

Sensorik und Bildverarbeitung für Landmarken-basierte Navigation

Dissertation

zur Erlangung des Grades eines
Doktors der Naturwissenschaften
der Fakultät für Mathematik und Physik
der Eberhard-Karls-Universität Tübingen

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



vorgelegt von
Wolfgang Stürzl
aus München

2004

Tag der mündlichen Prüfung: 20. April 2004

Dekan: Prof. Dr. H. Müther

1. Berichterstatter: Prof. Dr. W. Kley / Prof. Dr. H. A. Mallot

2. Berichterstatter: Prof. Dr. K. Werner

Die vorliegende Dissertation ist das Ergebnis von Forschungsarbeiten, die an der Abteilung „Cognitive and Computational Psychophysics“ des Max-Planck-Instituts für biologische Kybernetik und am Lehrstuhl Kognitive Neurowissenschaft der Universität Tübingen durchgeführt wurden. Betreut wurde sie von Prof. Dr. Hanspeter A. Mallot, Kognitive Neurowissenschaft, Zoologisches Institut der biologischen Fakultät, und Prof. Dr. Wilhelm Kley, Computational Physics, Institut für Astronomie und Astrophysik der Fakultät für Mathematik und Physik an der Universität Tübingen.

An dieser Stelle möchte ich mich bei allen bedanken, die zum Entstehen dieser Arbeit beigetragen haben:

bei Prof. Dr. Hanspeter A. Mallot und Prof. Dr. Wilhelm Kley für die inspirierende wissenschaftliche und menschliche Betreuung der Arbeit, die Bereitstellung von Mitteln u.a. für Roboter, Rechner und Kameras und für den Freiraum bei der Entwicklung und Durchführung eigener Ideen,

bei Dr. Hansjürgen Dahmen für viele anregende Diskussionen über Optik und andere interessante Themen,

bei Wolfgang Hübner für das Lösen hartnäckiger Softwareprobleme und für zahlreiche Diskussionen an unserer Kreidetafel und in Tübinger Kneipen,

bei Amelie Schmolke, Lorenz Gerstmayr und Ulrich Hoffmann für die freundschaftliche Zusammenarbeit im Robotik-Labor,

bei Herrn Stahl, Herrn Weber und Herrn Schneider von den Mechanikwerkstätten des Max-Planck-Instituts und der Tierphysiologie der Universität für das Anfertigen zuverlässiger Stereosysteme,

und bei allen KollegInnen in der Kognitiven Neurowissenschaft und der Psychophysikabteilung des Max-Planck-Instituts für die freundliche und anregende Team-Atmosphäre.

Tübingen, 12. November 2003

Wolfgang Stürzl

Inhaltsverzeichnis

Einleitung und Überblick	1
Themenbereiche und Ergebnisse	4
Zur Verwendung mathematischer Symbole	5
Überblick zum Heimfinden mit Ortssignaturen	6
1 Visuelles Heimfinden	7
1.1 Panoramischer Stereosensor für Khepera	8
1.1.1 Berechnung der Spiegelgeometrie	8
1.1.2 Justierung des Stereosensors	12
1.1.3 Abbildungsfehler eines Kegelspiegels	12
1.2 Bestimmung vertikaler Disparitäten	14
1.2.1 Varianzwert für Disparitäten	15
1.2.2 Disparitätssignatur eines Ortes	19
1.2.3 Berechnung von Objektdistanzen	19
1.3 Disparitäts-basiertes Heimfinden	20
1.3.1 Algorithmus	20
1.3.2 Berechnung erwarteter Disparitätssignaturen	22
1.3.3 Ähnlichkeitsmaß für Disparitätssignaturen	25
1.4 Vergleich von Disparitäts- und Bildsignaturen	27
1.4.1 Bild-basiertes Heimfinden	27
1.4.2 Kombination von Disparitäts- und Bildsignatur	28
1.4.3 Datenbank mit panoramischen Stereobildern	29
1.4.4 „Average Homeward Component“	31
1.4.5 Fangbereiche	33
1.5 Genauigkeit des visuellen Heimfindens	38
1.5.1 Bild-basiertes Heimfinden	38
1.5.2 Fehler durch „Equal Distance Assumption“	43
1.5.3 Disparitäts-basiertes Heimfinden	48
1.6 Erlernen der Bildtransformation	52
1.6.1 Trainieren des neuronalen Netzes	52
1.6.2 Verwendung der erlernten Bildtransformation	53

2	Fourier-basierter Bildvergleich	59
2.1	Bildvergleich mit Korrelation und SSD	60
2.1.1	Korrelation panoramischer Bilder	60
2.1.2	Fourierdarstellung panoramischer Bilder	62
2.1.3	SSD und Korrelation im Fourier-Raum	65
2.1.4	Bestimmung der optimalen Bildverschiebung	65
2.1.5	Algorithmus zur Orientierungsbestimmung	67
2.2	Vergleich mit „Standard“-Korrelation	68
2.3	Sinus-Cosinus-Form panoramischer Bilder	70
2.4	Fehler bei der Orientierungsbestimmung	71
2.4.1	Positionsabweichungen	71
2.4.2	Statistische Fehler	73
2.5	Fourier-basiertes Heimfinden	74
2.5.1	Berechnung erwarteter Bilder	75
2.5.2	Bestimmung des Heimvektors	76
2.5.3	Resultate	78
3	Panoramische Bildsensoren	83
3.1	Mathematische Beschreibung	83
3.2	Fehlstellungen der Kamera	85
3.2.1	Verschiebung des Kameraknotenpunkts	85
3.2.2	Neigung der optischen Achse	88
3.3	Konstanter Vergrößerungsfaktor	89
3.4	Panoramischer Stereosensor für Koala	91
3.4.1	Kaustiken des Stereosensors	93
3.4.2	Stereobilder bei Kameraverschiebung	96
3.4.3	Vergrößerungsfaktor	97
3.4.4	Justierung und Bildtransformation	98
3.5	Bestimmung der optimalen Fokusebene	100
3.5.1	Berechnung virtueller Bildflächen	101
3.5.2	Größe des Unschärfebereichs im Kamerabild	105
4	Binokulares Stereo	109
4.1	Binokulares Energiemodell	109
4.2	Implementierung des Energiemodells	112
4.2.1	Monokulare Normalisierung	114
4.2.2	Kombination der Frequenzkanäle	114
4.2.3	Steuerung der Vergenzstellung	116
4.2.4	Bestimmung von Disparitäten	118
4.2.5	Beispiel für die Funktionsweise des Systems	119

4.3	Wahrscheinlichkeitsdichten der neuronalen Aktivität	122
4.3.1	Nicht-normalisierte Zellen	122
4.3.2	Normalisierte Zellen	131
4.3.3	Wahrscheinlichkeit für falsche Detektion	140
Erweiterungen und Ausblick		143
Anhang		149
A.1	Erwartungswert und Varianz von Disparitäten	149
A.2	Varianz erwarteter Disparitätswerte	151
A.3	Lokale Bildveränderung	153
A.4	Begründung für die Näherung (1.96)	154
A.5	PCA auf rotationssymmetrischem Datensatz	156
A.6	Relative Orientierung von 2D-Bilder	158
A.7	Berechnung der Koeffizienten	162
A.8	Bestimmung des Parameters λ^*	164
A.9	Reflektierender Spiegelbereich	167
A.10	Hauptkrümmungen	169
A.11	Beweis der Gleichungen (4.121)–(4.123)	171
Literaturverzeichnis		175

Einleitung und Überblick

Die Fähigkeit, bekannte Orte – beispielsweise ein Nest oder eine Nahrungsquelle – wiederzufinden, ist für viele Tiere überlebenswichtig. Auch von autonomen mobilen Robotern wird erwartet, dass sie bestimmte Positionen zuverlässig anfahren können. Ein Beispiel ist das rechtzeitige selbstständige Aufsuchen einer Akkuladestation. In einfacher Form zeigten Grey Walters „Turtles“ (durch direkte Reiz-Reaktions-Koppelung gesteuerte elektromechanische Roboter) dieses Verhalten bereits Anfang der 50er Jahre [48]. Versuche, das Verhalten von Tieren bei der Navigation, ihre Sensorik und deren Verarbeitung zu verstehen, können zur Entwicklung effizienter Sensoren und Algorithmen beitragen, die mobilen Robotern eine zuverlässige Navigation in komplexen Umgebungen ermöglichen. Als Beispiel sei die Stabilisierung eines autonomen Modellflugzeugs mit Hilfe eines den Nebenaugen (Ocelli) von Libellen nachempfundenen Horizontsensors genannt [16].

Umgekehrt können Modelle zu Navigationsmechanismen von Tieren, die aus Verhaltensexperimenten abgeleitet wurden, durch Implementierung auf einem mobilen Roboter getestet werden. Der Unterschied zu einer Computersimulation besteht vor allem darin, dass Roboter sich in natürlichen Umgebungen bewegen können, in denen sie trotz unterschiedlicher „Hardware“ ähnliche Probleme wie Tiere zu bewältigen haben: Rauschen kann Sensorwerte stören, Beleuchtungsveränderung ein Wiedererkennen erschweren, das Handeln anderer Agenten die Umgebung maßgeblich beeinflussen etc. Eine bei einem simulierten Agenten erforderliche, möglicherweise komplexe Modellierung der Umwelt kann somit entfallen. Zudem sind die Auswirkungen von expliziten oder impliziten Annahmen und Vereinfachungen, die einer Simulation zwangsläufig zugrunde liegen, oft schwer abzuschätzen. Unter Umständen deckt erst der Versuch, einen Roboter mit einer bestimmten Fähigkeit zu versehen, die Schwierigkeiten auf, die Tiere in ihrem Alltag zu bewältigen haben. Simulationen hingegen bieten die Möglichkeit, auch die „Umwelt“ vollständig zu kontrollieren und somit verschiedene Algorithmen unter identischen Bedingungen zu vergleichen. Gleichzeitig können sehr aufwendige und detailreiche Modelle untersucht werden, da im Normalfall weniger Rücksicht auf die benötigte Rechenzeit genommen werden muss bzw. mehr Rechenleistung zur Verfügung steht.

Bei der Rückkehr zu einem Ort können zwei prinzipielle Strategien unterschieden werden: Zum einen können aus der aktuell vorliegenden Sensorinformation Merkmale extrahiert und mit einer gespeicherten Repräsentation des Zielorts verglichen werden. Diese von der Orientierung und der Position in der Umgebung, aber auch von den zur Verfügung stehenden Sensoren und der anschließenden Verarbeitung abhängigen Merkmale („external cues“ [32]) werden im Folgenden unter dem Begriff „Landmarke“ zusammengefasst. Abweichend vom alltäglichen Sprachgebrauch wird deshalb in dieser Arbeit unter einer „Landmarke“ nicht vorrangig ein einzelnes Objekt wie z.B. ein weithin sichtbarer Kirchturm verstanden, sondern meist die lokal verfügbare Sensorinformation, die nach Verarbeitung als Repräsentation bzw. „Signatur“ des Ortes gespeichert wird.

Zum anderen besteht die Möglichkeit, durch ständiges Aufsummieren der Eigenbewegung die aktuelle Position relativ zum Ziel auch ohne externe Bezugspunkte zu bestimmen. Dies ist vor allem in merkmalsarmen Umgebungen hilfreich, in denen die Unterscheidung verschiedener Positionen beschränkt ist, sowie bei starken, nicht vorhersagbaren Veränderungen, die ein Wiedererkennen bekannter Orte erschweren. Dieses Verhalten, das mathematisch als Integration infinitesimaler Vektoren beschrieben werden kann (siehe Abschnitt 11.1 in [59]), wird als Wegintegration bezeichnet und ist für Insekten [19, 68], Säugetiere [31, 66] und auch beim Menschen nachgewiesen [39]. Speziell in der Robotik wird hierfür der Begriff „Odometrie“ verwendet, siehe z.B. [10]. Die Messung der Einzelbewegungen kann durch Auswertung des durch die Eigenbewegung induzierten optischen Flusses erfolgen. Diese Möglichkeit wurde vor allem für die Honigbiene eingehend untersucht [87, 89]. Aber auch ohne externe Reize ist durch Messung von Beschleunigungen (z.B. durch das vestibuläre System) und Wahrnehmung von Gelenkstellungen („Propriozeption“) eine Bestimmung der momentanen Bewegung möglich [59]. Letzteres entspricht bei mobilen Robotern beispielsweise einer einfachen Messung der Radumdrehungen. Desweiteren kann das Bewegungskommando selbst als so genannte „Efferenzkopie“ verwendet werden. Ein großer Nachteil der Wegintegration besteht darin, dass sich kleine Fehler in den Einzelmessungen aufaddieren und folglich nach einer gewissen Zeit eine zuverlässige Positionsbestimmung nicht mehr möglich ist. Deshalb muss eine regelmäßige „Rekalibrierung“ durch einen Bezug zur externen Welt erfolgen.

Infolge ihrer komplementären Eigenschaften ermöglicht eine Kombination von Wegintegration und Landmarkeninformation eine deutliche Steigerung der Navigationsleistung und wird bei praktisch allen Robotersystemen eingesetzt, siehe z.B. [98]. Einfache Modelle zur Integration beider Prinzipien wurden aufgrund von Navigationsexperimenten mit Insekten [21, 89, 106] und Säugetieren [32, 63] entwickelt. So gibt es Hinweise, dass Bienen mit der an

einem Ort verfügbaren Sensorinformation einen Vektor (oder auch nur eine Richtung) zum Nest oder zum nächsten Ort einer Route assoziieren [20]. Ungenauigkeiten der Wegintegration können dann nach „Ablaufen des Vektors“ mit Hilfe visueller Merkmale korrigiert werden („visuelle Zielanfahrt“).

Werden mit verschiedenen Orten mehrere Vektoren – beispielsweise zu den benachbarten Orten – verknüpft, so ist eine flexible (vom gewählten Ziel abhängige) Navigation möglich. Das Ortsgedächtnis kann dann in Form eines Graphen dargestellt werden, wobei die Knoten Orte symbolisieren, an denen Landmarkeninformation gespeichert wurde, und die Kanten mit dem Differenzvektor zwischen den Nachbarknoten verknüpft sind [50]. Derartige „metrisch-topologische Karten“ sind für mobile Roboter von Interesse, da sie eine an die Umgebung und an die Bedürfnisse des Agenten angepasste Repräsentation der Umwelt ermöglichen. Im Vergleich zu Besetzungsgitterkarten besitzen sie üblicherweise einen deutlich geringen Speicherbedarf und erlauben eine deutlich einfachere Wegplanung [97]. Allerdings ist nicht zu jedem Zeitpunkt eine direkte Landmarken-basierte Lokalisation möglich, da (verarbeitete) Sensorinformationen nur an den Knoten gespeichert sind.

Die vorliegende Arbeit beschäftigt sich mit der Verwendung von Landmarken für Navigationsaufgaben. Hierbei wird speziell das Wiederfinden bekannter Orte anhand visueller Information untersucht, das als „visuelles Heimfinden“ (engl. „visual homing“) oder „visuelle Zielanfahrt“ bezeichnet wird.¹

Welche Information aus Bildern extrahiert und für Raumorientierung benutzt wird, ist nur bei wenigen Tierarten verstanden. Navigationsexperimente mit Bienen [12] und (Wüsten-)Ameisen [23] deuten jedoch darauf hin, dass diese ein nur geringfügig verarbeitetes Bild der Umgebung verwenden, um zu wichtigen Orten (z.B. Bienenstock, Futterquellen) zurück zu finden. Nagetiere hingegen vertrauen hauptsächlich auf geometrische Merkmale, siehe z.B. [18, 22]. So suchten Ratten in einer rechteckigen Box an den beiden geometrisch äquivalenten Orten nach Futter – unter Missachtung eindeutiger visueller und olfaktorischer Reize [18]. Aber auch Kleinkinder bevorzugen im Gegensatz zu Erwachsenen geometrische Merkmale, um einen Ort wiederzufinden [46]. Geometrische Landmarkeninformation hat gegenüber reiner Bildinformation den Vorteil größerer Invarianz gegenüber Beleuchtungsvariationen und Veränderungen infolge der verschiedenen Jahreszeiten. Die Analyse und der Vergleich beider Ortsrepräsentationen mit Hilfe mobiler Roboter ist ein zentraler Bestandteil dieser Arbeit. Hierfür wurden geeignete Sensoren und Algorithmen für das visuelle Heimfinden entwickelt.

¹Einen Überblick über weitere Navigationsmechanismen und deren Klassifizierung geben Franz und Mallot [36] und Trullier et al. [100].

Themenbereiche und Ergebnisse der Arbeit

Im Einzelnen untergliedert sich die Arbeit in folgende Themenbereiche:

In Kapitel 1 („Visuelles Heimfinden mit panoramischen Bildern“) wird zunächst ein Sensor für einen kleinen Tischroboter („Khepera“) vorgestellt, der mit Hilfe kegelförmiger Spiegelflächen die Aufnahme panoramischer Stereobilder ermöglicht. Durch Auswertung der Stereobilder können Disparitätssignaturen extrahiert werden, die – da Disparitäten² leicht in Distanzen zu umgebenden Objekten umgerechnet werden können – unmittelbar mit der Geometrie der Aufnahmeorte verknüpft sind. Anschließend wird ein Algorithmus zum Heimfinden anhand von Disparitätssignaturen vorgestellt. Im Vergleich mit dem von Franz et al. [38] entwickelten Bild-basierten Heimfinden besitzt das Verfahren eine wesentlich größere Invarianz gegenüber Beleuchtungsveränderungen, jedoch kleinere „Fangbereiche“³ Um die Vorteile beider Repräsentationen zu erhalten, wird außerdem ein Verfahren beschrieben und getestet, das Bild- und Disparitätssignaturen kombiniert. Die anschließenden Untersuchungen zur Genauigkeit von Bild- und Disparitäts-basiertem Heimfinden unter Berücksichtigung von Bildstörungen lassen u.a. Aussagen darüber zu, an welchen Positionen Ortssignaturen gespeichert werden sollten, um ein Wiederfinden mit großer Präzision zu erreichen. Das Kapitel wird abgeschlossen durch die Beschreibung eines neuronalen Netzes, das ein durch die Bewegung des Roboters induziertes mittleres Flussfeld bestimmt und somit eine Vorhersage erwarteter Bilder ermöglicht. Durch Vergleich mit dem in [38] beschriebenen Verfahren wird gezeigt, dass die erlernte Bildtransformation die Basis für ein Bild-basiertes Heimfinden bilden kann.

Die Verwendung der Fouriertransformation für den Vergleich panoramischer Bilder wird in Kapitel 2 untersucht. Es wird ein Algorithmus zur effizienten Bestimmung der Relativorientierung in Sub-Pixel-Auflösung beschrieben, dessen Komplexität linear von der Zahl der Fourierkoeffizienten abhängt (im Vergleich zur quadratischen Abhängigkeit von der Zahl der Bildpixel bei herkömmlichen Methoden). Darauf aufbauend wird eine Methode für Bild-basiertes Heimfinden entwickelt und mit dem in Kapitel 1 und [38] beschriebenen Verfahren verglichen. Das Fourier-basierte Heimfinden zeichnet sich durch geringen Speicherbedarf und Rechenaufwand aus, besitzt jedoch kleinere Fangbereiche.

Kapitel 3 gibt eine mathematische Beschreibung panoramischer Stereosensoren mit nur einer Kamera für axialsymmetrische Spiegelflächen. Dabei werden sowohl die Auswirkungen von Kamerafehlstellungen als auch die

²Als Disparitäten werden Verschiebung zwischen korrespondierenden Bereichen des Stereobildes bezeichnet.

³Der Fangbereich einer Ortssignatur ist der Teil der Umgebung, innerhalb dessen eine Rückkehr zum Zielort möglich ist.

Abhängigkeit der Abbildungsqualität von der verwendeten Spiegelgeometrie untersucht. Hierfür wird die Lage der virtuellen Bild- und Kameraknotenpunkte berechnet. Als Spezialfall wird ein Stereosensor für einen autonomen mobilen Kleinroboter („Koala“) vorgestellt, der im Vergleich zu dem in Kapitel 1 beschriebenen Sensor eine deutlich größere Stereobasis und bessere Bildqualität besitzt.

In Kapitel 4 wird ein biologisch motiviertes aktives Stereokamerasystem vorgestellt, das mit Hilfe von Vergenzbewegungen die globale Bildkorrelation maximiert und anschließend die Stereobilder lokal auswertet. Im Vergleich zu panoramischen Stereosensoren wird infolge des begrenzten Sichtfeldes und der verbesserten Auflösung eine deutlich realistischere Modellierung des Stereosehens bei Säugetieren möglich. Dem System liegt eine Erweiterung des binokularen Energiemodells zugrunde, das auf elektrophysiologischen Erkenntnissen zu Disparitäts-selektiven Zellen beruht [70, 77]. Die Auswirkung der verwendeten zusätzlichen Normalisierung wird durch Berechnung von Wahrscheinlichkeitsdichten der neuronalen Aktivität für „Random-Dot“-Stereo-Stimuli untersucht. Die Normierung der Komplexzellantwort erweist sich dabei als entscheidender Vorteil für eine vom Bildkontrast weitgehend unabhängige Bestimmung der Disparitäten. Es wird u.a. gezeigt, dass die normalisierten Neurone vergleichbare Tuningkurven (bezüglich der Stimulusdisparität) besitzen wie nicht-normalisierte Komplexzellen, die Varianz der Zellaktivität jedoch deutlich geringer ausfällt.

Zur Verwendung mathematischer Symbole

Um das Lesen der mathematischen Formeln zu erleichtern, wurden Symbole nach einheitlichen Regeln vergeben. Gekennzeichnet sind:

- Vektoren durch fettgedruckte Buchstaben, z.B. \mathbf{I} , \mathbf{x} ,
- Matrizen durch fettgedruckte Buchstaben mit „Dach“, z.B. $\hat{\Sigma}$, $\hat{\mathbf{g}}$,
- Fourierkoeffizienten durch kalligraphische bzw. Fraktur-Buchstaben, z.B. \mathcal{I} , \mathfrak{a} .

$\delta(x)$ bezeichnet die Dirac'sche Delta-Distribution mit den Eigenschaften

$$\begin{aligned}\delta(x - x_0) &= 0, \quad x \neq x_0, \\ \int f(x)\delta(x - x_0)dx &= f(x_0),\end{aligned}$$

$\delta_{i,j}$ hingegen das diskrete Kronecker-Delta, d.h.

$$\delta_{i,j} := \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}.$$

Überblick zum Heimfinden mit Ortssignaturen

Der in dieser Arbeit verwendete Ansatz zum visuellen Heimfinden besitzt drei zentrale Komponenten:

1. Die Ortssignatur als Repräsentation des zu erreichenden Ortes,
2. eine Ähnlichkeits- bzw. Distanzmaß für Ortssignaturen,
3. ein Verfahren, das aus der aktuellen Signatur berechnet, welche Ortssignaturen für verschiedene Bewegungen zu erwarten sind.

In den einzelnen Schritten bei der Annäherung zum Ziel wird jeweils die Bewegungsentscheidung (φ, l) bzw. (x, y) gewählt, für die die berechnete Signatur die größte Ähnlichkeit zu der gespeicherten am Zielort aufweist.

Im Rahmen dieser Arbeit werden drei verschiedene Ortssignaturen für das visuelle Heimfinden verwendet: Panoramische Disparitäts- und Bildsignaturen in Kapitel 1, Abschnitte 1.3 und 1.4, sowie die Fourierkoeffizienten panoramischer Bilder in Kapitel 2, Abschnitt 2.5. Eine Zusammenstellung der jeweiligen Komponenten ist in Tabelle 1 zu sehen.

1. Ortssignatur	2. Ähnlichkeitsmaß	3. Vorausberechnung
Panoramische Disparitäts- und Varianzwerte $[\mathbf{d}, \mathbf{v}] = \{(d_i, v_i)\}_{i=0, 1, \dots, N-1}$ → Abschnitt 1.3	$E_d([\mathbf{d}, \mathbf{v}]^h, [\mathbf{d}, \mathbf{v}]^e)$, Glg. (1.40)	$d^e(\phi'(\phi \varphi, l)) = d'(\phi \varphi, l)$, $d'(\phi \varphi, l)$ aus Glg. (1.31) $\phi'(\phi \varphi, l)$ aus Glg. (1.32)
Panoramisches 1D-Bild, Grauwerte $\mathbf{I} = \{I_i\}_{i=0, 1, \dots, N-1}$ → Abschnitt 1.4.1	$E_I(\mathbf{I}^h, \mathbf{I}^e)$, Glg. (1.42)	$I^e(\phi'(\phi \varphi, l)) = I(\phi)$, $\phi'(\phi \varphi, l)$ aus Glg. (1.30) mit $r(\phi) = R_{\text{typ}} = \text{const}$
Fourierkoeffizienten eines panoramischen 1D-Bildes $\{(\mathbf{a}_k, \mathbf{b}_k)\}$ (Sinus-Cosinus-Form) bzw. $\{(C_k , \varphi_k)\}_{k=0, 1, \dots, K-1}$ (Betrag und Phase) → Abschnitt 2.5	$E_I(\{(\mathbf{a}_k^h, \mathbf{b}_k^h)\}, \{(\mathbf{a}_k^e, \mathbf{b}_k^e)\})$ Glg. (2.81)	$\mathbf{a}_k^e(x, y) = \mathbf{a}_k + \mathbf{a}_k^x x + \mathbf{a}_k^y y$, $\mathbf{b}_k^e(x, y) = \mathbf{b}_k + \mathbf{b}_k^x x + \mathbf{b}_k^y y$, Gln. (2.76) und (2.77)
	wird durch Optimierung der Gln. (2.86) – (2.88) zusammengefasst	

Tabelle 1: Überblick über die in dieser Arbeit für das visuelle Heimfinden verwendeten Ortssignaturen, sowie die zugehörigen Gleichungen für Ähnlichkeitsmaße und zur Vorhersage erwarteter Signaturen. Eine genaue Beschreibung erfolgt in den angegebenen Abschnitten.

Kapitel 1

Visuelles Heimfinden mit panoramischen Bildern¹

Ein von Cartwright und Collett [12, 13] vorgeschlagenes Modell zum visuellen Heimfinden von Honigbienen geht davon aus, dass diese an wichtigen Orten (z.B. Bienenstock, Futterquellen) einen so genannten „Schnappschuss“ (engl. „snapshot“), d.h. ein nur geringfügig verarbeitetes panoramisches Bild speichern. Durch ständigen Vergleich des aktuellen mit dem gespeicherten Bild kann dann die Richtung bestimmt werden, in die geflogen werden muss, um die Bildähnlichkeit zu erhöhen. Dieses Verfahren ist auf einen bestimmten Bereich um den Ort, an dem ein Bild aufgenommen wurde, beschränkt. Die Größe dieses „Fangbereichs“ (engl. „catchment area“) ist abhängig von der jeweiligen Umgebung.

Das „Schnappschuss-Modell“ wurde in verschiedenen Variationen auf mobilen Robotern implementiert und in Simulationen untersucht [38, 55, 67, 105]. Um Navigation in Umgebungen zu ermöglichen, die nicht innerhalb eines einzelnen Fangbereichs liegen, wurde in [37] eine Graphstruktur als Ortsgedächtnis des Roboters verwendet: Orte, an denen Bilder gespeichert wurden, werden als Knoten repräsentiert, wohingegen Kanten anzeigen, dass sich der Roboter mittels lokaler visueller Navigation zwischen den benachbarten Knoten bewegen kann.

Im folgenden Abschnitt wird der Aufbau eines passiven panoramischen Stereosensors beschrieben. Anschließend wird gezeigt, wie dieser (zusätzlich zum Bild-basierten) ein Disparitäts²-basiertes Heimfinden ermöglicht, das die geometrische Struktur eines Ortes berücksichtigt und deshalb von Beleuchtungsveränderungen weitgehend unabhängig ist.

¹Teile der Abschnitte 1.1–1.4 wurden bereits in [91] und [93] veröffentlicht.

²Bei Stereosystemen mit zwei Kameras wird mit „Disparität“ die relative Verschiebung von korrespondierenden Punkten in den beiden Kamerabildern bezeichnet. Die im Kamerabild des Stereosensors auftretenden Disparitäten werden in Abschnitt 1.2 beschrieben.

1.1 Panoramischer Stereosensor für Khepera

Zahlreiche Tierarten besitzen ein sehr großes Sichtfeld, das in horizontaler Richtung oftmals 300° und mehr beträgt. Neben offensichtlichen Vorteilen, z.B. für das Erkennen von Predatoren, bietet ein großer Sehbereich optimale Voraussetzungen für die Bestimmung der Eigenbewegung mit Hilfe des optischen Flusses [25] und das Wiederfinden bekannter Orte [12].

Aus ähnlichen Gründen werden auch in der Robotik, beispielsweise in [1, 79, 80, 108], Kamerasysteme eingesetzt, die panoramische Bilder der Umgebung aufnehmen. Diese ermöglichen einen von der Orientierung des Roboters unabhängigen Bildvergleich und damit eine vergleichsweise einfache Lokalisierung, siehe z.B. [38] und Kapitel 2. Große Sichtfelder werden sehr häufig durch so genannte katadioptrische Systeme erzielt, die aus einer Kombination von Linsen und Spiegelflächen bestehen [5]. Wie im Folgenden dargestellt wird, kann durch eine geeignete Anordnung zweier Spiegelflächen auch panoramisches Stereo auf begrenztem Raum realisiert werden. Neben anderen Mechanismen, wie der Auswertung der Bewegungsparallaxe oder „Shape from Shading“ (siehe [59] für eine genaue Beschreibung), bietet Stereosehen eine relativ einfache Möglichkeit, Distanzen zu umgebenden Objekten und damit die Geometrie eines Ortes visuell zu bestimmen.

Abbildung 1.1 zeigt den Stereosensor des Khepera-Roboters, der panoramische Stereobilder mit Hilfe einer einzelnen Kamera aufnimmt, die senkrecht nach oben auf einen zweiteiligen kegelförmigen Spiegel gerichtet ist. Der „Stereospiegel“ wurde aus Aluminium gedreht und anschließend poliert. Die beiden Spiegelteile besitzen leicht unterschiedliche Öffnungswinkel von $\gamma_1 = 97^\circ$ und $\gamma_2 = 109^\circ$. Durch die gewählte Anordnung wird eine effektive Stereobasis von $b \approx 8$ mm erzielt.

1.1.1 Berechnung der Spiegelgeometrie

Mit Hilfe des Reflexionsgesetzes erhält man aus Abbildung 1.2 für den Elevationswinkel

$$\varepsilon = \theta - \gamma + 90^\circ \quad , \quad (1.1)$$

wobei der Kamerawinkel θ über

$$-\frac{r}{f} = \tan \theta \quad (1.2)$$

mit dem Bildpunkt mit Abstand r von der optischen Achse verknüpft ist (f ist die Brennweite der Kamera). Für eine Kegelfläche ist somit der Bereich des Elevationswinkels durch den halben Öffnungswinkel der Kamera begrenzt (Spiegelgeometrien mit $\frac{\partial \varepsilon}{\partial \theta} \neq 1$ werden in Kapitel 3 behandelt).

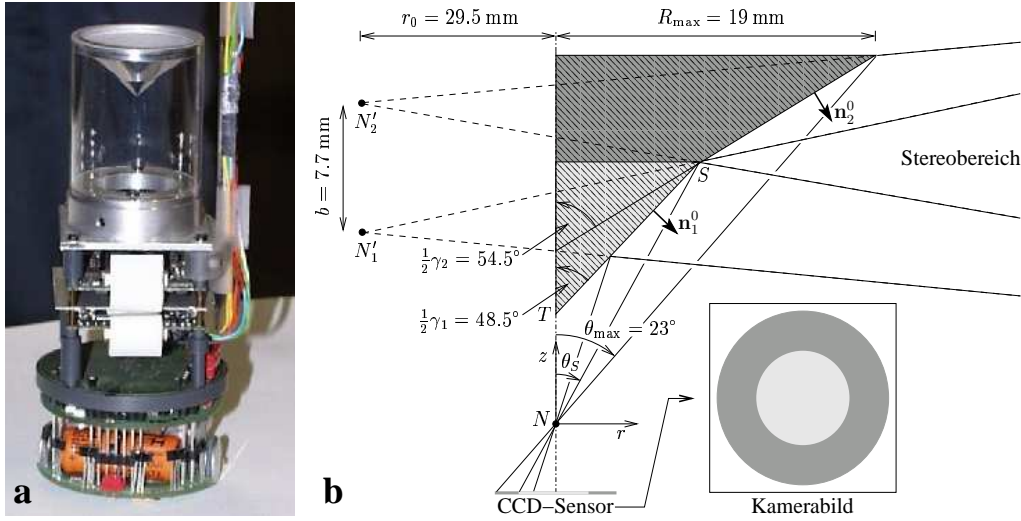


Abbildung 1.1: **a:** Khepera-Roboter mit Sensor für panoramisches Stereo (Durchmesser $\approx 5.5 \text{ cm}$, Höhe $\approx 13 \text{ cm}$). Rechts ist die Kabelverbindung zum externen Rechner zu erkennen. **b:** Nicht-maßstabgetreuer Axialschnitt durch den zweiteiligen Kegelspiegel. Die Stereoabbildung wird durch die vertikal übereinanderliegenden Spiegelpunkte N'_1 und N'_2 des Kameraknotenpunkts N verdeutlicht. Die Beschreibung der weiteren Größen erfolgt im Text. Das Kamerabild ist unten rechts schematisch dargestellt: die innere Kreisfläche (hellgrau) bzw. der äußere Ring (dunkelgrau) stellen die Bereiche dar, auf die über den unteren bzw. oberen Kegelspiegel abgebildet wird.

Der äußere Durchmesser des Kegelspiegels beträgt $2R_{\max} = 38.0 \text{ mm}$ (bedingt durch die Größe des Kheperas und der verfügbaren Plexiglasröhre, durch die der Spiegel oberhalb der Kamera angebracht ist), der effektiv verwendete Öffnungswinkel der Kamera $2\theta_{\max} = 46^\circ$, siehe Abb. 1.1 b. Für den Abstand Kameraknotenpunkt-Kegelgrundfläche gilt somit

$$z_{\max} := \frac{R_{\max}}{\tan \theta_{\max}} \approx 44.8 \text{ mm} . \quad (1.3)$$

Der Ort, an dem die beiden Kegel ineinander übergehen, ist der Schnittpunkt der Geraden

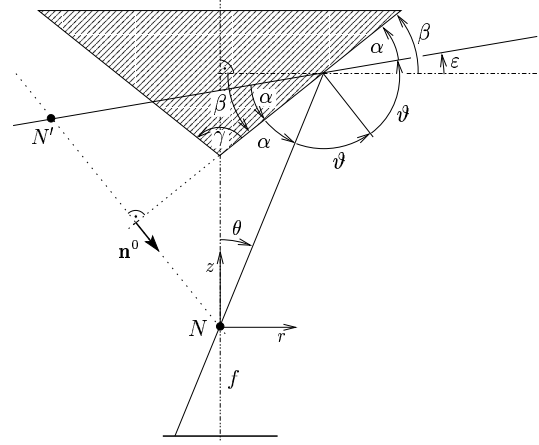
$$\mathbf{x}_{\theta_S} = \lambda \begin{pmatrix} \sin \theta_S \\ \cos \theta_S \end{pmatrix}, \quad \lambda \in \mathbb{R} \quad \text{und} \quad \mathbf{n}_2^0 \left(\begin{pmatrix} R_{\max} \\ z_{\max} \end{pmatrix} - \mathbf{x} \right) = 0 , \quad (1.4)$$

wobei mit $\mathbf{n}_i^0 := (\cos(\frac{1}{2}\gamma_i), -\sin(\frac{1}{2}\gamma_i))^T$ die Normalenvektoren der beiden Mantelflächen ($i = 1, 2$) bezeichnet werden. Der Index '1' bezieht sich auf den unteren Kegelspiegel, der auf den Kamerawinkelbereich $0 \leq \theta < \theta_S$

Abbildung 1.2: Zur Herleitung der Gleichung (1.1) für einen Kegelspiegel mit Öffnungswinkel γ : Aus der nebenstehenden Skizze liest man die Beziehungen

$$\begin{aligned}\alpha + \varepsilon &= \beta, \\ \beta + \frac{1}{2}\gamma &= 90^\circ, \\ \alpha + \beta + \theta &= 90^\circ\end{aligned}$$

ab und löst unter Elimination von α und β nach ε auf.



abbildet, Index '2' auf den oberen mit $\theta_S \leq \theta < \theta_{\max}$. Der Schnittpunkt hat bezüglich des Kameraknotenpunkts die Koordinaten

$$\mathbf{S} = \begin{pmatrix} r_S \\ z_S \end{pmatrix} := \overline{NS} \begin{pmatrix} \sin \theta_S \\ \cos \theta_S \end{pmatrix}, \quad (1.5)$$

$$\overline{NS} := \frac{R_{\max}}{\sin \theta_{\max}} \frac{\sin(\frac{1}{2}\gamma_2 - \theta_{\max})}{\sin(\frac{1}{2}\gamma_2 - \theta_S)}. \quad (1.6)$$

Virtuelle Kameraknotenpunkte

Betrachtet man eine Ebene, die durch die optische Achse (=Kegelachse) verläuft, so kann das Abbildungsverhalten des Stereosensors dadurch veranschaulicht werden, dass man sich zwei vertikal übereinanderliegende Kameras mit Knotenpunkten N'_1, N'_2 vorstellt, siehe Abb. 1.1 b. Die Orte dieser „virtuellen Kameraknotenpunkte“ erhält man, wie in den Abbildung 1.2 dargestellt, indem man den Kameraknotenpunkt N an der jeweiligen Kegelfläche spiegelt. Da S auf beiden Spiegelflächen liegt, gilt

$$\mathbf{N}'_i = 2(\mathbf{n}_i^0 \mathbf{S}) \mathbf{n}_i^0 = 2\overline{NS} \sin(\frac{1}{2}\gamma_i - \theta_S) \begin{pmatrix} -\cos(\frac{1}{2}\gamma_i) \\ \sin(\frac{1}{2}\gamma_i) \end{pmatrix}. \quad (1.7)$$

Abhängigkeit der Stereobasis von den Spiegelparametern

Aus (1.7) folgt für den Differenzvektor zwischen den virtuellen Kameraknotenpunkten

$$\Delta \mathbf{N} := \mathbf{N}'_2 - \mathbf{N}'_1 = 2\overline{NS} \sin[\frac{1}{2}(\gamma_2 - \gamma_1)] \begin{pmatrix} -\cos[\frac{1}{2}(\gamma_1 + \gamma_2) - \theta_S] \\ \sin[\frac{1}{2}(\gamma_1 + \gamma_2) - \theta_S] \end{pmatrix}.$$

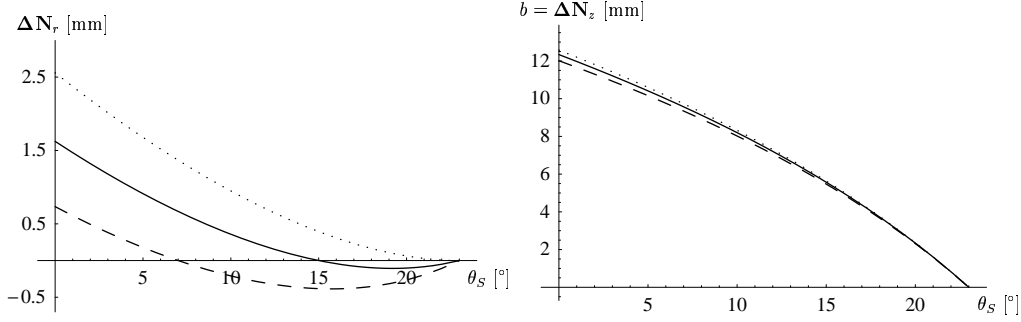


Abbildung 1.3: r - und z -Komponente von $\Delta\mathbf{N}$ (für $\theta_{\max} = 23^\circ$ und $R_{\max} = 19$ mm) in Abhängigkeit von θ_S für $\varepsilon_{\max} = 0^\circ$ (punktiert), 4° (kontinuierliche Kurve) und 8° (gestrichelt). Eine möglichst große Stereobasis $b := \Delta\mathbf{N}_z$ ist für die Genauigkeit der Distanzmessung mittels Stereo von Bedeutung. Zu beachten ist außerdem, dass infolge der geringen Auflösung im Bildzentrum der Bereich $0^\circ \leq \theta \lesssim 3^\circ$ nicht verwendbar ist.

Die Größe $b := \mathbf{N}'_{2z} - \mathbf{N}'_{1z}$ wird im Folgenden als „vertikale Stereobasis“ bezeichnet. Ein möglichst großes b ist vorteilhaft für eine gute Entfernungsauflösung mittels Stereo, vgl. Gleichung (1.25) in Abschnitt 1.2.3.

Für den in Abschnitt 1.2 beschriebenen Stereoalgorithmus wird für die maximalen Elevationswinkel der beiden Spiegelflächen $\max[\varepsilon_1] = \max[\varepsilon_2]$ gefordert, d.h. $\varepsilon_1(\theta_S) = \varepsilon_2(\theta_{\max})$. Aus (1.1) folgt mit $\varepsilon_{\max} := \max[\varepsilon_1] = \max[\varepsilon_2]$

$$\gamma_1 = 90^\circ - \varepsilon_{\max} + \theta_S, \quad (1.8)$$

$$\gamma_2 = 90^\circ - \varepsilon_{\max} + \theta_{\max}, \quad (1.9)$$

so dass nur noch θ_S und ε_{\max} geeignet gewählt werden müssen. Der Elevationsbereich des unteren Kegelspiegels ist durch $\varepsilon_{\max} - \theta_S \leq \varepsilon_1 \leq \varepsilon_{\max}$ festgelegt, für den oberen gilt $\varepsilon_{\max} - (\theta_{\max} - \theta_S) \leq \varepsilon_2 \leq \varepsilon_{\max}$.

In Abbildung (1.3) ist die r - und z -Komponente von $\Delta\mathbf{N}$ in Abhängigkeit von θ_S für $\varepsilon_{\max} = 0^\circ, 4^\circ, 8^\circ$ dargestellt. Für den Stereosensor des Kheperas wurde $\varepsilon_{\max} = 4^\circ$, $\theta_S = 11^\circ$ und somit $b \approx 7.7$ mm verwendet. Man erkennt, dass für diese Wahl $\Delta\mathbf{N}_r$ nur unwesentlich von Null verschieden ist. Bei einem benutzbaren Bereich des Kamerawinkels von $3^\circ \leq \theta \leq \theta_{\max} = 23^\circ$ erhält man für die Elevationsbereiche $-4^\circ \leq \varepsilon_1 \leq 4^\circ$ und $-8^\circ \leq \varepsilon_2 \leq 4^\circ$. Aus (1.8) und (1.9) folgt für den Öffnungswinkel des unteren Kegelspiegels $\gamma_1 = 97^\circ$ und für den oberen $\gamma_2 = 109^\circ$. Somit werden die beiden Horizonte nach Gleichung (1.1) auf die Kamerawinkel $\theta(\varepsilon_1 = 0^\circ) = \gamma_1 - 90^\circ = 7^\circ$ und $\theta(\varepsilon_2 = 0^\circ) = \gamma_2 - 90^\circ = 19^\circ$ abgebildet. Der Schnittpunkt S befindet sich gemäß Gleichung (1.5) bei (7.0 mm, 36.2 mm). Der Abstand der Kegelspitze T vom Kameraknotenpunkt beträgt $z_T = z_S - r_S \cot(\frac{1}{2}\gamma_1) \approx 30.0$ mm. Für

die virtuellen Kameraknotenpunkte folgt aus Gleichung (1.7): $\mathbf{N}'_1 \approx (-29.8, 33.7)^\top$ und $\mathbf{N}'_2 \approx (-29.5, 41.4)^\top$. Wegen der Rotationssymmetrie der Anordnung befinden sich die virtuellen Projektionspunkte auf zwei in etwa vertikal übereinander liegenden Kreisen mit Radien von ungefähr 3 cm.

Infolge kleiner Abweichungen von der berechneten Spiegelgeometrie im Außen- und vor allem im Übergangsbereich zwischen den beiden Spiegelteilen reduzieren sich die effektiv verwendbaren Bereiche. Um dennoch mittels Stereo auch geringe Entfernungen bestimmen zu können, wurde der Abstand des Spiegels zur Kamera gegenüber dem berechneten Wert leicht erhöht. Die abgebildeten Elevationsbereiche betragen ungefähr $\varepsilon_1 \in [-3^\circ, +2^\circ]$ und $\varepsilon_2 \in [-8^\circ, +2^\circ]$, siehe auch Abbildung 1.5.

Bemerkung: Der größere ε_2 -Bereich ist sinnvoll, da der Stereoalgorithmus (siehe Abschnitt 1.2) nach dem besten „Match“ des Bildes des unteren Elevationsbereichs im Bild des oberen sucht, siehe Abschnitt 1.2. Eine Umkehr der Verhältnisse, d.h. $-4^\circ \leq \varepsilon_1 \leq 8^\circ$ und $-4^\circ \leq \varepsilon_2 \leq 4^\circ$, hätte $\theta_S = 15^\circ$, $\gamma_1 = 101^\circ$ ($\gamma_2 = 109^\circ$ wie bisher) und die kleinere Stereobasis $b \approx 5.6$ mm zur Folge.

1.1.2 Justierung des Stereosensors

Die Position und Neigung des Kegelspiegels solange verändert, bis im Kamerabild die vertikalen Streifen eines zylinderförmigen Kalibrierungsmusters als durchgehende, von einem Zentrum radial nach außen verlaufende Linien und die horizontalen Streifen als konzentrische Kreise mit annähernd konstantem Abstand zu sehen waren. Anschließend wurden experimentell die Bildpunkte bestimmt, auf die die Horizontbereiche der beiden Spiegelflächen abgebildet werden. Eine grobe Bestimmung der (vertikalen) Winkelauflösung des Kamerabildes wurde ebenfalls mit Hilfe der horizontalen Streifen des Kalibrierungsmuster vorgenommen. Sie beträgt näherungsweise 0.22° je Pixel.

1.1.3 Abbildungsfehler eines Kegelspiegels

Für die Berechnung der Spiegelgeometrie wurde in Abschnitt 1.1.1 das Modell einer Lochkamera verwendet, d.h. es wurde nur der Strahlengang des Mittelpunktstrahls (Strahl, der durch den Kameraknotenpunkt N verläuft) betrachtet. Mit Hilfe der Kameralinse sollen jedoch benachbarte Strahlen auf einen Punkt fokussiert werden. Betrachtet man ein quasi-paralleles Strahlenbündel, das von einem sehr weit entfernten Objektpunkt ausgeht und an einem konischen Spiegel reflektiert wird, so erkennt man, dass nur die Strahlen, die vertikal versetzt verlaufen, nach der Reflexion parallel bleiben. Alle

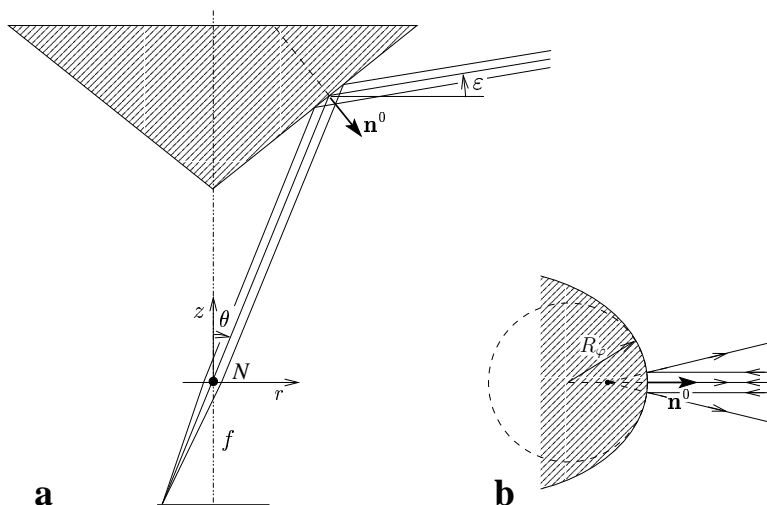


Abbildung 1.4: Abbildungsfehler des Kegelspiegels. **a:** Vertikal parallel versetzte Strahlen verlaufen auch nach Reflexion am Kegelspiegel parallel (Krümmung $\kappa_\theta = 0$). Alle anderen Strahlen eines parallelen Strahlenbündels werden jedoch unterschiedlich stark aufgefächert. Dies ist in **b** schematisch dargestellt für eine senkrechte Projektion auf den Normalenschnitt entlang der gestrichelten Linie (Krümmung $\kappa_\varphi = 1/R_\varphi \neq 0$).

anderen werden jedoch bei der Spiegelung unterschiedlich stark aufgefächert, siehe Abbildung 1.4. Dies hat, da nicht alle der ursprünglich parallelen Strahlen auf einen Bildpunkt fokussiert werden können, ein unscharfe Kamerabild zur Folge. Vertikale und horizontale Strukturen der Umgebung können nicht gleichzeitig scharf abgebildet werden – vergleichbar dem Astigmatismus bei optischen Linsen [44]. Besonders ausgeprägt ist dieser Effekt nahe der Kegelspitze, da dort sehr kleine Krümmungsradien R_φ auftreten.

Die Abbildungsqualität kann sowohl durch größere Spiegelradien und Abstände der Spiegel zur Kamera verbessert werden – mit dem Nachteil eines größeren Stereosensors – als auch durch konvex gekrümmte Spiegelflächen (mit $\frac{\partial \varepsilon}{\partial \theta} > 1$). Eine weiterführende Diskussion der Eigenschaften axialsymmetrischer katadioptrischer Systeme erfolgt in Kapitel 3.

Desweiteren hängt die Abbildungsqualität von der Blendenöffnung ab. Diese ist in Innenräumen bei schlechten Lichtverhältnissen weit zu öffnen, so dass sich Abbildungsfehler dort stärker auswirken als beispielsweise im Freien. Da jedoch die horizontale Auflösung der Bilder für das in den folgenden Abschnitten beschriebene visuelle Heimfinden mit 72 Pixel (dies entspricht einer horizontalen Auflösung von 5°) sehr niedrig gewählt wurde, stellt der Abbildungsfehler des Kegelspiegels keine ernsthafte Beeinträchtigung dar.

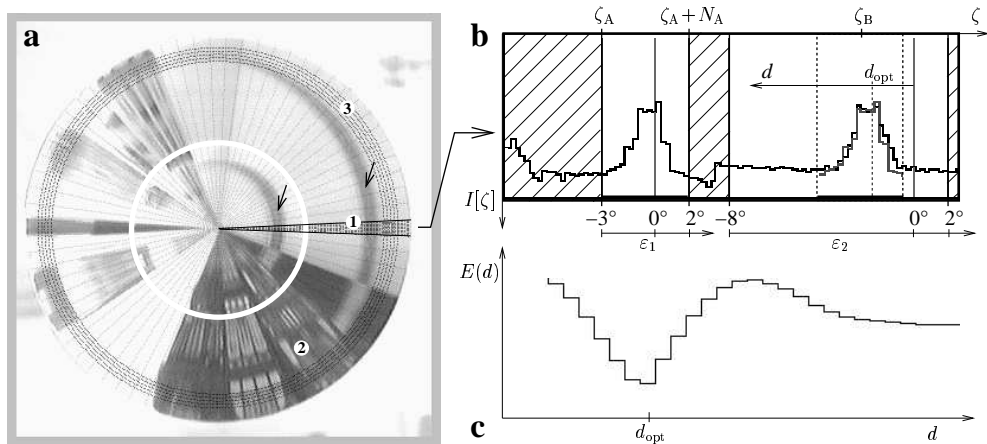


Abbildung 1.5: **a:** Unverarbeitetes Panorama-Stereobild. Eine horizontale Linie an den Wänden der Modellhausarena (Umgebung des Khepera-Roboters) wird in der Nähe des markierten Sektors (1) zweimal abgebildet (Pfeile). Bilder von Modellhäusern sind unten rechts zu erkennen (2). Die hervorgehobenen Pixel auf fünf konzentrischen Kreisen (3) werden für die Berechnung der in Abschnitt 1.4.1 beschriebenen Bildsignatur verwendet. **b:** Verlauf der Grauwerte innerhalb des markierten Sektors. Durch Suche nach der maximalen Korrelation zwischen dem inneren und äußeren Bildbereich wird die Bilddisparität ermittelt (siehe Fehlerfunktion in **c** und Text). Die schraffierten Bereiche werden wegen der geringen Auflösung im Bildzentrum (links) und infolge von Abbildungsfehlern im Übergangsbereich zwischen den beiden Teilen des Spiegels (Mitte) nicht verwendet.

1.2 Bestimmung vertikaler Disparitäten

Ein mit dem Stereosensor aufgenommenes panoramisches Stereobild ist in Abbildung 1.5 a zu sehen. Die Größe radialer Verschiebungen von korrespondierenden Bildausschnitten zwischen innerem und äußerem Bildbereich ist vom Abstand der abgebildeten Objekten abhängig. Da die Abbildung durch den Stereosensor – wie im letzten Abschnitt beschrieben – analog zu zwei vertikal versetzten rotierenden Kameras ist, werden die radialen Bildverschiebungen des Panoramabildes im Folgenden als „vertikale Disparitäten“³ bezeichnet. Deren Bestimmung wird in diesem Abschnitt beschrieben.

Wie in Abbildung 1.5 a dargestellt, werden die Bilder⁴ des Stereosensors in $N = 72$ Sektoren (5° -Bereiche horizontal) unterteilt. Jeder dieser Sektoren wird durch Mittelung über die unterschiedlichen Bogenlängen in ein

³Diese sind nicht zu verwechseln mit den vertikalen Disparitäten, die bei Stereosystemen mit horizontaler Basis (für nicht-parallele Kameraachsen) zusätzlich zu horizontalen Bildverschiebungen auftreten können, siehe Kapitel 6 in [59].

⁴Die verwendeten Bilder bestehen aus 8-Bit-Grauwerten $\in \{0, 1, \dots, 255\}$.

Grauwert-Array $I[\zeta]$, $\zeta = 0, 1, \dots, 99$, umgerechnet. Aufgrund des axialsymmetrischen Aufbaus des Stereosensors müssen sich korrespondierende Punkte der beiden Bildbereiche im gleichen Sektor befinden – die Epipolarlinien [99] verlaufen also radial nach außen.

Mit Hilfe eines einfachen korrelationsbasierten Stereoalgorithmus wird die mittlere Bildverschiebung zwischen den beiden Bildbereichen (Disparität) relativ zu den Horizontlinien $\varepsilon_1 = 0^\circ$ und $\varepsilon_2 = 0^\circ$ ermittelt. Die optimale Disparität d_{opt} folgt aus dem Minimum der Fehlerfunktion

$$E_m(d) := \sum_{\zeta=0}^{N_A-1} \left(I_A[\zeta] - I_B[\zeta - d] \right)^2 \quad (1.10)$$

$$:= \sum_{\zeta=0}^{N_A-1} \left(I[\zeta_A + \zeta] - I[\zeta_B + \zeta - d] \right)^2 . \quad (1.11)$$

$N_A = 20$ ist die Breite des verwendeten inneren Bildbereichs und ζ_B der Punkt im äußeren Bildbereich, der Disparität Null bezüglich ζ_A (Start des inneren Bildbereichs) besitzt, siehe Abb. 1.5 b. Die Bestimmung des globalen Minimums von $E_m(d)$ erfolgt durch Test aller möglichen ganzzahligen Disparitätswerte $d = 0, 1, \dots, N_d - 1$, wobei $N_d = 30$ die Anzahl der möglichen Disparitäten angibt. Es wurde nicht versucht, durch Interpolation eine Sub-Pixel-Auflösung der Disparität zu erhalten. Wie im Anhang A.1 gezeigt wird, ergibt die Minimierung von (1.10) einen mit den lokalen Bildkontrasten $\left(\frac{\partial}{\partial \zeta} I[\zeta]\right)^2$ gewichteten Mittelwert der lokalen Bilddisparitäten.

1.2.1 Varianzwert für Disparitäten

Zusätzlich wird zur jeder Disparität eine Größe v (im Folgenden als „Varianzwert“ bezeichnet) ermittelt, die angibt, wie genau und eindeutig das Minimum von $E_m(d)$ bestimmt werden konnte:

$$v := \max_d \left[(d - d_{\text{opt}})^2 \frac{2\sigma_n^2}{E_m(d) - E_m(d_{\text{opt}}) + 2\sigma_n^2} \right] , \quad (1.12)$$

wobei für die Varianz des Pixelrauschens der Wert $\sigma_n^2 = 9$ gewählt wurde (siehe Bemerkung weiter unten). Der ermittelte Disparitätswert wird dann als besonders zuverlässig angenommen (v ist klein), wenn alternative Verschiebungen einen höheren „Matching“-Fehler aufweisen, d.h. $E_m(d) - E_m(d_{\text{opt}})$ groß ist, und/oder diese nahe bei dem ermittelten Wert liegen, d.h. $(d - d_{\text{opt}})^2$ klein ausfällt. Der maximale Wert von v wird erreicht, wenn $E_m(d^*) - E_m(d_{\text{opt}}) = 0$ für ein d^* mit größtmöglicher Distanz zu d_{opt} , so dass $v = (d^* - d_{\text{opt}})^2$ gilt. Somit folgt für die obere Schranke $v \leq (N_d - 1)^2$. Die Berechnung eines Varianzwertes nach (1.12) wird im Folgenden begründet.

Maximum-Likelihood-Schätzung zur Berechnung des Varianzwerts

Nimmt man an, dass der Ausschnitt der Umgebung, der im inneren Bildbereich eines Sektors abgebildet wird, gleichzeitig im äußeren Bereich mit einer Verschiebung d erscheint, so gilt $\bar{I}_A[\zeta] = \bar{I}_B[\zeta - d]$. Infolge von Störungen weichen die Grauwerte im Kamerabild von den „wahren“ ab, d.h.

$$I_{A/B}[\zeta] = \bar{I}_{A/B}[\zeta] + \Delta I_{A/B}[\zeta] . \quad (1.13)$$

Betrachtet man mittelwertfreies Rauschen der Varianz σ_n^2 , das durch die Normalverteilung

$$P_n(\Delta I; \sigma_n^2) = \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(-\frac{\Delta I^2}{2\sigma_n^2}\right) , \quad (1.14)$$

beschrieben wird, so folgt für die Wahrscheinlichkeitsdichte der Differenz $\Delta I_{AB}[\zeta, d] := I_A[\zeta] - I_B[\zeta - d]$, unter der Voraussetzung, dass d die korrekte Disparität ist, d.h. für $\bar{I}_A[\zeta] = \bar{I}_B[\zeta - d]$:

$$\begin{aligned} P_\Delta(\Delta I_{AB}[\zeta, d]|d) &= \int P_n(I_A[\zeta] - \bar{I}_A[\zeta]; \sigma_A^2) P_n(I_A[\zeta] - \Delta I_{AB}[\zeta, d] - \bar{I}_B[\zeta - d]; \sigma_B^2) dI_A[\zeta] \\ &= \frac{1}{\sqrt{2\pi(\sigma_A^2 + \sigma_B^2)}} \exp\left(-\frac{(\Delta I_{AB}[\zeta, d] - (\bar{I}_A[\zeta - d] - \bar{I}_B[\zeta]))^2}{2(\sigma_A^2 + \sigma_B^2)}\right) \end{aligned} \quad (1.15)$$

$$= \frac{1}{\sqrt{2\pi(\sigma_A^2 + \sigma_B^2)}} \exp\left(-\frac{\Delta I_{AB}[\zeta, d]^2}{2(\sigma_A^2 + \sigma_B^2)}\right) . \quad (1.16)$$

Falls die Varianzen gleich sind, d.h. $\frac{1}{2}(\sigma_A^2[\zeta] + \sigma_B^2[\zeta]) =: \sigma_n^2 = \text{const} \forall \zeta$ gilt, und für unterschiedliche Pixel das Rauschen statistisch unabhängig ist, so erhält man aus (1.16) durch einfache Multiplikation die gemeinsame Wahrscheinlichkeitsdichte der Differenzen $\{\Delta I_{AB}[\zeta, d]\}_{\zeta=0,1,\dots,N_A-1}$:

$$P_\Delta(\{\Delta I_{AB}[\zeta, d]\}|d) = \frac{1}{\sqrt{2\pi} 2\sigma_n^{2N}} \exp\left(-\frac{\sum_\zeta \Delta I_{AB}[\zeta, d]^2}{2 2\sigma_n^2}\right) \quad (1.17)$$

$$= \frac{1}{\sqrt{2\pi} 2\sigma_n^{2N}} \exp\left(-\frac{E_m(d)}{2 2\sigma_n^2}\right) . \quad (1.18)$$

Bei einer Maximum-Likelihood-Schätzung, die auf der Bayes-Formel beruht [26, 76], folgt nun umgekehrt für die Wahrscheinlichkeit, dass d die gesuchte Disparität ist, bei gegebenen Differenzen $\{\Delta I_{AB}[\zeta, d]\}$:

$$P_d(d|\{\Delta I_{AB}[\zeta, d]\}) = K P_\Delta(\{\Delta I_{AB}[\zeta, d]\}|d) , \quad (1.19)$$

$$K := \left(\sum_d P_\Delta(\{\Delta I_{AB}[\zeta, d]\}|d)\right)^{-1} . \quad (1.20)$$

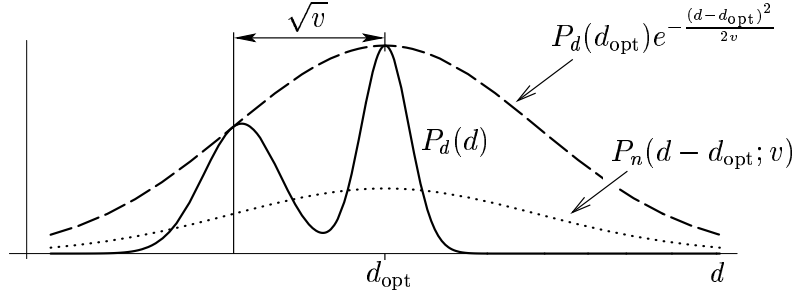


Abbildung 1.6: Näherung einer bimodalen Verteilung durch eine Normalverteilung mit Erwartungswert d_{opt} und Varianz v , die nach Gleichung (1.12) berechnet wird.

Wegen (1.19) und (1.18) entspricht einer Maximierung von $P_d(d|\{\Delta I_{\text{AB}}[\zeta, d]\})$ eine Minimierung von $E_m(d)$ bezüglich d . Im Folgenden wird zur Abkürzung $P_d(d|\{\Delta I_{\text{AB}}[\zeta, d]\})$ mit $P_d(d)$ bezeichnet.

Approximation durch Normalverteilung

Theoretisch wäre es möglich, als geometrische Repräsentation eines Ortes alle Wahrscheinlichkeiten $\{P_{d_i}(d)\}_{i=0,1,\dots,N_d-1}^{d=0,1,\dots,N_d-1}$ zu verwenden. Zur Reduktion der Daten werden jedoch nur die Disparitäten mit minimalen „Matching“-Fehlern $\{d_{\text{opt},i}\}$ und ihre Varianzwerte $\{v_i\}$, also nur $2N$ Werte (anstelle von $N_d \times N$) gespeichert. Dies kann, wie in Abbildung 1.6 dargestellt, als Näherung von $P_{d_i}(d)$ durch eine Normalverteilung $P_n(d - d_{\text{opt},i}; v_i)$ interpretiert werden. Die Verteilung nur lokal um das Optimum zu nähern, d.h. mögliche alternative Korrespondenzen (lokale Minima von $E_m(d)$) unberücksichtigt zu lassen, ist nicht ratsam: Man denke an periodische Strukturen, wie sie z.B. bei den in der Modellhausarena verwendeten Fachwerkhäusern auftreten können. Bei leicht veränderter Position des Roboters kann das globale Minimum dann durchaus bei einer alternativen Disparität liegen. Ein zu kleiner Varianzwert würde dann beim Heimfinden, das auf Vergleich von Disparitätssignaturen beruht (siehe Abschnitt 1.3.3), zu geringe Ähnlichkeiten zur Folge haben.

Unter der Annahme, dass $P_d(d)$ eine Normalverteilung mit Erwartungswert d_{opt} und Varianz σ_d^2 ist, gilt $P_d(d) = P_d(d_{\text{opt}}) \exp(-\frac{(d-d_{\text{opt}})^2}{2\sigma_d^2})$ und die Varianz kann durch

$$\sigma_d^2 = \frac{(d - d_{\text{opt}})^2}{-2 \ln\left(\frac{P_d(d)}{P_d(d_{\text{opt}})}\right)} \stackrel{(1.19)}{=} \frac{(d - d_{\text{opt}})^2}{-2 \ln\left(\frac{P_{\Delta}(\{\Delta I_{\text{AB}}[\zeta]\}d)}{P_{\Delta}(\{\Delta I_{\text{AB}}[\zeta]\}d_{\text{opt}})}\right)} \quad (1.21)$$

$$\stackrel{(1.18)}{=} \frac{(d - d_{\text{opt}})^2 2\sigma_n^2}{E_m(d) - E_m(d_{\text{opt}})} \quad (1.22)$$

berechnet werden. Obwohl $P_d(d)$ gewöhnlich höchstens näherungsweise eine Normalverteilung darstellt, wurde das Maximum von Gleichung (1.22) bezüglich d generell als Unsicherheitsmaß verwendet. Zusätzlich wurde – wie man durch Vergleich mit Gleichung (1.12) erkennt – noch eine Konstante $2\sigma_n^2$ hinzugefügt, um einen verschwindenden Nenner zu vermeiden. Diese ist dann von Bedeutung, wenn benachbarte Disparitäten den gleichen minimalen „Matching“-Fehler aufweisen: liegt beispielsweise die „wahre“ Disparität bei 1.5, so werden $E_m(1)$ und $E_m(2)$ in etwa gleich groß sein – ein Varianzwert $v \rightarrow \infty$ wäre in diesem Fall nicht sinnvoll.

Für Disparitäten in der Nähe des ermittelten Optimums ($d \approx d_{\text{opt}}$) erhält man näherungsweise

$$\begin{aligned}
E_m(d) &= \sum_{\zeta} (I_A[\zeta] - I_B[\zeta - d])^2 = \sum_{\zeta} (I_A[\zeta] - I_B[\zeta - d_{\text{opt}} - d + d_{\text{opt}}])^2 \\
&\approx \sum_{\zeta} \left(I_A[\zeta] - (I_B[\zeta - d_{\text{opt}}] + \partial I_B[\zeta - d_{\text{opt}}](d_{\text{opt}} - d)) \right)^2 \\
&= E_m(d_{\text{opt}}) + (d_{\text{opt}} - d)^2 \sum_{\zeta} \partial I_B[\zeta - d_{\text{opt}}]^2 . \tag{1.23}
\end{aligned}$$

Für die letzte Umformung wurde verwendet, dass $2 \sum_{\zeta} (I_A[\zeta] - I_B[\zeta - d_{\text{opt}}]) \times \partial I_B[\zeta - d_{\text{opt}}] = \frac{\partial}{\partial d} E_m(d_{\text{opt}}) = 0$ gilt, da d_{opt} die Fehlerfunktion $E_m(d)$ minimiert. Eingesetzt in (1.22) folgt somit

$$\sigma_d^2 \approx \frac{2\sigma_n^2}{\sum_{\zeta} \partial I_B[\zeta - d_{\text{opt}}]^2} \approx \frac{2\sigma_n^2}{\sum_{\zeta} \partial I_A[\zeta]^2} , \tag{1.24}$$

sofern der gefundene „Match“ korrekt ist, d.h. $I_B[\zeta - d_{\text{opt}}] \approx I_A[\zeta]$ gilt. Dies entspricht genau dem durch Varianzschätzung in Anhang A.1 ermittelten Wert $\langle \Delta d^2 \rangle$. Für unimodale Verteilungen gilt also näherungsweise $v \approx \langle \Delta d^2 \rangle$.

Bemerkung zu σ_n^2 : Für die Varianz eines Pixelwertes des Kamerabildes wurde bei üblicher Zimmerbeleuchtung $\sigma_{n,\text{cam}}^2 = 12$ gemessen. Durch die Mittelung in den Sektorbereichen reduziert sich dieser Wert. Da die Zahl der Pixel, über die gemittelt wird, mit zunehmendem Abstand vom Bildmittelpunkt zunimmt, ist $\sigma_n^2[\zeta] = \frac{1}{2}(\sigma_A^2[\zeta] + \sigma_B^2[\zeta])$ genau genommen keine Konstante (im inneren Bildbereich wurde $7 < \sigma_n^2 < 15$ gemessen, im äußeren $1 < \sigma_n^2 < 4$). Zum Kamerarauschen addieren sich weitere nicht-statistische Fehler, wie Bildverzerrungen durch ungenaue Justierung des Spiegels, Abbildungsfehler der Kamera und des Spiegels, nicht-konstante Bildverschiebung, diskrete Pixelabtastung etc. Zur Vereinfachung der Berechnungen wurde deshalb ein konstanter Wert $\sigma_n^2 = 9$ bzw. $\sigma_n = 3$ gewählt.

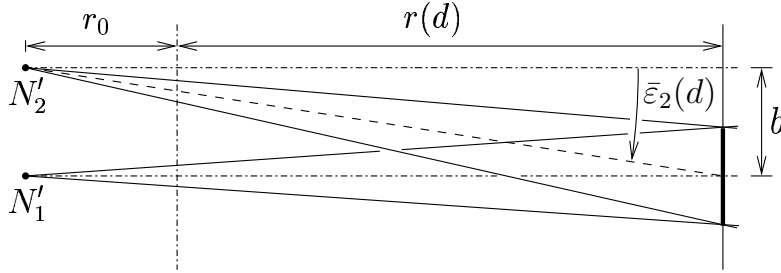


Abbildung 1.7: Berechnung der Objektdistanz r aus der Bildverschiebung d , siehe Gleichung (1.25). Das von N'_1 aus „gesehene“ Bild liegt von N'_2 aus betrachtet um Winkel $\bar{\epsilon}_2$ unterhalb des Horizonts. $\bar{\epsilon}_2$ hängt in guter Näherung linear von d ab: $|\frac{\partial \bar{\epsilon}_2}{\partial d}| \approx 3.8 \times 10^{-3}$ (in Radiant).

1.2.2 Disparitätssignatur eines Ortes

Nach der Stereobildauswertung können die ermittelten 72 Disparitäten⁵ und zugehörigen Varianzwerte, $[\mathbf{d}, \mathbf{v}] = \{(d_i, v_i), i = 0, 1, \dots, N - 1\}$, als Repräsentation des aktuellen Ortes verwendet werden. $[\mathbf{d}, \mathbf{v}]$ wird im Folgenden als Disparitätssignatur bezeichnet. Infolge der Verdeckung durch das Kabel, das den Khepera-Roboter mit einem externen Rechner verbindet (zu sehen in den Abbildungen 1.1 a und 1.8 a), können in einem Bereich von etwa 15° keine Disparitäten berechnet werden. Die drei entsprechenden Varianzwerte werden auf $\max[v] = (N_d - 1)^2$ gesetzt.

1.2.3 Berechnung von Objektdistanzen

Wie in Abbildung 1.7 dargestellt, kann aus der Disparität d (gemessen in Pixel) der Abstand zu einem Objekt bestimmt werden durch

$$r(d) \approx \frac{b}{\tan(-\bar{\epsilon}_2(d))} - r_0 \approx \frac{b}{\tan(|\frac{\partial \bar{\epsilon}_2}{\partial d}| d)} - r_0 \approx \frac{\alpha}{d} - r_0, \quad (1.25)$$

wobei für die Parameter $\alpha = b|\frac{\partial \bar{\epsilon}_2}{\partial d}|^{-1} \approx 2010$ mm/Pixel und $r_0 = -N'_{2z} \approx 29.5$ mm gilt. $\bar{\epsilon}_2 (\leq 0)$ ist der Elevationswinkel, um den ein im unteren Spiegel abgebildeter Gegenstand im oberen verschoben erscheint.

Der minimale Abstand (zur Symmetrieachse des Roboters), der mit dem Stereosensor gemessen werden kann, ist $\min[r] = r(N_d - 1) \approx 4$ cm (entspricht $\bar{\epsilon}_2 \approx -6^\circ$). Da der Durchmesser des Kheperas etwa 5.5 cm beträgt und die Infrarotsensoren, die sich auf der Sockelplatte des Kheperas befinden, bei

⁵Der Index „opt“ wird im Folgenden weggelassen.

einem Sensorabstand von ungefähr 2 cm die Hindernisvermeidung des Roboters aktivieren, treten in der Modellhausarena, die keine hervorstehenden Objekte enthält, praktisch nur Entfernungen größer als $\min[r]$ auf.

Für die Genauigkeit der Distanzmessung folgt aus (1.25)

$$\Delta r = \left| \frac{\partial r}{\partial d} \right| \Delta d = \frac{\alpha}{d^2} \Delta d = \frac{(r + r_0)^2}{\alpha} \Delta d \approx \frac{r^2}{\alpha} \Delta d . \quad (1.26)$$

Der Fehler nimmt also mit dem Quadrat der Entfernung zu.

Abbildung 1.8 zeigt Überlagerungen von Distanzen, die nach Bestimmung der Disparitäten mit Hilfe von Gleichung (1.25) berechnet wurden. Die geringere Auflösung für große Abstände ist sichtbar. Die Zahl der falschen Entfernungsschätzungen kann durch Beschränkung auf Disparitäten mit kleinen Varianzwerten v_i deutlich reduziert werden (Abb. 1.8 c). Der Grundriss der Modellhausarena ist nun klar zu erkennen. Dies zeigt, dass der Aufbau von Karten der Umgebung, wie beispielsweise in [57] oder [97] beschrieben, mit Hilfe des Stereosensors anstelle von Ultraschall-Sensoren oder eines Laser-Range-Finders möglich ist.

1.3 Disparitäts-basiertes Heimfinden

In diesem Abschnitt wird ein Algorithmus vorgestellt, der es dem Roboter ermöglicht, mit Hilfe der beschriebenen Disparitätssignaturen zu bekannten Orten zurückzufinden. Um einen Vergleich mit einem rein Bild-basierten Verfahren zu erleichtern, wurde der in [38] beschriebene Algorithmus für die Verwendung von Disparitäten erweitert.

1.3.1 Algorithmus

Um die richtigen Bewegungsentscheidungen zu ermitteln, werden aus der Signatur $[\mathbf{d}, \mathbf{v}]$ an der momentanen Position mehrere Signaturen $[\mathbf{d}, \mathbf{v}]_k^e := [\mathbf{d}^e(\varphi_k, l_k), \mathbf{v}^e(\varphi_k, l_k)]$, $k = 0, 1, \dots, N_e - 1$ berechnet (siehe Abschnitt 1.3.2), wie sie nach einer Bewegung (φ_k, l_k) des Roboters (Rotation um Winkel φ_k gefolgt von Translation der Länge l_k) zu erwarten sind, und mit der Disparitätssignatur am Zielort $[\mathbf{d}, \mathbf{v}]^h$ verglichen.⁶ Es werden $N_e = 132$ Positionen verwendet, die auf einem hexagonalen Gitter innerhalb eines Maximalabstands von $l_{\max} = 30$ cm liegen.

⁶Da vor allem große Entfernungen nur mit geringer Genauigkeit bestimmt werden können, siehe Gleichung (1.26), ist es sinnvoll, anstelle von Distanzen direkt Disparitätswerte zu vergleichen. Sollen dennoch Distanzen verwendet werden, sollte die entfernungsabhängige Genauigkeit durch entsprechende Gewichtung berücksichtigt werden.

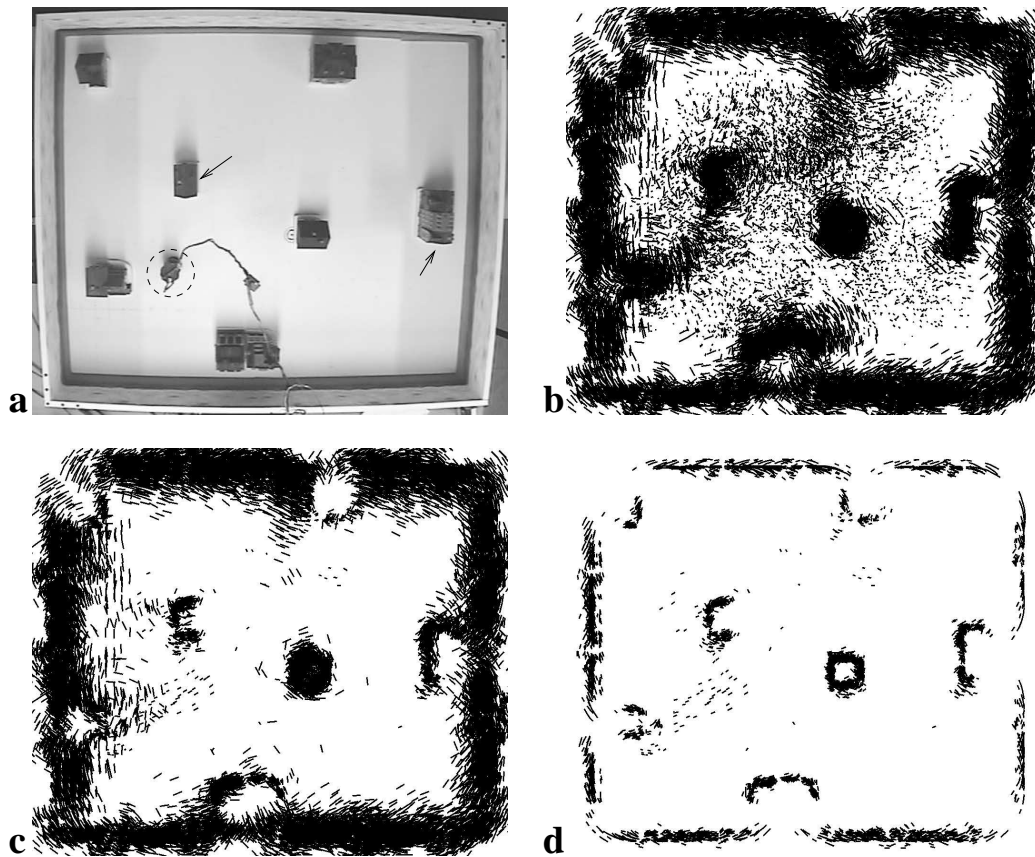
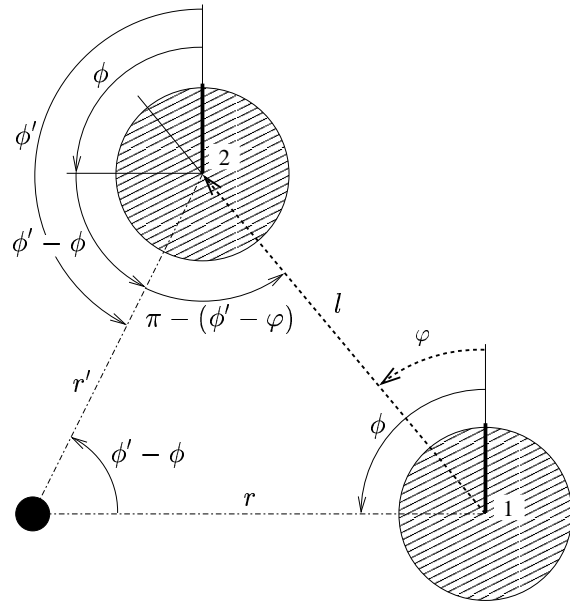


Abbildung 1.8: **a:** Modellhausarena (Größe etwa $140\text{ cm} \times 120\text{ cm}$) aus Sicht der Tracking-Kamera an der Decke. Der Khepera-Roboter ist durch einen Kreis markiert; Pfeile kennzeichnen Wände von Modellhäusern mit fehlender Struktur, für die der Stereoalgorithmus Disparitäten nur mit großer Unsicherheit bestimmen kann. **b–d:** Überlagerung von Distanzen, die für alle Aufnahmepositionen in der Bilddatenbank (siehe Abschnitt 1.4.3) nach Gleichung (1.25) berechnet wurden. **b:** Gezeigt sind alle Entfernungen $r(d) < 40\text{ cm}$. **c:** Nur die Abstände mit Varianzwerten $v < 1$ sind dargestellt. Die Zahl fehlerhafter Distanzen (infolge falscher Disparitäten) ist nun deutlich reduziert. Die Streuung infolge der Pixel-Auflösung der Disparitäten liegt für $r = 40\text{ cm}$ nach Gleichung (1.26) bei ungefähr $\Delta r(r = 40\text{ cm}, \Delta d = 1) \approx 8\text{ cm}$. **d:** Eine Beschränkung auf Abstände $r(d) \leq 20\text{ cm}$ verringert die Streuung, die durch die schlechte Auflösung für große Distanzen verursacht wird.

Anschließend führt der Roboter die Motorkommandos $(\varphi^{\text{est}}, l^{\text{est}})$ aus, für die die berechnete Signatur die größte Ähnlichkeit besitzt bzw. die in Abschnitt 1.3.3 beschriebene Fehlerfunktion $E_d([\mathbf{d}, \mathbf{v}]^h, [\mathbf{d}, \mathbf{v}]_k^e)$ den kleinsten

Abbildung 1.9: Ein Objekt mit Polarkoordinaten (ϕ, r) im lokalen Bezugssystem des Roboters (schraffierte Kreisscheibe, vertikale Linie zeigt Orientierung) besitzt nach einer Translation l in Richtung φ (bei unveränderter Orientierung des Roboters) die neuen Koordinaten (ϕ', r') .



Wert annimmt.⁷ Die Translation ist dabei auf maximal 5 cm beschränkt, um die Auswirkung einer einzelnen falschen Bewegungsentscheidung zu reduzieren.

Die beschriebenen Schritte werden solange wiederholt, bis der Ort größter Ähnlichkeit nur unwesentlich von der momentanen Position abweicht, d.h. $l^{\text{est}} < 5 \text{ mm}$ gilt⁸. Werden mehr als 50 Iterationen benötigt, wird das Heimfinden abgebrochen, da dann angenommen wird, dass sich der Startpunkt außerhalb des Fangbereichs der Signatur des Zielortes befand.

1.3.2 Berechnung erwarteter Disparitätssignaturen

Wie in Abbildung 1.9 dargestellt, kann die Position eines Objektes nach Translation des Roboters um l in Richtung φ bei konstanter Orientierung⁹ aus den ursprünglichen Koordinaten (ϕ, r) berechnet werden. Es gilt

$$r \begin{pmatrix} \cos \phi \\ \sin \phi \end{pmatrix} = l \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} + r' \begin{pmatrix} \cos \phi' \\ \sin \phi' \end{pmatrix} . \quad (1.27)$$

⁷Der Vektor $\mathbf{h}^{\text{est}} := l^{\text{est}}(\cos \varphi^{\text{est}}, \sin \varphi^{\text{est}})^{\top}$ wird im Folgenden auch als Schätzung des Heimvektors bezeichnet.

⁸Alternativ kann durch Überschreiten eines gewissen Schwellwerts bei der Ähnlichkeit der Signaturen das Erreichen des Zielorts festgestellt werden. Zusätzlich ändert sich die Richtung des Heimvektors in der Nähe des Zielortes stark.

⁹Der Khepera kann diese Endposition und -orientierung durch die Bewegungsabfolge „Rotation um φ , Geradeausfahrt der Länge l , Rotation um $-\varphi$ “ erreichen.

Für den neuen Abstand und Winkel folgt aus (1.27)

$$r' = \sqrt{r^2 - 2lr \cos(\phi - \varphi) + l^2} , \quad (1.28)$$

$$\tan(\phi' - \phi) = \frac{\frac{l}{r} \sin(\phi - \varphi)}{1 - \frac{l}{r} \cos(\phi - \varphi)} \implies \quad (1.29)$$

$$\phi'(\phi|\varphi, l) = \phi + \arctan \left[\frac{\frac{l}{r(\phi)} \sin(\phi - \varphi)}{1 - \frac{l}{r(\phi)} \cos(\phi - \varphi)} \right] . \quad (1.30)$$

Mit Hilfe von Gleichung (1.25) können die Abstände durch die entsprechenden Disparitäten ersetzt werden:

$$d'(\phi|\varphi, l) = \frac{\alpha}{r_0 + \sqrt{\left(\frac{\alpha}{d(\phi)} - r_0\right)^2 - 2l\left(\frac{\alpha}{d(\phi)} - r_0\right) \cos(\phi - \varphi) + l^2}} , \quad (1.31)$$

$$\phi'(\phi|\varphi, l) = \phi + \arctan \left[\frac{\frac{ld(\phi)}{\alpha - r_0 d(\phi)} \sin(\phi - \varphi)}{1 - \frac{ld(\phi)}{\alpha - r_0 d(\phi)} \cos(\phi - \varphi)} \right] . \quad (1.32)$$

Da Gleichung (1.30) bzw. (1.32) im Allgemeinen nicht nach ϕ aufgelöst werden kann, ist eine explizite Berechnung der erwarteten Signatur durch $d^e(\phi') = d'(\phi(\phi'))$ nicht möglich.¹⁰ Im Folgenden wird deshalb ein indirekter Weg zur Bestimmung von $d^e(\phi')$ beschrieben.

Verdeckungen

Wird durch die begrenzenden Objekte kein konvexer Bereich definiert, so können Verdeckungen auftreten. Diese sind nach Lu und Milios [58], die mit Laser-Range-Finder aufgenommene Entfernungsdaten vergleichen, in zwei Kategorien einzuteilen (siehe Abb. 1.10):

- (a) neu auftretende Verdeckungen, die aufgrund der vorhandenen Entfernungswerte zu erkennen sind. Hier wird für den betreffenden Abschnitt die kleinste Entfernung (größte Disparität) verwendet.
- (b) Bereiche, die vom aktuellen Standpunkt nicht einzublicken sind und für die entweder
 1. keine Distanzen vorliegen (markiert durch '?') oder
 2. sich die Reihenfolge benachbarter Entfernungswerte umkehrt (markiert durch '!').

¹⁰Für $l/r(\phi) \ll 1$ ist jedoch eine näherungsweise Lösung möglich, siehe Abschnitt 1.5.3.

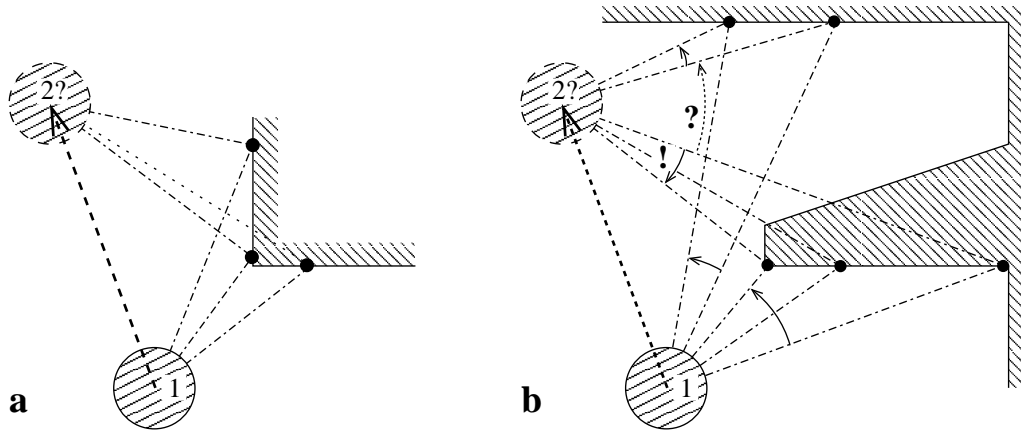


Abbildung 1.10: Berücksichtigung von Verdeckungen in nicht-konvexen Umgebungen bei der Berechnung von erwarteten Disparitätssignaturen: Im Gegensatz zu Verdeckungen, die aufgrund der aktuellen Disparitätswerte (in Position '1') erkannt werden können (a), ist für Bereiche, die vom aktuellen Standpunkt aus nicht einzusehen sind, eine korrekte Berechnung nicht möglich (b). Für die vorausberechnete Signatur (an der zu testenden Position '2?') liegen dann entweder überhaupt keine Werte für diesen Winkelbereich vor ('?') oder die Reihenfolge der Werte dreht sich um ('!').

In beiden Fällen ist keine Aussage über die zu erwartenden Distanz bzw. Disparitätswerte möglich. Dies wird dadurch berücksichtigt, dass die zugehörigen Varianzwerte auf den Maximalwert gesetzt werden.

Bestimmung von $d^e(\varphi, l)$ und $v^e(\varphi, l)$

Im Einzelnen erfolgt die Berechnung einer erwarteten Disparitätssignatur für eine mögliche Bewegung (φ, l) wie folgt (der Index $j = 0, 1, \dots, N - 1$ läuft über alle $N = 72$ Werte der zu berechnenden Signatur):

Zur Initialisierung werden alle Disparitätswerte der zu berechnenden Signatur auf $d_j^e(\varphi, l) = 0$ und alle Varianzwerte auf $v_j^e(\varphi, l) = \max[v] = (N_d - 1)^2$ gesetzt.

Für jeden Disparitätswert der aktuellen Signatur wird der erwartete Disparitätswert gemäß Gleichung (1.31), $d' = d'(\phi_j | \varphi, l)$, bestimmt. Der entsprechende Winkelbereich in der zu berechnenden Signatur wird mit Hilfe von Gleichung (1.32) ermittelt: $\phi'_{\min} = \phi'(\phi_j - \frac{1}{2}\Delta\phi | \varphi, l)$ und $\phi'_{\max} = \phi'(\phi_j + \frac{1}{2}\Delta\phi | \varphi, l)$, wobei $\phi_j = j\Delta\phi$, $\Delta\phi := \frac{2\pi}{N}$. Falls $(\phi'_{\max} - \phi'_{\min}) \bmod 2\pi > \pi$ gilt, wird angenommen, dass sich die Grenzen vertauscht haben und dieser Bereich somit von „hinten“ betrachtet wird (siehe '!' in Abbildung 1.10 b). v' erhält in diesem Fall den Wert $\max[v]$, andernfalls v_j zugewiesen (zur Vari-

anz eines erwarteten Disparitätswerts siehe Anhang A.2). Für alle Sektoren j , die mit mehr als zur Hälfte in den durch ϕ'_{\min} und ϕ'_{\max} definierten Bereich fallen, wird $d_j^e(\varphi, l) = d'$ und $v_j^e(\varphi, l) = v'$ gesetzt, falls d' größer ist als eventuell bereits eingetragene Werte, d.h. falls der neue Abstand kleiner ist als die bisher getesteten.

Um zu vermeiden, dass sehr unsichere Disparitäten in der berechneten Signatur zu falschen Werten führen, werden Disparitäten mit $v_j > 4$ von der Berechnung ausgeschlossen. Dies ist vor allem für große Disparitäten in Bewegungsrichtung von Bedeutung, die, da sie kleinen Distanzen entsprechen, einen großen Winkelbereich abdecken.

1.3.3 Ähnlichkeitsmaß für Disparitätssignaturen

Signaturwerte \mathbf{s} weisen infolge von Störungen Schwankungen auf. Sind die Abweichungen normalverteilt mit Kovarianzmatrix $\hat{\Sigma}$, so gilt für die Wahrscheinlichkeitsdichte von \mathbf{s} , wenn $\bar{\mathbf{s}}$ die „wahre“ Signatur ist:

$$P_{\mathbf{n}}(\mathbf{s} - \bar{\mathbf{s}}; \hat{\Sigma}) := \frac{1}{\sqrt{(2\pi)^N \det \hat{\Sigma}}} \exp\left(-\frac{1}{2}(\mathbf{s} - \bar{\mathbf{s}})^\top \hat{\Sigma}^{-1}(\mathbf{s} - \bar{\mathbf{s}})\right). \quad (1.33)$$

Somit folgt analog zur Herleitung in 1.2.1 für die Wahrscheinlichkeitsdichte der Differenzen $\Delta \mathbf{s}_{AB} := \mathbf{s}_A - \mathbf{s}_B$, falls \mathbf{s}_A und \mathbf{s}_B zur gleichen Signatur gehören ($\bar{\mathbf{s}}_A = \bar{\mathbf{s}}_B$):

$$\begin{aligned} P_{\Delta}(\Delta \mathbf{s}_{AB} | \bar{\mathbf{s}}_A = \bar{\mathbf{s}}_B) \\ = \left(\prod_{i=0}^{N-1} \int ds_{A,i} \right) P_{\mathbf{n}}(\mathbf{s}_A - \bar{\mathbf{s}}_A; \hat{\Sigma}_A) P_{\mathbf{n}}(\mathbf{s}_A - \Delta \mathbf{s}_{AB} - \bar{\mathbf{s}}_B; \hat{\Sigma}_B) \end{aligned} \quad (1.34)$$

$$= \frac{\exp\left(-\frac{1}{2} \Delta \mathbf{s}_{AB}^\top (\hat{\Sigma}_A + \hat{\Sigma}_B)^{-1} \Delta \mathbf{s}_{AB}\right)}{\sqrt{(2\pi)^N \det(\hat{\Sigma}_A + \hat{\Sigma}_B)}}. \quad (1.35)$$

Soll nun unter verschiedenen Signaturen \mathbf{s}_k , $k = 0, 1, \dots$ diejenige bestimmt werden, die die größte Ähnlichkeit zu einer gespeicherten Signatur \mathbf{s}^h besitzt, so kann dies formuliert werden als Suche nach der Signatur, für die $P_{\Delta}(\mathbf{s}^h - \mathbf{s}_k | \bar{\mathbf{s}}_h = \bar{\mathbf{s}}_k)$ den größten Wert annimmt. Maximieren von (1.35) entspricht einer Minimierung der negativen „Log-Likelihood“, siehe z.B. [76]:

$$\begin{aligned} -\ln P_{\Delta}(\mathbf{s}^h - \mathbf{s}_k | \bar{\mathbf{s}}^h = \bar{\mathbf{s}}_k) \\ = \frac{1}{2}(\ln[(2\pi)^N \det(\hat{\Sigma}^h + \hat{\Sigma}_k)] + (\mathbf{s}^h - \mathbf{s}_k)^\top (\hat{\Sigma}^h + \hat{\Sigma}_k)^{-1}(\mathbf{s}^h - \mathbf{s}_k)) \end{aligned} \quad (1.36)$$

Für den Vergleich von Disparitätssignaturen $[\mathbf{d}, \mathbf{v}]$ wird zur Vereinfachung angenommen, dass $\langle \Delta d_i \Delta d_j \rangle \approx v_i \delta_{i,j}$ gilt¹¹, d.h. dass die Disparitäten unkorreliert sind und v_i eine korrekte Schätzung der Varianz ist. Aus Gleichung (1.36) folgt dann

$$\begin{aligned} & -\ln P_{\Delta}(\mathbf{d}^h - \mathbf{d}_k^e | \bar{\mathbf{d}}^h = \bar{\mathbf{d}}_k^e) \\ & \approx \frac{1}{2} \sum_i \left(\ln[2\pi(v_i^h + v_{k,i}^e)] + \frac{(d_i^h - d_{k,i}^e)^2}{v_i^h + v_{k,i}^e} \right) . \end{aligned} \quad (1.37)$$

Besitzt man keine Kenntnis über die Orientierung des Roboters (und damit der aktuellen Signatur) relativ zur Signatur am Zielort, so müssen im einfachsten Fall alle möglichen Orientierungen $s = 0, 1, \dots, N - 1$ getestet werden¹²:

$$\min_s \left[\sum_i \ln[2\pi(v_{i[s]}^h + v_{k,i}^e)] + \frac{(d_{i[s]}^h - d_{k,i}^e)^2}{v_{i[s]}^h + v_{k,i}^e} \right] , \quad (1.38)$$

$$i[s] := (i + s) \bmod N . \quad (1.39)$$

Durch z.B. parabolische oder trigonometrische Interpolation kann Sub-Pixel-Genauigkeit bezüglich s erreicht werden. Ein einfacher „Gradientenabstieg in s “ ist, da die Fehlerfunktion lokale Minima besitzen kann, nicht zu empfehlen, sofern nicht eine Schätzung (z.B. durch Odometriemessungen) über die ungefähre Relativorientierung und somit für den Startwert vorliegt.

Da die Auswertung der \ln -Funktion in (1.38) vergleichsweise aufwendig ist, wurde als Ähnlichkeitsmaß für Disparitätssignaturen eine gewichtete Summe der quadratischen (Disparitäts-)Differenzen verwendet, d.h.

$$E_d([\mathbf{d}, \mathbf{v}]^h, [\mathbf{d}, \mathbf{v}]^e) := \min_s \left[\frac{\sum_i w(i, s) (d_{i[s]}^h - d_i^e)^2}{\sum_i w(i, s)} \right] , \quad (1.40)$$

$$w(i, s) := (v_{i[s]}^h + v_i^e)^{-1} . \quad (1.41)$$

Die Fehlerfunktion (1.40) führt bezüglich des Fangbereichs (getestet auf Bild-datenbank der Modellhausarena) zu sehr ähnlichen Resultaten wie (1.38), siehe Abschnitt 1.4.5 und Abbildung 1.13.

¹¹ $\langle x \rangle$ bezeichnet den Erwartungswert der Größe x .

¹²Ein schneller Bildvergleich unter Benutzung der Fouriertransformation wird in Abschnitt 2.1.5 vorgestellt.

1.4 Vergleich von Disparitäts- und Bildsignaturen

Im in diesem Abschnitt wird das beschriebene Disparitäts-basierte Heimfinden untersucht und mit einem bestehenden Bild-basierten Ansatz verglichen. Eine Möglichkeit beide Methoden zu kombinieren wird ebenfalls beschrieben.

1.4.1 Bild-basiertes Heimfinden

Das Heimfinden mit Bildsignaturen wurde bis auf unwesentliche Änderungen (72 statt 78 Pixel, andere Abbruchbedingung) wie in [38] beschrieben realisiert: Der Horizontbereich (siehe Markierungsnummer 3 in Abbildung 1.5 a) wird durch Mittelung und Histogrammausgleich in ein eindimensionales Panoramabild bestehend aus $N = 72$ Grauwert-Pixel konvertiert, wobei die Werte in dem vom Kabel verdeckten Bereich durch lineare Interpolation aus den benachbarten Grauwerten bestimmt werden.

Die Berechnung erwarteter Bilder erfolgt analog zu Abschnitt 1.3.2, wobei in Gleichung (1.30) gemäß der Annahme gleicher Abstände zu allen umgebenden Objekten („Equal Distance Assumption“ [38]) $r(\phi) = R = \text{const}$ gesetzt wird.¹³ Aus Gleichung (1.30) folgt, dass nur das Verhältnis $\lambda := l/R$ von Bedeutung ist. Ein Wert für R muss deshalb nicht festgelegt werden. Allerdings kann dann nur die Richtung zum Zielort bestimmt, jedoch keine Aussage über die Entfernung gemacht werden. Der Grauwert eines Objektes verändert sich – im Gegensatz zu Entfernungen – bei einer Bewegung des Roboters normalerweise kaum (wenn man von Schattenbildung, Spiegelungen o.ä. absieht), so dass für Bildsignaturen Gleichung (1.31) durch $I'(\phi|\varphi, l) = I(\phi)$ ersetzt wird. Dies entspricht der Annahme Lambert'scher Reflexion, siehe Kapitel 2 in [59].

Da alle Pixel der Bildsignatur vergleichbare Rauschpegel besitzen, wird für den Bildvergleich anstelle von Gleichung (1.38) bzw. (1.40) das einfachere Fehlermaß

$$E_1(\mathbf{I}^h, \mathbf{I}^e(\varphi_k, l_k)) := \min_s \sum_{i=0}^{N-1} \left(I_{i[s]}^h - I_i^e(\varphi_k, l_k) \right)^2 \quad (1.42)$$

verwendet. Mit diesen Anpassungen kann der in Abschnitt 1.3.1 beschriebene Algorithmus auch auf Bildsignaturen angewendet werden. Als typischer Abstand wird $R_{\text{typ}} = 45 \text{ cm}$ angenommen, so dass $\lambda \leq l_{\text{max}}/R_{\text{typ}} = 2/3$ gilt (vgl. auch Abbildung 1.23 b).

¹³Das in 1.3.2 beschriebene Verfahren eignet sich auch für die Berechnung von Bildsignaturen für variierende Entfernungswerte, siehe 1.4.5 und Abb. 1.13 (SSD*). Eine Möglichkeit, die Bildtransformation zu erlernen, ist in Abschnitt 1.6 beschrieben.

Bemerkung: Eine schnelle Berechnung erwarteter Bildsignaturen für konstante Distanzen R ist wie folgt möglich: Aus Gleichung (1.27) erhält man für $l \ll r(\phi) = R$:

$$\phi = \phi' - \arcsin\left[\frac{l}{r(\phi)} \sin(\phi' - \varphi)\right] \quad (1.43)$$

$$\approx \phi' - \frac{l}{R} \sin(\phi' - \varphi) \quad (1.44)$$

$$= \phi' - x \sin\phi' + y \cos\phi' , \quad (1.45)$$

wobei $x := \frac{l}{R} \cos\varphi$ und $y := \frac{l}{R} \sin\varphi$ substituiert wurde. Ist das aktuelle Bild $I(\phi)$, so folgt für das erwartete Bild an der Position x, y :

$$I^e(\phi' | x, y) \approx I(\phi' - x \sin\phi' + y \cos\phi') . \quad (1.46)$$

Wird zusätzlich $\sin\phi$ und $\cos\phi$ für die gewünschte Bildauflösung in einer Look-up-Tabelle gespeichert, so sind nur elementare Operationen erforderlich.

Die Näherung (1.44) erhält man auch, wenn man gleiche Abstände nicht für den aktuellen Ort sondern am Zielort postuliert, d.h. $r'(\phi') = R$. Aus Gleichung (1.27) folgt durch Eliminieren von r :

$$\phi = \phi' - \arctan\left[\frac{\frac{l}{r'(\phi')} \sin(\phi' - \varphi)}{1 + \frac{l}{r'(\phi')} \cos(\phi' - \varphi)}\right] \quad (1.47)$$

$$\approx \phi' - \frac{l}{r'(\phi')} \sin(\phi' - \varphi) \quad \text{falls } l \ll r'(\phi') \quad (1.48)$$

$$\approx \phi' - \frac{l}{R} \sin(\phi' - \varphi) . \quad (1.49)$$

1.4.2 Kombination von Disparitäts- und Bildsignatur

Zusätzlich wurde ein einfaches auf beiden Ortssignaturen basierendes Verfahren getestet. Die Berechnung der erwarteten Signaturen erfolgt jeweils getrennt, wie oben beschrieben. Die gesuchte Bewegungsentscheidung wird ermittelt gemäß

$$(\varphi^{\text{est}}, l^{\text{est}}) := \arg \max_{(\varphi_k, l_k)} \Pi\left(\frac{\tilde{E}_d(\varphi_k, l_k)}{\eta_d}\right) \Pi\left(\frac{\tilde{E}_I(\varphi_k, l_k)}{\eta_I}\right) , \quad (1.50)$$

$$\Pi(x) := (1 - \epsilon) \frac{1}{1 + x^2} + \epsilon , \quad (1.51)$$

$$\tilde{E}_d(\varphi, l) := E_d([\mathbf{d}^h, \mathbf{v}^h], [\mathbf{d}^e(\varphi, l), \mathbf{v}^e(\varphi, l)]) , \quad (1.52)$$

$$\tilde{E}_I(\varphi, l) := E_I(\mathbf{I}^h, \mathbf{I}^e(\varphi, l)) . \quad (1.53)$$

Für die in den nächsten Abschnitten beschriebenen Resultate wurde $\eta_d = 1$, $\eta_I = 10^4$ und $\epsilon = 0.01$ verwendet. Eine Anpassung der Konstanten (z.B. an den lokalen Bildkontrast) kann zu einer Verbesserung der Ergebnisse führen, wurde jedoch nicht untersucht.

$\Pi(\tilde{E}_d(\varphi_k, l_k)/\eta_d)$ bzw. $\Pi(\tilde{E}_I(\varphi_k, l_k)/\eta_I)$ können als (nicht-normierte) Wahrscheinlichkeiten dafür, dass (φ_k, l_k) bezüglich der Disparitäts- bzw. Bildsignatur eine korrekte Schätzung für den Heimvektor darstellt, interpretiert werden. Als Beispiel sei der Fall einer starken Beleuchtungsveränderung betrachtet. Im Normalfall gilt dann, da die aktuellen Bildern von den gespeicherten stark abweichen, $\tilde{E}_I(\varphi_k, l_k) \gg \eta_I$ für alle (φ_k, l_k) und folglich $\Pi(\tilde{E}_I(\varphi_k, l_k)/\eta_I) \approx \epsilon = \text{const}$, so dass die Disparitätssignatur „entscheidet“.

1.4.3 Datenbank mit panoramischen Stereobildern

Um visuelles Heimfinden mit Bild- und Disparitätssignaturen systematisch testen und die Auswirkungen von Beleuchtungsveränderungen untersuchen zu können, wurden zwei Bilddatensätze der Modellhausarena aufgenommen, die jeweils aus 1250 panoramischen Stereobildern bestehen. Die Aufnahmeorte in der Modellhausarena (siehe Abbildung 1.11 a) liegen näherungsweise auf einem 44×36 Gitter mit einer Zellgröße von $2.5 \times 2.5 \text{ cm}^2$ und besitzen einen Mindestabstand von etwa 6 cm zu den Modellhäusern und etwa 15 cm zu den Wänden der Arena. Letztere sind mit einer periodischen Landschaftstextur versehen. Unterschiedliche Beleuchtungsbedingungen für die beiden Bilddatensätze wurden durch eine Änderung der Position der Lichtquelle realisiert, siehe Abb. 1.12 a, c.

Um eine automatische Aufnahme durch den Khepera zu ermöglichen, ohne eine vollständige Wegplanung zur Vermeidung der Hindernisse zu benötigen, wurde das Verhalten des Roboters wie folgt gesteuert: Vor jeder Bewegung des Kheperas wird mit Hilfe des Tracking-Systems (Beschreibung siehe unten) die aktuelle Position des Roboters ermittelt. Weist die gemessene Position eine Abweichung von weniger als 3 mm zum Zielort auf (Genauigkeit des Trackingsystems liegt bei etwa 2 mm), wird ein Stereobild aufgenommen, in der Datenbank gespeichert und danach eine neue Zielposition ausgewählt. Anschließend versucht der Khepera, den Differenzvektor zwischen Ziel und aktueller Position abzufahren. Wird dabei (durch die aktiven Infrarot-Sensoren) in Fahrtrichtung ein Hindernis detektiert, so wird dieses „der-Wand-entlang“ umfahren, bis die Zielrichtung wieder verfolgt werden kann. Eine „Dead-Lock“-Situation, d.h. ein ständiges Wiederholen der gleichen Bewegungen ohne dem Ziel näher zu kommen, tritt nicht auf, da ausschließlich konvexe Hindernisse (Modellhäuser) vorhanden sind.

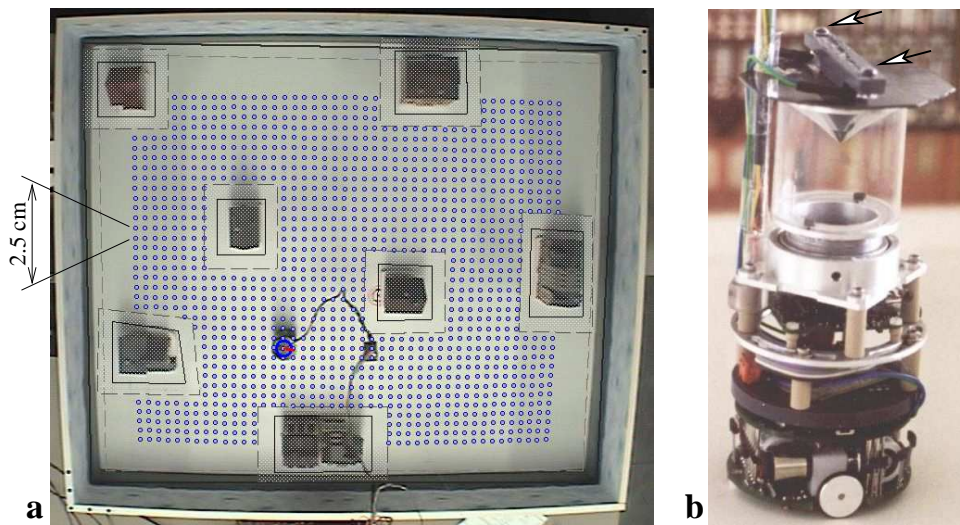


Abbildung 1.11: a: Blick von der Tracking-Kamera auf die Modellhausarena (Größe etwa $140\text{ cm} \times 120\text{ cm}$). Überlagert sind die 1250 Orte, an denen panoramische Stereobilder für die Bilddatenbank aufgenommen wurden. Die vom Tracking-System bestimmte aktuelle Position des Kheperas ist durch einen blauen Kreis markiert, seine Orientierung durch eine rote Linie. Dies wird durch die beiden Leuchtdioden ermöglicht (siehe Pfeile in **b**), die zum Zeitpunkt der Messung kurz aktiviert werden und somit leicht detektiert werden können.

Tracking-System

Um die absolute Position und Orientierung des Roboters messen zu können, wurde eine Farb-Videokamera an der Decke über der Roboterarena befestigt. Jedesmal wenn die Koordinaten des Kheperas bestimmt werden sollen, werden zwei farbige LEDs (grün und rot, Abstand 5 cm) am „Kopf“ des Kheperas (siehe Abb. 1.11 b) für die Zeitdauer von mindestens einem Vollbild (40 ms) ein- und anschließend wieder ausgeschaltet. Durch die Verwendung von Differenzbildern kann der Suchbereich für die LED-Positionen stark eingeschränkt werden. Außerdem ist die Bestimmung der Koordinaten mehrerer Roboter infolge der aktiven Steuerung der LEDs leicht zu realisieren.

Zur Umrechnung der Bildkoordinaten in Positionen wird ein regelmäßiges Punktmuster verwendet (Abstand zwischen benachbarten Punkten ist 5.5 cm), das auf Höhe des Kheperas ($\approx 13\text{ cm}$) in die Arena gelegt werden kann. Durch automatische Detektion der Punkte wird eine Beziehung zwischen Bild- und Raumkoordinaten festgelegt. Zwischen den Punkten können Positionen durch lineare Interpolation bestimmt werden, so dass eine explizite Bestimmung der Kameraparameter entfallen kann. Die Genauigkeit des

Systems beträgt ungefähr 2 mm für die Positionsmessung und etwa 2° für die Orientierungsbestimmung.

1.4.4 „Average Homeward Component“

Aus dem iterativen Vorgang beim Heimfinden (Abschnitt 1.3.1) folgt, dass die einzelnen Schätzungen für den Heimvektor nicht exakt sein müssen, da Abweichungen in den nachfolgenden Schritten korrigiert werden können. Entscheidend ist allerdings, dass der geschätzte Heimvektor \mathbf{h}^{est} (zumindest in der Mehrzahl der Schritte) eine Komponente in die Richtung des Ziels besitzt. Geht man von einer infinitesimalen Länge der Einzelschritte aus, so folgt, dass sich der Agent dem Ziel nähert, solange der Winkel zwischen \mathbf{h}^{est} und dem exakten Richtungsvektor \mathbf{h}^{true} kleiner als 90° ist. Als Maß für die Güte der berechneten Heimvektoren wird deshalb häufig die so genannte „Average Homeward Component“ [7] in Abhängigkeit vom Abstand \bar{l} zum Zielort betrachtet,

$$\text{AHC}(\bar{l}) := \left\langle \frac{\mathbf{h}^{\text{est}}(\bar{l})\mathbf{h}^{\text{true}}}{|\mathbf{h}^{\text{est}}(\bar{l})||\mathbf{h}^{\text{true}}|} \right\rangle = \langle \cos(\varphi^{\text{est}}(\bar{l}) - \varphi^{\text{true}}) \rangle, \quad (1.54)$$

wobei φ^{est} bzw. φ^{true} die Richtungswinkel von \mathbf{h}^{est} bzw. \mathbf{h}^{true} bezeichnen.

Abbildung 1.12 zeigt die „Average Homeward Component“ der berechneten Heimvektoren für unterschiedliche Beleuchtungsbedingungen und Ortssignaturen in Abhängigkeit vom Abstand \bar{l} zum Zielort. Alle möglichen 1250 Zielpositionen der Bilddatenbank wurden dafür ausgewertet, Breite der Bins beträgt 2.5 cm. Der starke Abfall für kleine \bar{l} ist vor allem darauf zurückzuführen, dass erwartete Signaturen nur für diskrete Positionen auf einem hexagonalen Gitter berechnet und somit für kleine Abstände nur wenige Winkel getestet werden. Auch ist für kleine Distanzen zum Zielort die Richtung des Heimvektors sehr empfindlich bezüglich Rauschen, siehe Gleichungen (1.106) und (1.120) in Abschnitt 1.5.

Bleibt die Beleuchtung zwischen der Aufnahme der Signatur am Zielort und der Berechnung des Heimvektors im Wesentlichen unverändert („konstante Beleuchtung“: simuliert durch Verwendung des gleichen Bilddatensatzes), so besitzen die mit Hilfe der Bildsignaturen berechneten Heimvektoren für nicht zu große Abstände eine etwas größere Komponente zum Zielort als bei Verwendung von Disparitätssignaturen.

Falls jedoch die Richtung zum Zielort bei stark veränderten Beleuchtungsverhältnissen (Abb. 1.12 c) erfolgen soll (simuliert durch Verwendung unterschiedlicher Datensätze für die Signaturen an den Zielpositionen und an den Testpositionen), erweisen sich die Bildsignaturen als sehr unzuverlässig. Die Bestimmung der Heimvektoren mit Hilfe von Disparitätssignaturen hingegen

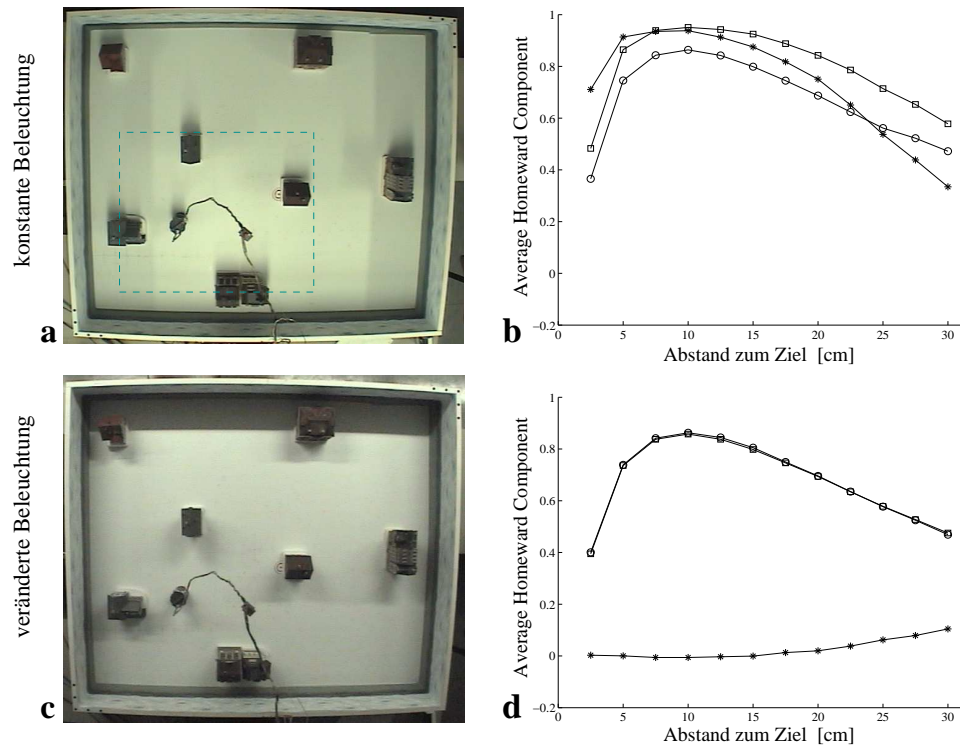


Abbildung 1.12: Abhängigkeit der AHC von der Beleuchtung für verschiedene Ortssignaturen: **a:** Beleuchtungsverhältnisse in der Modellhausarena bei der Aufnahme der Ortssignaturen (das gestrichelte Rechteck markiert den in Abbildung 1.14 dargestellten Bereich der Arena). **b:** AHC in Abhängigkeit vom Abstand \bar{l} zum Zielort bei unveränderten Beleuchtungsverhältnissen für Bildsignaturen (*), Disparitätssignaturen (○) und deren Kombination (□). **c:** Beleuchtungsverhältnisse nach Veränderung der Position der zentralen Lichtquelle in „Nord“-Richtung. **d:** AHC nach Veränderung der Beleuchtung, d.h. die aktuelle Signatur und die Signatur am Zielort werden bei unterschiedlichen Beleuchtungsbedingungen aufgenommen. Die Heimvektoren für Bildsignaturen besitzen in diesem Fall fast keine Komponente in Richtung des Zielorts, wohingegen die AHC für Disparitätssignaturen im Wesentlichen unverändert bleibt. Bei der kombinierten Methode bestimmen daher die Disparitätssignaturen die Richtung des Heimvektors.

ist – wie aufgrund der impliziten Repräsentation der Geometrie eines Ortes erwartet – sehr robust in Bezug auf Beleuchtungsveränderungen.

Die in Abschnitt 1.4.2 beschriebene Kombination beider Signaturen kann im ersten Fall das Niveau relativ zu den reinen Disparitätssignaturen anheben, ohne bei starken Beleuchtungsveränderungen unter das Niveau der Disparitätssignaturen zu fallen.

1.4.5 Fangbereiche

Eine entscheidende Größe für Landmarken-basierte Navigation ist der Bereich, in dem eine Landmarke erkannt und zur Orientierung verwendet werden kann. Das Heimfinden mit Hilfe der in dieser Arbeit verwendeten Ortssignaturen ist üblicherweise auf ein (einfach) zusammenhängendes Gebiet der Umgebung beschränkt. Dieses wird als „Fangbereich“ bzw. „catchment area“ [13] der jeweiligen Signatur bezeichnet.

Je größer die Fangbereiche desto weniger Signaturen sind nötig, um ein vorgegebenes Gebiet abzudecken. Soll ein Agent allein durch visuelles Heimfinden zu einer bestimmten Anzahl von Punkten zurückfinden können, so ist entscheidend, dass an Orten zwischen diesen Punkten eine kontinuierliche Kette von Signaturen derart gespeichert wird, dass die Signaturen jeweils im Fangbereich der beiden Nachbarn liegen. Die Zahl der zusätzlich benötigten Ortssignaturen sinkt mit zunehmender Größe der Fangbereiche.

Neben der Größe des Fangbereichs ist auch dessen Invarianz gegenüber möglichen Veränderungen von entscheidender Bedeutung. Beispiele sind Veränderungen der Lichtverhältnisse durch unterschiedlichen Sonnenstand oder Wolkenbewegungen, siehe [109]. Für Tiere, die Vorräte für die Wintermonate anlegen, sind zusätzlich saisonale Unterschiede von Bedeutung.

Aus den Resultaten zur „Average Homeward Component“ im letzten Abschnitt ist zu erwarten, dass die Fangbereiche der in Abschnitt 1.2.2 eingeführten Disparitätssignaturen – im Gegensatz zur Verwendung von Bildsignaturen – weitgehend invariant gegenüber Beleuchtungsveränderungen sind. Dies wird im Folgenden untersucht.

Simulationen zur Bestimmung der Fangbereiche

Mit Hilfe der Bilddatenbank können „Homing-Versuche“ in großer Anzahl simuliert werden. Diese wurden für die im Folgenden beschriebenen Ergebnisse für alle möglichen 1250 Zielpositionen von allen 1250 Positionen aus gestartet – eine Anzahl die mit dem Khepera kaum realisierbar ist.

Da die Datenbank nur Bilder für diskrete Positionen innerhalb der Roboterarena enthält, wurde der in Abschnitt 1.3.1 beschriebene Algorithmus angepasst.¹⁴ Jedesmal wenn der simulierte Agent die dem Heimvektor entsprechende Bewegungen ausgeführt hat, wird der Abstand zur nächstliegenden Aufnahmeposition bestimmt. Ist dieser größer als 2.1 cm, dann hat sich

¹⁴Alternativ können Bilder zwischen Aufnahmepositionen durch (z.B. bilineare) Interpolation berechnet werden. Dies verlangsamt allerdings selbst bei Bildern der Größe 72×100 die Simulation erheblich und führt zu Ausschmierungen von Objektgrenzen. Auch kann bei genauer Kenntnis der Abstände ein 3D-Modell der Arena erstellt werden und damit Bilder für beliebige Positionen generiert werden.

der Agent aus dem Bereich der Aufnahmepositionen entfernt¹⁵ und ist auf ein Hindernis (Modellhäuser) „gestoßen“ oder bewegt sich auf die Arena-wand zu. Dies würde beim realen Roboter über Infrarotsensoren detektiert werden und eine Hindernisvermeidungsreaktion auslösen. In der Simulation wird in diesem Fall das Heimfinden abgebrochen. Ebenfalls zum Abbruch führt der Fall, in dem die nächstliegende Position mit der letzten Aufnahme-position (der Ort, an der der aktuelle Heimvektor berechnet wurde) identisch ist. In diesem Fall wird angenommen, dass die vermutete Zielposition bereits erreicht ist.¹⁶ In beiden Fällen wird der Abstand zum tatsächlichen Zielort gespeichert. Anderenfalls wird der Agent auf die nächstliegende Aufnahme-position gebracht und ein neuer Heimvektor berechnet.

Um Rechenzeit zu sparen, wird der Abstand zur exakten Zielposition nach Beendigung des simulierten „Homing-Versuchs“ in jeder zuvor besuchten Aufnahme-position gespeichert. Somit kann, wenn in den nachfolgenden Durchläufen (mit anderen Startpositionen) eine bereits bekannte Position angefahren wird, dieser Wert einfach übernommen werden.

Resultate

Abbildung 1.13 vergleicht die mittlere Größe von Fangbereichen für unterschiedliche Ortssignaturen. Die Größe eines Fangbereichs wurde durch die Zahl der Startpunkte bestimmt, für die der Zielort mit einem maximalen Fehler von $l_{CA} = 10$ cm erreicht werden konnte.

Für das Disparitäts-basierte Heimfinden werden drei verschiedene Ähnlichkeitsmaße verglichen: „ML“: Gleichung (1.38), die in 1.3.3 mit Hilfe einer „Maximum-Likelihood“-Schätzung abgeleitet wurde. „wSSD“: gewichtete SSD („Sum of Squared Differences“), Gleichung (1.40). Diese erlaubt einen schnelleren Vergleich von Signaturen (da sie keine Auswertung der In-Funktion benötigt) und besitzt nur geringfügig kleinere Fangbereiche. „SSD“: einfache SSD, wobei Disparitäten in Winkelbereichen, in denen keine Disparitäten gemessen (durch Kabel verdeckter 15° -Bereich) bzw. berechnet werden können, durch lineare Interpolation aus den benachbarten Disparitäten bestimmt werden. Die deutlich verringerte Größe der Fangbereiche rechtfertigt die Verwendung der Varianz-Werte v_j in „ML“ und „wSSD“.

Für das Bild-basierte Heimfinden wird als Ähnlichkeitsmaß die einfache SSD verwendet, Gleichung (1.42). Neben der in 1.4.1 beschriebenen Berech-

¹⁵Dieser Wert ergibt sich aus dem maximalen Abstand von $\frac{5}{4}\sqrt{2}(\approx 1.77)$ cm zu einem Gitterpunkt innerhalb eines regelmäßigen Gitters mit Zellgröße 2.5×2.5 cm² und der Positionstoleranz der Aufnahme-punkte von 3 mm.

¹⁶Dies hat zur Folge, dass kleine Heimvektoren, die, wenn sie wiederholt auftreten, den realen Roboter näher zum Ziel führen können, unberücksichtigt bleiben.

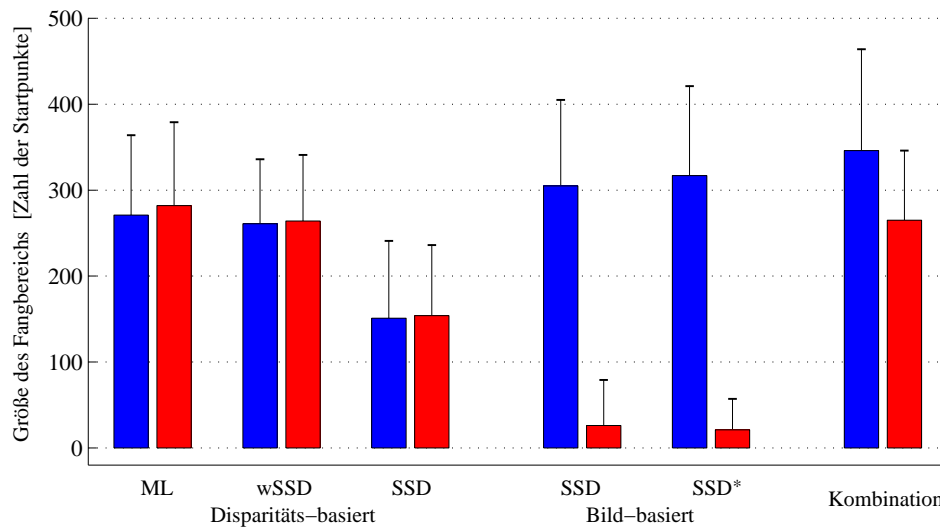


Abbildung 1.13: Größenvergleich der Fangbereiche für das Heimfinden mit Hilfe von Disparitätssignaturen (links), Bildsignaturen (Mitte) und kombinierten Signaturen (rechts) für unterschiedliche Beleuchtungsbedingungen (blaue Balken: „konstante Beleuchtung“, rote Balken: „veränderte Beleuchtung“ im Vergleich zur Aufnahme der Signatur am Zielort). Dargestellt sind die Mittelwerte (über alle 1250 Zielpositionen) und Standardabweichungen (Fehlerbalken). Weitere Erläuterungen im Text.

nung erwarteter Bilder unter der Annahme gleicher Abstände $r(\phi) = R$, wurde ein weiteres Verfahren getestet (SSD*): Abstände zu benachbarten Objekten werden aus den Disparitäten mit Varianzwert $v_j \leq 4$ ($j = 0, 1, \dots, N-1$) berechnet. Für $v_j > 4$ wird der Abstand $R_{\text{typ}} = 45$ cm verwendet. Für Bildbereiche, in denen infolge von Verdeckungen keine Grauwerte berechnet werden können, werden diese durch lineare Interpolation aus den benachbarten Werten bestimmt. Allerdings rechtfertigt der nur geringfügig größere Mittelwert der Fangbereiche im Vergleich zur einfachen Berechnung mit $r(\phi) = R$ den zusätzlichen Aufwand kaum.

Als Beispiel zeigt Abbildung 1.14 einen Vergleich von Fangbereichen für Disparitäts- und Bild-basiertes Heimfinden bei unterschiedlichen Beleuchtungsbedingungen. Die Zielposition befindet sich innerhalb des in Abb. 1.12 a markierten Bereichs.

Bei konstanter Beleuchtung können durch Bildsignaturen im Mittel größere Fangbereiche erzielt werden als mit Disparitätssignaturen. Hierbei ist allerdings zu berücksichtigen, dass (vor allem für Disparitätssignaturen) die Größe des Fangbereichs davon abhängt, in welchem Bereich um den aktuellen Standpunkt und in welcher Anzahl erwartete Signaturen berechnet und

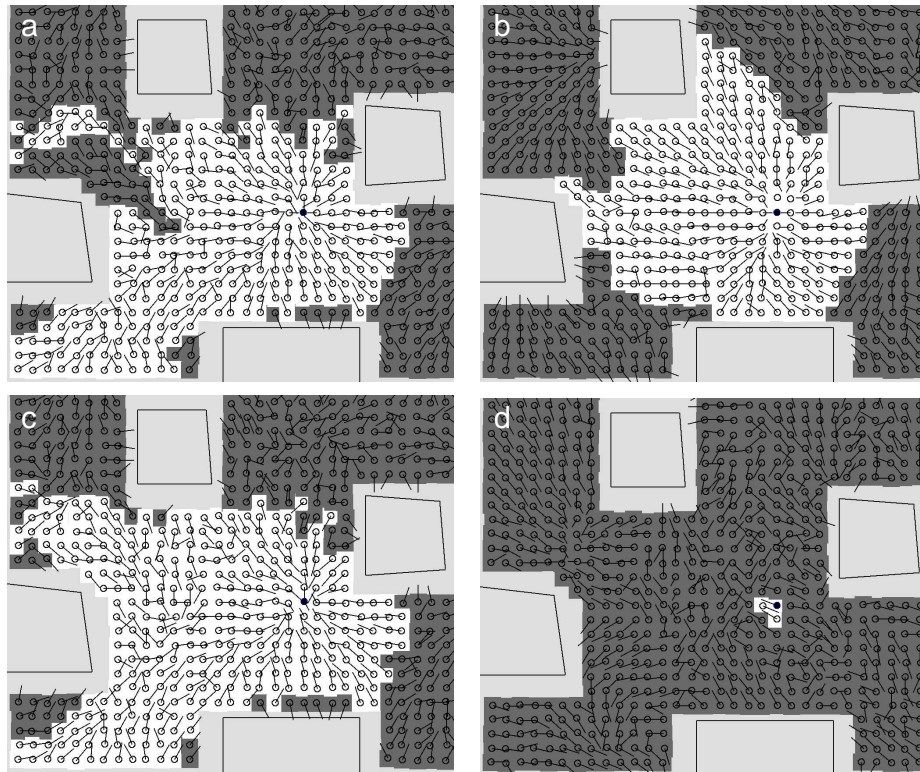


Abbildung 1.14: Beispiele von Fangbereichen (weiß unterlegte Flächen) für eine Zielposition (schwarzer Punkt) im linken unteren Bereich der Modellhausarena. Gezeigt ist jeweils der in Abbildung 1.12 a markierte Ausschnitt. Obere Reihe: Gleiche Beleuchtung bei Aufnahme der Signatur am Zielort und den „Homing-Versuchen“ mit (a) Disparitäts- bzw. (b) Bildsignaturen. Untere Reihe: Veränderte Beleuchtung beim Heimfinden mit (c) Disparitäts- bzw. (d) Bildsignaturen (Fangbereich besteht lediglich aus drei benachbarten Zellen). Polygone kennzeichnen die Positionen der Modellhäuser. Die Aufnahmeorte der panoramischen Stereobilder in den Bilddatenbanken sind durch kleine Kreise markiert. Die von diesen ausgehenden Linien geben die Richtung und die (aus Gründen der Übersicht auf 3 cm beschränkte) Länge der berechneten Heimvektoren an.

mit der Signatur am Zielort verglichen werden. Desweiteren gilt diese Aussage nur für die untersuchte Umgebung. Infolge des vertikalen Versatzes der virtuellen Kameraknotenpunkte ist der Stereoalgorithmus vorrangig auf horizontale Strukturen angewiesen. Das auf Bildsignaturen basierende Heimfinden benötigt jedoch Helligkeitsunterschiede entlang des Horizonts.

Bei starker Veränderung der Beleuchtung zwischen dem Zeitpunkt der Aufnahme der Signatur am Zielort und dem „Homing-Versuch“ (in den Si-

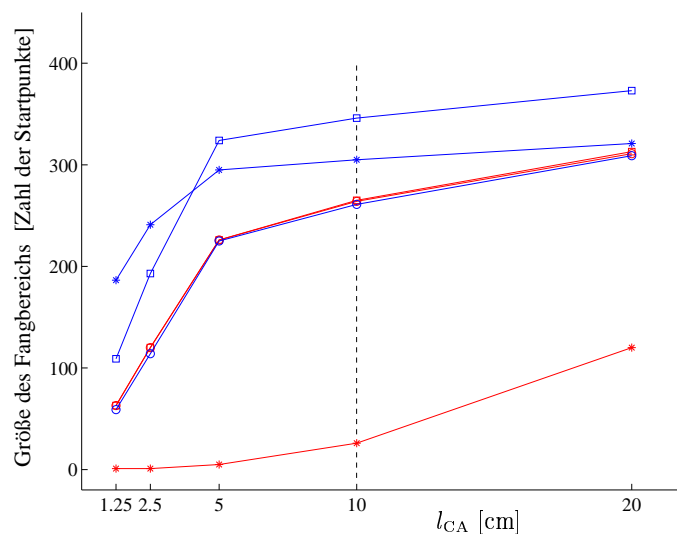


Abbildung 1.15: Mittlere Größe des Fangbereichs in Abhängigkeit von l_{CA} (maximal erlaubter Abstand zur Zielposition) unter beiden Beleuchtungsbedingungen (blau: „konstante Beleuchtung“, rot: „veränderte Beleuchtung“) für verschiedene Ortssignaturen: Bildsignaturen (*, mit Annahme gleicher Abstände \leftrightarrow „SSD“ in Abb. 1.13), Disparitätssignaturen (O, mit gewichteter SSD als Ähnlichkeitsmaß \leftrightarrow „wSSD“ in Abb. 1.13) und deren Kombination (□). (Die Werte für Disparitätssignaturen unter beiden Beleuchtungsbedingungen und für kombinierte Signaturen bei „veränderter Beleuchtung“ liegen dicht beieinander). Die gestrichelte Linie markiert den für die Ergebnisse in Abb. 1.13 verwendeten Wert $l_{CA} = 10$ cm.

mulationen werden die Signaturen am Zielort und die jeweils aktuelle Signatur aus unterschiedlichen Datenbanken entnommen) bleibt die Größe der Fangbereiche für Disparitäts-basiertes Heimfinden praktisch unverändert. Im Gegensatz dazu ist Heimfinden mit Bildsignaturen nicht mehr möglich. Die geringe Größe der Fangbereiche (etwa 550 Zielpositionen besitzen überhaupt keinen Fangbereich, d.h. konnten nicht bis auf $l_{CA} = 10$ cm erreicht werden) ist dadurch zu erklären, dass die berechneten Heimvektoren sehr häufig von der Zielposition weg zeigen.

Die Abhängigkeit der Größe des Fangbereichs von der maximalen erlaubten Abweichung zur Zielposition l_{CA} ist in Abbildung 1.15 dargestellt. Vor allem für kleine l_{CA} schneidet bei „konstanter Beleuchtung“ das Bild-basierte Heimfinden besser ab. Das bedeutet, dass viele Ziele genauer angefahren werden können, was infolge der höheren AHC für kleine Abstände (Abb. 1.12) zu erwarten war.

1.5 Genauigkeit des visuellen Heimfindens

In diesem Abschnitt soll untersucht werden, mit welcher Genauigkeit Heimvektoren berechnet und Zielpositionen durch visuelles Heimfinden mit Bild- bzw. Disparitätssignaturen erreicht werden können. Dies ist u.a. dann von Bedeutung, wenn weitere Messungen beispielsweise durch Wegintegration vorliegen: Durch Gewichtung der unterschiedlichen Werte entsprechend ihrer Sicherheit kann eine verbesserten Schätzung für die aktuelle Position bestimmt werden.

1.5.1 Bild-basiertes Heimfinden

Um die prinzipielle Genauigkeit zu bestimmen, mit der durch Bild-basiertes Heimfinden ein Ort, an dem ein Bild gespeichert wurde, erreicht werden kann, wird ein Bild in der Nachbarschaft des Zielorts betrachtet. Dieses kann eine höhere Ähnlichkeit zu dem gespeicherten Bild aufweisen als das aktuelle Bild am Zielort selbst, falls beispielsweise Beleuchtungsveränderungen oder Bildstörungen auftreten. Die zu erwartende Abweichung vom Zielort wird im Folgenden berechnet.

Für ein Bild an einer um $(\Delta x, \Delta y)$ versetzten Position (in gleicher Orientierung) gilt näherungsweise

$$\bar{I}(\phi|\Delta x, \Delta y) \approx \bar{I}(\phi) + \partial_x \bar{I}(\phi) \Delta x + \partial_y \bar{I}(\phi) \Delta y, \quad (1.55)$$

wobei die Abkürzungen $\partial_x \bar{I}(\phi) := \left. \frac{\partial \bar{I}(\phi|\mathbf{x})}{\partial x} \right|_{\mathbf{x}=\mathbf{x}_0}$ und $\partial_y \bar{I}(\phi) := \left. \frac{\partial \bar{I}(\phi|\mathbf{x})}{\partial y} \right|_{\mathbf{x}=\mathbf{x}_0}$ verwendet wurden (\mathbf{x}_0 bezeichne den Zielort). Die aufgenommenen Bilder besitzen Abweichungen $\Delta I(\phi)$ bzw. $\Delta I'(\phi)$ von den ungestörten Bildern $\bar{I}(\phi)$ bzw. $\bar{I}(\phi|\Delta x, \Delta y)$, d.h. es gilt $I(\phi) = \bar{I}(\phi) + \Delta I(\phi)$ bzw. $I(\phi|\Delta x, \Delta y) = \bar{I}(\phi|\Delta x, \Delta y) + \Delta I'(\phi)$. Es wird dabei angenommen, dass $\Delta I'$ unabhängig von der Verschiebung $(\Delta x, \Delta y)$ ist.

Wird als Ähnlichkeitsmaß die Summe der quadratischen Differenzen (SSD: „Sum of Squared Differences“) betrachtet, so minimiert das Bild mit Positionsverschiebung $(\Delta x, \Delta y)$, das die größte Ähnlichkeit zum gespeicherten Bild am Zielort besitzt, die Funktion

$$E_I(\Delta x, \Delta y) := \sum_{\phi} \left(I(\phi|\Delta x, \Delta y) - I(\phi) \right)^2 \quad (1.56)$$

$$\approx \sum_{\phi} \left(\Delta I'(\phi) - \Delta I(\phi) + \partial_x \bar{I}(\phi) \Delta x + \partial_y \bar{I}(\phi) \Delta y \right)^2. \quad (1.57)$$

Die Lösung von $\nabla E_I(\Delta x, \Delta y) = \mathbf{0}$ lautet:

$$\Delta x = \frac{\sum_{\phi} (\partial_y \bar{I}(\phi))^2 \sum_{\phi} \partial_x \bar{I}(\phi) (\Delta I(\phi) - \Delta I'(\phi))}{NN} - \frac{\sum_{\phi} \partial_x \bar{I}(\phi) \partial_y \bar{I}(\phi) \sum_{\phi} \partial_y \bar{I}(\phi) (\Delta I(\phi) - \Delta I'(\phi))}{NN}, \quad (1.58)$$

$$\Delta y = \frac{\sum_{\phi} (\partial_x \bar{I}(\phi))^2 \sum_{\phi} \partial_y \bar{I}(\phi) (\Delta I(\phi) - \Delta I'(\phi))}{NN} - \frac{\sum_{\phi} \partial_x \bar{I}(\phi) \partial_y \bar{I}(\phi) \sum_{\phi} \partial_x \bar{I}(\phi) (\Delta I(\phi) - \Delta I'(\phi))}{NN}, \quad (1.59)$$

$$NN := \sum_{\phi} (\partial_x \bar{I}(\phi))^2 \sum_{\phi} (\partial_y \bar{I}(\phi))^2 - \left(\sum_{\phi} \partial_x \bar{I}(\phi) \partial_y \bar{I}(\phi) \right)^2. \quad (1.60)$$

Aus Gleichung (1.48) folgt mit $\Delta x = l \cos \varphi$, $\Delta y = l \sin \varphi$ und den Ersetzungen $\phi' \rightarrow \phi$, $r' \rightarrow r$:

$$\bar{I}(\phi | \Delta x, \Delta y) \approx \bar{I}\left(\phi - \frac{\Delta x}{r(\phi)} \sin \phi + \frac{\Delta y}{r(\phi)} \cos \phi\right) \quad (1.61)$$

$$\implies \partial_x \bar{I}(\phi) \approx -\frac{\partial \bar{I}(\phi)}{r(\phi)} \sin \phi, \quad \partial_y \bar{I}(\phi) \approx \frac{\partial \bar{I}(\phi)}{r(\phi)} \cos \phi. \quad (1.62)$$

Nimmt man an, dass die Bildstörungen statistisch unabhängig sind und ihre Erwartungswerte verschwinden, d.h. $\langle \Delta I(\phi) \Delta I(\phi') \rangle = \sigma_N^2 \delta_{\phi, \phi'}$ („weißes Rauschen“), so gilt $\langle \Delta x \rangle = \langle \Delta y \rangle = 0$ und für die Kovarianzmatrix folgt

$$\hat{\Sigma}_I := \begin{pmatrix} \langle \Delta x^2 \rangle & \langle \Delta x \Delta y \rangle \\ \langle \Delta x \Delta y \rangle & \langle \Delta y^2 \rangle \end{pmatrix} = 2\sigma_N^2 \hat{\mathbf{S}}_I^{-1}, \quad (1.63)$$

$$\hat{\mathbf{S}}_I := \begin{pmatrix} \sum_{\phi} (\partial_x \bar{I}(\phi))^2 & \sum_{\phi} \partial_x \bar{I}(\phi) \partial_y \bar{I}(\phi) \\ \sum_{\phi} \partial_x \bar{I}(\phi) \partial_y \bar{I}(\phi) & \sum_{\phi} (\partial_y \bar{I}(\phi))^2 \end{pmatrix}. \quad (1.64)$$

Für die Fläche der „Fehlerellipse“, die durch

$$\mathbf{x}^\top \hat{\Sigma}_I^{-1} \mathbf{x} = 1 \quad (1.65)$$

gegeben ist, erhält man

$$A_I = \pi \sqrt{\det \hat{\Sigma}_I} = \frac{2\sigma_N^2}{\sqrt{\det \hat{\mathbf{S}}_I}} \pi = \frac{2\sigma_N^2}{\sqrt{NN}} \pi = \Delta l_{\text{eq}}^2 \pi, \quad (1.66)$$

wobei $\Delta l_{\text{eq}} := \left(\frac{2\sigma_N^2}{\sqrt{NN}}\right)^{\frac{1}{2}}$ als der Radius eines Kreises gleicher Fläche definiert wurde. Dieser kann als (skalares) Maß für die Ungenauigkeit des bildbasierten Heimfindens verwendet werden. Wegen (1.62) gilt näherungsweise

$$\Delta l_{\text{eq}}^2 \approx \frac{2\sigma_N^2}{\sqrt{\sum_{\phi} \frac{\partial \bar{I}(\phi)^2}{r(\phi)^2} \cos^2 \phi \sum_{\phi} \frac{\partial \bar{I}(\phi)^2}{r(\phi)^2} \sin^2 \phi - \left(\sum_{\phi} \frac{\partial \bar{I}(\phi)^2}{r(\phi)^2} \sin \phi \cos \phi\right)^2}} \quad .(1.67)$$

Neben einem geringen Rauschpegel begünstigen demnach ein hoher Bildkontrast und geringe Abstände zu umgebenden Objekten ein genaues Wiederfinden von Orten. Allerdings können kleine Abstände infolge von Verdeckungen die Größe des Fangbereichs stark limitieren.

Lokale Bildveränderung „liv“

Anhand der Gleichungen (1.58), (1.59) und (1.66) wird deutlich, dass die Genauigkeit des visuellen Heimfindens entscheidend von der Größe NN abhängt. Diese ist unmittelbar mit der in [92] definierten lokalen Bildveränderung „liv“ („local image variation“) verknüpft:

$$\text{liv}(\mathbf{x}) := \sqrt{\left(\frac{\partial \bar{\mathbf{I}}(\mathbf{x})}{\partial x}\right)^2 \left(\frac{\partial \bar{\mathbf{I}}(\mathbf{x})}{\partial y}\right)^2 - \left(\frac{\partial \bar{\mathbf{I}}(\mathbf{x})}{\partial x} \frac{\partial \bar{\mathbf{I}}(\mathbf{x})}{\partial y}\right)^2} \quad (1.68)$$

$$= \sqrt{\det \hat{\mathbf{g}}(\mathbf{x})} \quad , \quad (1.69)$$

$$\hat{\mathbf{g}}(\mathbf{x}) := \begin{pmatrix} \left(\frac{\partial \bar{\mathbf{I}}(\mathbf{x})}{\partial x}\right)^2 & \frac{\partial \bar{\mathbf{I}}(\mathbf{x})}{\partial x} \frac{\partial \bar{\mathbf{I}}(\mathbf{x})}{\partial y} \\ \frac{\partial \bar{\mathbf{I}}(\mathbf{x})}{\partial x} \frac{\partial \bar{\mathbf{I}}(\mathbf{x})}{\partial y} & \left(\frac{\partial \bar{\mathbf{I}}(\mathbf{x})}{\partial y}\right)^2 \end{pmatrix} \quad , \quad (1.70)$$

wobei $\mathbf{x} = (x, y)^{\top} \in \mathbb{R}^2$ eine Parametrisierung der Bildmannigfaltigkeit $\mathbf{I}(\mathbf{x})$ ist. Die Größe $\hat{\mathbf{g}}$ wird in der Differentialgeometrie als „metrischer Tensor“ bezeichnet [11].

Mit $\bar{\mathbf{I}}(\mathbf{x}) = (\bar{I}(\phi_0|\mathbf{x}), \bar{I}(\phi_1|\mathbf{x}), \dots, \bar{I}(\phi_{N-1}|\mathbf{x}))^{\top}$ ist Gleichung (1.68) äquivalent zu

$$\text{liv}(\mathbf{x}) = \sqrt{\sum_{\phi} \left(\frac{\partial \bar{I}(\phi|\mathbf{x})}{\partial x}\right)^2 \sum_{\phi} \left(\frac{\partial \bar{I}(\phi|\mathbf{x})}{\partial y}\right)^2 - \left(\sum_{\phi} \frac{\partial \bar{I}(\phi|\mathbf{x})}{\partial x} \frac{\partial \bar{I}(\phi|\mathbf{x})}{\partial y}\right)^2} \quad .$$

Somit gilt

$$\text{liv}(\mathbf{x})^2 = NN(\mathbf{x}) \quad . \quad (1.71)$$

Um eine möglichst genaue Lokalisierung zu erreichen, ist es deshalb sinnvoll, Bilder vor allem an Orten mit hohen liv-Werten aufzunehmen.

Mit Hilfe der Eigenwerte von $\hat{\mathbf{g}}$, $\lambda_{g,1}$ und $\lambda_{g,2}$, kann eine alternative Form der lokalen Bildveränderung angegeben werden:

$$\text{liv}(\mathbf{x}) = \sqrt{\lambda_{g,1} \lambda_{g,2}} \quad (1.72)$$

$$\begin{aligned} &= \frac{1}{2} \left(\left[\sum_{\phi} \left(\frac{\partial \bar{I}(\phi|\mathbf{x})}{\partial x} \right)^2 + \left(\frac{\partial \bar{I}(\phi|\mathbf{x})}{\partial y} \right)^2 \right]^2 - \left[\sum_{\phi} \left(\frac{\partial \bar{I}(\phi|\mathbf{x})}{\partial x} \right)^2 \right. \right. \\ &\quad \left. \left. - \left(\frac{\partial \bar{I}(\phi|\mathbf{x})}{\partial y} \right)^2 \right]^2 - \left[2 \sum_{\phi} \frac{\partial \bar{I}(\phi|\mathbf{x})}{\partial x} \frac{\partial \bar{I}(\phi|\mathbf{x})}{\partial y} \right]^2 \right)^{\frac{1}{2}} . \quad (1.73) \end{aligned}$$

Da häufig der erste Term dominiert, gilt näherungsweise

$$\text{liv}(\mathbf{x}) \lesssim \frac{1}{2} \sum_{\phi} \left(\frac{\partial \bar{I}(\phi|\mathbf{x})}{\partial x} \right)^2 + \left(\frac{\partial \bar{I}(\phi|\mathbf{x})}{\partial y} \right)^2 . \quad (1.74)$$

Nach Einsetzen von (1.62) folgt aus (1.73) mit Hilfe der Identitäten $\cos^2\phi - \sin^2\phi = \cos(2\phi)$ und $2 \sin\phi \cos\phi = \sin(2\phi)$:

$$\begin{aligned} \text{liv}(\mathbf{x}) &\approx \frac{1}{2} \left(\left(\sum_{\phi} \frac{\partial \bar{I}(\phi|\mathbf{x})^2}{r(\phi|\mathbf{x})^2} \right)^2 - \left[\left(\sum_{\phi} \frac{\partial \bar{I}(\phi|\mathbf{x})^2}{r(\phi|\mathbf{x})^2} \cos(2\phi) \right)^2 \right. \right. \\ &\quad \left. \left. + \left(\sum_{\phi} \frac{\partial \bar{I}(\phi|\mathbf{x})^2}{r(\phi|\mathbf{x})^2} \sin(2\phi) \right)^2 \right] \right)^{\frac{1}{2}} \quad (1.75) \end{aligned}$$

$$= \frac{N}{2} \sqrt{|\mathcal{C}_0(\mathbf{x})|^2 - |\mathcal{C}_2(\mathbf{x})|^2} , \quad (1.76)$$

wobei die Fourierkoeffizienten $\mathcal{C}_k := \frac{1}{N} \sum_{\phi} \frac{\partial \bar{I}(\phi|\mathbf{x})^2}{r(\phi|\mathbf{x})^2} e^{-ik\phi}$ substituiert wurden. Für $|\mathcal{C}_2(\mathbf{x})| \ll |\mathcal{C}_0(\mathbf{x})|$ gilt die Näherung (1.74) und Gleichung (1.67) vereinfacht sich zu

$$\Delta l_{\text{eq}}^2(\mathbf{x}) = \frac{2\sigma_N^2}{\text{liv}(\mathbf{x})} \approx \frac{4\sigma_N^2}{\sum_{\phi} \frac{\partial \bar{I}(\phi|\mathbf{x})^2}{r(\phi|\mathbf{x})^2}} . \quad (1.77)$$

Um liv-Werte näherungsweise aus Bildern, die an benachbarten Orten aufgenommen werden, zu berechnen, wird zunächst ein Flächenelement im Bildraum betrachtet, das einem kleinen Bereich $D \subset \mathbb{R}^2$ im Ortsraum entspricht. Es gilt

$$S(D) = \iint_D \text{liv}(\mathbf{x}) dx dy \approx \text{liv}(\mathbf{x}_S) \iint_D dx dy = \text{liv}(\mathbf{x}_S) D \quad (1.78)$$

$$\implies \text{liv}(\mathbf{x}_S) \approx \frac{S(D)}{D} , \quad (1.79)$$

wobei \mathbf{x}_S den Schwerpunkt von D bezeichnet. Drei Bildvektoren $\mathbf{I}(\mathbf{x}_1)$, $\mathbf{I}(\mathbf{x}_2)$, $\mathbf{I}(\mathbf{x}_3)$ an benachbarten Orten definieren ein Dreieck im Bildraum mit der Fläche

$$S_\Delta = \frac{1}{2} |\Delta \mathbf{I}_{12}| |\Delta \mathbf{I}_{13}| \sin \angle(\Delta \mathbf{I}_{12}, \Delta \mathbf{I}_{13}) \quad (1.80)$$

$$= \frac{1}{2} \sqrt{\Delta \mathbf{I}_{12}^2 \Delta \mathbf{I}_{13}^2 - (\Delta \mathbf{I}_{12} \Delta \mathbf{I}_{13})^2} \quad (1.81)$$

$$\Delta \mathbf{I}_{1a} := \mathbf{I}(\mathbf{x}_a) - \mathbf{I}(\mathbf{x}_1), \quad a = 2, 3 \quad (1.82)$$

Die entsprechende Fläche in \mathbb{R}^2 ist

$$D_\Delta = \frac{1}{2} |\Delta \mathbf{x}_{12}| |\Delta \mathbf{x}_{13}| \sin \angle(\Delta \mathbf{x}_{12}, \Delta \mathbf{x}_{13}) \quad (1.83)$$

$$= \frac{1}{2} \sqrt{\Delta \mathbf{x}_{12}^2 \Delta \mathbf{x}_{13}^2 - (\Delta \mathbf{x}_{12} \Delta \mathbf{x}_{13})^2} \quad (1.84)$$

$$\Delta \mathbf{x}_{1a} := \mathbf{x}_a - \mathbf{x}_1, \quad a = 2, 3 \quad (1.85)$$

Folglich gilt näherungsweise

$$\text{liv}(\mathbf{x}_S) \approx \frac{S_\Delta}{D_\Delta} = \sqrt{\frac{\Delta \mathbf{I}_{12}^2 \Delta \mathbf{I}_{13}^2 - (\Delta \mathbf{I}_{12} \Delta \mathbf{I}_{13})^2}{\Delta \mathbf{x}_{12}^2 \Delta \mathbf{x}_{13}^2 - (\Delta \mathbf{x}_{12} \Delta \mathbf{x}_{13})^2}} \quad (1.86)$$

$$\mathbf{x}_S = \frac{1}{3} (\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3) \quad (1.87)$$

Abbildung 1.16 zeigt die lokale Bildveränderung innerhalb der Modellhausarena, berechnet mit Hilfe der Bilddatenbank (bestehend aus 1250 ein-dimensionalen Panoramabildern). Wie erwartet, treten hohe liv-Werte vor allem in der Nähe von Modellhäusern mit kontrastreicher Textur auf.

Ein Beispiel für die Abhängigkeit der Genauigkeit des Bild-basierten Heimfindens vom liv-Wert zeigt Abbildung 1.17. Hierfür wurden zwei Bilddatensätze der Roboterarena verwendet, die Bilder enthalten, die an den gleichen Positionen, jedoch in zwei unterschiedlichen Orientierungen ($\approx 0^\circ$, $\approx 90^\circ$) aufgenommen wurden. Für jedes Bild des 0° -Datensatzes wurde der Positionsunterschied zu dem Bild des 90° -Datensatzes ausgewertet, das die größte Ähnlichkeit aufweist. Bei der Bestimmung der Bildähnlichkeit ist neben Kamerarauschen, begrenzter Auflösung ($N = 72$ Pixel) und eventuell auftretenden leichten Beleuchtungsveränderungen auch die Verdeckung durch den Kabelstrang zum externen Rechner von Bedeutung. Da die Aufnahmepositionen sich auf dem in Abschnitt 1.4.3 beschriebenen Gitter befinden, für das der minimale Positionsabstand 2.5 cm beträgt, besitzen die meisten Bilder der beiden Datensätze die größte Ähnlichkeit an der gleichen Position, d.h. für $l = 0$ cm. Bei kleinen liv-Werten kommt es jedoch häufiger als bei hohen liv-Werten vor, dass sich das ähnlichste Bild an einer benachbarten Position befindet. Das Heimfinden wird folglich öfters in größerer Distanz zum Ziel abgebrochen und ist somit ungenauer.

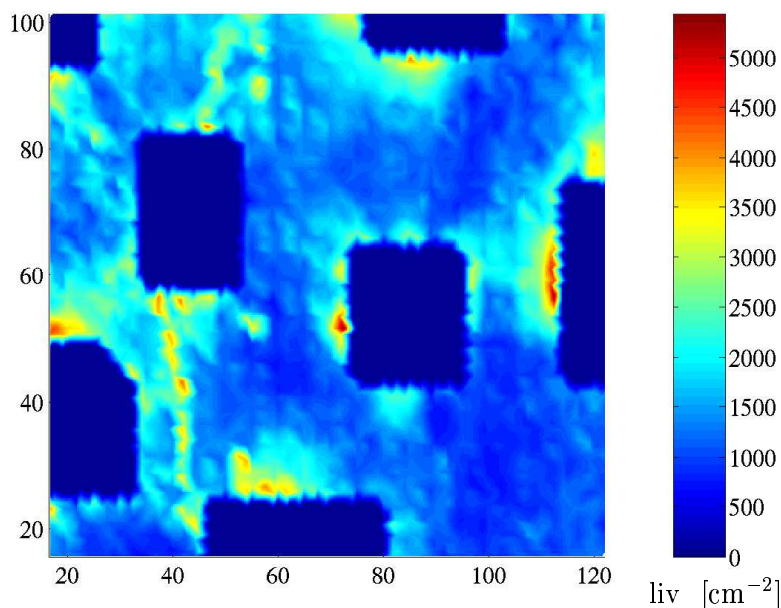


Abbildung 1.16: Lokale Bildveränderung berechnet für die Bilddatenbank der Modellhausarena. Hohe liv-Werte (siehe Farbtafel) treten hauptsächlich in der Nähe von Modellhäusern mit deutlicher Struktur an den Wänden auf (die Positionen der Modellhäuser sind durch kontinuierliche dunkle Bereiche gekennzeichnet).

1.5.2 Fehler durch „Equal Distance Assumption“

Im Folgenden wird untersucht, welche Fehler bei der Bestimmung des Heimvektors durch das in [38] beschriebene und in Abschnitt 1.4 verwendete Verfahren zu erwarten sind. Die Schätzung der Richtung φ und der relativen Orientierung ψ zum Zielort erfolgt durch Vergleich von „erwarteten“ Bildern $I^e(\phi|\lambda, \varphi, \psi)$, die unter der Annahme konstanter Abstände berechnet werden, und dem gespeicherten Bild an der Zielposition $I^h(\phi)$. Als Ähnlichkeitsmaß wird die Summe der quadratischen Differenzen verwendet, d.h.

$$E_1(\lambda, \varphi, \psi) := \sum_{\phi''} (I^h(\phi'') - I^e(\phi''|\lambda, \varphi, \psi))^2 . \quad (1.88)$$

Da $I^e(\phi''|\lambda, \varphi, \psi)$ nicht-linear von den Parametern abhängt, muss (1.88) normalerweise numerisch minimiert werden. In Abschnitt 1.4 erfolgte die Minimierung durch diskrete Abtastung des Suchraums, wobei der Bereich für λ auf $0 \leq \lambda \leq \frac{2}{3}$ beschränkt ist. Für (im Vergleich zu den Objektdistanzen) kleine Abstände \bar{l} von der Zielposition ist jedoch eine analytische Abschätzung für Positions- und Orientierungsfehler möglich, die infolge der „Equal Di-

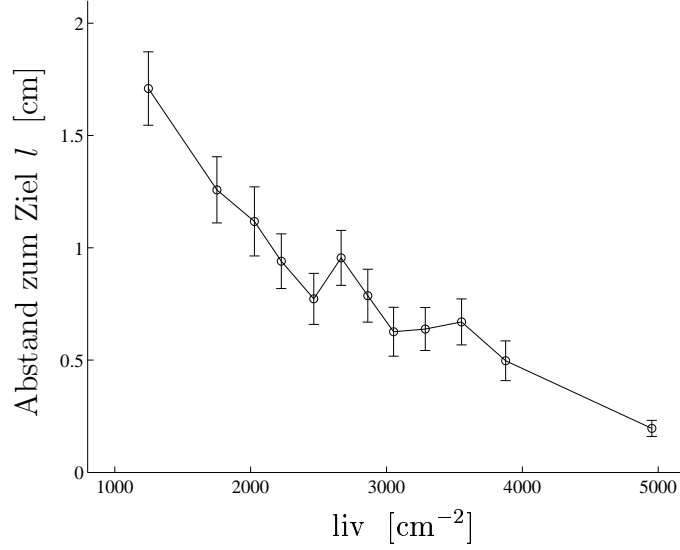


Abbildung 1.17: Abhängigkeit der „Homing“-Genauigkeit vom liv-Wert. Die Bins wurden so gewählt, dass sie jeweils $N_{\text{bin}} = 100$ Daten enthalten. Kreise kennzeichnen die Mittelwerte $\langle l_{\text{bin}} \rangle$, Fehlerbalken geben den Standardfehler des Mittelwertes an, d.h. $\sqrt{\frac{1}{N_{\text{bin}}(N_{\text{bin}}-1)} \sum_i^{N_{\text{bin}}} (l_{\text{bin},i} - \langle l_{\text{bin}} \rangle)^2}$.

stance Assumption“ und zusätzlicher Störungen (z.B. Sensorrauschen) zu erwarten sind. Abbildung 1.18 illustriert die in der folgenden Berechnung verwendeten Symbole.

Sind die exakten Werte für die Entfernung, Richtung und relative Orientierung zum Zielort durch \bar{l} , $\bar{\varphi}$ und $\bar{\psi}$ gegeben, so kann für $\bar{l} \ll r'(\phi')$ mit Gleichung (1.48) und der Substitution $\phi'' = \phi' - \bar{\psi}$ das Bild an der Zielposition aus dem Bild an der aktuellen Position näherungsweise bestimmt werden. Es gilt

$$\bar{I}^h(\phi'') \approx \bar{I}(\phi'' + \bar{\psi} - \frac{\bar{l}}{r'(\phi'' + \bar{\psi})} \sin(\phi'' + \bar{\psi} - \bar{\varphi})) , \quad (1.89)$$

wobei \bar{I}^h und \bar{I} die ungestörten Bilder bezeichnen. Das gespeicherte Bild I^h besitzt z.B. infolge beschränkter Auflösung und Sensorrauschen Abweichungen von \bar{I}^h :

$$I^h(\phi'') = \bar{I}^h(\phi'') + \Delta I^h(\phi'') . \quad (1.90)$$

Für die mit Hilfe der „Equal Distance Assumption“ aus dem aktuellen Bild $I(\phi)$, das ebenfalls Abweichungen $\Delta I(\phi)$ vom exakten Bild $\bar{I}(\phi)$ besitzt, berechneten Bilder erhält man aus (1.44) mit $\lambda := l/R$ und $\phi'' = \phi' - \bar{\psi}$:

$$I^e(\phi''|\lambda, \varphi, \psi) \approx I(\phi'' + \psi - \lambda \sin(\phi'' + \psi - \varphi)) . \quad (1.91)$$

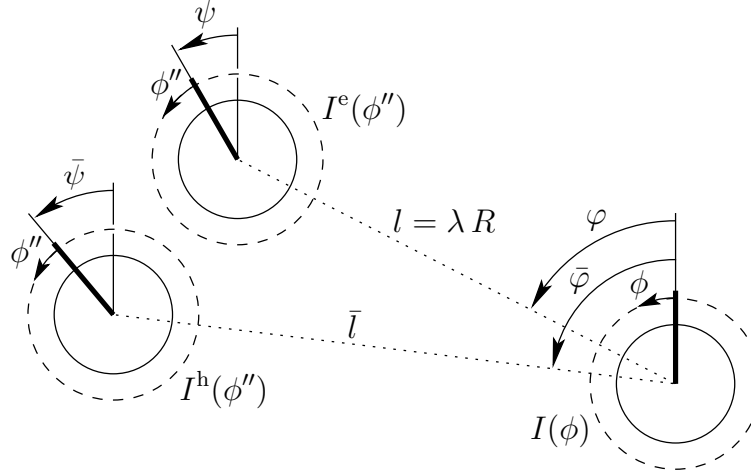


Abbildung 1.18: Überblick über die verwendeten Parameter: Bezüglich des aktuellen Bildes $I(\phi)$ befindet sich die Zielposition in Richtung $\bar{\varphi}$ mit Abstand \bar{l} . Das zugehörige Bild $I^h(\phi'')$ hat die Relativorientierung $\bar{\psi}$ (dicke Linien kennzeichnen die Orientierung der Bilder). Das unter der Annahme konstanter Abstände berechnete Bild $I^e(\phi'')$ besitzt eine davon abweichende Position ($\varphi, l = R\lambda$) und Orientierung ψ .

Mit den Definitionen

$$D(\phi) := \frac{1}{r'(\phi)} \quad , \quad \Delta\varphi := \varphi - \bar{\varphi} \quad , \quad \Delta\psi := \psi - \bar{\psi} \quad (1.92)$$

folgt für $\bar{l}D(\phi), \lambda, \Delta\psi, \Delta\varphi \ll 1$,

$$\begin{aligned} I^h(\phi'') &= \bar{I}^h(\phi'') + \Delta I^h(\phi'') \\ &\approx \bar{I}(\phi'' + \bar{\psi} - \bar{l}D(\phi'' + \bar{\psi}) \sin(\phi'' + \bar{\psi} - \bar{\varphi})) + \Delta I^h(\phi'') \\ &\approx \bar{I}(\phi'' + \bar{\psi}) + \Delta I^h(\phi'') \\ &\quad - \partial \bar{I}(\phi'' + \bar{\psi}) \bar{l}D(\phi'' + \bar{\psi}) \sin(\phi'' + \bar{\psi} - \bar{\varphi}) \quad , \end{aligned} \quad (1.93)$$

$$\begin{aligned} I^e(\phi'') &= I(\phi'' + \bar{\psi} + \Delta\psi - \lambda \sin(\phi'' + \bar{\psi} - \bar{\varphi} + \Delta\psi - \Delta\varphi)) \\ &\approx \bar{I}(\phi'' + \bar{\psi}) + \Delta I(\phi'' + \bar{\psi}) + \partial \bar{I}(\phi'' + \bar{\psi}) \\ &\quad \times (\Delta\psi - \lambda \sin(\phi'' + \bar{\psi} - \bar{\varphi} + \Delta\psi - \Delta\varphi)) \\ &\approx \bar{I}(\phi'' + \bar{\psi}) + \partial \bar{I}(\phi'' + \bar{\psi}) [\Delta\psi - \lambda \sin(\phi'' + \bar{\psi} - \bar{\varphi}) \\ &\quad - \lambda (\Delta\psi - \Delta\varphi) \cos(\phi'' + \bar{\psi} - \bar{\varphi})] + \Delta I(\phi'' + \bar{\psi}) \\ &\approx \bar{I}(\phi'' + \bar{\psi}) + \partial \bar{I}(\phi'' + \bar{\psi}) [\Delta\psi - \lambda \sin(\phi'' + \bar{\psi} - \bar{\varphi}) \\ &\quad + \Delta\varphi \lambda \cos(\phi'' + \bar{\psi} - \bar{\varphi})] + \Delta I(\phi'' + \bar{\psi}) \quad . \end{aligned} \quad (1.94)$$

Durch Einsetzen von (1.93) und (1.94) in (1.88) erhält man mit $\mu := \lambda\Delta\varphi$ und $\phi'' + \bar{\psi} \rightarrow \phi$ die genäherte Fehlerfunktion

$$\begin{aligned} \tilde{E}_I(\lambda, \mu, \Delta\psi) &\approx \sum_{\phi} \left(\Delta I^h(\phi - \bar{\psi}) - \Delta I(\phi) - \partial \bar{I}(\phi) \bar{l} D(\phi) \sin(\phi - \bar{\varphi}) \right. \\ &\quad \left. + \partial \bar{I}(\phi) [\lambda \sin(\phi - \bar{\varphi}) - \Delta\psi - \mu \cos(\phi - \bar{\varphi})] \right)^2. \end{aligned} \quad (1.95)$$

Minimieren von (1.95), d.h. Lösen der Gleichung $\nabla \tilde{E}_I(\lambda, \mu, \Delta\psi) = \mathbf{0}$, führt zu einem linearen Gleichungssystem in λ , μ und $\Delta\psi$. Um das Ergebnis zu vereinfachen wird angenommen, dass

$$\sum_{\phi} A(\phi) \{\sin\phi, \cos\phi, \sin\phi \cos\phi\} \ll \sum_{\phi} A(\phi) \{\cos^2\phi, \sin^2\phi\} \quad (1.96)$$

gilt, wobei $A(\phi)$ für $\partial \bar{I}(\phi)^2$ bzw. $\partial \bar{I}(\phi)^2 D(\phi)$ steht. Für eine Begründung dieser Annahme siehe Anhang A.4. Es folgt dann

$$\lambda \approx \langle \lambda \rangle - \frac{\sum_{\phi} \partial I(\phi) [\Delta I^h(\phi - \bar{\psi}) - \Delta I(\phi)] \sin(\phi - \bar{\varphi})}{\sum_{\phi} \partial I(\phi)^2 \sin^2(\phi - \bar{\varphi})}, \quad (1.97)$$

$$\mu \approx \langle \mu \rangle + \frac{\sum_{\phi} \partial I(\phi) [\Delta I^h(\phi - \bar{\psi}) - \Delta I(\phi)] \cos(\phi - \bar{\varphi})}{\sum_{\phi} \partial I(\phi)^2 \cos^2(\phi - \bar{\varphi})}, \quad (1.98)$$

$$\Delta\psi \approx \langle \Delta\psi \rangle + \frac{\sum_{\phi} \partial I(\phi) [\Delta I^h(\phi - \bar{\psi}) - \Delta I(\phi)]}{\sum_{\phi} \partial I(\phi)^2}, \quad (1.99)$$

$$\Delta\varphi = \frac{\mu}{\lambda} \approx \frac{\langle \mu \rangle}{\langle \lambda \rangle}. \quad (1.100)$$

Für die Erwartungswerte gilt (sofern die Mittelwerte von $\Delta I^h(\phi)$ und $\Delta I(\phi)$ vernachlässigbar sind):

$$\langle \lambda \rangle \approx \bar{l} \langle D \rangle, \quad \langle D \rangle := \frac{\sum_{\phi} \partial \bar{I}(\phi)^2 D(\phi) \sin^2(\phi - \bar{\varphi})}{\sum_{\phi} \partial \bar{I}(\phi)^2 \sin^2(\phi - \bar{\varphi})}, \quad (1.101)$$

$$\langle \mu \rangle \approx \bar{l} \frac{\sum_{\phi} \partial \bar{I}(\phi)^2 [\langle D \rangle - D(\phi)] \sin(\phi - \bar{\varphi}) \cos(\phi - \bar{\varphi})}{\sum_{\phi} \partial \bar{I}(\phi)^2 \cos^2(\phi - \bar{\varphi})}, \quad (1.102)$$

$$\langle \Delta\psi \rangle \approx \bar{l} \frac{\sum_{\phi} \partial \bar{I}(\phi)^2 [\langle D \rangle - D(\phi)] \sin(\phi - \bar{\varphi})}{\sum_{\phi} \partial \bar{I}(\phi)^2}, \quad (1.103)$$

$$\langle \Delta\varphi \rangle \approx \frac{\langle \mu \rangle}{\langle \lambda \rangle} \approx \frac{\sum_{\phi} \partial \bar{I}(\phi)^2 [\langle D \rangle - D(\phi)] \sin(\phi - \bar{\varphi}) \cos(\phi - \bar{\varphi})}{\langle D \rangle \sum_{\phi} \partial \bar{I}(\phi)^2 \cos^2(\phi - \bar{\varphi})}. \quad (1.104)$$

Falls die Fehler ΔI^h und ΔI unabhängig in jedem Bildpixel auftreten und um

Mittelwert Null mit Standardabweichung σ_N verteilt sind, folgt aus (1.97)–(1.104) für die Varianzen

$$\text{Var}[\lambda] \approx \frac{2\sigma_N^2}{\sum_{\phi} \partial \bar{I}(\phi)^2 \sin^2(\phi - \bar{\varphi})} , \quad (1.105)$$

$$\text{Var}[\Delta\varphi] \approx \frac{2\sigma_N^2}{\bar{l}^2 \langle D \rangle^2 \sum_{\phi} \partial \bar{I}(\phi)^2 \cos^2(\phi - \bar{\varphi})} , \quad (1.106)$$

$$\text{Var}[\Delta\psi] \approx \frac{2\sigma_N^2}{\sum_{\phi} \partial \bar{I}(\phi)^2} . \quad (1.107)$$

$\langle \Delta\psi \rangle$, $\langle \Delta\varphi \rangle$ stellen die systematischen Fehler dar, die infolge von Abweichungen der tatsächlichen Abstände von einem konstanten Wert auftreten. Für $r'(\phi) \rightarrow R = \text{const}$, erhält man $D(\phi) \rightarrow \langle D \rangle \rightarrow 1/R$, $\langle \lambda \rangle \rightarrow \bar{l}/R$ und somit $\langle \Delta\psi \rangle$, $\langle \Delta\varphi \rangle \rightarrow 0$.

Aus (1.97) und (1.103) folgt, dass – auch in Fällen, in denen die Annahme gleicher Abstände nur eine grobe Näherung darstellt – mit abnehmender Distanz zum Zielort sowohl λ als auch der Fehler in den Orientierungsschätzung $\Delta\psi$ kleiner werden und ihre Erwartungswerte am Ziel verschwinden ($\langle \lambda \rangle$, $\langle \Delta\psi \rangle \rightarrow 0$ für $\bar{l} \rightarrow 0$). $\langle \Delta\varphi \rangle$ ist in erster Näherung von der Distanz zum Zielort \bar{l} unabhängig, für sehr kleine Abstände gilt jedoch $\Delta\varphi \propto 1/\bar{l}$. Dies hat u.a. zur Folge, dass die „Average Homeward Component“ für $\bar{l} \rightarrow 0$ stark abnimmt.

Bemerkungen: Die Varianz des Bildrauschens beträgt (bei üblicher Zimmerbeleuchtung) für die Kamera des Kheperas näherungsweise $\sigma_N^2 \approx 1$. Dies ist ein kleinerer Wert als in 1.2.1 angegeben ($\sigma_n^2 \approx 9$), da – wie in Abbildung 1.5 a dargestellt – für jeden der 72 Grauwerte der Bildsignatur über mehr Pixel gemittelt wird als für die 72×100 Grauwerte zur Bestimmung der Disparitäten (siehe Abschnitt 1.2).

Aus (1.105) folgt, dass das Heimfinden spätestens dann abgebrochen werden kann, wenn λ^2 in die Größenordnung von $\frac{\sigma_N^2}{\sum \partial I(\phi)^2}$ kommt. Dies bedeutet gleichzeitig, dass Ziele in einer kontrastreichen Umgebung, d.h. $\sum_{\phi} \partial I(\phi)^2$ groß, genauer angefahren werden können (bei vergleichbaren Abständen). Ist ein typischer Landmarkenabstand R_{typ} bekannt, so kann mit $\sqrt{\text{Var}[\lambda]} R_{\text{typ}}$ eine Abschätzung für die Genauigkeit, mit der eine Zielposition erreicht werden kann, angegeben werden. Dies folgt auch direkt aus (1.77) mit $r(\phi) \approx R_{\text{typ}}$ bzw. $\sum_{\phi} \frac{\partial I(\phi|\mathbf{x})^2}{r(\phi|\mathbf{x})^2} \approx R_{\text{typ}}^{-2} \sum_{\phi} \partial I(\phi|\mathbf{x})^2$.

Für nahe Ziele sind im Allgemeinen die Fehler $\Delta\varphi$ und $\Delta\psi$ klein, so dass aus dem Minimieren von (1.88) sinnvolle Schätzungen für $\bar{\varphi}$ und $\bar{\psi}$ gewonnen werden können, ohne dass die Position angefahren werden muss. Eine

zusätzliche Schätzung für den Abstand erhält man aus λ , wenn ein typischer Landmarkenabstand bekannt ist, d.h. $l = R_{\text{typ}}\lambda$. In [51] wurden Landmarkenabstände durch Vergleich mit Odometriewerten ermittelt, $R_{\text{typ}} = l_{\text{odo}}/\lambda$. Anschließend konnten dann Winkel und Distanzen zu benachbarten Orten, an denen Bilder gespeichert wurden, bestimmt werden, ohne diese anzufahren. Unter Verwendung einer Ausgleichsrechnung [50], die jedem Ort eine eindeutige 2D-Koordinate zuweist, konnte mit diesem Verfahren im Vergleich zu reiner Odometriemessung eine deutlich verbesserte Repräsentation der Umgebung erreicht werden.

1.5.3 Disparitäts-basiertes Heimfinden

Die Berechnung der Entfernung l , der Richtung φ und der relativen Orientierung ψ zum Zielort erfolgt für das Disparitäts-basierte Heimfinden nach Abschnitt 1.3 durch Minimieren der Fehlerfunktion

$$E_d(l, \varphi, \psi) := \frac{\sum_{\phi''} w(\phi''|l, \varphi, \psi) (d^h(\phi'') - d^e(\phi''|l, \varphi, \psi))^2}{\sum_{\phi''} w(\phi''|l, \varphi, \psi)}, \quad (1.108)$$

$$w(\phi''|l, \varphi, \psi) := (v^h(\phi'') + v^e(\phi''|l, \varphi, \psi))^{-1}. \quad (1.109)$$

Um die Genauigkeit, mit der ein Zielort anhand von Disparitätssignaturen erreicht werden kann, abschätzen zu können, wird im Folgenden angenommen, dass $l \ll r(\phi) \approx \alpha/d(\phi)$ gilt.

Aus Gleichung (1.27) folgt

$$r' = \sqrt{r^2 - l^2 \sin^2(\phi' - \varphi)} - l \cos(\phi' - \varphi). \quad (1.110)$$

Mit $r = \alpha/d - r_0$, Gleichung (1.25), erhält man aus (1.110),

$$\begin{aligned} d' &= \frac{d(\phi)}{\sqrt{(1 - \frac{r_0 d(\phi)}{\alpha})^2 - (\frac{ld(\phi)}{\alpha})^2 \sin^2(\phi' - \varphi) - \frac{ld(\phi)}{\alpha} \cos(\phi' - \varphi) + \frac{r_0 d(\phi)}{\alpha}}} \\ &\approx d(\phi) + \frac{1}{\alpha} ld(\phi)^2 \cos(\phi' - \varphi), \end{aligned} \quad (1.111)$$

wobei $\frac{1}{\alpha} ld \ll 1$ angenommen wurde. Aus der Näherung (1.48) folgt

$$\phi(\phi') \approx \phi' - \frac{1}{\alpha} ld'(\phi') \sin(\phi' - \varphi). \quad (1.112)$$

Eingesetzt in (1.111) erhält man

$$\begin{aligned} d'(\phi'|l, \varphi, \psi) &\approx d[\phi' - \frac{1}{\alpha} ld'(\phi') \sin(\phi' - \varphi)] + \frac{1}{\alpha} ld(\phi')^2 \cos(\phi' - \varphi) \\ &\approx d[\phi' - \frac{1}{\alpha} ld(\phi') \sin(\phi' - \varphi)] + \frac{1}{\alpha} ld(\phi')^2 \cos(\phi' - \varphi). \end{aligned} \quad (1.113)$$

Nach einer zusätzlichen Änderung der Orientierung, d.h. Rotation um ψ ($\phi'' := \phi' - \psi$), folgt

$$\begin{aligned} d''(\phi''|l, \varphi, \psi) &\approx d[\phi'' + \psi - \frac{1}{\alpha}ld(\phi'' + \psi) \sin(\phi'' + \psi - \varphi)] \\ &\quad + \frac{1}{\alpha}ld(\phi'' + \psi)^2 \cos(\phi'' + \psi - \varphi) . \end{aligned} \quad (1.114)$$

Im Gegensatz zur Bildsignatur, siehe z.B. Gleichung (1.89), ändert sich für Disparitätssignaturen nicht nur die Position ($\phi \rightarrow \phi''$) des Signaturwerts, sondern auch der Wert selbst.

Die Disparitätssignatur am Zielort kann mit (1.114) näherungsweise aus der aktuellen Signatur bestimmt werden ($\bar{d}(\phi)$ bezeichnet die „wahren“ Disparitäten),

$$\begin{aligned} \bar{d}^h(\phi'') &\approx \bar{d}[\phi'' + \bar{\psi} - \frac{1}{\alpha}\bar{l}\bar{d}(\phi'' + \bar{\psi}) \sin(\phi'' + \bar{\psi} - \bar{\varphi})] \\ &\quad + \frac{1}{\alpha}\bar{l}\bar{d}(\phi'' + \bar{\psi})^2 \cos(\phi'' + \bar{\psi} - \bar{\varphi}) . \end{aligned} \quad (1.115)$$

Die gespeicherte Signatur am Zielort weist jedoch Fehler auf, z.B. infolge von Bildrauschen, d.h.

$$d^h(\phi'') = \bar{d}^h(\phi'') + \Delta d^h(\phi'') . \quad (1.116)$$

Die Minimierung von E_d liefert Schätzungen, die gegenüber den exakten Werten \bar{l} , $\bar{\varphi}$, $\bar{\psi}$ die Abweichungen $\Delta l := l - \bar{l}$, $\Delta \varphi := \varphi - \bar{\varphi}$, $\Delta \psi := \psi - \bar{\psi}$ aufweisen. Diese lassen sich mit Hilfe der Fehlerfunktion

$$\begin{aligned} &\tilde{E}_d(\Delta l, \Delta \varphi, \Delta \psi) \\ &:= E_d(\bar{l} + \Delta l, \bar{\varphi} + \Delta \varphi, \bar{\psi} + \Delta \psi) \\ &\approx \frac{\sum_{\phi''} w(\phi''|\bar{l}, \bar{\varphi}, \bar{\psi})(d^h(\phi'') - d^e(\phi''|\bar{l} + \Delta l, \bar{\varphi} + \Delta \varphi, \bar{\psi} + \Delta \psi))^2}{\sum_{\phi''} w(\phi''|\bar{l}, \bar{\varphi}, \bar{\psi})} \end{aligned} \quad (1.117)$$

näherungsweise ermitteln. Die erwarteten Signaturen werden auf Basis der gemessenen Disparitäten berechnet:

$$\begin{aligned} &d^e(\phi''|\bar{l} + \Delta l, \bar{\varphi} + \Delta \varphi, \bar{\psi} + \Delta \psi) \\ &\approx \bar{d}''(\phi''|\bar{l} + \Delta l, \bar{\varphi} + \Delta \varphi, \bar{\psi} + \Delta \psi) \\ &\quad + \Delta d[\phi'' + \bar{\psi} - \frac{1}{\alpha}\bar{l}\bar{d}(\phi'' + \bar{\psi}) \sin(\phi'' + \bar{\psi} - \bar{\varphi})] . \end{aligned}$$

Somit folgt für die Differenz der erwarteten und der gespeicherten Signatur:

$$\begin{aligned} &d^h(\phi'') - d^e(\phi''|\bar{l} + \Delta l, \bar{\varphi} + \Delta \varphi, \bar{\psi} + \Delta \psi) \\ &\approx \Delta d^h(\phi'') - \Delta d[\phi'' + \bar{\psi} - \frac{1}{\alpha}\bar{l}\bar{d}(\phi'' + \bar{\psi}) \sin(\phi'' + \bar{\psi} - \bar{\varphi})] \\ &\quad - \Delta \psi \partial \bar{d}(\phi'' + \bar{\psi}) - \Delta \varphi \frac{1}{\alpha} \bar{l} \bar{d}(\phi'' + \bar{\psi}) \left(\partial \bar{d}(\phi'' + \bar{\psi}) \cos(\phi'' + \bar{\psi} - \bar{\varphi}) \right. \\ &\quad \left. + \bar{d}(\phi'' + \bar{\psi}) \sin(\phi'' + \bar{\psi} - \bar{\varphi}) \right) + \frac{1}{\alpha} \Delta l \bar{d}(\phi'' + \bar{\psi}) \\ &\quad \times \left(\partial \bar{d}(\phi'' + \bar{\psi}) \sin(\phi'' + \bar{\psi} - \bar{\varphi}) - \bar{d}(\phi'' + \bar{\psi}) \cos(\phi'' + \bar{\psi} - \bar{\varphi}) \right) . \end{aligned}$$

Da $w(\phi''|\bar{l}, \bar{\varphi}, \bar{\psi})$ nicht von Δl , $\Delta\varphi$ oder $\Delta\psi$ abhängt, kann bei der Optimierung von (1.117) der Nenner unberücksichtigt bleiben. Mit weiteren Näherungen analog zu (1.96) erhält man schließlich

$$\begin{aligned}\Delta l &\approx \frac{-\alpha \sum_{\phi} \tilde{w}(\phi) \Delta \tilde{d}(\phi) \bar{d}(\phi) [\partial \bar{d}(\phi) \sin(\phi - \bar{\varphi}) - \bar{d}(\phi) \cos(\phi - \bar{\varphi})]}{\sum_{\phi} \tilde{w}(\phi) \bar{d}(\phi)^2 [\partial \bar{d}(\phi) \sin(\phi - \bar{\varphi}) - \bar{d}(\phi) \cos(\phi - \bar{\varphi})]^2} , \\ \Delta \varphi &\approx \frac{\sum_{\phi} \tilde{w}(\phi) \Delta \tilde{d}(\phi) \bar{d}(\phi) [\partial \bar{d}(\phi) \cos(\phi - \bar{\varphi}) + \bar{d}(\phi) \sin(\phi - \bar{\varphi})]}{\frac{1}{\alpha} \bar{l} \sum_{\phi} \tilde{w}(\phi) \bar{d}(\phi)^2 [\partial \bar{d}(\phi) \cos(\phi - \bar{\varphi}) + \bar{d}(\phi) \sin(\phi - \bar{\varphi})]^2} , \\ \Delta \psi &\approx \frac{\sum_{\phi} \tilde{w}(\phi) \Delta \tilde{d}(\phi) \partial \bar{d}(\phi)}{\sum_{\phi} \tilde{w}(\phi) \partial \bar{d}(\phi)^2} , \\ \tilde{w}(\phi) &:= w(\phi - \bar{\psi}|\bar{l}, \bar{\varphi}, \bar{\psi}) \approx (v^h(\phi - \bar{\psi}) + v[\phi - \frac{1}{\alpha} \bar{l} \bar{d}(\phi) \sin(\phi - \bar{\varphi})])^{-1} , \\ \Delta \tilde{d}(\phi) &:= \Delta d^h(\phi - \bar{\psi}) - \Delta d[\phi - \frac{1}{\alpha} \bar{l} \bar{d}(\phi) \sin(\phi - \bar{\varphi})] .\end{aligned}$$

Nimmt man desweiteren an, dass die Fehler $\Delta d^h(\phi)$ und $\Delta d(\phi)$ den Erwartungswert Null und die Varianzen $v^h(\phi)$ bzw. $v(\phi)$ besitzen¹⁷ und zudem unkorreliert sind, d.h.

$$\begin{aligned}\langle \Delta d^h(\phi) \rangle &= \langle \Delta d(\phi) \rangle = 0 , \\ \langle \Delta d^h(\phi_1) \Delta d^h(\phi_2) \rangle &\approx v^h \delta_{\phi_1, \phi_2} , \\ \langle \Delta d(\phi_1) \Delta d(\phi_2) \rangle &\approx v \delta_{\phi_1, \phi_2} , \\ \langle \Delta d(\phi_1) \Delta d^h(\phi_2) \rangle &= 0 ,\end{aligned}\tag{1.118}$$

so folgt $\langle \Delta l \rangle, \langle \Delta \varphi \rangle, \langle \Delta \psi \rangle \approx 0$. Unter diesen Annahmen verschwinden die systematischen Fehler. Da die Entfernungen bzw. Disparitäten direkt bestimmt werden können, ist (im Gegensatz zum Bild-basierten Heimfinden) eine Annahme über die Abstände der umgebenden Objekte nicht notwendig. Für die Varianzen gilt

$$\langle \Delta l^2 \rangle \approx \frac{\alpha^2}{\sum_{\phi} \tilde{w}(\phi) \bar{d}(\phi)^2 [\partial \bar{d}(\phi) \sin(\phi - \bar{\varphi}) - \bar{d}(\phi) \cos(\phi - \bar{\varphi})]^2} ,\tag{1.119}$$

$$\langle \Delta \varphi^2 \rangle \approx \frac{\alpha^2}{\bar{l}^2 \sum_{\phi} \tilde{w}(\phi) \bar{d}(\phi)^2 [\partial \bar{d}(\phi) \cos(\phi - \bar{\varphi}) + \bar{d}(\phi) \sin(\phi - \bar{\varphi})]^2} ,\tag{1.120}$$

$$\langle \Delta \psi^2 \rangle \approx \frac{1}{\sum_{\phi} \tilde{w}(\phi) \partial \bar{d}(\phi)^2} .\tag{1.121}$$

Die statistischen Fehler sind – wie zu erwarten war – klein, wenn die Varianzwerte $v^h(\phi)$ und $v(\phi)$ klein ($\tilde{w}(\phi)$ ist dann groß) und die Disparitätswerte

¹⁷Dies ist nach Abschnitt 1.2.1 nur für eindeutige Disparitätswerte mit schmaler Verteilung $P_d(d)$ eine gute Näherung.

unterschiedlich sind ($\partial\bar{d}(\phi)^2$ groß). Dies ist analog zu den Resultaten für Bild-basiertes Heimfinden, siehe Gleichungen (1.105)–(1.107). Für die Bestimmung von φ und l ist es zudem vorteilhaft, wenn die Disparitäten nicht zu klein, d.h. die umgebenden Objekte nicht zu weit entfernt sind. Dies kann bei einer reinen Rotation keine Rolle spielen, weshalb in $\langle\Delta\psi^2\rangle$ kein entsprechender Term auftritt. Auch für Disparitäts-basiertes Heimfinden begrenzt Rauschen für kleine Entfernungen zum Zielort die Genauigkeit, mit der die Richtung des Heimvektors bestimmt werden kann, so dass die „Average Homeward Component“ für kleine \bar{l} verschwindet.

Bemerkung: Eine Berechnung der Genauigkeit, mit der durch Disparitäts-basiertes Heimfinden ein Ort wiedergefunden werden kann, ist analog zu 1.5.1 möglich, wobei Gleichung (1.56) durch ein gewichtetes Ähnlichkeitsmaß zu ersetzen ist, vgl. (1.108). Außerdem gilt, anstelle von (1.62), $\partial_x\bar{d}(\phi) \approx \frac{1}{\alpha}\bar{d}(\phi)[- \partial\bar{d}(\phi) \sin\phi + \bar{d}(\phi) \cos\phi]$ und $\partial_y\bar{d}(\phi) \approx \frac{1}{\alpha}\bar{d}(\phi)[\partial\bar{d}(\phi) \cos\phi + \bar{d}(\phi) \sin\phi]$, wie aus (1.113) mit $\Delta x = l \cos \varphi$, $\Delta y = l \sin \varphi$ und $\phi' \rightarrow \phi$ abgeleitet werden kann. Unter den Annahmen (1.118) erhält man dann mit $\bar{l} = \bar{\psi} = 0$:

$$\Delta l_{\text{eq}}^2 \approx \frac{1}{\frac{1}{2} \sum_{\phi} \frac{1}{2v(\phi)} [(\partial_x\bar{d}(\phi))^2 + (\partial_y\bar{d}(\phi))^2]} \quad (1.122)$$

$$\approx \frac{4\alpha^2}{\sum_{\phi} \frac{1}{v(\phi)} \bar{d}(\phi)^2 [\partial\bar{d}(\phi)^2 + \bar{d}(\phi)^2]} \quad (1.123)$$

$$\approx \frac{4}{\sum_{\phi} \frac{\partial\bar{d}(\phi)^2 + \bar{d}(\phi)^2}{v(\phi) r(\phi)^2}} \quad (1.124)$$

Ersetzt man desweiteren den Varianzwert durch die Näherung (1.24), d.h.

$$v(\phi) \approx \frac{2\sigma_n^2}{\sum_{\zeta=0}^{N_A-1} (\partial_{\zeta} I_A[\phi, \zeta])^2} \quad (1.125)$$

so folgt

$$\Delta l_{\text{eq}}^2 \approx \frac{8\sigma_n^2}{\sum_{\phi} \frac{\sum_{\zeta} (\partial_{\zeta} I_A[\phi, \zeta])^2 [\partial\bar{d}(\phi)^2 + \bar{d}(\phi)^2]}{r(\phi)^2}} \quad (1.126)$$

Durch Vergleich mit (1.77) wird deutlich, dass im Gegensatz zum Bild-basierten Heimfinden, für das ein hoher Bildkontrast entlang des Horizonts vorteilhaft ist, das Disparitäts-basierte Verfahren für eine hohe Genauigkeit einen guten (bezogen auf das Kamerabild) radialen bzw. (bezogen auf die Umgebung) vertikalen Bildkontrast benötigt. In diesem Fall können Disparitäten zuverlässig bestimmt werden, sofern keine mehrdeutige Korrespondenzen (beispielsweise infolge periodischer Strukturen) auftreten.

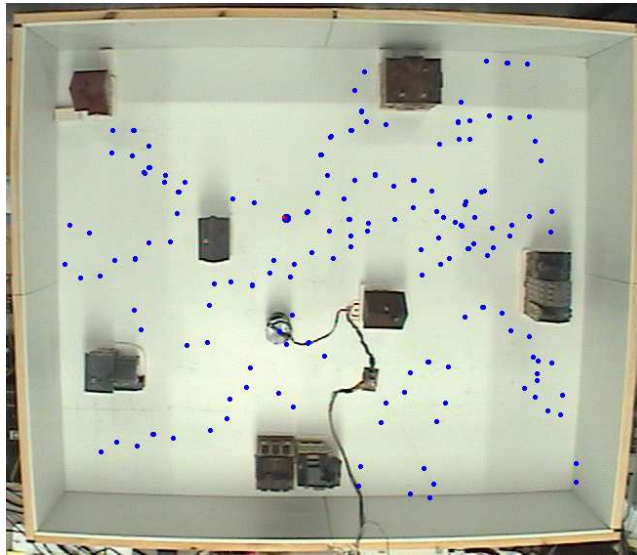


Abbildung 1.19: Einem Bild der Arena (aus Sicht der Tracking-Kamera) vor Beginn des Lernens sind die Positionen des Roboters (blaue Punkte) während eines etwa 15-minütigen „Random-Walks“ überlagert. Es wurden ca. 350 Translationen von 5 cm Länge und zufällige Rotationen um $\pm 45^\circ$ ausgeführt. Gezeigt sind nur die Anfangs- und Endpunkte der Translationen, zwischen denen keine Hindernisvermeidung ausgeführt werden musste. Nicht dargestellt sind außerdem Positionen, die vom Trackingsystem aufgrund (durch das Kabel) verdeckter LEDs nicht bestimmt werden konnten.

1.6 Erlernen der Bildtransformation

In Abschnitt 1.4 wurde das in [38] beschriebene Verfahren zum Bild-basierten Heimfinden verwendet. Um die Richtung zum Ziel bestimmen zu können, werden erwartete Bilder unter Annahme gleicher Abstände aller umgebenden Objekte berechnet und mit dem am Zielort gespeicherten Bild verglichen. Im Folgenden soll gezeigt werden, wie die dafür notwendige Bildtransformation durch Erfahrung erlernt werden kann. Dies wird durch ein einfaches neuronales Netz mit unüberwachtem assoziativem Lernen realisiert.

1.6.1 Trainieren des neuronalen Netzes

Zum Erlernen der Bildtransformation führt der Khepera einen „Random-Walk“ in der Modellhausarena durch, der aus zufälligen Rotationen um $+45^\circ$ oder -45° und Translationen von 5 cm Länge besteht, siehe Abbildung 1.19. Ein Lernschritt erfolgt nur dann, wenn eine Translation ausgeführt werden

konnte, ohne dass die Hindernisvermeidung aktiviert wurde. Dieser ist untergliedert in (siehe auch Abb. 1.20):

1. Vergleich einzelner Pixel des aktuellen Bildes¹⁸ \mathbf{I} mit denen des gespeicherten Bildes vor der Translationsbewegung \mathbf{I}^o (die Variable $t = 1, 2, \dots$ bezeichnet die Nummer des Lernschritts):

$$\text{corr}_{lk}(t) := \frac{1}{1 + \gamma |I_l(t) - I_k^o(t)|} , \quad (1.127)$$

wobei $\gamma = \frac{1}{4}$ verwendet wurde. Die genaue Form des Ähnlichkeitsmaßes ist dabei nicht von Bedeutung.

2. Anpassen der Gewichte $\forall l \in \{0, 1, \dots, N-1\}, k \in U_l$:

$$w_{lk}(t) := (1 - \alpha)w_{lk}(t-1) + \alpha w_{lk}^0(t) \quad (1.128)$$

$$w_{lk}^0(t) := \frac{\text{corr}_{lk}(t)}{N_U^{-1} \sum_{j \in U_l} \text{corr}_{lj}(t)} \quad (1.129)$$

Die Konstante α wurde für die weiter unten beschriebenen Resultate gleich $\frac{1}{30}$ gesetzt. Zu Beginn werden die Gewichte mit zufälligen Werten initialisiert, d.h. $w_{lk}(0) \in \{0, 0.1\}$.

Die Normierung in (1.129), d.h. die Division durch den Mittelwert der Korrelation, reduziert den Einfluss von Bildbereichen mit niedrigem Kontrast, die nur eine unsichere Bestimmung des Flussfeldes erlauben. Zur Beschleunigung des Algorithmus wird die Berechnung der Korrelation und die Anpassung der Gewichte auf benachbarte Pixel beschränkt, d.h.

$$U_l = \{l - n_U, l - n_U + 1, \dots, l + n_U\} , \quad (1.130)$$

wobei $n_U = 5$ verwendet wurde. Somit folgt $N_U = 2n_U + 1$ in Gleichung (1.129). Dies ist jedoch deutlich größer als die in der erlernten Transformation auftretenden maximalen Verschiebungen von $|k(l) - l| = 2$, siehe z.B. Lernschritte #20, #50 in Abbildung 1.21 a.

1.6.2 Verwendung der erlernten Bildtransformation

Die mit Hilfe der Gleichungen (1.127)–(1.129) angepassten Gewichte können dazu verwendet werden, das jeweils aktuelle Bild in ein erwartetes Bild (für

¹⁸Auch hier werden eindimensionale panoramische Grauwert-Bilder mit $N = 72$ Pixel verwendet. Benutzt man Bilder mit deutlich höherer Auflösung, dann ist es möglich, statt einzelner Pixel kleine Bildausschnitte zu vergleichen.

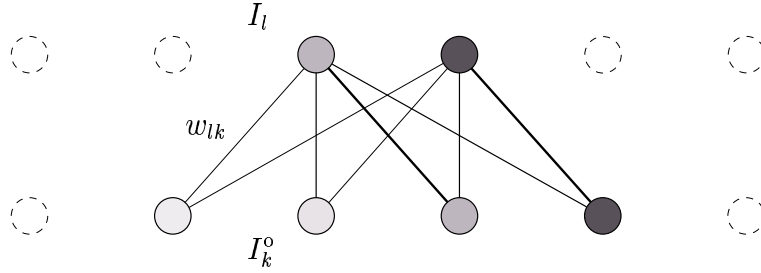


Abbildung 1.20: Netzwerktopologie in der Lernphase: Besitzen Pixel des aktuellen Bildes (obere Reihe) und Pixel des Bildes vor der Translation (untere Reihe) ähnliche Grauwerte ($I_l \approx I_k^o$), so werden die entsprechenden Gewichte w_{lk} gemäß Gleichungen (1.127)–(1.129) erhöht (größere Strichdicke deutet eine stärkere Gewichtung der jeweiligen Verbindung an). Für die Berechnung eines erwarteten Bildes wird nur die Verbindung mit dem stärksten Gewicht verwendet, siehe Gleichungen (1.131) und (1.132).

eine 5 cm-Translation) gemäß

$$I_l^e := I_{k(l)}, \quad l \in \{0, 1, \dots, N-1\} \quad (1.131)$$

$$k(l) := \arg \max_{k \in U_l} w_{lk} \quad (1.132)$$

zu transformieren.

Abbildung 1.21 vergleicht das vor der Translation berechnete Bild mit dem tatsächlichen Bild danach zu verschiedenen Stadien des Lernens: Zu Beginn jeder Translation wird ein erwartetes Bild $\mathbf{I}^e(\mathbf{I}^o(t) | \hat{\mathbf{W}}(t-1))$ mit Hilfe der im vorangegangenen Lernschritt angepassten Gewichte $\hat{\mathbf{W}}(t) := (w_{lk}(t))_{l,k=0,1,\dots,N-1}$ nach Gleichungen (1.131) und (1.132) berechnet. Dieses kann nach der Translation mit dem aktuellen Bild verglichen werden. Bereits nach etwa $t = 20$ Lernschritten besitzt die Transformation eine geordnete Struktur – seitliche Bildpixel werden im erwarteten Bild entgegen der Bewegungsrichtung verschoben – die sich im weiteren Verlauf nur geringfügig ändert. Außerdem besitzt sie große Ähnlichkeit mit der durch die „Equal Distance Assumption“ festgelegten Transformation. Dies ist nicht erstaunlich, da für die Anpassung der Gewichte über Translationen in verschiedene Richtungen gemittelt wird.

Ein quantitativer Vergleich ist in Abbildung 1.22 dargestellt, wobei die Summe der absoluten Differenzen (SAD) als Ähnlichkeitsmaß verwendet wurde,

$$\text{SAD}[\mathbf{I}^A, \mathbf{I}^B] := \sum_k |I_k^A - I_k^B|. \quad (1.133)$$

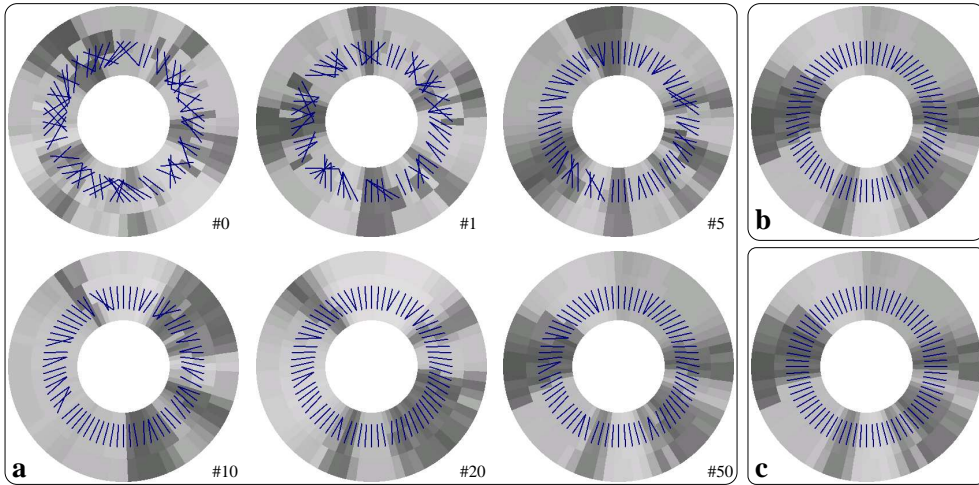


Abbildung 1.21: **a:** „Bildvorhersage“ durch Erlernen des mittleren Flussfeldes, dass einer Translation von 5 cm entspricht: Abgebildet ist jeweils im inneren Ring das Bild vor der Translation, im mittleren das mit Hilfe der erlernte Bildtransformation aus diesem berechnete Bild und im äußeren das aktuelle Bild nach der Translation. Die blauen Linien stellen die aktuelle Bildtransformation dar, d.h. zeigen an, welche Pixel des ursprünglichen Bildes für das berechnete Bild verwendet wurden. Angegeben ist zudem die Zahl der absolvierten Lernschritte ($\#t$). Ausgehend von einer zufälligen Verteilung der Gewichte (siehe $\#0$) stellt sich nach etwa 20 Schritten eine regelmäßige Struktur ein, die sich im weiteren Verlauf nicht mehr stark ändert. Zum Vergleich ist in **b** die Bildtransformation dargestellt, wie sie durch die Annahme gleicher Abstände mit $\lambda = 1/9$ (siehe Abschnitt 1.4.1) gegeben ist. **c** zeigt die „triviale Transformation“, d.h. das „transformierte“ Bild entspricht dem ursprünglichen Bild.

Verglichen wird in Abbildung 1.22 a der Fehler der erlernten Transformation,

$$\text{SAD}[t|\mathbf{W}(t-1)] := \text{SAD}[\mathbf{I}(t), \mathbf{I}^e(\mathbf{I}^o(t)|\hat{\mathbf{W}}(t-1))] , \quad (1.134)$$

mit dem Fehler infolge der „Equal Distance Assumption“ (EDA) mit $R = 45 \text{ cm}$ ($\leftrightarrow \lambda = 1/9$),

$$\text{SAD}[t|\text{EDA}] := \text{SAD}[\mathbf{I}(t), \mathbf{I}^e(\mathbf{I}^o(t)|\text{EDA})] , \quad (1.135)$$

und mit dem direkten Bildfehler $\text{SAD}[\mathbf{I}(t), \mathbf{I}^o(t)]$ für $1 \leq t \leq 50$. Bereits nach 6 Lernschritten gilt $\text{SAD}[t|\hat{\mathbf{W}}(t-1)] < \text{SAD}[\mathbf{I}(t), \mathbf{I}^o(t)]$ und nach etwa 15 Schritten $\text{SAD}[t|\hat{\mathbf{W}}(t-1)] \approx \text{SAD}[t|\text{EDA}]$.

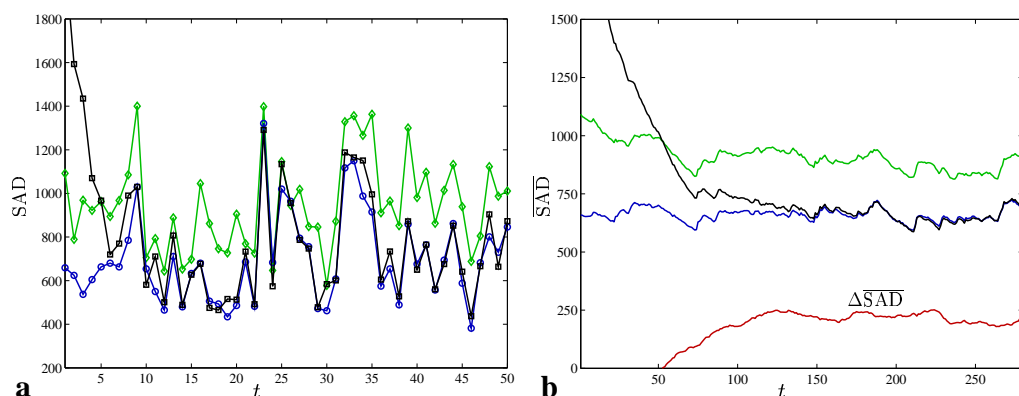


Abbildung 1.22: Fehler der erlernten Transformation $SAD[t|\mathbf{W}(t-1)]$ (schwarze Kurve) in Abhängigkeit von t im Vergleich zur „Equal Distance Assumption“ $SAD[t|EDA]$ (blau) und dem direkten Bildfehler $SAD[\mathbf{I}(t), \mathbf{I}^o(t)]$ (grün). **a:** gezeigt ist der Beginn der Lernphase $1 \leq t \leq 50$. **b:** Zeitgefilterte Fehler für $1 \leq t \leq 280$. Die rote Kurve zeigt die Größe $\overline{\Delta SAD}[t]$, die der Agent zur Kontrolle seines Lernerfolgs benutzen könnte (siehe Beschreibung im Text).

In Abbildung 1.22 b sind die entsprechenden Fehler für $1 \leq t \leq 280$ zur besseren Übersicht zeitlich gefiltert dargestellt, d.h.

$$\overline{SAD}[t] := (1 - \alpha)\overline{SAD}[t-1] + \alpha SAD[t] . \quad (1.136)$$

Als Zeitkonstante wurde wieder $\alpha = \frac{1}{30}$ gewählt. Auch hier zeigt sich, dass die erlernte Transformation zu annähernd identischen Resultaten führt wie die „Equal Distance Assumption“ und die Fehler deutlich unter dem direkten Bildvergleich bleiben.

Die rote Kurve stellt den Verlauf von $\overline{\Delta SAD}[t] := \overline{SAD}[\mathbf{I}(t), \mathbf{I}^o(t)] - \overline{SAD}[t|\hat{\mathbf{W}}(t-1)]$ dar. Diese Größe kann vom lernenden Agenten als leicht zugängliches Maß für den Lernerfolg verwendet werden. Sobald $\overline{\Delta SAD}[t] > 0$ gilt, ist eine sinnvolle Berechnung erwarteter Bilder möglich. Wenn sich $\overline{\Delta SAD}[t]$ nicht mehr stark verändert, könnte das Lernen – sofern die Umgebung weitgehend homogen ist – abgebrochen werden.

Der über alle Bildpaare ($N_t = 280$) gemittelte Fehler für die erlernte Transformation zur Zeit t , d.h.

$$\langle SAD \rangle [t] := \frac{1}{N_t} \sum_{t'=1}^{N_t} SAD[t'|\hat{\mathbf{W}}(t-1)] , \quad (1.137)$$

ist in Abbildung 1.23 a dargestellt. Dies ist eine Schätzung für den mittleren Fehler, der zu erwarten ist, wenn das Lernen zu Zeit $t-1$ abgebrochen wird.

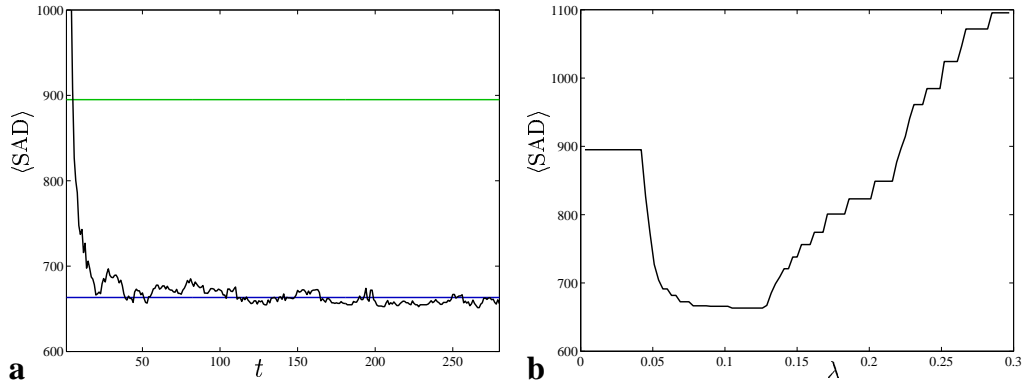


Abbildung 1.23: **a:** Über alle 280 Translationen gemittelter Fehler in Abhängigkeit der zur Zeit t erlernten Transformation (schwarz). Zum Vergleich sind der entsprechende Werte für die „Equal Distance Assumption“ mit $\lambda = 1/9$ (blau) und der mittlere Bildfehler $\frac{1}{N_t} \sum_t \text{SAD}[\mathbf{I}(t), \mathbf{I}^o(t)]$ eingezeichnet (grün). **b:** Mittleren Fehler für die „Equal Distance Assumption“ in Abhängigkeit von $\lambda (= 5 \text{ cm}/R)$.

Abbildung 1.23 b zeigt den mittleren Fehler für die „Equal Distance Assumption“ in Abhängigkeit des Parameters λ . Das Minimum liegt bei ungefähr $\lambda_{\text{opt}} \approx \frac{1}{9}$. Dies ist der für die Vergleiche mit der erlernten Transformation verwendete Wert. Bei einer Länge der Translationen von $l = 5 \text{ cm}$ entspricht diesem der Abstand $R = 5 \text{ cm}/\lambda_{\text{opt}} \approx 45 \text{ cm}$. Der stufenförmige Verlauf ist auf die geringe Auflösung und die Rundung auf ganze Pixel bei der Bildberechnung zurückzuführen. So gilt beispielsweise für den minimalen Wert von λ , für den sich das berechnete Bild von dem Originalbild unterscheiden kann, nach Gleichung (1.44)

$$\lambda \sin(\phi' - \varphi) \frac{N}{2\pi} \geq 0.5 \text{ [Pixel]} \quad (1.138)$$

$$\implies \lambda_{\text{min}} \approx 0.044 \text{ .} \quad (1.139)$$

Der Faktor $N/2\pi$ steht für die Umrechnung des Winkels in Pixel ($N = 72$).

Bemerkungen: Aus Gleichung (1.128) folgt

$$w_{lk}(t) = (1 - \alpha)^t w_{lk}(0) + \alpha \sum_{\tau=1}^t (1 - \alpha)^{t-\tau} w_{lk}^0(\tau) \quad (1.140)$$

$$\approx e^{-\alpha t} w_{lk}(0) + \alpha \sum_{\tau=1}^t e^{-\alpha(t-\tau)} w_{lk}^0(\tau) \quad \text{für } \alpha \ll 1 \text{ .} \quad (1.141)$$

Der Einfluss eines einzelnen Lernschritts fällt deshalb mit α^{-1} ab. Für $\alpha = \frac{1}{30}$ mittelt das System somit näherungsweise über 30 Schritte. Eine (evtl. auch von der Zahl der Lernschritte abhängige) geeignete Wahl von α ermöglicht eine Anpassung an die Umgebung und an die Qualität der einzelnen Flussfeldschätzungen. In der Modellhausarena ist – zumindest bei der verwendeten Bildauflösung von nur $N = 72$ Pixel – eine einzelne Bestimmung des optischen Flussfeldes unzureichend, siehe Abbildung 1.21 a für $t = 1$.

Zur Berechnung eines Bildes für eine beliebige Bewegungsrichtung – wie beispielsweise für das visuelle Heimfinden aus Abschnitt 1.4 erforderlich – muss das aktuelle Bild vor Verwendung der erlernten Transformation lediglich rotiert werden. Auch kann das beschriebene Modell leicht erweitert werden, um das Erlernen der Bildtransformation für Translationen und Rotationen unterschiedlicher Größe zu ermöglichen.

Kapitel 2

Fourier-basierter Bildvergleich¹

Müssen viele Bilder gespeichert und verglichen werden, beispielsweise für die Navigation in großen Umgebungen, so kann die Datenmenge schnell zum Problem werden. Eine Möglichkeit diese zu reduzieren, ist die Extraktion von Bildmerkmalen (engl. „features“) wie Kanten bestimmter Orientierung, Bildausschnitte mit hohem Kontrast oder besonderen Farbwerten, bis hin zum Erkennen einzelner Objekte.

Alternativ kann man versuchen, eine geeignete Transformation des ganzen Bildes vorzunehmen. Ein in den letzten Jahren häufig für die Lokalisation von Robotern verwendetes Verfahren [53, 73], ist die so genannte PCA („Principal Component Analysis“): Eine lineare Transformation, die aus Bildern der jeweiligen Umgebung berechnet wird und mit deren Hilfe dann anschließend Bilder durch wenige PCA-Koeffizienten repräsentiert werden können. Neben dem Speicherbedarf können dadurch gleichzeitig Störeinflüsse (z.B. Rauschen) und die Zahl der für einen Bildvergleich nötigen Operationen verringert werden. Es kann gezeigt werden, dass bezüglich des quadratischen Bildabstands – siehe Gleichung (2.1) – die PCA die lineare Transformation mit minimalem Rekonstruktionsfehler ist [28]. Für panoramische Bilder besitzt sie jedoch den Nachteil, dass die Koeffizienten von der Aufnahmerichtung abhängen und daher den Bildvergleich erschweren.

In diesem Kapitel wird als Alternative die Verwendung der Fouriertransformation untersucht. Diese besitzt den Vorteil, dass bei Rotation eines Bildes lediglich die Phasen der Fourierkoeffizienten variieren. Zunächst wird eine effiziente Bestimmung der relativen Bildorientierung hergeleitet, die einen beschleunigten Vergleich panoramischer Bilder ermöglicht. Nach der Diskussion möglicher Fehler bei der Orientierungsbestimmung wird ein Algorithmus für das Heimfinden mit Fourier-transformierten Bildern beschrieben. Eine

¹Teile der Abschnitte 2.1 und 2.2 wurden bereits in [92] veröffentlicht.

Verbindung zur PCA stellt Anhang A.5 her. Dort wird gezeigt, dass die Berechnung der PCA auf einem rotationssymmetrischen Datensatz, in dem Datenvektoren (z.B. eindimensionale Bilder) in allen möglichen Orientierungen zyklisch rotiert aufgenommen werden, auf eine Fouriertransformation führt.

2.1 Bildvergleich mit Korrelation und SSD

Ein einfaches Maß für die Ähnlichkeit zweier Bilder² stellt die Summe der quadratischen Differenzen (SSD: „Sum of Squared Differences“) dar:

$$\text{SSD}(\mathbf{I}, \mathbf{J}) := (\mathbf{I} - \mathbf{J})^2 = \sum_{j=0}^{N-1} (I_j - J_j)^2 \quad (2.1)$$

$$= \mathbf{I}^2 + \mathbf{J}^2 - 2\mathbf{I}^\top \mathbf{J} = \sum_j I_j^2 + J_j^2 - 2I_j J_j . \quad (2.2)$$

Bei gleicher Länge der Bildvektoren entspricht eine minimale SSD einem Maximum der Korrelation

$$\text{corr}(\mathbf{I}, \mathbf{J}) := \mathbf{I}^\top \mathbf{J} . \quad (2.3)$$

Da verschiedene Aufnahmen der gleichen Szene häufig leichte Helligkeitsverschiebungen aufweisen, verwendet man in der Praxis meist mittelwertfreie Bilder, siehe z.B. [33], d.h.

$$I_j - \frac{1}{N} \sum_l I_l \rightarrow I_j . \quad (2.4)$$

Um einen beschränkten Wertebereich zu erhalten, kann die normierte Korrelation verwendet werden³

$$\frac{\text{corr}(\mathbf{I}, \mathbf{J})}{|\mathbf{I}||\mathbf{J}|} \in [-1, 1] , \quad (2.5)$$

die darüberhinaus invariant gegenüber einer gleichmäßigen Veränderung des Kontrasts ($\mathbf{I} \rightarrow \alpha\mathbf{I}$) ist.

2.1.1 Korrelation panoramischer Bilder

Abbildung 2.1 zeigt die (normierte) Bildkorrelation relativ zum Startbild während einer vollständigen Rotation des Khepera-Roboters in Abhängigkeit

²Bilder werden im Folgenden (wie bereits in Kapitel 1) häufig als Vektoren $\mathbf{I} := (I_0, I_1, \dots, I_{N-1})^\top$ geschrieben.

³Eine auf $[0, 1]$ beschränkte SSD ist z.B. durch $\frac{1}{2} \frac{(\mathbf{I}-\mathbf{J})^2}{\mathbf{I}^2 + \mathbf{J}^2}$ zu erzielen.

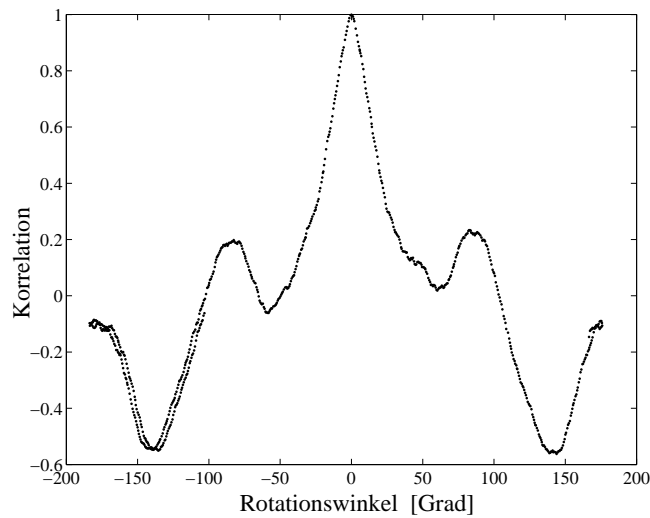


Abbildung 2.1: Normierte Bildkorrelation zum Startbild (bei 0°) während einer 450° -Rotation ($-190^\circ \rightarrow 260^\circ$) des Kheperas in Abhängigkeit vom Rotationswinkel in 0.5° -Schritten. Um Artefakte in der Korrelationskurve (z.B. leicht verschobene Kamera-Halbbilder) zu vermeiden, wurde die Rotation jeweils vor Aufnahme eines Bildes gestoppt. Durch das damit verbundene häufige Beschleunigen und Abbremsen weist die Winkelmessung über Odometrie eine leichte Verschiebung auf (zu sehen im Bereich $[-180^\circ, -100^\circ]$).

vom Rotationswinkel. Die eindimensionalen Panoramabilder bestehen (wie in Abschnitt 1.4.1 beschrieben) aus $N = 72$ Pixel, die aus den Kamerabildern extrahiert werden und einen 5° -Bereich um den Horizont repräsentieren. Wie zu erkennen ist, existieren neben einem ausgeprägten globalen Maximum bei 0° weitere lokale Maxima, deren Zahl von der Umgebung und dem Kamerarauschen abhängt. Dies erschwert den Vergleich panoramischer Bilder mit unbekannter Orientierung. Häufig werden deshalb alle möglichen relativen Orientierung getestet, d.h.

$$\text{corr}(\mathbf{I}, \mathbf{J}) := \max_s \text{corr}(\mathbf{I}, \mathbf{J}, s) \quad (2.6)$$

$$\text{corr}(\mathbf{I}, \mathbf{J}, s) := \sum_j I_j J_{(j+s) \bmod N} \quad (2.7)$$

Dies kann sehr zeitaufwendig sein, wenn beispielsweise ein Bild mit allen Bildern einer großen Datenbank verglichen werden muss. Für eindimensionale panoramische Bilder (bestehend aus N Pixeln) sind je Bildvergleich $O(N^2)$ Berechnungen notwendig, bzw. $O(N_x^2 N_y)$ Berechnungen für zweidimensionale Bilder mit $N_x \times N_y$ Pixel.

In den folgenden Abschnitten wird ein Algorithmus mit Komplexität $O(N)$ hergeleitet, der gleichzeitig eine Bestimmung der relativen Orientierung in Sub-Pixel-Genauigkeit ermöglicht.

2.1.2 Fourierdarstellung panoramischer Bilder

Ein trigonometrisches Polynom $I(\phi)$ von minimalem Grad, das an den äquidistanten Punkten $\phi_j = \frac{2\pi}{N}j, j = 0, 1, \dots, N-1$ die Werte I_j annimmt⁴, ist für gerades N durch

$$I(\phi) := \sum_{k=-N/2+1}^{N/2-1} \mathcal{I}_k e^{ik\phi} + \frac{1}{2} \mathcal{I}_{-N/2} (e^{i\frac{N}{2}\phi} + e^{-i\frac{N}{2}\phi}) \quad (2.8)$$

$$\mathcal{I}_k := \frac{1}{N} \sum_{j=0}^{N-1} I_j e^{-ik\phi_j} \quad (2.9)$$

gegeben⁵, siehe [64]. Die Form in (2.8) wurde so gewählt, dass für reelle Bilder $I_j \in \mathbb{R}$ auch $I(\phi)$ reellwertig ist. Aus $I_j^* = I_j$ folgt $\mathcal{I}_{-k} = \mathcal{I}_k^*$, $\mathcal{I}_{-N/2} = \frac{1}{N} \sum_j (-1)^j I_j$ und es gilt

$$I(\phi) = \mathcal{I}_0 + \sum_{k=1}^{N/2-1} [\mathcal{I}_k e^{ik\phi} + \mathcal{I}_k^* e^{-ik\phi}] + \frac{1}{2} \mathcal{I}_{-N/2} (e^{i\frac{N}{2}\phi} + e^{-i\frac{N}{2}\phi}) \quad (2.10)$$

$$= \mathcal{I}_0 + \sum_{k=1}^{N/2-1} 2\operatorname{Re}[\mathcal{I}_k e^{ik\phi}] + \operatorname{Re}[\mathcal{I}_{-N/2} e^{-i\frac{N}{2}\phi}] . \quad (2.11)$$

Vorteile der Fourierdarstellung

Aus (2.9) folgt unmittelbar, dass sich die Fourierkoeffizienten von panoramischen Bildern, die in unterschiedlicher Orientierung (unter idealen Bedingungen) aufgenommen wurden, nur in ihren Phasen unterscheiden (s Orientierungsdifferenz in Pixel):

$$\mathcal{I}_k(s) := \frac{1}{N} \sum_{j=0}^{N-1} I_{[(j+s) \bmod N]} e^{-ik\phi_j} \quad (2.12)$$

$$= \frac{1}{N} \sum_{j=0}^{N-1} I_j e^{-ik\frac{2\pi}{N}(j+s)} = e^{-i\frac{2\pi}{N}ks} \mathcal{I}_k . \quad (2.13)$$

⁴Im Folgenden werden nur eindimensionale Bilder betrachtet, eine Erweiterung auf zweidimensionale Bilder ist im Anhang A.6 zu finden.

⁵Für ungerades N gilt $I(\phi) = \sum_{k=-n}^n \mathcal{I}_k e^{ik\phi}$ mit $n := (N-1)/2$.

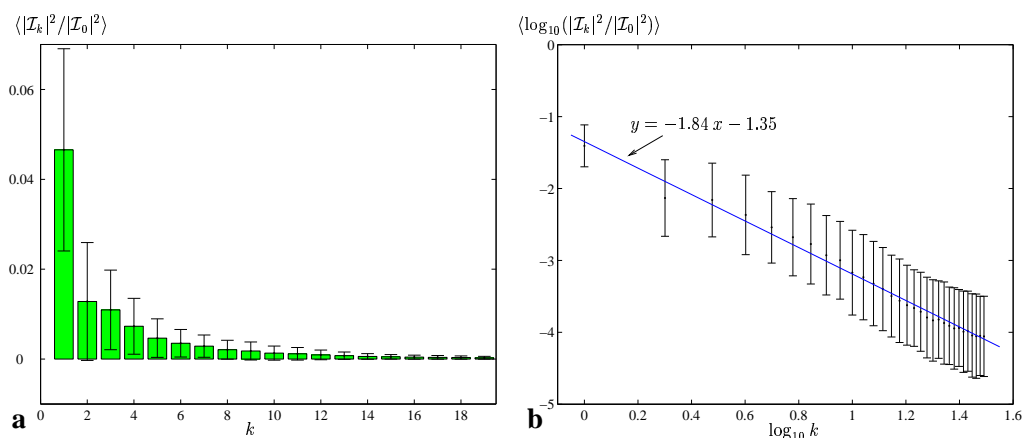


Abbildung 2.2: Frequenzspektrum der Bilddatenbank: **a:** Die Amplituden der Fourierkoeffizienten (gemittelt über 1250 panoramische Bilder) fallen schnell mit steigender Frequenz ab. **b:** In doppellogarithmischer Darstellung kann die Verteilung durch eine Gerade genähert werden. Fehlerbalken geben die Standardabweichung an.

Desweiteren ist bekannt, dass in natürlichen Bildern tiefe Frequenzen gewöhnlich dominieren: Das mittlere Frequenzspektrum kann in guter Näherung durch $\langle |\mathcal{I}_{\mathbf{k}}|^2 \rangle \propto |\mathbf{k}|^{-\alpha}$ beschrieben werden, wobei für den Exponenten α je nach Bilddatensatz Werte zwischen 1.5 und 2.5 ermittelt wurden [103]. Das Frequenzspektrum für die Bilddatenbank der Modellhausarena ist in Abbildung 2.2 dargestellt. Es gilt näherungsweise $\langle |\mathcal{I}_k|^2 \rangle \propto k^{-1.8}$. Deshalb können Bilder normalerweise durch wenige Fourierkoeffizienten approximiert werden, d.h.

$$I(\phi) \approx \sum_{k=-K}^K \mathcal{I}_k e^{ik\phi}, \quad (2.14)$$

mit $K \ll \frac{N}{2}$. Diese Eigenschaft wird u.a. bei der Komprimierung zweidimensionaler Bilder im JPEG-Format ausgenutzt.⁶

Um die Schreibweise zu vereinfachen, wird im Folgenden generell (2.14) anstelle der exakten Form (2.8) verwendet.

Bemerkung: Ist das Frequenzspektrum eines zweidimensionalen Bildes durch

$$|\mathcal{I}_{\mathbf{k}}|^2 \propto |\mathbf{k}|^{-\alpha} = (k_x^2 + k_y^2)^{-\frac{\alpha}{2}} \quad (2.15)$$

⁶Der JPEG-Algorithmus basiert auf einer zweidimensionalen diskreten Cosinustransformation, siehe z.B. [85].

gegeben, so kann das Spektrum eines aus diesem berechneten eindimensionalen Bildes nicht automatisch durch den gleichen Exponenten α beschrieben werden. Dies wird im Folgenden für den kontinuierlichen Grenzfall gezeigt.

Wird das eindimensionale Bild aus dem zweidimensionalen durch Faltung mit einer Filterfunktion bestimmt, d.h.

$$I(x) = \iint G(x - x', -y') I(x, y) dx' dy' , \quad (2.16)$$

so folgt für das Frequenzspektrum des eindimensionalen Bildes:

$$|\mathcal{I}(k)|^2 \propto \int \mathcal{G}(k, k_y)^2 |\mathcal{I}(k, k_y)|^2 dk_y , \quad (2.17)$$

wobei $\mathcal{G}(k_x, k_y)$ die Fouriertransformierte der Filterfunktion $G(x, y)$ bezeichnet.

Betrachtet man einen in x und y separierbaren Filter $G(x, y) = G_x(x)G_y(y)$ und nimmt für $G_y(y)$ eine Gaußfunktion mit Standardabweichung σ_y an, so folgt aus (2.15) und (2.17):

$$|\mathcal{I}(k)|^2 \propto \mathcal{G}_x(k)^2 \int \frac{\exp(-\sigma_y^2 k_y^2)}{(k^2 + k_y^2)^{\frac{\alpha}{2}}} dk_y \quad (2.18)$$

$$= \frac{\mathcal{G}_x(k)^2}{|k|^{\alpha-1}} \int \frac{\exp(-(\sigma_y k)^2 u^2)}{(1 + u^2)^{\frac{\alpha}{2}}} du \quad (u := k_y/|k|) . \quad (2.19)$$

Gleichung (2.19) besitzt die Grenzfälle:

1. $\sigma_y k \ll 1$: Es gilt näherungsweise

$$|\mathcal{I}(k)|^2 \propto \frac{\mathcal{G}_x(k)^2}{|k|^{\alpha-1}} \int \frac{du}{(1 + u^2)^{\frac{\alpha}{2}}} = \frac{\mathcal{G}_x(k)^2}{|k|^{\alpha-1}} \frac{\sqrt{\pi} \Gamma(\frac{\alpha-1}{2})}{\Gamma(\frac{\alpha}{2})} . \quad (2.20)$$

2. $\sigma_y k \gg 1$: $\exp(-(\sigma_y k)^2 u^2)$ kann durch $\sqrt{\pi} \sigma_y |k| \delta(u)$ genähert werden. Somit folgt

$$|\mathcal{I}(k)|^2 \propto \frac{\mathcal{G}_x(k)^2}{|k|^\alpha} \frac{\sqrt{\pi}}{\sigma_y} . \quad (2.21)$$

Je nach betrachtetem Frequenzbereich ist – ohne den Einfluss der Filterung in x -Richtung zu berücksichtigen – ein Exponent zwischen $\alpha - 1$ und α bzw. ein Übergang zwischen beiden Grenzfällen zu erwarten.

2.1.3 SSD und Korrelation im Fourier-Raum

Für zwei mittelwertfreien Bilder \mathbf{I} , \mathbf{J} mit Fourierkoeffizienten $\{\mathcal{I}_k = |\mathcal{I}_k|e^{i\psi_k}\}$ bzw. $\{\mathcal{J}_k = |\mathcal{J}_k|e^{i\varphi_k}\}_{k=0, \pm 1, \dots, \pm K}$ folgt für die orientierungsabhängige SSD aus Gleichung (2.2) mit $I_j = I_j^*$, $J_j = J_j^*$ (reelle Grauwerte) und (2.13):

$$\begin{aligned} \text{SSD}(\mathbf{I}, \mathbf{J}, s) &= \mathbf{I}^2 + \mathbf{J}^2 - 2 \text{corr}(\mathbf{I}, \mathbf{J}, s) \\ &= \left(\sum_{k,l=-K}^K \mathcal{I}_k^* \mathcal{I}_l + \mathcal{J}_k^* \mathcal{J}_l - 2 \mathcal{I}_k^* e^{-i\frac{2\pi}{N}l s} \mathcal{J}_l \right) \sum_j e^{i\frac{2\pi}{N}(l-k)j} \end{aligned} \quad (2.22)$$

$$\stackrel{(*)}{=} N \sum_{k=-K}^K |\mathcal{I}_k|^2 + |\mathcal{J}_k|^2 - 2 \mathcal{I}_k^* \mathcal{J}_k e^{-i\frac{2\pi}{N}k s} \quad (2.23)$$

$$= 2N \sum_{k=1}^K |\mathcal{I}_k|^2 + |\mathcal{J}_k|^2 - 2 \text{Re}[\mathcal{I}_k^* \mathcal{J}_k e^{-i\frac{2\pi}{N}k s}] \quad (2.24)$$

$$= 2N \sum_{k=1}^K |\mathcal{I}_k|^2 + |\mathcal{J}_k|^2 - 2|\mathcal{I}_k||\mathcal{J}_k| \cos(\varphi_k - \psi_k - \frac{2\pi}{N}k s) . \quad (2.25)$$

Für (*) wurde $\sum_{j=0}^{N-1} e^{i\frac{2\pi}{N}(l-k)j} = N\delta_{lk}$ verwendet. Die (orientierungsabhängige) Korrelation lautet somit im Frequenzraum, vgl. (2.2) und (2.3):

$$\text{corr}(\mathbf{I}, \mathbf{J}, s) = 2N \sum_{k=1}^K |\mathcal{I}_k||\mathcal{J}_k| \cos(\varphi_k - \psi_k - \omega_k s) , \quad (2.26)$$

wobei $\omega_k := \frac{2\pi}{N}k$ definiert wurde. Gleichung (2.26) ist in Abbildung 2.3 anhand zweier Bilder veranschaulicht, die am gleichen Ort in unterschiedlicher Orientierung aufgenommen wurden. Wie zu erkennen ist, besitzen die einzelnen Summanden Maxima nahe dem globalen Maximum der Bildkorrelation. Dies wird im Folgenden genauer untersucht.

2.1.4 Bestimmung der optimalen Bildverschiebung

Um die optimale relative Orientierung s^* zu ermitteln, wird die Nullstelle der Ableitung der Korrelation bestimmt:

$$0 = \frac{\partial}{\partial s} \text{corr}(\mathbf{I}, \mathbf{J}, s) \iff \quad (2.27)$$

$$0 = \sum_k |\mathcal{I}_k||\mathcal{J}_k| \sin(\varphi_k - \psi_k - \omega_k s) \omega_k \quad (2.28)$$

$$\begin{aligned} &= \sum_k |\mathcal{I}_k||\mathcal{J}_k| \omega_k [\sin(\varphi_k - \psi_k) \cos(\omega_k s) \\ &\quad - \cos(\varphi_k - \psi_k) \sin(\omega_k s)] . \end{aligned} \quad (2.29)$$

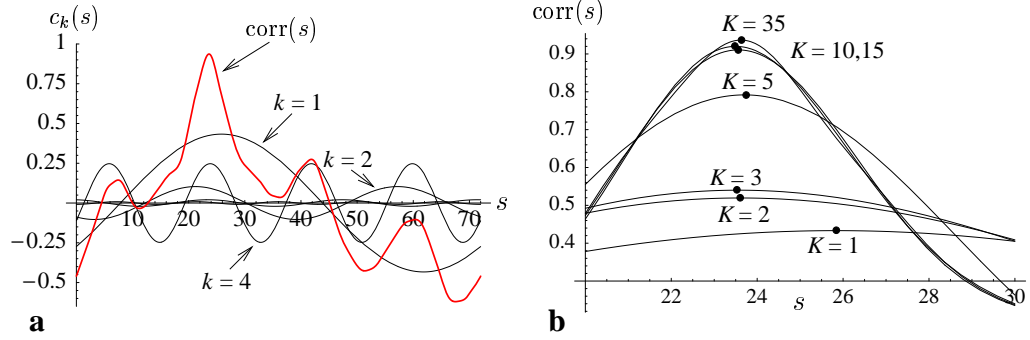


Abbildung 2.3: **a:** Die ersten fünf Summanden $c_k(s) := 2N|\mathcal{I}_k||\mathcal{J}_k| \cos(\varphi_k - \psi_k - \frac{2\pi}{N}ks)$ von Gleichung (2.26) und die Summe aller Terme mit $K = 35$ (rote Kurve) für zwei Panoramabilder mit $N = 72$ Pixel, die eine relative Orientierung von ungefähr $120^\circ (= 24 \times 5^\circ)$ haben. $c_3(s)$ und $c_5(s)$ besitzen sehr kleine Amplituden. **b:** Maximum der Korrelation in Abhängigkeit von der Zahl der Fourierkoeffizienten. Für $K = 1$ weist das Maximum noch eine Abweichung von mehr als 2 Pixel ($= 10^\circ$) auf.

Falls **I** (im Wesentlichen) lediglich eine unterschiedlich orientierte Version von **J** ist, so folgt $|\mathcal{I}_k| \approx |\mathcal{J}_k|$. Außerdem gibt es wegen (2.13) eine Verschiebung $s \in [0, N[$ und ganze Zahlen $n_k \in \mathbb{Z}$, so dass gilt:

$$\psi_k \approx \varphi_k - \omega_k s + 2\pi n_k, \quad \forall k \quad (2.30)$$

$$\iff 0 \approx \varphi_k - \psi_k - \omega_k s + 2\pi n_k \quad (2.31)$$

$$\iff s \approx s_k := \frac{\varphi_k - \psi_k + 2\pi n_k}{\omega_k} = \frac{N}{k} \left(\frac{\varphi_k - \psi_k}{2\pi} + n_k \right). \quad (2.32)$$

Da die n_k nicht fest sind, gibt es nur für $k = 1$ eine eindeutige Lösung mit $s_k \in [0, N[$. Wegen (2.31) kann

$$\sin(\varphi_k - \psi_k - \omega_k s) = \sin(\varphi_k - \psi_k - \omega_k s + 2\pi n_k) \quad (2.33)$$

$$\approx \varphi_k - \psi_k - \omega_k s + 2\pi n_k \quad (2.34)$$

genähert werden.⁷ Eingesetzt in (2.28) folgt

$$0 \approx \sum_k |\mathcal{I}_k||\mathcal{J}_k| \omega_k (\varphi_k - \psi_k - \omega_k s + 2\pi n_k) \quad (2.35)$$

$$\iff s \approx \bar{s} := \frac{\sum_k |\mathcal{I}_k||\mathcal{J}_k| \omega_k (\varphi_k - \psi_k + 2\pi n_k)}{\sum_k |\mathcal{I}_k||\mathcal{J}_k| \omega_k^2} \quad (2.36)$$

$$\stackrel{(2.32)}{=} \frac{\sum_k |\mathcal{I}_k||\mathcal{J}_k| \omega_k^2 s_k}{\sum_k |\mathcal{I}_k||\mathcal{J}_k| \omega_k^2} = \frac{\sum_k \alpha_k s_k}{\sum_k \alpha_k}, \quad (2.37)$$

wobei $\alpha_k := |\mathcal{I}_k||\mathcal{J}_k| \omega_k^2$ substituiert wurde.

⁷An den alternativen Nullstellen bei $\varphi_k - \psi_k - \omega_k s + 2\pi n_k + \pi$ befinden sich die Minima der Korrelation, da $\cos(\pi + 2\pi n) = -1$ ($n \in \mathbb{Z}$).

2.1.5 Algorithmus zur Orientierungsbestimmung

Betrachtet man (2.32) und (2.37) und berücksichtigt, dass die n_k unbekannt sind, so bietet sich folgender „coarse-to-fine“-Algorithmus zur Bestimmung der relativen Orientierung s an:

$$\bar{s}_1 = s_1 = \left(\frac{\varphi_1 - \psi_1}{\omega_1} \right) \bmod N, \quad \alpha_1 = |\mathcal{I}_1| |\mathcal{J}_1| \omega_1^2, \quad (2.38)$$

for $k = 2, 3, \dots, K$ do:

$$n_k = \text{rint} \left[\frac{\psi_k - \varphi_k + \omega_k \bar{s}_{k-1}}{2\pi} \right] \in \mathbb{Z}, \quad (2.39)$$

$$s_k = \frac{\varphi_k - \psi_k + 2\pi n_k}{\omega_k}, \quad \alpha_k = |\mathcal{I}_k| |\mathcal{J}_k| \omega_k^2, \quad (2.40)$$

$$\bar{s}_k = \frac{\sum_{l=1}^k \alpha_l s_l}{\sum_{l=1}^k \alpha_l} = \frac{(\sum_{l=1}^{k-1} \alpha_l) \bar{s}_{k-1} + \alpha_k s_k}{\sum_{l=1}^{k-1} \alpha_l + \alpha_k} \quad (2.41)$$

endfor

'rint[]' in (2.39) bedeutet ein Runden zur nächstgelegenen ganzen Zahl.

Die ermittelte Verschiebung $s^{\text{est}} := \bar{s}_K \bmod N$ kann anschließend zur Berechnung der (normierten) Bildkorrelation oder der SSD verwendet werden, siehe Gleichung (2.25).

Bemerkungen: Die in (2.37) angegebenen Gewichte α_k der verschiedenen „Frequenzkanäle“ wurden direkt aus der Maximierung der Bildkorrelation abgeleitet. Bei Kenntnis der Spektren von Störeinflüssen (z.B. Rauschen) kann jedoch eine alternative Gewichtung zur Abschwächung bestimmter Frequenzen sinnvoll sein.

Der vorgestellte Algorithmus ermöglicht die Bestimmung der relativen Orientierung und damit der Bildkorrelation mit $O(N)$ Berechnungen bzw. – da im Normalfall bereits wenige Fourierkoeffizienten genügen (Zahl der verwendeten Koeffizienten $K \ll$ Zahl der Pixel im 1D-Panoramabild N) – mit $O(K)$ Operationen. Wegen (2.39) ist eine näherungsweise richtige Bestimmung von \bar{s}_1 entscheidend. Kritisch sind sehr kleine Beträge von \mathcal{I}_1 und \mathcal{J}_1 (relativ zum Rauschpegel), die eine korrekte Bestimmung der Phasen verhindern (siehe auch Abschnitt 2.4.2). Obwohl dies – wegen der bereits erwähnten Dominanz tiefer Frequenzen – in natürlichen Umgebungen selten auftritt, wird in der verwendeten Implementierung des Algorithmus geprüft, ob α_1 einen kritischen Schwellwert unterschreitet. In diesem Fall wird eine zweite Bestimmung von s nach obigem Schema, jedoch mit Startwert $\bar{s}'_1 = (\bar{s}_1 + N/2) \bmod N$, durchgeführt. Derjenige der beiden Verschiebungswerte, der einen höheren Korrelationswert liefert, wird dann als Schätzung

der Relativorientierung verwendet. Falls zusätzlich zu α_1 weitere, unmittelbar nachfolgende Gewichte α_k , $k = 2, 3, \dots$ sehr kleine Werte aufweisen, müssten sogar noch mehr mögliche Startwerte berücksichtigt werden. Für die Bilddatenbank der Modellhausarena ist dieser Fall jedoch nicht aufgetreten. Einzelne kleine Werte α_k ($k > 1$) hingegen sind unkritisch, da \bar{s}_k nach Gleichung (2.41) mittels einer gewichteten Summe berechnet wird.

In dem Fall, dass die zu vergleichenden Bilder nicht im Wesentlichen nur verdrehte Versionen desselben Bildes sind, sondern völlig verschieden, ist es möglich, dass die ermittelte Schätzung von s zu einem geringeren Korrelationswert führt als man mit den Gleichungen (2.6) und (2.7) erhalten würde. Dies sollte jedoch für die meisten Anwendungen unkritisch sein.

Der Nachteil, eine (diskrete) Fouriertransformation der Bilder durchführen zu müssen, kann dadurch reduziert werden, dass man zum einen die Zahl der verwendeten Fourierkoeffizienten möglichst klein wählt. Zum anderen kann die Fouriertransformation bereits während des Aufbaus der Bilddatenbank bzw. der Exploration erfolgen.

Verbindung zu anderen Arbeiten: Der vorgestellte Algorithmus kann in gewisser Hinsicht als Erweiterung der in [72] beschriebenen „Zero-Phase“-Darstellung betrachtet werden: Panoramische Bilder werden vor dem Vergleich so gedreht, dass die Phase des ersten Fourierkoeffizienten \mathcal{I}_1 (bzw. $\mathcal{I}_{1,0}$ für 2D-Panoramabilder) verschwindet. Durch diese Festlegung der Orientierung muss die Bildkorrelation im Idealfall ebenfalls nur einmal berechnet werden. Wie jedoch anhand von Abbildung 2.3 b zu erkennen ist und in Abschnitt 2.4 genauer untersucht wird, kann die Verwendung von nur einer Fourierphase zu deutlichen Orientierungsfehlern und damit zu falschen Korrelationswerten führen.

Verschiedene Phasen-basierte Stereoalgorithmen verwenden die Phasen lokalisierter Gaborkoeffizienten, um Bildverschiebungen (Disparitäten) analog zu Gleichung (2.31) zu bestimmen, siehe z.B. [83, 96] und Abschnitt 4.1. Die Verwendung der Fouriertransformation bietet sich jedoch gerade wegen ihrer nicht-lokalen Struktur für die Orientierungsbestimmung von Panoramabildern an.

2.2 Vergleich mit „Standard“-Korrelation

Für Bilder, die während einer Rotation des Kheperas aufgenommen wurden (gleicher Datensatz wie für Abbildung 2.1), wird die relative Orientierung zum Startbild sowohl mit Hilfe des vorgestellten Algorithmus als auch mit dem „Standardverfahren“ d.h. den Gleichungen (2.6) und (2.7), ermittelt und

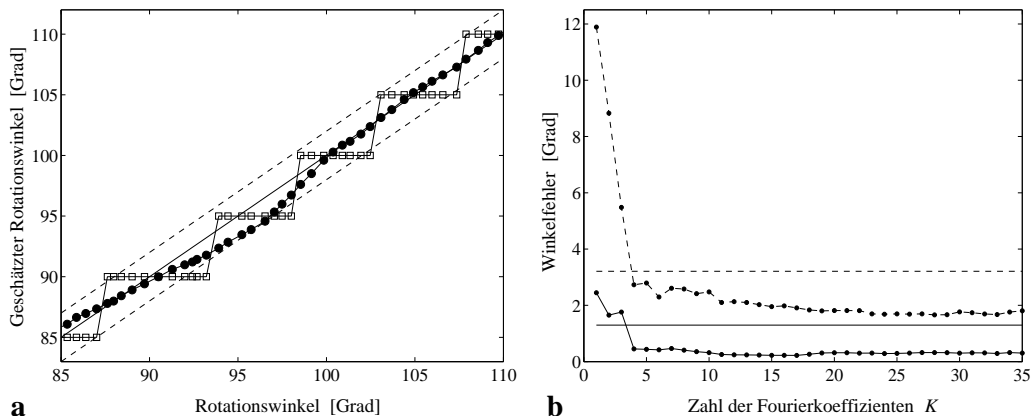


Abbildung 2.4: **a:** Rotationswinkel ermittelt mit Hilfe des beschriebenen Algorithmus unter Verwendung von 15 Fourierkoeffizienten (●) und der „Standard“-Korrelationsmethode (□): Die Gerade entspricht dem korrekten Winkel, gemessen mit der Odometrie des Roboters, ψ^{odo} . Gestrichelte Linien kennzeichnen $\psi^{\text{odo}} \pm 2^\circ$. Gezeigt ist nur der Bereich $[85^\circ, 110^\circ]$, indem der größte Fehler von ungefähr 2° auftritt. **b:** Maximum (gestrichelt) und Mittelwert (kontinuierliche Kurven) des Winkelfehlers $|\psi^{\text{odo}} - \psi^{\text{est}}|$, in Abhängigkeit von der Zahl der verwendeten Fourierkoeffizienten. Horizontale Geraden zeigen die entsprechenden Werte bei Verwendung des „Standard“-Verfahrens.

mit dem gemessenen Rotationswinkel verglichen. Die Messung des Rotationswinkels erfolgte mit Hilfe der Odometrie des Roboters. Um die in Abb. 2.1 zu erkennende Verschiebung zu korrigieren, wurden die Messwerte mit Hilfe der Bildkorrelation (nach einer vollständigen Rotation) linear skaliert. Die resultierenden Winkel werden mit ψ^{odo} bezeichnet. Wegen der geringen Bildauflösung und einem durch die Kabelverbindung zum externen Computer verdeckten 15° -Bereich (Grauwerte in diesem Bereich werden durch Interpolation aus den beiden angrenzenden Pixeln bestimmt) sind Abweichungen zu erwarten. Positionsfehler (siehe Abschnitt 2.4.1) spielen jedoch keine Rolle, da die Positionsabweichungen während der Rotation unterhalb der Messgenauigkeit des Tracking-Systems von ungefähr 3 mm lag.

Wie Abbildung 2.4 a zeigt, ermöglicht der Algorithmus die Bestimmung des Rotationswinkels $\psi = \frac{2\pi}{N}s$ in Sub-Pixel-Genauigkeit im Vergleich zur Verwendung der „Standard“-Korrelation, die (bei 72 Pixel) auf 5° -Schritte beschränkt ist.⁸ Wie in Abbildung 2.4 b dargestellt ist, liegen der Mittelwert $\approx 0.45^\circ$ und der Maximalwert $\approx 2.73^\circ$ des Fehlers bereits mit 4 Fourierkoeffizienten unterhalb denen der „Standard“-Korrelationsmethode ($\approx 1.30^\circ$

⁸Durch z.B. parabolische Interpolation kann natürlich auch für die „Standard“-Korrelation Sub-Pixel-Auflösung erreicht werden.

bzw. $\approx 3.21^\circ$). Der Wert von $\langle |\psi^{\text{odo}} - \psi^{\text{est}}| \rangle \approx 1.30^\circ$ entspricht ziemlich genau dem infolge der „Quantisierung“ (auf Vielfache von 5°) zu erwartenden Fehler. Bezeichnet ψ den korrekten Wert, so gilt für den mittleren Fehler des Schätzwertes ψ^{est} :

$$\langle |\psi - \psi^{\text{est}}| \rangle \approx \langle |\psi - \text{rint}[\psi/5^\circ]5^\circ| \rangle \quad (2.42)$$

$$= 1/5^\circ \int_{-5^\circ/2}^{5^\circ/2} |\Delta\psi| d\Delta\psi = 1.25^\circ . \quad (2.43)$$

2.3 Sinus-Cosinus-Form panoramischer Bilder

Um im Folgenden die Notation für (reellwertige) panoramische Bilder zu vereinfachen, werden reelle Koeffizienten \mathbf{a}_k , \mathbf{b}_k verwendet, die mit den komplexen Fourierkoeffizienten $\mathcal{I}_k = |\mathcal{I}_k|e^{i\psi_k}$ verknüpft sind. Es gilt:

$$\begin{aligned} \mathbf{a}_k &:= \mathcal{I}_k + \mathcal{I}_k^* , & \mathbf{b}_k &:= i(\mathcal{I}_k - \mathcal{I}_k^*) \\ \implies \mathcal{I}_k &= \frac{1}{2}(\mathbf{a}_k - i\mathbf{b}_k) , & |\mathcal{I}_k|^2 &= \frac{1}{4}(\mathbf{a}_k^2 + \mathbf{b}_k^2) , \\ \sin \psi_k &= \frac{-\mathbf{b}_k}{\sqrt{\mathbf{a}_k^2 + \mathbf{b}_k^2}} , & \cos \psi_k &= \frac{\mathbf{a}_k}{\sqrt{\mathbf{a}_k^2 + \mathbf{b}_k^2}} . \end{aligned} \quad (2.44)$$

Damit folgt aus Gleichung (2.11) die Sinus-Cosinus-Form (für gerade Pixelzahl N), siehe [64],

$$I(\phi) = \frac{1}{2}\mathbf{a}_0 + \sum_{k=1}^{n-1} \mathbf{a}_k \cos(k\phi) + \mathbf{b}_k \sin(k\phi) + \frac{1}{2}\mathbf{a}_n \cos(n\phi) , \quad (2.45)$$

wobei $n := N/2$. Die Koeffizienten können aus dem Bildvektor gemäß

$$\mathbf{a}_k := \frac{2}{N} \sum_j I_j \cos(k\phi_j) , \quad (2.46)$$

$$\mathbf{b}_k := \frac{2}{N} \sum_j I_j \sin(k\phi_j) \quad (2.47)$$

berechnet werden.

2.4 Fehler bei der Orientierungbestimmung

Kleine Störungen in den Koeffizienten⁹ $\mathcal{I}_k = \bar{\mathcal{I}}_k + \Delta\mathcal{I}_k = \frac{1}{2}(\mathbf{a}_k - i\mathbf{b}_k) = \frac{1}{2}[\bar{\mathbf{a}}_k + \Delta\mathbf{a}_k - i(\bar{\mathbf{b}}_k + \Delta\mathbf{b}_k)]$ verursachen einen Fehler in der Phase ψ_k :

$$\sin\Delta\psi_k = \sin(\psi_k - \bar{\psi}_k) = \sin\psi \cos\bar{\psi}_k - \cos\psi_k \sin\bar{\psi}_k \quad (2.48)$$

$$= \frac{-\mathbf{b}_k}{\sqrt{\mathbf{a}_k^2 + \mathbf{b}_k^2}} \frac{\bar{\mathbf{a}}_k}{\sqrt{\bar{\mathbf{a}}_k^2 + \bar{\mathbf{b}}_k^2}} - \frac{\mathbf{a}_k}{\sqrt{\mathbf{a}_k^2 + \mathbf{b}_k^2}} \frac{-\bar{\mathbf{b}}_k}{\sqrt{\bar{\mathbf{a}}_k^2 + \bar{\mathbf{b}}_k^2}} \quad (2.49)$$

$$= -\frac{\bar{\mathbf{a}}_k\Delta\mathbf{b}_k - \bar{\mathbf{b}}_k\Delta\mathbf{a}_k}{\sqrt{(\bar{\mathbf{a}}_k^2 + \bar{\mathbf{b}}_k^2)(\mathbf{a}_k^2 + \mathbf{b}_k^2)}} \quad (2.50)$$

$$= -\frac{\bar{\mathbf{a}}_k\Delta\mathbf{b}_k - \bar{\mathbf{b}}_k\Delta\mathbf{a}_k}{(\bar{\mathbf{a}}_k^2 + \bar{\mathbf{b}}_k^2)\sqrt{1 + \frac{2(\bar{\mathbf{a}}_k\Delta\mathbf{a}_k + \bar{\mathbf{b}}_k\Delta\mathbf{b}_k) + \Delta\mathbf{a}_k^2 + \Delta\mathbf{b}_k^2}{\bar{\mathbf{a}}_k^2 + \bar{\mathbf{b}}_k^2}}} \quad (2.51)$$

$$\approx -\frac{\bar{\mathbf{a}}_k\Delta\mathbf{b}_k - \bar{\mathbf{b}}_k\Delta\mathbf{a}_k}{\bar{\mathbf{a}}_k^2 + \bar{\mathbf{b}}_k^2} . \quad (2.52)$$

Somit gilt

$$\Delta\psi_k \approx -\frac{\bar{\mathbf{a}}_k\Delta\mathbf{b}_k - \bar{\mathbf{b}}_k\Delta\mathbf{a}_k}{\bar{\mathbf{a}}_k^2 + \bar{\mathbf{b}}_k^2} = \frac{\text{Im}[\bar{\mathcal{I}}_k^* \Delta\mathcal{I}_k]}{|\bar{\mathcal{I}}_k|^2} . \quad (2.53)$$

Der Algorithmus zur Orientierungsbestimmung war unter der Annahme abgeleitet worden, dass die zu vergleichenden Bilder bis auf eine Rotation im Wesentlichen identisch sind. Allerdings können auch in einer Umgebung mit statischen Objekten Abweichungen von diesem Idealzustand auftreten, z.B. infolge von Fehlern der abbildenden Optik (kein völlig rotationssymmetrischer Spiegel, Fehljustierung der Kamera, etc.), Verdeckungen durch Kabelverbindung zum externen Computer oder Beleuchtungsvariationen. Daneben können natürlich auch die Positionen, an denen die Bilder aufgenommen wurden, differieren. Zunächst soll der Einfluss einer räumlichen Verschiebung auf die Orientierungsbestimmung untersucht werden. Im Anschluss daran werden nicht-systematische Fehler (Rauschen) diskutiert.

2.4.1 Positionsabweichungen

Mit Hilfe der Gleichungen (1.55) und (1.62) kann ein Bild an einer um $(\Delta x, \Delta y) = \bar{l}(\cos\bar{\varphi}, \sin\bar{\varphi})$ versetzten Position näherungsweise aus dem „Originalbild“ $\bar{I}(\phi)$ berechnet werden:

$$\bar{I}(\phi|\bar{\varphi}, \bar{l}) \approx \bar{I}(\phi) - \frac{\partial\bar{I}(\phi)}{r(\phi)}\bar{l}\sin(\phi - \bar{\varphi}) . \quad (2.54)$$

⁹Ungestörte Größen werden im Folgenden – analog zu Abschnitt 1.5 – durch einen horizontalen Strich ($\bar{}$) gekennzeichnet.

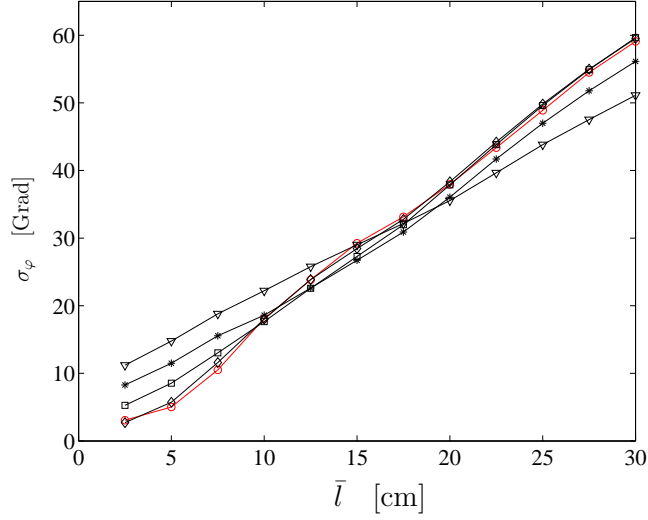


Abbildung 2.5: Zirkulare Standardabweichung der geschätzten relativen Orientierung $\frac{2\pi}{N}s_K$ von der gemessenen Orientierungsdifferenz $\psi_i^1 - \psi_i^2$ in Abhängigkeit vom Abstand der zu vergleichenden Bilder. Für die einzelnen Kurven wurde die Zahl der verwendeten Fourierkoeffizienten variiert: $K = 1$ (∇), $K = 2$ (*), $K = 5$ (\square) und $K = 10$ (\diamond) (Kurven für $K > 10$ entsprechen im Wesentlichen dem Verlauf für $K = 10$). Die rote Kurve zeigt $\sigma_\varphi(\bar{l})$ bei Verwendung der „Standard“-Korrelation, Gleichungen (2.6) und (2.7).

Aus (2.54) folgt mit (2.46) und (2.47), dass die Veränderungen der Koeffizienten $|\Delta\mathbf{a}_k|$, $|\Delta\mathbf{b}_k|$ in erster Näherung linear mit steigendem Abstand vom Zielort zunehmen. Wegen (2.53) ist dies auch für s_k , Gleichung (2.40), und $\bar{s}_K = (\sum \alpha_k s_k) / (\sum \alpha_k)$, Gleichung (2.41), zu erwarten.

Abbildung 2.5 zeigt die zirkulare Standardabweichung [62],

$$\sigma_\varphi(\bar{l}) := \sqrt{-2 \ln R(\bar{l})} , \quad (2.55)$$

$$R(\bar{l}) := \sqrt{\left(\frac{1}{N_I(\bar{l})} \sum_i \cos \Delta\psi_i(\bar{l})\right)^2 + \left(\frac{1}{N_I(\bar{l})} \sum_i \sin \Delta\psi_i(\bar{l})\right)^2} , \quad (2.56)$$

$$\Delta\psi_i(\bar{l}) := \frac{2\pi}{N} \bar{s}_{Ki} - (\psi_i^1 - \psi_i^2) , \quad (2.57)$$

in Abhängigkeit von der Distanz \bar{l} und der Zahl der Fourierkoeffizienten (Breite der Bins ist 2.5 cm, $N_I(\bar{l})$ ist die Zahl der Werte im jeweiligen Bin). ψ_i ist die mit Hilfe des Tracking-Systems (Winkelauflösung $\approx 2^\circ$) gemessene Orientierung der Bilder. Für die Summen in (2.56) wurden alle relativen Orientierungsbestimmungen, die die Bilddatenbank ermöglicht, ausgewertet. Wie zu erkennen ist, wächst der Orientierungsfehler näherungsweise linear mit dem Abstand an, wobei die Steigung leicht mit der Zahl der Fourierkoeffizienten K

zunimmt. Für kleine Abstände dominieren jedoch andere Fehlerquellen, wie z.B. Rauschen (siehe folgender Abschnitt), so dass ein größeres K vorteilhaft ist.

Bestimmt man die relative Orientierung mit Hilfe der „Standard“-Korrelation, so ist deren Standardabweichung in Abhängigkeit vom Abstand mit den Werten für $K = 10$ Fourierkoeffizienten nahezu identisch (die maximale Abweichung der roten Kurve in Abb. 2.5 von der Kurve für $K = 10$ beträgt ca. 1.5°). Dies bedeutet einerseits, dass die Sub-Pixel-Auflösung (siehe Abschnitt 2.2) nur bei sehr kleinen Abständen von Vorteil ist, da sonst der Fehler infolge von Positionsabweichungen dominiert. Andererseits folgt daraus, dass bei Verwendung der schnelleren Fourier-basierten Orientierungsbestimmung bei Positionsabweichungen keine (im Vergleich zum „Standardverfahren“) erhöhten Fehler zu erwarten sind.

2.4.2 Statistische Fehler

Im Folgenden wird mittelwertfreies und unabhängiges Rauschen betrachtet, d.h. für die Grauwerte gilt $I_j = \bar{I}_j + \Delta I_j$, $\langle \Delta I_j \rangle = 0$ und $\langle \Delta I_j \Delta I_k \rangle = \sigma_n^2 \delta_{jk}$. Für die Abweichungen in den Koeffizienten folgt mit (2.46) und (2.47):

$$\langle \Delta \mathbf{a}_k \rangle = \langle \Delta \mathbf{b}_k \rangle = 0, \quad (2.58)$$

$$\langle \Delta \mathbf{a}_0 \Delta \mathbf{a}_l \rangle = \frac{4}{N^2} \sigma_n^2 \sum_j \cos(l\phi_j) = \frac{4}{N} \sigma_n^2 \delta_{0l}, \quad (2.59)$$

$$\langle \Delta \mathbf{a}_k \Delta \mathbf{a}_l \rangle = \frac{4}{N^2} \sigma_n^2 \sum_j \cos(k\phi_j) \cos(l\phi_j) = \frac{2}{N} \sigma_n^2 \delta_{kl} \quad (k \geq 1), \quad (2.60)$$

$$\langle \Delta \mathbf{b}_k \Delta \mathbf{b}_l \rangle = \frac{4}{N^2} \sigma_n^2 \sum_j \sin(k\phi_j) \sin(l\phi_j) = \frac{2}{N} \sigma_n^2 \delta_{kl} \quad (k \geq 1), \quad (2.61)$$

$$\langle \Delta \mathbf{a}_k \Delta \mathbf{b}_l \rangle = \frac{4}{N^2} \sigma_n^2 \sum_j \cos(k\phi_j) \sin(l\phi_j) = 0. \quad (2.62)$$

Dies führt nach Gleichung (2.53) zu einem Fehler in der Phase. Berechnung des Mittelwerts und der (Ko-)Varianz ergibt:

$$\langle \Delta \psi_k \rangle = 0, \quad (2.63)$$

$$\begin{aligned} \langle \Delta \psi_k^2 \rangle &\approx \frac{\bar{\mathbf{b}}_k^2 \langle \Delta \mathbf{a}_k^2 \rangle - 2\bar{\mathbf{a}}_k \bar{\mathbf{b}}_k \langle \Delta \mathbf{a}_k \Delta \mathbf{b}_k \rangle + \bar{\mathbf{a}}_k^2 \langle \Delta \mathbf{b}_k^2 \rangle}{(\bar{\mathbf{a}}_k^2 + \bar{\mathbf{b}}_k^2)^2} \\ &= \frac{2}{N} \frac{\sigma_n^2}{\bar{\mathbf{a}}_k^2 + \bar{\mathbf{b}}_k^2} = \frac{1}{2N} \frac{\sigma_n^2}{|\bar{\mathcal{I}}_k|^2}, \end{aligned} \quad (2.64)$$

$$\langle \Delta \psi_k \Delta \psi_l \rangle = 0 \quad \text{für } k \neq l. \quad (2.65)$$

Für zwei Kamerabilder **I** und **J** mit gleichem Rauschpegel, die am selben Ort in unterschiedlicher Orientierung aufgenommen wurden, d.h. für deren Fourierkoeffizienten $\bar{\mathcal{J}}_k = \bar{\mathcal{I}}_k e^{-i\frac{2\pi}{N}k s}$ gilt, erhält man mit $\omega_k := \frac{2\pi}{N}k$ und $\psi_k = \omega_k s_k$ aus (2.63)–(2.65):

$$\langle \Delta s_k \rangle = 0, \quad \langle \Delta s_k \Delta s_l \rangle \approx \frac{1}{\omega_k^2} \frac{\sigma_n^2}{N |\bar{\mathcal{I}}_k|^2} \delta_{kl}. \quad (2.66)$$

Falls das Frequenzspektrum durch $|\bar{\mathcal{I}}_k|^2 \propto k^{-2}$ approximiert werden kann, dann ist $\langle \Delta s_k^2 \rangle$ näherungsweise unabhängig von k .

Vernachlässigt man in Gleichung (2.32) einen möglichen Fehler in n_k und nimmt folglich $\langle s_k \rangle = s$ an, so gilt

$$\left. \frac{\partial \bar{s}_K}{\partial \alpha_l} \right|_{\{s_k = \langle s_k \rangle\}, \{\alpha_k = \bar{\alpha}_k\}} \stackrel{(2.37)}{=} \left. \frac{(\sum_i \bar{\alpha}_i) s_l - (\sum_i \bar{\alpha}_i s_i)}{(\sum_i \bar{\alpha}_i)^2} \right|_{s_i = s_l = s} = 0. \quad (2.67)$$

Dies vereinfacht die Berechnung von $\langle \Delta \bar{s}_K^2 \rangle$ und es folgt

$$\langle \Delta \bar{s}_K^2 \rangle \approx \frac{\sum_{k=1}^K \bar{\alpha}_k^2 \langle \Delta s_k^2 \rangle}{(\sum_{k=1}^K \bar{\alpha}_k)^2} = \frac{\sigma_n^2 \sum_{k=1}^K |\bar{\mathcal{I}}_k|^4 \omega_k^4 \frac{1}{\omega_k^2 |\bar{\mathcal{I}}_k|^2}}{N (\sum_{l=k}^K |\bar{\mathcal{I}}_k|^2 \omega_k^2)^2} \quad (2.68)$$

$$= \frac{\sigma_n^2}{N \sum_{k=1}^K |\bar{\mathcal{I}}_k|^2 \omega_k^2}. \quad (2.69)$$

Falls $|\bar{\mathcal{I}}_k|^2 \omega_k^2 \approx \text{const}$ (unabhängig von k) gilt, erhält man

$$\langle \Delta \bar{s}_K^2 \rangle \approx \frac{\langle \Delta s_1^2 \rangle}{K} \propto \frac{\sigma_n^2}{NK}. \quad (2.70)$$

Der durch Rauschen verursachte Fehler nimmt also mit zunehmender Zahl K von Fourierkoeffizienten ab. Dies ist in Abbildung 2.5 für kleine Abstände zu erkennen.

2.5 Fourier-basiertes Heimfinden

In diesem Abschnitt wird gezeigt, wie ein effizientes Heimfinden mit Hilfe der Fouriertransformation realisiert werden kann. Diese ermöglicht eine Beschränkung auf wenige Koeffizienten und eine schnelle Orientierungsbestimmung durch Verwendung des in 2.1.5 beschriebenen Algorithmus. Zusätzlich ist von Vorteil, dass die Ableitung der Bildfunktion leicht aus der interpolierenden Funktion bestimmt werden kann.

2.5.1 Berechnung erwarteter Bilder

Unter der Annahme konstanter Abstände $r(\phi) = R$ kann das Bild, das bei einer Translation der Länge $l \ll R$ in Richtung φ zu erwartenden ist, nach Gleichung (1.46) mit $x := \frac{l}{R} \cos\varphi$ und $y := \frac{l}{R} \sin\varphi$ durch

$$I^e(\phi|x, y) \approx I(\phi - x \sin\phi + y \cos\phi) \quad (2.71)$$

aus dem aktuellen Bild $I(\phi)$ berechnet werden. Für $x, y \ll 1$ kann (2.71) in x und y linearisiert werden:

$$I^e(\phi|x, y) \approx I(\phi) - (x \sin\phi - y \cos\phi) \partial I(\phi) . \quad (2.72)$$

Wird das aktuelle Bild unter Verwendung der ersten K Summanden genähert, d.h.

$$I(\phi) \approx \frac{1}{2} \mathbf{a}_0 + \sum_{k=1}^K \mathbf{a}_k \cos(k\phi) + \mathbf{b}_k \sin(k\phi) , \quad (2.73)$$

so folgt für die Ableitung

$$\partial I(\phi) := \frac{\partial}{\partial \phi} I(\phi) \approx \sum_{k=1}^K \mathbf{b}_k k \cos(k\phi) - \mathbf{a}_k k \sin(k\phi) . \quad (2.74)$$

Setzt man (2.74) und (2.73) in Gleichung (2.72) ein und vergleicht dies mit der Sinus-Cosinus-Form des erwarteten Bildes,

$$I^e(\phi|x, y) \approx \frac{1}{2} \mathbf{a}_0^e(x, y) + \sum_{k=1}^K \mathbf{a}_k^e(x, y) \cos(k\phi) + \mathbf{b}_k^e(x, y) \sin(k\phi) , \quad (2.75)$$

so folgt – wie in Anhang A.7 gezeigt wird – für die Koeffizienten:

$$\mathbf{a}_k^e(x, y) = \mathbf{a}_k + \mathbf{a}_k^x x + \mathbf{a}_k^y y , \quad \mathbf{b}_k^e(x, y) = \mathbf{b}_k + \mathbf{b}_k^x x + \mathbf{b}_k^y y , \quad (2.76)$$

$$\begin{aligned} \mathbf{a}_k^x &:= \frac{1}{2} [-(k-1) \mathbf{a}_{k-1} + (k+1) \mathbf{a}_{k+1}] , \\ \mathbf{a}_k^y &:= \frac{1}{2} [(k-1) \mathbf{b}_{k-1} + (k+1) \mathbf{b}_{k+1}] , \\ \mathbf{b}_k^x &:= \frac{1}{2} [-(k-1) \mathbf{b}_{k-1} + (k+1) \mathbf{b}_{k+1}] , \\ \mathbf{b}_k^y &:= \frac{1}{2} [-(k-1) \mathbf{a}_{k-1} - (k+1) \mathbf{a}_{k+1}] . \end{aligned} \quad (2.77)$$

2.5.2 Bestimmung des Heimvektors

Wird die Bildähnlichkeit durch die SSD bestimmt, so gilt mit (2.14) und (2.44) für die Ähnlichkeit zweier Bilder \mathbf{I}^A und $\mathbf{I}^B \in \mathbb{R}^N$:

$$\begin{aligned} & \sum_{\phi} (I^A(\phi) - I^B(\phi))^2 \\ & \approx N \sum_{k=-K}^K |\mathcal{I}_k^A - \mathcal{I}_k^B|^2 \end{aligned} \quad (2.78)$$

$$= N |\mathcal{I}_0^A - \mathcal{I}_0^B|^2 + 2N \sum_{k=1}^K |\mathcal{I}_k^A - \mathcal{I}_k^B|^2 \quad (2.79)$$

$$= \frac{N}{4} (\mathbf{a}_0^A - \mathbf{a}_0^B)^2 + \frac{N}{2} \sum_{k=1}^K (\mathbf{a}_k^A - \mathbf{a}_k^B)^2 + (\mathbf{b}_k^A - \mathbf{b}_k^B)^2 . \quad (2.80)$$

Eine Schätzung des Heimvektors $(x, y)^\top$ und der relativen Orientierung s zwischen dem aktuellen Bild und dem Bild an der Zielposition kann durch Minimieren der Fehlerfunktion

$$E_I(x, y, s) := \frac{N}{2} \sum_{k=1}^K (\mathbf{a}_k^h(s) - \mathbf{a}_k^e(x, y))^2 + (\mathbf{b}_k^h(s) - \mathbf{b}_k^e(x, y))^2 \quad (2.81)$$

berechnet werden. In (2.81) wurde der Mittelwert von den Bildern abgezogen, so dass $\mathbf{a}_0^h = \mathbf{a}_0^e = 0$ gilt. $\mathbf{a}_k^e(x, y)$ und $\mathbf{b}_k^e(x, y)$ werden gemäß (2.76) aus den Koeffizienten des aktuellen Bildes berechnet. Für die Koeffizienten des um s gedrehten Bildes an der Zielposition folgt aus (2.13) mit $\omega_k = \frac{2\pi}{N}$:

$$\mathbf{a}_k^h(s) = \cos(\omega_k s) \mathbf{a}_k^h + \sin(\omega_k s) \mathbf{b}_k^h , \quad (2.82)$$

$$\mathbf{b}_k^h(s) = \cos(\omega_k s) \mathbf{b}_k^h - \sin(\omega_k s) \mathbf{a}_k^h . \quad (2.83)$$

Existiert bereits eine Schätzung s_0 für die relative Orientierung, d.h. $s = s_0 + \Delta s$, so gilt mit $\sin(x_0 + \Delta x) \approx \sin x_0 + \Delta x \cos x_0$ und $\cos(x_0 + \Delta x) \approx \cos x_0 - \Delta x \sin x_0$ ($\Delta x \ll 1$):

$$\mathbf{a}_k^h(s) \approx \mathbf{a}_k^h(s_0) + \mathbf{b}_k^h(s_0) \omega_k \Delta s , \quad (2.84)$$

$$\mathbf{b}_k^h(s) \approx \mathbf{b}_k^h(s_0) - \mathbf{a}_k^h(s_0) \omega_k \Delta s . \quad (2.85)$$

Man beachte, dass die Bedingung $\omega_k \Delta s = k \frac{2\pi}{N} \Delta s \ll 1$ die Verwendung hoher Frequenzen einschränkt. Einsetzen von (2.76), (2.84) und (2.85) in (2.81) ergibt

$$E_I(x, y, s_0 + \Delta s) \approx E_1(s_0) + E_2(x, y, \Delta s; s_0) , \quad (2.86)$$

$$E_1(s_0) := \frac{N}{2} \sum_k [\mathbf{a}_k^h(s_0) - \mathbf{a}_k]^2 + [\mathbf{b}_k^h(s_0) - \mathbf{b}_k]^2, \quad (2.87)$$

$$\begin{aligned} E_2(x, y, \Delta s; s_0) &:= \frac{N}{2} \sum_k 2[\mathbf{a}_k^h(s_0) - \mathbf{a}_k][\omega_k \mathbf{b}_k^h(s_0) \Delta s - \mathbf{a}_k^x x - \mathbf{a}_k^y y] \\ &\quad + 2[\mathbf{b}_k^h(s_0) - \mathbf{b}_k][-\omega_k \mathbf{a}_k^h(s_0) \Delta s - \mathbf{b}_k^x x - \mathbf{b}_k^y y] \\ &\quad + [\omega_k \mathbf{b}_k^h(s_0) \Delta s - \mathbf{a}_k^x x - \mathbf{a}_k^y y]^2 \\ &\quad + [-\omega_k \mathbf{a}_k^h(s_0) \Delta s - \mathbf{b}_k^x x - \mathbf{b}_k^y y]^2. \end{aligned} \quad (2.88)$$

Mit $\mathbf{a}_k = \sqrt{\mathbf{a}_k^2 + \mathbf{b}_k^2} \cos \psi_k$ und $\mathbf{b}_k = -\sqrt{\mathbf{a}_k^2 + \mathbf{b}_k^2} \sin \psi_k$ kann (2.87) umgeformt werden zu

$$\begin{aligned} E_1(s_0) &= \frac{N}{2} \sum_k \mathbf{a}_k^h(s_0)^2 + \mathbf{b}_k^h(s_0)^2 + \mathbf{a}_k^2 + \mathbf{b}_k^2 - 2(\mathbf{a}_k^h(s_0)\mathbf{a}_k + \mathbf{b}_k^h(s_0)\mathbf{b}_k) \end{aligned} \quad (2.89)$$

$$\begin{aligned} &= \frac{N}{2} \left(\sum_k (\mathbf{a}_k^h)^2 + (\mathbf{b}_k^h)^2 + \mathbf{a}_k^2 + \mathbf{b}_k^2 \right. \\ &\quad \left. - 2 \sum_k \sqrt{(\mathbf{a}_k^h)^2 + (\mathbf{b}_k^h)^2} \sqrt{\mathbf{a}_k^2 + \mathbf{b}_k^2} \cos(\psi_k^h - \psi_k + \frac{2\pi}{N} k s_0) \right). \end{aligned} \quad (2.90)$$

Gleichung (2.90) entspricht (2.25) mit $\mathcal{I}_k \rightarrow \frac{1}{2}(\mathbf{a}_k^h - i\mathbf{b}_k^h)$ und $\mathcal{J}_k \rightarrow \frac{1}{2}(\mathbf{a}_k - i\mathbf{b}_k)$.

Mit diesen Umformungen kann die Optimierung von (2.81) effizient durchgeführt werden:

1. Minimiere $E_1(s_0)$ (bezüglich s_0) mit Hilfe des Algorithmus zur Orientierungsbestimmung, Abschnitt 2.1.5. Dies resultiert in s_0^{opt} .
2. Minimiere $E_2(x, y, \Delta s; s_0^{\text{opt}})$ bezüglich x, y und Δs . Da (2.88) nur Terme bis zur 2. Ordnung in x, y und Δs enthält, führt die Bestimmung der Nullstelle des Gradienten von E_2 auf ein lineares Gleichungssystem.¹⁰ Dessen Lösung, die lediglich die Invertierung einer 3×3 -Matrix erfordert, ergibt eine Schätzung des Heimvektors $(x^{\text{opt}}, y^{\text{opt}})^\top$ und Δs^{opt} . Für die Orientierungsdifferenz folgt $s^{\text{opt}} = s_0^{\text{opt}} + \Delta s^{\text{opt}}$.

Das Verfahren kann iteriert werden, indem in einem zusätzlichen Schritt

$$\mathbf{a}_k + \mathbf{a}_k^x \bar{x}^{\text{opt}} + \mathbf{a}_k^y \bar{y}^{\text{opt}} \rightarrow \mathbf{a}_k, \quad (2.91)$$

$$\mathbf{b}_k + \mathbf{b}_k^x \bar{x}^{\text{opt}} + \mathbf{b}_k^y \bar{y}^{\text{opt}} \rightarrow \mathbf{b}_k \quad (2.92)$$

bestimmt wird. Dabei sind $\bar{x}^{\text{opt}} := \sum_{i=1}^{N_{\text{it}}-1} x_i^{\text{opt}}$ und $\bar{y}^{\text{opt}} := \sum_{i=1}^{N_{\text{it}}-1} y_i^{\text{opt}}$ die Summen aller in den vorangegangenen Durchläufen berechneten Verschiebungen.

¹⁰Alternativ kann eine numerische Optimierung von $E_I(x, y, s_0^{\text{opt}} + \Delta s)$, Gleichung (2.81), durch Gradientenabstieg in $x, y, \Delta s$ vorgenommen werden.

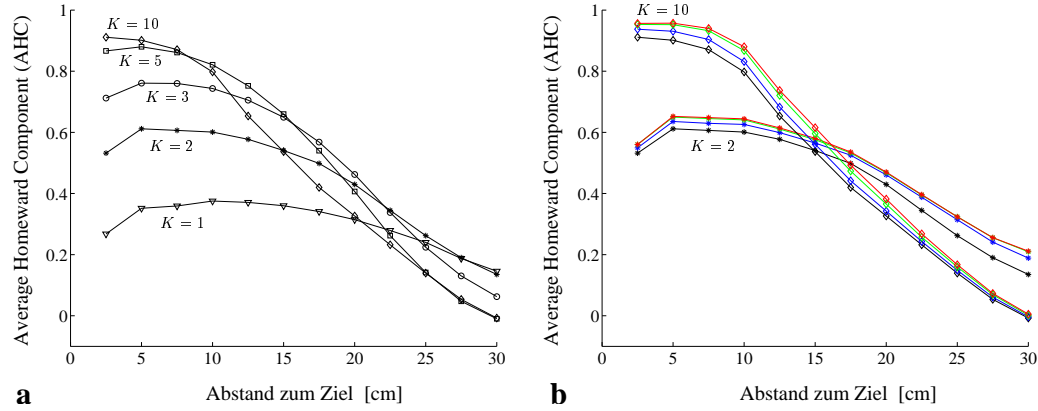


Abbildung 2.6: Average Homeward Component des berechneten Heimvektors in Abhängigkeit vom Abstand zum Zielort (Breite der Bins ist 2.5 cm). **a:** Abhängigkeit von der Zahl der Fourierkoeffizienten $K = 1$ (∇), $K = 2$ (*), $K = 3$ (\circ), $K = 5$ (\square) und $K = 10$ (\diamond). **b:** Abhängigkeit von der Zahl der Durchläufe (für $K = 2$ und $K = 10$) $N_{it} = 1$ (schwarz), $N_{it} = 2$ (blau), $N_{it} = 5$ (grün) und $N_{it} = 10$ (rot). (Kurven für $K > 10$ entsprechen denen für $K = 10$ oder besitzen eine geringere AHC und sind deshalb nicht dargestellt).

2.5.3 Resultate

Das beschriebene Verfahren wurde mit Hilfe der Bilddatenbank der Modellhausarena getestet.

In Abbildung 2.6 ist die „Average Homeward Component“ (AHC) des berechneten Heimvektors als Funktion des Abstands vom Zielort dargestellt. Dabei wurde sowohl die Zahl der Fourierkoeffizienten K als auch der Iterationen N_{it} variiert. Für Abstände $\bar{l} \lesssim 10$ cm erlaubt ein größeres K eine genauere Bestimmung der Richtung zum Zielort, da benachbarte Bilder besser unterschieden werden können. Mit $K = 1, 2$ berechnete Heimvektoren besitzen jedoch für große Abstände eine größere AHC. In beiden Fällen kann durch wenige Iterationen das Ergebnis noch etwas verbessert werden. Der stärkere Abfall der AHC für großes K war zu erwarten, da die Linearisierung sowohl der trigonometrischen Funktionen in (2.84) und (2.85) als auch der Koeffizienten \mathbf{a}_k^e und \mathbf{b}_k^e in (2.76), und (2.77) mit steigendem k problematischer wird.

Die Näherung (2.72) dürfte hauptsächlich dafür verantwortlich sein, dass die AHC für große Abstände zum Zielort deutlich kleiner ausfällt als für das in Abschnitt 1.4.1 beschriebene bild-basierte Heimfinden. Dies ist daran zu erkennen, dass $\Delta\phi \propto l/R$ bereits für $l/R \approx 0.1$ die Größe eines Pixels

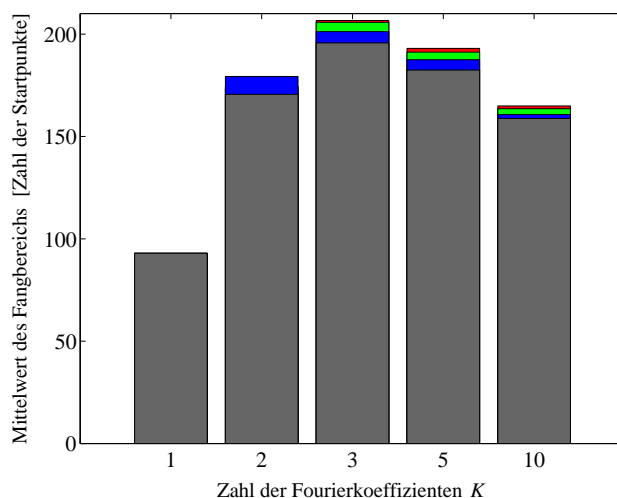


Abbildung 2.7: Mittlere Größe des Fangbereichs in Abhängigkeit von der Zahl der Fourierkoeffizienten K und der Zahl der Durchläufe $N_{it} = 1$ (grau), $N_{it} = 2$ (blau), $N_{it} = 5$ (grün) und $N_{it} = 10$ (rot). Nur die Iterationen, die zu einer Vergrößerung des Fangbereichs führen, sind dargestellt.

erreicht:

$$\frac{l}{R} \frac{N}{2\pi} \geq 1 \iff \frac{l}{R} \geq \frac{2\pi}{72} \approx 0.087. \quad (2.93)$$

Zusätzlich wird die Schätzung der Orientierung s_0 im 1. Schritt mit zunehmendem Abstand \bar{l} ungenauer (siehe Abschnitt 2.4.1). Im Nahbereich¹¹ ($\bar{l} \lesssim 5$ cm) ermöglicht jedoch das Fourier-basierte Verfahren (mit $K = 10$), das keine diskrete Abtastung des Parameterraums erfordert, eine deutlich genauere Bestimmung der Richtung zum Zielort: $\text{AHC}(\bar{l} = 2.5 \text{ cm}) \approx 0.96$ im Vergleich zu $\text{AHC}(\bar{l} = 2.5 \text{ cm}) \approx 0.71$ in Abb. 1.12 b.

Abbildung 2.7 zeigt, wie sich die Zahl der verwendeten Fourierkoeffizienten und Iterationen auf die Größe des Fangbereichs (Definition und Berechnung wie in Kapitel 1 mit $l_{CA} = 10$ cm und $R_{\text{typ}} = 45$ cm) auswirkt. Das Maximum bei $K = 3$ ist in Hinblick auf Abbildung 2.6 als Kompromiss zu sehen, der für alle Abstände eine hinreichend große AHC ermöglicht. Der im Vergleich zum Bild-basierten Heimfinden aus Abschnitt 1.4.1 deutlich kleinere Fangbereich (≈ 200 im Vergleich zu ≈ 300 in Abb. 1.13) ist – wie bereits bei den Resultaten für die AHC erwähnt – hauptsächlich auf die Linearisierung (2.72) zurückzuführen.

¹¹Bei einem größeren Landmarkenabstand z.B. in offenen Umgebungen kann diesem ein durchaus beachtliches Areal entsprechen.

Neben der deutlich schnelleren Berechnung der Heimvektoren ist der geringe Speicherbedarf ein Vorteil des vorgestellten Fourier-basierten Heimfindens: Für $K = 3$ werden als Repräsentation eines Ortes nur 6 Zahlen ($\mathbf{a}_k, \mathbf{b}_k, k = 1, 2, 3$) benötigt.

Abbildung 2.8 zeigt Beispiele von Heimvektoren für unterschiedliche Zielpositionen in der Modellhausarena. Neben dem jeweiligen globalen Optimum der Bildkorrelation existieren zusätzliche lokale Maxima, die zum Teil höheren Korrelationswerte besitzen als Positionen, die zum Fangbereich der Zielposition gehören. Dies erschwert die Wahl einer geeigneten Schwelle für die Bildähnlichkeit, ab der ein „Homing“-Versuch gestartet wird: Entweder wird für eine hohe Schwelle ein zu großer Bereich ausgeschlossen oder bei einer niedrigen Schwelle (falls keine zusätzliche Information über die aktuelle Position vorhanden ist) unter Umständen ein falsches lokales Maximum angefahren.¹² Im letzteren Fall kann dieser Irrtum allerdings durch Vergleich des Bildes am vermeintlichen Zielort mit dem gesuchten Bild erkannt werden.

¹²Dies ist ein generelles Problem und somit keine Besonderheit des Fourier-basierten Heimfindens.

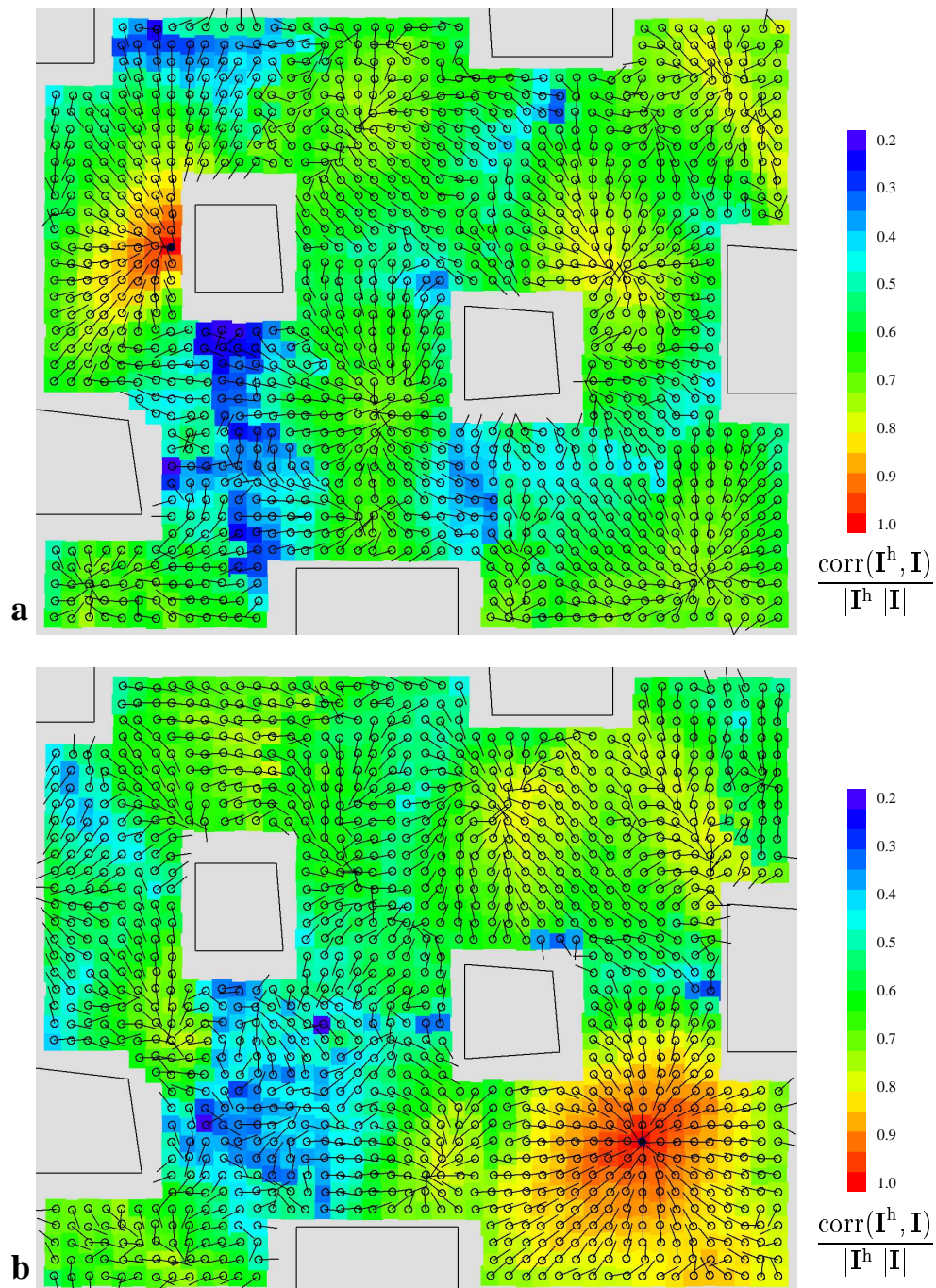


Abbildung 2.8: Heimvektoren und normierte Korrelation (farbcodiert), berechnet mit Fourier-transformierten Panoramabildern der Datenbank (Zahl der verwendeten Fourierkoeffizienten $K = 10$, 3 Iterationen). Die durch den schwarzen Punkt markierte Zielposition befindet sich in (a) in der Nähe eines Modellhauses (Abstand ≈ 6 cm) und in (b) in einem „offenen“ Bereich der Modellhausarena.

Kapitel 3

Panoramische Bildsensoren

In diesem Kapitel wird die Konstruktion von Spiegelflächen für die Aufnahme panoramischer Bilder diskutiert. Der Schwerpunkt liegt dabei auf axialsymmetrischen Spiegelflächen mit konstanter vertikaler Winkelvergrößerung. Als Spezialfall wird ein panoramischer Stereosensor für einen mobilen Kleinroboter vorgestellt, der eine größere Stereobasis und bessere Abbildungsqualität (und allerdings auch größere Abmessungen) besitzt als der in Abschnitt 1.1 beschriebene. Mit Hilfe einer mathematischen Beschreibung in Kugelkoordinaten werden Abbildungsfehler untersucht, die aufgrund der gewählten Geometrie des Spiegels oder infolge von Fehlstellungen der Kamera relativ zum Koordinatensystem der Spiegelfläche auftreten.

3.1 Mathematische Beschreibung rotations-symmetrischer Spiegelflächen

Das sphärische Koordinatensystem mit dem Kameraknotenpunkt als Ursprung sei definiert durch

$$\mathbf{e}(\theta, \varphi) := \begin{pmatrix} \sin \theta \cos \varphi \\ \sin \theta \sin \varphi \\ \cos \theta \end{pmatrix}, \quad (3.1)$$

$$\mathbf{e}_\theta(\theta, \varphi) := \frac{\partial}{\partial \theta} \mathbf{e}(\theta, \varphi) = \begin{pmatrix} \cos \theta \cos \varphi \\ \cos \theta \sin \varphi \\ -\sin \theta \end{pmatrix}, \quad (3.2)$$

$$\mathbf{e}_\varphi(\theta, \varphi) := \frac{1}{\sin \theta} \frac{\partial}{\partial \varphi} \mathbf{e}(\theta, \varphi) = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}. \quad (3.3)$$

Im Folgenden wird angenommen, dass rotationssymmetrische Spiegelflächen durch $\mathbf{x}(\theta, \varphi) := r(\theta) \mathbf{e}(\theta, \varphi)$ beschrieben werden. Da nur Bereiche von Inter-

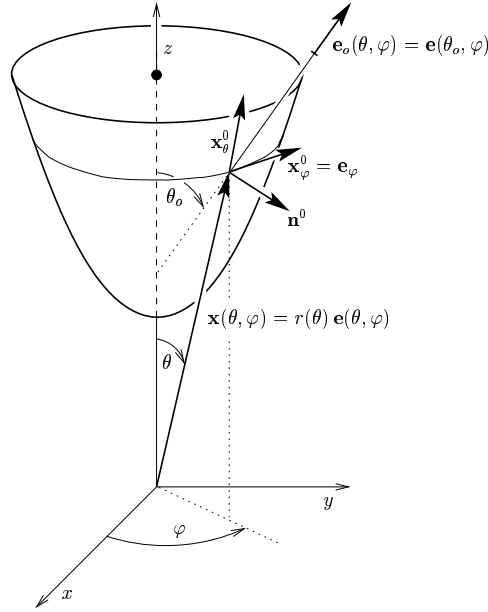


Abbildung 3.1: Koordinatensystem für die Beschreibung der optischen Abbildung an einer (bezüglich der z -Achse) rotationsymmetrischen Spiegelfläche. Der Knotenpunkt der Kamera befindet sich im Ursprung.

esse sind, die vom Kameraknotenpunkt aus einsehbar sind, ist $r(\theta)$ in jedem Fall eindeutig. Wie in Abbildung 3.1 dargestellt, wird ein vom Ursprung in Richtung $\mathbf{e}(\theta, \varphi)$ ausgehender Strahl am Punkt $\mathbf{x}(\theta, \varphi)$ der Spiegelfläche reflektiert. Das lokale orthonormale Koordinatensystem wird definiert durch \mathbf{x}_θ^0 , \mathbf{x}_φ^0 und den Normalenvektor \mathbf{n}^0 (der hochgestellte Index '0' kennzeichnet Vektoren mit Norm 1). Es gilt

$$\mathbf{x}_\theta(\theta, \varphi) := \frac{\partial}{\partial \theta} \mathbf{x}(\theta, \varphi) = r(\theta) \mathbf{e}_\theta(\theta, \varphi) + r'(\theta) \mathbf{e}(\theta, \varphi), \quad (3.4)$$

$$\mathbf{x}_\theta^0(\theta, \varphi) := \frac{\mathbf{x}_\theta(\theta, \varphi)}{|\mathbf{x}_\theta(\theta, \varphi)|} = \frac{1}{\sqrt{r(\theta)^2 + r'(\theta)^2}} \mathbf{x}_\theta(\theta, \varphi), \quad (3.5)$$

$$\mathbf{x}_\varphi(\theta, \varphi) := \frac{\partial}{\partial \varphi} \mathbf{x}(\theta, \varphi) = r(\theta) \sin \theta \mathbf{e}_\varphi(\theta, \varphi), \quad (3.6)$$

$$\mathbf{x}_\varphi^0(\theta, \varphi) := \frac{\mathbf{x}_\varphi(\theta, \varphi)}{|\mathbf{x}_\varphi(\theta, \varphi)|} = \mathbf{e}_\varphi(\theta, \varphi), \quad (3.7)$$

$$\mathbf{n}^0(\theta, \varphi) := \mathbf{x}_\varphi^0(\theta, \varphi) \times \mathbf{x}_\theta^0(\theta, \varphi) = -\frac{r(\theta) \mathbf{e}(\theta, \varphi) - r'(\theta) \mathbf{e}_\theta(\theta, \varphi)}{\sqrt{r(\theta)^2 + r'(\theta)^2}} \quad (3.8)$$

$$= -\frac{1}{\sqrt{r(\theta)^2 + r'(\theta)^2}} \begin{pmatrix} (r(\theta) \sin \theta - r'(\theta) \cos \theta) \cos \varphi \\ (r(\theta) \sin \theta - r'(\theta) \cos \theta) \sin \varphi \\ r(\theta) \cos \theta + r'(\theta) \sin \theta \end{pmatrix}. \quad (3.9)$$

Die Richtung des reflektierten Strahls berechnet sich mit Hilfe des Reflexionsgesetzes zu

$$\mathbf{e}_o(\theta, \varphi) = \mathbf{e}(\theta, \varphi) - 2(\mathbf{e}(\theta, \varphi) \mathbf{n}^0(\theta, \varphi)) \mathbf{n}^0(\theta, \varphi) \quad (3.10)$$

$$= \begin{pmatrix} b(\theta) \cos \varphi \\ b(\theta) \sin \varphi \\ a(\theta) \end{pmatrix} = \mathbf{e}(\theta_o, \varphi) , \quad (3.11)$$

$$\sin \theta_o = b(\theta) := \frac{-r(\theta)^2 \sin \theta + 2r(\theta)r'(\theta) \cos \theta + r'(\theta)^2 \sin \theta}{r(\theta)^2 + r'(\theta)^2} , \quad (3.12)$$

$$\cos \theta_o = a(\theta) := \frac{-r(\theta)^2 \cos \theta - 2r(\theta)r'(\theta) \sin \theta + r'(\theta)^2 \cos \theta}{r(\theta)^2 + r'(\theta)^2} . \quad (3.13)$$

Für die Richtungswinkel folgt damit:

$$\varphi_o = \varphi , \quad (3.14)$$

$$\theta_o = \Theta_o(\theta) := \operatorname{arccot}\left(\frac{a(\theta)}{b(\theta)}\right) \in [0, \pi] . \quad (3.15)$$

3.2 Fehlstellungen der Kamera

Befindet sich die Kamera nicht im Ursprung des Spiegelkoordinatensystems oder ist ihre Achse nicht parallel zur Symmetrieachse des Spiegels ausgerichtet, so wird die Abbildung nicht mehr exakt durch Gleichungen (3.12)–(3.15) beschrieben. Korrekturterme für kleine Störungen werden im Folgenden berechnet.

3.2.1 Verschiebung des Kameraknotenpunkts

Abweichungen der Kameraposition vom Ursprung führen zu Abbildungsfehlern. Wird die x -Achse entlang der zur Symmetrieachse orthogonalen Verschiebung der Kamera gelegt (d.h. $\Delta y = 0$), so trifft ein von der Kamera in Richtung $\mathbf{e}(\theta_i, \varphi_i)$ auslaufender Strahl die Spiegelfläche in einem Punkt, der durch die Koordinaten θ, φ definiert ist:

$$\mathbf{x}(\theta, \varphi) = \begin{pmatrix} \Delta x \\ 0 \\ \Delta z \end{pmatrix} + \lambda \begin{pmatrix} \sin \theta_i \cos \varphi_i \\ \sin \theta_i \sin \varphi_i \\ \cos \theta_i \end{pmatrix} , \quad \lambda \in \mathbb{R} . \quad (3.16)$$

Für $\Delta x = \Delta z = 0$ gilt $\theta = \theta_i, \varphi = \varphi_i$. In linearer Näherung, d.h.

$$\mathbf{x}(\theta, \varphi) = \mathbf{x}(\theta_i + \Delta\theta, \varphi_i + \Delta\varphi) \quad (3.17)$$

$$\approx \mathbf{x}(\theta_i, \varphi_i) + \mathbf{x}_\theta(\theta_i, \varphi_i)\Delta\theta + \mathbf{x}_\varphi(\theta_i, \varphi_i)\Delta\varphi , \quad (3.18)$$

erhält man

$$\Delta\theta = \Delta\Theta(\theta_i, \varphi_i) \approx \frac{\cos\theta_i \cos\varphi_i}{r(\theta_i)} \Delta x - \frac{\sin\theta_i}{r(\theta_i)} \Delta z, \quad (3.19)$$

$$\Delta\varphi = \Delta\Phi(\theta_i, \varphi_i) \approx -\frac{\sin\varphi_i}{r(\theta_i) \sin\theta_i} \Delta x. \quad (3.20)$$

Aus (3.20) ist ersichtlich, dass die lineare Näherung für kleine θ_i problematisch ist.

Für die Richtung des reflektierten Strahls folgt¹

$$\begin{aligned} \tilde{\mathbf{e}}_o(\theta_i, \varphi_i) &= \mathbf{e}(\theta_i, \varphi_i) - 2(\mathbf{e}(\theta_i, \varphi_i) \mathbf{n}^0(\theta_i + \Delta\theta, \varphi_i + \Delta\varphi)) \\ &\quad \times \mathbf{n}^0(\theta_i + \Delta\theta, \varphi_i + \Delta\varphi) \end{aligned} \quad (3.21)$$

$$\begin{aligned} &\approx \mathbf{e}_i - 2 \left((\mathbf{e}_i \mathbf{n}_i^0) \mathbf{n}_i^0 + \left(\mathbf{e}_i \frac{\partial \mathbf{n}^0}{\partial \theta} \Big|_i \right) \mathbf{n}_i^0 \Delta\theta + \left(\mathbf{e}_i \frac{\partial \mathbf{n}^0}{\partial \varphi} \Big|_i \right) \mathbf{n}_i^0 \Delta\varphi \right. \\ &\quad \left. + (\mathbf{e}_i \mathbf{n}_i^0) \frac{\partial \mathbf{n}^0}{\partial \theta} \Big|_i \Delta\theta + (\mathbf{e}_i \mathbf{n}_i^0) \frac{\partial \mathbf{n}^0}{\partial \varphi} \Big|_i \Delta\varphi \right) \end{aligned} \quad (3.22)$$

$$\begin{aligned} &= \mathbf{e}_o(\theta_i, \varphi_i) - 2 \frac{\partial}{\partial \theta} \left((\mathbf{e}_i \mathbf{n}^0(\theta, \varphi)) \mathbf{n}^0(\theta, \varphi) \right) \Big|_i \Delta\theta \\ &\quad - 2 \frac{\partial}{\partial \varphi} \left((\mathbf{e}_i \mathbf{n}^0(\theta, \varphi)) \mathbf{n}^0(\theta, \varphi) \right) \Big|_i \Delta\varphi, \end{aligned} \quad (3.23)$$

wobei zur Abkürzung $i := (\theta_i, \varphi_i)$ verwendet wurde. Einsetzen von (3.9) ergibt

$$\begin{aligned} \tilde{\mathbf{e}}_o(\theta_i, \varphi_i) &\approx \begin{pmatrix} b(\theta_i) \cos\varphi_i \\ b(\theta_i) \sin\varphi_i \\ a(\theta_i) \end{pmatrix} + 2\eta(\theta_i) \begin{pmatrix} a(\theta_i) \cos\varphi_i \\ a(\theta_i) \sin\varphi_i \\ -b(\theta_i) \end{pmatrix} \Delta\theta \\ &\quad - 2\zeta(\theta_i) b(\theta_i) \begin{pmatrix} -\sin\varphi_i \\ \cos\varphi_i \\ 0 \end{pmatrix} \Delta\varphi, \end{aligned} \quad (3.24)$$

$$\eta(\theta) := \frac{r(\theta)^2 + 2r'(\theta)^2 - r(\theta)r''(\theta)}{r(\theta)^2 + r'(\theta)^2}, \quad (3.25)$$

$$\zeta(\theta) := \frac{r(\theta)^2 \sin\theta - r(\theta)r'(\theta) \cos\theta}{-r(\theta)^2 \sin\theta + 2r(\theta)r'(\theta) \cos\theta + r'(\theta)^2 \sin\theta}. \quad (3.26)$$

$b(\theta)$ und $a(\theta)$ wurden bereits in (3.12) und (3.13) definiert. Die Richtungswinkel des reflektierten Strahls,

$$\tilde{\theta}_o := \theta_o + \Delta\theta_o = \Theta_o(\theta_i) + \Delta\Theta_o(\theta_i, \varphi_i), \quad (3.27)$$

$$\tilde{\varphi}_o := \varphi_o + \Delta\varphi_o = \varphi_i + \Delta\Phi_o(\theta_i, \varphi_i), \quad (3.28)$$

¹Größen, die infolge von Kamerafehlstellungen Abweichungen von den Sollwerten haben, werden im Folgenden durch eine Tilde ($\tilde{}$) gekennzeichnet.

können durch Vergleich von (3.24) mit

$$\mathbf{e}(\tilde{\theta}_o, \tilde{\varphi}_o) = \mathbf{e}(\Theta_o(\theta_i) + \Delta\theta_o, \varphi_i + \Delta\varphi_o) \quad (3.29)$$

$$\begin{aligned} &\approx \mathbf{e}(\Theta_o(\theta_i), \varphi_i) + \mathbf{e}_\theta(\Theta_o(\theta_i), \varphi_i)\Delta\theta_o \\ &\quad + \sin\Theta_o(\theta_i) \mathbf{e}_\varphi(\Theta_o(\theta_i), \varphi_i)\Delta\varphi_o \end{aligned} \quad (3.30)$$

bestimmt werden. Man erhält:

$$\Delta\theta_o = \Delta\Theta_o(\theta_i, \varphi_i) \approx 2\eta(\theta_i)\Delta\Theta(\theta_i, \varphi_i) \quad (3.31)$$

$$\approx 2\eta(\theta_i) \frac{\cos\theta_i \cos\varphi_i}{r(\theta_i)} \Delta x - 2\eta(\theta_i) \frac{\sin\theta_i}{r(\theta_i)} \Delta z, \quad (3.32)$$

$$\Delta\varphi_o = \Delta\Phi_o(\theta_i, \varphi_i) \approx -2\zeta(\theta_i)\Delta\Phi(\theta_i, \varphi_i) \quad (3.33)$$

$$\approx 2\zeta(\theta_i) \frac{\sin\varphi_i}{r(\theta_i) \sin\theta_i} \Delta x. \quad (3.34)$$

Gibt man die Richtungswinkel θ_o und φ_o vor, beispielsweise durch ein Kalibrierungsmuster (siehe Abschnitt 3.4.2), so sind die Kamerawinkel $\tilde{\theta}_i$ und $\tilde{\varphi}_i$ in Abhängigkeit von θ_o und φ_o von Interesse. Diese erhält man durch Auflösen der Gleichungen

$$\Theta_o(\tilde{\theta}_i) + \Delta\Theta_o(\tilde{\theta}_i, \tilde{\varphi}_i) = \theta_o, \quad (3.35)$$

$$\tilde{\varphi}_i + \Delta\Phi_o(\tilde{\theta}_i, \tilde{\varphi}_i) = \varphi_o \quad (3.36)$$

nach $\tilde{\theta}_i$ und $\tilde{\varphi}_i$. (3.35) und (3.36) folgen aus (3.27) und (3.28) durch die Ersetzungen $(\tilde{\theta}_o, \tilde{\varphi}_o) \rightarrow (\theta_o, \varphi_o)$ und $(\theta_i, \varphi_i) \rightarrow (\tilde{\theta}_i, \tilde{\varphi}_i)$. Kann aus (3.15) die Umkehrfunktion Θ_o^{-1} bestimmt werden, so folgt mit $\tilde{\theta}_i = \theta_i + \Delta\theta_i = \Theta_o^{-1}(\theta_o) + \Delta\Theta_i(\theta_o, \varphi_o)$ und $\tilde{\varphi}_i = \varphi_i + \Delta\varphi_i = \varphi_o + \Delta\Phi_i(\theta_o, \varphi_o)$ nach Linearisierung,

$$\theta_o = \Theta_o(\tilde{\theta}_i) + \Delta\Theta_o(\tilde{\theta}_i, \tilde{\varphi}_i) \quad (3.37)$$

$$= \Theta_o(\Theta_o^{-1}(\theta_o) + \Delta\theta_i) + \Delta\Theta_o(\Theta_o^{-1}(\theta_o) + \Delta\theta_i, \varphi_o + \Delta\varphi_i) \quad (3.38)$$

$$\approx \theta_o + \frac{\partial\Theta_o}{\partial\theta}(\Theta_o^{-1}(\theta_o))\Delta\theta_i + \Delta\Theta_o(\Theta_o^{-1}(\theta_o), \varphi_o), \quad (3.39)$$

$$\varphi_o = \tilde{\varphi}_i + \Delta\Phi_o(\tilde{\theta}_i, \tilde{\varphi}_i) \quad (3.40)$$

$$= \varphi_o + \Delta\varphi_i + \Delta\Phi_o(\Theta_o^{-1}(\theta_o) + \Delta\theta_i, \varphi_o + \Delta\varphi_i) \quad (3.41)$$

$$\approx \varphi_o + \Delta\varphi_i + \Delta\Phi_o(\Theta_o^{-1}(\theta_o), \varphi_o). \quad (3.42)$$

Für die Abweichungen der Kamerawinkel erhält man aus (3.39) und (3.42):

$$\Delta\theta_i = \Delta\Theta_i(\theta_o, \varphi_o) \quad (3.43)$$

$$\approx -\left(\frac{\partial\Theta_o}{\partial\theta}(\Theta_o^{-1}(\theta_o))\right)^{-1} \Delta\Theta_o(\Theta_o^{-1}(\theta_o), \varphi_o), \quad (3.44)$$

$$\Delta\varphi_i = \Delta\Phi_i(\theta_o, \varphi_o) \quad (3.45)$$

$$\approx -\Delta\Phi_o(\Theta_o^{-1}(\theta_o), \varphi_o). \quad (3.46)$$

3.2.2 Neigung der optischen Achse

Eine Neigung der optischen Achse (und damit der Bildebene) um $\Delta\theta_t$ gegenüber der Symmetrieachse des Spiegels ist vergleichsweise unkritisch, da diese im Wesentlichen lediglich eine Verschiebung der Abbildung im Kamerabild zur Folge hat. Es gilt näherungsweise:

$$\tilde{x}_{\text{img}} = x_{\text{img}} + \Delta x_{\text{img}} \approx f(\theta_i \cos \varphi_i - \Delta\theta_t \cos \varphi_t) , \quad (3.47)$$

$$\tilde{y}_{\text{img}} = y_{\text{img}} + \Delta y_{\text{img}} \approx f(\theta_i \sin \varphi_i - \Delta\theta_t \sin \varphi_t) . \quad (3.48)$$

wobei f die Brennweite der Kamera ist und φ_t die Richtung der Neigung angibt.

Für die exakte Berechnung werden zwei zu dem Normalenvektor der Bildebene $\mathbf{n}_{\text{img}}^0 = \mathbf{e}(\Delta\theta_t, \varphi_t)$ orthonormale Vektoren \mathbf{u}^0 und \mathbf{v}^0 so gewählt, dass $\mathbf{u}^0 = \mathbf{e}_x$ und $\mathbf{v}^0 = \mathbf{e}_y$ für den Grenzfall $\Delta\theta_t = 0$ gilt:

$$\mathbf{u}^0 = \frac{1}{\sqrt{\cos^2 \Delta\theta_t + \cos^2 \varphi_t \sin^2 \Delta\theta_t}} \begin{pmatrix} \cos \Delta\theta_t \\ 0 \\ -\cos \varphi_t \sin \Delta\theta_t \end{pmatrix} , \quad (3.49)$$

$$\mathbf{v}^0 = \frac{1}{\sqrt{\cos^2 \Delta\theta_t + \cos^2 \varphi_t \sin^2 \Delta\theta_t}} \begin{pmatrix} -\sin \varphi_t \cos \varphi_t \sin^2 \Delta\theta_t \\ \cos^2 \Delta\theta_t + \cos^2 \varphi_t \sin^2 \Delta\theta_t \\ -\sin \varphi_t \sin \Delta\theta_t \cos \Delta\theta_t \end{pmatrix} . \quad (3.50)$$

Bestimmung des Schnittpunkts der Geraden

$$\mathbf{x}_g = \lambda \mathbf{e}(\theta_i, \varphi_i) \quad (3.51)$$

mit der Bildebene

$$\mathbf{x}_E = f \mathbf{e}(\Delta\theta_t, \varphi_t) + \tilde{x}_{\text{img}} \mathbf{u}^0 + \tilde{y}_{\text{img}} \mathbf{v}^0 \quad (3.52)$$

ergibt die Lösung

$$\tilde{x}_{\text{img}} = \frac{1}{K_t} f (\cos \Delta\theta_t \cos \varphi_i \sin \theta_i - \sin \Delta\theta_t \cos \varphi_t \cos \theta_i) , \quad (3.53)$$

$$\begin{aligned} \tilde{y}_{\text{img}} = \frac{1}{K_t} f & (\cos^2 \Delta\theta_t \sin \varphi_i \sin \theta_i \\ & + \sin^2 \Delta\theta_t \cos \varphi_t \sin \theta_i \sin(\varphi_i - \varphi_t) \\ & - \sin \Delta\theta_t \cos \Delta\theta_t \sin \varphi_t \cos \theta_i) , \end{aligned} \quad (3.54)$$

$$\begin{aligned} K_t := & \sqrt{\cos^2 \Delta\theta_t + \cos^2 \varphi_t \sin^2 \Delta\theta_t} \\ & \times (\cos \Delta\theta_t \cos \theta_i + \cos(\varphi_i - \varphi_t) \sin \Delta\theta_t \sin \theta_i) . \end{aligned} \quad (3.55)$$

Mit $\cos \Delta\theta_t \approx 1$, $\sin \Delta\theta_t \approx \Delta\theta_t$ und $\tan \theta_i \approx \theta_i$ für $\theta_i \ll 1$ folgen die Näherungen (3.47) und (3.48).

3.3 Spiegelflächen mit konstantem Vergrößerungsfaktor

Chahl und Srinivasan beschreiben in [15] spezielle Spiegelflächen mit konstantem vertikalen Winkelvergrößerungsfaktor

$$\alpha := -\frac{\partial\Theta_o(\theta)}{\partial\theta} = \text{const} \iff \theta_o = \Theta_o(\theta) = \Theta_o(0) - \alpha\theta . \quad (3.56)$$

Die Berechnung der die Spiegelform bestimmenden Funktion $r(\theta)$ kann alternativ zur Herleitung in [15] mit Hilfe der Gleichungen (3.12)–(3.15) durchgeführt werden: Auflösen von

$$\cot\theta_o = \frac{a(\theta)}{b(\theta)} , \quad (3.57)$$

nach $\frac{r'(\theta)}{r(\theta)}$ ergibt

$$\frac{r'(\theta)}{r(\theta)} = \frac{\cos(\theta_o - \theta) + 1}{\sin(\theta_o - \theta)} = \cot\left[\frac{1}{2}\Theta_o(0) - \frac{1}{2}(\alpha + 1)\theta\right] . \quad (3.58)$$

Dies kann integriert werden und führt zu dem in [15] angegebenen Ergebnis,

$$r(\theta) = r(0) \frac{\sin(\frac{1}{2}\Theta_o(0))^k}{\sin(\frac{1}{2}\Theta_o(0) - \theta/k)^k} , \quad (3.59)$$

wobei $k := \frac{2}{\alpha+1}$ definiert wurde. Führt man die in Abbildung 3.2 dargestellten Parameter γ_S , θ_{\max} und R_{\max} ein (γ_S ist der Winkel relativ zur x -Achse an der „Spitze“ der Spiegelfläche und θ_{\max} der Kamerawinkel, der dem maximalen Radius R_{\max} der Spiegelfläche entspricht), so gilt mit $\Theta_o(0) = \pi - 2\gamma_S$:

$$r(\theta) = \frac{K}{\cos(\theta/k + \gamma_S)^k} , \quad (3.60)$$

$$K := r(0) \cos^k \gamma_S = r(\theta_{\max}) \cos(\theta_{\max}/k + \gamma_S)^k \quad (3.61)$$

$$= \frac{R_{\max}}{\sin \theta_{\max}} \cos(\theta_{\max}/k + \gamma_S)^k . \quad (3.62)$$

Der Spezialfall $\alpha = 1$ ($\leftrightarrow k = 1$) entspricht (für $\gamma_S > 0$) einem kegelförmigen Spiegel, siehe Abschnitt 1.1. Für den Winkel relativ zum Horizont (Elevation) gilt

$$\varepsilon = \frac{\pi}{2} - \theta_o = \alpha\theta + 2\gamma_S - \frac{\pi}{2} . \quad (3.63)$$

Mit (3.60) folgt aus (3.25) und (3.26)

$$\eta(\theta) = \frac{k-1}{k} = -\frac{1}{2}(\alpha-1) , \quad (3.64)$$

$$\zeta(\theta) = \frac{1}{2} \left(\frac{\sin \theta}{\sin(\alpha\theta + 2\gamma_S)} - 1 \right) . \quad (3.65)$$

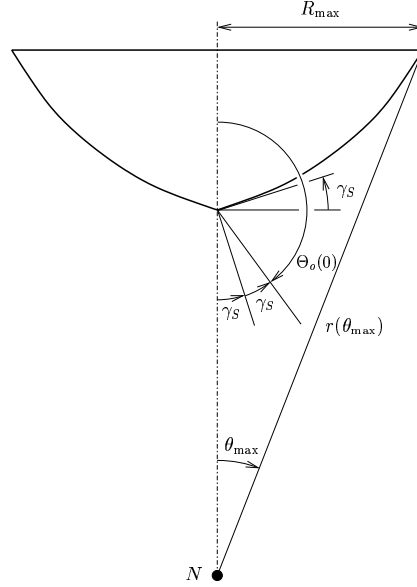


Abbildung 3.2: Darstellung der in den Gleichungen (3.60) und (3.62) verwendeten Parameter γ_S , θ_{\max} und R_{\max} , die neben dem Vergrößerungsfaktor α die Spiegelgeometrie beschreiben. Infolge des Reflexionsgesetzes gilt $\Theta_o(0) = \pi - 2\gamma_S$. N bezeichnet den Ort des Kameraknotenpunkts.

Für die Winkelfehler (3.19) und (3.20) infolge einer Verschiebung der Kamera gilt somit

$$\Delta\theta_o = \Delta\Theta_o(\theta_i, \varphi_i) \approx (\alpha - 1) \left(\frac{\sin \theta_i}{r(\theta_i)} \Delta z - \frac{\cos \theta_i \cos \varphi_i}{r(\theta_i)} \Delta x \right), \quad (3.66)$$

$$\Delta\varphi_o = \Delta\Phi_o(\theta_i, \varphi_i) \approx - \left(\frac{1}{\sin \theta_i} - \frac{1}{\sin(\alpha\theta_i + 2\gamma_S)} \right) \frac{\sin \varphi_i}{r(\theta_i)} \Delta x. \quad (3.67)$$

Für den Vergrößerungsfaktor bedeutet dies

$$\tilde{\alpha} = - \frac{\partial(\Theta_o(\theta_i) + \Delta\Theta_o(\theta_i, \varphi_i))}{\partial\theta_i} = \alpha - \frac{\partial\Delta\Theta_o(\theta_i, \varphi_i)}{\partial\theta_i} \quad (3.68)$$

$$\approx \alpha - (\alpha - 1) \left(\Delta z \frac{\partial \sin \theta_i}{\partial\theta_i r(\theta_i)} - \Delta x \cos \varphi_i \frac{\partial \cos \theta_i}{\partial\theta_i r(\theta_i)} \right) \quad (3.69)$$

$$= \alpha - (\alpha - 1) K^{-1} \cos(\theta_i/k + \gamma_S)^{k-1} \times \left(\cos(\theta_i + \theta_i/k + \gamma_S) \Delta z + \cos \varphi_i \sin(\theta_i + \theta_i/k + \gamma_S) \Delta x \right). \quad (3.70)$$

Für $\alpha = 1$, d.h. kegelförmige ($\gamma_S > 0$) bzw. plane ($\gamma_S = 0$) Spiegelflächen, ändert sich der Vergrößerungsfaktor bei einer Verschiebung der Kamera nicht. Auch für $\alpha \neq 1$ ist eine Verschiebung Δz vergleichsweise unkritisch, da zwar

$\tilde{\alpha} \neq \alpha$ gilt, $\tilde{\alpha}$ jedoch (wegen $\cos x \approx 1$ für $x \ll 1$) nur schwach von θ_i abhängt, d.h. näherungsweise konstant ist. Für eine Abweichung in x -Richtung nimmt $|\tilde{\alpha} - \alpha|$ stetig mit θ_i zu. Resultate numerischer Berechnungen für eine spezielle Spiegelgeometrie sind in [15] zu finden.

Aus $\Theta_o(\theta) = \pi - 2\gamma_S - \alpha\theta$ folgt

$$\theta = \Theta_o^{-1}(\theta_o) = -\frac{1}{\alpha}(\theta_o + 2\gamma_S - \pi) , \quad (3.71)$$

so dass die Winkelabweichungen (3.44) und (3.46) relativ zu einem externen Bezugssystem berechnet werden können:

$$\Delta\theta_i = \Delta\Theta_i(\theta_o, \varphi_o) \approx \frac{1}{\alpha}\Delta\Theta_o(-\frac{1}{\alpha}(\theta_o + 2\gamma_S - \pi)) \quad (3.72)$$

$$= -\frac{(\alpha - 1)}{\alpha} \left(\frac{\sin(\frac{1}{\alpha}(\theta_o + 2\gamma_S - \pi))}{r(-\frac{1}{\alpha}(\theta_o + 2\gamma_S - \pi))} \Delta z + \frac{\cos(\frac{1}{\alpha}(\theta_o + 2\gamma_S - \pi)) \cos \varphi_o}{r(-\frac{1}{\alpha}(\theta_o + 2\gamma_S - \pi))} \Delta x \right) , \quad (3.73)$$

$$\Delta\varphi_i = \Delta\Phi_i(\theta_o, \varphi_o) \approx -\Delta\Phi_o(-\frac{1}{\alpha}(\theta_o + 2\gamma_S - \pi)) \quad (3.74)$$

$$= -\left(\frac{1}{\sin(\frac{1}{\alpha}(\theta_o + 2\gamma_S - \pi))} + \frac{1}{\sin \theta_o} \right) \frac{\sin \varphi_o}{r(-\frac{1}{\alpha}(\theta_o + 2\gamma_S - \pi))} \Delta x , \quad (3.75)$$

wobei $r(-\frac{1}{\alpha}(\theta_o + 2\gamma_S - \pi)) = K \sin(\frac{\alpha+1}{2\alpha}\theta_o - \frac{1}{\alpha}(\frac{\pi}{2} - \gamma_S))^{-k}$ gilt. Beispiele für Anwendungen der Gleichungen (3.70), (3.73) und (3.75) werden in den Abschnitten 3.4.2 und 3.4.3 gegeben.

3.4 Panoramischer Stereosensor für Koala

Um die Exploration von Bürouräumen zu ermöglichen, wurde ein mobiler, autonomer Kleinroboter „Koala“ (siehe Abbildung 3.3) mit einem Stereosensor ausgestattet, der im Vergleich zum Sensor des Kheperas (Abschnitt 1.1) eine deutlich größere Stereobasis und bessere Abbildungseigenschaften besitzt. Ein kompakter Aufbau wurde durch eine veränderte Anordnung der Spiegelflächen realisiert, siehe Abbildung 3.4: Zwei getrennte Spiegelflächen mit identischem Durchmesser wurden so angebracht, dass im Zentrum des Kamerabildes die Umgebung über den oberen Spiegel abgebildet wird. Dazu wurde ein konzentrisches Loch in den unteren Spiegel gefräst.

Für den Abstand der beiden Spiegelflächen muss ein Kompromiss gefunden werden, zwischen einer möglichst großen Stereobasis und der begrenzten Tiefenschärfe der Videokamera. Um weit entfernte Objekte nach Reflexion am Spiegel scharf abbilden zu können, muss die Kamera auf eine Ebene in der Nähe der Spiegeloberfläche fokussiert werden, siehe Abschnitt 3.5.

Abbildung 3.3: Autonomer Koala-Roboter mit panoramischen Stereosensor (Größe etwa $35 \text{ cm} \times 30 \text{ cm}$, Höhe incl. Stereosensor ca. 40 cm). Die Steuerung des Roboters und die Bildauswertung erfolgt mit Hilfe eines PC/104-Rechners (befindet sich unterhalb des Kameraaufbaus). Die Kommunikation mit anderen Computern ist mit Hilfe eines WLAN-Adapters möglich. Der Aufbau erfolgte im Rahmen einer Studienarbeit [40].



Der Stereosensor wurde für einen maximalen Kamerawinkel $\theta_{\max} = 25^\circ$ und einen maximalen Spiegelradius $R_{\max} = 23 \text{ mm}$ (\approx Innenradius des Glaszylinders, in dem die Spiegel fixiert sind) berechnet. Um über beide Spiegel sowohl den Horizont als auch den z.B. für die Erkennung von Hindernissen relevanten Bereich darunter abbilden zu können, wurde als Vergrößerungsfaktor $\alpha = 4$ gewählt. Aufgrund der geringen effektiven Auflösung im Bildzentrum wurde für den oberen Spiegel, der auf den inneren Bildbereich abbildet, mit $0 \leq \theta_1 \leq \theta_{\max,1} = 15^\circ$ ein größerer Winkelbereich gewählt als für den unteren mit $15^\circ \leq \theta_2 \leq \theta_{\max,2} = 25^\circ$. Durch die Wahl von $\gamma_S = 0^\circ$ für den unteren Spiegel und $\gamma_S = 17.5^\circ$ für den oberen, folgt

$$85^\circ \leq \theta_{o,1} \leq 145^\circ, \quad (3.76)$$

$$80^\circ \leq \theta_{o,2} \leq 120^\circ, \quad (3.77)$$

bzw. für die Elevation nach Gleichung (3.63),

$$5^\circ \geq \varepsilon_1 \geq -55^\circ, \quad (3.78)$$

$$10^\circ \geq \varepsilon_2 \geq -30^\circ, \quad (3.79)$$

wobei nur der Bereich $85^\circ \leq \theta_{o,1} \leq 125^\circ$ bzw. $5^\circ \geq \varepsilon_2 \geq -35^\circ$ ausgewertet wird. Die nutzbaren vertikalen Winkelbereiche betragen somit jeweils etwa 40° . Der Verlauf der Strahlen ist in Abbildung 3.5 dargestellt.

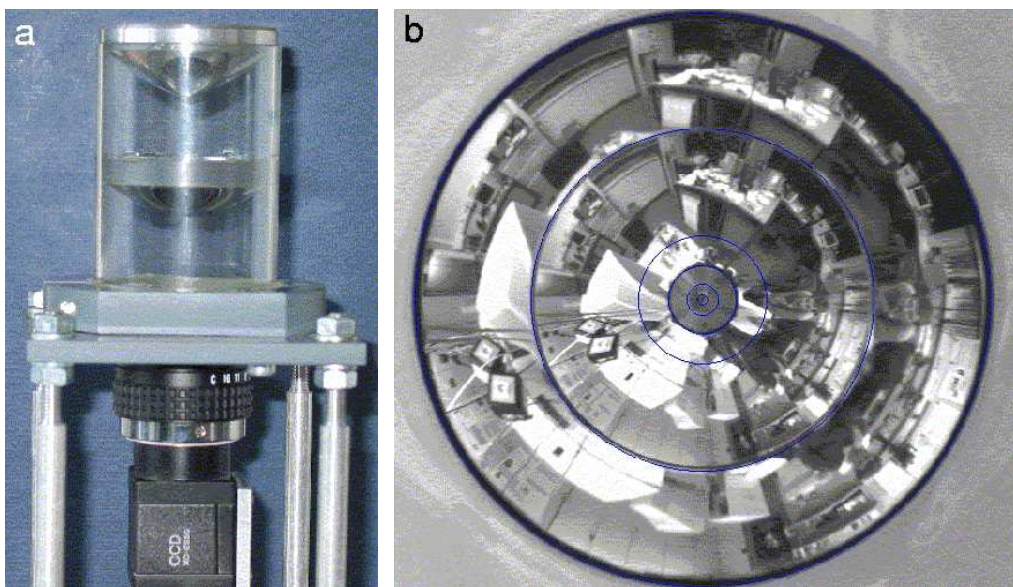


Abbildung 3.4: Panoramischer Stereosensor für Koala-Roboter: **a:** Durch ein konzentrisches Loch im unteren Spiegel können Objekte mit Hilfe des oberen Spiegels abgebildet werden. Diese Anordnung ermöglicht im Vergleich zum Stereosensors des Kheperas eine deutlich größere Stereobasis. **b:** Beispiel eines Stereobildes (Größe etwa 500×500 Pixel). In einem kleinen Bereich um das Bildzentrum wird ein Teil der „Rückseite“ des unteren Spiegels nach Reflexion am oberen abgebildet.

3.4.1 Kaustiken des Stereosensors

Als Kaustik wird die Einhüllende der Strahlen bezeichnet, die so einfallen, dass sie sich nach Reflexion an der Spiegelfläche im Kameraknotenpunkt schneiden [94]. Diese Strahlen verlaufen somit tangential zur Kaustik, siehe Abbildung 3.5. Wegen der Rotationssymmetrie kann die mathematische Beschreibung auf die x - z -Ebene und damit auf den Winkel θ beschränkt werden. Für einen Punkt der Kaustik gilt

$$\mathbf{c}(\theta, \lambda) = \mathbf{x}(\theta) + \lambda \mathbf{e}(\Theta_o(\theta)) , \quad (3.80)$$

wobei $\lambda = \lambda(\theta) \in \mathbb{R}$ zu bestimmen ist. In zwei Dimensionen gilt für den Richtungsvektor $\mathbf{e}(\theta) = (\sin \theta, \cos \theta)^\top$, die Spiegelfläche wird durch $\mathbf{x}(\theta) = r(\theta)(\sin \theta, \cos \theta)^\top$ beschrieben. Die Tangente in jedem Punkt der Kaustik muss parallel (kollinear) zum Richtungsvektor des zugehörigen Strahls sein, d.h.

$$\frac{d}{d\theta} \mathbf{c}(\theta, \lambda) = \frac{\partial}{\partial \theta} \mathbf{c}(\theta, \lambda) + \frac{\partial}{\partial \lambda} \mathbf{c}(\theta, \lambda) \frac{\partial \lambda}{\partial \theta} = k \mathbf{e}(\Theta_o(\theta)) . \quad (3.81)$$

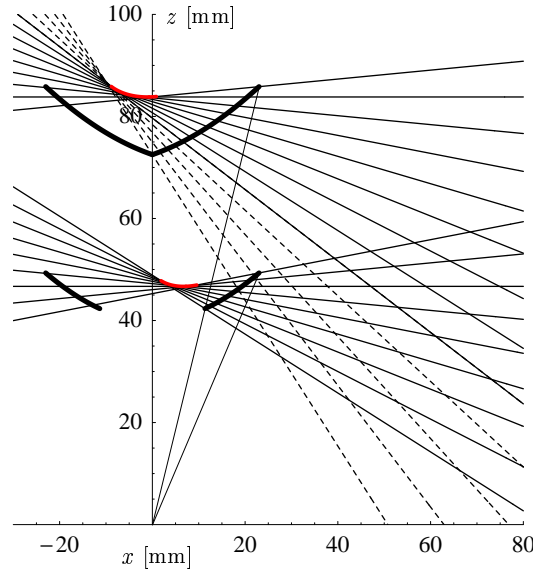


Abbildung 3.5: Strahlverlauf und Kaustiken (rot) des panoramischen Steirosensors. Die Strahlen sind für $\theta_1^l = 1.25^\circ l$, $l = 0, 1, \dots, 12$ und $\theta_2^l = 15^\circ + 1.25^\circ l$, $l = 0, 1, \dots, 8$ dargestellt.

Aus (3.80) folgt $\mathbf{e}(\Theta_o(\theta)) = \frac{\partial}{\partial \lambda} \mathbf{c}(\theta, \lambda)$, so dass (3.81) umgeformt werden kann in

$$\frac{\partial}{\partial \theta} \mathbf{c}(\theta, \lambda) = \left(k - \frac{\partial \lambda}{\partial \theta}\right) \frac{\partial}{\partial \lambda} \mathbf{c}(\theta, \lambda) = \tilde{k} \frac{\partial}{\partial \lambda} \mathbf{c}(\theta, \lambda), \quad (3.82)$$

mit $\tilde{k} := k - \frac{\partial \lambda}{\partial \theta} \in \mathbb{R}$. $\frac{\partial}{\partial \theta} \mathbf{c}(\theta, \lambda)$ und $\frac{\partial}{\partial \lambda} \mathbf{c}(\theta, \lambda)$ sind kollinear, falls

$$\det\left(\frac{\partial}{\partial \theta} \mathbf{c}(\theta, \lambda), \frac{\partial}{\partial \lambda} \mathbf{c}(\theta, \lambda)\right) = 0 \quad (3.83)$$

gilt, d.h. für eine verschwindende Jacobi-Determinante [94]. Es gilt

$$\begin{aligned} \frac{\partial}{\partial \theta} \mathbf{c}(\theta, \lambda) &= r(\theta) \begin{pmatrix} \cos \theta \\ -\sin \theta \end{pmatrix} + r'(\theta) \begin{pmatrix} \sin \theta \\ \cos \theta \end{pmatrix} \\ &\quad + \lambda \begin{pmatrix} \cos \Theta_o(\theta) \Theta_o'(\theta) \\ -\sin \Theta_o(\theta) \Theta_o'(\theta) \end{pmatrix}, \end{aligned} \quad (3.84)$$

$$\frac{\partial}{\partial \lambda} \mathbf{c}(\theta, \lambda) = \begin{pmatrix} \sin \Theta_o(\theta) \\ \cos \Theta_o(\theta) \end{pmatrix}. \quad (3.85)$$

Damit erhält man aus (3.83)

$$\lambda(\theta) = -\frac{1}{\Theta_o'(\theta)} \left(r(\theta) \cos(\theta - \Theta_o(\theta)) + r'(\theta) \sin(\theta - \Theta_o(\theta)) \right). \quad (3.86)$$

Mit $r(\theta)$ aus Gleichung (3.60) und $\Theta_o(\theta) = \pi - 2\gamma_S - \alpha\theta$ folgt $\lambda(\theta) = -r(\theta)/\alpha$ und eingesetzt in (3.80) die Kaustik

$$\mathbf{c}(\theta) := \mathbf{c}(\theta, \lambda(\theta)) \quad (3.87)$$

$$= r(\theta) \begin{pmatrix} \sin \theta \\ \cos \theta \end{pmatrix} - \frac{r(\theta)}{\alpha} \begin{pmatrix} \sin \theta_o(\theta) \\ \cos \theta_o(\theta) \end{pmatrix} \quad (3.88)$$

$$= r(\theta) \left[\begin{pmatrix} \sin \theta \\ \cos \theta \end{pmatrix} - \frac{1}{\alpha} \begin{pmatrix} \sin(\alpha\theta + 2\gamma_S) \\ -\cos(\alpha\theta + 2\gamma_S) \end{pmatrix} \right] . \quad (3.89)$$

Für den Stereosensor des Koala-Roboters ($\alpha = 4$) sind die Kaustiken für $5^\circ \leq \theta_1 \leq 15^\circ$ und $15^\circ \leq \theta_2 \leq 25^\circ$ in Abbildung 3.5 dargestellt (rote Kurven). Da die Kaustiken lokal begrenzt sind, können sie näherungsweise durch die Punkte

$$\langle \mathbf{c}_k \rangle = \frac{1}{\theta_{\max,k} - \theta_{\min,k}} \int_{\theta_{\min,k}}^{\theta_{\max,k}} \mathbf{c}_k(\theta) d\theta , \quad k = 1, 2 \quad (3.90)$$

beschrieben werden. Mit $\theta_{\min,1} = 0^\circ$, $\theta_{\max,1} = 15^\circ$ gilt für den oberen Spiegel $\langle \mathbf{c}_1 \rangle \approx (-5.0 \text{ mm}, 84.5 \text{ mm})$ und mit $\theta_{\min,2} = 15^\circ$, $\theta_{\max,2} = 25^\circ$ für den unteren Spiegel $\langle \mathbf{c}_2 \rangle \approx (5.0 \text{ mm}, 47 \text{ mm})$. Die Stereobasis beträgt folglich näherungsweise $b = \langle \mathbf{c}_1 \rangle_z - \langle \mathbf{c}_2 \rangle_z \approx 37.5 \text{ mm}$.

Die Kaustik des einfachen Kegelspiegels (siehe Abschnitt 1.1.1) ist unabhängig von θ , da mit $\alpha = 1$ aus (3.89)

$$\mathbf{c}(\theta) = \frac{K}{\cos(\theta + \gamma_S)} \begin{pmatrix} \sin \theta - \sin(\theta + 2\gamma_S) \\ \cos \theta + \cos(\theta + 2\gamma_S) \end{pmatrix} \quad (3.91)$$

$$= 2K \begin{pmatrix} -\sin \gamma_S \\ \cos \gamma_S \end{pmatrix} \quad (3.92)$$

folgt. Die Kaustik besteht folglich aus einem einzelnen Punkt (bzw. aus einem Ring in drei Dimensionen). (3.92) ist äquivalent zu Gleichung (1.7) mit $\frac{1}{2}\gamma_k = \frac{\pi}{2} - \gamma_{S,k}$.

Bemerkung: Wird die Rotationssymmetrie z.B. durch eine Verschiebung Δx der Kamera gestört, so muss die Kaustik in drei Dimensionen berechnet werden – parametrisiert durch zwei Variablen. Analog zu Gleichung (3.80) gilt

$$\mathbf{c}(\theta, \varphi, \lambda) = \mathbf{x}(\theta, \varphi) + \lambda \mathbf{e}(\tilde{\Theta}_o(\theta, \varphi), \tilde{\Phi}_o(\theta, \varphi)) , \quad (3.93)$$

wobei $\lambda = \lambda(\theta, \varphi) \in \mathbb{R}$ zu bestimmen ist, siehe [95]. Die Tangentialebene, die durch $\mathbf{x}(\theta, \varphi)$ und die Richtungsvektoren $\frac{d}{d\theta}\mathbf{c}(\theta, \varphi, \lambda)$, $\frac{d}{d\varphi}\mathbf{c}(\theta, \varphi, \lambda)$ definiert wird, muss parallel zu $\mathbf{e}(\tilde{\Theta}_o, \tilde{\Phi}_o)$ sein. Dies bedeutet, dass $\frac{d}{d\theta}\mathbf{c}(\theta, \varphi, \lambda)$,

$\frac{d}{d\varphi}\mathbf{c}(\theta, \varphi, \lambda)$ und $\mathbf{e}(\tilde{\Theta}_o, \tilde{\Phi}_o) = \frac{\partial}{\partial\lambda}\mathbf{c}(\theta, \varphi, \lambda)$ linear abhängig sein müssen:

$$0 = l \frac{d}{d\theta}\mathbf{c}(\theta, \varphi, \lambda) + m \frac{d}{d\varphi}\mathbf{c}(\theta, \varphi, \lambda) + n \mathbf{e}(\tilde{\Theta}_o, \tilde{\Phi}_o) \quad (3.94)$$

$$\begin{aligned} &= l \left(\frac{\partial}{\partial\theta}\mathbf{c}(\theta, \varphi, \lambda) + \frac{\partial}{\partial\lambda}\mathbf{c}(\theta, \varphi, \lambda) \frac{\partial\lambda}{\partial\theta} \right) \\ &\quad + m \left(\frac{\partial}{\partial\varphi}\mathbf{c}(\theta, \varphi, \lambda) + \frac{\partial}{\partial\lambda}\mathbf{c}(\theta, \varphi, \lambda) \frac{\partial\lambda}{\partial\varphi} \right) + n \frac{\partial}{\partial\lambda}\mathbf{c}(\theta, \varphi, \lambda) \end{aligned} \quad (3.95)$$

$$= l \frac{\partial}{\partial\theta}\mathbf{c}(\theta, \varphi, \lambda) + m \frac{\partial}{\partial\varphi}\mathbf{c}(\theta, \varphi, \lambda) + \left(l \frac{\partial\lambda}{\partial\theta} + m \frac{\partial\lambda}{\partial\varphi} + n \right) \frac{\partial}{\partial\lambda}\mathbf{c}(\theta, \varphi, \lambda) \quad (3.96)$$

$$= l \frac{\partial}{\partial\theta}\mathbf{c}(\theta, \varphi, \lambda) + m \frac{\partial}{\partial\varphi}\mathbf{c}(\theta, \varphi, \lambda) + \tilde{n} \frac{\partial}{\partial\lambda}\mathbf{c}(\theta, \varphi, \lambda) \quad , \quad (3.97)$$

mit $(l, m, \tilde{n}) \neq (0, 0, 0)$. Dies ist – wie im zweidimensionalen Fall – äquivalent zur Forderung nach einer verschwindenden Jacobi-Determinante, d.h.

$$\det \left(\frac{\partial}{\partial\theta}\mathbf{c}(\theta, \varphi, \lambda), \frac{\partial}{\partial\varphi}\mathbf{c}(\theta, \varphi, \lambda), \frac{\partial}{\partial\lambda}\mathbf{c}(\theta, \varphi, \lambda) \right) = 0 \quad . \quad (3.98)$$

3.4.2 Stereobilder bei Kameraverschiebung

Mit Hilfe der Gleichungen (3.73) und (3.75) können die Verzerrungen des Kamerabildes für ein weit entferntes externes Referenzmuster berechnet werden, d.h. $\tilde{\theta}_i = \Theta_o^{-1}(\theta_o) + \Delta\Theta_i(\theta_o, \varphi_o)$ und $\tilde{\varphi}_i = \varphi_o + \Delta\Phi_i(\theta_o, \varphi_o)$. Im Folgenden wird ein gitterartiges Muster betrachtet, dessen horizontalen Strukturen durch $\theta_o^k(\varphi_o) = 80^\circ + 5^\circ k$, $k = 0, 1, \dots, 10$, $0^\circ \leq \varphi_o < 360^\circ$ gegeben sind und dessen vertikalen Linien durch $\varphi_o^l(\theta_o) = 15^\circ l$, $l = 0, 1, \dots, 23$, $80^\circ \leq \theta_o \leq 130^\circ$ beschrieben werden.

Abbildung 3.6 zeigt berechnete Kamerabilder für unterschiedliche Verschiebungen $(\Delta x, \Delta z)$ der Kamera, wobei für die Spiegel eine perfekte Justierung angenommen wurde. Im Zentrum (d) ist das fehlerfreie Bild dargestellt. Durch eine Fehlstellung in x -Richtung verschwinden Bildbereiche des oberen Spiegels, die durch den unteren verdeckt werden, siehe (a), (f), (g). Für $\Delta x = 5$ mm, $\Delta z = 0$ mm (g) sind im Bildzentrum starke Verzerrungen zu erkennen. In diesem Bereich sind die Gleichungen (3.73) und (3.75) nur als grobe Näherungen zu betrachten. Ist $\Delta z > 0$ (e), so entsteht im Kamerabild ein „freier“, ringförmiger Bereich zwischen den Spiegeln. Im Gegensatz dazu werden für $\Delta z < 0$ (c) Winkelbereiche in Horizontnähe „verschluckt“. In (b) ist $\Delta x = 1.25$ mm mit $\Delta z = 5$ mm kombiniert dargestellt. Durch Vergleich mit $\Delta x = 1.25$ mm, $\Delta z = 0$ mm (a) ist zu erkennen, dass fast der gesamte abzubildende Winkelbereich im Kamerabild erscheint – wenn auch nicht mit den korrekten θ_i -Werten. Sofern die Kamera nicht perfekt bezüglich Δx justiert werden kann, ist folglich eine kleine Abweichung Δz sinnvoll, da dann eine Korrektur durch eine geeignete Bildtransformation möglich ist.

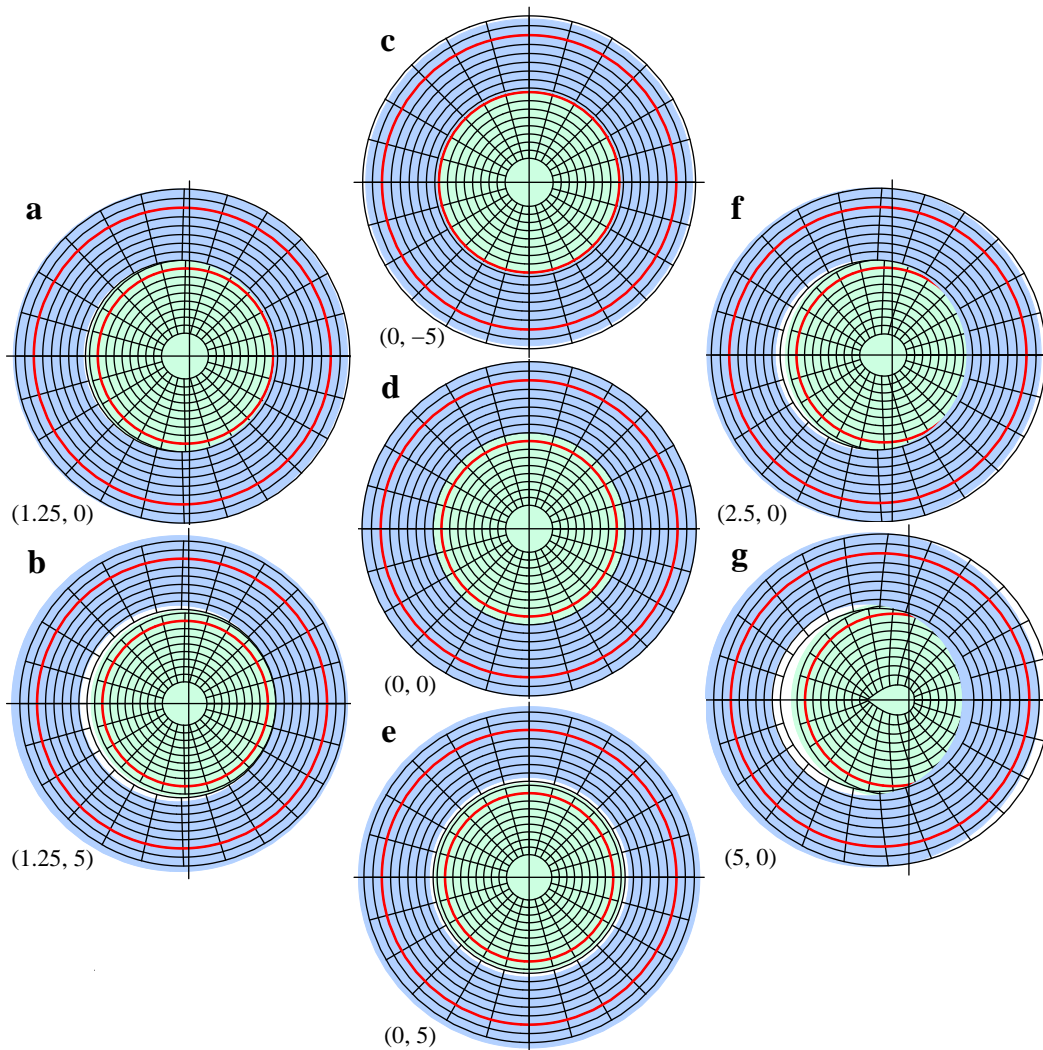


Abbildung 3.6: Berechnete Kamerabilder für das im Text beschriebene Kalibrierungsmuster bei unterschiedlichen Verschiebungen der Kamera. Angegeben ist jeweils unten links die Verschiebung $(\Delta x, \Delta z)$ in mm. Eingezeichnet ist das Bildkoordinatensystem, d.h. die optische Achse definiert den Ursprung. Grün bzw. blau unterlegt sind die Bilder des oberen bzw. unteren Spiegels. Rot markiert sind die Bilder des Horizonts $\theta_o = 90^\circ$.

3.4.3 Vergrößerungsfaktor

Abbildung 3.7 zeigt Veränderung des Winkel-Vergrößerungsfaktors $\tilde{\alpha}$ für $\Delta z = 1.25$ mm, 5 mm und $\Delta x = -1.25$ mm, -5 mm, berechnet nach Gleichung (3.70). Für kleine Verschiebungen ist die Abweichung $|\tilde{\alpha} - 4|$ in dem für den Stereosensor wichtigen θ_i -Bereichen unkritisch. Zu beachten ist jedoch,

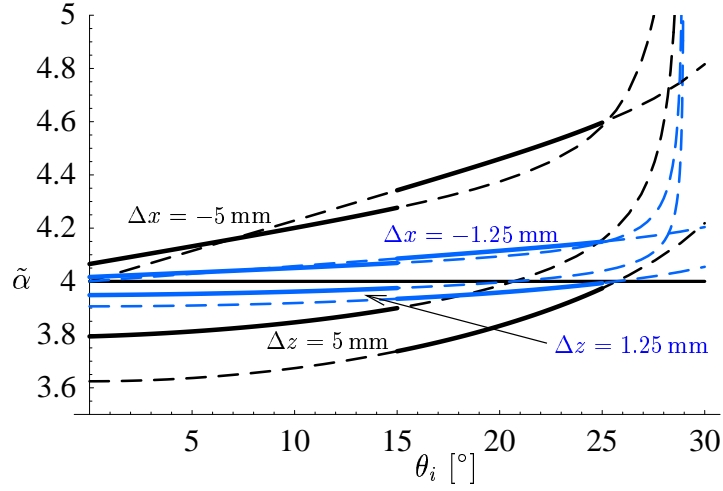


Abbildung 3.7: Vergrößerungsfaktor $\tilde{\alpha}$ berechnet in linearer Näherung für $\Delta z = 1.25, 5 \text{ mm}$ und $\Delta x = -1.25, -5 \text{ mm}$ (bei $\varphi_i = 0^\circ$, d.h. entlang der positiven x -Achse). Die entsprechenden Kurven für negatives Δz bzw. positives Δx ergeben sich durch Spiegelung an der Geraden $\tilde{\alpha} = 4$. Die für den Stereosensor relevanten Bereiche mit $0^\circ \leq \theta_i \leq 15^\circ$ für den oberen Spiegel bzw. $15^\circ \leq \theta_i \leq 25^\circ$ für den unteren entsprechen den kontinuierlichen Kurvenabschnitten.

dass $|\tilde{\alpha} - 4|$ für $\Delta x \neq 0$ von φ_i abhängt. Die Stereobildauswertung vereinfacht sich, wenn die Vergrößerungsfaktoren der korrespondierenden Bildausschnitte vergleichbar sind. Deshalb ist – wie man anhand von Abbildung 3.7 erkennt – eine Verschiebung Δx ungünstiger und schwieriger durch eine geeignete Bildtransformation (siehe nächster Abschnitt) zu korrigieren.

3.4.4 Justierung und Bildtransformation

Das Stereobild der Kamera (siehe Abb. 3.4 b) wird vor der Stereoauswertung in zwei ($k = 1, 2$) panoramische Bilder gleicher Auflösung ($N_i \times N_j = 360 \times 100$) transformiert ($i = 0, 1, \dots, N_i - 1, j = 0, 1, \dots, N_j - 1$):

$$I_k^{\text{uw}}(i, j) := (K_k^{ij})^{-1} \sum_{x,y} w_k^{ij}(x, y) I[\bar{x}_k(i, j) + x, \bar{y}_k(i, j) + y] , \quad (3.99)$$

$$K_k^{ij} := \sum_{x,y} w_k^{ij}(x, y) , \quad (3.100)$$

$$\bar{x}_k(i, j) := x_0 + (R_k + j\Delta R) \cos\left(\frac{2\pi}{N_i}i\right) , \quad (3.101)$$

$$\bar{y}_k(i, j) := y_0 + (R_k + j\Delta R) \sin\left(\frac{2\pi}{N_i}i\right) . \quad (3.102)$$



Abbildung 3.8: Mit Hilfe von Gleichung (3.99) transformierte Bilder des panoramischen Stereobildes aus Abbildung 3.4 (Auflösung ist jeweils 360×100 Pixel, für die Darstellung wurde in x -Richtung um Faktor 2 skaliert). **a:** Bild des oberen Spiegels. **b:** Bild des unteren Spiegels. **c:** Überlagerung von **a** und **b**: Korrespondierende Bildbereiche weisen nur vertikale Verschiebungen auf, so dass die Bestimmung der Disparitäten auf eine eindimensionale Maximierung der lokalen Bildkorrelation reduziert werden kann.

„uw“ steht dabei für engl. „unwarped“ (\approx entzerrt). In der gegenwärtigen Implementierung wird in (3.99) mit konstanter Gewichtung über ein Sektor-element mit Zentrum bei $(\bar{x}_k(i, j), \bar{y}_k(i, j))$ summiert.

Abbildung 3.8 zeigt die transformierten Bilder und deren Überlagerung für das Kamerabild aus Abb. 3.4. Für beide transformierten Bilder beträgt die Auflösung in θ_o (d.h. in vertikaler Richtung) $\rho_\theta = N_j / (\alpha \Delta\theta) \approx 100 \text{ Pixel} / 40^\circ = 2.5 \text{ Pixel} / ^\circ$ und ist damit deutlich höher als in φ (d.h. in horizontaler Richtung) $\rho_\varphi = N_i / 360^\circ \approx 360 \text{ Pixel} / 360^\circ = 1 \text{ Pixel} / ^\circ$.

Unter der Annahme einer perfekten Anordnung der Spiegel relativ zueinander wurden die Kamera und das Bildzentrum (x_0, y_0) per „Auge und Hand“ solange justiert bis in beiden transformierten Bilder korrespondierende Bildbereiche den gleichen Spaltenindex i besitzen, vgl. Abb. 3.8c. Eine weitere Korrektur von Fehlstellungen kann durch eine Verfeinerung der Bildtransformation erfolgen, beispielsweise durch Ersetzen des konstanten Parameters ΔR durch eine von j (und evtl. auch von i) abhängige radiale Schrittweite.

Da für die in Kapitel 1 beschriebenen Verfahren zum visuellen Heimfinden vor allem der Horizontbereich von Interesse ist, wurde dieser in den transformierten Bildern wie folgt bestimmt: Die Höhe eines Streifens der Breite $b = 37.5 \text{ mm}$ (entspricht der in Abschnitt 3.4.1 bestimmten Stereobasis) wurde solange variiert, bis seine Oberkante im transformierten Bild des obe-

ren Spiegels I_1^{uw} und seine Unterkante im transformierten Bild des unteren Spiegels I_2^{uw} für alle Entfernung des Spiegels von der Wand näherungsweise gleich bleiben. Durch schrittweise Rotation des Koala-Roboters wurden die Bildindizes j_1^l, j_2^l der Streifenober- bzw. unterkante (die dem Horizont für den oberen bzw. unterem Spiegel entsprechen) für $N_o = 12$ Orientierungen $\varphi^l = 360^\circ/N_o l, l = 0, 1, \dots, N_o - 1$ ermittelt. Anschließend wurden die 6 Parameter $\bar{j}_k^{\text{H}}, \Delta j_k^{\text{H}}$ und φ_k^{H} durch Minimieren der Fehlerfunktionen ($k = 1, 2$)

$$E_k^{\text{H}}(j_k^{\text{H}}, x_k^{\text{H}}, y_k^{\text{H}}) := \sum_l \left(\bar{j}_k^{\text{H}} + x_k^{\text{H}} \cos\left(\frac{2\pi}{N_o} j\right) + y_k^{\text{H}} \sin\left(\frac{2\pi}{N_o} j\right) - j_k^l \right)^2$$

bezüglich $\bar{j}_k^{\text{H}}, x_k^{\text{H}} := \Delta j_k^{\text{H}} \cos \varphi_k^{\text{H}}$ und $y_k^{\text{H}} := \Delta j_k^{\text{H}} \sin \varphi_k^{\text{H}}$ bestimmt. Die Lösung ergibt Mittelwert und ersten Fourierkoeffizienten, d.h.

$$\bar{j}_k^{\text{H}} = \frac{1}{N_o} \sum_l j_k^l, \quad (3.103)$$

$$\Delta j_k^{\text{H}} \cos \varphi_k^{\text{H}} = \frac{2}{N_o} \sum_l j_k^l \cos\left(\frac{2\pi}{N_o} l\right), \quad (3.104)$$

$$\Delta j_k^{\text{H}} \sin \varphi_k^{\text{H}} = \frac{2}{N_o} \sum_l j_k^l \sin\left(\frac{2\pi}{N_o} l\right). \quad (3.105)$$

Für die Positionen des Horizonts in den transformierten Bilder gilt dann näherungsweise

$$j_k^{\text{H}}(i) = \bar{j}_k^{\text{H}} + \Delta j_k^{\text{H}} \cos\left(\frac{2\pi}{N_o} i - \varphi_k^{\text{H}}\right). \quad (3.106)$$

3.5 Bestimmung der optimalen Fokusebene

Für eine scharfe Abbildung in beiden Bildbereichen eines panoramischen Stereosensors ist es notwendig, dass die von den Spiegelflächen erzeugten virtuellen Bilder möglichst im gleichen Abstand zum Kameraknotenpunkt entstehen. Allerdings wird im Normalfall bereits für eine einzelne Spiegelfläche ein parallel einfallendes Strahlenbündel nicht einen einzelnen virtuellen Bildpunkt besitzen. Dies wird im folgenden Abschnitt untersucht. Anschließend wird die daraus resultierende Unschärfe im Kamerabild berechnet. Für den Stereosensor des Koala-Roboters wird außerdem die Brennweite bestimmt, für die über beide Spiegelflächen eine befriedigende Abbildung erreicht werden kann.

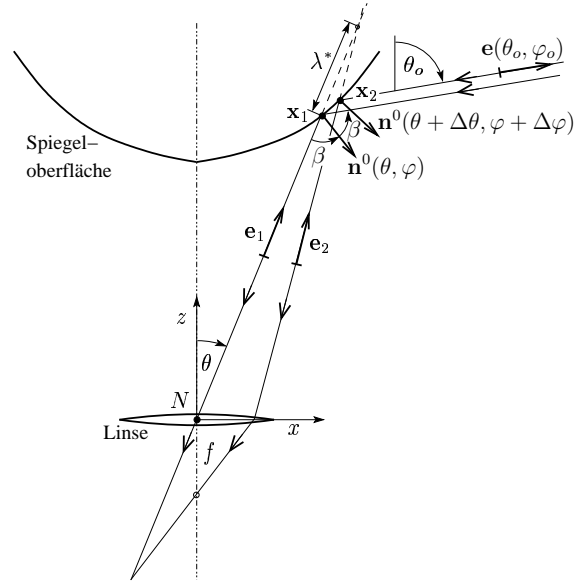


Abbildung 3.9: Berechnung virtueller Bildpunkte: Zwei vertikal versetzte, parallele Strahlen, die aus Richtung $-\mathbf{e}(\theta_o, \varphi_o)$ einfallen, werden an der Spiegelfläche in den Punkten \mathbf{x}_1 bzw. \mathbf{x}_2 reflektiert. Der entsprechende virtuelle Bildpunkt befindet sich am Schnittpunkt der reflektierten Strahlen (mit Richtungsvektoren \mathbf{e}_1 bzw. \mathbf{e}_2) und ist um λ^* von Punkt \mathbf{x}_1 entfernt. Für in andere Richtung versetzte, parallele Strahlen muss kein Schnittpunkt existieren, da der Verlauf der Strahlen im Allgemeinen nicht auf eine Ebene beschränkt ist.

3.5.1 Berechnung virtueller Bildflächen

Wie in Abbildung 3.9 dargestellt, werden zwei parallel einfallende Strahlen mit Richtungsvektor $-\mathbf{e}(\theta_o, \varphi_o)$, die von einem weit entfernten Objekt ausgehen und an benachbarten Punkten \mathbf{x}_1 und \mathbf{x}_2 die Spiegelfläche treffen, in Richtung $-\mathbf{e}_1$ und $-\mathbf{e}_2$ reflektiert. Die Geraden

$$\mathbf{g}_1(\lambda) = \mathbf{x}_1 + \lambda \mathbf{e}_1, \quad \lambda \in \mathbb{R}, \quad (3.107)$$

$$\mathbf{g}_2(\mu) = \mathbf{x}_2 + \mu \mathbf{e}_2, \quad \mu \in \mathbb{R} \quad (3.108)$$

haben ihren minimalen Abstand $|\mathbf{g}_1(\lambda) - \mathbf{g}_2(\mu)|^2$ in den Punkten, die durch

$$\lambda^* = \frac{(\mathbf{x}_2 - \mathbf{x}_1)((\mathbf{e}_1 \mathbf{e}_2) \mathbf{e}_2 - \mathbf{e}_1)}{(\mathbf{e}_1 \mathbf{e}_2)^2 - 1}, \quad (3.109)$$

$$\mu^* = \frac{(\mathbf{x}_1 - \mathbf{x}_2)((\mathbf{e}_1 \mathbf{e}_2) \mathbf{e}_1 - \mathbf{e}_2)}{(\mathbf{e}_1 \mathbf{e}_2)^2 - 1} \quad (3.110)$$

gegeben sind. In dem durch diese Punkte definierten Bereich entsteht das virtuelle Bild des Objekts.

Wird für $\mathbf{g}_1(\lambda)$ der Zentralstrahl durch den Kameraknotenpunkt betrachtet, so gilt $\mathbf{x}_1 = \mathbf{x}(\Theta_o^{-1}(\theta_o), \varphi_o)$ und $\mathbf{e}_1 = \mathbf{e}(\Theta_o^{-1}(\theta_o), \varphi_o)$. Im Folgenden wird $\theta := \Theta_o^{-1}(\theta_o)$ und $\varphi = \varphi_o$ gesetzt. Die Gerade $\mathbf{g}_2(\mu)$ treffe die Spiegelfläche in einem zu $\mathbf{x}_1 = \mathbf{x}(\theta, \varphi)$ benachbarten Punkt $\mathbf{x}_2 = \mathbf{x}(\theta + \Delta\theta, \varphi + \Delta\varphi)$. Für den Vektor \mathbf{e}_2 , der die Richtung des in diesem Punkt gespiegelten Strahls angibt, gilt

$$\mathbf{e}_2 = \mathbf{e}(\theta_o, \varphi) - 2[\mathbf{n}^0(\theta + \Delta\theta, \varphi + \Delta\varphi)\mathbf{e}(\theta_o, \varphi)]\mathbf{n}^0(\theta + \Delta\theta, \varphi + \Delta\varphi) .$$

Durch Entwicklung von \mathbf{x}_2 und \mathbf{e}_2 nach $\Delta\theta$ und $\Delta\varphi$ kann der Parameter λ^* näherungsweise berechnet werden. Wie in Anhang A.8 gezeigt wird, gilt:

$$\lambda^* \approx \lambda_{\theta\varphi}^*(\theta, \Delta\theta, \Delta\varphi) \quad (3.111)$$

$$:= -\frac{1}{2}r(\theta) \left[1 + \left(\frac{r'(\theta)}{r(\theta)} \right)^2 \right] \frac{\Delta\theta^2 g(\theta) + \Delta\varphi^2 h(\theta) \sin \theta}{\Delta\theta^2 g(\theta)^2 + \Delta\varphi^2 h(\theta)^2} , \quad (3.112)$$

$$g(\theta) := 1 + 2 \left(\frac{r'(\theta)}{r(\theta)} \right)^2 - \frac{r''(\theta)}{r(\theta)} , \quad (3.113)$$

$$h(\theta) := \sin \theta - \frac{r'(\theta)}{r(\theta)} \cos \theta . \quad (3.114)$$

Gleichung (3.112) besitzt die Extremwerte²

$$\lambda_{\theta}^*(\theta) := \lambda_{\theta\varphi}^*(\theta, \Delta\theta, \Delta\varphi = 0) = -\frac{r(\theta)[1 + (\frac{r'(\theta)}{r(\theta)})^2]}{2[1 + 2(\frac{r'(\theta)}{r(\theta)})^2 - \frac{r''(\theta)}{r(\theta)}]} , \quad (3.115)$$

$$\lambda_{\varphi}^*(\theta) := \lambda_{\theta\varphi}^*(\theta, \Delta\theta = 0, \Delta\varphi) = -\frac{r(\theta)[1 + (\frac{r'(\theta)}{r(\theta)})^2] \sin \theta}{2[\sin \theta - \frac{r'(\theta)}{r(\theta)} \cos \theta]} . \quad (3.116)$$

In Anhang A.10 wird gezeigt, dass $\mathbf{x}_{\theta}(\theta, \varphi)$ und $\mathbf{x}_{\varphi}(\theta, \varphi)$ die Hauptkrümmungsrichtungen angeben. Mit Hilfe der entsprechenden Hauptkrümmungsradien $R_{\theta}(\theta, \varphi)$ und $R_{\varphi}(\theta, \varphi)$ können (3.115) und (3.116) umgeformt werden zu

$$\lambda_{\theta}^*(\theta) = \frac{R_{\theta}(\theta)}{2\sqrt{1 + (\frac{r'(\theta)}{r(\theta)})^2}} = -(\mathbf{n}^0 \mathbf{e}) \frac{R_{\theta}(\theta)}{2} = \cos \beta \frac{R_{\theta}(\theta)}{2} , \quad (3.117)$$

²Dies kann mit Hilfe der Substitutionen $l \cos \alpha := \Delta\theta$, $l \sin \alpha := \Delta\varphi$ und $\chi := \cos^2 \alpha$ verdeutlicht werden. Es gilt dann $\frac{\Delta\theta^2 g(\theta) + \Delta\varphi^2 h(\theta) \sin \theta}{\Delta\theta^2 g(\theta)^2 + \Delta\varphi^2 h(\theta)^2} = \frac{\cos^2 \alpha g(\theta) + \sin^2 \alpha h(\theta) \sin \theta}{\cos^2 \alpha g(\theta)^2 + \sin^2 \alpha h(\theta)^2} = \frac{\chi g(\theta) + (1-\chi)h(\theta) \sin \theta}{\chi g(\theta)^2 + (1-\chi)h(\theta)^2}$. Der letzte Term besitzt keine Extrema und (da der Nenner positiv ist, sofern nicht $g(\theta) = h(\theta) = 0$) keine Polstellen in $\chi \in]0, 1[$. Die Extremwerte werden somit am Rand, d.h. für $\chi = 0$ und $\chi = 1$ angenommen.

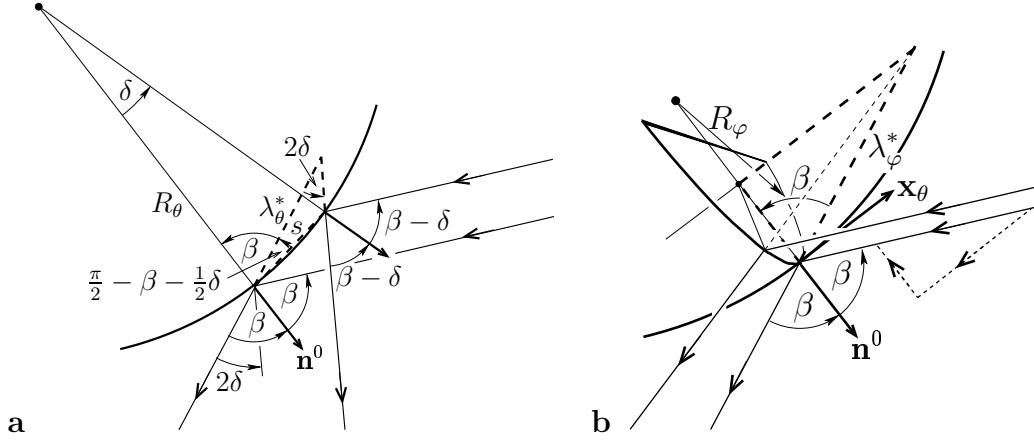


Abbildung 3.10: **a:** Die Größe λ_θ^* kann durch Betrachtung einer Reflexion an einem Ellipsoid mit Krümmungsradius R_θ in θ -Richtung ermittelt werden. **b:** Zur Bestimmung von λ_φ^* werden die Richtungsvektoren der parallel einfallenden Strahlen bezüglich des Normalenvektors in kollineare und orthogonale Komponenten zerlegt. Betrachtet man den Normalschnitt in Richtung \mathbf{x}_φ , so erfolgt die Reflexion näherungsweise an einer Kugel mit Radius R_φ .

$$\lambda_\varphi^*(\theta) = \frac{R_\varphi(\theta) \sqrt{1 + \left(\frac{r'(\theta)}{r(\theta)}\right)^2}}{2} = -\frac{1}{\mathbf{n}^0 \mathbf{e}} \frac{R_\varphi(\theta)}{2} = \frac{R_\varphi(\theta)}{2 \cos \beta}, \quad (3.118)$$

wobei β den Reflexionswinkel am Punkt \mathbf{x}_1 bezeichnet, siehe Abbildung 3.9. Gleichungen (3.117) und (3.118) können als Reflexion nicht-senkrecht einfallender Strahlen an einem Rotationsellipsoid mit lokalen Krümmungsradien R_θ und R_φ interpretiert werden:³ Für vertikal versetzte Strahlen kann λ_θ^* aus dem in Abbildung 3.10 a dargestellten gestrichelten Dreieck bestimmt werden. Für die Strecke s gilt $s = 2R_\theta \tan \frac{\delta}{2}$. Der λ_θ^* gegenüberliegende Winkel beträgt $\pi - (\frac{\pi}{2} - \beta - \frac{1}{2}\delta) - 2\delta = \frac{\pi}{2} + \beta - \frac{3}{2}\delta$. Aus dem Sinussatz folgt dann

$$\lambda_\theta^* = \frac{2R_\theta \tan \frac{\delta}{2}}{\sin(2\delta)} \sin\left(\frac{\pi}{2} + \beta - \frac{3}{2}\delta\right) \quad (3.119)$$

$$= \frac{2R_\theta \tan \frac{\delta}{2}}{\sin(2\delta)} \cos\left(\beta - \frac{3}{2}\delta\right) \stackrel{\delta \ll 1}{\approx} \frac{R_\theta}{2} \cos \beta. \quad (3.120)$$

Zur Berechnung von λ_φ^* wird ein Normalschnitt in \mathbf{x}_φ -Richtung betrachtet, siehe Abbildung 3.10 b. Für die Komponenten der einfallenden Strahlen, die parallel zum Normalenvektor verlaufen, kann die Abbildung näherungsweise als eine (zentrale) Reflexion an einer Kugel (bzw. Kreis) mit Radius R_φ

³Ein derartiges Ellipsoid ist beispielsweise durch die spezielle Wahl $a = \sqrt{R_\varphi R_\theta}$, $b = c = R_\varphi$ für die Hauptachsen zu erreichen.

beschrieben werden. Achsnahe Strahlen werden dabei so reflektiert, dass sie von einem Punkt mit Abstand $\frac{1}{2}R_\varphi$ vom Kreismittelpunkt auszugehen scheinen. Die Länge λ_φ^* erhält man dann durch Projektion dieses Punktes auf den verlängerten Hauptstrahl entlang des Lots auf dem Normalenschnitt, d.h. entlang von \mathbf{x}_θ (siehe gestricheltes Dreieck):

$$\lambda_\theta^* \approx \frac{\frac{1}{2}R_\varphi}{\cos \beta} . \quad (3.121)$$

Für Spiegel mit konstanter Winkelvergrößerung erhält man durch Einsetzen von $r(\theta)$ aus Gleichung (3.60) in (3.115) und (3.116):

$$\lambda_\theta^*(\theta) = \frac{r(\theta)}{\alpha - 1} , \quad (3.122)$$

$$\lambda_\varphi^*(\theta) = \frac{r(\theta) \sin \theta}{\sin(\alpha\theta + 2\gamma_S) - \sin \theta} . \quad (3.123)$$

Die entsprechenden virtuellen Bildflächen sind gegeben durch

$$\mathbf{f}_\theta(\theta, \varphi) := \mathbf{x}(\theta, \varphi) + \lambda_\theta^*(\theta)\mathbf{e}(\theta, \varphi) , \quad (3.124)$$

$$\mathbf{f}_\varphi(\theta, \varphi) := \mathbf{x}(\theta, \varphi) + \lambda_\varphi^*(\theta)\mathbf{e}(\theta, \varphi) . \quad (3.125)$$

Für $\gamma_S \neq 0$ berührt $\mathbf{f}_\varphi(\theta, \varphi)$ die Spiegelfläche bei $\theta = 0$, da $\lambda_\varphi^*(0) = 0$. Für $\gamma_S = 0$ berühren sich die beiden virtuellen Bildflächen bei $\theta = 0$, da $\lim_{\theta \rightarrow 0} \lambda_\varphi^*(\theta) = \lambda_\varphi^*(0)$. Die Abbildung kann dort in diesem Fall durch eine Reflexion an einer Kugel beschrieben werden.

Speziell für $\alpha = 1$ (Kegelspiegel) gilt $\lambda_\theta^* \rightarrow \infty$. $\lambda_\varphi^* = \frac{r(\theta) \sin \theta}{\sin(\theta + 2\gamma_S) - \sin \theta}$ ist jedoch endlich (mit Ausnahme von $\gamma_S = 0$, d.h. planer Spiegel). Dies entspricht dem in Abschnitt 1.1.3 beschriebenen Abbildungsfehler.

Die Vernachlässigung von Termen höherer Ordnung in $\Delta\theta$, $\Delta\varphi$ und die Betrachtung der Grenzfälle in (3.115) und (3.116) ist motiviert durch die Tatsache, dass für jede Einfallsrichtung nur ein begrenzter Bereich der Spiegelflächen zur Abbildung beitragen kann⁴: Nur ein kleiner Ausschnitt eines parallel einfallenden Strahlenbündels wird so am Spiegel reflektiert, dass es in die Blendenöffnung der Kamera fällt. Bereiche unscharfer Abbildung können normalerweise durch Vergleich von \mathbf{f}_θ und \mathbf{f}_φ erkannt werden. Für eine genauere Analyse muss jedoch der gesamte zur Abbildung beitragende Spiegelausschnitt berücksichtigt werden. Dieser wird in Anhang A.9 näherungsweise berechnet.

⁴Dies hat auch zur Folge, dass die Annahme parallel einfallender Strahlen bereits für relativ geringe Objektabstände berechtigt ist.

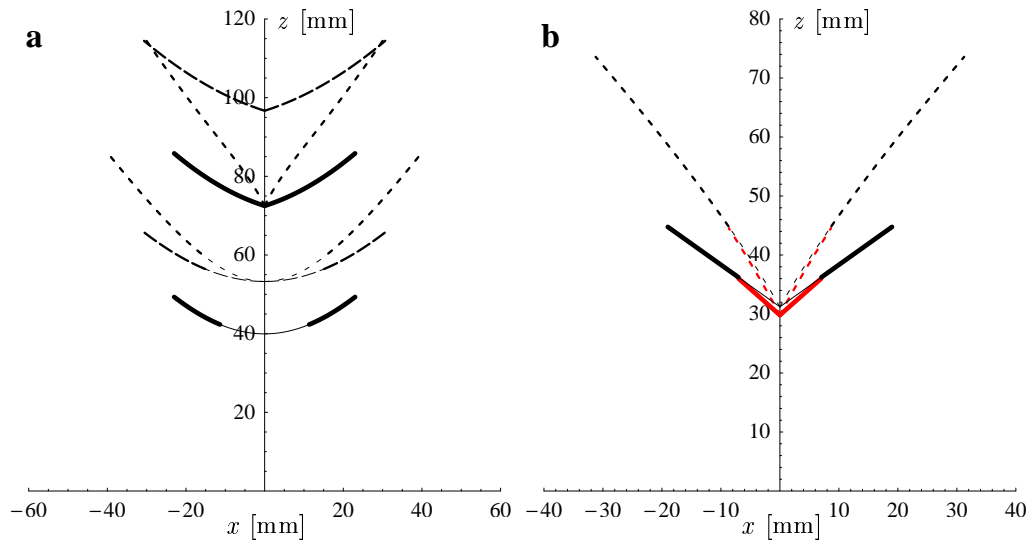


Abbildung 3.11: Virtuelle Bildflächen der verwendeten panoramischen Stereosensoren. Kontinuierliche Kurven stellen die Spiegelflächen dar. Gezeigt sind jeweils $\mathbf{f}_\theta(\theta, 0)$ (gestrichelt) und $\mathbf{f}_\varphi(\theta, 0)$ (punktiert). Dünn eingezeichnet sind Fortsetzungen von Kurven, die nicht zur Abbildung beitragen (da entsprechende Spiegelflächen nicht vorhanden oder verdeckt sind). **a:** Stereospiegel mit $\alpha = 4$ für Koala-Roboter. **b:** Stereospiegel mit $\alpha = 1$ für Khepera-Roboter. Zur besseren Unterscheidbarkeit ist der untere Kegel rot markiert. Für $\alpha = 1$ verläuft $\mathbf{f}_\theta(\theta, 0)$ im „Unendlichen“, siehe Gleichung (3.122).

Abbildung 3.11 zeigt die virtuellen Bildflächen der beiden verwendeten Stereosensoren mit $\alpha = 1$ für den Khepera und $\alpha = 4$ für den Koala-Roboter. Falls ein Objektpunkt nicht auf einen einzelnen virtuellen Bildpunkt abgebildet wird ($\mathbf{f}_\theta \neq \mathbf{f}_\varphi$), so hat dies eine unscharfe Abbildung im Kamerabild zur Folge. Für den Koala-Stereosensor sind vor allem die unterschiedlichen virtuellen Bildflächen der beiden Spiegelteile problematisch. Das Ausmaß der Unschärfe hängt dabei vom verwendeten Objektiv und der eingestellten Blende ab. Dies wird im folgenden Abschnitt untersucht.

3.5.2 Größe des Unschärfebereichs im Kamerabild

Wie in Abbildung 3.12 dargestellt, ändert sich in Abhängigkeit von der Bildweite b die Größe des Unschärfebereichs $\Delta\xi$. Ist die Bildweite auf b eingestellt, so gilt

$$\frac{\Delta\xi}{D} = 1 - \frac{b}{b_z} . \quad (3.126)$$

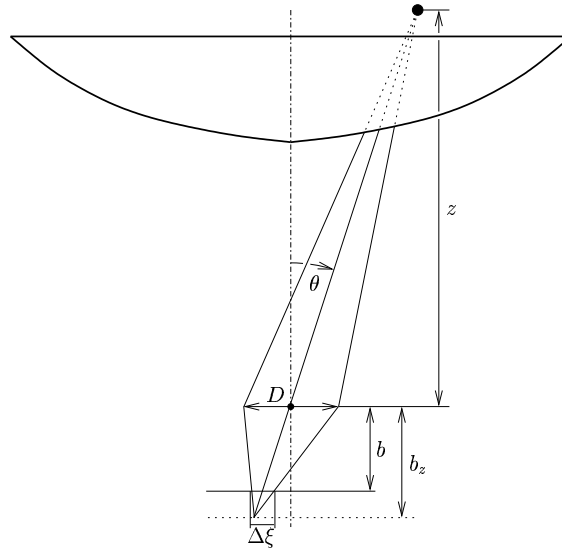


Abbildung 3.12: Weicht die eingestellte Bildweite b von der Bildweite b_z eines virtuellen Bildpunkts (mit Entfernung z vom Kameraknotenpunkt) ab, so führt dies zu einer unscharfen Abbildung in einem Bildbereich der Größe $\Delta\xi$, der linear von der Blendenöffnung D anhängt.

D ist der Durchmesser der Blendenöffnung; b_z die Bildweite, die der Entfernung z eines virtuellen Bildpunkts zum Kameraknotenpunkt entspricht. Mit Hilfe der Linsengleichung (für dünne Linsen) gilt

$$\frac{1}{z} + \frac{1}{b_z} = \frac{1}{f} \quad \text{und} \quad \frac{1}{g} + \frac{1}{b} = \frac{1}{f}, \quad (3.127)$$

wobei g die der eingestellten Bildweite entsprechende Gegenstandsweite bezeichnet. Eingesetzt in (3.126) folgt

$$\frac{\Delta\xi}{D} = 1 - \left(\frac{b}{f} - \frac{b}{z} \right) = \frac{f(g-z)}{z(g-f)}. \quad (3.128)$$

Bezeichnet $\Delta z := g - z$ die Abweichung von der fokussierten Ebene, so gilt für $f \ll g$:

$$\frac{\Delta\xi}{D} \approx \frac{f\Delta z}{zg}. \quad (3.129)$$

Eine kurze Brennweite ist somit von Vorteil, erschwert aber wegen $|\Delta(\frac{\Delta\xi}{D})| \approx \frac{|\Delta b|}{f}$ eine genaue Einstellung für eine spezielle Gegenstandsweite.

Wird z in (3.128) gleich der z -Komponente von \mathbf{f}_θ bzw. \mathbf{f}_φ gesetzt, so erhält man die Größe des Unschärfbereichs in θ - bzw. φ -Richtung, $\Delta\xi_\theta$ bzw.

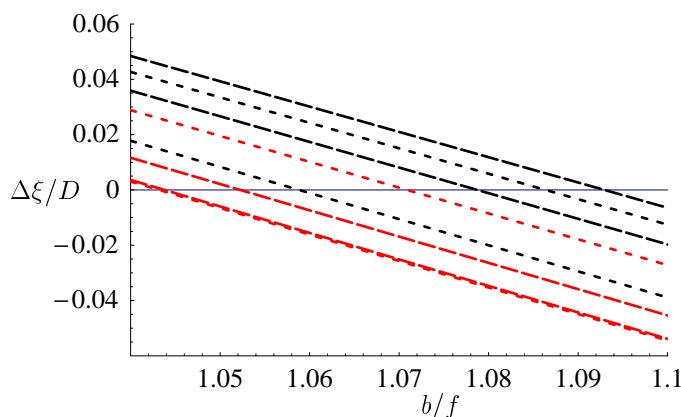


Abbildung 3.13: Größe der Unschärfebereiche $\Delta\xi_\theta$ (gestrichelt) und $\Delta\xi_\varphi$ (punktiert) für beide Spiegelflächen des Koala-Stereosensors in Abhängigkeit von der eingestellten Bildweite b (für den oberen Spiegel zusätzlich rot markiert). Von links nach rechts sind dargestellt: $D^{-1}\Delta\xi_{\varphi,1}(b, \theta = 15^\circ)$ und (davon kaum zu unterscheiden) $D^{-1}\Delta\xi_{\theta,1}(b, \theta = 15^\circ)$, $D^{-1}\Delta\xi_{\varphi,1}(b, \theta = 0^\circ)$, $D^{-1}\Delta\xi_{\varphi,2}(b, \theta = 25^\circ)$, $D^{-1}\Delta\xi_{\varphi,1}(b, \theta = 0^\circ)$, $D^{-1}\Delta\xi_{\theta,2}(b, \theta = 25^\circ)$, $D^{-1}\Delta\xi_{\varphi,2}(b, \theta = 15^\circ)$, $D^{-1}\Delta\xi_{\theta,2}(b, \theta = 15^\circ)$.

$\Delta\xi_\varphi$. Für den Stereosensor des Koala-Roboters zeigt Abbildung 3.13 $\Delta\xi_\theta$ und $\Delta\xi_\varphi$ für beide Spiegelflächen in Abhängigkeit der eingestellten Bildweite b für verschiedene Werte des Kamerawinkels θ . Die Brennweite der Kameralinse beträgt $f = 4.8$ mm. Die der „Spitze“ des oberen Spiegels ($z \approx 72.5$ mm) entsprechende Bildweite ist $b \approx 1.071f$. Dies ist näherungsweise gleich dem Wert $b^* = 1.067f$, an dem $D^{-1}\Delta\xi_{\theta,2}(b^*, \theta = 15^\circ) = -D^{-1}\Delta\xi_{\varphi,1}(b^*, \theta = 15^\circ)$ gilt, so dass für $b \approx b^*$ ein minimaler Unschärfebereich zu erwarten ist. Für preisgünstige Fixfokus-Objektive (wie das verwendete Kameraobjektiv), die üblicherweise auf „unendliche“ Objektweite, d.h. Bildweite $b = f$, eingestellt sind, bedeutet dies, dass der Abstand zur Bildebene z.B. durch einen dünnen Zwischenring vergrößert werden muss.

Bei einer Blendenöffnung von $D = 1$ mm entspricht $D^{-1}\Delta\xi_{\theta,2}(b^*, \theta = 15^\circ) \approx 0.023$ in etwa $\Delta\xi \approx 23 \mu\text{m}$. Dies ist bereits deutlich größer als die CCD-Zellgröße der verwendeten Videokamera, die bei ca. $10 \mu\text{m}$ liegt. Trotz der deutlich verbesserten Abbildungsqualität im Vergleich zum Stereosensor des Khepera-Roboters ist in dunklen Innenräumen, in denen die Blende weit geöffnet werden muss, eine geringfügige Bildunschärfe zu erkennen.

Bei der Einstellung der Brennweite ist zu berücksichtigen, dass für die Stereobildauswertung beide Bildbereiche auf die gleiche Auflösung transformiert werden (vgl. Abbildung 3.8), und deshalb eine leichte Unschärfe im äußeren Bildbereich besser toleriert werden kann.

Bemerkung: Mit Hilfe von Gleichung (3.128) erhält man eine Abschätzung für die Fläche des Unschärfebereichs:

$$\Delta A \approx |\Delta\xi_\theta| |\Delta\xi_\varphi| \stackrel{g \gg f}{\approx} \frac{D^2 f^2 |g - z_\varphi| |g - z_\theta|}{z_\varphi z_\theta g^2}, \quad (3.130)$$

wobei z_θ und z_φ die z -Komponenten von \mathbf{f}_θ und \mathbf{f}_φ bezeichnen. Variiert man die Objektweite g , so findet man zwei Minima von ΔA , in denen $|\Delta\xi_\theta(g_\theta)| = 0$ bzw. $|\Delta\xi_\varphi(g_\varphi)| = 0$ gilt. Die Form der Kurve ist in guter Übereinstimmung mit den in [5] und [69] angegebenen Resultaten von numerischen Berechnungen für eine hyperbolische Spiegelfläche.

Kapitel 4

Binokulares Stereo¹

In zahlreichen neurophysiologischen Studien wurden im visuellen Kortex von Katzen und Affen binokulare Nervenzellen untersucht, deren Aktivität eindeutig von der Disparität des visuellen Stimulus abhängt („Disparitäts-selektive Neurone“), siehe z.B. [3, 4, 49]. Zur Erklärung der experimentell gefundenen Eigenschaften dieser Zellen wurde das so genannte „binokulare Energiemodell“ entwickelt [70, 77]. In diesem Kapitel wird ein aktives Stereokamerasystem vorgestellt, das basierend auf Disparitäts-selektiven Energieneuronen Vergenzbewegungen ausführt und lokale Disparitäten bestimmt. Im Vergleich zu den panoramischen Stereosensoren aus Kapitel 1 und 3 wird infolge des begrenzten Sichtfeldes und der verbesserten Auflösung eine deutlich realistischere Modellierung des Stereosehens bei Säugetieren möglich. In Hinblick auf Landmarken-basierte Navigation ist die Zielsetzung, nach Implementierung des Systems auf einem mobilen Roboter die Aktivität der Neuronenpopulation als Ortssignatur zu verwenden und damit eine Rückkehr zu bekannten Orten zu ermöglichen.

Zunächst soll eine kurze Einführung in die Modellierung Disparitäts-selektiver Neurone gegeben werden.

4.1 Binokulares Energiemodell

Die im visuellen Kortex (vor allem in den Bereichen V1 und V2, bzw. entsprechenden Arealen 17 und 18 im Großhirn der Katze) gefundenen binokularen Einfachzellen (engl. „simple cells“) besitzen monokulare rezeptive Felder, die näherungsweise durch Gaborfunktionen (Produkt aus trigonometrischer und Exponential-Funktion) beschrieben werden können. Ein Phasenversatz $\varphi_l - \varphi_r$ zwischen dem linken und rechten Feld kann auftreten, ihre Mitten-

¹Teile der Abschnitte 4.1 und 4.2 wurden bereits in [90] veröffentlicht.

oder Zentralfrequenzen sind jedoch vergleichbar ($\nu_l = \nu_r$). Für unterschiedliche Zellen können Zentralfrequenz und Vorzugsorientierung variieren. Infolge des seitlichen Versatzes der beiden Augen bzw. Kameras treten in den Bildern horizontale Disparitäten (Bildverschiebungen) auf. Bei nicht-parallelen Blickachsen findet man zusätzlich vertikale Disparitäten. Da diese (bei sich schneidenden Achsen) vergleichsweise kleine Beträge besitzen, werden im Folgenden nur vertikal orientierte Felder betrachtet, die für die Bestimmung horizontaler Disparitäten am Besten geeignet sind. Als weitere Vereinfachung werden ausschließlich kreisförmige rezeptive Felder verwendet, deren Ausdehnung durch einen einzelnen Parameter σ (Breite der Gaußkurve) bestimmt wird.

Die mathematische Beschreibung für das linke und rechte rezeptive Feld einer *Einfachzelle* mit Zentralfrequenz ν lautet unter diesen Voraussetzungen:

$$f_{al\nu}(x, y) := \text{rf}(x, y, \nu, \varphi_l) , \quad (4.1)$$

$$f_{ar\nu}(x, y) := \text{rf}(x, y, \nu, \varphi_r) , \quad (4.2)$$

$$\text{rf}(x, y, \nu, \varphi) := \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \cos(2\pi\nu x + \varphi) . \quad (4.3)$$

Durch Hinzunahme einer weiteren Einfachzelle mit um 90° phasenverschobenen rezeptiven Feldern,

$$f_{bl\nu}(x, y) := \text{rf}(x, y, \nu, \varphi_l - \frac{\pi}{2}) , \quad (4.4)$$

$$f_{br\nu}(x, y) := \text{rf}(x, y, \nu, \varphi_r - \frac{\pi}{2}) , \quad (4.5)$$

entsteht ein so genanntes Quadraturpaar. Die Aktivität der Einfachzellen ist durch die lineare Filterung der Bilder gegeben, d.h.

$$S_{a\nu} := \iint f_{al\nu}(\xi, \eta) I_l(\xi, \eta) + f_{ar\nu}(\xi, \eta) I_r(\xi, \eta) d\xi d\eta , \quad (4.6)$$

$$S_{b\nu} := \iint f_{bl\nu}(\xi, \eta) I_l(\xi, \eta) + f_{br\nu}(\xi, \eta) I_r(\xi, \eta) d\xi d\eta . \quad (4.7)$$

Wie Abbildung 4.1 zeigt, quadrieren und addieren binokulare *Komplexzellen* die Ausgabe der Einfachzellen:

$$C_\nu := S_{a\nu}^2 + S_{b\nu}^2 . \quad (4.8)$$

Da biologische Neurone keine negativen Aktivitäten besitzen, kann Gleichung (4.8) durch insgesamt vier Zellen ($S_{a\nu}^+$, $S_{a\nu}^-$ und $S_{b\nu}^+$, $S_{b\nu}^-$) mit nicht-negativen Aktivitäten realisiert werden, wobei zwei Zellen jeweils komplementäre rezeptive Felder (d.h. exzitatorische und inhibitorische Bereiche sind vertauscht) besitzen, so dass gilt:

$$S_{a/b\nu} = S_{a/b\nu}^+ - S_{a/b\nu}^- \implies S_{a/b\nu}^2 = (S_{a/b\nu}^+)^2 + (S_{a/b\nu}^-)^2 . \quad (4.9)$$

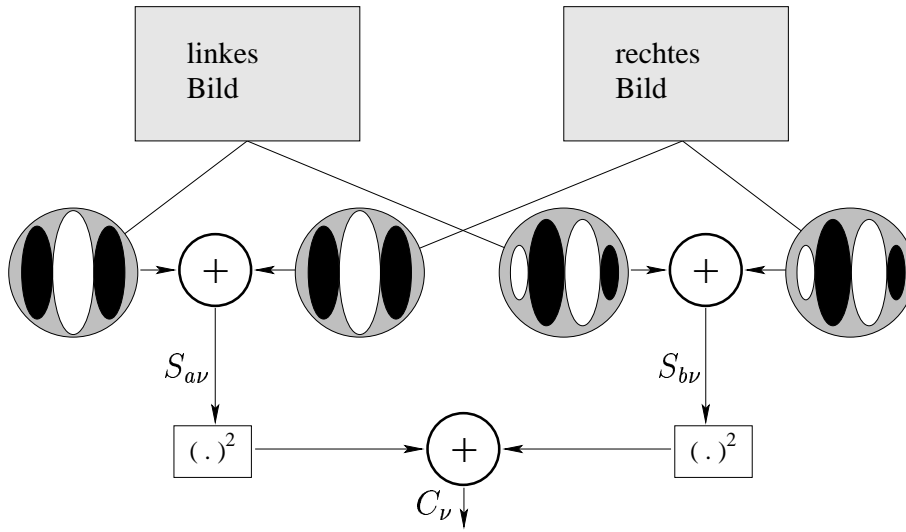


Abbildung 4.1: Schematische Darstellung des binokularen Energiemodells: Die beiden Bilder werden zunächst jeweils mit zwei orthogonalen Gaborfunktionen mit 90° Phasenverschiebung gefiltert. Die Aktivität der binokularen Einfachzellen, die die geraden bzw. ungeraden Filterantworten aufsummieren, werden nach Gleichrichtung (Quadrieren) durch eine Komplexzelle zusammengefasst, siehe Gleichung (4.8).

($S_{av}^+ S_{av}^- = S_{bv}^+ S_{bv}^- = 0$, da Zellen mit komplementären rezeptiven Felder nie gleichzeitig aktiv sind).

Durch die Definition der in der Literatur häufig verwendeten komplexwertigen monokularen Filterfunktionen

$$q_{lv}(x, y, \varphi_l) := f_{alv}(x, y) + i f_{blv}(x, y) \quad (4.10)$$

$$= \exp\left(-\frac{x^2 + y^2}{2\sigma^2} + i(2\pi\nu x + \varphi_l)\right), \quad (4.11)$$

$$q_{rv}(x, y, \varphi_r) := f_{arv}(x, y) + i f_{brv}(x, y) \quad (4.12)$$

$$= \exp\left(-\frac{x^2 + y^2}{2\sigma^2} + i(2\pi\nu x + \varphi_r)\right), \quad (4.13)$$

kann die Antwort einer Komplexzelle am Ort (x, y) formuliert werden als

$$C_\nu(x, y, \varphi_l, \varphi_r) = \left(\operatorname{Re}[Q_{lv}(x, y, \varphi_l)] + \operatorname{Re}[Q_{rv}(x, y, \varphi_r)]\right)^2 + \left(\operatorname{Im}[Q_{lv}(x, y, \varphi_l)] + \operatorname{Im}[Q_{rv}(x, y, \varphi_r)]\right)^2 \quad (4.14)$$

$$= |Q_{lv}(x, y, \varphi_l) + Q_{rv}(x, y, \varphi_r)|^2, \quad (4.15)$$

$$Q_{l\nu}(x, y, \varphi_l) := \iint q_{l\nu}(x + \xi, y + \eta, \varphi_l) I_l(\xi, \eta) d\xi d\eta, \quad (4.16)$$

$$Q_{r\nu}(x, y, \varphi_r) := \iint q_{r\nu}(x + \xi, y + \eta, \varphi_r) I_r(\xi, \eta) d\xi d\eta. \quad (4.17)$$

Um eine Komplexzelle mit Vorzugsdisparität D zu beschreiben, existieren zwei unterschiedliche Modelle, die auch kombiniert werden können [2, 35]:

- Phasen-Shift-Modell: linkes und rechtes rezeptives Feld haben unterschiedliche Phasen $\varphi_l - \varphi_r = 2\pi\nu D$, so dass gilt (φ beliebig):

$$C_{D\nu}(x, y) = |Q_{l\nu}(x, y, \varphi + 2\pi\nu\frac{D}{2}) + Q_{r\nu}(x, y, \varphi - 2\pi\nu\frac{D}{2})|^2. \quad (4.18)$$

- Positions-Shift-Modell: linkes und rechtes rezeptives Feld haben leicht versetzte Positionen, d.h.

$$C_{D\nu}(x, y) = |Q_{l\nu}(x + \frac{D}{2}, y, \varphi) + Q_{r\nu}(x - \frac{D}{2}, y, \varphi)|^2. \quad (4.19)$$

Wie z.B. in [77] diskutiert wird, besitzen Komplexzellen vom Phasen-Shift-Typ einen eingeschränkten Disparitätsbereich $D \in [-\frac{\pi}{\nu}, \frac{\pi}{\nu}]$. Folglich können nur Neurone mit niedriger Zentralfrequenz ν große Disparitäten eindeutig detektieren. Psychophysische Untersuchungen [60] zeigen jedoch, dass auch hohe Frequenzen bei der Wahrnehmung großer Bildverschiebungen beteiligt sind, ohne dass eine so genannte „coarse-to-fine“-Strategie angewandt wird. Für die im Folgenden beschriebene Anwendung wurde deshalb das Positions-Shift-Modell gewählt.

Die Verwendung von Gaborfiltern hat im Vergleich zu „klassischen“ Korrelations-basierten Stereoalgorithmen den Vorteil, dass sie sich vergleichsweise robust gegenüber kleinen Fehlstellungen der Kameraachsen und Bildverzerrungen verhalten [83]. Ebenso wird eine explizite Lösung des Korrespondenzproblems umgangen.

Im nächsten Abschnitt werden Einzelheiten der Implementierung des binokularen Energiemodells sowie Erweiterungen um zusätzliche Normalisierungstufen dargestellt. Die Berechnung von Bilddisparitäten aus den lokal aktivsten Komplexzellen wird in 4.2.4 beschrieben.

4.2 Implementierung des Energiemodells²

Es werden Filter $q_k(x_j, y_j)$ für vier Frequenzkanäle ($k = 1, 2, 3, 4$) mit einer Bandbreite von jeweils zwei Oktaven verwendet, siehe Abbildung 4.2. Mit

²Die in diesem Abschnitt beschriebene Implementierung des binokularen Energiemodells auf ein Stereokamerasystem wurde zum Teil im Rahmen einer Diplomarbeit durchgeführt [47].

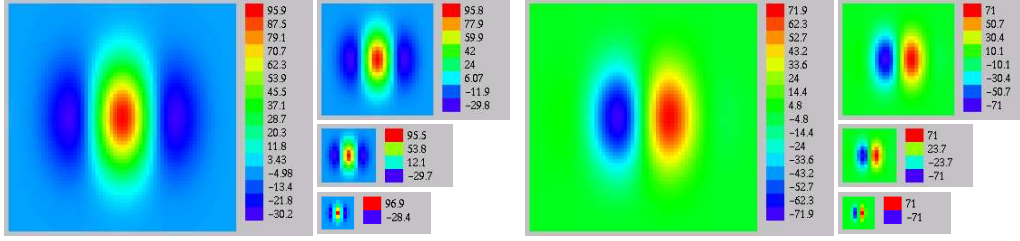


Abbildung 4.2: Cosinus- und Sinus-Gaborfilter mit einer Bandbreite von 2 Oktaven für vier Zentralfrequenzen $\nu_k \in \{5\nu_0, 10\nu_0, 20\nu_0, 40\nu_0\}$.

der Definition der Bandbreite,

$$\text{bw} := \log_2\left(\frac{\nu + \varsigma}{\nu - \varsigma}\right) \iff \frac{\varsigma}{\nu} = \frac{2^{\text{bw}} - 1}{2^{\text{bw}} + 1}, \quad (4.20)$$

bedeutet dies $\frac{\varsigma_k}{\nu_k} = \frac{3}{5}$. Die Größe ς_k definiert im Frequenzraum die Breite der um ν_k verschobenen Gaußfunktion $\frac{1}{2\pi\varsigma_k} \exp[-\frac{1}{2\varsigma_k^2}((\nu_x - \nu_k)^2 + \nu_y^2)]$, die die Fouriertransformierte der (komplexen) Gaborfunktion $\exp[-\frac{1}{2\sigma_k^2}(x^2 + y^2) + i2\pi\nu_k x]$ darstellt. Die Ausdehnung der rezeptiven Felder ist über $\sigma_k = (2\pi\varsigma_k)^{-1}$ festgelegt. Die gewählten Zentralfrequenzen sind $\nu_k = 2^{k-1}5\nu_0$ mit $\nu_0 := N_x^{-1}$, wobei N_x die Bildbreite ist. Die Frequenzabhängigkeit der Filterfunktionen ist in Abbildung 4.3 gezeigt.

Die Filterkoeffizienten werden nach Abtastung der Gaborfunktion so angepasst, dass ihre Antwort für ein konstantes Bild verschwindet, d.h.

$$\sum_{x,y} q_k(x,y) = 0.$$

Da die Gaborfunktionen für die gewählte vertikale Orientierung in x und y separierbar sind, kann die Berechnung von (4.16) und (4.17) als Verkettung von zwei eindimensionalen Filterungen dargestellt werden:

$$Q_\nu(x,y) = \int \exp\left(-\frac{(x+\xi)^2}{2\sigma^2} + i2\pi\nu(x+\xi)\right) \tilde{I}(\xi,y) d\xi, \quad (4.21)$$

$$\tilde{I}(\xi,y) := \int \exp\left(-\frac{(y+\eta)^2}{2\sigma^2}\right) I(\xi,\eta) d\eta. \quad (4.22)$$

Im diskreten Fall reduziert dies die Komplexität von $O(N_x N_y M_x M_y)$ auf $O[N_x N_y (M_x + M_y)]$ Multiplikationen und Additionen bei einem Bild von $N_x \times N_y$ Pixel und einer Filtergröße von $M_x \times M_y$. Für die Beschleunigung von Filteroperationen mit nicht separierbaren Filtern siehe [74].

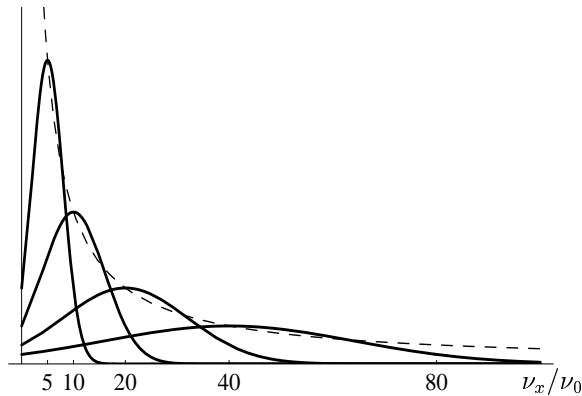


Abbildung 4.3: Die Gaborfilter entsprechen im Frequenzraum Gaußfunktionen der Breite $\varsigma_k = (2\pi\sigma)^{-1} = \frac{3}{5}\nu_k$.

Bemerkung: Der Filterungsvorgang kann weiter beschleunigt werden, indem – statt viermaliger Filterung desselben Bildes – der Filter niedrigster Auflösung ($k = 4$, höchste Frequenz) auf vier skalierte Versionen des Bildes angewendet wird. Allerdings können dann Bildverschiebungen von binokularen Komplexzellen mit tieferen Zentralfrequenzen ($k = 1, 2, 3$) nicht bzw. nur durch Interpolation in voller Auflösung bestimmt werden.

4.2.1 Monokulare Normalisierung

Um den Einfluss „interokularer“ Kontrastdifferenzen (z.B. infolge unterschiedlicher Blendeneinstellung der beiden Kameras) zu verringern und die Stärke der Filterantworten der unterschiedlichen Frequenzkanäle anzugleichen, werden die Filterausgänge vor Berechnung der Komplexzellantwort normalisiert (N_x, N_y Bildbreite und -höhe):

$$NQ_k(x, y) := \frac{Q_k(x, y)}{(N_x N_y)^{-1} \sum_{x,y} |Q_k(x, y)|} . \quad (4.23)$$

4.2.2 Kombination der Frequenzkanäle

Für die Komplexzellaktivität nach dem Positions-Shift-Modell gilt, vgl. (4.19),

$$C_{Dk}(x, y) = |NQ_{lk}(x + \frac{D}{2}, y) + NQ_{rk}(x - \frac{D}{2}, y)|^2 . \quad (4.24)$$

Diese hängt nicht nur von den Disparitäten der Bildbereiche $d(x, y)$ sondern auch vom lokalen Kontrast ab. Bei der Kombination von Komplexzellen aus den unterschiedlichen Frequenzkanälen mit gleicher Vorzugsdisparität wird

deshalb eine weitere (lokale) Normalisierung durchgeführt. Die Größe

$$NC_D(x, y) := \frac{\sum_k C_{Dk}(x, y)}{\sum_k (|NQ_{lk}(x + \frac{D}{2}, y)|^2 + |NQ_{rk}(x - \frac{D}{2}, y)|^2 + \epsilon_k)} \quad (4.25)$$

wird im Folgenden als normalisierte Komplexzellantwort bezeichnet. Die Konstante ϵ_k ist nur für den Fall sehr geringe lokaler Bildkontraste $\sum_k |NQ_{lk}|^2 + |NQ_{rk}|^2 \approx 0$ relevant (in der aktuellen Implementierung wurde $\epsilon_k = 1$ gesetzt). Aus (4.24) und (4.25) folgt unmittelbar $NC_D \geq 0$. Durch Einsetzen von (4.24) in (4.25) kann zudem $NC_D \leq 2$ gezeigt werden:

$$NC_D = \frac{\sum_k |NQ_{lk} + NQ_{rk}|^2}{\sum_k (|NQ_{lk}|^2 + |NQ_{rk}|^2 + \epsilon_k)} \quad (4.26)$$

$$= \frac{\sum_k 2|NQ_{lk}|^2 + 2|NQ_{rk}|^2 - |NQ_{lk} - NQ_{rk}|^2}{\sum_k (|NQ_{lk}|^2 + |NQ_{rk}|^2 + \epsilon_k)} \quad (4.27)$$

$$= \frac{2 \sum_k |NQ_{lk}|^2 + |NQ_{rk}|^2}{\sum_k (|NQ_{lk}|^2 + |NQ_{rk}|^2 + \epsilon_k)} \left(1 - \frac{\sum_k |NQ_{lk} - NQ_{rk}|^2}{2 \sum_k |NQ_{lk}|^2 + |NQ_{rk}|^2} \right) \quad (4.28)$$

$$\approx 2 \left(1 - \frac{\sum_k |NQ_{lk} - NQ_{rk}|^2}{2 \sum_k |NQ_{lk}|^2 + |NQ_{rk}|^2} \right) \text{ für } \epsilon_k \ll |NQ_l|^2 + |NQ_r|^2. \quad (4.29)$$

Anhand von (4.28) bzw. (4.29) wird deutlich, dass NC_D seinen Maximalwert erreicht, falls $NQ_{lk} = NQ_{rk}$ für alle k gilt. Für nur einen Frequenzkanal bedeutet $NQ_{lk} \approx NQ_{rk}$ nicht automatisch, dass die linken und rechten rezeptiven Felder korrespondierende Bildbereiche abtasten, da in seltenen Fällen unterschiedliche Bilder zu ähnlichen NQ_k -Werten führen können. Durch Kombination der verschiedenen Frequenzkanäle wird jedoch die Wahrscheinlichkeit einer „falschen“ Antwort weiter reduziert.

Die Tuningkurve³ einer normalisierten Komplexzelle in Abhängigkeit von der wahren Bilddisparität d zeigt Abbildung 4.4 a. Das deutliche Maximum und die geringe Varianz im Bereich der Vorzugsdisparität ($d \approx D = 0$) erlauben eine zuverlässige Detektion kleiner Bildverschiebungen. Dies ist für die lediglich Frequenz-gemittelten Komplexzelle (Abb. 4.4 b),

$$C_D(x, y) := \sum_k C_{Dk}(x, y), \quad (4.30)$$

wesentlich schwieriger. Eine weiterführende Untersuchung normalisierter und nicht-normalisierter Komplexzellantworten erfolgt in Abschnitt 4.3.

³Die Tuningkurve beschreibt die Abhängigkeit der Aktivität („Feuerrate“) eines Neurons von einem Stimulusparameter.

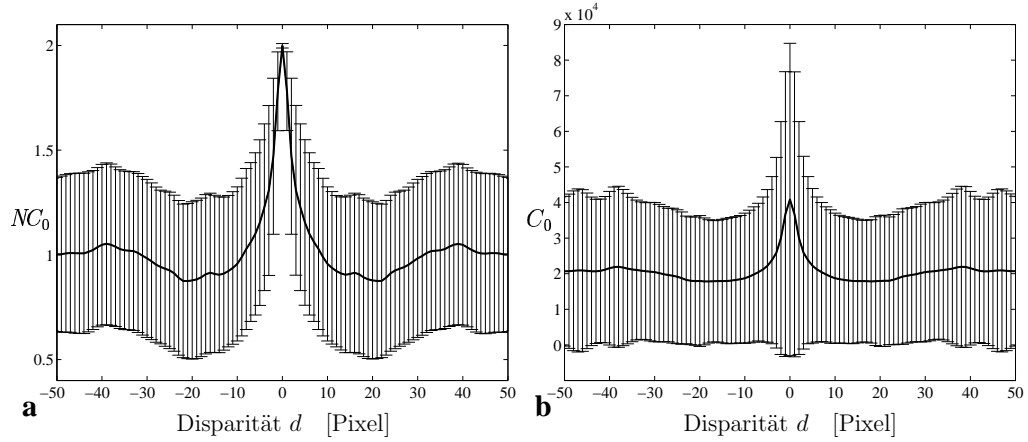


Abbildung 4.4: Tuningkurven einer normalisierten (a) und nicht-normalisierten Komplexzelle (b), Gleichungen (4.25) und (4.30) in Abhängigkeit von der Bilddisparität d , Vorzugsdisparität ist $D = 0$. Gezeigt sind die Mittelwerte (kontinuierliche Kurven) und Standardabweichungen (Fehlerbalken) der Zellantworten für jeweils 100000 Stereostimuli. Diese wurden generiert, indem aus 10 natürlichen Bildern an unterschiedlichen Positionen um d verschobene Ausschnitte extrahiert wurden.

4.2.3 Steuerung der Vergenzstellung

Das verwendete Stereokamerasystem ist in Abbildung 4.5 dargestellt. Es kann den Winkel zwischen den Kameraachsen (die so genannte „Vergenz“) aktiv einstellen. Die Vergenzsteuerung erfolgt so, dass die globale Bilddisparität minimiert wird. Dies ist in Übereinstimmung mit experimentellen Studien beim Menschen [61].

Die globale Bilddisparität wird durch die maximale Aktivität von Einheiten bestimmt, die im Folgenden als Vergenz-steuernde Neuronen bezeichnet werden:

$$D_{\text{glob}}^{\text{est}} := \arg \max_D VC_D, \quad (4.31)$$

$$VC_D := (N_x N_y)^{-1} \sum_{x,y} NC_D(x,y). \quad (4.32)$$

Für die Vorzugsdisparitäten der Vergenz-steuernden Neuronen wurden $D \in \{0, \pm 1, \pm 2, \pm 3, \pm 5, \pm 7, \pm 10, \pm 14, \pm 19, \pm 25, \pm 32, \pm 40, \pm 49, \pm 59 \text{ Pixel}\}$ verwendet, mit feinerer Abtastung für kleine Disparitäten.

Solange der geschätzte Wert $D_{\text{glob}}^{\text{est}}$ von Null abweicht, wird durch einen Schrittmotor, der eines der zwei ineinander greifenden Zahnräder antreibt, die Vergenzstellung symmetrisch verändert ($|\varphi_l| = \varphi_r = \frac{1}{2}\alpha$, siehe Abbildung 4.5). Damit der Basisabstand des Stereosystems b bei einer Vergenzbewegung

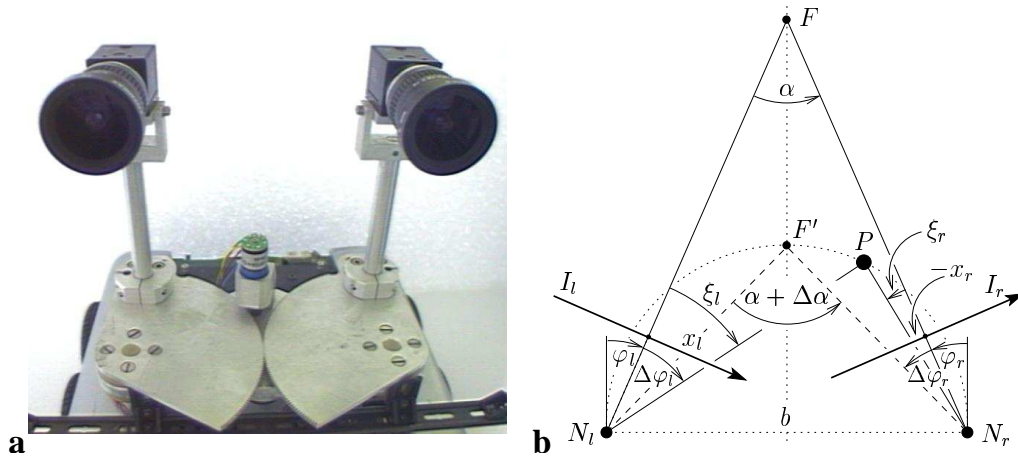


Abbildung 4.5: **a:** Stereokamerakopf mit symmetrischer Vergenzsteuerung. Die Vergenzstellung der beiden Monochrom-Kameras (Öffnungswinkel $\approx 80^\circ$) kann über einen Schrittmotor, der das linke Zahnrad antreibt, symmetrisch geändert werden. Die Stereobasis b des Systems beträgt 14.5 cm. **b:** Der Vergenzwinkel α ist durch den Schnittwinkel der Kameraachsen festgelegt. Nach einer (symmetrischen) Vergenzbewegung $F \rightarrow F'$ besitzen alle Punkte mit einer anfänglichen (Winkel-)Disparität $\delta = \xi_r - \xi_l$ die Disparität $\delta' \approx 0^\circ$. Die Positionen dieser Punkte befinden sich auf dem so genannten Vieth-Müller-Kreis (punktiert), der durch die Knotenpunkte der Kameras ($N_{l/r}$) und den Fixationspunkt F' verläuft, siehe [59].

unverändert bleibt, wurden die Kameras so befestigt, dass ihre Knotenpunkte über dem jeweiligen Drehpunkt zu liegen kommen. Die Änderung des Vergenzwinkels ist näherungsweise proportional zur globalen Bilddisparität:

$$\Delta\alpha \approx k_\alpha D_{\text{glob}}^{\text{est}} . \quad (4.33)$$

Für das Stereokamerasystem mit einer verwendeten Bildauflösung von 192×124 Pixel ist $k_\alpha \approx 0.5^\circ$ (je Pixel Disparität). Da die Vergenz fortwährend geregelt wird, ist eine exakte Bestimmung von k_α nicht nötig.

Gleichung (4.33) kann wie folgt begründet werden, siehe Abb. 4.5 b: In der anfänglichen Vergenzstellung α mit Fixationspunkt F besitzt der Punkt P die Winkeldisparität $\delta = \xi_r - \xi_l$. Alle Punkte, die auf demselben Vieth-Müller-Kreis liegen und somit auch der Punkt auf der Mittelsenkrechten F' , haben die gleiche Disparität [59]. Es gilt deshalb $\Delta\varphi_r - \Delta\varphi_l = \delta$. Um für F' und P die Disparität $\delta' = 0^\circ$ zu erreichen, muss folglich eine Vergenzbewegung $\Delta\alpha = \Delta\varphi_r - \Delta\varphi_l = \delta$ durchgeführt werden. Da der Kamerawinkel durch $\xi_{l/r} = -\arctan(\frac{k_p}{f} x_{l/r})$ aus der Pixelposition bestimmt werden kann (f ist die

Brennweite der Kameralinsen, k_p eine Konstante zur Umrechnung der Pixelzahl in Längeneinheiten), folgt $\Delta\alpha = \xi_r - \xi_l = \arctan(\frac{k_p}{f}x_l) - \arctan(\frac{k_p}{f}x_r) \approx \frac{k_p}{f}(x_l - x_r) = \frac{k_p}{f}d$.

Man beachte, dass Bilddisparitäten häufig mit entgegengesetztem Vorzeichen definiert werden. Für parallele Kameraachsen ($\alpha = 0$) treten bei der hier verwendete Definition, $d := x_l - x_r$, ausschließlich positive Disparitäten auf.

4.2.4 Bestimmung von Disparitäten

Der Vorteil der Vergenzsteuerung, die durch Minimieren der globalen Bilddisparität erfolgt, besteht für das Stereosystem darin, dass die Bestimmung der lokalen Bilddisparitäten auf kleine Werte beschränkt werden kann.⁴ Als Vorzugsdisparitäten wurden deshalb

$$D \in \{0, \pm 1, \pm 2, \dots, \pm 10\} \quad (4.34)$$

gewählt.

Aus biologischer Sicht ist eine explizite Bestimmung der Disparitäten unnötig, da in der Aktivität der Disparitäts-selektiven Neurone die für weitere Verarbeitungsschritte notwendige Information vollständig enthalten ist. Dies ist in manchen Fällen auch nicht sinnvoll, z.B. bei transparenten Schichten unterschiedlicher Tiefe, für die keine eindeutige Disparität angegeben werden kann.

Will man dennoch eine zweidimensionale Disparitätskarte aus der Neuronenpopulation extrahieren, etwa aus Gründen der besseren Darstellbarkeit oder um eine Reduktion der Datenmenge zu erreichen, kann dies durch Bestimmung der Neuronen mit der größten lokalen Aktivität geschehen,

$$D_{\text{loc}}^{\text{est}}(x, y) := \arg \max_D NC_D(x, y) \quad \forall x, y . \quad (4.35)$$

Zusätzlich kann als Maß für die Eindeutigkeit und Zuverlässigkeit, mit der die Disparität bestimmt werden konnte, ein Konfidenzwert gemäß

$$c(x, y) := \frac{\max_D NC_D(x, y)}{N_D^{-1} \sum_D (NC_D(x, y) + \epsilon_D)} \quad (4.36)$$

berechnet werden. $N_D = 21$ ist die Zahl der verwendeten Vorzugsdisparitäten und ϵ_D ist – analog zu ϵ_k in Gleichung – nur für den Fall sehr kleiner Komplexzellantworten von Bedeutung und wurde auf 1/4 der maximalen Zellantwort,

⁴Beim Menschen wird durch Vergenzbewegungen ein zu betrachtendes Objekt in den zentralen Sehbereich höchster Auflösung (Fovea) gebracht. Dieser könnte durch Verwendung einer „Log-Polar“-Transformation modelliert werden, siehe z.B. [9].

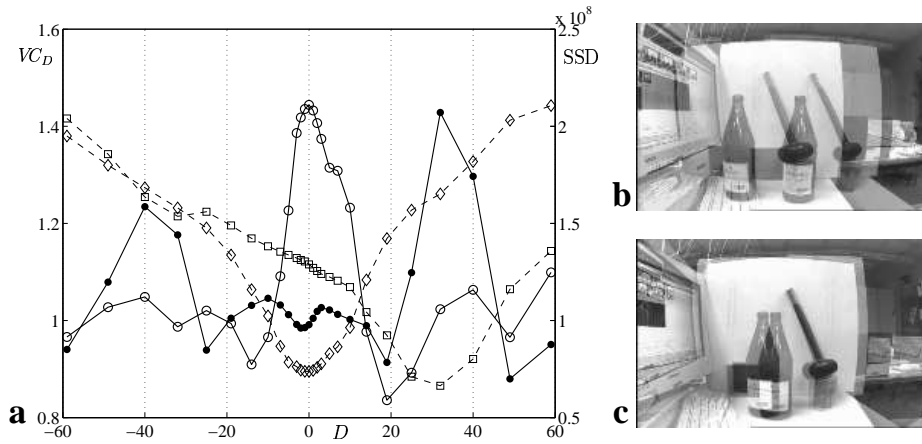


Abbildung 4.6: Beispiel einer Vergenzsteuerung: **a:** Aktivitätsverteilung der Vergenzsteuernden Neurone vor (markiert durch Punkte) und nach der Vergenzeinstellung (Kreise). Zum Vergleich ist als Maß für die Bildähnlichkeit die Summe der quadratischen Differenzen in Abhängigkeit von der globalen Disparität, $SSD(D) := \sum_{x,y} (I_l(x + \frac{1}{2}D, y) - I_r(x - \frac{1}{2}D, y))^2$, dargestellt (Quadrate bzw. Rauten). In der Ausgangssituation (siehe überlagertes linkes und rechtes Kamerabild in (b)) mit parallelen Kameraachsen ($\alpha = 0^\circ$) liegt die maximale Zellantwort (und die minimale SSD) bei $D_{\text{glob}}^{\text{est}} = 32$. Nach Einstellung des Vergenzwinkels auf circa 16° sind, wie die überlagerten Kamerabilder in (c) zeigen, die Bilddisparitäten deutlich reduziert und VC_0 ist maximal.

d.h. $\epsilon_D = 0.5$ gesetzt (dies ist mit dem Wertebereich in Abbildung 4.4 a zu vergleichen). Bildbereiche, in denen die wahre Disparität außerhalb des betrachteten Bereichs (4.34) liegt oder eine Bestimmung der Disparität (z.B. infolge von Verdeckungen) gar nicht möglich ist, werden besonders niedrige Konfidenzwerte aufweisen. $\frac{NC_D}{\sum_D NC_D}$ kann als Schätzung für die Wahrscheinlichkeit betrachtet werden, dass D die wirkliche Disparität ist.

4.2.5 Beispiel für die Funktionsweise des Systems

Abbildung 4.6 zeigt ein Beispiel für den Verlauf der Vergenzsteuerung. Die Ausgangssituation ist als Überlagerung des linken und rechten Kamerabildes dargestellt. Es sind große Disparitäten zu erkennen, die nach einer Vergenzbewegung von ungefähr 16° deutlich verringert sind und nun zum großen Teil in den Arbeitsbereich der lokalen Disparitätsdetektoren fallen.

Die mit Hilfe von (4.35) und (4.36) aus der Population normierter Komplexzellen extrahierten Karten der lokalen Disparitäten und Konfidenzwerte sind in den Abbildungen 4.7 c, d dargestellt. Eine einfach zu realisierende

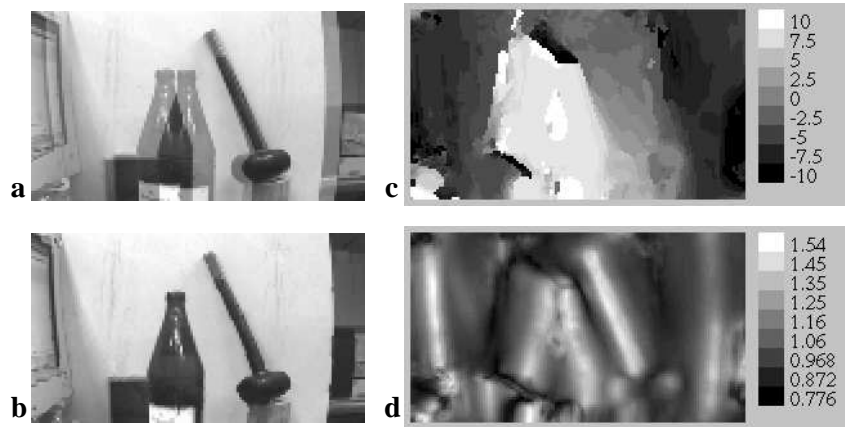


Abbildung 4.7: **a:** Überlagertes linkes und rechtes Kamerabild nach Einstellung des optimalen Vergenzwinkels von $\alpha \approx 16^\circ$ (entspricht Abb. 4.6 c). **b:** Mit Hilfe der Disparitätskarte (c) und Gleichung (4.37) können fast alle Doppelbilder fusioniert werden. **d:** Die Konfidenzkarte weist hohe Werte vor allem an kontrastreichen Konturen auf.

Möglichkeit, die Qualität der ermittelten Disparitätswerte zu beurteilen, bietet die Fusion des linken und rechten Kamerabildes mit Hilfe der Disparitätskarte, d.h.

$$I_{\text{fused}}(x, y) = \frac{1}{2} \left(I_l \left(x + \frac{1}{2} D_{\text{loc}}^{\text{est}}(x, y) \right) + I_r \left(x - \frac{1}{2} D_{\text{loc}}^{\text{est}}(x, y) \right) \right) \quad . \quad (4.37)$$

Vergleicht man das resultierende Bild (Abb. 4.7 b) mit dem überlagertem linken und rechten Bild nach der Vergenzeinstellung (Abb. 4.7 a), so erkennt man, dass fast alle Doppelbilder fusioniert werden konnten.

Nach Aufnahme der beiden Kamerabilder (8-Bit-Grauwerte) benötigt das System bei einer effektiven Bildauflösung von 192×124 Pixel etwa 400 ms auf einem AMD Athlon 700 MHz Prozessor zur Berechnung der normierten Komplexzellantworten und anschließender Bestimmung der Disparitätskarte.

Bemerkungen: In der vorgestellten Implementierung wurden Anzahl und Werte der Parameter (z.B. Zahl der Frequenzkanäle, Zentralfrequenzen, Vorzugsdisparitäten, etc.) fest vorgegeben („hard-wired“). Dies könnte durch Verwendung von Optimierungs- bzw. Lernalgorithmen verbessert werden. Neben einer größeren Effizienz ließe sich damit auch eine automatische Anpassung an die Hardware erreichen (z.B. Korrektur kleiner Fehlstellungen der Kameraachsen). Dies ist vor allem von Vorteil, wenn der Stereokamerakopf auf einem mobilen Kleinroboter („Koala“) eingesetzt wird. Als interne Re-

präsentationen bekannter Orte können dann Aktivitätsverteilungen der Disparitäts-selektiven Neuronenpopulation verwendet werden. Durch Vergleich der aktuellen Aktivitätsverteilung mit der im Ortsgedächtnis gespeicherten, soll – analog zu dem in Kapitel 1 beschriebenen Verfahren – ein Wiedererkennen und -finden bereits besuchter Orte realisiert werden. Die Effizienz der Zielfahrten könnte dann auch als Maß für die Güte der gewählten Parameter herangezogen werden.

Die im Vergleich zu Panoramabildern deutlich höhere Auflösung bietet kombiniert mit Tiefenschätzungen die Möglichkeit, einzelne Objekte zu definieren, (wieder-)zuerkennen und ggf. auch zu manipulieren. Die Verwendung einzelner Objekte im Gegensatz zu Ortssignaturen kann für Navigationsaufgaben von Vorteil sein, da sie zusätzliche Invarianzeigenschaften beispielsweise bezüglich des Standpunkts und der Blickrichtung besitzen und zudem einer verbalen bzw. symbolischen Beschreibung leicht zugänglich sind. Eine alternative Vergenzsteuerung, z.B. basierend auf so genannten „Saliency Maps“ [24], könnte sich dabei als hilfreich erweisen.

Außerdem kann das vorgestellte Kamerasystem im Gegensatz zu panoramischen Stereosensoren als Basis für ein Modellsystem des Stereosehens bei Säugetieren dienen. Vergleiche mit psychophysischen Experimenten bieten sich an. Eine Erweiterung um zusätzliche Verarbeitungsstufen ist hierfür sinnvoll. Desweiteren sind in der vorliegenden Implementierung keine Top-Down-Einflüsse modelliert worden, wenn man von der Vergenzsteuerung absieht, die nur indirekt über das veränderte Bild tieferliegende Verarbeitungsprozesse beeinflussen kann.

4.3 Wahrscheinlichkeitsdichten der Aktivität von Einfach- und Komplexzellen

Im Folgenden werden Wahrscheinlichkeitsdichten für normierte und nicht-normierte Komplexzellen des Positions-Shift-Modells berechnet. Es wird u.a. gezeigt, dass beide Zelltypen ähnliche Tuningkurven besitzen, die Varianz ihrer Aktivität jedoch sehr unterschiedlich ist. Die zugrunde liegende Frage ist dabei, inwieweit die neuronalen Antworten einzelner Zellen Rückschlüsse auf die tatsächliche Bilddisparität zulassen.

4.3.1 Nicht-normalisierte Zellen

Einfachzellen

Für die Aktivität der Einfachzellen gilt nach Gleichungen (4.6) und (4.7) für eine feste Zentralfrequenz (Index ν wird weggelassen):

$$S_a := \iint f_{al}(\xi, \eta) I_l(\xi, \eta) + f_{ar}(\xi, \eta) I_r(\xi, \eta) d\xi d\eta , \quad (4.38)$$

$$S_b := \iint f_{bl}(\xi, \eta) I_l(\xi, \eta) + f_{br}(\xi, \eta) I_r(\xi, \eta) d\xi d\eta . \quad (4.39)$$

Im Folgenden wird angenommen, dass die Generierung der Bilder durch einen Zufallsprozess beschrieben werden kann. Ist die Ausdehnung der rezeptiven Felder wesentlich größer als die Kohärenzlänge der Bilder, so sind S_a und S_b wegen des zentralen Grenzwertsatzes (siehe z.B. [11]) näherungsweise normalverteilt⁵. Folglich gilt für die gemeinsame Verteilungsdichte von S_a und S_b :

$$P_{S_a S_b}(S_a, S_b) = \frac{1}{2\pi \sqrt{\det \hat{\Sigma}_{ab}}} \exp\left[-\frac{1}{2}(\tilde{S}_a, \tilde{S}_b) \hat{\Sigma}_{ab}^{-1} (\tilde{S}_a, \tilde{S}_b)^\top\right] , \quad (4.40)$$

wobei die mittelwertfreien Größen $\tilde{S}_a := S_a - \langle S_a \rangle$, $\tilde{S}_b := S_b - \langle S_b \rangle$ definiert wurden. Die Kovarianzmatrix ist gegeben durch

$$\hat{\Sigma}_{ab} := \begin{pmatrix} \sigma_a^2 & \sigma_{ab}^2 \\ \sigma_{ab}^2 & \sigma_b^2 \end{pmatrix} , \quad (4.41)$$

⁵Dies gilt exakt für die weiter unten verwendeten Zufallsstereogramme. Für natürliche Bilder kann die Kohärenzlänge jedoch sehr groß werden [81]. In diesem Fall können Abweichungen von einer Normalverteilung auftreten.

mit $\sigma_a^2 := \langle \tilde{S}_a^2 \rangle = \langle S_a^2 \rangle - \langle S_a \rangle^2$, $\sigma_b^2 := \langle \tilde{S}_b^2 \rangle = \langle S_b^2 \rangle - \langle S_b \rangle^2$ und $\sigma_{ab}^2 := \langle \tilde{S}_a \tilde{S}_b \rangle = \langle S_a S_b \rangle - \langle S_a \rangle \langle S_b \rangle$. Ihre Inverse lautet

$$\hat{\Sigma}_{ab}^{-1} = \frac{1}{\det \hat{\Sigma}_{ab}} \begin{pmatrix} \sigma_b^2 & -\sigma_{ab}^2 \\ -\sigma_{ab}^2 & \sigma_a^2 \end{pmatrix} = \frac{1}{\sigma_a^2 \sigma_b^2 - (\sigma_{ab}^2)^2} \begin{pmatrix} \sigma_b^2 & -\sigma_{ab}^2 \\ -\sigma_{ab}^2 & \sigma_a^2 \end{pmatrix} .$$

Die Wahrscheinlichkeitsdichte einer einzelnen Einfachzelle S_a ergibt sich aus (4.40) durch „Abintegrieren“ von S_b ,

$$P_a(S_a) = \int P_{ab}(S_a, S_b) dS_b = \frac{1}{\sqrt{2\pi \sigma_a^2}} \exp\left[-\frac{(S_a - \langle S_a \rangle)^2}{2\sigma_a^2}\right] .$$

Komplexzellen

Nach dem binokularen Energiemodell ist die Antwort der Komplexzelle gegeben durch

$$C = S_a^2 + S_b^2 , \quad (4.42)$$

wobei die Einfachzellen ein Quadraturpaar bilden, siehe (4.1) – (4.5). Die quadratische Nichtlinearität ist in guter Übereinstimmung mit experimentellen Befunden [3]. In [65] wird jedoch gezeigt, dass diese auch bei Annahme einer linearer Übertragung mit Schwellwert durch neuronales Rauschen begründet werden kann. Ohne Rauschen wäre Gleichung (4.42) dann durch $C = [|S_a| - \theta_a]^+ + [|S_b| - \theta_b]^+$ zu ersetzen, wobei $[x]^+ := x$ für $x \geq 0$ und $[x]^+ := 0$ für $x < 0$ gilt und $\theta_{a/b}$ die jeweiligen Schwellwerte darstellen.

Im Folgenden wird angenommen, dass der „Offset“ (Erwartungswert) der Einfachzellantworten sehr klein ($\langle S_{a/b} \rangle \approx 0$) bzw. für die Aktivität der Komplexzelle irrelevant ist und somit näherungsweise

$$C = \tilde{S}_a^2 + \tilde{S}_b^2 = (S_a - \langle S_a \rangle)^2 + (S_b - \langle S_b \rangle)^2 \quad (4.43)$$

gilt. Berechnung des Erwartungswertes ergibt dann

$$\langle C \rangle = \sigma_a^2 + \sigma_b^2 . \quad (4.44)$$

Durch Diagonalisierung der Kovarianzmatrix mit Hilfe einer Orthogonalmatrix $\hat{\mathbf{O}}$ ($\hat{\mathbf{O}}^\top \hat{\mathbf{O}} = \hat{\mathbf{1}}$),

$$\hat{\Lambda}_{ab} := \text{diag}(\lambda_a^2, \lambda_b^2) = \hat{\mathbf{O}} \hat{\Sigma}_{ab} \hat{\mathbf{O}}^\top \quad (4.45)$$

folgt für die Wahrscheinlichkeitsdichte der Größen T_a, T_b , die durch $(T_a, T_b)^\top := \hat{\mathbf{O}} (\tilde{S}_a, \tilde{S}_b)^\top$ definiert sind:

$$\begin{aligned} P_{T_a T_b}(T_a, T_b) &= \frac{1}{2\pi \sqrt{\det \hat{\Lambda}_{ab}}} \exp\left[-\frac{1}{2}(T_a, T_b) \hat{\Lambda}_{ab}^{-1} (T_a, T_b)^\top\right] \\ &= \frac{1}{2\pi \lambda_a \lambda_b} \exp\left[-\frac{T_a^2}{2\lambda_a^2} - \frac{T_b^2}{2\lambda_b^2}\right] = P_{T_a}(T_a) P_{T_b}(T_b) . \end{aligned}$$

Mit

$$\det \hat{\mathbf{\Lambda}}_{ab} = \lambda_a^2 \lambda_b^2 = \det \hat{\mathbf{\Sigma}}_{ab} = \sigma_a^2 \sigma_b^2 - (\sigma_{ab}^2)^2, \quad (4.46)$$

$$\text{tr } \hat{\mathbf{\Lambda}}_{ab} = \lambda_a^2 + \lambda_b^2 = \text{tr } \hat{\mathbf{\Sigma}}_{ab} = \sigma_a^2 + \sigma_b^2 \quad (4.47)$$

erhält man für die Eigenwerte

$$\lambda_{a/b}^2 = \frac{1}{2} \left(\sigma_a^2 + \sigma_b^2 \pm \sqrt{(\sigma_a^2 - \sigma_b^2)^2 + 4(\sigma_{ab}^2)^2} \right). \quad (4.48)$$

Wegen $\hat{\mathbf{O}}^\top \hat{\mathbf{O}} = \hat{\mathbf{1}}$ gilt

$$C = \tilde{S}_a^2 + \tilde{S}_b^2 = T_a^2 + T_b^2. \quad (4.49)$$

Da T_a und T_b unabhängig normalverteilt sind, lässt sich aus (4.49) die Varianz von C unter Verwendung von (4.46), (4.47) und (4.44) berechnen:

$$\langle C^2 \rangle = \langle (\tilde{S}_a^2 + \tilde{S}_b^2)^2 \rangle = \langle T_a^4 \rangle + 2\langle T_a^2 \rangle \langle T_b^2 \rangle + \langle T_b^4 \rangle \quad (4.50)$$

$$= 3\lambda_a^4 + 2\lambda_a^2 \lambda_b^2 + 3\lambda_b^4 = 3(\text{tr } \hat{\mathbf{\Lambda}}_{ab})^2 - 4 \det \hat{\mathbf{\Lambda}}_{ab} \quad (4.51)$$

$$= 3(\sigma_a^2 + \sigma_b^2)^2 - 4[\sigma_a^2 \sigma_b^2 - (\sigma_{ab}^2)^2], \quad (4.52)$$

$$\text{Var}[C] = \langle C^2 \rangle - \langle C \rangle^2 = 2(\sigma_a^2 + \sigma_b^2)^2 - 4[\sigma_a^2 \sigma_b^2 - (\sigma_{ab}^2)^2] \quad (4.53)$$

$$= (\sigma_a^2 + \sigma_b^2)^2 + (\sigma_a^2 - \sigma_b^2)^2 + 4(\sigma_{ab}^2)^2 \quad (4.54)$$

$$= 2(\sigma_a^2)^2 + 2(\sigma_b^2)^2 + 4(\sigma_{ab}^2)^2. \quad (4.55)$$

Mit der Substitution $\sqrt{C}(\cos\phi, \sin\phi) := (T_a, T_b)$ folgt für die Wahrscheinlichkeitsdichte von C

$$P_C(C) = \int_{-\pi}^{\pi} P_{T_a T_b}(\sqrt{C} \cos\phi, \sqrt{C} \sin\phi) \det \left(\frac{\partial(T_a, T_b)}{\partial(C, \phi)} \right) d\phi \quad (4.56)$$

$$= \int_{-\pi}^{\pi} \frac{1}{2\pi \lambda_a \lambda_b} \exp \left[-\frac{C \cos^2\phi}{2\lambda_a^2} - \frac{C \sin^2\phi}{2\lambda_b^2} \right] \frac{1}{2} d\phi \quad (4.57)$$

$$= \frac{1}{2\lambda_a \lambda_b} \exp \left[-\frac{C}{4} \left(\frac{1}{\lambda_a^2} + \frac{1}{\lambda_b^2} \right) \right] \mathfrak{J}_0 \left[\frac{C}{4} \left(\frac{1}{\lambda_a^2} - \frac{1}{\lambda_b^2} \right) \right], \quad (4.58)$$

wobei mit $\mathfrak{J}_0[x] := \sum_{k=0}^{\infty} [k! \Gamma(k+1)]^{-1} \left(\frac{x}{2}\right)^{2k} = 1 + \frac{x^2}{4} + O(x^4)$ die „modifizierte Besselfunktion erster Gattung 0-ter Ordnung“ [11] bezeichnet wurde. Aus (4.48) folgt $\lambda_a^2 - \lambda_b^2 = \sqrt{(\sigma_a^2 - \sigma_b^2)^2 + 4(\sigma_{ab}^2)^2}$ und somit

$$P_C(C) = \frac{1}{2\lambda_a \lambda_b} \exp \left[-\frac{C(\lambda_a^2 + \lambda_b^2)}{4\lambda_a^2 \lambda_b^2} \right] \mathfrak{J}_0 \left[\frac{C(\lambda_a^2 - \lambda_b^2)}{4\lambda_a^2 \lambda_b^2} \right] \quad (4.59)$$

$$= \frac{1}{2\sqrt{\det \hat{\mathbf{\Sigma}}_{ab}}} \exp \left[-\frac{C(\sigma_a^2 + \sigma_b^2)}{4 \det \hat{\mathbf{\Sigma}}_{ab}} \right] \mathfrak{J}_0 \left[\frac{C\sqrt{(\sigma_a^2 - \sigma_b^2)^2 + 4(\sigma_{ab}^2)^2}}{4 \det \hat{\mathbf{\Sigma}}_{ab}} \right]. \quad (4.60)$$

Führt man die Größen $\varepsilon^2 := \lambda_b^2 - \lambda_a^2 = \sqrt{(\sigma_a^2 - \sigma_b^2)^2 + 4(\sigma_{ab}^2)^2}$ und $\bar{\sigma}^2 := \frac{1}{2}(\sigma_a^2 + \sigma_b^2)$ ein, so gilt $\det \hat{\Sigma}_{ab} = (\bar{\sigma}^2)^2 - \frac{1}{4}(\varepsilon^2)^2$. Für $\varepsilon^2/\bar{\sigma}^2 \ll 1$ kann $P_C(C)$ durch

$$\tilde{P}_C(C) := \frac{1}{2\bar{\sigma}^2} \exp\left[-\frac{C}{2\bar{\sigma}^2}\right] = \frac{1}{\sigma_a^2 + \sigma_b^2} \exp\left[-\frac{C}{\sigma_a^2 + \sigma_b^2}\right] \quad (4.61)$$

genähert werden. Es gilt $P_C(C) = \tilde{P}_C(C)(1 + O[(\varepsilon^2/\bar{\sigma}^2)^2])$. Die genäherte Verteilungsdichte ist bereits normiert, d.h. es gilt $\int_0^\infty \tilde{P}_C(C) dC = 1$, und besitzt den gleichen Erwartungswert wie die (exakte) Verteilungsdichte $P_C(C)$:

$$\int_0^\infty C \tilde{P}_C(C) dC = \sigma_a^2 + \sigma_b^2 . \quad (4.62)$$

Mit

$$\int_0^\infty C^2 \tilde{P}_C(C) dC = 2(\sigma_a^2 + \sigma_b^2)^2 \quad (4.63)$$

erhält man für die Varianz $(\sigma_a^2 + \sigma_b^2)^2$. Dies entspricht (4.55) für $\epsilon \rightarrow 0$.

Modellierung eindimensionaler Stereobilder

Zur Vereinfachung werden im Folgenden eindimensionale Bilder und rezeptive Felder betrachtet. Dies stellt jedoch für rezeptive Felder, die durch eine separierbare Funktion $f(x, y) = f_x(x)f_y(y)$ beschrieben werden, keine entscheidende Einschränkung dar. In diesem Fall kann ein eindimensionales Bild aus dem zweidimensionalen durch Faltung in vertikaler Richtung erzeugt werden:

$$I_{l/r}(x) := \int f_y(y) I_{l/r}(x, y) dy . \quad (4.64)$$

Für $f_y(y) = \exp(-\frac{1}{2}y^2/\sigma^2)$ entspricht dies Gleichung (4.22).

Korrespondierende Ausschnitte im rechten und linken Bild mit horizontaler Disparität d werden durch

$$I_l(x) = \bar{I}(x) + \epsilon_l(x) , \quad (4.65)$$

$$I_r(x) = B[\bar{I}(x - d)] + \epsilon_r(x) = \alpha \bar{I}(x - d) + \beta + \epsilon_r(x) , \quad (4.66)$$

modelliert.⁶ Dabei bezeichnet $\bar{I}(x)$ das „wahre“ (ungestörte) Bild, $\epsilon_{l/r}$ beschreibt den Einfluss von Kamerarauschen und $B[x] := \alpha x + \beta$ ist ein einfaches Modell [54] für Kontrast- und Helligkeitsunterschiede zwischen den

⁶Dies bedeutet, dass zumindest innerhalb der rezeptiven Felder eine konstante Disparität vorausgesetzt wird. Tiefenvariationen und Verdeckungen werden im Folgenden nicht berücksichtigt.

Bildern (z.B. infolge unterschiedlicher Einstellung der Blendenöffnungen). $\epsilon_{l/r}(x)$ kann analog zu Gleichung (4.64) aus $\epsilon_{l/r}(x, y)$ bestimmt werden. Fehlstellungen der Kameraachsen werden nicht berücksichtigt.

Die rezeptiven Felder für binokulare Einfachzellen vom Positions-Shift-Typ mit Zentren $(0, 0)$ im linken und $(D, 0)$ im rechten Bild lauten

$$f_{a/bl}(\xi) = f_{a/b}(\xi) , \quad (4.67)$$

$$f_{a/br}(\xi) = f_{a/b}(\xi - D) . \quad (4.68)$$

In der symmetrischen Schreibweise von Gleichung (4.19) entspricht dies $f_{a/bl}(\xi) = f_{a/b}(x + \frac{1}{2}D + \xi)$ und $f_{a/br}(\xi) = f_{a/b}(x - \frac{1}{2}D + \xi)$, betrachtet am Ort $x = -\frac{1}{2}D$.

Es wird angenommen, dass $\int f_{a/b}(x)dx \approx 0$ gilt, so dass eine konstante Helligkeitsverschiebung – modelliert durch Parameter β in (4.66) – keine Rolle spielt. Für biologische Zellen können zusätzlich Abweichungen zwischen linkem und rechtem rezeptiven Feld auftreten. Auch neuronales Rauschen wird im Folgenden nicht berücksichtigt.

Unter diesen Voraussetzungen folgt aus (4.65) und (4.67) bzw. (4.66) und (4.68) für die Aktivität der Einfachzellen

$$\begin{aligned} S_{a/bl} &:= \int I_l(x) f_{a/bl}(x) dx \\ &= \int \bar{I}(x) f_{a/b}(x) dx + \int \epsilon_l(x) f_{a/b}(x) dx , \end{aligned} \quad (4.69)$$

$$\begin{aligned} S_{a/br} &:= \int I_r(x) f_{a/br}(x) dx \\ &= \alpha \int \bar{I}(x - d) f_{a/b}(x - D) dx + \int \epsilon_r(x) f_{a/b}(x - D) dx \\ &= \alpha \int \bar{I}(x) f_{a/b}(x + \tilde{d}) dx + \int \epsilon_r(x) f_{a/b}(x - D) dx , \end{aligned} \quad (4.70)$$

$$\begin{aligned} S_{a/b} &= S_{a/bl} + S_{a/br} \\ &= \int \bar{I}(x) [f_{a/b}(x) + \alpha f_{a/b}(x + \tilde{d})] dx \\ &\quad + \int \epsilon_l(x) f_{a/b}(x) dx + \int \epsilon_r(x) f_{a/b}(x - D) dx , \end{aligned} \quad (4.71)$$

wobei $\tilde{d} := d - D$ definiert wurde. Für Komplexzellen ist \tilde{d} die Abweichung der Bilddisparität von der Vorzugsdisparität des Neurons.

Wird desweiteren angenommen, dass (über viele Bilder gemittelt) $\langle \bar{I}(x) \rangle = \text{const}$ gilt und $\langle \bar{I}(x) \bar{I}(x') \rangle$ nur eine Funktion des Abstands $|x - x'|$ ist, so folgt

für weißes mittelwertfreies Rauschen, d.h. $\langle \epsilon_{l/r}(x) \rangle = 0$ und $\langle \epsilon_u(x)\epsilon_v(x') \rangle = \sigma_u^2 \delta(x-x')\delta_{u,v}$, $u, v \in \{l, r\}$:

$$\langle S_{a/b} \rangle \propto \int f_{a/b}(x) dx \approx 0, \quad (4.72)$$

$$\begin{aligned} \langle S_{a/b}^2 \rangle &\approx (1 + \alpha^2) \iint \langle \bar{I}(x)\bar{I}(x') \rangle f_{a/b}(x)f_{a/b}(x') dx dx' \\ &\quad + (\sigma_l^2 + \sigma_r^2) \int f_{a/b}(x)^2 dx \\ &\quad + 2\alpha \iint \langle \bar{I}(x)\bar{I}(x') \rangle f_{a/b}(x + \tilde{d})f_{a/b}(x') dx dx'. \end{aligned} \quad (4.73)$$

Wegen (4.72) gilt $\sigma_a^2 \approx \langle S_a^2 \rangle$, $\sigma_b^2 \approx \langle S_b^2 \rangle$ und $\sigma_{ab}^2 \approx \langle S_a S_b \rangle$. Für den Mittelwert des Produkts $S_a S_b$ erhält man

$$\begin{aligned} \langle S_a S_b \rangle &= (1 + \alpha^2) \iint \langle \bar{I}(x)\bar{I}(x') \rangle f_a(x)f_b(x') dx dx' + (\sigma_l^2 + \sigma_r^2) \int f_a(x)f_b(x) dx \\ &\quad + \alpha \iint \langle \bar{I}(x)\bar{I}(x') \rangle [f_a(x + \tilde{d})f_b(x') + f_a(x)f_b(x' + \tilde{d})] dx dx'. \end{aligned} \quad (4.74)$$

Falls f_a eine gerade und f_b eine ungerade Funktion (z.B. Cosinus- und Sinus-Gaborfunktion) ist, so kann $\langle S_a S_b \rangle = \sigma_{ab}^2 = 0$ gezeigt werden.

Gaborfilter und Zufallsstereogramme

Im Folgenden werden die rezeptiven Felder mit Gaborfunktionen modelliert, d.h. $f_a(x) = \exp(-\frac{x^2}{2\sigma^2}) \cos(\omega x)$, $f_b(x) = \exp(-\frac{x^2}{2\sigma^2}) \sin(\omega x)$. Als Stimuli werden Zufalls- („Random-Dot“-)Stereogramme betrachtet. Für diese gilt im kontinuierlichen Grenzfall

$$\langle \bar{I}(x)\bar{I}(x') \rangle = I_0^2 + \sigma_I^2 \delta(x-x'). \quad (4.75)$$

Der konstante Offset I_0 der Bilder kann wegen $\int f_{a/b}(x) dx \approx 0$ unberücksichtigt bleiben.⁷

$$\begin{aligned} \text{Mit } \int f_{a/b}(x)^2 dx &= \frac{\sqrt{\pi}}{2} \sigma (1 \pm \exp[-(\omega\sigma)^2]) \quad \text{und} \\ \int f_{a/b}(x + \tilde{d})f_{a/b}(x) dx &= \frac{\sqrt{\pi}}{2} \sigma \exp[-\frac{\tilde{d}^2}{4\sigma^2}] (\cos(\omega\tilde{d}) \pm \exp[-(\omega\sigma)^2]) \end{aligned}$$

⁷Exakt gilt $\int \exp(-\frac{x^2}{2\sigma^2}) \sin(\omega x) dx = 0$, $\int \exp(-\frac{x^2}{2\sigma^2}) \cos(\omega x) dx = \sqrt{2\pi}\sigma \exp[-\frac{1}{2}(\omega\sigma)^2]$.

folgt aus (4.73):

$$\begin{aligned}
\langle S_{a/b}^2(\tilde{d}) \rangle & \approx [(1 + \alpha^2)\sigma_I^2 + \sigma_l^2 + \sigma_r^2] \int f_{a/b}(x)^2 dx + 2\alpha\sigma_I^2 \int f_{a/b}(x + \tilde{d})f_{a/b}(x) dx \\
& = \frac{\sqrt{\pi}}{2}\sigma \left((1 + \alpha^2)\sigma_I^2 + 2\alpha\sigma_I^2 \exp\left[-\frac{\tilde{d}^2}{4\sigma^2}\right] \cos(\omega\tilde{d}) + \sigma_l^2 + \sigma_r^2 \right) \\
& \quad \pm \exp[-(\omega\sigma)^2] \left[(1 + \alpha^2)\sigma_I^2 + 2\alpha\sigma_I^2 \exp\left[-\frac{\tilde{d}^2}{4\sigma^2}\right] + \sigma_l^2 + \sigma_r^2 \right] . \quad (4.76)
\end{aligned}$$

Für die im Kortex experimentell gefundenen Bandbreiten von ein bis zwei Oktaven [27] gilt⁸ $\frac{5}{3} < \omega\sigma < 3$, so dass näherungsweise $\exp[-(\omega\sigma)^2] \approx 0$ gesetzt werden kann.

Tuningkurve: Für den Erwartungswert der Komplexzellantwort erhält man

$$\begin{aligned}
\langle C(\tilde{d}) \rangle & = \langle S_a^2(\tilde{d}) \rangle + \langle S_b^2(\tilde{d}) \rangle \\
& \approx \sqrt{\pi}\sigma \left((1 + \alpha^2)\sigma_I^2 + 2\alpha\sigma_I^2 \exp\left[-\frac{\tilde{d}^2}{4\sigma^2}\right] \cos(\omega\tilde{d}) + \sigma_l^2 + \sigma_r^2 \right) \\
& = \sqrt{\pi}\sigma\sigma_I^2 \left((1 + \alpha^2) + 2\alpha \exp\left[-\frac{\tilde{d}^2}{4\sigma^2}\right] \cos(\omega\tilde{d}) + \frac{\sigma_l^2 + \sigma_r^2}{\sigma_I^2} \right) .
\end{aligned}$$

Mit $r := \frac{\sigma_l^2 + \sigma_r^2}{2\sigma_I^2}$ folgt

$$\begin{aligned}
\langle C(\tilde{d}) \rangle & = \sqrt{\pi}\sigma\sigma_I^2(1 + \alpha^2 + 2r) \left(1 + \frac{2\alpha}{(1 + \alpha^2 + 2r)} \exp\left[-\frac{\tilde{d}^2}{4\sigma^2}\right] \cos(\omega\tilde{d}) \right) \\
& = \langle C(\infty) \rangle \left(1 + \kappa \exp\left[-\frac{\tilde{d}^2}{4\sigma^2}\right] \cos(\omega\tilde{d}) \right) , \quad (4.77)
\end{aligned}$$

wobei $\langle C(\infty) \rangle := \sqrt{\pi}\sigma\sigma_I^2(1 + \alpha^2 + 2r)$ und die Modulationstiefe $\kappa := \frac{2\alpha}{(1 + \alpha^2 + 2r)}$ definiert wurden. Wegen $\alpha \geq 0$ und $r \geq 0$ gilt $0 \leq \kappa \leq 1$. Die maximale Modulationstiefe $\kappa = 1$ wird für $\alpha = 1$ (keine Kontrastveränderung zwischen den Kameras) und $r = 0$ (kein Rauschen) erzielt. Für $\kappa = 1$ entspricht (4.77) den Resultaten in [101] und [110].

Varianz der Zellaktivität: Mit $\langle S_a S_b \rangle = 0$,

$$\begin{aligned}
\langle S_a \rangle^2 - \langle S_b \rangle^2 & = \sqrt{\pi}\sigma \exp[-(\omega\sigma)^2] \left((1 + \alpha^2)\sigma_I^2 + 2\alpha\sigma_I^2 \exp\left[-\frac{\tilde{d}^2}{4\sigma^2}\right] + \sigma_l^2 + \sigma_r^2 \right)
\end{aligned}$$

⁸Nach Einsetzen von $\omega = 2\pi\nu$ und $\sigma = (2\pi\zeta)^{-1}$ folgt aus Gleichung (4.20): $\omega\sigma = \frac{2^{\text{bw}} + 1}{2^{\text{bw}} - 1}$.

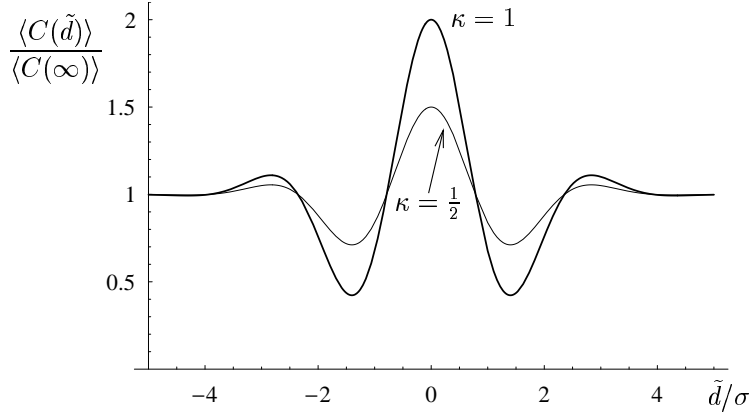


Abbildung 4.8: Tuningkurven $\langle C(\tilde{d}) \rangle$ der Komplexzellen mit $\omega\sigma = 2$ für Random-Dot-Stimulus bei unterschiedlicher Modulationstiefe κ , siehe Gleichung (4.77). Wegen (4.78) besitzen die zugehörigen Standardabweichungen $(\text{Var}[C(\tilde{d})])^{\frac{1}{2}} \approx \langle C(\tilde{d}) \rangle$ näherungsweise den gleichen Verlauf.

und $\sigma_a^2 \approx \langle S_a^2 \rangle$, $\sigma_a^2 \approx \langle S_a^2 \rangle$, $\sigma_{ab}^2 \approx \langle S_a S_b \rangle$ folgt aus (4.54) für die Varianz der Komplexzellantwort:

$$\begin{aligned}
 \text{Var}[C(\tilde{d})] &\approx (\langle S_a \rangle^2 + \langle S_b \rangle^2)^2 + (\langle S_a \rangle^2 - \langle S_b \rangle^2)^2 + 4\langle S_a S_b \rangle^2 \\
 &= \langle C(\tilde{d}) \rangle^2 + (\sqrt{\pi}\sigma)^2 \exp[-2(\omega\sigma)^2] \\
 &\quad \times \left((1 + \alpha^2)\sigma_l^2 + 2\alpha\sigma_l^2 \exp\left[-\frac{\tilde{d}^2}{4\sigma^2}\right] + \sigma_l^2 + \sigma_r^2 \right)^2 + 0 \\
 &\approx \langle C(\tilde{d}) \rangle^2
 \end{aligned} \tag{4.78}$$

In Abbildung 4.8 sind Tuningkurven für $\kappa = 1$ und $\kappa = \frac{1}{2}$ dargestellt. Die zugehörigen Wahrscheinlichkeitsdichten, berechnet nach Gleichung (4.60), zeigt Abbildung 4.9. Wegen der Symmetrie zu $\tilde{d} = 0$ ist nur der Bereich $\tilde{d}/\sigma \geq 0$ dargestellt. Die Maxima bei $C = 0$, $\tilde{d}/\sigma \approx 1.5$ entsprechen den Minima der Tuningkurven in Abbildung 4.8. Die beiden Extremfälle $\tilde{d} = 0$ und $\tilde{d} \rightarrow \infty$ sind nochmals in Abbildung 4.10 dargestellt. Wie zu erkennen ist, lässt eine einzelne Antwort kaum Rückschlüsse darüber zu, ob die Stimulusdisparität nahe bei Vorzugsdisparität des Neurons liegt. Ursache hierfür ist, dass die Aktivität der Komplexzelle neben der Disparität stark vom lokalen Stimuluskontrast abhängt. Dessen Einfluss könnte durch Vergleich mit den Antworten benachbarter Zellen mit abweichender Vorzugsdisparität herausgerechnet bzw. durch die im nächsten Abschnitt beschriebene Normalisierung stark verringert werden.

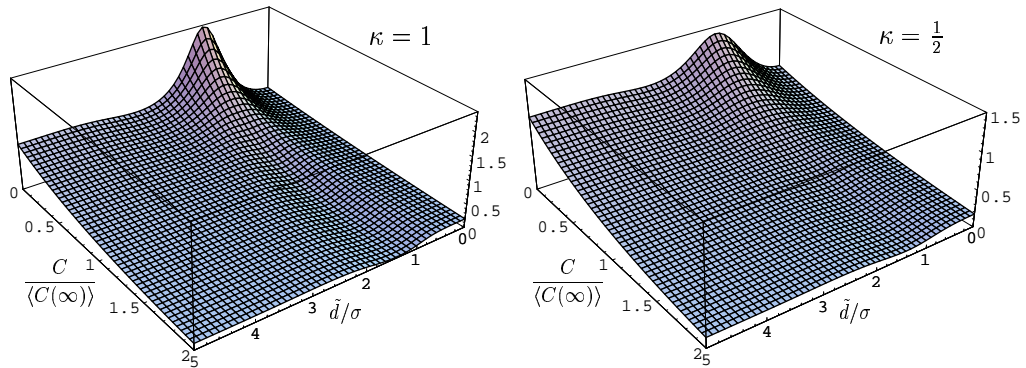


Abbildung 4.9: Wahrscheinlichkeitsdichten $P_C(C|\tilde{d})$ mit $\omega\sigma = 2$ bei unterschiedlicher Modulationstiefe κ .

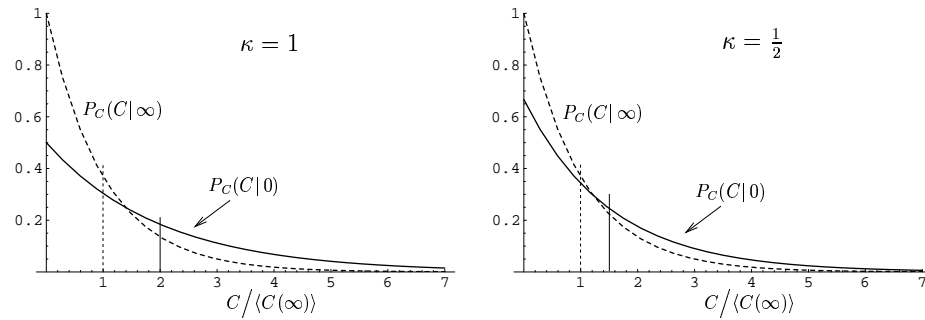


Abbildung 4.10: Vergleich der Wahrscheinlichkeitsdichten für $\tilde{d} = 0$ (kontinuierliche Kurven) und $\tilde{d} \rightarrow \infty$ (gestrichelte Kurven) bei unterschiedlicher Modulationstiefe κ . Die vertikalen Striche markieren die jeweiligen Erwartungswerte $\langle C(0) \rangle = 2\langle C(\infty) \rangle$ bzw. $\langle C(0) \rangle = 1.5\langle C(\infty) \rangle$ und $\langle C(\infty) \rangle$.

Bemerkung: Die Korrelationsfunktion für zwei Positionen innerhalb eines Bildes, $h(\mathbf{x} - \mathbf{x}') := \langle \bar{I}(\mathbf{x})\bar{I}(\mathbf{x}') \rangle$, steht über die Fouriertransformation mit dem Stimulusspektrum $\langle |\mathcal{I}(\mathbf{k})|^2 \rangle$ in Beziehung. Es gilt, siehe z.B. [56]:

$$h(\mathbf{z}) = \frac{1}{(2\pi)^2} \iint \langle |\mathcal{I}(\mathbf{k})|^2 \rangle e^{i\mathbf{kz}} dk_x dk_y . \quad (4.79)$$

Für weißes Rauschen, $\langle |\mathcal{I}(\mathbf{k})|^2 \rangle = \text{const}$, folgt dann Gleichung (4.75) nach Reduktion auf eine Dimension. Für natürliche Stimuli gilt näherungsweise $\langle |\mathcal{I}(\mathbf{k})|^2 \rangle \propto \frac{1}{\mathbf{k}^2}$, wobei die Singularität bei $|\mathbf{k}| = 0$ durch $\langle |\mathcal{I}(\mathbf{k})|^2 \rangle \propto \frac{1}{\gamma^2 + \mathbf{k}^2}$ behoben werden kann [26]. Die daraus mit (4.79) zu berechnende Korrelationsfunktion (vgl. [82]) erschwert analytische Berechnungen. Für sehr breite rezeptive Felder gilt jedoch Gleichung (4.75) näherungsweise.

4.3.2 Normalisierte Zellen

In diesem Abschnitt wird die Wahrscheinlichkeitsdichte für die Aktivität normalisierter Komplexzellen berechnet, wie sie in dem in Abschnitt 4.2 beschriebenen Stereokamerasystem verwendet werden. Allerdings wird im Vergleich zu Gleichung (4.25) nur ein einzelner Frequenzkanal berücksichtigt. Auch wird keine explizite monokulare Normalisierung der Einfachzellen nach Gleichung (4.23) durchgeführt, da diese im Wesentlichen durch $\alpha = 1$ modelliert werden kann, siehe (4.66). Da die verwendete binokulare Normalisierung lediglich eine vereinfachte Version des in [34] und [45] vorgeschlagenen Normalisierungsmodells für Disparitäts-selektiven Neuronen ist, sind die Resultate auch in Hinblick auf „physiologische Neurone“ von Interesse.

Einfachzellen

Analog zur Argumentation in 4.3.1 kann die gemeinsame Wahrscheinlichkeitsdichte von $S_{al}, S_{ar}, S_{bl}, S_{br}$, siehe Gleichungen (4.69) und (4.70), näherungsweise durch eine vierdimensionale Normalverteilung beschrieben werden. Mit der Definition $\mathbf{S} := (S_{al}, S_{ar}, S_{bl}, S_{br})^\top$ gilt

$$P_{\mathbf{S}}(\mathbf{S} | \delta) := \frac{1}{\sqrt{(2\pi)^4 \det \hat{\Sigma}(\tilde{d})}} \exp\left(-\frac{1}{2} \mathbf{S}^\top \hat{\Sigma}(\tilde{d})^{-1} \mathbf{S}\right), \quad (4.80)$$

mit Kovarianzmatrix

$$\hat{\Sigma}(\tilde{d}) := \begin{pmatrix} \sigma_{al}^2 & \sigma_{alar}^2 & \sigma_{albl}^2 & \sigma_{albr}^2 \\ \sigma_{alar}^2 & \sigma_{ar}^2 & \sigma_{arbl}^2 & \sigma_{arbr}^2 \\ \sigma_{albl}^2 & \sigma_{arbl}^2 & \sigma_{bl}^2 & \sigma_{blbr}^2 \\ \sigma_{albr}^2 & \sigma_{arbr}^2 & \sigma_{blbr}^2 & \sigma_{br}^2 \end{pmatrix}. \quad (4.81)$$

Hierbei wurde (wiederum) angenommen, dass der Erwartungswert $\langle \mathbf{S} \rangle$ verschwindet (kein DC-Offset). Desweiteren sei f_a eine gerade bzw. f_b eine ungerade Funktion und $\langle \bar{I}(x) \bar{I}(x') \rangle$ nur vom Abstand $|x - x'|$ abhängig. Um die Berechnungen zu vereinfachen, wird darüberhinaus $\sigma_l^2 = \sigma_r^2 =: \sigma_n^2$ gesetzt und die monokulare Normalisierung – wie bereits erwähnt – durch $\alpha = 1$ modelliert. Für die Modulationstiefe gilt somit $\kappa = \frac{1}{1+r} = \frac{\sigma_I^2}{\sigma_I^2 + \sigma_n^2}$. Unter diesen Annahmen folgt für weißes Rauschen, $\langle \epsilon_u(x) \epsilon_v(x') \rangle_n = \sigma_n^2 \delta(x - x') \delta_{u,v}$, $u, v \in \{l, r\}$:

$$\begin{aligned} \sigma_{al}^2 = \sigma_{ar}^2 = a & := \iint \langle \bar{I}(x) \bar{I}(x') \rangle f_a(x) f_a(x') dx dx' \\ & + \sigma_n^2 \int f_a(x)^2 dx, \end{aligned} \quad (4.82)$$

$$\sigma_{bl}^2 = \sigma_{br}^2 = b := \iint \langle \bar{I}(x)\bar{I}(x') \rangle f_b(x)f_b(x') dx dx' + \sigma_n^2 \int f_b(x)^2 dx , \quad (4.83)$$

$$\sigma_{alar}^2 = c := \iint \langle \bar{I}(x)\bar{I}(x') \rangle f_a(x)f_a(x' + \tilde{d}) dx dx' , \quad (4.84)$$

$$\sigma_{blbr}^2 = d := \iint \langle \bar{I}(x)\bar{I}(x') \rangle f_b(x)f_b(x' + \tilde{d}) dx dx' , \quad (4.85)$$

$$\sigma_{albr}^2 = -\sigma_{arbl}^2 = e := \iint \langle \bar{I}(x)\bar{I}(x') \rangle f_a(x)f_b(x' + \tilde{d}) dx dx' , \quad (4.86)$$

$$\sigma_{albl}^2 = \sigma_{arbr}^2 = 0 . \quad (4.87)$$

Für die Kovarianzmatrix (4.81) gilt damit

$$\hat{\Sigma}(\tilde{d}) = \begin{pmatrix} a & c & 0 & e \\ c & a & -e & 0 \\ 0 & -e & b & d \\ e & 0 & d & b \end{pmatrix} . \quad (4.88)$$

Mit Hilfe der orthonormalen Transformation

$$\hat{\mathbf{O}} := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix} = \hat{\mathbf{O}}^{-1} = \hat{\mathbf{O}}^\top \quad (4.89)$$

folgt für die Wahrscheinlichkeitsdichte von $\mathbf{u} := (u_a, v_a, u_b, v_b)^\top = \hat{\mathbf{O}}\mathbf{S} = \frac{1}{\sqrt{2}}(S_{al} + S_{ar}, S_{al} - S_{ar}, S_{bl} + S_{br}, S_{al} - S_{ar})^\top$

$$P_{\mathbf{u}}(\mathbf{u} | \tilde{d}) = \frac{1}{\sqrt{(2\pi)^4 \det \hat{\Lambda}(\tilde{d})}} \exp(-\frac{1}{2} \mathbf{u}^\top \hat{\Lambda}(\tilde{d})^{-1} \mathbf{u}) . \quad (4.90)$$

Die zugehörige Kovarianzmatrix ist gegeben durch

$$\hat{\Lambda}(\tilde{d}) = \hat{\mathbf{O}} \hat{\Sigma}(\tilde{d}) \hat{\mathbf{O}}^\top = \begin{pmatrix} a+c & 0 & 0 & -e \\ 0 & a-c & e & 0 \\ 0 & e & b+d & 0 \\ -e & 0 & 0 & b-d \end{pmatrix} , \quad (4.91)$$

ihre Inverse lautet mit $N_{1/2} := (a \pm c)(b \mp d) - e^2$

$$\hat{\Lambda}(\tilde{d})^{-1} = \begin{pmatrix} \frac{b-d}{N_1} & 0 & 0 & \frac{e}{N_1} \\ 0 & \frac{b+d}{N_2} & \frac{-e}{N_2} & 0 \\ 0 & \frac{-e}{N_2} & \frac{a-c}{N_2} & 0 \\ \frac{e}{N_1} & 0 & 0 & \frac{a+c}{N_1} \end{pmatrix} . \quad (4.92)$$

Wegen $S_{a/b} = \sqrt{2}u_{a/b}$ kann aus

$$\begin{aligned} P_{u_a u_b}(u_a, u_b | \tilde{d}) &= \int dv_a \int dv_b P_{\mathbf{u}}(\mathbf{u} | \tilde{d}) \\ &= \frac{1}{\sqrt{2\pi(a+c)}\sqrt{2\pi(b+d)}} \exp\left(-\frac{u_a^2}{2(a+c)} - \frac{u_b^2}{2(b+d)}\right) \end{aligned}$$

die gemeinsame Wahrscheinlichkeitsdichte der Einfachzellen S_a und S_b berechnet werden:

$$\begin{aligned} P_{S_a S_b}(S_a, S_b | \tilde{d}) &= P_{u_a u_b}\left(\frac{S_a}{\sqrt{2}}, \frac{S_b}{\sqrt{2}} | \tilde{d}\right) \det\left(\frac{\partial(S_a, S_b)}{\partial(u_a, u_b)}\right)^{-1} = \frac{1}{2} P_{u_a u_b}\left(\frac{S_a}{\sqrt{2}}, \frac{S_b}{\sqrt{2}} | \tilde{d}\right) \\ &= \frac{1}{\sqrt{2\pi}2(a+c)\sqrt{2\pi}2(b+d)} \exp\left(-\frac{S_a^2}{2 \cdot 2(a+c)} - \frac{S_b^2}{2 \cdot 2(b+d)}\right) \\ &= \frac{1}{\sqrt{2\pi}\sigma_a^2\sqrt{2\pi}\sigma_b^2} \exp\left(-\frac{S_a^2}{2\sigma_a^2} - \frac{S_b^2}{2\sigma_b^2}\right), \end{aligned}$$

wobei $\sigma_a^2 := 2(a+c)$ und $\sigma_b^2 := 2(b+d)$ substituiert wurde. Dies entspricht Gleichung (4.40) mit $\sigma_{ab}^2 = \langle S_a S_b \rangle = 0$.

Komplexzellen

Aus Gleichungen (4.10), (4.12), (4.16) und (4.17) folgt $|Q_l|^2 + |Q_r|^2 = \text{Re}[Q_l]^2 + \text{Im}[Q_r]^2 + \text{Re}[Q_r]^2 + \text{Im}[Q_l]^2 = S_{al}^2 + S_{bl}^2 + S_{ar}^2 + S_{br}^2$. Da die monokulare Normalisierung (4.23) durch $\alpha = 1$ modelliert wird, gilt $NQ_{l/r} \approx Q_{l/r}$. Wird nur ein einzelner Frequenzkanal betrachtet, vereinfacht sich Gleichung (4.25) weiter und für die Aktivität einer normalisierten Komplexzelle folgt:

$$\begin{aligned} NC &:= \frac{C}{S_{al}^2 + S_{ar}^2 + S_{bl}^2 + S_{br}^2 + \epsilon} = \frac{S_a^2 + S_b^2}{S_{al}^2 + S_{ar}^2 + S_{bl}^2 + S_{br}^2 + \epsilon} \\ &= \frac{(S_{al} + S_{ar})^2 + (S_{bl} + S_{br})^2}{\mathbf{S}^2 + \epsilon} = \frac{2u_a^2 + 2u_b^2}{\mathbf{u}^2 + \epsilon} \\ &= \frac{2(u_a^2 + u_b^2)}{u_a^2 + v_a^2 + u_b^2 + v_b^2 + \epsilon}. \end{aligned} \quad (4.93)$$

Mit den Substitutionen $\sqrt{w}(\cos \phi, \sin \phi) := (u_a, u_b)$ und $\sqrt{z}(\cos \psi, \sin \psi) := (v_a, v_b)$ erhält man

$$NC = \frac{2x}{w + z + \epsilon}. \quad (4.94)$$

Mit $\det\left(\frac{\partial(u_a v_a u_b v_b)}{\partial(w \phi z \psi)}\right) = \frac{1}{4}$ folgt für die Wahrscheinlichkeitsdichte in (w, ϕ, z, ψ) :

$$P_{w\phi z\psi}(w, \phi, z, \psi | \tilde{d}) = \frac{1}{4} P_{\mathbf{u}}(\sqrt{w} \cos \phi, \sqrt{z} \cos \psi, \sqrt{w} \sin \phi, \sqrt{z} \sin \psi | \tilde{d}) . \quad (4.95)$$

Für Gaborfilter, $f_a(x) = \exp(-\frac{x^2}{2\sigma^2}) \cos(\omega x)$, $f_b(x) = \exp(-\frac{x^2}{2\sigma^2}) \sin(\omega x)$, mit Bandbreiten zwischen 1 und 2, d.h. $\frac{5}{3} < \omega\sigma < 3$ kann $\exp[-(\omega\sigma)^2] \approx 0$ genähert werden. Für Zufallsstereogramme mit der Eigenschaft $\langle \bar{I}(x) \bar{I}(x') \rangle = I_0^2 + \sigma_I^2 \delta(x - x')$, Glg. (4.75), als Stimuli gilt dann

$$b \approx a \approx (\sigma_I^2 + \sigma_n^2) \frac{\sqrt{\pi}}{2} \sigma , \quad (4.96)$$

$$d \approx c \approx \sigma_I^2 \frac{\sqrt{\pi}}{2} \sigma \exp\left[-\frac{\tilde{d}^2}{4\sigma^2}\right] \cos(\omega \tilde{d}) , \quad (4.97)$$

$$e = \sigma_I^2 \frac{\sqrt{\pi}}{2} \sigma \exp\left[-\frac{\tilde{d}^2}{4\sigma^2}\right] \sin(\omega \tilde{d}) . \quad (4.98)$$

Somit folgt aus (4.95) mit (4.90) für die Wahrscheinlichkeitsdichte in w, z

$$\begin{aligned} P_{wz}(w, z | \tilde{d}) &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} P_{w\phi z\psi}(w, \phi, z, \psi) d\psi d\phi \\ &\approx \frac{1}{4(2\pi)^2(a^2 - c^2 - e^2)} \\ &\quad \times \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \exp\left[-\frac{(a-c)w + (a+c)z - 2e\sqrt{wz} \sin(\phi - \psi)}{2(a^2 - c^2 - e^2)}\right] d\phi d\psi \\ &= \frac{1}{4(2\pi)^2(a^2 - c^2 - e^2)} \\ &\quad \times \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \exp\left[-\frac{(a-c)w + (a+c)z - 2e\sqrt{wz} \sin \phi'}{2(a^2 - c^2 - e^2)}\right] d\phi' d\psi' \\ &= \frac{1}{8\pi(a^2 - c^2 - e^2)} \\ &\quad \times \int_{-\pi}^{\pi} \exp\left[-\frac{(a-c)w + (a+c)z - 2e\sqrt{wz} \sin \phi}{2(a^2 - c^2 - e^2)}\right] d\phi . \end{aligned} \quad (4.99)$$

Wegen (4.94) erhält man die gesuchte Wahrscheinlichkeitsdichte für NC durch Berechnung von

$$P_{NC}(NC | \tilde{d}) = \int_0^{\infty} \int_0^{\infty} P_{wz}(w, z) \delta\left(NC - \frac{2w}{w+z+\epsilon}\right) dw dz . \quad (4.100)$$

Die Delta-Funktion stellt sicher, dass nur über den Bereich integriert wird, für den (4.94) gilt. Mit $\delta(f(x)) = \sum_i |f'(x_i)|^{-1} \delta(x - x_i)$, wobei x_i Nullstelle

von f ist [42], folgt

$$\delta\left(NC - \frac{2w}{w+z+\epsilon}\right) = \frac{2(z+\epsilon)}{(2-NC)^2} \delta\left(w - \frac{NC}{2-NC}(z+\epsilon)\right) . \quad (4.101)$$

Die Integration nach w in (4.100) kann somit ausgeführt werden:

$$P_{NC}(NC|\tilde{d}) = \frac{2}{(2-NC)^2} \int_0^\infty (z+\epsilon) P_{wz}\left(\frac{NC}{2-NC}(z+\epsilon), z\right) dz . \quad (4.102)$$

Mit $k := \frac{NC}{2-NC}$ und (4.99) folgt

$$\begin{aligned} & P_{NC}(NC|\tilde{d}) \\ &= \frac{1}{4\pi(a^2 - c^2 - e^2)(2-NC)^2} \exp\left[-\frac{(a-c)k\epsilon}{2(a^2 - c^2 - e^2)}\right] \\ & \quad \times \int_{-\pi}^{\pi} \int_0^\infty (z+\epsilon) \exp\left[\frac{2e\sqrt{k(z+\epsilon)z} \sin\phi - [(a-c)k + a + c]z}{2(a^2 - c^2 - e^2)}\right] dz d\phi \\ &\approx \frac{1}{4\pi(a^2 - c^2 - e^2)(2-NC)^2} \exp\left[-\frac{(a-c)k\epsilon}{2(a^2 - c^2 - e^2)}\right] \\ & \quad \times \int_{-\pi}^{\pi} \int_0^\infty (z+\epsilon) \exp\left[-\frac{[(a-c)k + a + c - 2e\sqrt{k} \sin\phi]z}{2(a^2 - c^2 - e^2)}\right] dz d\phi , \end{aligned}$$

wobei $\sqrt{k(z+\epsilon)z} \approx \sqrt{k}z$ genähert wurde. Die Integrationen können nun unter Verwendung der in [11] angegebenen Formeln,

$$\begin{aligned} \int_0^\infty z \exp(-Az) dz &= A^{-2} \quad \text{für } A > 0 , \\ \int_{-\pi}^{\pi} (A + B \sin\phi)^{-2} d\phi &= \frac{2\pi|A|}{(A^2 - B^2)^{\frac{3}{2}}} \quad \text{für } |A| > |B| , \end{aligned}$$

ausgeführt werden. Mit $g[NC] := (a + (1 - NC)c)^2 - e^2 NC(2 - NC)$ und $K := a^2 - c^2 - e^2$ folgt

$$\begin{aligned} & P_{NC}(NC|\tilde{d}) \\ &\approx e^{-\frac{(a-c)NC\epsilon}{2K(2-NC)}} \left[\frac{K(a + (1 - NC)c)}{2g[NC]^{\frac{3}{2}}} + \frac{\epsilon}{2(2-NC)g[NC]^{\frac{1}{2}}} \right] . \quad (4.103) \end{aligned}$$

Für $\epsilon = 0$ ist die Näherung exakt und mit $\kappa = \frac{1}{1+r}$ und $r = \frac{\sigma_n^2}{\sigma_I^2}$ gilt

$$P_{NC}(NC|\tilde{d}; \epsilon = 0) = \frac{(a^2 - c^2 - e^2)(a + (1 - NC)c)}{2[(a + (1 - NC)c)^2 - e^2 NC(2 - NC)]^{\frac{3}{2}}} \quad (4.104)$$

$$= \frac{(1 - \kappa^2 e^{-\frac{\tilde{d}^2}{2\sigma^2}})(1 + \kappa(1 - NC)e^{-\frac{\tilde{d}^2}{4\sigma^2}} \cos(\omega\tilde{d}))}{2[(1 + \kappa(1 - NC)e^{-\frac{\tilde{d}^2}{4\sigma^2}} \cos(\omega\tilde{d}))^2 - \kappa^2 NC(2 - NC)e^{-\frac{\tilde{d}^2}{2\sigma^2}} \sin^2(\omega\tilde{d})]^{\frac{3}{2}}} , \quad (4.105)$$

wobei (4.96)–(4.98) eingesetzt wurden. Gleichung (4.105) besitzt die Grenzfälle

$$\tilde{d} \gg \sigma : \quad P_{NC}(NC|\tilde{d} \gg \sigma; \epsilon = 0) \approx \frac{1}{2} , \quad (4.106)$$

$$\tilde{d} = 0 : \quad P_{NC}(NC|\tilde{d} = 0; \epsilon = 0) = \frac{1 - \kappa^2}{2[1 + \kappa^2(1 - NC)]^2} . \quad (4.107)$$

Ohne Rauschen, d.h. $\sigma_n^2 \rightarrow 0$ folgt

$$\lim_{\kappa \rightarrow 1} P_{NC}(NC|\tilde{d} = 0; \epsilon = 0) = \delta(2 - NC) , \quad (4.108)$$

da $\int_0^2 P_{NC}(NC|\tilde{d} = 0; \epsilon = 0) dNC = 1$ und $\lim_{\kappa \rightarrow 1} P_{NC}(NC|\tilde{d} = 0; \epsilon = 0) = 0$ für $0 \leq NC < 2$ gilt.

Abbildung 4.11 zeigt Wahrscheinlichkeitsdichten für unterschiedliche Modulationstiefe κ . Ohne Rauschen ($\sigma_n^2 = 0 \iff \kappa = 1$) treten für kleine \tilde{d} -Werte ($\tilde{d} \ll \sigma$) nur hohe Aktivitäten $NC \approx 2$ auf, so dass umgekehrt mit großer Sicherheit von einer hohen Aktivität auf eine kleine Abweichung der Bilddisparität von der Vorzugsdisparität geschlossen werden kann. Auch für hohe Störpegel ($\kappa = \frac{1}{2} \iff \sigma_n^2 = \sigma_I^2$) bewirken kleine \tilde{d} -Werte im Normalfall eine hohe Komplexzellaktivität und erlauben eine deutlich bessere Disparitätsdetektion als ohne Normalisierung (vgl. mit Abb. 4.9).

Tuningkurve: Für $\epsilon = 0$ gilt für den Erwartungswert der normalisierten Komplexzelle:

$$\begin{aligned} \langle NC(\tilde{d}) \rangle &:= \int_0^2 NC P_{NC}(NC|\tilde{d}; \epsilon = 0) dNC \\ &= 1 + \frac{ac}{c^2 + e^2} - \frac{c(a^2 - c^2 - e^2)}{2(c^2 + e^2)^{\frac{3}{2}}} \ln \left(\frac{-ac + c^2 + e^2 + (a - c)\sqrt{c^2 + e^2}}{-ac - c^2 - e^2 + (a + c)\sqrt{c^2 + e^2}} \right) \end{aligned}$$

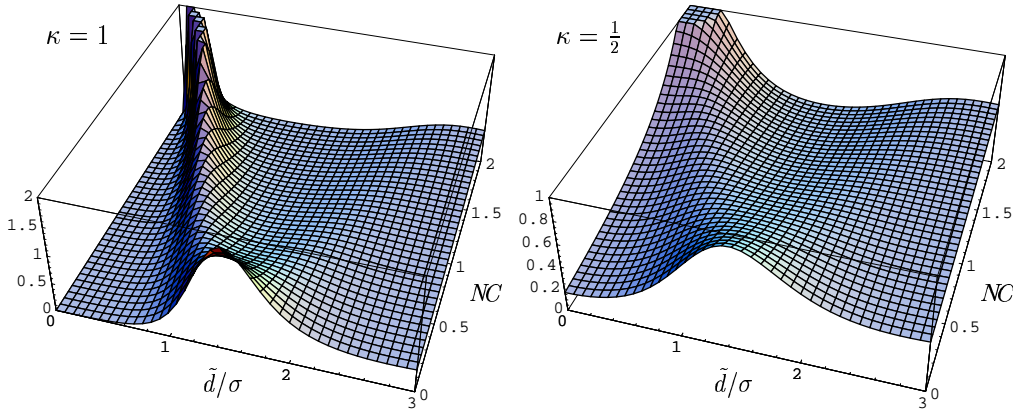


Abbildung 4.11: Wahrscheinlichkeitsdichten für die Aktivität normalisierter Komplexzellen berechnet nach Gleichung (4.105) mit $\omega\sigma = 2$ für unterschiedliche Störpegel: $\kappa = 1 \iff \sigma_n^2 = 0$ und $\kappa = \frac{1}{2} \iff \sigma_n^2 = \sigma_I^2$.

$$\stackrel{(*)}{=} 1 + \cos(\omega\tilde{d}) \left(\frac{e^{\frac{\tilde{d}^2}{4\sigma^2}}}{\kappa} - \left(\frac{e^{\frac{\tilde{d}^2}{2\sigma^2}}}{\kappa^2} - 1 \right) \operatorname{artanh}(\kappa e^{-\frac{\tilde{d}^2}{4\sigma^2}}) \right) \quad (4.109)$$

$$= 1 + \cos(\omega\tilde{d}) \kappa e^{-\frac{\tilde{d}^2}{4\sigma^2}} h[\kappa e^{-\frac{\tilde{d}^2}{4\sigma^2}}], \quad (4.110)$$

wobei die Funktion

$$h[u] := \frac{1}{u^2} - \left(\frac{1}{u^3} - \frac{1}{u} \right) \operatorname{artanh}(u) \quad (4.111)$$

definiert wurde. Für die Umformung (*) wurden die Gleichungen (4.96)–(4.98) eingesetzt und die Identität $\ln(\sqrt{\frac{1+u}{1-u}}) = \frac{1}{2} \ln(\frac{1+u}{1-u}) = \operatorname{artanh}(u)$ für $|u| < 1$ verwendet [11]. Unter Benutzung der Reihenentwicklung $\operatorname{artanh}(u) = \sum_{k=0}^{\infty} \frac{1}{2k+1} u^{2k+1}$ kann

$$h[u] = 1 - \frac{1}{3} + \sum_{k=1}^{\infty} \left(\frac{1}{2k+1} - \frac{1}{2k+3} \right) u^{2k}, \quad |u| < 1 \quad (4.112)$$

gezeigt werden.

Da $\frac{2}{3} < h[u] < 1$ für $0 < u = \kappa e^{-\frac{\tilde{d}^2}{4\sigma^2}} < 1$ gilt, folgt aus Gleichung (4.110):

$$|\langle NC(\tilde{d}) \rangle - 1| \leq |\cos(\omega\tilde{d})| \kappa e^{-\frac{\tilde{d}^2}{4\sigma^2}} = \left| \frac{\langle C(\tilde{d}) \rangle}{\langle C(\infty) \rangle} - 1 \right|. \quad (4.113)$$

Das Gleichheitszeichen in (4.113) gilt für $\tilde{d} = 0$, $\kappa = 1$, da $\lim_{u \rightarrow 1} h[u] = 1$. Außerdem erhält man

$$\langle NC(\tilde{d}) \rangle = \langle C(\infty) \rangle^{-1} \langle C(\tilde{d}) \rangle \quad (= 1) \quad (4.114)$$

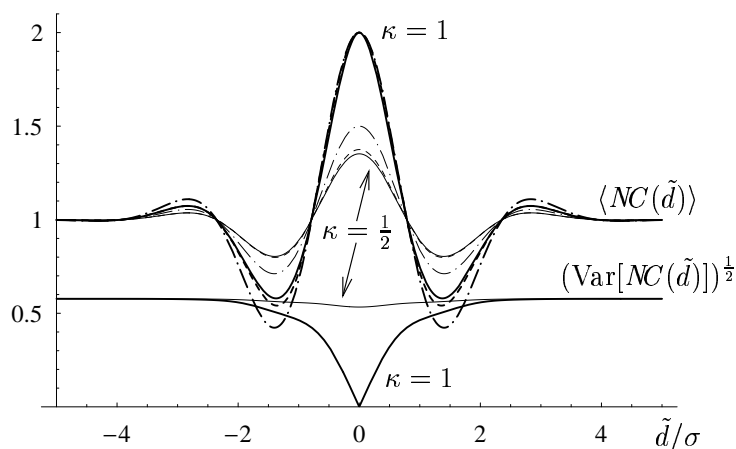


Abbildung 4.12: Tuningkurve $\langle NC(\tilde{d}) \rangle$ der normalisierten Komplexzellen (kontinuierliche Kurven), Gleichung (4.109), sowie deren Näherungen $1 + \kappa \cos(\omega\tilde{d})e^{-\frac{\tilde{d}^2}{4\sigma^2}} = \langle C(\tilde{d}) \rangle / \langle C(\infty) \rangle$ (strich-punktiert), $1 + \frac{2}{3}\kappa \cos(\omega\tilde{d})e^{-\frac{\tilde{d}^2}{4\sigma^2}} + \frac{1}{3}\kappa^3 \cos(\omega\tilde{d})e^{-\frac{3\tilde{d}^2}{4\sigma^2}}$ (gestrichelt), für $\omega\sigma = 2$ und unterschiedliche Störpegel $\sigma_n^2 = 0 \Leftrightarrow \kappa = 1$ (dicke Kurven) bzw. $\sigma_n^2 = \sigma_I^2 \Leftrightarrow \kappa = \frac{1}{2}$ (dünn). Die strich-punktierten Kurven sind mit denen in Abbildung 4.8 identisch. Zusätzlich dargestellt sind die zugehörigen Standardabweichungen $(\text{Var}[NC(\tilde{d})])^{\frac{1}{2}}$, siehe Gleichung (4.116) und Abb. 4.13.

für $\omega\tilde{d} = (\frac{1}{2} + n)\pi$, $n \in \mathbb{Z}$ und auch für $\tilde{d} \gg \sigma$. Trotz der unterschiedlichen Wahrscheinlichkeitsverteilungen, besitzen normalisierte und nicht-normalisierte Komplexzellen ähnliche Tuningkurven. Dies ist in Abbildung 4.12 dargestellt. Die Näherung $\langle NC(\tilde{d}) \rangle \approx 1 + \frac{2}{3}\kappa \cos(\omega\tilde{d})e^{-\frac{\tilde{d}^2}{4\sigma^2}} + \frac{1}{3}\kappa^3 \cos(\omega\tilde{d})e^{-\frac{3\tilde{d}^2}{4\sigma^2}}$ erhält man aus (4.110) und (4.112) durch Abbruch der Summe bei $k = 1$.

Varianz der neuronalen Aktivität: Aus

$$\begin{aligned}
 \langle NC(\tilde{d})^2 \rangle &:= \int_0^2 NC^2 P_{NC}(NC | \tilde{d}; \epsilon = 0) dNC \\
 &= \frac{a^2 (2c^2 - e^2) + 2ac(c^2 + e^2) + 2e^2(c^2 + e^2)}{(c^2 + e^2)^2} \\
 &\quad - \frac{(a^2 - c^2 - e^2)(a(2c^2 - e^2) + 2c(c^2 + e^2))}{2(c^2 + e^2)^{\frac{5}{2}}} \\
 &\quad \times \ln\left(\frac{-ac + c^2 + e^2 + (a-c)\sqrt{c^2 + e^2}}{-ac - c^2 - e^2 + (a+c)\sqrt{c^2 + e^2}}\right) \quad (4.115)
 \end{aligned}$$

folgt nach Einsetzen von (4.96)–(4.98) für die Varianz der Komplexzellantwort in Abhängigkeit von \tilde{d} :

$$\begin{aligned}\text{Var}[\mathcal{NC}(\tilde{d})] &= \langle \mathcal{NC}(\tilde{d})^2 \rangle - \langle \mathcal{NC}(\tilde{d}) \rangle^2 \\ &= (1 - \kappa^2 e^{-\frac{\tilde{d}^2}{2\sigma^2}}) H[\kappa^2 e^{-\frac{\tilde{d}^2}{4\sigma^2}}, \omega \tilde{d}] ,\end{aligned}\quad (4.116)$$

wobei die Funktion

$$\begin{aligned}H[u, v] &:= (\cos^2 v - \sin^2 v) \frac{1}{u^2} + \sin^2 v \frac{1}{u^3} \text{artanh}(u) \\ &\quad - \cos^2 v \left(\frac{1}{u^4} - \frac{1}{u^2} \right) \text{artanh}^2(u)\end{aligned}\quad (4.117)$$

definiert wurde. Aus der Reihenentwicklung von $\text{artanh}(u)$ erhält man

$$\text{artanh}^2(u) = \left(\sum_{k=0}^{\infty} \frac{1}{2k+1} u^{2k+1} \right)^2 = \sum_{n=0}^{\infty} c_n u^{2(n+1)} ,\quad (4.118)$$

$$c_n := \sum_{k=0}^n \frac{1}{2k+1} \frac{1}{2(n-k)+1} .\quad (4.119)$$

Mit (4.118) folgt aus (4.117):

$$H[u, v] = \frac{1}{3} + \sum_{n=1}^{\infty} \left(\frac{\sin^2 v}{2n+3} + \cos^2 v [c_n - c_{n+1}] \right) u^{2n} , \quad |u| < 1 .\quad (4.120)$$

In Anhang A.11 wird gezeigt, dass für $n \in \mathbb{N}_0$ gilt:

$$c_n = \frac{1}{n+1} \sum_{k=0}^n \frac{1}{2k+1} ,\quad (4.121)$$

$$c_n > c_{n+1} ,\quad (4.122)$$

$$c_n - c_{n+1} > c_{n+1} - c_{n+2} .\quad (4.123)$$

Aus (4.116), (4.120) und (4.122) folgt

$$\text{Var}[\mathcal{NC}(\tilde{d})] > \frac{1}{3} (1 - \kappa^2 e^{-\frac{\tilde{d}^2}{2\sigma^2}}) .\quad (4.124)$$

Für den Grenzfall $\tilde{d} = 0$, $\kappa = 1$ ($\implies u = 1$) gilt $\lim_{\tilde{d} \rightarrow 0} \text{Var}[\mathcal{NC}(\tilde{d})] = 0$.

Die Funktion $H[u, v]$, Gleichung (4.117), kann umgeformt werden zu

$$H[u, v] = H_1[u] + \sin^2 v H_2[u] ,\quad (4.125)$$

$$H_1[u] := \frac{1}{u^2} - \left(\frac{1}{u^4} - \frac{1}{u^2} \right) \text{artanh}^2(u) ,\quad (4.126)$$

$$H_2[u] := -2 \frac{1}{u^2} + \frac{1}{u^3} \text{artanh}(u) + \left(\frac{1}{u^4} - \frac{1}{u^2} \right) \text{artanh}^2(u) .\quad (4.127)$$

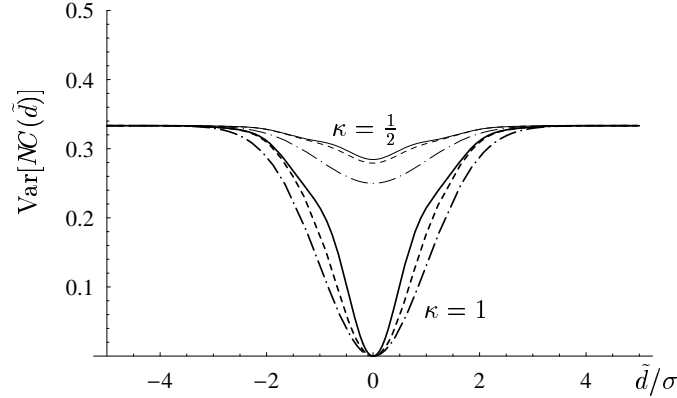


Abbildung 4.13: Varianz $\text{Var}[\mathcal{NC}(\tilde{d})]$ der normalisierten Komplexzellantwort (kontinuierliche Kurven) und Näherungen $\frac{1}{3}(1 - \kappa^2 e^{-\frac{\tilde{d}^2}{2\sigma^2}})$ (strichpunktirt), $(1 - \kappa^2 e^{-\frac{\tilde{d}^2}{2\sigma^2}}) \left(\frac{1}{3} + \kappa^2 e^{-\frac{\tilde{d}^2}{2\sigma^2}} \left[\frac{1}{5} \sin^2(\omega\tilde{d}) + \frac{7}{45} \cos^2(\omega\tilde{d}) \right] \right)$ (gestrichelt), für $\omega\sigma = 2$ und unterschiedliche Rauschpegel $\sigma_n^2 = 0 \Leftrightarrow \kappa = 1$ (dicke Kurven), $\sigma_n^2 = \sigma_I^2 \Leftrightarrow \kappa = \frac{1}{2}$ (dünn).

Da $H_2[u] > 0$ für $0 < u < 1$ gezeigt werden kann, folgt aus (4.116) und (4.125) für $u = \kappa e^{-\frac{\tilde{d}^2}{4\sigma^2}}$:

$$\begin{aligned} \text{Var}[\mathcal{NC}(\tilde{d})] &\leq (1 - u^2)(H_1[u] + H_2[u]) \\ &= (1 - u^2) \left(-\frac{1}{u^2} + \frac{1}{u^3} \text{artanh}(u) \right). \end{aligned} \quad (4.128)$$

Anhand von (4.128) kann leicht $\text{Var}[\mathcal{NC}(\tilde{d})] < \frac{1}{3}$ überprüft werden.

In Abbildung 4.13 ist $\text{Var}[\mathcal{NC}(\tilde{d})]$ für verschwindendes ($\sigma_n^2 = 0$) und starkes Rauschen ($\sigma_n^2 = \sigma_I^2$) gezeigt. Die Näherung $\text{Var}[\mathcal{NC}(\tilde{d})] \approx (1 - \kappa^2 e^{-\frac{\tilde{d}^2}{2\sigma^2}}) \left(\frac{1}{3} + \kappa^2 e^{-\frac{\tilde{d}^2}{2\sigma^2}} \left[\frac{1}{5} \sin^2(\omega\tilde{d}) + \frac{7}{45} \cos^2(\omega\tilde{d}) \right] \right)$ erhält man aus (4.116) und (4.120) durch Abbruch der Reihe bei $n = 1$. Die geringe Varianz – vor allem für kleine \tilde{d} – erleichtert die Detektion von Bilddisparitäten, die nahe bei der Vorzugsdisparität des Neurons liegen. Dies ist mit Abbildung 4.8 und $\text{Var}[C(\tilde{d})] \approx \langle C(\tilde{d}) \rangle^2$ zu vergleichen. Die Standardabweichung $(\text{Var}[\mathcal{NC}(\tilde{d})])^{\frac{1}{2}}$ wurde bereits in Abb. 4.12 dargestellt.

4.3.3 Wahrscheinlichkeit für falsche Detektion

Im Folgenden wird für normierte und nicht-normierte Komplexzellen die Wahrscheinlichkeit berechnet, dass ein Neuron, dessen Vorzugsdisparität d weit entfernt von der Stimulusdisparität D ist ($\tilde{d} := |d - D| \gg \sigma$), eine

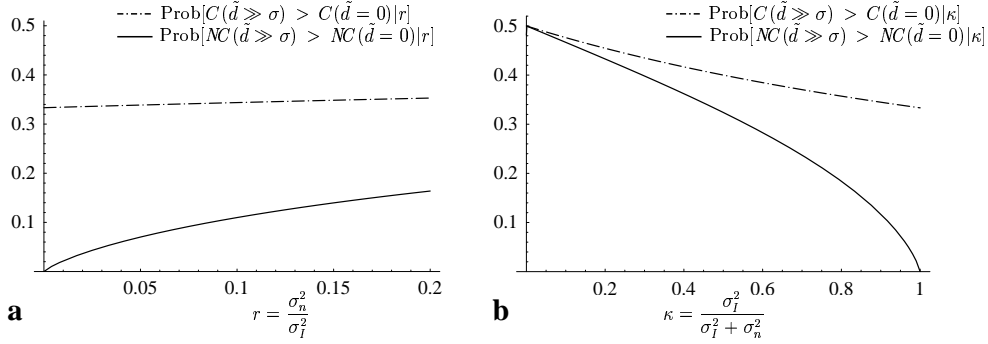


Abbildung 4.14: $\text{Prob}[\text{NC}(\tilde{d} \gg \sigma) > \text{NC}(\tilde{d} = 0)]$ und $\text{Prob}[C(\tilde{d} \gg \sigma) > C(\tilde{d} = 0)]$, Gleichungen (4.130) und (4.131), in Abhängigkeit von (a) Rauschpegel r und (b) Modulationstiefe κ .

höhere Aktivität zeigt als das Neuron mit passender Vorzugsdisparität ($\tilde{d} := |d - D| \approx 0$). Diese „Irrtums-Wahrscheinlichkeit“ kann mit $X = \{C, \text{NC}\}$, $l_{\text{NC}} = 2$, $l_C = \infty$ formuliert werden als

$$\begin{aligned} & \text{Prob}[X(\tilde{d} \gg \sigma) > X(\tilde{d} = 0)] \\ &= \int_0^{l_X} P_X(X|\tilde{d} = 0) \int_X^{l_X} P_X(X'|\tilde{d} \gg \sigma) dX' dX . \end{aligned} \quad (4.129)$$

Da die beiden Zellen (wegen $\tilde{d} \gg \sigma$) auf unterschiedliche Bildausschnitte „schauen“, sind die Antworten näherungsweise unabhängig.

Für die normierten Komplexzellen folgt mit (4.106) und (4.107)

$$\begin{aligned} & \text{Prob}[\text{NC}(\tilde{d} \gg \sigma) > \text{NC}(\tilde{d} = 0) | \kappa] \\ & \approx \frac{1 - \kappa}{4\kappa^2} \left((1 + \kappa) \ln \left[\frac{1 + \kappa}{1 - \kappa} \right] - 2\kappa \right) . \end{aligned} \quad (4.130)$$

Für die nicht-normierten Zellen gilt mit $P_C(C|\tilde{d} \gg \sigma) \approx \frac{1}{\langle C(\infty) \rangle} \exp(-\frac{C}{\langle C(\infty) \rangle})$ und $P_C(C|\tilde{d} = 0) \approx \frac{1}{(1+\kappa)\langle C(\infty) \rangle} \exp(-\frac{C}{(1+\kappa)\langle C(\infty) \rangle})$ jedoch

$$\text{Prob}[C(\tilde{d} \gg \sigma) > C(\tilde{d} = 0) | \kappa] \approx \frac{1}{2 + \kappa} . \quad (4.131)$$

Die „Irrtums-Wahrscheinlichkeit“ in Abhängigkeit von Rauschpegel $r = \frac{\sigma_n^2}{\sigma_l^2}$ und Modulationstiefe $\kappa = \frac{1}{1+r}$ ist in Abbildung 4.14 dargestellt. Für nicht zu starkes Rauschen diskriminieren die normierten Komplexzellen deutlich besser als die nicht-normierten. Für $\sigma_n^2 \rightarrow \infty \Leftrightarrow \kappa \rightarrow 0$ ist keine Unterscheidung möglich, d.h. $\text{Prob}[X(\tilde{d} \gg \sigma) > X(\tilde{d} = 0)] \rightarrow \frac{1}{2}$.

Navigation mit Landmarken: Erweiterungen und Ausblick

Im Folgenden werden zu den einzelnen Kapiteln Ideen und Vorschläge für weiterführende Arbeiten gegeben.

Zu Kapitel 1: „Visuelles Heimfinden mit panoramischen Bildern“

Effizienteres Heimfinden mit Disparitätssignaturen: Der Algorithmus zum Disparitäts-basierten Heimfinden wurde bewusst so gewählt, dass ein direkter Vergleich mit dem in [38] beschriebenen Bild-basierten Verfahren möglich wurde. Dabei konnte eine deutlich größere Invarianz von Disparitäts-Signaturen gegenüber Beleuchtungsveränderungen gezeigt werden. Der Algorithmus berechnet erwartete Disparitätswerte und vergleicht sie mit den am Zielort gespeicherten Disparitäten.

In der Robotik-Literatur existieren zahlreiche Arbeiten zur Positionsbestimmung von Robotern mit Hilfe von Entfernungsmessungen durch Laser-Range-Finder. Diese liefern im Normalfall Distanzwerte zu den umgebenden Objekten entlang des Horizonts in einem großen Winkelbereich (etwa 180° bis 360°). Durch Verwendung z.B. des IDC⁹-Algorithmus [58], der korrespondierende Punkte in zwei „Tiefenscans“ durch heuristische Annahmen festlegt, kann der Heimvektor sehr schnell bestimmt werden, da eine analytische Lösung für den Verschiebungsvektor und die Relativorientierung existiert. Die vergleichsweise aufwendige Berechnung von erwarteten Signaturen und der anschließende Vergleich kann dann wegfallen. Notwendig ist hierfür allerdings eine Umrechnung der Disparitäten in Entfernungswerte (siehe Abschnitt 1.2.3) und (beispielsweise aus Odometriemessungen) eine hinreichend genaue Schätzung für die relative Orientierung als Startwert (Fehler kleiner als 5° , siehe [58]). Da größere Entfernungen im Allgemeinen nur mit größerer Unsicherheit bestimmt werden können, bietet sich eine gewichtete Version des IDC-Algorithmus [75] an, bei der sichere Entfernungswerte stärker gewichtet werden können.

⁹Abkürzung für „Iterative Dual Correspondence“

Optische Flussfelder als Ortssignaturen: Zur Bestimmung geometrischer Merkmale der Umgebung kann das bestehende System durch Methoden erweitert werden, die eine größere biologische Plausibilität als panoramisches Stereo besitzen. Betrachtet man das große Gesichtsfeld vieler Insekten und Säugetiere, so liegt eine Auswertung der Bewegungsparallaxe nahe. Das Flussfeld, d.h. die Verschiebungen zwischen korrespondierenden Elementen in aufeinanderfolgenden Bildern kann dann direkt als Ortssignatur verwendet werden. Da die Flussvektoren von der Bewegungsrichtung abhängen, muss – im Gegensatz zur Verwendung der Disparitätssignatur – bei einem späteren Wiederfinden der Ort aus einer ähnlichen Richtung „angefahren“ werden. Um diese Beschränkung aufzuheben, ist eine Umrechnung in Entfernungswerte¹⁰ oder eine äquivalente Beschreibung nötig. Dabei ist zu berücksichtigen, dass in und entgegengesetzt zur Translationsrichtung die Flussvektoren verschwinden und somit in den angrenzenden Bildbereichen Entfernungsschätzungen sehr ungenau sind. Eine elegante Methode zur direkten Bestimmung von Entfernungen mit panoramischen Bildern ist der Interpolationsalgorithmus von Chahl und Srinivasan [17].

Heimfinden in dynamischen Umgebungen: In dieser Arbeit wurde mit Ausnahme von Beleuchtungsvariationen und Rauschen eine weitgehend statische Umwelt vorausgesetzt. In natürlichen Umgebungen treten jedoch weitere Veränderungen auf, z.B. Bewegungen von Personen und Fahrzeugen, die ein Wiederfinden von Orten beeinträchtigen können. Veränderungen in einem kleinen Winkelbereich sind bei einer panoramischen Ortssignatur vergleichsweise unkritisch. Bei größeren Störungen müssen diese jedoch als solche erkannt werden, damit ein Fehlverhalten vermieden werden kann. In diesem Fall könnte man versuchen, unveränderte Bereiche einer Ortssignatur zu detektieren, die dann für das Heimfinden verwendet werden. Sind alternative, zuverlässigere Ortsinformationen verfügbar z.B. durch Wegintegration, so sollte deren Einfluss verstärkt und – sofern eine Rückkehr zum gleichen Ort sichergestellt ist – die Veränderung einzelner Elemente einer Ortssignatur registriert werden. Algorithmen für den Kartenaufbau und die Lokalisierung in dynamischen Umgebungen für mobile Roboter mit Laser-Range-Finder existieren bereits [8, 43].

Zu Kapitel 2: „Fourier-basierter Bildvergleich“

Orientierungsschätzung und Heimfinden mit Fourier-transformierten Disparitätssignaturen: Die Verwendung der Fouriertransformation bietet sich auch für panoramische Disparitätssignaturen an. Allerdings sollte

¹⁰Ist die Geschwindigkeit unbekannt, so ist dies nur bis auf einen Skalierungsfaktor möglich [52].

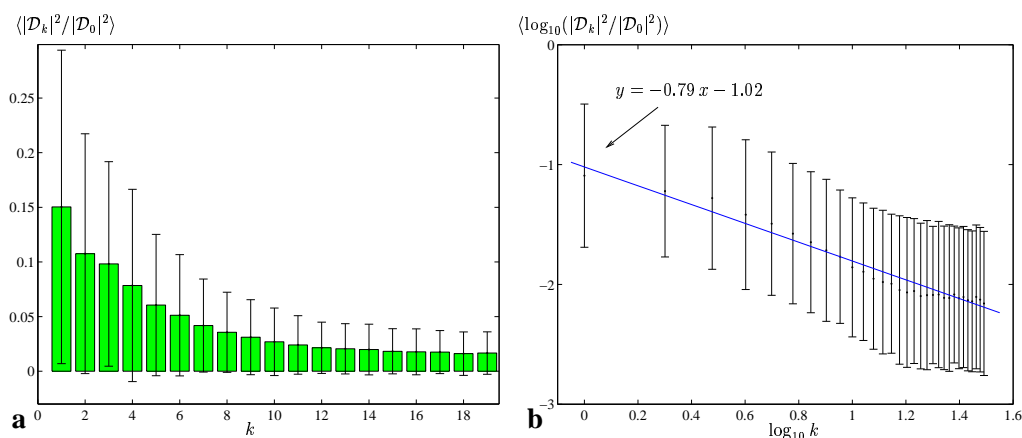


Abbildung E.1: Frequenzspektrum der Disparitätssignaturen der Datenbank: Disparitäten in dem durch den Kabelstrang verdeckten Bereich wurden durch Interpolation aus den benachbarten Werten bestimmt. **a:** Die Amplituden der Fourierkoeffizienten (gemittelt über 1250 Disparitätssignaturen) nehmen mit steigender Frequenz ab, jedoch ist der Abfall im Vergleich zu den Bildsignaturen (siehe Abb. 2.2) deutlich schwächer. **b:** In doppellogarithmischer Darstellung kann die Verteilung durch eine Gerade approximiert werden. Fehlerbalken geben die Standardabweichung an.

– zumal wenn nur wenige Koeffizienten benutzt werden – eine alternative Bestimmung der Fourierkoeffizienten gewählt werden, beispielsweise durch Minimierung der gewichteten Fehlerfunktion

$$\sum_i q_i \left(d_i - \frac{1}{2} \mathbf{a}_0 - \sum_{k=1}^K \mathbf{a}_k \cos(\omega_k i) + \mathbf{b}_k \sin(\omega_k i) \right)^2, \quad (\text{E.1})$$

wobei die Gewichte durch die Varianzwerte gegeben sein könnten, d.h. $q_i = \frac{1}{v_i}$. Abweichungen der trigonometrischen Interpolation von sicheren Disparitätswerten werden dadurch stärker „bestraft“.¹¹

Nach Bestimmung der Koeffizienten kann die Fourier-basierte Orientierungsbestimmung auf Disparitätssignaturen übertragen werden. Allerdings ist – zumindest für die Datenbank der Modellhausarena – sowohl die Dominanz tiefer Frequenzen nicht so ausgeprägt wie für Bildfrequenzen als auch die Streuung größer (siehe Abbildung E.1 im Vergleich zu Abb. 2.2), so dass unter Umständen mehr Koeffizienten verwendet werden müssen. Prinzipiell sollte auch eine Anpassung des Fourier-basierten Heimfindens an Fourier-transformierte Disparitätssignaturen möglich sein, wobei allerdings wie im

¹¹Falls alle Gewichte q_i gleich sind oder $K = \frac{N}{2}$ gilt, d.h. alle Fourierkoeffizienten verwendet werden, so folgt aus der Minimierung von (E.1) die „normale“ Fouriertransformation, vgl. Gleichungen (2.46), (2.47). Anderenfalls führt die Minimierung von (E.1) zu einem linearen Gleichungssystem in den Koeffizienten $\{\mathbf{a}_k\}, \{\mathbf{b}_k\}$.

Bild-basierten Fall wegen der verwendeten Näherungen kleinere Fangbereiche zu erwarten sind.

Zu Kapitel 3: „Panoramische Bildsensoren“

Verwendung zweidimensionaler Ortssignaturen: Mit Hilfe des in Abschnitt 3.4 beschriebenen panoramischen Stereosensors sollte infolge der verbesserten Abbildungsqualität und der damit verbundenen höheren Auflösung die Verwendung zweidimensionaler Bilder bzw. Disparitätssignaturen¹² für das visuelle Heimfinden möglich sein. Die Beschränkung auf den Horizontbereich hat sich für Disparitätssignaturen in einer Büroumgebung als ungenügend erwiesen. Mangelnde Strukturen z.B. auf Wänden und Türen in diesem Winkelbereich verhindern die Bestimmung von Disparitäten. Der etwas tieferliegende Übergang Wand-Boden bzw. die Bodenleiste würde jedoch für viele Richtungen eine Entfernungsschätzung ermöglichen. Eine Stereoskopische Erkennung von Hindernissen im Nahbereich dürfte ebenso realisierbar sein wie die Bestimmung von größeren Entfernungen durch Auswertung der Bewegungsparallaxe.

Die Verwendung zweidimensionaler Signaturen kann als Basis für ein visuelles Heimfinden in drei Dimensionen dienen, beispielsweise in unebenem Gelände. Die in Kapitel 1 und in den meisten Robotersystemen angenommene Beschränkung der Bewegung auf eine Ebene ist in Hinblick auf die natürliche Umgebung vieler Tiere unrealistisch. Allerdings erhöht sich die Zahl der zu bestimmenden Parameter im Extremfall von drei (in zwei Dimensionen) auf sechs. Neben den bekannten Größen (x, y) bzw. (φ, l) für die Position und ψ bzw. s (in Bildpixel) für die Orientierung kommen eine weitere Koordinate der Position und zwei Winkel hinzu. Letztere können beispielsweise so gewählt werden, dass ψ („yaw“) und ein zusätzlicher Neigungswinkel („pitch“) die „Vorwärts-Richtung“ des Roboters festlegen und ein weiterer Winkel die Drehung („roll“) um diese angibt.

Zu Kapitel 4: „Binokulares Stereo“

Ortsrepräsentation und Heimfinden bei begrenztem Sichtfeld: Wie bereits am Ende von Abschnitt 4.2 erwähnt, soll mit Hilfe des Stereokamerasystems ein visuelles Heimfinden realisiert werden. Ein Bild-basiertes Verfahren mit beschränktem Gesichtsfeld ist in [14] beschrieben, wobei allerdings vorausgesetzt wird, dass Start- und Zielposition die gleiche Orientierung

¹²Eine effiziente Berechnung zweidimensionaler Disparitätskarten mit Hilfe zweier panoramischer Kameras wird in [41] beschrieben.

besitzen. Ein auf Disparitätssignaturen (oder auch auf stereoptisch bestimmten Entfernungswerten) basierendes Verfahren ist jedoch nicht bekannt. Um einen Ort aus verschiedenen Richtungen wiederfinden zu können, ist im Vergleich zu einem panoramischen Bildsensor eine Repräsentation des Ortes mit mehreren Bildern nötig. Als Ortsgedächtnis bietet sich anstelle eines „Platzgraphen“, wie er beispielsweise in [29, 37] verwendet wird, ein so genannter „View-Graph“ an [84]. Ein Ort wird dann durch mehrere Knoten repräsentiert, in denen die unterschiedlichen Ansichten („views“) gespeichert sind. Die Kanten sind mit den Rotationen und (für Ansichten an unterschiedlichen Orten) mit den Translationen assoziiert, die nötig sind, um von einer Ansicht zur nächsten zu gelangen.

Gemeinsame Wahrscheinlichkeitsdichte Disparitäts-selektiver Zellen: In Bezug auf die neuronale Repräsentation von Bildverschiebungen ist die Berechnung der gemeinsamen Wahrscheinlichkeitsdichte mehrerer Neurone mit unterschiedlichen Vorzugsdisparitäten und leicht versetzten rezeptiven Feldern von Interesse. Gleichzeitig kann man versuchen, die binokulare Normalisierung in Gleichung (4.25) so zu erweitern, dass eine effizientere Codierung von Disparitäten erreicht wird. Dies könnte beispielsweise bedeuten, dass (bei gleichem Informationsgehalt) möglichst wenig Neurone gleichzeitig aktiv sind („Sparse Coding“ [71]). Simoncelli und Schwartz [88] konnten zeigen, dass im Vergleich zu linearen Modellen wie beispielsweise der „Independent Component Analysis“ (ICA) durch eine (nichtlineare) Normalisierung die statistische Abhängigkeit in den Antworten monokularer Zellen auf natürliche Stimuli weiter verringert werden kann. Durch Minimierung der statistischen Abhängigkeit kann eine optimale Wahl von Normalisierungsparametern berechnet werden, vgl. [86, 102].

Identifizierung einzelner Landmarken: Infolge der vergleichsweise hohen Auflösung kann versucht werden, einzelne Objekte als Landmarken zu identifizieren. Bei nicht zu großen Objektentfernungen (beispielsweise in Büroumgebungen) ist dabei die Tiefeninformation durch Stereo hilfreich. Die Fähigkeit einzelne Objekte als Landmarken zu verwenden, spielt bei der Navigation des Menschen eine wichtige Rolle, da sie auch in veränderlichen Umgebungen (solange das betreffende Objekt wahrnehmbar ist) eine Orientierung ermöglicht und eine verbale Mitteilung beispielsweise bei der Wegauskunft erleichtert. Bienen hingegen informieren andere Arbeiterinnen über eine Futterquelle mit Hilfe des „Schwänzeltanzes“ [104], womit sie die Richtung und die dem aufintegrierten optischen Fluss entsprechende Entfernung angeben [30, 87].

Anhang

A.1 Erwartungswert und Varianz von Disparitäten

In diesem Abschnitt werden Erwartungswert und Varianz von Disparitäten, die durch Minimieren der Fehlerfunktion

$$E_m(d) := \sum_{\zeta} w(\zeta) (I_A[\zeta] - I_B[\zeta - d])^2 \quad (\text{A.1})$$

bestimmt werden, untersucht. $I_A[\zeta]$ und $I_B[\zeta]$ sind eindimensionale Bildausschnitte auf korrespondierenden Epipolarlinien. Als „Fensterfunktion“ $w(\zeta)$ kann beispielsweise eine Gaußfunktion oder eine stückweise konstante Funktion,

$$w(\zeta) = w_A(\zeta) := \begin{cases} 1 & \text{für } 0 \leq \zeta < N_A \\ 0 & \text{sonst,} \end{cases} \quad (\text{A.2})$$

verwendet werden. In letzteren Fall entspricht (A.1) dann Gleichung (1.10). Werden die Bilder z.B. durch Kamerarauschen gestört, so gilt

$$I_A[\zeta] = \bar{I}_A[\zeta] + \Delta I_A[\zeta] \quad (\text{A.3})$$

$$I_B[\zeta] = \bar{I}_B[\zeta] + \Delta I_B[\zeta] , \quad (\text{A.4})$$

wobei $\bar{I}_{A/B}[\zeta]$ die ungestörten Bilder und $\Delta I_{A/B}[\zeta]$ die Bildstörungen bezeichnen. Sofern die „wahren“ Bilddisparitäten D_ζ , für die

$$\bar{I}_B[\zeta - D_\zeta] = \bar{I}_A[\zeta] \quad (\text{A.5})$$

gilt, innerhalb des „Fensterfunktion“ $w(\zeta)$ nicht zu stark variieren, kann nahe der optimalen Disparität $I_B[\zeta - d] = I_B[\zeta - D_\zeta + D_\zeta - d] \approx I_B[\zeta - D_\zeta] + \partial I_B[\zeta - D_\zeta](D_\zeta - d)$ approximiert werden, wobei $\partial I[\zeta] := \frac{\partial}{\partial \zeta} I[\zeta]$ definiert wurde. Für die Fehlerfunktion (A.1) folgt dann näherungsweise

$$E_m(d) \approx \sum_{\zeta} w(\zeta) \left(I_A[\zeta] - I_B[\zeta - D_\zeta] + \partial I_B[\zeta - D_\zeta](d - D_\zeta) \right)^2 . \quad (\text{A.6})$$

Minimierung bezüglich d liefert mit (A.3), (A.4) und (A.5)

$$d_{\text{opt}} \approx \frac{\sum_{\zeta} w(\zeta) \left(\partial I_{\text{B}}[\zeta - D_{\zeta}]^2 D_{\zeta} + (I_{\text{B}}[\zeta - D_{\zeta}] - I_{\text{A}}[\zeta]) \partial I_{\text{B}}[\zeta - D_{\zeta}] \right)}{\sum_{\zeta} w(\zeta) \partial I_{\text{B}}[\zeta - D_{\zeta}]^2} . \quad (\text{A.7})$$

Falls der Erwartungswert von $I_{\text{B}}[\zeta - D_{\zeta}] - I_{\text{A}}[\zeta] = \Delta I_{\text{B}}[\zeta - D_{\zeta}] - \Delta I_{\text{A}}[\zeta]$ (für alle ζ mit $w(\zeta) \neq 0$) verschwindet, dann gilt für den Erwartungswert von (A.7)

$$\langle d_{\text{opt}} \rangle \approx \left\langle \frac{\sum_{\zeta} w(\zeta) \partial I_{\text{B}}[\zeta - D_{\zeta}]^2 D_{\zeta}}{\sum_{\zeta} w(\zeta) \partial I_{\text{B}}[\zeta - D_{\zeta}]^2} \right\rangle \approx \frac{\sum_{\zeta} w(\zeta) \partial \bar{I}_{\text{A}}[\zeta]^2 D_{\zeta}}{\sum_{\zeta} w(\zeta) \partial \bar{I}_{\text{A}}[\zeta]^2} . \quad (\text{A.8})$$

Für eine gleichmäßige Bildverschiebung, d.h. $D_{\zeta} = D = \text{const}$, folgt aus (A.8)

$$\langle d_{\text{opt}} \rangle \approx D . \quad (\text{A.9})$$

Für unabhängige statistische Schwankungen, d.h. für $\langle \Delta I_{\text{B}}[\zeta] \Delta I_{\text{B}}[\eta] \rangle = \langle \Delta I_{\text{A}}[\zeta] \Delta I_{\text{A}}[\eta] \rangle = \sigma_n^2 \delta_{\zeta, \eta}$, erhält man für die Varianz von (A.7)

$$\langle \Delta d^2 \rangle \approx \left\langle \left(\frac{\sum_{\zeta} w(\zeta) (\Delta I_{\text{B}}[\zeta - D_{\zeta}] - \Delta I_{\text{A}}[\zeta]) \partial I_{\text{B}}[\zeta - D_{\zeta}]}{\sum_{\zeta} w(\zeta) \partial I_{\text{B}}[\zeta - D_{\zeta}]^2} \right)^2 \right\rangle \quad (\text{A.10})$$

$$\approx \frac{2 \sum_{\zeta, \eta} w(\zeta) w(\eta) \partial \bar{I}_{\text{A}}[\zeta] \partial \bar{I}_{\text{A}}[\eta] \langle \Delta I_{\text{A}}[\zeta] \Delta I_{\text{A}}[\eta] \rangle}{(\sum_{\zeta} w(\zeta) \partial \bar{I}_{\text{A}}[\zeta]^2)^2} \quad (\text{A.11})$$

$$= \frac{2 \sigma_n^2 \sum_{\zeta} w(\zeta)^2 \partial \bar{I}_{\text{A}}[\zeta]^2}{(\sum_{\zeta} w(\zeta) \partial \bar{I}_{\text{A}}[\zeta]^2)^2} . \quad (\text{A.12})$$

Falls die Fensterfunktion (A.2) verwendet wird, dann folgt aus (A.12)

$$\langle \Delta d^2 \rangle \approx \frac{2 \sigma_n^2}{\sum_{\zeta=0}^{N_{\text{A}}-1} \partial \bar{I}_{\text{A}}[\zeta]^2} . \quad (\text{A.13})$$

Nicht unerwartet erleichtern ein geringer Störpegel σ_n^2 und ein hoher (lokaler) Bildkontrast $\partial \bar{I}_{\text{A}}[\zeta]^2$ die Bestimmung von Bilddisparitäten.

A.2 Varianz erwarteter Disparitätswerte

Im Folgenden wird eine Schätzung für die Varianz von erwarteten Disparitätswerten $d^e(\phi')$ berechnet, deren Bestimmung in 1.3.2 beschrieben wurde.

Aus Gleichung (1.27) kann der Sinussatz,

$$\frac{l}{\sin(\phi' - \phi)} = \frac{r}{\sin(\phi' - \varphi)} , \quad (\text{A.14})$$

abgeleitet werden. Mit Hilfe von Gleichung (1.25), $r \approx \alpha/d - r_0 \approx \alpha/d$, folgt aus (A.14)

$$g(\phi, \phi', d(\phi)) := \frac{1}{\alpha} l d(\phi) \sin(\phi' - \varphi) - \sin(\phi' - \phi) \approx 0 . \quad (\text{A.15})$$

Vernachlässigt man eventuell auftretende Verdeckungen, so kann durch Lösen der impliziten Gleichung (A.15) für jeden Winkel ϕ' in der zu berechnenden Signatur der entsprechende Winkel in der aktuellen Signatur $\phi_0 = \phi(\phi')$ bestimmt werden. Abweichungen in $d(\phi)$ haben Abweichungen in ϕ_0 zur Folge, d.h.

$$\begin{aligned} 0 &\approx g(\phi_0 + \Delta\phi_0, \phi', d(\phi_0 + \Delta\phi_0) + \Delta d(\phi_0 + \Delta\phi_0)) \\ &\approx g(\phi_0, \phi', d(\phi_0)) + \left. \frac{\partial g}{\partial \phi} \right|_{\phi_0} \Delta\phi_0 + \left. \frac{\partial g}{\partial d} \frac{\partial d}{\partial \phi} \right|_{\phi_0} \Delta\phi_0 + \left. \frac{\partial g}{\partial d} \right|_{\phi_0} \Delta d(\phi_0) . \end{aligned}$$

Mit $g(\phi_0, \phi', d(\phi_0)) \approx 0$ folgt daraus

$$\begin{aligned} \Delta\phi_0 &\approx \left(- \left. \frac{\partial g}{\partial d} \right|_{\phi_0} \right) / \left(\left. \frac{\partial g}{\partial \phi} \right|_{\phi_0} + \left. \frac{\partial g}{\partial d} \frac{\partial d}{\partial \phi} \right|_{\phi_0} \right) \Delta d(\phi_0) , \\ \langle \Delta\phi_0^2 \rangle &\approx \left(- \left. \frac{\partial g}{\partial d} \right|_{\phi_0} \right)^2 / \left(\left. \frac{\partial g}{\partial \phi} \right|_{\phi_0} + \left. \frac{\partial g}{\partial d} \frac{\partial d}{\partial \phi} \right|_{\phi_0} \right)^2 \langle \Delta d(\phi_0)^2 \rangle \\ &= \left(\frac{\frac{1}{\alpha} l \sin(\phi' - \varphi)}{\frac{1}{\alpha} l \partial d(\phi_0) \sin(\phi' - \varphi) + \cos(\phi' - \phi_0)} \right)^2 \langle \Delta d(\phi_0)^2 \rangle , \quad (\text{A.16}) \end{aligned}$$

$$\begin{aligned} \langle \Delta\phi_0 \Delta d(\phi_0) \rangle &\approx - \left. \frac{\partial g}{\partial d} \right|_{\phi_0} / \left(\left. \frac{\partial g}{\partial \phi} \right|_{\phi_0} + \left. \frac{\partial g}{\partial d} \frac{\partial d}{\partial \phi} \right|_{\phi_0} \right) \langle \Delta d(\phi_0)^2 \rangle \\ &= - \frac{\frac{1}{\alpha} l \sin(\phi' - \varphi)}{\frac{1}{\alpha} l \partial d(\phi_0) \sin(\phi' - \varphi) + \cos(\phi' - \phi_0)} \langle \Delta d(\phi_0)^2 \rangle , \quad (\text{A.17}) \end{aligned}$$

wobei $\partial d(\phi) := \frac{\partial}{\partial \phi} d(\phi)$ definiert wurde.

Gleichung (1.27) hat für $r > l$ die (positive) Lösung

$$r'(\phi') = \sqrt{r(\phi)^2 - l^2 \sin^2(\phi' - \varphi) - l \cos(\phi' - \varphi)} . \quad (\text{A.18})$$

Aus (A.18) folgt mit $r \approx \alpha/d$ und der Lösung $\phi_0 = \phi(\phi')$

$$\begin{aligned}
d'(\phi') &= \frac{d(\phi_0)}{\sqrt{1 - \frac{1}{\alpha^2} l^2 d(\phi_0)^2 \sin^2(\phi' - \varphi) - \frac{1}{\alpha} l d(\phi_0) \cos(\phi' - \varphi)}} \\
&\approx d(\phi_0) \left(1 + \frac{1}{2\alpha^2} l^2 d(\phi_0)^2 \sin^2(\phi' - \varphi) + \frac{1}{\alpha} l d(\phi_0) \cos(\phi' - \varphi) \right) \\
&\approx d(\phi_0) + \frac{1}{\alpha} l d(\phi_0)^2 \cos(\phi' - \varphi) , \tag{A.19}
\end{aligned}$$

wobei $\frac{1}{\alpha} l d \ll 1$ angenommen wurde. Für die Varianz gilt

$$\begin{aligned}
\langle \Delta d'^2(\phi') \rangle &= \left(\frac{\partial d'}{\partial d} \right)^2 \langle \Delta d(\phi_0)^2 \rangle + 2 \frac{\partial d'}{\partial d} \frac{\partial d'}{\partial \phi_0} \langle \Delta \phi_0 \Delta d(\phi_0) \rangle + \left(\frac{\partial d'}{\partial \phi_0} \right)^2 \langle \Delta \phi_0^2 \rangle \\
&= \left(\frac{\partial d'}{\partial d} \right)^2 \langle \Delta d(\phi_0)^2 \rangle + 2 \frac{\partial d'}{\partial d} \frac{\partial d'}{\partial d} \partial d(\phi_0) \langle \Delta \phi_0 \Delta d(\phi_0) \rangle + \left(\frac{\partial d'}{\partial d} \partial d(\phi_0) \right)^2 \langle \Delta \phi_0^2 \rangle \\
&= \left(\frac{\partial d'}{\partial d} \right)^2 \left(\langle \Delta d(\phi_0)^2 \rangle + 2 \partial d(\phi_0) \langle \Delta \phi_0 \Delta d(\phi_0) \rangle + \partial d(\phi_0)^2 \langle \Delta \phi_0^2 \rangle \right) . \tag{A.20}
\end{aligned}$$

Einsetzen von (A.19), (A.17) und (A.16) in (A.20) ergibt für $\frac{1}{\alpha} l d \ll 1$ die Näherung

$$\begin{aligned}
\langle \Delta d'^2(\phi') \rangle &\approx \left(1 + 2 \frac{1}{\alpha} l d(\phi_0) \cos(\phi' - \varphi) - 2 \frac{1}{\alpha} l \partial d(\phi_0) \frac{\sin(\phi' - \varphi)}{\cos(\phi' - \phi_0)} \right) \langle \Delta d(\phi_0)^2 \rangle .
\end{aligned}$$

Für die in Abschnitt 1.3.2 beschriebene Berechnung erwarteter Disparitätssignaturen wurde dies weiter vereinfacht und $\langle \Delta d'^2(\phi') \rangle \approx \langle \Delta d(\phi_0)^2 \rangle$ verwendet.

A.3 Lokale Bildveränderung mit Fourier-transformierten Bildern

Werden Bilder durch ihre Fourierkoeffizienten \mathcal{I}_k , $k = 0, 1, \dots, K$ dargestellt,

$$I(\phi|\mathbf{x}) = \sum_k e^{ik\phi} \mathcal{I}_k(\mathbf{x}) , \quad (\text{A.21})$$

so kann die in Gleichung (1.68) definierte lokale Bildveränderung „liv“ direkt, d.h. ohne inverse Fouriertransformation berechnet werden. Es gilt

$$\begin{aligned} \text{liv}(\mathbf{x})^2 &= \left(\frac{\partial \mathbf{I}(\mathbf{x})}{\partial x} \right)^2 \left(\frac{\partial \mathbf{I}(\mathbf{x})}{\partial y} \right)^2 - \left(\frac{\partial \mathbf{I}(\mathbf{x})}{\partial x} \frac{\partial \mathbf{I}(\mathbf{x})}{\partial y} \right)^2 \end{aligned} \quad (\text{A.22})$$

$$= \sum_{\phi} \left(\frac{\partial I(\phi|\mathbf{x})}{\partial x} \right)^2 \sum_{\phi} \left(\frac{\partial I(\phi|\mathbf{x})}{\partial y} \right)^2 - \left(\sum_{\phi} \frac{\partial I(\phi|\mathbf{x})}{\partial x} \frac{\partial I(\phi|\mathbf{x})}{\partial y} \right)^2 . \quad (\text{A.23})$$

Mit Hilfe der allgemeinen Parseval-Gleichung (siehe z.B. [64]),

$$\sum_{\phi} A(\phi) B(\phi)^* = N \sum_k \mathcal{A}_k \mathcal{B}_k^* , \quad (\text{A.24})$$

folgt für reelle Bilder $I(\phi|\mathbf{x}) = I(\phi|\mathbf{x})^*$

$$\begin{aligned} \text{liv}(\mathbf{x})^2 &= N^2 \left(\sum_k \left| \frac{\partial \mathcal{I}_k(\mathbf{x})}{\partial x} \right|^2 \sum_k \left| \frac{\partial \mathcal{I}_k(\mathbf{x})}{\partial y} \right|^2 - \left(\sum_k \frac{\partial \mathcal{I}_k(\mathbf{x})}{\partial x} \frac{\partial \mathcal{I}_k^*(\mathbf{x})}{\partial y} \right)^2 \right) . \end{aligned} \quad (\text{A.25})$$

Eine Näherung für die Bestimmung des liv-Wertes aus Fourier-transformierten Bildern, die an benachbarten Orten aufgenommen werden, kann analog zu (1.86) abgeleitet werden.

A.4 Begründung für die Näherung (1.96)

In diesem Abschnitt soll die Annahme

$$\sum_{\phi} A(\phi) \{\sin\phi, \cos\phi, \sin\phi \cos\phi\} \ll \sum_{\phi} A(\phi) \{\cos^2\phi, \sin^2\phi\} ,$$

Gleichung (1.96), begründet werden. Ist die Fourierdarstellung von $A(\phi)$ durch

$$A(\phi) = \sum_k \mathcal{A}_k e^{ik\phi} = \sum_k |\mathcal{A}_k| \exp[i(k\phi + \alpha_k)]$$

gegeben, so gilt

$$\begin{aligned} \sum_{\phi} A(\phi) \cos\phi &= N |\mathcal{A}_1| \cos\alpha_1 , \\ \sum_{\phi} A(\phi) \sin\phi &= N |\mathcal{A}_1| \sin\alpha_1 , \\ \sum_{\phi} A(\phi) \sin\phi \cos\phi &= \frac{1}{2} \sum_{\phi} A(\phi) \sin(2\phi) = \frac{1}{2} N |\mathcal{A}_2| \sin\alpha_2 , \\ \sum_{\phi} A(\phi) \cos^2\phi &= \frac{1}{2} \sum_{\phi} A(\phi) [1 + \cos(2\phi)] = \frac{1}{2} N (|\mathcal{A}_0| + |\mathcal{A}_2| \cos\alpha_2) , \\ \sum_{\phi} A(\phi) \sin^2\phi &= \frac{1}{2} \sum_{\phi} A(\phi) [1 - \cos(2\phi)] = \frac{1}{2} N (|\mathcal{A}_0| - |\mathcal{A}_2| \cos\alpha_2) . \end{aligned}$$

Gilt nun $|\mathcal{A}_1|, |\mathcal{A}_2| \ll |\mathcal{A}_0|$, so folgt unmittelbar (1.96). Dies wird im Folgenden für $A(\phi) = \partial I(\phi)^2$ untersucht.

Fourierkoeffizienten von $\partial I(\phi)^2$

Hat das Bild $I(\phi)$ die Fourierkoeffizienten \mathcal{I}_k , d.h.

$$I(\phi) = \sum_{k=0}^{N-1} \mathcal{I}_k e^{ik\phi} = \sum_k |\mathcal{I}_k| \exp[i(k\phi + \psi_k)] ,$$

so gilt

$$\begin{aligned} \partial I(\phi)^2 &= |\partial I(\phi)|^2 = \partial I(\phi)^* \partial I(\phi) = \frac{\partial}{\partial \phi} I(\phi)^* \frac{\partial}{\partial \phi} I(\phi) \\ &= \sum_{l=1}^{N-1} \sum_{m=1}^{N-1} |\mathcal{I}_l| |\mathcal{I}_m| l m \exp[i(m-l)\phi + i(\psi_m - \psi_l)] . \end{aligned}$$

Für die entsprechenden Fourierkoeffizienten folgt

$$\begin{aligned}
(\partial \mathcal{I}^2)_k &:= \frac{1}{N} \sum_{\phi} \partial I(\phi)^2 \exp(-ik\phi) \\
&= \sum_{l=1}^{N-1} \sum_{m=1}^{N-1} |\mathcal{I}_l| |\mathcal{I}_m| l m \exp[i(\psi_m - \psi_l)] \frac{1}{N} \sum_{\phi} \exp[i(m-l-k)\phi] \\
&= \sum_{l=1}^{N-1} \sum_{m=1}^{N-1} |\mathcal{I}_l| |\mathcal{I}_m| l m \exp[i(\psi_m - \psi_l)] \delta_{m,[(l+k) \bmod N]} \\
&= \sum_{l=1}^{N-1} |\mathcal{I}_l| |\mathcal{I}_{[(l+k) \bmod N]}| l [(l+k) \bmod N] \exp[i(\psi_{[(l+k) \bmod N]} - \psi_l)] .
\end{aligned}$$

Speziell gilt für $k = 0$

$$(\partial \mathcal{I}^2)_0 = \sum_{l=1}^{N-1} |\mathcal{I}_l|^2 l^2 .$$

Normalerweise sind die Phasen in natürlichen Bildern für unterschiedliche Frequenzen verschieden, so dass $|\partial \mathcal{I}_k^2| \ll \partial \mathcal{I}_0^2$ für $k > 0$ folgt. Betrachtet man zur Veranschaulichung den Fall $|\mathcal{I}_k| \propto k^{-1}$, so erhält man

$$\begin{aligned}
(\partial \mathcal{I}^2)_0 &\propto \sum_{l=1}^{N-1} 1 = N - 1 , \\
|(\partial \mathcal{I}^2)_k| &\propto \left| \sum_{l=1}^{N-1} \exp[i(\psi_{[(k+l) \bmod N]} - \psi_l)] \right| \ll N - 1 \quad \text{für } k > 0 .
\end{aligned}$$

A.5 PCA auf rotationssymmetrischem Datensatz

Im Folgenden wird ein Datensatz betrachtet, der M eindimensionale Panoramabilder (bestehend aus N Pixel) in allen N möglichen Orientierungen enthält, d.h. insgesamt $M \times N$ Bilder. Wird auf diesem eine PCA („Principal Component Analysis“, Hauptachsentransformation) berechnet so entspricht – wie gezeigt wird – die resultierende Transformation der Fouriertransformation, d.h. die Vektoren $\mathbf{v}^k = (v_0^k, v_1^k, \dots, v_{N-1}^k)^\top$, $v_x^k := e^{i\omega_k x}$, $\omega_k := \frac{2\pi}{N}k$ ($k = 0, 1, \dots, N-1$) sind Eigenvektoren und $N^2 \langle |\mathcal{I}_k|^2 \rangle := \frac{1}{M} \sum_m \sum_x I_x^m e^{i\omega_k x} \sum_y I_y^m e^{-i\omega_k y}$ die entsprechenden Eigenwerte der Kovarianzmatrix.

Um die Notation zu vereinfachen, werden die um j Pixel rotierten Bilder $\mathbf{I}_n^m := (I_{(0-n) \bmod N}^m, I_{(1-n) \bmod N}^m, \dots, I_{(N-1-n) \bmod N}^m)^\top$ im Anschluss an das „Originalbild“ $\mathbf{I}^m = (I_0^m, I_1^m, \dots, I_{N-1}^m)^\top = \mathbf{I}_0^m$ in die Datenmatrix aufgenommen:

$$\hat{\mathbf{X}} := (\mathbf{I}_0^0, \mathbf{I}_1^0, \dots, \mathbf{I}_{N-1}^0; \mathbf{I}_0^1, \mathbf{I}_1^1, \dots, \mathbf{I}_{N-1}^1; \dots; \mathbf{I}_0^{M-1}, \mathbf{I}_1^{M-1}, \dots, \mathbf{I}_{N-1}^{M-1}) , \quad (\text{A.26})$$

$$X_{nj} = I_{(n-j) \bmod N}^{\text{int}[j/N]}, \quad n = 0, 1, \dots, N-1, \quad j = 0, 1, \dots, MN-1 . \quad (\text{A.27})$$

Beispiel für $M = 2$ Panoramabilder bestehend aus $N = 3$ Pixel:

$$\hat{\mathbf{X}} = \begin{pmatrix} I_0^0 & I_2^0 & I_1^0 & I_0^1 & I_2^1 & I_1^1 \\ I_1^0 & I_0^0 & I_2^0 & I_1^1 & I_0^1 & I_2^1 \\ I_2^0 & I_1^0 & I_0^0 & I_2^1 & I_1^1 & I_0^1 \end{pmatrix} . \quad (\text{A.28})$$

Für die Kovarianzmatrix

$$\hat{\mathbf{C}} := \frac{1}{M} \hat{\mathbf{X}} \hat{\mathbf{X}}^\top , \quad (\text{A.29})$$

$$C_{yx} = \frac{1}{M} \sum_j X_{xj} X_{yj} \quad (\text{A.30})$$

$$= \frac{1}{M} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I_{(x-n) \bmod N}^m I_{(y-n) \bmod N}^m , \quad (\text{A.31})$$

lautet die Eigenwertgleichung ($x = 0, 1, \dots, N-1$)

$$\lambda_k v_x^k = \sum_{y=0}^{N-1} C_{xy} v_y^k \quad (\text{A.32})$$

$$= \frac{1}{M} \sum_{m=0}^{M-1} \sum_{y=0}^{N-1} \sum_{n=0}^{N-1} I_{(x-n) \bmod N}^m I_{(y-n) \bmod N}^m v_y^k . \quad (\text{A.33})$$

Mit $a := y - n$ folgt

$$\lambda_k v_x^k = \frac{1}{M} \sum_{m=0}^{M-1} \sum_{y=0}^{N-1} \sum_{a=y-(N-1)}^y I_{(x-y+a) \bmod N}^m I_{a \bmod N}^m v_y^k . \quad (\text{A.34})$$

Wegen der mod-Operationen und da über alle $a = 0, 1, \dots, N-1$ summiert wird, können die Summationsgrenzen geändert werden:

$$\lambda_k v_x^k = \frac{1}{M} \sum_{m=0}^{M-1} \sum_{y=0}^{N-1} \sum_{a=0}^{N-1} I_{(x-y+a) \bmod N}^m I_a^m v_y^k . \quad (\text{A.35})$$

Mit $b := x - y + a$ folgt

$$\lambda_k v_x^k = \frac{1}{M} \sum_{m=0}^{M-1} \sum_{a=0}^{N-1} I_a^m \sum_{b=x-(N-1)+a}^{x+a} I_{b \bmod N}^m v_{x-b+a}^k . \quad (\text{A.36})$$

Für $v_y^k := e^{i\omega_k y}$ mit $\omega_k := \frac{2\pi}{N}k$ gilt

$$\lambda_k v_x^k = \frac{1}{M} \sum_{m=0}^{M-1} \sum_{a=0}^{N-1} I_a^m \sum_{b=x-(N-1)+a}^{x+a} I_{b \bmod N}^m e^{i\omega_k(x-b+a)} . \quad (\text{A.37})$$

Da $e^{i\omega_k y} = e^{i\frac{2\pi}{N}ky}$ ebenfalls N -periodisch (in y) ist, können wieder die Summationsgrenzen verschoben werden:

$$\lambda_k e^{i\omega_k x} = e^{i\omega_k x} \frac{1}{M} \sum_{m=0}^{M-1} \sum_{a=0}^{N-1} I_a^m e^{i\omega_k a} \sum_{b=0}^{N-1} I_b^m e^{-i\omega_k b} . \quad (\text{A.38})$$

$\mathbf{v}^k = (v_0^k, v_1^k, \dots, v_{N-1}^k)^\top$ mit $v_x^k := e^{i\omega_k x}$ ist somit Eigenvektor von $\hat{\mathbf{C}}$. Für den entsprechenden Eigenwert folgt mit der Definition des (komplexen) Fourierkoeffizienten $\mathcal{I}_k^m := \frac{1}{N} \sum_{x=0}^{N-1} I_x^m e^{-i\omega_k x}$ für das m -te Bild

$$\lambda_k = \frac{1}{M} \sum_{m=0}^{M-1} N^2 (\mathcal{I}_k^m)^* \mathcal{I}_k^m = N^2 \frac{1}{M} \sum_{m=0}^{M-1} |\mathcal{I}_k^m|^2 = N^2 \langle |\mathcal{I}_k|^2 \rangle . \quad (\text{A.39})$$

Da Bilder reellwertig sind und deshalb $\mathcal{I}_{-k} = \mathcal{I}_k^*$ gilt, besitzen \mathbf{v}^k und \mathbf{v}^{-k} den gleichen Eigenwert $N^2 \langle |\mathcal{I}_k|^2 \rangle$. Somit sind auch $\mathbf{a}^k := \frac{1}{2}(\mathbf{v}^k + \mathbf{v}^{-k})$ und $\mathbf{b}^k := \frac{1}{2i}(\mathbf{v}^k - \mathbf{v}^{-k})$ Eigenvektoren mit den Komponenten $a_x^k = \frac{1}{2}(v_x^k + v_x^{-k}) := \cos(\omega_k x)$ bzw. $b_x^k = \frac{1}{2i}(v_x^k - v_x^{-k}) := \sin(\omega_k x)$, $x = 0, 1, \dots, N-1$.

A.6 Bestimmung der relativen Orientierung für zweidimensionale Bilder

Im Folgenden wird der in Abschnitt 2.1 hergeleitete Algorithmus zur Orientierungsbestimmung eindimensionaler Bilder auf zweidimensionale Panoramabilder (bestehend aus $N_x \times N_y$ Pixel) erweitert. Diese werden abkürzend durch die Matrix $\hat{\mathbf{I}} := (I_{j,m})_{\substack{j=0,1,\dots,N_x-1 \\ m=0,1,\dots,N_y-1}}$ bezeichnet. Für die Fourierkoeffizienten

$$\mathcal{I}_{k,l} := \frac{1}{N_x N_y} \sum_{j=0}^{N_x-1} \sum_{m=0}^{N_y-1} I_{j,m} e^{-i(\frac{2\pi}{N_x}kj + \frac{2\pi}{N_y}lm)} \quad (\text{A.40})$$

gilt für reelle Bilder $I_{j,m}^* = I_{j,m} \in \mathbb{R}$:

$$\begin{aligned} \mathcal{I}_{-k,l} &= \frac{1}{N_x N_y} \sum_j \sum_m I_{j,m} e^{-i(\frac{2\pi}{N_x}(-k)j + \frac{2\pi}{N_y}lm)} \\ &= \frac{1}{N_x N_y} \sum_j \sum_m I_{j,m} e^{i(\frac{2\pi}{N_x}kj + \frac{2\pi}{N_y}(-l)m)} \\ &= \frac{1}{N_x N_y} \left(\sum_j \sum_m I_{j,m} e^{-i(\frac{2\pi}{N_x}kj + \frac{2\pi}{N_y}(-l)m)} \right)^* \\ &= \mathcal{I}_{k,-l}^* . \end{aligned} \quad (\text{A.41})$$

Es müssen also maximal $\frac{1}{2}N_x \times N_y$ komplexe Koeffizienten berechnet werden (= $N_x \times N_y$ reelle Zahlen). Um die Notation zu vereinfachen wird im Folgenden die inverse Fouriertransformation analog zu Gleichung (2.14) durch

$$I_{jm} = \sum_{k=-K}^K \sum_{l=-L}^L \mathcal{I}_{kl} e^{i(\frac{2\pi}{N_x}kj + \frac{2\pi}{N_y}lm)} \quad (\text{A.42})$$

angegeben ($K \leq \frac{1}{2}N_x$, $L \leq \frac{1}{2}N_y$).

Für die orientierungsabhängige Korrelation zweier Bilder, vgl. Gleichung (2.26), folgt mit (A.42), (A.41) und $\sum_{j=0}^{N-1} e^{i\frac{2\pi}{N}(p-k)j} = N\delta_{pk}$:

$$\text{corr}(\hat{\mathbf{I}}, \hat{\mathbf{J}}, s) := \sum_{j=0}^{N_x-1} \sum_{m=0}^{N_y-1} I_{j,m} J_{[(j+s) \bmod N_x], m} \quad (\text{A.43})$$

$$= \sum_{j=0}^{N_x-1} \sum_{m=0}^{N_y-1} \sum_{p,q} \mathcal{I}_{p,q}^* e^{-i\frac{2\pi}{N_x}pj} e^{-i\frac{2\pi}{N_y}qm} \sum_{k,l} e^{-i\frac{2\pi}{N_x}ks} \mathcal{J}_{k,l} e^{i\frac{2\pi}{N_x}kj} e^{i\frac{2\pi}{N_y}lm} \quad (\text{A.44})$$

$$= N_x N_y \sum_{k,l} e^{-i\frac{2\pi}{N}ks} \mathcal{I}_{k,l}^* \mathcal{J}_{k,l} \quad (\text{A.45})$$

$$\begin{aligned} &= N_x N_y \left(\sum_{k=1}^K \sum_l [e^{-i\frac{2\pi}{N}ks} \mathcal{I}_{k,l}^* \mathcal{J}_{k,l} + e^{i\frac{2\pi}{N}ks} \mathcal{I}_{-k,l}^* \mathcal{J}_{-k,l}] + \sum_l \mathcal{I}_{0,l}^* \mathcal{J}_{0,l} \right) \\ &= N_x N_y \left(\sum_{k=1}^K \sum_l 2\text{Re}[e^{-i\frac{2\pi}{N}ks} \mathcal{I}_{k,l}^* \mathcal{J}_{k,l}] + \sum_l \mathcal{I}_{0,l}^* \mathcal{J}_{0,l} \right) . \end{aligned} \quad (\text{A.46})$$

Mit $\mathcal{I}_{k,l} = |\mathcal{I}_{k,l}|e^{i\psi_{k,l}}$, $\mathcal{J}_{k,l} = |\mathcal{J}_{k,l}|e^{i\varphi_{k,l}}$ und $\omega_k := \frac{2\pi}{N}k$ erhält man

$$\begin{aligned} &\text{corr}(\hat{\mathbf{I}}, \hat{\mathbf{J}}, s) \\ &= N_x N_y \left(\sum_l \mathcal{I}_{0,l}^* \mathcal{J}_{0,l} + 2 \sum_{k=1}^K \sum_l |\mathcal{I}_{k,l}| |\mathcal{J}_{k,l}| \cos(\varphi_{k,l} - \psi_{k,l} - \omega_k s) \right) . \end{aligned} \quad (\text{A.47})$$

Für den orientierungsunabhängigen Term gilt:

$$\sum_{l=-L}^L \mathcal{I}_{0,l}^* \mathcal{J}_{0,l} = |\mathcal{I}_{0,0}| |\mathcal{J}_{0,0}| + 2 \sum_{l=1}^L |\mathcal{I}_{0,l}| |\mathcal{J}_{0,l}| \cos(\varphi_{0,l} - \psi_{0,l}) , \quad (\text{A.48})$$

$$\mathcal{I}_{0,l} = \frac{1}{N_x N_y} \sum_m e^{-i\frac{2\pi}{N_y}lm} \sum_j I_{j,m} , \quad (\text{A.49})$$

$$\mathcal{I}_{0,0} = \frac{1}{N_x N_y} \sum_m \sum_j I_{j,m} \quad (\text{A.50})$$

$$= 0 \quad (\text{für Mittelwert-freie Bilder}) . \quad (\text{A.51})$$

Maximieren der Korrelation führt auf

$$0 = \frac{\partial}{\partial s} \text{corr}(\mathbf{I}, \mathbf{J}, s) \iff \quad (\text{A.52})$$

$$0 = \sum_{k=1}^K \sum_l |\mathcal{I}_{k,l}| |\mathcal{J}_{k,l}| \omega_k \sin(\varphi_{k,l} - \psi_{k,l} - \omega_k s) \quad (\text{A.53})$$

$$\begin{aligned} &= \sum_{k=1}^K \omega_k \sum_l |\mathcal{I}_{k,l}| |\mathcal{J}_{k,l}| \left(\sin(\varphi_{k,l} - \psi_{k,l}) \cos(\omega_k s) \right. \\ &\quad \left. - \cos(\varphi_{k,l} - \psi_{k,l}) \sin(\omega_k s) \right) \end{aligned} \quad (\text{A.54})$$

$$\begin{aligned} &= \sum_{k=1}^K \omega_k \left(\cos(\omega_k s) \sum_l |\mathcal{I}_{k,l}| |\mathcal{J}_{k,l}| \sin(\varphi_{k,l} - \psi_{k,l}) \right. \\ &\quad \left. - \sin(\omega_k s) \sum_l |\mathcal{I}_{k,l}| |\mathcal{J}_{k,l}| \cos(\varphi_{k,l} - \psi_{k,l}) \right) . \end{aligned} \quad (\text{A.55})$$

Durch Einführen der Größen $\mathcal{I}_k = |\mathcal{I}_k|e^{i\psi_k}$ und $\mathcal{J}_k = |\mathcal{J}_k|e^{i\varphi_k}$, die durch

$$|\mathcal{I}_k||\mathcal{J}_k| \sin(\varphi_k - \psi_k) := \sum_l |\mathcal{I}_{k,l}||\mathcal{J}_{k,l}| \sin(\varphi_{k,l} - \psi_{k,l}) \quad (\text{A.56})$$

$$= \text{Im} \left[\sum_l \mathcal{I}_{k,l}^* \mathcal{J}_{k,l} \right], \quad (\text{A.57})$$

$$|\mathcal{I}_k||\mathcal{J}_k| \cos(\varphi_k - \psi_k) := \sum_l |\mathcal{I}_{k,l}||\mathcal{J}_{k,l}| \cos(\varphi_{k,l} - \psi_{k,l}) \quad (\text{A.58})$$

$$= \text{Re} \left[\sum_l \mathcal{I}_{k,l}^* \mathcal{J}_{k,l} \right] \quad (\text{A.59})$$

definiert sind, kann Gleichung (A.55) auf die bekannte Form (2.29) zurückgeführt und somit der Algorithmus für eindimensionale Bilder verwendet werden. Mit Hilfe von (A.57) und (A.59) müssen nur das Produkt der Beträge $|\mathcal{I}_k||\mathcal{J}_k| = |\sum_l \mathcal{I}_{k,l}^* \mathcal{J}_{k,l}|$ und die Phasendifferenz $\varphi_k - \psi_k$ berechnet werden.

Da normalerweise in Bildern tiefe Frequenzen dominieren, d.h. $\langle |\mathcal{I}_{k,l}|^2 \rangle$, $\langle |\mathcal{J}_{k,l}|^2 \rangle \propto (k^2 + l^2)^{-1}$ können die Summen $\sum_l = \sum_{l=-L}^L$ auf wenige Terme beschränkt werden ($L \ll N_y$). Setzt man für den Betrag der Frequenz eine obere Schranke k_{\max} , so hängt L von k ab:

$$k^2 + l^2 \leq k_{\max}^2 \iff \quad (\text{A.60})$$

$$|l| \leq L_k := \text{rint}[\sqrt{k_{\max}^2 - k^2}], \quad k = 1, 2, \dots, K (\leq k_{\max}) . \quad (\text{A.61})$$

Alternativer Algorithmus

Ausgehend von Gleichung (A.53) kann eine alternative Bestimmung der relativen Orientierung hergeleitet werden. Analog wie im eindimensionalen Fall gibt es für zwei Bilder die am gleichen Ort in unterschiedlicher Orientierung aufgenommen wurden eine Verschiebung s , so dass für alle k, l $\varphi_{k,l} - \psi_{k,l} - \omega_k s + 2\pi n_{k,l} \approx 0$ mit $n_{k,l} \in \mathbb{Z}$ gilt. Aus Gleichung (A.53) folgt dann:

$$0 = \sum_{k=1}^K \sum_{l=-L}^L |\mathcal{I}_{k,l}||\mathcal{J}_{k,l}| \omega_k \sin(\varphi_{k,l} - \psi_{k,l} - \omega_k s) \quad (\text{A.62})$$

$$\approx \sum_k \sum_l |\mathcal{I}_{k,l}||\mathcal{J}_{k,l}| \omega_k (\varphi_{k,l} - \psi_{k,l} - \omega_k s + 2\pi n_{k,l}) \quad (\text{A.63})$$

$$\begin{aligned} &= \sum_k \omega_k \sum_l |\mathcal{I}_{k,l}||\mathcal{J}_{k,l}| (\varphi_{k,l} - \psi_{k,l} + 2\pi n_{k,l}) \\ &\quad - s \sum_k \omega_k^2 \sum_l |\mathcal{I}_{k,l}||\mathcal{J}_{k,l}| \implies \end{aligned} \quad (\text{A.64})$$

$$s \approx \frac{\sum_k \omega_k \sum_l |\mathcal{I}_{k,l}| |\mathcal{J}_{k,l}| (\varphi_{k,l} - \psi_{k,l} + 2\pi n_{k,l})}{\sum_k \omega_k^2 \sum_l |\mathcal{I}_{k,l}| |\mathcal{J}_{k,l}|} \quad (\text{A.65})$$

$$= \frac{\sum_k \omega_k^2 \sum_l |\mathcal{I}_{k,l}| |\mathcal{J}_{k,l}| s_{k,l}}{\sum_k \omega_k^2 \sum_l |\mathcal{I}_{k,l}| |\mathcal{J}_{k,l}|}, \quad (\text{A.66})$$

$$s_{k,l} := \frac{\varphi_{k,l} - \psi_{k,l} + 2\pi n_{k,l}}{\omega_k}. \quad (\text{A.67})$$

In Hinblick auf (A.66) und (A.67) lautet ein möglicher modifizierter Algorithmus:

for $k = 1, 2, \dots, K$ do:

if $k = 1$

$$\bar{s}_{1,0} = s_{1,0} = \left(\frac{\psi_{1,0} - \varphi_{1,0}}{\omega_1} \right) \bmod N_x, \quad (\text{A.68})$$

$$\alpha_{1,0} = |\mathcal{I}_{1,0}| |\mathcal{J}_{1,0}| \omega_1^2 \quad (\text{A.69})$$

else

$$n_{k,0} = \text{rint} \left[\frac{\psi_{k,0} - \varphi_{k,0} + \omega_k \bar{s}_{k-1, L_{k-1}/2}}{2\pi} \right] \in \mathbb{Z}, \quad (\text{A.70})$$

$$s_{k,0} = \frac{\varphi_{k,0} - \psi_{k,0} + 2\pi n_{k,0}}{\omega_k}, \quad \alpha_{k,0} = |\mathcal{I}_{k,0}| |\mathcal{J}_{k,0}| \omega_k^2, \quad (\text{A.71})$$

$$\bar{s}_{k,0} = \frac{\sum_{p=1}^{k-1} \sum_{q=-L_p}^{L_p} \alpha_{p,q} s_{p,q} + \alpha_{k,0} s_{k,0}}{\sum_{p=1}^{k-1} \sum_{q=-L_p}^{L_p} \alpha_{p,q} + \alpha_{k,0}} \quad (\text{A.72})$$

endif

for $l = 1, 2, \dots, L_k$ do:

$$n_{k,-l} = \text{rint} \left[\frac{\psi_{k,-l} - \varphi_{k,-l} + \omega_k \bar{s}_{k,-l-1}}{2\pi} \right] \in \mathbb{Z}, \quad (\text{A.73})$$

$$s_{k,-l} = \frac{\varphi_{k,-l} - \psi_{k,-l} + 2\pi n_{k,-l}}{\omega_k}, \quad \alpha_{k,-l} = |\mathcal{I}_{k,-l}| |\mathcal{J}_{k,-l}| \omega_k^2, \quad (\text{A.74})$$

$$\bar{s}_{k,-l} = \frac{\sum_{p=1}^{k-1} \sum_{q=-L_p}^{L_p} \alpha_{p,q} s_{p,q} + \sum_{q=-l}^{l-1} \alpha_{k,l} s_{k,l}}{\sum_{p=1}^{k-1} \sum_{q=-L_p}^{L_p} \alpha_{p,q} + \sum_{q=-l}^{l-1} \alpha_{k,l}}, \quad (\text{A.75})$$

$$n_{k,l} = \text{rint} \left[\frac{\psi_{k,l} - \varphi_{k,l} + \omega_k \bar{s}_{k,-l}}{2\pi} \right] \in \mathbb{Z}, \quad (\text{A.76})$$

$$s_{k,l} = \frac{\varphi_{k,l} - \psi_{k,l} + 2\pi n_{k,l}}{\omega_k}, \quad \alpha_{k,l} = |\mathcal{I}_{k,l}| |\mathcal{J}_{k,l}| \omega_k^2, \quad (\text{A.77})$$

$$\bar{s}_{k,l} = \frac{\sum_{p=1}^{k-1} \sum_{q=-L_p}^{L_p} \alpha_{p,q} s_{p,q} + \sum_{q=-l}^l \alpha_{k,l} s_{k,l}}{\sum_{p=1}^{k-1} \sum_{q=-L_p}^{L_p} \alpha_{p,q} + \sum_{q=-l}^l \alpha_{k,l}} \quad (\text{A.78})$$

endifor

endifor

Die ermittelte Verschiebung ist dann $s^{\text{est}} = \bar{s}_{K, L_K} \bmod N_x$.

A.7 Berechnung der Sinus- und Cosinus- koeffizienten erwarteter Bilder

In diesem Abschnitt erfolgt die Herleitung der Koeffizienten für die Berechnung erwarteter Bilder, Gleichungen (2.76) und (2.77). Mit Hilfe der Ableitung des aktuellen Bildes, Gleichung (2.74),

$$\partial I(\phi) := \frac{\partial}{\partial \phi} I(\phi) \approx \sum_{k=1}^K \mathfrak{b}_k k \cos(k\phi) - \mathfrak{a}_k k \sin(k\phi) ,$$

folgt

$$\begin{aligned} & -\frac{l}{R} \sin(\phi - \varphi) \partial I(\phi) \\ & \approx (y \cos \phi - x \sin \phi) \left(\sum_{k=1}^K \mathfrak{b}_k k \cos(k\phi) - \mathfrak{a}_k k \sin(k\phi) \right) \\ & = x \sum_{k=1}^K k \left(\mathfrak{a}_k \sin \phi \sin(k\phi) - \mathfrak{b}_k \sin \phi \cos(k\phi) \right) \\ & \quad + y \sum_{k=1}^K k \left(-\mathfrak{a}_k \sin(k\phi) \cos \phi + \mathfrak{b}_k \cos(k\phi) \cos \phi \right) \\ & = \sum_{k=1}^K \frac{1}{2} k \left[x \mathfrak{a}_k \left(\cos[(k-1)\phi] - \cos[(k+1)\phi] \right) - x \mathfrak{b}_k \left(-\sin[(k-1)\phi] \right. \right. \\ & \quad \left. \left. + \sin[(k+1)\phi] \right) + y \mathfrak{a}_k \left(-\sin[(k-1)\phi] - \sin[(k+1)\phi] \right) \right. \\ & \quad \left. + y \mathfrak{b}_k \left(\cos[(k-1)\phi] + \cos[(k+1)\phi] \right) \right] \\ & = \sum_{k=1}^K \left(\frac{1}{2} k (-\mathfrak{a}_k x + \mathfrak{b}_k y) \cos[(k+1)\phi] + \frac{1}{2} k (\mathfrak{a}_k x + \mathfrak{b}_k y) \cos[(k-1)\phi] \right. \\ & \quad \left. + \frac{1}{2} k (-\mathfrak{a}_k y - \mathfrak{b}_k x) \sin[(k+1)\phi] + \frac{1}{2} k (-\mathfrak{a}_k y + \mathfrak{b}_k x) \sin[(k-1)\phi] \right) . \end{aligned} \tag{A.79}$$

Durch Vergleich des erwarteten Bildes \mathbf{I}^e , berechnet aus dem aktuellen Bild \mathbf{I} für eine Positionsveränderung $(x, y) = l(\cos \varphi, \sin \varphi)$, d.h.

$$\begin{aligned} I^e(\phi) & \approx I(\phi) - \frac{l}{R} \sin(\phi - \varphi) \partial I(\phi) \\ & \stackrel{\text{(A.79)}}{=} \frac{1}{2} \mathfrak{a}_0 + \sum_{k=1}^K \mathfrak{a}_k \cos(k\phi) + \mathfrak{b}_k \sin(k\phi) \end{aligned} \tag{A.80}$$

$$\begin{aligned}
& + \sum_{k=2}^{K+1} \frac{1}{2}(k-1)(-\mathbf{a}_{k-1}x + b_{k-1}y) \cos(k\phi) \\
& + \sum_{k=0}^{K-1} \frac{1}{2}(k+1)(\mathbf{a}_{k+1}x + b_{k+1}y) \cos(k\phi) \\
& + \sum_{k=2}^{K+1} \frac{1}{2}(k-1)(-\mathbf{a}_{k-1}y - b_{k-1}x) \sin(k\phi) \\
& + \sum_{k=0}^{K-1} \frac{1}{2}(k+1)(-\mathbf{a}_{k+1}y + b_{k+1}x) \sin(k\phi) , \quad (\text{A.81})
\end{aligned}$$

mit der Fourierdarstellung von \mathbf{I}^e ,

$$I^e(\phi) \approx \frac{1}{2}\mathbf{a}_0^e + \sum_{k=1}^K \mathbf{a}_k^e \cos(k\phi) + \mathbf{b}_k^e \sin(k\phi) , \quad (\text{A.82})$$

erhält man die gesuchten Koeffizienten ($1 \leq k \leq \frac{1}{2}N - 1$):

$$\mathbf{a}_0^e = \mathbf{a}_0 + \mathbf{a}_1x + \mathbf{b}_1y , \quad (\text{A.83})$$

$$\mathbf{a}_k^e = \mathbf{a}_k + \frac{1}{2}(k-1)(-\mathbf{a}_{k-1}x + \mathbf{b}_{k-1}y) + \frac{1}{2}(k+1)(\mathbf{a}_{k+1}x + \mathbf{b}_{k+1}y) , \quad (\text{A.84})$$

$$\mathbf{b}_k^e = \mathbf{b}_k + \frac{1}{2}(k-1)(-\mathbf{a}_{k-1}y - \mathbf{b}_{k-1}x) + \frac{1}{2}(k+1)(-\mathbf{a}_{k+1}y + \mathbf{b}_{k+1}x) . \quad (\text{A.85})$$

Aus (A.83) - (A.85) folgt mit den Definitionen (2.77) Gleichung (2.76).

A.8 Bestimmung des Parameters λ^* , Gleichung (3.112)

Im Folgenden wird durch Entwicklung von

$$\mathbf{x}_2 = \mathbf{x}(\theta + \Delta\theta, \varphi + \Delta\varphi) , \quad (\text{A.86})$$

$$\mathbf{e}_2 = \mathbf{e}(\theta_o, \varphi) - 2[\mathbf{n}^0(\theta + \Delta\theta, \varphi + \Delta\varphi)\mathbf{e}(\theta_o, \varphi)]\mathbf{n}^0(\theta + \Delta\theta, \varphi + \Delta\varphi) \quad (\text{A.87})$$

nach $\Delta\theta$ und $\Delta\varphi$ die Näherung (3.112) abgeleitet. Infolge der Symmetrie bezüglich $\Delta\varphi = 0$, kann das Ergebnis nicht linear von $\Delta\varphi$ abhängen. Es werden deshalb Terme bis zur zweiten Ordnung in $\Delta\theta$, $\Delta\varphi$ berücksichtigt. Mit den Abkürzungen $\mathbf{x} := \mathbf{x}(\theta, \varphi)$ und $\mathbf{n}^0 := \mathbf{n}^0(\theta, \varphi)$ gilt:

$$\mathbf{x}_2 = \mathbf{x}(\theta + \Delta\theta, \varphi + \Delta\varphi) \approx \mathbf{x} + \Delta\mathbf{x} \quad [+ \Delta^2\mathbf{x}] , \quad (\text{A.88})$$

$$\Delta\mathbf{x} := \mathbf{x}_\theta(\theta, \varphi)\Delta\theta + \mathbf{x}_\varphi(\theta, \varphi)\Delta\varphi , \quad (\text{A.89})$$

$$\tilde{\mathbf{n}}^0 := \mathbf{n}^0(\theta + \Delta\theta, \varphi + \Delta\varphi) \approx \mathbf{n}^0 + \Delta\mathbf{n} + \Delta^2\mathbf{n} , \quad (\text{A.90})$$

$$\Delta\mathbf{n} := \frac{\partial\mathbf{n}^0}{\partial\theta}\Delta\theta + \frac{\partial\mathbf{n}^0}{\partial\varphi}\Delta\varphi , \quad (\text{A.91})$$

$$\Delta^2\mathbf{n} := \frac{1}{2}\frac{\partial^2\mathbf{n}^0}{\partial\theta^2}\Delta\theta^2 + \frac{\partial^2\mathbf{n}^0}{\partial\varphi\partial\theta}\Delta\theta\Delta\varphi + \frac{1}{2}\frac{\partial^2\mathbf{n}^0}{\partial\varphi^2}\Delta\varphi^2 . \quad (\text{A.92})$$

Für (A.87) folgt dann die Näherung:

$$\mathbf{e}_2 \approx \mathbf{e}_o - 2(\mathbf{n}^0\mathbf{e}_o + \Delta\mathbf{n}\mathbf{e}_o + \Delta^2\mathbf{n}\mathbf{e}_o)(\mathbf{n}^0 + \Delta\mathbf{n} + \Delta^2\mathbf{n}) \quad (\text{A.93})$$

$$\approx \mathbf{e}_o - 2[(\mathbf{n}^0\mathbf{e}_o)\mathbf{n}^0 + (\Delta\mathbf{n}\mathbf{e}_o)\mathbf{n}^0 + (\Delta^2\mathbf{n}\mathbf{e}_o)\mathbf{n}^0 + (\mathbf{n}^0\mathbf{e}_o)\Delta\mathbf{n} + (\Delta\mathbf{n}\mathbf{e}_o)\Delta\mathbf{n} + (\mathbf{n}^0\mathbf{e}_o)\Delta^2\mathbf{n}] \quad (\text{A.94})$$

$$\approx \mathbf{e} - 2[(\Delta\mathbf{n}\mathbf{e}_o)\mathbf{n}^0 + (\mathbf{n}^0\mathbf{e}_o)\Delta\mathbf{n}] - 2[(\Delta^2\mathbf{n}\mathbf{e}_o)\mathbf{n}^0 + (\Delta\mathbf{n}\mathbf{e}_o)\Delta\mathbf{n} + (\mathbf{n}^0\mathbf{e}_o)\Delta^2\mathbf{n}] \quad (\text{A.95})$$

$$=: \mathbf{e} + \Delta\mathbf{e} + \Delta^2\mathbf{e} , \quad (\text{A.96})$$

wobei $\mathbf{e}_o = \mathbf{e}(\theta_o, \varphi)$ und $\mathbf{e} := \mathbf{e}(\theta, \varphi)$ abgekürzt wurde.

Entwicklung des Zählers und Nenners von (3.109) bis zur zweiten Ordnung in $\Delta\theta$, $\Delta\varphi$ ergibt

$$(\mathbf{x}_2 - \mathbf{x}_1)((\mathbf{e}_1\mathbf{e}_2)\mathbf{e}_2 - \mathbf{e}_1) \approx (\Delta\mathbf{x}\Delta\mathbf{e}) + (\mathbf{e}\Delta\mathbf{x})(\mathbf{e}\Delta\mathbf{e}) , \quad (\text{A.97})$$

$$(\mathbf{e}_1\mathbf{e}_2)^2 - 1 \approx 2(\mathbf{e}\Delta\mathbf{e}) + 2(\mathbf{e}\Delta^2\mathbf{e}) + (\mathbf{e}\Delta\mathbf{e})^2 . \quad (\text{A.98})$$

Somit gilt

$$\lambda^* \approx \frac{(\Delta\mathbf{x}\Delta\mathbf{e}) + (\mathbf{e}\Delta\mathbf{x})(\mathbf{e}\Delta\mathbf{e})}{2(\mathbf{e}\Delta\mathbf{e}) + 2(\mathbf{e}\Delta^2\mathbf{e}) + (\mathbf{e}\Delta\mathbf{e})^2} . \quad (\text{A.99})$$

Mit $\mathbf{e}_o = \mathbf{e} - 2(\mathbf{n}^0\mathbf{e})\mathbf{n}^0$ folgt

$$\begin{aligned}
\Delta\mathbf{e} &= -2\Delta\theta\left[\left(\frac{\partial\mathbf{n}^0}{\partial\theta}\mathbf{e}_o\right)\mathbf{n}^0 + \frac{\partial\mathbf{n}^0}{\partial\theta}(\mathbf{e}_o\mathbf{n}^0)\right] - 2\Delta\varphi\left[\left(\frac{\partial\mathbf{n}^0}{\partial\varphi}\mathbf{e}_o\right)\mathbf{n}^0 + \frac{\partial\mathbf{n}^0}{\partial\varphi}(\mathbf{e}_o\mathbf{n}^0)\right] \\
&= -2\Delta\theta\left[\left(\frac{\partial\mathbf{n}^0}{\partial\theta}[\mathbf{e} - 2(\mathbf{n}^0\mathbf{e})\mathbf{n}^0]\right)\mathbf{n}^0 + \left([\mathbf{e} - 2(\mathbf{n}^0\mathbf{e})\mathbf{n}^0]\mathbf{n}^0\right)\frac{\partial\mathbf{n}^0}{\partial\theta}\right] \\
&\quad - 2\Delta\varphi\left[\left(\frac{\partial\mathbf{n}^0}{\partial\varphi}[\mathbf{e} - 2(\mathbf{n}^0\mathbf{e})\mathbf{n}^0]\right)\mathbf{n}^0 + \left([\mathbf{e} - 2(\mathbf{n}^0\mathbf{e})\mathbf{n}^0]\mathbf{n}^0\right)\frac{\partial\mathbf{n}^0}{\partial\varphi}\right] \\
&= -2\Delta\theta\left[\left(\frac{\partial\mathbf{n}^0}{\partial\theta}\mathbf{e}\right)\mathbf{n}^0 - \left(\frac{\partial\mathbf{n}^0}{\partial\theta}\mathbf{n}^0\right)(\mathbf{n}^0\mathbf{e})\mathbf{n}^0 - (\mathbf{n}^0\mathbf{e})\frac{\partial\mathbf{n}^0}{\partial\theta}\right] \\
&\quad - 2\Delta\varphi\left[\left(\frac{\partial\mathbf{n}^0}{\partial\varphi}\mathbf{e}\right)\mathbf{n}^0 - \left(\frac{\partial\mathbf{n}^0}{\partial\varphi}\mathbf{n}^0\right)(\mathbf{n}^0\mathbf{e})\mathbf{n}^0 - (\mathbf{n}^0\mathbf{e})\frac{\partial\mathbf{n}^0}{\partial\varphi}\right]. \quad (\text{A.100})
\end{aligned}$$

Ist nun die Spiegelfläche durch $\mathbf{x}(\theta, \varphi) = r(\theta)\mathbf{e}(\theta, \varphi)$ gegeben, so gilt $\frac{\partial\mathbf{n}^0}{\partial\theta}\mathbf{n}^0 = \frac{\partial\mathbf{n}^0}{\partial\varphi}\mathbf{n}^0 = 0$ und $\frac{\partial\mathbf{n}^0}{\partial\varphi}\mathbf{e} = 0$ für beliebiges (zweifach differenzierbares) $r(\theta)$. Somit folgt

$$\begin{aligned}
\Delta\mathbf{e} &= -2\Delta\theta\left[\left(\frac{\partial\mathbf{n}^0}{\partial\theta}\mathbf{e}\right)\mathbf{n}^0 - (\mathbf{n}^0\mathbf{e})\frac{\partial\mathbf{n}^0}{\partial\theta}\right] - 2\Delta\varphi\left[\left(\frac{\partial\mathbf{n}^0}{\partial\varphi}\mathbf{e}\right)\mathbf{n}^0 - (\mathbf{n}^0\mathbf{e})\frac{\partial\mathbf{n}^0}{\partial\varphi}\right] \\
&= -2\Delta\theta\left[\left(\frac{\partial\mathbf{n}^0}{\partial\theta}\mathbf{e}\right)\mathbf{n}^0 - (\mathbf{n}^0\mathbf{e})\frac{\partial\mathbf{n}^0}{\partial\theta}\right] + 2\Delta\varphi(\mathbf{n}^0\mathbf{e})\frac{\partial\mathbf{n}^0}{\partial\varphi} \quad (\text{A.101})
\end{aligned}$$

und

$$\mathbf{e}\Delta\mathbf{e} = 0. \quad (\text{A.102})$$

Gleichung (A.99) vereinfacht sich damit auf

$$\lambda^* \approx \frac{1}{2} \frac{\Delta\mathbf{x}\Delta\mathbf{e}}{\mathbf{e}\Delta^2\mathbf{e}}. \quad (\text{A.103})$$

Mit $\mathbf{x}_\theta\mathbf{n}^0 = 0$, $\mathbf{x}_\theta\frac{\partial\mathbf{n}^0}{\partial\varphi} = 0$, $\mathbf{x}_\varphi\mathbf{n}^0 = 0$ und $\mathbf{x}_\varphi\frac{\partial\mathbf{n}^0}{\partial\theta} = 0$, erhält man aus (A.89) und (A.101)

$$\Delta\mathbf{x}\Delta\mathbf{e} = 2(\mathbf{n}^0\mathbf{e})\left[\left(\mathbf{x}_\theta\frac{\partial\mathbf{n}^0}{\partial\theta}\right)\Delta\theta^2 + \left(\mathbf{x}_\varphi\frac{\partial\mathbf{n}^0}{\partial\varphi}\right)\Delta\varphi^2\right]. \quad (\text{A.104})$$

Mit (A.91), (A.92) und (A.95) sowie $\frac{\partial^2\mathbf{n}^0}{\partial\varphi\partial\theta}\mathbf{n}^0 = 0$ und $\frac{\partial\mathbf{n}^0}{\partial\varphi}\mathbf{e} = 0$ folgt nach Nebenrechnung:

$$\mathbf{e}\Delta^2\mathbf{e} = 4(\mathbf{n}^0\mathbf{e})^2(\mathbf{n}^0\Delta^2\mathbf{n}) - 2(\Delta\mathbf{n}\mathbf{e})^2 \quad (\text{A.105})$$

$$= 2(\mathbf{n}^0\mathbf{e})^2\left[\left(\mathbf{n}^0\frac{\partial^2\mathbf{n}^0}{\partial\theta^2}\right)\Delta\theta^2 + \left(\mathbf{n}^0\frac{\partial^2\mathbf{n}^0}{\partial\varphi^2}\right)\Delta\varphi^2\right] - 2\left(\frac{\partial\mathbf{n}^0}{\partial\theta}\mathbf{e}\right)^2\Delta\theta^2. \quad (\text{A.106})$$

Gleichung (A.103) ist somit äquivalent zu

$$\lambda^* \approx \frac{(\mathbf{n}^0 \mathbf{e}) \left[(\mathbf{x}_\theta \frac{\partial \mathbf{n}^0}{\partial \theta}) \Delta \theta^2 + (\mathbf{x}_\varphi \frac{\partial \mathbf{n}^0}{\partial \varphi}) \Delta \varphi^2 \right]}{2(\mathbf{n}^0 \mathbf{e})^2 \left[(\mathbf{n}^0 \frac{\partial^2 \mathbf{n}^0}{\partial \theta^2}) \Delta \theta^2 + (\mathbf{n}^0 \frac{\partial^2 \mathbf{n}^0}{\partial \varphi^2}) \Delta \varphi^2 \right] - 2(\frac{\partial \mathbf{n}^0}{\partial \theta} \mathbf{e})^2 \Delta \theta^2} . \quad (\text{A.107})$$

Mit

$$\mathbf{n}^0 \mathbf{e} = -\frac{1}{\sqrt{1 + \left(\frac{r'(\theta)}{r(\theta)}\right)^2}} , \quad (\text{A.108})$$

$$\frac{\partial \mathbf{n}^0}{\partial \theta} \mathbf{x}_\theta = r(\theta) (\mathbf{n}^0 \mathbf{e}) \left[1 + 2 \left(\frac{r'(\theta)}{r(\theta)}\right)^2 - \frac{r''(\theta)}{r(\theta)} \right] , \quad (\text{A.109})$$

$$\frac{\partial \mathbf{n}^0}{\partial \varphi} \mathbf{x}_\varphi = r(\theta) \sin \theta (\mathbf{n}^0 \mathbf{e}) \left[\sin \theta - \cos \theta \frac{r'(\theta)}{r(\theta)} \right] , \quad (\text{A.110})$$

$$\mathbf{n}^0 \frac{\partial^2 \mathbf{n}^0}{\partial \theta^2} = -(\mathbf{n}^0 \mathbf{e})^4 \left[1 + 2 \left(\frac{r'(\theta)}{r(\theta)}\right)^2 - \frac{r''(\theta)}{r(\theta)} \right]^2 , \quad (\text{A.111})$$

$$\mathbf{n}^0 \frac{\partial^2 \mathbf{n}^0}{\partial \varphi^2} = -(\mathbf{n}^0 \mathbf{e})^2 \left[\sin \theta - \cos \theta \frac{r'(\theta)}{r(\theta)} \right]^2 , \quad (\text{A.112})$$

$$\frac{\partial \mathbf{n}^0}{\partial \theta} \mathbf{e} = (\mathbf{n}^0 \mathbf{e})^3 \frac{r'(\theta)}{r(\theta)} \left[1 + 2 \left(\frac{r'(\theta)}{r(\theta)}\right)^2 - \frac{r''(\theta)}{r(\theta)} \right] \quad (\text{A.113})$$

folgt

$$\Delta \mathbf{x} \Delta \mathbf{e} = 2r \frac{\Delta \theta^2 \left[1 + 2 \left(\frac{r'}{r}\right)^2 - \frac{r''}{r} \right] + \Delta \varphi^2 \sin \theta \left[\sin \theta - \cos \theta \frac{r'}{r} \right]}{1 + \left(\frac{r'}{r}\right)^2} . \quad (\text{A.114})$$

Nach einigen Umformungen erhält man aus (A.106)

$$\mathbf{e} \Delta^2 \mathbf{e} = -2 \frac{\Delta \theta^2 \left[1 + 2 \left(\frac{r'}{r}\right)^2 - \frac{r''}{r} \right]^2 + \Delta \varphi^2 \left[\sin \theta - \cos \theta \frac{r'}{r} \right]^2}{\left[1 + \left(\frac{r'}{r}\right)^2 \right]^2} . \quad (\text{A.115})$$

Setzt man (A.115) zusammen mit (A.114) in (A.103) ein, folgt schließlich Gleichung (3.112).

A.9 Reflektierender Spiegelbereich

In 3.5.1 wurden die virtuellen Bildflächen einer rotationssymmetrischen Spiegelfläche näherungsweise berechnet. Für eine exakte Bestimmung der Orte der virtuellen Bilder muss der gesamte Ausschnitt der Spiegelfläche betrachtet werden, der zur Abbildung beiträgt, d.h. der parallel einfallende Strahlen so reflektiert, dass sie in die Blendenöffnung der Kamera fallen (siehe auch numerische Berechnung in [5]). Dieser Bereich, der von der Kamera aus gesehen die Winkelbereiche $2\Delta\theta_a$, $2\Delta\varphi_a$ abdeckt, wird im Folgenden näherungsweise berechnet.

Als Grenzfall wird zunächst ein planer Spiegel parallel zur x - y -Ebene betrachtet ($\alpha = 1$, $\gamma_S = 0$), bei dem ein parallel einfallendes Strahlenbündel nicht aufgefächert wird. Bei einer kreisförmigen Blende mit Durchmesser D gilt für diesen:

$$\begin{aligned} r(\theta)\Delta\theta_a &\approx \frac{1}{2}D \cos\theta &\iff \Delta\theta_a &\approx \frac{D \cos\theta}{2r(\theta)} , \\ r(\theta)\sin\theta\Delta\varphi_a &\approx \frac{1}{2}D &\iff \Delta\varphi_a &\approx \frac{D}{2r(\theta)\sin\theta} . \end{aligned} \quad (\text{A.116})$$

Für konvexe Spiegel mit $\alpha > 1$ sind die Spiegelbereiche in etwa um Faktor $1/\alpha$ reduziert. Es gilt näherungsweise

$$\Delta\theta_a \approx \frac{D \cos\theta}{2\alpha r(\theta)} , \quad (\text{A.117})$$

$$\Delta\varphi_a \approx \frac{D}{2r(\theta)\sin(\alpha\theta + 2\gamma_S)} . \quad (\text{A.118})$$

Dies kann wie folgt hergeleitet werden. Ein vom Punkt $\Delta\mathbf{x}$ der Blendenöffnung in Richtung θ_i , φ_i auslaufender Strahl trifft die Spiegelfläche am Ort $\mathbf{x}(\theta_m, \varphi_m)$. Damit er dort in Richtung θ_o , φ_o reflektiert wird, wobei θ_o , φ_o durch die Richtung des reflektierten Mittelpunktstrahls festgelegt ist ($\theta_o = \Theta_o(\theta)$, $\varphi_o = \varphi$), muss θ_i , φ_i folgende Bedingungen erfüllen:

$$\mathbf{x}(\theta_m, \varphi_m) = \Delta\mathbf{x} + \lambda \mathbf{e}(\theta_i, \varphi_i) \quad \lambda \in \mathbb{R} , \quad (\text{A.119})$$

$$\mathbf{e}(\Theta_o(\theta), \varphi) = \mathbf{e}(\theta_i, \varphi_i) - 2(\mathbf{e}(\theta_i, \varphi_i)\mathbf{n}^0(\theta_m, \varphi_m))\mathbf{n}^0(\theta_m, \varphi_m) . \quad (\text{A.120})$$

Linearisierung mit $\theta_i = \theta + \Delta\theta_i$, $\varphi_i = \varphi + \Delta\varphi_i$, $\theta_m = \theta + \Delta\theta$, $\varphi_m = \varphi + \Delta\varphi$ und $\lambda = r(\theta) + \varepsilon$ ergibt mit $\mathbf{x}(\theta, \varphi) = r(\theta)\mathbf{e}(\theta, \varphi)$, $\mathbf{e}(\Theta_o(\theta), \varphi) = \mathbf{e}(\theta, \varphi) - 2[\mathbf{e}(\theta, \varphi)\mathbf{n}^0(\theta, \varphi)]\mathbf{n}^0(\theta, \varphi)$ und $\mathbf{e}_\varphi(\theta, \varphi)\mathbf{n}^0(\theta, \varphi) = \mathbf{e}(\theta, \varphi)\mathbf{n}_\varphi^0(\theta, \varphi) = 0$,

$$\mathbf{x}_\theta\Delta\theta + \mathbf{x}_\varphi\Delta\varphi \approx \Delta\mathbf{x} + r\mathbf{e}_\theta\Delta\theta_i + r\mathbf{e}_\varphi\Delta\varphi_i + \varepsilon\mathbf{e} , \quad (\text{A.121})$$

$$\begin{aligned} 0 &\approx [\mathbf{e}_\theta - 2(\mathbf{e}_\theta\mathbf{n}^0)\mathbf{n}^0]\Delta\theta_i + \mathbf{e}_\varphi\Delta\varphi_i \\ &\quad - 2[(\mathbf{e}\mathbf{n}^0)\mathbf{n}_\theta^0 + (\mathbf{e}\mathbf{n}_\theta^0)\mathbf{n}^0]\Delta\theta - 2(\mathbf{e}\mathbf{n}^0)\mathbf{n}_\varphi^0\Delta\varphi . \end{aligned} \quad (\text{A.122})$$

In (A.121) und (A.122) sind alle Größen bei (θ, φ) auszuwerten. Auch im Folgenden wird zur Abkürzung die Koordinatenangabe weggelassen. Durch Multiplikation von (A.121) mit \mathbf{e}_θ und von (A.122) mit \mathbf{n}^0 erhält man ($\mathbf{x}_\varphi \mathbf{e}_\theta = \mathbf{e}_\varphi \mathbf{e}_\theta = \mathbf{e} \mathbf{e}_\theta = \mathbf{e}_\varphi \mathbf{n}^0 = \mathbf{n}_\theta^0 \mathbf{n}^0 = \mathbf{n}_\varphi^0 \mathbf{n}^0 = 0$),

$$\mathbf{x}_\theta \mathbf{e}_\theta \Delta\theta \approx \Delta\mathbf{x} \mathbf{e}_\theta + r \Delta\theta_i, \quad (\text{A.123})$$

$$0 \approx -(\mathbf{e}_\theta \mathbf{n}^0) \Delta\theta_i - 2(\mathbf{e} \mathbf{n}_\theta^0) \Delta\theta. \quad (\text{A.124})$$

Eliminieren von $\Delta\theta_i$ ergibt

$$\Delta\theta \approx \frac{\Delta\mathbf{x} \mathbf{e}_\theta}{\mathbf{x}_\theta \mathbf{e}_\theta + 2r \frac{\mathbf{e} \mathbf{n}_\theta^0}{\mathbf{e}_\theta \mathbf{n}^0}} = \frac{\Delta\mathbf{x} \mathbf{e}_\theta}{(1 + 2 \frac{\mathbf{e} \mathbf{n}_\theta^0}{\mathbf{e}_\theta \mathbf{n}^0})r}. \quad (\text{A.125})$$

Analog erhält man durch Multiplikation von (A.121) und (A.122) mit \mathbf{e}_φ ($\mathbf{x}_\theta \mathbf{e}_\varphi = \mathbf{e} \mathbf{e}_\varphi = \mathbf{n}_\theta^0 \mathbf{e}_\varphi = \mathbf{n}^0 \mathbf{e}_\varphi = \mathbf{e}_\theta \mathbf{e}_\varphi = 0$):

$$\mathbf{x}_\varphi \mathbf{e}_\varphi \Delta\varphi \approx \Delta\mathbf{x} \mathbf{e}_\varphi + r \Delta\varphi_i, \quad (\text{A.126})$$

$$0 \approx \Delta\varphi_i - 2(\mathbf{e} \mathbf{n}^0)(\mathbf{n}_\varphi^0 \mathbf{e}_\varphi) \Delta\varphi. \quad (\text{A.127})$$

Daraus folgt durch Eliminieren von $\Delta\varphi_i$:

$$\Delta\varphi \approx \frac{\Delta\mathbf{x} \mathbf{e}_\varphi}{\mathbf{x}_\varphi \mathbf{e}_\varphi - 2(\mathbf{e} \mathbf{n}^0)(\mathbf{e}_\varphi \mathbf{n}_\varphi^0)r} = \frac{\Delta\mathbf{x} \mathbf{e}_\varphi}{[\sin\theta - 2(\mathbf{e} \mathbf{n}^0)(\mathbf{e}_\varphi \mathbf{n}_\varphi^0)]r}. \quad (\text{A.128})$$

Einsetzen von $\Delta\mathbf{x} = \frac{1}{2}D(\cos\psi, \sin\psi, 0)^\top$ und r aus Gleichung (3.60) in (A.125) und (A.128) ergibt schließlich

$$\Delta\theta_a \approx \frac{D \cos\theta \cos(\psi - \varphi)}{2\alpha r(\theta)}, \quad (\text{A.129})$$

$$\Delta\varphi_a \approx \frac{D \sin(\psi - \varphi)}{2r(\theta) \sin(\alpha\theta + 2\gamma_S)}. \quad (\text{A.130})$$

Mit $\psi = \varphi$ bzw. $\psi = \varphi + \frac{1}{2}\pi$ ist dies identisch zu (A.117) und (A.118).

A.10 Hauptkrümmungen

Im Folgenden werden für rotationssymmetrische Flächen, die durch $\mathbf{x}(\theta, \varphi) = r(\theta)\mathbf{e}(\theta, \varphi)$ gegeben sind, die Hauptkrümmungen und die Hauptkrümmungsrichtungen berechnet. Diese sind die Eigenwerte und Eigenvektoren der Weingarten-Abbildung, siehe [6].

Es werden zwei Kurven,

$$\mathbf{c}_\theta(\lambda) := \mathbf{x}(\theta + \lambda, \varphi), \quad \lambda \in \mathbb{R}, \quad (\text{A.131})$$

$$\mathbf{c}_\varphi(\mu) := \mathbf{x}(\theta, \varphi + \mu), \quad \mu \in \mathbb{R}, \quad (\text{A.132})$$

betrachtet, die durch den Punkt $\mathbf{x}(\theta, \varphi)$ der Fläche verlaufen. In diesem gilt für die Ableitungen:

$$\left. \frac{d}{d\lambda} \mathbf{c}_\theta(\lambda) \right|_{\lambda=0} = \mathbf{x}_\theta(\theta, \varphi) = r(\theta)\mathbf{e}_\theta(\theta, \varphi) + r'(\theta)\mathbf{e}(\theta, \varphi), \quad (\text{A.133})$$

$$\left. \frac{d}{d\mu} \mathbf{c}_\varphi(\mu) \right|_{\mu=0} = \mathbf{x}_\varphi(\theta, \varphi) = r(\theta) \sin \theta \mathbf{e}_\varphi(\theta, \varphi). \quad (\text{A.134})$$

Der Einheitsnormalenvektor der Fläche im Punkt $\mathbf{x}(\theta, \varphi)$ ist gegeben durch

$$\mathbf{n}^0(\theta, \varphi) = -\frac{r(\theta)\mathbf{e}(\theta, \varphi) - r'(\theta)\mathbf{e}_\theta(\theta, \varphi)}{\sqrt{r(\theta)^2 + r'(\theta)^2}}. \quad (\text{A.135})$$

Für die Weingarten-Abbildung von $\mathbf{x}_\theta(\theta, \varphi)$ gilt, siehe [6]:

$$\begin{aligned} W(\mathbf{x}_\theta(\theta, \varphi)) &= \left. \frac{d}{d\lambda} n^0(\theta + \lambda, \varphi) \right|_{\lambda=0} = \frac{\partial}{\partial \theta} n^0(\theta, \varphi) \\ &= \frac{1}{2} \frac{(2rr' + 2r'r'')(r\mathbf{e} - r'\mathbf{e}_\theta)}{[r^2 + (r')^2]^{\frac{3}{2}}} \\ &\quad - \frac{[r^2 + (r')^2](r'\mathbf{e} + r\frac{\partial}{\partial \theta}\mathbf{e} - r''\mathbf{e}_\theta - r'\frac{\partial}{\partial \theta}\mathbf{e}_\theta)}{[r^2 + (r')^2]^{\frac{3}{2}}} \end{aligned} \quad (\text{A.136})$$

$$\stackrel{(*)}{=} -\left(r^2 + (r')^2\right)^{-\frac{3}{2}} \left(-r^2r'\mathbf{e} + r(r')^2\mathbf{e}_\theta - rr'r''\mathbf{e} + (r')^2r''\mathbf{e}_\theta + [r^2 + (r')^2][2r'\mathbf{e} + (r - r'')\mathbf{e}_\theta] \right) \quad (\text{A.137})$$

$$\begin{aligned} &= -\left(r^2 + (r')^2\right)^{-\frac{3}{2}} \left([-r^2r' - rr'r'' + 2r^2r' + 2(r')^3]\mathbf{e} \right. \\ &\quad \left. + [r(r')^2 + (r')^2r'' + r^3 - r^2r'' + r(r')^2 - (r')^2r'']\mathbf{e}_\theta \right) \end{aligned} \quad (\text{A.138})$$

$$= -\left(r^2 + (r')^2\right)^{\frac{3}{2}} [r^2 - rr'' + 2(r')^2] (r'\mathbf{e} + r\mathbf{e}_\theta) \quad (\text{A.139})$$

$$= -\frac{r(\theta)^2 - r(\theta)r''(\theta) + 2r'(\theta)^2}{[r(\theta)^2 + r'(\theta)^2]^{\frac{3}{2}}} \mathbf{x}_\theta(\theta, \varphi). \quad (\text{A.140})$$

Für (*) wurde $\frac{\partial}{\partial \theta}\mathbf{e} = \mathbf{e}_\theta$ und $\frac{\partial}{\partial \theta}\mathbf{e}_\theta = -\mathbf{e}$ verwendet.

Analog folgt für die Weingarten-Abbildung von $\mathbf{x}_\varphi(\theta, \varphi)$:

$$W(\mathbf{x}_\varphi(\theta, \varphi)) = \frac{d}{d\mu} n^0(\theta, \varphi + \mu) \Big|_{\mu=0} = \frac{\partial}{\partial \varphi} n^0(\theta, \varphi) \quad (\text{A.141})$$

$$= - \frac{r(\theta) \frac{\partial}{\partial \varphi} \mathbf{e}(\theta, \varphi) - r'(\theta) \frac{\partial}{\partial \varphi} \mathbf{e}_\theta(\theta, \varphi)}{\sqrt{r(\theta)^2 + r'(\theta)^2}} \quad (\text{A.142})$$

$$\stackrel{(**)}{=} - \frac{r(\theta) \sin \theta \mathbf{e}_\varphi(\theta, \varphi) - r'(\theta) \cos \theta \mathbf{e}_\varphi(\theta, \varphi)}{\sqrt{r(\theta)^2 + r'(\theta)^2}} \quad (\text{A.143})$$

$$= - \frac{r(\theta) \sin \theta - r'(\theta) \cos \theta}{r(\theta) \sin \theta \sqrt{r(\theta)^2 + r'(\theta)^2}} r(\theta) \sin \theta \mathbf{e}_\varphi(\theta, \varphi) \quad (\text{A.144})$$

$$= - \frac{r(\theta) \sin \theta - r'(\theta) \cos \theta}{r(\theta) \sin \theta \sqrt{r(\theta)^2 + r'(\theta)^2}} \mathbf{x}_\varphi(\theta, \varphi) . \quad (\text{A.145})$$

Für (**) wurde $\frac{\partial}{\partial \varphi} \mathbf{e} = \sin \theta \mathbf{e}_\varphi$ und $\frac{\partial}{\partial \varphi} \mathbf{e}_\theta = \cos \theta \mathbf{e}_\varphi$ verwendet.

Somit ist $\mathbf{x}_\theta(\theta, \varphi)$ Hauptkrümmungsrichtung mit Hauptkrümmung

$$\kappa_\theta = - \frac{r(\theta)^2 - r(\theta)r''(\theta) + 2r'(\theta)^2}{[r(\theta)^2 + r'(\theta)^2]^{\frac{3}{2}}} \quad (\text{A.146})$$

$$= - \frac{1 - \frac{r''(\theta)}{r(\theta)} + 2\left(\frac{r'(\theta)}{r(\theta)}\right)^2}{r(\theta)\left[1 + \left(\frac{r'(\theta)}{r(\theta)}\right)^2\right]^{\frac{3}{2}}} \quad (\text{A.147})$$

und $\mathbf{x}_\varphi(\theta, \varphi)$ Hauptkrümmungsrichtung mit Hauptkrümmung

$$\kappa_\varphi = - \frac{r(\theta) \sin \theta - r'(\theta) \cos \theta}{r(\theta) \sin \theta \sqrt{r(\theta)^2 + r'(\theta)^2}} \quad (\text{A.148})$$

$$= - \frac{\sin \theta - \frac{r'(\theta)}{r(\theta)} \cos \theta}{r(\theta) \sin \theta \sqrt{1 + \left(\frac{r'(\theta)}{r(\theta)}\right)^2}} . \quad (\text{A.149})$$

Die entsprechenden Hauptkrümmungsradien sind

$$R_\theta = \frac{1}{\kappa_\theta} = -r(\theta) \frac{\left[1 + \left(\frac{r'(\theta)}{r(\theta)}\right)^2\right]^{\frac{3}{2}}}{1 + 2\left(\frac{r'(\theta)}{r(\theta)}\right)^2 - \frac{r''(\theta)}{r(\theta)}} , \quad (\text{A.150})$$

$$R_\varphi = \frac{1}{\kappa_\varphi} = -r(\theta) \sin \theta \frac{\sqrt{1 + \left(\frac{r'(\theta)}{r(\theta)}\right)^2}}{\sin \theta - \frac{r'(\theta)}{r(\theta)} \cos \theta} . \quad (\text{A.151})$$

A.11 Beweis der Gleichungen (4.121) – (4.123)

Im Folgenden wird gezeigt, dass aus Gleichung (4.119),

$$c_n = \sum_{k=0}^n \frac{1}{(2k+1)(2(n-k)+1)},$$

für $n \in \mathbb{N}_0$ folgt:

$$c_n = \frac{1}{n+1} \sum_{k=0}^n \frac{1}{2k+1}, \quad (\text{A.152})$$

$$c_n > c_{n+1}, \quad (\text{A.153})$$

$$c_n - c_{n+1} > c_{n+1} - c_{n+2}. \quad (\text{A.154})$$

Beweis von (A.152):

$$\begin{aligned} & \sum_{k=0}^n \frac{1}{(2k+1)(2(n-k)+1)} \\ &= \sum_{k=0}^n \frac{1}{2(n+1)(2k+1)} + \sum_{k=0}^n \frac{1}{2(n+1)[2(n-k)+1]} \\ &= \frac{1}{2(n+1)} \sum_{k=0}^n \frac{1}{2k+1} + \frac{1}{2(n+1)} \sum_{l=0}^n \frac{1}{2l+1} \quad (l := n-k) \\ &= \frac{1}{n+1} \sum_{k=0}^n \frac{1}{2k+1} \quad \square \end{aligned}$$

Beweis von (A.153): Aus (A.152) folgt

$$\begin{aligned} c_{n+1} &= \frac{1}{n+2} \sum_{k=0}^{n+1} \frac{1}{2k+1} \\ &= \frac{1}{n+2} \left(\sum_{k=0}^n \frac{1}{2k+1} + \frac{1}{2(n+1)+1} \right) \\ &= \frac{1}{n+2} \left((n+1)c_n + \frac{1}{2n+3} \right). \quad (\text{A.155}) \end{aligned}$$

Mit Hilfe von (A.155) kann (A.153) gezeigt werden:

$$c_n > c_{n+1}$$

$$\begin{aligned}
\iff (n+2)c_n &> (n+2)c_{n+1} \\
&= (n+1)c_n + \frac{1}{2n+3} \\
\iff c_n &> \frac{1}{2n+3} .
\end{aligned} \tag{A.156}$$

Die Gültigkeit der Ungleichung (A.156) folgt aus

$$\begin{aligned}
\frac{1}{2n+3} &< \frac{1}{2n+1} \quad \text{für } n \in \mathbb{N}_0 \\
&= \frac{1}{n+1} (n+1) \frac{1}{2n+1} \\
&= \frac{1}{n+1} \sum_{k=0}^n \frac{1}{2n+1} \\
&< \frac{1}{n+1} \sum_{k=0}^n \frac{1}{2k+1} = c_n \quad \square
\end{aligned}$$

Beweis von (A.154): Gleichung (A.154) ist äquivalent zu

$$\begin{aligned}
0 &< c_n + c_{n+2} - 2c_{n+1} \\
\stackrel{(A.155)}{\iff} 0 &< c_n + \frac{1}{n+3} \left((n+2)c_{n+1} + \frac{1}{2n+5} \right) \\
&\quad - \frac{2}{n+2} \left((n+1)c_n + \frac{1}{2n+3} \right) \\
&= c_n + \frac{1}{n+3} \left((n+1)c_n + \frac{1}{2n+3} + \frac{1}{2n+5} \right) \\
&\quad - \frac{2}{n+2} \left((n+1)c_n + \frac{1}{2n+3} \right) \\
&= \frac{2}{(n+2)(n+3)} \left(c_n - \frac{3n+7}{(2n+3)(2n+5)} \right) \\
\iff c_n &> \frac{3n+7}{(2n+3)(2n+5)} .
\end{aligned} \tag{A.157}$$

Die Ungleichung (A.157) ist für $n = 0$ erfüllt, da $c_0 = 1 > \frac{7}{15}$. Durch vollständige Induktion kann leicht gezeigt werden, dass (A.152) äquivalent ist zu

$$c_n = \frac{1}{2(n+1)} \sum_{k=0}^{n-1} \frac{1}{k+1} + \frac{1}{n+1} \sum_{k=n}^{2n} \frac{1}{k+1} . \tag{A.158}$$

Mit Hilfe von (A.158) kann die Gültigkeit der Ungleichung (A.157) für $n \geq 1$ gezeigt werden. Es gilt:

$$\begin{aligned}
 c_n &= \frac{1}{2(n+1)} \sum_{k=0}^{n-1} \frac{1}{k+1} + \frac{1}{n+1} \sum_{k=n}^{2n} \frac{1}{k+1} \\
 &> \frac{1}{2(n+1)} n \frac{1}{n} + \frac{1}{n+1} (n+1) \frac{1}{2n+1} \\
 &= \frac{1}{2(n+1)} + \frac{1}{2n+1} \\
 &= \frac{4n+3}{2(n+1)(2n+1)} .
 \end{aligned}$$

Somit folgt (A.157) für $n \geq 1$, falls

$$\begin{aligned}
 \frac{4n+3}{2(n+1)(2n+1)} &> \frac{3n+7}{(2n+3)(2n+5)} \\
 \iff (4n+3)(2n+3)(2n+5) &> 2(3n+7)(n+1)(2n+1) \\
 \iff 4n^3 + 30n^2 + 60n + 31 &> 0 .
 \end{aligned}$$

Die letzte Zeile ist für $n \in \mathbb{N}$ erfüllt. \square

Bemerkung: Mit Hilfe einer Reihenentwicklung der Digamma-Funktion¹³, $\Psi(\frac{1}{2} + n) = -C_E - 2 \ln 2 + 2 \sum_{k=1}^n \frac{1}{2k-1}$ für $n \in \mathbb{N}$ [78] folgt

$$c_n = \frac{1}{2(n+1)} \left(C_E + 2 \ln 2 + \Psi\left(\frac{1}{2} + (n+1)\right) \right) , \quad (\text{A.159})$$

wobei $C_E \approx 0.577$ die Euler-Konstante ist.

¹³Die Digamma-Funktion ist als logarithmische Ableitung der Gamma-Funktion definiert, d.h. $\Psi(x) := \frac{\Gamma'(x)}{\Gamma(x)}$, siehe z.B. [78, 107].

Literaturverzeichnis

- [1] G. Adorni, S. Cagnoni, S. Enderle, G. K. Kraetzschmar, M. Mordonini, M. Plagge, M. Ritter, S. Sablatnög, A. Zell. Vision-based localization for mobile robots. *Robotics and Autonomous Systems*, 36:103–119, 2001.
- [2] A. Anzai, I. Ohzawa, R.D. Freeman. Neural mechanisms for encoding binocular disparity: Receptive field position versus phase. *Journal of Neurophysiology*, 82:874–890, 1999.
- [3] A. Anzai, I. Ohzawa, R.D. Freeman. Neural mechanisms for processing binocular information I. simple cells. *Journal of Neurophysiology*, 82:891–908, 1999.
- [4] A. Anzai, I. Ohzawa, R.D. Freeman. Neural mechanisms for processing binocular information II. complex cells. *Journal of Neurophysiology*, 82:909–924, 1999.
- [5] S. Baker, S.K. Nayar. A theory of catdioptric image formation. In S. Chandran, U. Desai (Hrsg.), *ICCV 1998*, S. 35–40, Bombay, 1998. IEEE Computer Society, Narosa Publishing House.
- [6] C. Bär. *Elementare Differentialgeometrie*. de Gruyter, Berlin, New York, 2001.
- [7] E. Batschelet. *Circular statistics in biology*. Academic Press, London, 1981.
- [8] O. Bengtsson, A.-J. BaerVELdt. Localization in changing environments by matching laser range scans. In *Third European Workshop on Advanced Mobile Robots (EURobot'99)*, S. 169–176, 1999.
- [9] A. Bernardino, J. Santos-Victor. A binocular stereo algorithm for log-polar foveated systems. In H.H. Bülthoff, S.-W. Lee, T.A. Poggio, C. Wallraven (Hrsg.), *BMCV 2002, LNCS* vol. 2525, S. 127–136. Springer, 2002.

-
- [10] J. Borenstein, L. Feng. Measurement and correction of systematic odometry errors in mobile robots. *IEEE Journal of Robotics and Automation*, 12:869–880, 1996.
- [11] N.I. Bronstein, K.A. Semendjajew. *Taschenbuch der Mathematik*. Harri Deutsch, Thun, 24. Auflage, 1989.
- [12] B.A. Cartwright, T.S. Collett. Landmark learning in bees: experiments and models. *J. Comp. Physiol. A*, 151:521–543, 1983.
- [13] B.A. Cartwright, T.S. Collett. Landmark maps for honeybees. *Biol. Cybern.*, 57:85–93, 1987.
- [14] R. Cassinis, D. Duina, S. Inelli, A. Rizzi. Unsupervised matching of visual landmarks for robotic homing using fourier-mellin transform. *Robotics and Autonomous Systems*, 40:131–138, 2002.
- [15] J.S. Chahl, M.V. Srinivasan. Reflective surfaces for panoramic imaging. *Applied Optics*, 36:8275–8285, 1997.
- [16] J.S. Chahl, S. Thakoorand, N.L. Bouffant, G. Stange, M.V. Srinivasan, B. Hine, S. Zornetzer. Bioinspired engineering of exploration systems: A horizon sensor/attitude reference system based on the dragonfly ocelli for mars exploration applications. *Journal of Robotic Systems*, 20(1):35–42, 2003.
- [17] J.S. Chahl, M.V. Srinivasan. Range estimation with a panoramic visual sensor. *J. Opt. Soc. Am. A*, 14(9):2144–2151, 1997.
- [18] K. Cheng. A purely geometric module in the rat’s spatial representation. *Cognition*, 23:149–178, 1986.
- [19] M. Collett, T.S. Collett. How do insects use path integration for their navigation? *Biological Cybernetics*, 83(3):245–259, 2000.
- [20] M. Collett, T.S. Collett. Memory use in insect visual navigation. *Nature Reviews Neuroscience*, 3:542–552, 2002.
- [21] M. Collett, D. Harland, T.S. Collett. The use of landmarks and panoramic context in the performance of local vectors by navigating honeybees. *J. Exp. Biol.*, 205:807–814, 2002.
- [22] T.S. Collett, B.A. Cartwright, B.A. Smith. Landmark learning and visuo-spatial memories in gerbils. *J. Comp. Physiol. A*, 170:435–442, 1992.

-
- [23] T.S. Collett, E. Dillmann, A. Giger, R. Wehner. Visual landmarks and route following in desert ants. *J. Comp. Physiol. A*, 158:835–851, 1986.
- [24] J. Conradt, P. Simon, M. Pescatore, P.F.M.J. Verschure. Saliency maps operating on stereo images detect landmarks and their distance. In J.R. Dorronsoro (Hrsg.), *ICANN 2002, LNCS* vol. 2415, S. 795–800. Springer, 2002.
- [25] H. Dahmen, R.W. Wüst, J. Zeil. Extracting egomotion parameters from optic flow: principal limits for animals and machines. In M.V. Srinivasan and S. Venkatesh (Hrsg.), *From living eyes to seeing machines*, S. 174–198. Oxford University Press, 1997.
- [26] P. Dayan, L.F. Abbot. *Theoretical Neuroscience*. MIT Press, 2001.
- [27] R.L. DeValois, K.K. DeValois. *Spatial Vision*. Oxford University Press, 1988.
- [28] K.I. Diamantaras, S.Y. Kung. *Principal Component Neural Networks: Theory and Applications*. Wiley, 1996.
- [29] T. Duckett, S. Marsland, J. Shapiro. Fast, on-line learning of globally consistent maps. *Autonomous Robots*, 12:287–300, 2002.
- [30] H.E. Esch, J.E. Burns. Distance estimation by foraging honeybees. *J. Exp. Biol.*, 199:155–162, 1996.
- [31] A.S. Etienne. The control of short-distance homing in the golden hamster. In P. Ellen, C. Thinus-Blanc (Hrsg.), *Cognitive Processes and Spatial Orientation in Animals and Man*, S. 233–251. Martinus Nijhoff, Boston, 1987.
- [32] A.S. Etienne, R. Maurer, V. Séguinot. Path integration in mammals and its interaction with visual landmarks. *J. Exp. Biol.*, 199:201–209, 1996.
- [33] O. Faugeras, B. Hotz, H. Mathieu, T. Viéville, Z. Zhang, P. Fua, E. Théron, L. Moll, G. Berry, J. Vuillemin, P. Bertin, C. Proy. Real time correlation based stereo: algorithm implementations and applications. Forschungsbericht 2013, INRIA, 1996.
- [34] D.J. Fleet, D.J. Heeger, H. Wagner. Modeling binocular neurons in primary visual cortex. In M. Jenkin, L. Harris (Hrsg.), *Computational and Biological Mechanisms of Visual Coding*, S. 103–130. Cambridge University Press, 1996.

-
- [35] D.J. Fleet, D.J. Heeger, H. Wagner. Neural encoding of binocular disparity: Energy model, position shifts and phase shifts. *Vision Research*, 36:1839–1857, 1996.
- [36] M.O. Franz, H.A. Mallot. Biomimetic robot navigation. *Robotics and Autonomous Systems*, 30:133–153, 2000.
- [37] M.O. Franz, B. Schölkopf, H.A. Mallot, H.H. Bülthoff. Learning view graphs for robot navigation. *Autonomous Robots*, 5:111–125, 1998.
- [38] M.O. Franz, B. Schölkopf, H.A. Mallot, H.H. Bülthoff. Where did I take that snapshot? Scene-based homing by image matching. *Biol. Cybern.*, 79:191–202, 1998.
- [39] N. Fujita, R.L. Klatzky, J.M. Loomis, R.G. Golledge. The encoding-error model of pathway completion without vision. *Geographical Analysis*, 25(4):295–314, 1993.
- [40] L. Gerstmayr. Implementation und Test visueller Homing-Strategien. Studienarbeit, Universität Tübingen, 2003.
- [41] J. Gluckman, S.K. Nayar, K.J. Thoresz. Real-time omnidirectional and panoramic stereo. In *Proceedings of the 1998 DARPA image understanding workshop*, S. 299–303, 1998.
- [42] W. Greiner. *Klassische Elektrodynamik*. Harri Deutsch, 5. Auflage, 1991.
- [43] D. Hähnel, R. Triebel, W. Burgard, S. Thrun. Map building with mobile robots in dynamic environments. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 2003.
- [44] E. Hecht. *Optik*. Addison-Wesley, 2. Auflage, 1992.
- [45] D.J. Heeger. Half-squaring in responses of cat striate cells. *Visual Neuroscience*, 9:427–443, 1992.
- [46] L. Hermer, E.S. Spelke. A geometric process for spatial reorientation in young children. *Nature*, 370:57–59, 1994.
- [47] U. Hoffmann. Biologisch motivierte Stereobildverarbeitung für Roboteravigation. Diplomarbeit, Universität Tübingen, 2001.
- [48] O. Holland. The Grey Walter online archive.
http://www.ias.uwe.ac.uk/extra_S./gwonline.html.

-
- [49] D.H. Hubel, T.N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cats's visual cortex. *Journal of Physiology*, 160:106–154, 1962.
- [50] W. Hübner, H.A. Mallot. Integration of metric place relations in a landmark graph. In J.R. Dorronsoro (Hrsg.), *ICANN 2002, LNCS* vol. 2415, S. 825–830. Springer, 2002.
- [51] W. Hübner, W. Stürzl, H.A. Mallot. Building metric network maps from image matching. In H.H. Bülthoff, K.R. Gegenfurtner, H.A. Mallot, R. Ulrich, F.A. Wichmann (Hrsg.), *Tübinger Wahrnehmungskonferenz*, S. 73. Knirsch Verlag, 2003.
- [52] J.J. Koenderink, A.J. van Doorn. Facts on optic flow. *Biol. Cybern.*, 56:247–254, 1987.
- [53] B. Kröse and R. Bunschoten. Probabilistic localization by appearance models and active vision. In *IEEE International Conference on Robotics and Automation (ICRA)*, S. 2255–2260, 1999.
- [54] S.-H. Lai. Robust image matching under partial occlusion and spatially varying illumination change. *Computer Vision and Image Understanding*, 78:84–98, 2000.
- [55] D. Lambrinos, R. Möller, T. Labhart, R. Pfeifer, R. Wehner. A mobile robot employing insect strategies for navigation. *Robotics and Autonomous Systems*, 30:39–64, 2000.
- [56] Y.W. Lee. *Statistical Theory of Communication*. John Wiley and Sons, 6. Auflage, 1967.
- [57] F. Lu, E. Milios. Globally consistent range scan alignment for environment mapping. *Autonomous Robots*, 4:333–349, 1997.
- [58] F. Lu, E. Milios. Robot pose estimation in unknown environments by matching 2D range scan. *Journal of Intelligent and Robotic Systems*, 18:249–275, 1997.
- [59] H.A. Mallot. *Computational Vision. Information Processing in Perception and Visual Behavior*. The MIT Press, Cambridge, MA, 2000.
- [60] H.A. Mallot, S. Gillner, P.A. Arndt. Is correspondence search in human stereo vision a coarse-to-fine process? *Biological Cybernetics*, 74:95–106, 1996.

-
- [61] H.A. Mallot, A. Roll, P.A. Arndt. Disparity-evoked vergence is driven by interocular correlation. *Vision Research*, 36(18):2925–2937, 1996.
- [62] K.V. Mardia, P.E. Jupp. *Directional Statistics*. Wiley, 2000.
- [63] B.L. McNaughton, C.A. Barnes, J.L. Gerrard, K. Gothard, M.W. Jung, J.J. Knierim, H. Kudrimoti, Y. Qin, W.E. Skaggs, M. Suster, K.L. Weaver. Deciphering the hippocampal polyglot: The hippocampus as a path integration system. *J. Exp. Biol.*, 199:173–185, 1996.
- [64] K. Meyberg, P. Vachenauer. *Höhere Mathematik*. Bd. 2. Springer Verlag, 1991.
- [65] K.D. Miller, T.W. Troyer. Neural noise can explain expansive, power-law nonlinearities in neural response functions. *Journal of Neurophysiology*, 87:653–659, 2002.
- [66] M.-L. Mittelstaedt, H. Mittelstaedt. Homing by path integration in a mammal. *Naturwissenschaften*, 67:566–567, 1980.
- [67] R. Möller. Insect visual homing strategies in a robot with analog processing. *Biol. Cybern.*, 83:231–243, 2000.
- [68] M. Müller, R. Wehner. Path integration in desert ants, *cataglyphis fortis*. *PNAS*, 85:5287–5290, 1988.
- [69] S.K. Nayar, S. Baker. A theory of catdioptric image formation. Forschungsbericht CUCS-015-97, Dept. of Computer Science, Columbia University, 1997.
- [70] I. Ohzawa, G.C. DeAngelis, R.D. Freeman. Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. *Science*, 249:1037–1041, 1990.
- [71] B. Olshausen, D. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381:607–609, 1996.
- [72] T. Pajdla, V. Hlavac. Zero phase representation of panoramic images for image based localization. In F. Solina, A. Leonardis (Hrsg.), *CAIP 1999*, LNCS vol. 1689, S. 550–557. Springer, 1999.
- [73] L. Paletta, S. Frintrop, J. Hertzberg. Robust localization using context in omnidirectional imaging. In *IEEE International Conference on Robotics and Automation (ICRA)*, S. 2072–2077, 2001.

-
- [74] P. Perona. Deformable kernels for early vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(5):488–499, 1995.
- [75] S.T. Pfister, K.L. Kreichbaum, S.I. Roumeliotis, J.W. Burdick. Weighted range sensor matching algorithms for mobile robot displacement estimation. In *IEEE International Conference on Robotics and Automation*, 2002.
- [76] W.H. Press, S.A. Teukolsky, W.T. Vetterlin, B.P. Flannery. *Numerical Recipes. The Art of Scientific Computing*. Cambridge University Press, 2. Auflage, 1999.
- [77] N. Qian. Computing stereo disparity and motion with known binocular cell properties. *Neural Computation*, 6:390–404, 1994.
- [78] Wolfram Research. Digamma function, specialized values. <http://functions.wolfram.com/GammaBetaErf/PolyGamma/03/01/>.
- [79] A. Rizzi, R. Cassinis. A robot self-localization system based on omnidirectional color images. *Robotics and Autonomous Systems*, 34(1):23–38, 2001.
- [80] T. Röfer. *Panoramic image processing and route navigation*. Dissertation, Univ. Bremen, 1998.
- [81] D.L. Ruderman. The statistics of natural images. *Network: Computation in Neural Systems*, 5:517–548, 1994.
- [82] D.L. Ruderman. Origins of scaling in natural images. *Vision Research*, 37:3385–3398, 1997.
- [83] T.D. Sanger. Stereo disparity computation using gabor filters. *Biological Cybernetics*, 59:405–418, 1988.
- [84] B. Schölkopf, H.A. Mallot. View-based cognitive mapping and path planning. *Adaptive Behavior*, 3(3):311–348, 1995.
- [85] M.D. Schroeder. JPEG algorithm and associated data structures. <http://people.aero.und.edu/~mschroed/jpeg.html>.
- [86] O. Schwartz, E.P. Simoncelli. Natural signal statistics and sensory gain control. *Nat. Neurosci.*, 4(8):819–825, 2001.
- [87] A. Si, M.V. Srinivasan, S. Zhang. Honeybee navigation: properties of the visually driven ‘odometer’. *J. Exp. Biol.*, 206:1265–127, 2003.

-
- [88] E.P. Simoncelli, O. Schwartz. Modeling surround suppression in V1 neurons with a statistically-derived normalization model. *Adv. Neural Inform. Process. Syst.*, 11:153–159, 1999.
- [89] M.V. Srinivasan, S.W. Zhang, N. Bidwell. Visually mediated odometry in honeybees. *J. Exp. Biol.*, 200:2513–252, 1997.
- [90] W. Stürzl, U. Hoffmann, H.A. Mallot. Vergence control and disparity estimation with energy neurons: Theory and implementation. In J.R. Dorronsoro (Hrsg.), *ICANN 2002, LNCS* vol. 2415, S. 1255–1260. Springer, 2002.
- [91] W. Stürzl, H.A. Mallot. Roboternavigation mit Panorama-Stereokamera. In G. Barattoff, H. Neumann (Hrsg.), *Dynamische Perzeption*, S. 183–186. Aka GmbH, Berlin, 2000.
- [92] W. Stürzl, H.A. Mallot. Fast phase-based orientation estimation for panoramic images. In R.P. Würtz, M. Lappe (Hrsg.), *Dynamic Perception*, S. 289–294. Aka GmbH, Berlin, 2002.
- [93] W. Stürzl, H.A. Mallot. Vision-based homing with a panoramic stereo sensor. In H.H. Bühlhoff, S.-W. Lee, T.A. Poggio, C. Wallraven (Hrsg.), *BMCV 2002, LNCS* vol. 2525, S. 620–628, 2002.
- [94] R. Swaminathan, M.D. Grossberg, S.K. Nayar. Caustics of catadioptric cameras. In *ICCV 2001*, S. 2–9, 2001.
- [95] R. Swaminathan, M.D. Grossberg, S.K. Nayar. Non-single view-point catadioptric cameras: Geometry and analysis. Forschungsbericht CUCS-004-01, Dept. of Computer Science, Columbia University, 2001.
- [96] W.M. Theimer, H.A. Mallot. Phase-based binocular vergence control and depth reconstruction using active vision. *Computer Vision Graphics and Image Processing: Image Understanding*, 60:343–358, 1994.
- [97] S. Thrun. Learning metric-topological maps for indoor mobile robot navigation. *Artificial Intelligence*, 99(1):21–71, 1998.
- [98] S. Thrun, D. Fox, W. Burgard, F. Dellaert. Robust monte carlo localization for mobile robots. *Artificial Intelligence*, 128(1-2):99–141, 2001.
- [99] E. Trucco, A. Verri. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall, 1998.

-
- [100] O. Trullier, S. I. Wiener, A. Berthoz, J.-A. Meyer. Biologically based artificial navigation systems: review and prospects. *Progress in Neurobiology*, 51:483–544, 1997.
- [101] J.J. Tsai, J.D. Victor. Reading a population code: a multi-scale neural model for representing binocular disparity. *Vision Research*, 43:445–466, 2003.
- [102] R. Valerio, R. Navarro. Optimal coding through divisive normalization models of V1 neurons. *Network: Comput. Neural Syst*, 14:579–593, 2003.
- [103] A. van der Schaaf. *Natural Image Statistics and Visual Processing*. Dissertation, Delft University of Technology, 1998.
- [104] K. von Frisch. *Dance Language and Orientation of Bees*. Harvard University Press, 1967.
- [105] K. Weber, S. Venkatesh, M.V. Srinivasan. Insect-inspired robotic homing. *Adaptive Behavior*, 7(1):65–96, 1999.
- [106] R. Wehner, B. Michel, P. Antonsen. Visual navigation in insects: Coupling of egocentric and geocentric information. *J. Exp. Biol.*, 199:129–140, 1996.
- [107] E.W. Weisstein. World of MATHEMATICS: Digamma function. <http://mathworld.wolfram.com/DigammaFunction.html>.
- [108] N. Winters, J. Gaspar, G. Lacey, J. Santos-Victor. Omni-directional vision for robot navigation. In *Proc. IEEE Workshop on Omnidirectional Vision - Omnivis00*, S. 21–28, 2000.
- [109] J. Zeil, M.I. Hofmann, J.S. Chahl. Catchment areas of panoramic snapshots in outdoor scenes. *Journal of the Optical society of America A*, 20(3):450–469, 2003.
- [110] Y. Zhu, N. Qian. Binocular receptive field models, disparity tuning, and characteristic disparity. *Neural Computation*, 8:1611–1641, 1996.