

Perzeptuelle Organisation von Objektgrenzen unter Verwendung anisotroper Regularisierungsmethoden

Zur Erlangung des akademischen Grades

DOKTORINGENIEUR (Dr.-Ing.)

von der Fakultät für Elektrotechnik, Informatik und Mathematik
der Universität Paderborn
genehmigte Dissertation
von

Dipl.-Math. Marcus Hund

aus Cuxhaven

Referentin: Prof. Dr.-Ing. Bärbel Mertsching
Korreferent: Prof. Dr.-Ing. Horst-Michael Groß

Tag der mündlichen Prüfung: 27.05.2009

Paderborn 2009

Diss. EIM-E//255

Danksagung

Die vorliegende Arbeit entstand im Rahmen meiner Tätigkeit im Fachgebiet Grundlagen der Elektrotechnik an der Universität Paderborn. An dieser Stelle möchte ich mich bei all denen sehr herzlich bedanken, die mich in vielfältiger Weise bei der Anfertigung der Arbeit unterstützt haben.

Zunächst sei Frau Prof. Dr.-Ing. Bärbel Mertsching genannt, die mir die Bearbeitung des vorliegenden Themas ermöglichte und die durch den mir gewährten kreativen Freiraum einen wesentlichen Grundstein für das Gelingen dieser Arbeit legte. Ich danke ihr außerdem für den außergewöhnlichen organisatorischen Einsatz, der in der Endphase dieser Arbeit zu einer erheblichen Beschleunigung des Promotionsverfahrens führte.

Herrn Prof. Dr.-Ing. Horst-Michael Groß von der TU Ilmenau danke ich für die bereitwillige Übernahme des Korreferats, für das damit gezeigte Interesse an meiner Arbeit und vor allem für die zügige Erstellung des Gutachtens.

Bedanken möchte ich mich ebenfalls bei den mittlerweile 21 wissenschaftlichen und nichtwissenschaftlichen Kollegen des Fachgebietes und bei allen Studenten, die im Laufe meiner Tätigkeit im Fachgebiet mit mir zusammengearbeitet haben. Die großartige Zusammenarbeit mit jedem Einzelnen war geprägt von einem starken Zusammenhalt, für den nicht zuletzt Dr.-Ing. Siegbert Drüe Dank gebührt. Ganz besonders bedanke ich mich auch bei Dr.-Ing. Ralf Stemmer, Björn Meyer und dem Stipendiaten Lic. Doreid Kheirbek, die diese Zeit auf besonders humorvolle Art und Weise zu einem positiven Erlebnis gemacht haben.

Ferner danke ich den studentischen Hilfskräften und Absolventen, die mich durch ihr engagiertes Mitwirken unterstützten. Hervorheben möchte ich

dabei M. Sc. Michael Feldmann, der im Rahmen seiner Bachelorarbeit wichtige Beiträge zu der in den Kapiteln 4.2 und 4.3 dieser Arbeit vorgestellten Konturvervollständigung durch Splineinterpolation leistete. Für das Korrekturlesen des Manuskriptes danke ich Dipl.-Ing. Frank Schmidtmeier, Dipl.-Ing. Markus Schmidt und meiner Frau Claudia.

Ohne die Liebe und das Verständnis meiner Frau Claudia wäre diese Arbeit so nicht möglich gewesen. Ihr gilt daher mein größter Dank, ebenso meinen Eltern, die mir das Studium in jeder Hinsicht überhaupt erst ermöglicht haben.

Inhaltsverzeichnis

1	Einleitung	1
2	Grundlagen	5
2.1	Die Wahrnehmung des Menschen	5
2.2	Tiefenwahrnehmung	8
2.3	Scheinkonturen	11
2.4	Optische Illusionen und Wahrnehmungsphänomene	14
2.4.1	Helligkeitswahrnehmung	14
2.4.2	Größenwahrnehmung	15
2.4.3	Scheinkonturen	18
2.4.4	Figur-Grund-Gliederung	20
2.5	Strategien zur perzeptuellen Organisation	22
3	Bildvorverarbeitung	29
3.1	Filterung	31
3.2	Detektion von Konturen und Kreuzungspunkten	37
3.2.1	Schwellwertfreie Kantendetektion	38
3.2.2	Detektion von Kreuzungspunkten	41
3.3	Anisotrope Regularisierung	45
3.3.1	Minimierung einer globalen Kostenfunktion	47
3.3.2	Diffusion	57
3.3.3	Selbstorganisation der Kanteninformation	60
4	Bestimmung möglicher Konturvervollständigungen	67
4.1	Konturvervollständigung durch Tensor Voting	67
4.1.1	Repräsentation von Konturelementen durch Tensoren	67
4.1.2	Tensorkommunikation	69
4.1.3	Probleme des Tensorvoting	74
4.2	Bewertung von Verbindungswahrscheinlichkeiten	76
4.2.1	Ausschlusskriterien	76
4.2.2	Bestimmung der Verbindungswahrscheinlichkeit	79
4.3	Konturvervollständigung durch Spline Funktionen	83
5	Perzeptuelle Organisation	89

5.1	Figur-Grund-Gliederung an Kreuzungspunkten	89
5.2	Dynamische Tiefenzuordnung	92
6	Experimentelle Ergebnisse	99
6.1	Anisotrope Regularisierung	99
6.2	Konturverstärkung	107
6.3	Kanten- und Eckendetektion	109
6.4	Konturvervollständigung	112
6.5	Perzeptuelle Organisation vervollständigter Konturen	115
6.6	Natürliche Bilder	119
6.7	Laufzeiten	124
7	Zusammenfassung	125
	Literaturverzeichnis	129

1 Einleitung

Sowohl Lebewesen als auch autonom agierende Roboter, welche in Interaktion mit ihrer Umwelt treten, benötigen zur Navigation und zur Manipulation von Objekten unter anderem Informationen über die Form und die Position von Objekten in ihrer Umgebung. Zur Entfernungsmessung haben sich in der Praxis sogenannte aktive Verfahren etabliert. Hierbei wird durch eine aktive Beeinflussung der Umgebung, z. B. mittels Ultraschall oder Laserlicht, die Entfernung der Sensoreinheit zu einem Objekt gemessen. Häufig ist die Aufgabe eines Robotersystems jedoch auch mit der Objekterkennung verbunden. Hierzu sind die Sensoreinheiten aktiver Systeme in der Regel nicht in der Lage. Sogenannte passive Verfahren nutzen zur Analyse der Umgebung allein die natürliche Beleuchtung aus. Die Sensoreinheiten sind hierbei im Allgemeinen durch ein oder mehrere Kamerasysteme gegeben. Kamerasysteme bieten zum einen den Vorteil, dass eine aufwändige Datenfusion der Ergebnisse von Objekterkennung und Tiefenwahrnehmung entfällt und zum anderen den Vorteil einer kostengünstigen Realisierbarkeit. Durch die Abbildung einer dreidimensionalen Szene auf ein zweidimensionales Kamerabild geht allerdings die Tiefeninformation der betrachteten Szene verloren. Auch die Form von Objekten beschränkt sich in der zweidimensionalen Darstellung auf eine Kontur, die darüberhinaus durch verdeckende Objekte unterbrochen sein kann.

Zur Rekonstruktion der durch die Kameraprojektion verlorenen Kontur- und Tiefeninformationen dienen Lebewesen, speziell der Mensch, als Vorbild. Eine naheliegende Methode zur Rekonstruktion der Tiefeninformation besteht in Analogie zum natürlichen Vorbild in der Auswertung zweier Stereobilder. Während aber das Gebiet der stereoskopischen Korrespondenzbestimmung weitgehend in der Literatur behandelt wurde und eine Vielzahl an

effizienten Algorithmen hervorgebracht hat, haben monokulare Tiefenkriterien bisher nur wenig Einzug in Bildverarbeitungsanwendungen gehalten. Häufig beschränken sich Beiträge zum Thema in der Beschreibung oder Modellierung menschlicher Wahrnehmungsphänomene oder optischer Illusionen, welche anhand von synthetischen Bildern demonstriert werden.

In der vorliegenden Arbeit soll eines der wichtigsten monokularen Tiefenkriterien, die Verdeckung, herangezogen werden, um eine betrachtete Szene sinnvoll dreidimensional interpretieren zu können und verdeckte Konturen zu vervollständigen. Hierbei soll besonderer Wert auf die gesamte Prozesskette, angefangen bei der Filterung natürlicher Bilder, bis hin zur perzeptuellen Organisation des Bildinhalts gelegt werden. Dazu wird mit der anisotropen Regularisierung eine Methodik vorgestellt, die entsprechend einer gewünschten Skalierung eine Reduktion des Eingangsbildes auf wesentliche Bildinhalte gewährleistet, wobei saliente, also wesentliche, Konturinformationen beibehalten werden.

Kapitel 2 erläutert die Grundlagen menschlicher Wahrnehmung und die Mechanismen perzeptueller Organisation, wobei auch die Ansätze zu einer Reproduktion dieser Mechanismen in Bildverarbeitungsanwendungen besprochen werden. In Kapitel 3 werden die für die perzeptuelle Organisation notwendigen Vorverarbeitungsschritte vorgestellt. Besondere Berücksichtigung findet hierbei die kantenerhaltende Glättung des Eingangsbildes durch anisotrope Regularisierung, welche eine Reduzierung des Informationsgehaltes im Allgemeinen komplexer Eingangsbilder gewährleistet. Grundlage perzeptueller Organisation ist die in Kapitel 4 beschriebene Vervollständigung unterbrochener Konturen oder Objektgrenzen. Mit der als Tensor Voting bezeichneten Methodik und einer Splineinterpolation werden zwei alternative Verfahren zur Konturvervollständigung vorgestellt. Kapitel 5 beschreibt die eigentliche perzeptuelle Organisation des vorliegenden Ansatzes, bestehend aus einer Figur-Grund-Gliederung an markanten Bildpositionen und einer daraus folgenden dreidimensionalen Interpretation des zweidimensionalen Eingangsbildes. Die vorgestellten Verarbeitungsschritte werden anhand ei-

ner Vielzahl von synthetischen und natürlichen Testbildern in Kapitel 6 visualisiert und bewertet. Schließlich wird in Kapitel 7 eine Zusammenfassung der vorliegenden Arbeit gegeben.

2 Grundlagen

2.1 Die Wahrnehmung des Menschen

Das Abbildungssystem des menschlichen Auges besteht vereinfacht dargestellt aus einer Lochblende, einer konvexen Linse und einer Rezeptorebene. In der Rezeptorebene, der Retina, leiten ca. 126 Millionen Photorezeptoren ihre Signale indirekt an ca. 1 Million Ganglienzellen weiter (siehe Abbildung 2.1). Hierbei findet bereits eine Vorverarbeitung der Signale statt [ST97]. So zeigen schon Ganglienzellen, die Ausgangsschicht der Retina, ein komplexes Erregungsmuster. Die Signale der Ganglienzellen werden über die Sehnerven, das Chiasma opticum und die seitlichen Kniehöcker zur Weiterverarbeitung in den visuellen Cortex weitergeleitet. Dort befindet sich eine Vielzahl verschiedener Neuronen, die auf spezifische Reize in ihren rezeptiven Feldern reagieren. So finden sich z. B. Neuronen, die auf Lichtbalken einer bestimmten Orientierung reagieren, während andere Neuronen orientierungs- und richtungsspezifisch auf Lichtreize reagieren. Es findet im visuellen Cortex also ein Übergang zu einer merkmalsbasierten Repräsentation der wahrgenommenen Umwelt statt.

Unter den Merkmalen, die sich aus dem Retinaabbild ableiten lassen, kommt den Konturen eine besondere Bedeutung zu, wie der Versuch von Krauskopf [Kra63, Kra67, Spi01] zeigt. Einer Versuchsperson wird dabei eine rote Scheibe, umgeben von einem grünen Ring dargeboten. Wird die Kante zwischen der roten Scheibe und der umgebenden grünen Fläche stabilisiert, so kann sich der Reiz nicht gegen das Rezeptormosaik verschieben. Die Neurone adaptieren und es kann keine Kante mehr wahrgenommen werden. Als Folge verschwindet für die Versuchsperson die Wahrnehmung Rot im Inneren des Musters, stattdessen wird die gesamte Fläche einheitlich als

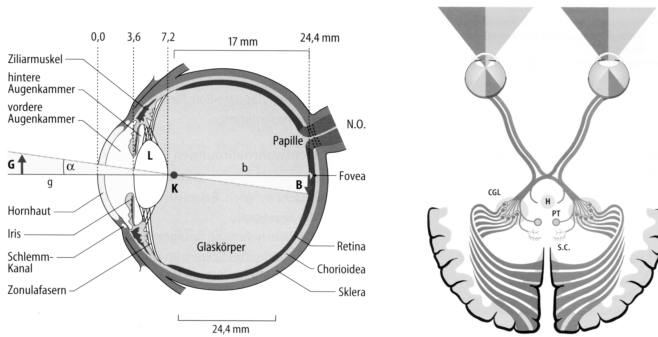


Abbildung 2.1: Sehbahn (nach [ST97]).

grün wahrgenommen. Ohne die Wahrnehmung der Begrenzung kann also auch die Fläche selbst nicht wahrgenommen werden.

Auch bei der Untersuchung sakkadischer Augenbewegungen zeigt sich, dass bei freier Betrachtung eines Bildes der Blick der Versuchsperson auf Unstetigkeitsstellen des dargebotenen Reizes, wie z. B. Kanten oder Ecken verweilt [Yar67]. Einheitliche Flächen werden dagegen kaum betrachtet (vgl. Abbildung 2.2). Flächen, die gleiche Helligkeit, Farbe und Struktur aufweisen, sind also im informationstheoretischen Sinne redundant. Die Wahrnehmung von Konturen ist damit offensichtlich der Wahrnehmung von Flächen vorangestellt [Spi01].

Der Merkmalsrepräsentation folgt eine Organisation des Wahrgenommenen zu komplexeren, handhabbaren Einheiten, sogenannten Perzepten. Ziel einer solchen perzeptuellen Organisation ist also die Abstraktion der Einzelinformationen bzw. Merkmale und der damit verbundene Aufbau einer inneren Repräsentation des Wahrgenommenen. Aus der Betrachtung der Art und Weise, wie das menschliche Gehirn eine Gruppierung von Teilfiguren oder Merkmalen vornimmt, haben sich in dem als Gestaltpsychologie bezeichneten Zweig der Wahrnehmungspsychologie Gruppierungsme-

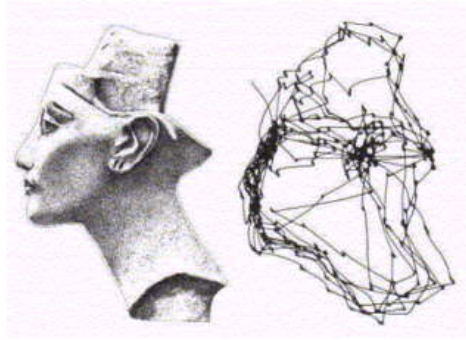


Abbildung 2.2: Muster der Augenbewegungen einer Versuchsperson beim freien Betrachten einer Abbildung der Nofretete-Büste (nach [Yar67]).

chanismen herauskristallisiert, welche als Gestaltgesetze bezeichnet werden [Wer23, Kof35, Pal99]. Zu diesen Gesetzen zählen unter anderem das Gesetz des gemeinsamen Schicksals, welches eine Gruppierung gemeinsam bewegter Elemente bezeichnet und das Gesetz der Ähnlichkeit, welches die Gruppierung von einander ähnlichen Elementen postuliert [Pal99]. Im Rahmen der vorliegenden Arbeit sind die Gesetze der Nähe und der guten Gestalt von Interesse. Das Gestaltgesetz der Nähe besagt, dass Merkmale, die einen geringeren Abstand zueinander aufweisen als andere, bevorzugt gruppiert werden, während das Gesetz der guten Gestalt eine Wahrnehmungsorganisation beschreibt, welche möglichst einfache, symmetrische oder regelmäßige Figuren bevorzugt. Korrespondierend hiermit besagt das Gesetz der Kontinuität, dass die Wahrnehmung unterbrochene Konturen so fortsetzt, dass sich kein abrupter Richtungswechsel ergibt und außerdem keine neuen Formen eingeführt werden müssen. Dies ist insbesondere für die in Kap. 2.3 beschriebene Scheinkonturwahrnehmung von Interesse. Weiterhin ist das Prinzip der Geschlossenheit zu nennen, das eine bevorzugte Wahrnehmung von geschlossenen Formen beschreibt. Perzepte werden in Bezug zum räum-

lichen und zum zeitlichen Kontext gesetzt und sind somit in ihrer Interpretation von diesen abhängig. Die als Wahrnehmungskonstanz bezeichnete Toleranz gegenüber geringen Veränderungen der Information auf der Sensorebene lässt sich neurologisch untermauern durch auf höheren Ebenen des Sehsystems entdeckte Zellen, welche denselben Reiz unabhängig von Entfernung, Beleuchtung und sogar partieller Verdeckung signalisieren [SVO93].

Neben Gruppierungsmechanismen ist ein wesentliches Merkmal perzeptueller Organisation die Gliederung des sensorischen Abbildes in Bereiche und die Einteilung der Bereiche in Figur und einen Grund, von dem sich die Figur als Vordergrund abhebt. Perzeptuelle Organisation beinhaltet somit auch eine dreidimensionale Repräsentation des Wahrgenommenen, was insbesondere auch dann der Fall ist, wenn nur ein Auge bzw. nur ein zweidimensionales sensorisches Abbild vorhanden ist.

2.2 Tiefenwahrnehmung

Hubel und Wiesel [HW70] fanden im visuellen Cortex Nervenzellen, sogenannte binokulare Neurone, die von beiden Augen stimuliert werden können. Unter diesen binokularen Neuronen existieren solche Zellen, die in Abhängigkeit von der Entfernung, auf die der Blick gerichtet ist, und der Entfernung eines Lichtreizes reagieren. Das Verarbeitungssystem verfügt somit über Mechanismen zur Tiefenwahrnehmung. Durch psychophysikalische Experimente konnte Bela Julesz [Jul60] nachweisen, dass der Mensch bereits allein aus der Information der Disparität, also der relativen Verschiebung korrespondierender Bildpunkte in den beiden Stereobildern, einen Tiefeneindruck erhält. Er zeigte dies durch die Einführung von Zufallspunktstereogrammen. Ein solches Zufallspunktstereogramm entsteht, indem zunächst zwei identische Bilder mit einem Muster zufällig verteilter weißer und schwarzer Bildpunkte erzeugt werden. In diese Bilder wird ein kleineres, ebenfalls zufällig erzeugtes Muster, z. B. in Form eines Rechtecks, versetzt eingefügt. Betrachtet man das in Bild 2.3 dargestellte Zufallspunktstereogramm mit

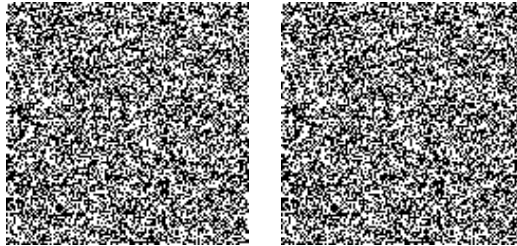


Abbildung 2.3: Zufallspunktstereogramm nach Julesz (nach [Jul60])

einem aufrecht zwischen die Bilder gestellten Blatt Papier, so fusionieren die Bilder zu einem. Es entsteht der Eindruck eines aus dem Hintergrund hervortretenden Rechtecks. Zufallspunktstereogramme enthalten keinerlei Informationen über die Form des Objekts, ebenso sind keine monokularen Tiefenkriterien anwendbar. Die einzige Information, die dem menschlichen Auge zur Verfügung steht, ist die der relativen Verschiebung des versetzt eingefügten Objektes. Dies zeigt, dass die Gewinnung der Tiefeninformation schon vor der Objekterkennung stattfinden und somit sogar zur Objektsegmentierung genutzt werden kann.

Beim Betrachten einer Szene mit nur einem Auge gewinnen wir bereits einen Tiefeneindruck durch sogenannte monokulare Tiefenkriterien. So beurteilen wir die Entfernung eines Objekts z. B. anhand der bekannten Größe dieses Objekts. Akkommodation, also der durch die Augenlinsen eingestellte Schärfebereich, und Konvergenzbewegungen der Augen, die sich auf ein Objekt ausrichten, liefern sensorische monokulare Tiefeninformationen. Ein weiteres Kriterium ist die Schattierung, die bestimmt, ob wir eine Struktur als Erhebung oder Vertiefung wahrnehmen, wobei das visuelle System des Menschen davon ausgeht, dass das Licht von oben einfällt.

Ein Beispiel für die Mehrdeutigkeit der Tiefenwahrnehmung durch Schattierung ist in Abbildung 2.4 gegeben. Während das menschliche Sehsystem

dazu tendiert, das Bild als einen großen und einen kleinen Krater wahrzunehmen, lässt sich diese Interpretation durch eine Drehung des Bildes um 180° in die Wahrnehmung zweier Vulkane verändern. Dargestellt ist in Wirklichkeit die 180° -Rotation der Abbildung hawaiianischer Aschekegel.



Abbildung 2.4: Mehrdeutigkeit zwischen Wahrnehmung eines Kraters oder eines Vulkans. Einfluss der Schattierung auf die Tiefenwahrnehmung (nach [PF05]).

In einem zeitlichen Kontext dient die sogenannte Bewegungsparallaxe als Tiefenkriterium, die berücksichtigt, dass weiter entfernte Objekte ihre Position auf der Sensorebene langsamer verändern als nähere Objekte, die sich mit der gleichen relativen Geschwindigkeit zum Beobachter bewegen. Die Konvergenz paralleler Strahlen am Horizont, wie dies z. B. bei Bahnschienen oder Straßen der Fall ist, wird als lineare Perspektive bezeichnet und wurde vielfach bereits in Renaissancebildern zur Erzeugung eines räumlichen Eindrucks eingesetzt. Ein weiteres Tiefenkriterium ist der Texturgradient, da gleichmäßig strukturierte Texturen mit zunehmender Entfernung zum Betrachter dichter erscheinen.

Das für die vorliegende Arbeit zentrale monokulare Tiefenkriterium ist die Verdeckung. Ein Großteil der Objekte, die sich in einer betrachteten Szene befinden, wird von anderen Objekten, die sich in größerer Nähe zum Betrachter befinden, verdeckt. Verdeckte Objekte erscheinen damit weiter

entfernt, während okkludierende Objekte als näher wahrgenommen werden. Die Objektgrenze eines okkludierten Objektes wird offensichtlich durch das Vordergrundobjekt teilweise unterbrochen. Für ein biologisches oder ein technisches Bildverarbeitungssystem kann dieser Sachverhalt herangezogen werden, um Objektgrenzen zu vervollständigen und über eine Figur-Grund-Gliederung eine dreidimensionale Repräsentation des Sensorabbildes zu erstellen.

2.3 Scheinkonturen

Scheinkonturen, auch als Scheinkanten oder illusionäre Konturen bezeichnet, sind wahrgenommene Konturen, für die sich kein korrespondierender Reiz auf der Rezeptorebene eines Sehsystems finden lässt. Neurophysiologische Untersuchungen haben gezeigt, dass die Wahrnehmung von Scheinkonturen sowohl bei Säugetieren wie Katzen, Affen und dem Menschen, bei Vögeln wie der Schleiereule und sogar bei Insekten, nämlich Honigbienen, nachweisen lässt [Nie02b, Hey03]. Die Gehirne und Sehsysteme dieser Lebewesen, die sich im Laufe der Evolution sehr unterschiedlich entwickelt haben, haben sich offensichtlich an gemeinsame visuelle Randbedingungen angepasst und versuchen, Wahrscheinlichkeitserwägungen über ihre Umgebung anzustellen, indem unterbrochene Begrenzungen vermeintlicher Objekte perzeptuell vervollständigt werden. Für das menschliche Sehsystem sind Scheinkonturen schon seit Beginn des zwanzigsten Jahrhunderts Gegenstand gestaltpsychologischer Untersuchungen [Kof35, ESS03, Wer23]. Zuerst wurde das Wahrnehmungsphänom um 1900 von Schuhmann erwähnt. Er beschrieb Scheinkonturen als Konturen, die nicht objektiv vorhanden sind [Sch00].

Abbildung 2.5(a) zeigt das sogenannte Kanizsa Dreieck [Kan79], das das Standardbeispiel für eine optische Illusion darstellt, welche auf Scheinkonturwahrnehmung basiert. Die sich hieraus ergebende Figur-Grund-Organisation leitet sich allein aus monokularen Kriterien ab (siehe [Kan79]). Der Begriff optische Illusion bezieht sich hierbei auf die Tatsache, dass das Kanizsa Bild

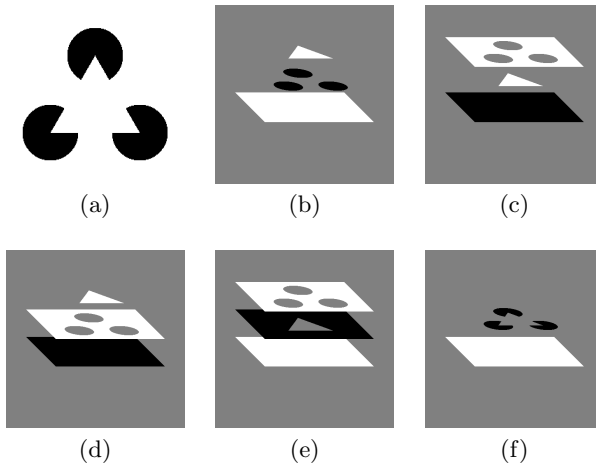


Abbildung 2.5: Das Kanizsa Dreieck (a) und unterschiedliche perspektivisch angedeutete dreidimensionale Konstellationen (b)-(f), die alle zum selben Retinaabbild (a) führen.

zweidimensional ist und sich lediglich aus drei schwarzen und einer weißen Fläche zusammensetzt, das menschliche Gehirn aber dazu tendiert, das Bild dreidimensional zu interpretieren. Da hier aber nicht von einer Fehlinterpretation gesprochen werden kann, wäre Wahrnehmungsphänomen in diesem Fall ein passenderer Begriff. Obwohl eine Vielzahl korrekter dreidimensionaler Interpretationen existiert (Abbildung 2.5(b)-(f)), welche alle zum selben Abbild auf der Sensorebene eines bildverarbeitenden Systems führen würden, besteht die präferierte Szeneinterpretation in einer perzeptuellen Organisation gemäß der Abbildung 2.5(b). D. h. es wird die Wahrnehmung eines weißen Dreiecks bevorzugt, das sich vor drei schwarzen Kreisen befindet, die sich wiederum vor einem weißen Hintergrund befinden. Die drei schwarzen Flächen bilden hierbei sogenannte induzierende Elemente, also Merkmale des betrachteten Bildes, die eine Scheinkonturwahrnehmung hervorrufen. Induzierende Elemente sind also an solchen Bildpositionen gegeben, wo auf-

grund verdeckter Flächen Diskontinuitäten im Konturverlauf auftreten, wie dies bei Eckpunkten, Endpunkten oder T-Kreuzungen (vgl. Kap. 3.2.2) der Fall ist. Es ist festzustellen, dass, obwohl beide Flächen weiß sind, das Dreieck gegenüber dem Hintergrund als heller erscheint, d. h. die Scheinkontur wird tatsächlich als ein Kontrastunterschied zwischen zwei Flächen wahrgenommen. Aufgrund der Mehrdeutigkeit der möglichen Interpretationen und dem damit schlecht gestellten Problem der perzeptuellen Organisation kann es nicht das Ziel eines Ansatzes zur perzeptuellen Organisation sein, die korrekte dreidimensionale Interpretation eines zweidimensionalen Bildes zu finden, sondern die wahrscheinlichste, wobei die Eigenschaften menschlicher Wahrnehmung hierbei als Anhaltspunkt dienen. Die Mehrdeutigkeit der Anordnung und der Einfluss höherer kognitiver Funktionen zeigen sich bei der Konzentration auf die Vorstellung einer weißen Fläche mit drei Löchern, hinter denen das Dreieck zu sehen ist, die tatsächlich zu einer Wahrnehmung entsprechend der Abbildung 2.5(c) führt und nun die Vervollständigung der Kreisflächen als Scheinkonturen präferiert. In diesem Zusammenhang ist zu unterscheiden zwischen sogenannter modaler und amodaler Vervollständigung. Modale Vervollständigung bezeichnet die Vervollständigung eines Vordergrundobjektes, was im Falle des Kanizsa-Dreiecks und der räumlichen Interpretation entsprechend der Abbildung 2.5(b) die Vervollständigung des Vordergrunddreiecks darstellt. Mit amodaler Vervollständigung wird dagegen die Vervollständigung verdeckter Konturen bezeichnet, was wiederum im Falle des Kanizsa-Dreiecks die Vervollständigung der schwarzen Kreisflächen bedeutet. In der Regel werden mit Scheinkonturen lediglich die modalen Vervollständigungen bezeichnet [Nie02b]. Neuere psychophysische Untersuchungen unterstützen die These, dass die Wahrnehmung modaler und amodaler Vervollständigungen auf einem gemeinsamen Mechanismus basiert [Nie02a, Nie02b], was auch der von Kellman begründeten Identitätshypothese [KGW01] und der in der vorliegenden Arbeit realisierten Konturvervollständigung entspricht. Die Identitätshypothese wird in [ASF02] in Zweifel gezogen, wobei sich die Gegenargumente allerdings auf stereoskopische induzierende Elemente und auf höhere, auf einer Wissens-

basis beruhende kognitive Funktionen beziehen. Weiterhin werden keine Hinweise darauf gegeben, wie die Konturvervollständigung stattfindet oder eine dreidimensionale Repräsentation gewonnen wird.

2.4 Optische Illusionen und Wahrnehmungsphänomene

Eine Vielzahl sogenannter optischer Illusionen basieren auf der Mehrdeutigkeit der dargebotenen Reize, so dass die Wahrnehmungsmechanismen, die den Beobachter zu einer bestimmten perzeptuellen Organisation des betrachteten Bildes führen, häufig im Widerspruch zu anderen Mechanismen oder der eigenen Erfahrung stehen. Solche Wahrnehmungsphänomene bieten wertvolle Hinweise auf die Funktionsweise menschlicher perzeptueller Organisation und können somit auch als Vorbild für effektive Bildverarbeitungs-Algorithmen herangezogen werden. Im Folgenden werden optische Illusionen und Wahrnehmungsphänomene präsentiert, welche einen gewissen Bezug zur vorliegenden Arbeit aufweisen. Neben der Helligkeitswahrnehmung, die in Bezug zur Filterung eines Eingangsbildes im Rahmen einer Vorverarbeitung steht und Skalierungseffekten sind hierbei insbesondere Phänomene zur Scheinkonturwahrnehmung und Figur-Grund-Gliederung von Interesse.

2.4.1 Helligkeitswahrnehmung

Die in Abbildung 2.6 dargestellte optische Illusion, benannt nach dem deutschen Physiologen Ludimar Hermann, wird als Hermann-Gitter bezeichnet. Die Bereiche, in denen sich die weißen Gitterlinien kreuzen, werden als dunkler im Vergleich zum Rest der Linie wahrgenommen, obwohl der Farbwert überall ein konstantes Weiß aufweist. Grund dafür ist die Zentrum-Umfeld Charakteristik rezeptiver Felder retinaler Ganglienzellen [GGM95], welche ein ähnliches Verhalten aufweisen wie die in Kap. 3 vorgestellten Filtermasken (vgl. Abbildung 3.6).

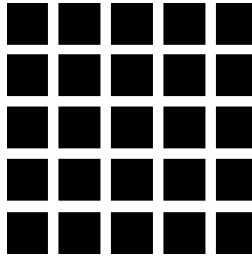


Abbildung 2.6: Hermann-Gitter (nach [Her70]).

Abbildung 2.7(a) ist ein Beispiel für die Farbkonstanz menschlicher Wahrnehmung, d. h. es werden zusammengehörige Flächen trotz eines Beleuchtungsunterschieds mit dem gleichen Farbwert wahrgenommen. So tendiert das menschliche Sehsystem, das Schachbrettmuster in Abbildung 2.7(a) in zwei Gruppen heller und dunkler Flächen einzuteilen. Die mit A und B bezeichneten Flächen werden daher mit unterschiedlichen Helligkeitswerten wahrgenommen, obwohl der dargebotene Reiz einen identischen Grauwert für die Flächen A und B aufweist. Dies wird in Abbildung 2.7(b) deutlich, wo dem ursprünglichen Bild zwei Balken mit dem Grauwert der Flächen A und B überlagert wurden. Offensichtlich haben Kontrastunterschiede für die menschliche Wahrnehmung also eine höhere Bedeutung als die eigentlichen Helligkeitswerte. Für eine Vielzahl praktischer Anwendungen, die sich auf natürliche Eingangsbilder beziehen, ist eine solche Beleuchtungsinvarianz auch für die digitale Bildverarbeitung von Bedeutung.

2.4.2 Größenwahrnehmung

Die in Abbildung 2.8(a) dargestellte optische Illusion wurde zuerst vom italienischen Psychologen Mario Ponzo (1882-1960) beschrieben, der vorschlug, dass das menschliche Gehirn die Größe eines Objektes in Relation zum Hintergrund beurteilt. Was in Abbildung 2.8(a) als Illusion bezeichnet wird, ist

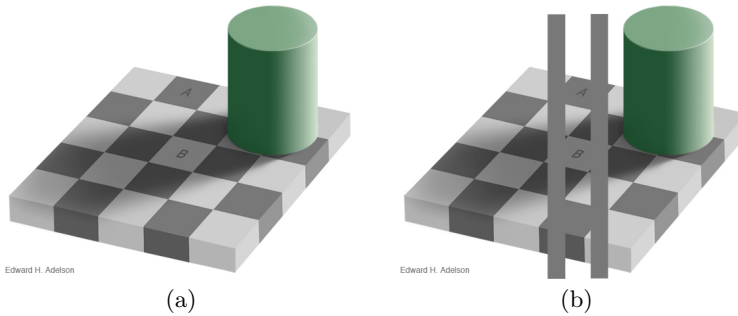


Abbildung 2.7: Schachbrett-Versuch von Edward H. Adelson aus dem Jahr 1995 (vgl. [Ade01]).

der Gegensatz zwischen der nachmessbaren Länge der beiden horizontalen Linien, welche beide gleich lang sind, und der menschlichen Wahrnehmung, die aufgrund einer dreidimensionalen Interpretation des dargebotenen Reizes die obere horizontale Linie als länger im Vergleich zur unteren Linie wahrnimmt. Das Tiefenkriterium, das dieser Interpretation zugrunde liegt, ist die bereits in Kap. 2.2 beschriebene lineare Perspektive, die besagt, dass parallele Linien sich zum Horizont hin verjüngen. Im Fall von Abbildung 2.8(a) werden also, ähnlich zu Straßenrändern oder Bahnschienen, die vier sich in einem Punkt treffenden dünnen Linien als perspektivische Darstellung von parallelen Linien interpretiert und die horizontalen Linien in diesen Kontext eingebettet, womit der oberen Linie eine größere Distanz zum Beobachter und damit auch eine höhere absolute Größe zugeordnet wird.

Ebenfalls auf dem Prinzip der linearen Perspektive basiert die in 2.8(b) dargestellte Müller-Lyer Illusion. Auch bei der zuerst vom deutschen Psychiater und Soziologen Franz Müller-Lyer beschriebenen optischen Illusion sind die vertikalen Linien gleich lang. Jedoch wird die linke vertikale Linie im Gegensatz zur rechten aufgrund der Einbettung in einen dreidimensionalen Kontext als kleiner wahrgenommen.

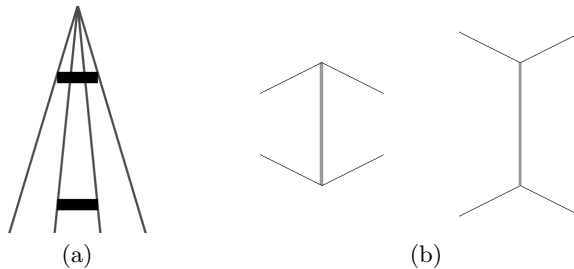


Abbildung 2.8: (a) Ponzo Illusion von 1913, (b) Müller-Lyer Illusion aus dem Jahr 1889.

Die in Abbildung 2.9(a) dargestellte Illusion wird ebenfalls als Ponzo-, zuweilen aber auch als Müller-Lyer Illusion bezeichnet. Offenbar enthält sie sowohl Elemente aus Abbildung 2.8(a) als auch aus Abbildung 2.8(b). Auch hier werden die auf Rezeptorebene gleich großen, dick gezeichneten vertikalen Balken als unterschiedlich groß wahrgenommen. Im Gegensatz dazu sind die im 1946 vom amerikanischen Augenarzt und Psychologen Adelbert Ames vorgestellten Ames-Raum abgebildeten Personen physikalisch betrachtet alle ähnlich groß (siehe Abbildung 2.9(b)), werden jedoch aufgrund ihrer Entfernung zur Kamera unterschiedlich groß abgebildet [Bau78]. Trotz des monokularen Tiefenkriteriums, das bekannten gleichartigen Objekten aufgrund der Größe ihrer Abbildung auf Rezeptorebene eine Tiefe zuordnet, tendiert die menschliche Wahrnehmung dazu, die Personen als unterschiedlich groß wahrzunehmen. Dies geschieht nicht nur auf Abbildungsebene, sondern auch bezogen auf die der Abbildung zugrunde liegenden Szene. Grund dafür ist der Raum, in dessen Kontext die Wahrnehmung der Personen eingebunden wird. Der Raum wird als rechtwinklig wahrgenommen, da die eigentlich trapezförmige Rückwand des Raumes nicht im rechten Winkel zu den Seitenwänden steht, sondern, bezogen auf den Betrachter, eine starke Neigung um die vertikale Achse aufweist. Dieser Eindruck wird zusätzlich verstärkt durch ebenfalls trapezförmige und unterschiedlich große Türen und Fenster,

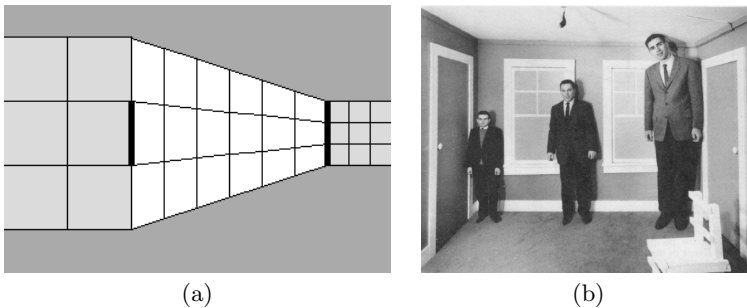


Abbildung 2.9: Gleiche Objektgröße auf der Ebene der sensoriellen Abbildung bei der Ponzo-Illusion in (a) und gleiche absolute Objektgröße in der physischen Realität beim Ames-Raum in (b).

deren Abmessungen so gewählt sind, dass für eine bestimmte Perspektive ihre Abbildung rechtwinklig erscheint.

2.4.3 Scheinkonturen

Abbildung 2.10 zeigt eine Modifikation des bereits in Kap. 2.3 vorgestellten Kanizsa-Dreiecks. Im Gegensatz zu dem in Kap. 2.3 vorgestellten Kanizsa-Dreieck besteht auf Rezeptorebene die Umgebung der drei schwarzen Flächen hier nicht aus einer weißen Fläche, sondern aus drei Flächen mit jeweils unterschiedlichen Grauwerten. Durch die Kanten, die von diesen Flächen gebildet werden, wird eine modale Vervollständigung der Kreise unterbunden und damit die Mehrdeutigkeiten bei der perzeptuellen Organisation verringert. Statt eines Vordergrund-Dreiecks tendiert in diesem Fall die menschliche Wahrnehmung aufgrund der unterschiedlich schattierten Flächen zu der räumlichen Interpretation einer Pyramide, die sich vor drei schwarzen Kreisen befindet.

Die nach dem deutschen Psychologen Walter Ehrenstein benannte Abbildung 2.11 erzeugt ebenfalls Scheinkonturen. In diesem Fall fungieren die

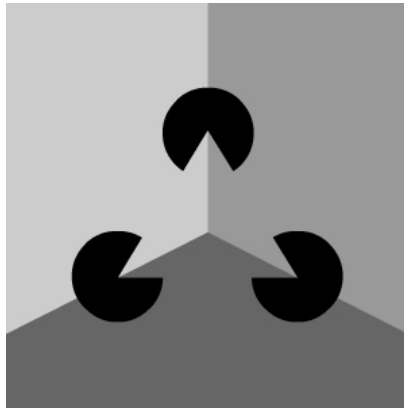


Abbildung 2.10: Modifikation des Kanizsa-Dreiecks zu einer Kanizsa Pyramide (nach [GG07]).

Linienenden als induzierende Elemente und erzeugen den Eindruck eines weißen Kreises, der sich vor den schwarzen Linien und vor einem weißen Hintergrund befindet, gegenüber dem er als wesentlich heller wahrgenommen wird. Der Eindruck einer Vordergrundfläche wird zerstört, wenn der Kreis mit einer schwarzen Linie umrandet wird.

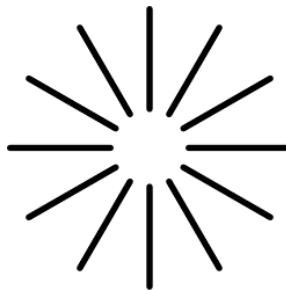


Abbildung 2.11: Ehrenstein-Täuschung.



Abbildung 2.12: Bild eines Dalmatiners von R. C. James (aus [Mar82]).

Die geschilderten Wahrnehmungsphänomene zur Scheinkonturwahrnehmung lassen sich alle mit Bottom-Up-Prozessen erklären, für die lediglich induzierende Elemente im sensoriellen Abbild vorhanden sein müssen. Dass aber auch eine Wissensbasis in Form erlernter Objektprototypen eine Rolle bei der Konturwahrnehmung spielt, zeigt Abbildung 2.12. Für die Wahrnehmung der Konturen des abgebildeten Dalmatiners ist Wissen über das Aussehen eines Hundes notwendig, die perzeptuelle Organisation muss in diesem Fall also von höheren kognitiven Funktionen beeinflusst sein.

2.4.4 Figur-Grund-Gliederung

Im Fall der Abbildung 2.13 bezieht sich der Begriff optische Illusion auf die Mehrdeutigkeit des Bildes in Bezug auf die Organisation der Flächen in einen Vorder- und einen Hintergrund. Wird die schwarze Fläche als Vordergrund betrachtet, so folgt die Wahrnehmung einer Vase vor weißem Hintergrund. Wird dagegen die schwarze Fläche als Hintergrund interpretiert, so ergibt

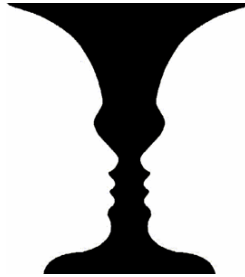


Abbildung 2.13: Rubin Vase, zuerst um 1915 vom dänischen Psychologen Edgar Rubin vorgestellt.

sich die Profilansicht zweier Gesichter, die sich vor dem schwarzen Hintergrund befinden.

Obwohl Erfahrung, zusammen mit eindeutigen induzierenden Merkmalen, wie z. B. der Stiel der Schaufel und die Hose des Mannes hinter dem Zaun, eine Interpretation der Abbildung 2.14 nahelegen, in der ein Mann und eine Frau sich hinter dem Zaun befinden, entsteht der Eindruck, dass der Oberkörper der Frau und der Unterkörper des Mannes sich hinter dem Zaun befinden, während der Oberkörper des Mannes und der Unterkörper der Frau sich vor dem Zaun befinden.

Hierbei fällt auf, dass die horizontalen Zaunlatten einen zum Rock der Frau und zur Jacke des Mannes identischen Grauwert aufweisen. Ausschlaggebend für die Wahrnehmung des Zaunes als Hintergrund in diesen Bereichen ist nach der Hypothese von Petter [Kan79] die Tatsache, dass der Zaun dünner ist als die anderen gleichfarbigen Bereiche und damit als weiter entfernt wahrgenommen wird.

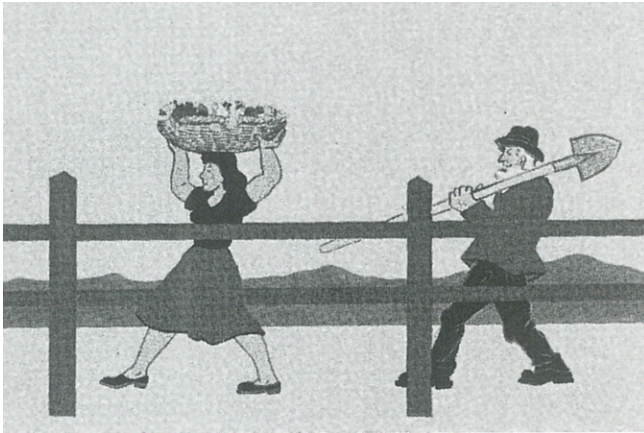


Abbildung 2.14: Mann und Frau mit wechselnden Figur-Grund-Attributen in Bezug zum Zaun (aus [Kan79]).

2.5 Strategien zur perzeptuellen Organisation

Grundsätzlich ist bei der Gruppierung von Konturelementen zwischen einer Verstärkung salienter Konturen und einer Konturvervollständigung zu unterscheiden. Die Verstärkung salienter Konturen hat in der Regel eine Aufbereitung der verrauchten Daten einer Kantendetektion zum Ziel und verstärkt somit ohnehin schon vorhandene Konturen [AB98, SU88]. Durch eine Konturvervollständigung werden dagegen Konturen rekonstruiert, die über weite Bereiche des Eingangsbildes aufgrund von Verdeckung oder fehlendem Kontrast zwischen Vorder- und Hintergrund nicht im Bild enthalten sind. Häufig wird diese Unterscheidung jedoch nicht getroffen, da eine Auffüllung von Konturlücken, wie sie im Rahmen einer Konturverstärkung stattfindet, sich oft auch über das Ende oder den Knick einer Kontur fortsetzen lässt und damit Phänomene wie z. B. die Konturwahrnehmung im Kanizsa-Dreieck erzeugt werden können. Hierbei ist jedoch zu beachten,

dass gerade amodale Vervollständigungen von Konturen, die hinter einem texturierten Vordergrund verlaufen, der seinerseits Konturen erzeugt, nicht ohne weiteres möglich sind. Die Mehrheit der Ansätze, welche sich mit der Thematik der Verstärkung salienter Konturen oder der Vervollständigung unterbrochener Konturen auseinandersetzen, nutzen ein bipolares Kommunikationsschema [HN01, HN04, PZ89, NM01], wie es von Grossberg und Mingolla [RGM00] vorgestellt wurde.

Eine solche Maske zur perzeptuellen Gruppierung besteht normalerweise aus ein oder zwei symmetrischen Flügeln, welche sowohl die Verbindungswahrscheinlichkeit zweier Konturelemente kodieren, als auch die Orientierung, welche ein Konturelement in Bezug auf ein betrachtetes Konturelement aufweisen sollte, um mit diesem gruppiert, d. h. verbunden zu werden (vgl. Abbildung 2.15).

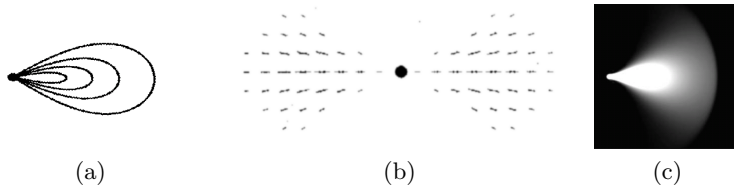


Abbildung 2.15: Kommunikationsschemata in verschiedenen Ansätzen: (a) Konturdarstellung des Verbindungsschemas von Heitger und van der Heydt [HHP⁺98], (b) bipolares Verbindungsschema nach Grossberg und Mingolla [NM01] und (c) modellierte Brown'sche Molekularbewegung nach Williams [ZW00] .

In [WT00] verglichen Williams und Thornber verschiedene Formen solcher Gruppierungsfelder. Eine detaillierte Übersicht zur Thematik der Konturverstärkung findet sich in [Han03] und [NM01], wobei in [NM01] Modelle betont werden, die eine Scheinkonturwahrnehmung implizieren. Zu nennen sind hier die Ansätze von Heitger et al. [HH93, PH01] und Neumann et

al. [HSN01] als neurophysisch motivierte Verfahren, sowie das Modell von Zweck und Williams [ZW00], welches die Brown'sche Molekularbewegung eines Partikels von einer Quelle ausgehend modelliert.

Ein weiterer Ansatz zur perzeptuellen Gruppierung von Konturelementen ist das von Medioni et al. begründete Tensor Voting Verfahren [GM96], welches auch in der vorliegenden Arbeit neben einer Spline-Interpolation als eine Möglichkeit der Konturvervollständigung genutzt und in Kap. 4 näher erläutert wird. In [MLT00] wurde Tensor Voting auf Konturverstärkungsprobleme in synthetischen bzw. binären Testbildern angewandt. Dieser Ansatz wurde in [MBM02b, Bab03] auf Grauwertbilder ausgeweitet, indem in einem Vorverarbeitungsschritt die Filterantworten sogenannter Gabor Filter (siehe auch [Gab46, Dau88, Lee96]) als für das Votingverfahren notwendige Eingangstensoren kodiert wurden. Das im Tensor Voting genutzte Kommunikationsschema wird als Stick-Voting Feld bezeichnet (vgl. Kap. 4), ist bipolar und bezieht sich damit auf Orientierungen mit Winkeln im Bereich von 0° bis 180° , im Gegensatz zu einem direktionalen, d. h. gerichteten Verbindungsschema, das Winkel im Bereich von 0° bis 360° unterscheidet. Für den Anwendungsbereich der Konturvervollständigung mag der Einsatz eines solchen bipolaren Voting-Feldes zwar im Kontext einer einheitlichen Behandlung aller Konturelemente auf einer dicht besetzten Merkmalskarte zwar Sinn machen, würde aber, wie schon erwähnt, zu Störungen bei der Organisation der Konturelemente in solche Bereichen führen, wo eine amodale Vervollständigung in einem stark texturierten Bildbereich gefunden werden soll. Für eine Konturvervollständigung, welche auch amodale Vervollständigungen mit einbezieht, wird ein gerichtetes Kommunikationsschema benötigt, welches ausgehend von induzierenden Elementen eine dünn besetzte Verbindungskarte erzeugt.

Ein solcher Ansatz wurde, bezogen auf das Tensor-Voting Verfahren, in [MM01b] vorgestellt, allerdings bestanden hierbei die Eingangsbilder aus Binärdateien und das vorgestellte monokulare Tiefenkriterium berücksichtigte lediglich die Länge der vervollständigten Konturen in sogenannten

„Spontaneously“ oder „Self Splitting Figures“, d. h. Abbildungen, in denen Vorder- und Hintergrundobjekt denselben Farb- bzw. Binärwert aufweisen (vgl. Kap. 5.1 und Abbildung 2.14). Darüber hinaus werden in [MM01b] induzierende Elemente in Form von Eckpunkten ebenfalls mit Hilfe des Tensor Voting Verfahrens bestimmt, was, wie in Kap. 4.1 näher erläutert wird, in komplexeren natürlichen Eingangsbildern zu hohen Fehlzuordnungen führen kann.

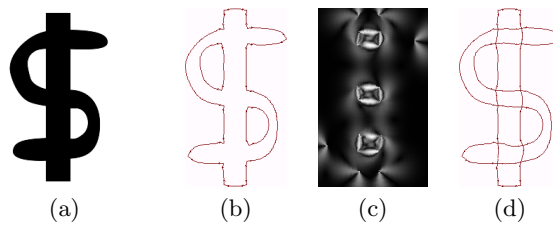


Abbildung 2.16: Konturvervollständigung nach [MM01b]: (a) binäres Eingangsbild, (b) detektierte Konturen, (c) Tensorkarte zur Konturvervollständigung und (d) detektierte und vervollständigte Konturen.

Zu den Ansätzen, welche nicht nur eine Scheinkonturwahrnehmung bzw. Konturvervollständigung realisieren, sondern dies auch mit einer Figur-Grund-Gliederung verbinden, zählt der Ansatz von Heitger et al. [HHP⁺98]. Allerdings wird hierbei lediglich an Eckpunkten eine reale Kontur, zusammen mit ihrer vervollständigten Fortsetzung, als Objektgrenze des Vordergrundes angesehen, während der Bereich, in dem sich die zweite Kontur des Eckpunktes befindet, dem Hintergrund zugeordnet wird (vgl. Abbildung 2.17(a)). Offensichtlich führt ein solches Vorgehen zu Fehlzuordnungen im Falle amodaler Vervollständigungen, wie z. B. den Kreisvervollständigungen im Kanizsa-Dreieck. Die Autoren bemerken selbstkritisch, dass ein solch simpler Mechanismus getäuscht werden kann [HHP⁺98] und führen die modalen Vervollständigungen eines Zebras vor weißem Hintergrund als Beispiel an. In

diesem Fall würde obiges Vorgehen den Körper des Zebras als Hintergrund klassifizieren.

In [GKP96] wird von Geiger et al. zunächst eine Karte definiert, welche die Wahrscheinlichkeit dafür angibt, dass der Bildinhalt im betreffenden Bereich dem Vordergrund zuzuordnen ist. Auf dieser Karte findet dann ein Diffusionsprozess statt, der von Figur-Grund-Unterscheidungen an Schlüsselpunkten ausgeht, wobei eine Diffusion über reale Kanten hinaus verhindert wird. Ein solcher Diffusionsprozess kann seiner Natur nach ein Bild nur in die zwei Kategorien von Vordergrund und Hintergrund unterteilen (vgl. Abbildung 2.17(c)). Eine Betrachtung komplexer Tiefenverhältnisse, wie Objekte, die ihren Zustand von Vorder- zu Hintergrund wechseln, ist nicht möglich. So werden auch nur synthetische, binäre Eingangsbilder betrachtet, welche in der wahrscheinlichsten dreidimensionalen Interpretation lediglich zwei Tiefenebenen aufweisen.

Ebenso werden im Ansatz von Williams [Wil94, WH96] nur binäre synthetische Eingangsbilder betrachtet (vgl. Abbildung 2.17(d)). Die Vervollständigung okkludierter Oberflächen leitet die Figur-Grund-Entscheidung für Konturkreuzungen aus einer kombinatorischen Betrachtung der Konstellation an diesen Schlüsselpunkten ab. Hierbei werden lokale Kriterien ausschließlich in die globale Organisation eingeschlossen, was die Betrachtung global inkonsistenter Konstellationen verhindert. Dieses Vorgehen wird dabei nicht mit dem in [ZW00] vorgestellten Gruppierungsverfahren kombiniert.

Auch das von Saund vorgestellte Verfahren basiert im Gegensatz zur vorliegenden Arbeit aus einer kombinatorischen Betrachtung der möglichen Figur-Grund-Entscheidungen an Kreuzungspunkten [Sau99a, Sau99b]. Die Konturvervollständigung geht also der Figur-Grund-Gliederung nicht voraus, sondern ist Folge des kombinatorischen Entscheidungsprozesses (vgl. Abbildung 2.17(b)). In [Sau99a] wird das präsentierte Verfahren auf ein natürliches Eingangsbild angewandt, wobei anhand von Fehlzuordnungen der Einfluss der Bildvorverarbeitung aufgezeigt wird.

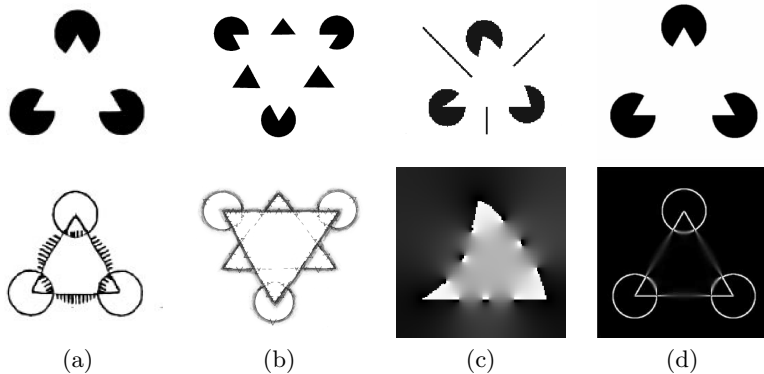


Abbildung 2.17: Kanizsadreiecke, obere Reihe: Eingangsbild und untere Reihe: Ergebnis des jeweiligen Verfahrens: (a) vervollständigte Konturen nach Heitger et al., kleine Linien zeigen in Richtung des Hintergrunds (aus [PH01]), (b) eine mögliche Konturvervollständigung nach Saund, Pfeile und Schattierung zeigen in Richtung des Hintergrunds, gestrichelte Linien deuten amodale Vervollständigungen an (aus [Sau99a]), (c) nach Geiger et al. durch Diffusion bestimmte Wahrscheinlichkeit, dass der Bildinhalt zum Vordergrund gehört (aus [GKP96]) und (d) stochastisches Vervollständigungsfeld nach Williams et al. (aus [WJ96])

Insgesamt lässt sich festhalten, dass Ansätze zur perzeptuellen Gruppierung, die natürliche Eingangsbilder präsentieren, diese zumeist nutzen, um die Grenzen des Verfahrens aufzuzeigen [HHP⁺98, Sau99a]. Die eigentliche Funktionsweise des Verfahrens wird in der Regel an einer optischen Illusion wie dem Kanizsa-Dreieck demonstriert. Hieraus wird die Bedeutung einer Bildvorverarbeitung ersichtlich, welche für einen Prozess der perzeptuellen Organisation in komplexen realen Bildern notwendig ist. Ebenso fällt auf, dass existierende Ansätze zur perzeptuellen Organisation lediglich eine binäre Figur-Grund-Gliederung an Konturkreuzungen vornehmen.

Eine Zuweisung relativer Tiefe entlang der gesamten Kontur, ob real oder vervollständigt, findet nicht statt.

3 Bildvorverarbeitung

Ziel der Bildvorverarbeitung ist es, nachfolgenden Verarbeitungsschritten möglichst genaue und vollständige Informationen über die Position und die Orientierung von Konturelementen bereitzustellen. Gerade für die vervollständigung von Konturen, die aufgrund einer Verdeckung oder fehlendem Kontrast zwischen Vorder- und Hintergrund unterbrochen sind, wird nicht nur die Position einer solchen Konturunterbrechung benötigt, sondern auch die Richtung, in der die unterbrochene Kontur fortgesetzt werden soll. Dies stellt ein wesentliches Kriterium für die Wahl der Filter, welche zur Kontur- bzw. Kantendetektion und zur Detektion von Ecken oder Konturkreuzungspunkten genutzt werden, dar. Für den Anwendungsbereich der Kantendetektion haben sich gradientenbasierte Filtermethoden etabliert [MS98, ZT98]. Die Filtermasken, die in diesem Zusammenhang genutzt werden, basieren in der Regel auf einer Diskretisierung eines Ableitungsoperators erster Ordnung (vgl. Abbildung 3.1). Kanten, d. h. orientierte Kontrastunterschiede im Grau- oder Farbwertverlauf eines Bildes, sind hierbei als lokale Extremstelle der Filterantwort gegeben.

Durch die Diskretisierung der Summe der beiden reinen zweiten Ableitungen erhält man einen Laplaceoperator, welcher Kanten durch Nullstellendurchgänge in der Filterantwort anzeigt. Da Kantendetektoren, welche auf Diskretisierungen von Gradienten basieren, häufig empfindlich auf Rauschen reagieren, ist es üblich, die Filtermaske mit einem Tiefpassfilter, welches häufig aus einer gaußförmigen Funktion besteht, zu kombinieren [GW08, JHG99]. Dies führt wiederum zu einer erhöhten Lokalisierungsungenauigkeit der Kantendetektion. Daher muss bei einem solchen Vorgehen immer ein Kompromiss zwischen Rauschunempfindlichkeit und Lokalisierungsgenauigkeit gefunden werden. Wie schon erwähnt, wird für die Erzeugung von Konturver-

-1	0
0	1

(a)

0	-1
1	0

(b)

-1	-2	-1
0	0	0
1	2	1

(c)

-1	0	1
-2	0	2
-1	0	1

(d)

0	1	0
1	-4	1
0	1	0

(e)

Abbildung 3.1: (a) bis (d): Koeffizienten für gradientenbasierte Filtermasken (a) und (b) Roberts Filter und (c) und (d) Sobel Filter, (e) Laplace Filter.

vollständigungen, wie sie in Kapitel 4 vorgestellt wird, nicht nur die Position von Kanten und Eckpunkten benötigt, sondern auch die Orientierung von Kanten und die Orientierung von möglichen Kantenfortsetzungen an Kreuzungspunkten. Weiterhin wird in Kapitel 4.2 ein Mechanismus zur Berücksichtigung von Hell- Dunkelübergängen an Kreuzungspunkten benötigt. Im Folgenden wird daher auf eine Kombination verschiedener Filtermasken zurückgegriffen, welche nicht nur die geforderten Informationen liefert, sondern darüber hinaus auch eine schwellwertunabhängige Kantendetektion ermöglicht. Weiterhin bietet sie eine im Vergleich zu reinen Gradientenfiltern gute Lokalisierung bei gleichzeitiger Toleranz gegenüber Rauschprozessen.

Eine natürliche Szene ist in der Regel derart komplex aufgebaut, dass eine Kanten- und Eckendetektion eines hochaufgelösten Bildes unzählige Ecken und Kanten liefert, die in ihrer Skalierung korrekt erkannt sein mögen, aber einem Betrachter, der seine subjektive Skalierung an die ihn interessierenden Objekte anpasst, als unsinnig erscheint. Es ist daher wünschenswert,

ein Eingangsbild zu simplifizieren, d. h. hochfrequente Bildinhalte entsprechend einer gewünschten Skalierung aus dem Bild zu entfernen und damit das Bild auf die wesentlichen Bildinhalte zu reduzieren. Die hier naheliegenden Funktionen, welche eine Tief- oder Bandpasscharakteristik aufweisen, würden neben hochfrequenten Bildanteilen allerdings auch Informationen über saliente Kanten löschen oder verändern. Ein Mechanismus, der eine Glättung des Eingangsbildes unter Beibehaltung salienter Kanten bewirkt, wird in Form eines anisotropen Regularisierungsverfahrens in Kapitel 3.3 vorgestellt.

3.1 Filterung

Alle im Folgenden vorgestellten Filter basieren auf einer Gaußfunktion

$$g(x, y) = w(x, y) \cdot e^{\frac{-\|(x, y)^T\|_2^2}{\sigma}}, \quad (3.1)$$

wobei $(x, y)^T$ die Bildposition innerhalb der Faltungsmaske angibt. Da im Anwendungsfall die Filtermaske einen begrenzten Radius r aufweist, wird der Parameter σ so gewählt, dass für einen Wert $\varepsilon > 0$ im Fall $w(x, y) \equiv 1$ die Gaußfunktion unter den Wert ε fällt:

$$\sigma = \frac{-r^2}{\ln(\varepsilon)} \quad (3.2)$$

Im Folgenden wird, um hinreichend niedrige Werte am Rand der Filtermaske zu gewährleisten, $\varepsilon = 0,01$ gewählt. Über die Gewichtungsfunktion w sollen drei verschieden geformte Klassen von Faltungsmasken erzeugt werden, von denen zwei orientierungsselektiv sind. Die Orientierungsselektivität zu einem gegebenen Winkel α wird durch eine Rotation der Bildposition $\mathbf{p} = (x, y)^T$ in den Prototyp g der jeweiligen Filtermaske erreicht. Für die orientierungsabhängige Filtermaske f ergibt sich somit

$$f(\mathbf{p}) = g(\mathbf{R} \cdot \mathbf{p}) \quad \text{mit } \mathbf{R} = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \quad (3.3)$$

Durch die Wahl der Vorzeichen der Gewichtungsfunktion wird, wie im Folgenden erläutert, die Form der Filtermaske bestimmt. Weiterhin wird die Gewichtungsfunktion w so gewählt, dass

$$\sum_{\mathbf{p}} f(\mathbf{p}) = 0 \quad (3.4)$$

gilt und die Filterantworten auf einen gemeinsamen maximalen Wert der Filterantwort normiert werden.

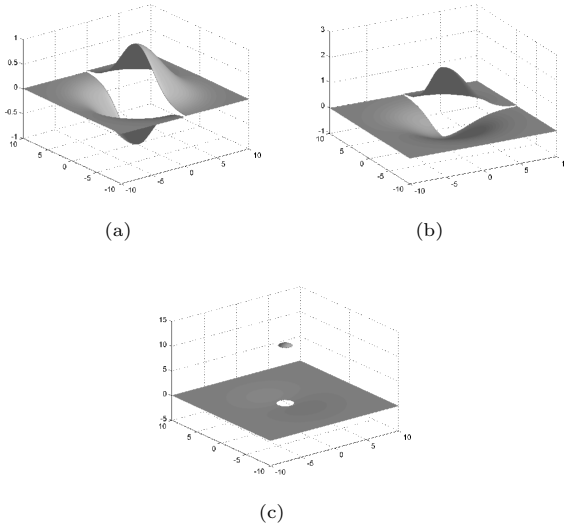


Abbildung 3.2: Kontinuierliche Darstellung der Filtermasken: (a) Kantenfilter (b) Eckfilter (c) Zentrum-Umfeld Filter.

Wird für die Gewichtungsfunktion

$$w(x, y) := w(x) = \text{sgn}(x) \quad (3.5)$$

gewählt, so ergibt sich als Prototyp die in Abbildung 3.2(a) dargestellte Filtermaske, die im Folgenden als Kantenfilter bezeichnet wird.

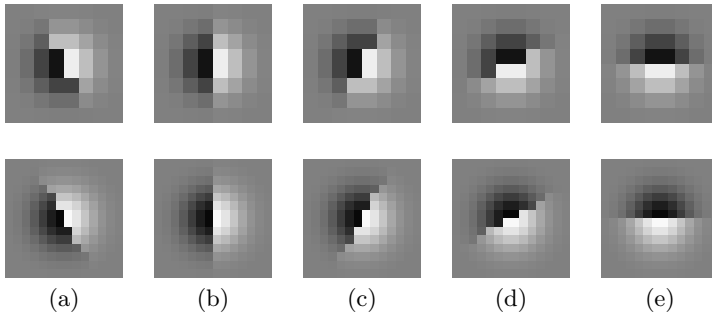


Abbildung 3.3: Orientierungsselektive gaußförmige Kantenfilterbank für Orientierungswinkel von: (a) 60, (b) 90, (c) 120, (d) 150 und (e) 180 Grad. Obere Reihe: Filterradius 4 Pixel und untere Reihe: Filterradius 7 Pixel.

Die hieraus folgenden diskreten Filtermasken sind für verschiedene Radien und Orientierungen in Abbildung 3.3 dargestellt. Dabei zeigen helle Grauwerte positive Vorzeichen der Filtermaske an, während dunkle Grauwerte negative Vorzeichen bedeuten. Da die Breite der Filtermaske immer dem doppelten Filterradius entspricht, ist der Funktionswert $x = 0$ der Signumfunktion in Gleichung (3.5) für die Bestimmung der Koeffizienten der diskreten Filtermaske uninteressant. Die hier präsentierten Kantenfilter weisen einen eindeutigen maximalen Ausschlag im Falle einer entsprechend orientierten stufenartigen Diskontinuität des Grauwertverlaufs auf, wie sie an Objektkanten auftritt. Im Gegensatz dazu würde ein Filter mit ungerader Zeilen- und Spaltenanzahl in einem solchen Fall zwei maximale Ausschläge erzeugen.



Abbildung 3.4: Ergebnisse der Filterung mit den in Abbildung 3.3 und 3.2(a) dargestellten Kantenfiltern, einem Filterradius von sieben Pixeln und 24 betrachteten Orientierungen. (a) Eingangsbild, (b) Filtermaximum und (c) Winkel der betragsgrößten Filterantworten.

Zu dem häufig in der Bildverarbeitung genutzten Testbild „Lena“ [Pla72] in Abbildung 3.4(a) ist in Abbildung 3.4(b) zu jedem Bildpunkt der jeweils höchste Betrag aller Kantenfilterantworten dargestellt. Hohe Beträge sind dabei hell dargestellt, während schwarz den Wert Null repräsentiert.

Die Filterantwort nach Faltung der Kantenfiltermaske mit dem Eingangsbild unterscheidet sich bei zwei Orientierungen α und $\alpha + 180^\circ$ lediglich durch ihr Vorzeichen. Es werden daher für die Kantenfilter nur Orientierungen im Bereich von 0° bis 180° betrachtet, wobei der Winkelbereich entsprechend der Anzahl der betrachteten Orientierungen in äquidistante Abschnitte eingeteilt wird. In Abbildung 3.4(c) ist zu jedem Bildpunkt der Winkel des Kantenfilters mit der betragsgrößten Filterantwort dargestellt. Um auf den für Digitalbilder üblichen Wertebereich von 0 bis 255 zu normieren, wird die Winkelinformation in Abbildung 3.4(c) zu gegebenem Winkel φ mittels $(\sin(\varphi) + 1) \cdot 127.5$ visualisiert.

Durch eine geeignete Wahl der Gewichtsfunktion w lässt sich der in Abbildung 3.2(b) dargestellte Verlauf der Filtermaske erzielen. Die Gewichte sind

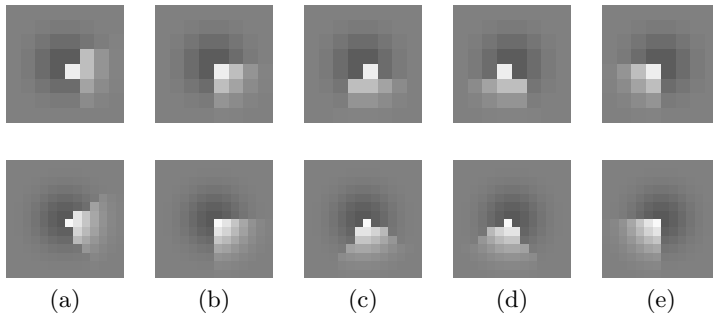


Abbildung 3.5: Orientierungsselektive gaußförmige Eckfilterbank für Orientierungswinkel von: (a) 60, (b) 90, (c) 120, (d) 150 und (e) 180 Grad. Obere Reihe: Filterradius 4 Pixel und untere Reihe: Filterradius 7 Pixel.

dabei so normiert, dass bei maximaler Filterantwort der Betrag des im Folgenden als Eckfilter bezeichneten Filters dem des Kantenfilters entspricht. Dies wird in Kap. 3.2 Bedeutung für die Vergleichbarkeit der Filterantworten haben. In Abbildung 3.5 sind wieder die diskreten Filtermasken für unterschiedliche Orientierungen und für zwei unterschiedliche Filterradien dargestellt. Da das Filter im Gegensatz zum Kantenfilter nicht spiegelsymmetrisch ist, wird hier die doppelte Anzahl an Orientierungen berücksichtigt, d. h. es werden Winkel im Bereich von 0° bis 360° betrachtet.

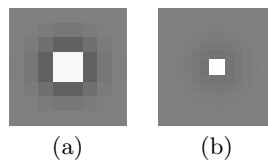


Abbildung 3.6: Zentrum-Umfeld Filter mit einem Filterradius von (a) 4 Pixel und (b) 7 Pixel.

Die in Abbildung 3.6 dargestellten diskreten Zentrum-Umfeld Filtermasken werden im Folgenden als ZU-Filter bezeichnet und entstehen durch eine Vorzeichenänderung der inneren vier Pixel der Filtermaske. Ähnlich dem Marr-Hildreth-Operator bzw. Laplacian of Gaussian (LoG) liefert das ZU-Filter durch Nulldurchgänge einen Hinweis auf Kanten, während die Maxima auf Linien oder einzelne Punkte schließen lassen [Mar82]. Auch hier ist der Betrag der maximalen Filterantwort auf den der bereits vorgestellten Filter normiert. Im Gegensatz zu diesen ist das ZU-Filter nicht orientierungsselektiv.

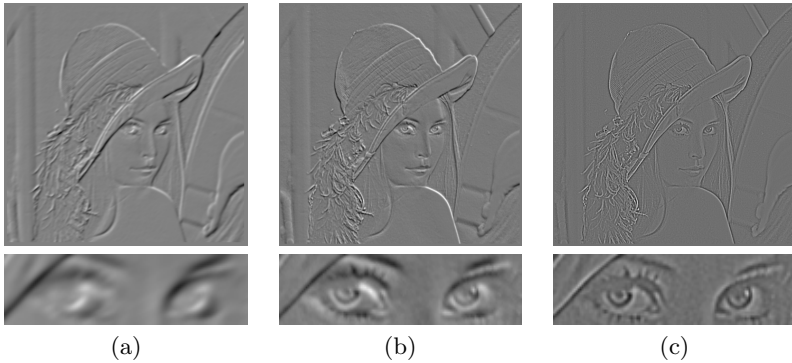


Abbildung 3.7: Filterergebnisse für das Lena Testbild in Abbildung 3.4.
 (a) Kantenfilter mit einer Orientierung von $\alpha = 150^\circ$, (b) Eckfilter, ebenfalls mit einer Orientierung von $\alpha = 150^\circ$ und (c) ZU-Filter. Der Filterradius der Filtermasken betrug jeweils 7 Pixel. Untere Reihe: Vergrößerung der Augen-Region.

Für die verschiedenen Filterklassen ist in Abbildung 3.7 jeweils ein Beispiel der Filterantwort gegeben, wobei wieder positive Werte hell dargestellt werden, negative Werte dunkel dargestellt werden und die Null auf einen mittleren Grauwert abgebildet wird. Für das Kantenfilter in Abbildung 3.7 (a) und das Eckfilter in Abbildung 3.7 (b) ist jeweils die Filterantwort des Filters mit einer Orientierung von 150° dargestellt (vgl. Abbildung 3.3(d)

und Abbildung 3.5(d)). Für das ZU-Filter existiert lediglich eine, orientierungsunabhängige Filterantwort, die in Abbildung 3.7(c) dargestellt ist.

Das unterschiedliche Verhalten der Filter wird in den Vergrößerungen der Augenregion in der unteren Reihe deutlich. Betrachtet man die Kante, d. h. die Diskontinuität des Grauwertverlaufs, welche sich im Originalbild im Bereich zwischen Hut und Gesicht befindet, so lässt sich feststellen, dass Kanten- und Eckfilter ein relatives Maximum der Filterantwort aufweisen, hier dunkel dargestellt, während das ZU-Filter, durch einen Hell-Dunkel-Wechsel dargestellt, im Konturbereich einen Vorzeichenwechsel und damit eine Nullstelle aufweist. Offensichtlich reagieren Eckfilter, genau wie Kantenfilter, auf Kanten. Aufgrund ihrer Konstruktion und Normierung fällt die Filterantwort an der Position einer idealen Kante allerdings im Vergleich zum Kantenfilter niedriger aus. Dieser Zusammenhang wird in Kap. 3.2 genutzt, um eine schwellwertfreie Kantendetektion zu realisieren.

3.2 Detektion von Konturen und Kreuzungspunkten

Eine auf gradientenbasierten Faltungsmasken (vgl. Abbildung 3.1) beruhende Kantendetektion ist in der Regel mit der Bildung eines Schwellwertes verbunden. Ansonsten würden auch solche Bildelemente als Kante deklariert, die zwar aufgrund von Rauschprozessen, sehr geringer Farbwertunterschiede oder Nichtlinearitäten im Grau- oder Farbwertverlauf eines Bildes ein lokales Extremum in der Filterantwort erzeugen, jedoch im eigentlichen Sinne keine Diskontinuität im Farb- oder Grauwertverlauf darstellen. Durch eine Schwellwertoperation wird die Kantendetektion allerdings stark abhängig von der Helligkeit des gegebenen Bildes [Mar82, Can86, Gab46]. Eine gewisse Abhilfe schafft hier die beim sogenannten Canny Filter häufig genutzte Vorgehensweise [Can86], bei der zunächst Konturen mit einem hohen Schwellwert extrahiert werden, um dann mit einem niedrigen Schwellwert vervollständigt zu werden, wobei die Beleuchtungsabhängigkeit aber weiterhin gegeben bleibt. Wird ferner die Größe der Faltungsmaske erhöht,

um den Einfluss von Rauschprozessen zu verringern, so führt dies zu Lokalisierungsungenauigkeiten sowohl bei der Kantendetektion, als auch bei der Bestimmung der Position von Eckpunkten. Um einen schwellwertfreien Ansatzpunkt für die Kantendetektion zu erhalten, wird im Folgenden zunächst nach einer idealen Kombination der in Kap. 3.1 vorgestellten Filter gesucht, um die hieraus gewonnenen Kanten anschließend ebenfalls anhand der Betrachtung der Kombinationen der Filterantworten zu vervollständigen. Die Betrachtung der Kombinationen von Filterantworten ermöglicht darüber hinaus, notwendige Bedingungen für Eckpunkte zu formulieren.

3.2.1 Schwellwertfreie Kantendetektion

Alle in Kap. 3.1 vorgestellten Filter basieren auf derselben Gaußfunktion, was mit einer entsprechenden Normierung den Vorteil bietet, dass die Filterwerte direkt miteinander vergleichbar sind. Dies ist in Abbildung 3.8 am Beispiel der Filterantworten im Bereich einer Kante und im Bereich einer Ecke dargestellt. Die in der Spalte (a) dargestellten Konstellationen zeigen jeweils eine Kante bzw. eine Ecke, die eine Grenze zwischen einem weißen und einem schwarzen Bildbereich darstellen. Entlang einer horizontalen Richtung werden die Filterantworten betrachtet, die sich aus der Faltung der Filtermaske mit dem Eingangsbild ergeben, wobei die graue Linie in jedem Bild den Bereich darstellt, in dem sich die mittleren beiden Zeilen der Filtermaske bei der Faltung befinden. Durch die Betrachtung der Filterantworten entlang der grau dargestellten Zeilen ergeben sich die in den Spalten (b) bis (d) dargestellten Verläufe, wobei in (b) die Filterantworten der in Kap. 3.1 vorgestellten Kantenfilter, in (c) die der Eckfilter und in (d) die Filterantworten der ZU-Filter darstellt sind. Für die orientierungsselektiven Kanten- und Eckfilter ist für jede Bildposition der Betrag der betragsgrößten Filterorientierung dargestellt, während die Filterantwort des ZU-Filters mit Vorzeichen dargestellt ist.

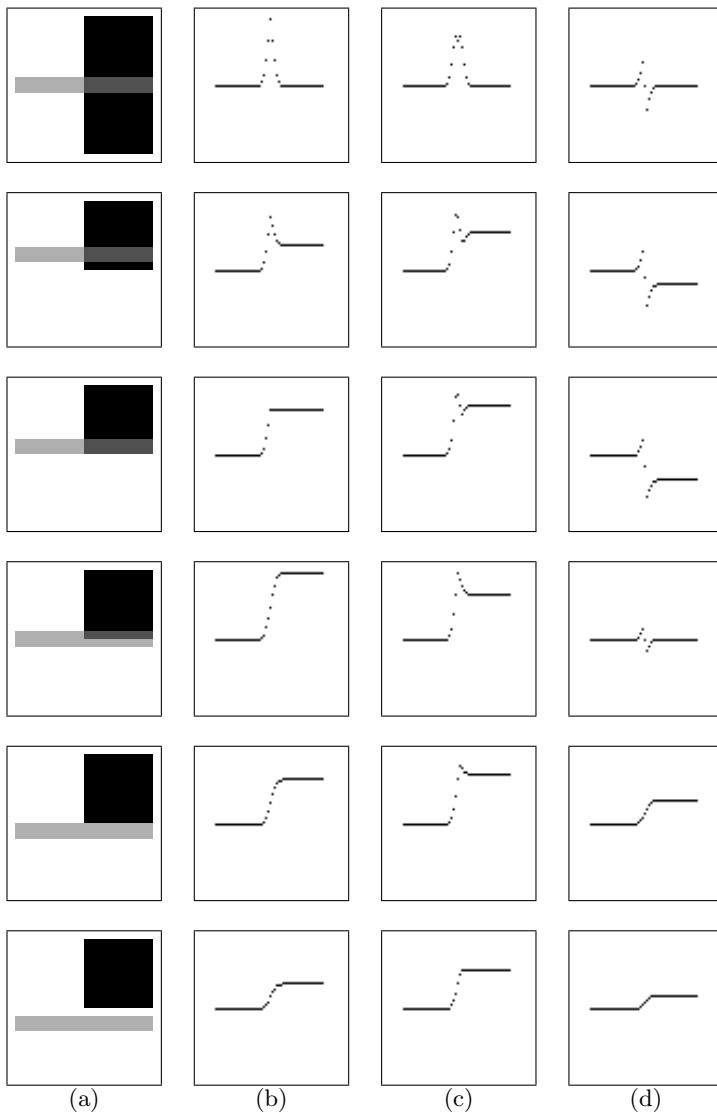


Abbildung 3.8: Charakteristika der Filterantworten an verschiedenen Bildpositionen, Erläuterung siehe Text.

Sämtliche Faltungen starten im weißen Bildbereich, sodass die Filterantworten alle auf der Nulllinie beginnen. Man beachte, dass der Radius der Filtermasken in diesem Beispiel sieben Pixel beträgt, was einer Filtermaske von 14×14 Pixel entspricht. Hierdurch werden auch für die Konstellation in der letzten Zeile von Abbildung 3.8 Filterantworten ungleich Null erzeugt, obwohl das Zentrum der Filtermaske sich immer im weißen Bildbereich befindet.

Die in der ersten Zeile von Abbildung 3.8 präsentierte Konstellation zeigt das Verhalten der Filter im Bereich einer Kante. Während die Kantenfilterbeträge ein eindeutiges Maximum aufweisen, weisen die Beträge der Eckfilter rechts und links von der Kante einen höheren Wert auf als an der Position der Kante selbst. Darüber hinaus sind die Eckfilterbeträge bis auf die Kantenposition an jeder Position höher als die der Kantenfilterbeträge. Weiterhin weist das ZU-Filter an der Position der Kante einen Nulldurchgang auf. Damit ist eine schwellwertfreie Kondition bestimmt für Bildpunkte, die in jedem Fall als Kanten deklariert werden können. Die Kombination eines Betragsmaximums des Kantenfilters, dessen Filterantwort betragsmäßig höher ist als die Eckfilters, verbunden mit einem Vorzeichenwechsel des ZU-Filters wird daher im Folgenden als ideale Kante bezeichnet. Experimentelle Ergebnisse zeigen, dass glatte Kantenverläufe im allgemeinen solche idealen Kanten erzeugen.

Ab der zweiten Zeile sind die Filterantworten für den Fall dargestellt, dass sich eine Ecke im Bereich der Filtermaske befindet. Statt eines Übergangs von einem weißen zu einem schwarzen Bildbereich sind jetzt permanent weiße Pixel im Bereich der Filtermaske enthalten. Nimmt, wie in der zweiten und der dritten Zeile dargestellt, der Einfluss der weißen Pixel zu und der der schwarzen Pixel ab, so ergibt sich eine veränderte Charakteristik der Filterantworten. Bewegt sich schließlich, wie in der vierten Zeile zu sehen ist, das Zentrum der Filtermaske direkt über die Position des Eckpunktes, so ergeben sich, im Vergleich zur idealen Kante, gänzlich andere Filterantworten.

ten. So liegt z. B. der maximale Ausschlag des Eckfilters über dem des Kantenfilters.

Die Kantendetektion der vorliegenden Arbeit beruht auf einem Tracking-Verfahren, das, ausgehend von einer noch nicht behandelten idealen Kante, diese fortzusetzen versucht. Hierbei wird zunächst nach einem benachbarten Bildpunkt gesucht, der ebenfalls eine ideale Kante darstellt. Liegt eine solche nicht vor, wird nach der nächstwahrscheinlichen Filterkombination gesucht, also nach einem Kantenfiltermaximum mit einem Nulldurchgang des ZU-Filters, wobei die Beträge des Eckfilters höher sind als die des Kantenfilters. Liegt auch eine solche Filterkombination nicht vor, werden die weiteren möglichen Kombinationen bis hin zu einem einfachen Nulldurchgang des ZU-Filters betrachtet. Die Kontursuche endet, wenn keine Fortsetzung gefunden wird oder das Tracking-Verfahren auf eine schon behandelte oder die aktuell behandelte Kante trifft. In dem Fall, dass das Tracking-Verfahren auf eine bereits behandelte Kante trifft, entsteht eine sogenannte T-Kreuzung, d. h. ein Kantenelement, das drei benachbarte Kantenelemente aufweist.

3.2.2 Detektion von Kreuzungspunkten

Um Kanten zu vervollständigen, die durch Verdeckung oder fehlenden Kontrast zwischen Vorder- und Hintergrund unterbrochen sind, werden als Ausgangspunkt der Vervollständigung die Positionen der Bildpunkte benötigt, die das Ende des realen Konturverlaufs markieren. Im Falle einer Verdeckung stellen diese Bildpunkte in der Regel den Kreuzungspunkt zweier Objektgrenzen dar. Solche Punkte, im Folgenden als Kreuzungspunkte bezeichnet, können auf Abbildungsebene sowohl zwei oder drei Konturen trennen oder den Endpunkt einer einzelnen Kontur darstellen. Verschiedene Konstellationen sind in Abbildung 3.9 exemplarisch dargestellt.

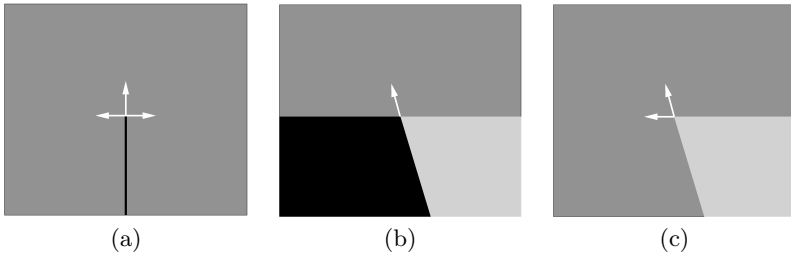


Abbildung 3.9: Verschiedene Kreuzungspunkte: (a) Endpunkt, (b) T-Kreuzung und (c) Eckpunkt.

Eine Linie wie in Abbildung 3.9(a) wird beim Kantentracking-Verfahren entsprechend Kap. 3.2 über relative Extrema der ZU-Filterantwort detektiert. Entsteht das dargestellte Linienende durch Verdeckung, so gibt es drei mögliche vom Kreuzungspunkt ausgehende Fortsetzungsrichtungen, die Fortsetzung der Linie selbst und die Objektgrenze des verdeckenden Objektes, die sich zu beiden Seiten des Endpunktes erstreckt. Die Linie ist in Abbildung 3.9(a) schwarz dargestellt, die Vervollständigungsrichtungen dagegen weiß. Die beiden Richtungen, die die Objektgrenze des verdeckenden Objektes repräsentieren, sind im eigentlichen Sinn keine Konturfortsetzungen, dienen aber der Konturvervollständigung des verdeckenden Objektes und damit der Scheinkonturwahrnehmung. Es fällt auf, dass diese beiden Vervollständigungsrichtungen im rechten Winkel zur realen Kontur gewählt sind, obwohl durchaus andere Neigungen denkbar sind. Auch wenn in [GM96] argumentiert wird, dass eine zur unterbrochenen Kontur rechtwinklige Objektgrenze als am wahrscheinlichsten anzusehen ist, stellt die rechtwinklige Wahl der Vervollständigungsrichtungen aber lediglich eine Approximation dar.

Die Richtungen der Konturvervollständigungen lassen sich aus der Richtung der Linie ableiten, deren Orientierung sich wiederum aus der Orientierung der betragsgrößten Kantenfilterantworten ergibt. Um die Orientierungsgenauigkeit zu erhöhen, wird hier, wie auch bei der folgenden Detektion von

T-Kreuzungen und Eckpunkten, eine Schwerpunktbildung zwischen der Filterantwort der betrachteten Orientierung und den Antworten der beiden benachbarten Orientierungen vorgenommen.

Bei der Überlagerung dreier Flächen mit unterschiedlichem Grauwert entsteht eine T-Kreuzung, wie sie in Abbildung 3.9(b) zu sehen ist. Eine solche T-Kreuzung wird während des in Kap. 3.2 beschriebenen Tracking-Verfahrens direkt als Folge der Kantendetektion erkannt. Allerdings liegt zunächst nur die Position eines Konturpunktes vor, der drei Konturen miteinander verbindet. Zur Ermittlung einer T-Kreuzung und damit einer Vervollständigungsrichtung müssen zwei der Konturen einander zugeordnet werden. Diese beiden Konturen stellen damit im Fall einer Verdeckung die Kontur der verdeckenden Fläche dar, während die dritte Kontur an der verdeckenden Fläche endet und durch ihre Richtung die Vervollständigungsrichtung vorgibt. Um zwei Konturen einander zuzuordnen, werden die durch die Kantenfilter gewonnenen Orientierungen aller drei Konturen miteinander verglichen. Die beiden Konturen, deren Orientierungen die geringste Winkeldifferenz aufweisen, wobei die Winkeldifferenz unter einem Schwellwert liegen muss, werden zu einer Kontur zusammengefasst. Der Schwellwert wird im Folgenden zu 30° gewählt, was eine gewisse Toleranz gegenüber Ungenauigkeiten in der Orientierungsbestimmung gewährleistet. Liegt keine der Winkeldifferenzen unter dem Schwellwert, kann keine sinnvolle Aussage über die Vervollständigungsrichtung getroffen werden. Es handelt sich in diesem Fall um eine sogenannte Y-Kreuzung.

Können Endpunkte und T-Kreuzungen noch direkt über die Anzahl benachbarter Konturelemente bestimmt werden, so ist dies für einen Eckpunkt, wie er in Abbildung 3.9(c) dargestellt ist, nicht ohne weiteres möglich, da sowohl der Eckpunkt als auch alle Konturelemente, die keinen Kreuzungspunkt darstellen, zwei benachbarte Konturelemente aufweisen. Darüber hinaus ist zu beachten, dass ab einer bestimmten Skalierung kein Unterschied mehr besteht zwischen einer Ecke und zwei diagonal benachbarten Konturelementen. Im Kontext dieser Arbeit sind Diskontinuitäten im Orientierungsverlauf von

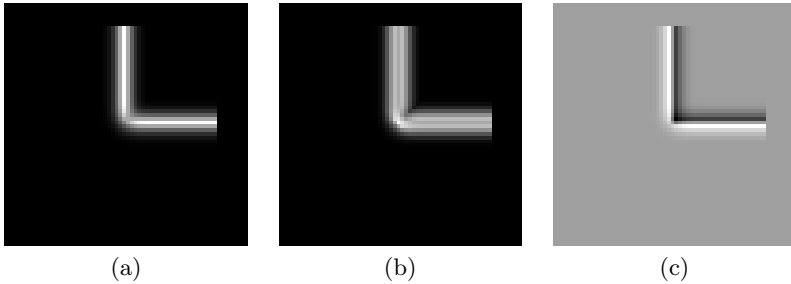


Abbildung 3.10: Filterantworten an einer rechtwinkligen Ecke entsprechend Abbildung 3.8: (a) Maximum der Beträge der Kantenfilterorientierungen, (b) Maximum der Beträge der Eckfilterorientierungen, und (c) vorzeichenbehaftete Filterantwort des ZU-Filters.

Interesse, die einen Hinweis auf unterbrochene Konturverläufe liefern können. Hierbei kommt wieder die Betrachtung der in Kap. 3.1 vorgestellten Filterantworten zum Tragen.

Betrachtet man die Abbildungen 3.10 und 3.8, so lässt sich feststellen, dass im Bereich der Ecke die Betragsmaxima der Eckfilter höher sind als die der Kantenfilter. Die Ecke liegt also nicht im Bereich einer idealen Kante, wie sie in Kap. 3.2 definiert wurde. Das intuitive Vorgehen zur Bestimmung der genauen Position der Ecke könnte eine Suche nach einem Betragsmaximum der Eckfilter entlang eines nichtidealen Kantenabschnitts sein. Es zeigt sich jedoch, dass ein solches Vorgehen bei Eckpunkten, die keinen rechten Winkel aufweisen, zu Lokalisierungsungenauigkeiten führt. Dagegen führt die Suche nach einem Betragsminimum des ZU-Filters entlang eines nichtidealen Kantenabschnitts unabhängig vom eingeschlossenen Winkel zu zuverlässigen Ergebnissen. Ein solches Vorgehen schließt jedoch solche stumpfen Winkel aus, deren Winkelinformation in Zusammenhang mit der Filterung zu einer idealen Kante führt. Daher findet in einem Nachbearbeitungsschritt zusätzlich eine Eckensuche über ideale Kantenelemente statt, welche eine

hohe Winkeldifferenz zu ihren beiden Nachbarn aufweisen, in deren weiterer Nachbarschaft sich aber hinreichend viele Kantenelemente mit niedriger Winkeldifferenz befinden. Auch hier ergeben sich die Orientierungen der Eckpunkte und die daraus folgenden Richtungen möglicher Fortsetzungen aus den Orientierungen der betragsgrößten Kantenfilter in den entsprechenden Bildpunkten.

Bezogen auf die Nachbarschaften eines Pixels sind offensichtlich auch Kreuzungspunkte denkbar, die bis zu acht Konturen miteinander verbinden. Ein solcher Fall ist allerdings nicht nur unwahrscheinlich, es ließen sich darüber hinaus auch keine sinnvollen Vervollständigungsrichtungen bestimmen. Im Rahmen dieser Arbeit werden daher nur Kreuzungspunkte betrachtet, die maximal drei Kanten miteinander verbinden.

3.3 Anisotrope Regularisierung

Die Idee anisotroper Diffusion in der Bildverarbeitung, also der Diffusion bzw. Glättung eines Bildes entlang einer bevorzugten Orientierung, lässt sich zurückführen auf einen zuerst von Perona und Malik vorgestellten Diffusionsprozess, dessen Diffusionsstärke vom Bildinhalt an der jeweiligen Position abhängt [PM90]. Das hiermit erzielte nicht homogene Verhalten, welches von Perona und Malik irreführenderweise als anisotrop bezeichnet wird, ermöglicht die Glättung eines Bildes mit einer gleichzeitigen Verstärkung salienter Konturen. Ziel ist es hierbei, im Gegensatz zu einer einfachen gaußförmigen Glättung, eine von der Skalierung, also der Stärke der Glättung, unabhängige Lokalisierung von Kanten zu erreichen.

Dieser Ansatz wird von Weickert aufgenommen [Wei96, Wei99], der wie Perona und Malik die Diffusionsgleichung

$$\frac{\partial u}{\partial t} = \nabla \cdot (D \nabla u) \quad (3.6)$$

nutzt, wobei er allerdings statt der von Perona und Malik vorgeschlagenen skalarwertigen Funktion den richtungsabhängigen Diffusionstensor D einführt, der aus einer Strukturtenorbeschreibung des Eingangsbildes gewonnen wird, um die lokale Bildstruktur zu beschreiben. Erst durch die Richtungsabhängigkeit des Diffusionstensors D und damit durch die Richtungsabhängigkeit der Diffusion selbst kann der Diffusionsprozess im eigentlichen Sinn als anisotrop bezeichnet werden. Ein wesentlicher Nachteil der Strukturtenorbeschreibung, speziell in Verbindung mit einer gaußförmigen Glättung, ist die hohe Orientierungsunsicherheit in der Nähe von Kreuzungspunkten.

Auch wenn Diffusionsfilter entsprechend Gleichung (3.6) und eine auf Regularisierungsmethoden basierende Diffusion, wie sie im Folgenden zur Realisierung einer anisotropen Diffusion entwickelt wird, eng miteinander verknüpft sind, besteht ein wesentlicher Unterschied darin, dass der Skalenraum, welcher die gefilterten Bilder für sämtliche Glättungsstärken umfasst, in Gleichung (3.6) durch die Zeit t aufgespannt wird und das Diffusionsfilterergebnis für eine hohe Skala bzw. mit zunehmenden Zeitwert gegen ein Bild konvergiert, das an jeder Bildposition den gleichen Grau- bzw. Farbwert aufweist und damit die maximale Glättung des Eingangsbildes darstellt. Dagegen wird der Skalenraum im folgenden Regularisierungsansatz zur anisotropen Diffusion durch den Glättungsparameter einer Energie- bzw. Kostenfunktion, die es zu minimieren gilt, aufgespannt.

Obwohl zahlreiche regularisierungsbasierte Methoden im Bereich der digitalen Bildverarbeitung existieren (siehe z. B. [IK99, Nor89, ABBFC97] und enthaltene Verweise), wird in der Regel weder ein klarer und einfach zu implementierender Ansatz zur anisotropen Diffusion präsentiert, noch ist die Problematik so formuliert, dass die Existenz und Eindeutigkeit der Lösung gewährleistet ist. In [ABBFC97] wird die Überlegenheit sogenannter kantenerhaltender Modelle gegenüber quadratischer Regularisierung postuliert, wie im Folgenden aber gezeigt wird, kann eine kantenerhaltende Diffusion durchaus als quadratische Regularisierung formuliert werden (vgl. [HM08]). Hierzu wird ein quadratischer Kostenterm definiert, der Bedingungen an den Diffusionsprozess enthält und dessen Minimum eine optimale Erfüllung dieser Bedingungen darstellt.

3.3.1 Minimierung einer globalen Kostenfunktion

Zunächst sei m die Anzahl der Pixel eines Bildes, welche von links oben nach rechts unten zeilenweise durchnummeriert werden. Weiterhin sei d die Anzahl der Farbkanäle des Bildes, d. h. $d = 1$ im Falle eines Grauwertbildes und $d = 3$ für ein RGB-Farbbild. Einem Bildpunkt $x \in [1, \dots, m]$ und einem Farbkanal $k \in [1, \dots, d]$ wird der Farbwert $\xi_{(x,k)}$ zugeordnet. Der initiale Farbwert des ursprünglichen Eingangsbildes sei mit $\xi_{(x,k)_0}$ bezeichnet. Weiter seien die Farbwerte in einem Vektor

$$\boldsymbol{\xi} = (\xi_{(1,1)}, \dots, \xi_{(m,1)}, \xi_{(1,2)}, \dots, \xi_{(m,d)})^T \quad (3.7)$$

zusammengefasst. Für das ursprüngliche Bild gilt analog $\boldsymbol{\xi}_0 \in \mathbb{R}^{m \cdot d}$ mit

$$\boldsymbol{\xi}_0 = (\xi_{(1,1)_0}, \dots, \xi_{(m,1)_0}, \xi_{(1,2)_0}, \dots, \xi_{(m,d)_0})^T \quad (3.8)$$

Wahl der Kostenfunktion

Eine kantenerhaltende Glättung des Eingangsbildes basiert auf der Annahme, dass benachbarte Bildpunkte, die gemeinsam auf einer Objektoberfläche liegen, annähernd gleiche Farbwerte aufweisen. Nur an Objektgrenzen treten Diskontinuitäten des Farbverlaufs auf. Eine solche Bedingung lässt sich durch einen ersten Kostenterm der Form

$$P_1(\xi) = \frac{c_1}{4} \cdot \sum_{k=1}^d \sum_{i=1}^m \sum_{j \in U_i} w_{ij} (\xi_{(i,k)} - \xi_{(j,k)})^2 \quad (3.9)$$

formulieren, der Kosten verursacht, falls benachbarte Pixel ungleiche Farbwerte aufweisen. Dabei stellt zu einem gegebenen Bildpunkt x die Menge U_x eine quadratische Nachbarschaft des Bildpunktes dar, die den Punkt x selbst nicht enthält. Die Konnektivität des Bildpunktes x mit einem benachbarten Bildpunkt i kann über den Gewichtungsfaktor $w_{xi} > 0$ gesteuert werden, wobei die Gewichte nicht notwendigerweise symmetrisch sein müssen, d. h. im Allgemeinen gilt $w_{xi} \neq w_{ix}$. Um später die numerische Stabilität des Iterationsverfahrens, das zur Lösung genutzt wird, zu gewährleisten, wird für die Summe der Gewichte der Umgebung U_x eines Bildpunktes x

$$\sum_{i \in U_x} w_{xi} \leq Q := \sum_{i \in U_x} 1 \quad (3.10)$$

gefordert. Der Wert Q stellt hierbei die Anzahl der in der Umgebung U_x enthaltenen benachbarten Bildpunkte dar. In Verbindung mit weiteren Kostentermen dient der Glättungsfaktor c_1 der Steuerung des Einflusses des Glättungsterms.

Betrachtet man den Kostenterm in Gleichung (3.9), so lässt sich feststellen, dass alle Bilder, die in jedem Bildpunkt den gleichen konstanten Farbwert aufweisen, den Kostenterm minimieren. Es existieren also mehrere Minima und damit keine eindeutig bestimmte Lösung. Um die Abhängigkeit des

geglätteten Bildes ξ vom Ursprungsbild ξ_0 in die Kostenformulierung zu integrieren, wird nun weiterhin gefordert, dass Kosten verursacht werden, falls ein Farbwert $\xi_{(x,k)}$ von seinem initialen Wert $\xi_{(x,k)_0}$ abweicht. Der entsprechende Kostenterm lautet

$$P_2(\xi) = c_2 \cdot \frac{1}{2} \cdot \sum_{k=1}^d \sum_{i=1}^m (\xi_{(i,k)} - \xi_{(i,k)_0})^2 \quad (3.11)$$

Auch hier dient der Faktor c_2 dazu, den Einfluss des Kostenterms zu steuern. Die Zusammenfassung der in Gleichung (3.9) und in Gleichung (3.11) geforderten Bedingungen ergibt nun eine globale Kostenfunktion P , deren Minimierung zu einem entsprechend der gewählten Gewichtungsfaktoren ge-glätteten Bild führt. Bei nur zwei Bedingungen kommt es lediglich auf das Verhältnis der Gewichtungsfaktoren c_1 und c_2 an, der Faktor c_2 kann also ohne Beschränkung der Allgemeinheit auf den Wert $c_2 = 1$ gesetzt werden. Übrig bleibt nur ein Skalierungs- bzw. Glättungsfaktor $c = c_1$.

$$\begin{aligned} P(\xi) &= P_1 + P_2 \\ &= \frac{1}{2} \cdot \sum_{k=1}^d \sum_{i=1}^m (\xi_{(i,k)} - \xi_{(i,k)_0})^2 \\ &\quad + \frac{c}{4} \cdot \sum_{k=1}^d \sum_{i=1}^m \sum_{j \in U_i} w_{ij} (\xi_{(i,k)} - \xi_{(j,k)})^2 \end{aligned} \quad (3.12)$$

Für die partielle Ableitung nach einem Farbwert $\xi_{(x,k)}$ gilt

$$\begin{aligned}
\frac{\partial}{\partial \xi_{(x,k)}} P(\xi) &= \xi_{(x,k)} - \xi_{(x,k)_0} \\
&+ \frac{c}{2} \left(\sum_{j \in U_x} w_{xj} (\xi_{(x,k)} - \xi_{(j,k)}) \right) \\
&+ \sum_{\{i|x \in U_i\}} w_{ix} (\xi_{(x,k)} - \xi_{(i,k)}) \quad (3.13)
\end{aligned}$$

Durch die Wahl der Potentialfunktion muss mindestens ein Minimum von P existieren. Ein notwendiges Kriterium an ein Minimum von P ist durch $\nabla P(\xi) = \mathbf{0}$, also

$$\frac{\partial}{\partial \xi_{(x,k)}} P(\xi) = 0 \quad \forall \xi_{(x,k)} \quad (3.14)$$

gegeben. Hieraus ergibt sich das lineare Gleichungssystem

$$\nabla P(\xi) = \mathbf{A} \cdot \xi - \xi_0 = \mathbf{0}. \quad (3.15)$$

Die Matrix \mathbf{A} in Gleichung (3.15) lässt sich darstellen durch

$\mathbf{A} = (a_{i,j})_{i,j} \in \mathbb{R}^{(m \cdot d) \times (m \cdot d)}$ mit

$$a_{i,j} = \begin{cases} 1 + \frac{c}{2} \cdot \left(\sum_k w_{ik} + \sum_k w_{ki} \right) & \text{für } i = j \\ -\frac{c}{2} \cdot (w_{ji} + w_{ij}) & \text{sonst} \end{cases} \quad (3.16)$$

Hierbei gilt $w_{ij} = 0$ für $j \notin U_i$. Weiterhin gilt $w_{ii} = 0$, $w_{ij} = w_{(i+d)(j+d)}$ und $j \notin U_i$, falls die Indizes i und j aus unterschiedlichen Farbkanälen stammen. Offensichtlich ist die Matrix \mathbf{A} symmetrisch und mit $\langle \mathbf{A}\xi, \xi \rangle > 0$ für $\xi \neq \mathbf{0}$ folgt die positive Definitheit von \mathbf{A} . Die Matrix ist damit regulär,

es existiert eine eindeutig bestimmte Lösung von Gleichung (3.15) und somit ein eindeutig bestimmtes Minimum von P [ZF84, Sch68].

Numerische Aspekte

Lösungen zur Bestimmung eines Minimums der Potentialfunktion P basieren im Allgemeinen auf der Idee des Gradienten- oder des Newtonverfahrens [Mei99]. Aufgrund der Regularität der Matrix \mathbf{A} in Gleichung (3.15) ließe sich die Lösung der Gleichung theoretisch auch durch ein direktes Verfahren bestimmen. Allerdings ist die Matrix \mathbf{A} schwachbesetzt, d. h. ein direktes Verfahren könnte die besondere Gestalt der Matrix nicht ausnutzen und würde vollbesetzte Zwischenmatrizen generieren, was den Speicherplatzbedarf im Vergleich zur Speicherung des Vektors $\boldsymbol{\xi}$ quadrieren würde. Die Menge der zu bestimmenden Elemente würde darüber hinaus eine erhöhte Berechnungszeit nach sich ziehen.

Für den hier vorgestellten Diffusionsprozess wird das Gradientenverfahren zur Bestimmung des Minimums herangezogen [Hun02, HM08]. Beim klassischen Gradientenverfahren wird die Minimallösung einer stetig differenzierbaren Funktion $P : \mathbb{R}^n \rightarrow \mathbb{R}$ bestimmt, indem, ausgehend von einem Startvektor, ein Versuchsvektor $\boldsymbol{\xi}_k$ iterativ in die lokal optimale Richtung \mathbf{d}_k korrigiert wird. In vielen iterativen Anwendungen wird keine theoretische Konvergenzbedingung, die ein Erreichen einer Lösung durch das Iterationsverfahren garantiert, angegeben. Bei jeder Änderung der Systemparameter muss daher in solchen Fällen die Schrittweite experimentell angepasst werden (vgl. [PM90, SW00, ABBFC97, Nor89]).

Die lokal optimale Richtung \mathbf{d}_k wird als Richtung des steilsten Abstiegs bezeichnet und ist durch den negativen Gradienten von P gegeben, $\mathbf{d}_k = -\nabla P(\boldsymbol{\xi}_k)$. Ist die Funktion P an einer Stelle $\boldsymbol{\xi}$ in eine Richtung \mathbf{d} differenzierbar und besitzt dort eine positive Richtungsableitung, so wird \mathbf{d} als Abstiegsrichtung von P im Punkt $\boldsymbol{\xi}$ bezeichnet [Kos93]. Lässt man hier auch andere Abstiegsrichtungen als die des Gradienten von P zu, so wird

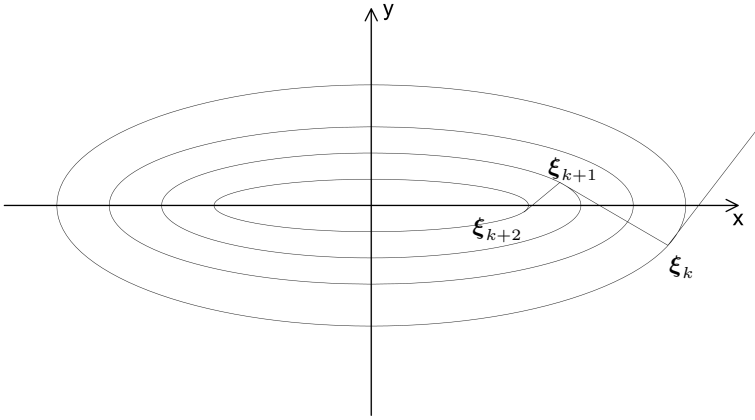


Abbildung 3.11: Höhen- bzw. Äquipotentiallinien einer Potentialfunktion P und qualitativer Konvergenzverlauf bei lokal optimaler Wahl der Schrittweite entsprechend Gleichung (3.18) (nach [Mei99, Hun02]).

das Verfahren als verallgemeinertes Gradientenverfahren bezeichnet. Für die Iteration des klassischen Gradientenverfahren gilt mit einer positiven Schrittweite λ_k

$$\xi_{k+1} = \xi_k - \lambda_k \nabla P(\xi_k), \quad \lambda_k > 0 \quad (3.17)$$

Für die Bestimmung der Schrittweite λ_k sind in der Literatur zahlreiche Ansätze zu finden (siehe z.B. [Mei99, Sch68, Kos93]). Besitzt die Funktion P auf der durch die Abstiegsrichtung \mathbf{d}_k bestimmten Halbgerade ein eindeutiges Minimum, so gewährleistet die Regel der optimalen Schrittweite eine Konvergenz des Verfahrens gegen ein Minimum von P . Ein besonderer Fall ist bei einem Gradienten entsprechend Gleichung (3.15) mit symmetrischer, positiv definiter Matrix \mathbf{A} gegeben. Die lokal optimale Schrittweite ergibt sich dann aus

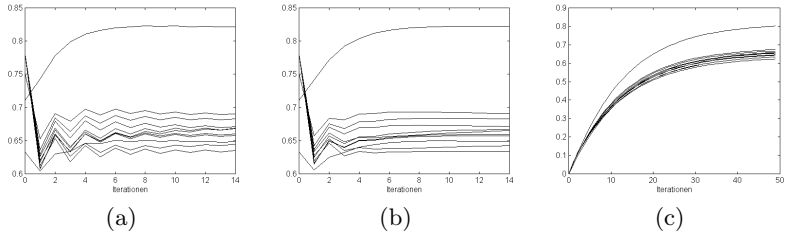


Abbildung 3.12: Dynamik einer Variablenschar mit (a) variabler Schrittweite, (b) konstanter Schrittweite und (c) konstanter Schrittweite mit dem Nullvektor als Startvektor (aus [Hun02])

$$\lambda_k = \frac{\langle \mathbf{r}_k, \mathbf{r}_k \rangle}{\langle \mathbf{A} \mathbf{r}_k, \mathbf{r}_k \rangle} \quad (3.18)$$

mit dem sogenannten Residuenvektor

$$\mathbf{r}_k = -\nabla P(\boldsymbol{\xi}_k) = \boldsymbol{\xi}_0 - \mathbf{A} \boldsymbol{\xi}_k \quad (3.19)$$

Für die Höhenlinien einer Potentialfunktion P ergibt sich bei einer solchen Wahl der Schrittweite der in Abbildung 3.11 dargestellte Verlauf.

Eine weitere Möglichkeit zur Wahl von λ_k besteht darin, eine konstante Schrittweite $\lambda_k = \lambda$ zu wählen. Wie in Abbildung 3.11 zu sehen ist, stellt die lokal optimale Schrittweite nicht unbedingt eine global optimale Schrittweite dar (vgl. [Mei99]). Darüber hinaus muss eine Schrittweite gemäß Gleichung (3.18) in jeder Iteration neu berechnet werden. Die Wahl einer konstanten Schrittweite besitzt daher ihre Berechtigung, was auch in den Abbildungen 3.12(a) und 3.12(b) festzustellen ist. Die in Abbildung 3.12(a) dargestellte Variablenschar, die die Dynamik der Werteverläufe bei Minimierung einer Kostenfunktion entsprechend Gleichung (3.12) mit einer lokal opti-

malen Schrittweite entsprechend Gleichung (3.18) darstellt, weist ein hohes Maß an Oszillation auf. Die mit einer konstanten Schrittweite bestimmten Variablenwerte in Abbildung 3.12(b) verändern sich dagegen bereits nach der achten Iteration nur noch unwesentlich.

Um eine konstante Schrittweite wählen zu können, muss jedoch sichergestellt sein, dass das Iterationsverfahren tatsächlich gegen das Minimum von P konvergiert. Hierzu werden zunächst einige Definitionen eingeführt [Mei99]. Betrachtet wird ein lineares Gleichungssystem der Form $\mathbf{Ax} = \mathbf{b}$.

Ein Iterationsverfahren $\phi : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ wird als linear bezeichnet, falls Matrizen $\mathbf{M}, \mathbf{N} \in \mathbb{R}^{n \times n}$ existieren, so dass

$$\phi(\mathbf{x}, \mathbf{b}) = \mathbf{M}\mathbf{x} + \mathbf{N}\mathbf{b} \quad (3.20)$$

gilt [Mei99, Hun02]. Das Verfahren heißt konsistent zur Matrix \mathbf{A} , wenn für alle $\mathbf{b} \in \mathbb{R}^n$ die Lösung $\mathbf{A}^{-1}\mathbf{b}$ ein Fixpunkt von ϕ zu \mathbf{b} ist. Ein lineares Iterationsverfahren ist genau dann konsistent zur Matrix \mathbf{A} , wenn

$$\mathbf{M} = \mathbf{I} - \mathbf{N}\mathbf{A} \quad (3.21)$$

gilt. Weiterhin ist das Verfahren genau dann konvergent, wenn der Spektralradius $\rho(\mathbf{M})$ der Iterationsmatrix \mathbf{M} die Bedingung

$$\rho(\mathbf{M}) < 1 \quad (3.22)$$

erfüllt. Ist ϕ ein konvergentes und zur Matrix \mathbf{A} konsistentes lineares Iterationsverfahren, so erfüllt das Grenzelement der Folge $\mathbf{x}_k = \phi(\mathbf{x}_{k-1}, \mathbf{b})$ für jeden Startvektor $\mathbf{x}_0 \in \mathbb{R}^n$ das Gleichungssystem $\mathbf{Ax} = \mathbf{b}$.

Betrachten wir nun das Gleichungssystem (3.15) und das Iterationsverfahren (3.17) mit konstanter Schrittweite $\lambda_k = \lambda$, so folgt mit $\mathbf{N} = \lambda \mathbf{I}$ und $\mathbf{M} = \mathbf{I} - \mathbf{N}\mathbf{A}$ die Linearität und die Konsistenz des Iterationsverfahrens zur Matrix \mathbf{A} .

Es bleibt zu zeigen, unter welchen Bedingungen die Schrittweite λ die Bedingung $\rho(\mathbf{M}) < 1$ erfüllt und somit die Konvergenz des Verfahrens sichert. Hierzu wird die Eigenschaft $\rho(\mathbf{M}) \leq \|\mathbf{M}\|$ für jede Matrixnorm $\|\cdot\|$ ausgenutzt [Mei99].

Sei

$$Q_i = \frac{1}{2} \sum_{k=1}^m (w_{ik} + w_{ki}), \quad (3.23)$$

wobei wieder $w_{ij} = 0$ für $j \notin U_i$ und für $i = j$ gelte. Sei ferner Q die maximale Anzahl der Pixel in einer Nachbarschaft U . Aus Gleichung (3.10) folgt damit $Q > Q_i \quad \forall i$.

Die Bedingung (3.22) ist erfüllt, wenn $\|\mathbf{M}\| < 1$ gilt. Unter Verwendung der Zeilensummennorm $\|\cdot\|_\infty$ ergibt sich die Bedingung

$$\sum_{k=1}^{m \cdot d} |m_{ik}| < 1 \quad \text{für jede Reihe } i \quad (3.24)$$

Hieraus folgt mit Beziehung (3.16)

$$|1 - \lambda(1 + c \cdot Q_i)| + \lambda \cdot c \cdot Q_i < 1 \quad (3.25)$$

und damit

$$0 < \lambda < \frac{2}{1 + 2 \cdot c \cdot Q_i} \quad (3.26)$$

Im Folgenden wird, um die Konvergenzbedingung (3.26) zu erfüllen, die Schrittweite λ zu

$$\lambda = \frac{1}{1 + c \cdot Q} \quad (3.27)$$

gewählt. Mit

$$\frac{1}{1 + c \cdot Q} < \frac{2}{1 + 2 \cdot c \cdot Q} \leq \frac{2}{1 + 2 \cdot c \cdot Q_i} \quad (3.28)$$

ist gewährleistet, dass die in Gleichung (3.26) gegebene Konvergenzbedingung erfüllt ist.

Damit ergibt sich eine konstante Schrittweite, die nur von der Skalierung bzw. dem Glättungsfaktor c und der Größe Q der betrachteten Pixelnachbarschaften abhängt. Hierbei ist der Unterschied zu beachten zwischen dem Startvektor des Iterationsverfahrens und dem ursprünglichen Bild ξ_0 . Die Lösung der Gleichung (3.15) ist unabhängig vom Startvektor, wie in Abbildung 3.12(c) zu sehen ist, der Startvektor hat lediglich Einfluss auf die benötigten Iterationen, die zum Erreichen der Lösung notwendig sind. Dagegen hängt die Lösung vom Eingangsbild ξ_0 ab, so dass, wie wiederum in Abbildung 3.12(c) dargestellt ist, die Lösung auch dann erreicht wird, wenn der Nullvektor als Startvektor genommen wird. Da aber offensichtlich das Eingangsbild eine gute Näherung an die gewünschte Lösung darstellt, ist es in der Regel sinnvoll, das Bild ξ_0 als Startvektor zu verwenden.

Der Vollständigkeit halber sei hier auf die Vielzahl weiterer iterativer Verfahren zur Lösung von Gleichung (3.15) hingewiesen, insbesondere ist in diesem Zusammenhang das Verfahren der konjugierten Gradienten zu nennen [Mei99, Kos89, Sch68].

3.3.2 Diffusion

Isotrope Diffusion

Im einfachsten Fall werden in einer quadratischen Nachbarschaft U_i eines Pixels i die Gewichtungsfaktoren symmetrisch zu

$$w_{ij} = w_{ji} = 1 \quad (3.29)$$

gewählt. Gleichung (3.13) vereinfacht sich dann zu

$$\begin{aligned} \frac{\partial}{\partial \xi_{(x,k)}} P(\xi) &= \xi_{(x,k)} - \xi_{(x,k)_0} \\ &+ c \sum_{j \in U_x} (\xi_{(x,k)} - \xi_{(j,k)}) \end{aligned} \quad (3.30)$$

Dies erzeugt eine ungerichtete, also linear isotrope Diffusion des Eingangsbildes und stellt somit den klassischen Fall einer einfachen Glättung dar.

Sowohl die isotrope Diffusion als auch ein Ansatz zur anisotropen Diffusion mit symmetrischen Gewichten, d. h. $w_{ij} = w_{ji}$, wurden bereits in [Hun02, HBD03, BHM04b, BHM04a, BHM05] im Kontext stereoskopischer Korrespondenzbestimmung genutzt. Dabei dienten Gleichung (3.30) bzw. (3.17) dazu, Zuordnungsmehrdeutigkeiten in einem stereoskopischen Disparitätsraum aufzulösen.

Abbildung 3.13 verdeutlicht die Eigenschaften linearer isotroper Diffusion, wiederum am Beispiel des Lena-Bildes. Mit zunehmenden Glättungsparameter c wird das Eingangsbild richtungsunabhängig geglättet, bis sämtliche Bilddetails verschwunden sind.



Abbildung 3.13: Linearer isotroper Skalenraum für das Lena-Bild: (a) $c=1$, (b) $c=3$, (c) $c=20$ und (d) $c=250$.

Anisotrope Diffusion

Ein wichtiger Vorteil der Formulierung des Diffusionsprozesses in Gleichung (3.12) ist die Tatsache, dass unter Beibehaltung der Bedingung (3.10) die Gewichtungsfaktoren beliebig gewählt werden können, ohne dass die Existenz oder die Eindeutigkeit der Lösung verloren geht. Um eine anisotrope Diffusion, also eine gerichtete Glättung des Eingangsbildes zu erreichen, wird die folgende Umfeldgewichtung eines gegebenen Pixels eingeführt. Hierzu werden zu jedem Pixel die Orientierungen betrachtet, die sich aus Anwendung der in Kap. 3.1 vorgestellten orientierungsselektiven Kantenfilter ergeben.

In Abhängigkeit des Orientierungswinkels eines gegebenen Pixel i findet eine Rotation des quadratischen Umfeldes U_i statt. Wird hierbei ein Punkt j des Umfeldes auf die Koordinaten (x, y) abgebildet, so ergibt sich der Gewichtungsfaktor w_{ij} des entsprechenden Umfeldpixels durch die folgende Funktion:

$$f(x, y) = \begin{cases} 0 & \text{für } y > 0.7 \\ 1.0 + \cos\left(\frac{\pi \cdot y}{0.7}\right) & \text{sonst} \end{cases} \quad (3.31)$$



Abbildung 3.14: Anisotroper Skalenraum für das Lena-Bild mit 20% unkorreliertem Rauschen für jedes Pixel des Eingangsbildes: (a) Eingangsbild, (b) $c=3$, (c) $c=20$ und (d) $c=250$.

Gleichung (3.31) gewährleistet eine strikte Diffusion entlang einer Richtung. Eine stärkere Kopplung von Bildpunkten, deren Verbindung nicht parallel zu einer durch den Orientierungswinkel charakterisierten Geraden verläuft, lässt sich erreichen, indem entweder die Funktion in Gleichung (3.31) in positive y -Richtung verschoben wird oder die Konstante 0.7 erhöht wird.

Es gilt

$$Q_i = \frac{1}{2} \sum_k (w_{ik} + w_{ki}) < Q, \quad (3.32)$$

womit entsprechend Gleichung (3.28) die Konvergenz des Verfahrens gesichert ist. Die Funktion (3.31) ist also wohldefiniert.

Abbildung 3.14 zeigt das Verhalten des anisotropen Diffusionsprozesses. Jetzt wird zu jedem Pixel des Eingangsbildes 20% unkorreliertes Rauschen addiert. Betrachtet man den Skalenraum, so fällt auf, dass für $c \rightarrow \infty$, anders als im Fall linearer isotroper Diffusion, eine lokale Nachbarschaft definiert werden kann, für die das Verfahren nicht gegen einen für alle Pixel identischen, konstanten Farbwert konvergiert.

3.3.3 Selbstorganisation der Kanteninformation

Wie schon erwähnt, werden die benötigten Orientierungswinkel zur Bestimmung der Gewichtungsfaktoren w_{ij} über die Filterergebnisse gaußförmiger Filtermasken gewonnen. Auch bei großen Filtermasken bleibt jedoch eine gewisse Empfindlichkeit gegenüber Rauschprozessen gegeben. Um diesem Problem Rechnung zu tragen, lässt sich die anisotrope Regularisierung, wie sie in Kap. 3.3.1 beschrieben wird, auch für die von den Filterantworten generierte Gradientenkarte nutzen [HM08].

Das bedeutet, dass Gleichung (3.12) für $d_{\text{neu}} = d + 2$ auf die Kanteninformation in Form einer Gradientenkarte angewandt wird. Die Gradientenkarte wird hierbei nach den eigentlichen Farbwerten in die letzten beiden Schichten, d. h. die durch $d + 1$ und $d + 2$ charakterisierten Schichten des Bildvektors

$$\xi = (\xi_{(1,1)}, \dots, \xi_{(m,d)}, \xi_{(1,d+1)}, \dots, \xi_{(m,d+1)}, \xi_{(1,d+2)}, \dots, \xi_{(m,d+2)})^T \quad (3.33)$$

geschrieben. Die Werte der Gradientenkarte werden durch

$$\xi_{(x,d+1)} = |\nabla g_x| \cdot \cos(2 \cdot \alpha_x) \quad (3.34)$$

und

$$\xi_{(x,d+2)} = |\nabla g_x| \cdot \sin(2 \cdot \alpha_x) \quad (3.35)$$

für eine Bildposition $x \in [1, \dots, m]$ beschrieben. Hierbei stellt $|\nabla g_x|$ das Betragsmaximum der Kantenfilter dar, während α_x die vom betragsgrößten Kantenfilter repräsentierte Orientierung an der Bildposition x darstellt. Die ursprünglichen Werte werden wieder in einem Vektor ξ_0 zusammengefasst. Da die Orientierungen der Kantenfilter hier lediglich in einem Bereich von

0 bis π betrachtet werden, gewährleistet der Faktor 2 in den Gleichungen (3.34) und (3.35), dass auch an den Bereichsgrenzen eine kontinuierliche Glättung der Gradientenkarte stattfindet.

Während des Iterationsprozesses wird nun Gleichung (3.12) in jedem Iterationsschritt neu formuliert, derart, dass zu jedem Bildpunkt x aus den Werten $\xi_{(x,d+1)}$ und $\xi_{(x,d+2)}$ aus den Gleichungen (3.34) und (3.35) der Winkel α_x neu gebildet wird. Gerade der Winkel α_x ist aber durch die Rotation der Umfeldkoordinaten und damit durch die Bestimmung der Gewichtungsfaktoren w_{ij} ausschlaggebend für die Form des koppelnden Umfelds eines Pixels. Es werden somit also in jedem Iterationsschritt die Gewichtungsfaktoren in Gleichung (3.12) in Abhängigkeit des Vektors ξ aktualisiert. Durch einen nachfolgenden Iterationsschritt des Gradientenverfahrens wird wiederum der Vektor ξ in Abhängigkeit der Gleichung (3.12) und damit in Abhängigkeit der Gewichtungsfaktoren aktualisiert.

Hierdurch ergibt sich ein dynamischer Selbstorganisationsprozess der Kanteninformation, der zu einer Verstärkung salienter Konturen führt. Dies ist in Abbildung 3.15 verdeutlicht. Die präsentierten Resultate wurden mit einem Kopplungsparameter $c = 6$, einem quadratischen Kopplungsumfeld von 7×7 Pixeln und 100 Iterationen berechnet. Das Eingangsbild 3.15(a), drei Kieselsteine auf Pflastersteinen, ist ein Testbild aus dem Kontext der Verstärkung salienter Konturen (vgl. [MWTX03]) und weist im Bereich der Pflastersteine einen stark texturierten Bereich mit hochfrequenten Bildanteilen auf. Die Filterergebnisse setzen sich aus 24 unterschiedlich orientierten Kantenfiltern mit einem Filterradius von jeweils sieben Pixeln zusammen, was einer Filtermaske von 14×14 Pixeln entspricht. Trotzdem lassen sowohl die Betragsmaxima der Kantenfilter in Abbildung 3.15(e) als auch die Winkelinformation in Abbildung 3.15(c) erkennen, dass die hochfrequenten Bildanteile ein korrektes Erkennen der Konturen der Kieselsteine und der Pflastersteine erschweren. Die Textur stellt hierbei im eigentlichen Sinne kein Rauschen dar, da die hochfrequenten Bildanteile durch sich in dem Beton befindliche kleinere Steine erzeugt werden und somit Teil des Signals sind.

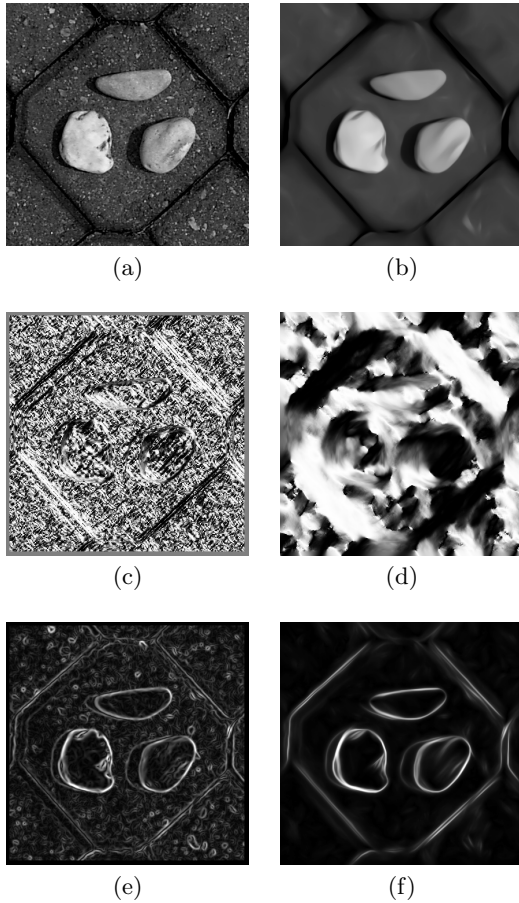


Abbildung 3.15: Testbild mit drei Kieselsteinen auf Pflastersteinen (aus [MWTX03]): (a) und (b) Grauwerte des Testbildes, (c) und (d) Winkelinformation und (e) und (f) Betragsmaximum $|\nabla g_x|$ der Kantenfilter, jeweils vor und nach der Selbstorganisation der Gradientenkarte.

Dennoch ist es in Bildverarbeitungsprozessen wünschenswert, Signalanteile auszublenden, die aufgrund des betrachteten Maßstabs bzw. entsprechend einer gewünschten Skalierung keinen Beitrag zu weiteren Verarbeitungsschritten leisten. Im Fall der Abbildung 3.15 ist es also sinnvoll, Bildinformation einer kleineren Skalierung zu eliminieren, um verlässliche Kanteninformationen über Objekte in der Größenordnung der drei Kieselsteine zu erhalten. Wie in Abbildung 3.15(f) zu sehen ist, führt der beschriebene Selbstorganisationsprozess der Gradientenkarte zu solch einem Ergebnis. Dargestellt ist zu jedem Bildpunkt x die euklidische Norm $\|(\xi_{(x,d+1)}, \xi_{(x,d+2)})\|_2$ der entsprechenden Werte der Gradientenkarte. Der Winkel, welcher sich nach dem Selbstorganisationsprozess zu jedem Bildpunkt aus den Werten $\xi_{(x,d+1)}$ und $\xi_{(x,d+2)}$ ergibt, ist in Abbildung 3.15(d) dargestellt. Auch die Grauwerte selbst verdeutlichen, dass das Bild auf wesentliche Elemente reduziert wird. In Abbildung 3.15(b), die die Bildwerte nach Selbstorganisationsprozess zeigt, sind die hochfrequenten Bildanteile herausgefiltert, die Position und Form wesentlicher Konturelemente hat sich im Gegensatz zum Ergebnis einfacher Glättungsfilter jedoch nicht verändert.

Da durch dieses Vorgehen die Systemmatrix \mathbf{A} durch die Winkelabhängigkeit der Umfeldkopplung vom Zustandsvektor $\boldsymbol{\xi}$ abhängt, $\mathbf{A} = \mathbf{A}(\boldsymbol{\xi})$, ergibt sich ein nichtlineares System, wobei allerdings aufgrund der Wahl der Schrittweite λ in Gleichung (3.27) in jedem Iterationsschritt eine Kontraktion vorliegt. Für den Spektralradius gilt nach wie vor für jede Iteration $\rho(\mathbf{M}(\boldsymbol{\xi})) < 1$, womit die Iteration gegen einen Fixpunkt konvergiert.

Um eine Verstärkung der Diskontinuitäten entlang von Konturverläufen zu erreichen, ist es möglich, die Gewichtungsfaktoren w_{ij} benachbarter Bildpunkte i und j zu Null zu setzen, falls diese sich in ihrem Farbwert mindestens um einen Schwellwert s unterscheiden, d. h. falls $|\xi_{(i,k)} - \xi_{(j,k)}| < s$ gilt. Es geht dabei allerdings die Eindeutigkeit der Lösung verloren, das Ergebnis des Iterationsverfahrens hängt stark vom Startvektor ab. Eine gute Wahl für den Startvektor ist der Nullvektor oder eine geglättete Version des Eingangsbildes. Wird dagegen das Ursprungsbild $\boldsymbol{\xi}_0$ als Startvektor

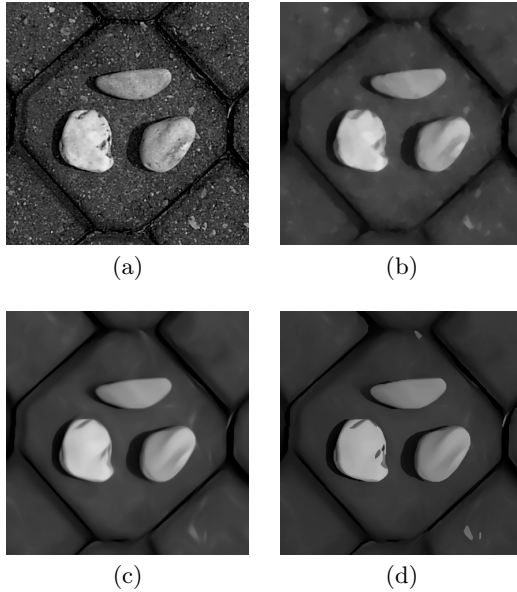


Abbildung 3.16: Steine-Testbild: (a) Eingangsbild (b) Ergebnis der anisotropen Regularisierung, (c) Ergebnis anisotroper Regularisierung mit Selbstorganisation der Gradientenkarte entsprechend Abbildung 3.15 und (d) Ergebnis anisotroper Regularisierung mit Selbstorganisation der Gradientenkarte und Unterdrückung der Kopplung bei Erreichen eines Schwellwerts von $s = 5$.

genommen, so ergibt sich für einen kleinen Schwellwert selbst bei hohem Koppungsparameter kaum ein Effekt. Der Einfluss des Schwellwertes ist in Abbildung 3.16 dargestellt. Während das Ergebnis der anisotropen Regularisierung in Abbildung 3.16(b) und das Ergebnis anisotroper Regularisierung mit Selbstorganisation der Gradientenkarte in Abbildung 3.16(c) mit den oben beschriebenen Parametern berechnet wurden, wurde für das Ergebnisschwellwertabhängiger Umfeldkopplung in Abbildung 3.16(d) der Nullvek-

tor als Startvektor mit einer Iterationsanzahl von 500 gewählt. Es ist zu erkennen, dass hier die Grenzen zwischen zwei Flächen im Bereich salienter Konturen durch eine stärkere Diskontinuität des Grauwertverlaufs noch klarer hervortreten. Weiterhin lässt sich feststellen, dass kleinere Gebiete, die einen starken Grauwertunterschied zu ihrer Umgebung aufweisen, speziell für kleine Schwellwerte s die Tendenz zeigen, nicht durch den Diffusionsprozess aufgelöst zu werden, was je nach Anwendung durchaus unerwünscht sein kann.

4 Bestimmung möglicher Konturvervollständigungen

Durch die in Kap. 3 beschriebene Vorverarbeitung stehen zu sämtlichen Kreuzungspunkten sowohl die Bildkoordinaten als auch die Richtungen zur Verfügung, in die die mögliche Fortsetzung einer unterbrochenen Kontur verläuft. Ziel ist es nun, Kreuzungspunkte, die nach den Gestaltprinzipien der Nähe und der guten Fortsetzung mit aller Wahrscheinlichkeit Teil derselben Kontur sind, in geeigneter Weise miteinander zu verbinden. Hierzu werden zwei Mechanismen betrachtet, zum einen das bereits in Kap. 2 angesprochene Tensor Voting Verfahren und zum anderen eine Spline-Interpolation des Verbindungsweges. Besondere Beachtung findet hierbei die Betrachtung der Bedingungen, unter denen zwei Kreuzungspunkte miteinander verbunden werden sollten. Ein Vergleich der beiden alternativen Vorgehensweisen findet in Kap. 6 statt.

4.1 Konturvervollständigung durch Tensor Voting

In [MLT00] beschreiben Medioni, Lee und Tang eine als Tensor Voting bezeichnete Methode zur Merkmalsverstärkung auf dünn besetzten und verrauschten Eingabedaten. Das wesentliche Merkmal ist hierbei die Beschreibung von Bild- bzw. Kantenelementen als Tensoren.

4.1.1 Repräsentation von Konturelementen durch Tensoren

Im zweidimensionalen Fall kann ein Tensor über \mathbb{R}^2 durch eine symmetrische 2×2 Matrix \mathbf{T} mit zwei zueinander rechtwinkligen Eigenvektoren $\mathbf{e}_1, \mathbf{e}_2$

und zwei entsprechenden reellen Eigenwerten $\lambda_1 > \lambda_2$ beschrieben werden. Ein Tensor, welcher ein Kurvenelement repräsentiert, lässt sich im zwei-dimensionalen Fall als eine Ellipse darstellen (vgl. Abbildung 4.1), wobei die Hauptachse die Richtung \mathbf{e}_1 der Tangente an die Kurve angibt und die Länge λ_1 die Salienz bzw. Stärke dieser Richtung repräsentiert. Die Länge λ_2 , welche dem zweiten Eigenvektor \mathbf{e}_2 zugeordnet ist, stellt ein Maß für die Orientierungsunsicherheit an der entsprechenden Bildposition dar, wie sie z.B. an Eckpunkten auftritt.

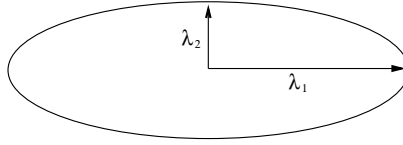


Abbildung 4.1: Visualisierung eines Tensors als Ellipse.

Die Definition der Salienzmaße ergibt sich aus der spektralen Dekomposition des Tensors durch

$$\mathbf{T} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^\top + \lambda_2 \mathbf{e}_2 \mathbf{e}_2^\top \quad (4.1)$$

bzw.

$$\mathbf{T} = (\lambda_1 - \lambda_2) \mathbf{e}_1 \mathbf{e}_1^\top + \lambda_2 (\mathbf{e}_1 \mathbf{e}_1^\top + \mathbf{e}_2 \mathbf{e}_2^\top). \quad (4.2)$$

Der Faktor $(\lambda_1 - \lambda_2)$ repräsentiert hierbei die Orientierungssicherheit in Richtung des Eigenvektors und wird daher als Kurvensalienz bzw. Sticksalienz bezeichnet. Bei maximaler Orientierungsunsicherheit gilt $\lambda_1 = \lambda_2$ und die Ellipse wird zu einem Kreis. Der Faktor λ_2 wird daher als Ballsalienz bezeichnet. Er gibt also ein Maß für die Orientierungsunsicherheit, speziell an Kreuzungspunkten, an. Es ist allerdings zu beachten, dass im Bereich von

Kreuzungspunkten die Tensorrepräsentation weder geeignet ist, die mindestens zwei unterschiedlichen auftretenden Orientierungen zu repräsentieren, noch stellt sie aufgrund der Orientierungsabhängigkeit ein verlässliches Maß für Kreuzungspunkte dar [MBM03], wie später erläutert wird.

4.1.2 Tensorkommunikation

Die Gruppierung von Merkmalselementen kann nun als Kombination der Elemente in Abhängigkeit ihrer Stick- bzw. Ballsalienz realisiert werden. Beim sogenannten Stick Voting wird von jedem orientierten Eingangselement ein als Stick-Votingfeld bezeichnetes Tensorfeld entsprechend der Richtung des Eigenvektors e_1 auf die Ergebniskarte addiert. Für die hier betrachtete Anwendung der Konturvervollständigung verdeckter Konturen bestehen die Eingangselemente aus Tensoren, die an der Position von Kreuzungspunkten entsprechend der in Kap. 3 bestimmten Vervollständigungsrichtung gerichtet sind, wobei für jede Vervollständigungsrichtung eines Kreuzungspunktes ein Stick-Votingfeld erzeugt und auf die Ergebniskarte addiert wird. Über diese Addition findet eine Überlagerung der Tensorfelder und damit eine Addition mehrerer Tensoren pro Bildposition statt. Damit ergibt sich in der Ergebniskarte die Orientierung eines Tensors an einer gegebenen Bildposition aus der Addition der aus den überlagerten Votingfeldern stammenden Matrizen und anschließender spektraler Dekomposition der Ergebnismatrix in Eigenvektoren und -werte.

Wie bereits in Kap. 2.5 erläutert, wird für eine sinnvolle Anwendung der Konturvervollständigung, die über eine einfache Konturverstärkung hinausgeht, ein gerichtetes Kommunikationsschema benötigt, welches im Gegensatz zu einem bipolaren Verbindungsschema nur in einer Richtung Werte ungleich Null im Stick-Votingfeld erzeugt (vgl. Abbildung 2.15). Ziel des Votingprozesses ist die Gruppierung benachbarter Elemente unter Beachtung des Gestaltprinzips der guten Fortsetzung. Daher wird die Orientierung jedes Tensors des Votingfeldes so gewählt, dass diese eine Tangente an

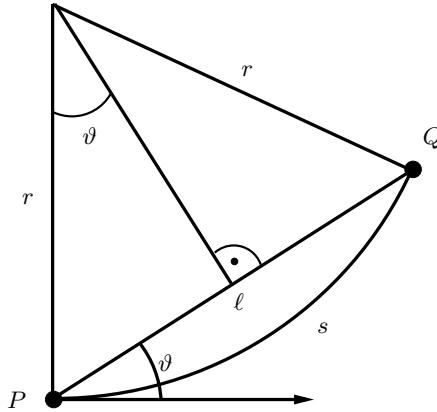


Abbildung 4.2: Skizze zur Bestimmung der Sticksalienz der Elemente des Votingfeldes und zur Bestimmung des Kreisbogens s und des Radius r .

den Kreis darstellt, der durch den betrachteten Punkt des Votingfeldes und durch das Zentrum des Feldes und damit durch den Punkt verläuft, der das Votingfeld erzeugt (siehe Abbildung 4.2).

Die Stärke des Feldes, d. h. der Wert des Eigenwertes λ_1 jedes Tensors der Maske, wobei $\lambda_2 = 0$ gilt, ergibt sich aus der folgenden Betrachtung. Gegeben sei ein Punkt P mit einer assoziierten Tangenten- bzw. Vervollständigungsrichtung. Weiter sei ein Punkt Q gegeben, wobei die Winkeldifferenz zwischen der Tangentenrichtung und der direkten Verbindung der Punkte P und Q mit ϑ bezeichnet sei. Sei ferner ℓ der Abstand der Punkte P und Q . Zwischen dem oben beschriebenen Winkel ϑ und dem Radius r des Kreises, der durch die Punkte P und Q verläuft und die mit P assoziierte Tangente aufweist, besteht der Zusammenhang

$$\sin(\vartheta) = \frac{\ell}{2r}. \quad (4.3)$$

Damit folgt für den Radius r

$$r = \frac{\ell}{2 \cdot \sin(\vartheta)} \quad (4.4)$$

Für die Länge des Kreisbogens s gilt im Bogenmaß

$$s = 2\pi r \cdot \frac{2\vartheta}{2\pi} = 2r\vartheta, \quad (4.5)$$

woraus mit Gleichung (4.4)

$$s = \frac{\ell \cdot \vartheta}{\sin(\vartheta)} \quad (4.6)$$

folgt.

Bestehende Ansätze zur Konturverstärkung oder Konturvervollständigung definieren die Verbindungsstärke V der Punkte P und Q und damit die Form des Kommunikationsschemas über eine Funktion der Form

$$V = V_d \cdot V_c \quad (4.7)$$

mit einem Term V_d zur Bewertung der Distanz und einem Term V_c zur Bewertung der Krümmung.

In [HH93], [HHP⁺98] und [PH01] benutzen Heitger et al.

$$V_{d1} = e^{-\frac{\ell^2}{2\sigma^2}} \quad (4.8)$$

und

$$V_{c1} = \begin{cases} \cos^k\left(\frac{\pi/2}{\alpha} \cdot \vartheta\right) & \text{für } |\vartheta| < \alpha \\ 0 & \text{sonst} \end{cases}, \quad (4.9)$$

wobei der Parameter σ in Gleichung (4.8) den Einfluss der Entfernung auf die Verbindungsstärke regelt, während durch den Parameter $k = 2n$, $n \in \mathbb{N}$ mit einem Öffnungswinkel $2\alpha = \pi$ die Öffnungsweite des Feldes geregelt wird. Hansen und Neumann benutzen ebenfalls V_{d1} und V_{c1} , allerdings mit $k = 1$ und $\alpha = 10^\circ$ [HN01]. In [MLT00] definieren Medioni et al. den Distanzbewertungsterm V_{d2} und den Krümmungsterm V_{c2} wie folgt:

$$V_{d2} = e^{-\frac{s^2}{2\sigma^2}} \quad (4.10)$$

und

$$V_{c2} = e^{-\frac{c \cdot \rho^2}{\sigma^2}} \quad \text{mit} \quad \rho = \frac{2 \sin(\vartheta)}{\ell} \quad (4.11)$$

Hier wird also statt der direkten Entfernung der Punkte P und Q die Strecke s entlang des Kreisbogens zur Distanzbewertung herangezogen. Für das Krümmungsmaß stellt c eine positive Konstante dar, während ρ , wie unmittelbar aus Gleichung (4.4) folgt, nichts anderes als der Kehrwert des Radius r ist (vgl. Abbildung 4.2). Hieraus folgt, dass bei gleichem Winkel ϑ und unterschiedlichen Abständen ℓ der Punkte P und Q das Krümmungsmaß unterschiedliche Werte liefert und damit stark von der Skalierung des betrachteten Bildes abhängt.

Für das Kommunikationsschema des Tensorvotings muss eine aus Tensoren bestehende Maske erzeugt werden, wobei, wie schon erwähnt, der Eigenvektor \mathbf{e}_1 in jedem Punkt der Maske die Tangentenrichtung an einen entsprechenden Kreis durch den Punkt und durch den Zentrumsunkt repräsentiert. Der Eigenwert λ_1 wird durch $\lambda_1 = V = V_d \cdot V_c$ gebildet. Um eine Abhängigkeit des Distanzmaßes von der Länge entlang eines potenziellen Kurvenstücks zu erzielen, wird für die in Kap. 6 vorgestellten Ergebnisse $V_d = V_{d2}$ gewählt. Um andererseits eine klare Trennung zwischen Distanz und Krümmung zu erzielen, wird eine skalierungsinvariante Krümmungsbewertung durch $V_c = V_{c1}$ mit $k = 1$ realisiert, wobei die Öffnungsweite des

Feldes direkt über den Öffnungswinkel $2 \cdot \alpha$ gesteuert wird [HM05]. Alle nachfolgenden Ergebnisse wurden mit $\alpha = 15^\circ$, d. h. einem Öffnungswinkel von 30° berechnet.

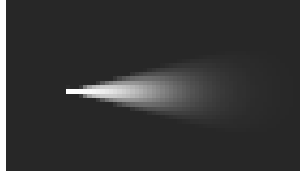


Abbildung 4.3: Sticksalienz für ein halbseitiges Votingfeld mit $V = V_{d2} \cdot V_{c1}$, $k = 1$ und $\alpha = 15^\circ$.

Die Sticksalienz eines solchen Votingfeldes, also der Wert $\lambda_1 - \lambda_2$ ist in Abbildung 4.3 dargestellt. Hierbei gilt aufgrund der idealen Orientierung innerhalb der Maske $\lambda_2 = 0$. Ausgehend von Kreuzungspunkten wird eine Konturvervollständigung nun realisiert, indem jede Vervollständigungsrichtung entsprechend ihrer Winkelinformation ein halbseitiges Votingfeld auf die Ergebnistensorkarte addiert (vgl. Abbildung 4.4). Ergibt sich durch eine Flankenmaximumsuche über die Ergebniskarte eine Verbindungslinie zwischen zwei Kreuzungspunkten, so wird diese als virtuelle Kontur den eigentlichen Konturen hinzugefügt.

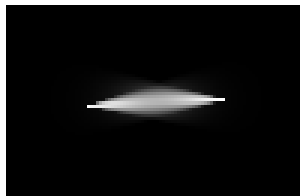


Abbildung 4.4: Sticksalienz der Überlagerung zweier Votingfelder.

4.1.3 Probleme des Tensorvoting

Durch die Repräsentation von Eingabeelementen als Tensor kann eine Vorzugsorientierung, deren Salienz und die Orientierungsunsicherheit, z. B. an Kreuzungspunkten dargestellt werden. So wird in [MBM02b, MBM02a, MBM03] vorgeschlagen, Kreuzungspunkte über eine hohe Ballsalienz der Tensorkarte nach einem Votingprozess auf der dichtbesetzten Eingangskarte zu erkennen.

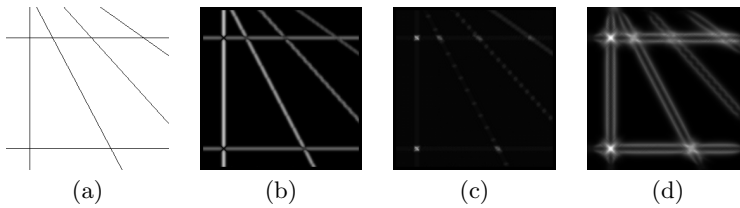


Abbildung 4.5: Tensor Voting auf einem Linienbild: (a) Eingangsbild, (b) Eingangssticksalienzen, (c) Ballsalienzen vor und (d) nach dem Stickvoting.

Die Eingangstensoren entstehen dabei für dichtbesetzte Karten allerdings durch eine Überlagerung von Tensoren, d. h. durch Tensoraddition. In [GM96] geschieht dies in Binärbildern durch Überlagerung von Kurvenelementen, in [MBM02b] durch die mit Gaborfilterantworten gewichtete Summe über ideal orientierte Tensoren. Ebenso entstehen die Ausgangstensoren durch Tensoraddition, d. h. durch einen Votingprozess. Bei einer Überlagerung von Tensoren, die die gleiche Salienz besitzen und in ihrer Orientierung eine Differenz von 90° aufweisen, erhält man einen idealen Balltensor und eine Sticksalienz von Null. Ist dagegen die Orientierungsdifferenz geringer, erhält man stattdessen eine geringere Ballsalienz und eine erhöhte Stick Salienz mit einer Orientierung, die aus der Mittelung der erzeugenden Tensoren entsteht.

Dies ist in Abbildung 4.5 dargestellt. Zu dem Eingangsbild in Abbildung 4.5(a) stellt Abbildung 4.5(b) die Sticksalienz der Tensorkarte dar, die sich aus der Multiplikation von Filterantworten orientierungsselektiver Filter mit entsprechend orientierten idealen Tensoren, d. h. $\lambda_2 = 0$, und anschließender Überlagerung sämtlicher Orientierungen ergibt. Die dazugehörige Ballsalienz ist in Abbildung 4.5(c) dargestellt und lässt bereits erkennen, dass sich für abnehmende Schnittwinkel zwischen den Geraden die Ballsalienzen verringern und eine Erkennung von Kreuzungspunkten über ein Maximum der Ballsalienz erschweren bis unmöglich machen. Gerade in Bildern mit hochfrequenten Bildinhalten wie z. B. Abbildung 3.15 ist also eine Detektion von Eck- oder Kreuzungspunkten mit Hilfe des Tensorvoting nicht ratsam. Auch nach erfolgtem Votingprozess bleibt eine Erkennung für nicht rechtwinklige Kreuzungspunkte schwierig, wie Abbildung 4.5(d) zeigt. Zu der winkelabhängigen Ballsalienz kommt hier auch eine Lokalisierungsungenauigkeit hinzu.

Weiterhin reicht eine Tensordarstellung nicht aus, die im Allgemeinen mehreren Orientierungen an Kreuzungspunkten darzustellen, hier muss zusätzlich ein Vektor oder ein Winkelwert pro Vervollständigungsrichtung gespeichert werden. Die Gestaltprinzipien der Nähe und guten Fortsetzung sind integraler Bestandteil des Votingprozesses, es erscheint daher für die hier beschriebene Konturvervollständigung sinnvoll, den Votingprozess, ausgehend von Kreuzungspunkten, auf einer einzigen Tensorkarte auszuführen (vgl. [Mas06a, Mas06b]) und nicht für jedes Paar von Vervollständigungsrichtungen eine separate Karte zu füllen, was den Berechnungsaufwand erheblich erhöhen würde. Wie aber bereits in Kap. 2.5 angesprochen wurde, besteht bei der Nutzung einer gemeinsamen Tensorkarte die offensichtliche Gefahr, dass sich die Tensoren einer modalen und einer amodalen Vervollständigung am Schnittpunkt der Konturen überlagern und damit die Linienerfolgung sowohl der modalen, als auch der amodalen Kontur unmöglich machen. Um speziell dem letzten Punkt Rechnung zu tragen, wird im Folgenden durch eine Splineinterpolation eine Alternative zum Tensorvoting vorgestellt.

4.2 Bewertung von Verbindungswahrscheinlichkeiten

Im Gegensatz zum Tensor Voting, wo das Prinzip der Nähe und der guten Fortsetzung integraler Bestandteil der Konturvervollständigung sind, muss für die in Kap. 4.3 beschriebene Splineinterpolation zunächst explizit bestimmt werden, welche Verbindungen sinnvoll sind, und welche Kreuzungspunkte nicht verbunden werden sollen.

4.2.1 Ausschlusskriterien

Es wird zunächst eine Reihe notwendiger Bedingungen an eine physikalisch sinnvolle Verbindung zweier Konturelemente betrachtet, aus der sich Ausschlusskriterien für eine solche Verbindung ableiten lassen (vgl. [Fel06]). Neben einer optionalen Skalierungsprüfung, welche eine Verbindung zu weit entfernter Kreuzungspunkte entsprechend einer gewünschten Skalierung ausschließt, ist hier die relative Lage zweier Verbindungsrichtungen zueinander von Interesse.

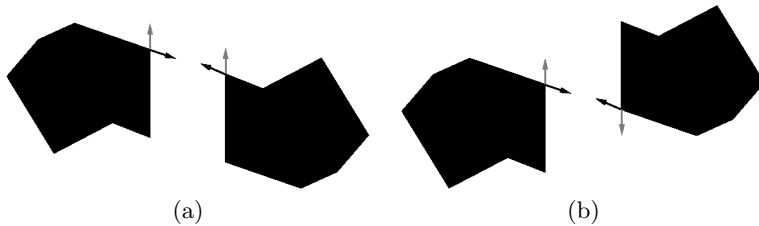


Abbildung 4.6: Unterschiedliche Nachbarorientierungen an Eckpunkten.

Betrachtet man in Abbildung 4.6 die schwarz dargestellten Richtungen möglicher Vervollständigungen, so lässt sich feststellen, dass diese sowohl in Abbildung 4.6(a) als auch in Abbildung 4.6(b) in idealer Weise zueinander ausgerichtet sind. Es ist allerdings unwahrscheinlich, dass Abbildung 4.6(b) durch eine Verdeckung entstanden ist und die betrachteten Eckpunkte damit

Teil derselben Kontur sind. Um einen Fall wie in Abbildung 4.6(b) auszuschließen, wird daher nun zum einen die Verbindungsgerade zwischen den beiden Eckpunkten betrachtet und zum anderen die hier grau dargestellte jeweilige Nachbarorientierung der betrachteten Vervollständigungsrichtung.

Liegen die in Abbildung 4.6 grau dargestellten Nachbarorientierungen auf unterschiedlichen Seiten der Verbindungsgerade, wird eine Verbindung der Kreuzungspunkte über die schwarz dargestellten Orientierungen ausgeschlossen.

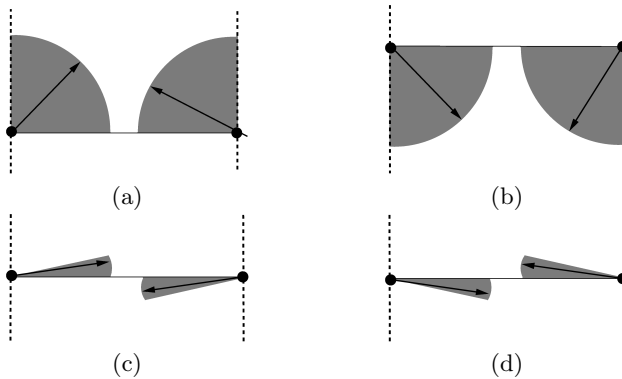


Abbildung 4.7: Erlaubte Bereiche für die Lage der Tangentenvektoren.

Das zweite Ausschlusskriterium erhält man über die Betrachtung des zulässigen Bereichs, in dem eine Verbindung zweier Vervollständigungsrichtungen Sinn macht. Diese müssen auf derselben Seite der Verbindungsgerade liegen und dabei einander zugewandt sein. Der hieraus folgende zulässige Bereich ist in den Abbildungen 4.7(a) und (b) grau dargestellt, wobei sich die Anordnungen nur um eine Spiegelung an der Verbindungsgeraden unterscheiden. Im praktischen Fall kann für die in Kap. 3 ermittelten Winkelwerte der Fall eintreten, dass die Vervollständigungsrichtungen bis auf eine geringe Abweichung in Richtung des jeweils anderen Kreuzungspunktes zeigen, aber auf

unterschiedlichen Seiten der Verbindungsgerade liegen (Abbildungen 4.7(c) und (d)). Hierzu wird der grau gezeichnete Toleranzbereich eingeführt, innerhalb dessen die Vervollständigungsrichtungen auch auf gegenüberliegende Seiten der Verbindungsgerade liegen können. Alle Verbindungswinkel, die die oben genannten Bedingungen nicht erfüllen, können somit ausgeschlossen werden.

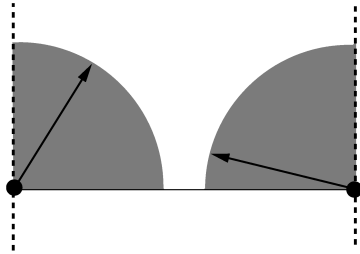


Abbildung 4.8: Ausschluss der Verbindung bei unsymmetrischen Verbindungsrichtungen.

Schließlich soll noch der Fall unterbunden werden, dass die Verbindung der gerichteten Konturelemente einen zu unsymmetrischen Verlauf aufweist, wie in Abbildung 4.8 dargestellt ist. Daher werden zu zwei Kreuzungspunkten i und j die Winkel γ_i und γ_j betrachtet, welche den jeweils kleinsten Winkel zwischen Vervollständigungsrichtung am Kreuzungspunkt und Verbindungsgerade der Kreuzungspunkte darstellen. Mit der Winkeldifferenz

$$\Delta\gamma = |\gamma_i - \gamma_j| \quad (4.12)$$

lässt sich nun das Ausschlusskriterium $\Delta\gamma > \Delta\gamma_{max}$ angeben, wobei $\Delta\gamma_{max}$, falls nicht anders angegeben, im Folgenden zu $\Delta\gamma_{max} = 45^\circ$ gesetzt wird, was der Anordnung in Abbildung 4.8 entspricht.

4.2.2 Bestimmung der Verbindungswahrscheinlichkeit

Zu einer gegebenen Vervollständigungsrichtung φ_i müssen sämtliche Vervollständigungsrichtungen φ_j , die anderen Kreuzungspunkten zugeordnet sind, daraufhin überprüft werden, ob eine Verbindung der Kreuzungspunkte über die Richtungen φ_i und φ_j sinnvoll ist. Für die Menge an Vervollständigungsrichtungen φ_j , welche nicht durch die Kriterien in Kap. 4.2.1 ausgeschlossen wurden, muss somit ein Maß p_{ij} definiert werden, welches die Bewertung der Plausibilität der Verbindung zweier Konturelemente über die Vervollständigungsrichtungen φ_i und φ_j ermöglicht. Ein Verbindungsweg über die Richtungen φ_i und φ_j wird dann generiert (vgl. Kap. 4.3), wenn sowohl

$$p_{ij} = \max_{k \in U_i} \{p_{ik}\} \quad (4.13)$$

als auch

$$p_{ji} = \max_{k \in U_j} \{p_{jk}\} \quad (4.14)$$

gilt, wenn also für jede der beiden Verbindungsrichtungen die Verbindung mit der jeweils anderen Richtung als am wahrscheinlichsten eingestuft wurde. Die Menge U_i bezeichnet hierbei die Menge aller potentieller zu φ_i korrespondierender Verbindungsrichtungen, welche nicht über die Kriterien in Kap. 4.2.1 ausgeschlossen wurden. Das Maß p_{ij} wird im Folgenden so definiert, dass für den Wertebereich

$$0 \leq p_{ij} \leq 1 \quad (4.15)$$

gilt. Da p_{ij} außerdem ein Maß für die Plausibilität der Verbindung über φ_i und φ_j darstellt, wird es im Folgenden als Verbindungswahrscheinlichkeit bezeichnet. Hierbei sei allerdings angemerkt, dass im Allgemeinen

$$\sum_{k \in U_i} p_{ik} \neq 1 \quad (4.16)$$

gilt. Die Verbindungswahrscheinlichkeit ist ihrer Natur nach symmetrisch, d. h. $p_{ij} = p_{ji}$. Außerdem soll sie die Gestaltprinzipien der Nähe und guten Fortsetzung berücksichtigen, welche im Gegensatz zum Tensor Voting bei der in Kap. 4.3 nicht direkter Bestandteil der Konturvervollständigung sind. Der Ansatz für p_{ij} lautet daher

$$p_{ij} = \xi_{ij} \cdot \frac{1}{1 + \lambda_{ij} \cdot \ell_{ij}} \quad (4.17)$$

Die Nähe zweier Kreuzungspunkte ist hierbei durch den Abstand ℓ_{ij} dieser Punkte gegeben. Die Symmetrie der Vervollständigungsrichtungen φ_i und φ_j kann über den linearen Zusammenhang

$$\lambda_{ij} = \Delta\gamma_{ij} \cdot \frac{\lambda_{end} - \lambda_{start}}{\Delta\gamma_{max}} + \lambda_{start} \quad (4.18)$$

berücksichtigt werden, wobei die Winkeldifferenz $\Delta\gamma_{ij}$ wieder entsprechend der Gleichung (4.12) gegeben ist. Der Einfluss des Gestaltprinzips der guten Fortsetzung lässt sich hierbei über die Parameter λ_{start} und λ_{end} und damit über den Wertebereich von λ_{ij} steuern.

Seien ψ_i und ψ_j die jeweils kleinsten Winkel zwischen der Verbindungsgerade der Kreuzungspunkte und den jeweiligen, durch die Vervollständigungsrichtungen φ_i und φ_j bestimmten Tangentenvektoren. Die Darstellung der Vervollständigungsrichtungen über die Winkel ψ_i und ψ_j hat gegenüber den Winkeln φ_i und φ_j den Vorteil, dass sie rotationsinvariant sind. Sei weiter der kleinere der beiden Winkel durch

$$\psi_{min_{ij}} = \begin{cases} \psi_i & \text{für } \psi_i \leq \psi_j \\ \psi_j & \text{für } \psi_i > \psi_j \end{cases} \quad (4.19)$$

gegeben. Ein weiterer Aspekt der guten Fortsetzung ist die in Kap. 4.2.1 bereits angesprochene nicht erwünschte Konstellation, in der die Verbindungsrichtungen ψ_i und ψ_j Werte von annähernd 90° annehmen. Für einen solchen Fall kann über den Term

$$\xi_{ij} = e^{-\psi_{min\,ij}^4} \quad (4.20)$$

eine Dämpfung der Verbindungswahrscheinlichkeit eingeführt werden.

Weiterhin ist es möglich, auch Polaritäten, also Helligkeitsinformationen und die Richtungen von Hell-Dunkelwechseln zu berücksichtigen, d. h. die in Abbildung 4.9(a) dargestellte Konstellation als wahrscheinlicher zu beurteilen als die in Abbildung 4.9(b) gegebene Konstellation.

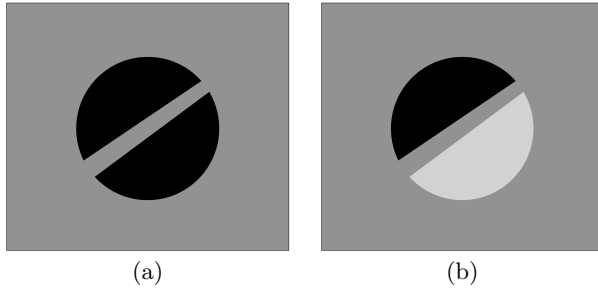


Abbildung 4.9: Unterschiedliche Polaritäten bei Bestimmung der Verbindungswahrscheinlichkeit.

Hierzu können die in Kap. 3.1 vorgestellten Eckfilter herangezogen werden, da die Filterantworten ein direktes Maß für die Polarität in einem Eckpunkt darstellen. Seien dazu die maximalen, auf den Bereich von -0.5 bis $+0.5$ normierten Eckfilterantworten an der Position der beiden Kreuzungspunkte durch f_i bzw. f_j gegeben. Einen Wahrscheinlichkeitsterm zur Berücksichtigung der Polarität erhält man durch $\delta_{ij} = \delta_1 \cdot \delta_2$ (vgl. [Hag07]), wobei

$$\delta_1 = e^{-(c_1 \cdot |f_i - f_j|)^4} \quad (4.21)$$

die Ähnlichkeit der beiden Filterantworten bewertet, während in

$$\delta_2 = 1 - e^{-(\text{sgn}(f_i) \cdot \text{sgn}(f_j) \cdot (|f_i| + |f_j|) - c_2)^4} \quad (4.22)$$

die Betonung auf gleiche Polarität der Filterantworten gelegt wird. Der Parameter c_1 in Gleichung (4.21) dient dabei der Einstellung der Flankensteilheit, während über den Parameter $c_2 < -1$ in Gleichung (4.22) der Graph der Kurve nach links verschoben werden kann, womit es möglich ist, den Toleranzbereich für unterschiedliche Polaritäten der Filterantworten zu verändern.

Gleichung (4.17) modifiziert sich in diesem Fall zu

$$p_{ij} = \xi_{ij} \cdot \delta_{ij} \cdot \frac{1}{1 + \lambda_{ij} \cdot \ell_{ij}} \quad (4.23)$$

Es ist bei dieser Vorgehensweise allerdings zu beachten, dass im Falle modaler Vervollständigungen Hintergrundflächen durchaus unzusammenhängend sein können und damit möglicherweise auch unterschiedliche Polaritätswerte liefern.

4.3 Konturvervollständigung durch Spline Funktionen

Ansatz für die Konturvervollständigung ist die Parameterdarstellung eines kubischen Splines K_S [Bar04, Loc92, Fel06, Häm91, Kno00]:

$$K_S: \quad \mathbf{S}(t) = \begin{pmatrix} S_x(t) \\ S_y(t) \end{pmatrix} = \begin{pmatrix} a \\ e \end{pmatrix} + t \begin{pmatrix} b \\ f \end{pmatrix} + t^2 \begin{pmatrix} c \\ g \end{pmatrix} + t^3 \begin{pmatrix} d \\ h \end{pmatrix} \quad (4.24)$$

mit $t \in [0, 1] \subset \mathbb{R}$. Zur Bestimmung der acht Unbekannten in Gleichung (4.24) werden acht linear unabhängige Gleichungen benötigt. Vier davon ergeben sich direkt aus dem Anfangspunkt $A(x_A, y_A)$ und dem Endpunkt $E(x_E, y_E)$ des Verbindungsweges:

$$\begin{aligned} S_x(0) &= x_A = a \\ S_y(0) &= y_A = e \\ S_x(1) &= x_E = a + b + c + d \\ S_y(1) &= y_E = e + f + g + h \end{aligned} \quad (4.25)$$

Die restlichen Bestimmungsgleichungen lassen sich durch die in Kap. 3 ermittelten Richtungen möglicher Konturfortsetzungen, welche zunächst als Winkel α_A für den Anfangspunkt bzw. α_E für den Endpunkt vorliegen, gewinnen. Hierzu werden die Tangentenvektoren

$$\mathbf{m}_A = \begin{pmatrix} m_{Ax} \\ m_{Ay} \end{pmatrix} = s \cdot \begin{pmatrix} \cos(\alpha_A) \\ \sin(\alpha_A) \end{pmatrix} \quad (4.26)$$

und

$$\mathbf{m}_E = \begin{pmatrix} m_{Ex} \\ m_{Ey} \end{pmatrix} = u \cdot \begin{pmatrix} -\cos(\alpha_E) \\ -\sin(\alpha_E) \end{pmatrix} \quad (4.27)$$

betrachtet, wobei die Längen $s = \|\mathbf{m}_A\|_2$ und $u = \|\mathbf{m}_E\|_2$ zunächst unbestimmt bleiben. Der Faktor -1 in der Definition des Vektors \mathbf{m}_E beruht auf der Tatsache, dass die ursprünglich mit den Kreuzungspunkten assoziierten Winkel in Richtung der jeweiligen Vervollständigung zeigen, was zu einer falschen Verbindung führen würde (vgl. Abbildung 4.10(a)). Stattdessen muss die Richtung des Tangentenvektors im Endpunkt eine entgegengesetzte Richtung aufweisen, um eine sinnvolle Tangente an die Splinefunktion darzustellen (vgl. Abbildung 4.10(b)).

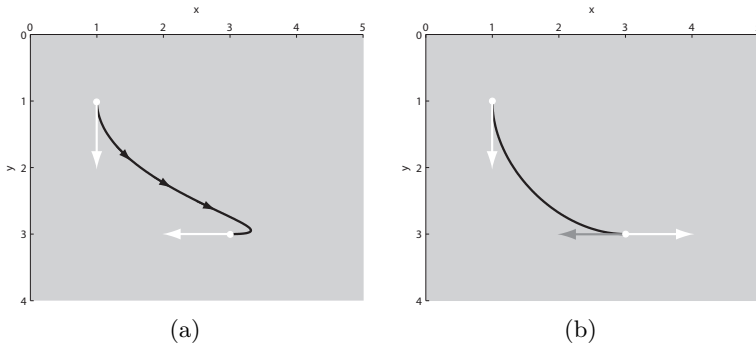


Abbildung 4.10: Generierter Verbindungsweg bei unterschiedlichen Richtungen des Tangentenvektors des Endpunktes. (a) Tangentenvektoren zeigen in Richtung der Vervollständigung und (b) der Tangentenvektor des Endpunktes wird um 180 Grad gedreht (nach [Fel06]).

Die in Gleichung (4.27) angegebenen Tangentenvektoren, welche sich aus der Richtung detektierter Konturen herleiten, stellen somit die Tangente an den kubischen Spline entsprechend Gleichung (4.24) dar. Betrachtet man nun also die Ableitungen

$$\begin{aligned} S'_x(t) &= b + 2ct + 3dt^2 \\ S'_y(t) &= f + 2gt + 3ht^2 \end{aligned} \quad (4.28)$$

so erhält man die restlichen Bestimmungsgleichungen:

$$\begin{aligned}
 S'_x(0) &= m_{Ax} = b \\
 S'_y(0) &= m_{Ay} = f \\
 S'_x(1) &= m_{Ex} = b + 2c + 3d \\
 S'_y(1) &= m_{Ey} = f + 2g + 3h
 \end{aligned} \tag{4.29}$$

Wird nun das Gleichungssystem, welches durch die Gleichungen (4.25) und (4.29) gegeben ist, nach den Koeffizienten der Splinedarstellung in Gleichung (4.24) aufgelöst, so ergibt sich:

$$\begin{aligned}
 a &= x_A \\
 b &= m_{Ax} \\
 c &= 3x_E - 3x_A - 2m_{Ax} - m_{Ex} \\
 d &= 2x_A - 2x_E + m_{Ax} + m_{Ex} \\
 e &= y_A \\
 f &= m_{Ay} \\
 g &= 3y_E - 3y_A - 2m_{Ay} - m_{Ey} \\
 h &= 2y_A - 2y_E + m_{Ay} + m_{Ey}
 \end{aligned} \tag{4.30}$$

Die Splinedarstellung des Verbindungsweges ist jetzt bis auf die Längen s und u der Tangentenvektoren eindeutig bestimmt. Setzt man nun $s = 1$ und $u = 1$, so lässt sich feststellen, dass der resultierende Verbindungsweg nicht skalierungsinvariant ist (Abbildung 4.11). Mit zunehmenden Abstand der Interpolationspunkte nähert sich die Splinedarstellung dem subjektiven Eindruck einer direkten Verbindung der beiden Punkte.

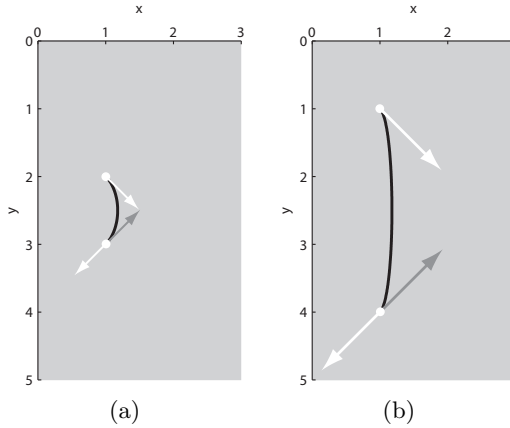


Abbildung 4.11: Skalierungsabhängigkeit des Verbindungsweges bei Wahl von Tangenteneinheitsvektoren. (nach [Fel06])

Abhilfe schafft hier die Betrachtung der Geradengleichungen

$$g: \quad \mathbf{g}(s) = \begin{pmatrix} x_A \\ y_A \end{pmatrix} + s \begin{pmatrix} \cos(\alpha_A) \\ \sin(\alpha_A) \end{pmatrix} \quad (4.31)$$

$$h: \quad \mathbf{h}(u) = \begin{pmatrix} x_E \\ y_E \end{pmatrix} + u \begin{pmatrix} \cos(\alpha_E) \\ \sin(\alpha_E) \end{pmatrix} \quad (4.32)$$

welche durch die Interpolationspunkte verlaufen und deren Richtungsvektoren den Tangentenvektoren entsprechen. Der Schnittpunkt der Geraden ist ein skalierungsunabhängiger Bezugspunkt [Fel06] und liefert Parameterwerte s und u , welche direkt in die Gleichungen (4.26) und (4.27) eingesetzt werden können. Um Einfluss auf das Krümmungsverhalten des Verbindungsweges zu haben, wird der skalierungsunabhängige Krümmungsparameter κ eingeführt.

Die Tangentenvektoren ergeben sich somit zu

$$\mathbf{m}_A = \begin{cases} \kappa \cdot s \cdot \begin{pmatrix} \cos(\alpha_A) \\ \sin(\alpha_A) \end{pmatrix} & \text{für } s > 0 \wedge u > 0 \\ \begin{pmatrix} \cos(\alpha_A) \\ \sin(\alpha_A) \end{pmatrix} & \text{sonst} \end{cases} \quad (4.33)$$

$$\mathbf{m}_E = \begin{cases} -\kappa \cdot u \cdot \begin{pmatrix} \cos(\alpha_E) \\ \sin(\alpha_E) \end{pmatrix} & \text{für } s > 0 \wedge u > 0 \\ \begin{pmatrix} -\cos(\alpha_A) \\ -\sin(\alpha_A) \end{pmatrix} & \text{sonst} \end{cases} \quad (4.34)$$

Mit der Fallunterscheidung in den Gleichungen (4.33) und (4.34) wird der Tatsache Rechnung getragen, dass bei der Bewertung der Verbindungswahrscheinlichkeiten in Kap. 4.2 eine gewisse Toleranz zugelassen wird gegenüber Verbindungsrichtungen, welche auf unterschiedlichen Seiten der Verbindungsgerade zwischen den Interpolationspunkten liegen. Dies kann zu Konstellationen mit parallelen Geraden und zu Konstellationen mit Werten $s < 0$ und $u < 0$ führen. Der Fall von sehr großen Werten für s und u braucht hier nicht weiter betrachtet zu werden, da dies einer Konstellation von annähernd parallelen Tangentenvektoren entspricht, welche einen rechten Winkel zur Verbindungsgerade zwischen den Interpolationspunkten aufweisen. Ein solcher Fall wird aber in Kap. 4.2 bereits ausgeschlossen.

Wie Abbildung 4.12 zeigt, kann über den eingeführten Krümmungsparameter κ das Krümmungsverhalten des Verbindungsweges beliebig modelliert werden. Im Kontext einer Detektion unterbrochener Objektgrenzen machen jedoch extreme Werte wie $\kappa = 0$, was einer direkten Verbindung der Kreuzungspunkte entspricht (vgl. Abbildung 4.12(a)), und hohe Werte von

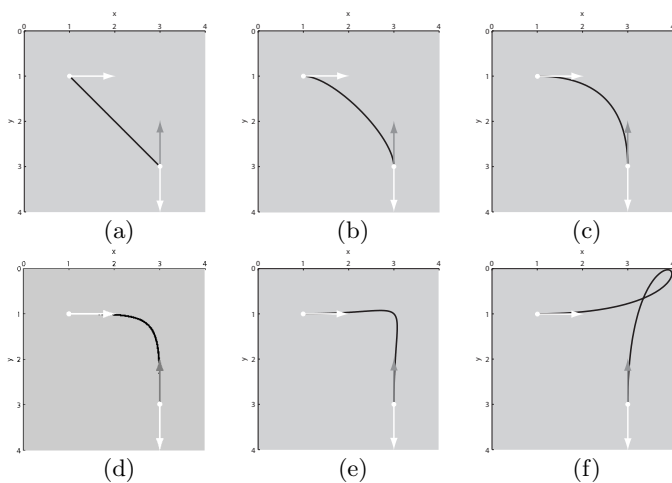


Abbildung 4.12: Verlauf des Verbindungsweges mit unterschiedlichen Krümmungsparametern (nach [Fel06]): (a) $\kappa = 0$ (b) $\kappa = 1$ (c) $\kappa = 2$ (d) $\kappa = 3$ (e) $\kappa = 4$ (f) $\kappa = 7,8$

κ keinen Sinn (vgl. Abbildung 4.12(f)). Im Folgenden wird daher, falls nicht anders angegeben, ein Krümmungswert von $\kappa = 2$ angenommen.

5 Perzeptuelle Organisation

Notwendiger Bestandteil perzeptueller Organisation ist neben einer Konturvervollständigung, wie sie in Kap. 4 vorgestellt wurde, eine Figur-Grund-Gliederung, d. h. eine Organisation des Bildes in Bereiche, welche sich als Figur vom restlichen Bereich, dem Hintergrund, abheben. Während in [WH96] und [Sau99b] versucht wird, die Gruppierung bzw. Vervollständigung von Konturen simultan zur Figur-Grund-Gliederung in einem kombinatorischen Ansatz durchzuführen, wird im vorliegenden Ansatz zunächst die Konturvervollständigung durchgeführt (siehe Kap. 4), um anschließend die gewonnenen Informationen als Kriterium zur Figur-Grund-Gliederung an Kreuzungspunkten nutzen zu können. Dieses Vorgehen korrespondiert mit der von Kellman et al. postulierten Identitätshypothese [KGW01], welche besagt, dass die Mechanismen für modale und amodale Konturvervollständigungen identisch sind und die Tiefenzuordnung vom Konturvervollständigungsprozess getrennt werden kann. Wichtig ist in diesem Zusammenhang eine Tiefenzuordnung, welche tolerant auf eine möglicherweise falsche initiale Figur-Grund-Gliederung an Kreuzungspunkten reagiert. Ein solcher Mechanismus wird in Kap. 5.2 vorgestellt.

5.1 Figur-Grund-Gliederung an Kreuzungspunkten

Da der vorliegende Ansatz von detektierten Konturen, also potenziellen Objektbegrenzungen ausgeht, findet auch die Zuordnung einer relativen Tiefe entlang der Kontur und damit entlang der Objektgrenze statt. Schlüsselpunkt für die Zuweisung von Vorder- oder Hintergrundeigenschaft ist der Kreuzungspunkt zweier Konturen. Falls hier, vom Kreuzungspunkt ausgehend, mindestens ein Teil der beteiligten Konturen eine vervollständigte

Kontur darstellt, kann davon ausgegangen werden, dass die gegebene Konstellation durch eine Verdeckung hervorgerufen wurde. Die unten vorgestellten Kriterien stellen Bedingungen dar, welche sich direkt aus dem physikalischen Zusammenhang bei der Verdeckung durch ein Vordergrundobjekt ableiten lassen. Es ist aber genauso möglich, weitere Kriterien zur Figur-Grund-Gliederung aus der Analyse der menschlichen Wahrnehmung zu gewinnen. So wird z. B. in [Wil94] eine Präferenz von schwarzen Objekten gegen einen weißen Hintergrund postuliert. In [MM01a] wird im Zusammenhang mit sich verdeckenden Objekten gleicher Farbe in binären Bildern angenommen, dass breite Objekte eher zum Vordergrund gehören, als dünne Objekte (vgl. Abb. 5.1(c)), d. h. diejenigen Objekte, welche die längeren Konturabschnitte vervollständigter Konturen enthalten, werden als Hintergrund angesehen. Angesichts der Tatsache, dass ähnliche Effekte durchaus in der menschlichen Wahrnehmung festzustellen sind (siehe Kap. 2.4.4), macht ein solches Vorgehen durchaus Sinn, allerdings folgt ein solches Kriterium zur Figur-Grund-Unterscheidung nicht direkt aus der gegenseitigen Verdeckung zweier Objekte und stellt somit auch kein notwendiges Kriterium dar. Weiterhin ist es ohne weiteres möglich, eine T-Kreuzung als Kriterium für eine Figur-Grund-Unterscheidung heranzuziehen, indem die durchgehende Kontur als Objektbegrenzung einer Vordergrundfläche deklariert wird, während die unterbrochene Kontur dem Hintergrund zugeordnet wird.

Die in Kap. 6 vorgestellten Ergebnisse resultieren allein aus den folgenden Figur-Grund-Kriterien an die Konstellation von realen und vervollständigten Konturen am Kreuzungspunkt zweier Konturen.

Falls am Kreuzungspunkt zweier Konturen eine Kontur ihren Zustand von real zu vervollständigt ändert und falls dies bei der kreuzenden Kontur nicht der Fall ist, gehört die erste Kontur zum Hintergrund, während die zweite Kontur die Begrenzung einer Vordergrundfläche darstellt.

Ein Beispiel für eine solche Konstellation ist in Abb. 5.1(a) gegeben.

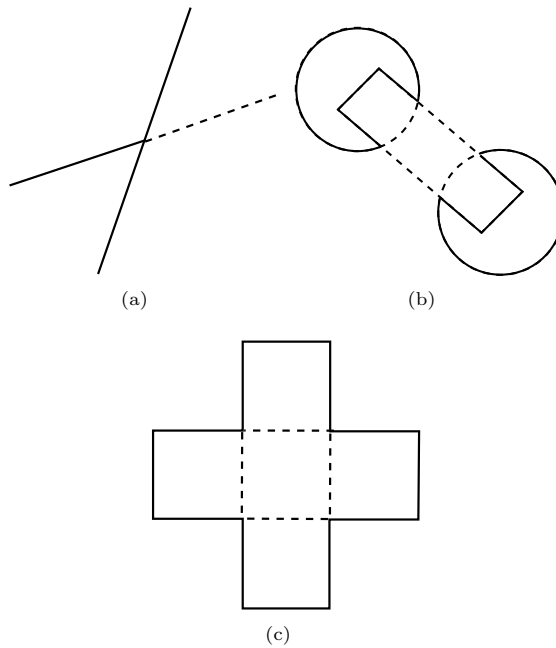


Abbildung 5.1: Illustrationen zu Figur-Grund-Kriterien an Kreuzungspunkten. Reale Konturen sind als durchgehende Linie dargestellt, während vervollständigte Konturen gestrichelt dargestellt sind. Erläuterung siehe Text.

Falls eine Kontur ihren Zustand von real zu vervollständigt ändert und falls der vervollständigte Teil in der Region liegt, welche von einer kreuzenden geschlossenen Kontur begrenzt wird, gehört die erste Kontur zum Hintergrund, während die zweite Kontur die Objektbegrenzung eines Vordergrundobjektes darstellt.

Eine solche Konstellation ist in Abb. 5.1(b) dargestellt. Die zu den beiden Kreisen gehörenden vervollständigten Konturen liegen im Inneren der

rechteckigen Kontur. Die Kreiskonturelemente im Bereich der Kreuzungspunkte werden damit als Hintergrund und die der rechteckigen Kontur als Vordergrund gekennzeichnet.

Wenn beide Konturen geschlossen sind und beide Konturen vervollständigte Anteile aufweisen, die in der von der jeweils anderen Kontur begrenzten Region liegen, wird keine Figur-Grund-Entscheidung getroffen.

Die beschriebene Situation ist in Abb. 5.1(c) dargestellt. Im Gegensatz zum Ansatz in [MM01a] wird hier also nicht die Länge der vervollständigten Konturen zur Entscheidung herangezogen, sondern keine Entscheidung getroffen. Damit können Fehlzuordnungen, wie Sie bei Objekten unterschiedlicher Größe nicht unwahrscheinlich sind (vgl. Abbildung 2.14), vermieden werden.

5.2 Dynamische Tiefenzuordnung

Ziel der Tiefenzuordnung ist es, die in Kap. 5.1 gewonnenen Tiefeninformationen an den Kreuzungspunkten zweier Konturen an den Rest der jeweiligen Kontur weiterzuleiten und eine sinnvolle dreidimensionale Interpretation des Konturverlaufs zu ermöglichen. Um zu gewährleisten, dass die Tiefenzuordnung einen kontinuierlichen dreidimensionalen Konturverlauf sogar für solche Konturen erzeugt, welche Flächen begrenzen, die vom Vorder- zum Hintergrund wechseln und umgekehrt, wird die Tiefenzuordnung in Form eines Diffusionsprozesses realisiert [HM06]. Im Gegensatz zur stereoskopischen Tiefenrekonstruktion ist es hier nicht möglich, eine quantitative Aussage zu treffen über die Entfernung eines Bildmerkmals zum Kamerasystem, die zugeordnete relative Tiefe gibt dagegen Aufschluss über die zum Kamerasystem relative Lage eines Objektes im Raum.

Die Kontur ist im Folgenden aufgeschlüsselt als Kettencode mit den zweidimensionalen Bildkoordinaten $(x_n, y_n)^T$ des n -ten Konturelements, $n \in [1, \dots, T]$, wobei $T \in \mathbb{N}$ die Länge der Kontur bzw. die Anzahl der

in der Kontur enthaltenen Pixel sei. Sei z_n die relative Tiefe des n -ten Konturelements mit den Nachbarn z_{n-2} , z_{n-1} , z_{n+1} und z_{n+2} . Tiefe eines Konturelements bedeutet in diesem Sinne die Tiefe der okkludierenden Oberfläche, falls die Kontur die Begrenzung einer Fläche darstellt. Ein höherer Tiefenwert bedeutet hier eine größere Entfernung zum Kamerasystem

Für die Glättungsparameter c_1 und c_2 modelliert der Kostenterm P_1 mit

$$\begin{aligned}
 P_1 = & \frac{1}{2} \sum_n \left(\frac{c_1}{2} \cdot \sum_{i=n-k}^{n+k} (z_n - z_{n+i})^2 \right. \\
 & + c_2 \cdot ((z_{n+1} - z_n) - (z_n - z_{n-1}))^2 \\
 & + \left. \begin{cases} (z_n - z_{n_0})^2 & z_n \text{ gehört zu einem Kreuzungspunkt} \\ 0 & \text{sonst} \end{cases} \right) \quad (5.1)
 \end{aligned}$$

die Forderung, dass benachbarte Pixel möglichst gleiche Tiefenwerte aufweisen sollen und die (diskretisierte) erste Ableitung benachbarter Pixel ebenfalls identisch sein soll. Die Größe der Nachbarschaft wird dabei durch den Parameter k beschrieben. Im Gegensatz hierzu werden im letzten Term Kosten erzeugt, falls der Tiefenwert z_n eines Kreuzungspunktes von dem vorgegebenen Wert z_{n_0} abweicht. Der Wert z_{n_0} wird im Verlauf des im Folgenden beschriebenen Iterationsprozesses aktualisiert, und zwar so, dass im Falle einer Konstellation, die der Figur-Grund-Gliederung entsprechend Kap. 5.1 widerspricht, der Wert z_{n_0} modifiziert wird. Wenn also die betrachtete Kontur im betrachteten Kreuzungspunkt als Vordergrund eingestuft wurde, die Summe des Tiefenwertes z_n und eines Offsets d aber größer ist als der Tiefenwert der anderen Kontur, so wird z_{n_0} aktualisiert zu $z_{n_0} = z_n - \Delta z$, wobei Δz die Korrekturschrittweite darstellt. Der Offset $d > 0$ bezeichnet hierbei den geforderten Mindestabstand zwischen zwei Konturen. Analog wird im Fall einer Hintergrundkontur verfahren, welche zu geringe Tiefenwerte aufweist. Tritt keiner dieser Fälle auf, wird nichts verändert, d. h. es wird $z_{n_0} = z_n$ gesetzt.

Durch Anwendung des Gradientenverfahrens zur Minimierung der in Gleichung (5.1) aufgestellten Kostenfunktion erhält man für die partielle Ableitung

$$\begin{aligned} \frac{\partial}{\partial z_n} P_1 &= c_1 \cdot \sum_{i=n-k}^{n+k} (z_n - z_{n+i}) \\ &+ c_2 \cdot (6z_n - 4z_{n+1} - 4z_{n-1} + z_{n+2} + z_{n-2}) \\ &+ \begin{cases} z_n - z_{n_0} & z_n \text{ gehört zu einem Kreuzungspunkt} \\ 0 & \text{sonst} \end{cases} \end{aligned} \quad (5.2)$$

Mit einer Schrittweite λ ergibt sich für den Wert z_n die folgende Iterationsvorschrift:

$$z_n^{(j+1)} = z_n^{(j)} - \lambda \cdot \frac{\partial}{\partial z_n} P_1 \quad (5.3)$$

Hierbei ist zu beachten, dass für die oben angegebenen Gleichungen außerdem eine Randbetrachtung berücksichtigt werden muss. Im Fall von nicht geschlossenen Konturen entfallen in Gleichung (5.1) Tiefenwerte, die außerhalb des Intervalls $[1, \dots, T]$ liegen und analog ändert sich Gleichung (5.2) an den entsprechenden Stellen. Im Fall geschlossener Konturen ist für die Nachbarn eines betrachteten Pixel zu beachten, dass die Nummerierung T-periodisch verläuft, d. h. es gilt z. B. $z_{T+1} = z_1$.

Wird nun für den Fall einer geschlossenen Kontur angenommen, dass die Kontur die Objektgrenze eines verhältnismäßig simplen Objektes darstellt und dass die Kontur sich außerdem auf einer Fläche befindet, die wie auch immer im Raum angeordnet sein kann, so lassen sich aus der Betrachtung der Tiefenwerte eines Kreises wertvolle Bedingungen für die Tiefenkriterien ableiten. Dies funktioniert auch für Objekte, die keine kreisförmige Begrenzung aufweisen, wie in Kap. 6 am Beispiel eines Bleistiftes gezeigt wird. Eine Bedingung, die sich aus einer solchen Einschränkung ableiten lässt, ist, dass

die Tiefenwerte sich gegenüberliegender Punkte dieselbe Höhendifferenz zur Tiefe M_z des Schwerpunktes $\mathbf{M} = (M_x, M_y, M_z)^T$ der Kontur aufweisen. Der Schwerpunkt \mathbf{M} ist hierbei gegeben durch

$$\mathbf{M} = \frac{1}{T} \sum_{n=1}^T (x_n, y_n, z_n)^T. \quad (5.4)$$

Es ergibt sich somit ein Kostenterm

$$P_2 = \frac{c_3}{4} \cdot \sum_n (z_n + z_{n+T/2} - 2M_z)^2. \quad (5.5)$$

Wird in jedem Iterationsschritt j der Schwerpunkt $\mathbf{M}^{(j)}$ entsprechend Gleichung (5.4) ermittelt, so folgt für die partiellen Ableitung

$$\frac{\partial}{\partial z_n} P_2 = c_3 \cdot (z_n + z_{n+T/2} - 2M_z^{(j)}). \quad (5.6)$$

Der Wert $T/2$ kann hierbei für ungerade T beliebig gerundet werden, da davon auszugehen ist, dass die Kontur bzw. der Wert von T hinreichend groß ist und die Position des gegenüberliegenden Pixels nur mit einer gewissen Toleranz bestimmt werden muss. Die in Gleichung (5.3) beschriebene Iterationsvorschrift ändert sich somit zu

$$z_n^{(j+1)} = z_n^{(j)} - \lambda \cdot \frac{\partial}{\partial z_n} (P_1 + P_2). \quad (5.7)$$

Abgesehen von der direkten Einbindung in den Kostenterm ist es auch möglich, nach erfolgter Kostenminimierung in einem Nachverarbeitungsschritt den dreidimensionalen Kurvenverlauf durch einen Kreis zu approximieren, die entsprechenden Kreisparameter zu bestimmen und die Tiefenwerte z_1, \dots, z_T eines solchen Kreises wieder mit den zweidimensionalen

Bildkoordinaten der Kontur zu verbinden. Hierzu wird zunächst die Vereinfachung angenommen, dass der Abstand zwischen zwei Konturelementen äquidistant ist. Anschließend muss der Kreismittelpunkt \mathbf{M} entsprechend der Gleichung (5.4) bestimmt werden. Der Radius R lässt sich durch

$$R = \frac{1}{T} \sum_{n=1}^T \|(x_n - M_x, y_n - M_y, z_n - M_z)^T\|_2 \quad (5.8)$$

approximieren. Betrachtet man ein Konturelement n und ein Konturelement $n + \frac{T}{4}$, wobei wiederum die T-Periodizität beachtet werden muss, so lässt sich zu den beiden Vektoren $(x_n - M_x, y_n - M_y, z_n - M_z)^T$ und $(x_{n+T/4} - M_x, y_{n+T/4} - M_y, z_{n+T/4} - M_z)^T$ ein Normvektor $\mathbf{n}^{(n)} = (n_x^{(n)}, n_y^{(n)}, n_z^{(n)})^T$ bestimmen, für den o. B. d. A. die Bedingung $n_z^{(n)} \geq 0$ gefordert sei. Der Neigungswinkel α (siehe Abbildung 5.2) lässt sich nun annähernd über

$$\alpha = \max_{n \in [1, \dots, T]} \left\{ \arccos \left(\frac{n_z^{(n)}}{\|\mathbf{n}^{(n)}\|_2} \right) \right\} \quad (5.9)$$

bestimmen.

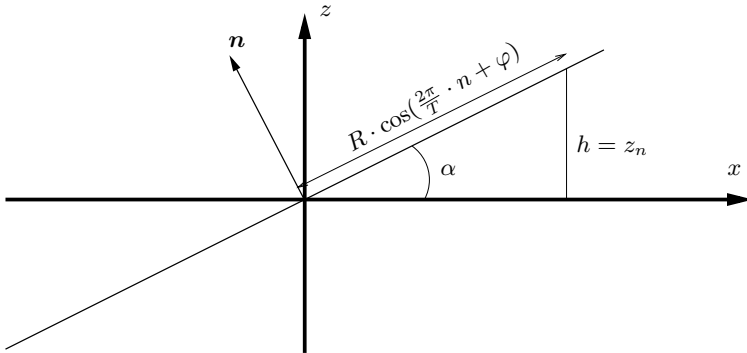


Abbildung 5.2: Seitenansicht eines Kreises mit Neigungswinkel α .

Um eine Approximation für den Phasenwinkel φ zu erhalten, der sowohl von der Rotation des um den Winkel α geneigten Kreises um die z -Achse, als auch vom Winkel des ersten Konturelements zur x -Achse abhängt, wird die gegebene Konstellation auf die x - y -Ebene projiziert (siehe Abbildung 5.3). Der Winkel φ_1 lässt sich nun über die Koordinaten (x_n, y_n) bestimmen [Mer99], während der Winkel φ_2 durch die Koordinaten der zweidimensionalen Projektion des Vektors $-\mathbf{n}$ bestimmt werden kann.

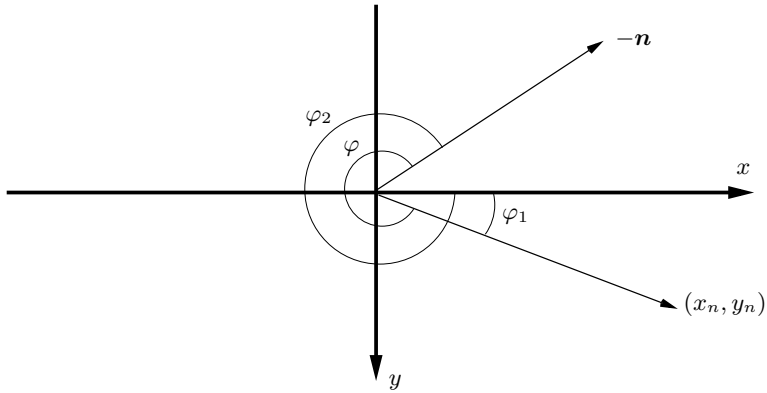


Abbildung 5.3: Bestimmung des Phasenwinkels φ .

Wie in Abbildung 5.2 zu erkennen ist, lässt sich nun mit gegebenem Phasenwinkel φ die Höhe eines Kreiskonturelements durch

$$z_n = M_z + \sin(\alpha) \cdot R \cdot \cos\left(\frac{2\pi}{T} \cdot n + \varphi\right) \quad (5.10)$$

bestimmen.

6 Experimentelle Ergebnisse

Während in den vorangegangenen Kapiteln bereits exemplarisch einige Ergebnisse gezeigt wurden, um die prinzipielle Funktionsweise einzelner Verfahrensschritte zu verdeutlichen, wird in diesem Kapitel die Leistungsfähigkeit des Verfahrens für jeden Verfahrensschritt anhand ausgesuchter Beispiele demonstriert.

6.1 Anisotrope Regularisierung

Die Abbildungen 6.1(a) bis (d) werden in [Wei96] als Beispiel für die Rauschunterdrückungseigenschaften verschiedener Diffusionsfilter vorgestellt und haben mittlerweile ihren Weg in die Standardliteratur digitaler Bildverarbeitung gefunden [Jäh02]. 70% der Bildpunkte des Eingangsbildes 6.1(a), das ein Dreieck und ein Rechteck enthält, sind durch Rauschen verfremdet. Abbildung 6.1(b) zeigt das Ergebnis einer linearen isotropen Diffusion, was einer einfachen Glättung des Eingangsbildes entspricht. Hierbei wird das Rauschen unterdrückt, allerdings werden auch die Kanten unscharf. In Abbildung 6.1(c) ist dagegen das Ergebnis nichtlinearer isotroper Diffusion dargestellt. Durch die Diffusion, deren Stärke von der des Gradienten abhängt, wird das Rauschen unterdrückt, ohne die Schärfe der Kanten zu beeinträchtigen. Allerdings findet im Bereich der Kanten selbst keine Diffusion statt, so dass diese verrauscht bleiben. Schließlich zeigt Abbildung 6.1(d) die von Weickert beschriebene nichtlineare anisotrope Diffusion (vgl. Kap. 3.3). Auch im Bereich von Kanten findet hier eine Diffusion statt, so dass auch für Kantenbereiche sowohl das Rauschen unterdrückt wird, als auch die Kanten scharf bleiben. Allerdings lässt sich eine Lokalisierungsungenauigkeit der Kanten aufgrund abgerundeter Ecken feststellen.

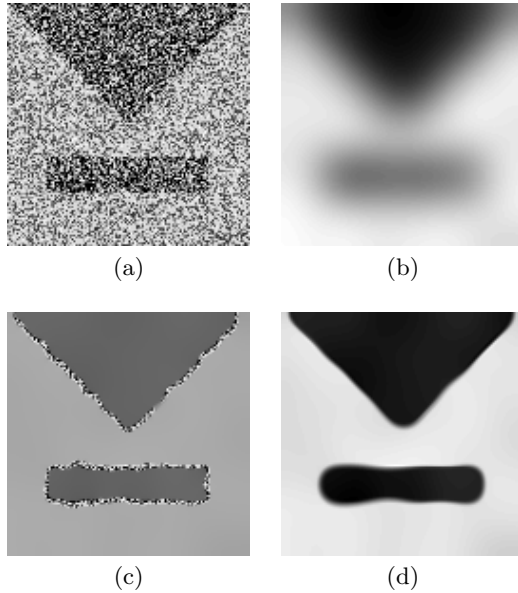


Abbildung 6.1: Ergebnisse unterschiedlicher Diffusionsprozesse (aus [Wei96] bzw. [Jäh02]): (a) Eingangsbild, (b) lineare Diffusion, (c) nichtlineare isotrope Diffusion und (d) nichtlineare anisotrope Diffusion.

In Abbildung 6.2(f) zeigt sich, verglichen mit den Filterergebnissen, deren betragsgrößte Filterbeträge in Abbildung 6.2(e) dargestellt sind, die Verbesserung der Kanteninformation nach anisotroper Regularisierung mit Selbstorganisation der Gradientenkarte entsprechend Kap. 3.3.3. Die salienten Kanten des Dreiecks und des Rechtecks verstärken sich und richten sich durch den Selbstorganisationsprozess aus, während die durch Rauschprozesse erzeugten Kanten unterdrückt werden. Abbildung 6.2(g) zeigt die Lösung des in Kap. 3.3.1 vorgestellten Regularisierungsterms mit anisotropen Gewichtungsfaktoren, die sich aus der selbstorganisierten Gradientenkarte ergeben. Das Gleiche ist in Abbildung 6.2(h) dargestellt, allerdings sind

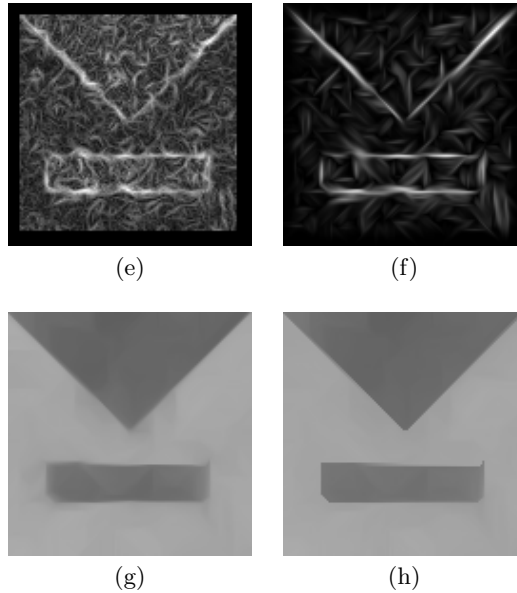


Abbildung 6.2: Diffusion durch anisotrope Regularisierung: (e) betragsgrößte Filterbeträge, (f) Betrag der Gradientenkarte nach erfolgtem Selbstorganisationsprozess, (g) Ergebnis der anisotropen Regularisierung mit Selbstorganisation der Gradientenkarte entsprechend Kap. 3.3.2 und (h) mit zusätzlichem Schwellwert.

hier die Gewichtungsfaktoren mit einem Schwellwert verbunden, der, wie in Kap. 3.3.3 beschrieben, eine Diffusion zwischen Bildpunkten, die sich in ihrem Grauwert zu stark unterscheiden, unterbindet. Um eine dem starken Rauschen entsprechende hohe Diffusion zu erzielen, wurde der Skalierungs- bzw. Glättungsparameter in beiden Fällen zu $c = 1000$ gesetzt. Der Vergleich mit Abbildung 6.1 zeigt, dass hier eine signifikante Verbesserung, speziell bei der Lokalisierungsgenauigkeit der Kantenerhaltung stattgefunden hat. Für den gesamten Bildbereich werden sowohl das Rauschen unter-

drückt, als auch die salienten Kanten scharf und ohne wesentliche Lokalisierungsungenauigkeiten abgebildet.

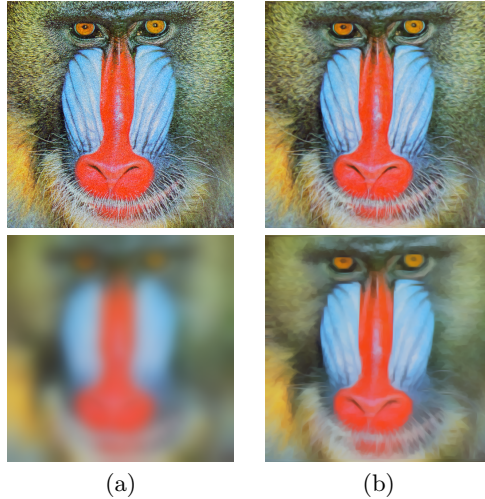


Abbildung 6.3: Mandrill-Testbild: (a) N - Eingangsbild mit 10% unkorreliertem Rauschen, (b) $R_{3,0_n}$ - anisotrope Diffusion mit Glättungsparameter $c = 3$, (c) R_{250_l} - lineare isotrope Diffusion mit Glättungsparameter $c = 250$ und (d) R_{250_n} - anisotrope Diffusion mit Glättungsparameter $c = 250$.

Der naheliegendste Anwendungsbereich kantenerhaltender Regularisierung ist die Rauschunterdrückung in digitalen Bildern. Im Folgenden findet daher eine Evaluation der anisotropen Glättung auf verrauschten Bildern statt. Die in Tabelle 6.1 dargestellten Resultate beziehen sich auf das in Abbildung 6.3 dargestellte Mandrill-Testbild. In Abbildung 6.3(a) wurde das ursprüngliche Bild in jedem Bildpunkt mit 10% Rauschen überlagert. Um zwei Bilder miteinander zu vergleichen, wird zunächst der prozentuale Anteil an Farb- bzw. Grauwerten betrachtet, die, bezogen auf die absolute Anzahl $m \cdot d$ an Farbwerten eines Bildes mit m Bildpunkten, für den betrachteten Farbwert

Ähnlichkeit	$s = 1$	$s = 6$	$s = 16$	$s = 21$
$N \wedge O$	6.19	36.38	79.41	90.36
$R_{0.1_l} \wedge O$	3.90	40.16	83.14	91.92
$R_{0.1_n} \wedge O$	4.04	40.88	85.60	94.40
$R_{0.2_n} \wedge O$	4.01	41.37	85.62	93.97
$R_{0.3_n} \wedge O$	3.99	41.13	84.40	92.75
$R_{0.5_n} \wedge O$	3.90	40.24	81.75	90.29
$R_{250_n} \wedge R_{250_o}$	29.18	98.83	99.97	99.99

Tabelle 6.1: Mandrill-Testbild: Ähnlichkeit zwischen verrauschtem Bild (N), Originalbild (O) und Ergebnis anisotroper (R_{c_n}) und isotroper (R_{c_l}) Diffusion.

mit gegebener Bildposition einen Differenzbetrag aufweisen, der unter einem Schwellwert s liegt. Ein hoher Wert bedeutet damit, dass die Bilder, bezogen auf den Schwellwert, eine hohe Ähnlichkeit aufweisen. In Tabelle 6.1 wird das Originalbild mit O bezeichnet, das verrauschte Eingangsbild mit N und zu gegebenen Glättungsparameter c bezeichnet R_{c_n} das Ergebnis der auf das verrauschte Eingangsbild angewandten anisotropen Diffusion, während für R_{c_o} das Originalbild als Eingangsbild dient.

Tabelle 6.1 stellt die oben beschriebene Ähnlichkeit für unterschiedliche Bilder und unterschiedliche Schwellwerte dar. Ein Vergleich zwischen dem Originalbild (O) und dem verrauschten Bild (N) zeigt, dass durch das additive Rauschen weniger als sieben Prozent der Bildpunkte pro Farbkanal ihren Farbwert beibehalten haben. Wie in Tabelle 6.1 zu sehen ist, führt im Vergleich zur Ähnlichkeit zwischen Original und verrauschtem Bild ($N \wedge O$) das beschriebene Regularisierungsverfahren ab einem betrachteten Schwellwert $s = 6$ zu einer höheren Ähnlichkeit zum Originalbild ($R_{c_n} \wedge O$), wobei diese Ähnlichkeit mit zunehmenden Glättungsparameter abnimmt. Die Erklärung hierfür liegt in den hochfrequenten Bildanteilen des Originalbilds, speziell im Bereich der Haare des Mandrills, die Teil des Signals und nicht des Rauschens sind, aber trotzdem mit zunehmenden Glättungsparameter

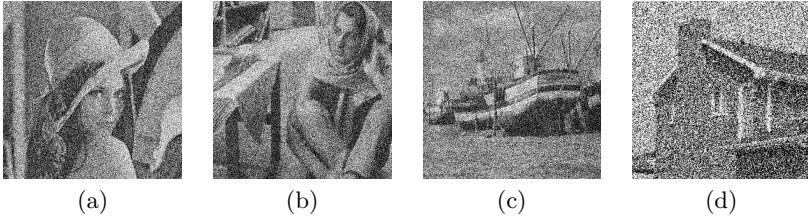


Abbildung 6.4: Testbilder für die Evaluierung in Tabelle 6.2: (a) Lena, (b) Barbara, (c) Boats und (d) House. Alle Bilder sind mit einem additiven gaußförmigen weißen Rauschen ($\sigma = 100$) überlagert (aus [Por08]).

herausgefiltert werden. Dies erklärt auch die besonders hohe Ähnlichkeit zwischen der Regularisierung des Originalbildes und der des verrauschten Eingangsbildes ($R_{250_n} \wedge R_{250_o}$) für einen hohen Glättungsparameter $c = 250$. Da die Eingangsbilder sich hier im Wesentlichen durch ihre hochfrequenten Bildanteile unterscheiden, führt die anisotrope Regularisierung für beide Bilder zu annähernd gleichen Ergebnissen. Wie zu erwarten, weist das Ergebnis anisotroper Diffusion eine höhere Ähnlichkeit zum Original auf, als das Ergebnis isotroper Diffusion des verrauschten Bildes ($R_{c_l} \wedge O$).

Das in [PSWS03, GCP05] von Portilla et al. entwickelte Verfahren zur Rauschunterdrückung in digitalen Bildern basiert auf einem statistischen Modell für die Koeffizienten der Multiskalenrepräsentation eines Bildes durch orientierungsselektive Filter im Rahmen einer Wavelet Transformation und wird in [PSWS03] mit gängigen Verfahren zur Rauschunterdrückung verglichen. Resultat der umfangreiche Evaluierung ist, dass das Verfahren von Portilla et al. gängige Verfahren bezüglich des PSNR-Wertes (Peak Signal to Noise Ratio) übertrifft.

$\sigma = 100$	Lena	Barbara	Boats	House
Portilla et al.	25.64	22.61	23.75	25.11
anisotrope Diffusion	25.65	22.07	23.13	24.92

Tabelle 6.2: Vergleich der Rauschunterdrückung durch anisotrope Diffusion mit den Ergebnissen von Portilla et al., die Ergebnisse sind als PSNR (Peak Signal to Noise Ratio) dargestellt. Die Eingangsbilder sind in Abbildung 6.4 dargestellt. Alle Ergebnisse der anisotropen Diffusion wurden mit identischen Parametern berechnet.

Um eine Einordnung der Rauschunterdrückungseigenschaften der hier vorgestellten anisotropen Regularisierung zu ermöglichen, wird diese im Folgenden mit dem Verfahren von Portilla et al. verglichen [HM08]. Hierbei werden mit den in Abbildung 6.4 gezeigten Bildern dieselben Testbilder verwendet wie in [PSWS03]. Zum Vergleich wird auch hier der PSNR-Wert herangezogen, der durch

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{255^2}{\sigma_e^2} \right) = 20 \cdot \log_{10} \left(\frac{255}{\sigma_e} \right) \quad (6.1)$$

definiert ist, wobei σ_e^2 den mittleren quadratischen Abstand zwischen Original und rekonstruiertem Bild darstellt. Der Wert 255 stellt den maximalen Farbwert bei gängiger 8-Bit Darstellung dar.

Bezogen auf den mittleren quadratischen Abstand zeigt die anisotrope Regularisierung einen wesentlichen Nachteil, sie zieht ein Zusammenschrumpfen des Histogramms nach sich, was sich unabhängig von der subjektiven Bildqualität unmittelbar auf den PSNR auswirkt. Um dennoch eine Vergleichbarkeit der Ergebnisse zu ermöglichen, basieren die in Tabelle 6.2 präsentierten Ergebnisse auf einem linearen Histogrammabgleich der anisotropen Regularisierungsergebnisse.

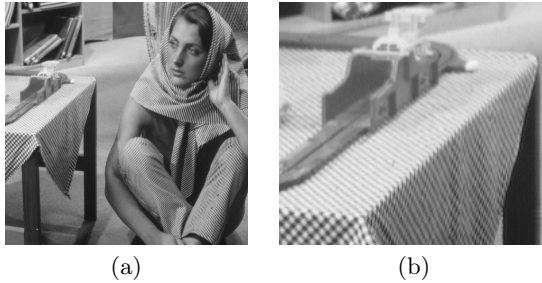


Abbildung 6.5: Barbara-Testbild: (a) Originalbild und (b) Vergrößerung der Tischregion.

Wie in Abbildung 6.6(c) zu sehen ist, verbleibt dennoch eine systematische Abweichung des Grauwertlevels der Bildregionen im Vergleich zum Originalbild 6.5(b). Für sämtliche Ergebnisse wurde ein Glättungsparameter von $c = 4.0$ mit 60 Iterationsschritten genutzt. Der Startvektor ist durch das verrauschte Eingangsbild gegeben und die Gewichtungsfaktoren der Umfeldkopplung resultieren aus einer Selbstorganisation der Gradientenkarte ent-

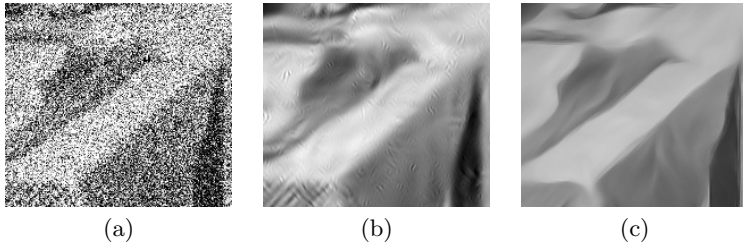


Abbildung 6.6: Barbara-Testbild: (a) Vergrößerung der Tischregion im verrauschten Eingangsbild mit $\sigma = 100$ und $\text{PSNR} = 8.13$, (b) Ergebnis der Rauschunterdrückung nach Portilla et. al. ($\text{PSNR} = 22.6$) und (c) Ergebnis der Rauschunterdrückung durch anisotrope Diffusion ($\text{PSNR} = 22.1$).

sprechend Kap. 3.3.3 mit einem Glättungsparameter von $c = 8.0$. Tabelle 6.2 zeigt, dass die PSNR-Werte des Verfahrens nach Portilla et al. [PSWS03] bis auf das Lena-Testbild höhere Werte aufweisen und damit die quantitativ bessere Rauschunterdrückung gewährleisten.

Betrachtet man allerdings die Ergebnisbilder (vgl. Abbildung 6.6(b)), so lässt sich feststellen, dass das Verfahren nach Portilla et al. für stark verrauschte Bilder Artefakte produziert, die im Ergebnis durch anisotrope Regularisierung nicht auftreten (Abbildung 6.6(c)). Gerade für die Zielsetzung der vorliegenden Arbeit, im Rahmen einer Bildvorverarbeitung eine verlässliche Detektion von Konturen und Kreuzungspunkten zu gewährleisten, ist dies jedoch ein entscheidender Punkt.

6.2 Konturverstärkung

Wie bereits in den Kapiteln 2.5 und 4.1.3 beschrieben wurde, lässt sich unter Verwendung eines bipolaren Verbindungsschemas das in Kap. 4.1 vorgestellte Tensorvotingverfahren zur Verstärkung salienter Konturen nutzen. Ein solcher Ansatz wird in [MBM02b, MBM02a, MBM03] verfolgt, wobei die Eingangstensoren aus einer Überlagerung unterschiedlich orientierter Tensoren mit den Filterantworten entsprechend orientierter Gaborfilter [Gab46, Dau88, Tra98, Tra96], resultieren.

Zum Testbild in Abbildung 3.15(a) aus Kap. 3.3.3 sind die Sticksalienen der Tensoren nach erfolgtem Stickvoting in Abbildung 6.7 dargestellt, wobei Abbildung 6.7(a) auf einem bipolaren Stickvotingfeld beruht, während der Votingprozess in Abbildung 6.7(b) zusätzlich inhibitorische Anteile aufweist. Der Vergleich mit dem Ergebnis der Selbstorganisation der Gradientenkarte in Abbildung 6.7(c) zeigt, dass die in Kap. 3.3.3 beschriebene Vorgehensweise sowohl in der Verstärkung salienter Konturen, der Lokalisierungsgenauigkeit, als auch der Unterdrückung von Rauschprozessen eindeutige Vorteile aufweist.

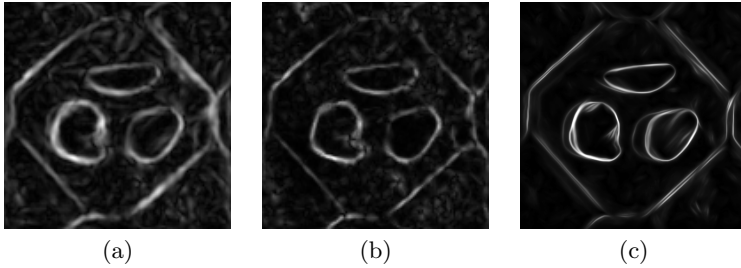


Abbildung 6.7: Ergebnis der Verstärkung salienter Konturen: (a) Ergebnis nach Stickvoting (aus [MBM02a]), (b) Votingprozess mit inhibitorischen Anteilen (aus [MBM02a]) und (c) Ergebnis der Selbstorganisation der Gradientenkarte nach Kap. 3.3.3.

Dies zeigt sich auch bei der Betrachtung der Kanten in Abbildung 6.8, die sich aus einer Kantendetektion nach erfolgter Konturverstärkung ergeben. Zum Vergleich ist in Abbildung 6.8(a) das Ergebnis einer Canny-Kantendetektion dargestellt.

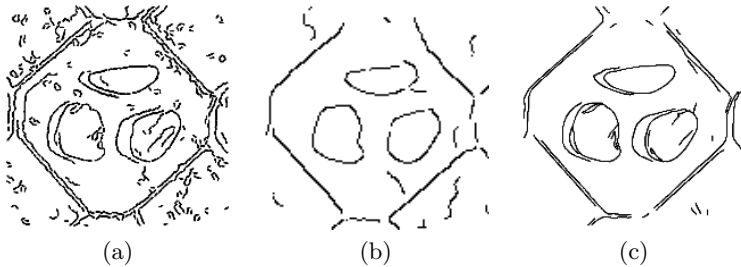


Abbildung 6.8: Detektierte Kanten zum Steine-Testbild: (a) Canny-Kantendetektion (aus [MTW99]), (b) Ergebnis nach Tensorvoting (aus [MM02]) und (c) Ergebnis der Kantendetektion nach Selbstorganisation der Gradientenkarte.

Auch hier zeigt sich, dass die vorgeschlagene Vorgehensweise Kanten lokalisierungsgenau detektiert und die Auswirkung der hochfrequenten Texturanteile auf die Kantendetektion unterbinden kann. Man beachte, dass im Gegensatz zu den in den Abbildungen 6.8(a) und (b) präsentierten Verfahren die hier vorgestellte Kantendetektion keinerlei Schwellwert benötigt.

6.3 Kanten- und Eckendetektion

Für die in Kap. 3.2.1 beschriebene schwellwertfreie Kantendetektion ist in Abbildung 6.8(c) bereits ein Beispiel gegeben. Die Funktionsweise soll nun anhand eines weiteren Testbildes erläutert werden. Zu der in Abbildung 6.9(a) dargestellten Laborszene sind in Abbildung 6.9(b) sämtliche Kandidaten einer Kantendetektion dargestellt, also relative Maxima der in Kap. 3.1 beschriebenen Kanten- und Eckenfilter und Nulldurchgänge der ZU-Filter.

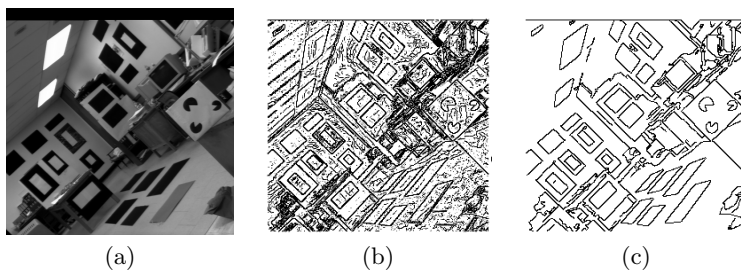


Abbildung 6.9: Laborszene: (a) Eingangsbild (aus [MS98]), (b) aus sämtlichen Filterergebnissen resultierende Kandidaten für Kanten und (c) detektierte Kanten nach Trackingverfahren.

Anstatt jedes Filtermaximum mit einer Schwellwertoperation zu versehen, wird nun, ausgehend von idealen Kombinationen der Filterantworten, wie in Kap. 3.2.1 beschrieben eine Kantenverfolgung realisiert. Das Ergebnis ist in

Abbildung 6.9(c) dargestellt und zeigt, dass aus der Menge aller möglichen Kanten die wesentlichen Kanten korrekt herausgebildet werden.

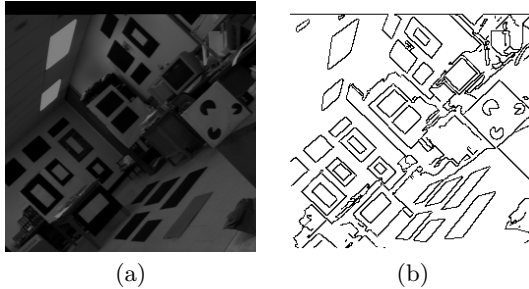


Abbildung 6.10: Laborszene: (a) Eingangsbild mit um 50% reduzierter Helligkeit und (b) detektierte Kanten.

Das Verwenden einer Schwellwertoperation, wie sie beim Cannydetektor oder gradientenbasierten Filtermasken genutzt wird, erzeugt eine hohe Empfindlichkeit dieser Verfahren bei einer Änderung der Helligkeit des Eingangsbildes, wobei dieser Effekt beim Cannydetektor durch Einsatz einer sogenannten Hysterese-Schwellwertbildung gemindert wird [ZT98, Can86]. Im Gegensatz dazu ist die in Kap. 3.2.1 beschriebene Kantenverfolgung relativ unabhängig von der Beleuchtung des Bildes, wie Abbildung 6.10 zeigt. Das Eingangsbild, Abbildung 6.10(a), zeichnet sich durch eine um 50% reduzierte Helligkeit gegenüber Abbildung 6.9(a) aus. Dagegen zeigen die detektierten Kanten in Abbildung 6.10(b) kaum Unterschiede zu denen in Abbildung 6.9(b) und demonstrieren damit die Beleuchtungsunabhängigkeit des vorgestellten Verfahrens.

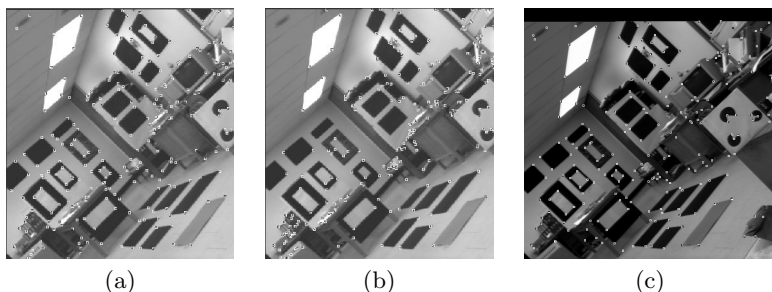


Abbildung 6.11: Vergleich von Verfahren zur Eckendetecktion (aus [MS98]): (a) CSS (Curvature Scale Space) [MS98], (b) SUSAN [SB97] und (c) Eckpunktdetecktion gemäß Kap. 3.2.2.

Abbildung 6.11(c) zeigt das Ergebnis der in Kap. 3.2.2 beschriebenen Detektion von Eckpunkten. Der Vergleich mit aktuellen Verfahren zur Eckendetecktion zeigt, dass die hier beschriebene Vorgehensweise zu vergleichbaren (Abbildung 6.11(a)) oder sogar genaueren Ergebnissen (Abbildung 6.11(b)) führt, wobei hier durch die bereits durchgeführten Vorverarbeitungsschritte kaum zusätzlicher Berechnungsaufwand entsteht.

Alle Ergebnisse zu Abbildung 6.9(a) wurden mit 18 Filterorientierungen und Filtermasken mit einer Größe von 14×14 Pixeln berechnet. Einige Kreuzungspunkte, die nicht in Abbildung 6.11(c) als Eckpunkt gekennzeichnet sind, wurden bereits im Verlauf der Kantendetektion als T-Kreuzung gekennzeichnet (vgl. Abbildung 6.9(c)). Für eine weiterführende Behandlung von Verfahren zur Detektion von Eckpunkten siehe [MS98, SB97, RT99, BKY02].

6.4 Konturvervollständigung

Das in Abbildung 6.12(a) dargestellte Eingangsbild zeigt wieder das Kanizsa Dreieck [Kan79], welches als klassisches Beispiel für Scheinkonturwahrnehmung in biologischen Sehsystemen gilt (vgl. Kap. 2.3). Abbildung 6.12(b) zeigt das Ergebnis der Bildvorverarbeitung. Über die orientierungsselektive Filterbank werden die im Bild vorhandenen Kanten extrahiert. Es werden außerdem sowohl Kreuzungspunkte, als auch die Richtungen möglicher Fortsetzungen detektiert, welche in Abbildung 6.12(b) heller als die detektierten Kanten dargestellt sind.

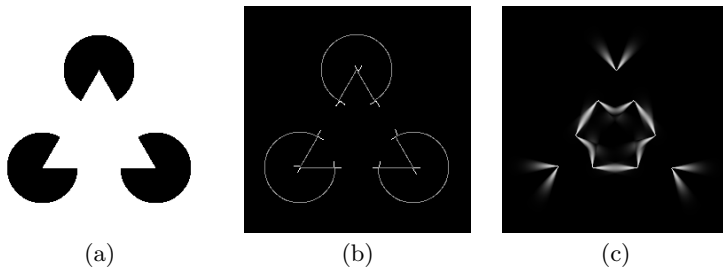


Abbildung 6.12: Kanizsa Dreieck: (a) Eingangsbild, (b) detektierte Kanten mit Fortsetzungsrichtung an Kreuzungspunkten und (c) Sticksalienz nach Votingprozess.

In Abbildung 6.12(c) ist das Ergebnis des Tensorvoting Prozesses visualisiert. Dargestellt ist die dünn besetzte Tensorkarte nach erfolgtem Votingprozess, wobei der Grauwert jedes Bildpunktes mit der Sticksalienz des jeweiligen Tensors korrespondiert. Falls eine sinnvolle Verbindung zwischen Kreuzungspunkten besteht, kann diese durch eine Flankenmaximumsuche gefunden werden.

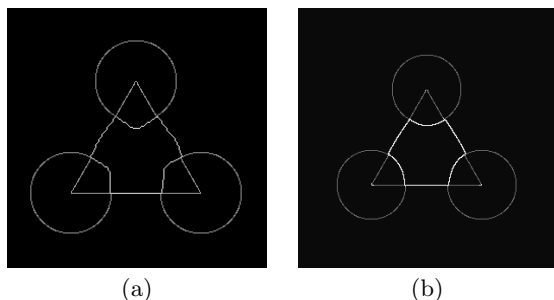


Abbildung 6.13: Konturvervollständigungen: (a) nach Tensorvoting und (b) nach Splineinterpolation.

Die derart bestimmten Konturvervollständigungen sind in Abbildung 6.13 zusammen mit den ursprünglich detektierten Konturen, dargestellt. Hierbei fällt auf, dass die durch das Tensorvoting erzeugte Konturvervollständigung in Abbildung 6.13(a) insbesondere für die Vervollständigung der Kreise eine ungleichmäßige Krümmung aufweist, was in der Skalierung der Votingfelder begründet ist, wie Abbildung 6.12(c) zeigt. Dagegen ist die in Abbildung 6.13(b) dargestellte Konturvervollständigung durch Splineinterpolation skalierungsunabhängig, die Krümmung kann hier über einen skalierungsunabhängigen Parameter eingestellt werden (vgl. Kap. 4.3). Dementsprechend kann hier unabhängig von der Größe des Eingangsbildes eine gleichmäßige Krümmung erzielt werden. Zusammenhängende Konturen werden in einem Kettencode gespeichert, darüber hinaus wird die Information festgehalten, ob es sich bei einem Konturabschnitt um eine real im Bild existierende Kontur handelt, oder ob der Konturabschnitt im Rahmen der Konturvervollständigung hinzugefügt wurde.

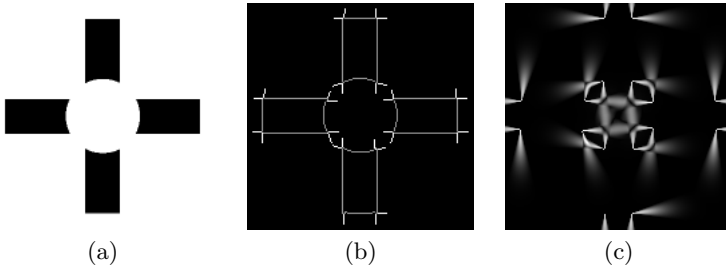


Abbildung 6.14: Überlagerung zweier Balken und einer Kreisfläche: (a) Eingangsbild, (b) detektierte Kanten mit Fortsetzungsrichtung an Kreuzungspunkten und (c) Sticksalienz nach Votingprozess.

Wie bereits in Kap. 4.1.3 erwähnt, kann eine Überkreuzung möglicher Vervollständigungen im Rahmen des Tensorvotingprozesses zu einer unerwünschten Überlagerung der Votingfelder führen. Dies ist in Abbildung 6.14(c) dargestellt. Die Vervollständigungsrichtungen, welche von den Konturunterbrechungen der schwarzen Balken in Abbildung 6.14(a) erzeugt werden, generieren Votingfelder, die sich mit rechtwinklig verlaufenden Feldern überlagern. Die Überlagerung führt zu Tensoren, deren zweiter Eigenwert λ_2 ansteigt und eine erhöhte Orientierungsunsicherheit signalisiert. Die Sticksalienz $\lambda_1 - \lambda_2$, welche die Orientierungssicherheit anzeigt, verschwindet dagegen stellenweise ganz, was zu einer fehlerhaften Konturvervollständigung führt, wie Abbildung 6.15 zeigt.

Während die Kreiskontur korrekt vervollständigt wird, ist dies für die Vervollständigungen der Balken nicht der Fall. Wie schon erwähnt, ließe sich dieses Problem umgehen, indem je zwei Vervollständigungsrichtungen gesondert betrachtet werden. Hierzu ist allerdings nicht nur eine separate Tensorkarte für jedes Verbindungspaar notwendig, es müsste zuvor auch eine Bewertung der Verbindungswahrscheinlichkeiten wie in Kap. 4.2 stattfinden. Dies widerspricht allerdings der Idee des Votingverfahrens, welches die

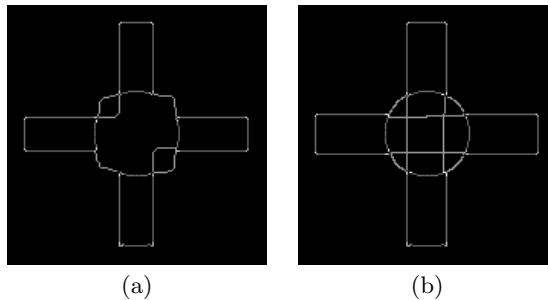


Abbildung 6.15: Konturvervollständigungen: (a) nach Tensorvoting und (b) nach Splineinterpolation.

Gestaltprinzipien der Nähe und guten Fortsetzung bereits durch Konstruktion des Votingfeldes berücksichtigt. Die Konturvervollständigung durch Splineinterpolation in Abbildung 6.15(b) führt dagegen erwartungsgemäß zu keinerlei Fehlzuordnungen, da hier die Verbindung der Kreuzungspunkte allein von der Position der Kreuzungspunkte und den Vervollständigungsrichtungen abhängt.

6.5 Perzeptuelle Organisation vervollständigter Konturen

Sowohl das Kanizsa Dreieck als auch die in Abbildung 6.16 dargestellten Ringe verdeutlichen die Zielsetzung perzeptueller Organisation, unterschiedliche Formen auch dann als Ganzes wahrzunehmen, wenn sie teilweise verdeckt sind. Für eine Tiefenzuordnung weisen die Ringe darüber hinaus die Problematik auf, dass keine binäre Entscheidung im Sinne von Vorder- oder Hintergrund getroffen werden kann, da jeder Ring in der sinnvollsten dreidimensionalen Interpretation sowohl dem Vorder- als auch dem Hintergrund des jeweils anderen Ringes zuzuordnen ist.

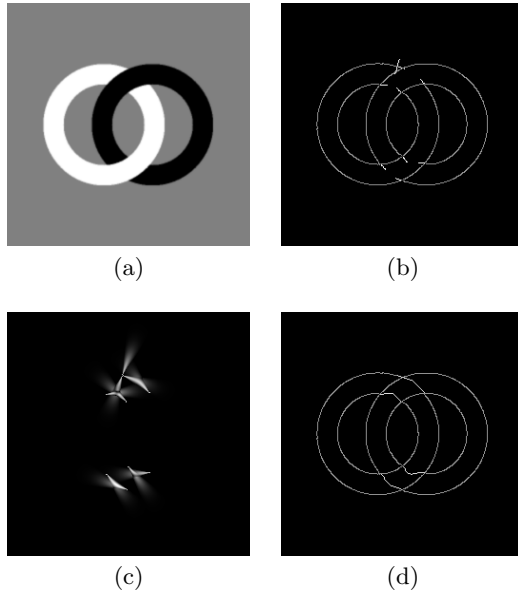


Abbildung 6.16: Ineinander verschlungene Ringe: (a) Eingangsbild, (b) detektierte Kanten mit Fortsetzungsrichtung an Kreuzungspunkten, (c) Sticksalienz nach Votingprozess und (d) Konturvervollständigungen nach Tensorvoting.

Zustandswechsel von realer zu vervollständigter Kontur geben im Bereich von Kreuzungspunkten Aufschluss darüber, ob eine Kontur die Objektbegrenzung eines Vordergrundobjekts darstellt, oder ob sie dem Hintergrund zuzuordnen ist. Durch Anwendung der in Kap. 5.1 beschriebenen Figur-Grund-Unterscheidung kann nun zunächst im Bereich von Kreuzungen zweier Konturen eine Unterscheidung von Vorder- und Hintergrund getroffen werden. Das Ergebnis des dynamischen Zuordnungsprozesses entsprechend Kap. 5.2, welcher die Tiefeninformation von den Kreuzungspunkten zum Rest der Kontur propagiert, ist in Abbildung 6.17 dargestellt. Nur Konturelementen ist hierbei ein Grauwert zugeordnet, sämtliche anderen Bild-

punkte sind schwarz dargestellt. Ein hoher, d. h. heller Grauwert repräsentiert eine relative Nähe zum Betrachter bzw. zum Kamerasystem, während dunkle Werte für eine höhere Entfernung zur Kamera stehen.

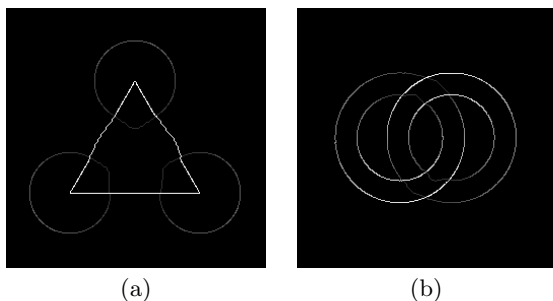


Abbildung 6.17: Tiefenzuordnung für (a) das Kanizsadreieck und (b) die Anordnung aus Abbildung 6.16(a).

Wie Abbildung 6.17 zeigt, kann mit der geschilderten Vorgehensweise eine Einteilung des Bildinhalts in überlappende Objektgrenzen und eine Extraktion monokularer Tiefeninformation aus den Eingabebildern realisiert werden. Darüber hinaus können auch komplexere Tiefenanordnungen, die einen Wechsel der Tiefenwerte entlang einer Kontur aufweisen, modelliert werden.

Die Problematik und Mehrdeutigkeit, welche mit einer Figur-Grund-Unterscheidung verbunden ist, ist in Abbildung 6.18 dargestellt, die das Verhalten der in Kap. 5.1 beschriebenen Figur-Grund-Unterscheidung demonstriert. Abbildung 6.18(a) weist starke Mehrdeutigkeiten auf, zu dem vervollständigten Kreis kann der weiße Balken sowohl als Vorder- als auch als Hintergrund interpretiert werden. In Analogie zur Musterung eines Zebras kann der Balken auch als Teil der Textur des Kreises aufgefasst werden. Da keine der in Kap. 5.1 beschriebenen Konditionen für eine Figur-Grund-Unterscheidung hier zutrifft, findet keine Tiefenzuordnung statt, die Konturen des Balkens erhalten die gleiche Tiefe wie die des Kreises. Auch Abbildung 6.18(b) weist

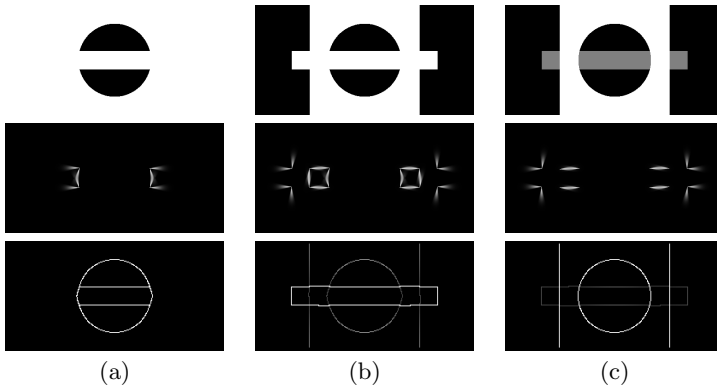


Abbildung 6.18: Obere Reihe: Eingangsbilder mit Kreis und Balken, mittlere Reihe: von Kreuzungspunkten entsprechend der Vervollständigungsrichtung generierte Votingfelder und untere Reihe: reale und vervollständigte Konturen mit zugeordneter relativer Tiefe.

immer noch Mehrdeutigkeiten auf, der weiße Balken kann sowohl vor einer schwarzen Kreisfläche als auch vor oder hinter einer weißen Fläche mit einem Loch liegen, die sich vor einem schwarzen Hintergrund befindet. Im Unterschied zu Abbildung 6.18(a) wird hier bereits durch die induzierenden Elemente und die Schließung der Kontur die Wahrnehmung des Balkens als ein Objekt forciert. Es gibt auch in diesem Fall keine eindeutig als richtig zu bewertende Interpretation, durch die in Kap. 5.1 formulierte Figur-Grund-Unterscheidung führt hier aber die Tatsache, dass sich die Kreisvervollständigungen im Inneren der geschlossenen Kontur des Balkens befinden, zu einer Interpretation, die die Objektgrenze des Balkens als Vordergrund bewertet. Schließlich lässt 6.18(c) den geringsten Spielraum für Mehrdeutigkeiten, da hier eine zusätzliche Farbinformation hinzukommt, welche die Wahrnehmung des Balkens als Hintergrund und des Kreises als Vordergrund forciert, wobei der Kreis hier sinnvoll als Objektgrenze eines Loches zu interpretieren ist. Die in Kap. 5.1 vorgestellte Figur-Grund-Unterscheidung

resultiert hier in einer eben solchen Interpretation, da die realen Konturen im Kreuzungsbereich von Kreis und Balken eine T-Kreuzung bilden, d. h. die Kontur des Balkens wechselt den Status von real zu vervollständigt, während die Kontur des Kreises durchgehend real bleibt.

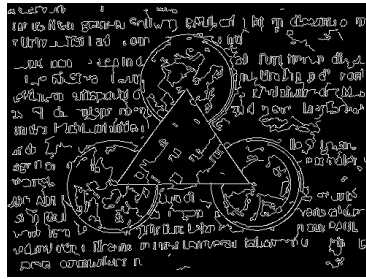
6.6 Natürliche Bilder

Das Ergebnis perzeptueller Organisation ist abhängig von der Qualität der Eckendetektion und der Orientierungsgenauigkeit von möglichen Konturf fortsetzungen. Während dies in den bisher behandelten synthetischen Bildern kein Problem darstellte, weisen natürliche Bilder eine Vielzahl an Eckpunkten auf, die zu falscher Detektion von Kreuzungspunkten und Vervollständigungsrichtungen führen können. Wie schon beschrieben, ist außerdem häufig eine Reduktion des Detailgrads des betrachteten Bildes wünschenswert, ohne wichtige Konturinformationen zu verändern. Durch die in Kap. 3.3 vorgestellte anisotrope Regularisierung kann allerdings einem Großteil dieser Problematik in einem weiteren Vorverarbeitungsschritt entgegengewirkt werden, wie Abbildung 6.19 zeigt.

Ohne den Vorverarbeitungsschritt der anisotropen Diffusion (vgl. Abb. 6.19(c)) liefert die Kanten- und Eckendetektion keine verlässlichen Informationen, wie Abbildung 6.19(b) zeigt. Grund hierfür sind die Schriftzeichen sowohl auf dem Papier, das den Hintergrund darstellt, als auch auf dem dem Stück Papier, das Teile der Münzen okkludiert und damit zu einer Tiefenkonstellation ähnlich dem Kanizsa Dreieck führt. Mit dem Vorverarbeitungsschritt der anisotropen Diffusion kann das Eingangsbild auf wesentliche Elemente reduziert werden, wodurch eine korrekte perzeptuelle Organisation gewährleistet wird (Abb. Abb. 6.19(d)).



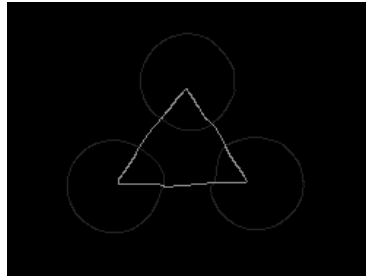
(a)



(b)



(c)



(d)

Abbildung 6.19: Drei Münzen auf beschriftetem Papier: (a) Originalbild, (b) Ergebnis der Kantendetektion auf dem Originalbild, (c) Eingangsbild nach anisotroper Diffusion mit schwellwert-behafteten Gewichtungsfaktoren und (d) Ergebnis perzeptueller Organisation mit vervollständigten Konturen und erfolgreicher Tiefenzuordnung.

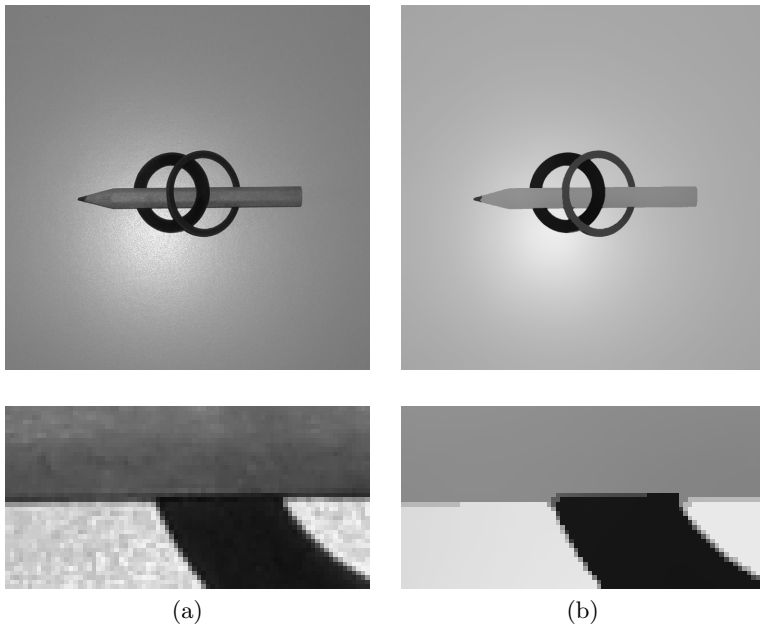


Abbildung 6.20: Bild eines Bleistiftes, der durch zwei Dichtungsringe gesteckt wurde: (a) Originalbild und (b) Ergebnis nach anisotroper Diffusion mit schwellwertbehafteten Gewichtungsfaktoren. Die untere Reihe zeigt eine Vergrößerung eines Bildbereiches mit sich kreuzenden Objektgrenzen.

Auch in Abbildung 6.20 erlaubt das durch die anisotrope Diffusion simplifizierte Bild eine korrekte Detektion von Kreuzungspunkten, die wiederum zu einer erfolgreichen Konturvervollständigung führt, wie Abbildung 6.21(a) zeigt. Auch hier besteht eine besondere Herausforderung in der Tiefenordnung der Objekte, die keinen konstanten Verlauf aufweist. Wie in Abbildung 6.21(b) zu erkennen ist, führt die Figur-Grund-Unterscheidung an Kreuzungspunkten und die anschließende dynamische Tiefenzuordnung zu einer korrekten dreidimensionalen Interpretation. Um dies weiter zu illu-

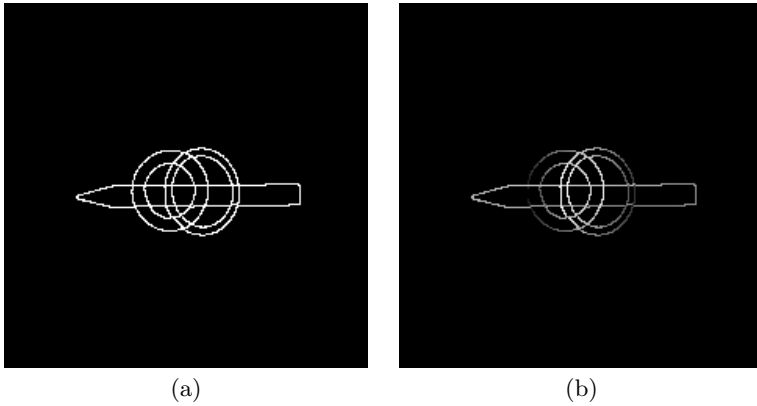


Abbildung 6.21: Bleistift-Testbild: (a) vervollständigte Kanten nach Splineinterpolation und (b) Tiefenwerte nach dynamischer Tiefenzuweisung.

strieren, sind in Abbildung 6.22 verschiedenen Seitenansichten der relativen Tiefe dargestellt.

Obwohl an den Kreuzungspunkten lediglich der Status des Vordergrundes oder des Hintergrundes für eine Kontur bestimmt wird, führt die dynamische Tiefenzuordnung, wie sie in Kap. 5.2 vorgestellt wird, zu einer dreidimensionalen Interpretation, welche die relative Lage der Dichtungsringe und des Bleistifts allein aufgrund des monokularen Tiefenkriteriums der Verdeckung qualitativ rekonstruieren kann. Die in Abbildung 6.22 dargestellten unteren beiden Tiefenwerte ergeben sich mit dem in Kap. 5.2 beschriebenen Nachverarbeitungsschritt, welcher auf der Interpretation von Objektgrenzen als Kreise auf einer geeigneten Ebene basiert. Wie Abbildung 6.22 zeigt, funktioniert dies selbst für ein Objekt wie den dargestellten Bleistift.

Abschließend ist zu bemerken, dass in dem zweidimensionalen Abbild einer realen Szene häufig eine Vielzahl von Objekten enthalten ist, die unterschiedliche Entfernungen zur Kamera aufweisen, aber über keine rekonstruierbaren

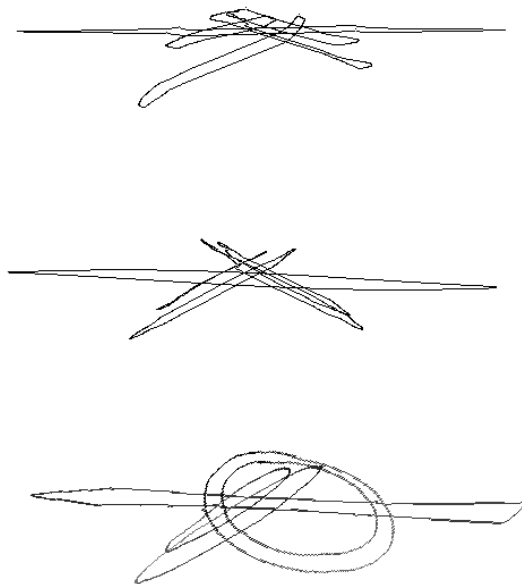


Abbildung 6.22: Bleistift-Testbild – Oben: Seitenansicht ohne Nachverarbeitungsschritt, Mitte: Seitenansicht mit Nachverarbeitungsschritt und Unten: perspektivische Seitenansicht der relativen Tiefenwerte mit Nachverarbeitungsschritt

verdeckten Objektgrenzen verfügen. Da in dem hier präsentierten Verfahren aber allein die rekonstruierbare Verdeckung als Tiefenkriterium genutzt wird, macht die Anwendung des Verfahrens in solchen Fällen ohne die Kombination mit weiteren Tiefenkriterien keinen Sinn.

6.7 Laufzeiten

Für das Grauwertbild der Abbildung 6.16(a) mit einer Größe von 256×256 Pixeln kann die Gesamtlaufzeit des Verfahrens auf unter einer Sekunde eingestellt werden. Die Berechnungen erfolgten dabei mit einer Intel Core 2 CPU 6600 mit 2,4 GHz und 4096 KB L2 Cache. Der größte Anteil der Laufzeit entfällt hierbei auf die Filterung. Mit einem Filterradius von 7 Pixeln und 12 betrachteten Orientierungen werden zur Filterung im Mittel 0,75 Sekunden benötigt. Bei einem Filterradius von 3 Pixeln und 5 Orientierungen reduziert sich die Berechnungszeit dagegen auf 0,12 Sekunden. Sowohl die Kanten- und Eckendetektion, die Konturvervollständigung durch Splineinterpolation, als auch die Figur-Grund-Unterscheidung benötigen mit weniger als 10 Millisekunden kaum Rechenzeit. Bei einem iterativen Prozess wie der dynamischen Tiefenzuweisung ist die Rechenzeit dagegen stark abhängig von der Parameterwahl, mit 20 Iterationen werden hier ca. 0,02 Sekunden benötigt. Ebenfalls rechenintensiv ist der iterative Vorverarbeitungsschritt der anisotropen Diffusion, der für Abbildung 6.16 nicht benötigt wird. Bei einer Umfeldgröße von 3×3 Pixeln und 20 Iterationen würde er 1,3 Sekunden beanspruchen, womit die anisotrope Diffusion rechenintensiver ist als der gesamte restliche Prozess der perzeptuellen Organisation.

7 Zusammenfassung

In Rahmen der vorliegenden Arbeit konnte eine Einteilung von Bildinhalten in überlappende Objektgrenzen realisiert werden, wobei Konturverläufe, die durch Verdeckung unterbrochen sind, erfolgreich rekonstruiert und zur Bestimmung der relativen Tiefenanordnung der gegebenen Konstellation herangezogen werden konnten. Die hier präsentierte Organisation von Objektgrenzen ermöglicht es, auch für Konturen von Objekten, die entlang des Konturverlaufs keinen konstanten Tiefenwert aufweisen, wie dies z. B. bei sich gegenseitig verdeckenden Objekten der Fall ist, eine sinnvolle dreidimensionale Interpretation der betrachteten Szene zu generieren. Es konnte somit gezeigt werden, dass eine Extraktion monokularer Tiefeninformation allein aufgrund von Verdeckungen möglich ist.

Grundlage für die perzeptuelle Organisation der Objektgrenzen ist die Detektion von Schlüsselpunkten, die durch eine Unterbrechung des Konturverlaufs durch ein verdeckendes Objekt im Falle amodaler Konturen oder durch ein Wechsel der Hintergrundfarbe zu einer dem Vordergrund identischen Farbe im Falle modaler Konturen entstanden sein können. Ausgehend von solchen Schlüsselpunkten findet eine Wahrscheinlichkeitserwägung statt, ob zwei Schlüssel- bzw. Kreuzungspunkte zueinander gruppiert werden können, und, falls dies der Fall ist, wie der Verbindungsweg dieser Punkte aussieht. Weiterhin wird unter Nutzung sowohl lokaler als auch globaler Kriterien die Anordnung von realen und vervollständigten Kanten an Schlüsselpunkten herangezogen, um sich kreuzende Konturen im Bereich von Schlüsselpunkten entweder dem Vorder- oder dem Hintergrund zuzuordnen. Aufgrund dieser Zuordnung findet durch einen Diffusionsprozess in einem abschließenden Schritt eine Propagation der Tiefeninformationen über die gesamte Kontur statt.

Das beschriebene Verfahren ist als Bottom-Up-Prozess realisiert und damit unabhängig von einer Wissensbasis. Dies hat zur Folge, dass die Wahrnehmungsprozesse in einem natürlichen Bild häufig in einem Detailgrad stattfinden, welcher nicht notwendigerweise der Seherfahrung eines Beobachters entspricht, der die Skalierung seiner Aufmerksamkeit von dem Wissen über die Größe und Anzahl der dargestellten Objekte abhängig macht. Darüber hinaus ist das Ergebnis abhängig von der Qualität der Eckendetektion und der Orientierungsgenauigkeit von möglichen Konturfortsetzungen. Während dies in synthetischen Bildern in der Regel kein Problem darstellt, weisen natürliche Bilder eine Vielzahl an Eckpunkten auf, die zu Störungen im Prozess der Scheinkonturwahrnehmung führen können. Um diesen Punkten Rechnung zu tragen und eine robuste Anwendung des beschriebenen Verfahrens auf natürlichen Bildern zu gewährleisten, wurde mit der vorgestellten anisotropen Regularisierung ein weiterer Vorverarbeitungsschritt realisiert, welcher es ermöglicht, ein Eingangsbild zu simplifizieren, d. h. hochfrequente Bildinhalte entsprechend einer gewünschten Skalierung aus dem Bild zu entfernen, ohne dabei Existenz und Lokalisierung markanter Konurverläufe zu verändern, wodurch das Bild auf die wesentlichen Bildinhalte reduziert wird.

Unter den im Verlauf der vorliegenden Arbeit berücksichtigten Themen der digitalen Bildverarbeitung, wie schwellwertfreier Kanten- und Eckendetektion, Rauschunterdrückung, Verstärkung salienter Konturen und Scheinkonturwahrnehmung, ist die anisotrope Regularisierung, insbesondere mit der vorgestellten Selbstorganisation der Gradientenkarte, von besonderem Interesse, da sie in allen gezeigten Anwendungen sehr vielversprechende Ergebnisse liefert und darüber hinaus eine Fülle weiterer Anwendungsmöglichkeiten, wie z. B. Inpainting, Bildvergrößerung oder Auflösung von Mehrdeutigkeiten in einem stereoskopischen Disparitätsraum bietet.

Die präsentierte Arbeit kann als ein erster Schritt in Richtung einer umfassenden maschinellen Tiefenwahrnehmung betrachtet werden, da gängige Verfahren zur Rekonstruktion der Tiefeninformation sich ausschließlich auf

die stereoskopische Korrespondenzbestimmung in Stereobildern, gegebenenfalls unter Berücksichtigung der zeitlichen Komponente, beschränken. Betrachtet man das visuelle System des Menschen, so liegt die Effizienz der Wahrnehmung nicht zuletzt in der Kopplung unterschiedlicher Tiefenkriterien begründet. Eine solche Kopplung von Tiefenkriterien ist also für maschinelle Bildverarbeitung durchaus denkbar und sinnvoll. So ließe sich die in der vorliegenden Arbeit vorgestellte Figur-Grund-Unterscheidung an Schlüsselpunkten ohne weiteres mit den Ergebnissen binokularer Tiefenwahrnehmung verknüpfen. Stereoskopische Tiefeninformation kann gerade an Schlüsselpunkten besonders gut zur Figur-Grund-Unterscheidung herangezogen werden, da Stereoalgorithmen in diesen Bereichen aufgrund des hohen Informationsgehaltes des Bildes besonders zuverlässige Aussagen treffen. Umgekehrt lässt sich die Konturinformation, welche durch die hier vorgestellte perzeptuelle Organisation auch teilweise verdeckte Objektgrenzen darstellt, sowohl zur Einschränkung des Suchraums stereoskopischer Korrespondenzbestimmung, als auch im Kontext von Objekterkennungsstrategien, die in der Regel auf eine vollständige Repräsentation von Objektgrenzen angewiesen sind, nutzen.

Literaturverzeichnis

- [AB98] ALTER, T. D.; BASRI, R.: Extracting Salient Curves from Images: An Analysis of the Saliency Network. In: *Int. J. Comput. Vision* 27 (1998), Nr. 1, S. 51–69. – ISSN 0920–5691
- [ABBFC97] AUBERT, G.; BARLAUD, M.; BLANC-FERAUD, L.; CHARBONNIER, P.: Deterministic edge-preserving regularization in computed imaging. In: *IEEE Trans. Imag. Process.* (1997). – 5(12)
- [Ade01] ADELSON, E. H.: On seeing stuff: the perception of materials by humans and machines. In: *Proceedings of the SPIE* Bd. 4299, Human Vision and Electronic Imaging VI, Bernice E. Rogowitz; Thrasyvoulos N. Pappas (Hrsg.), 2001, S. 1–12
- [ASF02] ANDERSON, B. L.; SINGH, M.; FLEMING, R. W.: The Interpolation of Object and Surface Structure. In: *Cognitive Psychology* 44 (2002), S. 148–190
- [Bab03] BABOS, Martin: *Anwendung des Tensor-Voting-Verfahrens auf Grauwertbilder*, Universität Hamburg, Diplomarbeit, Februar 2003
- [Bar04] BARTSCH, Hans-Jochen: *Taschenbuch Mathematischer Formeln*. 20. Auflage. Fachbuchverlag Leipzig im Carl Hanser Verlag, 2004
- [Bau78] BAUMGARTNER, G.: Physiologie des zentralen Sehsystems. In: *Gauer, Kramer, Jung (Hrsg.): Physiologie des Menschen* 13 (1978), S. 263–256
- [BHM04a] BROCKERS, Roland; HUND, Marcus; MERTSCHING, Bärbel: A fast cost relaxation stereo algorithm with occlusion detection for mobile robot applications. In: *Vision, Modeling, and Visualization 2004*. Stanford, USA, September 2004. – ISBN 3–89838–058–0, S. 47 – 53
- [BHM04b] BROCKERS, Roland; HUND, Marcus; MERTSCHING, Bärbel: Fast stereo vision for mobile robots by global minima of cost

- function. In: *6th IEEE Southwest Symposium on Image Analysis and Interpretation*. Lake Tahoe, Nevada, USA, 2004, S. 85 – 89
- [BHM05] BROCKERS, Roland; HUND, Marcus; MERTSCHING, Bärbel: stereo matching with occlusion detection using cost relaxation. In: *IEEE International Conference on Image Processing*, 2005
- [BKY02] BAE, Sun C.; KWEON, In-So; YOO, Choong D.: COP: a new corner detector. In: *Pattern Recognition Letters* 23 (2002), Nr. 11, S. 1349–1360
- [Can86] CANNY, J.: A computational approach to edge detection. In: *IEEE Trans. Pattern Anal. Mach. Intell.* 8 (1986), Nr. 6, S. 679–698. – ISSN 0162–8828
- [Dau88] DAUGMAN, J. G.: Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. In: *IEEE Trans. on Acoustics Speech and Signal Processing* 36 (1988), Nr. 7, S. 1169–1179
- [ESS03] EHRENSTEIN, Walter H.; SPILLMANN, Lothar; SARRIS, Viktor: Gestalt issues in modern neuroscience. In: *Axiomathes* 13 (2003), Nr. 3, S. 433–458
- [Fel06] FELDMANN, Michael: *Approximation des Verbindungsweges gerichteter Konturelemente durch analytische Funktionen*, Universität Paderborn, Baccalaureatsarbeit, 2006
- [Gab46] GABOR, D.: Theory of Communication. In: *Journal of the IEE* (1946), Nr. 93, S. 429–446
- [GCP05] GUERRERO-COLON, Jose A.; PORTILLA, Javier: Two-Level Adaptive Denoising using Gaussian Scale Mixtures in Overcomplete Oriented Pyramids. In: *IEEE International Conference on Image Processing (ICIP)* Bd. 1, 2005, S. 105–108
- [GG07] GUARDINI, Pietro; GAMBERINI, Luciano: *Zweiter Preis 2007 beim Best Visual Illusion of the Year Contest*. <http://illusioncontest.neuralcorrelate.com>, 2007 (Abruf: 01.04.2009)
- [GGM95] GOVE, A.; GROSSBERG, S.; MINGOLLA, E.: Brightness perception, illusory contours, and corticogeniculate feedback. In: *Visual Neuroscience* 12 (1995), Nr. 6, S. 1027–1052

- [GKP96] GEIGER, D.; KUMARAN, K.; PARIDA, L.: Visual organization for figureground separation. In: *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 1996, S. 155–160
- [GM96] GUY, Gideon; MEDIONI, Gerard: Inferring global perceptual contours from local features. In: *International Journal of Computer Vision* 20 (1996), Nr. 1-2, S. 113–133. – ISSN 0920–5691
- [GW08] GONZALEZ, Rafael C.; WOODS, Richard E.: *Digital Image Processing*. Prentice Hall, 2008. – ISBN 9780131687288
- [Hag07] HAGENLÜKE, Frank: *Bestimmung der Verbindungswahrscheinlichkeit gerichteter Konturelemente unter Berücksichtigung der Polarität orientierungsselektiver Filterantworten*, Universität Paderborn, Studienarbeit, 2007
- [Han03] HANSEN, Thorsten: *A neural model of early vision: Contrast, contours, corners and surfaces*, Universität Ulm, Diss., 2003
- [HBD03] HUND, Marcus; BROCKERS, Roland; DRÜE, Siegbert: Tiefenwahrnehmung für sichtgesteuerte mobile Roboter. (2003). – Vortrag, KI2003-Workshop Kognitive Robotik
- [Her70] HERMANN, L.: Eine Erscheinung simultanen Contrastes. In: *Pflügers Archiv für die gesamte Physiologie* 3 (1870), S. 13–15
- [Hey03] HEYDT, Rüdiger von d.: Image parsing mechanisms of the visual cortex. In: *The Visual Neurosciences* (2003), S. 1139–1150
- [HH93] HEITGER, F.; HEYDT, R. von d.: A computational model of neural contour processing: Figure-ground segregation and illusory contours. In: *International Conference on Computer Vision*, 1993, S. 32–40
- [HHP⁺98] HEITGER, Friedrich; HEYDT, Rüdiger von der; PETERHANS, Esther; ROSENTHALER, Lukas; KÜBLER, Olaf: Simulation of neural contour mechanisms: representing anomalous contours. In: *Image Vision Comput.* 16 (1998), Nr. 6-7, S. 407–421
- [HM05] HUND, Marcus; MERTSCHING, Baerbel: A computational approach to illusory contour perception based on the tensor voting technique. In: *10th Ibero-American Congress on Pattern Recognition (XCIARP)*, 2005

- [HM06] HUND, Marcus; MERTSCHING, Baerbel: Depth ordering and figure-ground segregation in monocular images derived from illusory contour perception. In: *17th European Conference on Artificial Intelligence (ECAI 2006)*, 2006
- [HM08] HUND, Marcus; MERTSCHING, Bärbel: Anisotropic diffusion by quadratic regularization. In: *3rd International Conference on Computer Vision Theory and Applications (VISAPP 2008)*, 2008, S. 101 – 107
- [HN01] HANSEN, Thorsten; NEUMANN, Heiko: Neural mechanisms for representing surface and contour features. In: *Emergent Neural Computational Architectures Based on Neuroscience - Towards Neuroscience-Inspired Computing*, Springer-Verlag, 2001. – ISBN 3-540-42363-X, S. 139–153
- [HN04] HANSEN, Thorsten; NEUMANN, Heiko: Neural mechanisms for the robust representation of junctions. In: *Neural Comput.* 16 (2004), Nr. 5, S. 1013–1037. – ISSN 0899-7667
- [HSN01] HANSEN, Thorsten; SEPP, Wolfgang; NEUMANN, Heiko: Recurrent long-range interactions in early vision. In: *Lecture Notes in Computer Science* Bd. 2036, 2001, S. 127
- [Hun02] HUND, Marcus: *Disparitätsbestimmung aus Stereobildern auf der Basis von Kostenfunktionen*, Universität Paderborn, Diplomarbeit, August 2002
- [HW70] HUBEL, D. H.; WIESEL, T. N.: Stereoscopic vision in macaque monkey. Cells sensitive to binocular depth in area 18 of the macaque monkey cortex. In: *Nature* 225 (1970), Nr. 5227, S. 41–42
- [Häm91] HÄMMERLIN, GÜNTHER UND HOFFMANN, KARL-HEINZ: *Numerische Mathematik*. 2. Auflage. Springer-Verlag, 1991
- [IK99] ITO, K.; KUNISCH, K.: An active set strategy based on the augmented Lagrangian formulation for image restoration. In: *M2AN Math. Model. Numer. Anal.* 33 (1999), S. 1–21
- [Jäh02] JÄHNE, Bernd: *Digital image processing (5th ed.): concepts, algorithms, and scientific applications*. Springer-Verlag, 2002
- [JHG99] JÄHNE, B. (Hrsg.); HAUSSECKER, H. (Hrsg.); GEISSLER, P. (Hrsg.): *Handbook Of Computer Vision And Applications*. Bd. 2. Academic Press, 1999

- [Jul60] JULESZ, Bela: Binocular depth perception of computer-generated patterns. In: *Bell System Technical Journal* 39 (1960), S. 1125–1162
- [Kan79] KANIZSA, G. (Hrsg.): *Organization in Vision*. Praeger, 1979
- [KGW01] KELLMAN, P.J.; GUTTMAN, S. E.; WICKENS, T. D.: Geometric and neural models of object perception. In: SHIPLEY, T. F. (Hrsg.); KELLMAN, P.J. (Hrsg.): *From fragments to objects: Segmentation and grouping in vision*, Oxford, UK: Elsevier Science, 2001
- [Kno00] KNOTT, Gary D.: *Interpolating Cubic Splines*. 18. Auflage. Birkhäuser Verlag, 2000
- [Kof35] KOFFKA, K.: *Principles of Gestalt psychology*. New York : Harcourt Brace, 1935
- [Kos93] KOSMOL, Peter: *Methoden zur numerischen Behandlung nicht-linearer Gleichungen und Optimierungsaufgaben*. 2. Auflage. Stuttgart : Teubner, 1993. – ISBN 3–519–12085–2
- [Kra63] KRAUSKOPF, J.: Effect of retinal image stabilization on the appearance of heterochromatic targets. In: *Journal of the Optical Society of America* 53 (1963), S. 741–744
- [Kra67] KRAUSKOPF, J.: Eeterochromatic stabilized images: A classroom demonstration. In: *American Journal of Psychology* 80 (1967), S. 634–637
- [Lee96] LEE, Tai S.: Image Representation Using 2D Gabor Wavelets. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18 (1996), Nr. 10, S. 959–971
- [Loc92] LOCHER, Franz: *Numerische Mathematik für Informatiker*. Springer-Verlag, 1992
- [Mar82] MARR, D.: *Vision: a computational investigation into the human representation and processing of visual information*. W. H. Freeman, San Francisco, 1982
- [Mas06a] MASSAD, A.: A perceptual grouping approach for visual interpolation between good continuation and minimal path using tensor voting. In: *BMVC06* Bd. 2, 2006, S. 639
- [Mas06b] MASSAD, A.: Visual interpolation for long-range contour completion by perceptual grouping based on tensor voting. In: *Pro-*

- ceedings of sixth IASTED International Conference for Visualization, Imaging, and Image Processing*, 2006, S. 537 – 542
- [MBM02a] MASSAD, A.; BABOS, M.; MERTSCHING, B.: Application of the tensor voting technique for perceptual grouping to grey-level images. In: GOOL, L. van (Hrsg.): *Pattern Recognition, 24th DAGM Symposium (DAGM2002)*, 2002, S. 306–313
- [MBM02b] MASSAD, A.; BABOS, M.; MERTSCHING, B.: Perceptual grouping in grey level images by combination of Gabor filtering and tensor voting. In: KASTURI, R. (Hrsg.); LAURENDEAU, D. (Hrsg.); SUEN, C. (Hrsg.): *ICPR Bd. 2*, 2002, S. 677–680
- [MBM03] MASSAD, A.; BABOS, M.; MERTSCHING, B.: Application of the tensor voting technique for perceptual grouping to grey-level images: quantitative evaluation, 2003. – Intl. Symposium on Image and Signal Processing and Analysis
- [Mei99] MEISTER, Andreas: *Numerik linearer Gleichungssysteme, Eine Einfuehrung in moderne Verfahren*. Braunschweig, Wiesbaden : Vieweg Verlag, 1999. – ISBN 3–528–03135–2
- [Mer99] MERZIGER, GERHARD UND WIRTH, THOMAS: *Repetitorium der Höheren Mathematik*. 4. Auflage. Binomi Verlag, 1999
- [MLT00] MEDIONI, G.; LEE, M.-S.; TANG, C.-K.: *A Computational Framework for Segmentation and Grouping*. Elsevier Science, 2000
- [MM01a] MASSAD, A.; MERTSCHING, B.: Segmentation of Spontaneously Splitting Figures into Overlapping Parts. In: RADIG, B. (Hrsg.); FLORCYK, S. (Hrsg.): *Pattern Recognition, 23rd DAGM Symposium*, 2001, S. 25 – 31
- [MM01b] MASSAD, Armin; MEDIONI, G.: 2-D Shape Decomposition into Overlapping Parts. In: ARCELLI, C. (Hrsg.); CORDELLA, L. (Hrsg.); BAJA, G. S. (Hrsg.): *Visual Form 2001, 4th International Workshop on Visual Form (IWVF 4)*. Capri, Italy, 2001, S. 398 – 409
- [MM02] MASSAD, A.; MERTSCHING, B.: Inferring salient features in images by perceptual grouping with inhibitory and excitatory tensor fields. In: WÜRTZ, R. (Hrsg.); LAPPE, M. (Hrsg.): *Dynamic Perception (DymPerc 2002)*. Bochum, 2002, S. 227 – 232

- [MS98] MOKHTARIAN, Farzin; SUOMELA, Riku: Robust image corner detection through curvature scale space. In: *IEEE Trans. Pattern Anal. Mach. Intell.* 20 (1998), Nr. 12, S. 1376–1381. – ISSN 0162–8828
- [MTW99] MAHAMUD, Shyjan; THORNER, Karvel K.; WILLIAMS, Lance R.: Segmentation of salient closed contours from real images. In: *IEEE International Conference on Computer Vision (ICCV)*, IEEE, 1999
- [MWTX03] MAHAMUD, Shyjan; WILLIAMS, Lance R.; THORNER, Karvel K.; XU, Kanglin: Segmentation of multiple salient closed contours from real images. In: *IEEE Trans. on Pattern Analysis and Machine Intelligence* 25 (2003), Nr. 4
- [Nie02a] NIEDER, Andreas: Seeing more than meets the eye: processing of illusory contours in animals. In: *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology* 188 (2002), Nr. 4, S. 249–260
- [Nie02b] NIEDER, Andreas: Die Wahrnehmung von Scheinkonturen - Wie sich das Gehirn Illusionen macht. In: *Neuroforum* (2002), Nr. 3, S. 210–217
- [NM01] NEUMANN, H.; MINGOLLA, E.: Computational neural models of spatial integration in perceptual grouping. In: SHIPLEY, T. (Hrsg.); KELLMAN, P. (Hrsg.): *From fragments to units: Segmentation and grouping in vision*. Oxford, UK : Elsevier Science, 2001, S. 353–400
- [Nor89] NORDSTROM, K. N.: Biased anisotropic diffusion—a unified regularization and diffusion approach to edge detection / EECS Department, University of California, Berkeley. 1989 (UCB/CSD-89-514). – Forschungsbericht
- [Pal99] PALMER, Stephen E.: *Vision Science: Photons to Phenomenology*. MIT Press, Cambridge, Massachusetts, 1999
- [PF05] PRADOS, E.; FAUGERAS, O.: Shape from shading: a well-posed problem? In: *Proceedings of CVPR'05*, IEEE, 2005
- [PH01] PETERHANS, Esther; HEITGER, Friedrich: Simulation of neuronal responses defining depth order and contrast polarity at illusory contours in monkey area V2. In: *Journal of Computational Neuroscience* 10 (2001), Nr. 2, S. 195–211

- [Pla72] Playboy's Playmate of the Month - Miss November. In: *Playboy Magazine, Playboy Enterprises, Inc.* 11/72 (1972)
- [PM90] PERONA, Pietro; MALIK, Jitendra: Scale-space and edge detection using anisotropic diffusion. In: *IEEE Trans. Pattern Anal. Mach. Intell.* 12 (1990), Nr. 7, S. 629–639
- [Por08] PORTILLA, Javier: *BLS-GSM denoising*. <http://decsai.ugr.es/~javier/denoise/>, 2008 (Abruf: 03.07.2008)
- [PSWS03] PORTILLA, J.; STRELA, V.; WAINWRIGHT, M.; SIMONCELLI, E.: Image denoising using scale mixtures of Gaussians in the wavelet domain. In: *IEEE Trans. Image Proc.*, 2003
- [PZ89] PARENT, Pierre; ZUCKER, Steven: Trace inference, curvature consistency, and curve detection. In: *IEEE Trans. Pattern Anal. Mach. Intell.* 11 (1989), Nr. 8, S. 823–839. – ISSN 0162–8828
- [RGM00] ROSS, W. D.; GROSSBERG, S.; MINGOLLA, E.: Visual cortical mechanisms of perceptual grouping: interacting layers, networks, columns, and maps. In: *Neural Netw.* 13 (2000), Nr. 6, S. 571–588. – ISSN 0893–6080
- [RT99] RUZON, Mark A.; TOMASI, Carlo: Corner detection in textured color images. In: *ICCV*, 1999, S. 1039–1045
- [Sau99a] SAUND, Eric: Perceptual organization of occluding contours generated by opaque surfaces. In: *Proceedings of the 1999 Conference on Computer Vision and Pattern Recognition*, 1999, S. 624–630
- [Sau99b] SAUND, Eric: Perceptual organization of occluding contours of opaque surfaces. In: *Comput. Vis. Image Underst.* 76 (1999), Nr. 1, S. 70–82. – ISSN 1077–3142
- [SB97] SMITH, Stephen M.; BRADY, J. M.: SUSAN - a new approach to low level image processing. In: *Int. J. Comput. Vision* 23 (1997), Nr. 1, S. 45–78. – ISSN 0920–5691
- [Sch00] SCHUMANN, F.: Beiträge zur Analyse der Gesichtswahrnehmungen. Erste Abhandlung. Einige Beobachtungen über die Zusammenfassung von Gesichtseindrücken zu Einheiten. In: *Zeitschrift für Psychologie und Physiologie der Sinnesorgane* 23 (1900), S. 1–32. – Englische Übersetzung durch A. Hogg (1987)

- in The perception of Illusory Contours, S. Petry, G.E. Meyer (New York: Springer) S. 40-49.
- [Sch68] SCHWARZ, H.: *Numerik symmetrischer Matrizen*. Stuttgart : Teubner, 1968
- [Spi01] SPILLMANN, Lothar: Gehirn und Gestalt II. Neuronale Mechanismen. In: *Kognitionswissenschaft* 9 (2001), Nr. 3, S. 122-143
- [ST97] SCHMIDT, R.F.; THEWS, G.: *Physiologie des Menschen*. 27. Auflage. Springer, 1997
- [SU88] SHA'ASHUA, Amnon; ULLMAN, Shimon: Structural saliency: the detection of globally salient structures using a locally connected network. 1988 (AIM-1061). – Forschungsbericht. – 15 S.
- [SVO93] SARY, G.; VOGELS, R.; ORBAN, G.A.: Cue-invariant shape selectivity of macaque inferior temporal neurons. In: *Science* 260 (1993), S. 995-997
- [SW00] SCHARR, Hanno; WEICKERT, Joachim: *An anisotropic diffusion algorithm with optimized rotation invariance*. 2000. – In G. Sommer, N. Kruger, C. Perwass (Hrsg.), Mustererkennung
- [Tra96] TRAPP, Ralph: Entwurf einer Filterbank auf der Basis neurophysiologischer Erkenntnisse zur orientierungs- und frequenzselektiven Dekomposition von Bilddaten. (1996). – Interner Bericht Nr. NAVIS 01/96
- [Tra98] TRAPP, Ralph: *Stereoskopische Korrespondenzbestimmung mit impliziter Detektion von Okklusionen*, Universität Paderborn, Fachbereich Elektrotechnik, Diss., September 1998
- [Wei96] WEICKERT, Joachim: *Anisotropic diffusion in image processing*, Universität Kaiserslautern, Diss., 1996
- [Wei99] WEICKERT, Joachim: Coherence-enhancing diffusion of colour images. In: *Image Vision Comput.* 17 (1999), Nr. 3-4, S. 201-212
- [Wer23] WERTHEIMER, Max: Untersuchungen zur Lehre von der Gestalt II. In: *Psychologische Forschung* 4 (1923), S. 301-350
- [WH96] WILLIAMS, Lance R.; HANSON, Allen R.: Perceptual completion of occluded surfaces. In: *Computer Vision and Image Understanding: CVIU* 64 (1996), Nr. 1, S. 1-20

- [Wil94] WILLIAMS, Lance R.: *Perceptual completion of occluded surfaces*, Dept. of Computer Science, University of Massachusetts at Amherst, Amherst, Mass., Diss., 1994
- [WJ96] WILLIAMS, Lance R.; JACOBS, David W.: Local parallel computation of stochastic completion fields. In: *Computer Vision and Pattern Recognition, CVPR* (1996), S. 161–168
- [WT00] WILLIAMS, Lance R.; THORNBUR, Karvel K.: A comparison of measures for detecting natural shapes in cluttered backgrounds. In: *International Journal of Computer Vision* 34(2/3) (2000), S. 81–96
- [Yar67] YARBUS, A. L. (Hrsg.): *Eye Movements and Vision*. Plenum, 1967
- [ZF84] ZURMÜHL, R.; FALK, S.: *Matrizen und ihre Anwendungen*. Berlin, Heidelberg, New York, Tokyo : Springer-Verlag, 1984
- [ZT98] ZIOU, D.; TABBONE, S.: Edge detection techniques: an overview. In: *International Journal on Pattern Recognition and Image Analysis* 8 (1998), Nr. 4, S. 537–559
- [ZW00] ZWECK, John W.; WILLIAMS, Lance R.: Euclidean group invariant computation of stochastic completion fields using shiftable-twistable functions. In: *ECCV* (2), 2000, S. 100–116