u \in [0, \Delta H]$ und $\theta \in [0, 2\pi]$.

Aus der y -Koordinate lässt sich direkt die Höhe u ablesen und somit analog zur Gleichung 2.3 die zugehörige Seitenhöhe, also der Radius der Polarkoordinaten in der parametrischen Gleichung der Mantelfläche bestimmen. Wir definieren uns hierfür eine Hilfsfunktion

$$\Sigma(y) := s + \frac{y}{\Delta H}(S - s) \quad (2.4)$$

Da $R, r, \Delta H$ und nun auch die Höhe bekannt sind, können wir den Winkel θ im Kegelstumpf ausrechnen. Anschließend muss dieser noch mit $\frac{R}{S}$ multipliziert werden um ihn auf $[0, \alpha]$ zu skalieren (siehe Gleichung 2.3). Wir wollen in der Gleichung 2.2 nach $\cos\phi$, beziehungsweise nach $\sin\phi$ umstellen und anschließend nach ϕ auflösen. Wir teilen also zunächst beide Gleichungen durch $r + \frac{y}{\Delta H}(R - r)$. Zu beachten ist nun, dass wir nicht die Umkehrfunktionen acos oder asin benutzen können, da \sin und \cos in einer Periode $[0, 2\pi]$ nicht injektiv sind. Die Wertebereiche der Umkehrfunktionen sind also eingeschränkt und wir erhalten somit im Allgemeinen nicht den korrekten Winkel in $[0, 2\pi]$. Das gleiche Problem zeigt sich bei $\tan\phi = \frac{\sin\phi}{\cos\phi}$ und der Umkehrfunktion atan . Um dieses Problem zu beheben muss man eine Fallunterscheidung durchführen und wir definieren die Funktion atan2

$$\text{atan2}(y, x) = \begin{cases} \text{atan} \frac{y}{x} & \text{für } x > 0 \\ \text{atan} \frac{y}{x} + \frac{\pi}{2} & \text{für } x < 0, y \geq 0 \\ \text{atan} \frac{y}{x} - \frac{\pi}{2} & \text{für } x < 0, y < 0 \\ \text{atan} + \frac{\pi}{2} & \text{für } x = 0, y > 0 \\ \text{atan} - \frac{\pi}{2} & \text{für } x = 0, y < 0 \\ 0 & \text{für } x = 0, y = 0, \end{cases}$$

die den Winkel ϕ im richtigen Quadranten, also in $[0, 2\pi)$ bestimmt.

Wir definieren wieder eine Hilfsfunktion

$$\Phi(x, y, z) := \frac{R}{S} \operatorname{atan2}\left(\frac{z}{r + \frac{y}{\Delta H}(R - r)}, \frac{x}{r + \frac{y}{\Delta H}(R - r)}\right). \quad (2.5)$$

Mit den beiden Hilfsfunktionen 2.4 und 2.5, sowie der Gleichung 2.3 ergibt sich insgesamt

$$\begin{aligned} \Psi: [r, R] \times [0, \Delta H] \times [r, R] &\rightarrow [s, S] \times [s, S] \\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} &\mapsto \begin{pmatrix} -\Sigma(y) \sin \Phi(x, y, z) \\ \Sigma(y) \cos \Phi(x, y, z) \end{pmatrix} \end{aligned} \quad (2.6)$$

Analog lässt die sich Umkehrabbildung konstruieren:

Sei ein Punkt $L(x, y)$ auf der Mantelfläche des Kegelstumpfs gegeben. Aus der parametrischen Form der Mantelfläche (siehe Gleichung 2.3) ergibt sich

$$L(x, y) = \left(-\left(s + \frac{u}{\Delta H}(S - s)\right) \sin \phi, \left(s + \frac{u}{\Delta H}(S - s)\right) \cos \phi\right)$$

für ein passendes $u \in [0, \Delta H]$ und $\phi \in [0, \alpha] \subseteq [0, 2\pi]$.

Da $L(x, y)$ in Polarkoordinaten gegeben ist, lässt sich der Radius durch $\sqrt{x^2 + y^2}$ bestimmen. Wir können den Winkel ϕ mit inverser Skalierung also analog zu oben durch folgende Hilfsfunktion bestimmen:

$$\Theta(x, y) := \frac{S}{R} \operatorname{atan2}\left(-\frac{x}{\sqrt{x^2 + y^2}}, \frac{z}{\sqrt{x^2 + y^2}}\right) \quad (2.7)$$

Die Höhe im Kegel und somit der Radius lässt sich nun als Umkehrabbildung zu 2.4 bestimmen:

$$H(x, y) := \frac{\left(\sqrt{x^2 + y^2}\right) - s}{S - s} \Delta H \quad (2.8)$$

Insgesamt ergibt sich:

$$\begin{aligned} \Psi^{-1}: [s, S] \times [s, S] &\rightarrow [r, R] \times [0, \Delta H] \times [r, R] \\ \begin{pmatrix} x \\ y \end{pmatrix} &\mapsto \begin{pmatrix} \left(r + \frac{H(x, y)}{\Delta H}(R - r)\right) \cos(\Theta(x, y)) \\ H(x, y) \\ \left(r + \frac{H(x, y)}{\Delta H}(R - r)\right) \sin(\Theta(x, y)) \end{pmatrix} \end{aligned} \quad (2.9)$$

2.2. Lineares Ausgleichsproblem

Gegeben seien m Messdaten $(x_1, y_1), \dots, (x_m, y_m)$, sowie eine lineare Funktion $\varphi(x) = \sum_{i=1}^n c_i \varphi_i(x)$ in x mit $c_1, \dots, c_n \in \mathbb{R}$, Funktionen $\varphi_1, \dots, \varphi_n$, $n \leq m$ und $n, m \in \mathbb{N}$.

Gesucht ist eine Lösung welche die mittlere Abweichung

$$\Delta_2 = \min_{c_1, \dots, c_n \in \mathbb{R}} \left(\sum_{j=1}^m (y_j - \varphi(x_j))^2 \right) = \min_{c_1, \dots, c_n \in \mathbb{R}} \left(\sum_{j=1}^m \left(\sum_{i=1}^n y_j - (c_i \varphi_i(x_j)) \right)^2 \right)$$

minimiert [Sto07].

Diese Problem wird als lineares Ausgleichsproblem oder als Methode der kleinsten Quadrate bezeichnet und lässt sich wie folgt formulieren:

$$\min_{c \in \mathbb{R}^n} \|y - Ac\|_2,$$

mit

$$\begin{aligned} c &= (c_1, \dots, c_n)^T \in \mathbb{R}^n \\ x &= (x_1, \dots, x_m)^T \in \mathbb{R}^m \\ y &= (y_1, \dots, y_m)^T \in \mathbb{R}^m \\ A &= (a_{ij}) \in \mathbb{R}^{m \times n} \quad \text{mit} \quad a_{ij} = \varphi_i(x_j) \end{aligned}$$

Das Problem lässt sich auch mit Hilfe der äquivalenten sogenannten Normalengleichung

$$(A^T A)c = A^T y$$

lösen, wobei mindestens eine Lösung existiert und bei mehreren Lösungen, die Lösung mit kleinsten 2-Norm ausgewählt wird.

Gilt darüber hinaus $\text{rang}(A) = n$ so ist $A^T A$ invertierbar und es gilt

$$c = \underbrace{(A^T A)^{-1} A^T}_{{A}^+} y. \quad (2.10)$$

Dabei ist c die eindeutige Lösung mit kleinsten 2-Norm. A^+ bezeichnet man auch als Pseudoinverse von A . Im Allgemeinen ist die Matrix $A^T A$ nicht invertierbar. Die Pseudoinverse kann somit nicht einfach mit $(A^T A)^{-1} A^T$ berechnet werden. Stattdessen lässt sie sich durch eine Singulärwertzerlegung (SVD) bestimmen [Sto11]. Die Matrizen $A \in \mathbb{R}^{m \times n}$ und $A^+ \in \mathbb{R}^{n \times m}$ können geschrieben werden als:

$$\begin{aligned} A &= U \Sigma V^T \\ A^+ &= V \Sigma^+ U^T \end{aligned}$$

mit orthogonalen Matrizen $U \in \mathbb{R}^{m \times m}$ und $V \in \mathbb{R}^{n \times n}$ und einer Diagonalmatrix $\Sigma \in \mathbb{R}^{m \times n}$ mit den absteigend sortierten Singulärwerten auf der Diagonale (aufgefüllt mit Nullen), sowie $\Sigma^+ \in \mathbb{R}^{n \times m}$ mit den jeweiligen reziproken Singulärwerten.

Die Lösung des Ausgleichsproblems kann also mittels SVD angeben werden als:

$$c = V \Sigma^+ U^T y$$

Zu beachten ist, dass dieses Vorgehen auch bei Matrizen A mit $\text{rang}(A) = n$ funktioniert, also wenn $A^T A$ invertierbar ist und darüberhinaus numerisch stabil implementiert werden kann [Sto11].

2.3. Kamerakalibrierung und projektive Geometrie

ist noch nicht fertig

Um die Parameter einer Kamera korrekt beschreiben zu können, ist ein Kameramodell notwendig. Das wohl bekannteste Kameramodell ist das Lochkamera-Modell. Die Lichtstrahlen der Szene gelangen dabei durch eine kleine Öffnung in die Kamera (siehe dazu Abbildung 2.6).

Das Bild das an der Rückseite der Kamera entsteht ist dabei um 180° gedreht. Damit man diese Rotation nicht betrachten muss, kann man die Bildebene virtuell vor die Lochblende setzen. Da sich der Abstand zur Blende somit nicht ändert, ändern sich auch die optischen Eigenschaften nicht.

Kamerakalibrierung wird benötigt, um eine Beziehung zwischen Punkten im dreidimensionalen Weltkoordinatensystem und den Punkten auf der zweidimensionalen Bildebene herstellen zu können. Konkret suchen wir eine projektive Abbildung

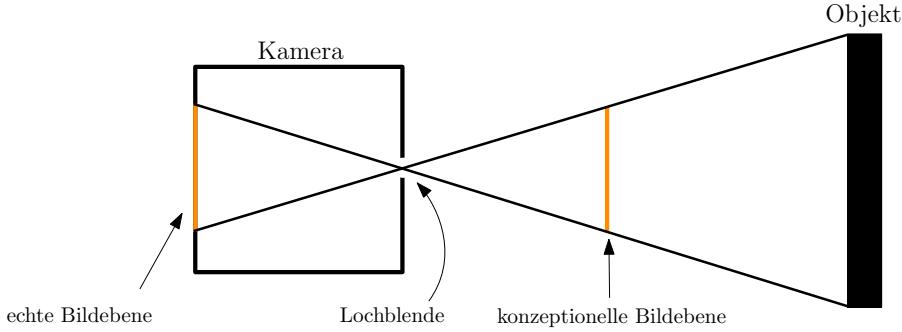


Abbildung 2.6.: Lochkameramodell

$$\begin{pmatrix} wu \\ wv \\ w \end{pmatrix} = \underbrace{\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ 1 & 1 & 1 & 1 \end{pmatrix}}_{=:A} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}, \quad (2.11)$$

die einen Punkt $P = (x, y, z)$ in homogenen Koordinaten $\tilde{P} = (x, y, z, 1)$ auf einen Punkt $\tilde{C} = (wu, wv, w)$ beziehungsweise nach der perspektivischen Division $C = (u, v)$ auf die Bildebene abbildet, wobei die a_{ij} von den intrinsischen und extrinsischen Parametern der Kamera abhängen [HS97].

Man kann die Matrix A mit gegebene Punktkorrespondenzen bestimmen. Wir benötigen dafür Punkte (x_k, y_k, z_k) , $k = 1, \dots, m$ im Weltkoordinatensystem und die korrespondierenden Punkte (u_k, v_k) auf der Bildebene. Wir stellen zunächst das Gleichungssystems 2.11 nach u und v um:

$$u_k = \frac{a_{11}x_k + a_{12}y_k + a_{13}z_k + a_{14}}{a_{31}x_k + a_{32}y_k + a_{33}z_k + a_{34}}$$

$$v_k = \frac{a_{21}x_k + a_{22}y_k + a_{23}z_k + a_{24}}{a_{31}x_k + a_{32}y_k + a_{33}z_k + a_{34}}.$$

Wir wollen nun die a_{ij} als Unbekannte betrachten und formulieren das Gleichungssystem noch einmal um zu:

$$\underbrace{\begin{pmatrix} x_1 & y_1 & z_1 & 1 & 0 & 0 & 0 & 0 & -u_1x_1 & -u_1y_1 & -u_1z_1 \\ 0 & 0 & 0 & 0 & x_1 & y_1 & z_1 & 1 & -v_1x_1 & -v_1y_1 & -v_1z_1 \\ \vdots & \vdots \\ x_m & y_m & z_m & 1 & 0 & 0 & 0 & 0 & -u_mx_m & -u_my_m & -u_mz_m \\ 0 & 0 & 0 & 0 & x_m & y_m & z_m & 1 & -v_mx_m & -v_my_m & -v_mz_m \end{pmatrix}}_{=:M} = \begin{pmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{14} \\ a_{21} \\ a_{22} \\ a_{23} \\ a_{24} \\ a_{31} \\ a_{32} \\ a_{33} \end{pmatrix} = \begin{pmatrix} u_1 \\ v_1 \\ \vdots \\ u_m \\ v_m \end{pmatrix}, \quad (2.12)$$

wobei a_{34} auf eins skaliert wird. Dies ist erlaubt, da wir in homogenen Koordinaten rechnen und dies durch die perspektivischen Division kompensiert wird. Es handelt sich hierbei für $m \geq 6$ um ein überbestimmtes lineares Gleichungssystem, was mittels der Methode der kleinsten Quadrate (siehe Kapitel 2.2) und durch eine Singulärwertzerlegung gelöst werden kann. Diese Verfahren wird auch als *Direct Linear Transformation* (DLT) bezeichnet.

Das beschriebene Verfahren vernachlässigt dabei Linsenverzerrungen. Solche Verzerrungen entstehen

entweder als Produktionsfehler günstiger Linsen, oder bewusst beispielsweise bei Weitwinkelkameras. Verzerrungen können in der Regel nicht linear modelliert werden. Der Ansatz über DLT funktioniert hier dementsprechend nicht.

Es wird grundsätzlich zwischen zwei Typen von Verzerrung unterschieden: radiale Verzerrung und tangentiale Verzerrung. Bei der Radiale Verzerrung entsteht dabei durch eine Skalierung des Abstand eines Bildpunktes zum Fokus. Wird der Abstand vergrößert spricht man von einer tonnenförmigen Verzerrung (siehe Abbildung 2.7(b)), wird er verkleinert von einer kissenförmigen Verzerrung (siehe Abbildung 2.7(c)). Radiale Verzerrung kann wie folgt modelliert werden:

$$\begin{aligned}\hat{x}^{(r)} &= x \left[1 + k_1 (x^2 + y^2) + k_2 (x^2 + y^2)^2 + \dots \right] \\ \hat{y}^{(r)} &= y \left[1 + k_1 (x^2 + y^2) + k_2 (x^2 + y^2)^2 + \dots \right],\end{aligned}$$

wobei k_1 und k_2 die Koeffizienten der radialen Verzerrung sind [Zhang2002].

Tangentiale Verzerrung ist auf einen fehlerhaften

$$\begin{aligned}\hat{x}^{(t)} &= [2p_1xy + p_2 ((x^2 + y^2) + 2x^2)] \\ \hat{y}^{(t)} &= [p_1 ((x^2 + y^2) + 2y^2) + 2p_2xy],\end{aligned}$$

wobei p_1 und p_2 die Koeffizienten der tangentialen Verzerrung sind [HS97].

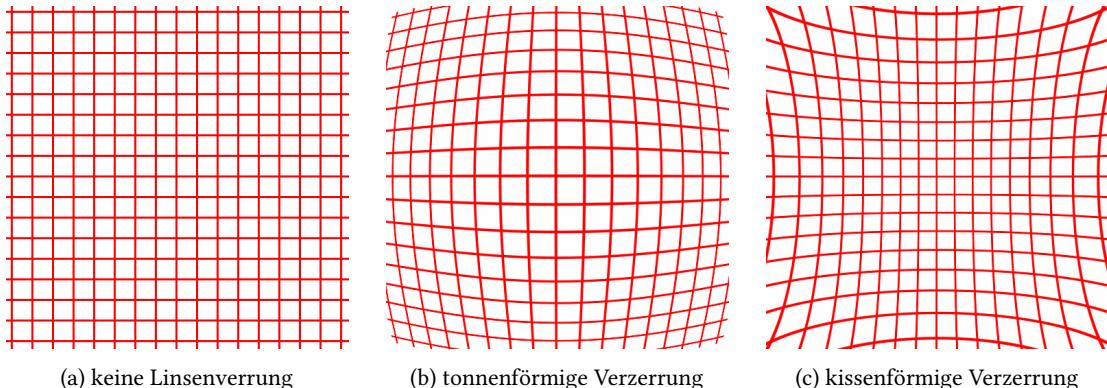


Abbildung 2.7.: Linsenverzerrungen

2.4. Blob-Detektor

2.4.1 Definition (Blob)

Ein Blob ist eine glatte zusammenhängende Region in einem Bild, die sich farblich von ihrer Umgebung abhebt [Lin93].

Ein Blob-Detektor ist also ein Verfahren zum Detektieren solcher Blobs. Die gefundenen Blobs können anschließend nach verschiedenen Kriterien, wie beispielsweise Größe, Farbe, Konvexität oder Rundheit gefiltert werden.

2.5. Canny

Der Canny-Algorithmus [Can86] ist ein Verfahren zur Kantendetektion. Im Gegensatz zu anderen Verfahren wie Sobel, oder Prewitt, versucht Canny die Fehlerrate der Kantendetektion minimal zu halten. Darüber

hinaus markiert Canny die Kanten möglichst exakt, minimiert also die Distanz eines markierten Punktes zum eigentlich Zentrum der Kante. Zuletzt gewährleistet Canny außerdem die Eindeutigkeit einer Kante. Das bedeutet, dass eine Kante nicht mehrmals markiert wird.

2.6. Hough-Transformation

Hough-Transformation ist ein Verfahren zur Detektion von beliebigen parametrisierbaren Konturen. In dieser Arbeit werden Hough-Transforamtionen benutzt um Geraden zu detektieren.

Eine beliebige zweidimensionale Gerade kann in Polarkoordinaten folgendermaßen implizit ausgedrückt werden:

$$x \cos \phi + y \sin \phi - d = 0, \quad (2.13)$$

wobei $\phi \in [0, 2\pi]$ der Winkel der Geraden mit der X-Achse und $d \geq 0$ der Radius, also der euklidische Abstand der Geraden zum Ursprung des Koordinatensystems ist.

Eine Gerade wird somit als ein Punkt (d, ϕ) in den Parameterraum (auch Hough-Raum) abgebildet.

Um eine Gerade eindeutig zu definieren benötigt man, wie auch bei der klassischen Defintionen $y = mx + b$, zwei Punkte. Nimmt man nur einen, so lässt sich jedoch die Auswahl von ϕ und d einschränken. Hat man beispielsweise einen Punkt (x_k, y_k) gegeben so lässt sich die Gleichung 2.13 nach d umstellen und man erhält eine sinusförmige Funktion in Abhängigkeit von ϕ .

Um nun beliebige Geraden detektieren zu können, werden ϕ und d zunächst passend diskretisiert:

$$\begin{aligned} \phi_i &= \phi_{min} + \frac{i}{n} \cdot (\phi_{max} - \phi_{min}) & \forall i \in [0, n] \\ d_j &= d_{min} + \frac{j}{m} \cdot (d_{max} - d_{min}) & \forall j \in [0, m] \end{aligned}$$

Es wird nun ein Akkumulator $\mathcal{H}(\phi, d)$ für alle ϕ_i und d_j auf null gesetzt.

Als Nächstes wird ein Kantentbild mittels Canny erzeugt und jene Pixel (x_k, y_k) betrachtet, die nicht null sind. Zu einem gegebenen Pixel wir nun für alle diskreten Winkel ϕ_i ein Wert für d ausgerechnet und auf den nächsten diskreten Wert d_j gerundet. Anschließend wird im Akkumulator \mathcal{H} der Wert an der Stelle (ϕ_i, d_j) erhöht. Für jeden Kantenpunkt werden sinusförmige diskrete Funktionen in den Hough-Raum abgebildet. Die Werte im Akkumulator werden oft auch als Votes bezeichnet. Am Ende des Verfahrens sucht man im Akkumulator nach Häufungspunkten. Jeder Häufungspunkt steht dort für einen Geradenkandidaten.

2.7. RANSAC

Random Sample Consensus (RANSAC) [FB81] ist ein nicht-deterministisches robustes Verfahren zur Parameterschätzung eines Modells bei einer, möglicherweise durch starke Ausreißer, gestörten Messreihe. Im Gegensatz zu Verfahren, wie der Methode der kleinsten Quadrate, die versuchen eine optimale Lösung für alle Messdaten zu bestimmen, nutzt RANSAC nur eine Teilmenge der Messreihe.

Wir wählen aus der Menge der Messdaten wiederholt zufällig die minimale Anzahl Messdaten aus, die nötig sind um das Modell eindeutig zu beschreiben und prüfen dann, wie gut das geschätzte Modell die restlichen Messdaten beschreibt. Die Güte des Modells wird im Allgemeinen durch ein Distanzmaß, wie zum Beispiel der euklidische Abstand, berechnet. Hat ein Messdatum eine vorher definierte Maximaldistanz zum geschätzten Modell nicht überschritten, wird es ins sogenannte Consensus Set des Modells aufgenommen. Das Modell mit dem größten Consensus Set wird schließlich ausgewählt.

Die Anzahl der Iterationen, die mindestens notwendig sind, um mit einer Wahrscheinlichkeit von $p \in [0, 1]$, bei einem relativen Ausreißeranteil von $\epsilon \in [0, 1]$ und einer Anzahl von k Daten, um das Modell eindeutig zu beschreiben, mindestens einmal eine ausreißerfreie Teilmenge der Messreihe zu erhalten, lässt sich berechnen mit [FB81]:

$$n_{min} = \frac{\log(1-p)}{\log\left(1-(1-\epsilon)^k\right)} \quad (2.14)$$

2.8. Ellipsen

2.8.1. Definition

2.8.1 Definition (Ellipse)

Eine Ellipse wird beschrieben durch diejenigen Punkte, dessen Summe der Abstände zu zwei gegebenen Punkten f_1 und f_2 (Brennpunkte) konstant sind. Der Mittelpunkt der Verbindungslinie der beiden Brennpunkte wird als Zentrum der Ellipse bezeichnet (siehe Abbildung 2.8). Als Hauptachse a bezeichnen wir die Verbindungslinie von Mittelpunkt durch einen der Brennpunkte bis zum Scheitel V_1 (bzw. V_3) der Ellipse. Die zu ihr rechtwinklige Verbindungslinie durch den Mittelpunkt bis zu einem der anderen Scheitelpunkte (V_2 oder V_4) nennen wir Nebenachse b .

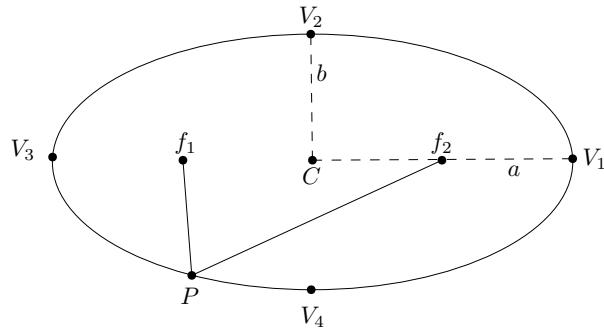


Abbildung 2.8.: Ellipse mit Brennpunkten f_1, f_2 , Zentrum C , Hauptachse a , Nebenachse b und Scheitelpunkten V_1 und V_2

In ihrer einfachsten Form liegt die Ellipse im Zentrum des Koordinatensystems und ihre Haupt- und Nebenachse a und b sind achsenausgerichtet. Das heißt ihre Hauptachse liegt auf der X-Achse und ihre Nebenachse auf der Y-Achse. Sie kann dann in der impliziten Form

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (2.15)$$

beschrieben werden.

Befindet sich die Ellipse nicht im Ursprung so muss eine Verschiebung beziehungsweise bei einer Rotation ein Drehwinkel (siehe Abbildung 2.9) ergänzt werden.

$$\frac{(x - x_0)^2}{a^2} + \frac{(y - y_0)^2}{b^2} = 1 \quad (2.16)$$

$$\frac{((x - x_0) \cos \theta + (y - y_0) \sin \theta)^2}{a^2} + \frac{((x - x_0) \sin \theta - (y - y_0) \cos \theta)^2}{b^2} = 1 \quad (2.17)$$

mit Ellipsenzentrum $(x_0, y_0) \in \mathbb{R}^2$, Hauptachsen $a, b \in \mathbb{R}^+$, sowie Drehwinkel $\theta \in [0, 2\pi)$ oder parametrisiert

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 + a \cos \phi \cos \theta - b \sin \phi \sin \theta \\ y_0 + a \cos \phi \sin \theta + b \sin \phi \cos \theta \end{pmatrix} \quad (2.18)$$

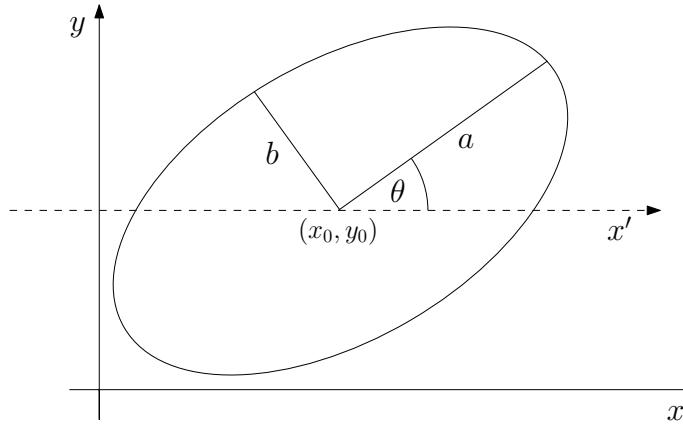


Abbildung 2.9.: Ellipse mit Zentrum (x_0, y_0) , Hauptachse a , Nebenachse b , sowie Drehwinkel θ

mit $\phi \in [0, 2\pi)$ und x_0, y_0, a, b, θ wie oben.

In ihrer allgemeinsten Form lässt sich eine Ellipse durch ein implizites Polynom zweiten Grades charakterisieren

$$ax^2 + by^2 + cxy + dx + ey + f = 0 \quad \text{mit} \quad c^2 - 4ab < 0 \quad (2.19)$$

mit $a, b, c, d, e, f \in \mathbb{R}$. Eine Ellipse lässt sich also durch sechs Punkte eindeutig beschreiben (fünf, wenn man f auf eins skaliert).

Die beiden Formen 2.17 und 2.19 sind äquivalent, falls die Ellipse nicht degeneriert ist (ohne Beweis) [Law72]. Die Umformung von 2.17 nach 2.19 ist mit Hilfe einer Hauptachsentransformation möglich. Da wir diese später brauchen, wird sie hier einmal exemplarisch vorgeführt.

Zunächst einmal fällt auf, dass der gemischte Term cxy genau dann null ist, wenn die Ellipse nicht rotiert wurde. Im ersten Schritt versuchen wir also die Rotation der Ellipse rückgängig zu machen, um den Rotationswinkel bestimmen zu können.

Die Gleichung 2.19 kann umgeformt werden zu:

$$\underbrace{(x \ y)}_{=:u^T} \underbrace{\begin{pmatrix} a & \frac{c}{2} \\ \frac{c}{2} & b \end{pmatrix}}_{=:M} \underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{=:u} + (d \ e) \underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{=:u} + f = 0$$

$$\Leftrightarrow u^T M u + (d \ e) u + f = 0$$

Der gemischte Term wird alleine durch $M = \begin{pmatrix} a & \frac{c}{2} \\ \frac{c}{2} & b \end{pmatrix}$ bestimmt. Da die Matrix M symmetrisch ist, ist sie orthogonal diagonalisierbar. Des Weiteren hat M zwei von null verschiedene Eigenwerte, denn

$$\det M = ab - \frac{c^2}{4}$$

ist nur dann gleich null, wenn $c^2 - 4ab = 0$, was ein Widerspruch zur Annahme in 2.19 ist. M hat somit eine von null verschiedene Determinante und somit vollen Rang, hat also zwei von null verschiedene Eigenwerte. Insbesondere gibt es also zwei Eigenvektoren von M , die zueinander orthogonal sind.

Es gilt $M = S^T DS$, wobei $S \in \mathbb{R}^{2 \times 2}$ eine orthogonale Matrix mit den normierten Eigenvektoren als Zeilen und $D = \text{diag}(\lambda_1, \lambda_2) \in \mathbb{R}^{2 \times 2}$ eine Diagonalmatrix mit den beiden Eigenwerten von M auf der Diagonalen ist. Ohne Beschränkung der Allgemeinheit gelte $\lambda_1 \leq \lambda_2$, andernfalls vertausche die Eigenvektoren in S .

Sei nun $v := Su$. So gilt:

$$\begin{aligned}
 & u^T (S^T D S) u + (d - e) \underbrace{(S^T S)}_{=1} u + f = 0 \\
 \Leftrightarrow & (S u)^T D (S u) + (d - e) S^T (S u) + f = 0 \\
 \Leftrightarrow & v^T D v + (d - e) S^T v + f = 0
 \end{aligned} \tag{2.20}$$

Man kann leicht nachrechnen, dass der gemischte Teil somit eliminiert wurde. Durch Anwenden der Transformation S wurde u also in das Koordinatensystem, in dem die Ellipse achsenausgerichtet ist, transformiert.

Eine Rotationsmatrix mit Rotationswinkel θ ist definiert durch:

$$R = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \tag{2.21}$$

Es gilt offenbar $S = R$ für ein geeignetes θ , da die Eigenvektoren normiert und orthogonal zueinander sind. θ kann also einfach ausgerechnet werden, denn es gilt:

$$\theta = \text{atan2}(\sin \theta, \cos \theta) = \text{atan2}(S_{2,1}, S_{1,1})$$

Multipliziert man nun 2.20 aus, ergibt sich:

$$\begin{aligned}
 & \lambda_1 v_1^2 + \lambda_2 v_2^2 + (d - e) \underbrace{S^T v}_{=: (d', e')} + f = 0 \\
 \Leftrightarrow & \lambda_1 v_1^2 + \lambda_2 v_2^2 + d' v_1 + e' v_2 + f = 0 \\
 \Leftrightarrow & (\lambda_1 v_1^2 + d' v_1) + (\lambda_2 v_2^2 + e' v_2) + f = 0 \\
 \Leftrightarrow & (\lambda_1 v_1^2 + d' v_1) + \left(\frac{d'^2}{4\lambda_1} - \frac{d'^2}{4\lambda_1} \right) + (\lambda_2 v_2^2 + e' v_2) + \left(\frac{e'^2}{4\lambda_2} - \frac{e'^2}{4\lambda_2} \right) + f = 0 \\
 \Leftrightarrow & \left[\lambda_1 \left(v_1^2 + \frac{2d'}{2\lambda_1} v_1 + \frac{d'^2}{4\lambda_1^2} \right) - \frac{d'^2}{4\lambda_1} \right] + \left[\lambda_2 \left(v_2^2 + \frac{2e'}{2\lambda_2} v_2 + \frac{e'^2}{4\lambda_2^2} \right) - \frac{e'^2}{4\lambda_2} \right] + f = 0 \\
 \Leftrightarrow & \underbrace{\lambda_1 (v_1 + \frac{d'}{2\lambda_1} v_1)^2}_{=-x'_0} + \underbrace{\lambda_2 (v_2 + \frac{e'}{2\lambda_2} v_2)^2}_{=-y'_0} - \underbrace{(\frac{d'^2}{4\lambda_1} + \frac{e'^2}{4\lambda_2} - f)}_{=:\sigma} = 0,
 \end{aligned} \tag{2.22}$$

da $\lambda_1, \lambda_2 \neq 0$.

Das Zentrum der transformierten Ellipse kann nun aus 2.22 einfach abgelesen werden. Um das Zentrum der eigentlichen Ellipse zu bestimmen, muss mit der inversen Rotation S^T multipliziert werden:

$$(x_0, y_0)^T = S^T (x'_0, y'_0)^T$$

Obige Gleichung lässt sich anschließend weiter vereinfachen:

$$\begin{aligned}
 & \lambda_1 (v_1 - x'_0)^2 + \lambda_2 (v_2 - y'_0)^2 = \sigma \\
 \Leftrightarrow & \frac{\lambda_1}{\sigma} (v_1 - x'_0)^2 + \frac{\lambda_2}{\sigma} (v_2 - y'_0)^2 = 1
 \end{aligned} \tag{2.23}$$

wobei $\sigma \neq 0$, wenn die Ellipse nicht zum Punkt entartet ist [Law72]. Vergleicht man nun 2.23 mit 2.16 so sieht man das folgendes gelten muss:

$$\begin{aligned} \frac{\lambda_1}{\sigma} = \frac{1}{a^2} &\quad \text{und} \quad \frac{\lambda_2}{\sigma} = \frac{1}{b^2} \\ \Leftrightarrow \sqrt{\frac{\sigma}{\lambda_1}} = a &\quad \text{und} \quad \sqrt{\frac{\sigma}{\lambda_2}} = b \end{aligned} \tag{2.24}$$

Es gilt wie erwartet $a \geq b$, da $\lambda_1 \leq \lambda_2$.

2.8.2. Abstand: Punkt zu Ellipse

Das hier beschriebene Verfahren zur Bestimmung der kürzesten euklidischen Distanz eines Punktes zu einer Ellipse stammt aus der Arbeit von David Eberly [Ebe13]. Wir betrachten nur Ellipsen im Ursprung, die achsenausgerichtet sind und darüber hinaus nur Punkte im ersten Quadranten. Ansonsten wir die Ellipse in den Ursprung verschoben und um ihren entgegengesetzten Drehwinkel rotiert. Da die Ellipse dann bezüglich der X- und Y-Achse symmetrisch ist, kann der Punkt einfach durch Spiegelung in den richtigen Quadranten transformiert werden. Der Abstand ändert sich dadurch nicht.

Wir bezeichnen von nun an $Q = (y_0, y_1)$ als eine Punkt, dessen Distanz zur Ellipse wir berechnen wollen und $E = (x_0, x_1)$ als denjenigen eindeutigen Punkt, welcher auf der Ellipse liegt und die kürzeste euklidische Distanz zum Punkt Q hat.

Aufgrund dieser Forderungen können wir ohne Beschränkung der Allgemeinheit folgende Aussagen treffen:

- Die Ellipse kann stets durch die implizite Gleichung

$$\frac{x_0^2}{a^2} + \frac{x_1^2}{b^2} = 1 \tag{2.25}$$

mit $a \geq b \geq 0$ beziehungsweise in der parametrischen Form

$$\mathcal{E}(\theta) = (a \cos \phi, b \sin \phi) \quad \phi \in [0, 2\pi)$$

beschrieben werden.

- Es gilt $y_0, y_1, x_0, x_1 \geq 0$.

Für die quadrierte Distanz von einem beliebigen Punkt Q zu einem Punkt $\mathcal{E}(\theta)$ auf der Ellipse gilt dann

$$F(\theta) = |\mathcal{E}(\theta) - Q|^2. \tag{2.26}$$

Wir wollen F minimieren und betrachten die Ableitung:

$$F'(\theta) = 2(\mathcal{E}(\theta) - Q) \cdot \mathcal{E}'(\theta) \tag{2.27}$$

F' wird null, wenn $(\mathcal{E}(\theta) - Q)$ und $\mathcal{E}'(\theta)$ zu einander orthogonal sind. Daraus folgt, dass der Vektor von Q zu E senkrecht zur Ellipse stehen muss. (siehe Abbildung 2.10).

Betrachten wir also die Funktion:

$$G(x_0, x_1) = \frac{x_0^2}{a^2} + \frac{x_1^2}{b^2} - 1. \tag{2.28}$$

(x_0, x_1) ist genau dann ein Punkt auf der Ellipse, wenn $G(x_0, x_1) = 0$. Der Gradient von G in (x_0, x_1) ist ein Normalenvektor auf der Ellipse und somit auch der halbe Gradient $\nabla G(x_0, x_1)/2$. Der Vektor von E zu Q muss dieselbe Richtung haben. Es gilt somit:

$$(y_0, y_1) - (x_0, x_1) = t \frac{\nabla G(x_0, x_1)}{2} = t \left(\frac{x_0}{a^2}, \frac{x_1}{b^2} \right) \tag{2.29}$$

für ein $t \in \mathbb{R}$.

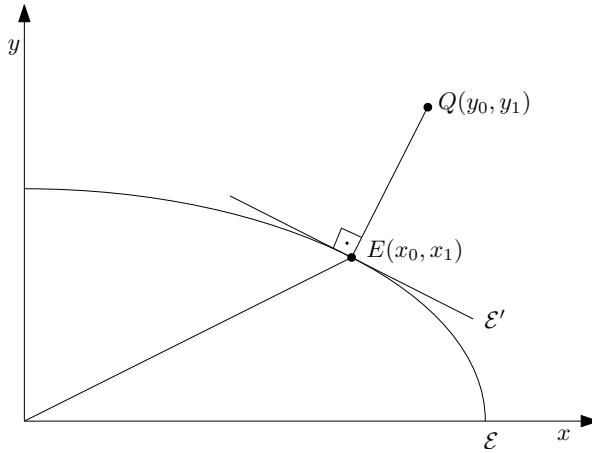


Abbildung 2.10.: Ellipsenausschnitt im ersten Quadranten mit Abfragepunkt Q und eingezeichneter kürzester Distanz zur Ellipse

Umgestellt nach y_0 und y_1 , beziehungsweise nach x_0 und x_1 ergibt sich:

$$y_0 = x_0 \left(1 + \frac{t}{a^2}\right), \quad y_1 = x_1 \left(1 + \frac{t}{b^2}\right) \quad (2.30)$$

$$x_0 = \frac{a^2 y_0}{t + a^2}, \quad x_1 = \frac{b^2 y_1}{t + b^2} \quad (2.31)$$

Man macht nun eine Fallunterscheidung:

1. Der einfachste Fall ist, wenn sich der Punkt Q auf der Y -Achse (außer $(0, 0)$) befindet, wenn also gilt $y_0 = 0, y_1 > 0$. Da die Hauptachse nach der X -Achse ausgerichtet ist und $a >= b$ gilt, ist der Punkt auf der Ellipse mit der kürzesten Distanz zu Q offenbar $E = (0, b)$ und für die Distanz gilt $d = |y_1 - b|$.
2. Als nächstes betrachten wir ein Q auf der X -Achse (einschließlich $(0, 0)$), wenn also gilt $y_0 \geq 0, y_1 = 0$. Es gilt also mit 2.30

$$y_0 = x_0 \left(1 + \frac{t}{a^2}\right), \quad 0 = x_1 \left(1 + \frac{t}{b^2}\right)$$

Wenn $x_1 = 0$ gilt, muss $x_0 = a$ gelten, damit $E(x_0, x_1)$ auf der Ellipse ist. Es gilt analog zum ersten Fall $E = (a, 0)$ mit $d = |y_0 - a|$

Gilt $x_1 \neq 0$, so können wir in der zweiten Gleichung durch x_1 teilen und es folgt $t = -b^2$ und somit $y_0 = x_0 \left(1 - \frac{b^2}{a^2}\right)$. Auf Grund der Krümmung der Ellipse gilt außerdem $x_0 < a$ und somit ergibt sich zusammen mit 2.31 die Ungleichung:

$$\begin{aligned} x_0 &= \frac{a^2 y_0}{a^2 - b^2} < a \\ \Leftrightarrow y_0 &< \frac{a^2 - b^2}{a} < a \end{aligned}$$

Für Punkte $Q(y_0, 0)$ mit $y_0 \geq \frac{a^2 - b^2}{a}$, ist der kürzeste Punkt also wieder $E(a, 0)$

Für Punkte $Q(y_0, 0)$ mit $y_0 < \frac{a^2 - b^2}{a}$, muss für $E(x_0, x_1)$ nach Umstellen der Ellipsengleichung 2.25 $x_1 = b \cdot \sqrt{1 - \left(\frac{x_0}{a}\right)}$ gelten. Die Distanz beträgt dann:

$$d^2 = (x_0 - y_0)^2 + x_1^2 = b^2 \left(1 - \frac{y_0^2}{a^2 - b^2} \right)$$

3. Der letzte Fall, den wir betrachten müssen, ist der allgemeinste Fall. Es gilt $y_0 > 0$, sowie $y_1 > 0$. Da wir uns nur im ersten Quadranten bewegen, gilt darüber hinaus $x_0, x_1 \geq 0$. Mit diesen Eigenschaften und 2.30 lässt sich folgende Einschränkung für t herleiten:

$$\begin{aligned} 0 < y_0 &= x_0 \left(1 + \frac{t}{a^2} \right) \\ \Leftrightarrow -1 \cdot a^2 &< t. \end{aligned} \tag{2.32}$$

Analog ergibt sich mit y_1 : $-b^2 < t$. Da $a \geq b$ gilt, reicht es, sich nur die zweite Ungleichung anzuschauen, da sie die erste impliziert. Setzt man nun 2.31 in 2.28 ein, erhält man:

$$F(t) = \left(\frac{ay_0}{t + a^2} \right)^2 + \left(\frac{by_1}{t + b^2} \right)^2 - 1 \tag{2.33}$$

Man kann nun zeigen, dass diese Funktion auf dem gesamten Intervall $[-b^2, \infty)$ monoton fällt und links gekrümmmt ist. Darüber hinaus gilt:

$$\lim_{t \searrow -b^2} F(t) = +\infty, \quad \lim_{t \rightarrow \infty} F(t) = 1.$$

Da F stetig ist, muss es also eine Nullstelle geben, die aufgrund des Monotonie- und Krümmungsverhaltens sogar eindeutig ist. Die Nullstelle lässt sich beispielsweise durch Intervallschachtelung oder Newton-Verfahren bestimmen.

2.9. Analytical Deformable Templates

Analytical Deformable Templates ist ein Verfahren zur Detektion von analytischen Kurven. Ein Template ist dabei definiert durch eine Menge von Parametern, die apriori-Wissen über die erwartete Form ermöglichen. Als Beispiel sei eine Ellipse genannt, die durch ihr Zentrum, ihren Achsen und den Drehwinkel (x_0, y_0, a, b, ϕ) definiert werden kann. Es wird dann eine Energiefunktion konstruiert, die Terme enthält, die das definierte Template

die mittels Optimierung einer Energiefunktion.

etwas mehr ausführen

3. Methodik

In diesem Kapitel stellen wir zwei Vorgehensweisen zur Entzerrung von Kegeloberflächen vor. Zunächst gehen wir auf das verwendete Kalibrierungsmuster ein, worauf hin die einzelnen Schritte der Entzerrung erläutert werden.

Die geometrischen Eigenschaften des Kegelstumpfs ($r, R, \Delta H$) können gemessen und somit als bekannt angenommen werden. Darüber hinaus nehmen wir an, dass sich das Zentrum des kleineren Kreises im links-händigen Weltkoordinatensystem an der Position $(0, 0, 0)$ (siehe Abbildung 2.3 in Kapitel 2.1) befindet. Durch diese Einschränkung gehen jegliche absolute Größenverhältnisse verloren. Die Larven können jedoch weiterhin relativ zu einander verglichen werden.

3.1. Kalibrierungsmuster

3.1.1. Aufbau des Kalibrierungsmusters

Um eine Beziehung zwischen Bildpunkten und Kegelpunkten herstellen zu können, ist ein Kalibrierungsmuster notwendig.

Die Wahl des Kalibrierungsmusters spielt dabei eine entscheidende Rolle bei der Robustheit und Präzision der Entfaltung. Es muss gewährleistet sein, dass die charakteristischen Merkmale des Musters auch bei leichten Abweichungen der Kamera vom Lot und schlechteren Beleuchtungssituationen zuverlässig erkannt werden. Das Muster muss darüber hinaus so entworfen sein, dass beim Zusammenlegen im Kegel, dessen geometrische Eigenschaften nicht verfälscht, sondern realitätsgerecht wiedergeben werden.

Wir haben uns für ein Muster entschieden, dass in äquidistanten Abständen ΔR , beginnend mit dem kleinen Radius r des Kegelstumpfs (siehe Abbildung 2.3) Kreislinien und in gleichen Winkelabständen $\Delta\alpha$ auf der Seitenhöhe Liniensegmente besitzt. Das zusammengelegte Muster ist in Abbildung 3.1 zu sehen, beziehungsweise das entfaltete in 3.2. Die Anzahl der Kreislinien wird mit n gekennzeichnet, die Anzahl sichtbarer Liniensegmente im Kegel mit m . Zu beachten ist, dass bedingt durch das Entfalten, in Abbildung 3.2 ein Liniensegment doppelt zu sehen ist. Die schwarzen Kreise bezeichnen wir als Samples.

Dadurch dass die Geometrie des Kegels bekannt ist, kann jedem Sample nun ein Punkt auf dem Kegel im Weltkoordinatensystem zugeordnet werden. Da ein Kegel beliebig um die Y-Achse rotiert werden kann, ist diese Zuordnung zunächst nicht eindeutig. Dazu nehmen wir an, dass das Liniensegment mit dem kleinsten Winkel zur X-Achse mit dem Kegelwinkel $\theta = 0$ korrespondiert (siehe 2.2).

Zur Konstruktion des Musters benötigt man ΔS und s , die man aus der Geometrie des Kegels errechnen kann und außerdem den Öffnungswinkel, der gegeben ist als $\alpha = 2\pi \frac{R}{S}$ (siehe 2.3 in Kapitel 2.1).

3.1.2. Anzahl der Samples

Die Anzahl der Samples sollte groß genug sein, um möglichst viel geometrische Informationen des Kegels zu erhalten, aber klein genug, dass eine Detektion der Samples problemlos möglich ist. Insbesondere auf dem innersten Kreis, macht sich eine zu hohe Sampleanzahl negativ bemerkbar, da der Abstand zueinander sehr klein wird, was eine Detektion erschwert. Des Weiteren sollte noch ein möglichst großer Teil der Kreislinien zu sehen bleiben, da diese für die Ellipsendetektion benötigt werden.

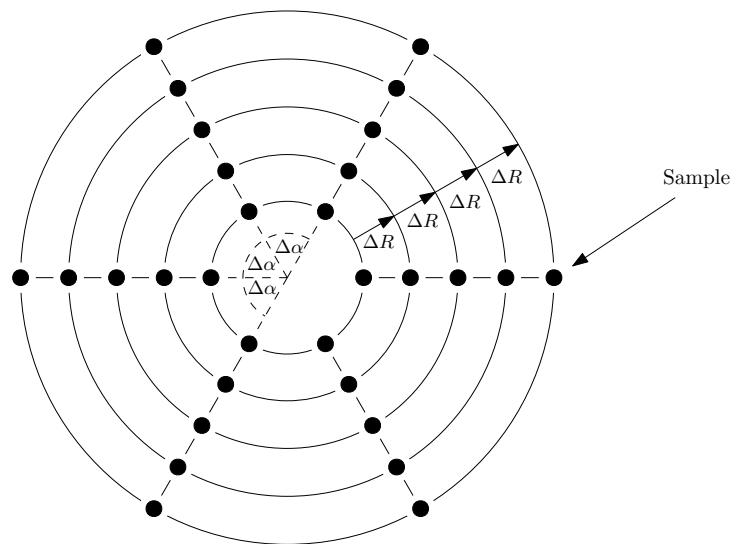


Abbildung 3.1.: Kalibrierungsmuster von oben mit $n = 5, m = 6$

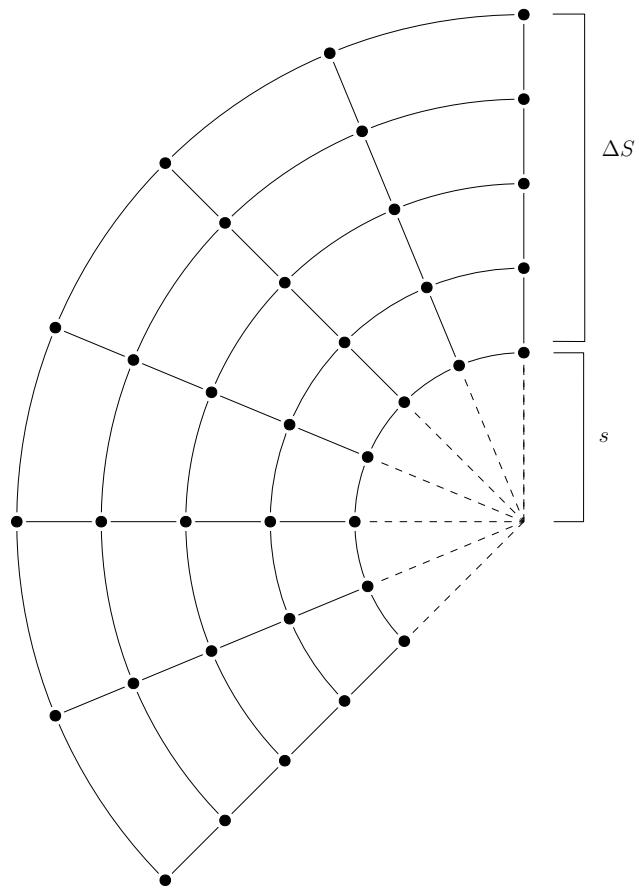


Abbildung 3.2.: Kalibrierungsmuster entfaltet mit $n = 5, m = 6$

3.2. Intrinsische Kamerakalibrierung

Bedingt durch die Wahl einer Weitwinkelkamera, enthält die Linse der Kamera eine starke tonnenförmige (nach außen gewölbte) Verzerrung (siehe Abbildung 3.3). Diese muss herausgerechnet werden, da sonst Abstände im Bild nicht mehr der Realität entsprechen und dadurch die Präzision der Entfaltung stark abnimmt (siehe Kapitel 5). Es wird also zunächst eine Kamerakalibrierung durchgeführt, wobei wir nur an den intrinsischen Parametern, inklusive Verzerrungskoeffizienten, interessiert sind, um das Bild anschließend entzerren zu können.

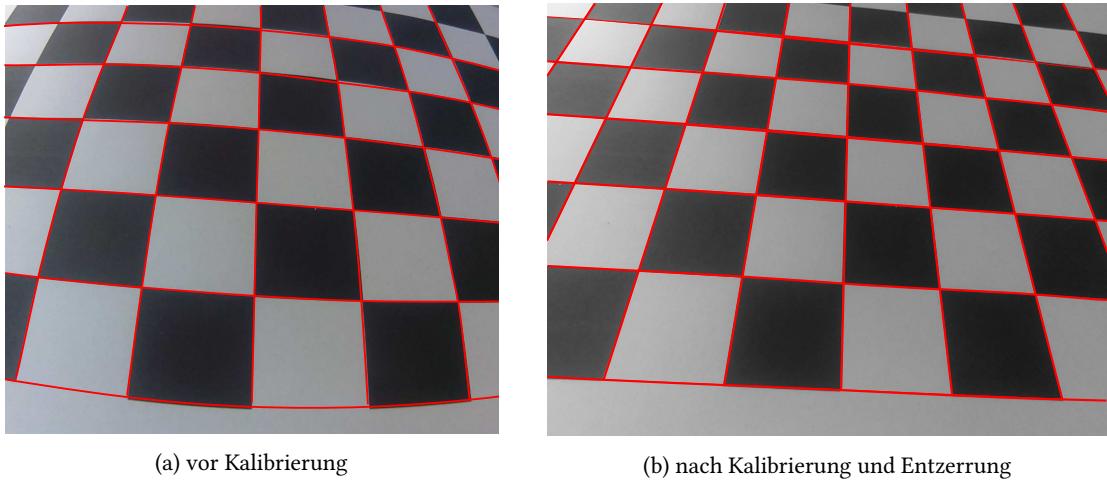


Abbildung 3.3.: Kamerakalibrierung

3.3. Detektion der charakteristischen Punkte

Nach der Kamerakalibrierung und entsprechender Entzerrung werden die Bildkoordinaten der Samples bestimmt. Dazu wird ein Blob-Detektor (siehe Kapitel 2.4) benutzt. Um ein Sample korrekt detektieren zu können, muss sich der Punkt farblich stark von seiner Umgebung abheben (siehe Definition: Blob 2.4.1). Insbesondere dürfen die Kreislinien und Liniensegmente des Kalibrierungsmusters also nicht durchgezogen sein. Nach der Detektion werden die Blobs nach folgenden Kriterien gefiltert:

- **Fläche:** zu kleine Blobs werden verworfen
- **Rundheit:** zu unrude Blobs werden verworfen. Rundheit ist hier definiert als $circ = \frac{4\pi \cdot \text{Fläche}}{(\text{Umfang})^2} \in [0, 1]$, wobei also ein Kreis mit $circ = 1$ maximal rund ist.
- **Konvexität:** zu unkonvexe Blobs werden verworfen. Konvexität ist hier definiert als $conv = \frac{\text{Fläche Blob}}{\text{Fläche konvexe Hülle}}$

In Abbildung 3.4 ist beispielhaft links ein Grauwertbild und rechts die detektierten Blobs (grün) auf dem gleichen Bild nach der Kameraentzerrung zu sehen.

3.4. Ellipsen-Detektion

3.4.1. RANSAC

Nachdem die Sample-Positionen bestimmt wurden, muss für jeden Sample entschieden werden, auf welcher der Kreislinien er liegt. Da die Kreise, bedingt durch perspektivische Verzerrung, zu Ellipsen werden, wird eine Verfahren benötigt, dass Ellipsen erkennt.

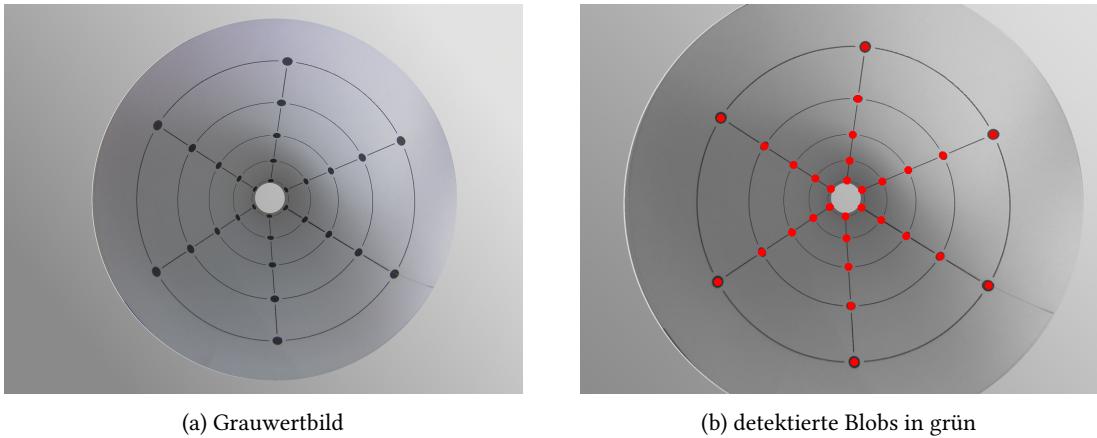


Abbildung 3.4.: Detektion der Samples

Zunächst werden die Kanten mit Hilfe von Canny (siehe Kapitel 2.5) detektiert (siehe Abbildung 3.5).

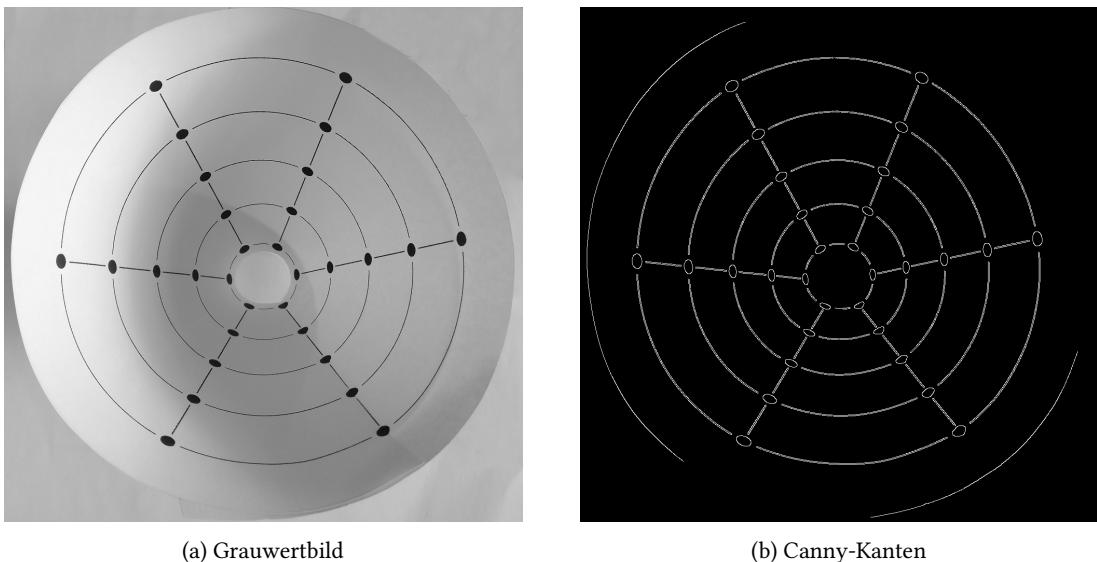


Abbildung 3.5.: Canny-Kantendetektion auf Grauwertbild

Anschließend versuchen wir möglichst genau das Zentrum der innersten Ellipsen zu schätzen. Wir benutzen dafür Hough-Transformationen (siehe Kapitel 2.6), um Linien im Kantenbild zu detektieren. Es werden anschließend die Schnittpunkte aller Liniensegmente bestimmt. Bedingt durch Ungenauigkeiten beim Ausschneiden und Zusammenlegen im Kegel und perspektivischer Verzerrung, schneiden sich nicht alle Liniensegmente in einem Punkt. Darüber hinaus werden, auf Grund der Liniendicke auf dem Kalibrierungsmuster, durch Canny viele Linien als doppelt Linien gekennzeichnet¹. Auch ein inhomogener Hintergrund, erschwert die Schnittpunktsbestimmung. Um also möglichst robust einen Kandidaten auszuwählen, wird zuerst der Median der x -Koordinaten der Schnittpunkte und dann der Median der y -Koordinaten bestimmt. Die erhaltenen Koordinaten bilden den Schnittpunkt (siehe Abbildung 3.6).

Von diesem Schnittpunkt aus werden, in einer vorher definierte Anzahl, gleichmäßig, in alle Richtungen

¹Dies ist kein Widerspruch zur Eindeutigkeit von Canny-Kanten (siehe 2.5). Stellt man sich ein relativ breites Liniensegment vor, so gibt es einmal den Übergang vom Hintergrund auf die Linie, sowie den Übergang von der Linie wieder auf den Hintergrund

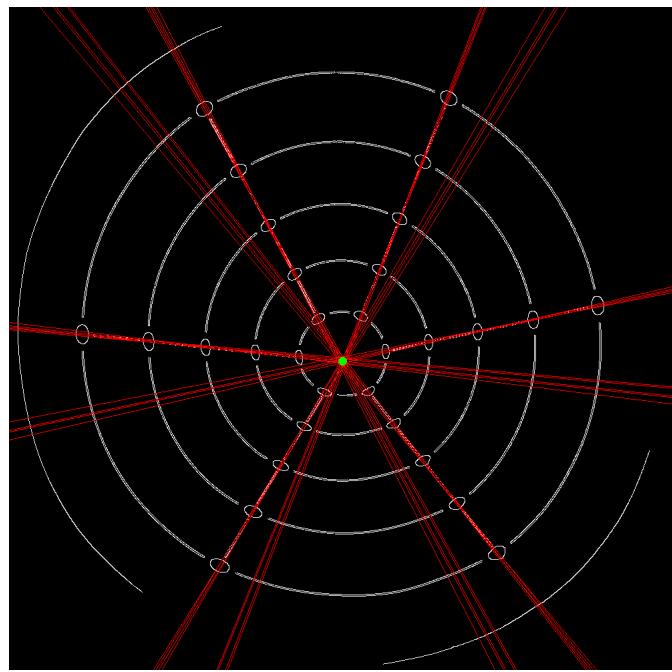
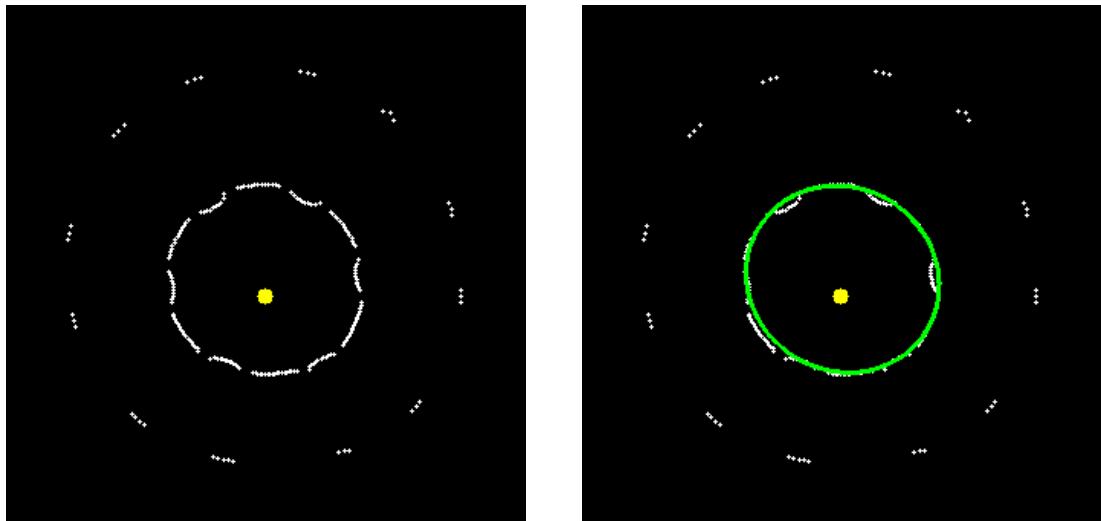


Abbildung 3.6.: Hough-Transformation zur Linien-Detektion (in rot gekennzeichnet) und bestimmter Schnittpunkt (in grün)

Strahlen ausgesendet. Trifft ein Strahl ein weißes Pixel, wird dessen Position gekennzeichnet, trifft er den Rand des Bildes, wird er ignoriert. In Abbildung 3.7(a) sind die getroffenen weißen Pixel und der zugehörige Aussendepunkt eingezeichnet.



(a) bestimme Pixel-Positionen

(b) bestimme Ellipse (grün)

Abbildung 3.7.: Ellipsendetektion: bestimme Pixel-Positionen (weiß), Aussendepunkt (gelb)

Mit Hilfe der Positionen der weißen Pixel, wird anschließend durch RANSAC (siehe Kapitel 2.7) eine Ellipse geschätzt.

Es wird konkret für sechs zufällig ausgewählte Punkte das lineare Gleichungssystem

$$\begin{pmatrix} x_1^2 & y_1^2 & x_1y_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_6^2 & y_6^2 & x_6y_6 & x_6 & y_6 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

gelöst, was auf der Gleichung 2.19 aus Kapitel 2.8 basiert. Nach dem Lösen wird geprüft, ob es sich tatsächlich um eine Ellipse handelt und mittels Hauptachsentransformation (siehe Kapitel 2.8) in die Ellipsenform (x_0, y_0, a, b, ϕ) umgewandelt. Wir verwerfen anschließend Ellipsen mit, dessen Hauptachse wesentlich größer ist, als ihre Nebenachse, da wir eher wenig gestauchte Ellipsen erwarten.

Um die, für RANSAC benötigte, Distanz zu berechnen, wird das Verfahren aus Kapitel 2.8.2 genutzt, was die exakte euklidische Distanz eines Punktes zu einer Ellipse approxmiert. Ein Verfahren wie das Verfahren der kleinsten Quadrate anstelle von RANSAC funktioniert hier nicht, da die weißen Pixel bezüglich einer zu bestimmenden Ellipse, ausreißerbehaftet sind. Wird zum Beispiel auf Grund schlechter Lichtverhältnisse eine Kreislinie nicht deutlich aufgenommen, kann es in dem Kantenbild (Abbildung 3.5) zu „Löchern“ in den Kreislinien kommen und folglich treffen die ausgesendeten Strahlen die nächst äußere Kreislinie (siehe Abbildung 3.8).

Da die Laufzeit nicht im Vordergrund steht, kann eine großzügige Schätzung des Fehleranteils von $\epsilon = 0.4$ mit einer gewünschten Wahrscheinlichkeit $p = 0.9999$ gewählt werden, was zu einer Mindestanzahl an Iterationen von circa 200 führt (siehe Kapitel 2.7). Die letztendlich bestimmten Ellipsen sind beispielhaft in Abbildung 3.9 zu sehen.

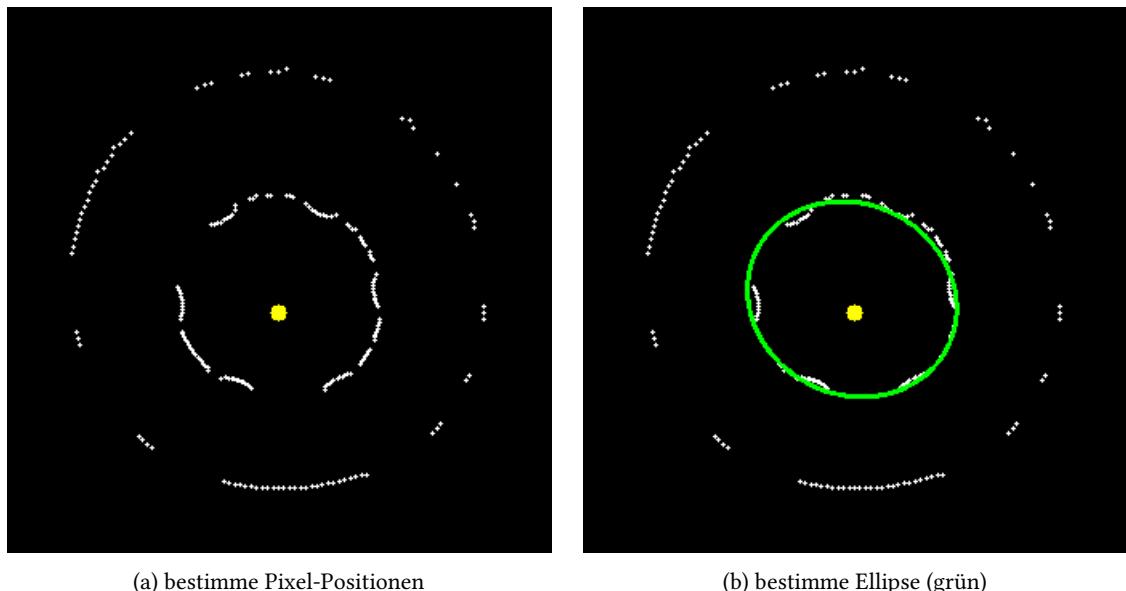


Abbildung 3.8.: Ellipsendetektion bei Ausreißern

3.4.2. Analytical Deformable Templates

Als Alternative zu RANSAC kann man die Ellipsen auch mittel *Analytical Deformable Templates* (siehe Kapitel 2.9) detektieren. Wir betrachten dazu folgende Funktion, deren Nullstellen wie in Kapitel 2.8, eine

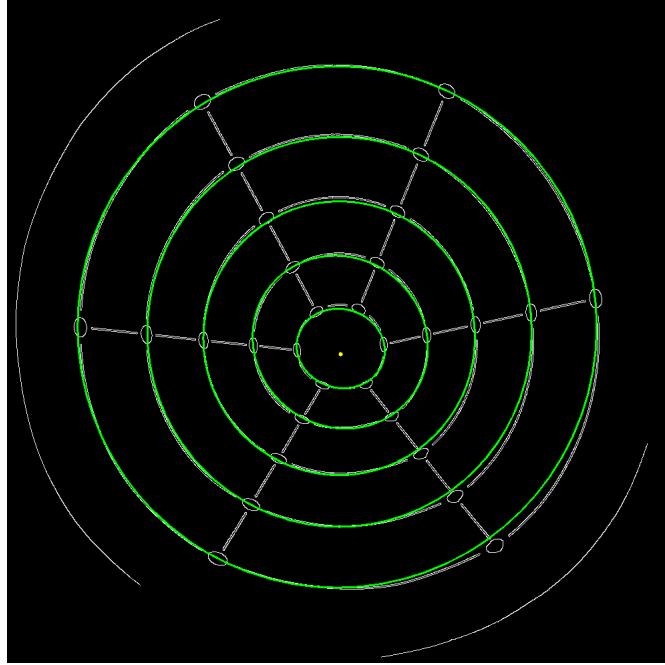


Abbildung 3.9.: detektierte Ellipsen

um θ gedrehte Ellipse, mit Zentrum (x_0, y_0) und Haupt- und Nebenachse a und b beschreiben.

$$G(x, y) = \frac{((x - x_0) \cos \theta + (y - y_0) \sin \theta)^2}{a^2} + \frac{((x - x_0) \sin \theta - (y - y_0) \cos \theta)^2}{b^2} - 1$$

Wir konstruieren eine Energiefunktion:

$$E = E_M + E_A + E_S,$$

die sich zusammensetzt aus einem Term E_M , der die Kantenstärke auf dem Ellipsenrand maximiert, sowie einem Term E_A der die Winkelähnlichkeit zwischen Ellipsennormale und Gradientenrichtung maximiert. Darüber hinaus fügen wir einen Term E_S hinzu, der die Ellipse schrumpfen lässt.

Genauer sehen die Terme wie folgt aus

$$\begin{aligned} E_M &= -\alpha \frac{1}{n} \sum_{i=0}^{n-1} I_M(p_i) \\ E_A &= \beta \frac{1}{n} \sum_{i=0}^{n-1} \left(I_O(p_i) - \text{atan2}\left(\frac{\partial G}{\partial y}(p_i), \frac{\partial G}{\partial x}(p_i)\right) \right)^2 \\ E_S &= \gamma \frac{1}{n} \sum_{i=0}^{n-2} |p_i - p_{i+1}|, \end{aligned}$$

wobei n die Anzahl der Punkte auf dem Ellipsenrand sind. p_i ist definiert als der i -te Punkt auf der Ellipse (siehe parametrische Darstellung einer Ellipse). Man wählt dabei als Startwert eine Ellipse, deren Zentrum dem Bildzentrum entspricht und deren Hauptachse und Nebenachse möglichst groß sind und bestimmt dann numerisch, beispielsweise durch Gradient Descent, ein Minimum der Funktion. Der Term E_M wird minimal, wenn entlang der Ellipse die Kantenstärke (*Magnitude*) groß ist. Der Term E_A wird minimal wenn

die Winkel der Normalenvektoren ähnlich zu denen der Kanten sind (*Angle*) und E_S wird minimal, wenn die Distanz aufeinanderfolgender Ellipsepunkte klein wird, also die Ellipse als ganzes klein wird (*Size*). Der Term lässt die Ellipse also schrumpfen. α , β und γ steuern hierbei den Einfluss der einzelnen Terme. Das Minimum der Funktion sollte die äußerste Ellipse im Kalibrierungsmuster sein. Anschließend wird diese Ellipse vom Bild entfernt und man sucht wiederholt Minima der Funktion bis alle Ellipse detektiert wurden.

3.5. Zuordnung der Punkte

Nach der Bestimmung der Ellipsen muss jede Sample-Positionen der zugehörigen Kreislinie, sowie Liniensegment zugeordnet werden, um seine Position auf dem Kegel bestimmen zu können. Zunächst wird für jeden Punkt diejenige Kreislinie ausgewählt, dessen zugehörige Ellipse die kürzeste Distanz zu ihm hat (siehe Abbildung 3.11(a)).

Mit Hilfe dieser Zuordnung können die Ellipsen aus Kapitel 3.4 erneut geschätzt werden. Diesmal wird das Verfahren der kleinsten Quadrate genutzt, da nur die ausreißerfreien Samples als Messdaten dienen und wir eine optimale Lösung für alle Samples anstreben.

Um nun die Samples auch ihren Liniensegmenten zuzuordnen, wird zunächst der Mittelpunkt der Samples auf der innersten Ellipse bestimmt. Anschließend werden die Samples auf der innersten Ellipse nach dem Winkel der Verbindungslinien zwischen Sample und Mittelpunkt mit der X-Achse sortiert. Die restlichen Samples können nicht nach dem gleichen Schema sortiert werden, da der bestimmte Mittelpunkt nicht der genaue Schnittpunkt aller Liniensegmente ist. Der Mittelpunkt liegt also nicht auf der Verlängerung eines Liniensegments. Die Winkel der Verbindungslinien der Samples (auf einem gemeinsamen Liniensegment) und dem Mittelpunkt mit der X-Achse sind nicht identisch (siehe Abbildung 3.10).

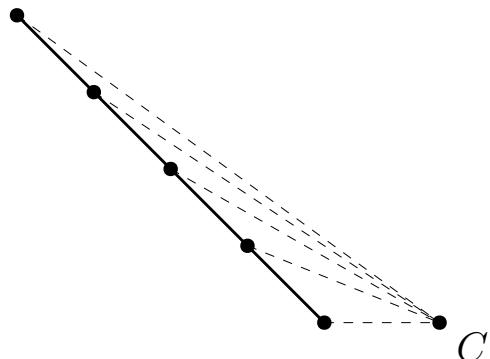


Abbildung 3.10.: Winkel zwischen Samples und Mittelpunkt C sind nicht identisch

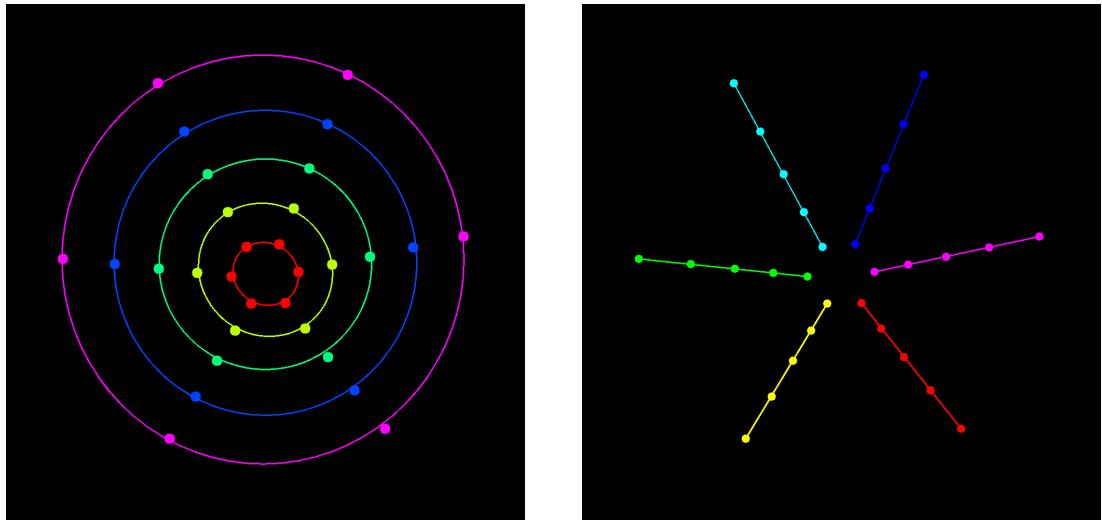
Stattdessen wird für jedes Sample auf den darauffolgenden Ellipsen das Sample auf der vorherigen Ellipse mit der kürzesten Distanz bestimmt. Die Samples können nun entsprechend sortiert werden. Die zugeordneten Liniensegmente sind exemplarisch in Abbildung 3.11(b) zu sehen.

3.6. Weltkoordinaten bestimmen

Mit Hilfe der Parametrisierung des Kegelstumpfs aus Kapitel 2.5 können wir die 3D-Koordinaten eines Samples folgendermaßen angeben:

Ohne Beschränkung der Allgemeinheit, seien die Ellipsen $i = 0, \dots, n - 1$ aufsteigend nach ihrer „Größe“² sortiert, so wie es das Verfahren in 3.4 beschreibt. Außerdem seien die Liniensegmente $j = 0, \dots, m - 1$ aufsteigend nach Winkel mit der X-Achse, wie in 3.5 beschrieben, sortiert. Eine Sample kann also eindeutig

²Etwas formaler, könnte man die Ellipsen hier nach ihrem Flächeninhalt sortieren. Für Ellipsen $E_0(x_0, y_0, a_0, b_0, \theta_0)$ und $E_1(x_1, y_1, a_1, b_1, \theta_1)$ gilt $E_0 \leq E_1$ g.d.w. $\pi \cdot a_0 \cdot b_0 \leq \pi \cdot a_1 \cdot b_1$



(a) Zuordnung von Punkten zu Ellipsen

(b) Zuordnung von Punkten zu Liniensegmenten

Abbildung 3.11.: Zuordnung von Punkten zu Ellipsen (links) und Liniensegmenten (rechts)

durch ein Tupel $(i, j) \in [0, n-1] \times [0, m-1]$ identifiziert werden und (x_{ij}, y_{ij}, z_{ij}) bezeichne seine Koordinaten im Weltkoordinatensystem.

Analog zur parametrischen Darstellung von Kegelstümpfen (Gleichung 2.2) in Kapitel 2.1 ergibt sich:

$$\begin{aligned} x_{ij} &= r_i \cos \theta_j \\ y_{ij} &= h_i \\ z_{ij} &= r_i \sin \theta_j \end{aligned}$$

$\forall (i, j) \in [0, n-1] \times [0, m-1]$ mit

$$\begin{aligned} r_i &= r + \frac{i}{n} \cdot (R - r) & \forall i \in [0, n-1] \\ h_i &= \frac{i}{n} \cdot \Delta H & \forall i \in [0, n-1] \\ \theta_j &= \frac{j}{m-1} \cdot 2\pi & \forall j \in [0, m-1] \end{aligned}$$

3.7. Entfaltung

Die eigentliche Entfaltung des Kegels kann mit zwei unterschiedlichen Ansätzen realisiert werden. Die erste Möglichkeit ist die *Vorwärtsentfaltung*. Hierbei wird für jedes Pixel auf dem Kegelbild eine 3D-Koordinate durch geeignete Interpolation bestimmt und dann auf die Mantelfläche abgebildet. Beim zweiten Ansatz, der *Rückwärtsentfaltung*, wird ein Punkt von der Mantelfläche zurück auf den Kegel abgebildet und von dort mit einer Projektionsmatrix auf die Bildebene projiziert und dann interpoliert.

Im folgenden wird genauer auf beide Verfahren eingegangen, sowie deren Probleme erläutert.

3.7.1. Vorwärtsentfaltung

Bei der *Vorwärtsentfaltung* muss wie oben erwähnt zu jedem Pixel die zugehörige 3D Koordinate im Weltkoordinatensystem berechnet werden. Da bisher jedoch nur die Positionen der Samples bekannt sind muss hier

Zunächst betrachten wir diejenigen Pixel, die sich weder auf einer Kreislinie, noch auf einem Liniensegment befinden. Es gibt zu einem Pixel P also immer vier Sample-Nachbarn (bl, br, tr, tl). Diese Situation ist in Abbildung 3.12 illustriert.

Nachdem die vier Nachbarn bestimmt wurden, können im ersten Schritt die Abstände d_1 und d_2 zu inneren Ellipse E_b , respektive äußerer Ellipse E_t berechnet werden. Mithilfe dieser Abständen kann nun eine *Interpolationsellipse* E_{int} definiert werden als

$$E_{int} = \left(\frac{d_1}{d_1 + d_2} \right) \cdot E_t + \left(\frac{d_2}{d_1 + d_2} \right) E_b,$$

wobei eine Multiplikation mit einem Skalar alle Charakteristika einer Ellipse skaliert. Der Drehwinkel θ wird hierbei bei 2π umgebrochen. Eine Addition geschieht elementweise. Im nächsten Schritt wird der Schnittpunkt L mit dem Liniensegment \overline{bltl} , sowie der Schnittpunkt R mit dem Liniensegment \overline{brtr} bestimmt.

Da sich tl, L und bl nun auf einem gemeinsamen Liniensegment befinden, kann bezüglich der Weltkoordinaten linear interpoliert werden. Analoges gilt für tr, R und br .

Die drei Punkte (L, P, R) befinden sich auf der Interpolationsellipse und somit können die Winkel (ϕ_L, ϕ_P, ϕ_R) bezüglich des gemeinsamen Ellipsenkoordinatensystems bestimmt werden.

Analog zu oben werden die 3D-Koordinaten von P als lineare Interpolation zwischen den gerade bestimmten 3D-Koordinaten von L und R bestimmt. Als Interpolationsfaktor benutzen wir $\frac{\phi_L - \phi_P}{\phi_L - \phi_R}$

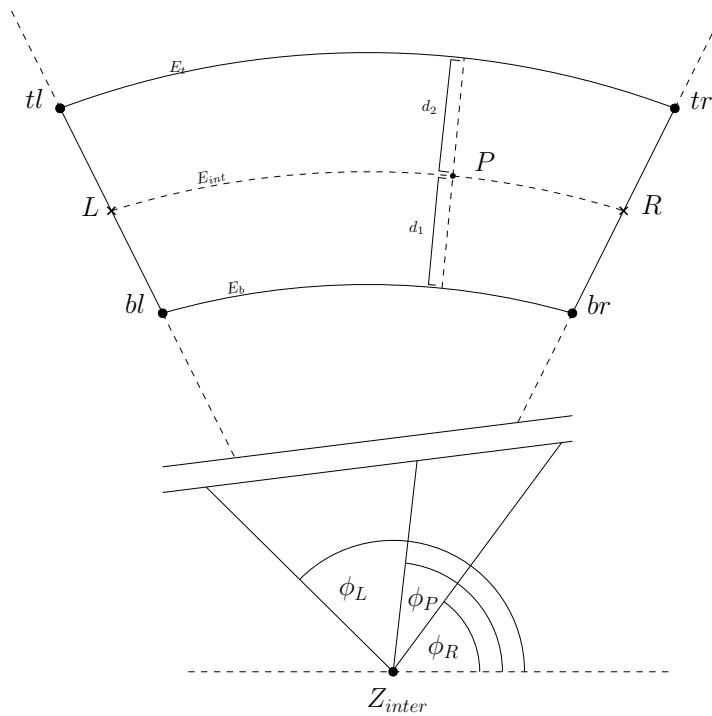


Abbildung 3.12.: Interpolation der 3D-Koordinaten

Analog werden die auf Liniensegmenten befindlichen Punkte einfach linear interpoliert. Die Punkte, die sich auf Kreislinien befinden werden über Winkel interpoliert.

Jedes Pixel hat nun 3D-Koordinaten im Kegel erhalten, die beispielhaft in Abbildung 3.13 zu sehen sind. Anhand der Abbildung lässt sich das erste Problem bei der Vorwärtsentfaltung feststellen. Zwischen den Samples auf einer Kreislinie sollten die interpolierte Positionen nach außen gewölbt sein, da ein Kegel rund ist. Da wir jedoch Interpolieren und nicht Extrapolieren, gehen die Werte nie über die zur Interpolation genutzten Werte hinaus. Konkreter kann eine interpolierte Position keine größeren oder kleineren X und Z

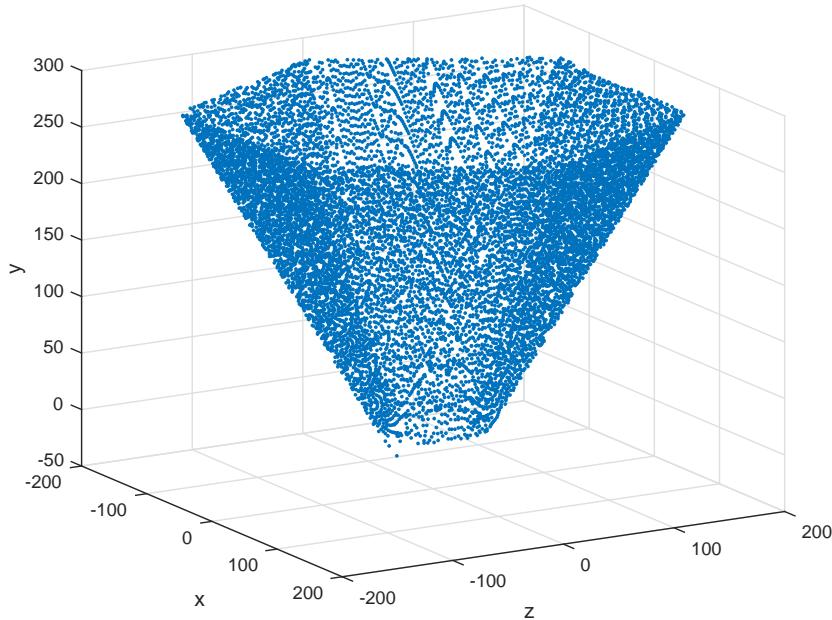


Abbildung 3.13.: interpolierter Kegel

Koordinaten erhalten, als die der genutzten Samples. Diese wäre jedoch notwendig, sodass die Rundung des Kegels erhalten bleibt. Die Oberfläche des Kegels scheint eckig.

Neben der außerdem sehr hohen Laufzeit, bedingt durch die komplexe Interpolation, ergibt sich noch ein weiteres Problem, dass sich erst bei der Entfaltung ergibt.

Die eigentliche Entfaltung geschieht über die Abbildung 2.6, die in Kapitel 2.1 konstruiert wurde. Die erhaltenen Werte müssen anschließend skaliert werden, da im entfalteten Bild sonst 1 mm einem Pixel entspreche. Es ergibt sich das entfaltete Bild, wie in Abbildung 3.14 dargestellt. Auch schon bei kleinen Skalierungen entstehen auffällige „Löcher“ im entfalteten Bild. Grund dafür ist, dass jedes Pixel aus dem Ursprungsbild auf eine 2D-Koordinate der Mantelfläche abgebildet wird. Auch nach einer Skalierung handelt es sich bei diesen Werten im Allgemeinen nicht um ganzzahlige Werte. Es muss im entfalteten Bild gerundet werden. Auch wenn Rundungsfehler nicht entstanden, fehlte es einfach an genügend Informationen, gerade in den inneren, kleineren Regionen des Ursprungsbild.

Da wir von dem Ursprungsbild aus auf das entfaltete Bild abbilden, ist außerdem eine Interpolation auf dem Ursprungsbild nicht möglich, da wir vorher nicht genau wissen, wo Löcher entstehen. Man müsste als entweder auf dem resultierendem Bild interpolieren (siehe Kapitel 6), oder zu gegebenen Löchern über die Umkehrabbildung auf dem Ursprungsbild interpolieren.

Es bietet sich dann jedoch an, direkt die Umkehrabbildung zu nutzen, was die Motivation hinter der Rückwärtsentfaltung ist.

3.7.2. Rückwärtsentfaltung

Bei der Rückwärtsentfaltung gehen wir von dem entzerrten Bild aus, dessen geometrischen Eigenschaften aus der parametrischen Form der Mantelfläche des Kegels bekannt sind und versuchen rückwärts Pixelkandidaten im Ursprungsbild zu bestimmen.

Zunächst bestimmen wir zu den gegebenen 3D-Korrespondenzen der Samples eine Projektionsmatrix wie in Kapitel 2.3 mittels *Direct Linear Transformation*. Wir machen also eine Kamerakalibrierung, wobei die

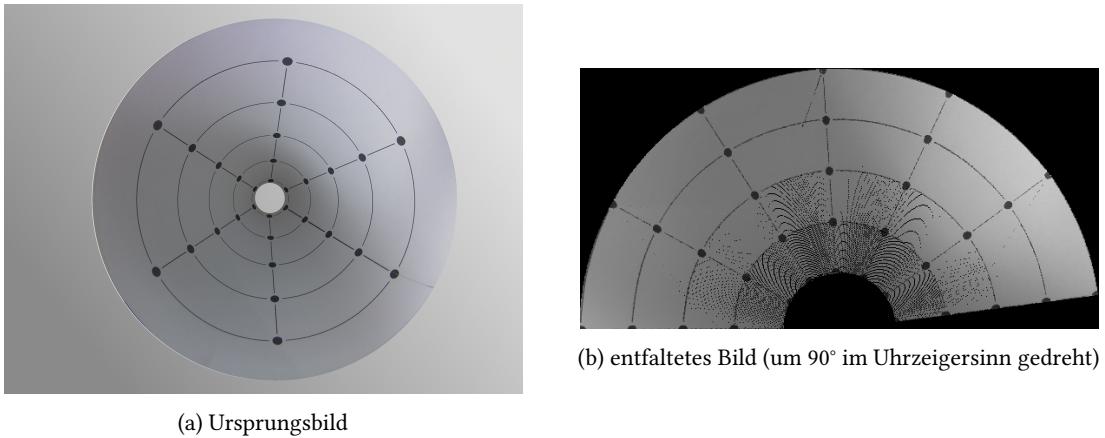


Abbildung 3.14.: Vorwärtsentfaltung

Kameraverzerrungen schon herausgerechnet wurden (siehe Kapitel 3.2). Wir erhalten somit eine Abbildung die aus den 3D-Koordinaten im Kegel auf Bildkoordinaten abbildet.

Wir benutzen anschließend für jedes Pixel auf dem entzerrten Bild die Umkehrabbildung 2.9 aus Kapitel 2.1. Wir erhalten wieder 3D-Kegelkoordinaten. Die Projektionsmatrix kann nun einfach auf diese Punkte angewandt werden und wir erhalten Bildpunkte im Ursprungsbild.

Die resultierenden Bildpunkte sind im Allgemeinen nicht ganzzahlig. Im Gegensatz zur Vorwärtsentfaltung kann hier jedoch einfach interpoliert werden.

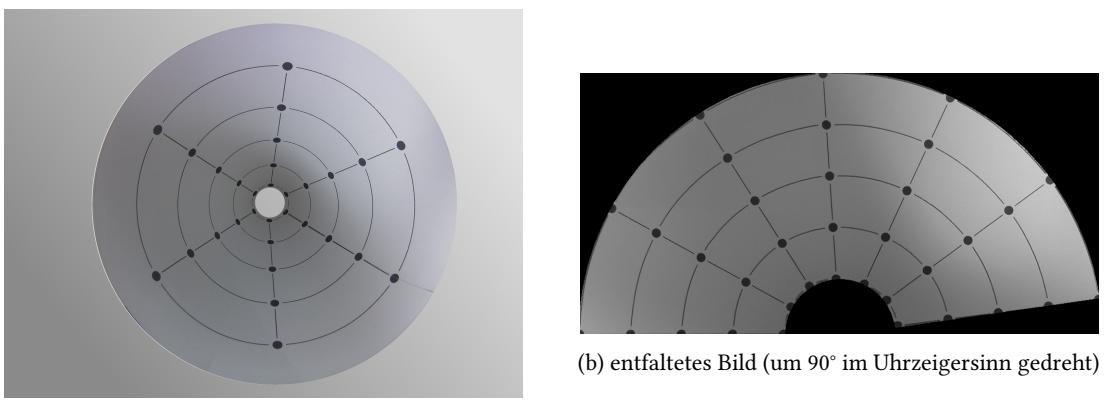


Abbildung 3.15.: Rückwärtsentfaltung

4. Implementierung

Zwecks Vereinfachung des Kalibrierungsprozesses wurde ein Assistent in Form einer graphischen Benutzeroberfläche programmiert.

Im ersten Schritt des Assistenten wird dabei optional eine intrinsische Kamerakalibrierung durchgeführt (siehe Abbildung 4.1).

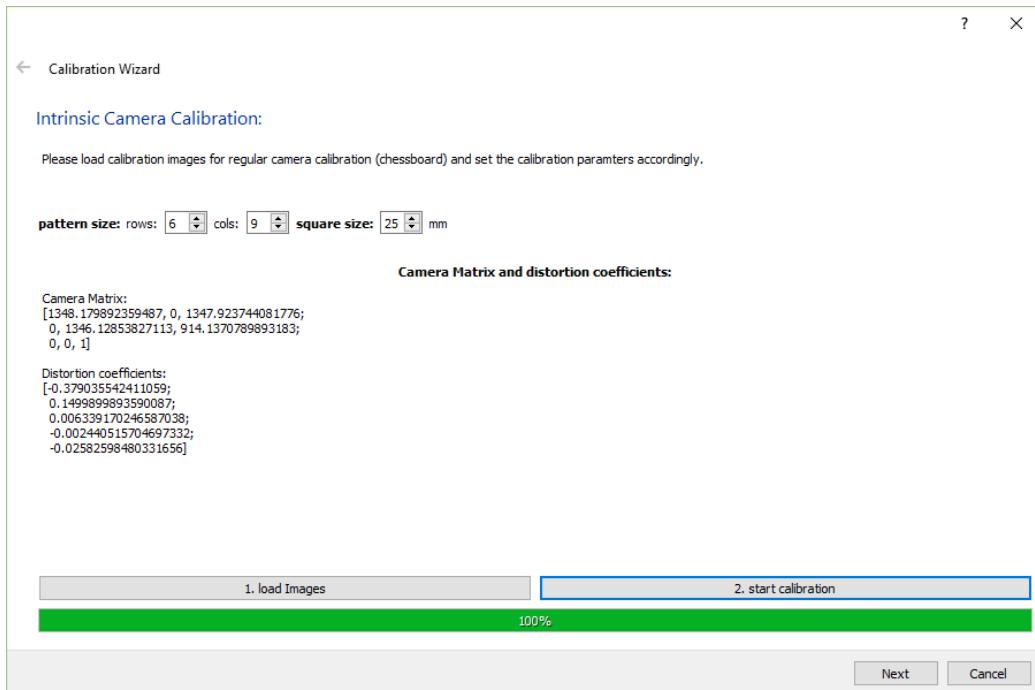


Abbildung 4.1.: Kalibrierungsassistent: intrinsische Kamerakalibrierung

Anschließend wird die eigentliche Kegelkalibrierung ausgeführt. Dabei wird zunächst das Kalibrierungsbild geladen und gegebenenfalls nach einer stattgefundenen intrinsischen Kalibrierung entzerrt. Es werden nun die Sample-Positionen detektiert und der Nutzer hat die Möglichkeit zu überprüfen, ob alle Positionen korrekt detektiert wurden und andernfalls fehlerhafte Punkte zu entfernen und / oder Punkte hinzuzufügen. Im Anschluss werden die Ellipsen und Liniensegmente bestimmt und Punktcorrespondenzen hergestellt (siehe Abbildung 4.2).

Im letzten Schritt kann zwischen beiden Entfaltungsverfahren gewählt werden. Anschließend können die Einstellungen in eine XML-Datei exportiert werden, in der, neben der Kamera-Matrix und Verzerrungskoeffizienten, auch die zwei Abbildungsmatrizen des ausgewählten Verfahrens gespeichert werden.

Die Abbildungsmatrizen sind dabei wie folgt aufgebaut. Als dst bezeichnen wir das entfaltete Bild. src ist das Ursprungsbild.

Bei der Vowärtsentfaltung gilt:

$$dst(map_x(x, y), map_y(x, y)) = src(x, y),$$

wobei map_x und map_y die Abbildungsmatrizen sind und die gleiche Größe wie das Ursprungsbild haben.

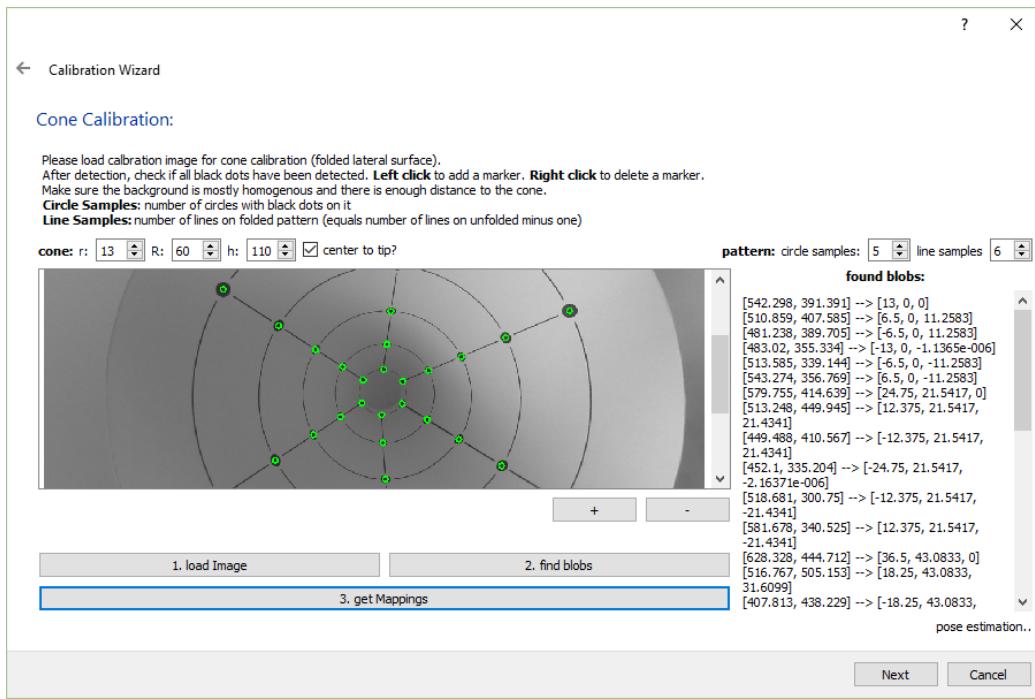


Abbildung 4.2.: Kalibrierungsassistent: Kegelkalibrierung

Da wird bei der Entfaltung die Größe des Ergebnisbildes benötigen, und wir diese bei der Erstellung der Abbildungsmatrizen berechnet haben, sind Breite und Höhe, in $map_x(0, 0)$, respektive $map_y(0, 0)$, kodiert. Wir verlieren damit die Information $dst(0, 0)$ im Ergebnisbild. Dieses Pixel ist aber ohnehin null (siehe zum Beispiel Abbildung 3.14 in Kapitel 3.7).

Bei der Rückwärtsentfaltung gilt:

$$dst(x, y) = src((map_x(x, y), map_y(x, y)),$$

wobei map_x und map_y wieder die Abbildungsmatrizen sind und hier die gleiche Größe wie das entfaltete Bild haben. Eine Kodierung wie bei der Vowärtsentfaltung ist hier also nicht notwendig.

5. Analyse

In diesem Kapitel...

Zunächst werden die beiden Verfahren zur Entzerrung gegenübergestellt. Anschließend untersuchen wir verschiedene Einflüsse auf der Qualität der Entzerrung, woraufhin wir die Laufzeit ermitteln. Zum Schluss Ransac...

Die Laufzeitanalysen wurden auf einem Raspberry Pi 2 Model B durchgeführt mit folgenden Eigenschaften:

- 900MHz ARM Cortex-A7 CPU
- 1GB RAM
- GCC-4.9.2
- OpenCV 2.4.13

5.1. Vergleich Vorwärtsentfaltung und Rückwärtsentfaltung

In Kapitel 3 haben wir festgestellt, dass das entfaltete Bild mittels Vorwärtsentfaltung Löcher enthält. Es ist hierbei von Interesse wie sich die Anzahl der Löcher bei einer Änderung der Auflösung verhält. Der Einfluss der Ausgabeauflösung auf die Anzahl der Löcher lässt sich in Abbildung 5.1 ablesen.

Wie erwartet verhält sich die Anzahl der Löcher quadratisch zur gewählten Ausgabeauflösung. Der Informationsgehalt (Pixelanzahl) des Ursprungsbildes bleibt natürlich trotz erhöhter Ausgabeauflösung konstant. Da sich die Auflösung der Seitenhöhe quadratisch zur Gesamtanzahl der Pixel verhält, ist auch ein quadratisches Wachstum der Anzahl der Löcher zu erwarten.

Als Reprojektionsfehler einer Abbildung wird die Distanz zwischen einem gemessenen Punkt und einem korrespondierendem projizierten Punkt bezeichnet.

Im Falle der Rückwärtsentfaltung sind die gemessenen Punkte die Bildpositionen der Samples im Ursprungsbild. Da die Geometrie des Kegels bekannt ist, wissen wir wo sich die Samples auf dem entfalteten Bild befinden müssen (siehe Parametrisierung der Mantelfläche 2.3 in Kapitel 2). Wir können nun diese Positionen mit Hilfe der Abbildung zur Entfaltung und der Projektionsmatrix zurück auf das Ursprungsbild abbilden. Diese Punkte sind die projizierten Punkte. Da die Abbildung von der Mantelfläche zur Kegeloberfläche exakt ist, ist der Reprojektionsfehler der Rückwärtsentfaltung alleine durch die Projektionsmatrix bestimmt.

Bei der Vorwärtsentfaltung sind die gemessenen Punkte gegeben durch die bekannten Sample-Positionen auf der Mantelfläche. Die Projizierten erhält man, nach der Abbildung der detektierten Sample-Positionen des Ursprungsbild auf die Mantelfläche. Der Reprojektionsfehler ist also alleine durch die Genaugkeit der Sample-Detektion definiert und somit bei diesem Verfahren immer nahe null. Obwohl das Endergebnis, bedingt durch die Löcher, bei der Vorwärtsentfaltung optisch schlechter ist, ist der Reprojektionsfehler also bei der Vorwärtsentfaltung immer kleiner. Als Vergleich zwischen den beiden Verfahren eignet sich der Reprojektionsfehler also nicht.

Wir entscheiden uns optisch und auf Grund der schlechteren Laufzeit bei der Vorwärtsentfaltung (siehe Abbildung 5.2) für die Rückwärtsentfaltung. Alle weiteren Auswertungen beziehen sich von nun an auf die Rückwärtsentfaltung.

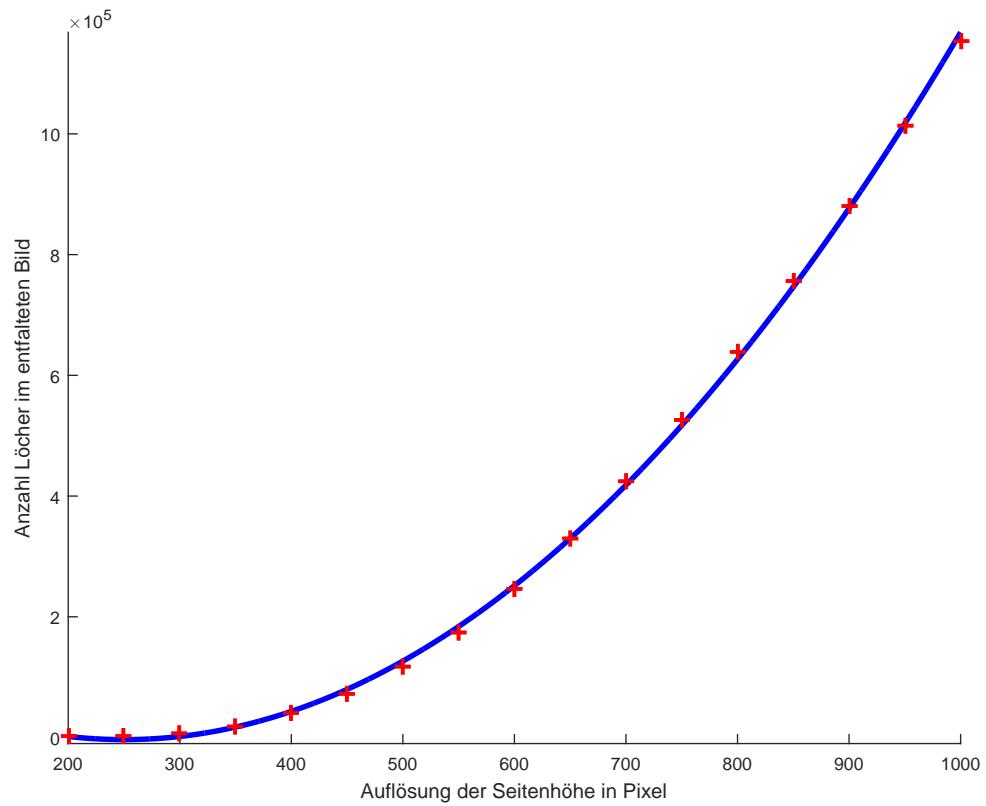


Abbildung 5.1.: Einfluss der Ausgabeauflösung auf die Anzahl der Löcher

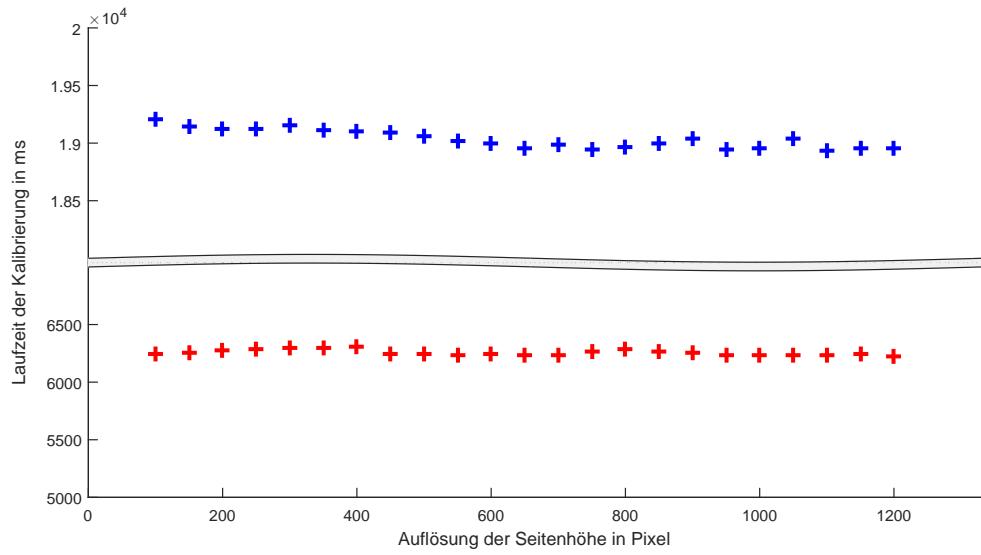


Abbildung 5.2.: Laufzeitvergleich zwischen Vorwärtsentfaltung (blau) und Rückwärtsentfaltung (rot)

5.2. Einfluss der intrinsischen Kalibrierung

Ein wichtiger Einflussfaktor auf die Qualität der Entfaltung ist die intrinsische Kamerakalibrierung, die vor der eigentlichen Kegelkalibrierung stattfindet. Ihre Hauptaufgabe besteht darin, die Linsenverzerrungen der Kamera herauszurechnen.

Wir messen den Einfluss der Kamerakalibrierung mit Hilfe des Reprojektionsfehlers. Dazu betrachten wir fünf verschiedene Bilder, die ein Mal mit und ein Mal ohne intrinsische Kalibrierung entzerrt werden. Bei beiden Gruppen wird anschließend jeweils der Reprojektionsfehler pro Sample bestimmt und verglichen. Die Ergebnisse sind in Abbildung 5.3 zu sehen. In der linken Abbildung sind hierbei die Fehler im kalibrierten Fall zu erkennen, im Linken im Unkalibrierten. In der Abbildung ist dabei ein Kreuz bei (u, v) , falls die Abweichung des projizierten Punktes in x -Richtung u , sowie in y -Richtung v beträgt.

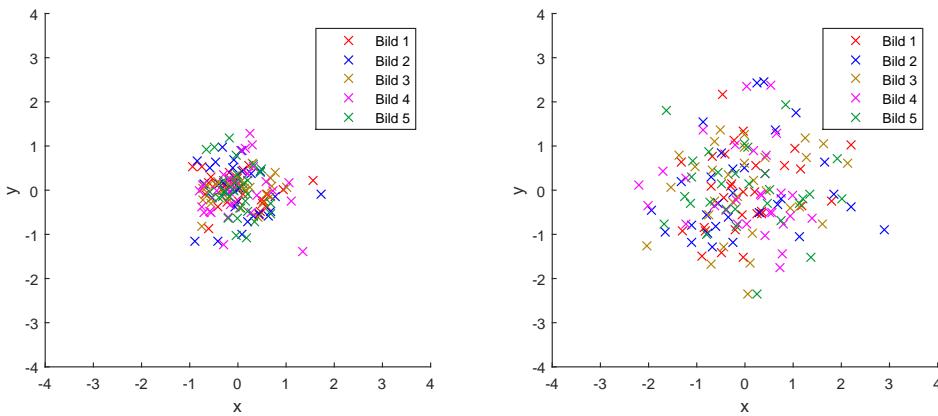


Abbildung 5.3.: Einfluss der intrinsischen Kalibrierung auf den Reprojektionsfehler

Es ist klar zu erkennen, dass die Reprojektionsfehler bei den Bildern ohne intrinsische Kamerakalibrierung wesentlich größer ist. Der starke Einfluss kommt unter Anderem daher, dass wir eine Weitwinkelkamera mit starker tonnenförmiger Verzerrung eingesetzt haben. Ohne eine Modellierung der Linsenverzerrungen weichen die Abstände zwischen den Sample-Positionen stark von der Realität ab. Die Projektionsmatrix wird mit fehlerhaften Daten bestimmt.

5.3. Einfluss der Rotation der Kamera

Um den Einfluss der Rotation der Kamera untersuchen zu können, wurde der Kegel mit Kalibrierungsmuster gerendert, da die Kameraposition dann exakt bekannt ist und äußere Faktoren wie Lichtverhältnisse und inhomogene Hintergründe kontrolliert werden können.

Es werden anschließend Bilder erzeugt, in denen in 1° Schritten die Kamera von 0° bis 12° um die X-Achse rotiert wird. Für jedes dieser Bilder wird anschließend eine Rückwärtsentfaltung durchgeführt und der durchschnittliche Reprojektionsfehler ermittelt. Abbildung 5.5 zeigt, dass der Reprojektionsfehler relativ rotationsinvariant ist.

5.4. Laufzeit der Entfaltung

Durch den Kalibrierungsprozess erhalten wir zwei Abbildungsmatrizen (siehe Kapitel 4), mit denen wir anschließend eine Reihe von Bildern entzerrten möchten. Wir untersuchen dabei die Laufzeit pro gewählter Auflösung, in 50er Schritten, und ermitteln jeweils die Laufzeit gemittelt über 200 Bilder. Untersucht wird die Laufzeit bikubischer, sowie mit linearer Interpolation.

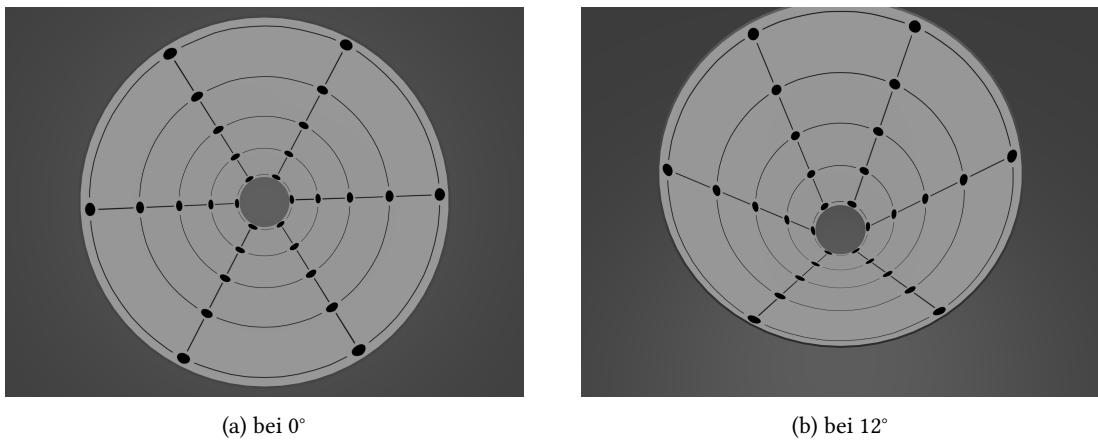


Abbildung 5.4.: gerendeter Kegel mit Kalibrierungsmuster in Blender aus verschiedenen Blickrichtungen

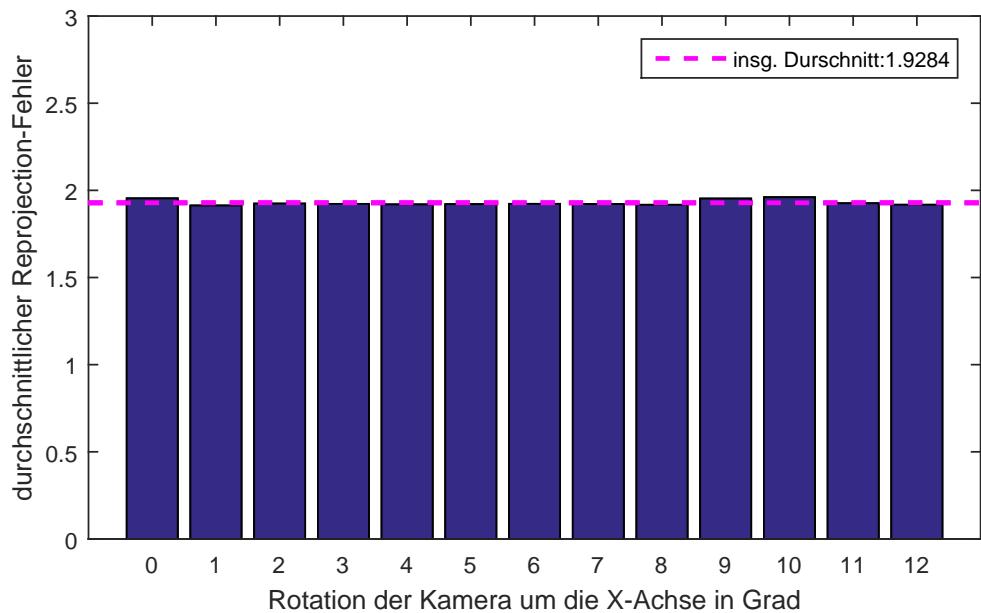


Abbildung 5.5.: Einfluss der Rotation der Kamera

5.5. Evaluierung Deformable Templates

Das Problem bei diesem Ansatz, ist dass die Funktion, auch nach starker Gaussglättung, schnell in kleine lokale Minima läuft, obwohl ein wesentlicher stärkeres Minimum in näherer Umgebung wäre. Darauf hinaus muss für ein robustes Optimierungsverfahren der Gradient und im besten Fall sogar die Hesse-Matrix zur Verfügung gestellt werden.

In Abbildung 5.7 sieht man beispielhaft einen Auschnitt der Energiefunktion, wobei in diesem Fall zwecks Veranschaulichung nur die Haupt- und Nebenachse der Ellipse variable sind. Das Zentrum bleibt das Zentrum des Bildes und der Winkel θ ist konstant null. Trotz dieser Einschränkungen lässt sich gut erkennen, dass circa bei $a = b = 300$ ein starkes Minimum liegt. Es gibt jedoch neben diesem Minimum in näherer Umgebung kleine irrelevante Minima, wie beispielsweise in (360, 260). Solche „Becken“ erschweren die Suche nach der

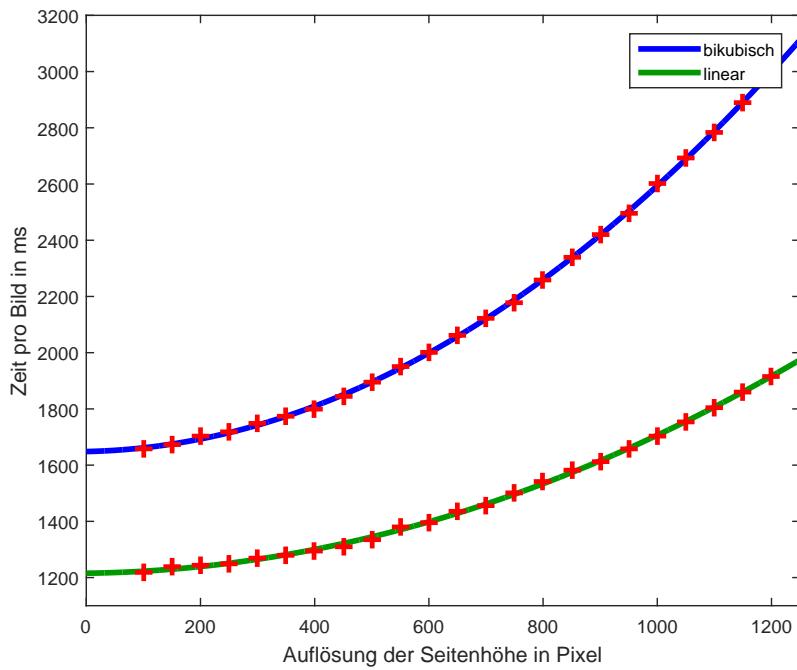


Abbildung 5.6.: Einfluss der Ausgabeauflösung auf die Laufzeit der Entfaltung

nächsten Ellipse erheblich.

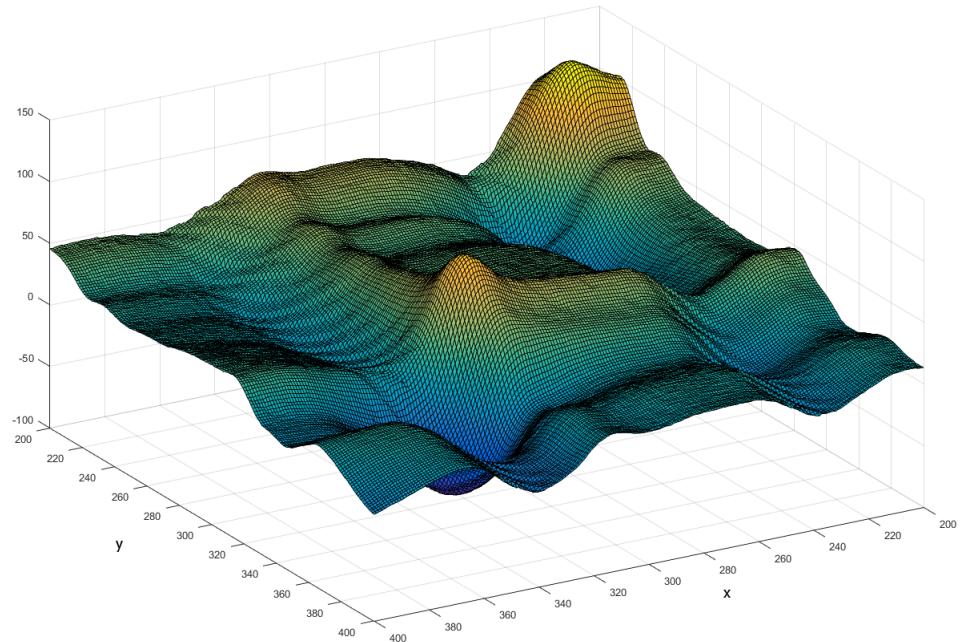


Abbildung 5.7.: Deformable Templates: Ausschnitt der Energiefunktion für variable Haupt- und Nebenachse

5.6. Evaluierung des RANSAC-Verfahrens zur Ellipsendetektion

Die robuste Ellipsendetektion ist ein wichtiger Schritt bei der Entfaltung. Bei beiden Verfahren werden die bestimmten Ellipsen genutzt um Korrespondenzen zwischen den Sample-Positionen und Punkten auf dem Kegel im Weltkoordinatensystem herzustellen. Bei der Vorwärtsentfaltung werden die Ellipsen darüber hinaus benötigt, um für die Pixel geeignete 3D-Koordinaten interpolieren zu können. Es ist also von großer Bedeutung wie gut die Ellipsendetektion mittels RANSAC funktioniert.

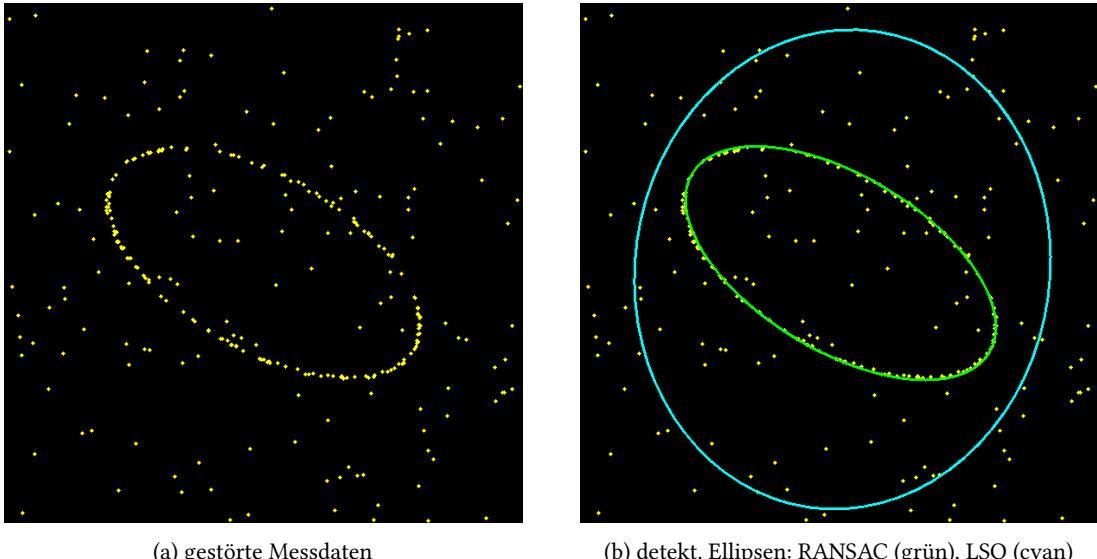


Abbildung 5.8.: Vergleich RANSAC und LSQ bei gleichverteilten Ausreißern $\epsilon = 0.5, p = 0.99$

In unserem Verfahren ist eine Gleichverteilung der Ausreißer unwahrscheinlich. Viel wahrscheinlicher ist es, dass eine der nächst äußeren Ellipsen frühzeitig sichtbar wird (siehe Kapite 3.4).

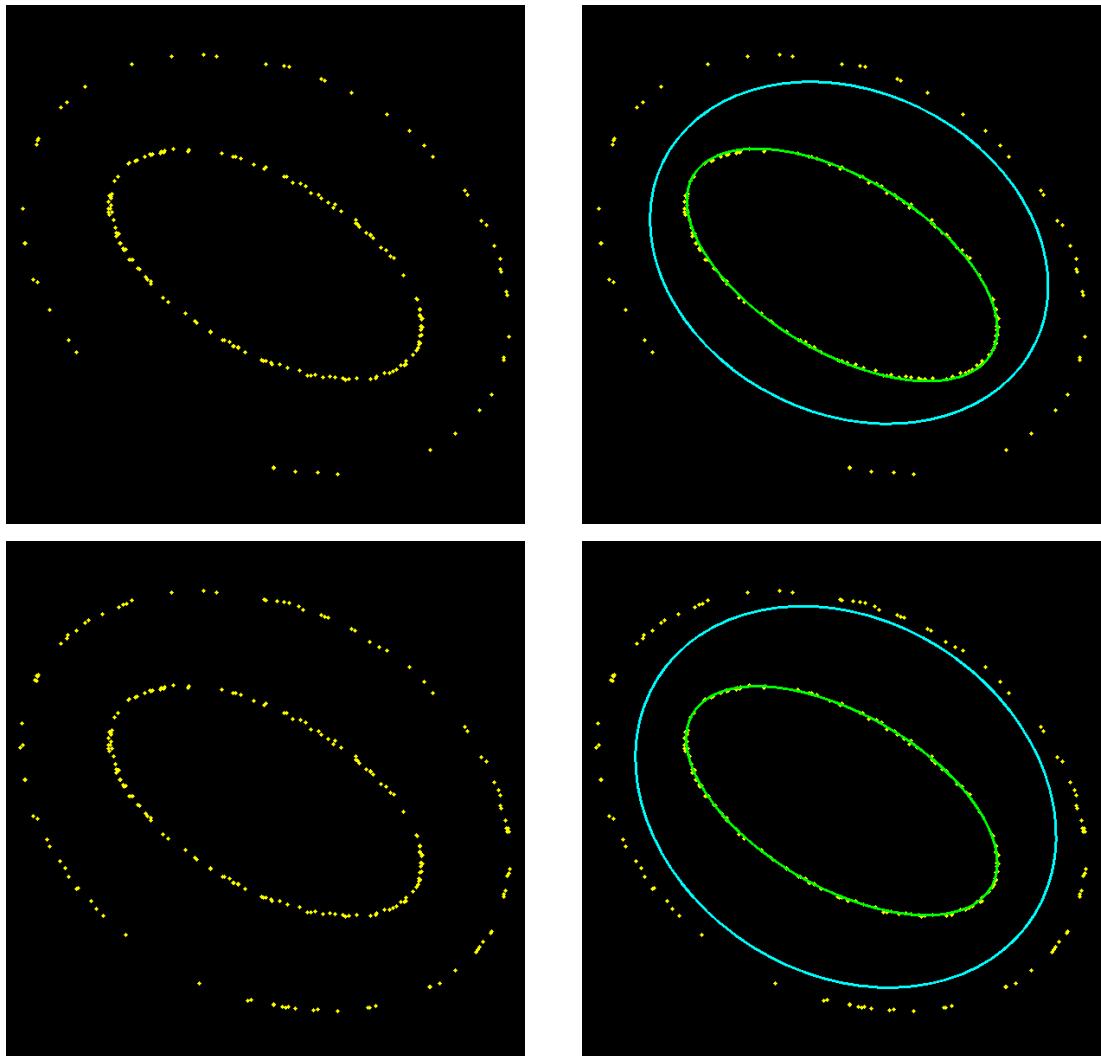


Abbildung 5.9.: Vergleich RANSAC und LSQ bei Schattenellipsen mit $p = 0.99$ und $\epsilon = 0.25$ (oben), $\epsilon = 0.4$ unten, links gestörte Messdaten, rechts detektierte Ellipsen RANSAC (grün), LSQ (cyan)

6. Fazit und Ausblick

6.1. Parallelisierung

eigentlich quatsch, weil auf raspberry pi

6.2. Verbesserung der Linien-Detektion

Um die Linien-Detektion etwas genauer zu machen, könnte man die Gradientenrichtung des Kantenbilds miteinbeziehen. Dies reduziert die Anzahl falscher Votes und verbessert darüber hinaus die Laufzeit [OC76].
probabilistic hough?

6.3. Verbesserung der Vorwärtsentfaltung

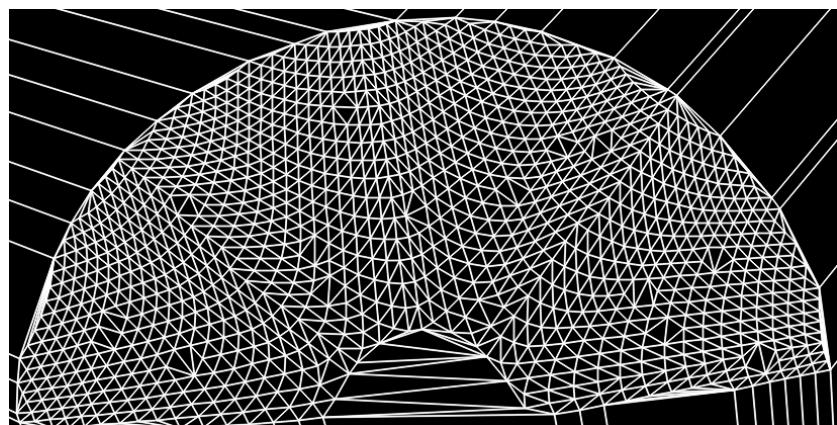
Das Hauptproblem der Vorwärtsentfaltung sind die Löcher auf dem entfalteten Bild. Um diese schließen zu können, könnte man eine Delaunay-Triangulation durchführen. Da das Verfahren jedoch ohne Triangulation schon relativ rechenintensiv ist, und die Rückwärtsentfaltung sehr gute Ergebnisse liefert wurde dieses Möglichkeit nicht weiter untersucht. Solch eine Triangulation ist beispielhaft in Abbildung 6.1 abgebildet. Zum Zwecke der Veranschaulichung wurde hierbei nur ein Teil der Daten benutzt.

6.4. Verbesserung der Rückwärtsentfaltung

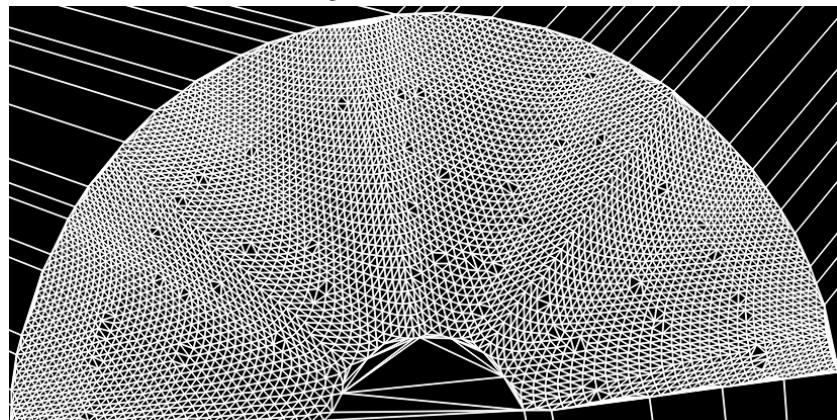
Aktuell wird die Projektionsmatrix bei der Rückwärtsentfaltung durch *Direct Linear Transformation* bestimmt. Dieses Verfahren minimiert jedoch nicht den Reprojektionsfehler. Besser wäre hier deshalb ein iteratives Verfahren, wie der Levenberg–Marquardt-Algorithmus (konkret für Projektionsmatrizen beschrieben in [HZ00]). Im Gegensatz dazu könnte auch ein RANSAC-Ansatz verwendet werden. Statt eine optimale Lösung für alle detektierte Samples zu bestimmen, berechnet man wiederholt für sechs Punkte¹ eine Projektionsmatrix und untersucht dann den Reprojektionsfehler für alle anderen Punkte und wählt schließlich die Projektionsmatrix mit dem größten Consensus Set aus.

Literatur überprüfen, insbesondere Seitenangaben

¹sechs Punkte sind mindestens notwendig um eine Projektionsmatrix bestimmen zu können, da es elf unbekannte gibt (siehe Kapitel 2.3).



(a) Triangulation mit 10% der Punkte



(b) Triangulation mit 40% der Punkte

Abbildung 6.1.: Delaunay-Triangulation

A. Ergebnisse

muss das hier
überhaupt
rein?

res	200	250	300	350	400	450	500	550	600	650	700	750	800	850	900	950	1000
holes	1623	2721	7799	19337	40168	72349	116676	173543	245216	329354	423877	527094	638029	756185	881294	1013880	1153448

Tabelle A.1.: test123

Auflösung	100	150	200	250	300	350	400	450	500	550	600	650	700	750	800	850	900	950	1000	1050	1100	1150	1200
vorwärts	6245	6256	6276	6290	6294	6300	6305	6244	6240	6239	6243	6236	6236	6270	6235	6266	6232	6238	6237	6237	6239	6240	6224
rückwärts	19202	19147	19126	19127	19154	19114	19102	19092	19056	19013	18993	18952	18989	18949	18964	19002	19034	18949	18956	19043	1893	18951	18954

Tabelle A.2.: test445

Abbildungsverzeichnis

2.1.	Gerader Kreiskegel	3
2.2.	Kegelstumpf und Ergänzungskegel	4
2.3.	Kegelstumpf	5
2.4.	Kegelmantelfläche	5
2.5.	Abbildung der Kegelstumpfhöhe auf die Seitenhöhe	6
2.6.	Lochkameramodell	9
2.7.	Linsenverzerrungen	10
2.8.	Ellipse mit Brennpunkten f_1, f_2 , Zentrum C , Hauptachse a , Nebenachse b und Scheitelpunkten V_1 und V_2	12
2.9.	Ellipse mit Zentrum (x_0, y_0) , Hauptachse a , Nebenachse b , sowie Drehwinkel θ	13
2.10.	Ellipsenausschnitt im ersten Quadranten mit Abfragepunkt Q und eingezeichneter kürzester Distanz zur Ellipse	16
3.1.	Kalibrierungsmuster von oben mit $n = 5, m = 6$	20
3.2.	Kalibrierungsmuster entfaltet mit $n = 5, m = 6$	20
3.3.	Kamerakalibrierung	21
3.4.	Detektion der Samples	22
3.5.	Canny-Kantendetektion auf Grauwertbild	22
3.6.	Hough-Transformation zur Linien-Detektion (in rot gekennzeichnet) und bestimmter Schnittpunkt (in grün)	23
3.7.	Ellipsendetektion: bestimme Pixel-Positionen (weiß), Aussendepunkt (gelb)	23
3.8.	Ellipsendetektion bei Ausreißern	24
3.9.	detektierte Ellipsen	25
3.10.	Winkel zwischen Samples und Mittelpunkt C sind nicht identisch	26
3.11.	Zuordnung von Punkten zu Ellipsen (links) und Liniensegmenten (rechts)	27
3.12.	Interpolation der 3D-Koordinaten	28
3.13.	interpolierter Kegel	29
3.14.	Vorwärtsentfaltung	30
3.15.	Rückwärtsentfaltung	30
4.1.	Kalibrierungsassistent: intrinsische Kamerakalibrierung	31
4.2.	Kalibrierungsassistent: Kegelkalibrierung	32
5.1.	Einfluss der Ausgabeauflösung auf die Anzahl der Löcher	34
5.2.	Laufzeitvergleich zwischen Vorwärtsentfaltung (blau) und Rückwärtsentfaltung (rot)	34
5.3.	Einfluss der intrinsischen Kalibrierung auf den Reprojektionsfehler	35
5.4.	gerendeter Kegel mit Kalibrierungsmuster in Blender aus verschiedenen Blickrichtungen	36
5.5.	Einfluss der Rotation der Kamera	36
5.6.	Einfluss der Ausgabeauflösung auf die Laufzeit der Entfaltung	37
5.7.	Deformable Templates: Ausschnitt der Energiefunktion für variable Haupt- und Nebenachse	37
5.8.	Vergleich RANSAC und LSQ bei gleichverteilten Ausreißern $\epsilon = 0.5, p = 0.99$	38
5.9.	Vergleich RANSAC und LSQ bei Schattenellipsen mit $p = 0.99$ und $\epsilon = 0.25$ (oben), $\epsilon = 0.4$ unten, links gestörte Messdaten, rechts detektierte Ellipsen RANSAC (grün), LSQ (cyan)	39
6.1.	Delaunay-Triangulation	42

Literatur

- [Can86] John Canny. „A Computational Approach to Edge Detection“. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-8.6* (1986), S. 679–698. ISSN: 01628828. DOI: 10.1109/TPAMI.1986.4767851 (siehe S. 10).
- [Ebe13] David Eberly. „Distance from a Point to an Ellipse, an Ellipsoid, or a Hyperellipsoid“. In: (2013), S. 1–13 (siehe S. 15).
- [FB81] Martin a Fischler und Robert C Bolles. „Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography“. In: *Communications of the ACM* 24.6 (1981), S. 381–395. ISSN: 0001-0782. DOI: 10.1145/358669.358692. URL: <http://dx.doi.org/10.1145/358669.358692> (siehe S. 11).
- [HS97] Janne Heikkilä und Olli Silvén. „A Four-step Camera Calibration Procedure with Implicit Image Correction.“ In: *Cvpr* (1997), S. 1106–1112. ISSN: 1063-6919. DOI: 10.1109/CVPR.1997.609468. URL: <http://dx.doi.org/10.1109/CVPR.1997.609468> (siehe S. 9, 10).
- [HZ00] R. I. Hartley und A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049, 2000 (siehe S. 41).
- [Law72] J D Lawrence. *A Catalog of Special Plane Curves*. Dover Publications, 1972, S. 62–63 (siehe S. 13, 14).
- [Lin93] Tony Lindeberg. „Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention“. In: *International Journal of Computer Vision* 11.3 (1993), S. 283–318. ISSN: 09205691. DOI: 10.1007/BF01469346 (siehe S. 10).
- [OC76] F. O’Gorman und M. B. Clowes. „Finding Picture Edges Through Collinearity of Feature Points“. In: *IEEE Transactions on Computers* C-25.4 (Apr. 1976), S. 449–456. ISSN: 0018-9340. DOI: 10.1109/TC.1976.1674627 (siehe S. 41).
- [Sto07] Burlisch Stoer. *Numerische Mathematik 1*. Springer, 2007, S. 249–253 (siehe S. 7).
- [Sto11] Burlisch Stoer. *Numerische Mathematik 2*. Springer, 2011, S. 21–23 (siehe S. 8).

Plagiatsklärung

Hiermit versichere ich, dass die vorliegende Arbeit über

Entzerrung von Kegeloberflächen aus einer Einkameraansicht basierend auf projektiver Geometrie

selbstständig verfasst worden ist, dass keine anderen Quellen und Hilfsmittel als die angegebenen benutzt worden sind und dass die Stellen der Arbeit, die anderen Werken – auch elektronischen Medien – dem Wortlaut oder Sinn nach entnommen wurden, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht worden sind.

Lars Haalck, Münster, 14. September 2016

Ich erkläre mich mit einem Abgleich der Arbeit mit anderen Texten zwecks Auffindung von Übereinstimmungen sowie mit einer zu diesem Zweck vorzunehmenden Speicherung der Arbeit in eine Datenbank einverstanden.

Lars Haalck, Münster, 14. September 2016