3D Objektrepräsentation basierend auf dem Strukturmultivektor

Diplomarbeit am Lehrstuhl für Kognitive Systeme
des Instituts für Informatik und Praktische Mathematik
der Christian–Albrechts–Universität zu Kiel
Prof. Dr. Gerald Sommer

vorgelegt von

Dipl.-Met. Martin Pörksen

Betreuer :
Dr. Norbert Krüger
Dipl.-Ing. Michael Felsberg

August 2001

Inhaltsverzeichnis

Kapitel 1

Einleitung

1.1 Motivation

In dieser Arbeit wird eine neue Ähnlichkeitsfunktion zur Korrespondenzfindung in Stereobildern basierend auf dem Strukturmultivektor [Fe00] entwickelt. Diese Art der Vorverarbeitung wird in den Akkumulationsalgorithmus von Ackermann [Ack00] zur Erzeugung von Objektrepräsentationen integriert. Bei der Verarbeitung von Stereobildern mit kalibrierten Kameras kann man aus den Korrespondenzen von 2D Punkten die zugehörigen 3D Punkte rekonstruieren. Die Position des 3D Punktes kann durch den Schnitt der beiden Halbgeraden bestimmt werden, die von den 2D Punkten in den Projektionsebenen durch die optischen Zentren der Kamera gehen [Fa93]. Weil die meisten informationstragenden Strukturen in einem Bild intrinsisch eindimensionale Linien bzw. Kanten sind [ZB90], ist es sinnvoll, bei der Rekonstruktion auch die Orientierung zu berücksichtigen. Aus 2D Punkten mit deren assoziierten Orientierungen werden 2D Liniensegmente erzeugt, d.h. die Rekonstruktion wird anhand von 2D Liniensegmentmittelpunkten durchgeführt. Aus korrespondierenden 2D Punkten mit deren assoziierten Orientierungen kann ein 3D Punkt mit assoziierter 3D Orientierung berechnet werden. Jedes der 2D Liniensegmente spannt mit dem optischen Zentrum der Kamera eine Ebene auf. Die Schnittgerade von der linken und rechten Ebene bestimmt die 3D Orientierung des 3D Segmentes.

Die Suche nach korrespondierenden Strukturen in Stereobildern ist ein großes Problem [KK96]. Dieselbe Szene wird im linken und rechten Bild aus verschiedenen Perspektiven betrachtet, was zur Folge hat, dass sich die Bildstrukturen in den Projektionen unterscheiden. Nicht nur der unterschiedliche Blickwinkel, sondern auch eine evtl. andere Reflektanz des einfallenden Lichtes der zu betrachtenden Objekte bereiten Schwierigkeiten. Außerdem sind die Positionen der korrespondierenden 2D Punkte und deren 2D Orientierungen verschieden. Durch diese Unterschiede wird die Rekonstruktion jedoch erst ermöglicht, d.h. man möchte einerseits Korrespondenzen durch Ähnlichkeiten in den Bildern finden, andererseits benötigt man für die Rekonstruktion vor allem die Unterschiede in der Orientierung. Die Positionsunterschiede können durch eine Verschiebung auf der Epipolargeraden ausgeglichen werden. Der Positionsunterschied, Disparität genannt, ist direkt, aber nicht linear mit der Tiefenschätzung verbunden. Die Disparität ist umgekehrt proportional zur Tiefe, d.h. bei großer Disparität ist die Tiefe gering. Bei zunehmender Tiefe nimmt die Genauigkeit der Rekonstruktion anhand der Disparität stark ab [HS93]. Wenn die Tiefe eines 3D Segmentes größer wird, verringert sich die mittlere Orientierungsdifferenz zwischen den 2D Segmenten im linken und rechten Bild. Die Genauigkeit der Abschätzung der 3D Orientierung hängt daher erheblich von der Tiefe des 3D Segmentes ab.

In der klassischen Stereobildverarbeitung wird oft ein Template-Matching zur Korrespondenzfindung durchgeführt [KK96]. Es wird dabei eine Ähnlichkeitsfunktion durch die Untersuchung der Grauwertdifferenzen zwischen $n \times m$ - Fenstern im linken und rechten Bild definiert. Das Template-Matching funktioniert jedoch nur, wenn die 2D Orientierungsdifferenzen nicht zu groß sind. In dieser Arbeit wird eine neue Ähnlichkeitsfunktion basierend auf dem monogene Signal [FS00a] entwickelt, das eine orthogonale Identitätszerlegung des Signals in eine lokale Amplitude als energetische Information, eine lokale Orientierung als geometrische Information und eine lokale Phase als strukturelle Information zuläßt. Der Strukturmultivektor [Fe00], aus dem sich diese drei Komponenten berechnen lassen, bildet die Grundlage der Bildvorverarbeitung. Eine intrinsische 1D Struktur kann in einem Grauwertbild durch die geometrische Information, der Orientierung, und durch eine strukturelle Information, der Phase, beschrieben werden. Die Phase charakterisiert die Kantenart bzw. Linienart, d.h. die Grauwertübergänge im Bild.

Die Bildvorverarbeitung basierend auf dem Strukturmultivektor wird in den Algorithmus von Ackermann [Ack00] integriert. Ackermann hat die Houghtransformation [Ja97] angewandt und anschließend die Houghgeraden in einzelne 2D Liniensegmente diskretisiert. Die korrespondierenden 2D Liniensegmente im linken und rechten Stereokamerabild wurden bei ihm durch das photometrische Matching und der Schnittpunktbestimmung der Epipolargerade mit den Houghgeraden bestimmt. Im Gegensatz zu Ackermann erfolgt in dieser Arbeit keine Houghtransformation, sondern eine Vorverarbeitung mit dem Strukturmultivektor und die Korrespondenzfindung basiert nun auf der Orientierungs- und Phasendifferenz zwischen den 2D Punkten in beiden Bildern. Die generierten 3D Liniensegmente bilden eine 3D Bildrepräsentation. Durch mehrfache Objektdrehung kann aus den jeweils erzeugten 3D Bildrepräsentationen eine 3D Objektrepräsentation gewonnen werden, in der die Konfidenz der 3D Segmente nach jeder Objektdrehung aktualisiert wird. Dabei entspricht die Konfidenz der Annahme des Systems, dass ein zu betrachtendes Segment vorhanden ist. Weil die Entitäten über die Zeit akkumuliert werden, kann man eine robuste 3D Objektrepräsentation aus fehlerbehafteten 3D Bildrepräsentationen gewinnen.

Bei der Bildvorverarbeitung von Ackermann wurde ein globales Verfahren benutzt, das nachträglich lokalisiert wurde, was zu Problemen z.B. bei kleinen Strukturen im Grauwertbild führte. Bei der Houghtransformation erfolgt eine Einschränkung auf Geraden im Grauwertbild, dagegen können durch die lokale Vorverarbeitung durch den Strukturmultivektor auch krummlinige Konturen erfasst werden. Das Ziel der vorliegenden Arbeit ist es, durch die Verwendung eines lokalen Operators anstatt der Houghtransformation und durch die Verwendung einer neuen Ähnlichkeitsfunktion beim Stereomatching eine bessere 3D Objektrepräsentation zu gewinnen.

1.2 Vorgehensweise

In **Kapitel 2** erfolgt die Bildvorverarbeitung mit dem Strukturmultivektor. Die Erhaltung der Identitätszerlegung des monogenen Signals gestattet eine Informationsaufspaltung in eine lokale Amplitude, eine lokale Phase und eine lokale Orientierung. Anschließend wird die Phasenkongruenz erzeugt, um für das linke Kamerabild die 2D Liniensegmentmittelpunkte zu bestimmen.

Das **Kapitel 3** behandelt das Korrespondenzproblem. Zu einem sich im linken Bild befindenden 2D Liniensegmentmittelpunkt möchte man den zugehörigen 2D Punkt im rechten Bild finden. Dazu wird eine neu entwickelte Ähnlichkeitsfunktion vorgestellt, die auf der Phasen- und Orientierungsdifferenz zwischen den 2D Punkten im linken und rechten Bild beruht. Es wird eine tiefenabhängige Funktion eingeführt, die die Gewichtung zwischen der Phasen- und Orientierungsabweichung bei der Korrespondenzfindung bestimmt. Diese tiefenabhängige Funktion wird durch statistische Untersuchungen aus den Varianzen der Orientierungs- und Phasendifferenzen bestimmt. Als Ergebnis der Ähnlichkeitsfunktion erhält man das korrespondierende 2D Punktepaar.

In **Kapitel 4** wird die Erzeugung von 3D Bildrepräsentationen und 3D Objektrepräsentationen im Detail behandelt. Dabei werden Vergleiche mit der Vorgehensweise von Ackermann durchgeführt und es erfolgt eine direkte Gegenüberstellung der Ergebnisse.

Zum Schluss wird in **Kapitel 5** die Phasenkonsistenz der 2D Liniensegmente bei mehrfacher Objektdrehung untersucht, um festzustellen, ob die 2D Phase bei Rotations- und Translationsbewegungen stabil bleibt. Dabei wird innerhalb einer Monosequenz der Kamerabilder die Erhaltung der Phase von 2D Kantenmittelpunkten betrachtet. Danach wird bei einer Stereosequenz die Phasenabweichung von korrespondierenden 2D Punkten untersucht. Um die strukturelle Information in einer 3D Repräsentation zu kodieren, wird aus den 2D Phasen korrespondierender 2D Punkte eine gemittelte Phase des zugehörigen 3D Punktes assoziiert. Dadurch enthalten die 3D Objektrepräsentationen sowohl geometrische als auch die strukturelle Informationen. Die Phasenkonsistenz von 3D Punkten bei der Objektdrehung wird betrachtet, was durch einen Phasenvergleich beim Aktualisieren einer 3D Objektrepräsentation mit einer 3D Bildrepräsentation möglich ist.

Kapitel 2

Bildvorverarbeitung

Für die Korrespondenzfindung bei Stereokamerabildern wird zunächst eine Merkmalsextraktion durchgeführt. Das Ziel ist es, die Objektkanten im Bild zu lokalisieren, um die Korrespondenzfindung in einem sparsifizierten Raum durchzuführen, in welchem die wesentlichen Bildinformationen enthalten sind. Im Gegensatz zu Ackermann [Ack00] wird jedoch keine Houghtransformation zur Kantendetektion durchgeführt. Es erfolgt stattdessen eine Bildvorverarbeitung mit dem Strukturmultivektor zur lokalen Kantenbestimmung. Die theoretische Grundlagen des Strukturmultivektors werden zu Beginn dieses Kapitels dargestellt. Anhand von Grauwerttestbildern wird die 2D Signalzerlegung in die lokale Amplitude, die lokale Orientierung und die lokale Phase demonstriert. Durch die Verwendung von zwei Bandpassfiltern ist es möglich, eine Phasenkongruenz über verschiedene Frequenzen zu erzeugen. Zum Schluss des Kapitels werden nach einer Schwellwertbildung auf die Phasenkongruenz die 2D Liniensegmentmittelpunkte bestimmt, die für die Korrespondenzfindung in Kapitel 3 und 4 relevant sind.

2.1 Strukturmultivektor

2.1.1 Theoretische Grundlagen

Zunächst erfolgt eine Darstellung der theoretischen Grundlagen des Strukturmultivektors, wie es in den Arbeiten von Felsberg und Sommer [Fe00], [FS00a], [FS00b] und [FS00c] im Detail zu entnehmen ist. Der Strukturmultivektor verbindet durch eine geeignete algebraische Darstellung, z.B. in Quaternionen, die Konzepte des Strukturtensors, siehe Jähne [Ja97], und des 1D analytischen Signals. Daher werden zuerst diese beiden Konzepte näher erläutert.

Der **Strukturtensor** ist eine geeignete Methode, um die Existenz und die Orientierung von lokalen, intrinsischen 1D und 2D Nachbarschaften zu erkennen. Aus dem Tensorfeld kann im 1D Fall das Orientierungsvektorfeld erzeugt werden. Der Strukturtensor beinhaltet die Amplitude, die als Maß für die Existenz einer lokalen Struktur gesehen werden kann, außerdem die Orientierung der lokalen Struktur und drittens nach Jähne [Ja97] zusätzlich die Kohärenz, d.h. eine Abschätzung der Abweichung von einer intrinsischen 1D Nachbarschaft. Die zuletzt genannte Eigenschaft ist eng verknüpft mit der Varianz der Orientierung, die eine quadratische Eigenschaft ist und somit nicht in einem linearen Konzept wie dem Strukturmultivektor behandelt werden kann. Den Strukturtensor kann man durch Quadraturfilter berechnen, die gleich näher erläutert werden. Der Tensor selber besitzt jedoch nicht die typischen Eigenschaften eines Quadraturfilters. Besonders die Linearität und die im Abschnitt 2.1.4. zu behandelnde Identitätszerlegung des Signals in eine lokale Amplitude, lokale Phase und lokale Orientierung gehen verloren, weil die Phase vernachlässigt wird. Die Phase enthält jedoch die lokale strukturelle Information des Signals.

Das **1D analytische Signal** $f_A(x) := f(x) - f_H(x)i$ ist ein komplexes Signal, dessen Realteil $f(x)$ durch das reelle Signal und dessen Imaginärteil $f_H(x)$ durch die Hilberttransformierte des Signals gegeben ist. In der polaren Darstellung besteht es aus einem Betrag und einem Argument. Das Argument, das als lokale Phase interpretiert werden kann, und der Betrag, der als lokale Amplitude gesehen werden kann, sind voneinander unabhängig. Real- und Imaginärteil bilden ein um $\frac{\pi}{2}$-phasenverschobenes Hilbertpaar bzw. Quadraturpaar.

Weil in diesem Kapitel quaternionwertige Funktionen verwendet werden, erfolgt eine kurze Einführung in die Algebra der **Quaternionen**, die mit $\mathbb{H}$ bezeichnet werden :

$$\mathbb{H} = \{q = a + bi + cj + dk \mid a, b, c, d \in \mathbb{R}\} \tag{2.1}$$

Die Quaternionen bilden eine assoziative Algebra über dem linearen 4D Vektorraum. Dieser 4D Vektorraum wird von der Basis $\{1, i, j, k\}$ aufgespannt. Das Produkt der Algebra, das Quaternionenprodukt genannt wird, ist definiert durch $i^2 = j^2 = -1$ und $ij = -ji = k$. Es folgen die Multiplikationsregeln :

$$k^2 = -ijji = -1 \qquad ik = iij = -j = -ki \qquad jk = -ijj = i = -kj \tag{2.2}$$

Eine Quaternion $q = a+bi+cj+dk$ läßt sich vereinfacht darstellen durch $q = s+v = [s, v]$, wobei der skalare Teil, d.h. der Realteil, mit s bezeichnet wird und der vektorielle Teil, d.h. die reine Quaternion, mit v bezeichnet wird. Die Summation zweier Quaternionen ergibt sich in der Quaternionenalgebra durch :

$$q_3 = q_1 + q_2 = [s_1 + s_2, v_1 + v_2] = [s_3, v_3] \tag{2.3}$$

Die Multiplikation zweier Quaternionen erfordert 16 reelle Multiplikationen und 12 reelle Additionen und ist nicht kommutativ.

$$q_3 = q_1q_2 = [(s_1s_2 - v_1 \cdot v_2), (s_1v_2 + s_2v_1 + v_1 \times v_2)] = [s_3, v_3] \tag{2.4}$$

Für die Erkennung lokaler Merkmale von Signalen sind **Quadraturfilter** gut geeignet. Sie bestehen aus einem geraden und einem ungeraden Bandpassfilter. Die Impulsantwort und der Output eines Quadraturfilterpaares sind ein analytisches Signal. Im 1D Fall besteht ein Quadraturfilterpaar aus einem Bandpassfilter und der Hilberttransformierten des Bandpassfilters.

Bisher konnten Quadraturfilter im 2D Fall nur bzgl. einer Vorzugsrichtung entworfen werden. Es gab bislang kein lineares Verfahren, um ein **isotropes, 2D analytisches Signal** zu erhalten, weil es keinen ungeraden 2D Filter mit isotroper Energie gibt. Als Ausweg kämen 1D Quadraturfilter mit bestimmter Vorzugsrichtung in Betracht. Bei skalarwertigen 2D Filtern gibt es keine ungeraden Filter mit isotroper Energie, aber bei vektorwertigen 2D Filtern sind die gewünschten Eigenschaften erfüllbar. Die Ursache dafür, dass ein 2D analytisches Signal bisher nicht entworfen werden konnte, liegt in der Beschränkung auf die Algebra der komplexen Zahlen. In dieser Arbeit werden Quaternionen zur Darstellung des Signals benutzt.

Eine multidimensionale Verallgemeinerung von den Quadraturfiltern sind die **sphärischen Quadraturfilter**, kurz SQF genannt. Die sphärischen Quadraturfilter zeichnen sich unter anderem dadurch aus, dass sie eine isotrope, d.h. richtungsunabhängige, Energie haben, linear sind und einen geringen Berechnungsaufwand aufweisen. Das Ergebnis der Filterung eines Signals mit einem sphärischen Quadraturfilter ist ein quaternionwertiges Feld, das sogenannte Strukturmultivektorfeld, welches im Folgenden kurz **Strukturmultivektor** genannt wird. Dieser Sachverhalt kann folgendermaßen dargestellt werden. Ein Operator $\mathcal{O}$, der die Eigenschaft eines sphärischen Quadraturfilters hat, wird auf ein Grauwertbild f angewandt. Das Ergebnis ist der Strukturmultivektor SMV, der ein Vektorfeld $SMV : \mathbb{R}^2 \to \mathbb{R}^3$ darstellt.

$$\mathcal{O}_{SQF}\{f\} = SMV \tag{2.5}$$

Die Bildvorverarbeitung mit dem Strukturmultivektor ist linear, fast ohne Informationsverlust und von geringer Komplexität, weil nur drei Faltungen durchzuführen sind. Aus der Filterantwort lassen sich die lokale Phase, die lokale Orientierung und die lokale Amplitude berechnen.

Weil die Bildvorverarbeitung in dieser Arbeit im Frequenzraum durchgeführt wird, erfolgt zunächst eine kurze Darstellung des **Faltungssatzes der Fouriertransformation**. Er besagt, dass eine Faltung im Ortsraum einer Multiplikation im Frequenzraum entspricht. Umgekehrt entspricht einer Multiplikation im Ortsraum eine Faltung im Frequenzraum. Seien $g(x)$ und $h(x)$ zwei Signale und $G(u)$ und $H(u)$ die zugehörigen Fouriertransformierten, dann gilt :

$$g(x) * h(x) \quad \circ\!\!-\!\!\bullet \quad G(u) \cdot H(u) \qquad\qquad g(x) \cdot h(x) \quad \circ\!\!-\!\!\bullet \quad G(u) * H(u) \tag{2.6}$$

Wenn h ein vektorwertiger Operator bestehend aus den Komponenten $h_r : \mathbb{R}^2 \to \mathbb{R}$, $h_i : \mathbb{R}^2 \to \mathbb{R}$ und $h_j : \mathbb{R}^2 \to \mathbb{R}$ und $\mathcal{F}$ die Fouriertransformation sind, dann ist die Filterantwort eines Operators $h \circ\!\!-\!\!\bullet H$ gegeben durch folgende Gleichung :

$$h * f = \mathcal{F}^{-1}\{H \cdot \mathcal{F}\{f\}\} \tag{2.7}$$

Durch die Verwendung der schnellen Fouriertransformation lassen sich somit Faltungen effizient berechnen. Die **Riesztransformierte** $F_R(\mathbf{u})$ eines Signals f ergibt sich zu :

$$F_R(\mathbf{u}) = \frac{iu + jv}{|\mathbf{u}|} F(\mathbf{u}) \tag{2.8}$$

Dabei ist $F(\mathbf{u})$ die Fouriertransformierte des Signals f, wobei $\mathbf{u} = (u, v)^T$ ein Vektor mit Betrag $|\mathbf{u}| = \sqrt{u^2 + v^2}$ ist.

Das **monogene Signal** ist als quaternionwertiges 2D Signal eine Verallgemeinerung des 1D analytischen Signals. Das monogene Signal $F_M(\mathbf{u})$ basiert auf der Riesztransformation, die für die 1D Hilberttransformation eine Verallgemeinerung auf höhere Dimensionen darstellt. Das so entstandene 2D analytische Signal setzt sich zusammen aus dem 2D Signal und der zugehörigen Riesztransformierten.

$$F_M(\mathbf{u}) = F(\mathbf{u}) - F_R(\mathbf{u})i \tag{2.9}$$

Das Konzept des monogenen Signals beinhaltet die lokale Phase, die lokale Amplitude und im Vergleich zum 1D analytischen Signal zusätzlich die lokale Orientierung. Das Einsetzen der Gleichungen und das Anwenden der Multiplikationsregeln für Quaternionen ergibt :

$$F_M(\mathbf{u}) = F(\mathbf{u}) - \frac{iu + jv}{|\mathbf{u}|} iF(\mathbf{u}) = F(\mathbf{u}) + \frac{u + kv}{|\mathbf{u}|} F(\mathbf{u}) \tag{2.10}$$

Eine komponentenweise Aufspaltung in zwei Übertragungsfunktionen liefert :

$$F_M(\mathbf{u}) = \left(1 + \frac{u}{|\mathbf{u}|}\right) F(\mathbf{u}) + k\frac{v}{|\mathbf{u}|} F(\mathbf{u}) \tag{2.11}$$

Diese Gleichung wird bei der Berechnung des Strukturmultivektors im Folgenden benutzt.

2.1.2 Radialer Bandpassfilter

Weil die zu untersuchenden 2D Signale aus mehreren Teilsignalen zusammengesetzt sind und somit verschiedene lokale Phasen auf verschiedenen Skalen auftreten, werden die zu betrachtenden Signale mit einem Bandpass gefiltert, um andere Teilsignale auszublenden. Die Oktave Ω bestimmt die Filterskala, d.h. je höher die Oktave gewählt wird, desto feiner ist die Skala. Die relative Bandbreite $b = 0.75$ entspricht der Bandbreite des Filters von einer Oktave und die relative Bandbreite $b = 0.55$ entspricht zwei Oktaven. Als radialer Bandpass $B(|\mathbf{u}|)$ wird der Lognormal-Filter verwendet. Er wird im Frequenzraum benutzt und ergibt sich durch :

$$B(|\mathbf{u}|) = \exp\left(-\frac{\left(\log \frac{|\mathbf{u}|}{2^{\Omega}}\right)^2}{2 \cdot (\log(b))^2} \right) \tag{2.12}$$

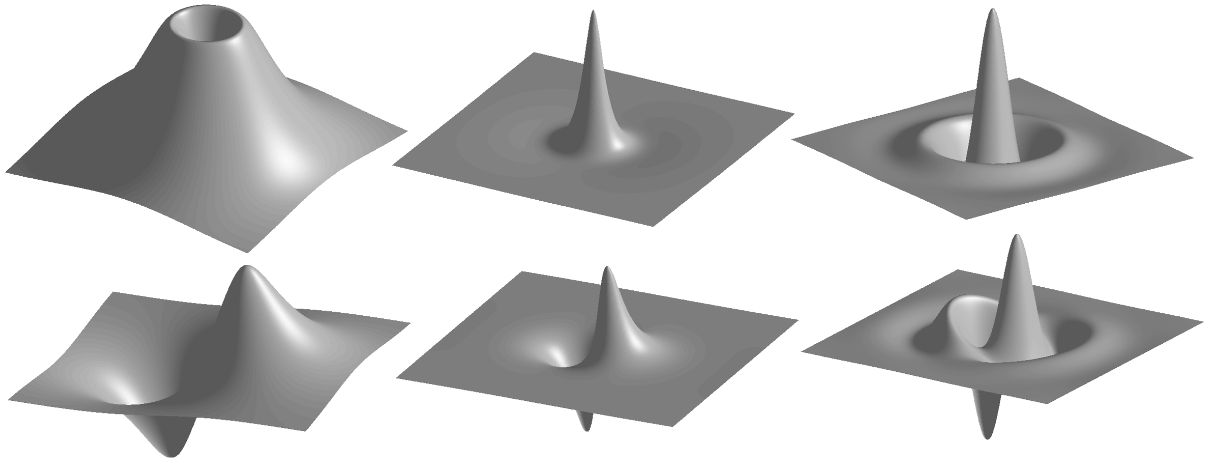


Abbildung 2.1: Sphärische Quadraturfilter basierend auf dem Lognormal-Filter. In der oberen Reihe sind die geraden Filter und in der unteren Reihe die ungeraden Filter dargestellt. In der linken Spalte sind die Filter im Frequenzraum abgebildet, die anderen sind Ortsraumfilter mit verschiedenem $\frac{b}{\Omega}$ - Verhältnis. Quelle : Felsberg und Sommer [FS00c]

2.1.3 Berechnung des Strukturmultivektors

Das Bild f wird im Ortsraum durch die Fouriertransformation $\mathcal{F}$ auf das Spektrum F im Frequenzraum abgebildet, um das Spektrum des Eingabebildes zu ermitteln.

$$F = \mathcal{F}\{f\} \tag{2.13}$$

Die Fouriertransformierten der Komponenten des sphärischen Quadraturfilters werden berechnet, indem das Spektrum des Bildes F mit dem Bandpass $B(|\mathbf{u}|)$ und den drei reellwertigen Übertragungsfunktionen 1, $\frac{u}{|\mathbf{u}|}$ und $\frac{v}{|\mathbf{u}|}$ multipliziert wird.

$$F_1 = F \cdot B(|\mathbf{u}|) \qquad F_2 = \frac{u}{|\mathbf{u}|} \cdot F \cdot B(|\mathbf{u}|) \qquad F_3 = \frac{v}{|\mathbf{u}|} \cdot F \cdot B(|\mathbf{u}|) \tag{2.14}$$

Durch die inverse Fouriertransformation $\mathcal{F}^{-1}$ erhält man dann die Realkomponente f_r, die i-Komponente f_i und die j-Komponente f_j des Strukturmultivektors. Die k-Komponente f_k wird nicht betrachtet, weil sie grundsätzlich null ist.

Für einen vektorwertigen Operator h gelten die folgenden drei Gleichungen. Die i-Komponente und j-Komponente ergeben sich durch $h_i = h_r * r_1$ und $h_j = h_r * r_2$ mittels der Riesztransformation. Dabei gelten $R_1 = \frac{u}{r}$ und $R_2 = \frac{v}{r}$.

$$h_r * f = \mathcal{F}^{-1}\{H_r \cdot \mathcal{F}\{f\}\} \tag{2.15}$$

$$h_i * f = h_r * f * r_1 = \mathcal{F}^{-1}\{H_r \cdot R_1 \cdot \mathcal{F}\{f\}\} \tag{2.16}$$

$$h_j * f = h_r * f * r_2 = \mathcal{F}^{-1}\{H_r \cdot R_2 \cdot \mathcal{F}\{f\}\} \tag{2.17}$$

Bezüglich der Gleichung 2.14 erhält man :

$$f_r = \mathcal{F}^{-1}\{F_1\} \qquad f_i = -\mathcal{F}^{-1}\{F_2\}i \qquad f_j = -\mathcal{F}^{-1}\{F_3\}i \tag{2.18}$$

Dabei spart man eine inverse Fouriertransformation, wenn man eine Aufspaltung in zwei Übertragungsfunktionen verwendet wie in Gleichung 2.11, was aus der Hermitesymmetrie der Fouriertransformation folgt :

$$f_r + if_i = \mathcal{F}^{-1}\{F_1 + F_2\} \tag{2.19}$$

Falls Eingabebilder verwendet werden, deren Ausmaße einer Zweierpotenz entsprechen, kann man die schnelle Fouriertransformation unmittelbar benutzen. Andernfalls erfolgt eine Einbettung in ein größeres Bild, dessen Pixelanzahl in horizontaler und vertikaler Richtung der nächsthöheren Zweierpotenz entspricht.

Um die Eigenschaften des Strukturmultivektors zu veranschaulichen, werden Grauwerttestbilder herangezogen. Es werden im Folgenden zwei Testbilder verwendet, ein synthetisch erzeugter Quader und eines der Stereokamerabilder bei der Aufnahme eines Kartons. Um später die Kantenextraktion erfolgreich durchzuführen, ist es erforderlich, dass die 3D Oberflächendiskontinuitäten korrespondierende Grauwertkanten im Bild haben.

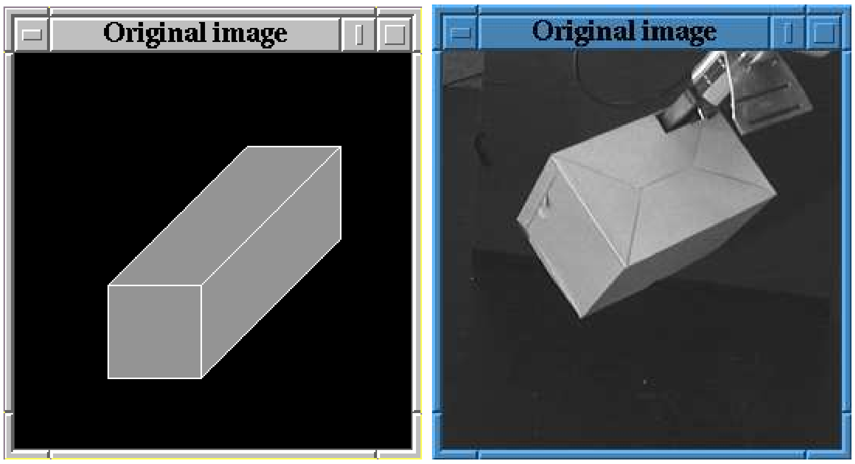


Abbildung 2.2: Eingabetestbilder : Links ein synthetischer Quader und rechts eines der Bilder der Stereokamera bei der Aufnahme eines Kartons.

Man erkennt in der Abbildung 2.3 bei dem synthetisch erzeugten Quader die komplementären Filterantworten in der i-Komponente und der j-Komponente besonders gut.

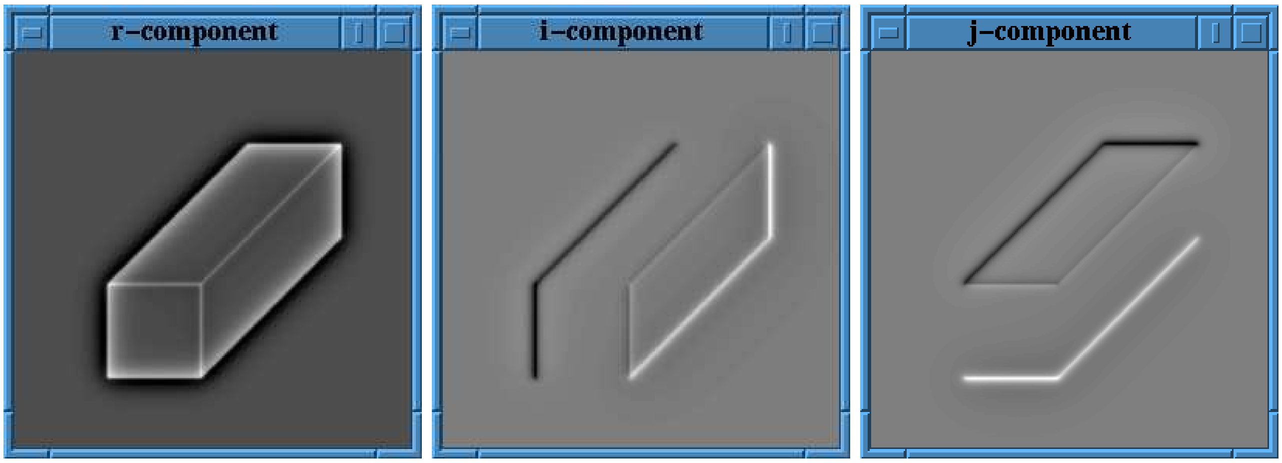


Abbildung 2.3: Links Realteil, in der Mitte i-Komponente und rechts j-Komponente des Quaders bei Oktave $\Omega = 4$

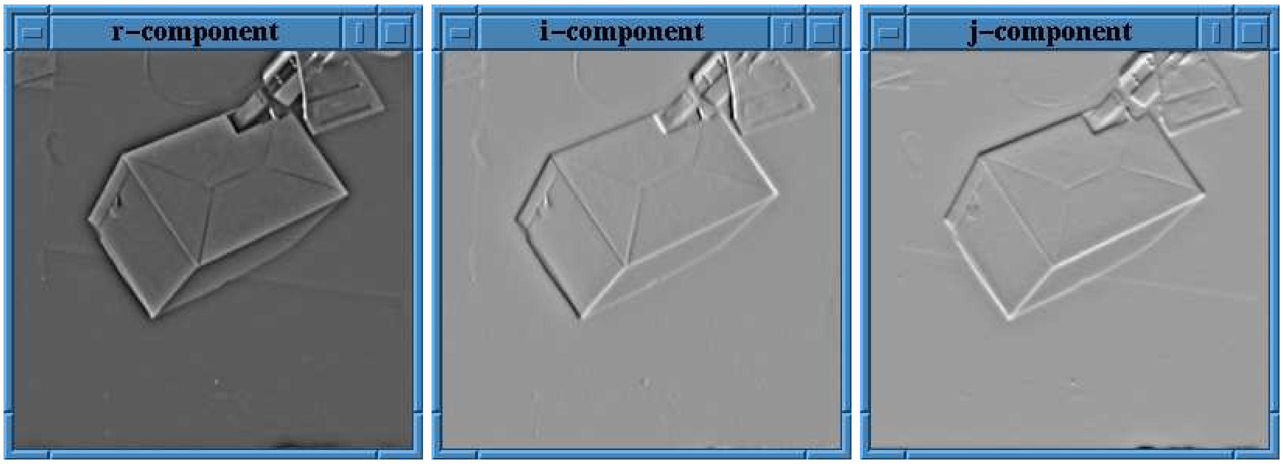


Abbildung 2.4: Links Realteil, in der Mitte i-Komponente und rechts j-Komponente des Kartons bei Oktave $\Omega = 5$

2.1.4 Zerlegung der Identität

Eine wichtige Eigenschaft des 1D analytischen Signals ist die Identifikation seines Betrages als lokale Amplitude und seines Argumentes als lokale Phase. Bei der polaren Darstellung des komplexen Signals dient die Amplitude als Quantitätsmaß und die Phase als Qualitätsmaß. Beim 2D Signal kommt noch die lokale Orientierung der Struktur als geometrische Information hinzu. Aus den drei Komponenten des Strukturmultivektors lassen sich die lokale Amplitude, die lokale Orientierung und die lokale Phase berechnen. Die Amplitude M ergibt sich aus :

$$M = \sqrt{f_r^2 + f_i^2 + f_j^2} \tag{2.20}$$

Die lokale Orientierung θ entspricht dem Winkel zwischen der i-Komponente f_i und der j-Komponente f_j modulo π. Sie nimmt Werte im Intervall $[0; \pi[$ an.

$$\theta = \frac{\arg((f_i + i f_j)^2)}{2} \tag{2.21}$$

Die lokale Phase φ ist folgendermaßen definiert :

$$\varphi = \begin{cases} \arg\left(f_r + i\sqrt{f_i^2 + f_j^2}\right) & \text{falls} \quad f_j \geq 0 \\ -\arg\left(f_r + i\sqrt{f_i^2 + f_j^2}\right) & \text{falls} \quad f_j < 0 \end{cases}$$

Die lokale Phase nimmt dabei Werte im Intervall $]-\pi; \pi]$ an. In der Abbildung 2.5 ist die Interpretation des Strukturmultivektors zu erkennen, die sich aus seiner Darstellung in sphärischen Koordinaten ergibt.

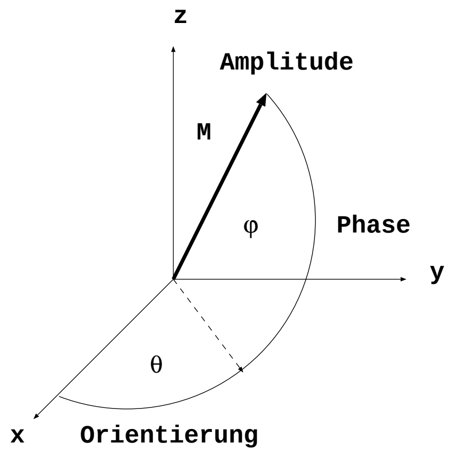


Abbildung 2.5: Komponenten des Strukturmultivektors : Darstellung in Kugelkoordinaten.

Die Phase gibt die strukturelle Information, d.h. die Kanten- bzw. Linienart an. Es können vier prototypische Fälle im Grauwertkontinuum auftreten, die in der Abbildung 2.6 dargestellt sind. In der ersten Reihe handelt es sich um eine weiße Linie, in der zweiten Reihe um eine Schwarz-weiß-Kante, in der dritten Reihe um eine schwarze Linie und in der vierten Reihe um eine Weiß-schwarz-Kante. Im nachfolgenden (φ, θ)-Diagramm erkennt man die Rotation der verschiedenen Linien- und Kantenarten, wobei die Orientierung θ auf der Abzisse und die Phase φ auf der Ordinate aufgetragen ist.

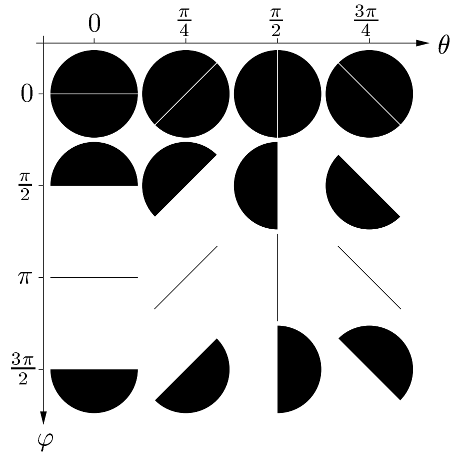


Abbildung 2.6: Kanten- und Linienarten in Abhängigkeit von der Orientierung und der Phase. Quelle : Felsberg und Sommer [FS01]

Bei der Orientierung und der Phase sind beim synthetischen Quader im Gegensatz zum Karton kaum Störungen in Form von Rauschen zu erkennen, wie es die Abbildung 2.7 und Abbildung 2.8 verdeutlichen. Besonders hervorzuheben ist bei den Bildern vom Karton, dass an den Stellen mit hoher lokaler Amplitude der Rauschanteil im Phasenbild und im Orientierungsbild sehr gering ist.

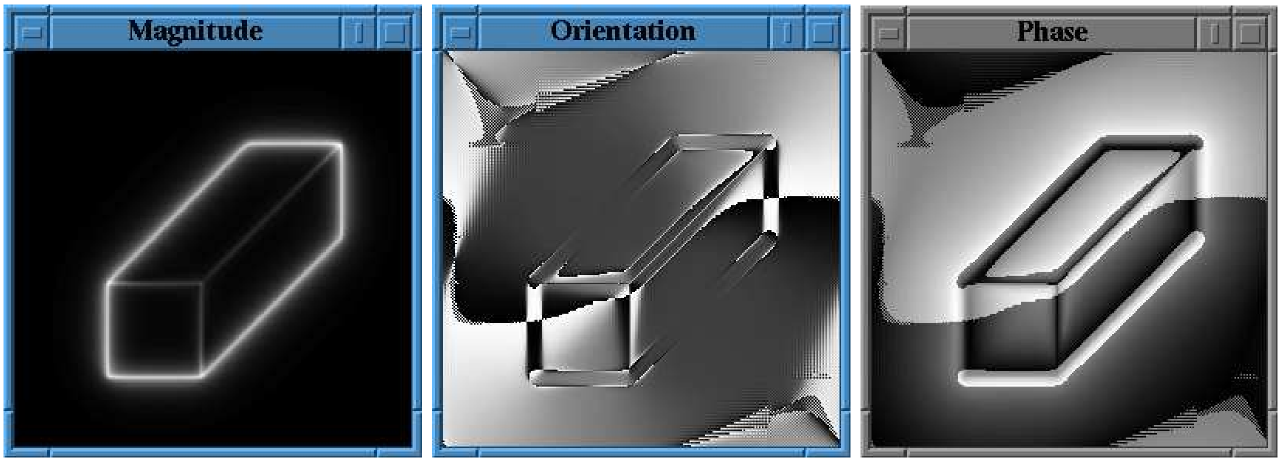


Abbildung 2.7: Links Amplitude, in der Mitte Orientierung und rechts Phase des Quaders bei Oktave $\Omega = 4$

Abbildung 2.8: Links Amplitude, in der Mitte Orientierung und rechts Phase des Kartons bei Oktave $\Omega = 5$

Zuletzt erfolgt eine Überprüfung des Testbildes mit dem Siemensstern in Abbildung 2.9. Die Ergebnisse stimmen mit Felsberg [FS00b] überein. Man kann erkennen, dass die Amplitude unabhängig vom Winkel ist. Die Orientierung nimmt mit dem Winkel zu und der Phasenverlauf ist bis auf den Ursprung für alle Orientierungen gleich.

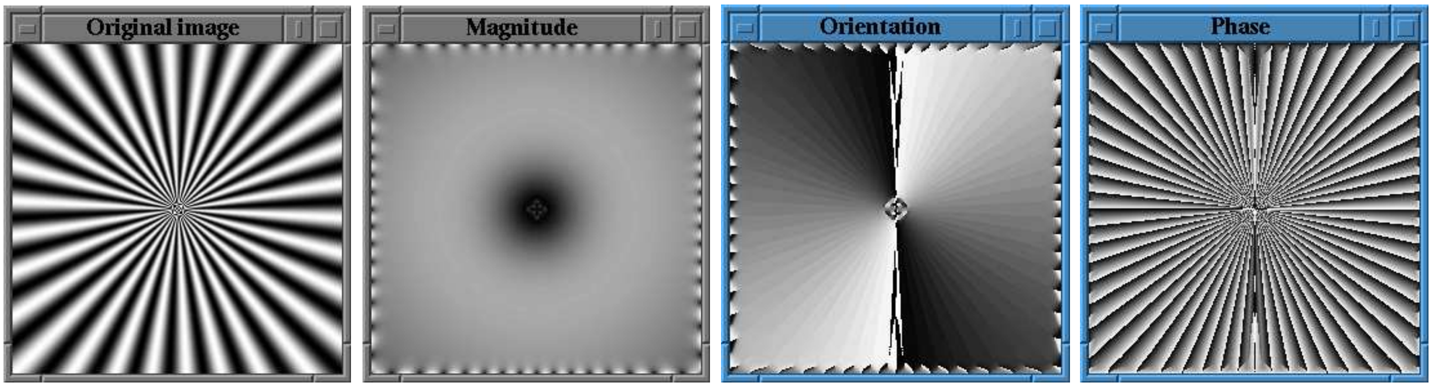


Abbildung 2.9: Siemensstern : Links Originalbild, in der Mitte links Amplitude, in der Mitte rechts Orientierung und rechts Phase bei Oktave $\Omega = 4$

2.2 Phasenkongruenz und Kantendetektion

Es wird in diesem Abschnitt die Phasenkongruenz ermittelt, die zur anschließenden Kantendetektion nach einem Algorithmus von Kovesi [Ko99] verwendet werden kann. Außerdem wird die Phasenkongruenz später zur Liniensegmentation benötigt. Es werden sphärische Quadraturfilter zur Berechnung der Phasenkongruenz angewandt, deren Bandpasscharakteristik drei Oktaven Bandbreite aufweist. Als Mittenfrequenzen wählt man die Nyquistfrequenz und ein Achtel der Nyquistfrequenz, d.h. einen Abstand von drei Oktaven. Für die Nyquistoktave, die dem oberen Band Π_{upper} entspricht, und für das untere Band Π_{lower} werden die Mittenfrequenzen ermittelt, wobei in der folgenden Gleichung der Parameter s die Bildgröße in Pixel darstellt :

$$\Pi_{upper} = \frac{\log(s)}{\log(2)} - 1 \tag{2.22}$$

Es gilt : $\Pi_{lower} = \Pi_{upper} - 3$ und es wird eine 3-Oktaven-Bandbreite mit dem Wert $b = 0.41$ in der Bandpassgleichung benutzt.

Die Rauschenergie N wird in der folgenden Gleichung durch den oberen Bandpass abgeschätzt, wobei nach Anwendung des Logarithmus der Mittelwert von allen Pixelwerten genommen wird. Dabei ist $SMV(\Pi_{upper})$ der Strukturmultivektor für das obere Band und $SMV(\Pi_{lower})$ repräsentiert den Strukturmultivektor für das untere Band. Die Rauschenergie N wird durch die Amplitude M des Strukturmultivektors für das obere Band bestimmt. Zur Notation ist zu bemerken, dass eine Mittelwertbildung des Argumentes der Exponentialfunktion durchgeführt wird, was durch den Überstrich symbolisiert wird.

$$N = \exp\left(\overline{\log\left(M(SMV(\Pi_{upper})) + 1\right)}\right)^2 \tag{2.23}$$

Die Energie E ergibt sich aus dem Produkt der Amplituden vom oberen und unteren Strukturmultivektor :

$$E = M(SMV(\Pi_{lower})) \cdot M(SMV(\Pi_{upper})) \tag{2.24}$$

Das entrauschte Skalarprodukt T der Strukturmultivektoren berechnet sich durch folgenden Zusammenhang, wobei der Wert null als zweites Argument der Maximumfunktion das Nullbild darstellt.

$$T = \max\left(SMV(\Pi_{lower}) \cdot SMV(\Pi_{upper}) - N, 0\right) \tag{2.25}$$

Zuletzt wird die Phasenkongruenz ϕ ermittelt. Zu beachten ist, dass im Nenner die Amplitude des Kreuzproduktes der Strukturmultivektoren eingeht. Für den Regularisierungsparameter ϵ, der eine Division durch null verhindert, gilt : $\epsilon = 0.01$.

$$\phi \approx \frac{T}{M(SMV(\Pi_{lower}) \times SMV(\Pi_{upper})) + E + \epsilon} \tag{2.26}$$

Danach wird zur Kantendetektion die Hysteresis-Threshold-Funktion von P. Kovesi [Ko99] angewandt, die die Phasenkongruenz ϕ, einen oberen Schwellwertparameter mit dem Wert 0.75 und einen unteren Schwellwertparameter mit dem Wert 0.35 benötigt. Als Ausgabe liefert die Funktion das Kantenbild. Diese Funktion markiert alle Pixel als Kantenpixel, wenn der Pixelwert der Phasenkongruenz über dem oberen Schwellwert liegt. Danach werden jeweils alle acht umgebenden Pixel der bereits markierten Pixel untersucht, um zu überprüfen, ob deren Wert über dem unteren Schwellwert liegt. Diejenigen Nachbarpixel, die dieses Kriterium erfüllen, werden dann ebenfalls markiert. Der Algorithmus ist rekursiv, so dass die nachträglich markierten Pixel ebenso behandelt werden wie die Pixel, deren Wert über dem oberen Schwellwert liegen. Somit werden Kanten sukzessive verlängert und Lücken gefüllt. Es ist offensichtlich, dass bei zu hohen Schwellwerten weniger Kanten detektiert werden und daher das Objekt nicht vollständig erkennbar ist, aber dafür ist der Rauschanteil im Bild relativ gering. Dagegen erhält man bei zu niedrigen Schwellwerten zwar alle Kanten, aber auch viel Rauschen. Zu bemerken ist, dass für die spätere Ermittlung der 2D Liniensegmentmittelpunkte eine Schwellwertbildung auf der Phasenkongruenz ϕ durchgeführt wird. Geeignete obere Schwellwerte für die Phasenkongruenz liegen zwischen 0.5 und 0.8. Bei kontrastreichen und rauscharmen Bildern ist ein hoher Schwellwert von 0.8 angemessen, bei kontrastarmen Bildern sollte man zu dem Wert 0.5 tendieren. In der Abbildung 2.10 sind beim Phasenkongruenzbild des Quaders Echokanten zu erkennen. Dieses Phänomen tritt nur bei synthetischen Bildern aufgrund der Endlichkeit des Bildträgers und den damit verbundenen periodischen Wiederholungen der Phasenkongruenz auf.

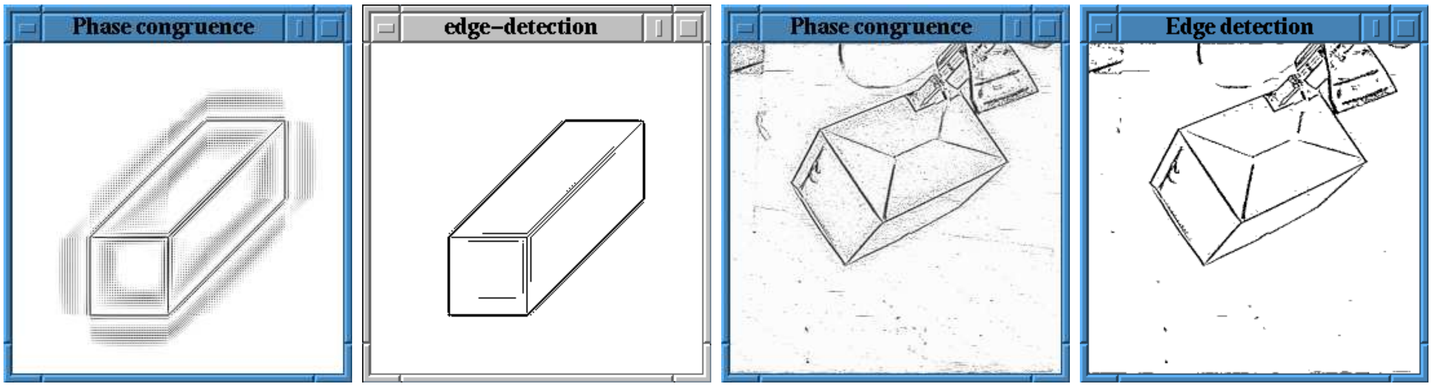


Abbildung 2.10: Von links nach rechts : Phasenkongruenz und Kantendetektion des synthetischen Quaders. Danach Phasenkongruenz und Kantendetektion des Kartons.

2.3 Verbesserte Orientierung

Die Orientierungen für das obere und untere Band sind instabil, wenn die Phase Werte nahe null oder $\pm\pi$ annimmt. Eine wesentliche Verbesserung läßt sich durch die folgende phasenadaptive Interpolation erreichen. Man erzeugt ein komplexes Bild $Img_\theta = (f_j, f_i)$, dessen Realteil die j-Komponente und dessen Imaginärteil die i-Komponente des Strukturmultivektors sind. Dann wird dieses komplexe Bild durch die Amplitude M des Strukturmultivektors dividiert.

$$Img_{\theta,M} := \frac{Img_\theta}{M} \tag{2.27}$$

Anschließend wird das Quadrat von $Img_{\theta,M}$ mit einem 3×3 Boxfilter F_1 bestehend aus Einsen gefaltet.

$$\widehat{Img_\theta} := (Img_{\theta,M})^2 * F_1 \tag{2.28}$$

Die bisherige Orientierungsabschätzung lautete :

$$\theta = \frac{\arg(f_i + if_j)^2}{2} \tag{2.29}$$

Jetzt aber wird die Orientierung durch $\widehat{Img_\theta}$ berechnet, wobei $\widehat{Img_{\theta,real}}$ dessen Realteil und $\widehat{Img_{\theta,imag}}$ dessen Imaginärteil sind.

$$\theta = \frac{\left(\pi - \arctan\frac{\widehat{Img_{\theta,imag}}}{\widehat{Img_{\theta,real}}}\right)}{2} \tag{2.30}$$

Wie man in der Abbildung 2.11 erkennt, ist eine Filterung mit einem 3×3 Filter durchaus angemessen. Bei zu großen Filtermasken werden die Orientierungswerte zu sehr gemittelt, was einen Informationsverlust bedeutet, wenn sich die Orientierung im lokalen Umfeld stark ändert. Verwendet wird die 3×3 Filterung der Orientierung für das untere Band.

Abbildung 2.11: Links Orientierung des unteren Bandes (Oktave $\Omega = 5$) ohne Filterung bei einem Karton, in der Mitte Orientierung nach Anwendung eines 3×3 und rechts eines 9×9 Filters.

2.4 Oktavenabhängigkeit des Bandpassfilters

Bei der Implementierung des Bandpassfilters, des Strukturmultivektors und der Phasenkongruenz muss man auf eine geeignete Wahl der Parameter achten, damit die gewünschten Ergebnisse wie z.B. die Liniensegmentation qualitativ zufriedenstellend sind. Von entscheidender Bedeutung ist vor allem die Wahl der Oktave Ω, die beim Bandpassfilter auftritt. Man erkennt in der Abbildung 2.12, dass bei einem hohen Oktavenwert die Strukturen des Kartons deutlich besser zu erkennen sind.

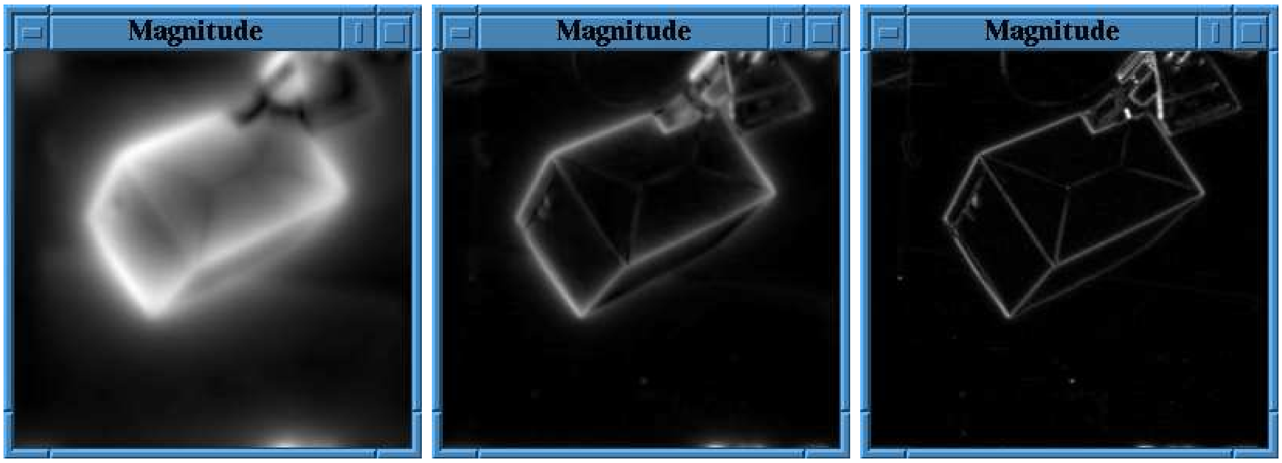


Abbildung 2.12: Amplitude eines Kartons bei der Oktave 2 links, Oktave 4 in der Mitte und Oktave 6 rechts.

Im Phasenbild sieht man dagegen bei einem hohen Oktavenwert deutlich mehr Störungen. Das hat zur Folge, dass sich die Phase innerhalb weniger Pixel erheblich verändert, was zu Lokalisationsproblemen führen kann, wenn man Werte aus dem Phasenbild entnehmen möchte. Dieses Problem wird in Kapitel 5 von Bedeutung sein, wo man die Phasenkonsistenz innerhalb einer Bildsequenz überprüft.

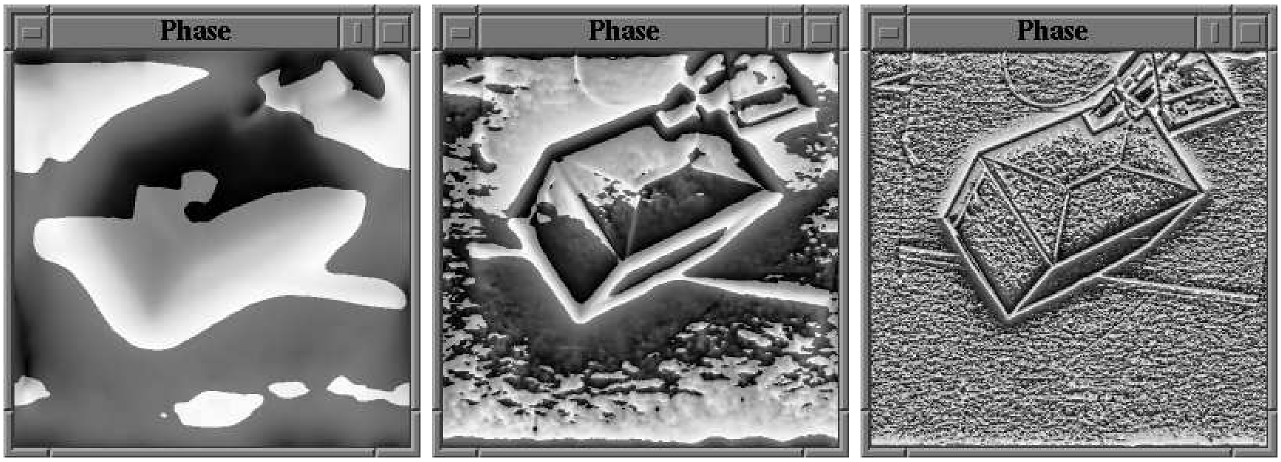


Abbildung 2.13: Phase eines Kartons bei der Oktave 2 links, Oktave 4 in der Mitte und Oktave 6 rechts.

Die gleichen Schwierigkeiten hat man beim Orientierungsbild. Auch hier verändert sich bei einem hohen Oktavenwert die Orientierung innerhalb weniger Pixel schnell.

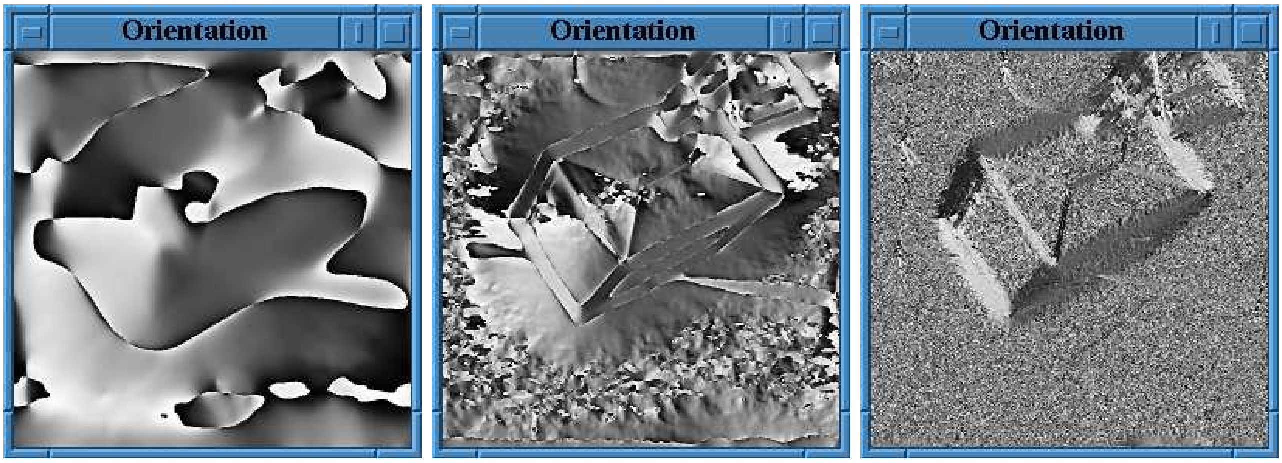


Abbildung 2.14: Orientierung eines Kartons bei der Oktave 2 links, Oktave 4 in der Mitte und Oktave 6 rechts.

2.5 Ergebnisse

Die Phasenkongruenz liefert eine Konfidenz bzw. Identität mit Werten im reellen Intervall $[0;1]$. Sie wird mit einem Schwellwert behaftet, um die Objektkanten zu extrahieren, die im Vergleich zum Rauschen in der Regel eine hohe Konfidenz besitzen. Anschließend wird ein $n \times n$ Gitter über die schwellwertbehaftete Phasenkongruenz gelegt. In jeder Gitterzelle wird dann das Maximum der Phasenkongruenz ermittelt, um die 2D Liniensegmentmittelpunkte zu erhalten. Das Maximum in einer Gitterzelle liegt somit jeweils im reellen Intervall zwischen dem vorgegebenen Schwellwert und dem Wert eins. Wenn das Gitter sehr engmaschig gewählt wird, steigt der Rechenaufwand, wobei sich die Ergebnisse in der Regel gleichzeitig verbessern. Gute Ergebnisse erhält man bei einem 5×5 Gitter, welches im Folgenden verwendet wird.

Die 2D Liniensegmentmittelpunkte und die zugehörigen Orientierungen θ, die sich im Intervall von 0 bis π befinden, sind somit bekannt, um 2D Liniensegmente zu erzeugen. Für Bilder mit relativ vielen geradlinigen Objektbegrenzungen ist ein 10×10 Gitter auf die schwellwertbehaftete Phasenkongruenz ausreichend, um die 2D Liniensegmentmittelpunkte zu ermitteln. Bei detailreichen Bildern mit vielen krummlinigen Konturen bietet sich auf jeden Fall ein 5×5 Gitter an, um lokale Strukturen besser zu erkennen.

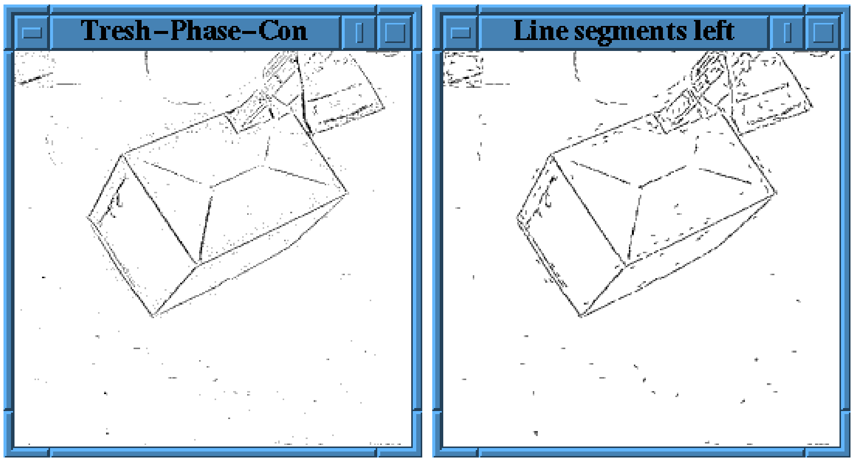


Abbildung 2.15: Links Phasenkongruenz des Kartons nach Schwellwertbildung von 0.7 und rechts Liniensegmente bei einem 5 x 5 Gitter.

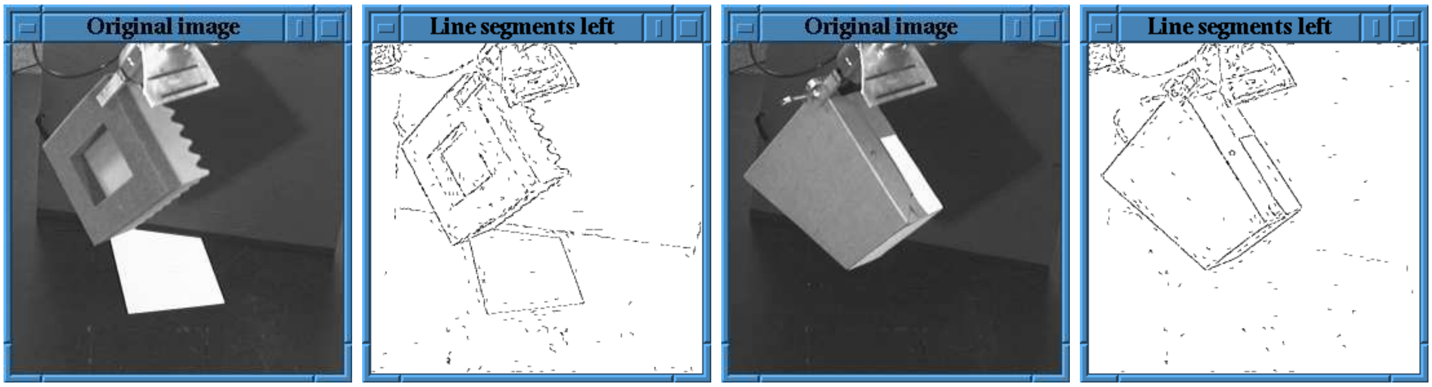


Abbildung 2.16: Von links nach rechts : Eines der beiden Bilder der Stereokamera bei der Aufnahme eines Schaumstoffteils und Liniensegmente bei einem 5 x 5 Gitter nach Schwellwertbildung von 0.7 auf der Phasenkongruenz. Analog für einen Ordner.

Exemplarisch wird ein Bild, das nicht unter Laborbedingungen aufgenommen wurde, in den Abbildungen 2.17 und 2.18 verarbeitet. Je kleiner die Strukturen im Bild sind, desto schlechter ist eine Approximation durch 2D Liniensegmente möglich. Entscheidend für die Qualität der Phasenkongruenz und der daraus resultierenden Liniensegmentation ist der lokale Signal-Rauschabstand im Eingabebild. Man erhält jedoch auch 2D Liniensegmente an Stellen, wo das Rauschen im Originalbild eine Mindestschwelle überschritten hat.

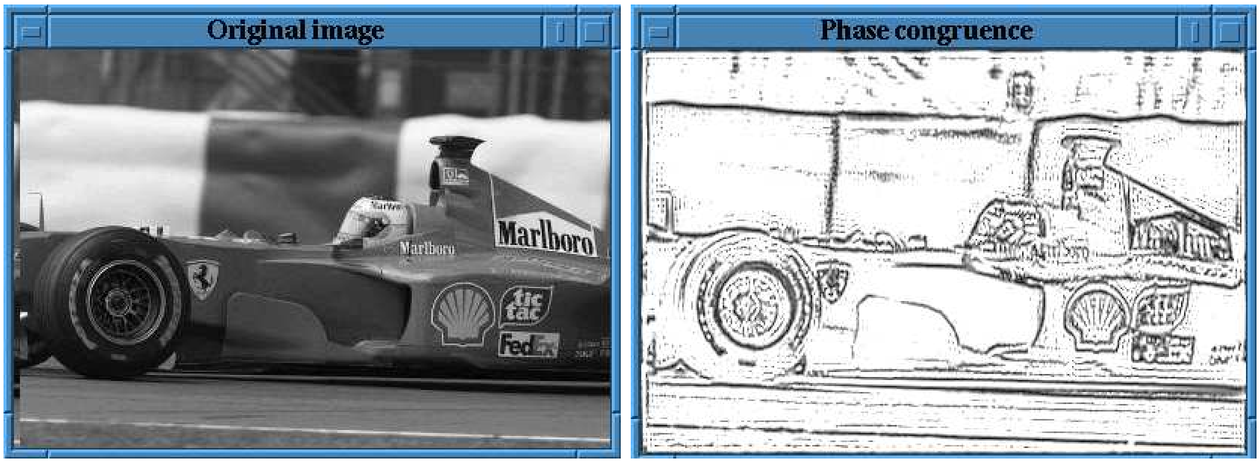


Abbildung 2.17: Links Originalbild und rechts Phasenkongruenz

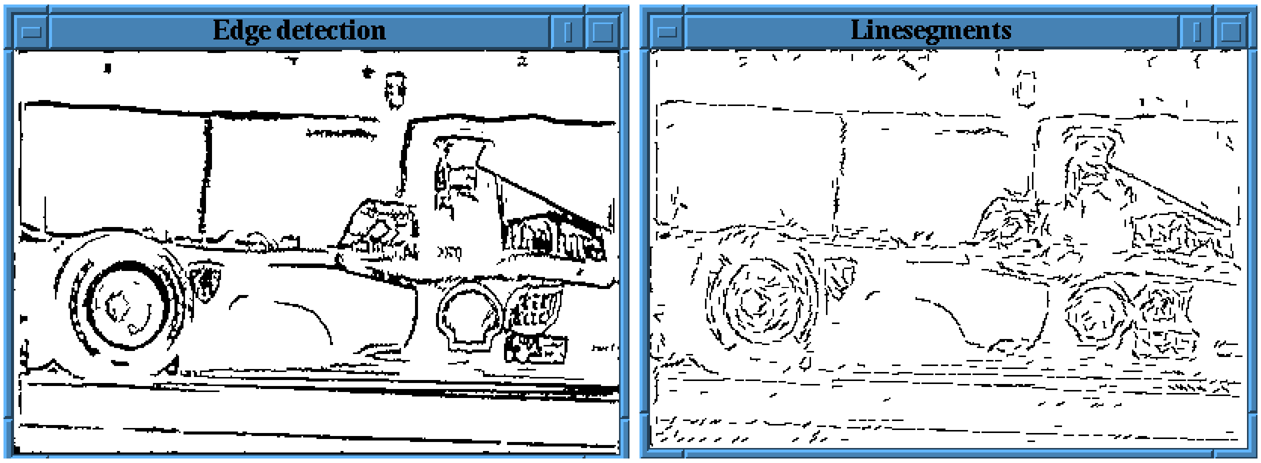


Abbildung 2.18: Links Kantendetektion und rechts 2D Liniensegmente bei einem 5 x 5 Gitter nach einer Schwellwertbildung von 0.6 auf der Phasenkongruenz.

Insgesamt gesehen sind die erzielten Ergebnisse zufriedenstellend. Die Bildvorverarbeitung mit dem Strukturmultivektor und die anschließende Erzeugung einer schwellwertbehafteten Phasenkongruenz stellen eine geeignete Methode dar, um Objektkanten im Eingabebild in 2D Liniensegmente zu zerlegen. Die Größe des gewählten Gitters, das auf die Phasenkongruenz gelegt wird, die Wahl des Schwellwertes und die lokalen Signal-Rauschverhältnisse im Eingabebild haben erheblichen Einfluss auf die Qualität der Liniensegmentation.

Kapitel 3

Ähnlichkeitsfunktion für das Stereomatching

3.1 Motivation

In diesem Kapitel wird das Korrespondenzproblem betrachtet. Man möchte zu einem vorgegebenen 2D Liniensegmentmittelpunkt im linken Stereokamerabild den korrespondierenden 2D Liniensegmentmittelpunkt im rechten Bild bestimmen. Für die Lösung des Korrespondenzproblems wird ein Ansatz basierend auf der Epipolargeometrie gewählt, daher liegt der korrespondierende Punkt auf der Epipolargeraden im rechten Bild. Der korrespondierende Punkt wird durch eine in dieser Arbeit neu entwickelte Distanz- bzw. Ähnlichkeitsfunktion ermittelt. Die Distanzfunktion berechnet die Phasen- und Orientierungsdifferenz zwischen einem 2D Punkt im linken und einem 2D Punkt rechten Bild. Das Ergebnis der Distanzfunktion geht in die Ähnlichkeitsfunktion ein, mit deren Hilfe das korrespondierende 2D Punktepaar ermittelt wird, bei dem die Konfidenz des zugehörigen 3D Punktes maximal wird. Die Konfidenz eines 3D Punktes, die im Intervall $[0;1]$ liegt, entspricht wie im 2D Fall der Annahme des Systems, dass das zugehörige 3D Liniensegment auf einer Objektkante bzw. Objektlinie liegt.

Die Abbildung 3.1 veranschaulicht die Abhängigkeit der Orientierung und Phase bzgl. der verschiedenen Kanten- und Linienarten. Zum besseren Verständnis der Positionsbeschreibungen auf der abgebildeten Kugel wird die Terminologie aus der Kartographie verwendet. Betrachtet wird eine Kugeloberfläche, auf deren Längengraden sich die Phasenwerte ändern, d.h. am Nordpol hat die Phase φ den Wert null, am Äquator hat sie je nach Durchlaufrichtung den Wert $\pm\frac{\pi}{2}$ und am Südpol den Wert $\varphi = \pm\pi$. Die durchgezogenen Pfeile und schwarz markierten Punkte liegen auf der dem Betrachter zugewandten Seite. Die gestrichelten Pfeile und die weiß markierten Punkte befinden sich auf der Rückseite. Auf den Breitengraden wird das Orientierungsintervall $[0;\pi]$ zweimal durchlaufen. Für die verschiedenen Kanten- und Linienarten hat dieser Sachverhalt folgende Auswirkungen. An den Polen laufen alle Längengrade zusammen, daher existieren in Polnähe alle Orientierungen. In Nordpolnähe können alle Orientierungen einer weißen Linie auf dunklem Hintergrund und in Südpolnähe können alle Orientierungen einer schwarzen Linie auf hellem Hintergrund auftreten. Direkt an den Polen ist keine Orientierung definiert. Entlang des Äquators ist der Übergang von einer Schwarz-weiß-Kante zu einer Weiß-schwarz-Kante und umgekehrt durch eine jeweilige Drehung um π dargestellt. Auf der Nordhalbkugel befindet sich im Grauwertkontinuum der Übergang von einer Kante zu einer weißen Linie und auf der Südhalbkugel von einer Kante zu einer schwarzen Linie.

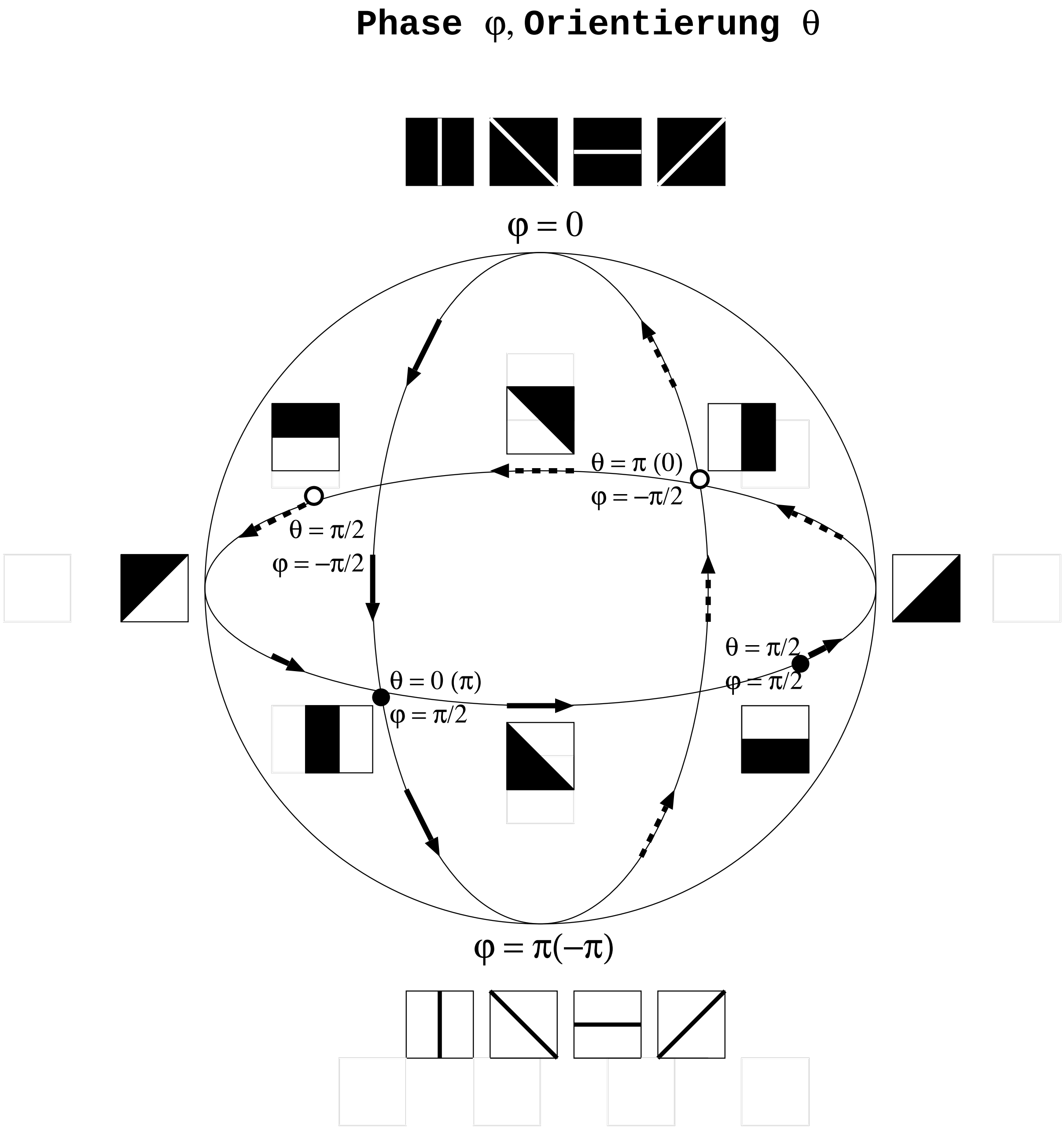


Abbildung 3.1: Kanten- und Linienarten bei verschiedener Orientierung und Phase

Die beiden Intervalle der Orientierung $\theta \in [0; \pi]$ und das Phasenintervall $\varphi \in [-\pi; \pi]$ auf der Kugeloberfläche haben eine Mehrdeutigkeit zur Folge, die auftritt, wenn man eine lokale Orientierung in eine lokale Richtung überführt. Ein Punkt entspricht entweder dem Tupel (θ, φ) oder $(\theta + \pi, -\varphi)$. Wenn sich die Orientierung um π ändert, wechselt das Vorzeichen der Phase. Umgekehrt gilt auch, dass sich bei einer Vorzeichenänderung der Phase die Orientierung um π ändert. Dieser Zusammenhang entsteht einerseits dadurch, dass auf einer Kugelseite die Phasenwerte positiv und auf der anderen negativ sind. Andererseits hat die Orientierung zwei Längengrade als Diskontinuitäten, wo sich der Orientierungswert $\theta = 0$ zu $\theta = \pi$ ändert und umgekehrt. Die erwähnte Mehrdeutigkeit würde nicht auftreten, wenn es nur ein Orientierungsintervall $[0; \pi]$ gäbe.

Wenn man zu einem 2D Punkt im linken Kamerabild, der die Orientierung θ_{left} und die Phase φ_{left} hat, den korrespondierenden 2D Punkt im rechten Bild mit dem Tupel $(\theta_{right}, \varphi_{right})$ bestimmen will, dann läßt sich das Problem mit Hilfe der Abbildung 3.1 folgendermaßen lösen. Der 2D Punkt im linken Kamerabild und ein zu untersuchender 2D Punkt auf der zugehörigen Epipolargeraden im rechten Kamerabild entsprechen zwei Punkte auf der Kugeloberfläche. Der kürzeste Abstand der beiden Punkte läßt sich durch eine Distanzfunktion d berechnen. Für diese Distanzfunktion wird eine tiefenabhängige Funktion $\alpha(z)$ benötigt, die die Gewichtung zwischen der Phasen- und Orientierungsabweichung bestimmt. Das Ergebnis der Distanzfunktion geht in die Ähnlichkeitsfunktion ein, mit der das korrespondierende 2D Punktepaar und die Konfidenz des zu erzeugenden 3D Liniensegmentes ermittelt wird. Die **Distanzfunktion** d ist als Minimum der beiden Funktionswerte d_1 und d_2 definiert, d.h. es gilt : $d := \min(d_1, d_2)$ mit :

$$d_1 = \sqrt{\alpha(z)(\operatorname{atan2}(\sin(\varphi_{left} - \varphi_{right}), \cos(\varphi_{left} - \varphi_{right})))^2 + (1 - \alpha(z))(\theta_{left} - \theta_{right})^2} \tag{3.1}$$

$$d_2 = \sqrt{\alpha(z)(\operatorname{atan2}(\sin(\varphi_{left} + \varphi_{right}), \cos(\varphi_{left} + \varphi_{right})))^2 + (1 - \alpha(z))(\pi - |\theta_{left} - \theta_{right}|)^2} \tag{3.2}$$

Dabei ist $\alpha(z)$ für eine bestimmte Tiefe z ein dimensionsloser Funktionswert im Intervall $[0; 1]$. Die Minimumbildung ist erforderlich, weil auf der Kugeloberfläche stets der kürzeste Abstand bzgl. der modifizierten Mahalanobisdistanz zwischen zwei Punkten zu nehmen ist. Die Funktion atan2 ist definiert durch die folgende Gleichung. Sie liefert Funktionswerte im Intervall $[-\pi; \pi]$, wie man in Abbildung 3.2 erkennen kann. Die Funktion $\arctan \frac{|y|}{x}$ liefert dagegen Funktionswerte im Intervall $[0; \pi[$. Der Wert der Signumfunktion ist für die Rotationsrichtung in der 2D Ebene entscheidend.

$$\operatorname{atan2}(y, x) := \operatorname{sign}(y) \cdot \arctan\left(\frac{|y|}{x}\right) \tag{3.3}$$

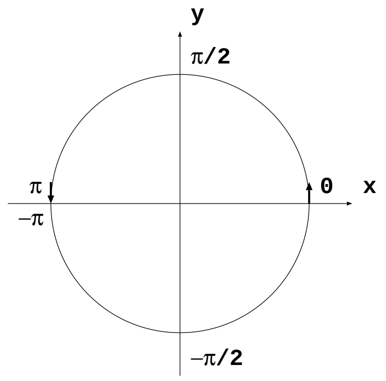


Abbildung 3.2: Funktion atan2(y, x)

In Abbildung 3.3 ist ein Beispiel dargestellt, um die Minimumbildung aus d_1 und d_2 zu verdeutlichen. Es gibt auf einem Großkreis auf der Kugeloberfläche zwei Wege, um vom Punkt A zum Punkt B zu gelangen. Zu erkennen sind außerdem die Orientierungs- und Phasenwerte für die Punkte A und B, die auf die beiden Großkreise in Abbildung 3.1 projiziert wurden, um deren Position und somit deren Differenz besser zu erkennen. Wie in Abbildung 3.1 sind die Punkte, die auf der Vorderseite der Kugel liegen, schwarz markiert. Analog sind die Punkte auf der Kugelrückseite weiß gekennzeichnet.

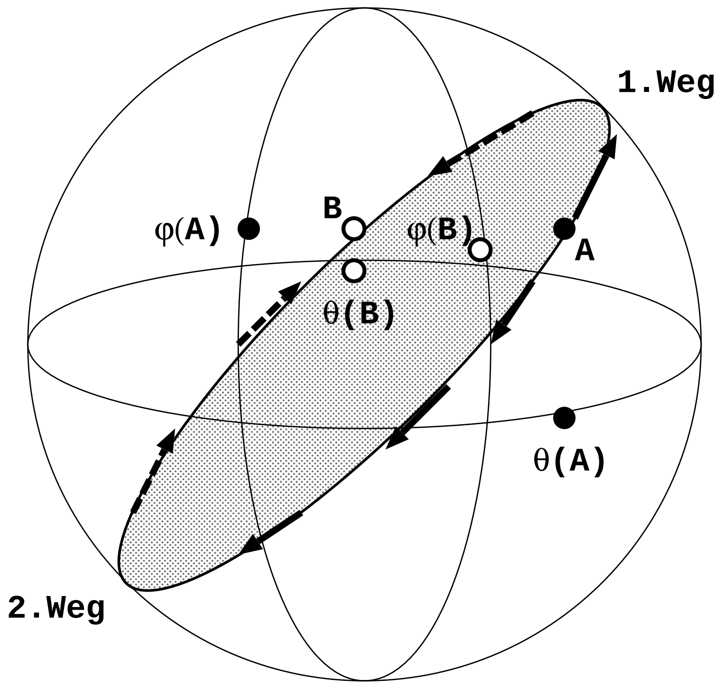


Abbildung 3.3: Beispiel für die kürzeste Distanz zweier Punkte auf der Kugeloberfläche

Im Folgenden werden die möglichen Fälle erläutert, wie zwei Punkte A und B auf einer Kugel platziert sein können. In der Abbildung 3.4 kann man in der oberen Reihe den Fall erkennen, dass beide Punkte im selben Orientierungsintervall und zugleich im selben

Phasenintervall liegen. Die untere Reihe beinhaltet den anderen Fall, nämlich dass sich die beiden Punkte in verschiedenen Orientierungsintervallen befinden, daher müssen die Phasen auch in verschiedenen Intervallen liegen.

Aus der bereits behandelten Abbildung 3.1 geht hervor, dass zwei Punkte, die im selben Orientierungsintervall liegen, dasselbe Phasenvorzeichen haben müssen. Daher entfällt die Betrachtung der beiden übrigen Fälle, nämlich dass sich die beiden Punkte im selben Orientierungsintervall, aber in verschiedenen Phasenintervallen befinden bzw. dass sie sich im selben Phasenintervall, aber in verschiedenen Orientierungsintervallen befinden.

In Abbildung 3.4 erkennt man, wie die Terme in der Distanzfunktion entstehen. Die obere Reihe repräsentiert die Gleichung 3.1, d.h. es wird die direkte Differenz der Orientierungen $\Delta\theta = \theta_{left} - \theta_{right}$ gebildet, weil sie im selben Intervall liegen. Die untere Reihe veranschaulicht die Gleichung 3.2, bei der die Orientierungen in verschiedenen Intervallen liegen, daher wird $\pi - |\theta_{left} - \theta_{right}|$ gebildet, um den kleineren Winkel zwischen den Punkten A und B zu ermitteln. Im Phasenterm :

$$\operatorname{sign}(\sin(\varphi_{left} \pm \varphi_{right})) \cdot \arctan \frac{|\sin(\varphi_{left} \pm \varphi_{right})|}{\cos(\varphi_{left} \pm \varphi_{right})} \tag{3.4}$$

werden die Subtraktionen beim Fall d_1 und die Additionen im Fall d_2 durchgeführt.

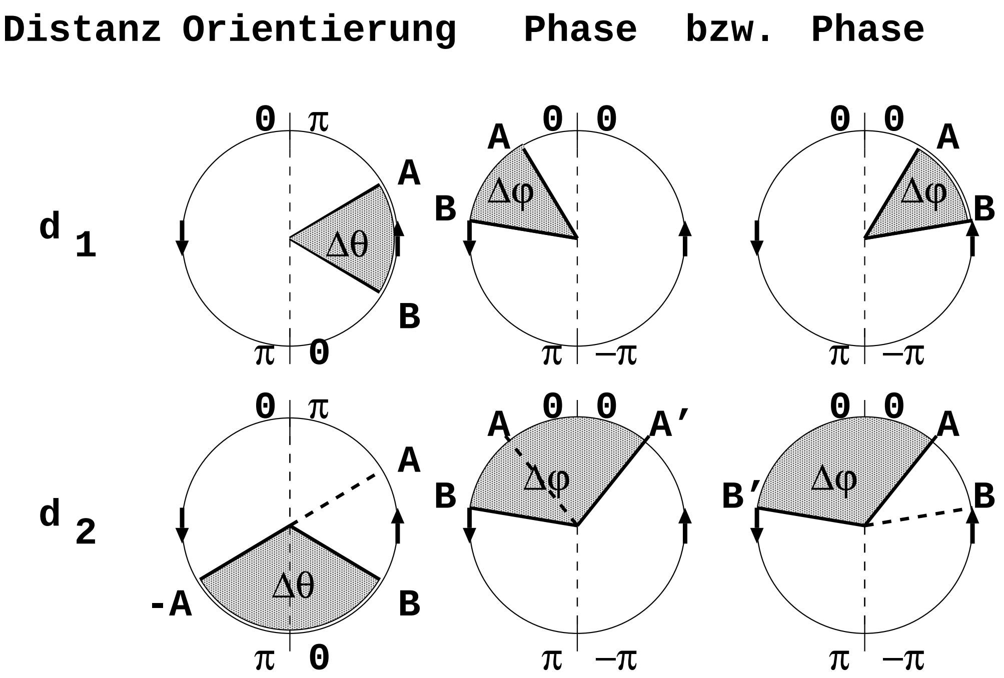


Abbildung 3.4: Erläuterung der Distanzfunktion

Bei dem Korrespondenzproblem wird in der hier behandelten Ähnlichkeitsfunktion zu einem 2D Punkt im linken Bild derjenige 2D Punkt auf der Epipolargeraden im rechten Bild gesucht, bei dem die Konfidenz des zugehörigen 3D Punktes maximal wird. Wenn man ein korrespondierendes 2D Punktepaar gefunden hat, kann man einen 3D Punkt rekonstruieren, dessen Konfidenz sich durch eine **Ähnlichkeitsfunktion** ergibt, die folgendermaßen definiert wird :

$$conf_{3D} := \left(\frac{\pi - d}{\pi}\right) \cdot conf_{2D,left} \cdot conf_{2D,right} \tag{3.5}$$

Man erkennt, dass die Konfidenz $conf_{3D}$ des 3D Punktes maximal wird, wenn die Distanzfunktion d einen sehr geringen Wert annimmt und die Konfidenzen der 2D Punkte $conf_{2D,left}$ und $conf_{2D,right}$ im linken bzw. rechten Bild hoch sind. Die Konfidenzen der 2D Punkte, die im Intervall $[0;1]$ liegen, ergeben sich aus den Phasenkongruenzen des linken bzw. rechten Bildes. Von entscheidender Bedeutung in der Distanzfunktion ist die tiefenabhängige Funktion $\alpha(z)$.

Vorab werden drei Spezialfälle der Funktion $\alpha(z)$ und deren Wirkung auf die Distanzfunktion kurz erläutert. Der erste Fall ist $\alpha(z) = 0$, d.h. die Phasendifferenz ist bedeutungslos und nur die Orientierungsdifferenz geht in die Distanzfunktion ein.

$$\alpha(z) = 0 \qquad \Longrightarrow \qquad \begin{cases} d_1 = \theta_{left} - \theta_{right} \\ d_2 = \pi - |\theta_{left} - \theta_{right}| \end{cases}$$

Der zweite Fall beinhaltet, dass $\alpha(z) = 1$ gilt, d.h. die Orientierungsdifferenz wird vernachlässigt und nur die Phasendifferenz wird bei der Korrespondenzfindung betrachtet.

$$\alpha(z) = 1 \qquad \Longrightarrow \qquad \begin{cases} d_1 = \mathrm{atan2}(\sin(\varphi_{left} - \varphi_{right}), \cos(\varphi_{left} - \varphi_{right})) \\ d_2 = \mathrm{atan2}(\sin(\varphi_{left} + \varphi_{right}), \cos(\varphi_{left} + \varphi_{right})) \end{cases}$$

Der dritte Fall tritt ein, wenn $\alpha(z) = \frac{1}{2}$ gilt. Dadurch sind die Orientierungs- und die Phasendifferenz in der Distanzfunktion gleichberechtigt. Dies entspricht der Wahl des kürzeren Bogens in Abbildung 3.3.

In den nächsten Abschnitten dieses Kapitels wird zuerst die Orientierungsabweichung zwischen den beiden 2D Projektionen bei vorgegebenen 3D Liniensegmenten ermittelt. Dabei werden vereinfachte Projektionsmatrizen zugrundegelegt. Anschließend wird die Phasenabweichung bei 2D Schwarz-weiß-Kanten auf einem quadratischen Pixelgitter ermittelt. Aus den berechneten Standardabweichungen bzw. Varianzen der Orientierungs- und Phasendifferenzen wird eine Gleichung für die tiefenabhängige Funktion $\alpha(z)$ mit Hilfe der Mahalanobisdistanz hergeleitet. Die Grundlagen der projektiven Geometrie und der Kamerakalibration kann man bei Faugeras [Fa93] nachlesen.

3.2 Orientierungsabweichung zwischen 2D Projektionen

3.2.1 Projektionsmatrix

Das Ziel ist es, zu einem vorgegebenen 3D Liniensegment die Differenz der Orientierungen in den beiden Projektionsebenen zu bestimmen. Um die mittlere Orientierungsabweichung zu ermitteln, werden durch einen Zufallsgenerator viele 3D Liniensegmente in einem Quader mit $\Delta z \ll \Delta x$ und $\Delta z \ll \Delta y$ erzeugt und die jeweiligen Orientierungsdifferenzen in einem Histogramm dargestellt. Abbildung 3.5 veranschaulicht den Versuchsaufbau. Der 3D Segmentanfangspunkt und der 3D Segmentendpunkt werden auf die Projektionsebenen abgebildet und es wird die Orientierungsdifferenz der projizierten Segmente berechnet.

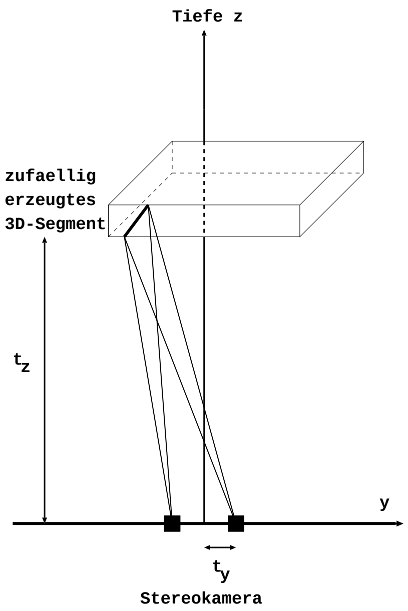


Abbildung 3.5: Versuchsaufbau zur Bestimmung der Orientierungsabweichung

Die Abbildung eines 3D Punktes auf eine 2D Ebene geschieht durch eine Projektionsmatrix, die intrinsische Parameter für den Kameraaufbau und extrinsische Parameter für die Orientierung und Position der Kamera zum Objekt enthält. Zu den intrinsischen Parametern zählen k_u, k_v, u_0 und v_0. Der Vektor $(u_0, v_0)^T$ gibt die Koordinaten an, bei denen die optische Achse die Bildebene schneidet. Die Parameter k_u^{-1} und k_v^{-1} stellen die horizontale und vertikale Größe der Pixel in Meter dar. In der folgenden Gleichung ist $(U, V, H)^T$ der Projektionsvektor und f die Brennweite der Kamera. Die Größe H bezieht sich auf den Koordinatenursprung in der Projektionsebene. Es ergibt sich somit die Projektionsgleichung ohne Berücksichtigung von Rotation und Translation :

$$\begin{pmatrix} U \\ V \\ H \end{pmatrix} = \begin{pmatrix} -fk_u & 0 & u_0 & 0 \\ 0 & -fk_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \tag{3.6}$$

Damit erhält man die beiden folgenden Abbildungsgleichungen :

$$u = \frac{U}{H} = -fk_u \frac{x}{z} + u_0 \qquad v = \frac{V}{H} = -fk_v \frac{y}{z} + v_0 \tag{3.7}$$

Der zu projizierende Vektor ist $(x, y, z)^T$ und $(u, v)^T$ seine Projektion. Die Brennweiten in Pixeln sind gegeben durch den Parameter ξ, wobei $\xi_u = -fk_u$ und $\xi_v = -fk_v$ gelten. Mit einer 3×3 - Rotationsmatrix R und einem Translationsvektor $\vec{t} = (t_x, t_y, t_z)^T$, der die Verschiebung vom Ursprung des Weltkoordinatensystems zur Linsenmitte bewirkt, läßt sich durch die folgende Matrizenmultiplikation die 3×4 - Projektionsmatrix P aufstellen, die im nächsten Abschnitt benötigt wird :

$$P = \begin{pmatrix} \xi_u & 0 & u_0 & 0 \\ 0 & \xi_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} R & \vec{t} \\ \vec{O}^T & 1 \end{pmatrix} \tag{3.8}$$

Dabei ist der Vektor $\vec{O}$ der dreidimensionale Nullvektor.

3.2.2 Berechnungen

Zur möglichst einfachen Bestimmung der Orientierungsdifferenzen der Projektionen werden für die Projektionsmatrix einige Annahmen gemacht. Zuerst ersetzt man die Rotationsmatrix durch die Identitätsabbildung, d.h. es wird keine Rotation durchgeführt.

$$P = \begin{pmatrix} \xi_u & 0 & u_0 & 0 \\ 0 & \xi_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} R & \vec{t} \\ \vec{O}^T & 1 \end{pmatrix} = \begin{pmatrix} \xi_u & 0 & u_0 & 0 \\ 0 & \xi_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.9}$$

Außerdem werden die Parameter folgendermaßen gesetzt : $\xi_u = \xi_v = 1$ und $u_0 = v_0 = 0$. Die Matrizenmultiplikation ergibt somit :

$$P = \begin{pmatrix} \xi_u & 0 & u_0 & \xi_u t_x + u_0 t_z \\ 0 & \xi_v & v_0 & \xi_v t_y + v_0 t_z \\ 0 & 0 & 1 & t_z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \end{pmatrix} \tag{3.10}$$

Eine weitere Vereinfachung betrifft den Translationsvektor $\vec{t}$. Für die linke Projektionsmatrix wird $t = (0, -|t_y|, t_z)^T$ gewählt und für die rechte Projektionsmatrix setzt man $t = (0, |t_y|, t_z)^T$. Die Komponente t_z stellt die Tiefe dar, also den Abstand der Projektionsebenen zum Quader, in dem die 3D-Segmente erzeugt werden. Die Komponente t_y entspricht dem halben Abstand der Kameras. Damit erhält man die beiden folgenden Projektionsmatrizen, die sich nur in dem Vorzeichen der t_y-Komponente unterscheiden :

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \pm|t_y| \\ 0 & 0 & 1 & t_z \end{pmatrix} \tag{3.11}$$

Somit nimmt das Gleichungssystem für die Projektion diese Gestalt an :

$$\begin{pmatrix} U \\ V \\ H \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \pm|t_y| \\ 0 & 0 & 1 & t_z \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \tag{3.12}$$

Die folgende Gleichung wird für beide Projektionsebenen sowohl für den Segmentanfangspunkt als auch für den Segmentendpunkt erstellt, daher wird sie viermal angewandt.

$$u = \frac{x}{z + t_z} \tag{3.13}$$

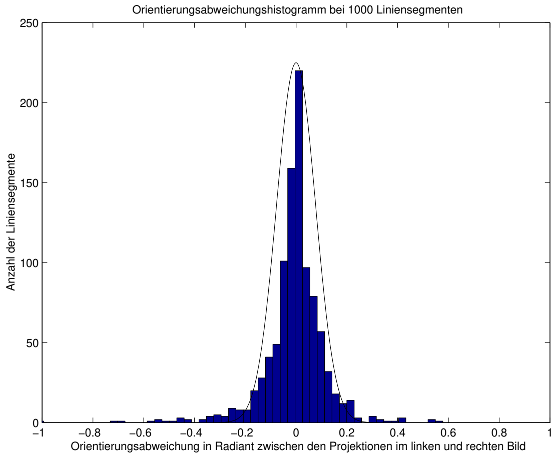
Die nächste Gleichung wird ebenso viermal benötigt, wobei für das Segment in der rechten Projektionsebene $+|t_y|$ und in der linken Projektionsebene $-|t_y|$ zu wählen ist.

$$v = \frac{y \pm |t_y|}{z + t_z} \tag{3.14}$$

Die Orientierung in beiden Projektionsebenen ergibt sich durch die folgende Gleichung, die sowohl für die linke als auch für die rechte Projektionsebene benutzt wird. Die Indizes *start* und *end* bedeuten dabei Start- und Endpunkt des 2D-Segmentes. Die Orientierungsabweichung der Projektionen berechnet sich anschließend durch deren Differenz : $\theta_{right} - \theta_{left}$

$$\theta = \operatorname{atan2}\left(\frac{v_{start} - v_{end}}{u_{start} - u_{end}}\right) \tag{3.15}$$

Als Ergebnis der Orientierungsabweichung zwischen den Projektionen bei eintausend zufällig erzeugten 3D Segmenten in einem Quader erhält man in Abbildung 3.6 Histogramme, die durch eine Gaußkurve approximiert werden. Man erkennt, dass sich bei betragsmäßig zunehmender y-Komponente des Translationsvektors und damit bei zunehmendem Abstand zwischen den Kameras die Streuung der Orientierungsabweichung verstärkt.



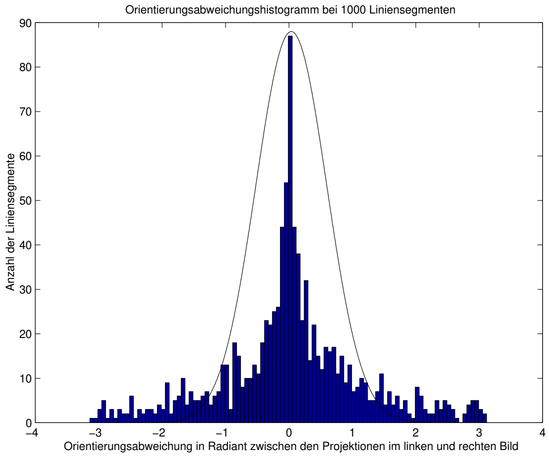


Abbildung 3.6: Links : Histogramm der Orientierungsabweichungen in Radiant zwischen 2D Projektionen mit passender Gaussapproximation bei eintausend 3D Segmenten. Der Betrag der y-Komponente des Translationsvektors ist 10 und seine z-Komponente ist 150. Rechts : Analog für y-Komponente : 100 und z-Komponente : 150.

Für eintausend zufällig erzeugte 3D Liniensegmente ergeben sich somit folgende Tabellen der Standardabweichungen, wenn man die y-Komponente und die z-Komponente des Translationsvektors bei der linken und rechten Projektionsmatrix gleichmäßig variiert. Man erkennt, dass sich die Standardabweichung bei betragmäßig zunehmender y-Komponente erhöht, aber bei zunehmender z-Komponente, also zunehmender Tiefe, verringert.

z-Komp. Translationsvektor	100	100	100	100	100	100
y-Komp. Translationsvektor	± 1	± 5	± 10	± 20	± 50	± 100
Mittelwert (rad)	-0.0003	-0.0017	-0.0016	-0.0131	-0.0016	0.0350
Standardabweichung (rad)	0.0284	0.1175	0.2123	0.3956	0.8780	1.3759

z-Komp. Translationsvektor	150	150	150	150	150	150
y-Komp. Translationsvektor	± 1	± 5	± 10	± 20	± 50	± 100
Mittelwert (rad)	0.0001	0.0013	-0.0009	-0.0081	0.0057	-0.0371
Standardabweichung (rad)	0.0215	0.0843	0.1533	0.2967	0.6385	1.1182

z-Komp. Translationsvektor	200	200	200	200	200	200
y-Komp. Translationsvektor	± 1	± 5	± 10	± 20	± 50	± 100
Mittelwert (rad)	0.0004	0.0043	0.0063	-0.0059	0.0042	-0.0068
Standardabweichung (rad)	0.0138	0.0619	0.1071	0.2520	0.5289	0.8770

In der Abbildung 3.7 sind die soeben ermittelten Standardabweichungen grafisch dargestellt.

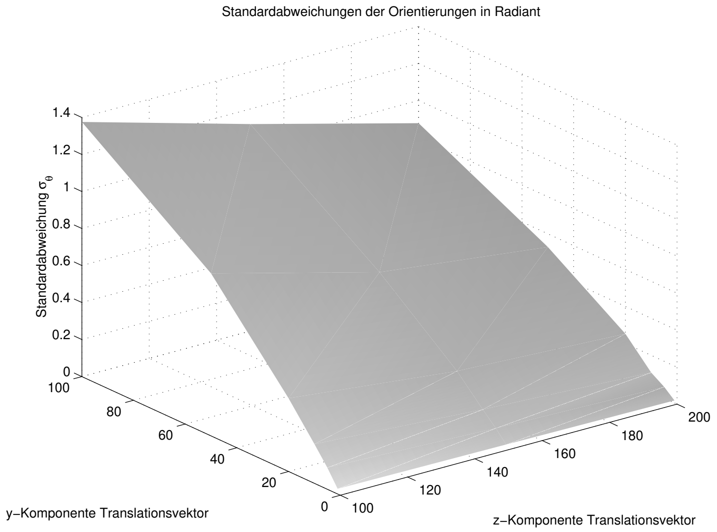


Abbildung 3.7: Standardabweichungen der vorherigen drei Tabellen als 3D Grafik. Die Komponenten des Translationsvektors sind auf der Grundfläche aufgetragen und auf der Ordinate steht die Standardabweichung.

Um die später in Kapitel 4 verwendete Aufnahmegeometrie der Stereobildkamera im Labor zu simulieren, wird der Translationsvektor $\vec{t}$ auf $(0, \pm 12.5, 200)^T$ gesetzt, weil der Kameraabstand 25 cm beträgt und der Objektabstand ca. 2 Meter groß ist. Die Brennweite der verwendeten TRC-Stereokamera beträgt 50 mm, die Zellengröße ist 11×11 Mikrometer bei einer Chipgröße von 10×8.2 mm. Die sich ergebende Standardabweichung wurde durch ein weiteres Zufallsexperiment mit den Werten $t_y = 12.5$ und $t_z = 200$ ermittelt. In der Projektionsmatrix wurden dabei die Werte ξ_u und ξ_v nicht mehr auf eins gesetzt, sondern mit den tatsächlichen Werten belegt. Die berechnete Standardabweichung beträgt : $\sigma_\theta = 0.1823$. Dieser Wert wird später zur konkreten Berechnung von $\alpha(z)$ für die benutzte Aufnahmegeometrie der Laborbilder benötigt.

3.3 Phasenabweichung bei 2D Schwarz-weiß-Kanten

3.3.1 Einführung

In diesem Unterkapitel wird der Phasenverschiebungsfehler auf einem quadratischen Pixelgitter in der Projektionsebene betrachtet. Eine Objektkante im Bild geht meistens nicht direkt durch die Pixelmittelpunkte. Der maximale Phasenverschiebungsfehler bei einer exakt horizontal verlaufenden Objektkante in der Projektionsebene beträgt $\frac{b}{2}$, wobei b die Pixelbreite ist, wie es in Abbildung 3.8 zu erkennen ist. Das ist der Fall, wenn die Kante direkt auf einer Gitterlinie zwischen den Pixeln verläuft. Bei Objektkanten, die diagonal verlaufen, ist der maximale Fehler größer als $\frac{b}{2}$, nämlich $\frac{\sqrt{2}}{2} \cdot b \approx 0.707 \cdot b$. Dieser Fall tritt ein, wenn die Diagonalkante nur eine der vier Ecken des Pixels schneidet.

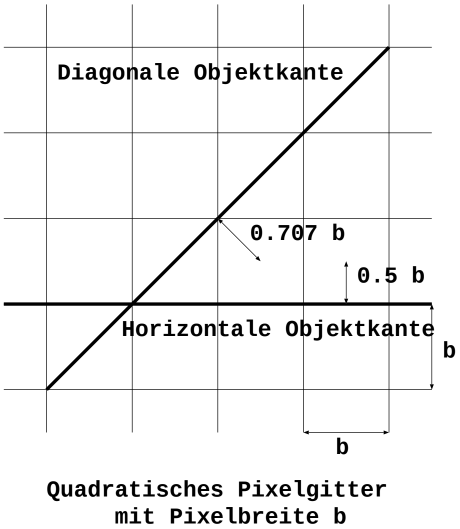


Abbildung 3.8: Phasenverschiebungsfehler von Objektkanten in der Projektionsebene

Zu beachten ist, dass aufgrund der Wahl des Bandpassfilters acht Abtastpunkte auf einer 2π-Periode der linearen Phase vorhanden sind. Der maximale Ortsfehler entlang des Gitters beträgt $\frac{1}{2}$, daher gilt die Gleichung für den maximalen Fehler : $\Delta\varphi_{max} = \frac{2\pi}{8} \cdot \frac{1}{2} = \frac{\pi}{8}$. Die Abbildung 3.9 veranschaulicht diesen Sachverhalt. Entlang der Diagonale ist der maximale Ortsfehler : $\Delta\varphi_{max} = \frac{1}{\sqrt{2}} \cdot \frac{\pi}{4}$.

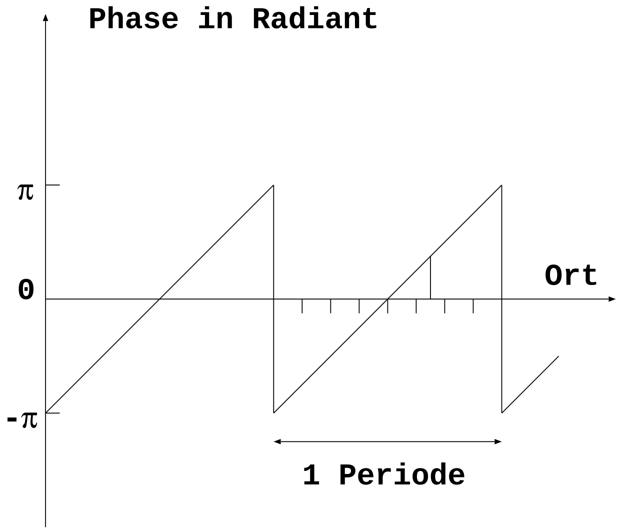


Abbildung 3.9: Maximaler Phasenfehler entlang des Gitters

3.3.2 Berechnungen

1.) Erzeugung von 2D Schwarz-weiß-Kanten durch einen Zufallsgenerator :

Es werden nur Betrachtungen im 2D Fall durchgeführt, d.h. es wird kein 3D Objekt auf die Projektionsebenen der Stereokamera projiziert. Eine Gerade bzw. Kante wird dadurch erzeugt, indem man den Normalenvektor $\vec{n} = (n_x, n_y)$ zu dieser Geraden bzw. Kante bildet. Der Normalenvektor definiert eine Gerade in der Hesseschen Normalform. Ein Zufallsgenerator bestimmt die Orientierung θ des Normalenvektors aus dem Intervall $[0; 2\pi]$. Somit gilt :

$$\theta \in [0, 2\pi] \qquad n_x = \cos\theta \qquad n_y = \sin\theta \tag{3.16}$$

2.) Bestimmung des Grauwertes für jedes Bildpixel :
Das Skalarprodukt zweier Vektoren kann zur Projektion eines Vektors auf einen anderen verwendet werden. Mit Hilfe der Hesseschen Normalform läßt sich der Abstand eines Punktes Q von der Geraden bestimmen. Es wird für alle Punkte Q mit Ortsvektor $\vec{q} = (q_x, q_y)$ folgende Grauwertfunktion $\varrho(\vec{q})$ definiert, wie es in Abbildung 3.10 zu sehen ist :

$$\varrho(\vec{q}) := \begin{cases} 0 & \text{falls} \quad \vec{n}\cdot\vec{q} \geq 0 \\ 1 & \text{falls} \quad \vec{n}\cdot\vec{q} < 0 \end{cases}$$

Der Punkt Q liegt auf der Kante, wenn $\vec{n}\cdot\vec{q} = 0$ gilt. Er liegt im schwarzen Gebiet, wenn $\vec{n}\cdot\vec{q} < 0$ erfüllt ist und er liegt im weißen Gebiet, wenn $\vec{n}\cdot\vec{q} > 0$ gilt.

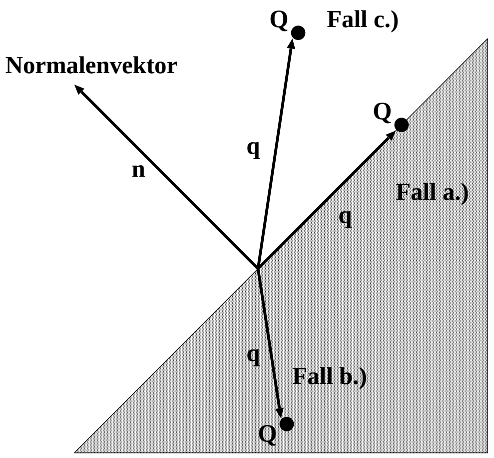


Abbildung 3.10: Positionszuordnung aller Ortsvektoren bzgl. der Kante

3.) Berechnung des Phasenbildes des Strukturmultivektors :
Mit dem niederfrequenten Teil des Bandpassfilters wird, wie in Kapitel 2 beschrieben, das zugehörige Phasenbild berechnet, um für jedes Pixel des Bildes die Phaseninformation zu erhalten.

4.) Berechnung der Phasendifferenz :
Für alle Pixel auf der erzeugten Kante gilt folgende Gleichung, in der für ganzzahlige q_x das zugehörige reelle q_y ermittelt wird :

$$n_x \cdot q_x + n_y \cdot q_y = 0 \quad \Leftrightarrow \quad q_y = -\frac{n_x \cdot q_x}{n_y} \tag{3.17}$$

Nach der Berechnung der reellen q_y-Werte erfolgt eine Rundung auf ganzzahlige q_y-Werte, damit das Koordinatenpaar (q_x,q_y) einen Kantenpixel darstellt. Anschließend wird für das Pixel (q_x,q_y) die Phase $\varphi(q_x, q_y)$ aus dem Phasenbild des Strukturmultivektors entnommen. Die Phasendifferenz zwischen einem Kantenpixel, das nach dem soeben beschriebenen Verfahren ermittelt wurde, und einem tatsächlichen Kantenpunkt wird berechnet durch :

$$\Delta\varphi = \begin{cases} \frac{\pi}{2} - \varphi(q_x, q_y) & \text{falls} \quad \left|\frac{\pi}{2} - \varphi(q_x, q_y)\right| < \left|-\frac{\pi}{2} - \varphi(q_x, q_y)\right| \\ -\frac{\pi}{2} - \varphi(q_x, q_y) & \text{falls} \quad \left|\frac{\pi}{2} - \varphi(q_x, q_y)\right| \geq \left|-\frac{\pi}{2} - \varphi(q_x, q_y)\right| \end{cases}$$

Dabei ist $\pm\frac{\pi}{2}$ die korrekte Phase einer Schwarz-weiß-Kante bzw. einer Weiß-schwarz-Kante, wie man in Abbildung 3.1 erkennen kann.

5.) Erzeugung von Histogrammen :

Für eine Gesamtanzahl von 20 und 80 Normalenvektoren bzw. Kanten wurde jeweils ein Histogramm erzeugt. Im linken Bild der Abbildung 3.11 wurden 20 Kanten betrachtet. Der Mittelwert liegt bei $6.1 \cdot 10^{-4}$ Radiant und die Standardabweichung bei $\sigma_\varphi = 0.1050$. Im rechten Histogramm von Abbildung 3.11 wurden 80 Kanten betrachet. Der Mittelwert beträgt $1.5653 \cdot 10^{-4}$ Radiant und die Standardabweichung ist $\sigma_\varphi = 0.1032$. Zu bemerken ist, dass die Standardabweichung der Phasendifferenzen im Vergleich zur Standardabweichung der Orientierungsdifferenzen kameraunabhängig ist. Die zugehörigen Gaußkurven sind dargestellt.

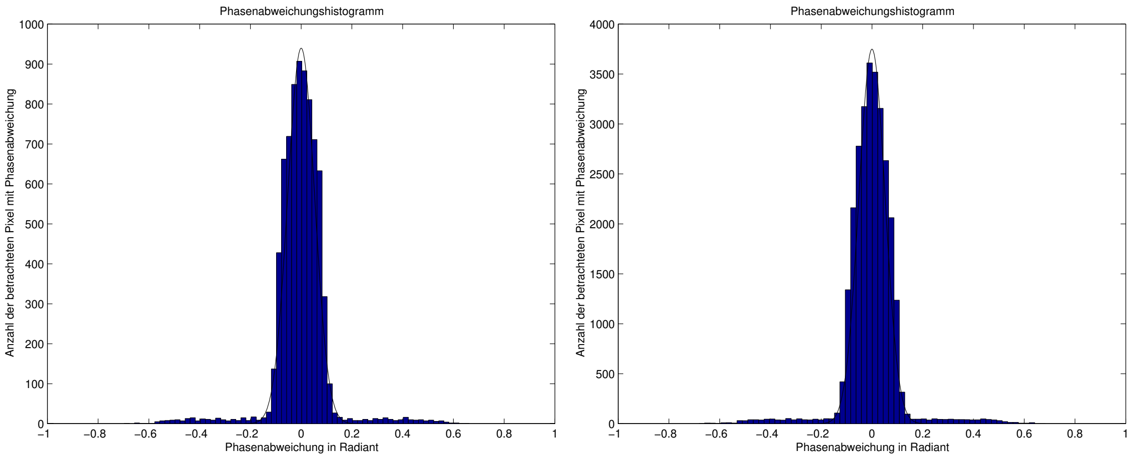


Abbildung 3.11: Histogramme der Phasenabweichungen in Radiant bei 20 Kanten im linken Bild und 80 Kanten im rechten Bild mit passender Gaussapproximation. Die Kanten wurden durch einen Zufallsgenerator erzeugt. Auf der Ordinate ist die Anzahl der Kantenpixel abzulesen.

3.4 Funktion $\alpha(\mathbf{z})$

Im Folgenden wird die tiefenabhängige Funktion $\alpha(z)$ hergeleitet und anschließend der Funktionswert für die verwendeten Laboraufnahmen ermittelt. Zur Ableitung von $\alpha(z)$ wird die Wahrscheinlichkeitsdichte der Normalverteilung benötigt :

$$p(\mathbf{x}) = \frac{1}{2\pi\sqrt{|\Sigma|}} \cdot e^{-\frac{1}{2}(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)} \tag{3.18}$$

Dabei läßt sich die Kovarianzmatrix Σ mit Determinante $|\Sigma|$ durch folgenden Erwartungswert E ausdrücken.

$$\Sigma = E[(\mathbf{x}-\mu)(\mathbf{x}-\mu)^T] \tag{3.19}$$

Der Mittelwert μ entspricht dem Erwartungswert von $\mathbf{x}$.

$$\mu = E[\mathbf{x}] = \int_{-\infty}^{\infty} \mathbf{x}p(\mathbf{x})d\mathbf{x} \tag{3.20}$$

In der Gleichung für die Wahrscheinlichkeitsdichte $p(\mathbf{x})$ wird der folgende Term als Mahalanobisdistanz Δ^2 von $\mathbf{x}$ nach μ bezeichnet :

$$\Delta^2 = (\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu) \tag{3.21}$$

Die Kovarianzmatrix Σ und ihre Inverse Σ^{-1} lauten für die hier betrachteten Verteilungen der Phase φ und der Orientierung θ :

$$\Sigma = \begin{pmatrix} \sigma_\theta^2 & 0 \\ 0 & \sigma_\varphi^2 \end{pmatrix} \qquad \Sigma^{-1} = \begin{pmatrix} \frac{1}{\sigma_\theta^2} & 0 \\ 0 & \frac{1}{\sigma_\varphi^2} \end{pmatrix} \tag{3.22}$$

Die Phase φ und die Orientierung θ sind voneinander unabhängig, d.h. sie sind unkorreliert. Zur Berechnung der Mahalanobisdistanz Δ^2 werden die beiden Terme für die Phase und die Orientierung berechnet, wobei θ_M und φ_M die jeweiligen Erwartungswerte darstellen.

$$(\theta-\theta_M, 0) \cdot \begin{pmatrix} \frac{1}{\sigma_\theta^2} & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \theta-\theta_M \\ 0 \end{pmatrix} = \frac{(\theta-\theta_M)^2}{\sigma_\theta^2} \tag{3.23}$$

$$(0, \varphi-\varphi_M) \cdot \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{\sigma_\varphi^2} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \varphi-\varphi_M \end{pmatrix} = \frac{(\varphi-\varphi_M)^2}{\sigma_\varphi^2} \tag{3.24}$$

Durch Summation dieser beiden Terme ergibt sich die Mahalanobisdistanz insgesamt zu :

$$\Delta^2 = \frac{\sigma_\varphi^2(\theta - \theta_M)^2 + \sigma_\theta^2(\varphi - \varphi_M)^2}{\sigma_\theta^2 \sigma_\varphi^2} \tag{3.25}$$

Eine Linie mit fester Mahalanobisdistanz entspricht einer Ellipse im Euklidischen Raum. Im Mahalanobisraum entspricht es einem Kreis.

$$\frac{(\theta - \theta_M)^2}{\sigma_\theta^2} + \frac{(\varphi - \varphi_M)^2}{\sigma_\varphi^2} = r^2 \tag{3.26}$$

Durch Normierung dieser Kreisgleichung und Einsetzen der Einheitsdifferenz erhält man folgende Gleichung mit einer Konstanten c :

$$c\left(\frac{1}{\sigma_\theta^2} + \frac{1}{\sigma_\varphi^2}\right) = 1 \qquad \Longrightarrow \qquad c = \left(\frac{\sigma_\theta^2 + \sigma_\varphi^2}{\sigma_\theta^2 \sigma_\varphi^2}\right)^{-1} \tag{3.27}$$

Sei α ein Gewichtungsparameter für die Phase φ und entsprechend $1 - \alpha$ eine Gewichtung für die Orientierung θ. Dann gilt mit $\alpha = \frac{c}{\sigma_\varphi^2}$:

$$\alpha = \frac{\sigma_\theta^2}{\sigma_\varphi^2 + \sigma_\theta^2} \qquad\qquad (1 - \alpha) = \frac{\sigma_\varphi^2}{\sigma_\varphi^2 + \sigma_\theta^2} \tag{3.28}$$

Weil die Standardabweichung bzw. Varianz der Orientierung tiefenabhängig ist, gilt :

$$\alpha(z) = \frac{\sigma_\theta^2(z)}{\sigma_\varphi^2 + \sigma_\theta^2(z)} \tag{3.29}$$

In den vorangegangenen Abschnitten wurden die Werte $\sigma_\varphi = 0.1032$ und $\sigma_\theta = 0.1823$ für die im Labor aufgenommenen Stereokamerabilder ermittelt. Die Tiefe z beträgt dabei $2m$. Es ergibt sich $\alpha(2m) = 0.757$, d.h. dass bei der Korrespondenzfindung die Phasendifferenz zwischen zwei 2D Punkten mehr Gewicht als deren Orientierungsdifferenz hat. In Kapitel 4 wird der ermittelte Funktionswert $\alpha(z)$ bei der Erzeugung von Repräsentationen bei vorgegebenen Stereobildsequenzen benötigt.

Kapitel 4

3D Repräsentationen

Bei der Korrespondenzfindung mit der in Kapitel 3 entwickelten Ähnlichkeitsfunktion werden korrespondierende 2D Liniensegmentmittelpunkte in einem Stereobildpaar ermittelt. Aus jedem korrespondierenden 2D Punktepaar wird ein 3D Liniensegmentmittelpunkt erzeugt. Die 3D Orientierung ergibt sich aus den beiden 2D Orientierungen und die 3D Konfidenz berechnet sich aus der Ähnlichkeitsfunktion, wie es in Kapitel 3 dargestellt wurde. Wenn aus einem Stereobildpaar nach der Korrespondenzfindung die 3D Segmente erzeugt wurden, kann aus diesen 3D Segmenten eine 3D Bildrepräsentation gewonnen werden. Durch mehrfache Objektdrehung erhält man eine Sequenz von Stereobildpaaren, d.h. man erzeugt eine Sequenz von 3D Bildrepräsentationen. Die erste 3D Bildrepräsentation bildet die initiale 3D Objektrepräsentation. Jede weitere 3D Bildrepräsentation wird durch eine starre Körperbewegung, d.h. durch eine Koordinatentransformation bzgl. Rotation und Translation, in die Position der initialen 3D Objektrepräsentation zurückgedreht, um einen 3D Liniensegmentvergleich durchzuführen. Die Konfidenz der 3D Segmente wird bei jedem Aktualisierungsschritt der 3D Objektrepräsentation mit einer 3D Bildrepräsentation neu ermittelt. Gegebenfalls werden neue 3D Segmente in die 3D Objektrepräsentation eingefügt, wenn durch eine Objektdrehung neue Seitenflächen etc. sichtbar werden. Das Aktualisieren der 3D Objektrepräsentation mit 3D Bildrepräsentationen wird Akkumulation genannt. Die Akkumulation wird mit allen 3D Bildrepräsentationen, die aus der Stereobildsequenz erzeugt werden, durchgeführt, um als Endergebnis eine *robuste* 3D Objektrepräsentation zu erhalten. Zusammenfassend sei erwähnt, dass sich eine 3D Bildrepräsentation stets auf ein einziges Stereobildpaar und eine 3D Objektrepräsentation sich auf den akkumulierten Fall, also auf mehrere 3D Bildrepräsentationen bezieht. Im Folgenden wird häufig bei den Begriffen Bild- und Objektrepräsentation auf das "3D"-Präfix verzichtet, weil der Begriff Repräsentation in dieser Arbeit stets eine 3D Darstellung beinhaltet.

Die Erzeugung einer Bild- und Objektrepräsentation basiert auf einem Algorithmus von Ackermann [Ack00], daher wird sein Verfahren mit der Vorgehensweise in dieser Arbeit verglichen. Anhand von Beispielen wird auch der Unterschied in den Ergebnissen bei den Bild- und Objektrepräsentationen dargestellt.

4.1 Erzeugung einer 3D Bildrepräsentation

4.1.1 Vorgehensweise von Ackermann

Für ein Stereobildpaar wird eine Bildrepräsentation durch folgenden Algorithmus nach Ackermann [Ack00] erzeugt. Eine detaillierte Darstellung ist aus zwei Gründen relevant. Einerseits kann man die Unterschiede zwischen der vorliegenden Arbeit und der Arbeit von Ackermann besser erkennen, andererseits werden aber auch einige Methoden direkt angewendet.

- 1.) Bilderzeugung :
  Vor der Bilderzeugung ist eine Kamerakalibration und die Festlegung des Weltkoordinatensystems erforderlich. Dann erzeugt eine Stereokamera von einem im Greifer des Roboterarms eingespannten Objekt ein linkes und rechtes Bild.

- 2.) Gradientenbetrag und Gradientenorientierung :
  Beide Bilder werden mit einem 3×3 Binomialfilter, d.h. einem diskretem Gaußfilter, geglättet. Dann werden für beide Bilder das Gradientenbetragsbild und das Gradientenwinkelbild, auch Gradientenorientierungsbild genannt, berechnet. Anschließend wird eine Kontrastnormalisierung durchgeführt, d.h. der Gradientenbetrag wird für jeden Punkt auf das reelle Intervall $[0; 1]$ normalisiert. Danach wird das Gradientenbetragsbild mit einem Binomialfilter geglättet.

- 3.) Houghtransformation :
  Auf das linke und rechte Bild wird die Houghtransformation angewandt. Sie ist robust gegenüber Lücken in den Objektkanten und gegenüber Rauschen, allerdings geht bei der Houghtransformation die Lokalität verloren, d.h. der genaue Anfangs- und Endpunkt einer Objektkante lassen sich schwer bestimmen. Problematisch bei der Houghtransformation ist, dass bei Objektkanten, die mehrere Pixel breit sind, mehrere parallele Houghgeraden erzeugt werden. Dieses Problem wird gelöst durch Unterdrückung von Nichtmaxima entlang des Gradienten. Das Ergebnis der Houghtransformation ist eine Menge von Houghgeraden, die teilweise mit den Objektkanten zusammenfallen.

- 4.) 2D Liniensegmente :
  Im linken Kamerabild werden die Houghgeraden in 2D Liniensegmente konstanter Länge diskretisiert. Dadurch wird nachträglich die global wirkende Houghtransformation lokalisiert. Das ist in der Bildvorverarbeitung der entscheidende Unterschied zur Verwendung des Strukturmultivektors. Ein 2D Liniensegment (x, y, θ, l, c) ist eindeutig bestimmt durch den Segmentmittelpunkt (x, y), durch die Orientierung θ, durch die Länge l und durch die Konfidenz c.

- 5.) Konfidenzbestimmung :
  Für jedes 2D Liniensegment geschieht im linken Kamerabild die Konfidenzbestimmung auf der Basis vom Gradientenbetrag und der Gradientenorientierung. Die Kon-

fidenz ist proportional zum Gradientenbetrag, aber umgekehrt proportional zur Orientierungsabweichung zwischen der Houghgerade, aus der das 2D Segment erzeugt wurde, und dem Gradienten. Der Wert für die Konfidenz liegt im Intervall $[0; 1]$ und gibt die Wahrscheinlichkeit für das Zusammenfallen von einem 2D Liniensegment und der zugehörigen Objektkante an. Die Konfidenz ist daher ein Maß für die Sicherheit bzw. der Glaubwürdigkeit der Information. Anschließend wird eine Schwellwertoperation bzgl. der Konfidenz durchgeführt. Ein 2D Liniensegment kann somit nur existieren, wenn es einen globalen Support durch die Houghtransformation und einen lokalen Support durch die Konfidenzbewertung hat.

- 6.) Korrespondierende 2D Liniensegmente :
  Ziel ist es, zu einem 2D Liniensegment im linken Bild das korrespondierende Segment im rechten Bild zu finden. Dazu wird im linken Bild der 2D Liniensegmentmittelpunkt mit einem $n \times m$ Fenster umgeben. Dieses Fenster wird mit allen möglichen $n \times m$ Fenstern im rechten Bild verglichen, wobei die Fenstermittelpunkte auf der Epipolargeraden sind. Dabei wird jeweils die Quadratsumme der Grauwertdifferenzen zwischen den Fenstern berechnet, was man als photometrisches Matching bzw. als Template-Matching bezeichnet. Die Fenstermittelpunkte im rechten Bild mit einer kleinen Quadratsumme weisen die größte Ähnlichkeit in der Umgebung auf und zählen zu den Kandidaten des korrespondierenden Punktes.

  Der 2D Liniensegmentmittelpunkt im linken Bild entspricht also einer Epipolargeraden im rechten Bild, auf der der korrespondierende Punkt liegt. Durch die Nutzung der Epipolargeometrie wird somit das 2D Problem auf ein 1D Problem reduziert. Die Korrespondenzfindung beschränkt sich auf die 2D Liniensegmentmittelpunkte. Die Schnittpunkte der Epipolargeraden mit den Houghgeraden im rechten Bild liefern eine Kandidatenmenge für den korrespondierenden Punkt. Die Kandidatenmenge wird eingeschränkt durch das photometrische Matching und durch den Vergleich der Orientierung der Segmente. Außerdem darf eine maximale Disparität der Projektionen nicht überschritten werden.

- 7.) 3D Liniensegmente :
  Wenn zwei korrespondierende 2D Liniensegmente gefunden wurden, ist für die 3D Segmentberechnung eine Längenbestimmung, eine Tiefenschätzung und eine Orientierungsbestimmung möglich. Die Tiefenschätzung erfolgt anhand der bekannten Projektionsmatrizen. Das 3D Liniensegment (x,y,z,ψ_1,ψ_2,l,c) wird aus den Projektionen des Segmentes in beiden Kamerabildern berechnet. Der Mittelpunkt (x,y,z) wird aus den Mittelpunkten der 2D Segmente bestimmt. Die Orientierung (ψ_1,ψ_2) in Kugelkoordinaten ergibt sich aus den Orientierungen der 2D Segmente. Die Segmentlänge l ist im 3D Raum nicht mehr konstant, weil sie vom Blickwinkel abhängig ist. Die Konfidenz c ist gegeben durch die Konfidenzen der 2D Segmente und die durch die Kreuzkorrelation gegebene Korrespondenzkonfidenz. Die 3D Segmente werden folgendermaßen erzeugt :

$$(x_1, y_1, \theta_1, l_1, c_1) \times (x_2, y_2, \theta_2, l_2, c_2) \longmapsto (x, y, z, \psi_1, \psi_2, l, c) \tag{4.1}$$

- 8.) Bildrepräsentation :
  Zuletzt wird die berechnete Bildrepräsentation mit der zugehörigen Projektionsmatrix auf das linke Kamerabild projiziert und als Ergebnis angezeigt. Die erste Bildrepräsentation einer Stereobildsequenz ist zugleich die initiale Objektrepräsentation.

4.1.2 Neue Vorgehensweise

Die Unterschiede zwischen Ackermann und der Vorgehensweise in der vorliegenden Arbeit basieren auf zwei wesentlichen Punkten. Erstens verzichtet man auf die Houghtransformation und verwendet den Strukturmultivektor zur Bildvorverarbeitung und zweitens wird eine andere Ähnlichkeitsfunktion zur Korrespondenzfindung benutzt. Ackermann wählte die Houghtransformation, um eine robuste Information zu erhalten. Durch die neue Vorgehensweise möchte man aber eine möglichst genaue und vollständige Objektrepräsentation generieren.

- 1.) Bilderzeugung :
  Die Bilderzeugung erfolgt wie bei Ackermann.

- 2.) Strukturmultivektor :
  Wie bereits in Kapitel 2 näher erläutert wurde, werden die Eingabebilder durch die Fouriertransformation in den Frequenzraum abgebildet, um die Spektren der Eingabebilder zu ermitteln. Die Fouriertransformierten der Komponenten des sphärischen Quadraturfilters werden berechnet, indem die Spektren mit einem Bandpass und reellwertigen Übertragungsfunktionen multipliziert werden. Durch die anschließende inverse Fouriertransformation erhält man dann die Komponenten des Strukturmultivektors, aus denen die lokale Amplitude als energetische Information, die lokale Orientierung als geometrische Information und die lokale Phase als strukurelle Information ermittelt werden können. Die lokale Phase gibt dabei Aufschluss über die Kantenart bzw. Linienart.

- 3.) Phasenkongruenz :
  Die Phasenkongruenz wird aus der Amplitude der Strukturmultivektoren für das obere und untere Band erzeugt. Dabei verwendet man zwei Bandpassfilter mit jeweils drei Oktaven Bandbreite. Anschließend wird eine Schwellwertbildung auf die Phasenkongruenz durchgeführt, um das Bild auf die Kanteninformationen zu reduzieren. In der Abbildung 4.1 erkennt man Schwellwertoperationen mit den Werten 0.5, 0.7 und 0.8. Es muss abhängig von den Grauwertkontrasten im Bild ein angemessener Schwellwert gefunden werden, damit zugleich das Rauschen minimiert wird, aber trotzdem keine Kanteninformation verlorengeht. Vor allem kleinere Strukturen verschwinden schnell bei hohen Schwellwerten.

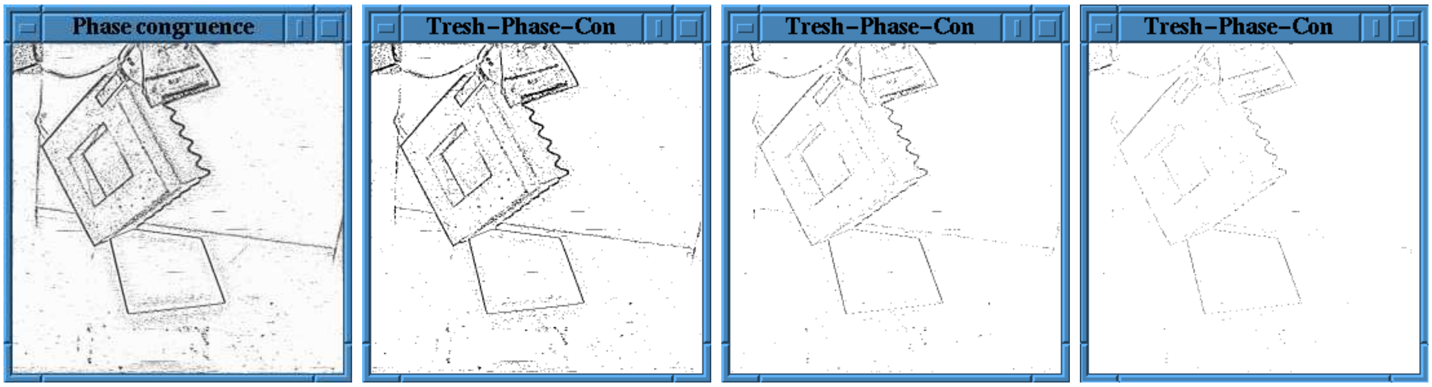


Abbildung 4.1: Von links nach rechts : Phasenkongruenz eines Schaumstoffteils mit Schwellwert 0.0, d.h. Originalbild, dann 0.5, 0.7 und 0.8

- 4.) 2D Liniensegmente :
  Über die Phasenkongruenz des linken Kamerabildes wird ein 5×5 Gitter gelegt und in jeder Gitterzelle wird das Maximum der Phasenkongruenz ermittelt. Die Maxima sind die 2D Liniensegmentmittelpunkte im linken Bild. Somit gibt es in jeder 5×5 Gitterzelle höchstens ein 2D Liniensegment. Falls der vorgegebene Schwellwert auf der Phasenkongruenz höher ist als die vorhandenen Werte in einer betrachteten Gitterzelle, dann gibt es in dieser Gitterzelle kein 2D Liniensegment. Die Orientierung dieser 2D Liniensegmentmittelpunkte wird nachträglich verbessert, wie es im Kapitel 2 beschrieben wurde.

- 5.) Konfidenzbestimmung :
  Die Konfidenzen der 2D Liniensegmentmittelpunkte im linken Bild und der 2D Punkte rechten Bild sind durch die Phasenkongruenz gegeben. Die Phasenkongruenzwerte liegen im Intervall $[0; 1]$.

- 6.) Korrespondierende 2D Liniensegmente : Distanz- und Ähnlichkeitsfunktion
  Gegeben sind die 2D Liniensegmentmittelpunkte im linken Stereokamerabild. Außerdem kennt man die Orientierung, die Phase und die Phasenkongruenz für das linke und das rechte Bild. Die Korrespondenzfindung erfolgt mittels der Epipolargeometrie und der Ähnlichkeit von Phase und Orientierung zwischen den Bildern, wie es in Kapitel 3 erläutert wurde. Im Gegensatz zu Ackermann [Ack00] entfallen das photometrische Matching und die Kandidatenbildung durch die Schnittpunktbestimmung der Epipolargerade mit den Houghgeraden.

  Aus numerischen Gründen werden zu allen 2D Liniensegmenten im linken Bild, die zu nahe am Rand liegen, keine korrespondierende Segmente gesucht, weil die Berechnung der Epipolargeraden im rechten Bild für Randpunkte im linken Bild Probleme bereitet. Für alle 2D Segmente im linken Bild, deren Abstand mindestens 5 - 10 Pixel vom Rand beträgt, wird die zugehörige Epipolargerade erzeugt. Die Kandidatenmenge bilden somit alle Punkte auf der Epipolargeraden, wobei jedoch durch die folgenden Kriterien diverse Punkte aus der Kandidatenmenge entfernt werden. Erstens werden alle Punkte entfernt, deren Orientierung mit der Epipolargeraden fast identisch ist, weil man sonst den 2D Liniensegmentmittelpunkt nicht genau bestimmen kann. Zweitens werden alle Punkte entfernt, die auf der Epipolargeraden zu nahe am Rand sind, weil die Phase nicht mehr eindeutig ist. Drittens bleiben nur diejenigen Punkte in der Kandidatenliste, die eine vorgegebene maximale Disparität nicht überschreiten, d.h. wenn (u_{left}, v_{left}) die Koordinaten eines Punktes im linken Bild und (u_{right}, v_{right}) die Koordinaten eines Punktes im rechten Bild sind, dann gilt für die Disparität λ :

$$\lambda = v_{right} - v_{left} = \frac{\kappa f}{z} \tag{4.2}$$

Dabei ist κ der Abstand der optischen Zentren und f der Abstand von den Projektionsebenen zu den optischen Zentren. Die Größe z stellt den Abstand vom 3D Punkt zur Verbindungsgerade zwischen den optischen Zentren dar. Wenn der 3D Punkt im Unendlichen liegt, ist die Disparität null. Falls der 3D Punkt in die Nähe eines optischen Zentrums gebracht wird, nimmt die Disparität sehr hohe Werte an.

Diese Kriterien schränken die möglichen Koordinaten des korrespondierenden 2D Liniensegmentmittelpunktes im rechten Bild ein. Alle jetzt noch vorhandenen Punkte in der Kandidatenmenge werden durch die im Kapitel 3 behandelte Distanz- und Ähnlichkeitsfunktion untersucht. Die Distanzfunktion d berechnet die Phasen- und Orientierungsdifferenz zwischen einem 2D Liniensegmentmittelpunkt im linken Bild und einem Punkt auf der zugehörigen Epipolargeraden im rechten Bild. Der Funktionswert $\alpha(z)$ gibt die Gewichtung zwischen der Orientierungs- und Phasendifferenz in der Distanzfunktion in Abhängigkeit von der Tiefe z an. Als Ergebnis erhält man durch die Ähnlichkeitsfunktion den korrespondierenden Punkt im rechten Bild, bei dem die Konfidenz des zugehörigen 3D Segmentes maximal wird.

- 7.) 3D Liniensegmente :
  Ein 3D Liniensegment wird nur dann erzeugt, wenn es zu einem gegebenen 2D Segment im linken Bild auch ein korrespondierendes 2D Segment im rechten Bild gibt. Es werden wie bei Ackermann der Weltpunkt und die Orientierung etc. des 3D Segmentes berechnet. Allerdings betrachtet Ackermann den Anfangs- und Endpunkt des 3D Segmentes, was numerische Probleme bereitet. In der vorliegenden Arbeit beziehen sich alle Berechnungen nur auf die Liniensegmentmittelpunkte. Die Konfidenz des 3D Segmentes ergibt sich aus der Ähnlichkeitsfunktion. Fehlkorrespondenzen sind durchaus zugelassen und fallen später bei der Akkumulation wegen geringer Konfidenz heraus, weil die zugehörigen 3D Segmente nicht durch andere Bildrepräsentationen bestätigt werden.

- 8.) Bildrepräsentation :
  Aus den 3D Liniensegmenten wird wie bei Ackermann eine Bildrepräsentation erzeugt.

- 9.) Qualitätskontrolle :
  Die Bildrepräsentation wird mit einem 3D Display-Tool auf deren Qualität untersucht. Mit diesem Tool werden die 3D Segmente grafisch dargestellt und es ist eine Drehung der Bildrepräsentation möglich. Dadurch lassen sich die Qualität der Kantendetektion, die Häufigkeit der Fehlkorrespondenzen und die korrekte Orientierung der 3D Segmente überprüfen.

Die nachfolgenden beiden Flussdiagramme in den Abbildungen 4.2 und 4.3 veranschaulichen die Vorgehensweise von Ackermann und den Algorithmus in dieser Arbeit, um einen schematischen Überblick zu gewinnen und einen direkten Vergleich zu ermöglichen. Die wesentlichen Unterschiede liegen in der Bildvorverarbeitung und in der Korrespondenzfindung. Bei der global wirkenden Houghtransformation, die Ackermann verwendet hat, werden im Gegensatz zur Vorverarbeitung mit dem Strukturmultivektor kurze Kanten und krummlinige Konturen nur schlecht oder gar nicht erfasst. Durch die Bildvorverarbeitung mit der Houghtransformation erhält man vor allem für längere Kantenstücke eine robuste 3D-Objektrepräsentation. Die Korrespondenzfindung beruht bei Ackermann auf einem Template-Matching. In dieser Arbeit werden jedoch die lokalen Phasen- und Orientierungsdifferenzen betrachtet.

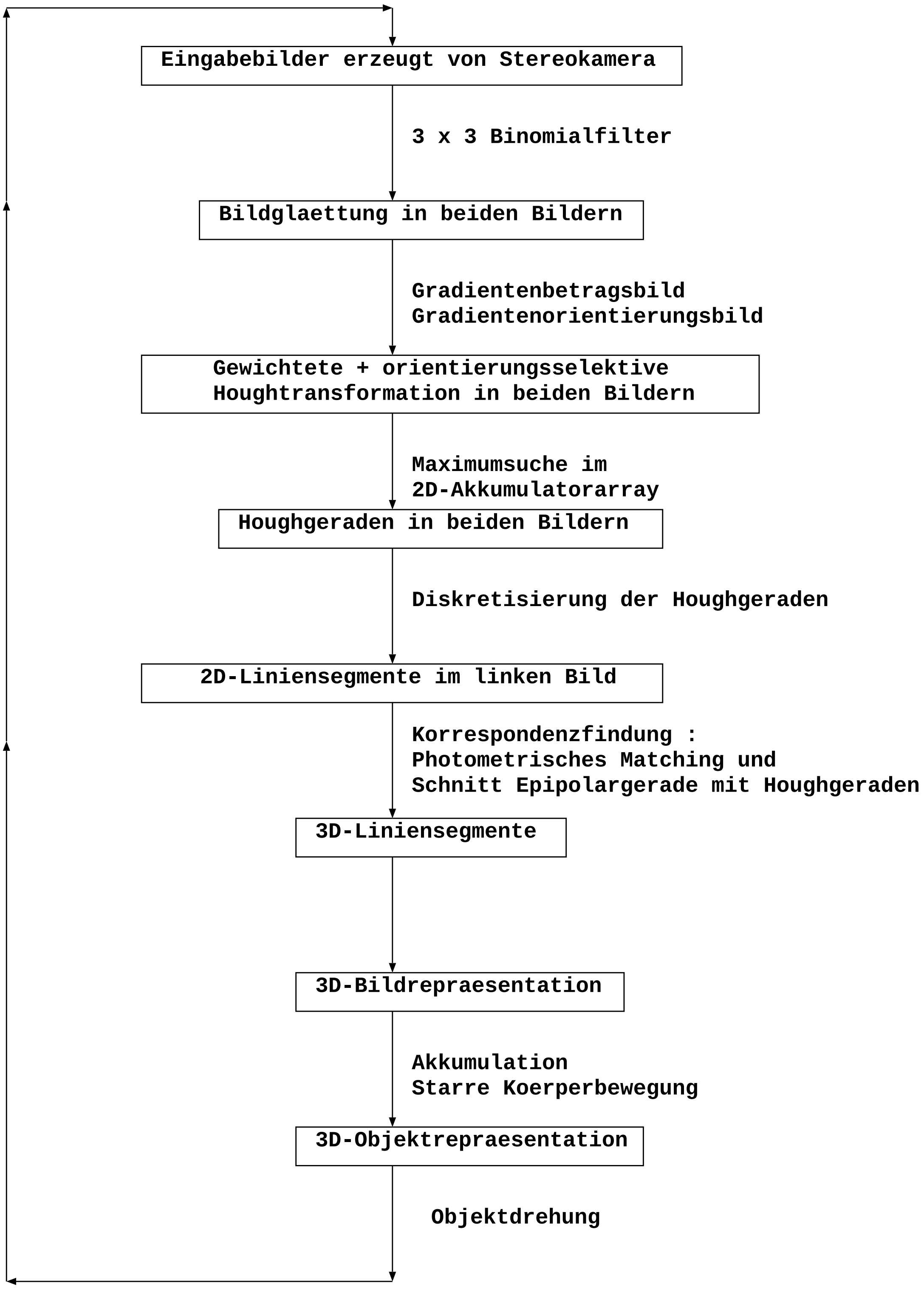


Abbildung 4.2: Vorgehensweise von Ackermann

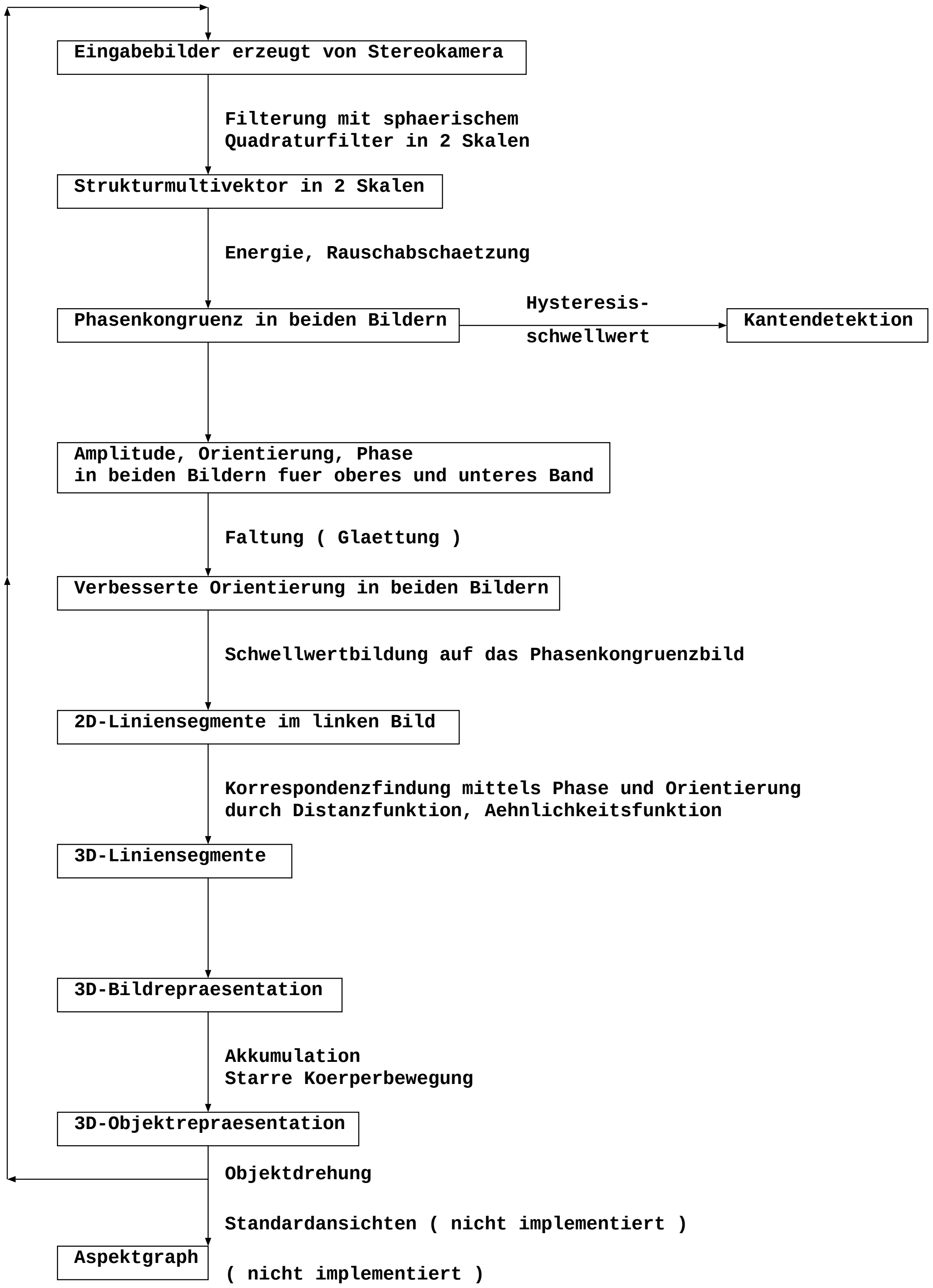


Abbildung 4.3: Vorgehensweise in dieser Arbeit

4.1.3 Ergebnisse der 3D Bildrepräsentation

In der Abbildung 4.4 ist zu erkennen, dass die Qualität der Bildrepräsentation von der Wahl des Schwellwertes auf die Phasenkongruenz abhängt. Exemplarisch sind für die Schwellwerte 0.6 und 0.8 zwei Bildrepräsentationen dargestellt. Der Vergleich mit dem Ergebnis von Ackermann zeigt, dass die Bildvorverarbeitung mit dem Strukturmultivektor und die neue Methode der Korrespondenzfindung genauer sind, weil man die untersuchten Objekte, siehe auch Abbildung 4.5, vollständiger erkennen kann. Besonders hervorzuheben ist die wellenförmige Objektberandung des Schaumstoffteils in Abbildung 4.4, die beim Verfahren in der vorliegenden Arbeit, aber nicht bei der global wirkenden Houghtransformation detektiert wird. Die Anwendung der Houghtransformation ist bei kurzen Geraden und bei krummlinigen Konturen problematisch, weil Informationen verloren gehen können.

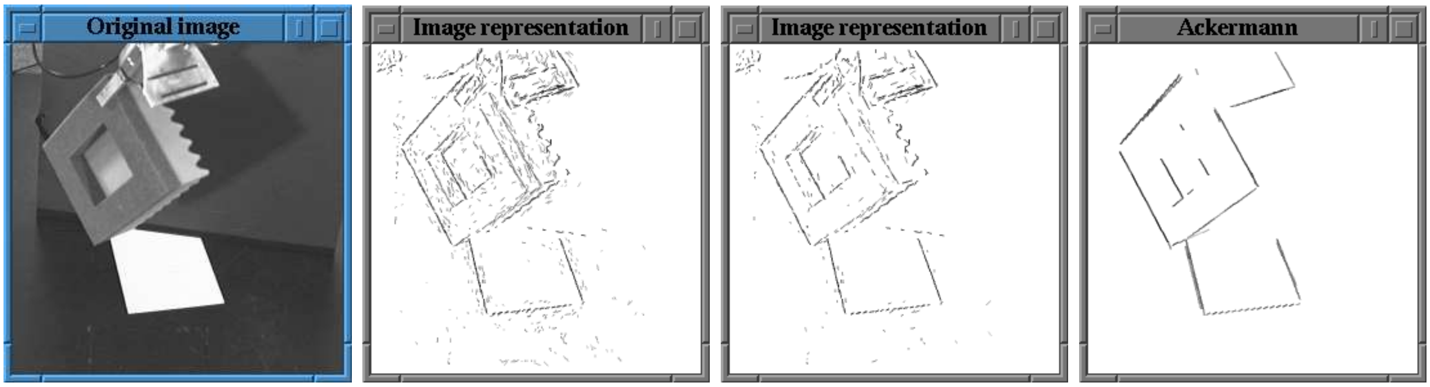


Abbildung 4.4: Links eines der Bilder der Stereokamera von einem Schaumstoffteil und in der Mitte links eine Bildrepräsentation mit einer vorherigen Schwellwertbildung von 0.6 auf die Phasenkongruenz. In der Mitte rechts analog mit einem Schwellwert von 0.8. Ganz rechts zum Vergleich die von Ackermann erzeugte Bildrepräsentation.

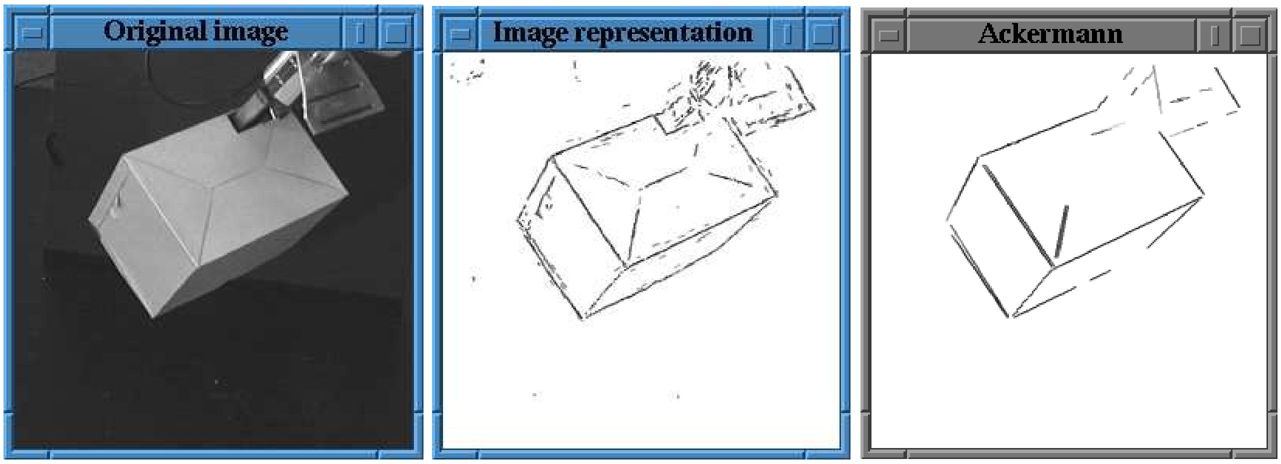


Abbildung 4.5: Links eines der Bilder der Stereokamera und in der Mitte die Bildrepräsentation beim Karton. Rechts zum Vergleich die von Ackermann erzeugte Bildrepräsentation.

Man erkennt in den Abbildungen 4.4 und 4.5, dass die Bildrepräsentationen zwar detailreicher als bei Ackermann sind, aber auch etwas mehr Rauschen enthalten. Im Vergleich zu dem Bild mit den 2D Liniensegmenten des Kartons in Abbildung 2.15 wurde bereits Rauschen unterdrückt, weil entweder die Konfidenz der Rauschsegmente im rechten Bild bei der Korrespondenzfindung zu gering war oder weil keine Korrespondenz gefunden wurde.

In der Abbildung 4.6 wird eine bereits erzeugte Bildrepräsentation mit einem 3D Displaytool untersucht, um zu überprüfen, ob die 3D Liniensegmente bezüglich ihrer Lage und Orientierung den Karton gut widerspiegeln. Abgesehen von einigen Fehlkorrespondenzen, die später im Laufe der Akkumulation wegen geringer Konfidenz herausfallen, ist die 3D Darstellung zufriedenstellend.

Abbildung 4.6: Darstellung einer Bildrepräsentation vom Karton mit einem 3D Displaytool aus verschiedenen Blickwinkeln.

4.2 Erzeugung einer 3D Objektrepräsentation

4.2.1 Vorgehensweise

Die Bildrepräsentation, die aus dem ersten Stereobildpaar einer Bildsequenz erzeugt wird, ist zugleich die initiale Objektrepräsentation. In diesem Abschnitt wird erläutert, wie durch Akkumulation von weiteren Bildrepräsentationen, die durch Objektdrehungen zustande kommen, die Objektrepräsentation aktualisiert wird. Bei der Akkumulation werden unsichere Informationen in der Objektrepräsentation, d.h. 3D Segmente mit geringer Konfidenz, durch weitere Bildrepräsentationen bestätigt oder eliminiert. Im Falle einer Bestätigung, d.h. wenn ein zu betrachtendes Segment auch in anderen Bildrepräsentationen vorhanden ist, erhöht sich die Konfidenz, andernfalls wird sie erniedrigt. Weil sich die Objektrepräsentation dadurch im Laufe der Zeit stabilisiert, verlieren die einzelnen Bildrepräsentationen und deren Fehler nach einigen Aktualisierungsschritten immer mehr an Gewicht.

Beim Aktualisieren einer Objektrepräsentation mit einer Bildrepräsentation wird eine starre Körperbewegung der aktuellen Bildrepräsentation auf die Roboterstellung der initialen

Objektrepräsentation durchgeführt, so dass ein 3D Liniensegmentvergleich zwischen beiden Repräsentationen möglich ist. Bei dem direkten Vergleich aller 3D Liniensegmente gibt es drei Fälle zu unterscheiden. Der erste Fall tritt ein, wenn ein zu betrachtendes 3D Segment in der Objektrepräsentation, aber nicht in der aktuellen Bildrepräsentation vorhanden ist. Die Konfidenz des 3D Segmentes wird dann in der Objektrepräsentation gesenkt. Der zweite Fall tritt ein, wenn das 3D Segment in der Bildrepräsentation, aber nicht in der Objektrepräsentation vorhanden ist. Das 3D Segment wird dann mit geringer Konfidenz in die Objektrepräsentation eingefügt. Der dritte Fall tritt ein, wenn eine Korrespondenz gefunden wurde. Das neue 3D Segment wird aus den korrespondierenden 3D Segmenten berechnet und seine Konfidenz erhöht. Der Algorithmus für die Akkumulation mit dem Zeitparameter t ist folgendermaßen :

- 1.) Zu Beginn wird ein Stereokamerabild in der Roboterstellung $\mathrm{G(t = 0)}$ erzeugt.
- 2.) Aus diesem Stereokamerabild entsteht die erste Bildrepräsentation $\mathrm{IR(t = 0)}$, die zugleich die initiale Objektrepräsentation $\mathrm{OR(t = 0)}$ ist.
- 3.) Der Roboter wird in eine Stellung $\mathrm{G}(t > 0)$ bewegt und es wird eine neue Bildrepräsentation $\mathrm{IR(t)}$ erzeugt.
- 4.) $\mathrm{IR(t)}$ wird durch eine starre Körperbewegung auf die ursprüngliche Roboterstellung $\mathrm{G(t = 0)}$ überführt und es wird die Akkumulation vollzogen, was durch den Operator $\oplus$ dargestellt ist : $OR(t) := OR(t-1) \oplus G_t^0(IR(t))$.
- 5.) Die Roboterstellung wird verändert und man fährt mit Schritt 3.) fort. Diese Iteration bis zum Ende der Bildsequenz der Stereokamera durchgeführt.

Benachbarte 3D Liniensegmente mit nahezu gleicher Orientierung werden nach dem letzten Akkumulationsschritt, also nach der Verarbeitung der gesamten Stereobildsequenz, in der Objektrepräsentation verschmolzen. Man erzeugt dabei aus vorhandenen 3D Liniensegmenten, die bezüglich ihrer Orientierungsabweichung einen bestimmten Schwellwert unterschreiten und deren 3D Abstände der Liniensegmentmittelpunkte nicht zu groß sind, neue 3D Liniensegmente. Dabei wird jedes vorhandene 3D Liniensegment in der Objektrepräsentation, die nach der Akkumulation aller Bildrepräsentationen entstanden ist, höchstens für einen Verschmelzungprozess herangezogen.

Das Hauptproblem bei dem Verschmelzungsprozess besteht darin, geeignete Schwellwerte für die maximale Orientierungsabweichung und die maximalen Abstände der zu verschmelzenden 3D Segmente zu finden. Die Verschmelzung wird mit einer Gewichtung der Konfidenzen der zu verschmelzenden Liniensegmente durchgeführt, d.h. die 3D Koordinaten, die Orientierung und die Länge etc. des neuen Segmentes werden aus den mit der Konfidenz gewichteten zu verschmelzenden Segmenten ermittelt. Im Folgenden wird eine Objektrepräsentation, bei der die Verschmelzung der Liniensegmente stattgefunden hat, als verschmolzene Objektrepräsentation bezeichnet.

4.2.2 Ergebnisse der 3D Objektrepräsentation

In diesem Abschnitt werden Objektrepräsentationen und deren Verschmelzungen exemplarisch bei zwei Objekten dargestellt. Mit einem 3D Displaytool werden die 3D Liniensegmente mit und ohne Konfidenzberücksichtigung visualisiert. Das in Abbildung 4.7 dargestellte sonderbare Objekt aus Metall wird im Folgenden aufgrund seiner Form mit dem Begriff Schaufel bezeichnet.

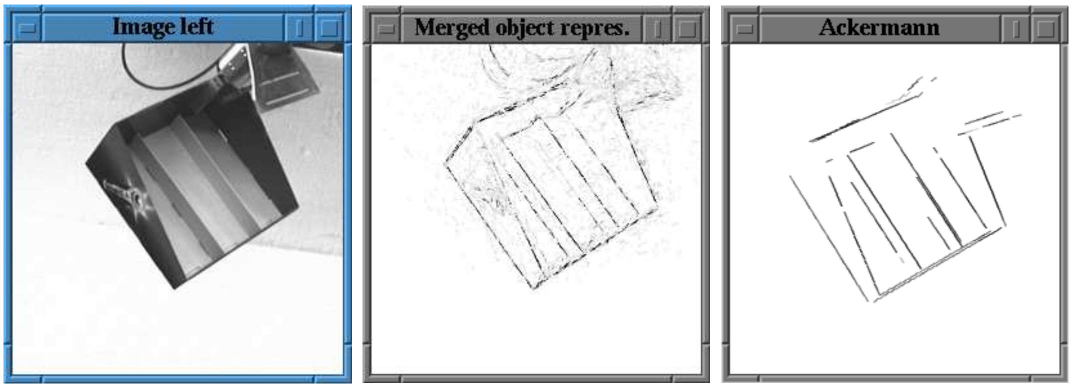


Abbildung 4.7: Links eines der beiden Bilder der Stereokamera von einer Schaufel, in der Mitte die verschmolzene Objektrepräsentation nach 10 Stereobildpaaren und rechts zum Vergleich die verschmolzene Objektrepräsentation von Ackermann

In Abbildung 4.7 ist zu erkennen, dass die verschmolzene Objektrepräsentation ebenso wie die Bildrepräsentation im vorherigen Abschnitt mehr Rauschen enthält, aber dafür mehr Detailinformationen widerspiegelt als bei Ackermann. Eine Verschmelzung nach jedem anstatt nur nach dem letzten Akkumulationsschritt führt in der Regel zu schlechteren Ergebnissen, weil diverse 3D Segmente mit relativ geringer Konfidenz bereits in einem frühen Akkumulationsstadium verschmolzen werden, ohne eine ausreichende Bestätigung durch weitere Bildrepräsentationen zu erfahren.

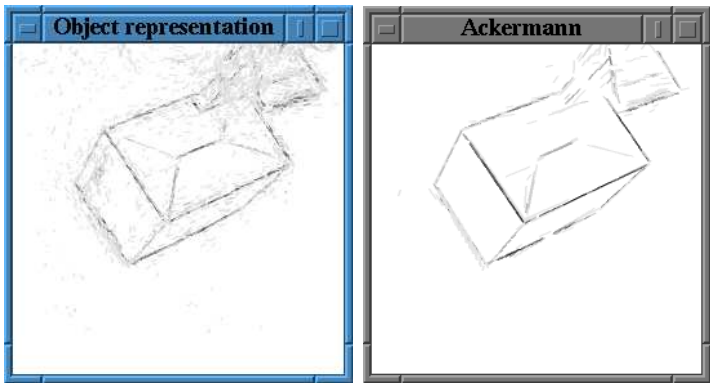


Abbildung 4.8: Links Objektrepräsentation beim Karton nach 10 Stereobildpaaren und rechts zum Vergleich das Ergebnis von Ackermann. Dabei wurde noch keine Verschmelzung durchgeführt.

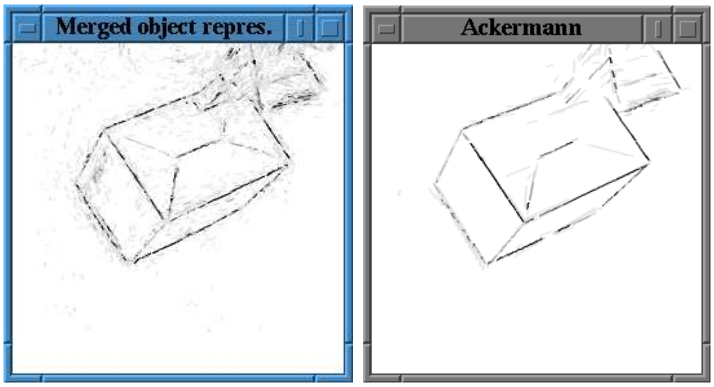


Abbildung 4.9: Links verschmolzene Objektrepräsentation nach 10 Stereobildpaaren und rechts zum Vergleich das Ergebnis von Ackermann

Auch in den Abbildungen 4.8 und 4.9 erkennt man denselben Sachverhalt. Die lokalen Strukturen sind bei Ackermann nicht so deutlich zu erkennen. Die Abbildungen 4.10 und 4.11 zeigen die verschmolzene Objektrepräsentation des Kartons dargestellt mit dem 3D Displaytool. In der Abbildung 4.10 erkennt man alle 3D Segmente ohne Graustufenberücksichtigung ihrer Konfidenz, daher erscheinen die Störungen genauso schwarz wie die Objektkanten. In der Abbildung 4.11 werden die Konfidenzgraustufen berücksichtigt, daher sind die Störungen wegen geringer Konfidenz kaum sichtbar.

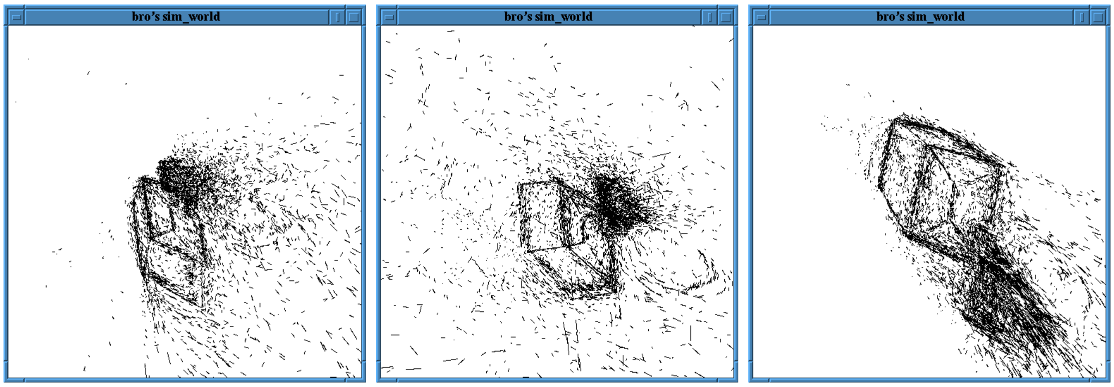


Abbildung 4.10: Verschmolzene Objektrepräsentation beim Karton nach 10 Stereobildpaaren mit einem Displaytool ohne Berücksichtigung der Konfidenz durch das Graustufenspektrum

Abbildung 4.11: Analog wie Abbildung 4.10, aber mit Berücksichtigung der Graustufen in Abhängigkeit der Konfidenz

Zusammenfassend kann man feststellen, dass die Objektkanten eindeutig erkannt werden, aber das Ziel weiterer Arbeiten wird es sein, Verfahren zu finden, um das Rauschen noch stärker zu reduzieren und die Konfidenz der Objektkanten weiter zu erhöhen. In den Abbildungen 4.10 und 4.11 kann man außerdem erkennen, dass die 3D Liniensegmente, die die Objektkanten darstellen, nicht besonders gut verschmolzen werden, daher muss auch der Verschmelzungsprozess noch optimiert werden.

Kapitel 5

Phasenkonsistenz und Aspektgraph

In diesem Kapitel wird die Erhaltung der lokalen Phase bei Objektbewegungen untersucht, um festzustellen, ob sie stabil unter einer Rotation und Translation ist. Dadurch wird überprüft, ob die lokale Phase als strukturelle Information invariant bei einer starren Körperbewegung ist und zur Kodierung von Objekten dienen kann. Die Phasenerhaltung wird sowohl für 2D Punkte als auch für 3D Punkte untersucht, wobei eine Phase für einen 3D Punkt assoziiert wird, die sich durch die Mittelung aus den 2D Phasen korrespondierender 2D Punkte ergibt. Dabei bezieht sich die Korrespondenz auf die in Kapitel 3 behandelte Ähnlichkeitsfunktion. Zu Beginn des Kapitels wird bei einem Stereobildpaar die Phasendifferenz von korrespondierenden 2D Punkten näher betrachtet. Danach wird bei mehrfacher Objektdrehung die Konsistenz der Phase von 2D Kantenmittelpunkten bei Monosequenzen der linken Stereokamerabilder untersucht. Anschließend wird die 3D Phasenkonsistenz beim Aktualisieren einer 3D Objektrepräsentation mit einer 3D Bildrepräsentation überprüft. Zum Schluss des Kapitels wird ein Konzept zur Erstellung von Aspektgraphen näher erläutert. Ein Aspektgraph besteht aus einer Minimalanzahl charakteristischer Objektansichten, die die Knoten im Graph darstellen. An den Kanten im Graph stehen die Rotationsbeziehungen, um von einer Ansicht zur anderen zu gelangen. Es wird ein einfacher Algorithmus zur Aspektgrapherzeugung beschrieben und die Ergebnisse anhand elementarer Beispiele diskutiert.

5.1 2D Phasenkonsistenz

Zuerst werden die Phasendifferenzen von korrespondierenden 2D Punkten in einem Stereobild untersucht. Aufgrund der Aufnahmegeometrie der behandelten Stereosequenzen vom Karton und der Schaufel etc. ist der Funktionswert $\alpha(z)$ von null und eins verschieden, sodass die folgenden Sonderfälle nicht eintreten, dass entweder nur die Orientierung oder nur die Phase bei der Korrespondenzfindung relevant ist. Aufgrund der in Kapitel 3 behandelten Ähnlichkeitsfunktion ist es naheliegend, dass korrespondierende 2D Punkte vermutlich eine ähnliche Phase haben, was im Folgenden überprüft wird. Die Phasenbilder in Abbildung 5.1 werden, wie in Kapitel 2 beschrieben, erzeugt.

Abbildung 5.1: Links eines der Kamerabilder von der Schaufel, in der Mitte das linke Phasenbild und rechts das rechte Phasenbild.

Es werden exemplarisch zwei Histogramme erzeugt, die in der Abbildung 5.2 zu sehen sind. Im linken Histogramm sind die Phasenabweichungen korrespondierender 2D Punkte beim ersten Stereobildpaar der Bildsequenz von der Schaufel dargestellt. Die Korrespondenz bezieht sich dabei auf die im Kapitel 3 behandelte Ähnlichkeitsfunktion. Im rechten Histogramm erkennt man analog die Phasenabweichungen für das erste Stereobildpaar der Kartonaufnahmen. Die Ordinate gibt die Anzahl der korrespondierenden 2D Punkte wieder und auf der Abzisse ist die Phasendifferenz : $\Delta\varphi = ||\varphi_{left}| - |\varphi_{right}||$ in Radiant aufgetragen. Wenn man das jeweilige Vorzeichen der Phasendifferenz berücksichtigen würde, ergäbe sich in beiden Histogrammen aufgrund der Symmetrie eine Gaußverteilung. Man erkennt, dass korrespondierende 2D Punkte in der Regel eine sehr geringe Phasenabweichung haben.

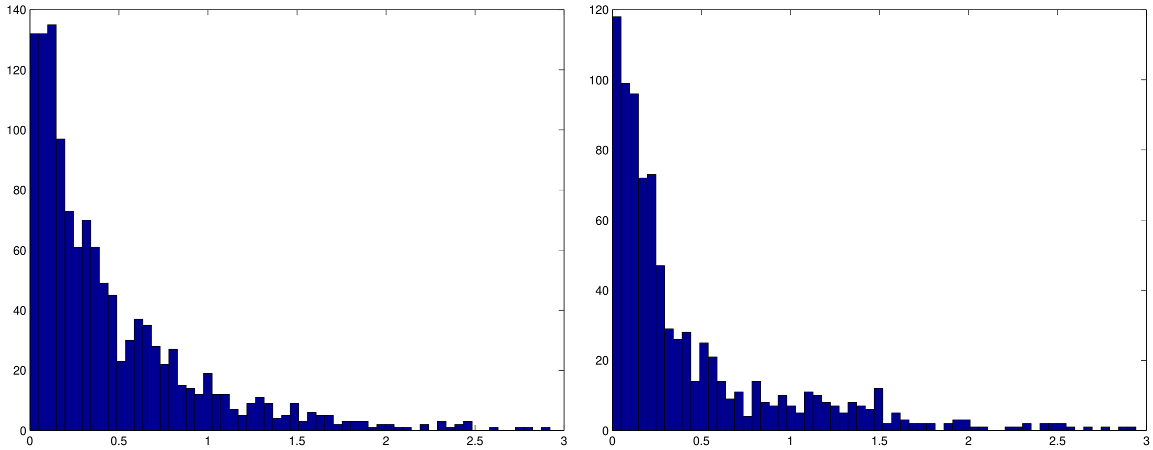


Abbildung 5.2: Phasendifferenzen zwischen korrespondierenden 2D Punkten. Linkes Histogramm : Erstes Stereobildpaar der Schaufel. Rechtes Histogramm : Erstes Stereobildpaar des Kartons. Auf der Ordinate steht die Anzahl der korrespondierenden 2D Punkte und auf der Abzisse die Phasenabweichung in Radiant.

Im Folgenden werden die Phasen von Kantenmittelpunkten bei Monosequenzen der linken Kamerabilder untersucht. Weil eine direkte Korrespondenzfindung innerhalb einer Monosequenz mit dem bisher verwendeten Stereokorrespondenzverfahren nicht möglich ist, werden die Pixelkoordinaten der Kantenmittelpunkte aus den Bildern ermittelt und dann die zugehörigen Phasenwerte ausgegeben. In der Abbildung 5.3 sind die untersuchten Kantenmittelpunkte markiert.

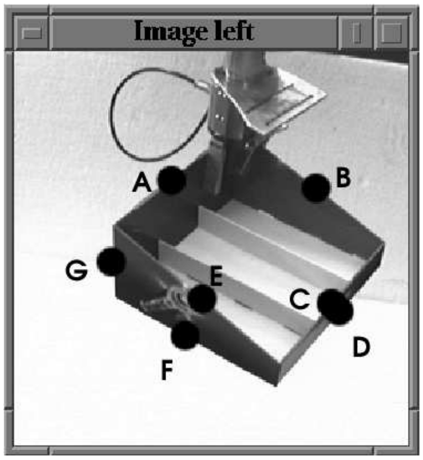


Abbildung 5.3: Bei der Monosequenz der linken Stereokamerabilder von der Schaufel werden die Kantenmittelpunkte A bis G betrachtet.

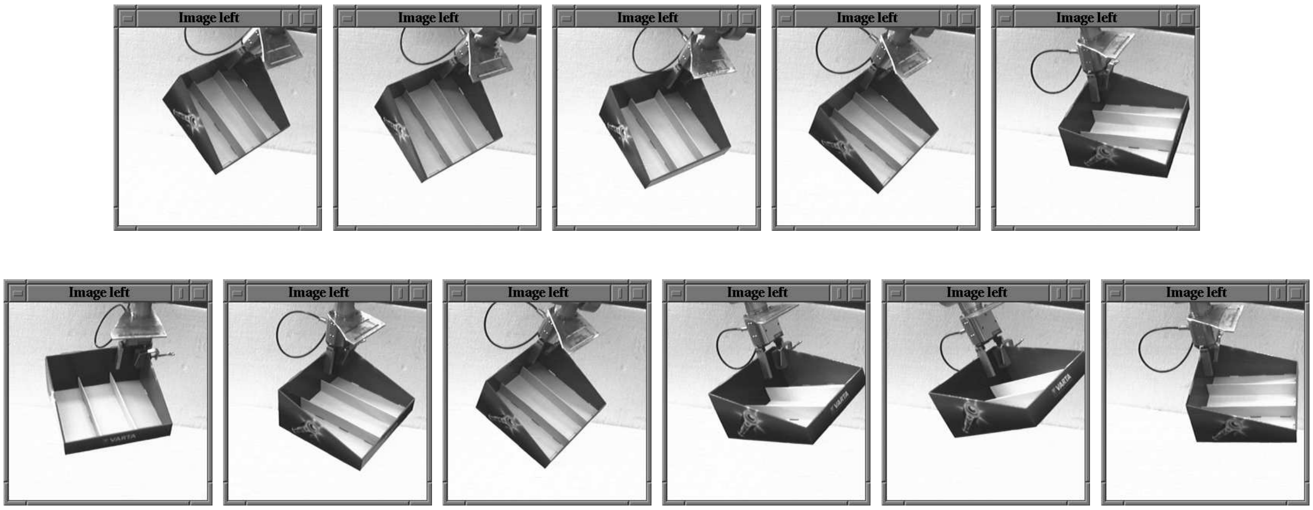


Abbildung 5.4: Monosequenz der linken Stereokamerabilder von der Schaufel : Bild 1 bis 5 in der oberen Reihe und Bild 6 bis 11 in der unteren Reihe.

Die nachfolgende Tabelle enthält die Phasenwerte in Radiant im Intervall $[-\pi;\pi]$ an den sieben Kantenpunkten A bis G bei der Monosequenz in Abbildung 5.4. Wenn ein Kantenmittelpunkt in einer Ansicht nicht erkennbar ist, wird kein Phasenwert, sondern das Symbol $\otimes$ in die Tabelle eingetragen.

Bildnr.	Punkt A	Punkt B	Punkt C	Punkt D	Punkt E	Punkt F	Punkt G
1	2.39	1.53	-1.07	$\otimes$	2.08	-1.31	2.31
2	2.12	1.57	-1.24	$\otimes$	2.21	-1.12	2.39
3	2.15	1.43	2.87	-1.70	2.23	-1.23	2.26
4	2.33	1.49	-1.13	$\otimes$	1.65	-1.25	2.25
5	2.05	1.66	2.06	-1.63	1.53	-1.61	-1.45
6	2.16	1.62	1.94	-1.74	2.97	$\otimes$	$\otimes$
7	2.26	1.79	2.36	-1.71	1.71	-1.28	-1.54
8	2.38	1.70	-1.19	$\otimes$	1.86	-1.42	2.01
9	2.27	1.78	2.12	-1.82	1.75	-1.55	-1.46
10	2.36	1.69	2.21	-1.93	-2.96	-1.74	-1.53
11	2.09	1.64	2.52	$\otimes$	1.20	-1.54	-1.48

Die Phasenwerte der ausgewählten Kantenmittelpunkte A bis G können aufgrund der in Kapitel 3 dargestellten Kanten- und Linienarten in Abbildung 3.1 folgendermaßen interpretiert werden :

- Der Punkt A ist in allen Abbildungen gut zu erkennen und die Phase ist nahezu konstant, zumal der Kantenhintergrund stets gleich ist. Bei der Bildnummer 3, 6, 9 und 10 befindet sich der Kantenmittelpunkt in der Nähe vom Steuerkabel des Greifers, was jedoch einen geringen Einfluss auf die Phase hat. Weil die Kante durch Abnutzung blank erscheint, handelt es sich nicht um eine ideale Weiß-schwarz-Kante bzw. Schwarz-weiß-Kante mit dem Phasenwert $\varphi = \pm\frac{\pi}{2} \approx \pm 1.57$, zumal der Hintergrund im oberen Bildteil hellgrau anstatt weiß ist.

- Den Punkt B kann man ebenso in allen Ansichten gut lokalisieren und die Phase ist auch nahezu konstant wie bei Punkt A. Die Phase von Punkt B liegt fast ideal bei $\frac{\pi}{2}$. Je nach Orientierung zur Lichtquelle erscheint die Kante bei Punkt B manchmal etwas heller, was den Kontrast zum hellen Hintergrund vermindert. Dieser Effekt ist auch bei den anderen Kanten gelegentlich erkennbar.

- Die Punkte C und D liegen sehr dicht zusammen. Wenn man den Punkt D erkennen kann, ist seine Phase konstant bei $\varphi \approx -\frac{\pi}{2}$. Der Punkt C ändert seine Phase in Abhängigkeit davon, ob der Punkt D sichtbar ist. Problematisch ist u.a., dass sich die Helligkeit der Schaufelinnenseite in Abhängigkeit von der Position zur Lichtquelle ändert. Dadurch handelt es sich nicht mehr um eine ideale Kante mit $\varphi = \pm\frac{\pi}{2}$. Bei Bildnr. 11 schaut man direkt auf die Kantenseite.

- Bei dem Punkt E ändert sich die Phase in Abhängigkeit vom Hintergrund der Kante. Bei Bildnr. 6 schaut man direkt von oben auf die Kante, daher unterscheidet sich der Phasenwert von den Aufnahmen mit der Nummer 1 bis 5 und 7 bis 9. Die letzten beiden Bilder mit der Nummer 10 und 11 sind Spezialfälle. Bei Bildnr. 10 ändert sich der Hintergrund von hell auf dunkel. Beim letzten Bild wie auch schon bei Bildnr. 5 trifft eine helle Kante der Schaufelinnenseite direkt auf den Kantenmittelpunkt E.

- Der Punkt F hat eine nahezu konstante Phase bei $\varphi \approx -\frac{\pi}{2}$. Der Hintergrund ändert sich nicht und außer bei Bildnr. 6 ist der Punkt F stets sichtbar.
- Bei dem Punkt G ändert sich die Phase mit wechselndem Vorzeichen, weil durch die Rotation aus einer Schwarz-weiß-Kante eine Weiß-schwarz-Kante wird. Bildnr. 6 ist ein Spezialfall, weil man den Punkt nicht gut lokalisieren kann.

Zusammenfassend läßt sich feststellen, dass die 2D Phase im Wesentlichen erhalten bleibt. Natürlich sind die Phasenwerte der Objektberandungskanten vom Hintergrund abhängig. Daher werden später Kantenmittelpunkte von einer Monosequenz der Kartonaufnahmen untersucht, wo der Hintergrund keinen Einfluss auf die Kanten hat. Zu beachten ist bei der Monosequenz der Aufnahmen von der Schaufel, dass gelegentlich eine Kantenverdeckung aufgetreten ist.

In den Abbildungen 5.5 und 5.6 werden die Phasenwerte beim linken Schaufelbild Nr.1 in den horizontalen Querschitten $Quer2$ und $Quer3$ und im vertikalen Querschnitt $Quer1$ dargestellt. Im rechten Bild der Abbildung 5.5 sind die ersten ca. 160 Pixel und die letzten ca. 60 Pixel Rauschen, weil in diesen Bereichen keine Objektkanten vorhanden sind. Im linken Bild der Abbildung 5.6 erkennt man analog in den ersten ca. 70 und in den letzten ca. 130 Pixeln keine Objektkanten. Die beiden Phasenquerschnitte $Quer1$ und $Quer2$ sollen einen Eindruck vermitteln, wie stark das Rauschen im Vergleich zu den Objektkanten ist. Außerdem erkennt man, dass sich die Phase innerhalb weniger Pixel stark ändern kann.

Um die Periode des Phasenrauschens zu bestimmen, werden die ersten ca. 70 Pixel des linken Bildes der Abbildung 5.6 näher betrachtet, weil dort keine Objektkanten auftauchen. Die Minima treten bei den Pixeln 21, 36, 55 und 73 auf, das entspricht einer Periode von ca. 15 bis 19 Pixeln. Zu beachten ist außerdem, dass sich die Phase nach dem Erreichen eines Minimums bzw. eines Maximums schnell ändert. Zum Beispiel hat Pixel 21 den Phasenwert -2.65 und Pixel 22 bereits den Wert $+0.38$. Entsprechendes gilt für die anderen Minima und Maxima.

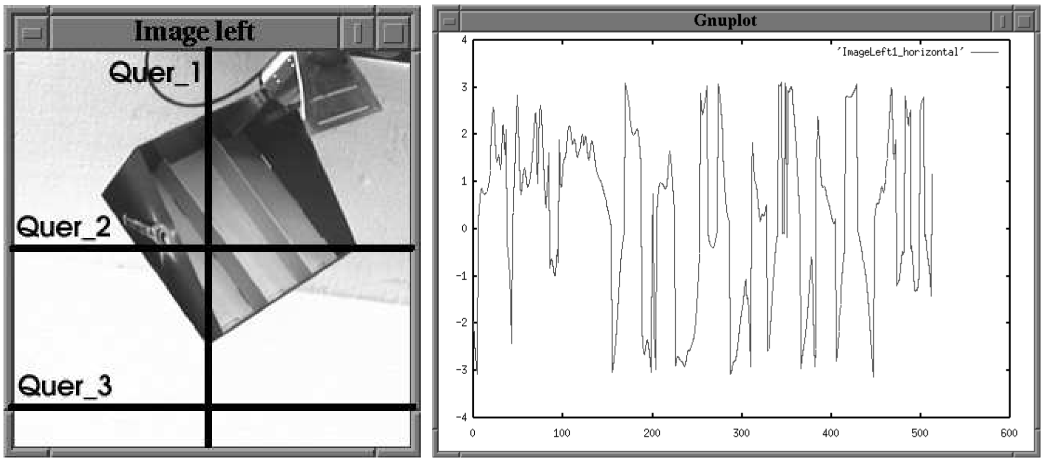


Abbildung 5.5: Links ein Bild von der Schaufel mit drei untersuchten Querschnitten. Rechts der Phasenverlauf durch den horizontalen Querschnitt $Quer2$. Auf der Abzisse sind die 512 Pixel des Querschnittes aufgetragen und auf der Ordinate steht die Phase in Radiant.

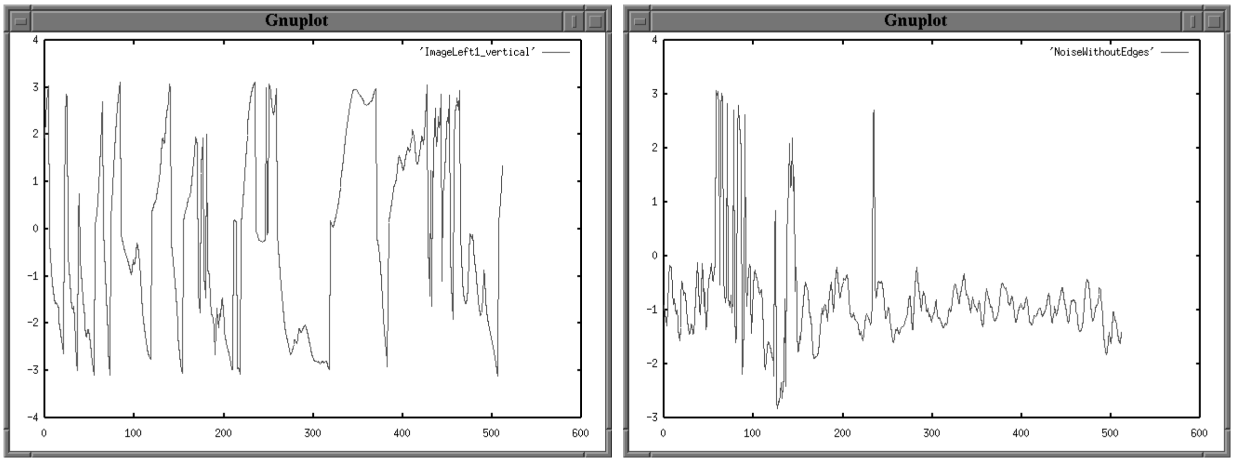


Abbildung 5.6: Links der Phasenverlauf durch den vertikalen Querschnitt $Quer1$ und rechts das Phasenrauschen im horizontalen Querschnitt $Quer3$. Auf der Abzisse sind die 512 Pixel der Querschnitte und auf der Ordinate ist die Phase in Radiant aufgetragen.

Das rechte Bild in der Abbildung 5.6 zeigt den horizontalen Querschnitt $Quer3$. In diesem Querschnitt, der horizontal am unteren Bildrand verläuft, existieren keine Objektkanten, daher enthält diese Abbildung nur das Phasenrauschen. Die Störungen im Phasenrauschen werden größer, wenn der Querschnitt $Quer3$ parallel zur Abzisse nach oben verlagert wird, weil sich dann der Einfluss der Objektkanten bemerkbar macht, auch wenn die Kanten noch nicht durch den Querschnitt getroffen werden.

Analog zur Phasenuntersuchung bei der Monosequenz der linken Stereobilder der Schaufel erfolgt dasselbe Verfahren bei den Aufnahmen vom Karton, wobei hier zu beachten ist, dass auf der Kartonoberfläche zusätzlich Linien vorhanden sind, wie in Abbildung 5.7 zu erkennen ist. Ausserdem liegen die betrachteten Punkte nicht auf den Objektberandungskanten, so dass der Hintergrund auf die Phase keinen Einfluss hat.

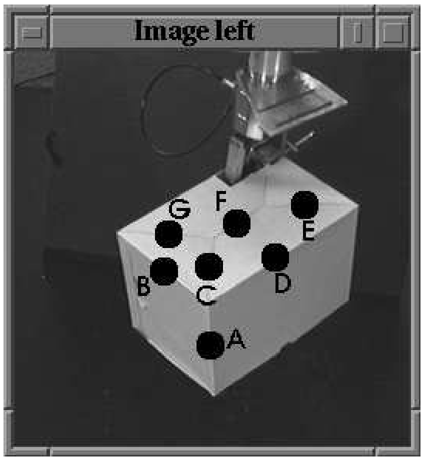


Abbildung 5.7: Bei der Monosequenz der linken Stereokamerabilder des Kartons werden die Punkte A bis G betrachtet.

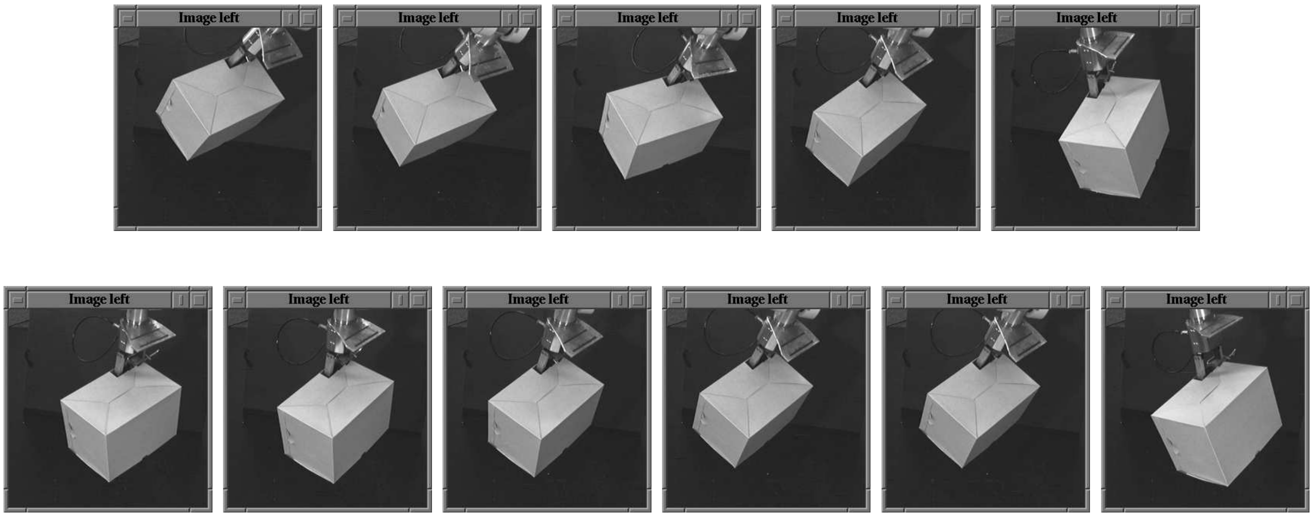


Abbildung 5.8: Monosequenz der linken Stereokamerabilder vom Karton : Bild 1 bis 5 in der oberen Reihe und Bild 6 bis 11 in der unteren Reihe.

Die folgende Tabelle enthält die Phasenwerte in Radiant bei den Punkten A bis G für die Monosequenz in Abbildung 5.8.

Bildnr.	Punkt A	Punkt B	Punkt C	Punkt D	Punkt E	Punkt F	Punkt G
1	2.31	0.45	-0.31	2.29	-0.26	-0.46	-0.17
2	2.22	0.37	-0.14	2.34	-0.29	-0.44	2.88
3	2.15	0.41	-0.41	2.19	-0.15	-0.32	2.81
4	2.04	0.50	-0.14	2.23	-0.31	-0.51	3.05
5	-2.07	1.61	-0.26	2.06	-1.53	-0.43	-0.51
6	-1.26	0.46	-0.55	2.02	-0.11	-0.25	2.87
7	-1.33	0.55	-0.38	2.11	-0.41	-0.53	-0.15
8	2.09	0.67	-0.08	2.19	-0.49	-0.48	-0.34
9	2.10	0.39	-0.47	2.12	-0.47	-0.22	3.00
10	2.03	0.31	-0.27	2.25	-0.10	-0.14	-0.02
11	-2.41	1.48	$\otimes$	2.31	$\otimes$	-0.51	$\otimes$

Die Phasenwerte der Punkte A bis G können folgendermaßen interpretiert werden :

- Der Punkt A liegt auf einer Kante. Weil die Kartonseitenflächen je nach Orientierung zur Lichtquelle in verschiedenen Grautönen erscheinen, ist die Phase nicht konstant. Hinzu kommt, dass die Kante selber viel Licht reflektiert und dadurch teilweise heller als die Seitenflächen ist.

- Der Punkt B hat abgesehen von den Bildnummern 5 und 11 eine konstante Phase. Bei diesen beiden Aufnahmen, bei denen der Karton eine vergleichbare Position hat, ändert sich die Beleuchtung. Der Punkt B liegt auf einer Kante, die sehr viel Licht reflektiert, daher kann man schon fast von einer weißen Linie sprechen. Eine ideale weiße Linie auf schwarzem Hintergrund hat die Phase $\varphi = 0$. Weil hier keine Idealbedingungen vorliegen, beträgt die Phase $\varphi \approx 0.4$.

- Die Punkte C und D haben eine nahezu konstante Phase.

- Der Linienpunkt E hat abgesehen von Bildnr. 5 eine unveränderte Phase. Bei Bildnr. 5 ist ein Schatten vom Greifer auf der Linie sichtbar und bei Bildnr. 11 ist die Beleuchtung zu hell. Daher ist dort die Linie nicht erkennbar.

- Der Linienpunkt F hat auch eine konstante Phase. Die Punkte C, E und F liegen auf der Kartonoberfläche und haben eine Phase $\varphi \approx -0.3$.

- Der Linienpunkt G liegt auch auf der Kartonoberfläche. Bei Bildnr. 11 reflektiert die Kartonoberfläche sehr viel Licht. Die Phase ist jedoch nicht konstant, was u.a. auch auf die Ableseungenauigkeit der Pixelkoordinaten des teilweise schwach erkennbaren Punktes G zurückzuführen ist.

Zusammenfassend kann auch hier festgestellt werden, dass die 2D Phasenkonsistenz im Wesentlichen gewährleistet ist. Störende Einflüsse wie z.B. erhöhte Lichtreflexion und Änderung des Hintergrundes aufgrund der Objektdrehung erschweren die Interpretation der Phasenwerte etwas. Im Vergleich zu den Bildern der Schaufel wurden hier beim Karton auch Linien betrachtet.

5.2 3D Phasenkonsistenz

Aus den assoziierten Phasen von 2D Punkten, die nach der in Kapitel 3 behandelten Ähnlichkeitsfunktion korrespondieren, wird durch eine Mittelung der 2D Phasen eine assoziierte Phase für den zugehörigen 3D Punkt ermittelt, um die strukturelle Information in einer 3D Repräsentation zu kodieren. Bisher war nur die 3D Orientierung aus den 2D Orientierungen als geometrische Information in einer 3D Repräsentation vorhanden. Die Berechnung der Phase φ_{3D} der 3D Liniensegmente basiert auf den beiden Orientierungen $\theta_{2D,left}$ bzw. $\theta_{2D,right}$ und auf den Phasen $\varphi_{2D,left}$ bzw. $\varphi_{2D,right}$ der korrespondierenden 2D Liniensegmente. Im Folgenden ist $\Delta\theta = |\theta_{2D,left} - \theta_{2D,right}|$ die Orientierungsdifferenz. Wenn zwei korrespondierende 2D Punkte gefunden wurden, ergibt sich die 3D Phase des zugehörigen 3D Segmentes durch :

$$\varphi_{3D} = \begin{cases} \operatorname{atan2}[\sin(\varphi_{2D,left}) + \sin(\varphi_{2D,right}), \cos(\varphi_{2D,left}) + \cos(\varphi_{2D,right})] & \text{falls} \quad \Delta\theta \leq \frac{\pi}{2} \\ \operatorname{atan2}[\sin(\varphi_{2D,left}) - \sin(\varphi_{2D,right}), \cos(\varphi_{2D,left}) + \cos(\varphi_{2D,right})] & \text{falls} \quad \Delta\theta > \frac{\pi}{2} \end{cases}$$

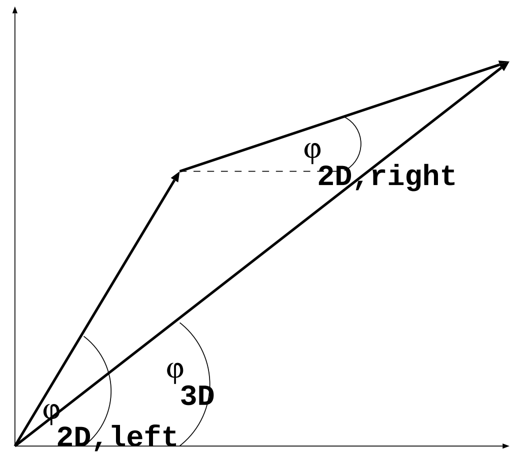


Abbildung 5.9: Erzeugung der Phase für 3D Liniensegmentmittelpunkte durch Vektoraddition

Die Phase für 3D Segmente wird durch die Vektoraddition der 2D Phasen im linken Bild und rechten Bild erzeugt. Abbildung 5.9 verdeutlicht den Sachverhalt für $\Delta\theta \leq \frac{\pi}{2}$. Falls $\Delta\theta > \frac{\pi}{2}$, dann steht zwischen den Sinustermen in der obigen Gleichung für φ_{3D} ein Minuszeichen.

Jede Bildrepräsentation wird durch eine starre Körperbewegung auf die Ursprungsansicht zurückgedreht, um einen 3D Liniensegmentvergleich durchzuführen, wie es in Kapitel 4 dargestellt wurde. Bei diesem Vergleich werden die Phasendifferenzen zwischen jedem 3D Segment in der aktuellen Bildrepräsentation und dem zugehörigen 3D Segment in der Objektrepräsentation festgestellt. Die so erhaltenen 3D Phasendifferenzen aller 3D Segmente werden in einem Histogramm in Abbildung 5.10 visualisiert. Man erkennt, dass die Phase im 3D Fall sehr stabil ist, weil Phasenabweichungen nur bis ca. 0.3 Radiant nennenswert sind. In den Histogrammen bleibt die Phase für 3D Segmente im Rahmen einer Gaußverteilung erhalten.

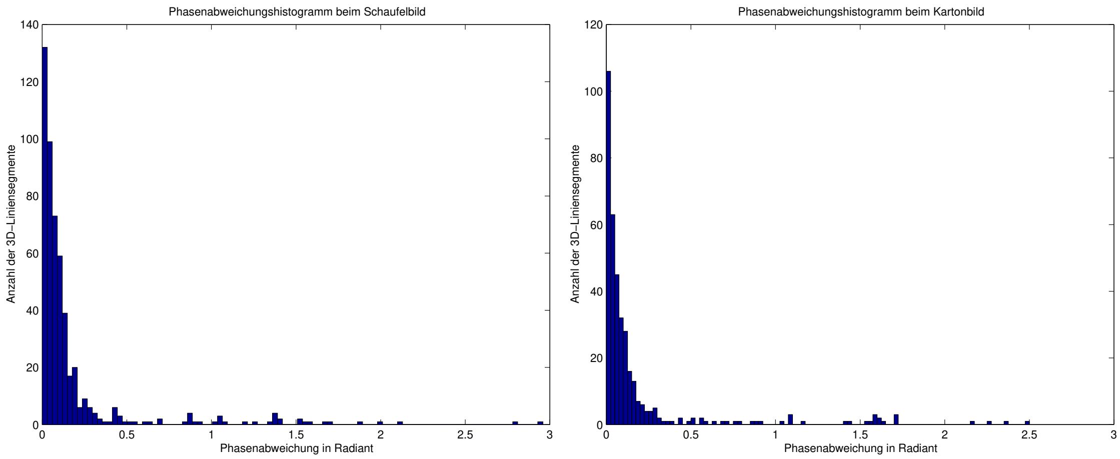


Abbildung 5.10: Histogramme der Phasendifferenzen aller 3D Segmente zwischen einer Bildrepräsentation und einer Objektrepräsentation. Das linke Histogramm bezieht sich auf das Bild der Schaufel beim Aktualisieren der Objektrepräsentation, das rechte Histogramm analog auf den Karton. Auf der Ordinate steht die Anzahl der 3D Segmente und auf der Abzisse die 3D Phasenabweichung in Radiant.

Beim Aktualisieren der Objektrepräsentation gilt die folgende Phasengleichung für jedes 3D Segment. Diese Gleichung hat dieselbe Gestalt wie die Aktualisierungsgleichungen für die Konfidenz und andere Parameter.

$$\varphi_{OR} := \frac{\gamma \cdot \varphi_{OR} + \varphi_{IR}}{1 + \gamma} \tag{5.1}$$

Der Index OR steht für Objektrepräsentation und IR steht für Bildrepräsentation. Der Parameter γ ist ein Gewicht, das angibt, wie häufig das zu betrachtende 3D Segment schon aktualisiert wurde. Bei jedem Aktualisierungsschritt des 3D Segmentes erhöht sich der Wert von γ um eins.

Ein Problem bereitet die Phasenuntersuchung bei der Verschmelzung der 3D Segmente nach dem letzten Akkumulationsschritt. Beim Verschmelzen wird aus einer Liste von 3D Segmenten, deren Orientierungsdifferenzen und deren 3D Abstände vorgegebene Schwellwerte unterschreiten, ein neues 3D Segment erzeugt, wie es in Kapitel 4 beschrieben wurde. Als Lösung dieses Problems ist folgende Gleichung implementiert worden, die besagt, dass sich die Phase $\tilde{\varphi}$ des verschmolzenen Segmentes aus den konfidenzgewichteten Phasenwerten der zu verschmelzenden Segmente ergibt.

$$\tilde{\varphi} = \frac{\sum_{i=1}^{N} \varphi_i c_i}{\sum_{i=1}^{N} c_i} \tag{5.2}$$

Dabei ist N die Anzahl der zu verschmelzenden Segmente und c_i die Konfidenz des i-ten Segmentes im Verschmelzungsprozess.

5.3 Aspektgraph

Die in Kapitel 4 behandelten 3D Objektrepräsentationen wurden durch die Akkumulation von 3D Bildrepräsentationen erzeugt. Dabei wurde stets eine starre Körperbewegung der aktuellen 3D Bildrepräsentation auf die Ursprungsansicht durchgeführt, um einen 3D Liniensegmentvergleich zwischen den beiden 3D Repräsentationen zu ermöglichen. Wenn eine vorliegende Stereobildsequenz eine vollständige Drehung des Objektes beinhaltet und nicht nur eine geringfügige Drehung wie z.B. in den behandelten Monosequenzen in Kapitel 5, dann bereitet eine starre Körperbewegung auf die Ursprungsansicht Probleme. Das Ziel ist es, eine Stereobildsequenz, die eine vollständige Objektdrehung beinhaltet, zu zerlegen. Jedes Stereobildpaar gehört zu einer Klasse, innerhalb der sich die zugehörigen 3D Bildrepräsentationen durch eine starre Körperbewegung mit einer 3D Objektrepräsentation vergleichen lassen. Man möchte mit möglichst wenigen Klassen, d.h. Ansichten, eine vollständige 3D Beschreibung von Objekten finden. Es wird dabei nicht die Geonenstruktur von Biedermann behandelt, der mit 36 geometrischen Grundbausteinen alle denkbaren Objekte erzeugt. Stattdessen wird auf die Aspektgrapherzeugung eingegangen.

Ein Aspektgraph besteht aus einer endlichen Anzahl von charakteristischen Ansichten eines Objektes, den **Standardansichten**, die in der Literatur auch mit **Aspekten** bezeichnet werden. Eine Standardansicht eines Objektes hat die Eigenschaft, dass eine größtmögliche Drehung ohne eine Kantenverdeckung durchführbar ist. Ein Aspektgraph wird aus einer Minimalanzahl von Ansichten erzeugt, die nötig sind, um ein 3D Objekt vollständig zu beschreiben. Ein Aspektgraph enthält eine Menge von Standardansichten und deren Beziehung zueinander durch Rotation des Objektes. Im Aspektgraph bilden die Standardansichten die Knoten. An den Kanten im Graphen stehen jeweils Tripel aus Rotationswinkeln, um von einer Standardansicht zur nächsten zu kommen, d.h. an jedem Knoten stehen drei Winkelangaben (μ_x, μ_y, μ_z) für die Blickrichtung und an jeder Kante stehen die drei Winkelangaben ($\Delta\mu_x$, $\Delta\mu_y$, $\Delta\mu_z$) für die jeweilige 3D Drehung.

Bei der Betrachtung eines Objektes aus verschiedenen Perspektiven muss man entscheiden, wann man eine neue Repräsentation für eine Ansicht braucht, d.h. wann eine neue Standardansicht erzeugt werden muss. Außerdem stellt sich die Frage, wieviele Standardansichten zur vollständigen 3D Beschreibung eines vorgegebenen Objektes mindestens nötig sind. Das Ziel ist es, eine Klassifikation von Stereokamerabildern in Standardansichten durchzuführen und einen Aspektgraphen zu erzeugen. Die einzelnen Aspekte eines Objektes können definiert werden als **Äquivalenzklassen** unter der Äquivalenzrelation der starren Körperbewegung, d.h. innerhalb einer Äquivalenzklasse von Rotations- und Translationsbewegungen tritt keine Kantenverdeckung auf. Die erzeugten Bildrepräsentationen innerhalb einer Äquivalenzklasse zeigen dieselben Kanten aus leicht unterschiedlichen Blickwinkeln. Eine Standardansicht wird aus den Bildrepräsentationen einer Äquivalenzklasse generiert. Bei jeder Objektdrehung mit gleichzeitiger Erzeugung von Stereokamerabildern sind zwei Fälle zu unterscheiden :

- 1. Fall : Die Anwendung der starren Körperbewegung ist erfolgreich, d.h. man befindet sich noch innerhalb derselben Äquivalenzklasse. Dann verbessert man die bisher gewonnene Standardansicht durch das Akkumulationsverfahren für die Objektrepräsentation, wie es im Kapitel 4 beschrieben wurde.
- 2. Fall : Die Anwendung der starren Körperbewegung ist nicht erfolgreich, d.h. man befindet sich nicht mehr in derselben Äquivalenzklasse wie zuvor. Das ist der Fall, wenn aufgrund der Objektdrehung neue Objektkanten erkennbar werden, daher wird eine neue Standardansicht erzeugt.

Es erfolgt somit eine Objektzerlegung in Äquivalenzklassen bzw. Standardansichten. Aus dem Wissen der Bewegung bei der Objektdrehung folgt die Lösung des Korrespondenzproblems zwischen den Bildrepräsentationen. Die Kontrolle über das Objekt ermöglicht eine effiziente aktive Strategie zur Erzeugung und zum Lernen von Ansichten.

Im Folgenden wird ein einfacher **Algorithmus** angegeben, um zu einem vorgegebenen Objekt einen Aspektgraph zu erzeugen, der mindestens zwei Aspekte enthält. Das zu betrachtende Objekt befindet sich im Greifer des Roboterarms und wird gedreht. Die Blickrichtung $\beta \in [0; 2\pi[$ ist variabel und es werden nach jeder Drehung um $\Delta\beta$ neue Stereokamerabilder aufgenommen. Das Objekt befindet sich zu Anfang in der Startposition $\beta_0 = 0$ Grad und der Algorithmus ist beendet, wenn die Ursprungsansicht wiedererkannt wird. Der Zähler für die Anzahl der Standardansichten ist S.

- 1.) Initialisierung : $\beta_0 = 0$, $S = 0$
- 2.) Die Anzahl der Standardansichten wird erhöht : $S = S + 1$
- 3.) Man erzeugt beim Blickwinkel β eine Bildrepräsentation, die als Standardansicht S dient.
- 4.) Objektdrehung, d.h. Winkeländerung : $\beta \leftarrow \beta + \Delta\beta$
- 5.) Es wird wieder eine Bildrepräsentation erzeugt. Man vergleicht sie durch eine starre Körperbewegung mit der zuletzt erzeugten Standardansicht und der ersten Standardansicht ($\beta = 0$), sofern bereits $S \geq 2$ gilt. Falls alle Kanten für $S = 1$ gematched werden können, aber zur Zeit $S \geq 2$ gilt, dann wird der Algorithmus

beendet, weil die Ursprungsansicht wiedererkannt wurde. Falls alle Kanten mit der zuletzt erzeugten Standardansicht gematched werden können, dann wird wieder der Schritt 4 ausgeführt, weil man sich noch innerhalb derselben Äquivalenzklasse befindet. Falls eine Kantenverdeckung bzgl. der zuletzt erzeugten Standardansicht vorliegt, dann wird der Schritt 2 ausgeführt, d.h. die starre Körperbewegung war nicht erfolgreich.

Zu beachten ist bei der 3D Objektdrehung, dass auch die obere und untere Deckfläche von der Stereokamera erfasst werden, d.h. die Objektdrehung umfasst in der Regel mehrere Winkel.

Man kann sich elementare **Beispiele** ohne eine Bildsequenzverarbeitung veranschaulichen. Eine weiße Billiardkugel enthält nur eine Standardansicht, daher würde der Aspektgraph nur einen Knoten enthalten. Ein Quader dagegen hat sechs Flächen, acht Ecken und zwölf Kanten. Er hat höchstens acht Standardansichten, nämlich die Frontalansichten auf die acht Ecken, bei denen man jeweils drei Flächen zugleich sieht.

Ein Tetraeder hat vier Ecken und sechs Kanten. Nach Koenderink und van Doorn [KvD79] und nach Platinga und Dyer [PD90] enthält der zugehörige Aspektgraph vierzehn Knoten, nämlich vier Knoten, bei denen drei Flächen zugleich sichtbar sind, sechs Knoten, bei denen jeweils zwei Flächen zugleich sichtbar sind, und vier Knoten, bei denen jeweils nur eine Fläche sichtbar ist. Bei der Anwendung des Konzeptes der starren Körperbewegung kommt man jedoch mit den zuerst genannten vier Knoten aus. Daher wird hier ein Aspektgraph mit vier Standardansichten vorgestellt, nämlich die jeweiligen Ansichten auf die Ecken, so dass man bei jedem Aspekt drei Flächen zugleich erkennen kann. Die Abbildung 5.11 verdeutlicht die Situation. In der Mitte der Kanten sind die sogenannten Bifurkationsbereiche dick markiert. Dort bereitet das Matching auf eine Standardansicht Probleme, weil man z.B. direkt auf eine Kante des Tetraeders sieht. In den Bifurkationsbereichen reicht eine geringfügige Drehung aus, um in den Bereich einer anderen Standardansicht zu gelangen. Der Übersichtlichkeit wegen sind die Winkelangaben an den Knoten und die Winkeldrehungen an den Kanten weggelassen.

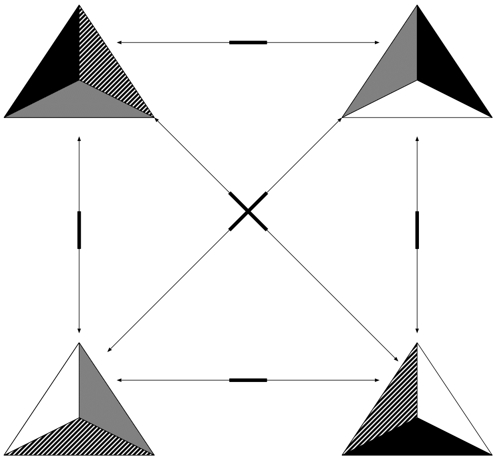


Abbildung 5.11: Aspektgraph eines Tetraeders mit Bifurkationsbereichen

Je höher der Grad der 3D Isotropie eines Objektes ist, desto geringer ist die Anzahl der Standardansichten. Es ist naheliegend, dass komplizierte Objekte mehr Aspekte als eine homogene Kugel erfordern. Die Abbildungen 5.12 und 5.13 zeigen eine alternative Darstellung von Aspektgraphen. Auf der Ordinate und der Abzisse sind die Winkelbereiche in Kugelkoordinaten, die bei der Objektdrehung auftreten, im Intervall $[-\pi; \pi]$ aufgetragen. Der Koordinatenursprung ist der Blick auf den Mittelpunkt des Bereiches einer Standardansicht. Für das jeweils zu betrachtende Objekt sind die Winkelbereiche zu lokalisieren, in denen eine starre Körperbewegung erfolgreich ist, d.h. man kann die Drehbereiche für die Standardansichten direkt eintragen. Dabei ist jedoch die Zyklizität des Winkelparameterraumes zu beachten, weil z.B. eine Drehung um π nach rechts identisch mit einer Drehung und $-\pi$ nach links ist. In Abbildung 5.12 erkennt man den Aspektgraphen einer Münze. Die innere Kreisscheibe ist der Winkelbereich für die Münzvorderseite und der äußere Kreisring ist der Bereich für die Münzrückseite.

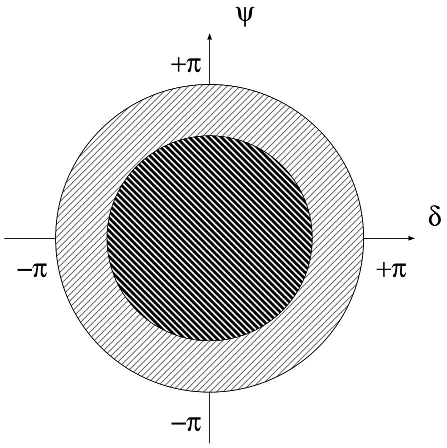


Abbildung 5.12: Aspektgraph einer Münze.

In Abbildung 5.13 erkennt man den prinzipiellen Aufbau von Aspektgraphen würfelähnlicher Objekte. Die innere Kreisscheibe entspricht der Standardansicht $S1$ der Vorderseite, der äußere Kreisring entspricht der Standardansicht $S6$ der Rückseite. Die anderen vier Bereiche bilden die Winkelregionen der Standardansichten $S2$, $S3$, $S4$ und $S5$ für die vier Seitenflächen.

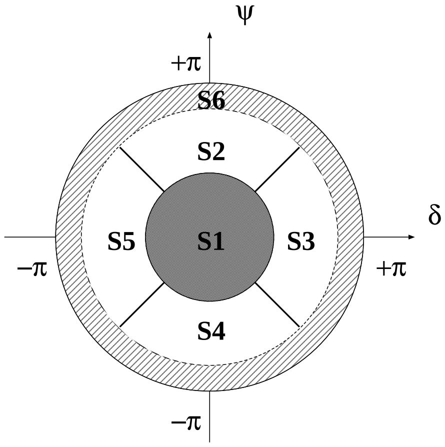


Abbildung 5.13: Aspektgraph eines würfelähnlichen Objektes

Kapitel 6

Zusammenfassung und Ausblick

In dieser Arbeit wurde ein neues Verfahren zur Korrespondenzfindung bei Stereokamerabildern entwickelt. Die Bildvorverarbeitung erfolgte mit dem von Felsberg [Fe00] behandelten Strukturmultivektor, der eine orthogonale Identitätszerlegung des 2D Signals in eine lokale Amplitude, eine lokale Phase und eine lokale Orientierung gestattet. Diese Zerlegung ermöglicht erst ein neues Verfahren für das Stereomatching, indem strukturelle und geometrische Information zusammen benutzt, aber getrennt behandelt werden können. Es wurde die Phasenkongruenz erzeugt, aus deren Maxima die 2D Liniensegmentmittelpunkte generiert wurden. Um zu einem 2D Liniensegmentmittelpunkt im linken Kamerabild den korrespondierenden 2D Punkt auf der Epipolargeraden im rechten Kamerabild zu ermitteln, wurde eine neue Ähnlichkeitsfunktion eingeführt, die die lokale Phasen- und Orientierungsdifferenz zwischen zwei Punkten berücksichtigt. Die Gewichtung zwischen der Phasen- und Orientierungsdifferenz bewirkt eine tiefenabhängige Funktion, die durch statistische Untersuchungen aus den Varianzen der Orientierungs- und Phasendifferenzen bestimmt wurde. Das Ergebnis der Distanzfunktion geht in die neu entwickelte Ähnlichkeitsfunktion ein, die das korrespondierende 2D Punktepaar bestimmt und die Konfidenz des zugehörigen 3D Segmentes berechnet.

Diese Bildvorverarbeitung und Korrespondenzfindung werden in den Algorithmus von Ackermann [Ack00] integriert. Durch die von ihm verwendete global wirkende Houghtransformation mit anschließender Diskretisierung der Houghgeraden in 2D Liniensegmente geht die Lokalität in der Bildanalyse verloren, daher erfolgte die Bildvorverarbeitung mit dem Strukturmultivektor. Das häufig verwendete Template-Matching zur Korrespondenzfindung, bei dem man Grauwertdifferenzen zwischen gleichgroßen Fenstern im linken und rechten Bild untersucht, wurde in der vorliegenden Arbeit nicht benutzt.

Aus korrespondierenden 2D Liniensegmentmittelpunkten mit deren 2D Orientierungen kann man somit ein 3D Segment mit zugehöriger 3D Orientierung rekonstruieren. Die generierten 3D Liniensegmente bilden eine 3D Bildrepräsentation. Durch mehrfache Objektdrehung kann aus den jeweils erzeugten 3D Bildrepräsentationen eine 3D Objektrepräsentation gewonnen werden, in der die Konfidenz der 3D Segmente nach jeder Objektdrehung aktualisiert wird. Nach der Verarbeitung einer längeren Stereobildsequenz erhält man eine robuste 3D Objektrepräsentation, weil hohe Konfidenzen nur bei denjenigen 3D Segmenten auftreten, die in vielen 3D Bildrepräsentationen bestätigt wurden.

Innerhalb von Monosequenzen der linken Stereokamerabilder wurde die 2D Phasenkonsistenz von Kanten- und Linienmittelpunkten bei verschiedenen Objekten untersucht, um festzustellen, ob die 2D Phase unter Rotation und Translation stabil bleibt. Abgesehen von den Objektberandungskanten, deren Phase hintergrundsabhängig ist, wurde die Stabilität der 2D Phase bei mehrfacher Objektbewegung bestätigt. Außerdem wurde die 2D Phasenabweichung zwischen Punkten, die nach der neu entwickelten Ähnlichkeitsfunktion korrespondieren, im linken und rechten Kamerabild untersucht, was zu einer Gaußverteilung führte. Um in einer 3D Repräsentation auch die strukturelle Information zu kodieren, wurde aus den 2D Phasen korrespondierender 2D Punkte eine gemittelte Phase des zugehörigen 3D Punktes assoziiert. Die Phasenkonsistenz von 3D Liniensegmenten wurde beim Aktualisieren einer 3D Objektrepräsentation mit einer 3D Bildrepräsentation überprüft. Auch hier konnte die Phasenkonsistenz bestätigt werden.

Zusammenfassend läßt sich feststellen, dass man durch die Bildvorverarbeitung mit dem Strukturmultivektor und der Einführung einer neuen Ähnlichkeitsfunktion bei der Korrespondenzfindung zu guten Ergebnissen bei 3D Repräsentationen von Objekten gelangen kann. Die Resultate sind genauer und vollständiger als bei Ackermann [Ack00], weil keine Information verloren geht. Bei der Bildvorverarbeitung mit der global wirkenden Houghtransformation, die Ackermann durchgeführt hat, fehlen in den 3D Repräsentationen teilweise Objektkanten oder sie weisen Kantenlücken auf. Sowohl die 3D Bildrepräsentationen als auch die akkumulierte 3D Objektrepräsentation nach der Verarbeitung von kompletten Stereobildsequenzen enthalten bei den getesteten Objekten lokal mehr Informationen als bei Ackermann.

Das in dieser Arbeit behandelte Verfahren zur Bildvorverarbeitung und zur Erzeugung von 3D Repräsentationen läßt sich auf vielfältige Art verbessern. Die Rauschunterdrückung ist noch nicht optimal, weil die Objektkanten in Grauwertbildern mit ungünstigen lokalen Signal-Rauschverhältnissen noch nicht hinreichend gut detektiert werden. Weiterhin wäre die Elimination des Greifers vom Roboterarm bei den im Labor erzeugten Stereokamerabildern sinnvoll, damit bei der Rekonstruktion und bei der Korrespondenzfindung nur das zu untersuchende Objekt betrachtet wird. Die Benutzung von 3D Repräsentationen könnte für Erkennungs- und Trackingaufgaben von Bedeutung sein. Das Lernen von Parametern wie z.B. der Konfidenzschwellwerte bei der Korrespondenzfindung ist ebenso ein weiteres Aufgabenfeld, daher ist ein gleitender Übergang zu bestimmten Charakteristika neuronaler Netze denkbar. Nach Seibert und Waxman [SW92] ist auch die Aspektklassifikation eines vorgegebenen Objektes mit neuronalen Netzen ein interessanter Gesichtspunkt zum Lernen und Erkennen von Objekten.

Kapitel 7

Anhang

7.1 Symbolverzeichnis

atan2	spezielle Arcustangensfunktion	26
b	relative Bandbreite	10
$B(\|\mathbf{u}\|)$	radialer Bandpassfilter	10f.
c	Konfidenz eines Liniensegmentes	42f.
$conf_{2D}$	Konfidenz eines 2D Liniensegmentes	29
$conf_{3D}$	Konfidenz eines 3D Liniensegmentes	29
d	Distanzfunktion	26ff.
E	Energie	17
$\mathcal{F}$	Fast Fouriertransformation. Inverse Transformation wird mit $\mathcal{F}^{-1}$ notiert	9ff.
F	Bild im Frequenzraum	10
f	Brennweite bei der Projektionsgeometrie	30f.
f	Grauwertbild im Ortsraum	9ff.
F_1	quadratischer Filter bestehend aus Einsen	18
f_i	i-Komponente der sphärischen Quadraturfilterantwort	11ff.
f_j	j-Komponente der sphärischen Quadraturfilterantwort	11ff.
$f(x)$	Realteil des analytischen Signals	7
$f_A(x)$	analytisches Signal	7
$f_H(x)$	Imaginärteil des analytischen Signals	7
$F(\mathbf{u})$	Fouriertransformierte von $f(\mathbf{x})$	9
$F_M(\mathbf{u})$	monogenes Signal	9
f_r	Realteil der sphärischen Quadraturfilterantwort	11ff.
$F_R(\mathbf{u})$	Riesztransformierte	9
$G(t)$	Roboterstellung zur Zeit t	51
H	Koordinatenursprung in der 2D Projektionsebene	30
$\mathbb{H}$	Menge der Quaternionen	7
h	Operator mit Komponenten h_r, h_i und h_j	9
i	imaginäre Einheit im Komplexen bzw. Quaternionischen	7ff.
Img_θ	komplexes Bild mit Realteil f_j und Imaginärteil f_i	18

$Img_{\theta,M}$	komplexes Bild zur genaueren Bestimmung der Orientierung	18
$\widehat{Img_\theta}$	gefaltetes komplexes Bild, um die Orientierung genauer zu bestimmen	18
$\widehat{Img_{\theta,imag}}$	Imaginärteil von $\widehat{Img_\theta}$	18
$\widehat{Img_{\theta,real}}$	Realteil von $\widehat{Img_\theta}$	18
IR	image representation (Bildrepräsentation)	51
j	imaginäre Einheit im Quaternionischen	7ff.
k	imaginäre Einheit im Quaternionischen	7
k_u, k_v	horizontale und vertikale Größe der Pixel im Meter	30
l	Länge eines Liniensegmentes	42
M	lokale Amplitude	13
n	Normalenvektor $(n_x, n_y)^T$ zu einer 2D Kante	36f.
N	Rauschenergie	16f.
OR	Objektrepräsentation	51
P	Projektionsmatrix	31f.
q	Ortsvektor $(q_x, q_y)^T$ zu einem beliebigen Punkt Q in der 2D Ebene	36f.
$\mathbb{R}$	Menge der reellen Zahlen	7f.
R	3×3 Rotationsmatrix für die Projektionsmatrix	31
s	Bildgröße in Pixel	16
S	Standardansicht im Aspektgraph	66ff.
SMV	Strukturmultivektor	8, 16f.
SQF	sphärischer Quadraturfilter	8
$\vec{t}$	Translationsvektor für die Projektionsmatrix mit Komponenten t_x, t_y, t_z	31ff.
T	entrauschtes Skalarprodukt der Strukturmultivektoren	17
u, U	Koordinate entlang der Abzisse in einer 2D Ebene	9ff., 31ff.
u_0	u-Koordinate, bei der die optische Achse die Bildebene schneidet	30f.
v, V	Koordinate entlang der Ordinate in einer 2D Ebene	9ff., 31ff.
v_0	v-Koordinate, bei der die optische Achse die Bildebene schneidet	30f.
x, y, z	x, y, z-Koordinate im kartesischen Koordinatensystem	30ff.
$\alpha(z)$	dimensionslose Funktion für Gewichtung zw. Phase und Orientierung	26ff.
β	Winkel im Algorithmus zur Aspektgrapherzeugung	66
γ	Gewicht für das Aktualisieren der Phase	64
ϵ	Regularisierungsparameter	17
θ	lokale Orientierung	13, 25ff.
κ	Abstand der optischen Zentren	45
λ	Disparität	45
μ	Winkel für eine Standardansicht im Aspektgraphen	65
$\sigma_\varphi, \sigma_\theta$	Standardabweichungen für Phase und Orientierung	34, 39
ϕ	Phasenkongruenz	17
φ	lokale Phase	13, 25ff.
ψ_1, ψ_2	Kugelkoordinatenwinkel	42f.
ξ_u, ξ_v	Brennweite in Pixeln als Elemente der Projektionsmatrix	31
Δ^2	Mahalanobisdistanz	38f.
Π_{lower}	unteres Band des Filters, analog Π_{upper} oberes Band	16f.
Ω	Oktave des Bandpassfilters	10

7.2 Implementation

Der Code wurde in der Programmiersprache C++ im Softwarepaket KiVi des Lehrstuhls für Kognitive Systeme geschrieben. Im Folgenden werden die Datei-, Prozedur- und Funktionsnamen mit deren Semantik kurz dargestellt.

Datei aspect.cpp : Hauptprogramm bzw. Demonstrationsprogramm

- Prozedur Main :
  Der Benutzer kann sich eine von vier Bildserien der Stereokamera auswählen, wobei jede Sequenz aus 11 Stereobildpaaren besteht. Zu den auswählbaren Objekten gehören ein Karton, ein Ordner, ein Schaumstoffteil und eine Schaufel. Neben den Prozeduraufrufen zur Bildvorverarbeitung mit dem Strukturmultivektor werden auch die grafischen Resultate angezeigt.

- Prozedur InitialRepresentation :
  Die erste Bildrepräsentation wird erzeugt, die zugleich die initiale Objektrepräsentation darstellt. Es wird die Prozedur ComputeWithoutHough aufgerufen.

- Prozedur UpdateOrRep :
  Nach dem Einlesen jedes weiteren Stereobildpaares wird die bisher gewonnene Objektrepräsentation aktualisiert. Dabei wird auch hier jedesmal die Prozedur ComputeWithoutHough aufgerufen.

- Prozedur Merge :
  Wenn das letzte Stereobildpaar verarbeitet wurde, werden benachbarte 3D Segmente in der akkumulierten Objektrepräsentation verschmolzen.

Datei imageRepresentation.cpp / .h : Korrespondenzfindung

- Prozedur ComputeWithoutHough :
  Es erfolgen Vorbereitungen wie z.B. das Einlesen der Kalibrationsmatrizen und die Initialisierung der Epipolargeometrie. Dann erfolgen die Prozeduraufrufe GetBestMatchWithoutHough und Create3DSegmentsWithoutHough.

- Prozedur GetBestMatchWithoutHough :
  Eine Überprüfung aller Pixel auf der Epipolargeraden bzgl. der vorgegebenen maximalen Disparität und der minimalen Konfidenz etc. wird durchgeführt. Entscheidend bei der Korrespondenzfindung ist der Prozeduraufruf StructureMultiVectorDistance.

- Prozedur Create3DSegmentsWithoutHough :
  Aus einem korrespondierenden 2D Punktepaar wird das zugehörige 3D Liniensegment mit Orientierung und Konfidenz etc. erzeugt.

Datei KMultiVecLib.cpp / .h : Funktionsbibliothek zum Strukturmultivektor

- Funktion EdgeDetectionByPhaseCongruence :
  Die Berechnung der sphärischen Quadraturfilterantworten wird mit Hilfe der Prozeduraufrufe CreateBandpass und UseFrequencyFilter durchgeführt. Danach werden für das obere und untere Band die Amplitude, die Phase und die Orientierung ermittelt. Zum Schluss wird die Phasenkongruenz erzeugt und die Prozedur HysteresisThreshold wird aufgerufen.

- Funktion HysteresisThreshold :
  Die Kantendetektion nach dem Algorithmus von Kovesi wird berechnet.

- Funktion LineSegment :
  Das Einzeichnen der 2D Liniensegmente für das linke Bild wird vollzogen.

- Funktion StructureMultiVectorDistance :
  Diese Funktion liefert zu einem vorgegebenen 2D Liniensegmentmittelpunkt im linken Bild den korrespondierenden 2D Punkt auf der Epipolargeraden im rechten Bild mittels der Distanz- und Ähnlichkeitsfunktion.

Datei KMultiVecImageProc.cpp / .h : Klasse KMultiVecImageProc

- Prozedur CreateBandpass :
  Es wird ein Bandpass erzeugt, der mit $(1 + \frac{u}{|\mathbf{u}|})$ bzw. mit $\frac{v}{|\mathbf{u}|}$ multipliziert wird.

- Funktionen Absolut, Phi und Theta :
  Die Amplitude M, die Phase φ und die Orientierung θ werden berechnet.

Datei KImageLibFFT.cpp / .h : Funktions- und Prozedurbibliothek zur FFT

- Prozedur UseFrequencyFilter :
  Die Fast Fourier Transformation (FFT), die Filterung mit dem Bandpass im Frequenzraum und die inverse Fouriertransformation werden berechnet. Für nicht quadratische Bilder wird zuvor eine Einbettung in die nächstgrößere Zweierpotenz durchgeführt.

Datei KCreateTestImageLib.cpp / .h : Funktionsbibliothek für synthetische Bilder

- Funktionen Circle, CircleFilled, Siemens, Rectangle, Quader und FilledQuader :
  Es werden diverse Testbilder wie z.B. Kreise, Rechtecke und Quader erzeugt.

- NotPow2Img :
  Ein nicht quadratisches Bild wird erzeugt, um die Bildeinbettung in eine Bildgröße, die der nächsthöheren Zweierpotenz entspricht, für die schnelle Fouriertransformation zu testen.

7.3 Hilfsprogramm zur Orientierungsabweichung

Datei : OrientHisto.m
Inhalt : Die Implementierung der Orientierungsabweichung zwischen 2D Projektionen wird im Matlab-Code dargestellt. Es werden durch einen Zufallsgenerator 3D Segmente erzeugt, die durch eine Stereokamera betrachtet werden. Die Winkelabweichungen zwischen den zugehörigen 2D Segmenten in den Projektionsebenen werden in einem Histogramm ausgegeben.

Parameter :

orientdiff	Vektor mit den Orientierungsabweichungen für Histogrammausgabe
orientstep	Maximalanzahl der Balken im Histogramm
segmente	Anzahl der betrachteten 3D Testsegmente
tY	y-Komponente des Translationsvektors
tZ	z-Komponente des Translationsvektors
xcoo2Dleftsegstart	x-Koordinate des 2D Segmentstartpunktes im linken Bild
xcoo2Drightsegstart	x-Koordinate des 2D Segmentstartpunktes im rechten Bild
xcoo3Dsegend	x-Koordinate des 3D Segmentendpunktes
xcoo3Dsegstart	x-Koordinate des 3D Segmentanfangspunktes
ycoo2Dleftsegend	y-Koordinate des 2D Segmentendpunktes im linken Bild
ycoo2Drightsegend	y-Koordinate des 2D Segmentendpunktes im rechten Bild
ycoo3Dsegend	y-Koordinate des 3D Segmentendpunktes
ycoo3Dsegstart	y-Koordinate des 3D Segmentanfangspunktes
zcoo3Dsegend	z-Koordinate des 3D Segmentendpunktes
zcoo3Dsegstart	z-Koordinate des 3D Segmentanfangspunktes

```
function orientdiff = OrientHisto();
```

Parameter werden gesetzt.

```
segmente = 1000; tY = 12.5; tZ = 150; orientdiff = []; random1 = -50;
random2 = 50;

for k=1:segmente;
```

Durch einen Zufallsgenerator werden der Anfangspunkt und Endpunkt des 3D Segmentes erzeugt.

```
xcoo3Dsegstart = unifrnd(random1,random2);
ycoo3Dsegstart = unifrnd(random1,random2);
zcoo3Dsegstart = unifrnd(random1,random2);
xcoo3Dsegend = unifrnd(random1,random2);
ycoo3Dsegend = unifrnd(random1,random2);
zcoo3Dsegend = unifrnd(random1,random2);
```

Die Projektionen des 3D Segmentes werden berechnet, d.h. die Koordinaten des projizierten Segmentanfangspunktes und des Endpunktes werden ermittelt.
```
ucoo2Dleftsegstart = xcoo3Dsegstart / ( zcoo3Dsegstart + tZ );
ucoo2Dleftsegend = xcoo3Dsegend / ( zcoo3Dsegend + tZ );
ucoo2Drightsegstart = xcoo3Dsegstart / ( zcoo3Dsegstart + tZ );
ucoo2Drightsegend = xcoo3Dsegend / ( zcoo3Dsegend + tZ );

vcoo2Dleftsegstart = ( ycoo3Dsegstart - tY ) / ( zcoo3Dsegstart + tZ );
vcoo2Dleftsegend = ( ycoo3Dsegend - tY ) / ( zcoo3Dsegend + tZ );
vcoo2Drightsegstart = ( ycoo3Dsegstart + tY ) / ( zcoo3Dsegstart + tZ );
vcoo2Drightsegend = ( ycoo3Dsegend + tY ) / ( zcoo3Dsegend + tZ );
```

Die Differenz der Koordinaten wird gebildet, um die 2D Segmentlänge zu ermitteln.
```
urightdiff = ucoo2Drightsegstart - ucoo2Drightsegend;
uleftdiff = ucoo2Dleftsegstart - ucoo2Dleftsegend;
vrightdiff = vcoo2Drightsegstart - vcoo2Drightsegend;
vleftdiff = vcoo2Dleftsegstart - vcoo2Dleftsegend;

if urightdiff ~= 0 & uleftdiff ~= 0
```

Die Orientierung der Projektionen wird berechnet.
```
zr = urightdiff + i*vrightdiff;
zr = zr / abs(zr);
zl = uleftdiff + i*vleftdiff;
zl = zl / abs(zl);
```

Die Differenz der Orientierungen wird ermittelt.
```
orientsinglediff = angle(zr/zl);
orientdiff = [orientdiff orientsinglediff];
end
end,
```

Das Histogramm mit den Winkelabweichungen wird ausgegeben.
```
orientstep = 100;
hist(orientdiff,orientstep);
title('Orientierungsabweichungshistogramm bei 1000 Liniensegmenten');
xlabel('Orientierungsabweichung in Radiant zwischen den Projektionen im
linken und rechten Bild');
ylabel('Anzahl der Liniensegmente');
```

Die Ausgabe des Mittelwertes und der Standardabweichung erfolgt.
```
mittel=mean(orientdiff)
stabw=std(orientdiff)
```

7.4 Hilfsprogramm zur Phasenabweichung

Datei : PhaseHisto.m
Inhalt : Die Implementierung der Phasenabweichung bei zufällig erzeugten 2D Schwarz-weiß-Kanten wird im Matlab-Code dargestellt.

Parameter :

AnzNormVec	Anzahl der zufällig erzeugten Normalenvektoren
deltaX	x-Komponente des Normalenvektors
deltaY	y-Komponente des Normalenvektors
PhaseDiffVec	Vektor für die Histogrammausgabe
PhaseImage	Phasenbild des Strukturmultivektors
PhaseStep	Maximalanzahl der Balken im Histogramm
theta	zufällige Orientierung des Normalenvektors
xCoo	x-Komponente eines zu überprüfenden Pixels
yCoo	y-Komponente eines zu überprüfenden Pixels

```
function phasediff = PhaseHisto();

AnzNormVec = 20; PhaseDiffVec = [];
```

Über die zufällig erzeugten Normalenvektoren wird eine Schleife gebildet.

```
for k=1:AnzNormVec;
```

Der Normalenvektor wird erzeugt. Die Größen Δx, Δy sind reell und können auch negativ sein.

```
theta = unifrnd(0,2*pi);
deltaX = cos(theta);
deltaY = sin(theta);
if deltaY == 0
deltaY = 0.005;
end
```

Für jedes Pixel wird bestimmt, ob es schwarz oder weiß ist.

```
[xCoo,yCoo] = meshgrid(-256:255,-256:255);
SkalarProd = xCoo*deltaX + yCoo*deltaY;
SkalarProd = min(SkalarProd,1);
SkalarProd = max(SkalarProd,-1);
```

Das Phasenbild des Strukturmultivektors wird erzeugt.

```
V1=vectorarray(SkalarProd,4);
PhaseImage = angle(V1);
```

Die Phasendifferenz wird berechnet. Für die Kantenpixel gilt : $x \cdot \Delta x + y \cdot \Delta y = 0$. Dabei ist die x-Koordinate ganzzahlig, aber die y-Koordinate reell, die dann gerundet wird, um das Pixel (x,y) zu bestimmen. Anschließend wird der Phasenwert aus dem Phasenbild des Strukturmultivektors bei Koordinaten (x,y) abgelesen und die Phasendifferenz wird bei den Kantenpixeln berechnet. Die Koordinatenumrechnung von [-256,255] nach [1,512] ist nötig, weil das Phasenbild die Größe [1,512] besitzt.

```
for x=-256:255;
y = -(x*deltaX)/deltaY;
if( y >= -256 & y <= 255 )
y = round(y);
yNew = y + 257;
xNew = x + 257;
PhaseValue = PhaseImage(yNew,xNew);
PhaseDiff1 = (pi/2) - PhaseValue;
PhaseDiff2 = (-pi/2) - PhaseValue;
PhaseSingleDiff = (abs(PhaseDiff1)<abs(PhaseDiff2))*PhaseDiff1
+ (abs(PhaseDiff1)>=abs(PhaseDiff2))*PhaseDiff2;
PhaseDiffVec = [PhaseDiffVec PhaseSingleDiff];
end
end,
end,
```

Die Histogrammausgabe erfolgt.

```
PhaseStep = 100;
hist(PhaseDiffVec,PhaseStep);
xlabel('Phasenabweichung in Radiant');
ylabel('Anzahl der betrachteten Pixel mit Phasenabweichung');
```

Die Ausgabe des Mittelwertes und der Standardabweichung erfolgt.

```
mittel=mean(PhaseDiffVec)
stabw=std(PhaseDiffVec)
```

7.5 Literaturverzeichnis und Index

[Ack00] **Ackermann, M.**
AKKUMULIEREN VON OBJEKTREPRÄSENTATIONEN IM WAHRNEHMUNGS-HANDLUNGS-ZYKLUS
Diplomarbeit am Institut für Informatik der Universität Kiel, 2000

[BB82] **Ballard, D.H., Brown, C.M.**
COMPUTER VISION
Prentice-Hall, Eaglewood Cliffs, 1982

[Fa93] **Faugeras, O.**
THREE-DIMENSIONAL COMPUTER VISION - A GEOMETRIC VIEWPOINT
MIT-Press, Cambridge, 1993

[Fe00] **Felsberg, M.**
STRUCTURE MULTIVECTOR FOR LOCAL ANALYSIS OF IMAGES
Bericht Nr. 2001 des Instituts für Informatik der Universität Kiel, 2000

[FS00a] **Felsberg, M., Sommer, G.**
THE MONOGENIC SIGNAL
Bericht Nr. 2009 des Instituts für Informatik der Universität Kiel, 2000

[FS00b] **Felsberg, M., Sommer, G.**
A NEW EXTENSION OF LINEAR SIGNAL PROCESSING FOR ESTIMATING LOCAL PROPERTIES AND DETECTING FEATURES
DAGM Symposium, Kiel, 2000

[FS00c] **Felsberg, M., Sommer, G.**
THE MULTIDIMENSIONAL ISOTROPIC GENERALIZATION OF QUADRATURE FILTERS IN GEOMETRIC ALGEBRA
AFPAC, Kiel, 2000

[FS01] **Felsberg, M., Sommer, G.**
THE STRUCTURE MULTIVECTOR
Applied Geometrical Algebras in Computer Science And Engineering, 2001

[Ha99] **Hahn, M.**
SEMIGLOBALE VERFAHREN ZUR GENERIERUNG VON ECKPUNKTHYPOTHESEN IN 2D UND 3D
Diplomarbeit am Institut für Informatik der Universität Kiel, 1999

[HS93] **Haralick, R.M., Shapiro, L.G.**
COMPUTER AND ROBOT VISION
Addison-Wesley, Reading, 1993

[Ja97] **Jähne, B.**
DIGITALE BILDVERARBEITUNG
Springer-Verlag, Berlin, 4.Auflage, 1997

[KK96] **Klette, R., Koschan, A., Schlüns, K.**
COMPUTER VISION
Vieweg Technik, 1996

[KvD79] **Koenderink, J.J., van Doorn, A.J.**
THE INTERNAL REPRESENTATION OF SOLID SHAPE WITH RESPECT TO VISION
Biological Cybernetics, 32, 211-216, 1979

[Ko99] **Kovesi, P.**
IMAGE FEATURES FROM PHASE INFORMATION
Videre : Journal of Computer Vision Research, Vol. 1, pp. 891-906, 1999

[Kr00] **Krüger, N., Ackermann, M., Sommer, G.**
ACCUMULATION OF 3D OBJECT REPRESENTATIONS WITHIN A PERCEPTION-ACTION-CYCLE
Conference Paper, EIS 2000

[PD90] **Plantinga, H., Dyer, C.R.**
VISIBILITY, OCCLUSION AND THE ASPECT GRAPH
International Journal of Computer Vision, 5:2, 137-160, 1990

[SW92] **Seibert, M., Waxman, A.M.**
ADAPTIVE 3D OBJECT RECOGNITION FROM MULTIPLE VIEWS
IEEE Transactions on pattern analysis and machine intelligence, Vol. 14, No. 2, 1992

[So99] **Sommer, G.**
THE GLOBAL ALGEBRAIC FRAME OF THE PERCEPTION-ACTION-CYCLE
Jähne et al., Handbook of Computer Vision and Applications, Vol. 3, pp. 221-264, 1999

[SM98] **Sommer, G., Mahlmeister, U.**
COMPUTER VISION I
Skript zur Vorlesung, 1998

[WB99] **Wallis, G., Bülthoff, H.**
LEARNING TO RECOGNIZE OBJECTS
Trends in Cognitive Sciences, Vol. 3, No. 1, 1999

[ZB90] **Zetsche, C., Barth, E.**
FUNDAMENTAL LIMITS OF LINEAR FILTERS IN VISUAL PROCESSING OF 2D SIGNALS
Vision Research, Vol. 30, 1990

Index

Abbildungsverzeichnis

7.6 Danksagung

In erster Linie danke ich Herrn Prof. Dr. Sommer für die Genehmigung, diese Diplomarbeit an seinem Lehrstuhl schreiben zu dürfen. Den Herren Dr. Norbert Krüger und Dipl.-Ing. Michael Felsberg danke ich für die sehr gute Betreuung. Nicht zuletzt bin ich allen Mitarbeitern und Studenten des Instituts für das angenehme Arbeitsklima und für die Hilfsbereitschaft dankbar.

7.7 Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Diplomarbeit selbstständig und nur unter der Anleitung der Herren Prof. Dr. Sommer, Dr. Krüger und Dipl.-Ing. Felsberg angefertigt habe. Als Hilfsmittel dienten die angegebene Literatur und das KiVi-Softwarepaket am Lehrstuhl.

Kiel, den 13. August 2001