

Perspektiveninvariante Inferenz von Eckpunkten in Packmustern von Kartonagen

Perspective invariant inference of corner points in packing patterns of cardboard boxes

Felix Endres, Lucas Reinhart, Tobias Kaupp und Volker Willert

University of Applied Sciences Würzburg-Schweinfurt
Institute Digital Engineering (IDEE)
Ignaz-Schön-Straße 11
97421 Schweinfurt

Zusammenfassung Diese Arbeit befasst sich mit der Inferenz von Eckpunkten von Kartonagen, die in einem regelmäßigen dichten Packmuster flächig angeordnet sind. Als Sensordaten werden ausschließlich 2D Kamerabilder und keine 3D Information benutzt. Die Kartonagen werden aus extremen Perspektiven betrachtet, wie sie typischerweise beim „Blick ins Regal“ für automatisierte Kommissionieraufgaben vorkommen. Ausgehend von vier Eckpunkten einer beliebigen Kartonage wird ein auf dem Doppelverhältnis basierendes Verfahren vorgestellt, das die Eckpunkte aller möglicher benachbarter Kartonagenanordnungen berechnen kann. Des Weiteren wird die Fehlerfortpflanzung unter der Annahme von Eckpunktmessungen mit normalverteiltem Rauschen betrachtet und aus der Fehlerverteilung ein parametrisches Modell für die ortsvarianten 2D Wahrscheinlichkeitsverteilungen aller abgeleiteter Eckpunkte ermittelt.

Schlüsselwörter Mustererkennung, Robotik, perspektivische Invarianten

Abstract This work deals with the inference of corner points of cardboard boxes, which are arranged two-dimensionally in a regular dense packing pattern. Only 2D camera images and no 3D information are used as sensor data. The cardboard boxes are viewed from extreme perspectives, typically encountered when “looking at the shelf” for automated picking tasks. Starting from

four corners of an arbitrary cardboard box, a method based on the crossratio is presented that can compute the corners of all possible neighboring box arrangements. Furthermore, the error propagation assuming corner point measurements with normally distributed noise is considered and a parametric model for the 2D probability distributions that vary across image location of all derived corner points is obtained from the error distribution

Keywords Pattern recognition, robotics, perspective invariants

1 Einleitung

Das dieser Arbeit zugrundeliegende Forschungsprojekt beschäftigt sich mit der Objekterkennung für Intralogistikanwendungen. Die hier behandelte Problemstellung ergibt sich aus einem Projekt mit einem Industriepartner zur Entwicklung eines mobilen *pick-and-place* Roboters zur automatisierten Kommissionierung diverser Warentypen. Der Roboter soll im Mischbetrieb mit menschlichen Arbeitskräften zur Kommissionierung von Mischpaletten eingesetzt werden, wodurch eine Instrumentierung der Umgebung nur eingeschränkt möglich ist. Daraus ergeben sich insbesondere für die Erkennung der Waren einige Herausforderungen. Im Kommissionierbereich sind sich stark ändernde Lichtverhältnisse durch Sonneneinstrahlung und Verschattung vorherrschend. Zusätzlich können sich die visuellen Objekteigenschaften durch Verschmutzung der Waren verändern. Insbesondere ist aber durch die Palettenhöhe und den daraus resultierenden Blickwinkel auf die Palette teilweise nur eine extreme Perspektive zur Objekterkennung vorhanden (siehe auch Abbildung 1). Da dies ein häufig auftretendes Problem bei der Objekterkennung ist, wurden in der Literatur bereits verschiedene Größen untersucht, die invariant bezüglich perspektivischer Verzeichnung sind [1]. Eine der untersuchten perspektivischen Invarianten ist das Doppelverhältnis, das sich als robustes und genaues Maß erwiesen hat [2]. Des Weiteren wurde das Prinzip des Doppelverhältnisses erweitert, um Flächeninvarianten unter projektiven Abbildungen zu erhalten [3].

Um die Produkte greifen zu können, muss ein Greifpunkt ermittelt werden. Dafür würde eine Rekonstruktion des Packmusters im Bild der



Abbildung 1: links: Originalaufnahme der Kamera, rechts: Auflösungsverlust durch Transformation in eine Draufsicht.

Palette gute Kandidaten für die Greifpunkte liefern, z. B. die Mitte der segmentierten Kartonage. Um eine Rekonstruktion des Packmusters zu erreichen, möchte man Informationen nutzen, die typischerweise verfügbar sind, z. B. Größe und Geometrie der Objekte (Palette, Kartonage, usw.). Eine Möglichkeit zur Rekonstruktion besteht darin, eine Draufsicht der Szene zu erstellen. Ein Beispiel dafür ist in Abbildung 1 zu sehen, wo ein deutlicher Auflösungsverlust und große Interpolationsartefakte für die Kartonagen im hinteren Bereich zu erkennen sind. Das kann zu Fehlern bei der Rekonstruktion des Packmusters der Palette führen. Zudem muss für die Transformation in eine Draufsicht die Pose der Kamera in Bezug zur Oberfläche des Packmusters bzw. die Homographie [4] aus den Bilddaten rekonstruiert werden.

Im folgenden Abschnitt 2 wird ein Ansatz vorgestellt, der direkt auf das Bild ohne vorherige Transformation angewendet werden kann. Unter Verwendung des Doppelverhältnisses, sowie der Breite und der Länge der Kartonagen, geben wir eine Formel zur Berechnung möglicher Eckpunkte von Kartonagen im Packmuster an. Abschnitt 3 zeigt die Ergebnisse unserer Methode, angewandt auf ein Beispielbild. Außerdem werden die Auswirkungen von Messungenauigkeiten bei der Extraktion der Eckpunktkoordinaten der Referenzkartonage auf die Genauigkeit der Eckenerkennung der Kartonagen untersucht. Abschnitt 4 fasst die Ergebnisse zusammen und gibt einen Ausblick auf zukünftige Arbeiten.

2 Herleitung der Inferenz von Eckpunkten

Ausgangspunkt ist eine Ansicht von oben auf eine einzelne Kartonage als Referenz. Die Koordinaten der Eckpunkte werden mit $x_a^{00}, x_b^{00}, x_c^{00}, x_d^{00}$ bezeichnet¹, wobei die hochgestellten Zahlen ein lokales Koordinatensystem für jeden Referenzeckpunkt darstellen. Wenn das Seitenverhältnis der Kartonage bekannt ist, lassen sich die Möglichkeiten berechnen, wie weitere Kartonagen angelegt werden können. Für das Seitenverhältnis 2:1 zeigt Abbildung 2(a) die drei Möglichkeiten, wie eine zweite Kartonage auf der rechten Seite angeordnet werden könnte und Abbildung 2(b) für die obere Seite.

Die Punkte x_i^{10}, x_i^{20} und x_i^{01}, x_i^{02} sind die möglichen Eckpunkte angrenzender Kartonagen in x und y Richtung vom Eckpunkt x_i . Sie werden im Folgenden *inferierte Eckpunkte erster Ordnung* genannt, da sie direkt mit dem Doppelverhältnis berechnet werden können. Die Punkte x_i^{11}, x_i^{12} und x_i^{21}, x_i^{22} werden *inferierte Eckpunkte zweiter Ordnung* genannt, da sie sich wiederum von den Punkten erster Ordnung ableiten lassen.

Daraus ergibt sich die in Abbildung 3 gezeigte Konfiguration von Punkten. Aus dieser Sicht lassen sich die nächsten Eckpunkte direkt berechnen, wenn man die Länge L , die Breite B und die Koordinaten der Eckpunkte x_i^{00} der Referenzkartonage kennt. Bei einer perspektivischen Ansicht sind jedoch die Distanzen zwischen den Eckpunkten von Kartonage zu Kartonage unterschiedlich und verändern sich zusätzlich bei der Veränderung der Perspektive. In diesem Fall kann zur Berechnung der Eckpunkte das Doppelverhältnis genutzt werden.

Doppelverhältnisse in Packmustern

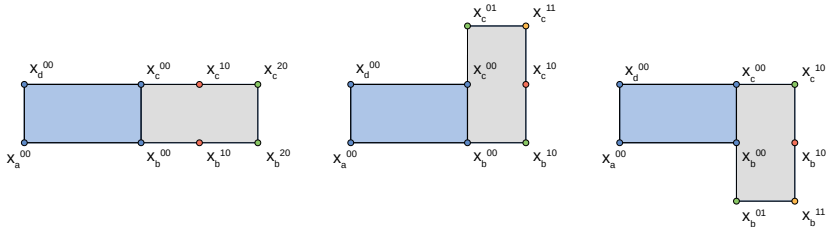
Das Doppelverhältnis $[A, B; C, D]$ von vier auf einer Geraden liegenden Punkten A, B, C, D ist definiert durch

$$[A, B; C, D] := \frac{\overline{AC} \cdot \overline{BD}}{\overline{BC} \cdot \overline{AD}}. \quad (1)$$

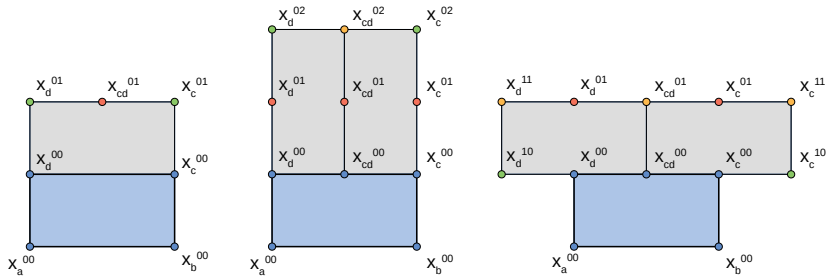
Das Doppelverhältnis ist eine projektive Invariante, d.h. es bleibt durch projektive Abbildungen unverändert [5]. Um die nächsten Eckpunkte

¹ Aus Übersichtsgründen werden an manchen Stellen bei den Punkten $x_a^{00}, x_b^{00}, x_c^{00}, x_d^{00}$ die hochgestellten Zahlen weggelassen.

Perspektiveninvariante Inferenz von Eckpunkten



(a) Anordnungen von zwei Kartonagen mit Seitenverhältnis 2:1, die nebeneinander liegen.



(b) Anordnungen von Kartonagen mit Seitenverhältnis 2:1, die übereinander liegen.

Abbildung 2: Grüne Punkte sind die korrekten inferierten Eckpunkte erster Ordnung. Rote Punkte sind Eckpunkte erster Ordnung aus anderen möglichen Konfigurationen. Gelbe Punkte sind inferierte Punkte zweiter Ordnung.

unter einer perspektivischen Ansicht abzuleiten, werden zuerst aus der bekannten Konfiguration (siehe Abbildung 4) mit beliebigen L, B und x_a, x_b, x_c, x_d die vier möglichen Doppelverhältnisse r_1, r_2, r_3, r_4 berechnet:

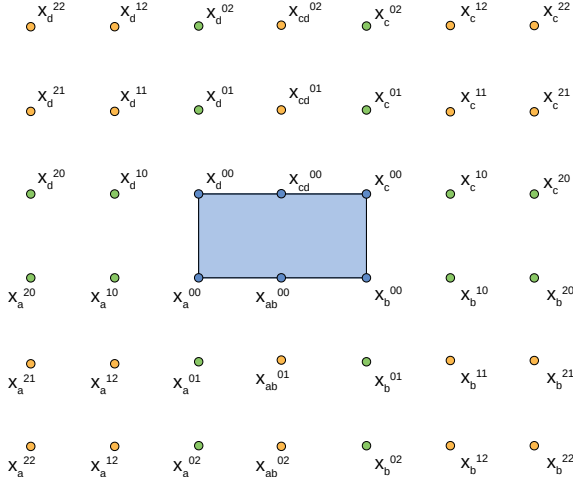


Abbildung 3: Eckpunkte für alle Konfigurationen von Kartonagen mit Seitenverhältnis 2:1. Blaue Punkte sind Teil der Referenzkartonage. Grüne Punkte sind inferierte Punkte erster Ordnung. Gelbe Punkte sind inferierte Punkte zweiter Ordnung.

$$\begin{aligned} r_1 &= [x_a^{00}, x_{ab}^{00}, x_b^{00}, x_b^{10}] = [x_b^{00}, x_{ab}^{00}, x_a^{00}, x_a^{10}] = [x_d^{00}, x_{cd}^{00}, x_c^{00}, x_c^{10}] \\ &= [x_c^{00}, x_{cd}^{00}, x_d^{00}, x_d^{10}] = \frac{L \cdot (L/2 + B)}{L/2 \cdot (L + B)} = \frac{L + 2B}{L + B}, \end{aligned} \quad (2)$$

$$\begin{aligned} r_2 &= [x_a^{00}, x_{ab}^{00}, x_b^{00}, x_b^{20}] = [x_b^{00}, x_{ab}^{00}, x_a^{00}, x_a^{20}] = [x_d^{00}, x_{cd}^{00}, x_c^{00}, x_c^{20}] \\ &= [x_c^{00}, x_{cd}^{00}, x_d^{00}, x_d^{20}] = \frac{L \cdot (L/2 + L)}{L/2 \cdot (L + L)} = \frac{3}{2}, \end{aligned} \quad (3)$$

$$\begin{aligned} r_3 &= [x_a^{00}, x_{da}^{00}, x_d^{00}, x_d^{01}] = [x_d^{00}, x_{da}^{00}, x_a^{00}, x_a^{01}] = [x_b^{00}, x_{bc}^{00}, x_c^{00}, x_c^{01}] \\ &= [x_c^{00}, x_{bc}^{00}, x_b^{00}, x_b^{01}] = \frac{B \cdot (B/2 + B)}{B/2 \cdot (B + B)} = \frac{3}{2}, \end{aligned} \quad (4)$$

$$\begin{aligned} r_4 &= [x_a^{00}, x_{da}^{00}, x_d^{00}, x_d^{02}] = [x_d^{00}, x_{da}^{00}, x_a^{00}, x_a^{02}] = [x_b^{00}, x_{bc}^{00}, x_c^{00}, x_c^{02}] \\ &= [x_c^{00}, x_{bc}^{00}, x_b^{00}, x_b^{02}] = \frac{B \cdot (B/2 + L)}{B/2 \cdot (B + L)} = \frac{B + 2L}{B + L}. \end{aligned} \quad (5)$$

Interessanterweise ergeben sich nur drei unterschiedliche Werte, wobei r_1 und r_4 von L und B abhängen und $r_2 = r_3 = 3/2$ identisch und unabhängig vom Seitenverhältnis der Kartonage sind.

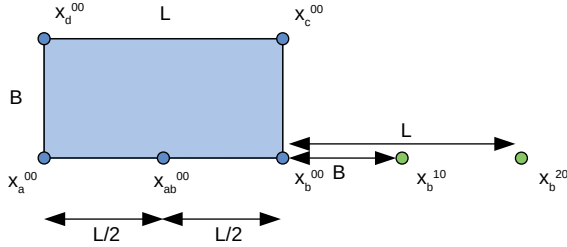


Abbildung 4: Mit bekannter Länge L und Breite B der Referenzkartonage können vier mögliche Doppelverhältnisse berechnet werden (siehe Text).

Eckpunktberechnungen

Als nächstes wird die Konfiguration aus einer beliebigen Perspektive betrachtet, wie in Abbildung 5 dargestellt. Unter Verwendung homogener Koordinaten können Hilfspunkte x_e, x_f und x_g aus den Gleichungen 6 - 8 berechnet werden. Hierbei steht \times für das dreidimensionale Kreuzprodukt und \mathbf{x}_i steht für den Koordinatenvektor des Punktes x_i . Jede Gerade durch die Punkte x_i und x_j wird durch einen Vektor \mathbf{l}_{ij} parametrisiert:

$$\mathbf{l}_{ac} = \mathbf{x}_a \times \mathbf{x}_c, \quad \mathbf{l}_{bd} = \mathbf{x}_b \times \mathbf{x}_d, \quad \mathbf{x}_e = \mathbf{l}_{ac} \times \mathbf{l}_{bd}, \quad (6)$$

$$\mathbf{l}_{ab} = \mathbf{x}_a \times \mathbf{x}_b, \quad \mathbf{l}_{cd} = \mathbf{x}_c \times \mathbf{x}_d, \quad \mathbf{x}_g = \mathbf{l}_{ab} \times \mathbf{l}_{cd}, \quad (7)$$

$$\mathbf{l}_{ad} = \mathbf{x}_a \times \mathbf{x}_d, \quad \mathbf{l}_{bc} = \mathbf{x}_b \times \mathbf{x}_c, \quad \mathbf{x}_f = \mathbf{l}_{ad} \times \mathbf{l}_{bc}. \quad (8)$$

Damit benachbarte Punkte über das Doppelverhältnis nur aus den vier Referenzpunkten inferiert werden können, werden die vier Seitenhalbierende $x_{ab}, x_{cd}, x_{bc}, x_{da}$ als Hilfspunkte eingeführt. Die Seitenhalbierenden x_{ij} , $i \neq j \in \{a, b, c, d\}$ werden konstruiert, indem die Gerade auf dem einer der Fluchtpunkte x_f, x_g und der Mittelpunkt der Referenz-

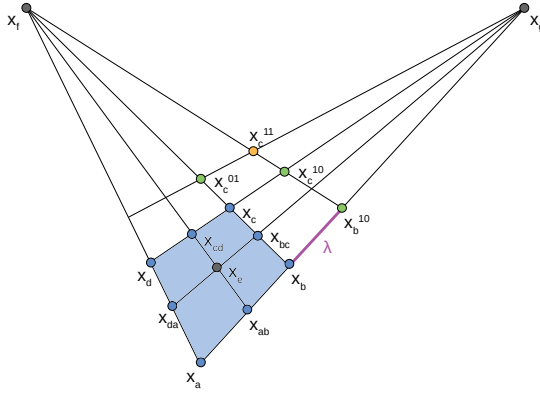


Abbildung 5: Konfiguration bei beliebiger Perspektive.

renzkartonage x_e liegt, mit einer Seite des Vierecks geschnitten wird:

$$x_{ab} = l_{fe} \times l_{ab}, \quad x_{bc} = l_{ge} \times l_{bc}, \quad (9)$$

$$x_{cd} = l_{fe} \times l_{cd}, \quad x_{da} = l_{ge} \times l_{da}. \quad (10)$$

Mit den Seitenhalbierenden und dem Doppelverhältnis können wir dann den Abstand λ zu einem Nachbarpunkt x_b^{10} berechnen:

$$r_1 = \frac{\|x_a - x_b\| \|x_{ab} - x_b^{10}\|}{\|x_{ab} - x_b\| \|x_a - x_b^{10}\|} = \frac{\|x_a - x_b\| (\|x_{ab} - x_b\| + \lambda)}{\|x_{ab} - x_b\| (\|x_a - x_b\| + \lambda)},$$

$$\Rightarrow \lambda = \frac{(1 - r_1) (\|x_{ab} - x_b\| \|x_a - x_b\|)}{r_1 \|x_{ab} - x_b\| - \|x_a - x_b\|}. \quad (11)$$

Schließlich ergeben sich die Koordinaten des Punktes x_b^{10} wie folgt:

$$x_b^{10} = x_b + \lambda \frac{x_b - x_a}{\|x_b - x_a\|}. \quad (12)$$

Auf die gleiche Weise ist es möglich alle anderen Eckpunkte erster Ordnung zu berechnen. Den Eckpunkt zweiter Ordnung x_c^{11} erhält man durch

$$l_{c^{10}f} = x_c^{10} \times x_f, \quad l_{c^{10}f} = x_c^{01} \times x_g \quad (13)$$

$$\Rightarrow x_c^{11} = l_{c^{10}f} \times l_{c^{01}g}. \quad (14)$$

Durch das Schneiden von Geraden, die aus den inferierten Eckpunkten erster Ordnung mit den Fluchtpunkten gebildet werden, können so auch alle anderen Eckpunkte zweiter Ordnung ermittelt werden.

3 Statistische Analyse und Parametrisches Modell

In der Abbildung 6 wurde die im Abschnitt 2 vorgestellte Methode auf ein Beispielbild angewandt. Dabei wurden die Eckpunkte x_a, x_b, x_c, x_d einer Kartonage im Bild als gemessene Referenzpunkte angenommen. Dann wurden auf Basis einer bivariaten Normalverteilung $\mathcal{N}(\mu, \Sigma)$ mit Mittelwertvektor $\mu = x_i$ und Kovarianzmatrix Σ , die der Einheitsmatrix entspricht, für jeden Eckpunkt 2000 verrauschte Eckpunkte berechnet. Anschließend wurden für alle 2000 Konfigurationen die Eckpunkte erster und zweiter Ordnung berechnet.

Daraufhin wurden mit Hilfe des Expectation–Maximization Algorithmus [6] die resultierenden Verteilungen der inferierten Eckpunkte durch eine bivariate Normalverteilung approximiert (siehe Abbil-

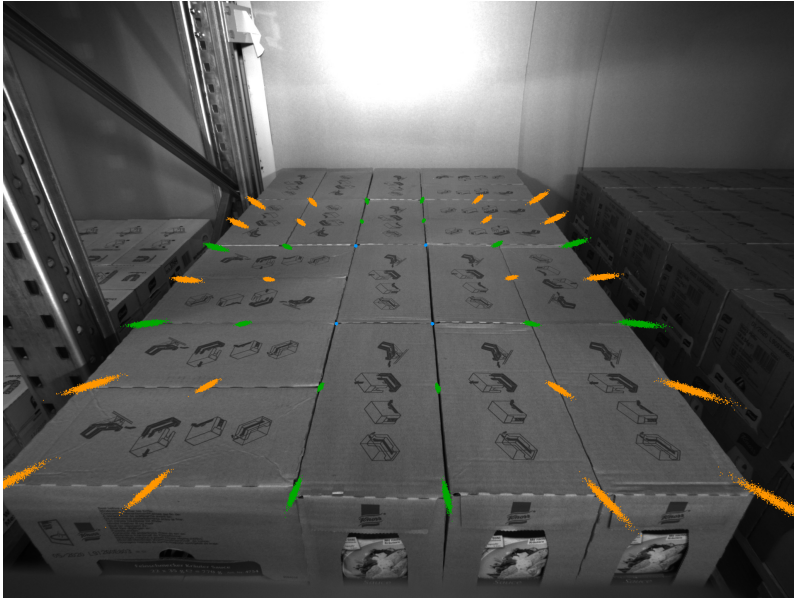


Abbildung 6: Mögliche Konfiguration von Kartonagen. Blau gefärbt sind die 2000 verauschten Referenzeckpunkte. Grün sind die daraus resultierenden inferierten Eckpunkte erster Ordnung und orange die inferierten Eckpunkte zweiter Ordnung.

dung 8). Es ist deutlich zu erkennen, dass die Orientierungen der Normalverteilungen sehr gut mit der Orientierung der Verbindungslinien zwischen dem Mittelpunkt der Referenzkartonage und dem jeweiligen inferierten Punkt übereinstimmen. Abbildung 7 zeigt die Standardabweichungen σ_1, σ_2 der ersten und zweiten Hauptkomponente aller bivariaten Normalverteilungen in Abhängigkeit von der Distanz des ursprünglichen inferierten Punktes zum Mittelpunkt der Kartonage x_c . Für die erste Hauptkomponente kann die Abhängigkeit durch ein Polynom zweiter Ordnung $p(x)$ approximiert werden, für die zweite Komponente ist die Korrelation durch eine Gerade $l(x)$ beschrieben. Hieraus lässt sich ein Modell für den Einfluss des normalverteilten Rauschens auf die inferierten Punkte ableiten. Sei x_i^{nm} ein von der ursprünglichen Konfiguration inferierter Punkt und \mathbf{u} der Vektor

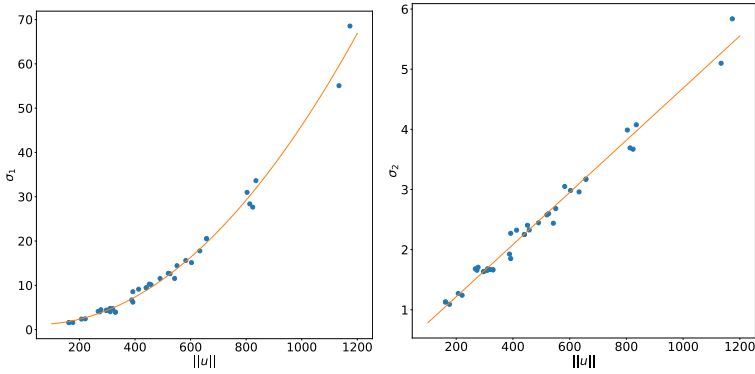


Abbildung 7: Abhängigkeit der Standardabweichungen σ_1, σ_2 der beiden Hauptkomponenten von der Distanz $\|u\|$ in Pixel.

von x_e zu x_i^{nm} , d.h. $\mathbf{u} = \mathbf{x}_i^{nm} - \mathbf{x}_e$. Als nächstes definiert man den Vektor $\mathbf{v} := (-u_2, u_1)$, sodass \mathbf{u} und \mathbf{v} senkrecht aufeinander stehen. Dann ergibt sich ein parametrisches Modell für die Verteilung von x_i^{nm} in Abhängigkeit von x_e , das einer bivariaten Normalverteilung $\mathcal{N}(\boldsymbol{\mu}_i^{nm}, \boldsymbol{\Sigma}_i^{nm})$ mit $\boldsymbol{\mu}_i^{nm} = \mathbf{x}_i^{nm}$ und

$$\boldsymbol{\Sigma}_i^{nm} = \begin{pmatrix} \frac{\mathbf{u}}{\|\mathbf{u}\|}, \frac{\mathbf{v}}{\|\mathbf{v}\|} \end{pmatrix} \begin{pmatrix} p(\|\mathbf{u}\|)^2 & 0 \\ 0 & l(\|\mathbf{u}\|)^2 \end{pmatrix} \begin{pmatrix} \frac{\mathbf{u}}{\|\mathbf{u}\|}, \frac{\mathbf{v}}{\|\mathbf{v}\|} \end{pmatrix}^T \quad (15)$$

entspricht. In Abbildung 8 sind beispielhaft für vier inferierte Punkte die Verteilungen mit den 2σ Konturen der Normalverteilung, welche durch den EM-Algorithmus berechnet wurde und der Normalverteilung, die mit dem Modell ermittelt wurde, dargestellt. Sowohl das Modell der Orientierung, als auch das lineare Modell der Standardabweichung entlang der zweiten Hauptkomponente passen sehr gut mit der normalverteilten Statistik der Datenpunkte überein. Bei der Standardabweichung der ersten Hauptkomponente ergeben sich teilweise vertretbare Abweichungen. In zukünftigen Arbeiten sollte das hier vorgestellte datengetriebene Modell durch eine stringente Berechnung der Fehlerfortpflanzung mittels der Formel (12) unter Annahme von normalverteilten Koordinaten der Referenzeckpunkte verifiziert werden.

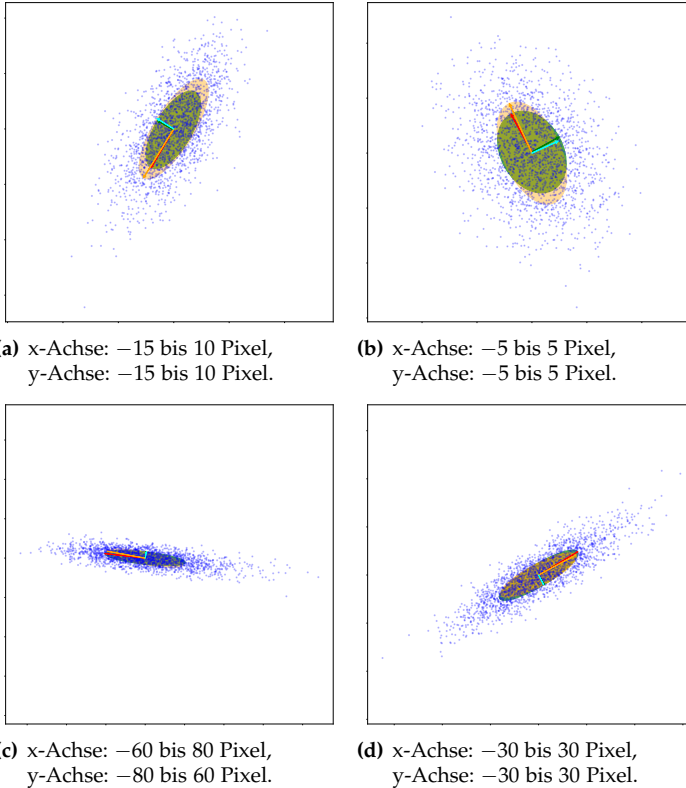


Abbildung 8: Verteilungen für vier inferierte Eckpunkte. Die grüne Ellipse ist die 2σ Kontur der gefitteten Normalverteilung, die orangene Ellipse ist die 2σ Kontur der aus dem Modell berechneten Normalverteilung. Der rote und grüne Vektor zeigt die Vorzugsrichtungen der EM-Normalverteilung, orange und cyan, die der Modell-Normalverteilung.

4 Zusammenfassung

Es konnte gezeigt werden, dass im Prinzip alle möglichen Eckpunkte rechteckiger Elemente, die zu einem regelmäßigen, dicht besetzten, flächigen Muster zusammengesetzt werden, ausgehend von vier gemessenen Eckpunkten eines beliebigen Elements aus dieser Anord-

nung, über das Doppelverhältnis und das Wissen über die Abmaße des Rechtecks berechnet werden können. Insbesondere bei extremer Perspektive können sich Vorteile ergeben, da trotz starker Abnahme der Bildauflösung der einzelnen projizierten Rechtecke mit zunehmender Distanz zur Kamera, die Lage der Eckpunkte genau berechenbar ist. Ungenauigkeiten bei der Messung der Eckpunktkoordinaten des Referenzrechtecks pflanzen sich weniger stark fort, je größer der Abstand und kleiner die Auflösung in der Projektion sind.

Damit kann dieses Verfahren nicht nur zur Lageerkennung von Kartonagen, sondern auch auf andere Muster, die aus rechteckigen Elementen bestehen adaptiert werden, wie beispielsweise Parkettböden, Mauerwerk oder schachbrettartige Kalibriermuster. Die Fehlerverteilung lässt sich ziemlich genau mit einer ortsvarianten zweidimensionalen Normalverteilung approximieren, wobei die ortsabhängigen Werte des Erwartungswertvektors und der Kovarianzmatrix in Abhängigkeit vom Mittelpunkt des Referenzrechtecks und der Koordinaten des zu inferierenden Eckpunktes berechnet werden können. Damit ergibt sich ein vollständiges probabilistisches Modell, das beispielsweise als Potentialfunktion in einem Markovschen Zufallsfeld zur Bestimmung des vollständigen Packmusters von Kartonagen benutzt werden kann.

Literatur

1. I. Weiss, "Geometric invariants and object recognition," *International Journal of Computer Vision*, vol. 10, no. 3, pp. 207–231, Jun. 1993.
2. D. Forsyth, J. Mundy, A. Zisserman, C. Coelho, A. Heller, and C. Rothwell, "Invariant descriptors for 3d object recognition and pose," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 10, pp. 971–991, 1991.
3. L. Nielsen and G. Sparr, "Projective area-invariants as an extension of the cross-ratio," *CVGIP: Image Understanding*, vol. 54, no. 1, pp. 145–159, 1991.
4. Y. Ma, S. Soatto, J. Kosecka, and S. S. Sastry, *An Invitation to 3-D Vision: From Images to Geometric Models*. Springer-Verlag, 2003.
5. J. N. Cederberg, *A Course in Modern Geometries*, 2nd ed., ser. Undergraduate Texts in Mathematics. Springer-Verlag New York, 2001.
6. C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2006.