

Gruppierung von Kontursegmenten
mit
Mechanismen perzeptiver Gruppierung

Dissertation zur Erlangung des akademischen Grades

DOKTOR-INGENIEUR

an der Technischen Fakultät der Universität Bielefeld vorgelegt
von

Anke Maßmann

Bielefeld — November 2000

Gedruckt auf alterungsbeständigem Papier ∞ ISO 9706

Danksagung

Ohne aufmunternde kollegiale Unterstützung ist eine Arbeit wie diese, die zum Großteil während einer Vollzeittätigkeit in der Industrie im “fernen” Hannover entstand, nicht zu leisten. Deshalb gebührt mein Dank hier an erster Stelle allen meinen Kollegen, die mich während dieser Zeit begleiteten. Stellvertretend nennen möchte ich Grit Herrmann, Bernd Hildebrandt und Reinhard Moratz von der Arbeitsgruppe für angewandte Informatik. Ihre Freundschaft und stete Bereitschaft mich fachlich und menschlich zu unterstützen werde ich nicht vergessen. Auf keinen Fall unerwähnt bleiben darf mein langjähriger Bürokollege und unermüdlicher Leser meiner schriftlichen Ausarbeitung Thomas Fuhr. Herrn Lothar Kehrer danke ich für anregende interdisziplinäre Diskussionsrunden zur Wahrnehmungspsychologie und gemeinsame Seminare. Bei Daniel Schlüter bedanke ich mich, daß er stets bereit war mich vor Ort in Bielefeld zu unterstützen. Meinen Kollegen Rainer Arz und Peter Reinig von der ke-Kommunikationselektronik und Manfred Backhaus, Christian Kühne und Christian Malorny vom Anwendungsentwicklungszentrum der Deutschen Post in Hannover gilt mein Dank für Ihr stetes Verständnis meiner besonderen Situation.

Herr Prof. Dr.-Ing Stefan Posch hat meine Arbeit von Anfang an begleitet und betreut. Er war stets mein aufmerksamster Leser und Diskussionspartner. Ihm und Herrn Prof. Dr. Helge Ritter danke ich ferner für die Übernahme der Begutachtung.

Der Deutschen Forschungsgemeinschaft (DFG) danke ich für die geleistete finanzielle Unterstützung im Rahmen des SFB 360.

Last but not least bedanke ich mich ganz besonders bei meinem Mann Hans-Peter ohne dessen liebevolle Rücksichtnahme diese Arbeit nie vollendet worden wäre.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Schwerpunkt Gestaltpsychologie	1
1.2	Schwerpunkt Bildverarbeitung	3
1.3	Zielsetzung der vorliegenden Arbeit	5
1.4	Die Arbeit im Überblick	6
2	Gestaltpsychologie	9
2.1	Die “klassische” Gestaltpsychologie	9
2.2	Gruppierungsgesetze	11
2.2.1	Gesetz der Nähe	12
2.2.2	Gesetz der Ähnlichkeit oder der Gleichheit	12
2.2.3	Gesetz der guten Kurve	13
2.2.4	Gesetz der Geschlossenheit	14
2.2.5	Gesetz der Symmetrie	15
2.2.6	Figur-Grund Beziehung	16
2.2.7	Gesetz der gemeinsamen Bewegung oder des gemeinsamen Schicksals	17
2.2.8	Gesetz des gemeinsamen Bereichs	17
2.2.9	Gesetz des Zusammenhangs	18
2.2.10	Prägnanz	18
3	Gestaltpsychologie in der Bildverarbeitung	23
3.1	Gestaltmotivierte Klassifikation für Verfahren der digitalen Bildverarbeitung	24

3.1.1	Verfahren perzeptiver Gruppierung für 2D-Bilddaten	26
3.2	Zweidimensionale Gruppierungsverfahren	28
3.2.1	David Lowe und sein Fahrrad	28
3.2.2	Perzeptive Organization für spezielle und allgemeine Domänen	30
3.2.3	Perzeptive Organization und digitale Bildverarbeitung	34
4	Generierung von Gruppenhypothesen mit Bereichen perzeptiver Aufmerksamkeit	41
4.1	Domäne, Vorverarbeitung, Primitiva	42
4.2	Gruppierungssystem: Eine Hierarchie für Gruppenhypothesen	43
4.3	Suche nach gruppierbaren Primitiva	46
4.3.1	Motivation	46
4.3.2	Ermittlung von Bereichen perzeptiver Aufmerksamkeit	47
4.3.2.1	Untersuchung der Stichprobe	49
4.3.2.2	Untersuchung der Stichprobe für kollineare Gruppen	51
4.3.2.3	Untersuchung der Stichprobe für kurvilineare Gruppen	52
4.3.2.4	Untersuchung der Stichprobe für Nähe-Gruppen	53
4.3.2.5	Untersuchung der Stichprobe für parallele Gruppen	54
4.3.2.6	Akkumulation der Anschlußpunkte	55
4.3.3	Ableitung von Suchbereichen für den Gruppierungsprozeß	58
4.4	Generierung der Gruppenhypothesen	62
4.4.1	Generierung von Gruppenhypothesen mit Suchbereichen	62
4.4.2	Generierung von Geschlossenheiten	64
5	Markov Random Fields	69
5.1	Markov Random Fields und Bildverarbeitung	70
5.2	Das Ising Modell	72
5.3	Das Labeling Problem	73
5.4	Die Markov Eigenschaft: Nachbarn und Cliques	74
5.5	Gibbs-Markov-Äquivalenz	77

5.6	Optimierung	78
5.7	Energieminimierung	81
5.7.1	Energieminimierung mit Highest Confidence First	83
5.7.1.1	Die erweiterte Label-Menge	83
5.7.1.2	Stabilität und lokale Energie	84
5.7.1.3	Algorithmus	85
6	Ein Markov Random Field für Gruppenhypothesen	89
6.1	Mehrdeutigkeit der Gruppenhypothesen	90
6.2	Modellierung von Kontextwissen in einem Markov Random Field	90
6.2.1	Abbildung in einen ungerichteten Graphen	92
6.2.2	Signifikanz der Hypothesen	97
6.2.2.1	Datengetriebene Signifikanz	97
6.2.3	Design der Clique-Potentiale	104
6.2.3.1	Potential einer ein-einelementigen Clique	106
6.2.3.2	Potential einer zweielementigen Clique	106
6.2.3.3	Potentialfunktion einer dreielementigen Clique	111
7	Resultate und Ausblick	113
7.1	Vorstellung der Resultate	113
7.1.1	Schema zur Repräsentation	114
7.1.2	Grauwertbilder	115
7.1.3	Initiale Primitiva	115
7.1.4	Resultierende Gruppenhypothesen	116
7.1.4.1	Runde- und winklige Elemente der Szene	117
7.1.4.2	Auflösung von Verdeckungen	117
7.1.4.3	Generierung von Geschlossenheiten	119
7.1.4.4	Zusammenfassung	122
7.1.5	Bewertung der Hypothesen durch das Markov Random Field	124
7.1.5.1	Kanten, Knoten und Iterationen	125

7.1.6	Auflösung mehrdeutiger Geschlossenheiten	128
7.2	Ausblick	132
7.2.1	Primitiva des konturbasierten Ansatzes	132
7.2.2	Einbeziehung weiterer Primitiva	134
7.2.3	Kombination der Gruppenhypthesen	134
7.2.4	Bewertung der Gruppenhypthesen	135
8	Zusammenfassung	137
	Literaturverzeichnis	139
A		145
A.1	Legende der verwendeten Gruppen	145
A.2	Verwendete Variable	146
A.3	Gewichte und Maxima	149
B		
	Ergebnisbilder	151
B.1	Bild1:	152
B.2	Bild2	154
B.3	Bild3	156
B.4	Bild4	158
B.5	Bild5	160
B.6	Bild6	162
B.7	Bild7	164
B.8	Bild8	166
B.9	Bild9	168
B.10	Laufzeit	170

1 Einleitung

Das Umfeld der vorliegenden Arbeit ist bestimmt von zwei sehr verschiedenen wissenschaftlichen Disziplinen: der Wahrnehmungspsychologie und der digitalen Bildverarbeitung. Das Bindeglied dieser zwei Disziplinen stellen Forschungsarbeiten der Gestaltpsychologie zur menschlichen Wahrnehmung dar. Vereinfacht gesagt wurde festgestellt, daß die menschliche Wahrnehmung die außerordentliche Fähigkeit besitzt, spontan die Dinge ihres Sehfeldes zu strukturieren, auch wenn keine Erwartung oder auch Erfahrung über das Gesehene besteht. Das Erstaunliche ist, daß dieser Prozeß bei unvoreingenommenen Beobachtern ebenso abläuft, wie bei Experten des Gesehenen. Diese Mechanismen der Strukturbildung verleihen dem visuellen System des Menschen die intuitive Fähigkeit die wichtige Information in einer Szene von ihren unwichtigen Bestandteilen zu trennen. Die Mechanismen dieser intuitiven Strukturbildung wiederum sind von großem Interesse für die digitale Bildverarbeitung. Auch hier gilt es den Inhalt einer Szene zu analysieren, diesmal mit Methoden der automatischen Bildanalyse. Eine stärkere Strukturierung des Szeneninhalts kann die Komplexität und damit den Rechenaufwand mindern, aber auch Mehrdeutigkeiten in der Interpretation der Bilddaten verringern. Die vorliegende Arbeit hatte demnach zwei wesentliche Schwerpunkte: zum einen aus den Forschungsergebnissen der Gestaltpsychologie geeignete Mechanismen des menschlichen visuellen Systems zur Strukturbildung abzuleiten. Zum anderen geeignete Methoden der digitalen Bildverarbeitung zu finden, um diese Mechanismen der menschlichen Wahrnehmung an ein automatisiertes System zu adaptieren.

1.1 Schwerpunkt Gestaltpsychologie

Die Grundlagen der Gestaltpsychologie liegen zu Beginn des 20. Jahrhunderts. Wie es Rock & Palmer in [73] beschreiben entstand sie, wie viele bedeutende wissenschaftliche Disziplinen, in Auflehnung zur vorherrschenden Lehrmeinung ihrer Zeit: in ihrem Fall der Theorie des Strukturalismus. Einer der wesentlichen Annahmen des Strukturalismus, dem

Elementarismus, stellte die Gestaltpsychologie ihre Vorstellung von einer ganzheitlichen Wahrnehmung entgegen. Erklärt der Strukturalismus, daß eine Gesamtwahrnehmung in einzelne Bausteine zerlegbar ist: also die Wahrnehmung eines Quadrats nichts anderes sei, als die Empfindung einer speziellen Menge von Punkten, von denen jeder seinen Reiz auf der Netzhaut auslöst, hält die Gestalttheorie entgegen, daß die Wahrnehmung nicht aus einer Folge von einzelnen Reizen entsteht, sondern eine Gesamtheit darstellt, deren Bestandteile die einzelnen Reize seien. Das Netzhautbild eines Quadrats könne sich nach Lage und Größe so verändern, daß völlig andere Reize ausgelöst würden, trotzdem bliebe aber die Wahrnehmung des Quadrats erhalten. Der Hauptlehrsatz der Gestaltpsychologie lautet deshalb: "Das Ganze ist verschieden von der Summe seiner Teile". Die Form eines Quadrats ist demnach die Qualität einer Gesamterfahrung oder auch "Gestalt", die den einzelnen Bestandteilen nicht zukommt.

Die Gliederung der Wahrnehmung in Bereiche wurde durch die Gestalttheorie anhand der Gestalt- oder auch Gruppierungsgesetze verdeutlicht. Sie beschreiben, wie sich "Figuren" oder Gestalten in der Wahrnehmung "gruppieren" und so von ihrer Umgebung, dem "Grund" unterscheiden. Allein bis zum Jahre 1933 wurden insgesamt 114 Gestaltgesetze beschrieben. Die wichtigsten Gestaltgesetze, wie sie in der Zwischenzeit auch in digitale Bildverarbeitung Eingang fanden, sind:

- Gesetz der Nähe
- Gesetz der Ähnlichkeit oder der Gleichheit
- Gesetz der guten Kurve
- Gesetz der Geschlossenheit
- Gesetz der Symmetrie
- Figur-Grund Beziehung
- Gesetz der gemeinsamen Bewegung oder des gemeinsamen Schicksals
- Gesetz des gemeinsamen Bereichs
- Gesetz des Zusammenhangs
- Prinzip der Prägnanz

Jedes dieser Gesetze beschreibt einen bestimmten Aspekt, wie beispielsweise der Nähe oder der Ähnlichkeit im Wahrgenommenen, nachdem die Prinzipien der Gruppierung wirksam



Abbildung 1.1: Die Gruppierungsgesetze der Gestaltpsychologie sind ein Modell, wie das visuelle System des Menschen die Dinge des Sehfeldes in separate Einheiten strukturiert und sich so in seiner Umgebung orientiert. Werden Gruppierungsgesetze kontraproduktiv eingesetzt, wie dies bei der Tarnfärbung vieler Tiere vorkommt, sind sie in der Lage die Gegenstände mit ihrer Umgebung zu verschmelzen und gewissermaßen unsichtbar werden zu lassen. Hier demonstriert an der Zeichnung eines Hirschen nach Rock [18].

werden. Diese Gesetze dürfen jedoch nicht als isoliert angesehen werden: in Abbildung 1.1 ist beispielsweise erkennbar, wie schwarze, weiße oder graue Regionen aufgrund des “Zusammenhangs” und der “Ähnlichkeit” gruppiert werden und gleichmäßig gemusterte Bereiche aufgrund von “Ähnlichkeit” zusammengefaßt werden. Der aufmerksame Beobachter wird sicherlich noch mehr Prinzipien entdecken. Das Ganze ist gleichzeitig eine beispielhafte Verdeutlichung, wie sich Tiere anhand ihrer Tarnfärbung diese Prinzipien zu Nutze machen, um so gleichsam mit ihrer Umgebung zu verschmelzen.

1.2 Schwerpunkt Bildverarbeitung

Ein Betrachten der in der menschlichen Wahrnehmung wirksamen Prinzipien aus informationstheoretischer Sicht wurde bereits von Attneave in [1] formuliert. Er schlägt vor, die Wahrnehmung als einen informationsverarbeitenden Prozeß anzusprechen. Unter diesem Gesichtspunkt können die von den Gestaltpsychologen postulierten Gruppierungsgesetze

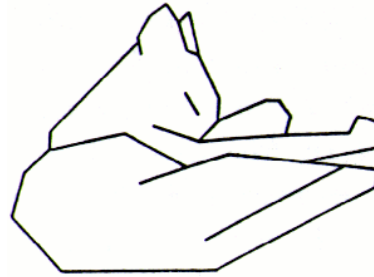


Abbildung 1.2: Für Attneave ist der Grad der Organisation in einer Szene gleichbedeutend mit Grad der Redundanz der enthaltenen Information. Umgekehrt konzentriert sich die Information an den Stellen der größten Änderung in der Szene, wie den Konturen und hier ganz besonders in den Stellen der stärksten Richtungsänderung. Er hat dies mit dem Bild einer schlafenden Katze demonstriert, indem er die Punkte maximaler Richtungsänderung durch Konturen verband und so zu der gezeigten Skizze kam.

als eine Reihe von Operationen zur Datenreduzierung betrachtet werden, welche die sensorische Information in eine Form kodieren, die ökonomischer ist als die Form, in der sie auf die Rezeptoren stößt. Er stellt beispielsweise die Hypothese auf, daß “gute Gestalten”, wie beispielweise ein Kreis oder ein Quadrat, genau deshalb gut sind, weil in ihnen weniger zu erinnernde Information enthalten ist. Umgekehrt ist für ihn die Information in einer Szene an den Stellen größter Änderung konzentriert: beispielweise in den Konturen von Regionen und ganz besonders an den Stellen, an denen eine Kontur ihre Richtung am stärksten ändert. Er hat dies anhand der Kontur einer schlafenden Katze visualisiert, indem er die Punkte maximaler Richtungsänderung durch gerade Striche verband. Die Figur ist trotzdem unverkennbar, wie in Abbildung 1.2 gezeigt.

Witkin und Tenenbaum beschreiben in ihrer Arbeit (siehe dazu auch [25], [46]), das von ihnen so genannte “principle of non-accidentalness”. Dieses Prinzip besagt, daß das Erscheinen von raum-zeitlicher Kohärenz oder Regularität im Bild, bzw. einer Bildfolge in der Regel nicht zufällig auftritt, sondern auf eine gemeinsame Ursache oder Prozeß in der Szene deutet. Sie motivieren mit ihrer Arbeit die grundlegende Beschäftigung mit der Analyse von Strukturen in der digitalen Bildverarbeitung. Will man nun die im letzten Abschnitt vorgestellten Gruppierungsgesetze an die Bedürfnisse der automatischen Bildanalyse adaptieren, sind zwei Dinge zu beachten: zum einen beschreiben die von der Gestaltpsychologie aufgestellten Gesetze Tendenzen wie beispielsweise Nähe oder Ähnlichkeit. Die genaue Wirkung dieser Faktoren ist aber keineswegs quantifiziert worden. Zum anderen ist zu beachten, daß sich alle Verfahren der digitalen Bildverarbeitung an der Untersu-

chung bestimmter Primitiva orientieren, wie beispielsweise Kanten oder Regionen. Auf die verschiedenen Primitiva sind in der Regel nicht alle der im Abschnitt zur Gestaltpsychologie beispielhaft vorgestellten Gruppierungsgesetze gleichmäßig anwendbar. Lowe stützt sich beispielsweise in seiner Arbeit (siehe dazu auch [15] und [16]) auf die Auswertung von Kontursegmenten und folgt damit der Empfehlung von Attneave. Die untersuchten Gestaltprinzipien sind Nähe, Kollinearität (“Gute Kurve”) und Parallelität (“Symmetrie”). Bereiche innerhalb derer diese Gestaltprinzipien wirksam sind, berechnet er aus Heuristiken über die Primitiva, bei denen die Länge des Primitivum eine wesentliche Rolle spielt. Da, wie oben ausgeführt, über den Prozeß der Gruppierung etliche Annahmen getroffen werden müssen, berechnet Lowe im Anschluß für jede der Gruppenhypothesen eine Signifikanz. Diese Signifikanz erlaubt eine Aussage darüber, ob die Bildung der Gruppe rein zufällig erfolgte oder eine gemeinsame Ursache zugrunde liegt.

1.3 Zielsetzung der vorliegenden Arbeit

Die vorliegende Arbeit beschäftigt sich mit einer Domäne die einem Holz-Konstruktionsspielzeug namens **Baufix** entstammt. Die Elemente dieser Domäne sind Holzbauteile, wie beispielsweise Leisten und Schrauben, aus denen sich in der untersuchten Domäne ein Flugzeug montieren läßt. Als initiale Primitiva des Gruppierungsprozesses dienen Kontursegmente, die in Geradenstücke und Ellipsensegmente unterschieden werden. Der Gruppierungsprozeß ist in drei hierarchisch angeordneten Ebenen eingeteilt: Die Ebenen sind in Richtung wachsender Komplexität mit **1D**, **2x1D** und **2D** benannt. Die erzeugten Gruppenhypothesen der 1D-Ebene sind Kollinearität, Kurvilinearität (Gute Kurve) und Nähe, der 2x1D-Ebene Parallelität (Symmetrie) und der 2D-Ebene Geschlossenheit. Die initialen Primitiva stehen auf allen Ebenen dem Gruppierungsprozeß zur Verfügung - die aus ihnen erzeugten Gruppenhypothesen, wie Kollinearität, gehen als Primitiva erneut in den Gruppierungsprozeß ein.

Das bereits im vorhergehenden Abschnitt angesprochene Problem der mangelnden Quantifizierbarkeit gestalttheoretischer Faktoren wurde hier durch die Untersuchung einer repräsentativen Stichprobe der Domäne angegangen. Ziel der Untersuchung war es, die gruppierende Wirkung eines Primitivum auf seine Umgebung zu bestimmen. Als Ergebnis entstanden Suchbereiche für kollineare, kurvilineare, Nähe- und parallele Gruppenhypothesen, die nicht auf willkürlichen Annahmen fußen, sondern aus den Charakteristika der verschiedenen Gruppentypen entstanden. Diese Bereiche wurden in Anlehnung an die Gestaltgesetze mit dem Begriff **Bereich perzeptiver Aufmerksamkeit** benannt.

Die Erzeugung der Gruppenthypothesen mit Hilfe der Bereiche perzeptiver Aufmerksamkeit ist lokal orientiert. Durch die lokale Betrachtungsweise entstehen mehrdeutige Gruppenthypothesen, die eine unerwünschte inkonsistente Interpretation der Szene bedeuten. Aus diesem Grund schließt sich eine Bewertung der gefundenen Gruppenthypothesen an den Gruppierungsprozeß an. Diese Verifikation der Hypothesen muß gleichzeitig die lokale Interaktion der benachbarten Gruppenthypothesen, sowie die Ermittlung einer global optimalen Interpretation berücksichtigen. Aus diesem Grund wurde der Formalismus eines Markov Random Field gewählt, weil er geeignet ist diese unterschiedlichen Anforderungen zu vereinen.

1.4 Die Arbeit im Überblick

Die vorliegende Arbeit wird mit Kapitel 2, einer Einführung in die Begriffswelt der Gestaltpsychologie, eingeleitet. Diese Einführung gliedert sich in zwei Teile: Der erste Abschnitt beschäftigt sich mit einem kurzen geschichtlichen Abriß der Gestaltpsychologie und stellt einige der bedeutenden Vertreter und ihrer Ideen vor. Daran anschließend werden die wichtigsten von der Gestaltpsychologie postulierten Gruppierungsgesetze vorgestellt. Das nachfolgende Kapitel 3 nimmt die eingeführten Begriffe aus Sicht der Bildverarbeitung erneut auf. Es beginnt mit einer Klassifikation der möglichen Ansätze, wie sie sich aus den verschiedenen Einsatzgebieten der Bildverarbeitung ergeben, orientiert an der Arbeit von Sakar & Boyer [54]. Ergänzt wird diese Klassifikation durch eine Übersicht beispielhafter Arbeiten zu den einzelnen Gebieten.

Im Anschluß an diese Klassifikation folgt eine vertiefte Vorstellung der Arbeiten von David Lowe, Rakesh Mohan und Sakar & Boyer, von denen die diese Arbeit maßgeblich beeinflußt wurde.

Der in der vorliegenden Arbeit verfolgte Ansatz wird mit Kapitel 4 eingeleitet. Eine kurze Einführung stellt dar, wie durch eine geeignete Vorverarbeitung die initialen Primitiva des Gruppierungsprozesses extrahiert werden und wie der Gruppierungsprozeß selbst in eine Hierarchie verschieden komplexer Level eingebettet ist. Der Schwerpunkt des Kapitels liegt auf der Vorstellung der Suchbereiche der verschiedenen Gruppenthypothesen, den Bereichen perzeptiver Aufmerksamkeit. Die Stichprobe der Domäne, ihre Auswertung und die Ableitung der Suchbereiche werden nacheinander vorgestellt. Abgeschlossen wird das Kapitel durch eine Vorstellung der Erzeugung der verschiedenen Gruppenthypothesen mit Hilfe dieser Suchbereiche.

Der Bewertung der Gruppenthypothesen durch ein Markov Random Field wurde aufgrund

der Komplexität ein eigenes Kapitel gewidmet. Vorangestellt ist Kapitel 5 welches eine Einleitung zum Thema Markov Random Fields darstellt. Hier wird auf die Eigenschaften von Markov Random Fields unter dem speziellen Blickwinkel ihrer Anwendung in der Bildverarbeitung eingegangen. Diese Einführung erklärt grundlegende Begriffe, wie Nachbarschaft, Nachbarschaftssystem, Labeling, Entität, Konfiguration, Clique und Cliquepotential. Es wird gezeigt, wie aus Entitäten und Nachbarschaftssystem ein ungerichteter Graph entsteht. Ein wichtiger Abschnitt dieses Kapitels ist die Vorstellung des Theorems von Hammersley und Clifford. Dieses Theorem beschreibt die Äquivalenz von Markov Random Fields und Gibbs-Verteilungen und ermöglicht damit eine Überführung der vielen lokalen Nachbarschaftssysteme in eine globale Betrachtung der Szene. Die Schätzung der Maximum a posteriori Wahrscheinlichkeit eines Markov Random Field ist unter Ausnutzung dieses Theorems gleichwertig durch die Minimierung einer Energiefunktion ermittelbar. Abgeschlossen wird dieses Kapitel mit der Vorstellung des Highest Confidence First eines Verfahrens zur Minimierung der Energiefunktion, welches im Rahmen dieser Arbeit zum Einsatz kam.

Die Bewertung der Gruppenhypothesen in Kapitel 6 nimmt die eingeführten Begriffe des Kapitels 5 erneut auf und erläutert die speziellen Anpassungen im Rahmen der vorliegenden Arbeit. Es wird einführend kurz vorgestellt, was mehrdeutige Gruppenhypothesen bedeuten und wie sie entstehen. Anschließend wird gezeigt, wie Gruppenhypothesen - unter Beachtung der verschiedenen Hierachiestufen - in einen ungerichteten Graphen abgebildet werden. Dabei werden die Nachbarschaftsrelationen Konkurrenz und Unterstützung vorgestellt. Daran anschließend erfolgt eine ausführliche Vorstellung der Bestimmung datengetriebener Label für die verschiedenen Gruppenhypothesen. Datengetriebene Label entstehen aus Heuristiken über die räumlich-geometrischen Eigenarten der Hypothesen. Eingesetzt werden datengetriebenen Label zur Berechnung einer "Stabilität" mit der der Prozeß der Ermittlung der Energieminimums mittels Highest Confidence First startet. Ein letzter wichtiger Punkt dieses Kapitels ist die Beschreibung des Designs geeigneter Potentialfunktionen zur Bestimmung der Clique-Potentiale. Das Design dieser Clique-Potentiale bedeutet die Einbringung von a priori Wissen in ein Markov Random Field.

Kapitel 7 ist der Vorstellung der Resultate gewidmet. Da sich sinnvolle Änderungen aus der Diskussion der Resultate ableiten lassen ist der Abschnitt "Ausblick" an dieser Stelle mit aufgenommen. Anhand von 9 Bildern, die repräsentativ für die untersuchte Domäne sind, wird die Bildung der Gruppenhypothesen und ihrer Bewertung näher untersucht. Diskutiert werden die Vor- und Nachteile des gewählten Ansatzes anhand ausgewählter Spezifika, wie der Auflösung von Verdeckungen oder der Generierung von Geschlossenheiten. Die Darstellung der Resultate erfolgt in einheitlicher Form. Visualisiert sind das

Grauwertbild, die initialen Primitiva des Gruppierungsprozesses, und die als signifikant bewerteten Gruppenhypothesen der verschiedenen Hierarchieebenen. Die bildliche Darstellung wird ergänzt durch Tabellen über Art und Anzahl der initialen Primitiva und der hypothetisierten Gruppen. Der Verlauf der Bewertungsprozesses durch das MRF ist durch Tabellen mit den relevanten Daten, wie z.B. der Anzahl der erzeugten Knoten und Kanten festgehalten.

Die sich anschließende Betrachtung sinnvoller Erweiterungen des vorliegenden Ansatzes wird unter den Schwerpunkten der Einbeziehung neuer Primitiva und weiterer Gestaltgesetze in den Gruppierungsprozeß angegangen.

2 Gestaltpsychologie

“Das Ganze ist mehr als die Summe seiner Teile”, dieser Max Wertheimer zugeschriebene Ausspruch ist in die Alltagssprache eingegangen, auch wenn die meisten Menschen die ihn benutzen nicht wissen in welchem Zusammenhang er gehört. Aus Teilen zusammengesetzte “Ganzheiten”, denen so eine neue Bedeutung zukommt, die den Teilen nicht gemein ist - ein Forschungsgegenstand der Gestaltpsychologie - der besonders im Bereich der Wahrnehmungsforschung seine Spuren hinterlassen hat.

Die in diesem Kapitel versuchte Einführung in die Gestaltpsychologie wird sich in zwei Teile gliedern: Der erste Abschnitt beschäftigt sich mit einem kurzen geschichtlichen Abriß der Gestaltpsychologie und stellt einige der bedeutenden Vertreter und ihrer Ideen vor. Daran anschließend werden die wichtigsten, von der Gestaltpsychologie postulierten, Gruppierungsgesetze eingeführt. Das nachfolgende Kapitel 3 wird sich dann der Anwendung dieser Gesetze im Bereich der digitalen Bildverarbeitung widmen.

2.1 Die “klassische” Gestaltpsychologie

Dieser Abriß über die Entstehungsgeschichte der Gestaltpsychologie stützt sich in der Hauptsache auf die Ausarbeitungen von Rock & Palmer [73], Jaeger [8] und Anzyferowa [31] zu diesem Thema. Die Geschichte der Gestaltpsychologie ist relativ kurz erzählt: die ersten Ideen entstanden um die Jahrhundertwende. Sie erlebte eine kurze Blütezeit in den zwanziger und beginnenden dreißiger Jahren unseres Jahrhunderts. Eng mit der Entstehungsgeschichte der Gestaltpsychologie verbunden ist die sogenannte “Berliner Schule der Gestaltpsychologie”. Durch sie kam die Gestaltpsychologie zu internationalem Ansehen. Die Emigration der führenden Vertreter gestaltpsychologischer Forschung brachte diese Disziplin in Deutschland fast vollständig zum Erliegen. Inzwischen gibt es wieder Zentren gestaltpsychologischer Forschung in den USA, Deutschland, Italien und Japan (siehe dazu auch [73]). Als die Hauptvertreter der “klassischen” Gestaltpsychologie gelten nach Anzyferowa die Herren: Wertheimer, Köhler, Koffka, Metzger, Lewin, Rubin, Dun-

cker, Katz und Michotte. Max Wertheimer (1880-1943), Wolfgang Köhler (1887-1967) und Kurt Koffka (1886-1941) werden allgemein hin als die Begründer angesehen. Sie erarbeiteten bereits vor dem ersten Weltkrieg am Psychologischen Institut der Akademie für Sozial- und Handelswissenschaften in Frankfurt/Main die grundlegenden Ideen. Für Rock & Palmer läßt sich der Beginn der Gestaltpsychologie genau datieren: Die Veröffentlichung eines Aufsatzes über Scheinbewegungen im Jahr 1912 durch Max Wertheimer. Eine **Scheinbewegung** entsteht, wenn durch zwei sukzessiv aufleuchtende Lichtquellen, vorausgesetzt der räumliche Abstand und die zeitliche Abfolge ist optimal gewählt, der Eindruck der Hin- und Herbewegung eines Objekts zwingend auftritt. Dieses Phänomen erscheint trotz des Wissens um die physikalischen Verhältnisse, und ohne daß dieser Eindruck von der tatsächlichen Bewegung eines Objekts unterscheidbar ist. Wertheimer bezeichnete diesen Effekt mit **Phi-Phänomen** und schloß aus dessen Auftreten, daß die Wahrnehmung des Ganzen von der Wahrnehmung seiner Teile völlig verschieden sein muß. Die Bewegung ist in diesem Sinne die **Gestalt**, die unterschieden ist, von den sie bildenden Teilen, den sukzessiv aufleuchtenden Lichtern. Als Vorläufer dieser Ideen gelten Ernst Mach, der in seiner "Analyse der Empfindungen" den Gedanken von der Gestaltqualität erläutert und Christian Ehrenfels, der diesen Gedanken von der Gestaltqualität als Grundeigenschaft der Wahrnehmung weiterentwickelt. Nach Ehrenfels werden einzelne Elemente, die im Bewußtsein existieren, durch Bewußtseinsakte zu Gestalten vereinigt. Im Unterschied zu Ehrenfels postulierte jedoch die spätere Gestaltpsychologie, daß das, was wir wahrnehmen, nicht die Summe oder Folge von Einzelempfindungen ist, sondern eine Gesamtheit (Gestalt), deren Bestandteile die Einzelempfindungen sind. Gleichzeitig kommt der Wahrnehmung der Gestalt eine Bedeutung zu, die den Bestandteilen nicht gemein ist. Diese Eigenschaft wird auch mit **Übersummativität** bezeichnet. Sie ist die Grundlage für den einleitend zitierten Ausspruch : "Das Ganze ist mehr als die Summe seiner Teile", dessen korrekte Wiedergabe nach Rock & Palmer [73] lautet: "Das Ganze ist verschieden von der Summe seiner Teile".

Gestalten besitzen neben der Übersummativität zwei wichtige Eigenschaften: Transponierbarkeit und Prägnanz. Mit dem Begriff **Transponierbarkeit** werden unveränderliche Eigenschaften in der Wahrnehmung einer Gestalt beschrieben: Die Veränderung absoluter Merkmale der Bestandteile verändert nicht die Wahrnehmung der Gestalt, sofern die Verhältnisse unter den Bestandteilen der Gestalt gleich bleiben. Solange z.B. alle Kanten eines Vierecks gleichlang sind, wird ein Quadrat wahrgenommen, unabhängig davon, ob sich Größe oder Lage des Quadrats ändert. Ein vielzitiertes Beispiel der Transponierung ist die unveränderte Wahrnehmung einer Melodie, die in unterschiedlichen Tonarten gespielt wird. Die Bestandteile der Melodie sind in diesem Fall die einzelnen Töne deren Abstand

trotz der Transponierung gleichbleibend ist.

Mit dem Begriff **Prägnanz** beschreibt die Gestaltpsychologie die Tendenz des visuellen Systems, gegen die regelmäßigste und symmetrische Wahrnehmung zu konvergieren. Dieses auch mit dem “Streben nach einer guten Gestalt” titulierte Verhalten äußert sich z.B. in der Wahrnehmung einer gestrichelten kreisförmigen Linie als Kreis. Die unsichtbaren Anteile einer teilweise abgedeckten Figur werden nach der gleichen Gesetzmäßigkeit stets regelmäßig und symmetrisch ergänzt.

Die hier beschriebenen Phänomene versuchten die frühen Gestaltpsychologen mit Hilfe der Isomorphie zu beschreiben. Isomorph, also gleichartig, seien die Wahrnehmung und die dadurch im Gehirn ausgelösten Prozesse. Dies bedeutet beispielsweise für die Wahrnehmung eines weißen Kreises auf grauem Untergrund, daß im Gehirn ein korrespondierendes kreisförmiges Gebiet existiert. Entlang der Umrandung des Gebietes bewegen sich relativ starke elektrische Ladungen, diese entsprechen der Kontur des Kreises. Um dieses Gebiet herum existiert ein Feld mit schwächeren Ladungen, dieses entspricht dem grauen Hintergrund. Wertheimer beschreibt die von ihm untersuchte Scheinbewegung zweier sukzessiv aufleuchtender Lichtquellen mit einer korrespondierenden Reizung zweier unterschiedlicher Stellen im Gehirn. Zwischen diesen Stellen fließen elektrische Ladungen hin und her, die den Eindruck der Scheinbewegung hervorrufen. Obwohl diese Theorien über die Arbeitsweise des Gehirns zwischenzeitlich von der modernen Neurophysiologie im wesentlichen widerlegt wurden (siehe dazu auch [73]), bleibt es unbestritten das Verdienst der Gestaltpsychologie wichtige Phänomene der Wahrnehmung entdeckt zu haben.

Die in diesem Abschnitt eingeführten Prinzipien der Übersummativität, Transponierbarkeit und Prägnanz gliedern die Wahrnehmung in Bereiche die vereinfachend mit “Figur” und “Grund” bezeichnet werden können. Die Wirkungsweise dieser Prinzipien wird durch die sogenannten Gestaltgesetze spezifiziert. Sie beschreiben, wie sich Figuren oder Gestalten in der Wahrnehmung von ihrer Umgebung, dem Grund, unterscheiden. Bis zum Jahre 1933 wurden insgesamt 114 Gestaltgesetze beschrieben. Mit einigen der wichtigsten, wird sich der folgende Abschnitt beschäftigen.

2.2 Gruppierungsgesetze

Die hier vorgestellten Gruppierungsgesetze werden die im vorhergehenden Abschnitt eingeführten grundlegenden Ideen der Gestaltpsychologie verdeutlichen. Neben den “klassischen” Gesetzmäßigkeiten, wie sie bereits durch die frühe Gestaltpsychologie beschrieben wurden, sind zwei durch Rock & Palmer entdeckte Zusammenhänge mit aufgenommen

worden: das “Gesetz des gemeinsamen Bereichs” und das “Gesetz des Zusammenhangs”. Sie beschreiben zwei wichtige Aspekte der Wahrnehmung, die nicht zuletzt auch für die digitale Bildverarbeitung von großem Interesse sind. Abschließend wird das Prinzip der Prägnanz noch einmal näher betrachtet, da es für die Bildung von Gruppen von elementarer Bedeutung ist.

2.2.1 Gesetz der Nähe

Für eine Anordnung gleichartiger Primitiva, wie in Abbildung 2.1 dargestellt, werden bevorzugt diejenigen Entitäten zusammengefaßt oder gruppiert, die dicht beieinander liegen. Der Begriff der **Nähe** ist relativ zu verstehen: sie zeichnet innerhalb der verschiedenen Anordnungen unter a), b) und c) von Abbildung 2.1 die jeweilig kleineren Abstände aus. Unter a) ist eine Anordnung von 10 Punkten dargestellt, die in zwei Gruppen von je 5 Punkten unterteilt wird. In den Abbildungen b) und c) wird die Anordnung in Spalten, bzw. Zeilen aufgrund der Abstände der Punkte gruppiert.

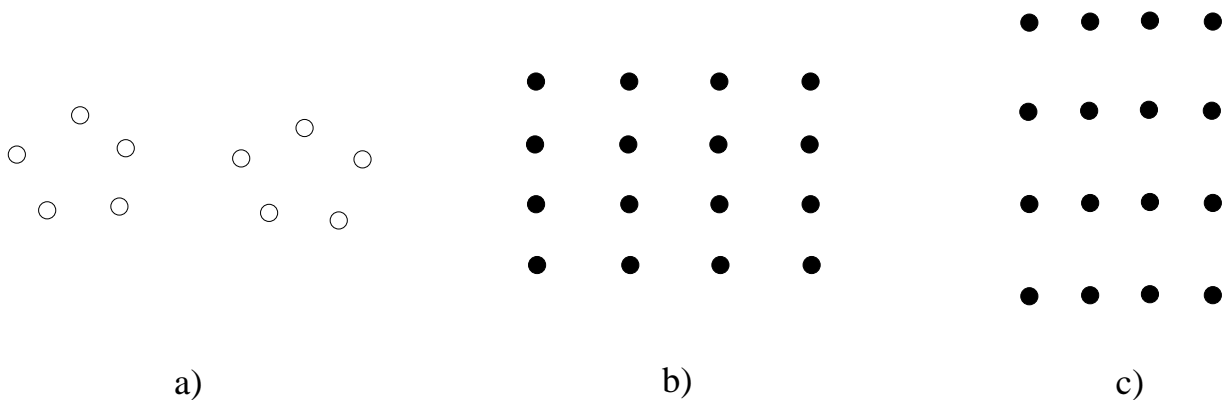


Abbildung 2.1: Gesetz der Nähe:

- a) Bildung zweier Gruppen von Punkten nach Köhler [12]
- b) und c) Gruppierung in Spalten, bzw. Zeilen nach Rock [19]

2.2.2 Gesetz der Ähnlichkeit oder der Gleichheit

Die Ähnlichkeit oder Gleichheit von Primitiva, z.B. in Bezug auf die Farbe, unterstützt die Bildung einer Gruppe. In Abbildung 2.2 ist unter a) dargestellt, wie eine Anordnung von - der Form nach - gleichartigen Primitiva durch unterschiedliche Färbung in zwei Gruppen zerfällt. In b) werden die Primitiva spaltenweise gruppiert. Unter c) ist dargestellt, wie

die Gruppierung nach Gleichheit der Farbe eine konkurrierende Gruppierung nach Nähe dominiert.

Gruppierung nach der Ähnlichkeit umfaßt Merkmale wie Form, Farbe, Helligkeit, Orientierung und Größe. Sie ist die Grundlage für Tests des Farbensehens: in einer Anordnung von Primitiva findet eine Gruppierung in bekannte Symbole wie Zahlen statt, die von Farbenblinden nicht gesehen werden. Die Tarnfärbungen von Tieren macht sich ebenfalls diese Gruppenbildung zunutze: Fische sind beispielsweise auf der Unterseite hell und auf der Oberseite dunkel. Je nach Betrachtungsrichtung verschimmen sie mit der hellen Farbe des Himmels oder der dunklen Farbe des Untergrundes.

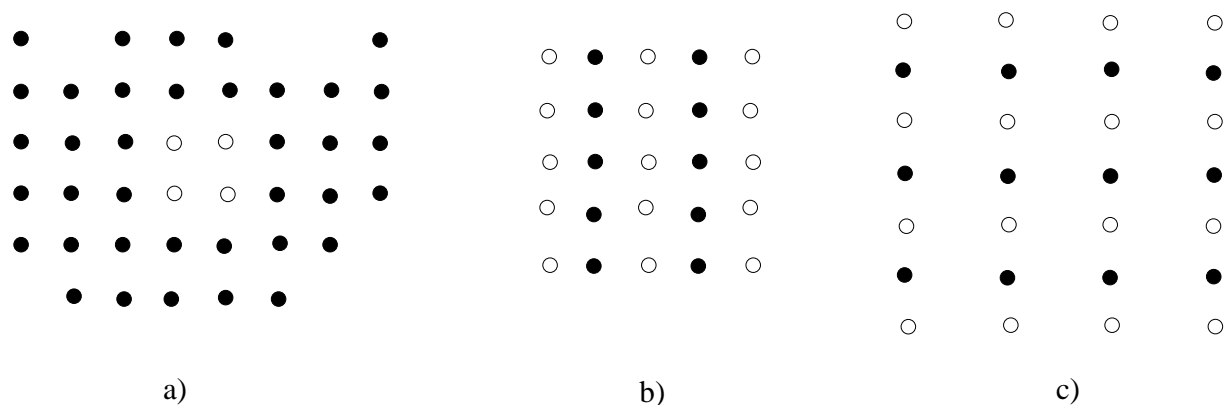


Abbildung 2.2: Gesetz der Ähnlichkeit:

- a) Gruppierung nach der Farbe nach Köhler [12]
- b) Gruppierung in Spalten entsprechend der Farbe nach Rock [19]
- c) Gruppierung in Reihen nach Gleichheit der Farbe, die eine Gruppierung nach Nähe unwirksam macht nach Rock [18]

2.2.3 Gesetz der guten Kurve

Die Teile eines Linienzuges oder einer Kurve werden als zusammenhängend wahrgenommen, bei denen eine stetige Fortsetzung ohne große Winkelabweichungen gegeben ist. Für das Beispiel in Abbildung 2.3 werden alle gestrichelten Fortsetzungen der Kurve unter a) als **gute Kurven** wahrgenommen. Im Gegensatz zu den Fortsetzungen unter b).

Unter c) und d) der gleichen Abbildung ist eine Anordnung gezeigt, die normalerweise vom Betrachter als sich kreuzende Linienzüge empfunden wird. Die unter d) durch die Umrandung angedeutete alternative Gruppierung der Anordnung wird normalerweise nicht wahrgenommen.

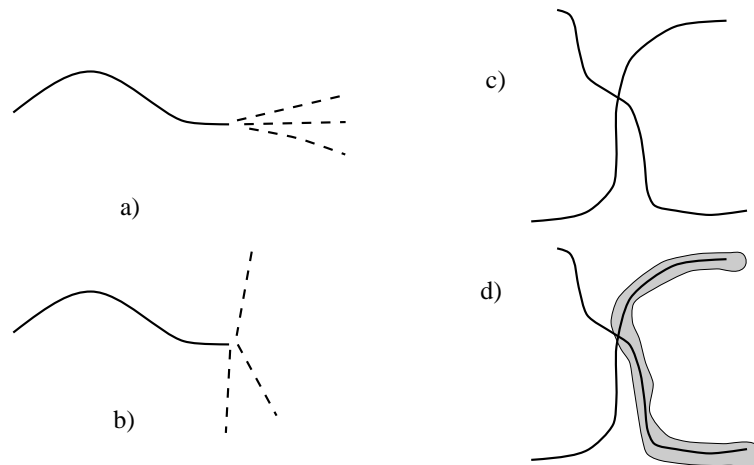


Abbildung 2.3: Gesetz der guten Kurve (siehe dazu auch [18]):

Die gestrichelten Linien unter a) stellen eine gute Fortsetzung der schwarzen Kurve dar. Im Gegensatz dazu das Beispiel unter b)

c) Die Figur wird als sich kreuzende Linienzüge wahrgenommen.

d) Die, durch die Umrandung der rechten Hälfte der Anordnung angedeutete, alternative Gruppierung wird normalerweise nicht gesehen.

Dieses Gesetz der guten Kurve hilft dabei Objekte "auseinander zu sehen". Oberflächenverläufe innerhalb der Teile eines Objekts sind normalerweise sanft, im Gegensatz zu Konturverläufen aneinander grenzender Objekte.

2.2.4 Gesetz der Geschlossenheit

Geschlossenheit bezeichnet die Tendenz, bevorzugt solche Primitiva zu gruppieren, die eine geschlossene Form darstellen. Die Linien die eine Fläche umschließen, werden leichter als Einheit aufgefaßt. In Abbildung 2.4 ist unter a) die Anordnung aus Abbildung 2.3 wieder aufgenommen, nur diesmal mit verbundenen Endpunkten. Als Ergebnis zerfällt jetzt die Anordnung in zwei geschlossene Entitäten. Die Gruppierung nach dem Prinzip der guten Kurve wird hier von der konkurrierenden Gruppierung nach Geschlossenheit dominiert. Unter b) und c) in Abbildung 2.3 zeigt ein weiteres Beispiel wie sich die Anordnung von b) durch Verbinden von Endpunkten c) neu gruppiert wird. Hier dominiert die Gruppierung nach Geschlossenheit eine konkurrierende Nähe-Gruppierung.

Wie diese Beispiele eindrucksvoll zeigen, ist das Gesetz der Geschlossenheit für die Gliederung einer beobachteten Szene von großer Bedeutung.

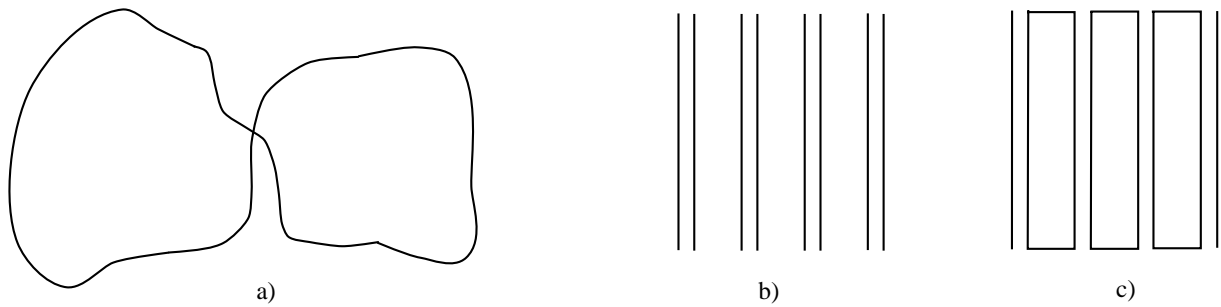


Abbildung 2.4: Gesetz der Geschlossenheit:

a) die Anordnung der kreuzenden Linien aus Abbildung 2.3 mit verbundenen Endpunkten: Die Figur zerfällt jetzt in zwei geschlossene Entitäten (nach Rock [18]).

Eine Anordnung von Linienzügen, die einmal durch die unterschiedliche Nähe b) gruppiert wird und unter c) durch Verbindung von Endpunkten eine Gruppierung nach Geschlossenheit erfährt (nach Katz [11]).

2.2.5 Gesetz der Symmetrie

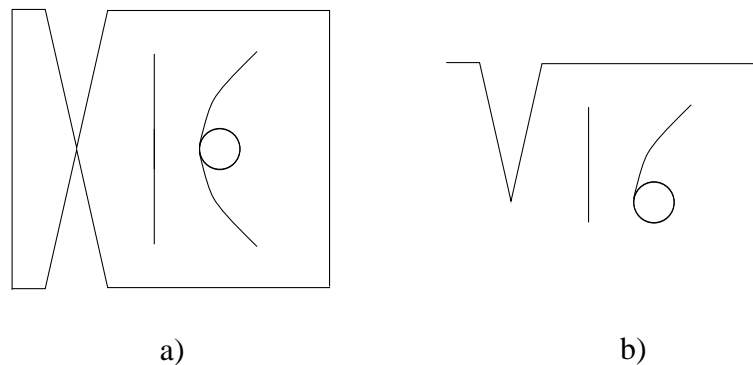


Abbildung 2.5: Gesetz der Symmetrie:

In a) ist eine Anordnung gezeigt, welche die Figur aus b) enthält. Diese ist jedoch spontan nicht erkennbar, da durch die symmetrische Ergänzung der Figur aus b) neue Zusammenhänge entstanden sind (nach Köhler [12]).

Linienzüge die sich symmetrisch ergänzen werden bevorzugt gruppiert. Diese Gruppierungen nach Symmetrie sind so nachhaltig, daß sie in der Lage sind, wohlbekannte Zusammenhänge zu überdecken. In Abbildung 2.5 ist unter a) eine Anordnung erkennbar, die die Figur aus b) enthält. Diese Figur wurde symmetrisch ergänzt und darüberhinaus an den Endpunkten verbunden. Ohne explizites Wissen über die ergänzte Figur, ist sie für einen

unvoreingenommenen Betrachter nicht erkennbar.

2.2.6 Figur-Grund Beziehung

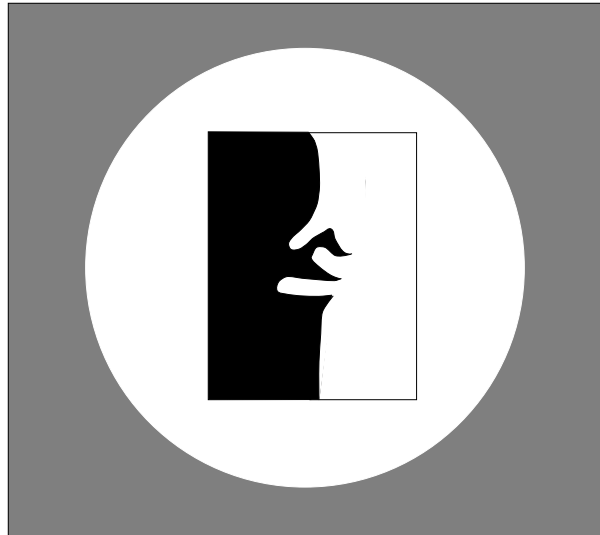


Abbildung 2.6: Figur-Hintergrund Trennung:

Die Anordnung zeigt ein Rechteck welches sich in eine weiße oder schwarze Figur unterteilt. Jede der beiden kann als Figur wahrgenommen werden, jedoch nie gleichzeitig beide (nach Rock [19]).

Dieser Prozeß ist grundlegend für jegliche Wahrnehmung von Objekten. Entscheidend sind die Grenzlinien zwischen den Flächen: die geschlossene Umrandung der kleineren Fläche wird als Figur vor einem Hintergrund wahrgenommen. Dieser Prozeß wurde 1921 von dem dänischen Psychologen Edgar Rubin beschrieben. Wie zentral in diesem Prozeß die Wahrnehmung der geschlossenen Umrandung ist, soll Abbildung 2.6 verdeutlichen. Gezeigt ist ein Rechteck, welches sich in eine weiße oder schwarze Region unterteilt. Jede der beiden Regionen kann als Figur wahrgenommen werden, jedoch niemals beide gleichzeitig. Die begrenzende Kontur zwischen den Regionen wirkt, je nachdem welche der beiden als Figur gesehen wird, als zu dieser gehörig. Die figuralen Wahrnehmungen beider Regionen sind trotz der gemeinsamen Kontur sehr unterschiedlich.

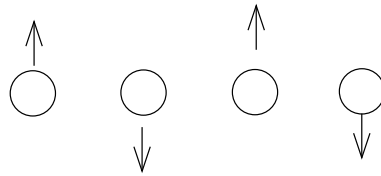


Abbildung 2.7: Gesetz der gemeinsamen Bewegung:

Der 1. und 3. Punkt, sowie der 2. und 4. Punkt bewegen sich jeweils paarweise aufwärts, bzw. abwärts. Sie werden entsprechend paarweise gruppiert (nach Rock [18]).

2.2.7 Gesetz der gemeinsamen Bewegung oder des gemeinsamen Schicksals

Primitiva, welche sich mit gleicher Geschwindigkeit in dieselbe Richtung bewegen, werden als Einheit empfunden. In Abbildung 2.7 sollen sich jeweils 2 der Punkte paarweise aufwärts, bzw. abwärts bewegen, angedeutet durch die Pfeile. Die Punkte werden durch einen Beobachter paarweise, entsprechend ihrer Bewegungsrichtung gruppiert, als ob unter ihnen eine unsichtbare Verbindung bestünde. Als weiteres Beispiel kann eine keilförmige Formation von Zugvögeln vor einem wolkenbedeckten Himmel dienen. Aufgrund der gemeinsamen Bewegung heben sie sich vor dem Hintergrund ab und werden als Einheit wahrgenommen.

2.2.8 Gesetz des gemeinsamen Bereichs

Ein Beobachter neigt dazu, Primitiva welche sich innerhalb eines abgeschlossenen Bereiches befinden, in Gruppen zusammenzufassen. In Abbildung 2.8 sind schwarze Punkte gezeigt, welche untereinander gleiche Abstände besitzen. Durch ihre paarweise Lage in einer gemeinsamen Region werden sie aufgrund dieser Zugehörigkeit gruppiert.

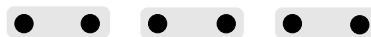


Abbildung 2.8: Gesetz des gemeinsamen Bereichs:

Die Punkte werden nach Rock & Palmer [73] paarweise durch die Bereiche gruppiert

2.2.9 Gesetz des Zusammenhangs

Einheitlich strukturierte, zusammenhängende Bereiche werden als Einheit gesehen. In Abbildung 2.9 sind in mehreren Reihen Anordnungen von Punkten dargestellt, die jeweils paarweise, durch eine Linie verbunden sind. Diese Gruppierung der Punkte nach ihrem

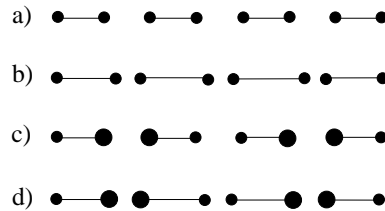


Abbildung 2.9: Gesetz des Zusammenhangs:

a)-b) Paarweise Gruppierung der Punkte durch eine Verbindungslinie konkurrierende Gruppierungen nach Nähe in b), Ähnlichkeit in c), oder einer Kombination aus beiden in d), werden durch den Zusammenhang dominiert (nach Rock & Palmer [44]).

Zusammenhang ist in der Lage, konkurrierende Gruppierungen nach Nähe in b), Ähnlichkeit in c) und einer Kombination beider in d), zu dominieren. Dies zeigt wie bedeutungsvoll dieser Zusammenhang für die Wahrnehmung einer Szene ist. Rock & Palmer messen der Gruppierung nach Zusammenhang eine fundamentale Bedeutung zu (siehe dazu auch [73], [44]). Die Eigenschaft eines Objektes ein zusammenhängendes Bild abzugeben, ist diejenige, die zu seiner Erkennung wesentlich ist. Die Existenz dieses Gesetzes ist daher nach Ansicht von Rock & Palmer derart naheliegend, daß Wertheimer dieses Prinzip aus diesem Grund nicht für erklärungsbedürftig hielt.

2.2.10 Prägnanz

Alle beschriebenen Gruppierungsgesetze waren nach der Auffassung von Wertheimer und anderen Gestaltpsychologen Manifestationen einer Tendenz zu bestimmten Gruppenbildungen, die sie mit den Adjektiven einfach, gut, stabil, regelmäßig, symmetrisch oder konsistent beschrieben (siehe dazu auch [18]). Sie wählten für diesen Zusammenhang das Wort **prägnant**. Prägnanz unter dem Aspekt der Einfachheit betrachtet, beschreibt die Bevorzugung der Gruppierung einfacher Formen. Der Begriff der Einfachheit ist jedoch nicht eindeutig und wird in der Regel intuitiv verstanden. In Abbildung 2.10 ist eine Anordnung von Linien dargestellt, die normalerweise als zwei übereinander liegende Rechtecke wahrgenommen werden. Die Anordnung ist jedoch mehrdeutig und könnte auch als ein Rechteck

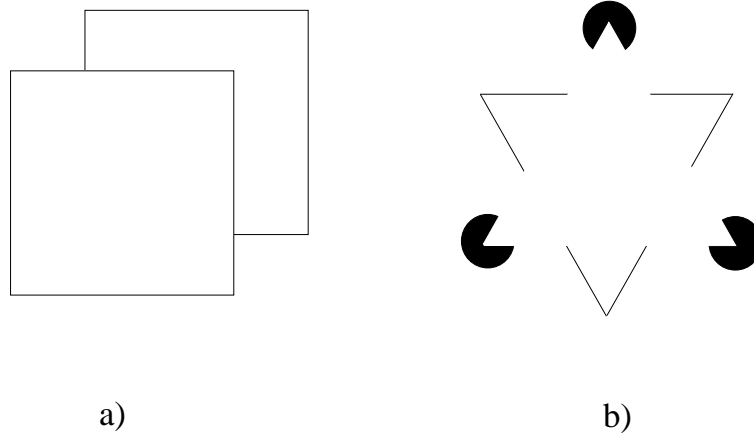


Abbildung 2.10: Tendenz einfacher Gruppenbildungen:

- a) Eine Anordnung, die als zwei sich überlagernde Rechtecke wahrgenommen wird. Die alternative Gruppierung in ein Rechteck und eine L-förmige Struktur - welche aneinander grenzen - wird nicht gesehen (nach Rock [18]).
- b) Eine Illusionsfigur, die als zwei Dreiecke und drei Kreisflächen gesehen wird, obwohl ein Dreieck objektiv nicht vorhanden ist (nach Kaniza [10]).

und eine L-förmige Struktur erscheinen, die aneinander grenzen. Diese Gruppierung wird jedoch normalerweise nicht gesehen. Eine Erklärung kann lauten: zwei Rechtecke sind eine einfachere Wahrnehmung, als zwei verschiedene Figuren. Eine weitere Erklärung kann lauten: Ein Rechteck ist eine einfachere Figur im Vergleich zu einer L-förmigen Struktur. Ein Rechteck ist regelmäßiger und symmetrischer und es hat weniger Seiten und weniger Winkel. Die Figur unter b) ist objektiv betrachtet, eine Ansammlung von drei winkligen Linienzügen und drei Kreisflächen mit einem fehlenden Segment. Sie wird jedoch als ein schwarz umrandetes Dreieck und drei Kreisen, die alle zusammen von einem weißen Dreieck überlagert werden, wahrgenommen. Die Konturen des weißen Dreiecks entstehen nur durch die geschickte Anordnung der schwarzen Elemente des Bildes, sie werden deshalb auch mit dem Ausdruck Scheinkonturen oder Illusionskanten bezeichnet (siehe dazu [10]). Auch hier wird die Tendenz zur Bildung einfacher Figuren deutlich, diesmal durch Ergänzung.

Die Bevorzugung symmetrischer Gruppierungen, wie sie durch das **Gesetz der Symmetrie** in Abbildung 2.5 dargestellt ist, kann in dem Sinn der Einfachheit als ein Aspekt von Prägnanz angesehen werden. Attneave erklärt in [1] diese Tendenz mit der Information, die zur Reproduktion einer Form gespeichert werden muß. Eine symmetrische Figur enthält

redundante Information: ist eine Hälfte der Figur bekannt, kann auf die zweite Hälfte geschlossen werden. Je einfacher und regelmäßiger die Form, desto weniger Information ist zu ihrer Reproduktion notwendig.

Ein weiterer Aspekt der Prägnanz ist in der internen Konsistenz einer Gestalt wirksam. Er kann so verstanden werden, daß eine solche Figur z.B. durch einen mathematischen

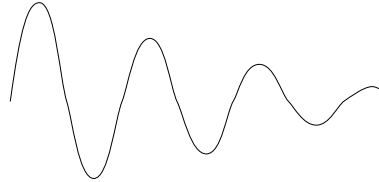


Abbildung 2.11: Ein Linienzug, der sich durch eine mathematische Funktion beschreiben läßt, verdeutlicht die Konsistenz einer Figur: wird ein Teil der Figur verdeckt, ist aus dem sichtbaren Rest gut zu ergänzen (nach Rock [18]).

Zusammenhang beschreibbar ist, wie die bedämpfte Sinusschwingung in Abbildung 2.11. Das Prinzip, welches diese Figur erzeugt, kann aus Teilen derselben abgeleitet werden und besitzt dennoch Gültigkeit für die gesamte Figur: Wird ein Teil der Figur verdeckt, kann sie aus dem sichtbaren Rest richtig ergänzt werden. Die Bevorzugung von **guten Kurven**, wie sie in Abbildung 2.3 dargestellt ist, kann in diesem Sinn als ein Aspekt von Prägnanz gedeutet werden.

Ein weiterer Aspekt von Prägnanz ist, daß sie besondere, herausragende Merkmale von Figuren beschreibt:

- von allen möglichen Orientierungen einer Linie sind horizontale, bzw. vertikale besonders gut unterscheidbar,
- von allen denkbaren Kurvenverläufen ist eine gerade Linie sehr gut unterscheidbar,
- von allen Winkeln die zwei Linien miteinander bilden können, ist der rechte Winkel besonders ausgezeichnet.

Die Präsenz solcher Merkmale ist wichtig für die Wahrnehmung einer Figur. Hat eine Figur z.B. eine horizontale Basis ist es besonders auffällig, wenn sie leicht um diese Basis gedreht wird.

Zusammenfassung: Dieses Kapitel beschäftigte sich mit den Begründern der Gestaltpsychologie und ihren grundlegenden Ideen und Prinzipien, wie sie in Abschnitt 2.1 vorgestellt

wurden. Die von der Gestaltpsychologie postulierten Prinzipien der Übersummativität, Transponierbarkeit und Prägnanz wurden einführend erklärt. Der Schwerpunkt dieses Kapitels lag auf dem zweiten Abschnitt 2.2, der die nachhaltigsten und einprägsamsten Ergebnisse gestaltpsychologischer Forschung vorstellte: die Gruppierungsgesetze. Diese Gruppierungsgesetze beschreiben die Gliederung von Wahrnehmungsinhalten in Bereiche die vereinfachend mit “Figur” und “Grund” bezeichnet werden können. Die Orientierung an diesen Gruppierungsgesetzen ist von großem Interesse für die automatische Analyse von Szenen. Wie diese aus der Untersuchung der menschlichen Wahrnehmung erkannten Zusammenhänge sinnvoll in der digitalen Bildverarbeitung eingesetzt werden können, damit wird sich das folgende Kapitel auseinandersetzen.

3 Gestaltpsychologie in der Bildverarbeitung

Der vorhergehende Abschnitt war den Begründern der Gestaltpsychologie und ihren grundlegenden Ideen gewidmet. Die von ihnen mittels der Gestaltgesetze beschriebenen Wahrnehmungsphänomene sind für die digitale Bildverarbeitung heute von großem Interesse, wie von Witkin & Tenenbaum [25], [46] und Lowe [15] herausgestellt. Menschen strukturieren spontan die Dinge des Wahrnehmungsfeldes, auch wenn sie keine Erwartung an das Gesehene haben. Dieser Prozeß funktioniert bei unvoreingenommenen Beobachtern einer Szene genauso gut, wie bei Experten des Dargestellten. Diese Strukturen bleiben in der Regel stabil, auch wenn der Beobachter die betrachtete Szene anschließend in einen semantischen Kontext einordnet. Wie Witkin und Tenenbaum feststellen, besitzt das visuelle System des Menschen die intuitive Fähigkeit, die wichtige Information in einer Szene von den unwichtigen Bestandteilen trennen zu können. Sie postulierten das sogenannte “principle of non-accidentalness”, welches besagt, daß das Erscheinen von raum-zeitlicher Kohärenz oder Regularität im Bild in der Regel nicht zufällig auftritt, sondern auf eine gemeinsame Ursache oder Prozeß in der Szene deutet. Man erhofft sich durch die Ausnutzung dieser in der menschlichen Wahrnehmung wirksamen Prinzipien auch in der digitalen Bildverarbeitung die Extraktion wichtiger Szeneninformation zu verbessern. Aus diesem Grund wird dieses Kapitel ebenfalls die Gestaltgesetze zum Thema haben, diesmal jedoch aus dem Blickwinkel der digitalen Bildverarbeitung. Dieser Bereich von Verfahren wird in der Bildverarbeitung üblicherweise mit “perzeptiver Gruppierung” oder auch “perzeptiver Organisation” angesprochen.

Alle Verfahren der digitalen Bildverarbeitung orientieren sich in der Regel an der Untersuchung bestimmter Primitiva, wie beispielsweise Kanten oder Regionen. Auf solche Primitiva sind in der Regel nicht alle im vorhergehenden Abschnitt 2.2 vorgestellten Gruppierungsgesetze anwendbar. Dieses Kapitel wird sich daher einleitend mit einer Klassifikation von verschiedenen Verfahren der Bildverarbeitung und der von ihnen verwendeten Primi-

tiva beschäftigen. Daran anschließend wird eine Anzahl verschiedener Ansätze vorgestellt, die entweder im Bereich meiner Arbeit liegen, bzw. von denen meine Arbeit maßgeblich beeinflusst wurde.

3.1 Gestaltmotivierte Klassifikation für Verfahren der digitalen Bildverarbeitung

Als Gruppierung kann in der Bildverarbeitung ganz allgemein jedes Vorgehen verstanden werden, welches in einer Szene extrahierte Primitiva zu einer neuen Einheit, der Gruppe, zusammenfaßt. Die Primitiva entsprechen dabei den Merkmalen der Szene, die durch eine gewählte Vorverarbeitung extrahiert wurden. Wurde beispielsweise für die Vorverarbeitung ein Kantenfilter und eine Konturpunktverktung eingesetzt, sind die Primitiva Ketten von Kantenpixeln, die eine Kante repräsentieren. Das anschließende Gruppierungsverfahren wird versuchen diese Primitiva beispielsweise in Geradenstücke zusammenzufassen. Analog zu der in Abschnitt 2.1 beschriebenen Übersummativität von Gruppen kommt der Gruppe "Geradenstück" eine Bedeutung zu, die den Bestandteilen, den Konturpunktketten nicht gemein ist. Ein Geradenstück kann durch Parameter wie Länge und Orientierung beschrieben werden und stellt so ein erheblich komplexeres Bildmerkmal dar. Die so gruppierten Geradenstücke können ihrerseits wieder die Primitiva eines erneuten Gruppierungsprozesses sein. Dies zeigt sehr gut, daß die von den unterschiedlichen Verfahren gruppierten Primitiva über eine sehr unterschiedliche Komplexität verfügen, anhand derer die verschiedenen Verfahren unterscheidbar sind. Es ist der Verdienst von Sakar & Boyer sich erstmals mit einer gründlichen Ausarbeitung diesem Thema gewidmet zu haben (siehe dazu auch [54]). Der nachfolgende Versuch einer Klassifikation verschiedener Verfahren der perceptiven Gruppierung, wird sich deshalb sehr eng an ihrer Arbeit orientieren. Abbildung 3.1 zeigt die Kriterien, anhand derer Sakar & Boyer ihre Klassifikation vornehmen: Die Spalten der abgebildeten Tabelle sind entsprechend der Bilddaten angeordnet. Dabei steigt die Komplexität der betrachteten Daten längs der x-Achsenrichtung von zwei- und dreidimensionalen Bilddaten (2D, 3D) zu zwei- und dreidimensionalen Bildsequenzen (2D + time, 3D + time) an. Die Zeilen der Tabelle sind entsprechend der betrachteten Primitiva in verschiedenen Schichten angeordnet, die Sakar & Boyer mit: Signal Level, Primitive Level, Structural Level und Assembly Level bezeichnen. Die Komplexität der gruppierbaren Primitiva ist längs der y-Achsenrichtung ansteigend. Jede so gebildete Zelle der Tabelle ist wiederum in zwei Zeilen unterteilt: Die obere Zeile beschreibt die Primitiva dieser Schicht und die untere Zeile beschreibt die möglichen Gruppierungen der Primitiva

Assembly Level	Structures found below organization of the structures arrangement of polygon	Structures found below Gestalt organizations of surface, clusters, polyhedral boxes, etc.	Structures found below Gestalt organization of flow element clusters	Structures found below organization of flows, groups of rigid bodies etc, articulated
Structural Level	Edge & Region primitives ribbons, corners, merges polygon, closed region 3D inferences, strands	Co-parametric surfaces, Edges useful surface combinations, polyhedral, cylindrical boxes, parallel surfaces	Optic Lines, Edge Primitives organization of flow elements, polygons, regions, etc.	Optic Lines, Surfaces flow elements, rigid bodies, non rigid motion
Primitive Level	Regions, Edge Chains region boundaries contour segmentation	Surface Patches, Clusters co-parametric surfaces, occlusion detection	Optic Flow Vectors swirls, vertices, sinks, sources	3D Optic Vectors Surface Patches optic lines, co-parametric surfaces
Signal Level	Dots, Interest Points dot clusters, edge chains regions, texture analysis	Range Images, Points in 3D-Space surface patches, discontinuities, point clusters	Moving Dots, Image Sequence optic flow, edge streaks	Dynamic 3D scenes optic vectors, surface segmentation
	2D	3D (2 1/2 D)	2D + time	3D + time

Abbildung 3.1: Klassifikation von Verfahren der perceptiven Gruppierung (nach Sakar & Boyer [54]).

in dieser Schicht. Für den **2D-Signal Level** sind die Primitiva beispielsweise Pixel die unter anderem zu Konturpunktketten gruppiert werden.

Sakar & Boyer haben im Sinne der Vollständigkeit ihrer Übersicht auch Gebiete mit vorgestellt, in denen noch keine Forschungsaktivitäten erfolgt sind. Dies gilt besonders für die Bereiche des **Structural-** und **Assembly-Level** zwei- und dreidimensionaler Bildsequenzen. Die von Ihnen zu den einzelnen Bereichen mit angegebenen Verweise auf beispielhafte Arbeiten zu den einzelnen Gebieten sind bei Interesse unter [54] nachzuschlagen. Der folgende Abschnitt wird sich in seiner Betrachtung der Tabelle aus Abbildung 3.1 auf zweidimensionale Bilddaten beschränken, da meine eigene Arbeit in diesem Umfeld angesiedelt ist. Eine detailliertere Einführung dreier Arbeiten aus dem Bereich des zweidimensionalen **Primitive-** und **Structural-Level**, von denen meine eigene Arbeit wesentlich beeinflusst wurde, wird danach dieses Kapitel abschließen.

3.1.1 Verfahren perzeptiver Gruppierung für 2D-Bilddaten

Die hier betrachteten Verfahren sind in der Spalte für zweidimensionale Bilddaten der Tabelle in Abbildung 3.1 einzuordnen. Auf der untersten, signalnahen Stufe dieser Spalte, dem **Signal Level**, entsprechen die Primitiva den Pixeln des Bildes. Diese werden beispielsweise in Konturpunktketten oder Regionen gruppiert. In den Bereich des Signal Level gehören die Verfahren der sogenannten niederen Bildverarbeitung, wie die Kantendetektion zur Extrahierung der Primitiva und die Konturpunktverkettung der gefundenen Primitiva als perzeptive Gruppierung. Diese Verfahren werden in der Regel nicht unter der Rubrik “perzeptives Gruppieren” in der Literatur beschrieben. Trotzdem können sie grundsätzlich diesem Bereich zugeordnet werden. Beschrieben sind sie in vielen grundlegenden Werken zur Bildverarbeitung, von denen hier - stellvertretend für viele - Ballard & Brown [27], Jähne[9] oder Besslich & Bässmann [42] genannt werden sollen.

Neben diesen eher “allgemeinen” Verfahren gibt es eine Reihe von Ansätzen in der niederen Bildverarbeitung, die aus dem speziellen Blickwinkel der perzeptiven Gruppierung entwickelt wurden, wie in [37], [59] und [53]. Viele Verfahren nutzen im wesentlichen eine Gruppierung nach den Prinzipien der Ähnlichkeit, Nähe oder des Zusammenhangs aus.

Die Primitiva des **Primitive Level** sind die Gruppen des Signal Level, beispielsweise Konturpunktketten oder Regionen konstanter Grauwerte. In diesen Bereich fallen viele Verfahren, die auf der Grenze der niederen zur mittleren Bildverarbeitung liegen. Ein typischer Vertreter solcher Verfahren ist die Approximation der verketteten Kanten-Pixel. Die auf der Bildmatrix beruhende Darstellung der verketteten Kanten-Pixel wird in eine geometrische Repräsentation der Daten, z.B. durch eine Approximation in Geradenstücke

oder Ellipsensegmente, wie in [69], [57], [75], [4], [45], [24], [20], [68], [30], [47] [63] beschrieben. Approximationen funktionieren in der Regel auf der Basis des Gesetzes der guten Kurve, indem sie Aussagen über Krümmungsverläufe oder Orientierung treffen. Eine typische Fragestellung der oben erwähnten Arbeiten ist daher die Art der notwendigen Aufspaltung der verketteten Kantenpixel durch die gewählte Approximation. Zu betrachten sind hier die gewählte Approximationsform wie Geradenstück und/oder Ellipsensegment, die die zugrunde liegenden Ketten fragmentieren. Die Unterteilung orientiert sich hier im wesentlichen an der Art des Krümmungs- und Steigungsverlaufs. Zum anderen ist die gewählte Granularität von Interesse: eine durch Rauschen zu einer Zickzackkurve verzerrte Kante kann in einer anderen Granularität durch eine Gerade repräsentiert werden. Fischler & Bolles [70] beschäftigen sich deshalb in ihrer Arbeit speziell mit diesem Problem der Aufspaltung von Linienzügen - unter Einbeziehung von umfangreichen Tests mit Versuchspersonen. Sie kommen zu dem Schluß, daß es keine einheitliche Definition für dieses Problem gibt, sondern immer eine problemabhängige Lösung gesucht werden muß.

Die Primitiva des **Structural Level** sind wiederum die Gruppen des **Primitive Level**. Die Gruppierung der Primitiva erzeugt zweidimensionale Strukturen. Der Großteil der Arbeiten zur perzeptiven Organisation in der Bildverarbeitung gehört in diesen Bereich, wie z.B. in [49], [48], [40], [35], [52], [61], [23], [72], [77], [56], [58], [41], [43], [74], [71], [76], [26], [36], [32], [33], [7], [5], [6], [38] und [39] beschrieben. Diese Ebene wird üblicherweise gemeint, wenn in der Bildverarbeitung von perzeptiver Organisation geredet wird. Der Großteil der Arbeiten ist konturbasiert. Typisch für diese Stufe der Gruppierung ist das Gesetz der Symmetrie, unter Einbeziehung der Parallelität als Sonderfall der Symmetrie. Die Prinzipien von Nähe, Zusammenhang und guter Kurve werden genutzt um Beschreibungen geschlossener Umrißkonturen zu erreichen. Die Gruppen dieser Ebene dienen bei vielen Ansätzen als Grundlage einer Objekterkennung.

Die letzte Stufe, der **Assembly Level**, gruppiert die Primitiva des **Structural Level** in größere Einheiten oder Aggregate. Dies kann z.B. durch eine Gruppierung von Rechtecken erfolgen, wie von Mohan und Nevatia in [17] und [65] beschrieben, um Umrisse von Häusern zu extrahieren. Im allgemeinen Fall können geschlossene Strukturen oder "Ribbons" gruppiert werden, um Oberflächen allgemeiner Objekte zu beschreiben, wie in [64], [67] und [66] dargestellt. Wollte man den Einsatz von Gestaltprinzipien interpretieren, könnte man das Gesetz des gemeinsamen Bereichs anführen, obwohl dies von den Verfassern der entsprechenden Arbeiten nicht immer getan wird. Die im letzten Level detektierten Primitiva werden entsprechend ihrer angenommenen Zugehörigkeit zu einem Objekt gruppiert.

Nach dieser Darstellung der Einordnung von Verfahren des perzeptiven Gruppierens in das

von Sakar & Boyer entworfene Schema, folgt eine detailliertere Darstellung dreier Arbeiten aus dem Bereich zweidimensionaler Szenen, von denen meine eigene Arbeit maßgeblich beeinflusst wurde.

3.2 Zweidimensionale Gruppierungsverfahren

Die im letzten Abschnitt vorgestellte Klassifikation von Verfahren perzeptiver Gruppierung in der Bildverarbeitung nach Sakar & Boyer wird hier mit einer eingehenden Betrachtung dreier Verfahren fortgesetzt. Diese Verfahren lassen sich entsprechend der Tabelle in Abbildung 3.1 dem zweidimensionalen Structural- bis Assembly Level zuordnen. Der Grund für die intensivere Beschäftigung ist, daß sie meine eigene Arbeit in vielen Dingen beeinflusst haben.

Im Sinne eines besseren Vergleichs der verschiedenen Ansätze werde ich eine einheitliche Einteilung in verschiedene methodische Schritte wählen. Dies entspricht nicht immer der von den Autoren selbst gewählten Form, bemüht sich jedoch im besten Sinn jedem der Verfasser gerecht zu werden. Der Vorteil liegt in besseren Vergleichsmöglichkeiten unter den verschiedenen Arbeiten, indem Gemeinsamkeiten und Unterschiede so leichter erkennbar werden. Die gewählten Schritte sind mit den folgenden Stichpunkten beschreibbar:

- Domäne, Vorverarbeitung, Bildprimitiva
- Suche in der Szene nach gruppierbaren Primitiva
- Repräsentation der Gruppen in einem System, z.B. in einer Hierarchie von Gruppen
- Bewertung, Auswertung, Auswahl der hypothetisierten Gruppen für eine Weiterverarbeitung in nachfolgenden Stufen

3.2.1 David Lowe und sein Fahrrad

David G. Lowe gehört zu den Pionieren der gestaltbasierten Bildverarbeitung. Seine sehr grundlegende Arbeit auf diesem Gebiet hat zweifellos auf viele ihren Einfluß hinterlassen. Berühmt ist sein "Fahrrad": eine Anordnung von schwarzen Strichen auf weißem Untergrund, die sich für einen Beobachter zu einem Fahrrad gruppieren.

Domäne, Vorverarbeitung, Primitiva Lowe benutzt grauwertbasierte Eingangsbilder mit gleichförmigen Gegenständen. In [15] und [16] ist an jeweils einem Beispielbild die Arbeitsweise des Systems gezeigt:

- Szene mit einer größeren Anzahl Einwegrasierer
- Szene mit zwei Papierheftern

Das System ist kanten-, bzw. konturbasiert. Die Kantendetektion erfolgt durch Kombination eines Marr-Hildreth und eines Sobel-Operators. Die gefundenen Kantenpixel werden verkettet und anschließend durch Geradenstücke approximiert. Diese Geradenstücke sind die Primitiva des Gruppierungsprozesses.

Gruppierungssystem Das von Lowe vorgestellte System (SCERPO, Spatial Correspondence, Evidential Reasoning, and Perceptual Organization) hat die Erkennung dreidimensionaler Objekte zum Ziel. Die hypothetisierten Gruppen dienen in diesem Rahmen als ein Indikator für die Präsenz eines Objekts. Dem System ist ein dreidimensionales Linienmodell des gesuchten Objekts bekannt. Dieses wird mit den Gruppenthypothesen verglichen und nach Möglichkeit in Übereinstimmung gebracht. In diesem Kontext sind für Lowe Strukturen von besonderem Interesse, die ihr Erscheinungsbild bei leichter Veränderung des Blickwinkels nicht verändern, wie beispielsweise Parallelitäten. Dieser Zusammenhang wird in [15] ausführlich diskutiert. Aus diesem Grund stützt Lowe sein System auf Gruppenthypothesen, welche nach den Relationen von Nähe, Parallelität und Kollinearität erzeugt werden. Eine Kombination dieser Gruppen ergibt neue Hypothesen mit einer höheren Komplexität (s.a. Abbildung 3.2). Für diese wird der Vergleich mit dem Linienmodell durchgeführt. Diese Kombination von Gruppenthypothesen entspricht einer Hierarchie der Hypothesen.

Suche nach gruppierbaren Primitiva Für eine effiziente Suche werden alle Endpunkte der Primitiva einer Szene in einer gitterartigen Datenstruktur, der "position matrix", angeordnet. Die Elemente dieser position matrix können gleichzeitig auch bezüglich Orientierung und Länge angesprochen werden. Lowe schränkt die Suche nach gruppierbaren Primitiva in einen Bereich um das betrachtete Primitivum ein. Dieser Bereich ist von der Länge des Primitivum abhängig.

Bewertung der Gruppen Lowe berechnet für jede der Gruppenthypothesen nach Nähe, Parallelität und Kollinearität eine Signifikanz. Diese erlaubt eine Aussage darüber, ob die

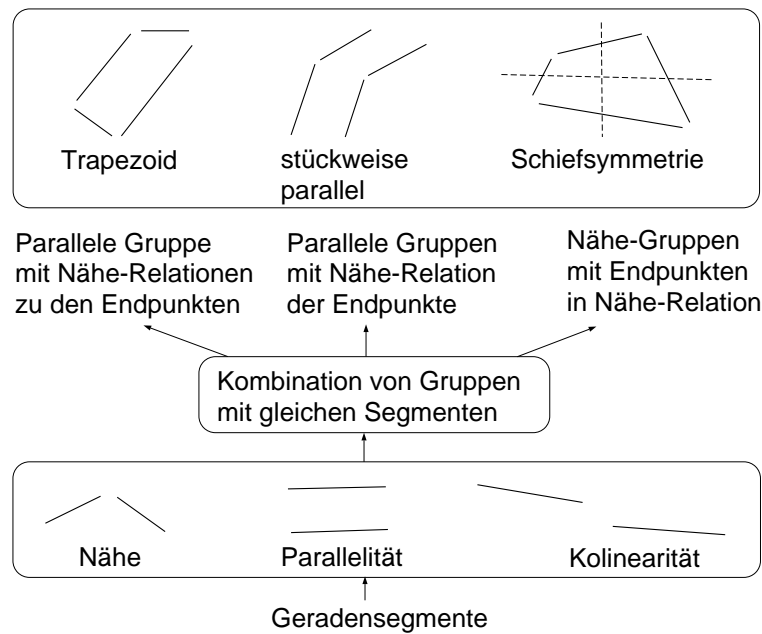


Abbildung 3.2: Hierarchie der Gruppenbildung nach Lowe [15]

Bildung der Gruppe rein zufällig erfolgte oder eine Regularität zugrunde liegt. Für die Berechnung der Signifikanz wird die Dichte von Endpunkten oder Primitiva in einem Bereich um das betrachtete Primitivum ausgewertet. Je geringer die Dichte im betrachteten Bereich, desto größer ist die Signifikanz der Hypothese.

Dieser Berechnung liegt die Annahme zugrunde, daß alle Primitiva der Szene in Bezug auf ihren Ort und ihre Orientierung gleichverteilt sind. Diese Gleichverteilung gilt weiterhin für unterschiedliche Skalierungen der Szene. Lowe führt in diesem Zusammenhang eine skalierungsinvariante Dichte D ein, die aufgrund der Gleichverteilung eine Konstante ist. Die Berechnung der Signifikanzen erfolgt rein datengetrieben. Für die Hypothesierung von Gruppen, in einer Kombination aus Nähe, Parallelität oder Kollinearität, wird die Signifikanz aus einer Multiplikation der Signifikanzen der zugrunde liegenden Gruppen berechnet.

3.2.2 Perzeptive Organization für spezielle und allgemeine Domänen

Rakesh Mohan beschreibt in seiner Arbeit [17] zwei Ansätze für die Generierung und Bewertung von Gruppenhypothesen. Beide Ansätze ergänzen sich gegenseitig in ihrer Zielsetzung:

A Der erste Ansatz untersucht Gruppierungen von geradenförmigen Primitiva ei-

ner sehr speziellen Domäne, in der bevorzugt rechtwinklige Strukturen auftreten (**specific shapes**). Aufgrund der eingeschränkten Domäne existiert ein a priori Wissen über die erwarteten Strukturen, welches den Gruppierungsprozeß steuert.

B Der zweite Ansatz untersucht eine Domäne mit allgemeinen Formen (**generic shapes**). Ein Vorwissen über die erwarteten Strukturen kann nicht eingebracht werden.

Die Darstellung wird sich schwerpunktmäßig auf den unter **A** beschriebenen Ansatz stützen, da er ein vollständiges System aus Erzeugung und Bewertung der Hypothesen darstellt. Die unter **A** entwickelten Ideen und Strategien treffen im Kern auch auf den Ansatz unter **B** zu.

Domäne, Vorverarbeitung, Primitiva Die Eingangsdaten des Systems unter **A** sind grauwertbasierte Stereobildpaare. Sie zeigen Luftbilder von Gebäuden. Da die Bildebene durch den gewählten Blickwinkel parallel zur Erdoberfläche liegt, sind die die Strukturen der Szene vorzugsweise rechtwinklig.

Die Eingangsdaten unter **B** sind grauwertbasierte Innenszenen mit verschiedenen Objekten wie Telefon, Spraydosen etc. in Großaufnahme. Allen Objekten ist gemeinsam, daß sie über abgerundete Oberflächen verfügen.

Beide Ansätze sind kanten, bzw. konturbasiert. Die Vorverarbeitung für den Ansatz unter **A** verwendet ein System welches auf dem "Nevatia-Babu line finder" basiert. Dieses liefert geradenförmige Kontursegmente als initiale Primitiva des Gruppierungsprozesses.

Der unter **B** verfolgte Ansatz detektiert die Kanten des Bildes mit dem Canny-Operator. Anschließend werden die Kantenpixel zu Konturpunktketten verbunden. Diese Konturpunktketten werden an Stellen mit starken Steigungsänderungen aufgetrennt, sodaß kurvenförmige Kontursegmente mit sanften Krümmungsverläufen als initiale Primitiva für den Gruppierungsprozeß entstehen.

Gruppierungssystem Die unter **A** und **B** beschriebenen Ansätze sind beide in ein hierarchisches System der Gruppierungshypothesierung eingebettet. Beginnend mit den Kanten des Bildes werden Schritt für Schritt Hypothesen mit wachsender Komplexität generiert. Neben diesem aufwärts gerichteten Prozeß, der Hypothesengenerierung gibt es einen abwärts gerichteten Prozeß, der die Unterstützung der Erzeugung von Hypothesen geringer Komplexität durch Hypothesen höherer Komplexität beschreibt. Für den Ansatz unter **A** unterstützt z.B. eine Parallelitätshypothese die Generierung einer kollinearen Gruppe durch ihre globalere Sicht auf die zugrunde liegende Struktur.

Die generierten Hypothesen sind im Anschluß an den Gruppierungsprozeß als Primitiva

für Stereo-matching Verfahren oder Objektsegmentierung nutzbar. Der unter **A** beschriebene Ansatz ist in ein Verfahren zur Erzeugung von 3D-Modellen für Gebäude integriert. In einem ersten Schritt werden dazu die hypothetisierten Gruppen, welche rechteckige Strukturen beschreiben, in einem constraint satisfaction Netz bewertet. Die so selektierten Gruppen dienen als Primitiva für eine Höhenschätzung mittels Stereo-matching und eine Umrißbeschreibung durch eine Objektsegmentierung.

Suche nach gruppierbaren Primitiva Für beide Ansätze gilt, daß für eine Suche nach gruppierbaren Primitiva Heuristiken über die räumliche Ausdehnung des betrachteten Gruppentyps herangezogen werden.

Für den unter **A** beschriebenen Ansatz sind die Primitiva des Gruppierungsprozesses Geradenstücke. Diese werden zuerst auf die Bildung kollinear Gruppenhypothesen untersucht. Betrachtet werden zu diesem Zweck Ansammlungen paralleler, sich überlappender Geradenstücke. Die Orientierung der Gruppenhypothese wird aus dem längengewichteten Mittel der gruppierten Geradenstücke ermittelt. Die Länge der Hypothese entspricht den Projektionen der Primitiva auf die Gruppe.

Alle hypothetisierten kollinearen Gruppen werden anschließend auf **L**- und **T**-förmige Kreuzungen längs ihrer Achsen untersucht. **T**-förmige Kreuzungen deuten auf eine Okklusion der zugrundeliegenden Strukturen hin. Diese Linien werden am Kreuzungspunkt aufgetrennt, sodaß zwei **L**-förmige Kreuzungen entstehen. Alle Gruppen werden bis zu den gefundenen Kreuzungspunkten verlängert.

Alle kollinearen Hypothesen stellen die Primitiva für die Generierung paralleler Hypothesen dar. Für die Erzeugung paralleler Gruppenhypothesen werden Randbedingungen, wie die Orientierungsdifferenz und die gegenseitige Überlappung der Primitiva, ausgewertet. Parallele Hypothesen unterstützen in einem Umkehrprozeß die Bildung kollinear Hypothesen: Für zwei parallele Primitiva, die sich nicht vollständig überlappen, werden mögliche, vorher nicht berücksichtigte Erweiterungen in eine erneute Betrachtung eingezogen.

Alle parallelen Hypothesen stehen als Primitiva für die Untersuchung **U**-förmiger Strukturen zur Verfügung. Liegen die Endpunkte der Primitiva einer parallelen Hypothese an mindestens einer Seite in einer Ebene rechtwinklig zu den Primitiva, bildet dies den Ausgangspunkt für eine **U**-förmige Gruppenhypothese. In diesem Fall werden Geradenstücke, die die Endpunkte der Parallelen verbinden, für die Hypothesisierung der Gruppe herangezogen. Analog zur Generierung paralleler Hypothesen können im Rahmen einer **U**-Struktur kollineare Gruppenhypothesen neu erzeugt werden, die vorher, aufgrund der Randbedingungen für diese Gruppen, nicht berücksichtigt wurden.

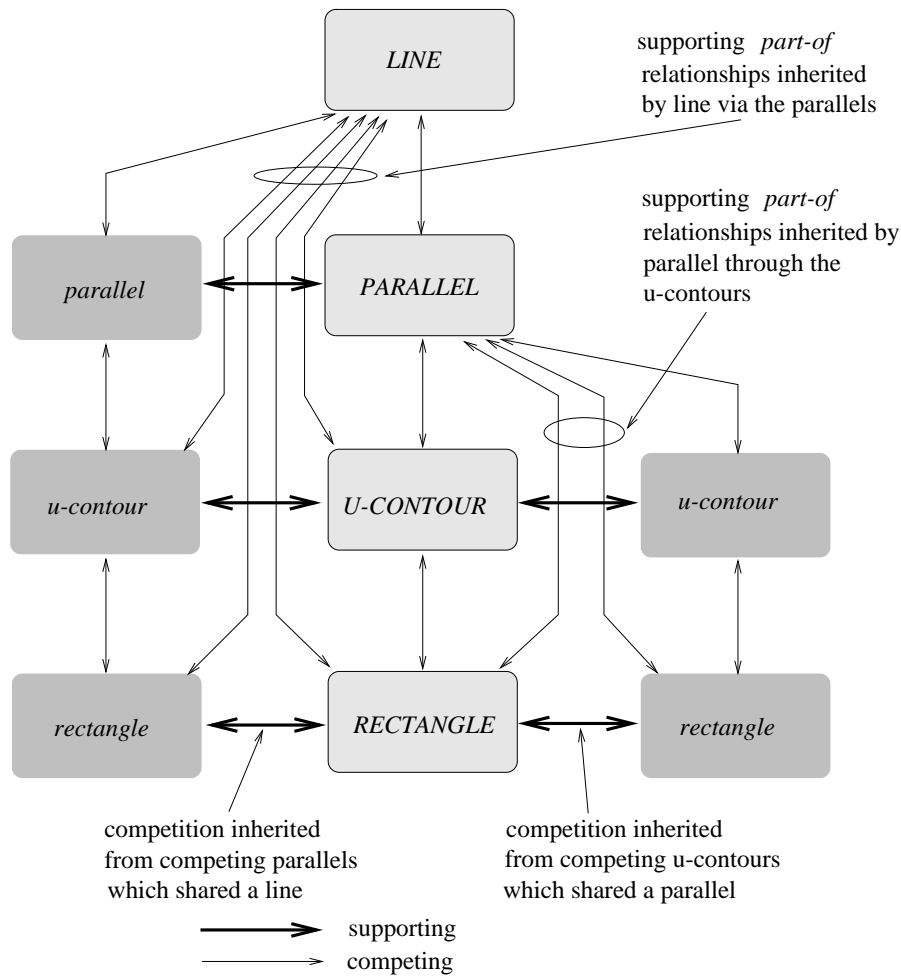


Abbildung 3.3: Abbildung der Gruppenhierarchie in die Knoten eines Netzes mit unterstützenden und konkurrierenden Relationen nach Mohan [17]

Die **U**-förmigen Gruppenthypothesen bilden wiederum die Primitiva für eine Untersuchung rechteckförmiger Strukturen: Rechteckige Hypothesen beruhen auf jeweils zwei **U**-förmigen Hypothesen derselben zugrundeliegenden parallelen Gruppe.

Das hier für den Ansatz **A** geschilderte Vorgehen ist im Kern auch auf den Ansatz **B** übertragbar. Die Komplexität der erzeugten Hypothesen ist dort jedoch höher, da anstatt von Parallelitäten, Symmetrien und anstatt von Rechtecken, allgemeine geschlossene Strukturen (**ribbon**) generiert werden.

Die so erzeugten Gruppenthypothesen können mehrdeutig sein: unterschiedliche Gruppenthypothesen verfügen zum Teil über die gleichen Primitiva. Aus diesem Grund findet eine Bewertung der Hypothesen in einem Netz statt, wie im folgenden Abschnitt beschrieben.

Bewertung der Gruppen Eine sich an den Gruppierungsprozeß anschließende Bewertung, soll die signifikanten Hypothesen herausfiltern. Rakesh Mohan benutzt zu diesem Zweck ein **constraint satisfaction** Netz.

Jede Gruppenhypothese wird auf einen Knoten des Netzes abgebildet, wie in Abbildung 3.3 veranschaulicht. Unter den Gruppenhypothesen existieren zwei verschiedene Relationen, die mit Bestandteil (**part-of**) und Konkurrenz (**competition**) charakterisiert sind. Diese Relationen werden als Verbindungen der Knoten des Netzes repräsentiert.

Bestandteilsrelationen entstehen zwischen Gruppenhypothesen unterschiedlicher Komplexität, welche sich die gleichen Primitiva teilen, wie z.B. eine **U**-Kontur, die auf einer Parallelität aufbaut. Bestandteilsrelationen werden durch positiv gewichtete Unterstützungen (**supporting link**) repräsentiert.

Hypothesen einer Ebene der Abbildung 3.3 verbindet eine Konkurrenz, wenn sie alternative Hypothesen einer zugrunde liegenden Struktur darstellen. Eine Konkurrenzrelation wird durch eine negativ gewichtete Verbindung (**conflicting link**) repräsentiert. Eine Konkurrenzrelation kann zwischen den verschiedenen Ebenen in Abbildung 3.3 vererbt werden: für zwei konkurrierende **U**-Konturen der Abbildung werden jeweils Rechtecke hypothetisiert. Diese erben die Konkurrenz der **U**-Konturen.

Jedem Knoten wird ein initialer Wert zugeordnet, der die zugrundeliegenden Daten beschreibt, z.B. die Überlappung und der Abstand der Primitiva einer Parallelität.

3.2.3 Perzeptive Organization und digitale Bildverarbeitung

Sakar & Boyer kommt der Verdienst zu, sich mit einer sehr gründlichen und fundierten Ausarbeitung dem Thema "perzeptives Gruppieren" in der digitalen Bildverarbeitung gewidmet zu haben.

Domäne, Vorverarbeitung, Primitiva: Sakar & Boyer benutzen grauwertbasierte Eingangsbilder mit recht unterschiedlichen Domänen. Die gezeigten Resultate erstrecken sich auf folgende Bereiche:

- Szene mit Bauklötzen
- Laborszene
- Innenraumszene mit viereckigen Strukturen
- Satellitenbild einer Rollbahn

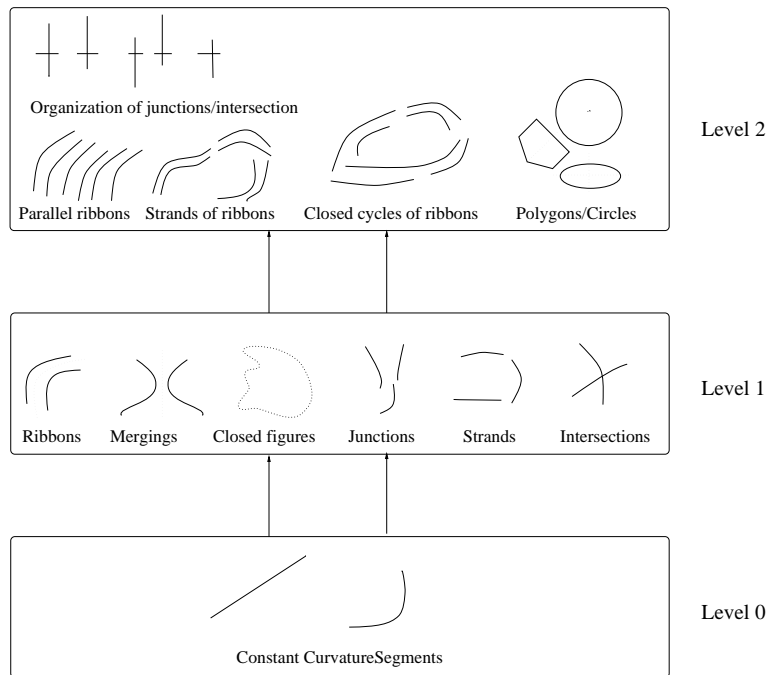


Abbildung 3.4: Einteilung der detektierten Strukturen in Hierarchieebenen, entsprechend ihrer wachsenden Komplexität nach Sakar & Boyer [55]

- Luftbild einer Grünanlage
- Stereoluftbildpaar einer Großstadt

Das Verfahren arbeitet kanten-, bzw. konturbasiert. Die Kanten-Detektion erfolgt mit Hilfe des “optimal zero crossing operator”. Die Kanten werden anschließend in Kontursegmente konstanter Krümmung approximiert. Diese Kontursegmente sind die initialen Primitiva des Gruppierungsprozesses

Gruppierungssystem: Sakar & Boyer betten den Verlauf des Gruppierungsprozesses in einen hierarchischen Ansatz ein. Diese Hierarchie orientiert sich an der wachsenden Komplexität der zu detektierenden Strukturen der Szene. Abbildung 3.4 zeigt beispielhaft die Strukturen dreier aufeinander folgender Stufen oder “Level”.

Innerhalb dieser Hierarchie sind “bottom up” und “top down” Komponenten vorgesehen: Bottom up dienen Hypothesen einer Stufe als Gruppierungsprimitiva der nachfolgenden Stufe. Top down wird aufgrund zusätzlicher Information über die Szene die Anwesenheit bestimmter Strukturen überprüft. In jeder Stufe dieser Hierarchie ist die gleiche Abfolge von Verfahren vorgesehen:

- Untersuchung der gruppierbaren Primitiva (“voting scheme”)
- Konstruktion der Gestaltgraphen
- Hypothesisierung der Gruppen durch Verknüpfung der Gestaltgraphen
- Bewertung der Hypothesen, bzw. Generierung neuer Hypothesen (“Perceptual Inference Network”)

In Analogie zu Begriffen aus der menschlichen Wahrnehmung unterteilen Sakar & Boyer die aufgelisteten Schritte in eine “prä-attentive” und eine “attentive” Phase. Prä-attentiv sind alle Verfahren die eine Strukturbildung - von der Untersuchung der gruppierbaren Primitiva bis zur Hypothesisierung der Gruppen - vornehmen. Attentiv ist die Bewertung der Gruppen und die Hypothesisierung neuer Gruppen aufgrund der verfügbaren Information im Perceptual Inference Network.

Suche nach gruppierbaren Primitiva Sakar & Boyer definieren einen mehrdimensionalen Parameterraum, der beispielsweise durch die räumlichen Koordinaten x und y , die Steigung θ und die Krümmung κ der Primitiva aufgespannt wird. Die Auswertung der örtlichen Nähe der Primitiva in diesem Raum erlaubt eine Aussage über ihre Gruppierbarkeit. Die Abbildung der Primitiva in diesen Raum bezeichnen Sakar & Boyer mit dem Begriff “votieren”: Ein Primitivum votiert entsprechend seiner Parameterwerte für einen Punkt des Raumes. Die örtliche Nähe der Punkte in diesem Raum erlaubt eine Aussage über die Gruppierbarkeit der zugrunde liegenden Primitiva. Diese Untersuchung der örtlichen Nähe kann durch eine Aufteilung des Raumes in einzelne Zellen vereinfacht werden: Alle Punkte in einer Zelle sind Nachbarn und werden gruppiert. Die Einteilung des Raumes in einzelne Zellen kann durch eine Quantisierung der Koordinaten der einzelnen Achsen erfolgen. Diese Untersuchung der Zellen des Parameterraumes wird beeinträchtigt durch:

- Rauschen, welches die Abbildung der Primitiva in die Zellen verändert
- Eine mehrdeutige Abbildung eines Primitivum auf eine Zellengrenze
- Punkte benachbarter Zellen können einen geringeren Abstand besitzen, als Punkte innerhalb Zelle

Um alle Punkte zu finden, die zusammengehörig sind, wird für das Quantisierungsraaster ein Offset definiert, welcher den Gitterabstand gleichmäßig teilt. Entsprechend dem gewählten Offset wird der Parameterraum sequentiell auf Gruppen untersucht. Die in den

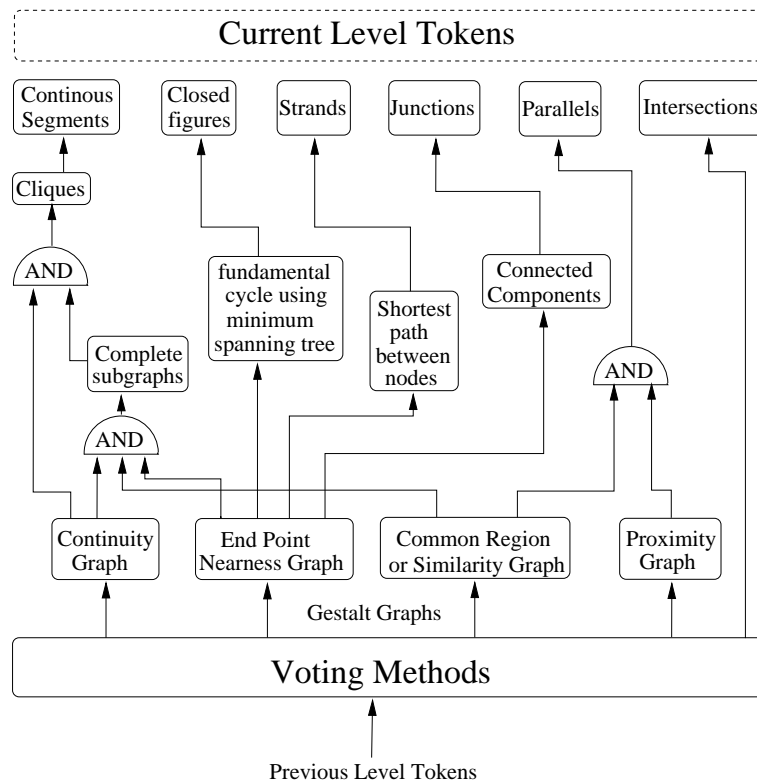


Abbildung 3.5: Organisation der Gestaltgraphen auf der untersten Hierarchieebene nach Sakar & Boyer [55]

Zellen des Parameterraumes gruppierten Primitiva werden im Sinne ihrer Zugehörigkeit zu verschiedenen Gruppentypen, wie sie in Abbildung 3.4 gezeigt sind, näher untersucht. Sakar & Boyer benutzen Graphen für diese Zuordnung, die sie Gestaltgraphen nennen. Fünf verschiedene Graphen werden konstruiert. Mit Ausnahme des Endpoint Nearness Graphen repräsentieren dabei die Knoten des Graphen die untersuchten Primitiva:

- **Proximity Graph:** Eine Kante verbindet zwei Knoten, wenn die Endpunkte der Primitiva in unmittelbarer Nachbarschaft liegen und die Orientierung der Primitiva gleich ist.
- **Endpoint Nearness Graph:** Die Knoten des Graphen repräsentieren die Endpunkte der Primitiva. Eine Kante verbindet zwei Knoten, wenn sie in unmittelbarer Nachbarschaft liegen, oder demselben Primitivum angehören. Die Kanten werden durch eine unterschiedliche Gewichtung in Verbindungen zwischen unterschiedlichen Primitiva und dem Primitivum selbst gekennzeichnet.
- **Continuity Graph:** Eine Kante verbindet zwei Knoten, wenn die Primitiva sich

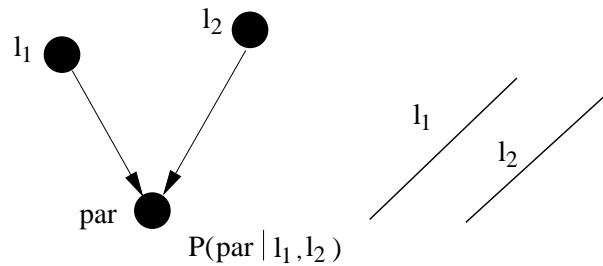


Abbildung 3.6: Die Struktur eines Bayes Netzes für zwei Geradenstücke l_1, l_2 die eine Parallelität bilden, nach Sakar & Boyer [55]

gegenseitig fortsetzen.

- **Similarity Graph:** Eine Kante verbindet zwei Knoten, wenn die Primitiva über gleichartige photometrische oder geometrische Eigenschaften verfügen.
- **Common Region Graph:** Eine Kante verbindet zwei Knoten, wenn die Primitiva in einer Region liegen

Eine geeignete Verknüpfung der Gestaltgraphen, wie in Abbildung 3.5 gezeigt, mündet in der Hypothesisierung der verschiedenen Gruppen.

Bewertung der Gruppen Die Bewertung der hypothesierten Gruppen wird in einem modifizierten Bayes Netz vorgenommen: dem Perceptual Inference Network (PIN). Die Knoten des Netzes repräsentieren die möglichen Hypothesen wie: Ellipse, Strand, Corner, etc. Verbunden werden sie durch gerichtete Kanten. Diese Kanten modellieren die Relationen der Primitiva, wie in Abbildung 3.6 beispielhaft für die Formulierung einer Parallelität demonstriert.

Der Nachrichtenaustausch zwischen den Knoten des Bayes Netzes erfolgt unabhängig von der Kantenrichtung in beiden Richtungen. Dies unterstützt die in diesem Abschnitt unter **Gruppierungssystem** eingeführten “bottom up” und “top down” Komponenten der Hierarchie. Sakar & Boyer strukturieren das Bayes Netz mit “Composite Nodes”. Composite Nodes fassen gleiche Gruppen-Hypothesen, wie z.B. Parallelität, unabhängig vom Ort ihres Auftretens zusammen. Diese Zusammenfassung der Hypothesen in Composite Nodes reduziert die Anzahl der erzeugten Knoten auf die Gruppenthypothesen-Typen des Systems, steigert jedoch gleichzeitig die Komplexität der einzelnen Knoten. Die Anordnung der Knoten erfolgt in einer Hierarchie, die die Komplexität der Hypothesen widerspiegelt. Ein Beispiel für eine solche Anordnung ist in Abbildung 3.7 gezeigt. In der Abbildung ist

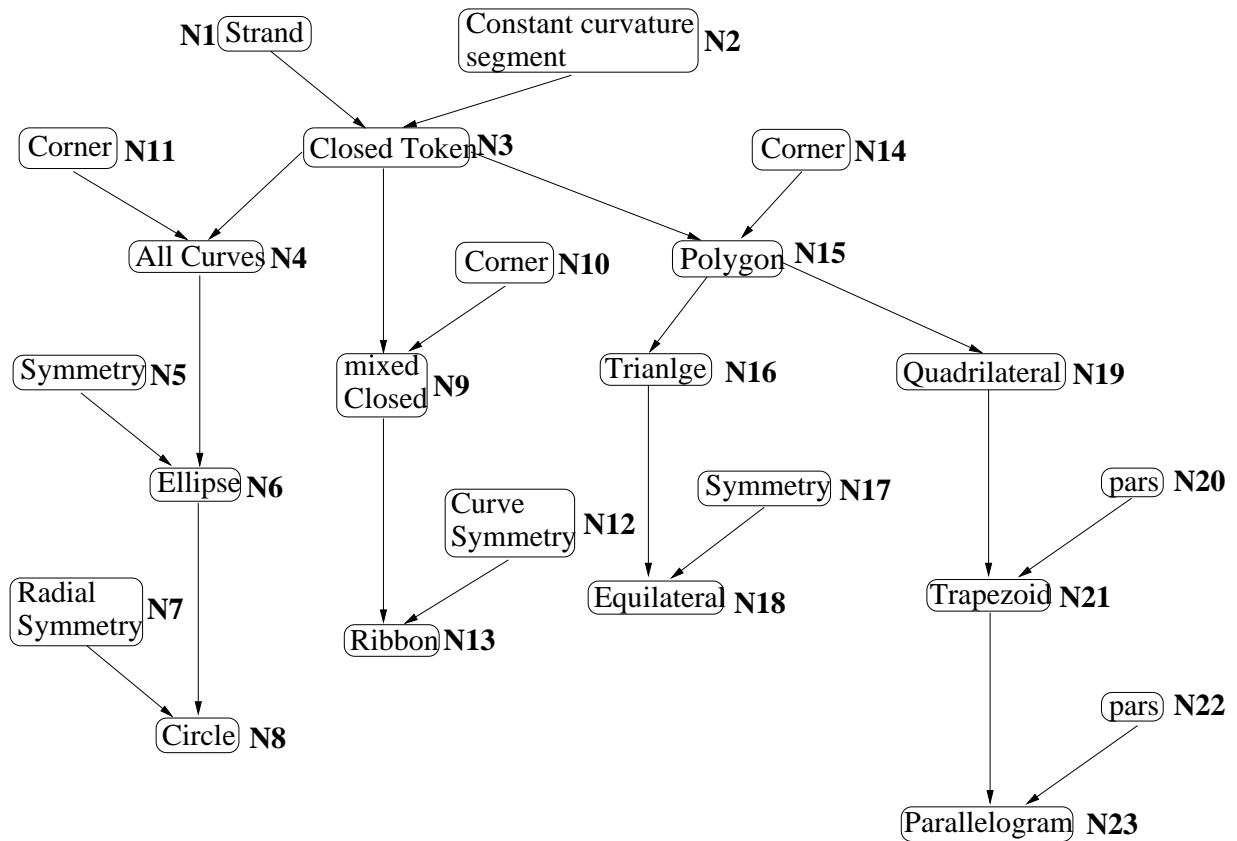


Abbildung 3.7: Struktur der Composite Nodes nach Sakar & Boyer [55]

erkennbar, daß bestimmte Knoten verdoppelt sind: Dies dient dazu, die in einem Bayes Netz nicht erlaubten Zyklen zu verhindern. Für die Instantiierung des Netzes erhalten zum einen die Composite Nodes eine apriori Wahrscheinlichkeit. Diese Wahrscheinlichkeit spiegelt die Häufigkeit des Vorkommens der durch den Knoten repräsentierten Hypothese in der Szene wider. Ist dieses apriori Wissen nicht verfügbar, werden alle Composite Nodes mit der gleichen Wahrscheinlichkeit instantiiert.

Zum anderen wird für die beobachtete Hypothese am Ort x, y eine Evidenz berechnet, die spezifische Eigenschaften der Gruppe beschreibt, wie z.B. der Approximationsfehler für Kontursegmente. Die Repräsentation dieses apriori Wissens erfolgt durch die Modellierung eines Scheinknotens Z , der dem Netzknoten X über seine Beobachtung mit $P(Z = z | x)$ berichtet.

Weist die generierte Szenenbeschreibung Mehrdeutigkeiten auf, kommen “special purpose modules” zum Einsatz. Sakar & Boyer beschreiben Mehrdeutigkeit mit der Erzeugung mehr als einer Hypothese für eine Anzahl von Primitiva. Mehrdeutige Hypothesen werden in einen “locational conflict graph” abgebildet. Die Knoten des Graphen entsprechen

den Hypothesen und die Kanten den existierenden “Konflikten”. Die im Konfliktfall eingesetzten special purpose modules untersuchen das Grauwertbild genauer auf bestimmte Merkmale. Diese Merkmale sind beispielsweise “Corner”, “Contour” oder “Region”. In Abbildung 3.7 sind dies z.B. die Composite Nodes **N10**, **N11** und **N12** für das Modul “Corner”, der Composite Node **N2** für das Modul “Contour” und der Composite Node **N3** für das Modul “Region”. Die so gewonnene, zusätzliche Information wird, bei gleichzeitiger Auswertung der Cliques des locational conflict graph, herangezogen, die Konflikte zu lösen. Eine eindeutige Beschreibung der Szene wird so ermöglicht.

Zusammenfassung Dieses Kapitel beschäftigte sich mit der Umsetzung gestaltpsychologischer Prinzipien in der digitalen Bildverarbeitung. Da der in der Bildverarbeitung verwendete Oberbegriff “perzeptives Gruppieren” oder “perzeptive Organisation” für diese Verfahren in ihrer Komplexität stark unterschiedliche Methoden umfaßt, wurde in Abschnitt 3.1 der von Sakar & Boyer vorgenommene Versuch zur Klassifikation solcher Verfahren vorgestellt. Daran anschließend wurden beispielhaft drei Verfahren zur Perzeptiven Organisation in den Abschnitten 3.2.1, 3.2.2 und 3.2.3 vorgestellt, von denen meine eigene Arbeit maßgeblich beeinflusst wurde. Die nachfolgenden drei Kapitel 4, 5 und 6 werden den von mir verfolgten Ansatz vorstellen. Dabei wird Kapitel 4 “Generierung von Gruppenthypothesen mit Bereichen perzeptiver Aufmerksamkeit” die Domäne, die Vorverarbeitung, die Primitiva und die Suche in der Szene nach geeigneten Primitiva beschreiben. Kapitel 5 “Markov Random Fields” dient der Einführung wesentlicher Grundlagen für Kapitel 6 “Ein Markov Random Field für Gruppenthypothesen”, welches dann die Formulierung der Gruppen in einem System und die Bewertung, Auswertung und Auswahl der Gruppen zum Thema hat.

4 Generierung von Gruppenhypothesen mit Bereichen perzeptiver Aufmerksamkeit

Dieses Kapitel beschreibt einen neuartigen Ansatz zur Generierung von Gruppenhypothesen mit Hilfe der Mechanismen perzeptiver Gruppierung. Die Gliederung dieses Kapitels lehnt sich an Abschnitt 3.2 des vorigen Kapitels 3: “Gestaltpsychologie in der Bildverarbeitung” an.

Der Verlauf des hier beschriebenen Ansatzes gliedert sich in die folgenden Teilbereiche:

- Die **Vorverarbeitung** beinhaltet Verfahren wie Kantendetektion und Konturapproximation. Das Ziel der Vorverarbeitung ist die Bereitstellung geeigneter Primitiva für den Gruppierungsprozeß.
- Die **Suche** geeigneter Gruppierungsprimitiva für die **Generierung der Gruppenhypothesen** schließt sich an. Diese Suche ist an die Erfordernisse der unterschiedlichen Gruppenhypothesen angepaßt.
- Die **Bewertung der Hypothesen** bildet den Abschluß. Das Ziel der Bewertung ist die Auflösung inkonsistenter Information, welche aus der Generierung mehrdeutiger Gruppenhypothesen entsteht, bei einer gleichzeitigen Reduktion der Hypothesenanzahl durch einen geeigneten Bewertungsmechanismus.

Die Beschreibung der Bewertung der Gruppenhypothesen ist sehr umfangreich, deshalb ist ihr ein eigenes Kapitel 6: “Ein Markov Random Field für Gruppenhypothesen” gewidmet. In Kapitel 5: “Markov Random Fields” werden die notwendigen Grundlagen für das Verständnis von Kapitel 6 eingeführt.

Der Schwerpunkt dieses Kapitels liegt auf der Vorstellung eines neuartigen Ansatzes zur Suche nach geeigneten Gruppierungsprimitiva, angepaßt an die unterschiedlichen Grup-

pentypen. Die Motivation für diesen Ansatz liegt in dem Bestreben, diese Suche durch eine Einschränkung auf einen Bereich in der Umgebung des betrachteten Primitivum zu vereinfachen: den **Bereich der perzeptiven Aufmerksamkeit** für dieses Primitivum. Neu an diesem Ansatz ist das Bestreben, diese Bereiche nicht aus willkürlichen Annahmen zu formen, sondern aus den Charakteristika der betrachteten Gruppentypen entstehen zu lassen. Zu diesem Zweck wurde eine repräsentativen Stichprobe der untersuchten Domäne unter Berücksichtigung der unterschiedlichen Gruppen ausgewertet. Aus dieser Auswertung entstanden Häufungsbilder. Die in diesen Häufungsbildern gespeicherte Information über die **gruppierende Wirkung** eines Primitivum auf seine Umgebung wird in geeigneter Art und Weise für den automatisierten Gruppierungsprozeß aufbereitet.

4.1 Domäne, Vorverarbeitung, Primitiva

Der hier vorgestellte Ansatz arbeitet mit Grauwertbildern. Die Bildinhalte entstammen einer einheitlichen Domäne: einem Holz-Konstruktionsspielzeug namens **Baufix**. Deshalb wird die Domäne auch verkürzt mit **Baufix-Szenario** bezeichnet. Wichtige Konstruktionselemente dieses Spielzeugs sind z.B. flache Holzleisten mit Schraubenlöchern für Längsverbindungen oder Holzwürfel mit Schraubenlöchern für Eckverbindungen ("Schraubwürfel"), sowie Holzschrauben und rautenförmige Muttern. Aus den Einzelteilen des Konstruktionsspielzeugs kann in der untersuchten Domäne ein Flugzeug konstruiert werden. Die Bildinhalte zeigen:

- Konstruktionsteile in Einzellage
- Konstruktionsteile, die sich überlappen
- teilmontierte Konstruktionsteile
- Flugzeug

Der von mir verfolgte Ansatz arbeitet kanten- bzw. konturbasiert. Die **Kanten-Detektion** erfolgt mit Hilfe eines Sobel-Operators. Ein sich anschließender Verdünnungs-Algorithmus auf Basis der **non-maxima suppression** und **Hystereseschwellwert** reduziert die detektierten Kanten auf Pixelbreite. Aus diesen verdünnten Kanten werden mit Hilfe eines von Leonardis [13] entwickelten Verfahrens Geradenstücke und Ellipsensegmente approximiert. Die Ellipsensegmente können alternativ durch Kreis- bzw. Ellipsengleichungen

beschrieben werden. In einem sich anschließenden Schritt findet die Beseitigung offensichtlicher Fragmentierungen statt. **Offensichtliche Fragmentierungen** werden für aufeinander folgende Geradenstücke angenommen, zwischen denen nur kleine Lücken bestehen. In der näheren Umgebung der Geradenstücke gibt es jeweils nur eine weitere Anschlußmöglichkeit, die eine passende Orientierung besitzt. In diesem Fall werden die Fragmente durch eine erneute Approximation in ein Geradenstück angenähert.

Die solchermaßen approximierten Geradenstücke und Ellipsensegmente stellen die **initialen Primitiva** des Gruppierungsprozesses dar.

4.2 Gruppierungssystem: Eine Hierarchie für Gruppenhypothesen

Der Ablauf des beschriebenen Gruppierungsprozesses ist in eine Hierarchie von Gruppenhypothesen eingebunden, wie sie in Abbildung 4.1 gezeigt ist. Diese Hierarchie der Gruppenhypothesen orientiert sich an der wachsenden Komplexität der gebildeten Strukturen: eine kollineare Gruppe ist z.B. die Ausgangsbasis für die Hypothesisierung einer parallelen Gruppe und eine parallele Gruppe ist z.B. die Ausgangsbasis für die Hypothesisierung einer Geschlossenheit. Die Ebenen werden - in Richtung wachsender Komplexität - mit **1D**, **2x1D** und **2D** benannt. Diese Benennung ist aus der räumlichen Struktur der Gruppenhypothesen der betreffenden Ebene abgeleitet. Der Gruppierungsprozeß startet mit der Untersuchung der Kontursegmente, den initialen Primitiva des Gruppierungsprozesses. Aus diesen werden die Gruppen der 1D-Ebene erzeugt, darauf folgen die Gruppen der 2x1D-Ebene und abschließend die Gruppen der 2D-Ebene. In jeder Hierarchieebene wird nach passenden Erweiterungen der betrachteten Primitiva dieser Ebene gesucht.

Jede Ebene stützt sich in der Betrachtung der Gruppenhypothesen auf alle darunter liegenden. Dies bedeutet, daß mit Ausnahme von Geschlossenheiten, welche die höchste Hierarchiestufe darstellen, alle hypothetisierten Gruppen wiederum als Primitiva dem Gruppierungsprozeß zur Verfügung stehen. Kollineare und kurvilineare Gruppenhypothesen stehen darüberhinaus auch in der eigenen Ebene als Primitiva zur Verfügung. Kontursegmente können in allen Ebenen für eine Generierung der Hypothesen benutzt werden.

Primitiva und Gruppen der 1D-Ebene: Die Gruppen der 1D-Ebene sind **eindimensional** in dem Sinn, daß ein vorhandenes, eindimensionales Primitivum, wie z.B. ein Geradenstück, durch ein zweites Primitivum verlängert wird. Die Gruppen dieser Ebene sind

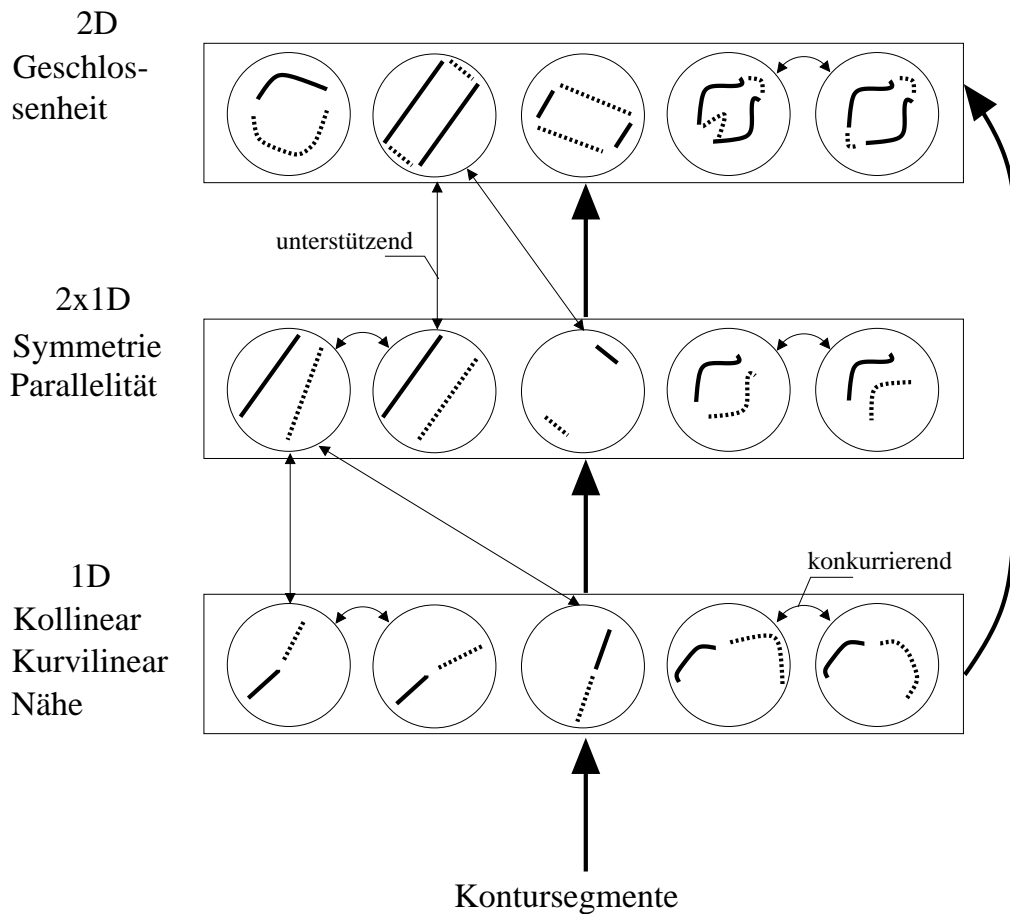


Abbildung 4.1: Hierarchischer Aufbau der Gruppenhypothesen:

Innerhalb der Kreise sind die verschiedenen Gruppen symbolisiert. Die Rechtecke grenzen die mit 1D, 2x1D und 2D benannten Hierarchieebenen ab. Die Generierung der Gruppenhypothesen erfolgt in der Reihenfolge der Ebenen, beginnend mit der 1D-Ebene. Hypothesierte Gruppen einer Ebene und Kontursegmente sind die Primitiva der folgenden Ebenen, angedeutet durch die dicken schwarzen Pfeile.

Die Gruppenhypothesen sind durch die Relationen Unterstützung und Konkurrenz verbunden. Diese Relationen sind beispielhaft durch die dünnen Pfeile angedeutet.

Kollinearität, Kurvilinearität und Nähe.

Eine **kollineare Gruppe** besteht aus mindestens zwei geradenförmigen Primitiva. Dabei kann es sich um die initialen Primitiva, als auch um bereits hypothesierte kollineare Gruppen handeln.

Eine **kurvilineare Gruppe** besteht aus mindestens zwei ellipsenförmigen Primitiva. Da-

bei kann es sich um die initialen Primitiva, als auch um bereits hypothetisierte kurvilineare Gruppen handeln.

Eine **Nähe-Gruppe** besteht aus genau zwei Primitiva. Dabei kann es sich um die initialen Primitiva, sowie hypothetisierte kollineare und kurvilineare Gruppen handeln.

Primitiva und Gruppen der 2x1D-Ebene: Die Gruppen der 2x1D-Ebene entstehen aus der Verknüpfung initialer Primitiva, kollinearere oder kurvilinearere Gruppen. Die Verknüpfung erfolgt derart, daß eine räumliche Komponente entsteht, im Gegensatz zu einer einfachen Verlängerung in der 1D-Ebene. Die Gruppen dieser Ebene sind Parallelität und Symmetrie, wobei die Parallelität als ein Sonderfall der Symmetrie angesehen wird.

Eine **parallele Gruppe** besteht aus genau zwei geradenförmigen Primitiva. Symmetrische Gruppen wurden im Rahmen dieser Arbeit nicht erzeugt.

Primitiva und Gruppen der 2D-Ebene: Gruppen dieser Ebene beschreiben geschlossene Konturen und besitzen damit eine räumliche Ausdehnung. In der 2D-Ebene werden Geschlossenheiten hypothetisiert. Eine **Geschlossenheit** kann initiale Primitiva, alle Gruppen der 1D-Ebene, sowie Parallelitäten der 2x1D-Ebene umfassen.

Unterstützung und Konkurrenz: Zwischen den vorgestellten Gruppenhypothesen existieren zwei unterschiedliche Relationen: Unterstützung und Konkurrenz. **Unterstützung** entsteht, wenn zwei Gruppenhypothesen, eine Bestandteilsrelation bilden. Eine **Bestandteilsrelation** besteht z.B. zwischen einer kollinearen Gruppenhypothese und ihren Bestandteilen, den geradenförmigen Primitiva. Dabei **unterstützen** die geradenförmigen Primitiva die Bildung der kollinearen Hypothese und umgekehrt **unterstützt** die kollineare Hypothese die Primitiva aus denen sie gebildet ist. Der Gruppe und ihren Bestandteilen wird so eine besondere Bedeutung zugebilligt. **Konkurrenz** entsteht zwischen Gruppenhypothesen einer Hierarchieebene, welche die gleichen Primitiva besitzen. Diese Gruppenhypothesen nehmen eine **mehrdeutige Interpretation** der Szene vor. Die Bildung der Konkurrenzrelation zwischen diesen Gruppen bringt diese Mehrdeutigkeit zum Ausdruck. Diese Relationen sind in Abbildung 4.1 beispielhaft gezeigt.

4.3 Suche nach gruppierbaren Primitiva

Die Suche nach geeigneten Primitiva für die Generierung der Gruppenhypothesen ist eine wichtige Komponente des vorgestellten Verfahrens. Diese Suche wird auf einen Bereich in der Umgebung eines Referenzpunktes des betrachteten Primitivums eingeschränkt. Ein **Referenzpunkt** kann z.B. der Endpunkt eines Primitivums sein. Für diese Suche wird ein neuartiger Ansatz verwendet. Die Besonderheit dieses Ansatzes liegt in der Art der für die Suche eingesetzten Bereiche: Sie sind nicht das Ergebnis willkürlicher Annahmen, sondern entstehen aus den Charakteristika der verschiedenen Gruppentypen. Die Ermittlung **spezifischer Bereiche** für einen Einsatz im automatisierten Gruppierungsprozeß ist das Thema des Abschnitts 4.3.2. Eine Beschreibung des Einsatzes dieser Bereiche und der Generierung der Gruppenhypothesen schließt sich in 4.4.1 an.

4.3.1 Motivation

Jeder Gruppierungsvorgang in der maschinellen Bildanalyse beginnt mit einer Suche nach passenden Erweiterungen der betrachteten Primitiva. Diese Suche ist ein sehr zentrales Moment des automatisierten Gruppierungsverfahrens: Der Aufwand für die Suche nach Erweiterungen nimmt in der Regel mit der Anzahl der betrachteten Primitiva erheblich zu. Eine örtliche Eingrenzung der Suche ist deshalb von grundsätzlichem Interesse. Gleichzeitig motiviert sich eine örtliche Beschränkung aus dem Umstand, daß die untersuchten Strukturen durch örtlich benachbarte Primitiva beschrieben werden. Es ist also sinnvoll, die Suche nach geeigneten Erweiterungen auf einen Bereich in der näheren Umgebung des betrachteten Primitivums einzuschränken.

Die Dimensionierung eines solchen Bereiches ist nicht trivial, wie das Beispiel in Abbildung 4.2 verdeutlichen soll. Die Abbildung zeigt eine Ansammlung verschiedener Geradenstücke: Jeweils zwei Segmente sind in exakt gleicher Richtung angeordnet und bilden ein Paar kollinearere Segmente. Die Länge der Lücken zwischen allen Segmentpaaren ist genau gleich. Der Einfluß dieser Lücke auf die **gruppierende** Wirkung eines Segmentpaares schwankt für einen Betrachter recht offensichtlich in Abhängigkeit zur Länge der Segmente. Aus diesem Grund kann für die kollinearen Segmentpaare in Abbildung 4.2 folgende Forderung abgeleitet werden:

- Die Länge der Segmente soll Einfluß auf die zwischen einem Segmentpaar akzeptierte Lückenlänge haben.

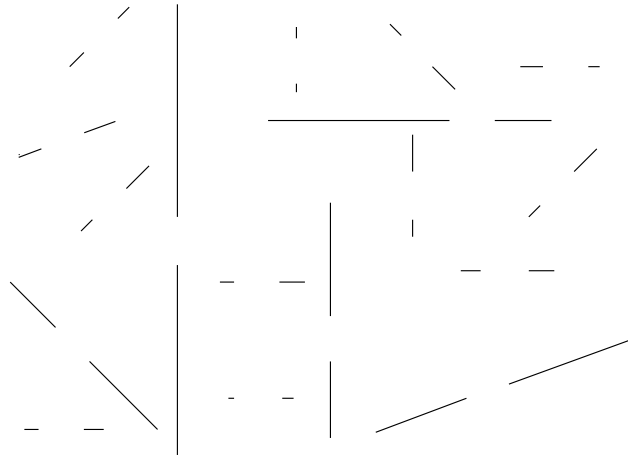


Abbildung 4.2: Die in der Abbildung gezeigten Linien bilden jeweils paarweise zueinander eine Fortsetzung. Die Lücke die jedes Linienpaar trennt, ist für alle Paare gleich groß. Abhängig von der Länge der Linien, werden die Linien eines Paares als mehr oder weniger stark verbunden wahrgenommen.

Diese Bereiche, in denen Primitiva in ihrer Nachbarschaft eine gruppierende Wirkung entfalten, nenne ich im folgenden **Bereiche perzeptiver Aufmerksamkeit**. Die Namensgebung soll die aktive Rolle eines betrachteten Primitivums in seiner Nachbarschaft betonen. Dem großem Interesse an der Einschränkung der Suche auf ausgewählte Bereiche in der computergestützten Anwendung perzeptiver Organisation stehen leider keine entsprechenden Experimente auf wahrnehmungspsychologischer Seite gegenüber, welche eine algorithmische Beschreibung erlauben würden. Dieser Ansatz stützt sich deshalb auf die statistische Auswertung einer Stichprobe, welche die wesentlichen Inhalte der verwendeten Domäne widerspiegelt. Dies wird im folgenden Abschnitt ausführlich vorgestellt.

4.3.2 Ermittlung von Bereichen perzeptiver Aufmerksamkeit

Die Suche nach gruppierbaren Primitiva wird auf Bereiche in der Umgebung ausgezeichneter Punkte, den **Referenzpunkten**, der untersuchten Primitiva und Gruppen eingeschränkt. Diese **Bereiche perzeptiver Aufmerksamkeit**, tragen den Charakteristika der verschiedenen Gruppentypen Rechnung. Für jeden Gruppentyp wurde deshalb ein spezifischer Bereich ermittelt. Die Untersuchung der Charakteristik des jeweiligen Gruppentyps erfolgte durch Auswertung von Datenmaterial der verwendeten Domäne (Baufix-Szenario). Die Vorgehensweise für die Gewinnung der Bereiche perzeptiver Aufmerksamkeit erfolgte nach dem folgenden Schema:

- Erstellung einer Stichprobe von Grauwertbildern für die untersuchte Domäne und die Durchführung der Vorverarbeitung
- Generierung von Gruppenhypothesen für die Primitiva dieser Stichprobe durch ein Labeling von Hand
- Ermittlung von Daten aus den so bestimmten Gruppen der Stichprobe, die geeignet sind, die Bereiche perceptiver Aufmerksamkeit zu beschreiben

Die durchgeführten Untersuchungen zu Bereichen perceptiver Aufmerksamkeit beruhen auf einer statistischen Auswertung einer Stichprobe von 40 Grauwertbildern. Diese Stichprobe spiegelt die wesentlichen Inhalte der verwendeten Domäne wider. Alle für den Bau des Spielzeugflugzeuges notwendigen Einzelteile fanden Berücksichtigung. Eine Anordnung dieser Teile in Einzellagen und Teilmontagen bis hin zum vollständigen Objekt bildet den Konstruktionsprozeß des Flugzeuges nach. Gleichzeitig ermöglichen differierende Brennweiten verschieden detaillierte Ansichten der Szene. Abbildung 4.3 zeigt Beispiele dieser Stichprobe.

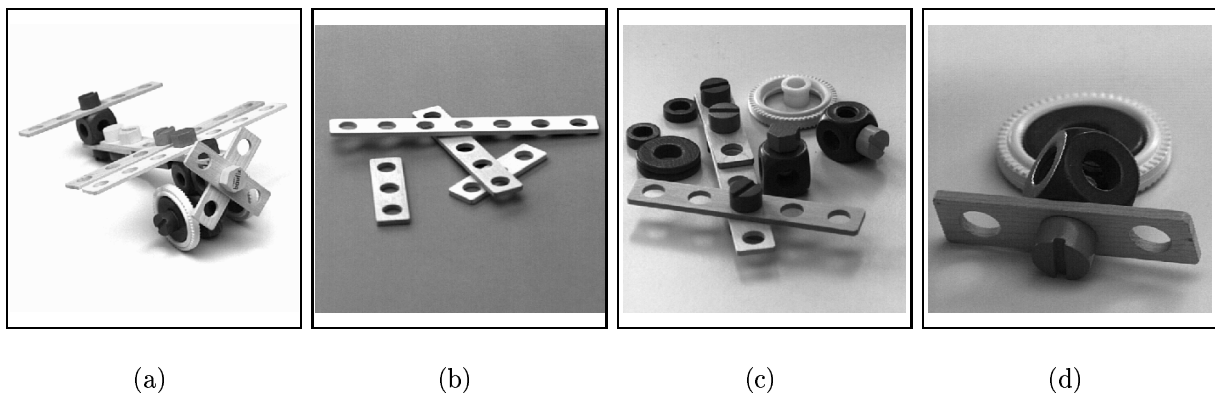


Abbildung 4.3: (a-d)Beispielbilder aus der erstellten Stichprobe

In einem ersten Schritt wurden diese Bilder mit den Verfahren der Kantendetektion bis zur Konturapproximation und der Beseitigung der offensichtlichen Fragmentierungen bearbeitet. Die so entstandenen symbolischen Beschreibungen in Form von Geradenstücken und Ellipsensegmenten bilden die initialen Primitiva des Gruppierungsprozesses. Den Gruppierungsprozeß selbst simulierte ein sich daran anschließendes - von Hand durchgeführtes - Labeling der verschiedenen Gruppentypen. In diesem Zusammenhang ist anzumerken, daß die Erstellung dieser manuellen Gruppenhypothesen pragmatisch in dem Sinne zu verstehen ist, daß genau solche Gruppen erzeugt wurden, die auch später durch das

System erstellt werden sollten. Auf diese Art entstanden repräsentative Stichproben für Kollinearitäten, Kurvilinearitäten, Nähe-Gruppen und Parallelitäten. Die Anzahl der Primitiva einer Gruppenthese und der Typ der erlaubten Primitiva entspricht den in Abschnitt 4.2 beschriebenen Zusammenhängen. Auf ein Labeling von Geschlossenheiten wurde verzichtet, da sich ihre Hypothesierung - wie noch zu sehen sein wird - vollständig auf alle bereits erzeugten Gruppen stützt.

4.3.2.1 Untersuchung der Stichprobe

Die Auswertung der Stichprobe erfolgte in zwei Teilschritten:

- Definition eines geeigneten Kriteriums für die Stichprobendaten, welches die gesuchten Bereiche charakterisiert
- Detektion eines charakteristischen Bereiches aus den Daten des jeweiligen Gruppentyps

Das gewählte Kriterium ist die relative räumliche Lage zweier ausgezeichnete Punkte eines Paares von Primitiva einer Gruppe. Ein Punkt gehört dem betrachteten Primitivum R oder auch **Referenzprimitivum** und wird mit **Referenzpunkt** P_R bezeichnet. Hier kann es sich z.B. um einen Endpunkt des Primitivums handeln. Der zweite Punkt gehört einem **Anschlußprimitivum** A , welches - entsprechend dem untersuchten Gruppentyp - eine gute Fortsetzung zum betrachteten Primitivum darstellt. Dieser Punkt wird mit **Anschlußpunkt** P_A bezeichnet. Der Anschlußpunkt kann z.B. der Endpunkt des zweiten Primitivum sein. Wie in Abbildung 4.4a) gezeigt, sind in diesem Fall beide Punkte so angeordnet, daß sie die Lücke zwischen den Primitiva begrenzen.

Damit alle aus der Stichprobe eines Gruppentyps extrahierten Merkmale miteinander vergleichbar sind, muß eine geeignete Normierung vorgenommen werden. Diese Normierung besteht aus zwei Komponenten: einer Längennormierung und einer Lagenormierung der Primitiva.

Die **Längennormierung** ist in Abbildung 4.4b) beispielhaft gezeigt: Die Länge L eines betrachteten geradenförmigen Primitivum R wird auf eine Normlänge L_N abgebildet: $\frac{L_N}{L}$. Unter der Annahme, daß die Normlänge $L_N = 1$ ist, kann die Lückenlänge D entsprechend dem Referenzprimitivum R normiert werden: $D_N = \frac{1}{L}D$. Durch diese Definition der Längennormierung ist die Untersuchung eines Paares von Primitiva nicht symmetrisch. Daher muß jedes Primitivum eines Paares für die Auswertung einmal als Referenzprimitivum betrachtet werden.

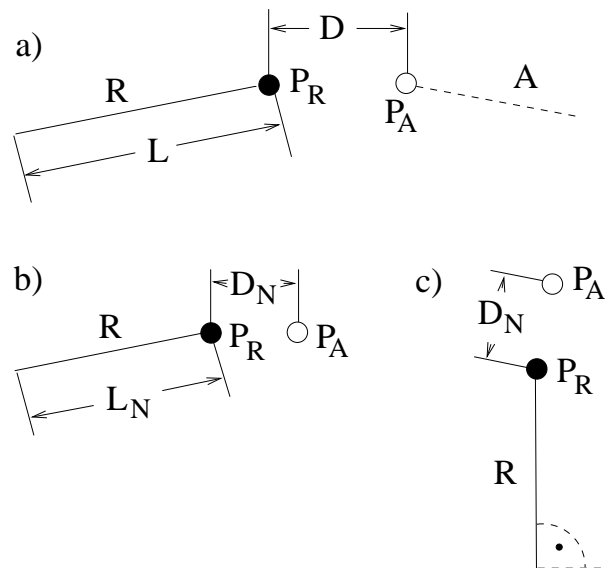


Abbildung 4.4: Beispiel für die paarweise Betrachtung von Gruppenprimitiva: Bestimmung von Referenzpunkt P_R (schwarz) und Anschlußpunkt P_A (weiß), sowie Normierung der Lückenlänge D und der Lage von R . Die relative Lage der Punkte P_R und P_A nach der Normierung dient als Merkmal zur Beschreibung des Bereichs perceptiver Aufmerksamkeit.

- a) Paar von Gruppenprimitiva mit dem betrachteten Primitivum R und dem Anschlußprimitivum A .
- b) Die Länge L_N von R und die Lücke D_N nach der Normierung.
- c) Drehung von R in eine 90° Normlage unter Beibehaltung der räumlichen Relation zwischen P_R und P_A .

Die **Lagenormierung** erlaubt die Zusammenfassung aller Referenzpunkte eines Gruppentyps in einem einzigen Punkt, den ich ebenfalls Referenzpunkt nenne. Wichtig ist, daß bei dieser Lagenormierung die relative Lage von Referenzpunkt und Anschlußpunkt erhalten bleiben. In der Umgebung dieses **Referenzpunktes** liegen alle Anschlußpunkte, die in der Stichprobe für diesen Gruppentyp gefunden wurden. Diese Anschlußpunkte ergeben die spezifische **Akkumulation** des untersuchten Gruppentyps. Eine dabei auftretende Überlagerung mehrerer Anschlußpunkte wird durch die Aufsummierung ihrer Anzahl berücksichtigt. In Abbildung 4.4c) ist diese Lagenormierung beispielhaft durch eine Drehung in eine 90° Normlage dargestellt.

Parallel zur Bestimmung der Bereiche perceptiver Aufmerksamkeit werden Schwellwerte δ_x und Maxima Δ_x für die verschiedenen Gruppentypen berechnet. Die **Schwellwerte** legen Grenzwerte für die Hypothesierung einer Gruppe fest. Sie werten im vorliegenden

Fall ausnahmslos die Orientierungsdifferenzen unter den Primitiva aus, und wurden so gelegt, daß 90% der Paare von Primitiva aus der Stichprobe eines Gruppentyps sie erfüllen. Die **Maxima** repräsentieren charakteristische Eigenschaften der unterschiedlichen Gruppentypen. Ein Maximum wird aus dem Histogramm der Daten der Stichprobe bestimmt: Es entspricht dem Wert an der Stelle der höchsten Trefferquote.

Die nachfolgenden Abschnitte beschreiben die notwendigen Schritte zur Ermittlung der relativen Lage von Referenzpunkt und Anschlußpunkt für die unterschiedlichen Gruppentypen. Gleichzeitig werden die berechneten Schwellwerte und Maxima vorgestellt.

4.3.2.2 Untersuchung der Stichprobe für kollineare Gruppen

Abbildung 4.5 zeigt die Vorgehensweise zur Bestimmung des Bereichs perzeptiver Aufmerksamkeit für kollineare Gruppen. Eine **kollineare Gruppe** besteht aus mindestens

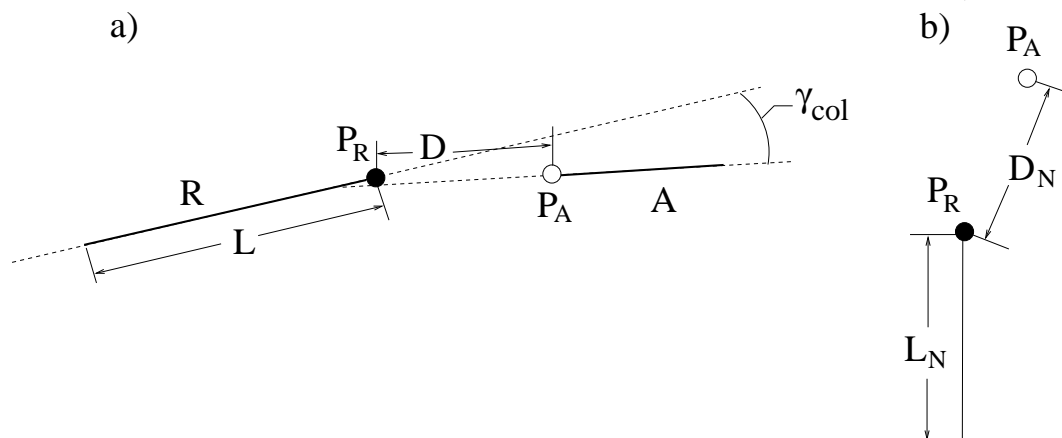


Abbildung 4.5: Untersuchung kollinearier Gruppen auf Bereiche perzeptiver Aufmerksamkeit:

a) Paar von Primitiva einer kollinearen Gruppe mit betrachtetem Primitivum R und Anschlußprimitivum A . Die Lückenlänge D wird durch den Referenzpunkt P_R und den Anschlußpunkt P_A begrenzt. γ_{col} bezeichnet die Orientierungsdifferenz der Primitiva.

b) Zustand nach der Normierung: L_N bezeichnet die normierte Länge des betrachteten Primitivums und D_N die normierte Lückenlänge. Das betrachtete Primitivum wurde in eine senkrechte Normlage gedreht unter Beibehaltung der relativen Lage von Referenzpunkt P_R und Anschlußpunkt P_A .

zwei geradenförmigen Primitiva. Das Paar untersuchter Primitiva entspricht der Gruppe

oder einer Untermenge von Primitiva der Gruppe. In Abbildung 4.5a) ist ein Paar von Primitiva der untersuchten Gruppe mit dem betrachtetem Primitivum R und dem Anschlußprimitivum A gezeigt. Die Lücke zwischen den Primitiva wird durch den Referenzpunkt P_R des betrachteten Primitivum und dem Anschlußpunkt P_A des Anschlußprimitivum begrenzt. Wie in 4.3.2.1 beschrieben wird für die Länge der Lücke D eine normierte Lückenlänge D_N entsprechend der Normierung des Referenzprimitivum berechnet.

Die Lagenormierung ist beispielhaft in Abbildung 4.5b durch eine senkrechte Normlage angedeutet. Die relative Lage von Referenzpunkt und Anschlußpunkt bleibt erhalten. Die Zusammenfassung aller Referenzpunkte der untersuchten Paare beschreibt den Bereich perceptiver Aufmerksamkeit für kollineare Gruppen.

Parallel zur Bestimmung der Bereiche perceptiver Aufmerksamkeit wird, entsprechend Abschnitt 4.3.2.1, aus den Orientierungsdifferenzen γ_{col} (vgl. Abbildung 4.5a) aller untersuchten Paare von Primitiva ein Schwellwert δ_{col} und ein Maximum für Orientierungsdifferenzen Δ_{col} festgelegt.

4.3.2.3 Untersuchung der Stichprobe für kurvilineare Gruppen

Die Bestimmung der Bereiche perceptiver Aufmerksamkeit für kurvilineare Gruppen erfolgt analog zu kollinearen Gruppen. Deshalb gelten die für kollineare Gruppen getroffenen Aussagen mit ein paar Erweiterungen, welche die speziellen Eigenschaften von ellipsenförmigen Primitiva berücksichtigen. Eine **kurvilineare Gruppe** besteht aus mindestens zwei ellipsenförmigen Primitiva, wie in Abbildung 4.6 gezeigt. Das Paar untersuchter Primitiva entspricht der Gruppe oder einer Untermenge von Primitiva der Gruppe.

Die Längennormierung kurvilinearere Gruppen nimmt für die Berechnung eine Fallunterscheidung vor, die auf der Auswertung des Öffnungswinkels ξ des Referenzprimitivum (vgl. Abbildung 4.6a)) beruht. Mit $L_N = 1$ berechnet sich die normierte Lückenlänge D_N zu:

$$D_N = \begin{cases} \frac{1}{L}D & : 0 < \xi \leq \pi \\ \frac{1}{u-L}D & : \pi < \xi \leq 2\pi \end{cases} \quad (4.1)$$

Der erste Teil von Gleichung 4.1 ermittelt den Normfaktor analog zu kollinearen Gruppen wie in Abschnitt 4.3.2.1 beispielhaft vorgestellt. Der zweite Teil der Gleichung dient der Bevorzugung der sich schließenden Ellipse, indem die für die Normierung betrachtete Länge aus der Differenz des Umfangs u der hypothetischen Ellipse, deren Bestandteil das Primitivum ist und der Länge L des Primitivum ermittelt wird.

Für die Lagenormierung ist die Orientierung der Krümmung des Referenzprimitivum zu beachten, da diese die Lage des Anschlußpunktes beeinflussen kann (es kann keine Annahme darüber gemacht werden, wie die Verteilung s-förmiger und ellipsenförmiger Strukturen

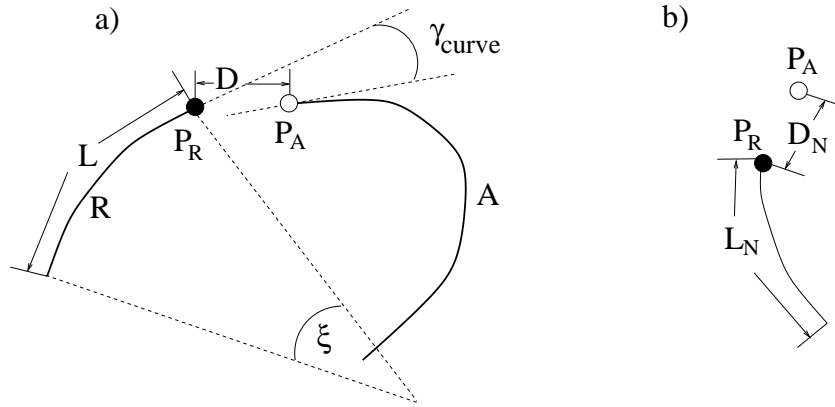


Abbildung 4.6: Untersuchung kurvilinearer Gruppen auf Bereiche perceptiver Aufmerksamkeit:

- a) Paar von Primitiva einer kurvilinearen Gruppe mit betrachtetem Primitivum R und Anschlußprimitivum A . Die Lückenlänge D wird durch den Referenzpunkt P_R und den Anschlußpunkt P_A begrenzt. γ_{curve} entspricht der Orientierungsdifferenz der Tangenten durch die Endpunkte P_R und P_A . b) Zustand nach der Normierung: L_N bezeichnet die normierte Länge des betrachteten Primitivum und D_N die normierte Lückenlänge. Das betrachtete Primitivum wurde in eine senkrechte Normlage mit rechtsweisender Krümmung gedreht unter Beibehaltung der relativen Lage von Referenzpunkt P_R und Anschlußpunkt P_A .

in der Domäne ist). In Abbildung 4.6c) ist beispielhaft eine senkrechte Normlage des Referenzprimitivum mit rechtweisender Krümmung angenommen.

Parallel zur Bestimmung des Bereichs perceptiver Aufmerksamkeit wird, entsprechend Abschnitt 4.3.2.1, aus den Orientierungsdifferenzen γ_{curve} der Tangenten in Referenz- und Anschlußpunkt (vgl. Abbildung 4.6) aller Paare ein Schwellwert δ_{curve} und ein Maximum für Orientierungsdifferenzen Δ_{curve} festgelegt.

4.3.2.4 Untersuchung der Stichprobe für Nähe-Gruppen

Die Bestimmung der Bereiche perceptiver Aufmerksamkeit für Nähe-Gruppen ist in Abbildung 4.7 gezeigt. **Nähe-Gruppen** bestehen aus genau zwei Primitiva, die sowohl geraden- als auch ellipsenförmig sein dürfen. Die Längennormierung und die Lagenormierung der Nähe-Gruppen wird entsprechend Abschnitt 4.3.2.1 vorgenommen. In Abbildung 4.7b) ist beispielhaft eine vertikale Normlage gezeigt.

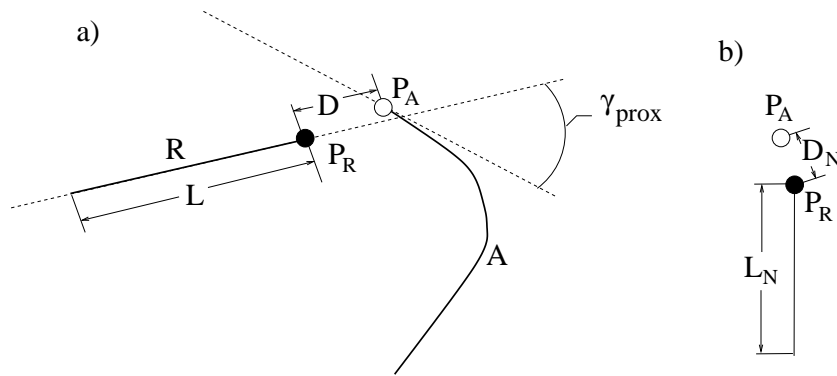


Abbildung 4.7: Untersuchung von Nähe-Gruppen auf Bereiche perceptiver Aufmerksamkeit:

- a) Nähe Gruppe mit betrachtetem Primitivum R und Anschlußprimitivum A . Die Lückenlänge wird durch den Referenzpunkt P_R und den Anschlußpunkt P_A begrenzt. γ_{prox} bezeichnet die Orientierungsdifferenz der Primitiva.
- b) Zustand nach der Normierung: L_N bezeichnet die normierte Länge des betrachteten Primitivum und D_N die normierte Lückenlänge. Das betrachtete Primitivum wurde in eine senkrechte Normlage gedreht unter Beibehaltung der relativen Lage von Referenzpunkt P_R und Anschlußpunkt P_A .

Parallel zur Bestimmung der Bereiche perceptiver Aufmerksamkeit wird für alle Gruppen die Orientierungsdifferenz γ_{prox} ermittelt. Für die Berechnung von γ_{prox} wird die Orientierung der Geradenstücke und die Orientierung der Tangente durch den betrachteten Endpunkt bei Ellipsensegmenten herangezogen. Entsprechend Abschnitt 4.3.2.1 wird aus allen γ_{prox} ein Maximum für Orientierungsdifferenzen Δ_{prox} von Nähe-Gruppen festgelegt.

4.3.2.5 Untersuchung der Stichprobe für parallele Gruppen

Die Bestimmung von Bereichen perceptiver Aufmerksamkeit für parallele Gruppen ist in Abbildung 4.8 gezeigt. Eine **parallele Gruppe** besteht aus genau zwei geradenförmigen Primitiva. Der Referenzpunkt P_R ist der Mittelpunkt des betrachteten Primitivum R . Der Anschlußpunkt P_A ist der Mittelpunkt des Anschlußprimitivum A (s.a. Abbildung 4.8a). Die Längennormierung und die Lagenormierung der parallelen Gruppen wird entsprechend Abschnitt 4.3.2.1 vorgenommen. In Abbildung 4.8b) ist beispielhaft eine horizontale Normlage des Referenzprimitivum gezeigt.

Parallel zur Bestimmung der Bereiche perceptiver Aufmerksamkeit wird, wie in Ab-

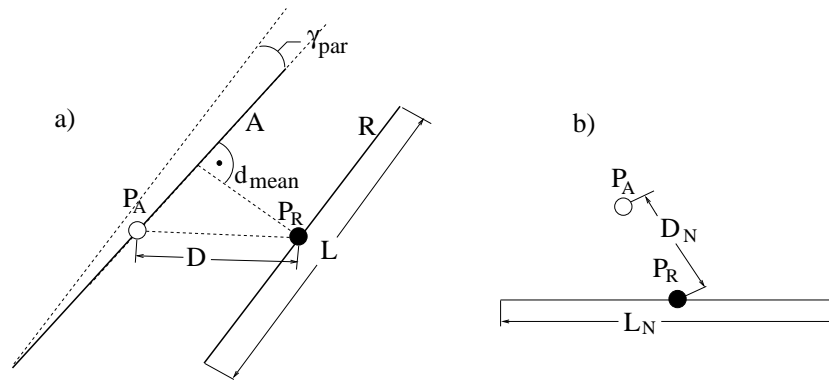


Abbildung 4.8: Untersuchung paralleler Gruppen auf Bereiche perzeptiver Aufmerksamkeit:

a) Parallele Gruppe mit betrachtetem Primitivum R und Anschlußprimitivum A . Die Lückenlänge D wird durch die Mittelpunkte der Primitiva: den Referenzpunkt P_R und den Anschlußpunkt P_A begrenzt. γ_{par} bezeichnet die Orientierungsdifferenz der Primitiva und d_{mean} die mittlere Distanz Primitiva.

b) Zustand nach der Normierung: L_N bezeichnet die normierte Länge des betrachteten Primitivum und D_N die normierte Lückenlänge. Das betrachtete Primitivum wurde in eine waagerechte Normlage gedreht unter Beibehaltung der relativen Lage von Referenzpunkt P_R und Anschlußpunkt P_A .

schnitt 4.3.2.1 beschrieben, aus den Orientierungsdifferenzen γ_{par} (vgl. Abbildung 4.8a) aller untersuchten parallelen Gruppen ein Schwellwert δ_{par} und ein Maximum für Orientierungsdifferenzen Δ_{col} festgelegt. Ein zweites Maximum Δ_{mean} charakterisiert die mittlere Distanz der Primitiva: Die mittlere Distanz d_{mean} wird berechnet, indem man vom Referenzpunkt das Lot auf das Anschlußprimitivum fällt. Anschließend wird d_{mean} mit dem Normfaktor des Referenzprimitivum multipliziert, um einen Vergleich unter allen mittleren Distanzen der Stichprobe zu ermöglichen.

4.3.2.6 Akkumulation der Anschlußpunkte

Die Abbildungen 4.9a,b) und 4.10a,b) zeigen die Akkumulationen der Anschlußpunkte als Resultat der Auswertung der handgelabelten Gruppen. Die Darstellung der Akkumulationen erfolgt in Form von Grauwertbildern. Der Default-Wert der Pixel ist mit 0 voreingestellt. Jeder gefundene Anschlußpunkt erhöht diesen Wert um 1. Fallen meh-

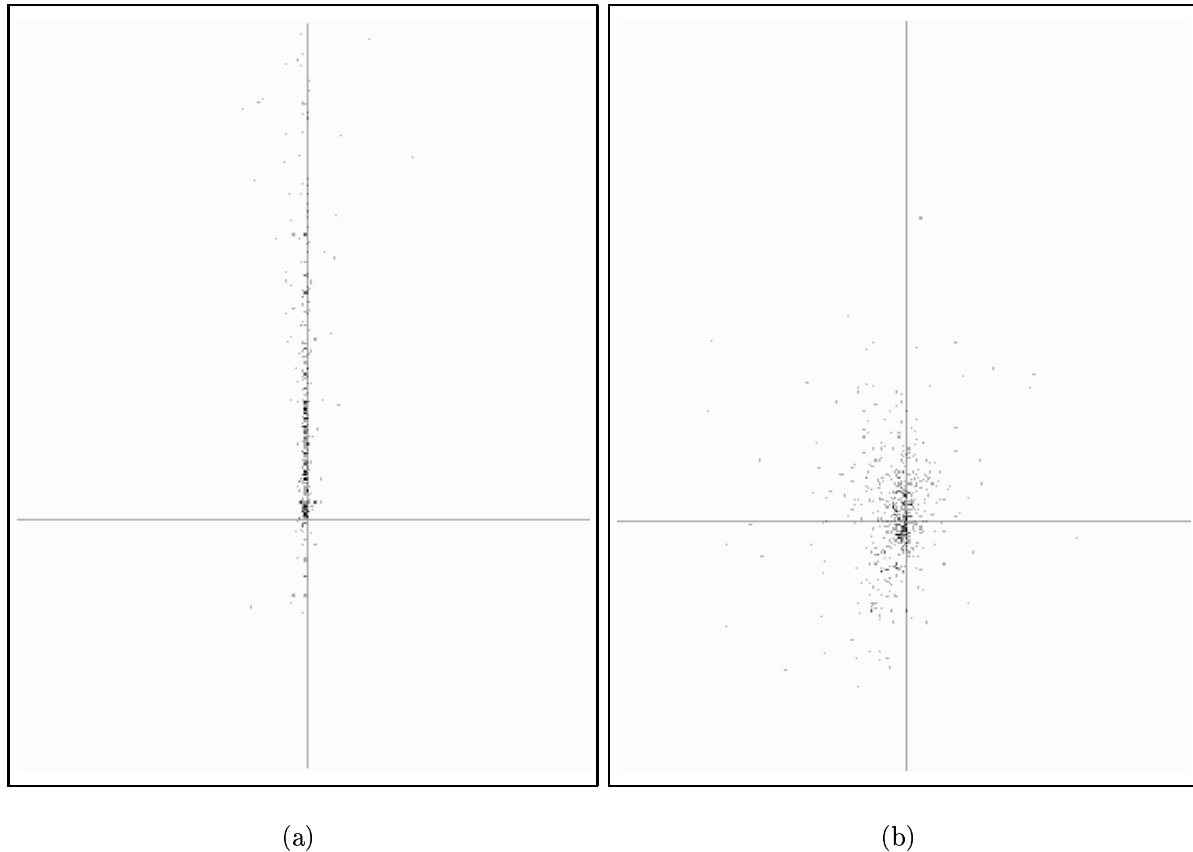


Abbildung 4.9: Akkumulation der Anschlußpunkte für a) kollineare- und b) kurvilineare Gruppen.

Der Referenzpunkt ist durch den Kreuzungspunkt der senkrechten Linienzüge gekennzeichnet.

rere Anschlußpunkte zusammen, entspricht also der Grauwert ihrer Anzahl. Für die in den Abbildungen 4.9a,b) und 4.10a,b) gezeigten Akkumulationen wurde ein einheitlicher Abbildungsmaßstab gewählt, sodaß ein Vergleich unter den verschiedenen Gruppentypen möglich ist. Aus Gründen der Visualisierung erfolgt die Darstellung der Grauwerte invers. Für alle Häufungsbilder ist der Referenzpunkt durch den Kreuzungspunkt zweier senkrechter Linien gekennzeichnet.

Für kollineare Gruppen unter 4.9a) zeigt sich eine Häufung in Richtung der Verlängerung des betrachteten Primitivum, mit geringen Abweichungen zu beiden Seiten. Kurvilineare Gruppen unter 4.9b) zeigen eine Häufung in der Umgebung des Referenzpunkts mit einer leicht vertikalen Vorzugsrichtung. Kollineare und kurvilineare Gruppen verfügen über bevorzugte Winkelbeziehungen unter den Primitiva der Gruppe. Dies führt zu der vertikalen

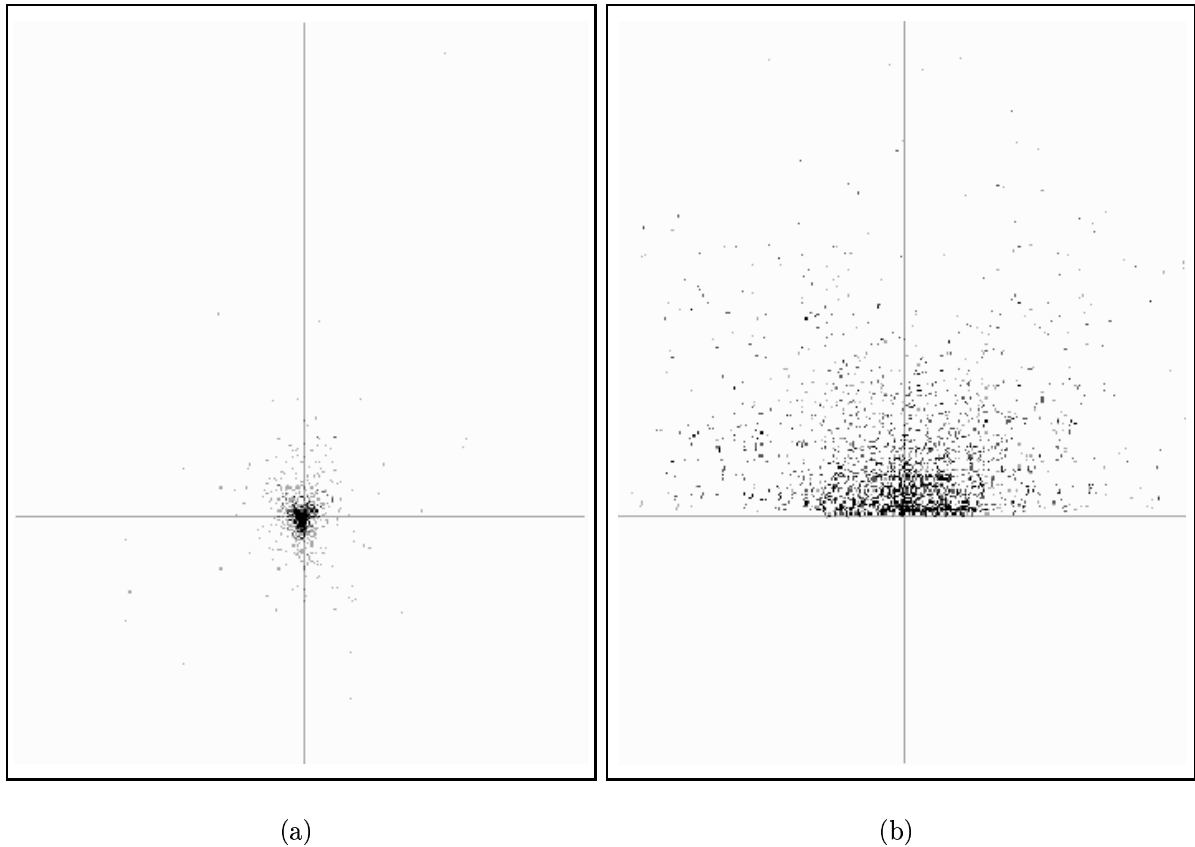


Abbildung 4.10: Akkumulation der Anschlußpunkte für a) Nähe-Gruppen und b) parallele Gruppen. Der Referenzpunkt ist durch den Kreuzungspunkt der senkrechten Linienzüge gekennzeichnet.

Ausprägung der Bereiche. Anschlußpunkte für Nähe-Gruppen in 4.10a) finden sich in einem kreisförmigen Bereich um den Referenzpunkt. Insgesamt ist dieser Bereich sehr viel kleiner im Vergleich zu allen anderen Gruppentypen. Für alle drei Gruppentypen gilt, daß im Häufungsbild unterhalb des Referenzpunktes Punkte erscheinen. Für kollineare und kurvilineare Gruppen deutet dies auf eine Überlappung der betrachteten Primitiva. Nähe-Gruppen verfügen über keine bevorzugten Winkelbeziehungen unter den Primitiva, sodaß Anschlüsse in allen Richtungen möglich sind. Die Anschlußpunkte paralleler Gruppen unter 4.10b), finden sich in einem halbkreisförmigen Bereich oberhalb des Referenzpunktes.

4.3.3 Ableitung von Suchbereichen für den Gruppierungsprozeß

In den vorhergehenden Abschnitten wurde beschrieben, wie aus den handgelabelten Stichproben der verschiedenen Gruppentypen Akkumulationen von Anschlußpunkten abgeleitet wurden. Diese Akkumulation speichert die Information über die gruppierende Wirkung eines Primitivum in seiner Umgebung. Diese Information soll in geeigneter Weise einem automatisierten Suchprozeß für Gruppen zugänglich gemacht werden, dessen Resultate der Generierung der Gruppenhypothesen dienen. Zu diesem Zweck wird aus den Akkumulationen der Anschlußpunkte für jeden Gruppentyp ein geschlossener Bereich segmentiert, den ich im folgenden mit **Suchbereich** bezeichne. Dieser Suchbereich wird auf die betrachtete Szene angewandt und erlaubt eine Aussage darüber, ob es sinnvoll ist, eine bestimmte Gruppenhypothese für zwei Primitiva zu erzeugen.

Die Bestimmung von Suchbereichen aus den Akkumulationen der Anschlußpunkte gliedert sich im wesentlichen in drei Schritte:

- Verschmelzen der Anschlußpunkte durch geeignete Filterung
- Segmentierung eines Bereiches
- Beschreibung dieses Bereichs durch einen zusammenhängenden Polygonzug

Das Verschmelzen eines Punktes mit seiner direkten Umgebung geschieht durch Anwendung von Gauß-Faltungskernen auf die Akkumulation der Anschlußpunkte. Die Entscheidung zugunsten der Gaußfunktion ist sinnvoll, da sie eine kreisförmige Ausweitung des Punktes mit einer gleichzeitig stark abklingenden Wirkung ermöglicht.

Jeder Punkt der Akkumulation besitzt einen spezifischen Grauwert. Die Höhe des Grauwerts entspricht der Anzahl der Anschlußpunkte. Ein Punkt mit einem hohen Grauwert soll auf seine Umgebung mehr Einfluß nehmen können, als Punkte mit einem niedrigen Grauwert. Dies wird durch die Variation der σ -Werte der Filterkurve erreicht. Je höher der Grauwert eines Punktes, desto höher der σ -Wert der Gaußfunktion und desto breiter der Einfluß des Punktes in seiner Nachbarschaft.

Abbildung 4.11a-g) zeigt den schematischen Verlauf des angewandten Verfahrens. In einem ersten Schritt wird die Akkumulation der Anschlußpunkte in Abbildung 4.11a) durch Schwellwertbildung in fünf Teilbilder in Abbildung 4.11b) aufgespalten. Der erste Schwellwert wird so gewählt, daß sich ca. 90% aller Punkte der Akkumulation im ersten Teilbild befinden. Für die restlichen vier Teilbilder werden die Schwellwerte durch lineare Teilung zwischen dem ersten Schwellwert und dem maximalen Grauwert gebildet.

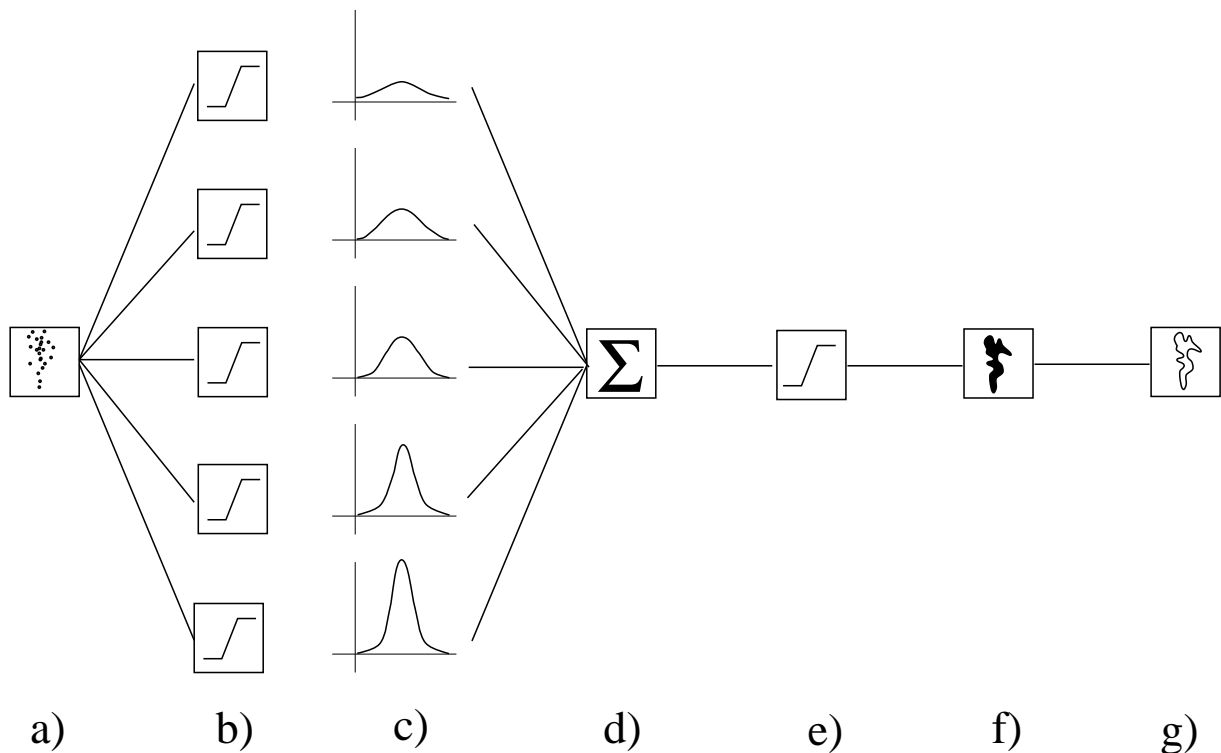


Abbildung 4.11: Schema für die Erstellung von Suchbereichsmasken:

von links nach rechts:

- a) Akkumulation der Anschlußpunkte eines Gruppentyps
- b) Schwellwertbildung entsprechend der maximalen Trefferquote
- c) Bearbeitung mit verschiedenen Gauß-Faltungskernen
- d) Addition der Filterbilder
- e) Bereichsegmentierung durch Binarisierung
- f) Glätten der Umrißkontur durch Closing
- g) Polygonzug des Bereichs

Jedes der Teilbilder wird mit dem passenden Gauß-Faltungskern in Abbildung 4.11c) bearbeitet. Jeweils fünf Faltungskerne bilden einen Filtersatz. Vier verschiedene Filtersätze mit den Größen: 15x15, 25x25, 35x35, 45x45 Pixel, wurden getestet. Die Größe des Filters wurde folgendermaßen auf die σ -Werte des Filtersatzes abgestimmt: Der größte σ -Wert jeden Satzes beträgt $\frac{1}{4}$ der Breite des Filters. Alle weiteren σ -Werte entsprechen $\frac{4}{5}$, $\frac{3}{5}$, $\frac{2}{5}$ und $\frac{1}{5}$ der Filterbreite.

Nach der Anwendung des Filtersatzes werden die fünf Teilbilder in Abbildung 4.11d) zu einem Ergebnisbild aufaddiert. Die addierten Filterbilder sind in den Abbildungen 4.12a)e) und 4.13a)e) für die untersuchten Gruppentypen gezeigt. Die in den folgenden Abbildun-

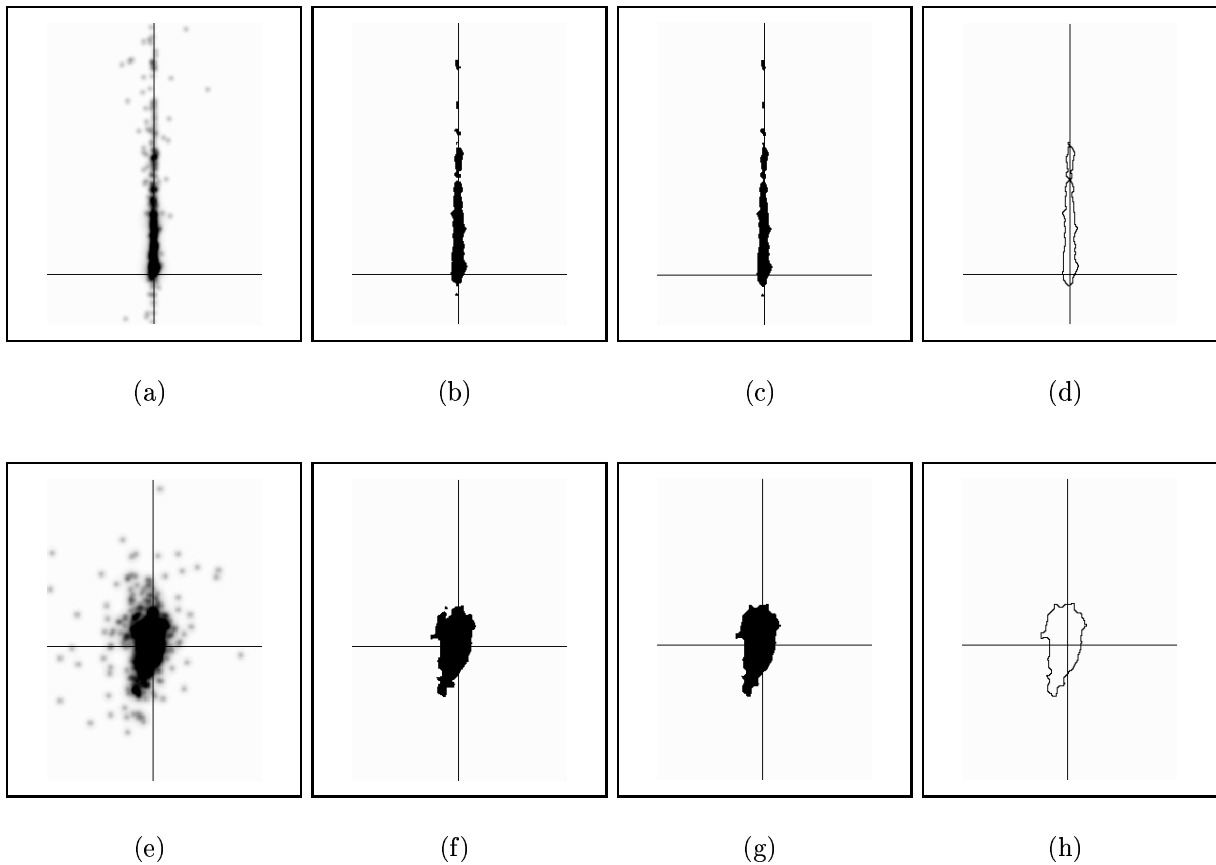


Abbildung 4.12: Bestimmung der Suchbereichsmaske für kollineare und kurvilineare Gruppen:

- a, e) Addition der Filterbilder (Filtergröße 45x45 Pixel)
- b, f) Binarisierung
- c, g) Closing
- d, h) Polygonzug
(Kennzeichnung des Referenzpunktes durch den Kreuzungspunkt der senkrechten Linienzüge)

gen gezeigten Ergebnisse sind für eine Filtergröße von 45x45 Pixel erstellt. Mit dieser Filtergröße erzeugte Polygonzüge wurden für die Implementation verwandt. Die Größe der segmentierten Bereiche erbrachte im automatischen Gruppierungsprozeß die besten Resultate. Gleichzeitig sind die Ränder der Bereiche durch die große Filtermaske relativ glatt.

Die Addition der Filterbilder wird anschließend binarisiert. Der für die Binarisierung gewählte Schwellwert entspricht der Hälfte des maximalen Grauwertes, des aktuellen Bil-

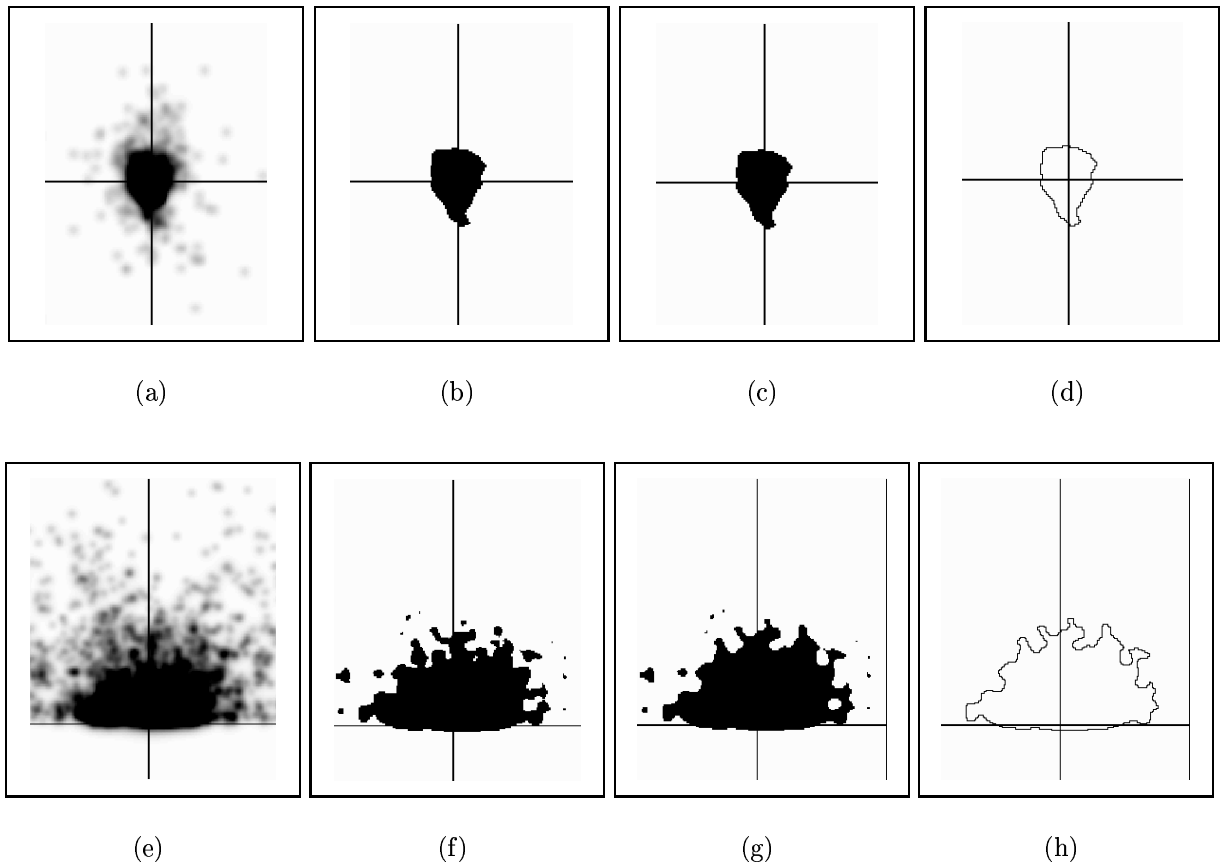


Abbildung 4.13: Bestimmung der Suchbereichsmaske für nahe und parallele Gruppen:

a, e) Addition der Filterbilder (Filtergröße 45x45 Pixel)

b, f) Binarisierung

c, g) Closing

d, h) Polygonzug

(Kennzeichnung des Referenzpunktes durch den Kreuzungspunkt der senkrechten Linienzüge)

des. Die durch die Binarisierung entstandenen Bereiche sind in den Abbildungen 4.12b)f) und 4.13b)f) gezeigt.

Die Glättung der Umrißkonturen in Abbildung 4.11f) erfolgt durch **Closing**, einer Zusammenfassung der morphologischen Verfahren **Dilatation** und **Erosion**. In einer dreifachen Iteration wurde ein Strukturelement eingesetzt, welches abwechselnd eine 4-er, bzw. 8-er Nachbarschaft untersuchte. Dies bewirkt eine richtungsinvariante Glättung der Konturen. Die Größe des Bereichs bleibt durch die Glättung im wesentlichen unbeeinflusst. Das Ergebnis des Closing zeigen die Abbildungen 4.12c)g) und 4.13c)g). Nach der Glättung der

Konturen wird der jeweils größte zusammenhängende Bereich in Abbildung 4.11g) durch einen Polygonzug umschrieben. Dieser Polygonzug kommt als **Suchbereich** in einem automatischen Gruppierungsprozeß zum Einsatz. Er stellt damit den Bereich der perceptiven Aufmerksamkeit dar. Die Polygonzüge der untersuchten Gruppen sind in den Abbildungen 4.12d)h) und 4.13d)h) gezeigt.

In diesem Abschnitt wurde gezeigt, wie aus den Akkumulationen der Anschlußpunkte Polygonzüge gewonnen werden, die als Suchbereiche in einem automatisierten Prozeß perceptiven Gruppierens zum Einsatz kommen. Diese Suchbereiche entsprechen dem Bereich der perceptiven Aufmerksamkeit. Wie sie für die verschiedenen Gruppen eingesetzt werden, wird im folgenden Abschnitt erläutert.

4.4 Generierung der Gruppenhypothesen

Die Erzeugung von Gruppenhypothesen gliedert sich in zwei Teilschritte:

- Die Suche nach einem Anschlußprimitivum für ein betrachtetes Primitivum
- Die Überprüfung der gefundenen Primitiva mit Hilfe weiterer Randbedingungen auf ihre Gruppenzugehörigkeit

Für die Suche nach gruppierbaren Anschlußsegmenten gibt es zwei verschiedene Strategien, die nach den betrachteten Gruppentypen unterschieden werden können. Für kollineare, kurvilineare, Nähe- und parallele Gruppenhypothesen wird die Suche auf der Basis der in Abschnitt 4.3.3 beschriebenen Suchbereiche durchgeführt. Im Gegensatz dazu erfolgt die Suche nach Geschlossenheiten mit Hilfe einer graphbasierten Auswertung von Nähe-Gruppen und parallelen Gruppen.

4.4.1 Generierung von Gruppenhypothesen mit Suchbereichen

Die Auswertung der Akkumulationen der Anschlußpunkte resultiert in Polygonzügen, die die Suchbereiche für kollineare, kurvilineare, Nähe- und parallele Gruppen repräsentieren. Diese Polygonzüge werden für die örtlich begrenzte Suche in der Umgebung des betrachteten Referenzprimitivum A eingesetzt. Der Einsatz dieser Suchbereiche geht folgendermaßen vor sich:

- Der Referenzpunkt des Suchbereichs wird mit dem Referenzpunkt des betrachteten Primitivum in Übereinstimmung gebracht.

- Die Größe des Polygonzuges wird invers mit der normierten Länge des Primitivum skaliert.
- Die Ausrichtung des Polygonzugs erfolgt entsprechend der Orientierung des betrachteten Primitivum.
- Untersuchung aller Primitiva, die mindestens einem Endpunkt innerhalb der Suchmaske besitzen, auf die Erzeugung einer Gruppenthypothese. Für diese Untersuchung werden Randbedingungen, wie z.B. die Orientierungsdifferenz der Primitiva, ausgewertet.

Kollineare Gruppenthypothesen Der Referenzpunkt des kollinearen Suchbereichs wird mit dem Referenzpunkt des betrachteten Primitivum in Übereinstimmung gebracht, und die Größe der Maske invers mit der normierten Länge des Primitivum skaliert. Die Orientierung der Suchmaske für kollineare Gruppen entspricht der Orientierung des geradenförmigen Referenzprimitivum. Beide Endpunkte des Referenzprimitivum werden nacheinander untersucht.

Alle Primitiva mit einem Anschlußpunkt innerhalb der Suchmaske werden auf die Bildung einer kollinearen Gruppenthypothese mit dem Referenzprimitivum untersucht.

Eine Randbedingung für die Generierung einer kollinearen Hypothese besagt, daß eine solche Gruppe nur aus geradenförmigen Primitiva bestehen darf. Diese Primitiva sollen längs ihrer Achsen eine gute Fortsetzung bilden, deshalb wird als zweite Randbedingung der Schwellwert der Orientierungsdifferenzen δ_{col} aus Abschnitt 4.3.2.2 angewendet. Ist die aktuelle Orientierungsdifferenz der beiden geradenförmigen Segmente kleiner als der Schwellwert, wird eine kollineare Hypothese generiert. Für die Repräsentation der kollinearen Hypothese werden die zugrundeliegenden Kantenelemente erneut durch ein Geradenstück approximiert. Eine kollineare Hypothese wird im weiteren Verlauf des Suchprozesses selbst wieder zu einem Primitivum, welches an seinen Endpunkten nach weiteren Anschlußprimitiva untersucht wird. Eine kollineare Hypothese ist so iterativ erweiterbar. Abbildung 4.14c) zeigt beispielhaft einen kollinearen Suchbereich in einer Szene mit Lochleisten.

Kurvilineare Gruppenthypothesen Das Vorgehen für den Einsatz der kurvilinearen Suchbereiche erfolgt analog zu kollinearen Suchbereichen. Eine Randbedingung für die Generierung einer kurvilinearen Hypothese besagt, daß eine solche Gruppe nur aus ellipsenförmigen Primitiva mit einem gleichartigen Krümmungsverlauf bestehen darf. Diese Primitiva

sollen längs ihrer Tangenten durch Referenzpunkt und Anschlußpunkt eine gute Fortsetzung bilden, deshalb wird als zweite Randbedingung der Schwellwert der Orientierungsdifferenzen δ_{curve} aus Abschnitt 4.3.2.3 angewendet. Ist die aktuelle Orientierungsdifferenz der beiden Tangenten durch Referenz- und Anschlußpunkt kleiner als der Schwellwert, wird eine kurvilineare Hypothese generiert. Analog zu kollinearen Hypothesen wird eine kurvilineare Hypothese durch eine erneute Approximation repräsentiert. Eine kurvilineare Hypothese wird im weiteren Verlauf des Suchprozesses selbst wieder zu einem Primitivum, welches an seinen Endpunkten nach weiteren Anschlußprimitiva untersucht wird. Eine kurvilineare Hypothese ist somit ebenfalls iterativ erweiterbar.

Nähe-Hypothesen Das Vorgehen für den Einsatz der Nähe-Suchbereiche erfolgt analog zu kollinearen Suchbereichen. Eine Nähe-Gruppe besteht aus exakt zwei Primitiva. Randbedingungen bezüglich der Art oder der Orientierung der Primitiva gibt es nicht. Zur Begrenzung der sehr großen Anzahl hypothetisierter Nähe-Gruppen wurde eine technische Randbedingung eingeführt: Zur Bildung einer Nähe-Hypothese muß die Gruppe wechselseitig von beiden Primitiva der Gruppe erzeugt werden können.

Parallele Hypothesen Das Vorgehen für den Einsatz der parallelen Suchbereiche erfolgt analog zu kollinearen Suchbereichen. Die Orientierung der Suchmaske für parallele Gruppen erfolgt senkrecht zur Orientierung ($\pm 90^\circ$) des Referenzprimitivum. Beide "Seiten" des Referenzprimitivum werden nacheinander untersucht. Alle Primitiva mit einem Anschlußpunkt innerhalb der Suchmaske werden auf die Bildung einer parallelen Gruppenhypothese untersucht. Eine Randbedingung für die Generierung einer parallelen Hypothese besagt, daß eine solche Gruppe nur aus geradenförmigen Primitiva bestehen darf. Als zweite Randbedingung wird der Schwellwert der Orientierungsdifferenzen δ_{par} aus Abschnitt 4.3.2.5 angewendet. Ist die aktuelle Orientierungsdifferenz der Primitiva kleiner als der Schwellwert, wird eine parallele Hypothese erzeugt.

4.4.2 Generierung von Geschlossenheiten

Die Generierung von Geschlossenheiten beruht auf der Auswertung bereits erzeugter Nähe-Hypothesen und paralleler Hypothesen. Zwei verschiedene Strategien zur Bildung von Geschlossenheiten werden untersucht:

- Die Bildung einer zyklischen Abfolge von Nähe-Gruppen.

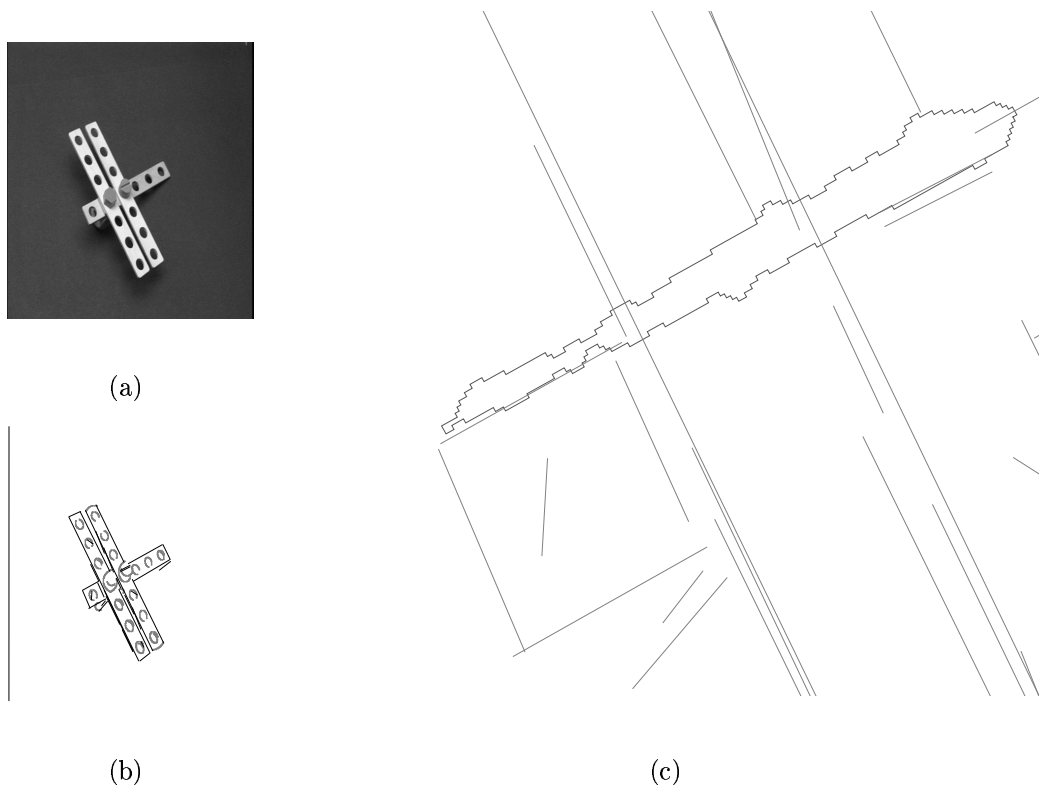


Abbildung 4.14: Suchbereich für kollineare Gruppen:

- a) Grauwertbild eines teilmontierten Flugzeugs (“Rumpf” und “Flügel”)
- b) Initiale Primitiva
- c) Ausschnitt aus b) mit Geradenstücken und einem kollinearen Suchbereich: Dieser Bereich ermöglicht die Gruppierung der “Rumpfleiste”, die von den beiden “Flügelleisten” verdeckt wird.

- Die Untersuchung auf schließende Hypothesen oder Primitiva an den gegenüberliegenden Endpunkten von parallelen Hypothesen.

Beide Ansätze basieren auf einem ungerichteten Graphen, dem **Proximitäts-Graphen**. In diesen Graphen werden die Gruppenthypothesen und Primitiva der 1-D Ebene abgebildet. Die Knoten des Graphen entsprechen den Endpunkten der kollinearen und kurvi-linearen Hypothesen, sowie den Endpunkten der Primitiva. Diese Knoten werden durch zwei unterschiedliche Kantentypen verbunden. Eine **Gruppierungskante** verbindet die Endpunkte einer Gruppenthypothese, bzw. eines Primitivum untereinander. Eine **Proximitätskante** verbindet die Endpunkte verschiedener Hypothesen oder Primitiva, die eine Nähe-Gruppe bilden. Für Geschlossenheiten, die nur auf der Auswertung von Nähe-Gruppen basieren, werden geschlossene Wege innerhalb des Graphen gesucht. Zwei wichtige Randbedingungen müssen erfüllt sein: Zum einen müssen die Kanten des geschlossenen

Weges immer abwechselnd einer Proximitäts-, bzw. Gruppierungskante entsprechen. Zum anderen dürfen unter den Kanten eines geschlossenen Weges keine Überschneidungen existieren, da dies zur Erzeugung ungewünschter Geschlossenheiten führt (es werden nur einfache, geschlossene Strukturen gesucht).

Für Geschlossenheiten, die auf der Auswertung von Parallelitäten beruhen, werden innerhalb des Proximitätsgraphen die kürzesten Wege gesucht, die die jeweils gegenüberliegenden Endpunkte miteinander verbinden. Auch für diese Wege gilt, daß die Kanten abwechselnd einer Proximitäts- oder Gruppierungskante entsprechen müssen.

Die oben beschriebene Art der Generierung von Geschlossenheiten durch die Auswertung von Nähe-Hypothesen und Parallelitäten, kann zu Geschlossenheitshypothesen führen, denen identische Kontursegmente zugrunde liegen. Um die Anzahl der erzeugten Geschlossenheiten zu begrenzen, wird in diesem Fall die Geschlossenheit auf Grundlage von Nähe-Hypothesen, zu Gunsten der Geschlossenheit auf Grundlage der Parallelität, vernachlässigt. Beide Geschlossenheiten beschreiben den gleichen Szeneninhalte, deshalb wird es nicht als sinnvoll angesehen, beide einer Bewertung durch das Markov Random Field zuzuführen. Die Auswahl fällt auf die Parallelität, weil diese im Vergleich zu Nähe-Hypothesen als signifikanter angesehen wird.

Zusammenfassung Das vorliegende Kapitel befaßte sich mit der Vorstellung eines neuartigen Ansatzes zur Suche nach geeigneten Gruppierungsprimitiva. Neu an diesem Ansatz ist das Bestreben, diese **Bereiche perzeptiver Aufmerksamkeit** nicht aus willkürlichen Annahmen zu formen, sondern aus den Charakteristika der betrachteten Gruppentypen entstehen zu lassen.

Die verschiedenen Gruppentypen dieses Ansatzes sind: Kollinearität, Kurvilinearität, Nähe, Parallelität und Geschlossenheit. Diese Gruppentypen sind in eine Hierarchie der Gruppierungshypothesen eingebunden, wie sie der Abschnitt 4.2 vorstellt. Diese Hierarchie trägt der wachsenden räumlichen Komplexität der Hypothesen Rechnung. Die Hierarchieebenen werden deshalb auch mit den Namen **1D**, **2x1D** und **2D** unterschieden. Der Gruppierungsprozeß startet mit der Untersuchung der Kontursegmente: den initialen Primitiva. Wie in Abbildung 4.1 gezeigt, werden dann - aufeinanderfolgend - die Gruppen der verschiedenen Hierarchieebenen erzeugt.

Für die Suche nach geeigneten Primitiva zur Generierung der Gruppenhypothesen werden spezifische Bereiche eingesetzt, welche den Charakteristika der unterschiedlichen Gruppentypen Rechnung tragen. Die Ermittlung dieser Bereiche ist in Abschnitt 4.3 ausführlich dargestellt. Für diese Untersuchung wurde eine repräsentative Stichprobe der Domäne erzeugt, wie sie in Abschnitt 4.3.2 beispielhaft vorgestellt ist. Die Auswertung der Stichprobe

erfolgt in drei Teilschritten: Definition eines geeigneten Merkmals; Definition einer Norm für das Merkmal; Detektion des charakteristischen Bereichs.

Wie diese Auswertung für die unterschiedlichen Gruppentypen aussieht, ist in den nachfolgenden Abschnitten 4.3.2.2 für kollineare Gruppen, 4.3.2.3 kurvilineare Gruppen, 4.3.2.4 und 4.3.2.5 parallele Gruppen ausgeführt. Für Geschlossenheiten wurden keine Suchbereiche betrachtet, da ihre Generierung auf einer spezifischen Auswertung von Nähe-Gruppen und Parallelitäten basiert, wie sie in Abschnitt 4.4.2 besprochen wird.

Aus dieser Auswertung entstehen Akkumulationen von Anschlußpunkten, wie sie der Abschnitt 4.3.2.6 zeigt. Die in diesen Akkumulationen gespeicherte Information über die gruppierende Wirkung eines Primitivum auf seine Umgebung wird in geeigneter Art und Weise für den automatisierten Gruppierungsprozeß aufbereitet. Die Vorgehensweise gliedert sich im wesentlichen in drei Schritte, die in Abschnitt 4.3.3 ausführlich vorgestellt werden: Verschmelzen der Punkte durch geeignete Filterung; Segmentierung des entstandenen Bereiches; Beschreibung dieses Bereiches durch einen zusammenhängenden Polygonzug. Mit Hilfe der Suchbereiche können nun in einem automatisierten Gruppierungsprozeß die verschiedenen Hypothesen generiert werden. Die Generierung der Hypothesen gliedert sich in zwei Teilschritte: Die Suche nach einem Anschlußprimitivum und die Überprüfung der gefundenen Primitiva auf ihre Gruppenzugehörigkeit mit Hilfe weiterer Randbedingungen. Dieses ist für die unterschiedlichen Hypothesen in Abschnitt 4.4.1 ausführlich dargestellt. Wie oben bereits erwähnt, stellt die Erzeugung von Geschlossenheiten in diesem Zusammenhang eine Ausnahme dar.

Die Erzeugung der Gruppenthypothesen mit Bereichen perzeptiver Aufmerksamkeit, wie in diesem Kapitel vorgestellt, basiert auf der Auswertung rein lokaler Information. Durch diese lokale Betrachtungsweise ist die Entstehung mehrdeutiger Gruppenthypothesen nicht ausgeschlossen. Die Auflösung dieser inkonsistenten Interpretation erfolgt mit Hilfe einer Bewertung der Gruppenthypothesen. Für ein optimales Ergebnis muß es möglich sein, globale Szeneninformation in den Bewertungsprozeß einfließen zu lassen. Wie dieses im einzelnen geschieht, wird das Thema der beiden folgenden Kapitel sein: In Kapitel 5 wird der grundlegende Mechanismus von Markov Random Fields eingeführt, auf dem die Bewertung beruht. Kapitel 6 wendet diesen Mechanismus auf die Bewertung der Gruppenthypothesen an.

5 Markov Random Fields

Die Generierung von Gruppenhypothesen mit den im vorangehenden Kapitel 4 vorgestellten **Bereichen perzeptiver Aufmerksamkeit** begrenzt die Betrachtung der Gruppenhypothesen auf einen lokalen Rahmen. Diese lokale Betrachtungsweise kann in einer Mehrdeutigkeit der generierten Gruppenhypothesen resultieren. Diese Mehrdeutigkeit tritt beispielsweise auf, wenn verschiedene Gruppenhypothesen für eine Anzahl betrachteter Primitiva eine inkonsistente Interpretation, z.B. über den Verlauf einer Linie, liefern. Zur Auflösung von Mehrdeutigkeiten ist eine Auswertung der Gruppenhypothesen im Kontext ihrer Nachbarschaft notwendig. Gleichzeitig muß der eingesetzte Mechanismus in der Lage sein, eine Aussage über die Bedeutung der untersuchten Hypothesen im Rahmen der betrachteten Szene zu erlauben. Für die Bewertung der Gruppenhypothesen wurde deshalb der Mechanismus eines **Markov Random Field** gewählt, weil er diese wichtigen Eigenschaften zur Verfügung stellt:

- **Nachbarschaftssysteme** ermöglichen die Betrachtung einer Gruppenhypothese im Kontext ihrer Nachbarn. Innerhalb dieser Nachbarschaftssysteme kann Wissen über die lokalen Relationen der Gruppenhypothesen modelliert werden.
- Aus allen **Konfigurationen** des Feldes ist eine globale Interpretation der betrachteten Szene ermittelbar, z.B. durch die Auswertung der **Maximum a posteriori** Wahrscheinlichkeit.

Die Betrachtung von Gruppenhypothesen im Rahmen von Markov Random Fields wird von mir in zwei aufeinander folgenden Kapiteln vorgestellt: Dieses Kapitel führt die Grundlagen für den Einsatz von Markov Random Fields in der Bildverarbeitung ein. Die typischen Anwendungsgebiete dieser Random Fields werden vorgestellt und die grundlegenden Begriffe geklärt. Das folgende Kapitel 6 stellt danach den speziellen Ansatz zur Bewertung der Gruppenhypothesen vor.

5.1 Markov Random Fields und Bildverarbeitung

Vielen in der Bildverarbeitung verfolgten Fragestellungen liegt eine räumlich-lokale Struktur zugrunde. In der Restauration verrauschter Bilder ist es beispielsweise sinnvoll, für die Ermittlung eines aktuellen Pixelwertes die Pixelwerte der direkten Umgebung des betrachteten Pixel zu berücksichtigen. Der Einsatz von Markov Random Fields für Fragestellungen dieser Art ist von besonderem Interesse, da Markov Random Fields Nachbarschaftssysteme besitzen, die eine kontextsensitive Betrachtung ermöglichen. Eine **Nachbarschaft** verbindet eine Anzahl von Entitäten i des Random Field zu Nachbarn mit Hilfe einer **Nachbarschaftsrelation**. Diese Entitäten können z.B. den Pixeln einer Bildmatrix entsprechen. Eine Nachbarschaftsrelation kann in diesem Fall so gewählt sein, daß sie alle Pixel in der näheren Umgebung des betrachteten Pixel aufgrund eines räumlichen Abstandsmaßes verbindet. Betrachtet man die Pixel als Knoten eines ungerichteten Graphen, dann werden die Kanten des Graphen durch die Nachbarschaftsrelation gebildet. Alle durch Kanten verbundenen Knoten sind benachbart und bilden ein **Nachbarschaftssystem** bezüglich der betrachteten Entität. Die Entität selbst gehört ihrem Nachbarschaftssystem nicht an. Jede Entität ist mit einer **Zufallsvariablen** F_i verbunden. Aufgrund der Betrachtung des Nachbarschaftssystems wird der Entität ein Wert zugeordnet, der in die Zufallsvariable abgebildet wird. Dieser Wert kann in Form eines **Label** l aus einer **Labelmenge** \mathcal{L} vergeben werden. Die gewählte Labelmenge muß geeignet sein, alle relevanten Zustände des untersuchten Prozesses zu repräsentieren. Der Vorgang der passenden Zuordnung eines Label aus der Labelmenge zu jeder Entität wird auch mit **Labeling Problem** bezeichnet. Ist allen Entitäten des Random Field ein Label zugeordnet, spricht man von einer **Konfiguration** f des Feldes. Die Konfiguration mit der größten Wahrscheinlichkeit¹ $P(f)$ stellt die gesuchte Lösung dar. Der zu untersuchende Lösungsraum $f \in \mathcal{F}$ wird gebildet aus der Menge aller möglichen Konfigurationen eines Random Field.

Die Bestimmung der Wahrscheinlichkeit $P(f)$ einer Konfiguration vereinfacht sich durch die Anwendung des **Theorems von Hammersley und Clifford** (siehe dazu auch Ge-man [29]). Dieses Theorem beschreibt die Äquivalenz von **Markov Random Fields** und **Gibbs-Verteilungen**: Die Bestimmung der Wahrscheinlichkeit der Konfiguration geschieht mittels der Auswertung einer Energiefunktion $U(f)$. Die Berechnung dieser Energiefunktion wird mit Hilfe von Cliques c durchgeführt: Cliques sind gegenseitig benachbarte Knoten eines ungerichteten Graphen. Sie stellen Teilmengen der Entitäten eines Nachbarschaftssystems dar. Für diese Cliques werden Potentialfunktionen V_c definiert,

¹Die Wahrscheinlichkeit einer Konfiguration $P(f)$ wird in der Literatur auch oft als Verbundwahrscheinlichkeit bezeichnet.

die **Clique-Potentiale**. Durch das Design dieser Clique-Potentiale ist es möglich, a priori Wissen in ein Markov Random Field einzubringen. Die Summierung aller Clique-Potentiale ergibt die Energiefunktion.

Die gesuchte Konfiguration des Random Field ergibt sich aus der Betrachtung der a posteriori Wahrscheinlichkeiten des Random Field. Dabei gilt folgender Zusammenhang: Die maximale a posteriori Wahrscheinlichkeit (MAP) entspricht der Minimierung der Energiefunktion $U(f)$.

Die schematische Vorgehensweise für eine gegebene Aufgabenstellung aus der Bildverar-

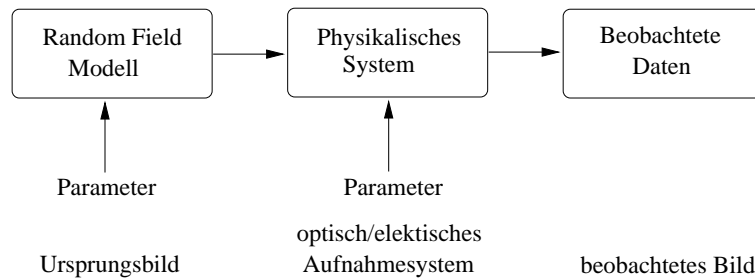


Abbildung 5.1: Schematische Vorgehensweise nach Bouman [2]

beitung ist in Abbildung 5.1 gezeigt: In der Abbildung besteht diese Aufgabenstellung z.B. in der Restauration verrauschter Bilddaten. Ein beobachtetes Bild entspricht nicht dem ursprünglichen Zustand der Szene. Die beobachteten Daten wurden in diesem Fall durch physikalische Randbedingungen, wie z.B. die Kameraparameter verfälscht. Geman [29] beschreibt diesen Prozeß durch eine Überlagerung von Rauschen, Verzerrungen, Transformation und einer invertierbaren mathematischen Operation, wie z.B. einer Addition.

Das gesuchte Ursprungsbild wird durch ein Markov Random Field Modell modelliert. Das Verhalten der Entitäten des Nachbarschaftssystems ist durch das Design der Clique-Potentiale festgelegt. Aus den beobachteten Daten kann die a posteriori Wahrscheinlichkeit für eine bestimmte Konfiguration errechnet werden. Ein Optimierungsprozeß bestimmt aus der Menge aller möglichen Konfigurationen die Konfiguration, für welche die a posteriori Wahrscheinlichkeit ein Maximum besitzt: die sogenannte **MAP-Schätzung**.

Die Anwendungen für Markov Random Fields erstrecken sich über alle Bereiche der digitalen Bildverarbeitung (siehe dazu auch Bouman [2] und Li [14]). Der Großteil der Realisationen entfällt dabei in den Bereich der niederen Bildverarbeitung. Die Themen der Arbeiten sind Bildrestauration, Kantenverfolgung oder Textursegmentation (siehe dazu auch [34], [51], [3], [28], [62], [29]).

Im Bereich der mittleren Bildverarbeitung können Markov Random Fields z.B. für das

perzeptive Gruppieren von Kontursegmenten eingesetzt werden (s.a. Kapitel 6).

Die höhere Bildverarbeitung verwendet Markov Random Fields in den Bereichen Objekt-Matching, Objekterkennung und Lagebestimmung (siehe dazu auch [14], [62], [60]).

In diesem Abschnitt wurde die Beschreibung von Systemen im Bereich der digitalen Bildverarbeitung durch den Formalismus von Markov Random Fields einführend vorgestellt. Die hier besprochenen Begriffe werden in den folgenden Abschnitten weiter vertieft.

5.2 Das Ising Modell

Im vorangehenden Abschnitt wurde die Beschreibung von Systemen im Bereich der digitalen Bildverarbeitung mit Hilfe von Markov Random Fields eingeführt. Die Grundlage zur Beschreibung solcher Systeme beruht unter anderem auf den Forschungen des Physikers Ernst Ising. Dieser versuchte empirische Beobachtungen an ferromagnetischen Materialien durch ein Modell auszudrücken. In der einfachsten Form geschieht dies mit einer eindimensionalen Folge von Punkten: Diese symbolisieren beispielweise Dipole, welche wie in Abbildung 5.2, jeweils einen - von zwei möglichen - Zuständen annehmen können.

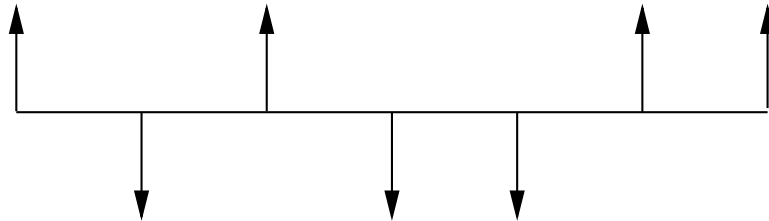


Abbildung 5.2: Anordnung der Dipole nach Kindermann und Snell [50]

Jeder der Dipole $i \in \mathcal{S}$ ist mit einer Zufallsvariable F_i verbunden. \mathcal{S} ist die Menge der vorhandenen Entitäten, hier der Dipole. Die Menge aller Zufallsvariablen F bildet ein eindimensionales Random Field.

Hat jeder Dipol, wie in Abbildung 5.2, einen Zustand eingenommen, stellt dies eine Realisation ($F = f$) der Menge der Zufallsvariablen, bzw. eine Konfiguration des Random Field dar.

Ising ordnete jeder Konfiguration ein spezifisches Energiemaß $U(f)$ zu. Dieses berechnet sich aus der Summierung der Interaktion von jeweils paarweise benachbarten Dipolen $i \in \mathcal{S}, j \in \mathcal{S}$, sowie der Wirkung eines externen Magnetfeldes auf jeden Dipol i . Für jeden Dipol i, j ist eine Funktion, der **Spin**, auszuwerten: ist der Spin am Dipol i positiv, gilt

$f_i = 1$. Für einen negativen Spin gilt entsprechend $f_i = 0$:

$$U(f) = -J \sum_{i,j} f_i f_j - mH \sum_i f_i \quad (5.1)$$

J und m sind Materialkonstanten und H steht für das externe Magnetfeld.

Ising wies den Konfigurationen f Wahrscheinlichkeiten zu, die dem Ausdruck $e^{-\frac{1}{T}U(f)}$ proportional sind und fand damit ein Wahrscheinlichkeitsmaß für die Konfigurationen f :

$$P(f) = \frac{1}{Z} e^{-\frac{1}{T}U(f)} \quad (5.2)$$

Dieser Ausdruck für $P(f)$ ist als **Gibbs-Verteilung** bekannt. Z ist eine Normalisierungskonstante, die als **partition function** bezeichnet wird. Die **Temperatur** T ist ebenfalls eine Konstante. Die Erweiterung der eindimensionalen Punktfolge auf ein zweidimensionales, regelmäßiges Gitter ist als **Ising-Modell** oder auch **Auto-logistisches Modell** bekannt.

5.3 Das Labeling Problem

Für die Repräsentation aller relevanten Zustände eines Prozesses, ist die Definition einer geeigneten Labelmenge \mathcal{L} erforderlich. Für das im letzten Abschnitt vorgestellte Ising Modell entsprach die Labelmenge \mathcal{L} der Anzahl der Orientierungen (positiv, negativ) des Spin. Die Definition der Labelmenge erfolgt also unter Berücksichtigung der Aufgabenstellung und der erwarteten Resultate: Für das Beispiel einer Kantensegmentierung ist - analog zum Ising Modell - im einfachsten Fall eine zweielementige Labelmenge $\mathcal{L} = \{Kante, \neg Kante\}$ zuzuordnen. Den Kanten eines Bildes können jedoch sehr unterschiedliche Ursachen zugrunde liegen. Sollen die Label auch diese Information abbilden, muß die Definition einer geeigneten Labelmenge entsprechend verändert werden. Je nach Anwendungsfall können Label symbolische oder numerische Werte abbilden. Dabei kann ein Label sowohl einwertig als auch vektorwertig definiert sein.

Ist jeder Entität ein Label zugeordnet, wird es möglich, die Entitäten des Feldes im Kontext ihrer jeweiligen Nachbarn zu betrachten. Wie solche Nachbarschaften entsprechend der Aufgabenstellung gewählt und ausgewertet werden, wird das Thema des folgenden Abschnitts sein.

5.4 Die Markov Eigenschaft: Nachbarn und Cliques

In den vorangehenden Abschnitten wurde eine sehr wichtige Eigenschaft von Markov Random Field bereits mehrfach erwähnt: die Möglichkeit zur Modellierung von Wissen über lokale Relationen. Die Begriffe Nachbarschaft und Clique sind in diesem Zusammenhang von großer Bedeutung. Eine **Nachbarschaft** \mathcal{N}_i verbindet eine Anzahl von Entitäten i des Random Field zu Nachbarn mit Hilfe einer Nachbarschaftsrelation. Diese Entitäten entsprechen z.B. den Pixeln einer Bildmatrix. Die Nachbarschaftsrelation verbindet in diesem Fall beispielsweise alle Pixel in der näheren Umgebung eines betrachteten Pixels aufgrund eines räumlichen Abstandsmaßes. Alle durch die gewählte Nachbarschaftsrelation verbundenen Entitäten bilden ein **Nachbarschaftssystem** \mathcal{N} bezüglich der betrachteten Entität i :

$$\mathcal{N} = \{\mathcal{N}_i \mid \forall i \in \mathcal{S}\} \quad (5.3)$$

Ein Nachbarschaftssystem ermöglicht die Betrachtung einer Entität im Kontext einer der Problemstellung angemessenen Anzahl von Entitäten.

Damit die Verteilung der Zufallsvariablen F auf den Entitäten $i \in \mathcal{S}$ - unter Berücksichtigung des Nachbarschaftssystems \mathcal{N} - ein Markov Random Field genannt werden kann, müssen die folgenden Bedingungen gelten:

$$P(f) > 0, \quad \forall f \in \mathcal{F} \quad (5.4)$$

$$P(f_i \mid f_{\mathcal{S}-\{i\}}) = P(f_i \mid f_{\mathcal{N}_i}) \quad (5.5)$$

Die in Gleichung 5.4 geforderte **Positivität** der Wahrscheinlichkeit der Konfiguration, ist eine technische Randbedingung, die nach Li [14] in der Regel immer erfüllt ist.

Die Wahrscheinlichkeit der Zuordnung eines Label f_i für i unter Berücksichtigung der Label aller anderen Entitäten $\mathcal{S} - \{i\}$ entspricht der Betrachtung der Label der Entitäten des Nachbarschaftssystems $f_{\mathcal{N}_i}$. Gleichung 5.5 wird deshalb auch die **lokale Charakteristik** eines Markov Random Field genannt.

Je nach Art der untersuchten Entitäten können Nachbarschaftssysteme ein sehr unterschiedliches Aussehen besitzen. Zwei wesentliche Eigenschaften sind jedoch allen gemeinsam:

1. Eine Entität ist kein Mitglied der eigenen Nachbarschaft (**Irreflexivität**):

$$i \notin \mathcal{N}_i \quad (5.6)$$

2. Gehört eine Entität i der Nachbarschaft der Entität i' an, ist i' Bestandteil der Nachbarschaft von i (**Symmetrie**):

$$i \in \mathcal{N}_{i'} \iff i' \in \mathcal{N}_i \quad (5.7)$$

Das Nachbarschaftssystem \mathcal{N} und die Entitäten \mathcal{S} konstituieren einen ungerichteten Graphen \mathcal{G} . Die Knoten des Graphen entsprechen den Entitäten und die Kanten der Nachbarschaftsrelation. Innerhalb eines Nachbarschaftssystems existieren Cliques c unterschiedlicher Komplexität. Eine **Clique** ist eine Teilmenge der Entitäten des ungerichteten Graphen, für die gilt, daß alle Entitäten jeweils paarweise benachbart sein müssen. Dabei ist die einfachste Form einer Clique die betrachtete Entität selbst.

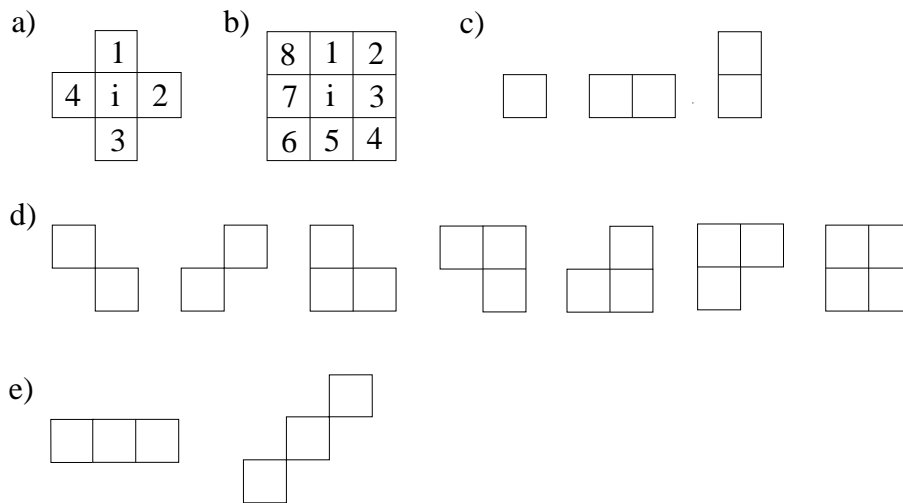


Abbildung 5.3: Nachbarschaftssysteme und Cliquesbildung für reguläre Entitäten nach Li [14].

- a) 4-er Nachbarschaft, b) 8-er Nachbarschaft, c) Formen der Cliques für die 4-er Nachbarschaft,
 d) zusätzliche Formen der Cliques für die 8-er Nachbarschaft,
 e) Gegenbeispiele: beide Formationen sind keine Cliques, da die jeweils äußeren Entitäten nicht benachbart sind.

Korrespondieren die betrachteten Entitäten mit den Pixeln der Bildmatrix, entspricht die Anordnung der Entitäten einem regelmäßigen Gitter (**reguläre Entitäten**). Für reguläre Entitäten faßt eine Nachbarschaftsrelation alle Nachbarn i' der betrachteten Entität i aufgrund der Auswertung der euklidischen Distanz und unter Berücksichtigung eines Radius r , zu einer Nachbarschaft zusammen:

$$\mathcal{N}_i = \{i' \in \mathcal{S} \mid [dist(Pixel_{i'}, Pixel_i)]^2 \leq r, i' \neq i\} \quad (5.8)$$

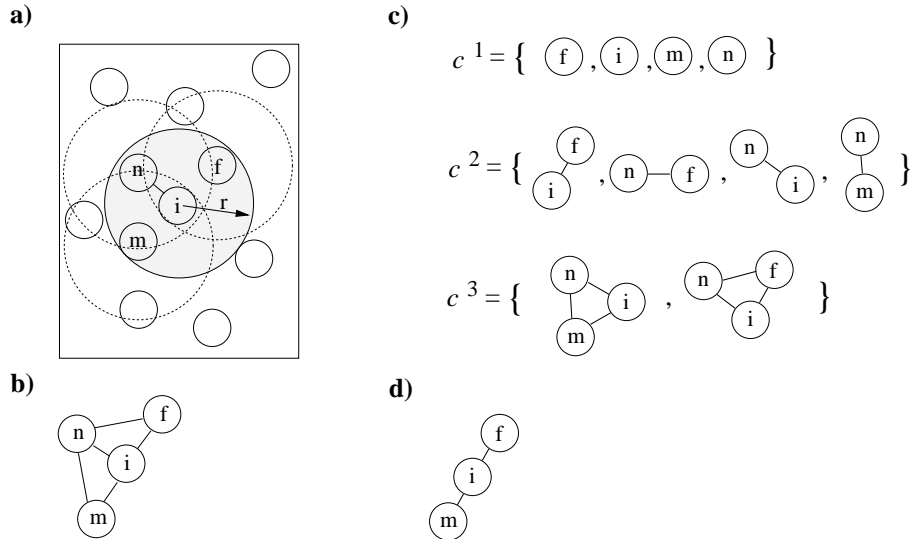


Abbildung 5.4: Nachbarschaftssystem und Cliquesbildung für irreguläre Entitäten nach Li [14]

- a) Nachbarschaftsrelation für die Abbildung der Bildprimitiva, in die Knoten des ungerichteten Graphen.
- b) Nachbarschaftssystem für den Knoten i .
- c) Die ein-, zwei- und dreielementigen Cliques des Nachbarschaftssystems.
- d) Gegenbeispiel: die Knoten $\{m, i, f\}$ bilden keine dreielementige Clique, da $\{m, f\}$ nicht benachbart sind.

Abbildung 5.3 zeigt die beiden kleinstmöglichen Nachbarschaften für ein Pixelraster: die sogenannte 4-er und 8-er Nachbarschaft. Die Formen aller möglichen Cliquesbildungen sind ebenfalls gezeigt.

Eine irreguläre Anordnung von Entitäten (**irreguläre Entitäten**) entsteht, wenn Bildmerkmale wie Eckpunkte oder Primitiva auf diese abgebildet werden. Bei den Primitiva kann es sich z.B. um die detektierten Geradenstücke einer Szene handeln. Die Bestimmung der Nachbarn \mathcal{N}_i der betrachteten Entität i kann analog zu Gleichung 5.8 mit dem euklidischen Abstand unter den Bildmerkmalen vorgenommen werden:

$$\mathcal{N}_i = \{i' \in \mathcal{S} \mid [dist(Feature_{i'}, Feature_i)]^2 \leq r, i' \neq i\} \quad (5.9)$$

Für die erwähnten Geradenstücke kann ein solches Abstandmaß z.B. über die Länge oder die Orientierung der Segmente gebildet werden: alle gleichlangen oder gleichorientierten Segmente sind Nachbarn. Abbildung 5.4 zeigt in Anlehnung an Li [14] die Organisation eines solchen Nachbarschaftssystems und die Formation der Cliques. Wie in der Abbildung

gezeigt, können Cliques anhand der Anzahl der Entitäten die sie umfassen typisiert werden in: einelementige Cliques c^1 , zweielementige Cliques c^2 usw.

In diesem Abschnitt wurden zwei sehr wichtige Begriffe in der Betrachtung von Markov Random Fields eingeführt: **Nachbarn** und **Cliques**. Die in Gleichung 5.5 vorgestellte **lokale Charakteristik** setzt die Betrachtung der Entitäten des Nachbarschaftssystems mit der Betrachtung aller Entitäten des Feldes gleich. Der folgende Abschnitt wird sich damit auseinandersetzen, wie aus den bedingten Wahrscheinlichkeiten der lokalen Charakteristik in Gleichung 5.5 ein Ausdruck für die Wahrscheinlichkeit einer Konfiguration $P(f)$ abgeleitet werden kann. Die in diesem Abschnitt eingeführten Cliques, werden in diesem Zusammenhang noch einmal näher betrachtet.

5.5 Gibbs-Markov-Äquivalenz

Im vorangehenden Abschnitt wurde eine grundlegende Forderung für Markov Random Fields vorgestellt, die sogenannte **lokale Charakteristik**. Sie beschreibt mittels einer bedingten Wahrscheinlichkeit, wie das Label f_i einer Entität i nur vom Zustand der Entitäten des Nachbarschaftssystems $f_{\mathcal{N}_i}$ abhängt. Für eine globale Betrachtung der Konfiguration des Feldes ist die Bestimmung der Wahrscheinlichkeit der Konfiguration $P(f)$ erforderlich. Die Ableitung der Wahrscheinlichkeit einer Konfiguration aus den bedingten Wahrscheinlichkeiten der Nachbarschaftssysteme ist jedoch schwierig, für ihre Berechnung bedient man sich deshalb in der Regel des **Theorems von Hammersley und Clifford**. Dieses Theorem beschreibt die Äquivalenz von **Gibbs-Verteilungen und Markov Random Fields**: Ein Satz von Zufallsvariablen F ist ein Gibbs Random Field (GRF) auf \mathcal{S} und \mathcal{N} , genau dann, wenn seine Konfigurationen f einer Gibbs-Verteilung gehorchen (siehe dazu auch Li [14]).

Wie in der Vorstellung des Ising-Modells in Abschnitt 5.2 beschrieben, wird die Wahrscheinlichkeit der Konfiguration $P(F = f)$ mit der Wahrscheinlichkeitsverteilung einer Gibbs-Verteilung gleichgesetzt: $P(f) = \frac{1}{Z} e^{-\frac{1}{T} U(f)}$ (siehe dazu Gleichung 5.2).

Die Energie der Konfiguration $U(f)$ wird aus der Summe der Clique-Potentiale V_C aller Cliques $c \in C$, mit C die Menge aller Cliques, berechnet:

$$U(f) = \sum_{c \in C} V_c(f) \quad (5.10)$$

Der praktische Wert des Theorems liegt in der Spezifizierung der Wahrscheinlichkeit der Konfiguration $P(f)$ durch das Design der **Clique-Potentiale** V_C . Diese Potentialfunktionen bilden den lokalen Charakter eines MRF auf das Gibbs Random Field ab und legen

dabei gleichzeitig das gewünschte Systemverhalten fest. Dieses Design der verschiedenen Potentialfunktionen bedeutet die Einbringung von a priori Wissen in das Verhalten des Random Field. Die Interaktion der gelabelten Entitäten wird festgelegt und damit bestimmt das Design der Potentialfunktionen die Form der Verteilungsfunktion.

Für die Spezifizierung einer Gibbs Verteilung auf ein Random Field ist es von Vorteil, wenn dieses über die Eigenschaften der **Homogenität** und **Isotropie** verfügt. Homogenität bedeutet, daß die Definition der bedingten Wahrscheinlichkeit $P(f_i | f_{N_i})$ bzw. des Clique-Potentials $V_c(f)$, unabhängig von der Lage der Entität, bzw. Clique in \mathcal{S} ist. Eine Isotropie des betrachteten Feldes liegt vor, wenn das berechnete Potential für eine Clique c , unabhängig von deren Orientierung ist.

Die genaue Bestimmung der Wahrscheinlichkeit der Konfiguration $P(f)$ für ein MRF erfordert nach Gleichung 5.2 die Berechnung der partition function Z . Dies bedeutet eine Summierung über alle Konfigurationen $f \in \mathcal{F}$. Da diese Form generell sehr aufwendig ist, macht man sich zu Nutze, daß Z eine Konstante ist: Ist nur die Bestimmung des Maximums, bzw. Minimums von Interesse, kann sie vernachlässigt werden, da sie lediglich als proportionaler Faktor in die Berechnung eingeht. Die Suche nach der optimalen Konfiguration des Feldes beruht daher in der Regel auf einer vergleichenden Betrachtung der Wahrscheinlichkeiten der Konfigurationen des MRF. Die Bestimmung der gesuchten Konfiguration des Feldes mit Hilfe eines Optimierungsprozesses, wird das Thema des folgenden Abschnitts sein.

5.6 Optimierung

Wie im vorausgehenden Abschnitt dargestellt, ermöglicht das Theorem von Hammersley und Clifford die Berechnung der Wahrscheinlichkeit der Konfiguration $P(f)$ eines Markov Random Field. Dieses Random Field modelliert das unbekannte Ursprungsbild. Die verfügbare Information über das Ursprungsbild sind die Daten o des beobachteten Bildes. Diese Daten sind jedoch z.B. durch ein physikalisches System verfälscht. Aus diesen Beobachtungen gilt es nun, auf den ursprünglichen Zustand zu schließen. Das beobachtete Bild kann ebenfalls als ein Markov Random Field modelliert werden, für das analog die Wahrscheinlichkeit der Konfiguration $P(f)$ mit Hilfe des Theorems von Hammersley und Clifford ermittelbar ist. Chou und Brown [28] bezeichnen dieses Markov Random Field mit **a posteriori Markov Random Field**, da es das beobachtete Ergebnis beschreibt. Die zugrundeliegende "wahre" Konfiguration ist $f_o \in \mathcal{F}$. Diese Konfiguration soll durch eine Konfiguration $f_o^* \in \mathcal{F}$, welche das gesuchte Ursprungsbild optimal beschreibt, an-

genähert werden. Für die Suche nach der Konfiguration f_o^* wird angenommen, daß die beobachteten Daten o das Ergebnis eines statistischen Prozesses sind. Auf Grundlage der beobachteten Daten o , wird die Entscheidung für eine Konfiguration $\hat{f}_o \in \mathcal{F}$ getroffen. Es werden Kosten K definiert, die als Entscheidungshilfe dienen: Die Kosten für falsche Entscheidungen sind hoch und Kosten für richtige Entscheidungen niedrig. Diese Kosten werden für jede Konstellation von $K(f_o, \hat{f}_o)$ berechnet.

Da die beobachteten Daten o Ergebnis eines statischen Prozesses sind, hängen die Kosten nicht nur von der tatsächlichen Konfiguration f_o ab. Aus diesem Grund wird der Erwartungswert der Kosten $\mathcal{E}\{K\}$ optimiert. Dabei orientiert sich der entscheidungstheoretische Lösungsansatz an einer Strategie, welche im Mittel die Kosten der getroffenen Entscheidungen minimiert. Mathematisch wird dies durch die folgende Gleichung beschrieben:

$$\mathcal{R} = \mathcal{E}\{K\} = \sum_o \sum_{f_o} K(f_o, \hat{f}_o) P(o, f_o) \quad (5.11)$$

Der Erwartungswert der Kosten $\mathcal{E}\{K\}$ wird auch als Risiko \mathcal{R} bezeichnet. Die Kostenfunktion $K(f_o, \hat{f}_o)$ hat die typischerweise die folgende Form:

$$K(f_o, \hat{f}_o) = \begin{cases} const & : \hat{f}_o \neq f_o \text{ mit } const > 0 \\ 0 & : \hat{f}_o = f_o \end{cases} \quad (5.12)$$

mit ein paar einfachen Umformungen, wie sie Schürmann [22] beschreibt, erhält man für Gleichung 5.11 den Ausdruck:

$$\mathcal{R} = \sum_o \mathcal{R}_o(\hat{f}_o) P(o) \quad (5.13)$$

Der Ausdruck $\mathcal{R}_o(\hat{f}_o)$ wird als **bedingtes Risiko** bezeichnet. Das bedingte Risiko wird durch die folgende Definition beschrieben:

$$\mathcal{R}_o(\hat{f}_o) := \sum_{f_o \in \mathcal{F}} K(f_o, \hat{f}_o) P(f_o | o) \quad (5.14)$$

Mit Hilfe der Definition des bedingten Risikos in Gleichung 5.14, wird die Minimierung des Risikos \mathcal{R} gleichwertig durch eine Minimierung des bedingten Risikos in Gleichung 5.13 für alle Beobachtungen o beschrieben:

$$f_o^* := \min_{\hat{f}_o \in \mathcal{F}} \mathcal{R}_o(\hat{f}_o) \quad (5.15)$$

Wird die Kostenfunktion folgendermaßen gewählt:

$$K_{Bayes} = \begin{cases} 1 & : \hat{f}_o \neq f_o \\ 0 & : \hat{f}_o = f_o \end{cases} \quad (5.16)$$

dann gilt für das bedingte Risiko:

$$\mathcal{R}_o(\hat{f}_o) = 1 - P(\hat{f}_o | o) \quad (5.17)$$

Mit Hilfe von Gleichung 5.17 folgt für die optimale Schätzung die Identität mit der maximalen a posteriori Wahrscheinlichkeit $P(f | o)$:

$$f_o^* := \arg \max_{f \in \mathcal{F}} P(f | o) \quad (5.18)$$

Gleichung 5.18 wird deshalb auch mit **maximum a posteriori Schätzung** bezeichnet.

Bei einer geringen Anzahl von Konfigurationen f , kann die maximale a posteriori Wahrscheinlichkeit effizient durch Ausrechnen und Vergleichen der a posteriori Wahrscheinlichkeiten für alle $f \in \mathcal{F}$ erfolgen. Für die sehr großen Konfigurationsmengen jedoch, die für Markov Random Fields normal sind, ist dieses “brute-force” Vorgehen nicht praktikabel. Aus diesem Grund kommen für die Bestimmung der optimalen Schätzung f_o^* Suchverfahren zum Einsatz. Diese Suchverfahren sind das Thema des folgenden Abschnitts 5.7. An dieser Stelle erfolgt zunächst eine Konkretisierung der Gleichung 5.18, die für das Verständnis des folgenden Abschnitts von Bedeutung ist. Unter Vernachlässigung von $P(o) = \textit{konstant}$ wird Gleichung 5.18 mit Bayes folgendermaßen umgeformt:

$$f_o^* := \arg \max_{f \in \mathcal{F}} P(o | f)P(f) \quad (5.19)$$

Die Definition von $P(f)$ in Gleichung 5.19 ist durch die Anwendung des Theorems von Hammersley und Clifford aus Gleichung 5.2 bestimmt.

Für die Berechnung von $P(o | f)$ sind die bedingten Wahrscheinlichkeiten $P(o_i | f_i)$ der Entitäten $i \in \mathcal{S}$ des Feldes zu untersuchen. Sie beschreiben die Wahrscheinlichkeit der Beobachtung o_i an einer Entität i unter der Voraussetzung, daß ein bestimmtes Label f_i vorliegt. Unter der Voraussetzung, daß räumlich getrennte Beobachtungen unabhängig sind, gilt für das Produkt dieser bedingten Wahrscheinlichkeiten:

$$P(o | f) = \prod_{i \in \mathcal{S}} P(o_i | f_i) \quad (5.20)$$

Bouman [2] bezeichnet diesen Ausdruck auch als **Datenmodell**.

Unter Berücksichtigung der Gleichungen 5.2 und 5.20 gilt für die a posteriori Wahrscheinlichkeit:

$$\begin{aligned} P(f | o) &= \prod_{i \in \mathcal{S}} P(o_i | f_i) \frac{1}{Z} e^{-\frac{1}{T}U(f)} \\ &= e^{\ln[\prod_{i \in \mathcal{S}} P(o_i | f_i)]} \frac{1}{Z} e^{-\frac{1}{T}U(f)} \\ &= \frac{1}{Z} e^{\ln[\prod_{i \in \mathcal{S}} P(o_i | f_i)] - \frac{1}{T}U(f)} \end{aligned} \quad (5.21)$$

Der Ausdruck im Exponenten von Gleichung 5.21 wird in einen Energieterm $E(f)$ zusammengefaßt:

$$P(f | o) = \frac{1}{Z} e^{E(f)} \quad (5.22)$$

Da die Energie $U(f)$ aus der Aufsummierung von Cliquepotentialen $V_c(f)$ berechnet wird kann mit Gleichung 5.10 für den Energieterm $E(f)$ geschrieben werden:

$$E(f) = -\frac{1}{T} \left(\sum_{c \in \mathcal{C}} V_c(f) - T \sum_{i \in \mathcal{S}} \ln P(o_i | f_i) \right) \quad (5.23)$$

Die partition function Z ist konstant und wird daher für die Betrachtung des Maximums, bzw. Minimums vernachlässigt. Unter Berücksichtigung der Gleichungen 5.18, 5.22, 5.23 gilt für die optimale Schätzung:

$$f_o^* = \min_{f \in \mathcal{F}} [E(f)] \quad (5.24)$$

Die optimale Schätzung f_o^* entspricht der Konfiguration f , für die der Energieterm $E(f)$ sein Minimum einnimmt. Das Minimum des Energieterms entspricht gleichzeitig der maximalen a posteriori Schätzung.

5.7 Energieminimierung

Wie in den beiden letzten Abschnitten gezeigt, ist die Schätzung der Maximum a posteriori Wahrscheinlichkeit eines Markov Random Field unter Ausnutzung des Theorems von Hammersley und Clifford gleichwertig durch die Minimierung des Energieterms $E(f)$ ermittelbar. Die Bestimmung des globalen Minimums der Energiefunktion ist nicht trivial. Dies gilt im Besonderen, wenn die Energiefunktion viele lokale Minima besitzt. Es ist daher sinnvoll, die entstehenden Kosten für eine korrekte Bestimmung des globalen Minimums, im Verhältnis zu einer näherungsweise Bestimmung, für die jeweilige Aufgabenstellung abzuschätzen.

Es stehen verschiedene Verfahren zur Auswahl, die sich grob mit den Begriffen **lokale Suche** und **globale Suche** in zwei Gruppen einteilen lassen. Die Begriffe lokal und global beziehen sich in diesem Zusammenhang auf den Fokus mit dem das Energieminimum bestimmt wird. Li [14] definiert zu diesem Zweck ein Nachbarschaftssystem von Konfigurationen: $\mathcal{N} = \{\mathcal{N}(f) | f \in \mathcal{F}\}$, mit $\mathcal{N}(f)$ die benachbarten Konfigurationen von f im Lösungsraum \mathcal{F} . Für den Fall diskreter Label ist eine **n-Nachbarschaft** bestimmt durch:

$$\mathcal{N}^n(f) = \{x \in \mathcal{F} | x \text{ weicht in maximal } n \text{ Labelwerten von } f \text{ ab}\} \quad (5.25)$$

Für dieses Nachbarschaftssystem existiert eine Konfiguration f^* mit minimaler Energie, sodaß gilt:

$$E(f^*) \leq E(f) \quad \forall f \in \mathcal{N}(f^*) \quad (5.26)$$

$E(f^*)$ ist ein lokales Minimum, wenn die Konfigurationen des gewählten Nachbarschafts-systems eine Teilmenge des Lösungsraums der Konfigurationen \mathcal{F} darstellen und ein globales Minimum, wenn die Konfigurationen des Nachbarschafts-systems alle Konfigurationen $f \in \mathcal{F}$ des Lösungsraums enthalten.

Eine lokale Suche startet mit einer initialen Konfiguration $f \in \mathcal{F}$. Innerhalb der Nachbarschaft $\mathcal{N}(f)$ wird eine weitere Konfiguration f' gesucht, für die gilt $E(f') < E(f)$. In diesem Fall wird f durch f' ersetzt. Der beschriebene Vorgang wiederholt sich solange, bis keine Konfiguration mit einem niedrigeren Energieniveau mehr gefunden werden kann, also gilt: $E(f') \geq E(f)$ für alle $f' \in \mathcal{N}(f)$. Verfahren, die eine lokale Suche verwenden, haben den Vorteil deterministisch zu sein. Das ermittelte Energieminimum ist allerdings im allgemeinen nicht mit dem globalen Energieminimum identisch.

Ein typisches Vorgehen zur Bestimmung neuer Konfigurationen f' ist z.B. das Gradientenabstiegsverfahren. Die von mir verwendete Anwendung einer lokalen Suche, das **Highest Confidence First**, nach Chou & Brown (siehe dazu auch [62] und [28]) wird am Ende dieses Kapitels noch näher vorgestellt.

Typischer Vertreter einer globalen Suche ist das **Simulated Annealing**. Das Simulated Annealing ahmt den physikalischen Abkühlungsprozeß einer geschmolzenen Materie nach. Die Bestimmung neuer Konfigurationen wird mit einer stochastische Suche durchgeführt. Dies führt dazu, daß jede neue Konfiguration f' nicht automatisch eine niedrigere Energie besitzt, sondern auch Konfigurationen mit einer höheren Energie akzeptiert werden. Ein **Einschwingen** in lokale Minima soll verhindert werden. Diese stochastische Suche wird durch die **Temperatur** T gesteuert: Der Wert für T wird im Verlauf des Verfahrens vermindert. Für hohe Werte von T (Anfangszustand) werden hohe Schwankungen in den Energiewerten der Konfigurationen akzeptiert. Diese Schwankungsbreite vermindert sich mit der Senkung der Temperatur, bis im Nullpunkt keine Erhöhung der Energie mehr möglich ist. Die Schrittweite der Temperaturminderung (**Annealing Schedule**: siehe auch Geman [29]) prägt den Ablauf entscheidend: Ein schneller **Abkühlprozeß** birgt die Gefahr des Einschwingens in lokale Minima in sich. Ein langsames Abkühlen führt mit hoher Wahrscheinlichkeit auf ein globales Minimum, eine wesentliche Erhöhung der Rechenzeit muß jedoch gleichzeitig in Kauf genommen werden.

5.7.1 Energieminimierung mit Highest Confidence First

Verfahren zur Bestimmung des Energieminimums bedienen sich entweder einer **globalen** oder **lokalen Suchmethode**, wie sie prinzipiell im letzten Abschnitt vorgestellt wurden. Dieser Abschnitt befaßt sich jetzt näher mit einem speziellen Ansatz zur Energieminimierung mittels einer lokalen Suche, wie er in dem von mir verfolgten Ansatz zur Bewertung von Gruppenhypothesen zum Einsatz kommt. Dieses Verfahren ist das sogenannte **Highest Confidence First** (HCF) (siehe dazu auch [28], [62]).

Der Name des Verfahrens deutet die grundlegende Strategie an: Eine Entscheidung für eine Entität, z.B. über die Verifikation einer hypothetisierten Gruppe, wird auf Basis möglichst gesicherter Information getroffen. Ist eine Entscheidung zum aktuellen Zeitpunkt nicht möglich, wird sie vertagt bis aus der Untersuchung der Umgebung der betrachteten Entität mehr Information für eine Entscheidung verfügbar ist. Getroffene Entscheidungen können revidiert werden, wenn die Auswertung der Umgebung dazu Anlaß gibt. Dieses Prinzip ist in der Bildverarbeitung auch als **least commitment** bekannt. Diese Entscheidungsfindung realisieren Chou & Brown in ihrem Ansatz durch die Einführung eines speziellen Label l_{uc} , welches den unentschiedenen Zustand einer Entität repräsentiert.

In jedem Iterationsschritt f wird eine Entität i betrachtet und der Wert des mit dieser Entität verbundenen Label f_i gegebenenfalls geändert. Dies bedeutet, daß die nacheinander untersuchten Konfigurationen (vergleiche Gleichung 5.25) sich nur in einem Labelwert unterscheiden.

Die Auswahl der betrachteten Entitäten wird gesteuert durch ein Stabilitätsmaß G . Jeder Entität des Markov Random Field ist, in Abhängigkeit von der Konfiguration, ein **Stabilitätsmaß** zugeordnet: dieses gibt Auskunft über den möglichen Energiegewinn, wenn sich der Labelwert der Entität ändert.

5.7.1.1 Die erweiterte Label-Menge

Chou & Brown definieren für ihren Ansatz eine Menge diskreter Label $l \in \mathcal{L}$. Diese Labelmenge \mathcal{L} wird um ein zusätzliches Label l_{uc} erweitert, sodaß eine neue Labelmenge $\mathcal{L}^+ = \mathcal{L} \cup \{l_{uc}\}$ entsteht. Chou & Brown bezeichnen dieses Label auch als **Null-Label**, da es einen unentschiedenen (**uncommitted**) Zustand repräsentiert. Infolge der erweiterten Labelmenge \mathcal{L}^+ entsteht ein erweiterter Lösungsraum für Konfigurationen \mathcal{F}^+ .

Initial besitzen alle Entitäten des Random Field das Label l_{uc} . Chou & Brown beeinflussen den Ablauf ihres Verfahrens wesentlich mit einer Regel die besagt, daß jede Entität i welche vom Label **unentschieden** $f_i = l_{uc} \in \mathcal{L}^+$ in einen Belegungszustand $f_i = l \in \mathcal{L}$ gewechselt hat, nie wieder den ursprünglichen Zustand einnehmen kann. Eine Änderung des

aktuellen Labelwertes der Entität f_i ist jedoch immer möglich. Die grundlegende Idee ist, eine Abfolge von Konfigurationen zu durchlaufen, mit einer Startkonfiguration in der alle Knoten unentschieden sind, und einer Endkonfiguration in der alle Knoten eine Belegung (**committed**) besitzen. Diese Konfiguration besitzt im Idealfall das minimale Energiemaß in Bezug auf den Lösungsraum der Konfigurationen \mathcal{F} .

Der Verlauf des Verfahrens ist iterativ. Pro Iteration wird eine Entität i mit ihrem Nachbarschaftssystem \mathcal{N}_i betrachtet und der Entität ein neues Label zugeordnet.

5.7.1.2 Stabilität und lokale Energie

Die Auswahl der jeweils betrachteten Entität erfolgt aufgrund der Auswertung eines Stabilitätsmaßes G : Ausgewählt wird die Entität mit der geringsten Stabilität, denn diese verspricht den größten Energiegewinn für eine Neuberechnung des Labelwertes.

Die Berechnung der Stabilität beruht auf der Ermittlung der **lokalen Energie** der betrachteten Entität. Sei $U_i(f)$ die lokale Energie einer Entität i und C_i die Menge der Cliques der Nachbarschaft \mathcal{N}_i der Entität i , dann folgt daraus für die lokale Energie:

$$U_i(f) = \sum_{c \in C_i} V_c(f) \quad (5.27)$$

Die Berechnung der lokalen Energie berücksichtigt nur die Cliques der Nachbarschaft \mathcal{N}_i , die bereits einen Belegungszustand besitzen.

Für die Berechnung der Stabilität wird eine Fallunterscheidung vorgenommen:

- Die betrachtete Entität besitzt einen Belegungszustand $f_i = l \in \mathcal{L}$, d.h. $f_i \neq l_{uc}$.
- Die betrachtete Entität ist unentschieden, d.h. $f_i = l_{uc} \in \mathcal{L}^+$.

Die Stabilität einer Entität i berechnet sich aus der Energiedifferenz wie folgt:

Der Ausdruck $subst(f, i, l)$ bezeichne die Konfiguration, welche entsteht, wenn die i -te Komponente f_i von f durch l ersetzt wird. Die minimale Energie $U_i^{min}(f)$ einer Entität (bei unveränderten Labeln in ihrer Nachbarschaft) wird mit diesem Ausdruck folgendermaßen definiert:

$$U_i^{min}(f) = \min_{l \in \mathcal{L}} [U_i(subst(f, i, l))] \quad (5.28)$$

Das Label $l^* \in \mathcal{L}$ bezeichnet das Argument, welches die Funktion $U_i(subst(f, i, l))$ minimiert.

Mit den eingeführten Definitionen für $U_i^{min}(f)$ und l^* wird die Stabilität wie folgt definiert:

$$G_i(f) = \min_{l \in \mathcal{L}, l \neq f_i} [U_i(subst(f, i, l)) - U_i(f)] \quad , \quad \text{falls } f_i \in \mathcal{L} \quad (5.29)$$

$$G_i(f) = - \min_{l \in \mathcal{L}, l \neq l^*} [U_i(subst(f, i, l)) - U_i^{min}(f)] \quad , \quad \text{falls } f_i = l_{uc} \quad (5.30)$$

Für die Berechnung der lokalen Energie U_i einer Entität i leisten nur die Cliques der Nachbarschaft \mathcal{N}_i einen Beitrag, die sich bereits entschieden haben, d.h. deren Entitäten einen Labelwert $f_i \in \mathcal{L}$ besitzen. Die Stabilität $G_i(f)$ einer Entität i mit einem Belegungszustand $f_i \in \mathcal{L}$ nach Gleichung 5.29 wird negativ, wenn sich die lokale Energie U_i durch eine Änderung des Labelwertes $f_i \in \mathcal{L}$ mindern läßt. Befindet sich die untersuchte Entität in einem unentschiedenen Zustand, d.h. $f_i = l_{uc}$, entspricht die Stabilität $G_i(f)$ der Differenz aus der zweitniedrigsten und der niedrigsten möglichen lokalen Energie U_i . Diese Stabilität ist immer negativ, wie in Gleichung 5.30 definiert.

Der Ablauf des Algorithmus folgt einer iterativen Auswertung der lokalen Energien, gesteuert durch die Stabilität. Dies wird im folgenden Abschnitt vorgestellt.

5.7.1.3 Algorithmus

Der Ablauf des Highest Confidence First (siehe dazu auch Abbildung 5.5) wird durch eine Prioritätswarteschlange gesteuert. In dieser Schlange werden die Entitäten i gemäß ihren Stabilitäten $G_i(f)$ in aufsteigender Reihenfolge einsortiert. Der Algorithmus startet in einer initialen Konfiguration f in der alle Entitäten in einem unentschiedenen Zustand $f_i = l_{uc} \in \mathcal{L}^+$ sind. Für jede Entität wird ein initialer Labelwert f_i^o ermittelt. Durch Negation dieses initialen Labels wird für jede Entität eine initiale Stabilität berechnet: $-f_i^o = G_i^{init}$. In der im folgenden Kapitel 6 vorgestellten Anwendung entspricht dieses initial ermittelte Label der Signifikanz, die aus den zugrunde liegenden Daten einer Gruppenthese ermittelt wird. Im Rahmen der Anwendung wird deshalb dieses initiale Label als **datengetriebenes** Label bezeichnet.

An der Spitze der Prioritätswarteschlange steht die Entität mit der jeweils niedrigsten Stabilität G_{top} . Der Labelwert f_i der Entität mit der Stabilität G_{top} wird untersucht und erhält mit **Change_State**(f_i) einen neuen Wert zugeordnet. Dann wird für die betrachtete Entität i durch **Update_G**($G_i(f)$) eine neue Stabilität berechnet. **Adjust_Heap**(i) sortiert die Entität entsprechend der neu berechnete Stabilität in die Prioritätswarteschlange ein. Im Anschluß daran muß die Stabilität aller Entitäten r der Nachbarschaft $r \in \mathcal{N}_i$ neu berechnet und diese ebenfalls einsortiert werden. Solange eine Stabilität $G_i(f)$ existiert, die einen Wert besitzt, welcher kleiner als Null ist, kann die globale Energie des Feldes vermindert werden. Andernfalls bricht der Algorithmus ab und die ermittelte Konfiguration stellt das Ergebnis der Energieminimierung dar.

Die Zuweisung der initialen Stabilitäten G_i^{init} bestimmt wesentlich den qualitativen Ablauf des Verfahrens. Entitäten mit einem hohen initialen Labelwert besitzen gleichzeitig eine sehr niedrige Stabilität und werden deshalb vorrangig untersucht. Für die im folgenden

```
 $f = ( 0, \dots, 0 );$   
 $top = \text{Create\_Heap}( f );$   
while (  $G_{top} < 0$  ) {  
     $i = top;$   
     $\text{Change\_State}( f_i );$   
     $\text{Update\_G}( G_i(f) );$   
     $\text{Adjust\_Heap}( i );$   
    foreach(  $r \in \mathcal{N}_i$  ) {  
         $\text{Update\_G } G_r(f);$   
         $\text{Adjust\_Heap}( r );$   
    }  
}  
return(  $f$  );
```

Abbildung 5.5: Pseudocode des HCF-Algorithmus nach [28]

Kapitel 6 vorgestellte Anwendung bedeutet dies, daß Entitäten, welche eine sehr hohe Unterstützung durch ihre zugrunde liegenden Daten besitzen, sich zuerst für die Belegung mit einem Labelwert f_i entscheiden. Diese Entitäten unterstützen im weiteren Verlauf wiederum Entitäten mit niedrigeren datengetriebenen Labels in der Entscheidung für einen Labelwert.

Durch die stetige Neuberechnung der Stabilitäten von untersuchter Entität und den Entitäten der Nachbarschaft kann eine Entität ihre Entscheidung für einen Labelwert im weiteren Verlauf des Algorithmus verändern. Dies trägt der zunehmenden Kontextinformation durch Entitäten mit einem Belegungszustand Rechnung.

Zusammenfassung In diesem Kapitel wurde der Einsatz von **Markov Random Fields** für typische Fragestellungen der Bildverarbeitung dargestellt. Diesen Fragestellungen liegt in der Regel eine räumlich lokale Struktur zugrunde, wie z.B. in der Restauration veräuschter Bilder. Markov Random Fields sind in diesem Rahmen von besonderem Interesse, da sie zum einen die Modellierung von Wissen über lokale Relationen erlauben. Zum anderen eine global optimale Beschreibung der Szene, durch eine geeignete Untersuchung der verschiedenen Konfigurationen des Feldes, ermöglichen. Die Modellierung der lokalen Relationen erfolgt auf der Grundlage der in Abschnitt 5.4 vorgestellten Nachbar-

schaftssysteme und Cliques. Jede Entität des Random Field ist mit einer Zufallsvariablen verbunden. Auf diese Zufallsvariable wird unter Berücksichtigung des zugeordneten Nachbarschaftssystems ein Wert abgebildet. Dieser Wert kann in Form eines in Abschnitt 5.3 eingeführten **Label** aus einer **Labelmenge** vergeben werden. Die Labelmenge wird so gewählt, daß sie alle relevanten Zustände des untersuchten Prozesses repräsentiert.

Eine **Konfiguration** des Feldes liegt vor, wenn allen Zufallsvariablen einen Zustand eingenommen haben. Die Bestimmung der Wahrscheinlichkeit einer Konfiguration wird durch das in Abschnitt 5.5 beschriebene **Theorem von Hammersley und Clifford** vereinfacht. Dieses Theorem beschreibt die Äquivalenz von **Gibbs-Verteilungen** und Markov Random Fields. Der praktische Wert des Theorems liegt in der Spezifizierung der Wahrscheinlichkeit der Konfiguration $P(f)$ durch das Design der **Clique-Potentiale**. Diese Potentialfunktionen bilden den lokalen Charakter eines MRF auf das Gibbs Random Field ab und legen dabei gleichzeitig das gewünschte Systemverhalten fest.

Die Ermittlung der Konfiguration mit der maximalen a posteriori Wahrscheinlichkeit entspricht, wie in Abschnitt 5.6 behandelt, der Minimierung einer Energiefunktion. Die Minimierung der Energiefunktion ist allgemein eine sehr komplexe Aufgabenstellung, da sie die Interaktion der Label aller Entitäten über die Menge aller möglichen Konfiguration berücksichtigen muß. Je nach der Aufgabenstellung kommen deshalb unterschiedliche Verfahren zum Einsatz, die sich durch die in Abschnitt 5.7 eingeführten Begriffe der **lokalen** und **globalen Suche** grob klassifizieren lassen. Lokale Verfahren haben den Vorteil deterministisch zu sein, dafür entspricht das ermittelte Minimum im allgemeinen nicht dem globalen Minimum. Typische Vertreter beider Ansätze wurden im Abschnitt 5.7 vorgestellt. Abschließend wurde in Abschnitt 5.7.1 das von mir verwendete **Highest Confidence First** nach Chou & Brown [28] beschrieben.

Nachdem in diesem Kapitel die typischen Anwendungsgebiete für Markov Random Fields in der Bildverarbeitung, sowie die grundlegenden Begriffe dargestellt wurden, folgt im nächsten Kapitel 6 die eingehende Beschreibung des von mir gewählten Ansatzes für die Bewertung der Gruppenhypothesen.

6 Ein Markov Random Field für Gruppenhypothesen

In Kapitel 4 wurde ein neuartiger Ansatz zur Generierung von Gruppenhypothesen mit Bereichen perzeptiver Aufmerksamkeit vorgestellt. Die Erzeugung der Gruppenhypothesen erfolgt durch die Auswertung dieser Bereiche perzeptiver Aufmerksamkeit auf einer lokalen Basis. Als Folge dieser lokale Betrachtungsweise entstehen mehrdeutige Gruppenhypothesen. Diese Mehrdeutigkeit bedeutet eine unerwünschte inkonsistente Interpretation der Szene, die eine Verifikation der Gruppenhypothesen erfordert. Aus diesem Grund schließt sich an den Vorgang der Erzeugung der Gruppenhypothesen eine Bewertung dieser Gruppenhypothesen im globalen Rahmen der betrachteten Szene an. Gleichzeitig erlaubt diese Bewertung eine Aussage über die Bedeutung der einzelnen Hypothesen im Szenenkontext und kann so helfen ihre Anzahl zu verringern. Diese Verringerung ist aufgrund der in der Regel sehr großen Anzahl der Hypothesen für ihre weitere Betrachtung durch nachfolgende Verfahren z.B. zur Objekterkennung wichtig.

Der hier vorgestellte Ansatz zur Bewertung der Gruppenhypothesen stützt sich auf den Formalismus von Markov Random Fields, wie er in Kapitel 5 eingeführt wurde. Dieses Kapitel wird sich schwerpunktmäßig mit den folgenden Themen beschäftigen:

- Motivation für die Verifikation der generierten Gruppenhypothesen durch einen geeigneten Bewertungsmechanismus.
- Abbildung der Gruppenhypothesen in einen **ungerichteten Graphen** und Formulierung einer geeigneten **Nachbarschaftsrelation** für Gruppenhypothesen.
- Verfahren für die Bestimmung **datengetriebener Label** der unterschiedlichen Gruppentypen, welche eine geeignete Aussage über die initiale Konfiguration der beobachteten Daten erlauben.
- Das Design geeigneter **Clique-Potentiale** für die Berechnung der Energiefunktion.

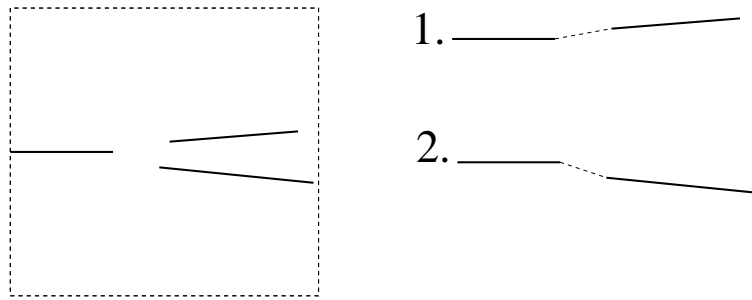


Abbildung 6.1: Beispiel einer mehrdeutigen Gruppenthypothese:

Für 3 Geradenstücke werden 2 kollineare Gruppen hypothetisiert.

6.1 Mehrdeutigkeit der Gruppenthypothesen

Mehrdeutigkeiten sind in der lokalen Betrachtungsebene der Bereiche perzeptiver Aufmerksamkeit nicht auflösbar, weil die Verifikation einer Gruppenthypothese die Einbeziehung globaler Zusammenhänge erfordert. Ein kleines Beispiel in Abbildung 6.1 soll das Anliegen verdeutlichen: es werden gleichzeitig zwei Kollinearitäten hypothetisiert denen ein Segment gemeinsam ist. Für den betrachteten Ausschnitt gilt, daß die erzeugten Gruppen für den Betrachter intuitiv gleichrangig sind. In Abbildung 6.2 sind diese Gruppen in einem größerem Zusammenhang dargestellt. Die szenenbestimmende Struktur ist jetzt mühelos erkennbar. Der globale Zusammenhang ist für einen menschlichen Betrachter im Sinne des Wortes "augenfällig". In der Bildverarbeitung muß jedoch zuerst ein "Kontext" geschaffen werden, der es ermöglicht, die lokal hypothetisierten Gruppen in einem globalen Zusammenhang zu bewerten. Wie ein solcher Kontext mit Hilfe von Markov Random Fields formuliert werden kann, wird im folgenden Abschnitt einführend vorgestellt.

6.2 Modellierung von Kontextwissen in einem Markov Random Field

Wie im vorausgehenden Abschnitt herausgearbeitet, ist für die Bestimmung signifikanter Strukturen einer Szene die Betrachtung der Gruppenthypothesen in einem globalen Kontext erforderlich. Die Verifikation der Gruppenthypothesen mittels eines geeigneten Bewertungsmechanismus muß gleichzeitig die lokale Interaktion der benachbarten Gruppenthypothesen, sowie die Ermittlung einer global optimalen Interpretation berücksichtigen. Als Formalismus wurde ein Markov Random Field gewählt (siehe dazu auch [21]), weil es gut

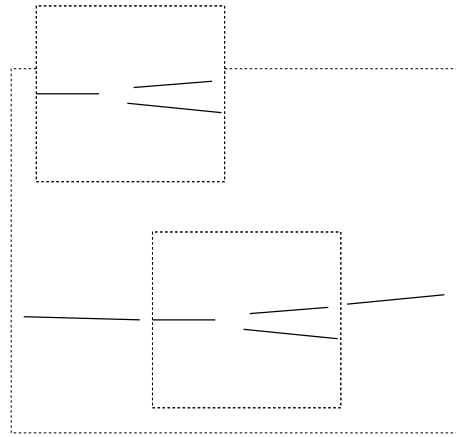


Abbildung 6.2: Mehrdeutigkeit und globaler Zusammenhang:

Die kollineare Gruppe 1. aus Abbildung 6.1 ist im globalen Zusammenhang signifikanter

geeignet ist die gestellten Anforderungen zu erfüllen.

In der hier betrachteten Anwendung entsprechen die Entitäten den einzelnen Gruppenthypothesen und ihren initialen Primitiva, den ellipsen- und geradenförmigen Segmenten. Bei dieser Abbildung von Bildmerkmalen in die Knoten des ungerichteten Graphen ist die Anordnung der Entitäten irregulär (**irreguläre Entitäten**), wie in Abschnitt 5.4 vorgestellt.

Die Entitäten sind durch zwei verschiedene Nachbarschaftsrelationen verbunden die in Abschnitt 6.2.1 mit den Begriffen Konkurrenz und Unterstützung unterschieden werden. Alle mit einer Entität i durch Unterstützung oder Konkurrenz verbundenen Entitäten bilden eine **Nachbarschaft** \mathcal{N}_i bezüglich i , wie in Abschnitt 5.4 beschrieben. Die Entitäten eines Markov Random Field sind mit Zufallsvariablen verbunden. Der Wert wird in der vorgestellten Anwendung in Form eines Label vergeben, wie in Abschnitt 5.3 eingeführt. Das Label erlaubt eine Aussage über die Bedeutung einer Hypothese im Szenekontext, es wird deshalb in Abschnitt 6.2.2 mit dem Begriff Signifikanz belegt. Innerhalb eines Nachbarschaftssystems existieren Cliques, wie in Abschnitt 5.4 vorgestellt. Wie in Abschnitt 5.5 beschrieben bedeutet das Design der Potentialfunktionen für diese Cliques eine Einbringung von a priori Wissen in das Random Field. Hier werden Potentialfunktionen für ein- und zweielementige Cliques definiert. Potentialfunktionen einelementiger Cliques bewerten die in den Knoten abgebildete Hypothese, bzw. das Primitivum. Die Potentialfunktionen zweielementiger Cliques unterscheiden sich, analog zur Definition der Nachbarschaftsrelation der Entitäten, in Unterstützung und Konkurrenz. Das Vorgehen zur Bestimmung einer global optimalen Interpretation der Szene entspricht den Abschnit-

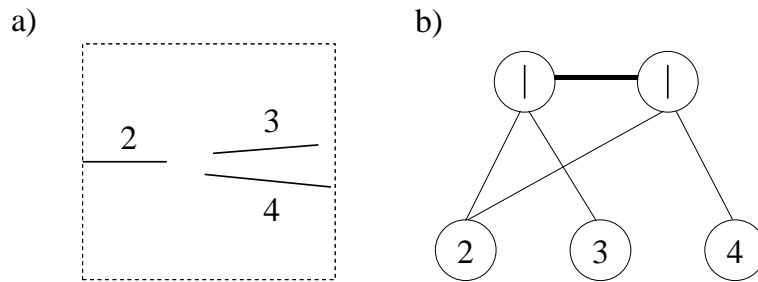


Abbildung 6.3: Gruppenthypothesen und ihre Abbildung in einen ungerichteten Graphen:

a) Aus den Segmenten 2 und 3, sowie 2 und 4 werden zwei kollineare Gruppenthypothesen hypothetisiert. Sie stellen eine mehrdeutige Interpretation der Szene dar.

b) Abbildung der Segmente und Gruppenthypothesen in die Knoten eines ungerichteten Graphen: Die Segmente und ihre kollinearen Hypothesen verbindet eine unterstützende Relation - die kollinearen Gruppenthypothesen eine konkurrierende Relation.

ten 5.5 bis 5.7 des letzten Kapitels.

Nach dieser kurzen Einordnung der Vorgehensweise des gewählten Ansatzes in den Rahmen des vorausgehenden Kapitels 5 und einem kurzen Ausblick auf den Inhalt dieses Kapitels, folgt jetzt eine detaillierte Beschreibung der einzelnen Schritte.

6.2.1 Abbildung in einen ungerichteten Graphen

Die Entitäten des Markov Random Field entsprechen im gewählten Ansatz den Gruppenthypothesen und ihren Primitiva. Ihre Repräsentation im Rahmen eines Markov Random Field erfolgt daher in Form **irregulärer Entitäten**. Eine Abbildung von Gruppenthypothesen und Segmenten in die Knoten eines ungerichteten Graphen zeigt Abbildung 6.3 für ein einfaches Beispiel mit drei Kontursegmenten und zwei kollinearen Gruppenthypothesen. Kreise symbolisieren die Knoten des Graphen. Jeder Gruppenthypothese und jedem initialen Kontursegment wird ein eigener Knoten zugeordnet. Knoten die Kontursegmente repräsentieren, sind durch die Nummern der zugehörigen Segmente kenntlich. Knoten welche Gruppenthypothesen repräsentieren, durch ihre Symbolik, hier Kollinearität (siehe dazu auch im Anhang Abbildung A.1, Tabelle A.1). Die räumliche Anordnung der Knoten erfolgt entsprechend den in Abschnitt 4.2 besprochenen Hierarchieebenen. Die Verbindungslinien zwischen den Knoten sind die Kanten des Graphen. Sie symbolisieren die Nachbarschaftsrelationen Unterstützung und Konkurrenz zwischen den Entitäten.

Unterstützung entsteht, wenn zwischen Gruppenhypothesen oder Gruppenhypothesen und initialen Segmenten, eine Bestandteilsrelation besteht. Eine kollineare Gruppe kann z.B. Bestandteil eine Parallelitätshypothese sein. In Abbildung 6.3 sind die linearen Kontursegmente 2, 3, 4 Bestandteile der kollinearen Hypothesen $\{2|3\}$, bzw. $\{2|4\}$. Die unterstützende Relationen wird durch die Verbindungslinien zwischen den Segmentknoten und zugehörigen kollinearen Gruppenknoten deutlich. Die Relation der Unterstützung ist in ihrer Richtung nicht festgelegt: Die Kontursegmente in Abbildung 6.3 unterstützen die kollinearen Gruppenhypothesen, deren Bestandteile sie sind, und die kollinearen Gruppenhypothesen unterstützen umgekehrt die Bestandteile aus denen sie gebildet wurden. Unterstützung ist typisch für die Relationen zwischen den in Abschnitt 4.2 eingeführten Hierarchieebenen. Dies ist das Resultat der geforderten Bestandteilsrelation. Unterstützung ist aber auch innerhalb einer Hierarchieebene möglich, wenn in der Ebene gebildete Gruppen Bestandteil neuer Gruppen werden, die der gleichen Ebene angehören. Dies ist z.B. der Fall, wenn aus zwei kollinearen Gruppenhypothesen eine Nähe-Gruppe hypothetisiert wird.

Konkurrenz entsteht zwischen Gruppenhypothesen, die eine mehrdeutige Interpretation der Szene zulassen: Für kollineare, kurvilineare, Nähe- und parallele Hypothesen ist dies der Fall, wenn zwei Hypothesen gemeinsame Segmente oder Hypothesen enthalten. Zwei Geschlossenheiten sind konkurrierend, wenn sie sich räumlich überlappen. In Abbildung 6.3 ist beiden kollinearen Gruppen $\{2|3\}$ und $\{2|4\}$ das Segment 2 gemeinsam. Dies bedeutet eine mehrdeutige Interpretation der Szene. Die Kante zwischen den Gruppenhypothesen ist in Abbildung 6.3 durch einen dicken schwarzen Strich hervorgehoben, um die Konkurrenz zu verdeutlichen. Konkurrierende Relationen treten nur zwischen gleichartigen Gruppenhypothesen einer Hierarchieebene (Abschnitt 4.2) auf. Nur in einem solchen Fall kann eine Mehrdeutigkeit in der Interpretation des Szeneninhalts stattfinden.

Abbildung 6.4 zeigt an einem hypothetischen Beispiel der Konturapproximation einer Leiste, wie die initialen Primitiva und Gruppenhypothesen in einen ungerichteten Graphen abgebildet werden. Dieses kleine Beispiel mit (nur) 19 Einzelsegmenten zeigt eindrucksvoll, wie schnell die Komplexität in einem solchem Graphen wachsen kann. Um die Struktur überschaubar zu halten, wurden in der Abbildung nur unterstützende Kanten berücksichtigt.

Hypothesen der 1-D Ebene: Die Gruppenhypothesen dieser Ebene sind in Abbildung 6.5a für das Beispiel aus Abbildung 6.4 gezeigt. Die kollinearen Hypothesen $\{7 | 8\}$ und $\{6 | 8\}$ (siehe dazu auch Abbildung 6.3) und die kurvilinearen Hypothesen $\{9 \subset 11\}$

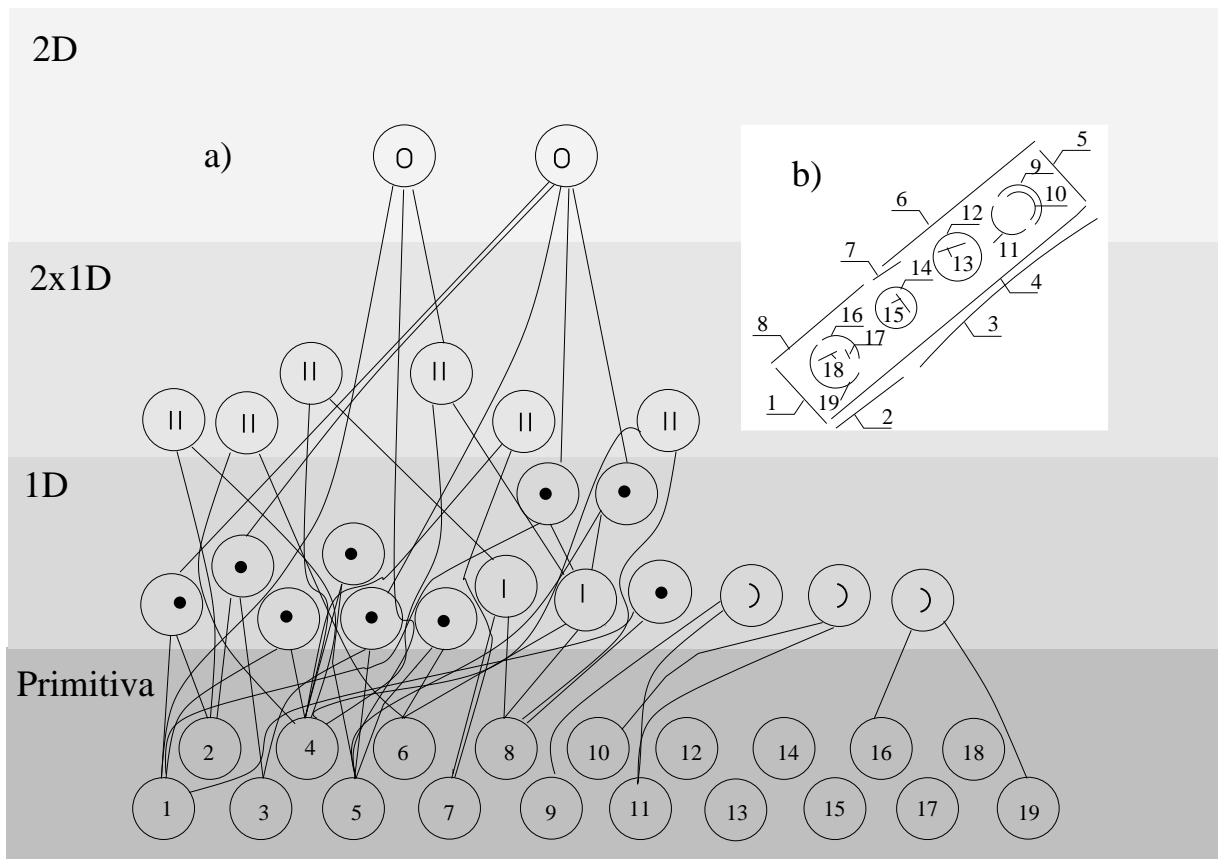


Abbildung 6.4: Beispiel für die Konturapproximation einer Leiste:

- a) ungerichteter Graph ohne konkurrierende Kanten
- b) hypothetisierte Konturapproximation einer Leiste

und $\{10 \subset 11\}$ verbindet eine konkurrierende Relation. Die obere und untere Umrandung des Loches führt im kurvilinearen Fall zu einem mehrdeutigen Verlauf der Kontursegmente.

Die Bildung der verschiedenen Nähehypthesen erfolgt auf Grundlage der initialen Segmente und bereits hypothetisierter Gruppen. Aus diesem Grund ist ein Teil der Gruppen leicht oberhalb der Ebene der 1-D Hypthesen angeordnet. Nähe-Gruppen bestehen immer aus genau zwei Primitiva und können im Gegensatz zu den anderen Gruppenthypothesen der 1-D Ebene nicht expandiert werden. Aus diesem Grund wird auf dieser Ebene keine Konkurrenzrelation für diese Gruppen formuliert.

Hypothesen der 2x1D Ebene: Die Gruppenthypothesen dieser Ebene sind in Abbildung 6.6a für das Beispiel aus Abbildung 6.4 gezeigt. Für die Bestimmung der Konkurrenzrelation zwischen den Gruppenthypothesen, wird die von den Gruppenthypothesen auf-

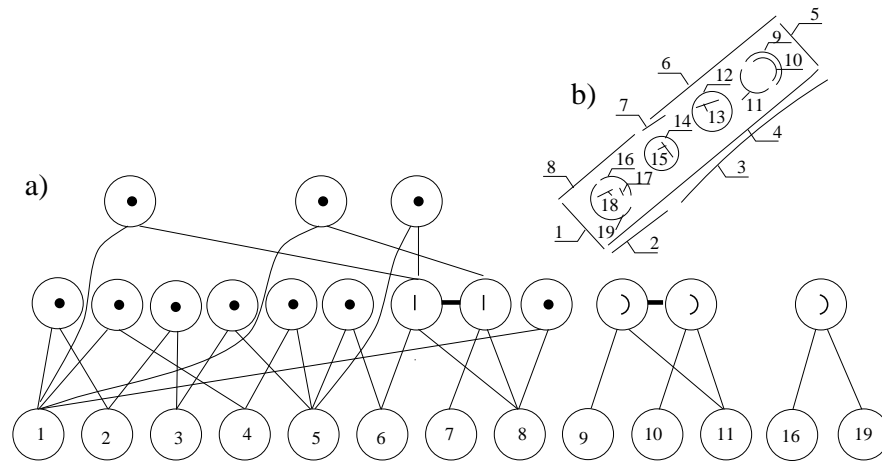


Abbildung 6.5: Konkurrierende und unterstützende Hypothesen der 1-D Ebene:

- a) ungerichteter Graph: Zahlen stehen für Kontursegmente, ein schwarzer Punkt für eine Nähe Gruppe, ein einfacher Strich für eine kollineare Gruppe und ein Halbkreis für eine kurvilineare Gruppe. Unterstützende Kanten sind durch eine einfache Linie, Konkurrenzen durch eine fette Linie gekennzeichnet.
- b) hypothetisierte Konturapproximation einer Leiste

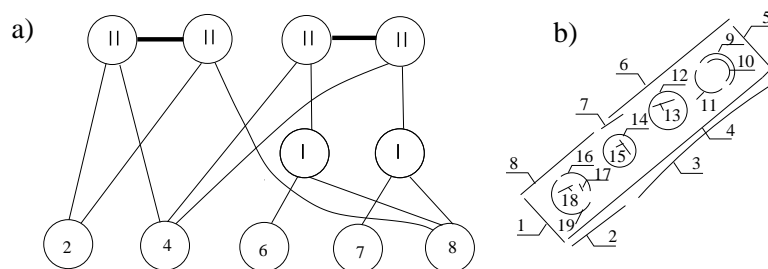


Abbildung 6.6: Konkurrierende und unterstützende Hypothesen der 2x1D Ebene:

- a) ungerichteter Graph: Zahlen stehen für Kontursegmente, ein einfacher Strich für eine kollineare Gruppe und ein doppelter Strich für eine parallele Gruppe. Unterstützende Kanten sind durch eine einfache Linie, Konkurrenzen durch eine fette Linie gekennzeichnet.
- b) hypothetisierte Konturapproximation einer Leiste

gespannte Fläche miteinander verglichen. Ist eine Fläche die Teilmenge einer anderen, ist dies eine mehrdeutige Interpretation der Szene und es entsteht Konkurrenz. Für die abgebildeten Parallelenhypothesen gilt, daß die Parallelenhypothesen $\{2 \mid 4\}$ und $\{2 \mid \{8 \parallel 6\}\}$,

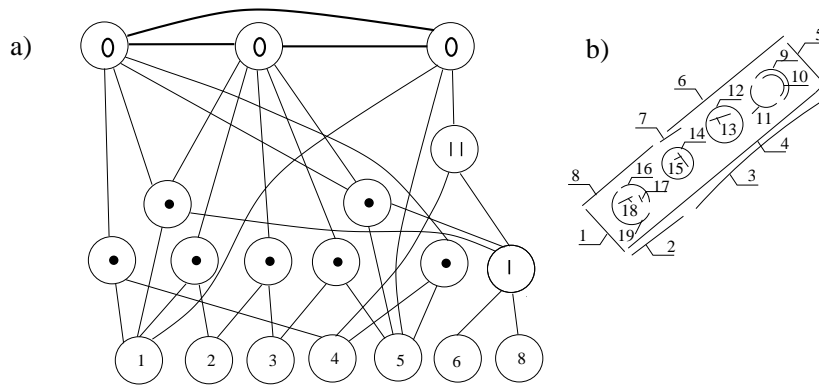


Abbildung 6.7: Konkurrernde und unterstützende Hypothesen der 2D Ebene:

- a) ungerichteter Graph: Zahlen stehen für Kontursegmente, ein schwarzer Punkt für eine Nähe-Gruppe, ein einfacher Strich für eine kollineare Gruppe, ein doppelter Strich für eine parallele Gruppe und eine Ellipse für eine Geschlossenheit. Unterstützende Kanten sind durch eine einfache Linie, Konkurrenzen durch eine fette Linie gekennzeichnet.
- b) hypothetisierte Konturapproximation einer Leiste

sowie $\{4 \mid \{8\parallel 6\}\}$ und $\{4 \mid \{8\parallel 6\}\}$ Konkurrenzrelationen besitzen.

Hypothesen der 2-D Ebene: Hypothesen über Geschlossenheiten stützen sich, wie in Abbildung 6.7 gezeigt, auf alle darunter liegenden Ebenen. Geschlossenheiten werden einerseits aus einer zyklischen Abfolge von Nähegruppen hypothetisiert. In der Abbildung bilden die hypothetisierten Nähe Gruppen zwei solcher zyklischen Abfolgen. Zum anderen bilden die Parallelitätshypothesen eine Ausgangsbasis für Geschlossenheiten. In der Abbildung wird die Parallelitätshypothese $\{6 \parallel 8\parallel 4\}$ durch die Segmente $\{1\}$ und $\{5\}$ geschlossen. Alle Hypothesen stehen in einer Konkurrenzrelation zueinander, da die von ihnen beschriebenen Flächen sich überlagern.

In diesem Abschnitt wurde dargestellt, wie die Entitäten des Ansatzes aussehen und wie sie mit Hilfe der verwendeten Nachbarschaftsrelationen in einen ungerichteten Graphen abgebildet werden. Der Aufbau dieses Graphen orientiert sich dabei an der in Abschnitt 4.2 vorgestellten Hierarchie der Gruppenthypothesen. Im folgenden Abschnitt wird ausführlich dargestellt, wie jeder dieser Entitäten ein Labelwert zugeordnet wird.

6.2.2 Signifikanz der Hypothesen

Für eine geeignete Repräsentation aller relevanten Zustände ist die Definition einer geeigneten Labelmenge \mathcal{L} erforderlich. Dies wird auch mit Labeling Problem bezeichnet (siehe dazu auch Abschnitt 5.3). Das einer Entität zugeordnete Label entspricht in der vorgestellten Anwendung dem Vertrauen, welches in die Konsistenz der Gruppenhypothese gesetzt wird. Damit bringt dieses Label die Bedeutung der betrachteten Hypothese im Szenenkontext zum Ausdruck. Aus diesem Grund wird es auch mit Signifikanzlabel oder kurz Signifikanz bezeichnet. Der Wertebereich der diskreten, numerischen Label ist auf das Intervall $[0,1]$ definiert, sodaß gilt:

$$\mathcal{L} = \{l \mid l \in \{0, k, 2k, \dots, nk = 1\}\}, \quad n \in \mathbb{N}, \quad k = \text{konstant} \quad (6.1)$$

Der Wert 1 entspricht dabei der maximal möglichen Signifikanz.

6.2.2.1 Datengetriebene Signifikanz

Für jeden Knoten des ungerichteten Graphen wird eine **datengetriebene** Signifikanz errechnet. Die Berechnung dieser datengetriebenen Signifikanzen erfolgt auf der Grundlage von Heuristiken über die räumlich-geometrischen Eigenarten der verschiedenen Hypothesen. Für eine kollineare Gruppe ist es z.B. erwünscht, wenn die Lücke zwischen den Primitiva der Gruppe gering und ihre Orientierung gleichartig ist. Aus diesem Grund ist es notwendig eine differenzierte Berechnung der datengetriebenen Signifikanzen für die verschiedenen Gruppentypen vorzunehmen. Die Berechnung der datengetriebenen Signifikanz einer Hypothese kann sich dabei aus der Berechnung mehrerer Teilsignifikanzen zusammensetzen. Die einzelnen Teilsignifikanzen können durch Gewichtungen α_x in ihrer Wirkung auf die Berechnung der datengetriebenen Signifikanz beeinflußt werden. Die Bestimmung der Gewichtungen ist empirisch (siehe dazu auch Anhang A.3). Wie in Abschnitt 5.7.1.3 eingeführt, wird aus der datengetriebenen Signifikanz einer Hypothese die initiale Stabilität G errechnet, die den Prozeß zur Ermittlung des Energieminimums mittels HCF steuert.

Kontursegmente: Die Berechnung der datengetriebenen Signifikanz l_{cs}^o für Kontursegmente cs , beruht auf Fehlerannahmen, wie sie bereits durch den Approximationsalgorithmus der Kontursegmente (siehe dazu auch [13]) vorgenommen werden. Die Daten sind einfach und schnell verfügbar und garantieren gleichzeitig die Konsistenz der ermittelten

datengetriebenen Signifikanzen mit der Konturapproximation. Die Berechnung des Approximationsfehlers ϵ_{cs} für Geradenstücke beruht auf der Berechnung der euklidischen Distanz zwischen den Kantenpixeln und den lotrechten Punkten der approximierten Gerade:

$$\epsilon_{cs} = \frac{1}{J} \sum_{j=1}^J \text{dist}(p_j, cs)^2 \quad p_j \in cs \quad (6.2)$$

Die Berechnung der datengetriebenen Signifikanz l_{cs}^o erfolgt unter Berücksichtigung des maximalen Approximationsfehler ϵ_{cs}^{max} und des minimalen Approximationsfehler ϵ_{cs}^{min} aller in der Szene vorhandenen Kontursegmente:

$$l_{cs}^o = 1 - \frac{\epsilon_{cs} - \epsilon_{cs}^{min}}{\epsilon_{cs}^{max}} \quad (6.3)$$

Diese Abbildungsvorschrift ordnet für den Fall, daß der aktuell beobachtete Fehler ϵ_{cs} gleich dem minimalen Fehler ϵ_{cs}^{min} ist, der datengetriebenen Signifikanz l_{cs}^o immer die maximale Signifikanz 1 zu.

Kollinearität: Die Berechnung der datengetriebenen Signifikanz einer kollinearen Hypothese setzt sich aus der Berechnung dreier Teilsignifikanzen zusammen:

- l_{col1}^o beschreibt den Approximationsfehler der neu entstandenen kollinearen Hypothese, im Vergleich zu den Kantenpixeln der zugrundeliegenden Kontursegmente
- l_{col2}^o vergleicht die Länge, der die Gruppe konstituierenden Segmente mit den Lücken zwischen diesen Segmenten
- l_{col3}^o vergleicht die Orientierungsdifferenzen der Kontursegmente mit einem aus der Stichprobe ermittelten Maximum für Orientierungsdifferenzen

Für eine kollineare Gruppenhypothese wird eine Repräsentation in Form eines Geradenstücks in der Gesamtlänge der Gruppe vorgenommen. Dazu werden die zugrundeliegenden Kantenpixel der Primitiva der Gruppe einer erneuten Konturapproximation unterzogen. Die erste Teilrechnung der datengetriebenen Signifikanz l_{col1}^o einer kollinearen Hypothese col erfolgt daher analog zur Berechnung der datengetriebenen Signifikanz eines Kontursegmentes. Der aktuelle Approximationsfehler einer kollinearen Hypothese ϵ_{col} wird mit dem minimalen $\epsilon_{\forall col}^{min}$ und dem maximalen Approximationsfehler $\epsilon_{\forall col}^{max}$ aller kollinearen Gruppenhypothesen verglichen:

$$l_{col1}^o = 1 - \frac{\epsilon_{col} - \epsilon_{\forall col}^{min}}{\epsilon_{\forall col}^{max}} \quad (6.4)$$

Das arithmetische Mittel über die Länge aller in der Hypothese enthaltenen Liniensegmente \bar{L}_{col} , sowie das arithmetische Mittel der Lückenlängen zwischen diesen Segmenten \bar{D}_{col} , bestimmen die Berechnung der zweiten Teilsignifikanz für kollineare Gruppen:

$$l_{col2}^o = \frac{\bar{L}_{col} - \bar{D}_{col}}{\bar{L}_{col}} \quad (6.5)$$

Ist das Ergebnis der Gleichung kleiner als Null, wird der Ausdruck zu Null gesetzt.

Die Berechnung der dritten Teilsignifikanz l_{col3} berücksichtigt das arithmetische Mittel der unter den Primitiva der Gruppe auftretenden Orientierungsdifferenzen $\bar{\gamma}_{col}$. Der für $\bar{\gamma}_{col}$ ermittelte Wert wird mit dem aus der Stichprobe für kollineare Gruppen ermittelten Maximum für Orientierungsdifferenzen (siehe dazu Abschnitt 4.3.2.2) Δ_{col} verglichen. Für die Berechnung der Teilsignifikanz gilt die folgende Fallunterscheidung:

$$l_{col3}^o = \begin{cases} 1 & : \bar{\gamma}_{col} \leq \Delta_{col} \\ \Delta_{col}/\bar{\gamma}_{col} & : \bar{\gamma}_{col} > \Delta_{col} \end{cases} \quad (6.6)$$

Die Art der Fallunterscheidung zeigt, daß hier Δ_{col} wie ein Schwellwert verstanden wird: Die ideale Orientierungsdifferenz unter den Segmenten ist Null. Tatsächlich hatten in der Stichprobe aber ein Großteil der Segmente eine Orientierungsdifferenz > 0 . Aus diesem Grund wird der Teilsignifikanz für das Intervall $[0, \Delta_{col}]$ der maximale Wert 1 zugeordnet. Die datengetriebene Signifikanz einer kollinearen Hypothese l_{col}^o ergibt sich aus der Aufsummierung der Teilsignifikanzen.

$$l_{col}^o = \alpha_1 l_{col1}^o + \alpha_2 l_{col2}^o + \alpha_3 l_{col3}^o \quad (6.7)$$

Kurvilinearität: Die datengetriebene Signifikanz einer Kurvilinearität wird analog der Kollinearität gebildet. Da alle Formeln und ihre Herleitungen identisch sind, wird auf eine ausführliche Darstellung verzichtet. Für die Berechnung der Orientierungsdifferenzen wird - analog zur Auswertung der Stichprobe für kurvilineare Gruppen - die Orientierung der Tangente im betrachteten Endpunkt berücksichtigt.

Nähe: Die Berechnung der datengetriebenen Signifikanz von Nähe-Hypothesen setzt sich aus der Berechnung zweier Teilsignifikanzen zusammen:

- l_{prox1}^o vergleicht die Länge, der die Gruppe konstituierenden Segmente mit den Lücken zwischen diesen Segmenten
- l_{prox2}^o vergleicht die Orientierungsdifferenzen der Kontursegmente mit einem aus der Stichprobe ermittelten Maximum für Orientierungsdifferenzen

Für die Berechnung der ersten Teilsignifikanz l_{prox1}^o wird die Länge der Segmente ins Verhältnis zur Lücke D_{prox} zwischen diesen gesetzt. Dies geschieht identisch zur Berechnung der zweiten Teilsignifikanz für kollineare Hypothesen l_{col1}^o . Im Unterschied zu Kollinearitäten gibt es nur eine Lücke, sodaß keine Mittelwertbildung nötig ist (Gleiches gilt für die unten folgende Berechnung der Orientierungsdifferenz γ_{prox}). Zur Erinnerung: Nähe-Hypothesen werden immer aus zwei Kontursegmenten gebildet.

Die erste Teilsignifikanz einer Nähe-Hypothese l_{prox1}^o ergibt sich damit zu:

$$l_{prox1}^o = \frac{\bar{L}_{prox} - D_{prox}}{\bar{L}_{prox}} \quad (6.8)$$

Ist das Ergebnis der Gleichung kleiner als Null, wird der Ausdruck zu Null gesetzt.

Die Berechnung der zweiten Teilsignifikanz l_{prox2}^o vergleicht die Orientierungsdifferenz der Kontursegmente γ_{prox} mit einem aus der Stichprobe für Nähe-Gruppen ermittelten Maximum für Orientierungsdifferenzen Δ_{prox} (siehe dazu Abschnitt 4.3.2.4). Für die Berechnung der Teilsignifikanz wird die folgende Fallunterscheidung vorgenommen:

$$l_{prox2}^o = \begin{cases} \gamma_{prox}/\Delta_{prox} & : \gamma_{prox} < \Delta_{prox} \\ 1 & : \gamma_{prox} = \Delta_{prox} \\ \Delta_{prox}/\gamma_{prox} & : \gamma_{prox} > \Delta_{prox} \end{cases} \quad (6.9)$$

Für das Handlabeling naher Gruppen fanden alle Paarungen von Primitiva Berücksichtigung, die weder eine kollineare noch kurvilineare Hypothese bildeten und deren Endpunkte dicht beieinander lagen. Dies traf vorzugsweise auf Segmente in einer annähernd rechtwinkligen Lage zu, da viele Elemente der untersuchten Domäne rechte Winkel aufweisen (Leisten, Würfel, etc.). Ein entsprechendes Maximum (siehe dazu Abschnitt 4.3.2.4) war das Ergebnis der Betrachtung der Orientierungsdifferenzen der untersuchten Primitiva. Im Gegensatz zur Signifikanzberechnung kollinearere und kurvilinearere Hypothesen wird hier das Maximum nicht wie ein Schwellwert betrachtet, da es intuitiv keinen idealen Wert für die Orientierungsdifferenz naher Gruppen gibt. Um jedoch das in dieser Domäne existierende Maximum für die Bewertung auszunutzen, wurde die Bewertung so gewählt, daß sie im Bereich des Maximums hoch ist und zu beiden Seiten abfällt.

Die datengetriebene Signifikanz einer Nähe-Hypothese l_{prox}^o ergibt sich aus der Aufsummierung der Teilsignifikanzen:

$$l_{prox}^o = \alpha_7 l_{prox1}^o + \alpha_8 l_{prox2}^o \quad (6.10)$$

Parallelität: Parallele Hypothesen werden mit Hilfe von initialen Geradenstücken, sowie kollinearen Hypothesen gebildet. Da kollineare Hypothesen wiederum durch geradenförmige Segmente approximiert werden, wird nachfolgend nur von Segmenten gesprochen. Die

Berechnung der datengetriebenen Signifikanz l_{par}^o für parallele Hypothesen untergliedert sich in die Ermittlung dreier Teilsignifikanzen:

- l_{par1}^o berücksichtigt die Orientierungsdifferenz zwischen den Segmenten der Hypothese
- l_{par2}^o wertet den mittleren Abstand zwischen den Segmenten aus
- l_{par3}^o beschreibt den Grad der Überlappung beider Segmente

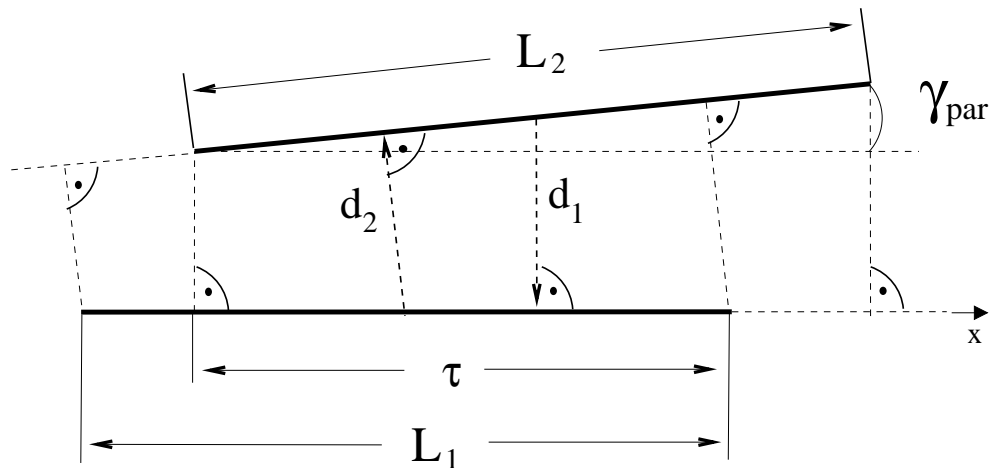


Abbildung 6.8: Zur Berechnung der Teilsignifikanzen einer Parallelität:

- 1.) γ_{par} ist die Orientierungsdifferenz der Segmente L_1 und L_2
- 2.) Aus dem arithmetischen Mittel der Lote d_1, d_2 wird der mittlere Abstand berechnet
- 3.) Die Länge τ beschreibt den Anteil der gegenseitigen Überlappung

Die Berechnung der ersten Teilsignifikanz l_{par1}^o vergleicht die Orientierungsdifferenz der parallelen Hypothese γ_{par} mit einem aus der Stichprobe für parallele Gruppen ermittelten Maximum für Orientierungsdifferenzen Δ_{par_orient} (siehe dazu Abschnitt 4.3.2.5). Für die Berechnung der Teilsignifikanz gilt die folgende Fallunterscheidung:

$$l_{par1}^o = \begin{cases} 1 & : \gamma_{par} \leq \Delta_{par_orient} \\ \Delta_{par_orient}/\gamma_{par} & : \gamma_{par} > \Delta_{par_orient} \end{cases} \quad (6.11)$$

Analog zu Kollinearitäten und Kurvilinearitäten wird hier das Maximum der Orientierungsdifferenzen wie ein Schwellwert eingesetzt.

Die zweite Teilsignifikanz l_{par2}^o bewertet den mittleren Abstand unter den Segmenten der Gruppenhypothese bezogen auf die Normlänge (siehe dazu Abschnitt 4.3.2). In einem ersten Schritt wird dafür der Normfaktor x_{par} für das arithmetische Mittel der Segmente

L_1 und L_2 aus Abbildung 6.8 ermittelt:

$$x_{par} = \frac{2 \cdot \text{Normlänge}}{L_1 + L_2} \quad (6.12)$$

Anschließend wird von den Mittelpunkten beider Segmente das Lot auf das jeweils gegenüberliegende Segment gefällt. Dies ergibt die beiden Längen d_1 und d_2 in Abbildung 6.8. Das arithmetische Mittel $\bar{d} = (d_1 + d_2)/2$ wird mit x_{par} multipliziert und stellt damit ein Maß für den mittleren Abstand dar. Dieses Maß wird mit einem aus der Stichprobe ermittelten Maximum Δ_{par_mean} für mittlere Segmentabstände verglichen (siehe dazu Abschnitt 4.3.2.5). Dabei wird die folgende Fallunterscheidung vorgenommen:

$$l_{par2_i}^o = \begin{cases} (\bar{d} x_{par}) / \Delta_{par_mean} & : \bar{d} x_{par} < \Delta_{par_mean} \\ 1 & : \bar{d} x_{par} = \Delta_{par_mean} \\ \Delta_{par_mean} / (\bar{d} x_{par_mean}) & : \bar{d} x_{par} > \Delta_{par_mean} \end{cases} \quad (6.13)$$

Diese Fallunterscheidung setzt voraus, daß es für eine "ideale" Parallelität ein bestimmtes Maß für das Verhältnis von Abstand und Segmentlängen gibt: das aus der Stichprobe ermittelte Maximum.

Zur Ermittlung des Anteils der gegenseitigen Überlappung werden die Endpunkte der Segmente auf eine Gerade entlang des jeweils gegenüberliegenden Segmentes projiziert. Anschließend wird für ein Segment die Strecke ausgewertet, die innerhalb dieser vier Punkte - den eigenen Endpunkten und den Projektionsendpunkten des anderen Segments - liegt. Abbildung 6.8 zeigt diesen Vorgang für das untere Segment. Die gesuchte Überlappung der Segmente ist mit τ gekennzeichnet. Für die Berechnung der dritten Teilsignifikanz wird τ durch das arithmetische Mittel der Segmentlängen \bar{L} geteilt:

$$l_{par3}^o = \frac{\tau}{\bar{L}_{par}} \quad (6.14)$$

Je höher der Grad der gegenseitigen Überlappung, desto besser die Bewertung der Hypothese.

Der Wert der datengetriebenen Signifikanz einer parallelen Hypothese, l_{par}^o , ergibt sich aus der Aufsummierung der drei Teilsignifikanzen:

$$l_{par}^o = \alpha_9 l_{par1}^o + \alpha_{10} l_{par2}^o + \alpha_{11} l_{par3}^o \quad (6.15)$$

Geschlossenheit: Hypothesen über Geschlossenheiten beziehen in ihre Auswertung alle bis jetzt vorgestellten Hypothesen ein. Wie in Abschnitt 4.4.2 vorgestellt, gibt es zwei Strategien zur Bildung von Geschlossenheiten: zum einen beruhend auf der Auswertung

von zyklischen Nähegruppen. Zum anderen unter Einbeziehung von Parallelitäten. Entsprechend können zwei verschiedene Bewertungen $l_{clos_prox}^o$ und $l_{clos_par}^o$ angestoßen werden. Für Geschlossenheiten, die auf Nähegruppen basieren, werden zwei Teilsignifikanzen ermittelt:

- l_{clos1}^o bewertet die Länge der Segmente und der Lücken zwischen diesen
- l_{clos2}^o berücksichtigt die Anzahl der in einer Geschlossenheitshypothese enthaltenen Segmente

Über die Länge aller in der Gruppe enthaltenen Segmente wird das arithmetische Mittel \bar{L}_{clos} gebildet. Gleiches geschieht mit den Lücken zwischen diesen Segmenten \bar{D}_{clos} . Mit diesen Werten ergibt sich die erste Teilsignifikanz zu:

$$l_{clos1}^o = \begin{cases} \frac{\bar{L}_{clos} - \bar{D}_{clos}}{L_{clos}} & : \lambda \leq 4 \\ \frac{\bar{L}_{clos} - \bar{D}_{clos}}{L_{clos}} \frac{1}{\lambda} & : \lambda > 4 \end{cases} \quad (6.16)$$

Ist das Ergebnis der Gleichung kleiner als Null, wird der Ausdruck zu Null gesetzt.

Es wird eine Fallunterscheidung vorgenommen, die Geschlossenheitshypothesen mit maximal vier Segmenten bevorzugt. Diese Unterscheidung beruht auf der Beobachtung, daß Gruppen mit einer großen Segmentanzahl λ in der Regel ein günstigeres ‐Länge-zu-Lücke‐ Verhältnis besitzen und damit die Bewertung dominieren.

Die zweite Teilsignifikanz beruht auf der Anzahl λ der in der Hypothese enthaltenen Segmente. Eine zu Gleichung 6.16 analoge Fallunterscheidung wird vorgenommen:

$$l_{clos2}^o = \begin{cases} 1 & : \lambda \leq 4 \\ \frac{4}{\lambda} & : \lambda > 4 \end{cases} \quad (6.17)$$

Hypothesen mit weniger Segmenten erfahren weniger Unterstützung bei der Berechnung der zweielementigen Clique-Potentiale. Die Auswertung der Segmentanzahl in einer Gruppe dient deshalb als Korrekturfaktor.

Die datengetriebene Signifikanz einer Geschlossenheit, basierend auf der Auswertung von Nähegruppen, wird aus der Aufsummierung der Teilsignifikanzen ermittelt.

$$l_{clos_prox}^o = \alpha_{12} l_{clos1}^o + \alpha_{13} l_{clos2}^o \quad (6.18)$$

Für Geschlossenheiten, die auf der Auswertung von parallelen Hypothesen beruhen, werden die oben eingeführten Teilsignifikanzen l_{clos1}^o und l_{clos2}^o analog berechnet. Zusätzlich findet eine dritte Teilsignifikanz, die die Bewertung der parallelen Hypothese aus Gleichung 6.15 berücksichtigt, Eingang in die Berechnung. Diese dritte Teilsignifikanz wird folgendermaßen ermittelt:

$$l_{clos3}^o = 0.5 + \frac{l_{par}^o}{2} \quad (6.19)$$

Parallele Hypothesen sind erheblich komplexer als Nähe Gruppen. Deshalb sollen auch Geschlossenheiten, die auf solchen Gruppen basieren, eine bessere Bewertung erfahren. Damit gleichzeitig sehr niedrig bewertete Parallelitäten, die Gesamtbewertung nicht negativ beeinflussen, ist das Intervall für die Berechnung der dritten Teilsignifikanz auf den Bereich $[0.5, 1]$ festgesetzt.

Die datengetriebene Signifikanz einer Geschlossenheit, basierend auf der Auswertung paralleler Hypothesen, wird aus allen drei eingeführten Teilsignifikanzen ermittelt:

$$l_{clos-par}^o = \alpha_{12}^* l_{clos1}^o + \alpha_{13}^* l_{clos2}^o + \alpha_{14}^* l_{clos3}^o \quad (6.20)$$

Die Gewichte α_{12}^* und α_{13}^* können bei Bedarf abweichend zu Gleichung 6.18 gewählt werden.

In diesem Abschnitt wurde die Berechnung der datengetriebenen Signifikanzen eingehend beschrieben. Für diese Berechnung werden Heuristiken über die räumlich-geometrischen Eigenarten der Hypothesen ausgenutzt. Dieses Vorgehen ist daher nur geeignet, eine "lokale Signifikanz" der Hypothesen zu bestimmen. Für die Suche nach einer global optimalen Konfiguration ist es im Gegensatz dazu notwendig, die Hypothesen im Kontext ihrer Nachbarn zu betrachten, wie in Abschnitt 6.2 beschrieben. Wesentlich für diesen Optimierungsprozeß ist die Auswertung der Clique-Potentiale der Nachbarschaftssysteme. Mit dem Design dieser Clique-Potentiale wird sich der folgende Abschnitt auseinandersetzen.

6.2.3 Design der Clique-Potentiale

Beim Aufbau eines ungerichteten Graphen entstehen Cliques, wie in Abschnitt 5.4 vorgestellt. Cliques können durch das Design der Clique Potentiale a priori Wissen abbilden (siehe dazu auch Abschnitt 5.5). In der hier vorgestellten Anwendung legen die Clique-Potentiale die Interaktion der Gruppenthypothesen fest:

- sie formulieren die Abhängigkeit zwischen der datengetriebenen Signifikanz einer Hypothese und deren Bewertung durch das Random Field
- sie ermöglichen die Betrachtung unterstützender- und konkurrierender Relationen unter den Gruppenthypothesen.

In Abbildung 6.9 sind in Anlehnung an Abbildung 6.3 Cliquesbildungen für zwei kollineare Hypothesen und ihre Segmente gezeigt. Die in der Abbildung gezeigten Cliques unterscheiden sich durch die Anzahl ihrer Elemente. Sie werden entsprechend als einelementige Cliques, zweielementige Cliques, etc. bezeichnet.

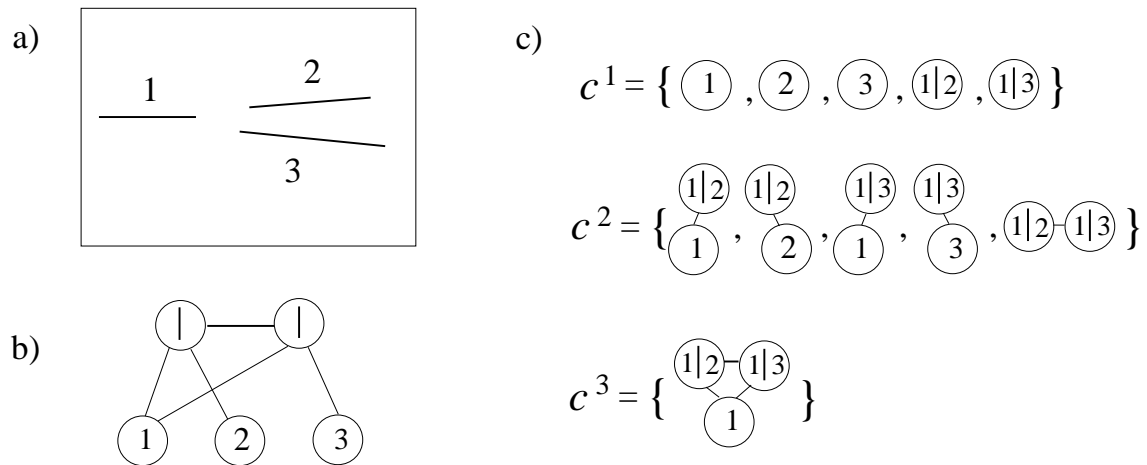


Abbildung 6.9: Mögliche Cliquenbildungen für kollineare Gruppenhypothesen:

- a) Szene mit drei Segmenten aus denen die kollinearen Gruppen $\{1 | 2\}$ und $\{1 | 3\}$ hypothesisiert werden
- b) Abbildung der Segmente und Gruppenhypothesen in einen ungerichteten Graphen
- c) Mengen einelementiger-, zweielementiger und dreielementiger Cliquen

In der vorgestellten Anwendung werden ein-, zwei- und dreielementige Cliquen berücksichtigt. Das Potential dreielementiger Cliquen wird vereinfachend auf einen konstanten Wert festgesetzt.

Die einelementige Clique: c^1 ist der Knoten, bzw. die Hypothese selbst. Daher entspricht die Anzahl der einelementigen Cliquen der Anzahl der Knoten des Netzes. Für jede einelementige Clique wird eine datengetriebene Signifikanz berechnet.

Zweielementige Cliquen: c^2 entstehen zwischen Nachbarknoten, die durch eine Relation verbunden sind. Diese Relationen formulieren Konkurrenz und Unterstützung.

Die dreielementige Clique: Sie entsteht, wenn Nachbarknoten in konkurrierender Relation von einer gemeinsamen Hypothese gestützt werden (siehe dazu auch Abbildung 6.3).

Das Basispotential: Für die unterschiedlichen Cliquentypen, sowie die Relationen Unterstützung und Konkurrenz, wurde ein Design spezieller Potentialfunktionen V_C^{nS} durchgeführt. Wobei mit n die Wertigkeit der Clique (einelementig etc.) und mit S die Relationen Unterstützung und Konkurrenz (*sup*, *comp*) angezeigt werden. Damit der spezifische

Einfluß der verschiedenen Potentialfunktionen in der Betrachtung der globalen Energie vergleichbar bleibt, wird für jede Potentialfunktion ein Basispotential V^n definiert (n entsprechend der Wertigkeit der Clique). Alle diese Basispotentiale sind einheitlich auf ein Intervall in den Grenzen $[-1,1]$ normiert. Das spezifische Potential V_C^{nS} wird aus der Gewichtung des Basispotentials mit einem Faktor β_x berechnet. Dieser Faktor kann entsprechend dem gewünschten Einfluß des spezifischen Potentials auf die Berechnung der globalen Energie gewählt werden:

$$V_C^{nS} = \beta_x V^n \quad (6.21)$$

Die Gewichtungen β_x wurden empirisch bestimmt.

6.2.3.1 Potential einer ein-einelementigen Clique

Das Potential einer einelementigen Clique V_C^1 bewertet die Übereinstimmung zwischen dem aktuell berechneten Label f_i einer Hypothese und dem Wert ihrer datengetriebenen Signifikanz f_i^o . Je größer die Übereinstimmung beider Labelwerte ist, desto größer ist das "Vertrauen" in die ermittelten Werte - da sie ja initial durch die Eingangsdaten und gleichzeitig durch die Auswertung der Nachbarschaftsbeziehungen gestützt werden. Dieses Vertrauen drückt sich in einem kleinen Energiebeitrag der betrachteten Clique zur Gesamtenergie $U(f)$ aus. Unter der Berücksichtigung der speziellen Gewichtung β_1 für einelementige Cliquen berechnet sich das spezifische Potential einer einelementigen Clique zu:

$$V_C^1 = \beta_1 V^1 = \beta_1 |f_i - f_i^o| \quad (6.22)$$

Die Abweichung zwischen den Labelwerten des Basispotentials $V^1 = f_i - f_i^o$ ist in beiden Richtungen gleichwertig. Aus diesem Grund wird das Intervall $[-1,+1]$ für Basispotentiale durch die Bildung des Betrages $|V^1|$ auf den Bereich $[0,+1]$ eingeschränkt.

6.2.3.2 Potential einer zweielementigen Clique

Entsprechend den Relationen Unterstützung und Konkurrenz sind unterschiedliche Potentiale definiert.

Unterstützung: In der Betrachtung der unterstützenden Relation von Hypothesen werden qualitativ vier Fälle unterschieden, wie sie auch in in der linken Hälfte von Abbildung 6.10 erkennbar sind:

- **Fall 1:** Beide Knoten besitzen eine hohe Signifikanz. Dies ist der Fall wenn die auf die Hypothesenknoten abgebildeten Label f_i, f_j einen Wert > 0.5 verfügen. Zwei solchermaßen signifikante Hypothesen, welche sich gegenseitig unterstützen, sind konsistent mit dem modellierten a priori Wissen. Der Beitrag einer zweielementigen Clique zur Energiefunktion $U(f)$ soll in diesem Fall negativ werden.
- **Fall 2+3:** Jeweils ein Hypothesenknoten verfügt über ein Label mit einem Wert > 0.5 , während der zweite Knoten ein Label mit einem Wert ≤ 0.5 besitzt. Beide Hypothesen widersprechen sich und liefern deshalb eine unerwünschte inkonsistente Interpretation des Bildinhaltes. Der Beitrag einer zweielementigen Clique zur Energiefunktion $U(f)$ soll in einem solchen Fall positiv sein.
- **Fall 4:** Beiden Hypothesenknoten sind Label zugeordnet, deren Wert ≤ 0.5 ist. Eine relevante Aussage über die Bedeutung beider Hypothesen für den Szeneninhalte ist daraus nicht ableitbar. Der Energiebeitrag einer solchen zweielementigen Clique ist in diesem Fall neutral.

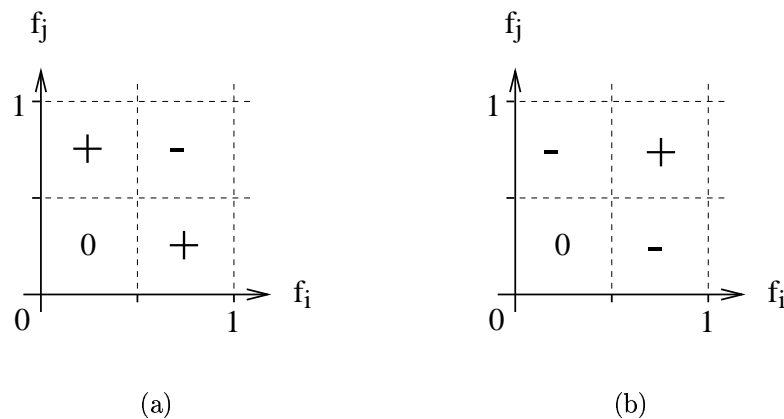


Abbildung 6.10: Qualitative Darstellung der zweielementigen Potentialfunktionen für: (a) Unterstützung und (b) Konkurrenz in Abhängigkeit von den Werten der mit den Hypothesen verknüpften Label.

Für die in der linken Hälfte von Abbildung 6.10 gezeigten Quadranten für unterstützende Relationen ist eine geeignete Abbildungsvorschrift zu definieren, die das gewünschte Verhalten auf das Basispotential für zweielementige Cliquen abbildet. Im einfachsten Fall kann jedem Quadranten ein fester Wert: $-1, +1, 0$ zugewiesen werden. Der einzige Vorteil einer solchen Abbildungsvorschrift liegt in ihrer Einfachheit: Die Dynamik der Label-Interaktion findet kaum Berücksichtigung und an den Grenzen zwischen den Quadranten entstehen

Sprungstellen. Die Entscheidung fiel daher auf eine Formulierung der Fälle 1, 2 und 3 mit Hilfe von Gaußfunktionen Γ :

$$\Gamma(f_i, f_j) = \frac{1}{2\pi\sigma^2} \exp \left\{ -\frac{1}{2} \left[\left(\frac{f_i - \mu_1}{\sigma} \right)^2 + \left(\frac{f_j - \mu_2}{\sigma} \right)^2 \right] \right\} \quad (6.23)$$

Der neutrale Energiebeitrag aus Fall 4 wird einheitlich auf Null abgebildet. Eine Approximation der Abbildungsvorschrift mit Hilfe von Gaußfunktionen bietet die Möglichkeit die Dynamik der Label-Interaktionen zu berücksichtigen und gleichzeitig glatte Übergänge an den Quadrantengrenzen zu schaffen. Damit ergibt sich der Verlauf des Basispotentials V^2 einer zweielementigen Clique, deren Entitäten durch eine unterstützende Relation verbunden sind, zu:

$$V^2 = \begin{cases} 0 & : 0 \leq f_i \leq 0.5 \quad \text{und} \quad 0 \leq f_j \leq 0.5 \\ \kappa \Gamma & : 0 \leq f_i \leq 0.5 \quad \text{und} \quad 0.5 < f_j \leq 1 \quad \text{mit} \quad \mu_1 = 0, \quad \mu_2 = 1 \\ -\kappa \Gamma & : 0.5 < f_i \leq 1 \quad \text{und} \quad 0.5 < f_j \leq 1 \quad \text{mit} \quad \mu_1 = 1, \quad \mu_2 = 1 \\ \kappa \Gamma & : 0.5 < f_i \leq 1 \quad \text{und} \quad 0 \leq f_j \leq 0.5 \quad \text{mit} \quad \mu_1 = 1, \quad \mu_2 = 0 \end{cases} \quad (6.24)$$

Der stetige Verlauf der Abbildungsvorschrift an den Grenzen der Quadranten geschieht durch eine geeignete Wahl des σ -Wertes. Die Erwartungswerte, μ_1 und μ_2 , dieser zweidimensionalen Gaußfunktion werden in den vier Quadranten in Abhängigkeit von f_i und f_j gewählt. Der Faktor κ normiert den Wertebereich der Funktion auf das Intervall $[-1, +1]$ für Basispotentiale. Den qualitativen Verlauf dieser zweidimensionalen Gaußfunktionen in den verschiedenen Quadranten zeigt Abbildung 6.11. Für die Berechnung des speziellen Potentials einer zweielementigen Clique V_C^{2sup} in unterstützender Relation wird das in Gleichung 6.24 berechnete Basispotential mit dem Faktor β_2^{sup} gewichtet:

$$V_C^{2sup} = \beta_2^{sup} V^2 \quad (6.25)$$

Konkurrenz: wirkt in beiden Richtungen gleichermaßen, da jede der untersuchten Hypothesen die Inkonsistenz verursachen kann. Analog zur Unterstützung können vier Fälle - unter "umgekehrten Vorzeichen" - qualitativ unterschieden werden. Sie sind in der rechten Hälfte von Abbildung 6.10 aufgezeigt:

- **Fall1:** Beide Knoten besitzen eine hohe Signifikanz. Analog zur Unterstützung besitzen damit beide Hypothesenknoten einen Labelwert > 0.5 . Der Fall, daß zwei gut bewertete Hypothesen eine inkonsistente Interpretation zulassen, widerspricht der Modellierung. Der Beitrag der zweielementigen Clique zur Energiefunktion soll in diesem Fall positiv werden.

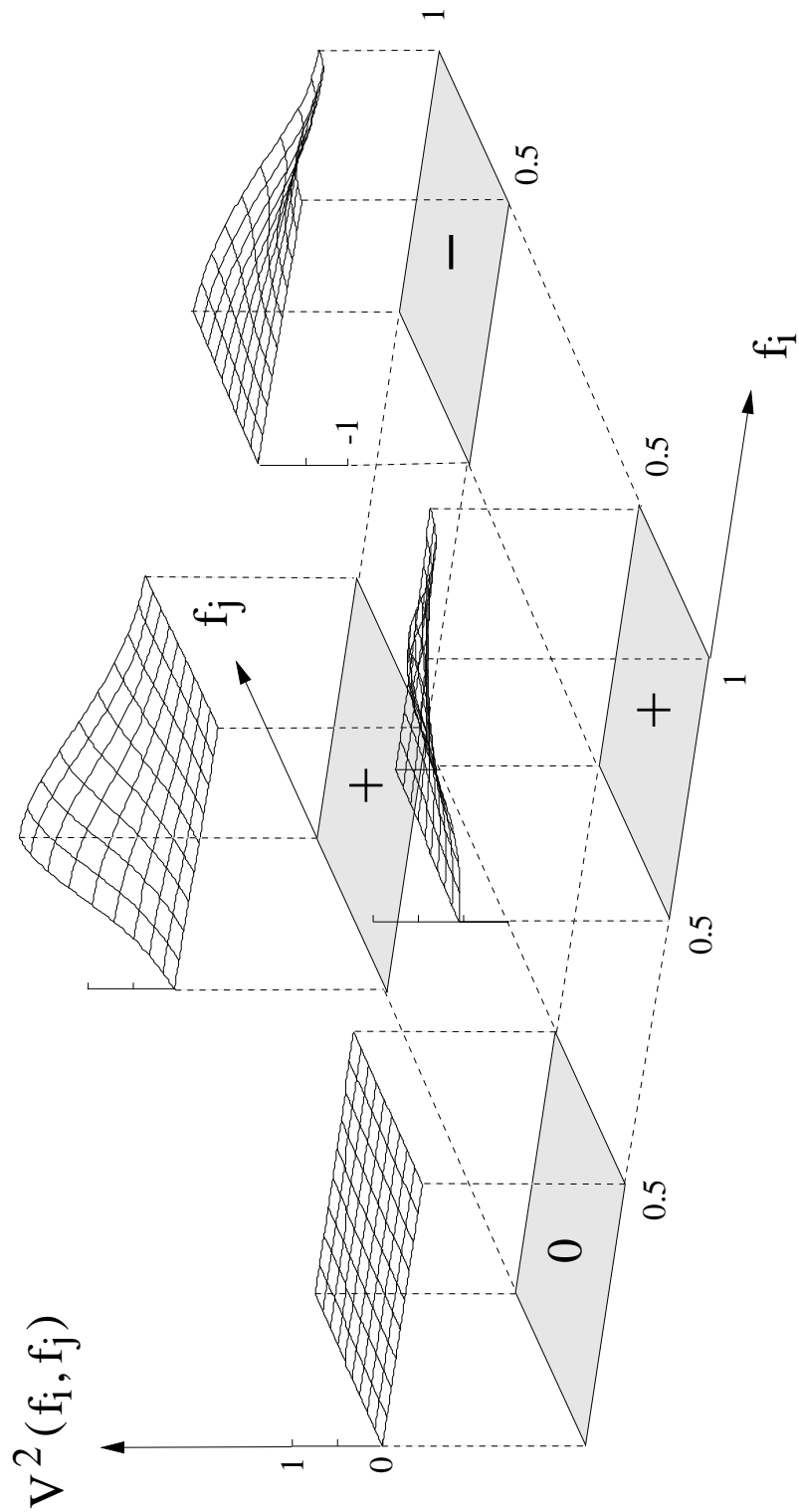


Abbildung 6.11: Basispotential einer zweielementigen Clique in unterstützender Relation

- **Fall 2+3:** Jeweils ein Knoten verfügt über ein Label mit einem Wert > 0.5 , während der zweite Knoten ein Label mit einem Wert ≤ 0.5 besitzt. Die Inkonsistenz der Szeneninterpretation, die sich in der Konkurrenz unter Gruppenthypothesen ausdrückt, ist durch die unterschiedliche Bewertung der Hypothesen auflösbar. Der Beitrag der zweielementigen Clique zur Energiefunktion $U(f)$ soll in diesem Fall negativ sein.
- **Fall4:** Beiden Knoten sind Label zugeordnet, deren Werte ≤ 0.5 sind. Analog zur Unterstützung ist auch in diesem Fall eine relevante Aussage nicht ableitbar. Der Energiebeitrag der zweielementigen Clique ist in diesem Fall ebenfalls neutral.

Der Verlauf des Basispotentials zweielementiger Cliquen in konkurrierender Relation entspricht der Negation des Basispotentials zweielementiger Cliquen V^2 in unterstützender Relation in Gleichung 6.24:

$$-V^2 = \begin{cases} 0 & : 0 \leq f_i \leq 0.5 \quad \text{und} \quad 0 \leq f_j \leq 0.5 \\ -\kappa \Gamma & : 0 \leq f_i \leq 0.5 \quad \text{und} \quad 0.5 < f_j \leq 1 \quad \text{mit} \quad \mu_1 = 0, \quad \mu_2 = 1 \\ \kappa \Gamma & : 0.5 \leq f_i \leq 1 \quad \text{und} \quad 0.5 < f_j \leq 1 \quad \text{mit} \quad \mu_1 = 1, \quad \mu_2 = 1 \\ -\kappa \Gamma & : 0.5 < f_i \leq 1 \quad \text{und} \quad 0 \leq f_j \leq 0.5 \quad \text{mit} \quad \mu_1 = 1, \quad \mu_2 = 0 \end{cases} \quad (6.26)$$

Für die Berechnung des speziellen Potentials einer zweielementigen Clique in konkurrierender Relation V_C^{2comp} wird das durch Gleichung 6.26 berechnete Basispotential mit dem Faktor β_2^{comp} gewichtet:

$$V_C^{2comp} = \beta_2^{comp}(-V^2) \quad (6.27)$$

Für die Untersuchung von Konkurrenzrelationen zwischen Geschlossenheitshypothesen gilt eine zusätzliche Gewichtung des Basispotentials mit einem Faktor $1 + \omega$. Der Faktor ω beschreibt den Anteil des Gebiets der Geschlossenheitshypothese mit der kleineren Gesamtfläche, welches von der Hypothese mit der größeren Gesamtfläche überdeckt wird. Das spezifische Potential $V_{C_{clos}}^{2comp}$ einer zweielementigen Clique deren Entitäten Geschlossenheitshypothesen in einer konkurrierenden Relation entsprechen, wird damit zu:

$$V_{C_{clos}}^{2comp} = (1 + \omega)\beta_2^{comp}(-V^2) \quad (6.28)$$

Diese zusätzliche Gewichtung trägt dem Umstand Rechnung, daß verschieden stark überlappte Flächenanteile von Geschlossenheitshypothesen auch eine unterschiedlich starke Konkurrenz beinhalten.

6.2.3.3 Potentialfunktion einer dreielementigen Clique

Für dreielementige Cliques wird keine Differenzierung in Konkurrenz oder Unterstützung vorgenommen. Ihr Beitrag zur Energiefunktion ist auf einen konstanten Wert festgelegt, da sich bisher alle relevanten Aspekte in der Bewertung der Gruppenthypothesen mit Hilfe einelementiger- und zweielementiger Potentialfunktionen modellieren ließen. Wenn es notwendig wird, zusätzliches Wissen in das MRF einzubringen, können diese dreielementigen Potentialfunktionen dazu dienen.

Zusammenfassung Das vorliegende Kapitel befaßte sich mit der Verifikation der generierten Gruppenthypothesen aus Kapitel 4. Als geeigneter Bewertungsmechanismus wurde von mir der Formalismus von Markov Random Fields gewählt, wie er in Kapitel 5 einführend dargestellt wurde.

Die untersuchten Gruppenthypothesen entsprechen den Merkmalen eines Bildes und stellen damit die **irregulären Entitäten** eines Markov Random Field dar. Sie werden durch die Knoten eines ungerichteten Graphen repräsentiert, wie in Abschnitt 6.2.1 beschrieben. Die Kanten des ungerichteten Graphen repräsentieren die **Nachbarschaftsrelationen** Konkurrenz und Unterstützung. **Unterstützung** entsteht, wie in Abschnitt 6.2.1 beschrieben, wenn die Entitäten eine Bestandteilsrelation verbindet. **Konkurrenz** entsteht zwischen Entitäten, die eine mehrdeutige, inkonsistente Interpretation der Szene darstellen. Dieser Sachverhalt ist in Abschnitt 6.2.1 beschrieben.

Die Repräsentation der Gruppenthypothesen durch den ungerichteten Graphen erfolgt im Rahmen der in Abschnitt 4.2 beschriebenen Hierarchie der Hypothesengenerierung. Der beispielhafte Aufbau eines solchen Graphen ist in Abschnitt 6.2.1 geschildert.

Für alle Gruppenthypothesen wird eine **datengetriebenen Signifikanz** ermittelt. Die Berechnung dieser Signifikanzen erfolgt, wie in Abschnitt 6.2.2.1 dargestellt, aufgrund von Heuristiken über die räumlich geometrischen Eigenschaften der unterschiedlichen Gruppenthypothesen. Aus der datengetriebenen Signifikanz wird eine initiale Stabilität errechnet, die den Optimierungsprozeß mittels HCF steuert (siehe dazu auch Abschnitt 5.7.1.3). Die Berechnung der Wahrscheinlichkeit einer Konfiguration des Markov Random Field wird durch die Anwendung des in Abschnitt 5.5 dargestellten **Theorems von Hammersley & Clifford** wesentlich vereinfacht. Der praktische Wert dieses Theorems liegt in der Spezifizierung der Wahrscheinlichkeit einer Konfiguration durch das Design geeigneter Potentialfunktionen, den **Clique Potentialen**. Das Design der Clique-Potentiale bedeutet die Einbringung von a priori Wissen in ein Markov Random Field.

Die von mir vorgestellte Anwendung definiert ein- und zweielementige Clique-Potentiale.

Das in Abschnitt 6.2.3.1 eingeführte einelementige Clique-Potential erlaubt eine Aussage über die Relation zwischen der datengetriebenen Signifikanz einer Hypothese und ihrer Bewertung durch das Random Field. Zweielementige Clique-Potentiale formulieren die unterstützenden und konkurrierenden Relationen unter den Hypothesen. Sie sind in Abschnitt 6.2.3.2 dargestellt.

Die maximal berücksichtigte Elementanzahl einer Clique liegt bei drei. Der Beitrag dreielementiger Cliques wurde jedoch auf einen konstanten Wert festgelegt, da sich alle relevanten Aspekte der Anwendung durch das Design der ein- und zweielementigen Potentiale modellieren ließen.

Die Darstellung des hier verfolgten Ansatzes zur Generierung und Bewertung von Gruppenhypothesen wird im nachfolgenden Kapitel 7 mit der Beschreibung konkreter Resultate fortgesetzt. Alle gezeigten Bildinhalte entstammen dem in Abschnitt 4.1 geschilderten **Baufix-Szenario**.

7 Resultate und Ausblick

Die Aufgabe dieses Kapitels ist es, die in den vorhergehenden Kapiteln, insbesondere Kapitel 4 “Generierung von Gruppenhypothesen mit Bereichen perzeptiver Aufmerksamkeit” und Kapitel 6 “Ein Markov Random Field für Gruppenhypothesen”, vorgestellten Verfahren anhand der erzielten Resultate noch einmal zusammenzuführen. Diese werden, wie könnte es in der Bildverarbeitung auch anders sein, hauptsächlich in Form von Abbildungen demonstriert. Die Darstellung der Resultate selbst wird im Anhang, Abschnitt B vorgenommen. Da sich aus der Vorstellung der Resultate die Weiterentwicklung des hier verfolgten Ansatzes gut ableiten läßt, wird ein Ausblick das Kapitel abrunden.

7.1 Vorstellung der Resultate

Für eine Beurteilung der hier gezeigten Resultate ist es angebracht, sich die Aufgabe der perzeptiven Gruppierung in der Bildverarbeitung noch einmal vor Augen zu führen. Diese Verfahren sind in der Regel eingebunden in Bildverarbeitungssysteme, die der Erkennung von Objekten einer Szene dienen. Das beispielsweise in Abschnitt 3.2.1 vorgestellte System “SCERPO”, von David Lowe, hat die Erkennung dreidimensionaler Objekte zum Ziel. Dem System ist ein dreidimensionales Linienmodell des gesuchten Objekts bekannt. Dieses wird mit den resultierenden Hypothesen des Gruppierungsprozesses verglichen. Die Aufgabe der perzeptiven Gruppierung kann im Rahmen eines solchen Verfahrens als “Dienstleistung” der nachfolgenden Objekterkennung verstanden werden, indem sie eine Reduktion der Eingangsdaten bei gleichzeitiger Extraktion signifikanter Strukturen der Szene vornimmt. Diese signifikanten Strukturen erfüllen idealerweise das von Witkin und Tenenbaum postulierte “principle of non-accidentalness” (siehe dazu auch Einleitung zum Kapitel 3). Dieses Prinzip besagt, daß das Erscheinen von raum-zeitlicher Kohärenz oder Regularität in der Regel nicht zufällig auftritt, sondern auf eine gemeinsame Ursache oder Prozeß in der Szene deutet. Eine Reduktion der Eingangsdaten bezieht sich in dem hier vorgestellten Ansatz zum einen auf eine wirksame Verringerung der initialen Primitiva und erzeugten

Gruppenhypothesen, zum anderen gilt es, die durch den Gruppierungsprozeß auftretenden Mehrdeutigkeiten zu reduzieren. Eine Art diese Reduktion zu erreichen, kann durch die Bildung resultierender Gruppenhypothesen hoher Komplexität erfolgen: im Fall des oben erwähnten Systems von Lowe handelt es sich bei den resultierenden Hypothesen beispielsweise um Trapezoide, stückweise Parallelitäten und Schiefsymmetrien (siehe dazu auch Abbildung 3.2). Für den hier vorgestellten Ansatz werden unter den resultierenden Hypothesen primär Geschlossenheiten verstanden. Geschlossenheiten repräsentieren im Idealfall die Umrisse eines Objekts, wie sie durch seine zweidimensionale Projektion in der Szene entstehen. Bei den nachfolgenden Betrachtungen wird in diesem Zusammenhang vereinfachend von Objekten oder Objektumrissen die Rede sein, soweit nicht anders erwähnt, ist in diesem Fall immer die zweidimensionale Projektion gemeint.

Die Bewertung der Gruppenhypothesen und initialen Primitiva ist eine weitere Art die oben erwähnte Reduktion zu erreichen. Durch eine Betrachtung der Gruppenhypothesen im globalen Zusammenhang der Szene sollen gleichzeitig im Gruppierungsprozeß entstandene Mehrdeutigkeiten aufgelöst werden, wie in Kapitel 6: “Ein Markov Random Field für Gruppenhypothesen”, beschrieben.

Die Beurteilung der Ergebnisse erfolgt hier in qualitativer Form, indem die gefundenen Resultate anhand der eigenen Wahrnehmung gemessen werden. Eine Vorgehensweise, wie sie auch von Li (siehe dazu auch [14], Seite 28) vorgeschlagen wird: Eine gefundene optimale Lösung kann nur so gut oder schlecht sein, wie das Programm, das zur Lösung eingesetzt wurde. Wenn eine Differenz in der maschinellen Lösung und der eigenen Wahrnehmung einer optimalen Lösung auftritt, rät Li deshalb dazu, der eigenen Wahrnehmung zu trauen.

7.1.1 Schema zur Repräsentation

Die Darstellung der Resultate erfolgt in einheitlicher Form im Anhang, Abschnitt B. Für jedes Bild ist die Darstellung auf zwei Seiten aufgeteilt. Die erste Seite zeigt das Grauwertbild und die initialen Primitiva des Gruppierungsprozesses. Eine Tabelle gibt Auskunft über Art und Anzahl der initialen Primitiva und der hypothetisierten Gruppen. Die zweite Seite zeigt drei Abbildungen, welche die als signifikant bewerteten Gruppenhypothesen der 1D-, 2x1D- und 2D-Ebene zeigen.

Die Daten zum Verlauf des Bewertungsprozesses, wie beispielsweise die Anzahl der erzeugten Knoten und Kanten, sind in drei Tabellen in Abschnitt 7.1.5.1 zusammengestellt.

7.1.2 Grauwertbilder

Die für die Vorstellung der Resultate ausgewählten Grauwertbilder entsprechen der in Abschnitt 4.1 vorgestellten Domäne: Die Abbildungen B.1, B.3, B.7, B.9 und B.11 zeigen die Konstruktionsteile in Einzellage. Die Abbildungen B.5, B.15 und B.17 enthalten Teilmontagen. Die Abbildung B.13 zeigt das komplette Flugzeug.

Die abgebildeten Objekte sind in ihrer Eignung für die Verarbeitung durch ein Bildverarbeitungssystem nicht so einfach, wie ihre Herkunft aus der Spielzeugwelt vordergründig suggeriert. Beispielsweise sind die Umrißkonturen aller Objekte nicht ideal rechteckig sondern abgerundet. Dies gilt insbesondere für den sogenannten Schraubwürfel, ein auf einer kugeligen Grundform aufgebautes Objekt mit 6 abgeflachten Seiten. Für den Schraubwürfel sind in der Regel nur die Ellipsen der abgeflachten Seiten vorhanden, wie in Abbildung B.7b erkennbar. Die schließenden Ecken des Würfels sind virtuell und nur für den Beobachter vorhanden, der um die Bedeutung des Objekts weiß. Die Objekte sind entweder lasiert oder naturfarben: in jedem Fall ist die inhärente Holzstruktur erhalten, wie es besonders gut in Abbildung B.17a, b beobachtet werden kann. Einige Objekte besitzen Texturen, wie beispielsweise die Schraube mit Außengewinde und die Rautenmutter mit Innengewinde in Abbildung B.9a, b oder der Reifen in Abbildung B.11a, b. Alle diese Texturen führen in der Regel zu einer großen Ansammlung kleiner Primitiva. Alle Bilder zeigen die Objekte leicht perspektivisch, sodaß Schattenkonturen entstehen, wie beispielsweise an den Seiten der Lochleisten oder im Inneren der Leistenlöcher (siehe dazu auch Abbildung B.1a, b). Für ein kantenbasiertes Verfahren, wie es hier eingesetzt wird, stellen die aufgezählten Eigenschaften der Objekte und der Aufnahme also durchaus realistische Anforderungen dar.

7.1.3 Initiale Primitiva

Die initialen Primitiva stellen die Eingangsdaten des Gruppierungsprozesses dar. Sie werden mit Hilfe des in Abschnitt 4.1 beschriebenen Vorgehens erzeugt. Der Inhalt der untersuchten Szenen kann grob in zwei Hauptgruppen unterschieden werden: solche die primär durch Geradenstücke approximiert werden und solche die primär durch Ellipsensegmente approximiert werden. Typische Vertreter der ersten Gruppe sind die Umrisse der Leisten oder der Rautenmutter. Für die zweite Gruppe sind die Löcher der Leisten oder der Schraubwürfel typisch. Für beide Hauptgruppen sind in den Abbildungen 7.1a, b Beispiele gezeigt. Diese beiden Hauptgruppen werde ich nachfolgend auch mit den Begriffen **winklige Elemente** und **runde Elemente** der Szene unterscheiden. Alle Objekte enthalten

sowohl winklige als auch runde Elemente: Je nachdem ob der Umriß eines Objekts eher zur Erzeugung von Geradenstücken (Leiste, Rautenmutter) oder eher zur Erzeugung von Ellipsensegmenten (Schraubwürfel, Unterlegscheibe) neigt, werde ich in Zukunft auch von **winkligen Objekten** im Unterschied zu **runden Objekten** sprechen.

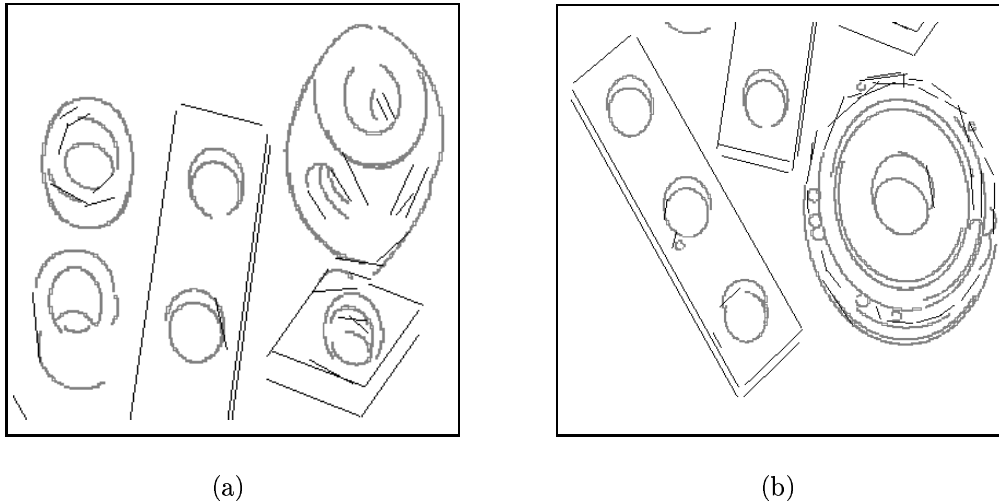


Abbildung 7.1: Bestimmte Szeneninhalte werden bevorzugt durch Geradenstücke, bzw. durch Ellipsensegmente approximiert. Zwei Ausschnitte aus Bild 6 in Abbildung B.11b sollen dies verdeutlichen:

- a) Die Umrisse von Leiste und Rautenmutter werden bevorzugt durch Geradenstücke approximiert. Die Umrisse von Schraubwürfel, Unterlegscheibe, Buchse, sowie die Löcher von Leiste und Rautenmutter, werden bevorzugt durch Ellipsensegmente approximiert.
- b) Die Umrisse der Leisten werden bevorzugt durch Geradenstücke approximiert, während das Rad und die Löcher der Leisten bevorzugt durch Ellipsensegmente approximiert werden.

7.1.4 Resultierende Gruppenhypothesen

Der Prozeß der Generierung der Gruppenhypothesen ist hierarchisch, wie in Abschnitt 4.2 vorgestellt. Die Bildung der Hypothesen erfolgt auf der Grundlage von Suchbereichen, wie in Abschnitt 4.3.3 beschrieben. Die Erzeugung der verschiedenen Gruppenhypothesen berücksichtigt die Auswertung weiterer Randbedingungen, wie in Abschnitt 4.4.1 dargestellt. Kollineare und kurvilineare Gruppenhypothesen werden durch eine erneute

Approximation der zugrundeliegenden Konturpunktketten repräsentiert (siehe auch Abschnitt 4.4). Alle anderen Gruppen werden durch die ihnen angehörenden Bestandteile, initiale Primitiva, kollineare und kurvilineare Hypothesen, dargestellt. Die Darstellung der Ergebnisse des Gruppierungsprozesses wurde entsprechend der Gruppierungshierarchie in Hypothesen der 1D-, 2x1D- und 2D-Ebenen vorgenommen. Alle dargestellten Gruppierungshypothesen besitzen eine Signifikanz die ≥ 0.8 ist. Dieser Wert wurde von mir auf Grundlage der Betrachtung der Resultate festgelegt. Die Abbildungen zeigen die resultierenden Gruppierungshypothesen der 1D-, 2x1D- und 2D-Ebene.

Um die Stärken und Schwächen des hier vorgestellten Ansatzes besser einschätzen zu können, sollen nachfolgend einige Eigenschaften einer näheren Betrachtung unterzogen werden.

7.1.4.1 Runde- und winklige Elemente der Szene

In Abschnitt 7.1.3 wurde beschrieben, wie der Inhalt der betrachteten Szenen in winklige und runde Elemente unterteilt werden kann. Bei der Betrachtung der resultierenden Gruppierungshypothesen aller Ebenen fällt auf, daß der Inhalt der Szene der den winkligen Elementen entspricht, sehr vollständig erscheint, während beispielsweise fast alle Löcher der Lochleisten fehlen. Löcher werden sehr oft durch fast geschlossene Ellipsenbögen approximiert. Durch das stetige Verkleinern des Suchbereiches ab der Hälfte des Umfangs des approximierten Ellipsensegments (siehe dazu auch Abschnitt 4.3.2.3) werden für fast geschlossene Bögen in der Regel keine kurvilinearen Hypothesen gebildet. In den Abbildungen 7.2a und 7.2b ist gut zu beobachten, wie die Löcher der Leisten und des Schraubwürfels, die in der Abbildung der initialen Primitiva gut zu erkennen sind, in der Abbildung der 1D-Ebene fast völlig fehlen. Diese Löcher kontrastieren sehr gut zu ihrer Umgebung, sei es durch scharfe Schattenkonturen oder durch den andersfarbigen, durchscheinenden Grund. Sie können sehr gut segmentiert und in einem Ellipsenbogen approximiert werden. Die Umrißkonturen des Würfels, der Unterlegscheibe oder der Buchse kontrastieren im Gegensatz dazu in den Abbildungen nicht so stark mit ihrer Umgebung. Ihre Approximation zerfällt deshalb in eine Reihe kleinerer Ellipsensegmente und Geradenstücke, die anschließend in kurvilinearen und Nähe-Hypothesen gruppiert werden.

7.1.4.2 Auflösung von Verdeckungen

Der hier vorgestellte Ansatz ist in der Lage Verdeckungen durch Überlagerung von Objekten aufzulösen, wie in Abbildung 7.3 gezeigt. Zum aktuellen Zeitpunkt findet eine solche

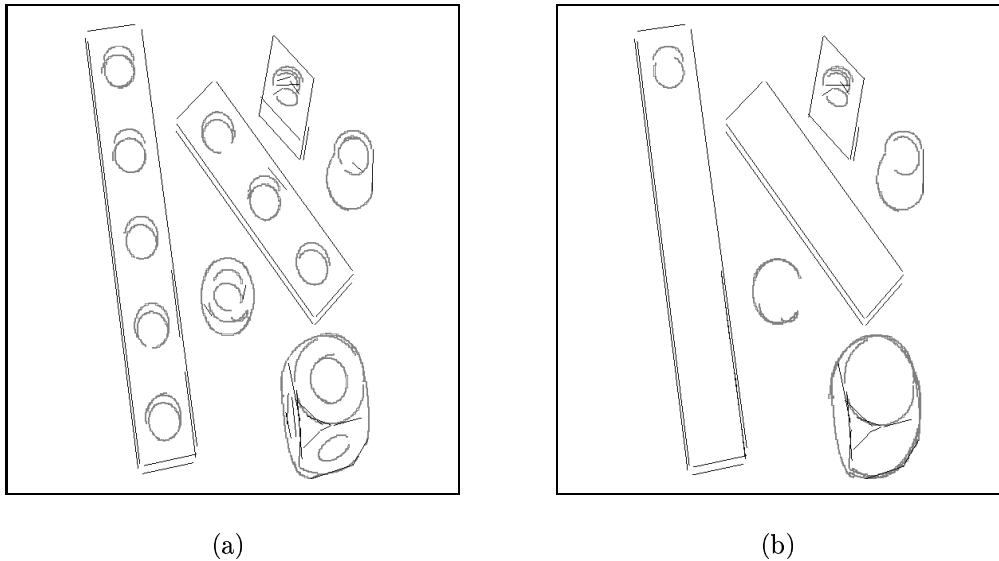


Abbildung 7.2: In den resultierenden Gruppenhypothesen der 1D-Ebene ist der Szeneninhalt der winkligen Elementen entspricht sehr vollständig vorhanden, im Gegensatz zu runden Elementen:

- a) Die initialen Primitiva zeigen deutlich die winkligen und runden Elemente der Szene.
- b) In der Darstellung der resultierenden Hypothesen der 1D-Ebene fehlen die Löcher der Leisten und des Schraubwürfels fast vollständig.

Auflösung von Verdeckungen nur für Leisten statt, wie die Abbildung beispielhaft demonstriert. Die Auflösung einer Verdeckung beruht auf der Tatsache, daß eine kollineare Gruppenhypothese generiert werden kann, die in der Lage ist die Lücke zu schließen. Dies ist sehr gut an der verdeckten Leiste in Abbildung 7.3b,c) zu erkennen. Wie in den Abbildungen 4.12d und 7.3d zu erkennen, zeichnen sich kollineare Suchbereiche durch eine starke Ausdehnung in vertikaler Richtung aus. Wie in Abschnitt 4.4 beschrieben, werden diese Suchbereiche zusätzlich in Abhängigkeit vom untersuchten Geradenstück skaliert. Dies bedeutet für Leisten, deren Umrisse in der Regel durch relativ lange Geradenstücke approximiert werden, daß eine Überbrückung fast immer erfolgt. Können, wie in Abbildung 7.3c alle Lücken einer Verdeckung durch kollineare Gruppen geschlossen werden, ist es möglich, daß auf dieser Basis eine Geschlossenheit generiert wird (siehe dazu auch Abbildung B.6c). Ein weiteres Beispiel für die Auflösung einer Verdeckung ist in Abbildung B.16a und c erkennbar. Dort gelingt es, beide Leistenumrisse des Propellers mit Geschlossenheiten zu beschreiben. Im Fall des Propellers in den Abbildungen B.13b und B.14a und c kann für

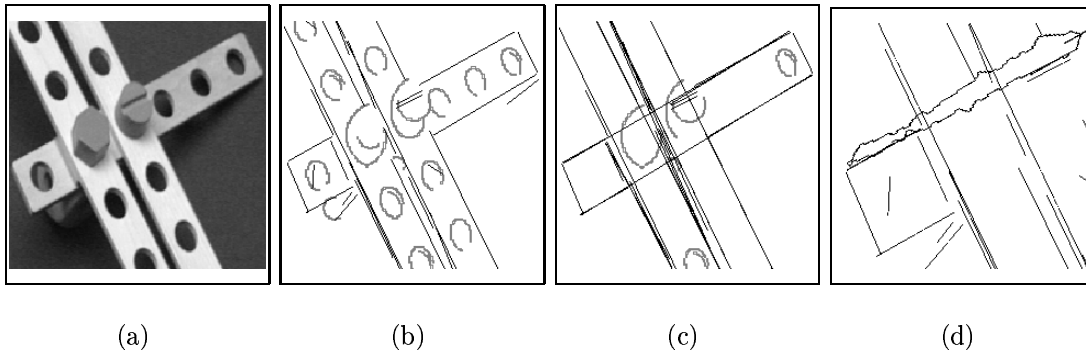


Abbildung 7.3: Auflösung von Verdeckungen am Beispiel eines Ausschnitts aus Bild3 in Abbildung B.5, der die Verschraubung der drei Leisten zeigt: Die untere Leiste ist durch die oberen verdeckt. Die oberen Leisten sind durch die Schraubenköpfe teilweise verdeckt.

- a) Grauwertbild
- b) Initiale Primitiva
- c) Resultierende Hypothesen der 1D-Ebene: für alle drei Leisten konnten kollineare Gruppen generiert werden, auf deren Grundlage Geschlossenheiten erzeugt wurden.
- d) Vergrößerung des Inhalts von a)-d) mit einem kollinearen Suchbereich für den oberen Umriß der verdeckten Leiste.

die obere Leiste, die Verdeckung durch den Schraubenkopf aufgelöst werden. Für die untere Leiste ist jedoch nur für die obere Längsseite eine Überbrückung durch eine kollineare Hypothese gefunden worden. Für die untere Längsseite schlägt dies fehl. In diesem Fall konnte keine Geschlossenheitshypothese generiert werden. In Abbildung B.18a und c ist erkennbar, wie beide Lücken der unteren Leiste durch kollineare Gruppen geschlossen wurden. Die Generierung der Geschlossenheit scheiterte hier aufgrund der großen Lücke im unteren rechten Bereich der Leiste.

7.1.4.3 Generierung von Geschlossenheiten

Geschlossenheiten stellen zur Zeit die komplexesten der resultierenden Hypothesen dar. Geschlossenheiten sind in der Lage, die Eingangsdaten wirksam zu reduzieren und gleichzeitig die signifikante Szeneninformation zu extrahieren: Die Abbildungen 7.4a und b zeigen einen Bildausschnitt mit dem Propeller des Flugzeuges. Durch die inhärente Holzstruktur wurden für beide Leisten viele kleine und kleinste Kontursegmente erzeugt. Die für die

obere Leiste generierte Geschlossenheit kennzeichnet diesen Bereich als signifikant für eine weitere Untersuchung, unter Vernachlässigung des Großteils der für diesen Bereich approximierten Kontursegmente.

Wie in Abschnitt 4.4.2 ausgeführt, beruht die Generierung von Geschlossenheiten auf der Auswertung erzeugter Nähe-Hypothesen und paralleler Hypothesen. Aus diesem Grund folgt eine getrennte Betrachtung beider Typen von Geschlossenheiten.

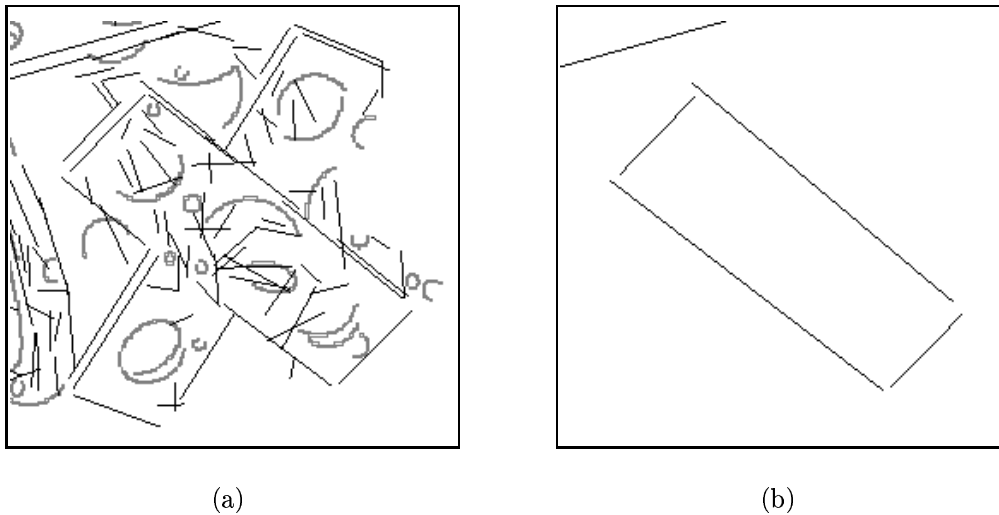


Abbildung 7.4: Die Reduktion der Eingangsdaten durch die Generierung von Geschlossenheiten, am Beispiel des Propellers aus Bild7:

- a) Initiale Primitiva des Propellers. Gut erkennbar ist, daß durch die inhärente Holzstruktur, viele kleine und kleinste Segmente erzeugt werden.
- b) Für die obere Leiste des Propellers wurde eine Geschlossenheitshypothese generiert, welche diesen Bereich als signifikant kennzeichnet.

Geschlossenheiten auf der Basis von Parallelitäten: Die Generierung von Geschlossenheiten auf der Basis von Parallelitäten tritt für alle Umriss auf, die sich zur Hypothesierung von Parallelitäten gut eignen. Dieses sind in erster Linie die Leisten. Für alle Leistenumrisse, für die eine Geschlossenheit hypothetisiert wurde, existiert jeweils mindestens eine auf Basis einer Parallelität. Neben den Leisten treten Geschlossenheiten auf der Basis von Parallelitäten nur bei den Rautenmuttern auf. Beispielsweise konnte für die Rautenmuttern in den Abbildungen B.8c und B.10c eine Geschlossenheit auf der Basis einer Parallelität hypothetisiert werden.

Geschlossenheiten auf der Basis von Nähe-Gruppen: Geschlossenheiten auf Grundlage von Nähe-Gruppen treten für die oben erwähnten Umrisse von Leisten und Rautenmuttern nur dann auf, wenn die zugrundeliegenden Kontursegmente nicht identisch mit denen einer Geschlossenheitshypothese auf Basis einer Parallelität sind. Treten Geschlossenheiten mit identischen Kontursegmenten auf, werden die Hypothesen auf der Basis von Nähe-Gruppen aussortiert, wie in Abschnitt 4.4.2 beschrieben. Geschlossenheiten auf Grundlage von Nähe-Gruppen treten für diese also nur dann auf, wenn ihnen ein andere Szeneninformation zu Grunde liegt. Dies ist beispielsweise für die Teilmontage aus Abbildung B.5a der Fall. Diese Teilmontage enthält drei Leisten, eine “Rumpfleiste” und zwei “Flügelleisten”. Für diese drei Leisten wird jeweils eine Geschlossenheit auf der Basis einer Parallelität erzeugt, wie in in Abbildung 7.5a gezeigt. Zusätzlich wird eine auf den erzeugten Nähe-Hypothesen der Geradenstücke basierende Geschlossenheit gebildet, welche die hintere “Flügelleiste” und einen Teil der “Rumpfleiste” umfaßt, wie in Abbildung 7.5b gezeigt. Auch wenn diese generierte Geschlossenheit bei näherer Betrachtung nicht wünschenswert ist, ist es trotzdem sinnvoll solche Geschlossenheiten zu erzeugen, da sie in der Lage sind andere Zusammenhänge zu detektieren.

Aufgrund der kleinen Suchbereiche für Nähe-Gruppen und der angewendeten Randbedingung (siehe dazu auch Abschnitt 4.4.1), können Geschlossenheiten auf der Basis von Nähegruppen nur hypothetisiert werden, wenn die einzelnen Kontursegmente der Geschlossenheit wie auf einer Perlenschnur “aufgereiht” sind. Dies trifft in der Regel auf Geradenstücke zu, wie sehr gut an den Leistenumrissen in Abbildung 7.2b erkennbar. Die approximierten Ellipsensegmente neigen im Gegensatz dazu zur Überlappung mit anderen Ellipsensegmenten oder Geradenstücken. Trotz der Überlappung liegen die Endpunkte der Segmente oft relativ weit entfernt, sodaß keine Nähe-Gruppen erzeugt werden. In Abbildung 7.6a ist für den Umriß der Raute aus Bild B.11 eine Abfolge von Geradenstücken und einem Ellipsensegment gezeigt: für das Ellipsensegment wird jedoch keine Nähe-Gruppe hypothetisiert (siehe dazu auch Abbildung 7.6b). Diese beschriebene Neigung zur Überlappung der Ellipsensegmente wirkt sich negativ auf die Bildung von Geschlossenheiten für die in Abschnitt 7.1.3 eingeführten runden Objekte der Szene aus. In den untersuchten Szenen wurden lediglich zwei Geschlossenheiten für Schraubwürfel und eine für das Loch einer Leiste gebildet. Wie in den Abbildungen B.10c, B.12c und B.18c erkennbar, sind hier die Segmente gut “aufgereiht”. Da im derzeitigen Stand keine Generierung von symmetrischen Gruppen, analog zu Parallelitäten, implementiert ist, gibt es keine weitere Unterstützung für die Bildung von Geschlossenheiten für runde Objekte.

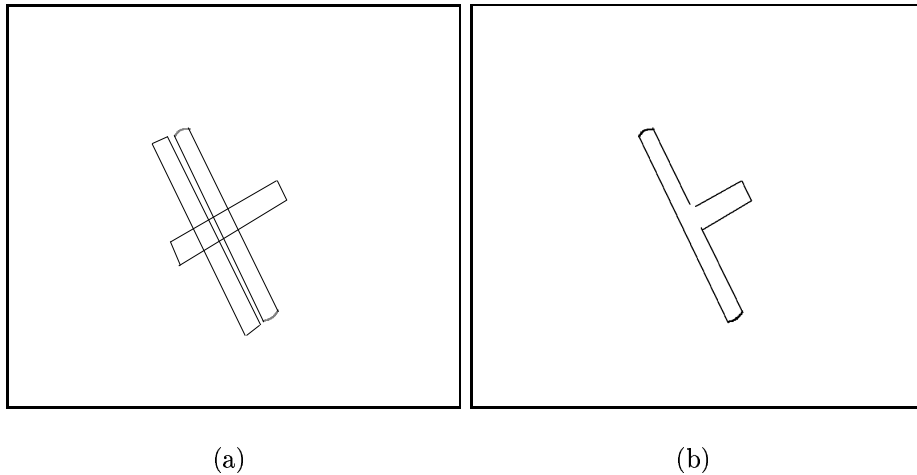


Abbildung 7.5: Hypothesierte Geschlossenheiten für die Teilmontage aus Abbildung B.5a:

- a) Drei auf der Basis von Parallelitäten hypothesierte Geschlossenheiten für die "Flügel" und den "Rumpf".
- b) Geschlossenheit auf der Basis von Nähegruppen für die hintere "Flügel- leiste" und einen Teil des "Rumpfes"

7.1.4.4 Zusammenfassung

Die in den letzten Abschnitten aufgeführten Stärken und Schwächen des gewählten Ansatzes sollen in der nachfolgenden Tabelle 7.1 noch einmal quantitativ zusammengefaßt werden. In der Tabelle sind für die einzelnen Bilder die gefundenen Geschlossenheitshypothesen mit einer Signifikanz ≥ 0.8 in Relation zu den im Bild vorhandenen Objekten aufgelistet. In Relation meint, ob für die umschreibende Kontur eines Objekts mindestens eine Hypothese generiert wurde, nicht ihre Anzahl. Mehrdeutige Hypothesen, wie sie gerade bei Leisten durch Schattenkonturen auftreten, finden hier keine Berücksichtigung. Deshalb ist die detektierte Anzahl der Geschlossenheitshypothesen der Resultatbilder oft höher (siehe dazu auch die Tabellen B.5, B.7, B.8 und B.9). Die durch die Geschlossenheitshypothesen beschriebenen Objekte sind in winklige Objekte und runde Objekte (siehe dazu auch Seite 7.1.3) unterschieden. Wenn Umrißkonturen trotz einer Überlagerung durch ein anderes Objekt ergänzt werden konnten, ist dies unter Überbrückung mit "ja" vermerkt. Für Bilder mit Objekten in Einzellage ist es leicht, die Anzahl der Objekte anzugeben. Bei Bildern mit Teilmontagen wurde die Anzahl der sichtbaren Objekte angegeben. Für das komplett montierte Flugzeug wurde auf diese Angabe verzichtet.

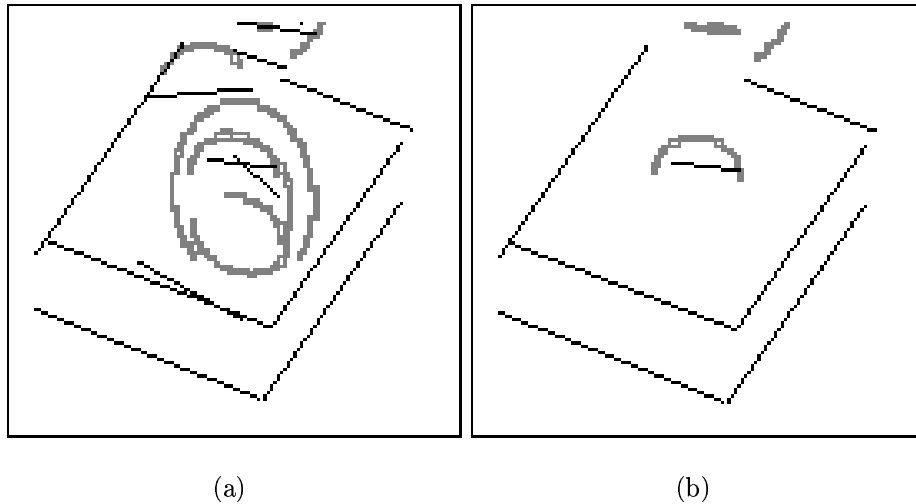


Abbildung 7.6: Vergrößerung der Rautenmutter aus Abbildung B.11:

- a) Initiale Primitiva der Rautenmutter. In der oberen Kante ist ein Ellipse-segment erkennbar, welches die Geradenstücke stark überlappt.
- b) Hypothetisierte 1-D Gruppen für denselben Bildausschnitt. In der oberen Ecke ist eine Lücke entstanden, weil keine Nähegruppen hypothetisiert werden konnten

Mit Ausnahme des Flugzeugs, kann für alle Bilder festgestellt werden, daß im Schnitt für die Hälfte der sichtbaren Objekte Geschlossenheitshypothesen, entsprechend ihrer Umrisse, erzeugt wurden. Betrachtet man die sichtbaren Oberflächen aller Objekte genauer, ist gut zu erkennen, daß der durch Geschlossenheitshypothesen beschriebene Anteil deutlich höher ist. Dies bedeutet, daß in der Regel für einen Großteil des Szeneninhalts eine Beschreibung generiert werden konnte, welche der eingangs aufgestellten Forderung nach höchstmöglicher Komplexität entspricht. Betrachtet man den Typ (winklig, rund) der Objekte näher, ist auffällig, daß winklige Objekte und hier insbesondere Leisten, gut detektiert werden. Die Ursachen für diese Bevorzugung wurden in Abschnitt 7.1.4.1 bei der Betrachtung der runden und winkligen Elemente eingeführt und für die Bildung von Geschlossenheiten auf der Basis von Nähe-Gruppen in Abschnitt 7.1.4.3 nochmals vertieft.

In den Abbildungen mit Teilmontagen in B.5a, B.15a, B.17a und mit einem vollständig montierten Flugzeug in Abbildung B.13a ist auffällig, daß Leisten besonders wichtigen Bauteilen des Flugzeuges entsprechen. Durch Leisten werden der Rumpf, die Flügel, der Propeller und das Leitwerk konstruiert. Deshalb kann abschließend festgestellt werden, daß der implementierte Mechanismus durchaus in der Lage ist, wesentliche Bereiche in einer Szene herauszufiltern. Die signifikanten Strukturen (siehe auch “principal of non-

Abb.	#Objekte	#Geschlossenheiten	winklige Objekte	runde Objekte	Überbrückung
Bilder mit Objekten in Einzellage					
B.1	4	3	Leiste, Rautenmutter		–
B.3	10	4	Leiste Rautenmutter		–
B.7	6	3	Leiste Rautenmutter		–
B.9	4	4	Leiste Rautenmutter	Schraubwürfel	–
B.11	6	3	Leiste	Schraubwürfel	–
Bilder mit Teilmontagen					
B.5	5	3	Leiste		ja
B.15	3	2	Leiste		
B.17	3	1	Leiste	Leistenloch	
Flugzeug					
B.13		4	Leiste		ja

Tabelle 7.1: Gegenüberstellung von Objektanzahl zu Geschlossenheitshypothesen mit einer Signifikanz ≥ 0.8 :

In der ersten Spalte ist die Anzahl der sichtbaren Objekte in der Szene angeführt. Die zweite Spalte gibt an, für wieviele der Objekte Geschlossenheitshypothesen erzeugt werden konnten. Die dritte und vierte Spalte, beschreibt für welche Art von Objekt, rund oder winklig, die Hypothesen generiert wurden. In der letzten Spalte wird erwähnt, ob ein teilweise verdecktes Objekt erfolgreich ergänzt werden konnte. Dies trifft selbstverständlich nur auf Bilder mit Teilmontagen oder das Flugzeug zu.

accidentalness”) wurden zu einem Großteil detektiert. Dabei ist der Mechanismus auch in der Lage verdeckte Konturen richtig zu ergänzen.

7.1.5 Bewertung der Hypothesen durch das Markov Random Field

Die Bestimmung des Energieminimums wird mit Hilfe des **Highest Confidence First** vorgenommen, wie in Abschnitt 5.7.1 ausführlich vorgestellt. Die Gruppenhypothesen und

ihre Primitiva bilden die Knoten eines ungerichteten Graphen, dessen Kanten den Relationen Unterstützung und Konkurrenz entsprechen. Der Aufbau dieses ungerichteten Graphen ist in Abschnitt 6.2.1 beschrieben.

Der Verlauf des Bewertungsprozesses kann anhand der gebildeten Kanten und Knoten und der durchlaufenen Iterationen betrachtet werden. Dies ist das Thema des folgenden Abschnitts. Im Anschluß daran wird die Auflösung von Mehrdeutigkeiten durch das Random Field beispielhaft für Geschlossenheitshypothesen diskutiert.

7.1.5.1 Kanten, Knoten und Iterationen

Kennzeichnend für den Verlauf des Bewertungsprozesses sind die generierten Knoten und Kanten des ungerichteten Graphen, sowie die Anzahl der durchlaufenen Iterationen. Diese Daten sind übersichtsartig für alle Bilder in den Tabellen 7.2, 7.3 und 7.4 zusammengefaßt. Zu den in den Tabellen angegebenen Prozentzahlen ist zu sagen, daß sie jeweils auf ganze Zahlen gerundet wurden. Sie sollen an der jeweiligen Stelle nur eine Vorstellung von der Größenordnung vermitteln.

Tabelle 7.2 listet die Anzahl der gebildeten Knoten und das Verhalten der Signifikanz

Bild	Gesamt #	$f_i^o = f_i$		$f_i^o < f_i$		$f_i^o > f_i$	
1.	87	39	45%	25	29%	23	26%
2.	122	49	36%	59	45%	14	11%
3.	154	69	45%	32	21%	53	34%
4.	144	69	48%	38	26%	37	26%
5.	128	59	46%	32	25%	37	29%
6.	247	84	34%	77	31%	86	35%
7.	433	181	42%	153	35%	99	23%
8.	194	83	43%	42	22%	69	35%
9.	757	254	38%	151	20%	322	35%

Tabelle 7.2: Die Gesamtanzahl der Knoten bezieht sich auf alle Gruppenhypothesen und initialen Primitiva, die Bestandteil einer Gruppenhypothese sind.

$f_i^o = f_i$ listet alle Knoten auf, deren Signifikanz durch die Bewertung unverändert blieb. Analog dazu ist $f_i^o < f_i$ die Anzahl der Knoten, die nach der Bewertung eine höhere und unter $f_i^o > f_i$ eine niedrigere Signifikanz besaßen.

am Knoten auf. Die Gesamtanzahl der Knoten in der zweiten Hauptspalte der Tabelle

bezieht sich auf alle Gruppenhypothesen und alle initialen Primitiva, die Bestandteil einer Gruppenhypothese sind. Initiale Primitiva, die kein Bestandteil einer Gruppe sind, besitzen keine Kanten zu Nachbarknoten und wurden aus diesem Grund nicht in die Betrachtung einbezogen: Sie verändern im Verlauf des Bewertungsprozesses weder ihre eigene Signifikanz (vergleiche dazu auch Abschnitt 6.2.3.1), noch wirken sie verändernd in ihrer Nachbarbarschaft. Die Gesamtanzahl der Knoten spiegelt direkt die Komplexität der untersuchten Szene wider. Sie reicht von 87 Knoten für das erste Bild, bis zu 757 Knoten für das neunte Bild. Diese Komplexität ist abhängig von der Anzahl, Größe und der Anordnung der initialen Primitiva: Im ersten Bild gibt es wenige, relativ lange Primitiva, die primär den Umrissen der Objekte entsprechen (siehe dazu auch Abbildung B.1b). Das neunte Bild enthält eine relativ hohe Anzahl kleiner, dicht beieinanderliegender Primitiva (siehe dazu auch Abbildung B.17b), die durch die Holztextur der Oberfläche erzeugt wurden. Diese neigen in stärkerem Maße zur Erzeugung von Gruppenhypothesen. Im ersten Bild wurden für 111 initiale Primitiva (in Tabelle B.1) 87 Knoten erzeugt, im neunten Bild dagegen für 353 initiale Primitiva (in Tabelle B.9) 757 Knoten.

Die Gesamtanzahl der Knoten wird in den nachfolgenden drei Spalten von Tabelle 7.2 unterteilt in solche, für die das datengetriebene Label bestätigt wurde ($f_i^o = f_i$) und in solche, für die ein höheres ($f_i^o < f_i$), bzw. niedrigeres Label ($f_i^o > f_i$) als das datengetriebene Label vergeben wurde. Im Schnitt wurde bei allen Bildern für ca. 42% der Knoten das datengetriebene Label bestätigt. In den restlichen Fällen wurde je zur Hälfte ein höheres, bzw. niedrigeres Label vergeben. Mehr als der Hälfte der Knoten war am Ende der Optimierung also ein neues Label zugeordnet. Dies zeigt, daß der durch die Modellierung in einem Markov Random Field geschaffene Kontext in der Lage ist, daß Resultat nachhaltig zu beeinflussen. Das Ziel, die Bewertung der lokal generierten Gruppenhypothesen in einem globalen Zusammenhang, wie in Abschnitt 6.1 gefordert, wurde erfolgreich umgesetzt.

Die Gesamtanzahl der Kanten in der zweiten Tabelle 7.3 bezieht sich auf die Gesamtanzahl der Knoten aus Tabelle 7.2. Sie spiegelt, analog zur Knotenanzahl, die Komplexität der untersuchten Szene direkt wider. Von größerem Interesse ist hier jedoch die durchschnittliche Anzahl der Kanten pro Knoten. Dieser Durchschnitt erlaubt eine Aussage über die "Vernetztheit" der Knoten des Graphen. Dieser Wert reicht von minimal 1.3 für das erste Bild, bis zu maximal 3.6 Kanten pro Knoten im neunten Bild. Vergleicht man dazu den Anteil der unterstützenden zu den konkurrierenden Kanten, so ist erkennbar, daß für diese Zunahme insbesondere die Bildung von Konkurrenzen verantwortlich ist. Je mehr Kanten mit einem Knoten verbunden sind, desto höher ist der Berechnungsaufwand der Energiebilanz am Knoten, da er direkt von der Anzahl der Cliques am Knoten abhängig ist.

Die letzte Tabelle 7.4 enthält nur zwei Einträge: die Gesamtanzahl der Iterationen und die

Bild	Gesamt #	Unterstützung	Konkurrenz	pro Knoten		
1.	116	107	92%	3	8%	1.3
2.	169	141	83%	28	17%	1.4
3.	411	252	61%	159	39%	2.7
4.	247	185	75%	62	25%	1.7
5.	194	164	85%	30	15%	1.5
6.	455	331	73%	124	27%	1.8
7.	872	606	69%	266	31%	2.0
8.	485	320	66%	165	34%	2.5
9.	2706	1274	47%	1432	53%	3.6

Tabelle 7.3: Die Gesamtanzahl der gebildeten Kanten wird unterteilt in die konkurrierenden und unterstützenden Kanten. Die durchschnittliche Anzahl der Kanten pro Knoten ist bezogen auf die Gesamtanzahl der Knoten aus Tabelle 7.2.

Bild	Gesamt #	pro Knoten
1.	192	2.2
2.	152	1.2
3.	246	1.6
4.	300	1.6
5.	275	2.1
6.	511	2.1
7.	952	2.2
8.	420	2.2
9.	1395	1.8

Tabelle 7.4: Die Gesamtanzahl der durchlaufenen Iterationen wird bezogen auf die durchschnittliche Anzahl der Iterationen pro Knoten. Die Anzahl der Knoten ist bezogen auf die Gesamtanzahl der Knoten aus Tabelle 7.2.

Anzahl der Iterationen pro Knoten. Die Gesamtzahl der Iterationen ist offensichtlich grob abhängig von der Anzahl der generierten Knoten und Kanten. Für einen Zusammenhang der Anzahl der Iterationen mit der Anzahl der nicht bestätigten datengetriebenen Label in der ersten Hauptspalte, oder mit dem Prozentsatz generierter Konkurrenzen, bzw. Kanten pro Knoten, ließ sich kein Hinweis finden. Die Anzahl der Iterationen pro Knoten läßt eine Aussage über die relative Dynamik des Optimierungsprozesses zu. Wie oft wurde

Hypothese	Bauteil	Art der Geschlossenheit	f_i^o	$f_i \geq 0.8$	Konkurrenz
1.	vorderer Flügel	parallel	0.9	ja	–
2.	hinterer Flügel	parallel	0.9	ja	4.
3.	Rumpf	parallel	0.75	ja	4.
4.	hinterer Flügel, Teil des Rumpfs	Nähe	0.49	nein	2., 3.

Tabelle 7.5: Daten der Geschlossenheitshypothesen aus Abbildung 7.7:

Die Nummer der ersten Spalte entspricht der Hypothese der Abbildung. Die zweite Spalte benennt das Bauteil, dessen Umriß die Geschlossenheit beschreibt. Die dritte Spalte gibt an, ob die Geschlossenheit auf der Auswertung von Nähe-Gruppen oder Parallelitäten beruht. Spalte vier listet die datengetriebene Signifikanz auf und Spalte fünf gibt an, ob die Geschlossenheit nach der Bewertung als signifikant eingestuft ist. Die letzte Spalte gibt an, ob eine Konkurrenz-Relation zu einer anderen Hypothese besteht.

ein Knoten wiederholt besucht und seine Energiebilanz erneut berechnet? Der Wert reicht von minimal 1.2 Besuchen pro Knoten für das zweite Bild, bis zu maximal 2.2 Besuchen pro Knoten für das erste, siebte und achte Bild. Ein Zusammenhang der Iterationen pro Knoten mit irgendeiner der anderen Daten der Tabelle ließ sich nicht feststellen. Sie ist offensichtlich nur von der jeweiligen Verteilung der Label abhängig

7.1.6 Auflösung mehrdeutiger Geschlossenheiten

Der Prozeß der Generierung der Gruppenhypothesen führt durch seine lokale Betrachtungsweise (siehe dazu auch Kapitel 4) zu mehrdeutigen Gruppenhypothesen. Diese Mehrdeutigkeiten bedeuteten eine inkonsistente Interpretation der Szene (siehe dazu auch Abschnitt 6.1). Aus diesem Grund lautete eine der Anforderungen an das Markov Random Field, die Verifikation der Gruppenhypothesen durch einen geeigneten Bewertungsmechanismus (siehe dazu auch Einleitung zu Kapitel 6).

Im folgenden soll die Behandlung mehrdeutiger Hypothesen durch das Markov Random Field anhand der Geschlossenheiten der Teilmontage in Abbildung B.5 und des kompletten Flugzeugs in Abbildung B.13 näher betrachtet werden. Zur besseren Übersicht sind in den Abbildungen 7.7 und 7.8 auf den folgenden Seiten nochmals alle generierten Geschlossenheiten, zusammen mit dem jeweiligen Grauwertbild gezeigt. Die Tabellen 7.5 und 7.6 listen

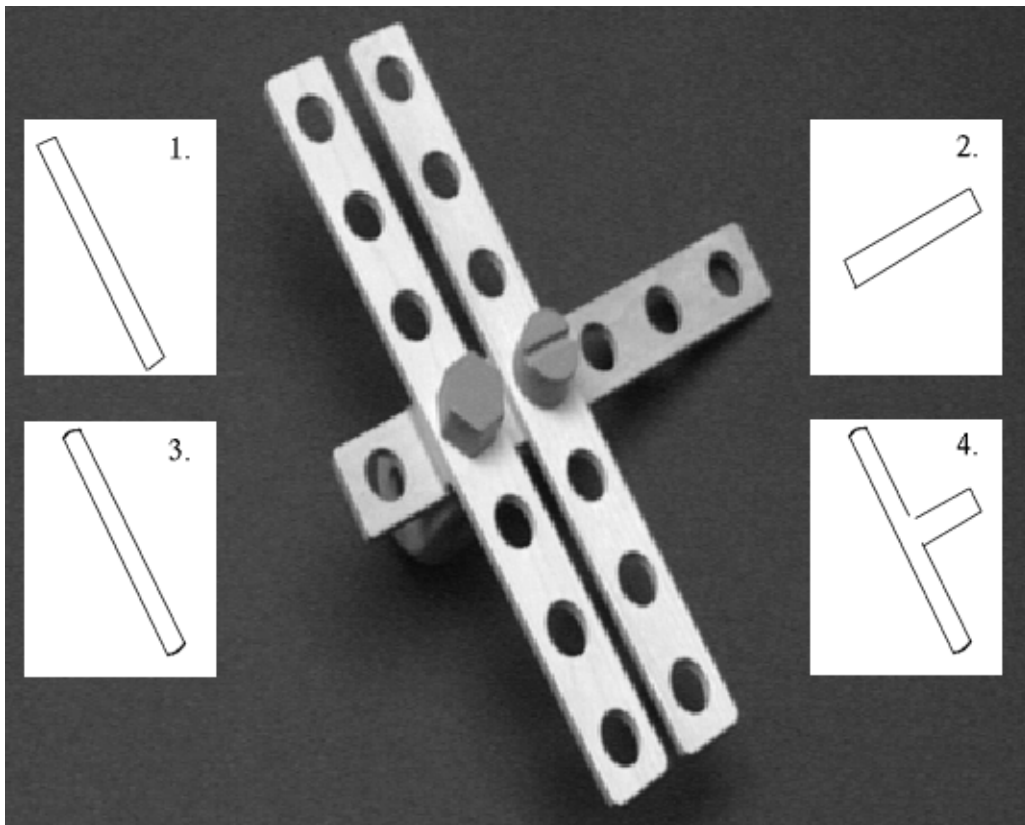


Abbildung 7.7: Die Teilmontage aus Abbildung B.5 mit allen erzeugten Geschlossenheiten.

die relevanten Daten zu den Geschlossenheiten der Abbildungen auf: Von besonderem Interesse sind hier das zugewiesene Label vor und nach der Bewertung, sowie die Existenz von Konkurrenz, da Gruppenthypothesen die eine mehrdeutige Interpretation zulassen, beim Aufbau des ungerichteten Graphen durch eine konkurrierende Kante verbunden werden (siehe dazu auch Abschnitt 6.2.1).

Für die Szene in Abbildung 7.7 werden vier Geschlossenheitshypothesen generiert. Drei Geschlossenheiten beschreiben die Umrisse der beiden Flügel und des Rumpfes. Die vierte Geschlossenheit beschreibt den Umriß des hinteren Flügels und eines Teils des Rumpfes. Diese vierte Gruppe steht in Konkurrenz zu Gruppe zwei (hinterer Flügel) und Gruppe drei (Rumpf). Die vierte Gruppe besitzt das niedrigste datengetriebene Label f_i^o (siehe dazu auch Tabelle 7.5), welches durch die Konkurrenz nach der Bewertung zu Null wird. Dieses Aussortieren von Gruppe vier entspricht durchaus auch dem Verhalten eines menschlichen Beobachters.

Für das Flugzeug in Abbildung 7.8 wurden insgesamt sechs Geschlossenheitshypothesen generiert. Die Geschlossenheiten beschreiben die Umrisse von Leitwerk, Propeller und

Hypothese	Bauteil	Art der Geschlossenheit	f_i^o	$f_i \geq 0.8$	Konkurrenz
1.	Leitwerk	parallel	0.9	ja	–
2.	vorderer Flügel	parallel	0.9	ja	5., 3.
3.	vorderer Flügel, Teil der Schraube	Nähe	0.55	ja	2., 4., 5.
4.	hinterer Flügel,	parallel	0.9	ja	3., 5.
5.	beide Flügel	parallel	0.3	nein	2., 3., 4.
6.	Propeller	parallel	0.75	ja	–

Tabelle 7.6: Daten der Geschlossenheitshypothesen aus Abbildung 7.8:

Die Nummer der ersten Spalte entspricht der Hypothese der Abbildung. Die zweite Spalte benennt das Bauteil, dessen Umriß die Geschlossenheit beschreibt. Die dritte Spalte gibt an, ob die Geschlossenheit auf der Auswertung von Nähe-Gruppen oder Parallelitäten beruht. Spalte vier listet die datengetriebene Signifikanz auf und Spalte fünf gibt an, ob die Geschlossenheit nach der Bewertung als signifikant eingestuft ist. Die letzte Spalte gibt an, ob eine Konkurrenz-Relation zu einer anderen Hypothese besteht.

beiden Flügeln. Für die Flügel wurden mehrdeutige Hypothesen generiert: die fünfte Hypothese umfaßt die Umrissbeiden Flügel. Die dritte Hypothese beschreibt im wesentlichen die Umrissbeiden Flügel unter Einbeziehung eines Teils des Umrisses des hinteren Flügels verbunden durch einen Teil des Schraubenumrisses. Hypothese zwei beschreibt den vorderen und Hypothese vier den hinteren Flügel. Für alle Hypothesen wurden entsprechende konkurrierende Kanten erzeugt (siehe dazu auch Tabelle 7.6). Die fünfte Hypothese besitzt das niedrigste datengetriebene Label f_i^o und konnte aufgrund der Konkurrenzrelation aussortiert werden. Dies gilt nicht für Hypothese drei, bei der die Unterstützung überwog, sodaß nach der Bewertung die maximale Signifikanz zugewiesen war. In diesem Fall wirkte sich eine spezielle Randbedingung zur Ermittlung des Potentials (siehe dazu auch Abschnitt 6.2.3.2, “Konkurrenz”) negativ aus, welche sich an der gegenseitig überdeckten Fläche orientiert. Betrachtet man die Flächen genauer, ist gut zu erkennen, daß Hypothese fünf alle konkurrierenden Hypothesen vollständig überdeckt, im Gegensatz zu Hypothese drei. Da die Gewichtung des konkurrierenden Potentials aufgrund der gegenseitig überdeckten Fläche grundsätzlich sinnvoll ist, ist anderweitig Abhilfe zu schaffen. Geht man davon aus, daß ein menschlicher Beobachter die Gruppen drei und fünf zugunsten von zwei und vier aussortiert hätte, ist das Ziel also nur zum Teil erreicht worden.

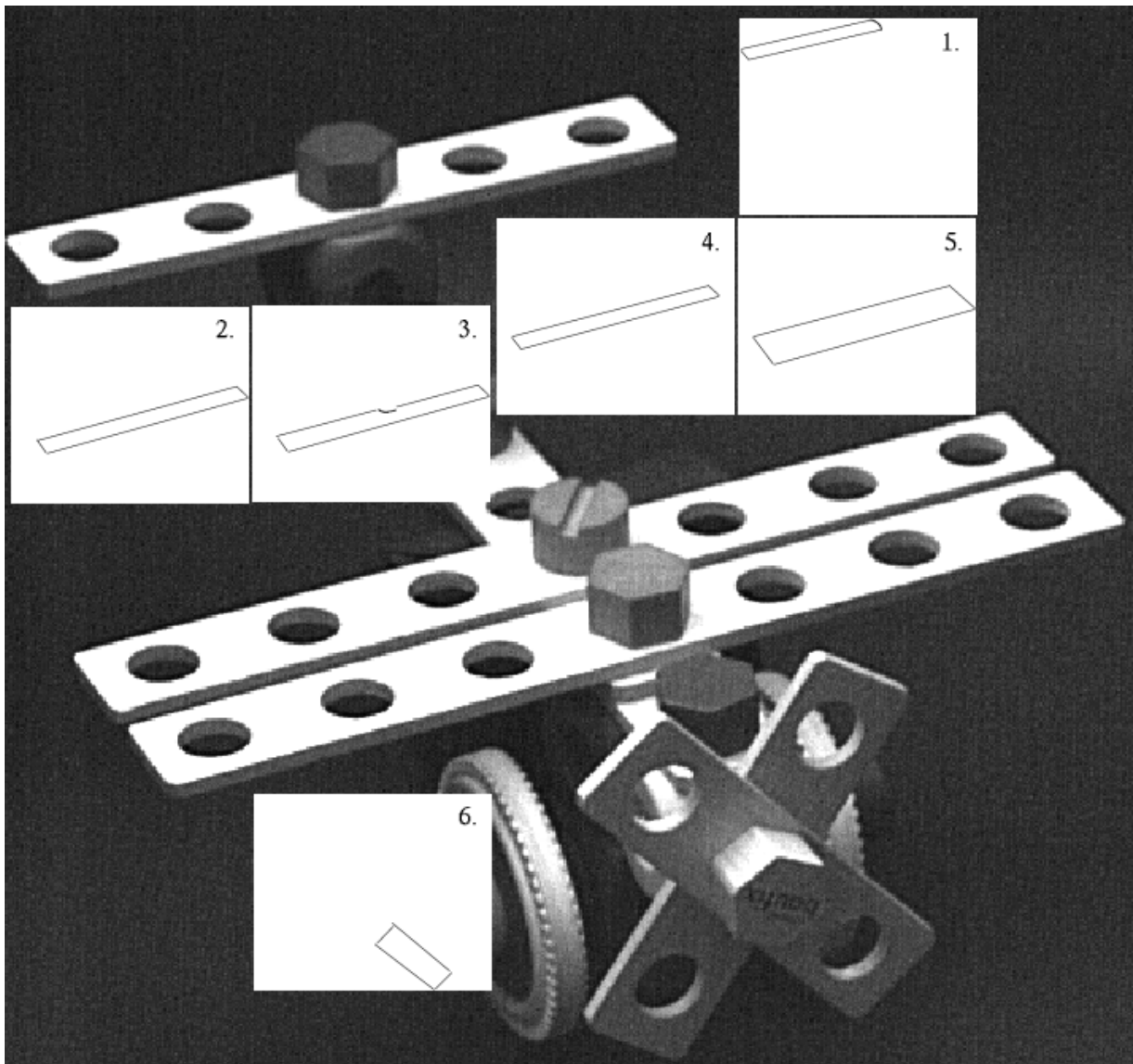


Abbildung 7.8: Die Teilmontage aus Abbildung B.13 mit allen erzeugten Geschlossenheits-hypothesen.

Abschließend kann festgehalten werden, daß eine Auflösung mehrdeutiger Gruppenthypothesen, hier am Beispiel von Geschlossenheiten, erfolgreich stattfindet. Dabei sind einzelne Mechanismen, und hier besonders die Bildung von Konkurrenzrelationen, noch zu optimieren.

7.2 Ausblick

Sinnvolle Erweiterungen des in den vorangehenden Kapiteln vorgestellten Ansatzes bestehen zum einen in direkten Verbesserungen, wie sie sich aus der Besprechung der Resultate ableiten lassen. Zum anderen sind grundsätzliche Erweiterungen im Einsatz der Gestaltgesetze denkbar. Diese können aus einem Vergleich der Gestaltgesetze in Kapitel 2 und ihrer Umsetzung in der Bildverarbeitung in Kapitel 3 (unter besonderer Berücksichtigung der Tabelle in Abbildung 3.1) mit dem von mir beschriebenen Ansatz in den Kapiteln 4 und 6 abgeleitet werden. Vor dem Hintergrund der Eingliederung meiner Arbeit, entsprechend der Tabelle in Abbildung 3.1, in den Bereich zweidimensionaler Bilddaten (siehe dazu auch Abschnitt 3.1), möchte ich mich im nachfolgenden auf die Erweiterungen beschränken, die innerhalb dieses Rahmens angesiedelt sind.

Diese Erweiterung des Bestehenden kann unter verschiedenen Schwerpunkten betrachtet werden: Der Einsatz der Gestaltgesetze kann erweitert werden, zum einen durch die Einbeziehung neuer Primitiva in den Gruppierungsprozeß. Zum anderen durch eine Erweiterung der Betrachtung für die bereits untersuchten Primitiva. Die Bewertung der Gruppenhypothesen ist ein weiterer Aspekt, der zu untersuchen ist. Die verschiedenen Maßnahmen, ermöglichen teilweise in ihrer Kombination die Einbeziehung weiterer Gestaltgesetze. Aus diesem Grund wird, nach einer einleitenden Aufzählung, eine entsprechende Diskussion stattfinden.

7.2.1 Primitiva des konturbasierten Ansatzes

Die hier vorliegende Arbeit beruht auf einem rein konturbasierten Ansatz. Aufgrund der Eigenschaften der untersuchten Primitiva (Geradenstücke und Ellipsensegmente), konnten nicht alle der in Kapitel 2 vorgestellten Gestaltgesetze zum Einsatz kommen. In der gegenwärtigen Form wurden die folgenden Gesetzmäßigkeiten als geeignet angesehen:

- **Gesetz der Nähe:** Auswertung der Nähe von Endpunkten benachbarter Primitiva, für die Generierung von Gruppenhypothesen.
- **Gesetz der Gleichheit:** Auswertung der Gleichheit der Primitiva für die Gruppenhypothesen Kollinearität, Kurvilinearität und Parallelität.
- **Gesetz der guten Kurve:** Untersuchung von Randbedingungen zur Orientierungsdifferenz für kollineare und kurvilineare Hypothesen.
- **Gesetz der Geschlossenheit:** Generierung von Geschlossenheitshypothesen.

- **Gesetz der Symmetrie:** Untersuchung des Spezialfalls der Parallelität zwischen Geradenstücken.
- **Gesetz des Zusammenhangs:** Gilt für alle Approximationen von Konturen und ihre Gruppierung in die Primitiva Geradenstück und Ellipsensegment.
- **Prägnanz:** Generierung von Geschlossenheitshypothesen für Projektionen verdeckter Objekte in der Szene.

Das **Gesetz der Symmetrie** wird derzeit, wie oben beschrieben, nur auf den Spezialfall der Parallelität angewandt. Eine Erweiterung auf die Untersuchung allgemeiner Symmetrien ist sinnvoll und wurde bereits in Abschnitt 7.1.4.3 im Zusammenhang mit der Generierung von Geschlossenheitshypothesen diskutiert. Die Untersuchung von Symmetrien für große, fragmentierte Ellipsensegmente, wie sie bei allen runden Objekten auftreten, kann helfen, diese als Geschlossenheitshypothesen zu generieren (siehe hierzu auch die Schraubwürfel in den Abbildungen B.7b und B.9b). Symmetrien beschreiben dieselben räumlichen Konstellationen wie Parallelitäten und können deshalb mit diesen in der 2x1D-Ebene angesiedelt werden.

Das **Gesetz der Gleichheit** und das **Gesetz der Nähe** sind auf fragmentierte, sich gegenseitig stark überlappende Ellipsensegmente anwendbar (siehe dazu auch die Radreifen in den Abbildungen B.3b und B.11b). Solche Ellipsensegmente sollten innerhalb einer zu definierenden Region in einer resultierenden Hypothese zusammengefaßt werden: Ausgehend von der jeweiligen Hypothese wird die nähere Umgebung auf weitere Ellipsensegmente untersucht, die einen ähnlichen Krümmungsverlauf besitzen. Ein ähnliches Vorgehen beschreibt Mohan in [17] auf den Seiten 46-47 für Geradenstücke. Die hier beschriebene Gruppierung kann analog zu Kurvilinearitäten in der 1D-Ebene angesiedelt werden.

Das **Gesetz der Gleichheit** kann zur Bildung von Aggregaten dienen. Als Aggregat wird hier eine Ansammlung von Primitiva oder Gruppenthypothesen verstanden, denen eine Anzahl zu definierender Merkmale gemein ist. Als Beispiel sollen hier die Leisten der Abbildung B.7a, b herangezogen werden. Die Löcher der Leisten werden durch eine Ansammlung von Ellipsenbögen beschrieben. Auf Gleichheit können folgende Merkmale untersucht werden: Typ der Primitiva, Größe der Primitiva, fast lineare Ausrichtung der Mittelpunkte der Primitiva. Die Bildung von Aggregaten bedeutet eine grundsätzliche Erweiterung der bisher eingesetzten Strategie zur Bildung von Gruppenthypothesen, die derzeit nur auf die Expansion der Primitiva und Gruppenthypothesen zu Geschlossenheiten ausgerichtet ist. Aggregate beschreiben analog zu Geschlossenheiten zweidimensionale Konstellationen und können mit diesen der 2D-Ebene zugeordnet werden.

7.2.2 Einbeziehung weiterer Primitiva

Der konturbasierte Ansatz ist durch die Einbeziehung regionenbasierter Verfahren mit hohem Nutzen erweiterbar. Die Objekte der untersuchten Domäne sind gut unterscheidbar gefärbt oder naturfarben. Mit Ausnahme der Räder (Felge, Radreifen) und der Schrauben (Kopf, Gewinde) sind alle Objekte einfarbig. Für Räder und Schrauben gilt jedoch, daß auch die Bestandteile jeweils einfarbig sind. Eine Gruppierung in Regionen auf der Basis einer Farbsegmentation ist vor diesem Hintergrund vielversprechend. Räder und Schrauben sind die einzigen Objekte der Domäne, deren Oberflächen teilweise texturiert sind (Reifenprofil und Schraubengewinde). In Abbildung B.9b kann für das Schraubengewinde und in Abbildung B.11b für das Reifenprofil beobachtet werden, daß diese Texturen durch eine große Ansammlung kleiner und kleinster Konturen abgebildet werden. Besonders für die Schraubengewinde ist eine Gruppierung in Regionen gleicher Textur sinnvoll, da sie von keinem der bisher beschriebenen Gruppierungsverfahren erfaßt werden konnten.

Das **Gesetz der Gleichheit** und das **Gesetz des Zusammenhangs** ist grundsätzlich bei allen regionenbasierten Verfahren wirksam, da die untersuchten Pixel aufgrund gemeinsamer Merkmale, wie Farbe oder Textur in einer Gruppierung über eine zusammenhängende Region zusammengefaßt werden. Hypothesen über Regionen besitzen eine zweidimensionale räumliche Ausdehnung (2D-Ebene) und sollten mit Hypothesen über Geschlossenheiten gleichgestellt werden.

7.2.3 Kombination der Gruppierungshypothesen

Gruppierungshypothesen über Regionen ergeben in der Kombination mit konturbasierten Gruppierungshypothesen neue und interessante Aspekte. Das **Gesetz des gemeinsamen Bereichs** ist geeignet die verschiedenen Hypothesen erneut zu gruppieren. Für eine Leiste, wie in Abbildung B.1a, b, kann sich mit den diskutierten Erweiterungen folgende Konstellation ergeben:

- Eine Geschlossenheitshypothese beschreibt den Umriß der Leiste.
- Für die Löcher der Leiste konnte ein Aggregat hypothetisiert werden.
- Eine Hypothese über eine Region gleicher Farbe, die der Ausdehnung der Leiste entspricht, wird generiert. Geschlossenheit und Aggregat lassen sich innerhalb der Region einordnen.

Besonders interessant sind solche Kombination für die eher “schwierigen” Objekte der Domäne, wie beispielweise der Schraubwürfel. Die Beschreibung des gemeinsamen Bereiches gleicher Farbe durch die Regionenhypothese ist in der Lage, die in diesem Bereich gefundenen Ellipsen als zusammengehörig zu kennzeichnen. Durch geeignete Untersuchung der Geometrie dieser Ellipsen, z.B. der räumlichen Anordnung innerhalb des detektierten Bereiches, ist die Hypothesisierung eines Aggregats möglich. Die Überlagerung des Randes der Region mit den innerhalb gefundenen Ellipsen ist in der Lage, die Hypothesisierung von Geschlossenheiten für Schraubwürfel zu verbessern.

Die Kombinationen von Gruppenthypothesen in einen gemeinsamen Bereich stellt eine neue Qualität von Information dar. In Anlehnung an die in Abschnitt 4.2 beschriebene Hierarchie der Gruppenthypothesen sollte für sie eine neue Hierarchieebene eingeführt werden, die oberhalb der 2D-Ebene angeordnet ist.

7.2.4 Bewertung der Gruppenthypothesen

Für alle in den vorangehenden Abschnitten 7.2.1 - 7.2.3 angesprochenen Erweiterungen in der Generierung von Gruppenthypothesen, sind geeignete Abbildungen in das Markov Random Field zu untersuchen. Die Erweiterung kann analog zu den bereits vorhandenen Mechanismen erfolgen: Untersuchung von Heuristiken für datengetriebene Signifikanzen und Abbildung in den ungerichteten Graphen (siehe dazu auch Kapitel 6). Dies gilt auch für die in Abschnitt 7.2.3 beschriebene Kombination verschiedener Hypothesen: Diese können in eine Hypothese über einen gemeinsamen Bereich abgebildet werden, der durch Gruppenthypothesen wie Geschlossenheit, Unterstützung erfährt.

Die Wirkung der Unterstützungsrelationen sollte auf eine Umsetzung der in Abschnitt 4.2 beschriebenen Hierarchie der Hypothesen näher untersucht werden. Die Wirkung der Unterstützung ist bisher gleichwertig, unabhängig von der Hierarchiestufe der sich unterstützenden Hypothesen. Wird beispielweise eine Geschlossenheitshypothese durch eine Anzahl Nähe-Gruppen und eine Parallelität unterstützt, ist es sinnvoll, daß die Parallelität einen höheren Einfluß nehmen kann, da sie eine signifikantere räumliche Anordnung beschreibt.

Der Aufbau des Markov Random Field sollte zukünftig dynamisch gestaltet werden. Für den hier vorgestellten Ansatz (siehe dazu auch Kapitel 6) war eine zeitliche Optimierung des Bewertungsprozesses nicht vordringlich. Aus diesem Grund konnte die Erzeugung des ungerichteten Graphen und die Bewertung der Hypothesen in einem einzigen Schritt erfolgen. Für eine Weiterentwicklung sollte dieser Prozeß in einzelne Schritte zerlegt werden. Diese Schritte sollten für ausgewählte Bereiche bereits im Vorfeld Hypothesen über eine

abschließende Bewertung liefern. Eine Auswahl dieser Bereiche kann auf der Grundlage der in Abschnitt 7.2.2 beschriebenen Hypothesen über Regionen erfolgen. Eine abschließende Untersuchung des gesamten Graphen wird dann diese Hypothesen bestätigen oder verändern.

Die Erzeugung neuer Knoten sollte situationsabhängig möglich sein. Wenn beispielsweise die Anwesenheit von Nähe-Gruppen und einer Parallelität das Vorhandensein einer Geschlossenheit befürworten, diese aber initial nicht erzeugt werden konnte, sollte es möglich sein, die Erzeugung einer Geschlossenheit nochmals mit größerer Aufmerksamkeit zu untersuchen. Bei Bedarf ist ein Knoten für diese Hypothese zu erzeugen.

8 Zusammenfassung

Das menschliche visuelle System besitzt Mechanismen der Strukturbildung, die eine intuitive und effiziente Wahrnehmung der Umgebung ermöglichen. Diese Mechanismen zergliedern eine Szene durch Gruppenbildung in ihre Bestandteile. Diese Bildung von Strukturen wurde durch die Gestaltpsychologie mit Hilfe von Gruppierungsgesetzen beschrieben. Die Zielsetzung der vorliegenden Arbeit war, diese Mechanismen für die digitale Bildverarbeitung nutzbar zu machen. Dabei konnte sich diese Arbeit zum einen auf die Forschung der Gestaltpsychologie stützen. Zum anderen dienten frühere Arbeiten im Bereich der digitalen Bildverarbeitung als Wegweiser.

Bei der Umsetzung der Gruppierungsgesetze ist zu beachten, daß die wirksamen Prinzipien, wie beispielsweise "Nähe" oder "Ähnlichkeit" in den Gestaltgesetzen nicht quantifiziert sind. Ihre Wirkung wurde im vorliegenden Fall durch die Untersuchung einer repräsentativen Stichprobe der Domäne angegangen. Als Ergebnis der Untersuchung entstanden Suchbereiche für die verschiedenen Gruppierungshypothesen, die charakteristische Eigenschaften dieser Hypothesen widerspiegeln. Diese Bereiche wurden in Anlehnung an die Gestaltgesetze mit dem Begriff **Bereiche perzeptiver Aufmerksamkeit** belegt.

Der Ansatz der vorliegenden Arbeit ist konturbasiert: die initialen Primitiva des Gruppierungsprozesses sind Geradenstücke und Ellipsensegmente, die durch eine geeignete Vorverarbeitung aus der Szene extrahiert werden. Die Generierung der Gruppierungshypothesen mit Hilfe der Bereiche perzeptiver Aufmerksamkeit ist eingebunden in eine dreistufige Hierarchie, die diese Gruppierungshypothesen entsprechend ihrer räumlichen Komplexität anordnet. Auf der untersten Stufe, der 1D-Ebene, werden die Hypothesen Kollinearität, Kurvilinearität und Nähe generiert. In der folgenden 2x1D-Ebene, werden Parallelitäten generiert und in der 2D-Ebene Geschlossenheiten. Jeder dieser Ebenen stehen die Ergebnisse der vorherigen, als auch die initialen Primitiva zur Verfügung.

Mit den Bereichen perzeptiver Aufmerksamkeit erfolgt die Generierung der Gruppierungshypothesen in einem lokalen Rahmen. Eine Folge davon ist die Erzeugung mehrdeutiger Gruppierungshypothesen, die eine unerwünschte inkonsistente Interpretation der Szene bedeuten.

Dies ist ein wesentlicher Grund, weshalb sich an den Gruppierungsprozeß eine Bewertung der Hypothesen anschließt. Diese Bewertung macht sich den Formalismus von Markov Random Fields zu Nutze, weil er eine lokale Interaktion der benachbarten Gruppenthypothesen erlaubt, als auch die Bestimmung einer global optimalen Interpretation der Szene ermöglicht. Die Interaktion der benachbarten Gruppenthypothesen wird gesteuert durch das Design der Cliquepotentiale, welches a priori Wissen in das Random Field einbringt. Mit Hilfe des Theorems von Hammersley und Clifford, welches die Äquivalenz von Markov Random Fields und Gibbs-Verteilungen beschreibt, wird die Schätzung der Maximum a posteriori Wahrscheinlichkeit in die Minimierung einer Energiefunktion überführt. Die Entitäten des Random Field entsprechen den einzelnen Gruppenthypothesen und ihren initialen Primitiva. Dieses bedeutet eine irreguläre Abbildung der Bildmerkmale in die Knoten des ungerichteten Graphen. Die Kanten des Graphen entsprechen den Nachbarschaftsrelationen Unterstützung und Konkurrenz, die zwischen den Gruppenthypothesen bestehen. Die Entitäten eines Markov Random Field sind mit Zufallsvariablen verbunden, deren Wert in der Anwendung in Form eines Label vergeben wird. Die Minimierung der Energiefunktion wird mit Hilfe des "Highest Confidence First" durchgeführt. Dafür werden im Vorfeld datengetriebene Label für die generierten Gruppenthypothesen bestimmt. Diese datengetriebenen Label gestatten eine Aussage über die "Stabilität" der Hypothese, mit der der Minimierungsprozeß startet.

Sinnvolle Erweiterungen des vorliegenden Ansatzes sind durch eine Einbeziehung neuer Primitiva und weiterer Gestaltgesetze in den Gruppierungsprozeß möglich.

Literaturverzeichnis

- [1] Fred Attneave. *Informationstheorie in der Psychologie*. Verlag Hans Huber Bern Stuttgart Wien, 1974.
- [2] Charles A. Bouman, 1995. Tutorial ThA1 – Workshop – IEEE International Conference on Image Processing.
- [3] U.B. Desai. Markov random field models for early vision problems. Technical report, Department of Electrical Engineering, Indian Institute of Technology – Bombay, 1993. Technical report: SPANN-LAB-93-3.
- [4] Martin A. Fischler. The perception of linear structure: A generic linker. In *DARPA, Image Understanding Workshop*, pages 1565–1579, 1994.
- [5] David W. Jacobs. Groper: A grouping based recognition system for two dimensional objects. In *IEEE Computer Society Workshop on Computer Vision*, pages 164–169, 1987.
- [6] David W. Jacobs. Grouping for recognition. Technical report, Massachusetts Institute of Technology Artificial Intelligence Laboratory, 1989. A.I. Memo No. 1177.
- [7] David W. Jacobs. Robust and efficient detection of salient convex groups. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(1):23–37, 1996.
- [8] Siegfried Jaeger. *Wolfgang Köhler und seine Zeit*. Ausstellungsführer der Freien Universität Berlin. Universitätsbibliothek der Freien Universität Berlin, 1990.
- [9] Friedrich Jähne. *Digitale Bildverarbeitung*. Springer, 1989.
- [10] Gaetano Kaniza. *Organization in Vision*. PRAEGER, 1979.
- [11] David Katz. *Gestaltpsychologie*. Schwabe, 1969.
- [12] Wolfgang Köhler. *Die Aufgabe der Gestaltpsychologie*. de Gruyter, 1971.
- [13] Aleš Leonardis. *Image Analysis Using Parametric Models*. PhD thesis, University of Ljubljana, 1993.
- [14] S.Z. Li. *Markov Random Field Modeling in Computer Vision*. Springer-Verlag Tokyo, 1995.

- [15] David G. Lowe. *Perceptual Organization and Visual Recognition*. Kluwer Academic Publisher, 1985.
- [16] David G. Lowe. Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence*, pages 355–395, 1987.
- [17] Rakesh Mohan. *Perceptual Organization for Computer Vision*. PhD thesis, University of Southern California, 1989.
- [18] Irvin Rock. *An Introduction to Perception*. Macmillan Publishing Co., Inc./ Collier-Macmillan Publishers, 1975.
- [19] Irvin Rock. *Perception*. W.H. Freeman und Company, 1984.
- [20] Eric Saund. Labeling of curvilinear structure across scales by token grouping. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR*, pages 257–263, 1992.
- [21] Daniel Schlüter. Bewertung konturbasierter gruppierungen mit hilfe von markov random fields. Master's thesis, Universität Bielefeld, 1995.
- [22] Jürgen Schürmann. Mustererkennung mit statistischen methoden. Technical report, Daimler-Benz AG, Forschungszentrum Ulm, Institut für Informationstechnik, Sommersemester 1992 und 1993. Vorlesungsskript.
- [23] Amnon Sha'shua. Structural saliency: The detection of globally salient structures using a locally connected network. Master's thesis, Dept. of applied math., Weizmann institute of science, 1988.
- [24] Qi-Gang Qao und A.K.C. Wong. Curve detection based on perceptual organization. *Pattern Recognition*, 26(7):1039–1047, 1993.
- [25] Jay M. Tenenbaum und Andrew P. Witkin. On the role of structure in vision. In Jacob Beck und Barbara Hope und Azriel Rosenfeld, editor, *Human and Machine Vision*, pages 481–543. Academic Press, 1983.
- [26] Vittorio Murino und Carlo S. Regazzoni und Gian Luca Foresti. Grouping as a searching process for minimum energy configurations of labelled random fields. *Computer Vision and Image Understanding*, 64(1):157–174, 1996.
- [27] Dana H. Ballard und Christopher M. Brown. *Computer Vision*. Prentice-Hall, 1982.
- [28] P.B. Chou und C.M. Brown. The theory and practice of bayesian image labeling. *Int. Journal of Computer Vision*, 4:185–210, 1990.
- [29] Stuart Geman und Donald Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, 1984.

-
- [30] John Dolan und Edward Riseman. Computing curvilinear structure by token-based grouping. In *Conference on Computer Vision and Pattern Recognition*, pages 264–270, 1992.
- [31] L.I. Anzyferowa und E.W. Schorochowa. *Behaviorismus, Gestaltpsychologie, Tiefenpsychologie*, volume Heft 11 of *Psychologische Beiträge*. Volk und Wissen, Volkseigener Verlag Berlin, 1969.
- [32] A. Ylä-Jääski und F. Ade. Line segment ribbons and their grouping. In *11th International Conference on Pattern Recognition (ICPR)*, pages 750–754, 1992.
- [33] A. Ylä-Jääski und F. Ade. Grouping symmetrical structures for object segmentation and description. *Computer Vision and Image Understanding*, 63(3):399–417, 1996.
- [34] Davi Geiger und Federico Girosi. Parallel and deterministic algorithms from mrf's: Surface reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(5):401–412, 1981.
- [35] Radu Horaud und Françoise Veillon und Thomas Skordas. Finding geometric and relational structures in an image. In *1. European Conference on Computer Vision (ECCV)*, pages 374–384, 1990.
- [36] Bruno Zoppis und Gae:elle Calvary und James L. Crowley. Perceptual grouping for scene interpretation in an active vision system. Technical report, Lifa Institut IMAG, Grenoble, 199?
- [37] Winfried A. Fellenz und Georg Hartmann. Preattentive grouping and attentive selection for early visual computation. In *International Conference on Pattern Recognition (ICPR)*, pages 340–345, 1996.
- [38] Gideon Guy und Gérard Medioni. Inferring global perceptual contours from local features. In *DARPA, Image Understanding Workshop*, pages 881–892, 1993.
- [39] Gideon Guy und Gérard Medioni. Inferring global perceptual contours from local features. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 786–787, 1993.
- [40] Parag Havaldar und Gérard Medioni und Fritjof Stein. Extraction of groups for recognition. In *Computer Vision-ECCV'94*, Lecture Notes in Computer Science, pages 251–261. Jan-Olof Eklundh, 1994.
- [41] Sandra Denasi und Giorgio Quaglia und Domenico Rinaudi. The use of perceptual organization in the prediction of geometric structures. *Pattern Recognition Letters*, 13(7):529–539, 1992.
- [42] Philipp Besslich und Henning Bässmann. *Bildverarbeitung Ad Oculos*. Springer, 1991.
- [43] Philipp W. Besslich und Henning Bässmann. Gestalt-based approach to robot vision. In T. Jordanides und B. Torby, editor, *Expert Systems and Robotics*, volume F 71 of *NATO ASI Series*, pages 1–34, 1991.

- [44] Stephen Palmer und Irvin Rock. Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin & Review*, 1(1):29–55, 1994.
- [45] Ingemar J. Cox und James M. Rehg und Sunita Hingorani. A bayesian multiple-hypothesis approach to edge grouping and contour segmentation. *International Journal of Computer Vision*, 11(1):5–24, 1993.
- [46] Andrew Witkin und Jay Tenenbaum. What is perceptual organization for? In *IJCAI*, pages 1023–1026, 1983.
- [47] James N. Huddleston und Jezekiel Ben-Arie. Grouping edgels into structural entities using circular symmetry, the distributed hough transform, and probabilistic non-accidentalness. *CVGIP Image Understanding*, 57(2):227–242, 1993.
- [48] H.Q. Lu und J.K. Aggarwal. Extracting geometric structures and a region of interest for man-made object detection. In *SPIE 20th AIPR Workshop, Computer Vision Applications*, volume 1623, pages 79–90, 1991.
- [49] H.Q. Lu und J.K. Aggarwal. Applying perceptual organization to the detection of man-made objects in non-urban scenes. *Pattern Recognition*, 25(8):835–853, 1992.
- [50] Ross Kindermann und J. Laurie Snell. *Markov Random Fields and their Applications*, volume 1. American Mathematical Society, 1980.
- [51] Sabine Urago und Josiane Zerubia und Marc Berthod. A markovian model for contour grouping. In *IEEE International Conference on Pattern Recognition*, pages 556–558, 1994.
- [52] George Reynolds und J. Ross Beveridge. Searching for geometric structure in images of natural scenes. In *DARPA, Image Understanding Workshop*, pages 257–271, 1987.
- [53] J. Brian Subirana-Vilanova und Kah Kay Sung. Perceptual organization without edges. In *DARPA, Image Understanding Workshop*, pages 289–298, 1992.
- [54] Sudeep Sakar und Kim L. Boyer. Perceptual organization in computer vision: A review and a proposal for a classificatory structure. *IEEE Transaction on Systems, Man, and Cybernetics*, pages 382–399, 1993.
- [55] Sudeep Sakar und Kim L. Boyer. *Computing Perceptual Organization In Computer Vision*. Machine Perception & Artificial Intelligence. World Scientific, 1994.
- [56] Philippe Montesinos und Laurent Alquier. Perceptual organization of thin networks with active contour functions applied to medical and aerial images. In *International Conference on Pattern Recognition (ICPR)*, pages 647–651, 1996.
- [57] Richard Weiss und Michael Boldt. Geometric grouping applied to straight lines. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR*, pages 489–495, 1986.

-
- [58] Deborah A. Trytten und Mihran Tuceryan. Segmentation and grouping of object boundaries using energy minimization. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR*, pages 730–731, 1991.
- [59] Narendra Ahuja und Mihran Tuceryan. Extraction of early perceptual structure in dot patterns: Integrating region, boundary, and component gestalt. *Computer Vision, Graphics, and Image Processing*, 48:304–356, 1989.
- [60] Helmut Kristen und Olaf Munkelt. Markov-feld-basierte bildinterpretation mit automatisch generierten datenbasen. In *Mustererkennung 92*, pages 50–57, 1992.
- [61] Daniel P. Huttenlocher und Peter C. Wayner. Finding convex edge groupings in an image. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 406–412, 1991.
- [62] P.B. Chou und P.R. Cooper und M.J. Swain und C.M. Brown und L.E. Wixson. Markov random fields, theory and application. In R. Chellappa und A. Jain, editor, *Probabilistic Network Inference for Cooperative High and Low Level Vision*, pages 211–243. Academic Press, 1993.
- [63] Laurent Héault und Radu Horaud. Figure-ground discrimination: A combinatorial optimization approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9):899–914, 1993.
- [64] Rakesh Mohan und Ramakant Nevatia. Perceptual organization for segmentation and description. In *DARPA, Image Understanding Workshop*, pages 415–424, 1989.
- [65] Rakesh Mohan und Ramakant Nevatia. Using perceptual organization to extract 3-d structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(11):1121–1139, 1989.
- [66] Rakesh Mohan und Ramakant Nevatia. Perceptual organization for scene segmentation and description. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(6):616–635, 1992.
- [67] Ronald Chung und Ramakant Nevatia. Use of monocular groupings and occlusion analysis in a hierarchical stereo system. *Computer Vision and Image Understanding*, 62(3):245–268, 1995.
- [68] John Dolan und Richard Weiss. Perceptual grouping of curved lines. In *DARPA, Image Understanding Workshop*, pages 1135–1145, 1989.
- [69] Michael Boldt und Richard Weiss und Edward Riseman. Token-based extraction of straight lines. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(6):1581–1594, 1989.
- [70] Martin A. Fischler und Robert C. Bolles. Perceptual organization and curve partitioning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(1):100–105, 1986.

- [71] Rebecca L. Castaño und Seth Hutchinson. A probabilistic approach to perceptual grouping. *Computer Vision and Image Understanding*, 64(3):399–419, 1996.
- [72] Amnon Sha'ashua und Shimon Ullman. Structural saliency: The detection of globally salient structures using a locally connected network. In *Second International Conference on Computer Vision (ICCV)*, pages 321–327, 1988.
- [73] Irvin Rock und Stephen Palmer. Das Vermächtnis der Gestaltpsychologie. *Spektrum der Wissenschaft*, pages 68–75, 1991.
- [74] Alain Jacot-Descombes und Thierry Pun. Asynchronous perceptual grouping: From contours to relevant 2d-structures. *Computer Vision and Image Understanding*, 66(1):1–24, 1997.
- [75] David G. Lowe und Thomas O. Bindford. Perceptual organization as a basis for visual recognition. In *Proceedings of the National Conference on Artificial Intelligence, AAAI*, pages 255–260, 1983.
- [76] Gianluca Foresti und Vittorio Murion und Carlo S. Regazzoni und Gianni Vernazza. Grouping of rectilinear segments by the labeled hough transform. *CVGIP Image Understanding*, 58(3):22–42, 1994.
- [77] Tuomo S. Vehkomäki. Image segmentation by contour grouping: Knowledge based search in attributed proximity graph. In *Acta Polytechnica Scandinavica, Mathematics and Computing in Engineering Series No.76*. Finnish Academy of Technology, 1995.

Anhang A

A.1 Legende der verwendeten Gruppen

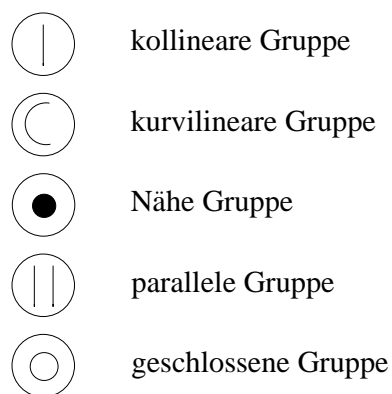


Abbildung A.1: Graphische Legende der verwendeten Gruppen

$\{1 2\}$	kollineare Gruppe mit den Segmenten 1 und 2
$\{1 \subset 2\}$	kurvilineare Gruppe mit den Segmenten 1 und 2
$\{1 \bullet 2\}$	Nähe Gruppe mit den Segmenten oder Gruppen 1 und 2
$\{1 2\}$	parallele Gruppe mit den Segmenten oder Gruppen 1 und 2
$\{1 \circ 2\}$	geschlossene Gruppe mit den Segmenten oder Gruppen 1 und 2

Tabelle A.1: Textuelle Legende der verwendeten Gruppen

A.2 Verwendete Variable

1D, 2x1D, 2D Hierarchieebenen für Gruppenhypothesen

A Anschlußprimitivum in einer betrachteten Gruppe

α Gewicht einer Teilsignifikanz

β Gewicht eines Basispotentials

c Clique

c^n n-elementige Clique

C Menge aller Cliques c

C_i Menge der Cliques der Nachbarschaft \mathcal{N}_i

cs Kontursegment

col kollinear

d Lot zwischen den Primitiva einer parallelen Gruppe

D Lückenlänge zwischen zwei Primitiva

D_N normierte Lückenlänge

\bar{D} arithmetisches Mittel von D

δ_x Schwellwert

Δ_x Maximum

$E(f)$ Energieterm

\mathcal{E} Erwartungswert

ϵ Approximationsfehler

F_i Zufallsvariable einer Entität i

F Menge aller Zufallsvariablen eines Markov Random Field

f eine Konfiguration des Markov Random Field

\mathcal{F} Menge aller möglichen Konfigurationen eines Markov Random Field

\mathcal{F}^+ erweiterter Lösungsraum für Konfigurationen

f_i aktuell einer Entität i zugeordnetes Label

f_o ursprüngliche Konfiguration des Markov Random Field

f_o^* optimale Annäherung der ursprünglichen Konfiguration f_o

f_i^o initiales Label einer Entität i

\hat{f}_o Schätzung für f_o

γ_x	Orientierungsdifferenz zweier Primitiva
G	Stabilitätsmaß
$G_i(f)$	lokale Stabilität
\mathcal{G}	ungerichteter Graph
γ	Orientierungsdifferenz zwischen zwei Primitiva
$\bar{\gamma}$	arithmetisches Mittel von γ
i, j	Entität des Markov Random Field
K	Kosten
L	Länge eines Primitivum
L_N	normierte Länge eines Primitivum
\bar{L}	arithmetisches Mittel von L
l	Label, Wert der einer Zufallsvariable zugeordnet wird
\mathcal{L}	Menge der Label l
\mathcal{L}^+	erweiterte Labelmenge
l°	datengetriebene Signifikanz
λ	Anzahl der Primitiva einer Gruppe
\mathcal{N}	Nachbarschaftssystem
\mathcal{N}_i	Nachbarschaft einer Entität i
$\mathcal{N}(f)$	benachbarte Konfigurationen f
N	Menge der natürlichen Zahlen
o	beobachtete Daten
P_A	Anschlußpunkt für die paarweise Betrachtung von Gruppenprimitiva
P_R	Referenzpunkt für die paarweise Betrachtung von Gruppenprimitiva
$P(f)$	Wahrscheinlichkeit für eine Konfiguration des Markov Random Field
$prox$	Nähe-Gruppe
p	Pixel
par	Parallelität
R	betrachtetes Primitivum einer Gruppe (Referenzprimitivum)
\mathcal{R}	Risiko
\mathcal{R}_o	bedingtes Risiko
\mathcal{S}	Menge der Entitäten i eines Markov Random Field
T	Konstante, Temperatur
τ	Grad der Überlappung der Primitiva eine parallelen Gruppe

u Umfang
 $U(f)$ Energiefunktion für eine Konfiguration
 $U_i(f)$ lokale Energie einer Entität

V_c Clique-Potential

ω Faktor für den Anteil gemeinsamer Fläche zweier Geschlossenheiten

x_N Normfaktor

ξ Öffnungswinkel eines ellipsenförmigen Referenzprimitivum

Z Normalisierungskonstante, partition function

A.3 Gewichte und Maxima

Kollinearität

α_1	0.5
α_2	0.25
α_3	0.25
Δ_{col}	0.05

Kurvilinearität

α_4	0.5
α_5	0.25
α_6	0.25
Δ_{curve}	0.62

Nähe

α_7	0.5
α_8	0.5
Δ_{prox}	1.07

Parallelität

α_9	0.4
α_{10}	0.4
α_{11}	0.2
Δ_{par_orient}	0.01
Δ_{par_mean}	3.5

Geschlossenheit

α_{12}	0.1
α_{12}^*	0.1
α_{13}	0.9
α_{13}^*	0.3
α_{14}	0.6

Clique-Potentiale

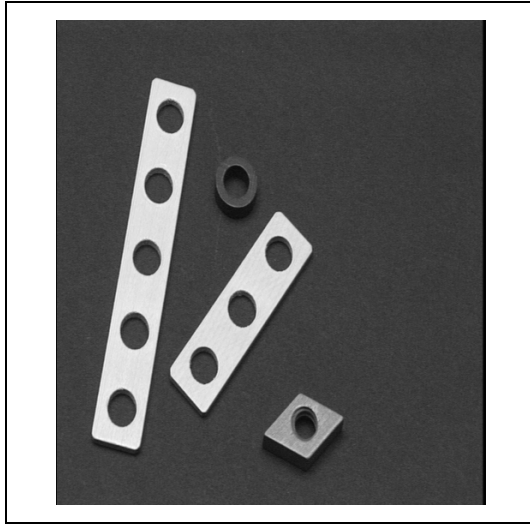
β_1	1
β_2^{sup}	1
β_2^{comp}	2

Anhang B

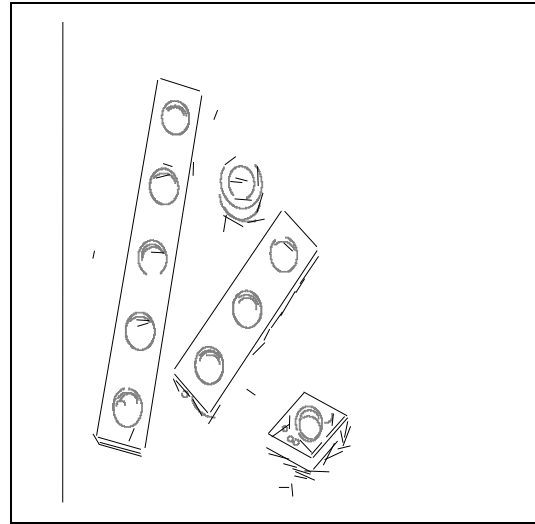
Ergebnisbilder

B.1 Bild1:

Bild1 in Abbildung B.1a) enthält insgesamt 4 Bauteile in Einzellege: 5-Lochleiste, 3-Lochleiste, Rautenmutter und Unterlegscheibe. Der Hintergrund ist einfarbig und dunkel.



(a)



(b)

Abbildung B.1: Bild1 :

- a) Grauwertbild
- b) Initiale Segmentierung: Geradenstücke und Ellipsensegmente

Hypothese	Typ	Gesamt #	Abbildung	# ≥ 0.8	Abbildung
Segment	Gerade	68	B.1(a)	27	–
Segment	Ellipse	43	B.1(b)	8	–
1D	Kollinear	4	–	4	B.2(a)
1D	Kurvilinear	1	–	1	B.2(a)
1D	Nähe	28	–	27	B.2(a)
2x1D	Parallel	13	–	9	B.2(b)
2D	Geschlossen	4	–	3	B.2(c)

Tabelle B.1: Hypothesierte Gruppen und initiale Primitiva

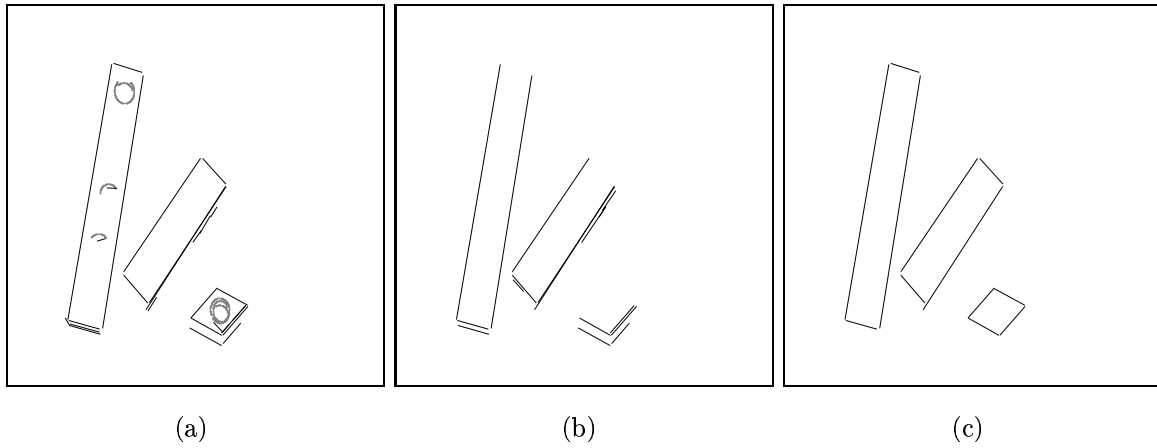
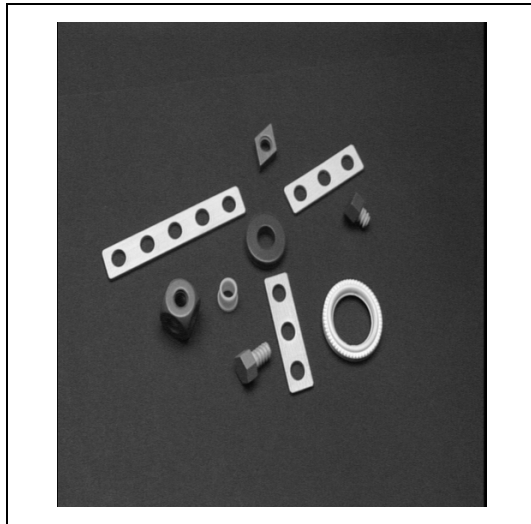


Abbildung B.2: Bild1:

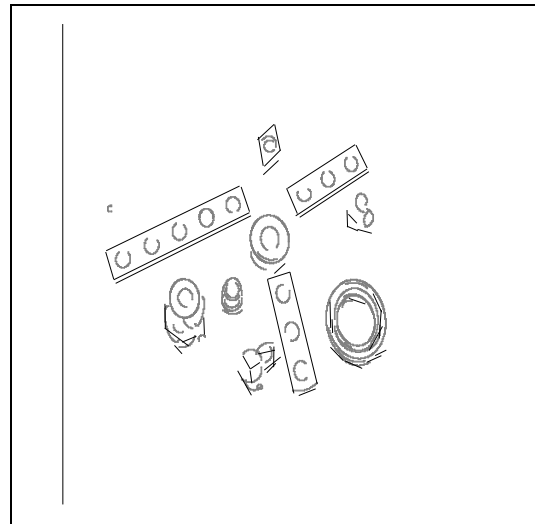
- a) Alle **1D**-Gruppen mit einer Bewertung ≥ 0.8
- b) Alle **2x1D**-Gruppen mit einer Bewertung ≥ 0.8
- c) Alle **2D**-Gruppen mit einer Bewertung ≥ 0.8

B.2 Bild2

Bild2 in Abbildung B.3a) enthält insgesamt 10 Bauteile in Einzellage. Der Hintergrund ist einfarbig und dunkel.



(a)



(b)

Abbildung B.3: Bild2 :

- a) Grauwertbild
- b) Initiale Segmentierung: Geradenstücke und Ellipsensegmente

Hypothese	Typ	Gesamt #	Abbildung	# ≥ 0.8	Abbildung
Segment	Gerade	48	B.3(a)	38	–
Segment	Ellipse	59	B.3(b)	17	–
1D	Kollinear	0	–	0	B.4(a)
1D	Kurvilinear	9	–	4	B.4(a)
1D	Nähe	39	–	37	B.4(a)
2x1D	Parallel	10	–	8	B.4(b)
2D	Geschlossen	6	–	4	B.4(c)

Tabelle B.2: Hypothesierte Gruppen und initiale Primitiva

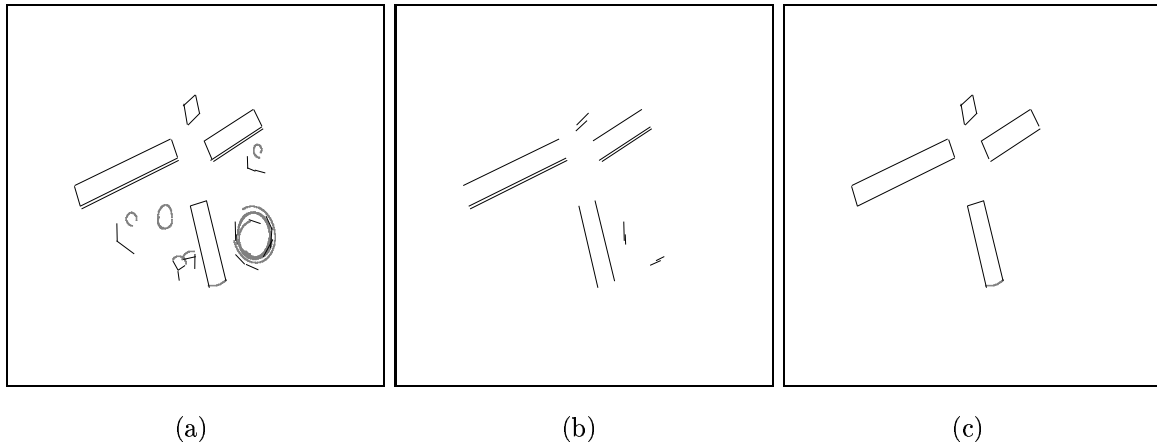
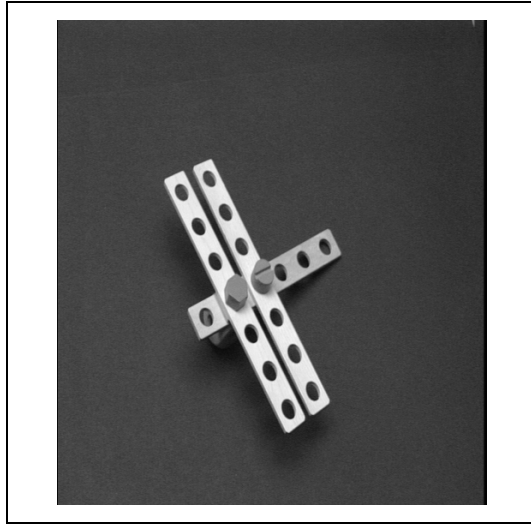


Abbildung B.4: Bild2:

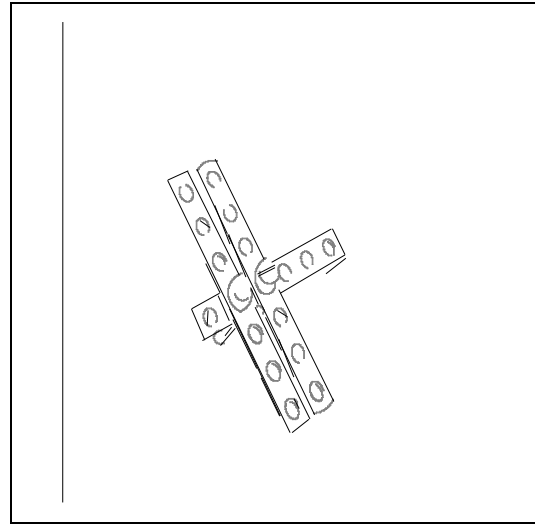
- a) Alle **1D**-Gruppen mit einer Bewertung ≥ 0.8
- b) Alle **2x1D**-Gruppen mit einer Bewertung ≥ 0.8
- c) Alle **2D**-Gruppen mit einer Bewertung ≥ 0.8

B.3 Bild3

Bild3 in Abbildung B.5a) zeigt eine Teilmontage mit “Rumpf” und “Flügeln”. Der Hintergrund ist einfarbig und dunkel.



(a)



(b)

Abbildung B.5: Bild3 :

- a) Grauwertbild
- b) Initiale Segmentierung: Geradenstücke und Ellipsensegmente

Hypothese	Typ	Gesamt #	Abbildung	# ≥ 0.8	Abbildung
Segment	Gerade	35	B.5(a)	20	–
Segment	Ellipse	32	B.5(b)	11	–
1D	Kollinear	16	–	10	B.6(a)
1D	Kurvilinear	0	–	0	B.6(a)
1D	Nähe	49	–	–	B.6(a)
2x1D	Parallel	49	–	23	B.6(b)
2D	Geschlossen	4	–	3	B.6(c)

Tabelle B.3: Hypothesierte Gruppen und initiale Primitiva

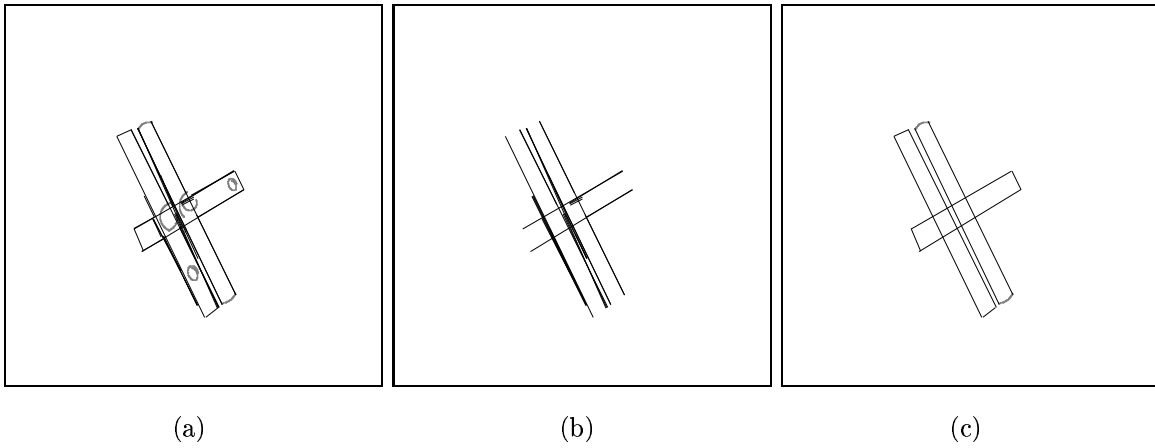


Abbildung B.6: Bild3:

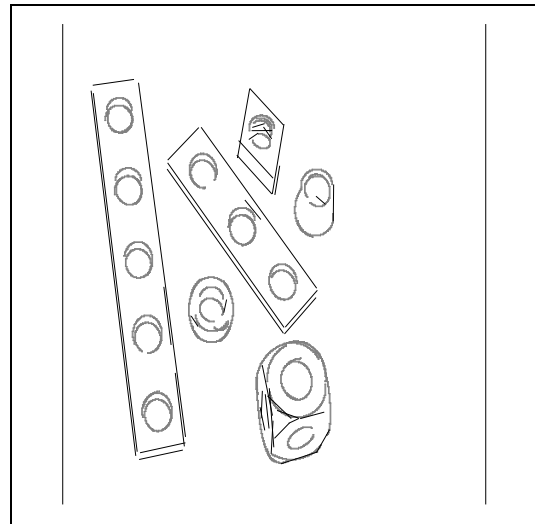
- a) Alle **1D**-Gruppen mit einer Bewertung ≥ 0.8
- b) Alle **2x1D**-Gruppen mit einer Bewertung ≥ 0.8
- c) Alle **2D**-Gruppen mit einer Bewertung ≥ 0.8

B.4 Bild4

Bild4 in Abbildung B.7a) enthält insgesamt 6 Bauteile in Einzellage. Der Hintergrund ist einfarbig und hell.



(a)



(b)

Abbildung B.7: Bild4 :

- a) Grauwertbild
- b) Initiale Segmentierung: Geradenstücke und Ellipsensegmente

Hypothese	Typ	Gesamt #	Abbildung	# ≥ 0.8	Abbildung
Segment	Gerade	44	B.7(a)	36	–
Segment	Ellipse	49	B.7(b)	15	–
1D	Kollinear	1	–	1	B.8(a)
1D	Kurvilinear	15	–	6	B.8(a)
1D	Nähe	52	–	48	B.8(a)
2x1D	Parallel	16	–	13	B.8(b)
2D	Geschlossen	3	–	3	B.8(c)

Tabelle B.4: Hypothesierte Gruppen und initiale Primitiva

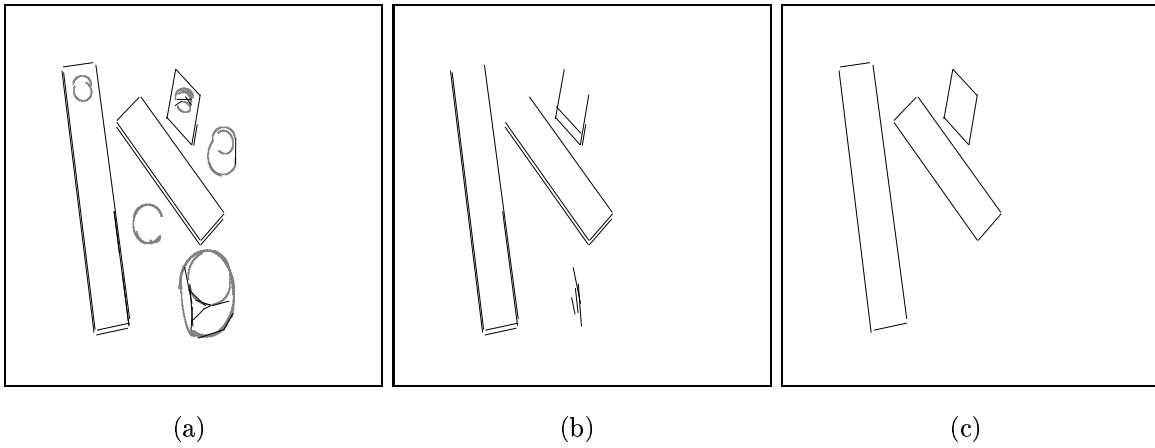
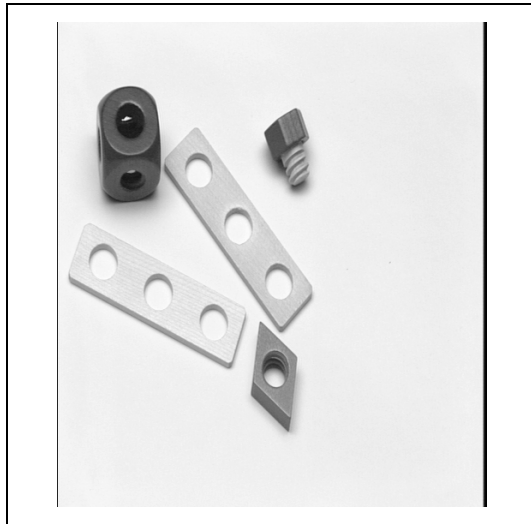


Abbildung B.8: Bild4:

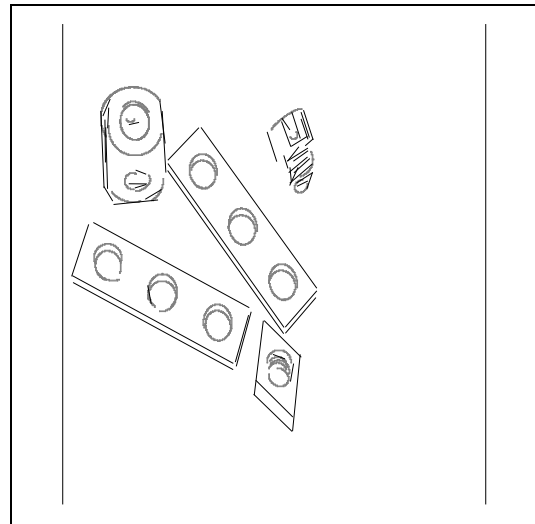
- a) Alle **1D**-Gruppen mit einer Bewertung ≥ 0.8
- b) Alle **2x1D**-Gruppen mit einer Bewertung ≥ 0.8
- c) Alle **2D**-Gruppen mit einer Bewertung ≥ 0.8

B.5 Bild5

Bild 5 in Abbildung B.9a) enthält insgesamt 6 Bauteile in Einzellage. Der Hintergrund ist einfarbig und hell.



(a)



(b)

Abbildung B.9: Bild5 :

- a) Grauwertbild
- b) Initiale Segmentierung: Geradenstücke und Ellipsensegmente

Hypothese	Typ	Gesamt #	Abbildung	# ≥ 0.8	Abbildung
Segment	Gerade	57	B.9(a)	35	–
Segment	Ellipse	33	B.9(b)	10	–
1D	Kollinear	0	–	0	B.10(a)
1D	Kurvilinear	4	–	4	B.10(a)
1D	Nähe	43	–	40	B.10(a)
2x1D	Parallel	24	–	15	B.10(b)
2D	Geschlossen	7	–	5	B.10(c)

Tabelle B.5: Hypothesierte Gruppen und initiale Primitiva

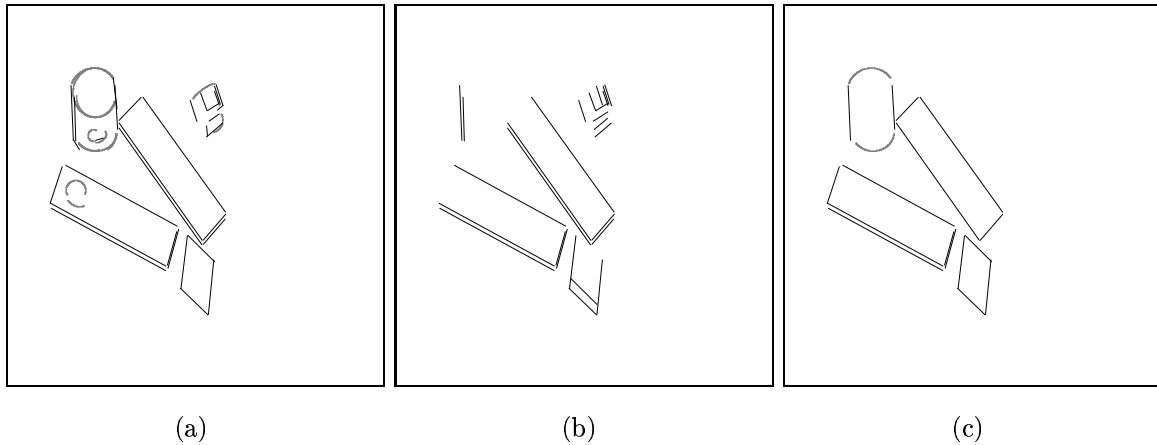


Abbildung B.10: Bild5:

- a) Alle **1D**-Gruppen mit einer Bewertung ≥ 0.8
- b) Alle **2x1D**-Gruppen mit einer Bewertung ≥ 0.8
- c) Alle **2D**-Gruppen mit einer Bewertung ≥ 0.8

B.6 Bild6

Bild 6 in Abbildung B.11a) enthält insgesamt 7 Bauteile in Einzellage. Der Hintergrund ist einfarbig und hell.

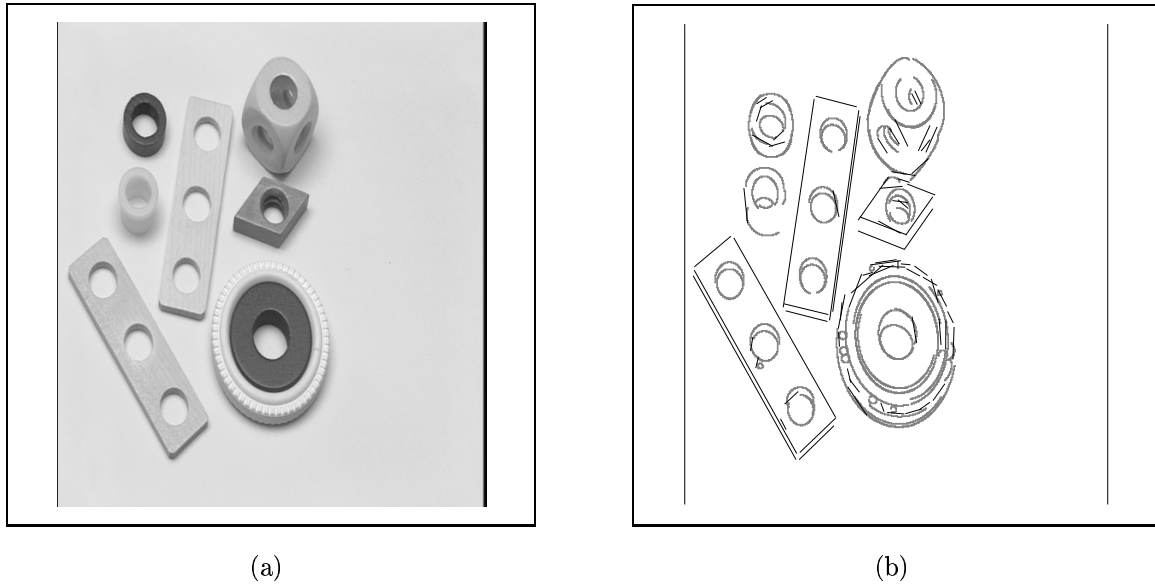


Abbildung B.11: Bild6 :

- a) Grauwertbild
- b) Initiale Segmentierung: Geradenstücke und Ellipsensegmente

Hypothese	Typ	Gesamt #	Abbildung	# ≥ 0.8	Abbildung
Segment	Gerade	77	B.11(a)	49	–
Segment	Ellipse	71	B.11(b)	31	–
1D	Kollinear	2	–	2	B.12(a)
1D	Kurvilinear	31	–	18	B.12(a)
1D	Nähe	106	–	79	B.12(a)
2x1D	Parallel	14	–	11	B.12(b)
2D	Geschlossen	4	–	3	B.12(c)

Tabelle B.6: Hypothesierte Gruppen und initiale Primitiva

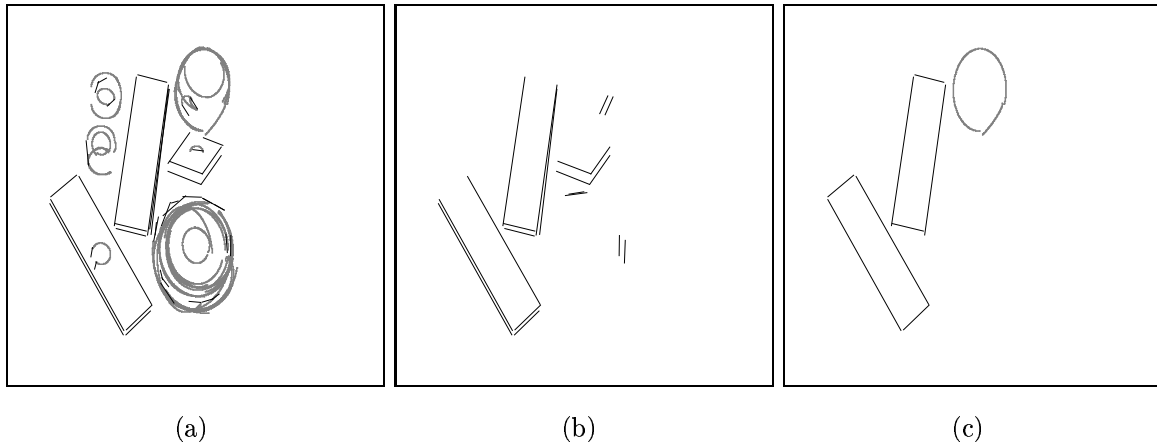
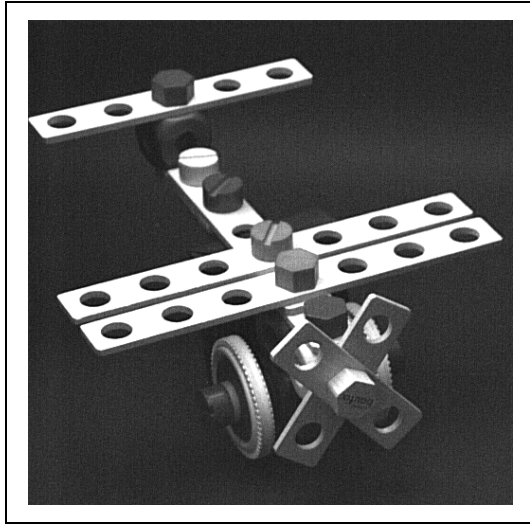


Abbildung B.12: Bild6:

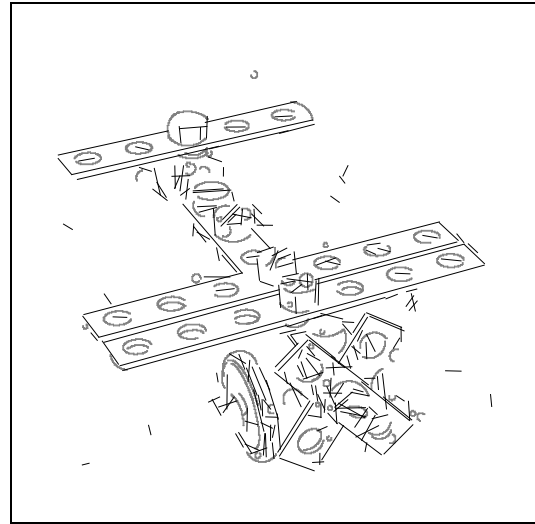
- a) Alle **1D**-Gruppen mit einer Bewertung ≥ 0.8
- b) Alle **2x1D**-Gruppen mit einer Bewertung ≥ 0.8
- c) Alle **2D**-Gruppen mit einer Bewertung ≥ 0.8

B.7 Bild7

Bild7 in Abbildung B.13a) zeigt ein vollständig montiertes Flugzeug. Der Hintergrund ist einfarbig und dunkel.



(a)



(b)

Abbildung B.13: Bild7 :

- a) Grauwertbild
- b) Initiale Segmentierung: Geradenstücke und Ellipsensegmente

Hypothese	Typ	Gesamt #	Abbildung	# ≥ 0.8	Abbildung
Segment	Gerade	218	B.13(a)	104	–
Segment	Ellipse	90	B.13(b)	30	–
1D	Kollinear	24	–	23	B.14(a)
1D	Kurvilinear	4	–	4	B.14(a)
1D	Nähe	161	–	161	B.14(a)
2x1D	Parallel	98	–	67	B.14(b)
2D	Geschlossen	6	–	5	B.14(c)

Tabelle B.7: Hypothesierte Gruppen und initiale Primitiva

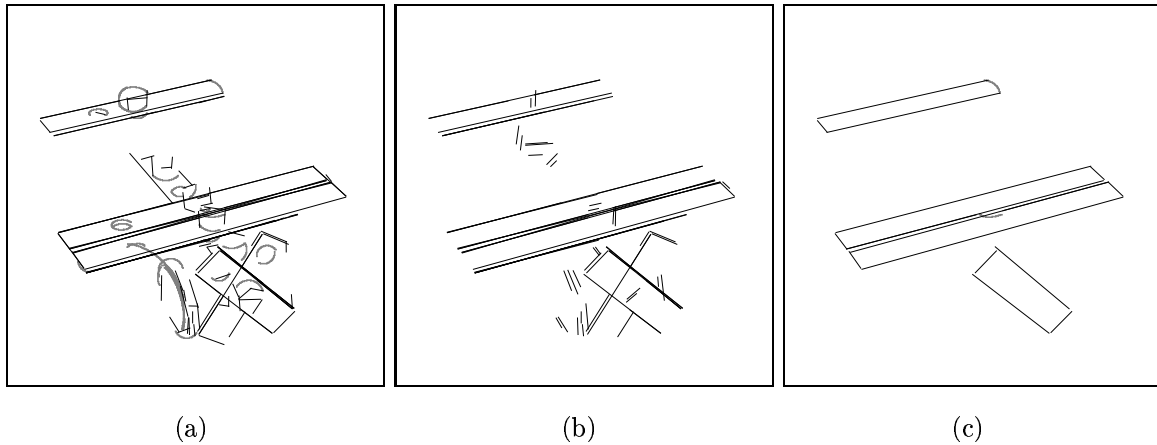
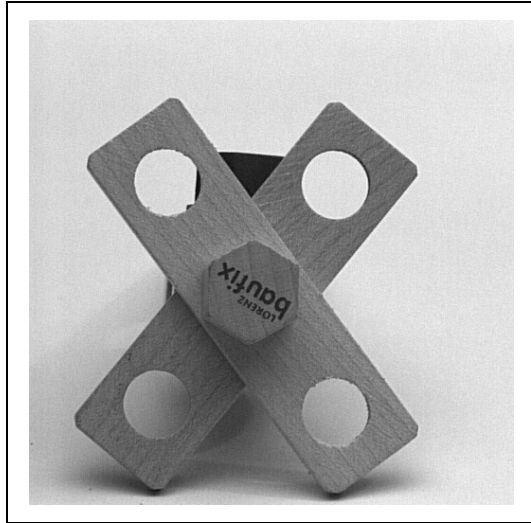


Abbildung B.14: Bild7:

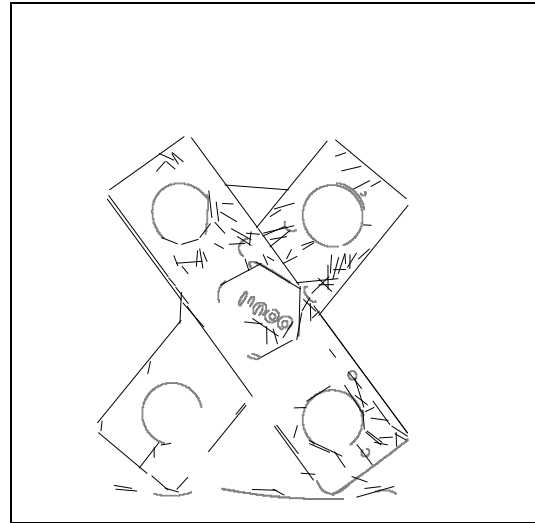
- a) Alle **1D**-Gruppen mit einer Bewertung ≥ 0.8
- b) Alle **2x1D**-Gruppen mit einer Bewertung ≥ 0.8
- c) Alle **2D**-Gruppen mit einer Bewertung ≥ 0.8

B.8 Bild8

Bild8 in Abbildung B.15a) eine Teilmontage den “Propeller”. Der Hintergrund ist einfarbig und hell.



(a)



(b)

Abbildung B.15: Bild8 :

- a) Grauwertbild
- b) Initiale Segmentierung: Geradenstücke und Ellipsensegmente

Hypothese	Typ	Gesamt #	Abbildung	# ≥ 0.8	Abbildung
Segment	Gerade	131	B.15(a)	35	–
Segment	Ellipse	29	B.15(b)	11	–
1D	Kollinear	16	–	15	B.16(a)
1D	Kurvilinear	2	–	2	B.16(a)
1D	Nähe	78	–	70	B.16(a)
2x1D	Parallel	25	–	17	B.16(b)
2D	Geschlossen	11	–	5	B.16(c)

Tabelle B.8: Hypothesierte Gruppen und initiale Primitiva

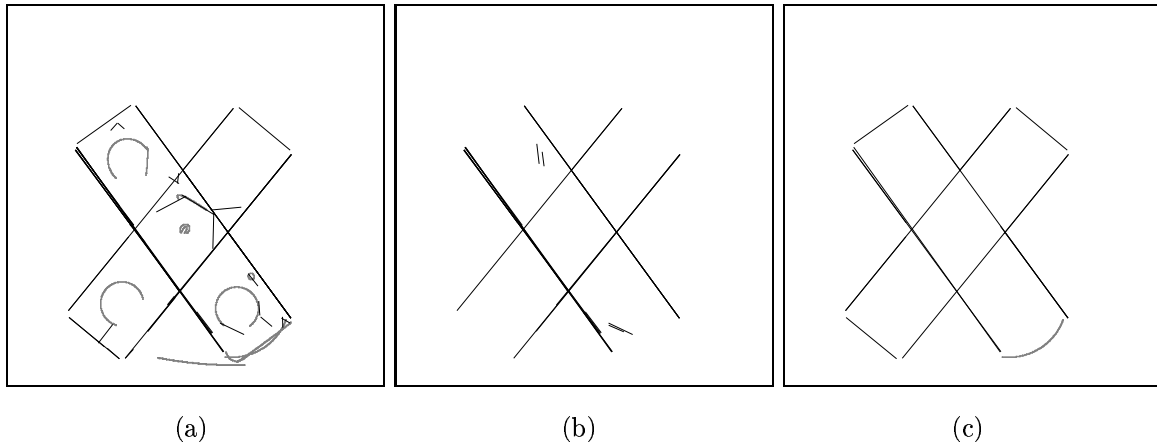
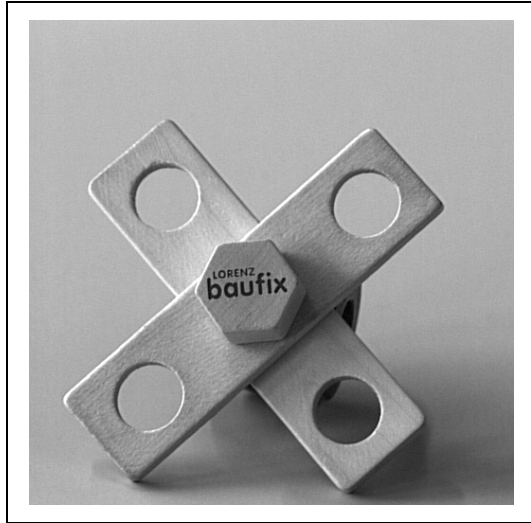


Abbildung B.16: Bild8:

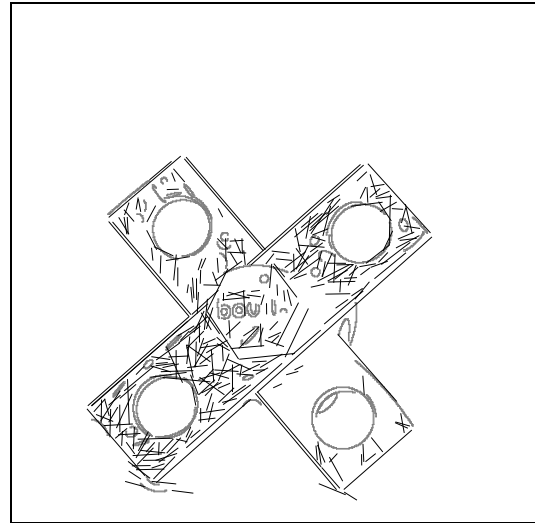
- a) Alle **1D**-Gruppen mit einer Bewertung ≥ 0.8
- b) Alle **2x1D**-Gruppen mit einer Bewertung ≥ 0.8
- c) Alle **2D**-Gruppen mit einer Bewertung ≥ 0.8

B.9 Bild9

Bild9 in Abbildung B.17a) eine Teilmontage den “Propeller”. Der Hintergrund ist einfarbig und hell.



(a)



(b)

Abbildung B.17: Bild9 :

- a) Grauwertbild
- b) Initiale Segmentierung: Geradenstücke und Ellipsensegmente

Hypothese	Typ	Gesamt #	Abbildung	# ≥ 0.8	Abbildung
Segment	Gerade	298	B.17(a)	135	–
Segment	Ellipse	55	B.17(b)	18	–
1D	Kollinear	76	–	49	B.18(a)
1D	Kurvilinear	4	–	4	B.18(a)
1D	Nähe	251	–	246	B.18(a)
2x1D	Parallel	227	–	111	B.18(b)
2D	Geschlossen	7	–	4	B.18(c)

Tabelle B.9: Hypothesierte Gruppen und initiale Primitiva

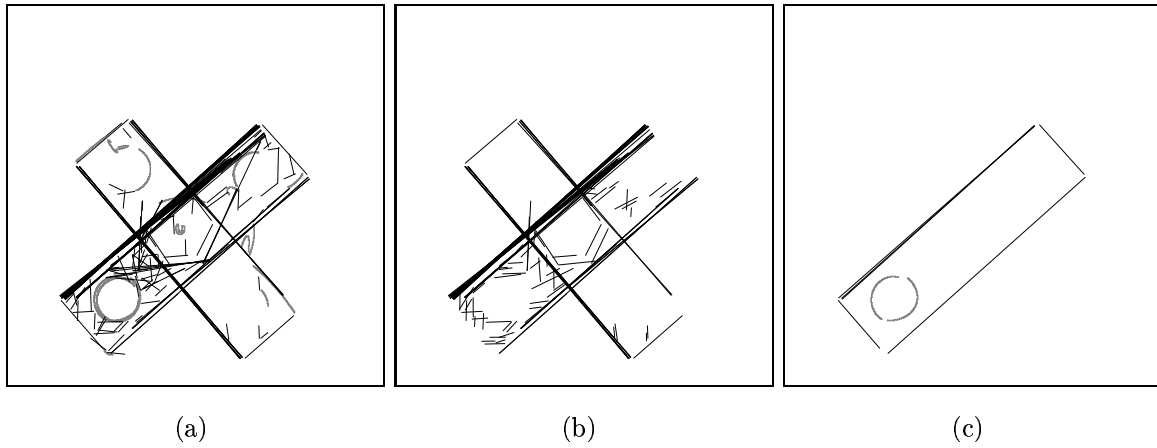


Abbildung B.18: Bild9:

- a) Alle **1D**-Gruppen mit einer Bewertung ≥ 0.8
- b) Alle **2x1D**-Gruppen mit einer Bewertung ≥ 0.8
- c) Alle **2D**-Gruppen mit einer Bewertung ≥ 0.8

B.10 Laufzeit

Die Laufzeit für die komplette Bearbeitung einer Szene lag im Batchbetrieb auf einer DEC-Alpha bei ca. einer Minute für die einfachen Bilder, bis zu fünf Minuten für die komplexen Szenen, wie das vollständig montierte Flugzeug.