

Formbasierte Identifikation zweidimensionaler Objekte

Dissertation

zur Erlangung des Doktorgrades
des Fachbereichs Mathematik
der Universität Hamburg

vorgelegt von
Rolf Lakämper
aus Gütersloh

Hamburg
2000

Als Dissertation angenommen vom Fachbereich Mathematik der
Universität Hamburg

auf Grund der Gutachten von Prof. Dr. U. Eckhardt
und PD Dr. L.J. Latecki

Hamburg, den 3.2.2000

Prof. Dr. H. Daduna
Dekan des Fachbereichs Mathematik

Inhaltsverzeichnis

1	Einleitung	7
1.1	Allgemeine Einführung	7
1.2	Motivation: Bilddatenbanken	9
1.3	Formbeschreibung	10
1.3.1	Klassische Ansätze	10
1.3.2	Formteile	12
1.3.3	Scale-Space Theorie	13
1.4	Inhalt dieser Arbeit	15
1.5	Vielen Dank	17
2	Grundlagen	19
2.1	Bezeichnungen aus der Bildverarbeitung	19
2.2	Aufbereitung des Bildmaterials	20
2.3	Der Tangentenraum	21
3	Diskrete Kurvenevolution	25
3.1	Einordnung	25
3.2	Motivation und Herleitung	26
3.3	Definition und Algorithmus	29
3.4	Der Stufenausgleich	30
3.5	Verschiedene Kostenfunktionen	33
3.6	Eigenschaften des Evolutionsverfahrens	36
3.6.1	Eigenschaften im Zusammenhang mit der Teilmengenrelation	37

3.6.2	Eigenschaften im Zusammenhang mit dem Stufenausgleich	40
3.7	Die Stetigkeit des Evolutionsverfahrens	60
3.8	Die topologierhaltende Evolution	68
3.9	Polygonevolution und digitale Strecken	74
3.10	Vergleich mit anderen Verfahren	82
3.10.1	Critical Point Detection	83
3.10.2	Scale Space Evolution	84
3.11	Einsatzgebiete	86
3.11.1	Signalverarbeitung	86
3.11.2	Vereinfachung von Graphen	86
3.11.3	Segmentierung zeitdiskreter mehrdimensionaler Signale	88
3.12	Ausblick	89
4	Formteile	91
4.1	Einleitung und Motivation	91
4.2	Formunterteilung durch Kurvenevolution	96
4.3	Ergebnisse	100
5	Formvergleich	103
5.1	Einleitung	103
5.2	Korrespondenz der visuellen Teile	105
5.2.1	Definition der Korrespondenz	105
5.2.2	Der Vergleich zweier Objekte	107
5.2.3	Mögliche Gruppierungen: ein Beispiel	108
5.2.4	Dynamisches Programmieren	111
5.3	Ähnlichkeitsmaß für Polygone	115
5.4	Ergebnisse der Korrespondenz	116
5.5	Ergebnisse des Formunterschieds	118
5.5.1	Vergleich mit Siddiqi et al.	118
5.5.2	Vergleich mit Mokhtarian et al.	119

<i>INHALTSVERZEICHNIS</i>	5
5.5.3 Vergleich mit Basri et al.	119
5.5.4 Vergleich mit Arkin et al.	124
5.6 Weitere Ergebnisse	125
5.6.1 Perspektivische Rotation	126
5.6.2 Perspektivische Verzerrung	126
5.6.3 Planare Verzerrung	126
5.6.4 Rauschen	126
6 Die Datenbank, MPEG-7	131
6.1 Die Datenbank	131
6.2 MPEG-7	132
6.2.1 Überblick	133
6.2.2 Test A: Robustheit gegenüber Skalierung und Rotation . .	134
6.2.3 Test B: Ähnlichkeit	135
6.2.4 Test C: Bewegung und Deformation	136
6.2.5 Gesamtergebnis	136
7 Zusammenfassung	139

Kapitel 1

Einleitung

1.1 Allgemeine Einführung

'Multimedia' und 'Interaktion' sind Schlagworte in diesem Jahrzehnt. Kein Produkt der Unterhaltungselektronik kann ohne den Hinweis auf Fähigkeiten in diesen Bereichen erfolgreich präsentiert werden. Auch wenn die Prädikate oftmals unzutreffend vergeben werden, sind doch die Wünsche und Vorstellungen dahinter äußerst interessant: durch die Interaktion, also dem Zusammenspiel zwischen Benutzereingabe und Programmreaktionen, sollen dem multimedialen System, also einem System, welches *'... durch die rechnergesteuerte, integrierte Erzeugung, Manipulation, Darstellung, Speicherung und Kommunikation von unabhängigen Informationen gekennzeichnet ist,...*' [32], Inhalte und Zusammenhänge dieser Informationen (meist Bild und Ton) entlockt werden können.

Doch die (multimediale) Interaktion stößt schnell an die Grenzen der Leistungsfähigkeit der angepriesenen Produkte, denn das Zusammenspiel des Nutzers und seines Programmes setzt – um interessant zu werden – eine *Deutung* der gegenseitig übermittelten Informationen voraus, die auf beiden Seiten zumindest gering übereinstimmt:

So ist ein digitaler Videoschnittplatz, der von der Programmseite her nur Anfragen der Art:

'Zeige das 475te Bild'

versteht zwar ein nützliches Werkzeug, die Codierung der Anfrage über den Bildindex jedoch unterfordert deutlich die Kommunikationsmöglichkeiten des Anwenders. Eine befriedigendere Anfrage, die das Prädikat 'interaktiv' mit Leben füllen könnte, wäre beispielsweise:

'Zeige das erste Bild, auf welchem ein Pferd vollständig zu sehen ist'.

Um Kommunikation mit dem Computer zu ermöglichen, muß dieser also die erzeugten, manipulierten oder dargestellten Informationen dem menschlichen Verständnis ähnlich deuten können.

Unterstützt von der rasanten Entwicklung insbesondere der digitalen Signalverarbeitung wird im Bereich der Bildverarbeitung dieser Schritt, nämlich der Übergang von der reinen Repräsentation bzw. Bearbeitung (z.B. Datenkompression, Verbesserung der Bildqualität usw.) hin zur Deutung, d.h. zur semantischen Codierung der Bilddaten, gerade vollzogen.

Bezeugt wird dieses z.B. durch die Entwicklung des Videostandards MPEG (=Motion Pictures Expert Group): war MPEG-2 noch ein 'passiver' Standard zur reinen Videokompression, also Bildübermittlung und Darstellung, so ist im nächsten Schritt, dem MPEG-4 Standard, durch die Möglichkeit getrennter Übertragung verschiedener Bildinhalte aktives Eingreifen in die Bilddaten möglich. Die aktuelle Entwicklung schließlich ist speziell auf Interaktion ausgerichtet: der MPEG-7 Standard unterstützt den Zugriff auf einzelne Bildinhalte, so daß das Videomaterial als Bilddatenbank dient.

Um einen solchen Standard erfüllen zu können, muß der Computer also Bildszenen nicht nur reproduzieren, sondern auch beschreiben können. Die Beschreibung der einzelnen Bildinhalte verlangt dabei eine Sprache, die der visuellen Wahrnehmung des Menschen angepaßt sein muß. Als objektbeschreibende Parameter kommen z.B. in Frage:

- Oberflächenbeschaffenheit (Textur)
- Farbe
- Bewegung
- Form

Verfahren, die auf die Analyse der drei erstgenannten Parameter zielen, gehören zur klassischen Bildverarbeitung und sind damit seit einiger Zeit erforscht.

Es hat sich in vielen Anwendungen herausgestellt, daß derart klassisch parametrisierte Bildszenen unzureichend beschrieben sind, da in bezug auf das menschliche Bildverständnis vor allem Formeigenschaften entscheidend für die Beurteilung von Bildinhalten sind.

Als Beispiel und als Motivation für die Notwendigkeit von Formbeschreibung sei eine typische interaktive Anwendung genannt: die Suche nach visuell ähnlichen Daten in Bilddatenbanken.

1.2 Motivation: Bilddatenbanken

Enorme Mengen an Bildmaterial stehen elektronisch zur Verfügung, nicht zuletzt das Internet boomt durch die technischen Möglichkeiten, Bilder schnell übermitteln zu können. Um diese Informationsflut ordnen zu können, bedarf es geeigneter Suchmaschinen, die gezielte, visuell basierte Anfragen an das Bildmaterial ermöglichen. Bisher existierende Datenbanken indizieren das Datenmaterial nicht bildobjektbasiert, sondern mittels lokaler, also bildpunktbasierter Eigenschaften (\rightarrow Textur) oder globaler Eigenschaften (\rightarrow Farbhistogramm). Ergebnisse solcher Verfahren verblüffen oftmals durch ihr Mißverhältnis zur menschlichen Bilddeutung.

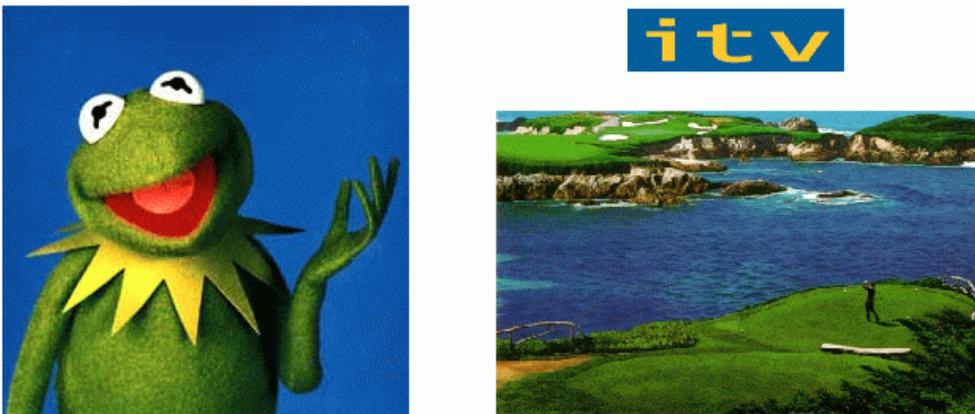


Abbildung 1.1: Drei als 'visuell ähnlich' eingestufte Bilder einer textur- und farbbasiert indizierten Bilddatenbank (<http://image.altavista.com/cgi-bin/avncgi>, Stand Juni 1999)

Abbildung 1.1 zeigt die typischen Schwächen einer textur- und farbbasierten Datenbank (hier die Suchmaschine <http://image.altavista.com/cgi-bin/avncgi>): die Objekte b) und c) sind das Ergebnis einer Anfrage nach 'visueller Ähnlichkeit' zu Objekt a). Die Merkmale 'glatte blaue Fläche' sowie 'rauhe grüne Fläche' sind die Verbindung zwischen den Bildern, eine Ähnlichkeit jedoch ist schwer auszumachen.

Abbildung 1.2 zeigt Ergebnisse einer formbasierten Suche: die Anfrage wird nur über den Umriß des gewünschten Objektes gestellt. Obgleich das Anfrage- und

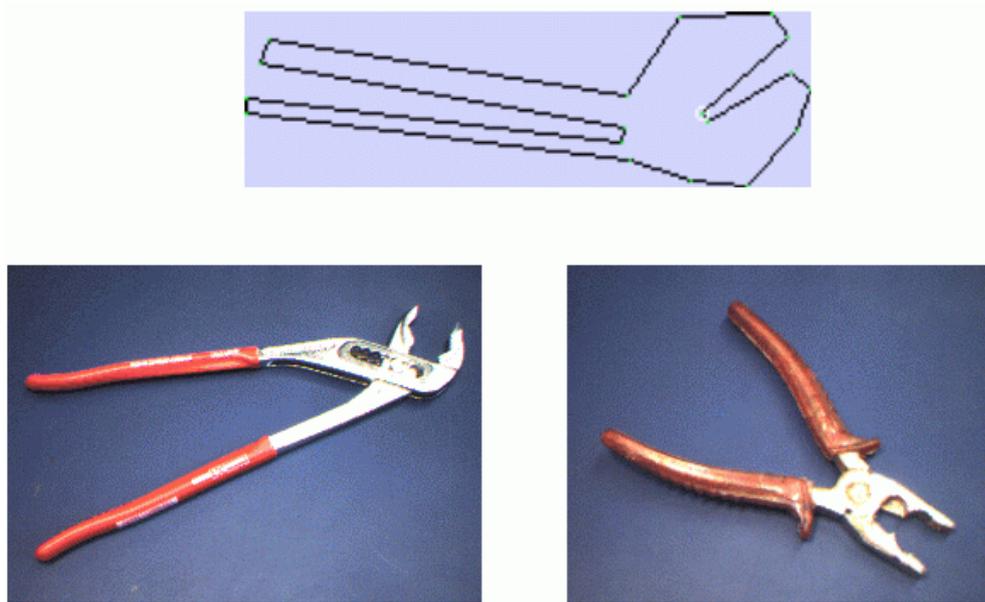


Abbildung 1.2: Drei als 'visuell ähnlich' eingestufte Objekte der formbasiert indizierten Bilddatenbank aus der vorliegenden Arbeit

Ergebnismaterial sich in Farb- und Texturinformationen stark unterscheidet und das Anfrageobjekt nur eine stark abstrahierte Skizze darstellt, d.h. die Formen nicht mathematisch identisch sind, erscheinen die Ergebnisse sinnvoll.

Es existieren beliebig viele weitere Einsatzmöglichkeiten eines Formerkennungssystems, genannt seien Anwendungen in der Autonavigation oder der optischen Qualitätskontrolle.

1.3 Formbeschreibung

1.3.1 Klassische Ansätze

Dieses Kapitel soll einen Überblick über wichtige herkömmliche Verfahren bieten, die der Formbeschreibung zugeordnet werden können. Auf Eigenschaften dieser Verfahren wird, so sie in Zusammenhang zu dieser Arbeit stehen, in späteren Kapiteln eingegangen.

In der Praxis war bisher Objektform von zweitrangigem Interesse, da die Voraussetzungen, nämlich die Zerlegung des Bildmaterials in die zu beschreibenden Formen, nicht gegeben war. Die Situation hat sich geändert:

- mit modernen Segmentierverfahren (z.B. Pauwels [8],) ist es in vielen Fällen möglich, vorhandenes Bildmaterial in Bildobjekte aufzuspalten
- Praxissituationen lassen durch Vorwissen über die Bildinhalte leichtes Segmentieren zu (z.B. Fließbandkontrolle)
- das Bildmaterial liegt segmentiert vor (→ Bluescreen- oder Virtual-Studio-Aufnahmen), der internationale Standard MPEG-4 unterstützt die segmentweise (→ MPEG-4 Layer) Codierung

Die Untersuchung von Formeigenschaften rückt somit auch in das Interesse der Praxis.

Bisherige Lösungen beschreiben Bildobjekte im wesentlichen anhand von globalen, einfachen geometrischen Eigenschaften. Hier sind zu nennen:

- Geraden- oder Kreisdetektion mittels Hough-Transformation
- Beschreibung durch Hauptachsen, Schwerpunkt, umfassende einfache Körper (→ bounding box)

Sie sind nur in sehr speziellen Gebieten einsetzbar, stellen aber den momentanen Stand praktisch eingesetzter Bildverarbeitungstechnik dar, z.B. in der optischen Kontrolle. Die vorhergehende Motivation legt eine *allgemeine* Formbeschreibung nahe, diese Aufgabe jedoch ist bisher unzureichend gelöst; dazu Kimia, Tannenbaum, Zucker [15], S.192:

'Unfortunately, standard geometries of traditional and modern mathematics do not satisfactorily address (..) aspects of shape for the purposes of object recognition. Topology is so general that bounding contours of nonfractal objects (planar, closed, and simple) are equivalent. On the other hand, Congruence geometries, such as Euclidian, affine, and projective geometries require an exact match, or some distance or area tolerance from it.'

Die allgemeine Formbestimmung stellt Fragen nach...

- ...der Definition formbestimmender Merkmale
- ...der mathematischen Beschreibung.

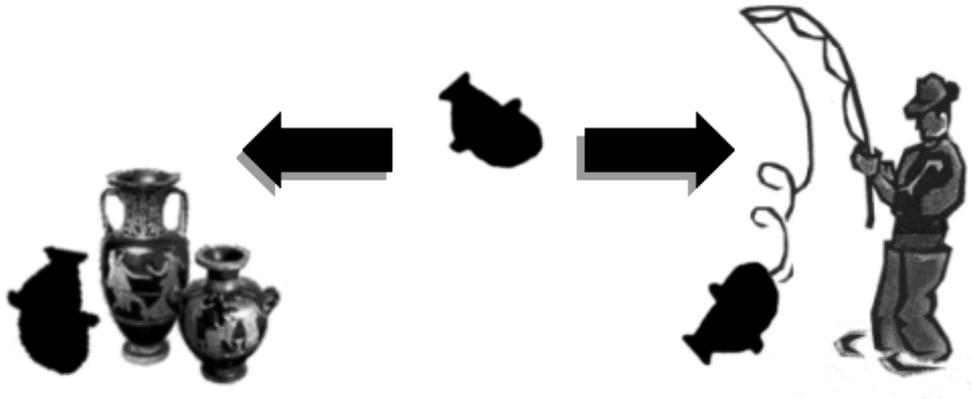


Abbildung 1.3: Kontextbezogene Formerkennung

Schon die erste Frage ist nicht eindeutig zu beantworten. Formverständnis ist vielfältiger Natur, so ist z.B. Form und Semantik von Objekten für menschliche Betrachter schwer zu trennen: das Objekt in Abbildung 1.3 a) wird als Amphore, in Abbildung 1.3 b) als Fisch bezeichnet werden, obgleich a) und b) formidentische Objekte zeigt. Andererseits können oftmals Objekte ohne jeglichen Kontext erkannt werden, siehe Abbildung 1.4. Die dort gezeigten Figuren besitzen weder Textur, Farbe, noch Bewegung, auch sind sie in keine definierende Umgebung gestellt.



Abbildung 1.4: Objekterkennung anhand der Form

Formen in semantischem Zusammenhang zu erkennen ist Aufgabe der KI-Forschung, die Beschreibung der Formen im Rahmen der Bildverarbeitung zielt auf formimmanente Eigenschaften, unabhängig vom umgebenden Rahmen.

1.3.2 Formteile

Ein wesentlicher Schritt, um der Natur des Formverständnisses näherzukommen, ist die Unterteilung der Form in einzelne Formteile, d.h. der Unterteilung des Randes oder der Fläche der Form. Die Frage nach der Definition formbestimmender

Merkmale beantwortet sich also mit der Definition von Merkmalen, durch welche eine Flächen- oder Randunterteilung gegeben ist.

Die mathematische Beschreibung dieser Trennungsmerkmale ist zumindest im Fall der Betrachtung des Objektrandes meist lokaler Natur: die Punkte der Randkurve werden anhand ihres Krümmungsverhaltens bewertet. So sind z.B. bei Hoffman und Richards [13] Trennungspunkte durch negative Krümmungsminima definiert, d.h. die Kurve wird in konvexe Bögen zerlegt.

Die Betrachtung lokaler Eigenschaften erzeugt zwei Probleme:

- Die mathematische Beschreibung greift auf Begriffe der Differentialgeometrie zurück, die zu untersuchenden Objekte aber liegen im Rechner digital vor. Das bedeutet, daß entweder die mathematische Beschreibung auf die digitale Natur des Objekts angepaßt, oder das Objekt kontinuierlich modelliert werden muß. Beide Richtungen sind problematisch: die Übersetzung der Differentialgeometrie scheitert prinzipiell an notwendigen Voraussetzungen zu Grenzwertprozessen, die im Digitalen nicht gegeben sind. Bekannte Lösungen, die diese Probleme umgehen (z.B. Definition der Krümmung durch Betrachtung von Nachbarpunkten in einer *festen* Umgebung), führen zu Definitionen, die sich äußerst anfällig gegenüber Störungen verhalten und damit in der Praxis zu erheblichen Schwierigkeiten führen. Die Überführung der digitalen Objekte in eine kontinuierliche Beschreibung erhöht zum einen den Rechenaufwand erheblich, zum anderen löst auch dieser Ansatz nicht ein prinzipielles Problem lokaler Berechnung:
- Verschiedenen z.B. durch konvexe Bögen definierten Teilen muß bei der Identifizierung der Gesamtform eine unterschiedliche Gewichtung zugeordnet werden: Einzelheiten von Objekten verlieren sich je nach Abbildungsgröße bzw. Entfernung zum Betrachter, in einem gewissen Rahmen aber bleibt die grobe Form erhalten und das Objekt weiterhin erkennbar, so ist nicht jeder einzelne Finger so entscheidend für die Identifizierung des Umrisses eines Menschen wie Arme und Beine. Lokale Definitionen können eine solche Bewertung nicht hergeben.

Dieses Problem führt zur *scale space theory*:

1.3.3 Scale-Space Theorie

Die *scale space theory* transformiert das betrachtete Objekt , — z.B. gegeben durch eine randbeschreibende Kurve — in eine kontinuierliche Folge von Objek-

ten $\{\gamma_t : t \geq 0\}$, wobei mit $t_1 > t_0$, γ_{t_1} eine vereinfachte Version von γ_{t_0} darstellt. Die Hauptidee dabei besteht darin, die Originalform $\gamma_t = \gamma_{t_0}$ mit steigendem Parameter t derart zu vereinfachen, daß zunächst Störungen (z.B. Rauschen) und visuell unwichtige Merkmale elimiert werden. Es ergibt sich eine *Auflösungspyramide*, an deren Ende eine grob vereinfachte Version des Originalobjektes steht, siehe Abbildung 1.5. Die Folge $\{\gamma_t : t \geq 0\}$ von Kurven wird als *Evolution* des Originals bezeichnet.

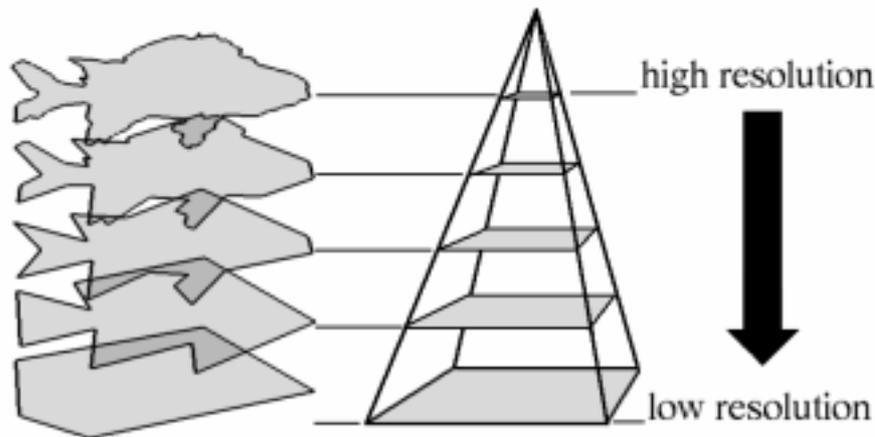


Abbildung 1.5: Auflösungspyramide: wichtige Formmerkmale bleiben bei starker Vereinfachung erhalten

Typischerweise basieren Berechnungen der Evolution auf linearen Diffusionsgleichungen (*Wärmeleitungsgleichung*) (z.B. Weickert [39]), deren Lösung sich auf Anwendung eines Gauß-Filters (Faltung der Ausgangsfunktion mit einer Gaußkurve) zurückführen läßt. Die Filterstärke — und damit die Evolutionsstufe — wird bestimmt durch die Breitenskalierung t der Gaußkurve. Das Verhalten dieses Filters wird in der klassischen Signaltheorie ausführlich behandelt.

Ein Beispiel:

Eine Evolution einer (kontinuierlichen) Randkurve $\gamma_t \subset \mathbb{R}^2$ mittels Gaußfaltung findet sich bei Mokhtarian et al. [31]. Abbildung 1.6) zeigt das Originalbild (a), die segmentierte (b) sowie die evolvierte (c) Randkurve. In [31] wird die Form anhand des Krümmungsverhaltens der Randkurve in verschiedenen Evolutionsstufen beschrieben, diese Beschreibung führt dort zu einem Ähnlichkeitsmaß von Formen.

Die Probleme der lokalen Betrachtungen werden gemindert, indem in jedem Evolutionsschritt lokale Punkte Informationen der sie umgebenden Punkte erhalten, d.h. mit jedem Evolutionsschritt repräsentiert jeder Punkt zunehmend globale



Abbildung 1.6: Gaußfilterung der Randkurve (die Bilder stammen von [30])

Information.

Zusätzlich ermöglicht die Verwendung von Scale-Space-Verfahren die Extraktion von Formmerkmalen auf verschiedenen visuellen Bedeutungsstufen. Für die oben genannte Aufteilung der Formen bedeutet dieses, daß grobe Teilungen in stark vereinfachten Strukturen gesucht werden, weitere Unterteilungen davon in Stufen der Evolution, die dem Original näher kommen.

1.4 Inhalt dieser Arbeit

Die vorliegende Arbeit beschäftigt sich mit der Beschreibung und Klassifizierung von statischen zweidimensionalen Objekten, wie sie z.B. durch Segmentierung aus Fotomaterial entstehen. Das Ziel dieser Arbeit steht in engem Zusammenhang mit dem schon in der Einleitung erwähnten internationalen Videostandard MPEG-7: dieser Standard wird im Laufe des Jahres 2000 verabschiedet. Ein wesentlicher Teil darin widmet sich der formbasierten Beschreibung von Bildobjekten und insbesondere der formbasierten Anfrage nach ihnen. Ziel dieser Arbeit ist es, in diesem Gebiet einen Beitrag zu liefern, der in diesen Standard miteinbezogen werden kann.

Ausgehend von dem segmentierten Bildmaterial wird ein System zur Beschreibung und zum Vergleich von zweidimensionalen digitalen Bildobjekten entwickelt, welches in einer in normaler Personalcomputerumgebung laufenden formbasierten Bilddatenbank Einsatz findet. Durch die Zielsetzung im Zusammenhang mit dem MPEG-7 Standard ist die Entwicklung dieses Systems stark praxisorientiert, und so ist die Kapitelaufteilung dieser Arbeit direkt an die Modularisierung des zugehörigen Computerprogrammes angelehnt. Die Anordnung der Kapitel entspricht ebenfalls der algorithmischen Reihenfolge.

Die Kapitel widmen sich folgenden Inhalten:

Im Gegensatz zu vielen klassischen, kontinuierlichen Verfahren wird stark auf die digitale Natur des zugrundeliegenden Datenmaterials geachtet: das Objekt wird

durch eine digitales Randpolygon $\subset \mathbb{Z}^2$ repräsentiert, alle Operationen verbleiben in dieser Darstellung. Kapitel 2 enthält dazu grundlegende Informationen und Definitionen.

Um die Robustheit der Bilderkennung zu gewährleisten, muß die extrahierte Randkurve geglättet und abstrahiert werden. Hierzu wird auf die Scale-Space-Theorie zurückgegriffen. Es wird eine neue Form der Kurvenevolution vorgestellt, welche sich grundsätzlich von klassischen Verfahren unterscheidet. Zunächst nur als glättende Filtervorstufe gedacht zeigte sich im Laufe der Entwicklung des Systems, daß dieses Evolutionsverfahren weitergehende Eigenschaften aufweist, welche sie zum Kernstück der Arbeit macht. Kapitel 3 beschäftigt sich daher sehr ausführlich mit diesem Evolutionsverfahren, zeigt die wichtigsten Eigenschaften auf und gibt einen Ausblick auf Einsatzmöglichkeiten auch außerhalb der Formbeschreibung.

Die Beschreibung der Form innerhalb der Bildverarbeitung zielt, wie schon in der Einleitung erwähnt, auf forminvariante Eigenschaften. Kapitel 4 zeigt Motivationen für die Existenz solcher 'naturegebener' Formmerkmale. In dem vorliegenden System wird, wie auch in klassischen Verfahren üblich, die Randkurve des Objektes in Teilkurven zerlegt. Die Definition der Aufteilung wird gegeben sowie der Einsatz des Evolutionsverfahrens in der Randunterteilung aufgezeigt.

Die durch Glättung, Abstraktion und Aufteilung der Randkurve entstandene Darstellung des Ausgangsobjektes ist Grundlage des Vergleiches des Objektes mit anderen Formen, z.B. Objekten in einer Bilddatenbank. Der Algorithmus des vorliegenden Systemes leistet dabei zweierlei: zum einen wird eine Zuordnung der einzelnen Teile der Vergleichsobjekte hergestellt, zum anderen wird die Ähnlichkeit der Objekte errechnet. Dabei wird auf Übereinstimmung mit menschlicher Intuition (außerhalb semantischer Deutungen) geachtet. Kapitel 5 zeigt die Vorgehensweise der Teilzuordnung, definiert das Vergleichsmaß und zeigt Eigenschaften auf.

Damit sind die Einzelkomponenten des Bilderkennungssystemes geschaffen, Kapitel 6 fügt die Module im Einsatzbeispiel der Bilddatenbank zusammen. Die Datenbank wird anhand des MPEG-7 Testmaterials zu anderen Verfahren, welche innerhalb des Standards vorgeschlagen wurden, verglichen.

Die Ergebnisse der Arbeit basieren auf [27] [20] [26] [23] [22] [21] [18] [25] [24], einige Ergebnisse sind als JAVA-Applets auf der Internetseite [28] zu sehen.

1.5 Vielen Dank

Dank gilt Herrn Prof. Dr U. Eckhardt für seine persönliche und fachliche Unterstützung, besonders aber für seinen Humor. PD Dr. Longin J. Latecki danke ich im besonderen für die hervorragende Zusammenarbeit und die Unmenge an guten Ideen !

Kapitel 2

Grundlagen

2.1 Bezeichnungen aus der Bildverarbeitung

Definition: Digitale Ebene

Die Bildverarbeitung bezeichnet \mathbb{Z}^2 als *digitale Ebene*.

Definition: 4,8-Nachbarschaft

Es sei p ein Punkt der digitalen Ebene.

Die Menge $\mathcal{N}_8(p) = \{q \in \mathbb{Z}^2 \mid \|p \leftrightarrow q\|_\infty = 1\}$ wird (aufgrund der Punktanzahl) als *8-Nachbarschaft* des Punktes p bezeichnet.

Die Menge $\mathcal{N}_4(p) = \{q \in \mathbb{Z}^2 \mid \|p \leftrightarrow q\|_1 = 1\} \subset \mathcal{N}_8$ wird entsprechend als *4-Nachbarschaft* des Punktes p bezeichnet, siehe Abbildung 2.1 a).

Definition: Digitale 4,8-Kurve

Eine *digitale Kurve* $\mathcal{D} = (d_0, d_1, \dots, d_n) \subset \mathbb{Z}^2$ ist eine Folge von $n + 1$ Punkten d_k der digitalen Ebene (versehen mit der Nachbarschaftsrelation $\mathcal{N}_{i=4,8}$) derart daß:
für $n < 2$: d_0 und d_1 \mathcal{N}_i -benachbart sind,

für $n \geq 2$: $d_{j=1, \dots, n-1}$ genau von d_{j-1} \mathcal{N}_i -benachbart ist.

($\mathcal{N}_4 \rightarrow$ 4-Kurve, $\mathcal{N}_8 \rightarrow$ 8-Kurve, $d_0 = d_n \rightarrow$ geschlossene Kurve, $d_0 \neq d_n \rightarrow$ offene Kurve), siehe Abbildung 2.1 b),c).

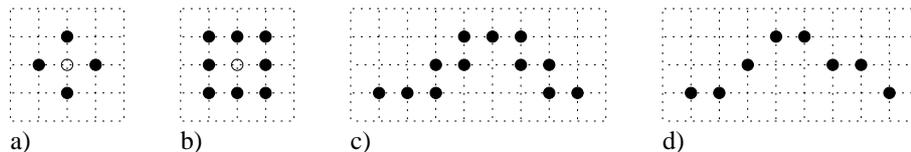


Abbildung 2.1: Nachbarschaft: a) \mathcal{N}_4 b) \mathcal{N}_8 c) 4-Kurve d) 8-Kurve

Definition: Randpolygon und zugeordneter Kantenzug

Die Elemente d_i der digitalen Randkurve (oftmals wegen der vereinfachten Darstellung bei Freeman [9]) als *Chain*-bzw. *Freemancode* bezeichnet) können als Eckpunkte eines Randpolygons, gedeutet werden siehe Abbildung 2.2.



Abbildung 2.2: Die digitale 8-Randkurve a) liefert die Eckpunkte des digitalen Randpolygons b)

Zu unterscheiden ist zwischen dem Randpolygon \bar{C} , und dem ihm zugeordneten *Kantenzug* \bar{C} :

\bar{C} ist definiert durch die Eckpunkte der digitalen Kurve, also $\bar{C} \subset \mathbb{Z}^2$, unter dem \bar{C} zugeordneten *Kantenzug* \bar{C} versteht man die durch \bar{C} definierten Strecken $(\overline{d_0 d_1}) \subset \mathbb{R}^2$, also: $\bar{C} = ((\overline{d_0 d_1}), (\overline{d_1 d_2}), \dots, (\overline{d_{n-1} d_n})) \subset \mathbb{R}^2$.

Definition: Digitales Grauwertbild

Unter einem *digitalen Grauwertbild* versteht man eine Abbildung einer Menge $\mathcal{I} \subset \mathbb{Z}^E$ in ein Intervall $G \subset \mathbb{Z}$ (z.B. Grauwertbild, typischerweise: $G = [0, 255]$) bzw. $F \subset \mathbb{Z}^3$ (z.B. RGB-Farbbild, typischerweise $f = [0, 255]^3$).

2.2 Aufbereitung des Bildmaterials

Alle in dieser Arbeit beschriebenen Operationen bearbeiten Randkurven, welche direkt aus den durch 8-Kurven dargestellten Rändern der Bildobjekte gewonnen werden. Dabei wird das Bildmaterial als Grauwertbild vorausgesetzt.

Um zu einer polygonalen Randrepräsentation zu gelangen, müssen die gewünschten Objekte aus dem Bild heraussegmentiert werden. Gängige Verfahren dazu sind z.B. Segmentierung durch Farb- und Texturclusteranalyse oder Wasserscheidenverfahren [19].

Man erhält eine binäre Bildmaske, welche noch verschiedene Filterstufen durchläuft (z.B. morphologische Filter zur Eliminierung ungewünschter Bildstörungen), bevor der Objektrand als digitale 8-Kurve extrahiert wird. Abbildung 2.3 zeigt diese Arbeitsschritte.

Diese 8-Kurve schließlich liefert direkt das gewünschte Randpolygon.

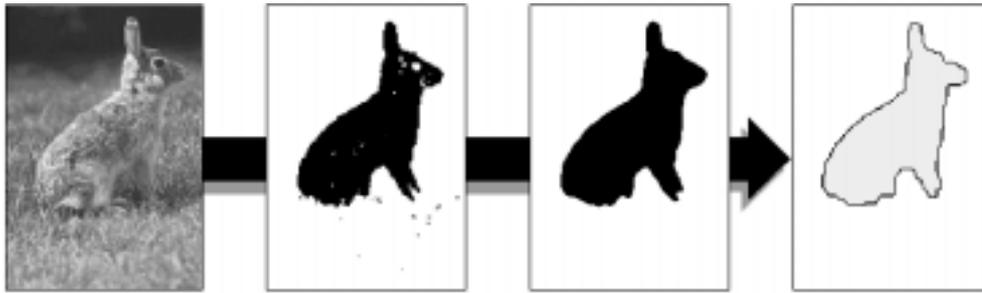


Abbildung 2.3: Von der Bildvorlage zur digitalen Randkurve

2.3 Der Tangentenraum

Die Darstellung des Randpolygons \mathcal{P} , durch seine Eckpunkte p_i in \mathbb{Z}^2 enthält Informationen, die in den folgenden Betrachtungen, insbesondere bei der Herleitung der Evolutionsvorschrift (Kapitel 3), nicht benötigt werden. Wird das Polygon um Translationsinformationen reduziert, so bleiben erhalten:

- der (euklidische) Abstand der Eckpunkte zueinander
- die Winkellage der Segmente $\overline{p_i p_{i+1}} \in \overline{\mathbb{Z}^2}$,

Genau diese Informationen sind in der *Tangentenraumdarstellung* enthalten.

Die folgende Darstellung geht der Einfachheit halber davon aus, daß \mathcal{P} , ein *geschlossenes* Polygon darstellt, mit leichten Änderungen gelten alle Angaben auch für nichtgeschlossene Polygone.

Zur Transformation des Polygons \mathcal{P} , in die Tangentenraumdarstellung wird $\overline{\mathbb{Z}^2}$ von $\overline{p_0 p_1}$ ausgehend durchlaufen, wobei die Länge sowie die Winkelausrichtung von $\overline{p_i p_{i+1}}$ extrahiert werden.

Formal stellt der Tangentenraum (bei geschlossenen Polygonen, von denen wir hier ausgehen) einen Torus $S_1 \times S_2$ dar, wobei S_1 einen Kreis mit Durchmesser 1, S_2 einen Kreis mit Durchmesser 2π benennt. An dieser Stelle jedoch soll der Tangentenraum –weniger formal als anschaulich– als Rechteck mit Seitenlängen 1 sowie 2π gezeigt werden, welches in üblicher Weise in ein Koordinatensystem in \mathbb{R}^2 eingebettet ist.

Definition: Tangentenraum \mathcal{T}

Gegeben sei ein Polygon \mathcal{P} , $\mathcal{P} = (p_0, p_1, \dots, p_n)$ sowie der zugehörige Kantenzug $\overline{\mathcal{P}}$. Die disjunkten Intervalle $\mathcal{I}_i, i = 0, \dots, n \Leftrightarrow 1$ seien eine Überdeckung des Intervalles $[0, 1] \subset \mathbb{R}$, $\mathcal{I}_i = [a_i, a_{i+1})$ mit $a_i = \sum_{j=0}^i |\overline{p_j p_{j+1}}|$.

Die Werte $w_i \in [0..2\pi]$ seien gegeben durch den Winkel der Strecke $\overline{p_{i+1} \leftrightarrow p_i}$ zur x -Achse, $\overline{}$ sei parametrisiert durch die normalisierte Bogenlänge t , ausgehend von p_0 .

Dann ist die Abbildung $\mathcal{T} : \mathbb{R} \supset [0, 1] \rightarrow \mathbb{R}$ gegeben durch:
 $t \rightarrow \mathcal{T}(t)$ mit $\mathcal{T}(t) = w_k, t \in \mathcal{I}_k$.

Ein Beispiel zeigt Abbildung 2.4. Auf der y -Achse wird die Winkellage, auf der x -Achse die Segmentlänge abgetragen, wobei die gesamte Umlauflänge des Polygons zu 1 normiert ist.

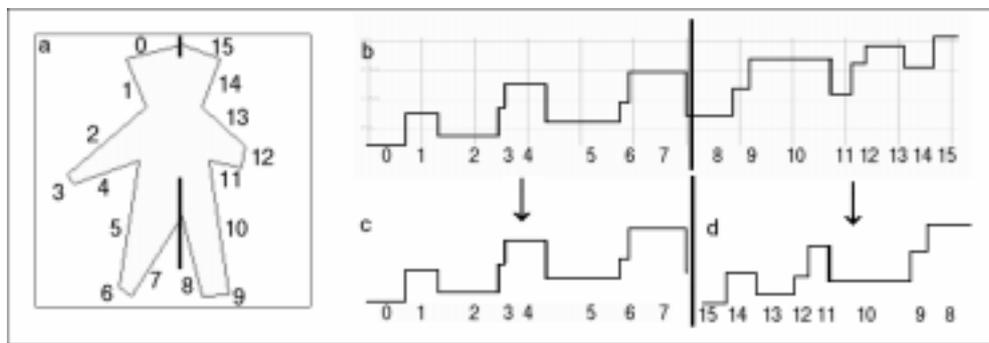


Abbildung 2.4: a) polygonale Randkurve und ihre Tangentenraumdarstellung b). An der Ähnlichkeit von c) und d) ist die Symmetrie des Objektes erkennbar

Zum intuitiveren Verständnis des Tangentenraumes seien einige Eigenschaften erwähnt:

- Die Höhendifferenz der Stufen an den Stellen a_i zeigt den Wendewinkel in p_i
- Konvexe (konkave) Bogenabschnitte in $\overline{}$ entsprechen genau monoton steigenden (fallenden) Intervallen im Tangentenraum
- Die Rotation von $\overline{}$ um den Winkel α entspricht einer Verschiebung der Stufenfunktion um α in y -Richtung
- Eine weitergehende Eigenschaft, die hier kurz erwähnt werden soll, im folgenden aber nicht mehr benutzt wird: Symmetrien der durch $\overline{}$ dargestellten Form sind im Tangentenraum gut sichtbar (siehe Abbildung 2.4): in $\overline{}$ verlaufe o.B.d.A. die Symmetrieachse durch p_0 und p_s (in Abbildung 2.4: p_0 und p_8). Sind \mathcal{C}_1 und \mathcal{C}_2 die Kurven über den Intervallen $[0, a_s)$ sowie $(a_s, a_{n-1}]$ (Abbildung 2.4 b), und bezeichnet \mathcal{C}'_2 die Spiegelung von \mathcal{C}_2 an der

y -Achse (Abbildung 2.4 d), so muß –intuitiv verstanden– \mathcal{C}_1 ähnlich \mathcal{C}'_2 verlaufen, insbesondere müssen beide Kurven ähnliches Monotonieverhalten aufweisen. Diese Eigenschaft kann zur Ermittlung der Symmetrien benutzt werden.

Eine weitergehende formale Beschreibung des Tangentenraumes für kontinuierliche Polygone findet sich in Zahn und Roskies [41], dort wird die Stufenfunktion im Tangentenraum *cumulative angular bend function* genannt. Zahn und Roskies benutzen die Tangentenraumtransformation nur als Zwischenschritt zur Beschreibung planarer Kurven durch Fourierdeskriptoren, nicht zur direkten Untersuchung der Kurve.

In der vorliegenden Arbeit wird der Tangentenraum zur Motivation des Stufenausgleiches bei der Polygonevolution (Kapitel 3) sowie bei der Herleitung des Vergleichsmaßes von Polygonzügen (Kapitel 5) verwendet.

Kapitel 3

Diskrete Kurvenevolution in digitalen Bildern

3.1 Einordnung

Um Robustheit des Formerkennungsverfahrens zu gewährleisten, muß eine Darstellung des Randpolygons ¹, erreicht werden, die nur 'sinnvolle', also visuell bedeutsame Informationen enthält. Es müssen zwei Arten überflüssiger Informationen herausgefiltert werden:

1. Die vorbereitenden Schritte zur Extraktion der Randkurve erzeugen bzw. übernehmen *Bildstörungen und verrauschte Strukturen*, siehe Abbildung 3.1.
2. Feinheiten in Formvorlagen sind für die Klassifizierung nicht notwendig, siehe Abbildung 3.1 und 3.2. Es gilt also, *relevante* von *nichtrelevanten* Randpunkten zu trennen.

Dieses Ziel kann als Approximationsaufgabe formuliert werden:

Es sei Ω die Menge aller digitalen Polygone, $p \in \Omega$. $\mathcal{I} : \Omega \rightarrow \mathbb{R}_0^+$ sei eine Abbildung, die den (Form-)Informationsgehalt eines Polygons modelliert. $\Delta : \Omega \times \Omega \rightarrow \mathbb{R}_0^+$ sei ein Ähnlichkeitsmaß, z.B. der Hausdorffabstand.

¹Die folgende Darstellung geht der Einfachheit halber davon aus, daß p ein *geschlossenes Polygon* darstellt, mit leichten Änderungen gelten alle Angaben auch für *nichtgeschlossene Polygone*. Alle Indizierungen in p sind modulo $\#$, zu verstehen.

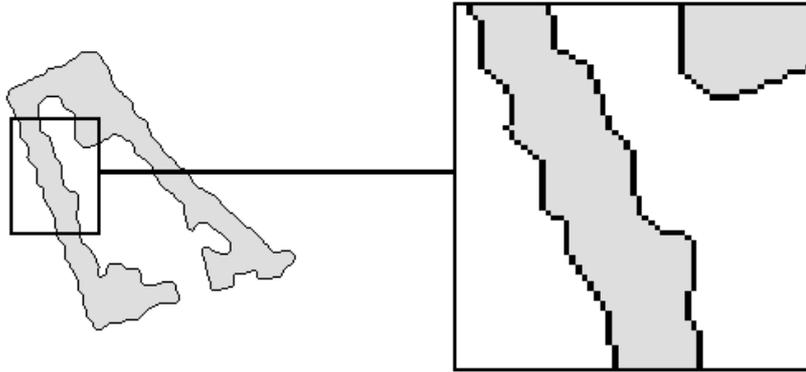


Abbildung 3.1: Randdarstellung: verrauschte Bildinformation. Obwohl die Randpunkte stark versetzt sind, ist es möglich, die Gesamtform zu erkennen.

Zu \mathcal{I} ist ein approximierendes Polygon \mathcal{I}' derart gesucht, daß $\Delta(\mathcal{I}, \mathcal{I}') < \epsilon$, $\epsilon > 0$ unter der Bedingung, daß $\mathcal{I}(\mathcal{I}') < \mathcal{I}(\mathcal{I})$, d.h. daß ein Polygon \mathcal{I}' gesucht ist, welches \mathcal{I} ähnlich ist, aber weniger Information enthält.

Klassische Verfahren ermitteln \mathcal{I}' durch lokales Versetzen der Eckpunkte. Daraus resultieren oftmals Probleme bei der Berechenbarkeit, zumindest aber dem Rechenaufwand solcher Verfahren.

3.2 Motivation und Herleitung

Der in dieser Arbeit vorgestellte Ansatz ist motiviert durch die Beobachtung, daß die aus dem Segmentationsprozeß erhaltene Randkurve zwar *Fehlinformation* (Rauschen und visuell unbedeutende Feinheiten) enthält, ihr jedoch *keine Information fehlt*, um klassifizierbar zu sein. Dabei spielt es keine Rolle, ob das durch Digitalisierung oder Segmentierung erzeugte Rauschen Randpunkte von ihrer exakten Position versetzt hat. Denn offenbar ist die Information der Gesamtform noch in der gegebenen Kontur enthalten, solange sie erkennbar ist.

Die meisten Standardansätze zur Evolution oder Approximation in der Bildverarbeitung versuchen die Störungen zu eliminieren, indem die Randpunkte auf ihre vermutete ursprüngliche Position zurückgesetzt werden. Dieses ist nur möglich, wenn die Formklasse, aus welcher das gegebene Objekt stammt, a priori bekannt ist (Beispiel: Suche von kreisförmigen Strukturen durch optimale Kreiseinpassung). Diese Vorinformation ist bei allgemeiner Formbeschreibung nicht gegeben.

Andererseits ist es nicht nötig, die exakte Position der Randpunkte der Ausgangs-

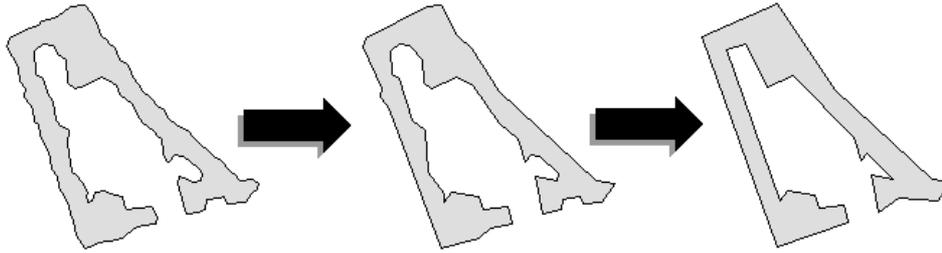


Abbildung 3.2: Entfernung redundanter Forminformation: Die Anzahl der polygonbestimmenden Eckpunkte wurde auf $\approx 3\%$ reduziert. Die Figur bleibt anhand ihrer Form erkennbar, obwohl sie nicht semantisch oder funktional klassifiziert werden kann.

form zu bestimmen, solange die Form auch mit versetzten Punkten erkennbar ist. Interpretiert man diese Aussage punktweise, so bedeutet sie, daß eine Teilmenge A einer gegebenen Punktmenge \mathcal{P} , ausreicht, um die Form zu repräsentieren, die restlichen Punkte $\mathcal{P} \setminus A$ sind redundant, siehe Abbildung 3.2. Kurz:

Jede durch ein Randpolygon gegebene Form enthält eine minimale Darstellung als Teilmenge aus visuell bedeutsamen Eckpunkten.

Dieses bedeutet nicht, daß die Punkte der repräsentativen Teilmenge A sich auf einer exakten Position (z.B. math. exakter Kreis oder Gerade) befinden, sie sind trotzdem ausreichend für die Bestimmung der Form.

Die in dieser Arbeit vorgestellte diskrete Kurvenevolution erlaubt es, die repräsentative Teilmenge A eines gegebenen Randpolygons \mathcal{P} , zu extrahieren.

Formal:

Definition: Naheste Abstraktion

Gegeben sei das Randpolygon $\mathcal{P} = (p_0, p_1, \dots, p_n) \subset \mathbb{Z}^2$, $\mathcal{I} : \mathcal{P}(\mathbb{Z}^2) \rightarrow \mathbb{R}_0^+$ sei ein Maß für den Informationsgehalt von \mathcal{P} . Als *naheste Abstraktion* $\mathcal{A}(\mathcal{P}) = \mathcal{P}'$ an \mathcal{P} , wird ein Polygon \mathcal{P}' bezeichnet, für welches gilt:

1. $\mathcal{P}' \subset \mathcal{P}$,
2. $\mathcal{I}(\mathcal{P}') \leq \mathcal{I}(\mathcal{P})$
3. $\forall \Upsilon \subset \mathcal{P} : \mathcal{I}(\Upsilon) \leq \mathcal{I}(\mathcal{P}')$

Das Maß $\mathcal{I}(\mathcal{P})$ für den Informationsgehalt wird dabei durch die Summe der In-

formationen $K_\Gamma : \mathcal{A} \rightarrow \mathbb{R}_0^+$ in den Eckpunkten $p_i \in \mathcal{A}$, gebildet:²

$$\mathcal{I}(\mathcal{A}) = \sum_{i=0}^{\#\Gamma-1} K_\Gamma(p_i)$$

$K_\Gamma(p_i)$ modelliert dabei die visuelle Bedeutung des Punktes p_i , es wird darauf in Kapitel 3.4 tiefer eingegangen. K_Γ wird im folgenden als *Kostenfunktion*, der Wert $K_\Gamma(p_i)$ als *Kosten* im Punkt p_i bezeichnet, er gibt den 'Preis' des Informationsverlustes der Abstraktion an.

Mit der Definition von \mathcal{I} und den Forderungen an \mathcal{A} folgt sofort:

$$\#\mathcal{A}' \Leftrightarrow \#\mathcal{A}, \quad \#\mathcal{A}' = 1, \quad \mathcal{A}' \setminus \mathcal{A} = \min_{p_i \in \Gamma} \{K_\Gamma(p_i)\}$$

d.h. die naehste Abstraktion \mathcal{A}' unterscheidet sich vom Ausgangspolygon genau um den Punkt geringster visueller Bedeutung. Dabei ist zu beachten, daB sich diese Definition nur auf die *Eckpunkte* des Polygons bezieht.

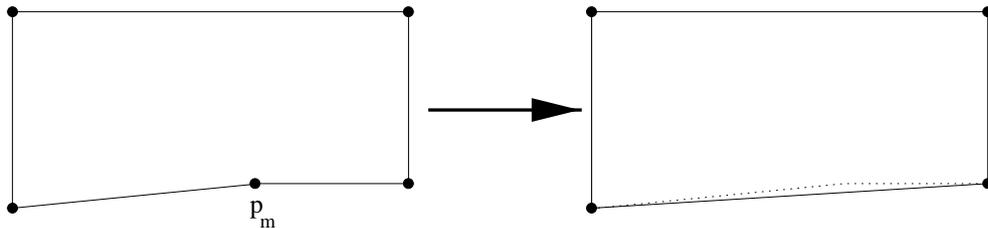


Abbildung 3.3: Naehste Abstraktion durch Polygonlinearisierung

Für den \mathcal{A} , zugeordneten Kantenzug $\overline{\mathcal{A}}$ bedeutet das:

Es sei m der Index des Punktes p_m mit $K_\Gamma(p_m) = \min_{p_i \in \Gamma} \{K_\Gamma(p_i)\}$

Dann ergibt sich mit $p'_i \in \mathcal{A}'$

$$\overline{\mathcal{A}'} = \bigcup_{i=0}^{\#\Gamma'-2} \overline{p'_i p'_{i+1}} = \overline{\mathcal{A}} \setminus (\overline{p_{m-1} p_m} \cup \overline{p_m p_{m+1}}) \cup \overline{p_{m-1} p_{m+1}}$$

d.h. die naehste Abstraktion entsteht durch geradliniges Verbinden der Nachbarpunkte des (entnommenen) Minimalpunktes. Dieser Vorgang wird als *Polygonlinearisierung* bezeichnet, siehe Abbildung 3.3.

²Mit der Summendefinition über K ist \mathcal{I} ein Maß im maßtheoretischen Sinn, als zugrundeliegende σ -Algebra kann der Einfachheit halber die Potenzmenge $\mathcal{P}(\mathbb{Z}^2)$ angesehen werden.

3.3 Definition und Algorithmus

Eine Evolution erhält man durch iteratives Abstrahieren des Ausgangspolygons \mathcal{P} , :

Definition: Diskrete Kurvenevolution

Gegeben sei wiederum das Randpolygon $\mathcal{P} = (p_0, \dots, p_n) \subset \mathbb{Z}^2$ mit $n \geq 3$. Die *diskrete Kurvenevolution* von \mathcal{P} ist eine Folge

$$\mathcal{E}_\Gamma = (\mathcal{P}^0, \mathcal{P}^1, \dots, \mathcal{P}^m), \quad m = \#, \Leftrightarrow 3$$

von Polygonen $\mathcal{P}^i \subset \mathcal{P}$, für welche gilt: $\mathcal{P}^{i+1} = \mathcal{A}(\mathcal{P}^i)$, wobei \mathcal{A} die nahesten Abstraktion bezeichnet.

Die Evolution läßt sich sehr einfach algorithmisch formulieren:

Diskrete Kurvenevolution (\mathcal{E}_Γ)

$k = 0$;

Do

· Suche in \mathcal{P}^k einen Punkt p_{i_m} derart, daß $K_{\Gamma^k}(p_{i_m})$ minimal ist;

· $\mathcal{P}^{k+1} = \mathcal{P}^k \setminus p_{i_m}$;

· $k = k + 1$;

until $\#, k = 3$.

Der Algorithmus terminiert, da $\#$, endlich ist, und in jedem Schritt die Anzahl der Punkte um 1 reduziert wird. Das Abbruchkriterium 'drei Eckpunkte' hat praktische Gründe: Formen werden durch geschlossene Polygonzüge beschrieben, die einfachste Form ist das Dreieck, nachfolgende Abstraktionen sind praktisch uninteressant.

Bemerkung:

Es ist zu beachten, daß sich der Minimalwert K_{Γ^k} auf das Polygon \mathcal{P}^k bezieht, d.h. es genügt nicht, alle Werte $K_\Gamma(p_i)$ zu betrachten und die Punkte p_i in aufsteigender Reihenfolge zu entfernen.

Ein einfaches Beispiel zeigt dazu Abbildung 3.4: Die linke Spalte zeigt die ersten drei Stufen des Evolutionsverlaufes nach obigem Algorithmus, also mit Neuberechnung der nächsten Evolutionsstufe \mathcal{P}^{k+1} durch die nahesten Abstraktion $\mathcal{A}(\mathcal{P}^k)$. Dabei seien die zu entfernenden Punkte visuell gewählt. Es ergibt sich ein intuitiver Abstraktionsverlauf.

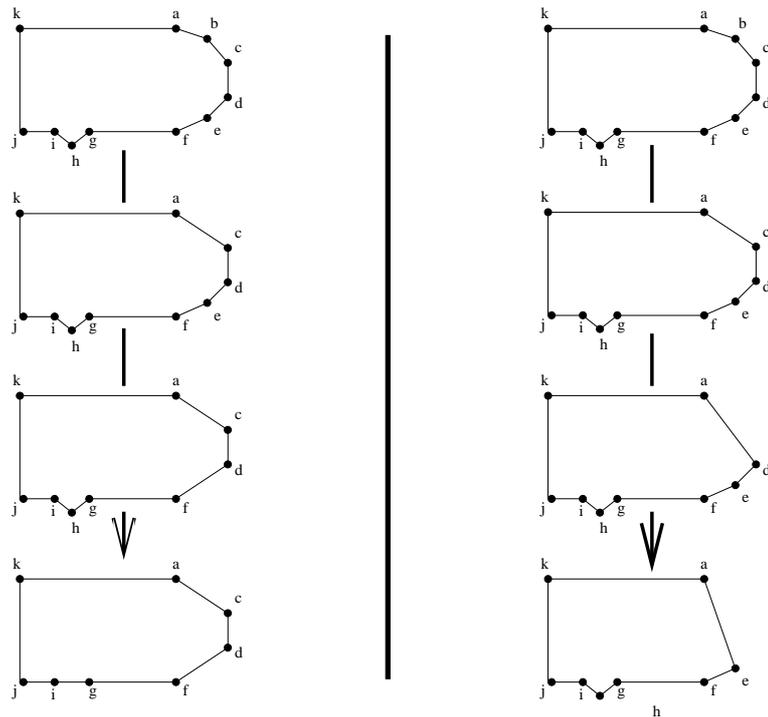


Abbildung 3.4: Evolution durch mehrfaches Anwenden der nahesten Abstraktion (linke Spalte), in der rechten Spalte dagegen wurde die Reihenfolge der zu entfernenden Punkte nur anhand des Ausgangspolygons festgelegt

Die Evolution der rechten Spalte dagegen entstand durch einmaliges (intuitives) Festlegen der visuelle Bedingungsreihenfolge der Eckpunkte des Ausgangspolygons. Dabei wurde die Reihenfolge (unwichtig \rightarrow wichtig) b,c,d,... angenommen, in dieser Reihenfolge werden die Punkte entfernt. Der Abstraktionsverlauf unterscheidet sich deutlich von der linken Spalte.

3.4 Der Stufenausgleich

So einfach der Algorithmus erscheint, so wichtig ist die Modellierung der visuellen Bedeutung der Eckpunkte durch die Kostenfunktion K_{Γ} . Sie gibt die visuelle Relevanz eines einzelnen Randpunktes für die Gesamtform an und ist der Schlüssel zur Übereinstimmung des Abstraktionsvorganges mit menschlich-intuitiver Vorgehensweise.

Die Kostenfunktion K modelliert die visuelle Relevanz eines Randpunktes p in Abhängigkeit vom Wendewinkel $\alpha(p)$ in p und der Länge der Polygonsegmente s_1 und s_2 , welche p einschließen.

		SEITENLÄNGE	
		kurz	lang
WINKEL	flach		
	spitz		

Abbildung 3.5: Motivation für die Kostenfunktion

Zur Motivation betrachte man Abbildung 3.5:

Die Abbildung zeigt vier Bilder eines rechteckig geformten Randpolygons (dünn gezeichnet) mit verschieden ausgeprägten Störungen (dick gezeichnet).

Offenbar erscheint die Form bei flachem Winkel und kurzer Segmentlänge am geringsten gestört, d.h. die visuelle Relevanz von p ist in Abbildung 3.5 oben links am geringsten. Sie steigt bei

- spitzerem Winkel
- längeren Nachbarsegmenten

Wie in Kapitel 2.3 erwähnt, extrahiert die Transformation eines Polygons in die *Tangentenraumdarstellung* genau die beachteten Polygonparameter Winkel und Segmentlänge. Wir betrachten daher zur Herleitung der Kostenfunktion K das Polygon γ , transformiert in den Tangentenraum $\mathcal{T}(\gamma)$.

Gegeben sei ein Polygon γ , parametrisiert durch die normierte Bogenlänge t , ein Eckpunkt $p = \gamma(t_p)$ sowie seine Nachbarecken $q = \gamma(t_q)$, $r = \gamma(t_r)$. Die Segmente $\overline{q_1 p}$ und $\overline{p q_2}$ seien mit s_1 und s_2 bezeichnet, ihre Winkellage sei durch $\alpha = \mathcal{T}(\frac{t_q + t_p}{2})$, $\beta = \mathcal{T}(\frac{t_p + t_r}{2})$, ihre Länge durch $l_1 = |s_1|$ und $l_2 = |s_2|$ gegeben, o.B.d.A $\alpha \geq \beta$, siehe Abbildung 3.6.

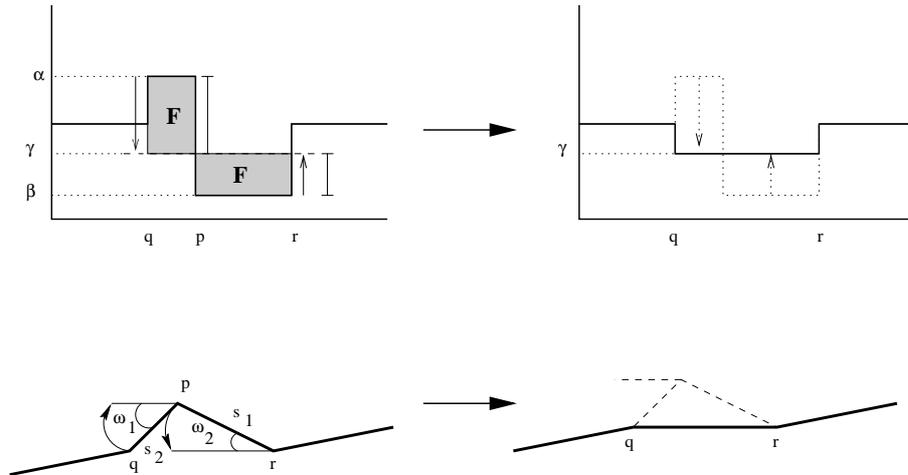


Abbildung 3.6: Herleitung der Kostenfunktion

Gesucht ist nun (im Tangentenraum) ein *Ausgleichswinkel* γ derart, daß:

$$l_1(\alpha \Leftrightarrow \gamma) = l_2(\gamma \Leftrightarrow \beta), \quad \alpha \geq \gamma \geq \beta$$

γ ist also genau derjenige Winkel, auf den die zu s_1 und s_2 gehörigen Stufen im Tangentenraum zueinander verschoben werden müssen, damit

- sie auf gleichem Winkel (nämlich γ) enden
- die bei der Verschiebung überstrichene Fläche beider Stufen betragsgleich ist.

Dieser Vorgang und damit die ganze Kostenfunktion wird *Stufenausgleich* genannt.

Einfache Berechnung ergibt:

$$\gamma = \frac{\alpha l_1 + \beta l_2}{l_1 + l_2}$$

Die Fläche F , welche von jeder der beiden Stufen im Tangentenraum überstrichen worden ist, beträgt:

$$F(q, p, r) = (\alpha \Leftrightarrow \gamma) l_1 = \frac{(\alpha \Leftrightarrow \beta) l_1 l_2}{l_1 + l_2}$$

Dieses Flächenmaß F wird als visuelle Relevanz des Punktes p angesehen, genau durch diese Fläche wird die Relevanz- bzw. Kostenfunktion des Polygons , im Punkt p definiert:

$$\boxed{K_{\Gamma}(p) = F(q, p, r)}$$

Der Stufenausgleich hat folgende Deutung auf dem Originalpolygon (siehe Abbildung 3.6):

- die vertikale *Verschiebung* einer Stufe im Tangentenraum entspricht einer *Rotation* des entsprechenden Segmentes im Originalpolygon
- die überstrichene Fläche $F(q, p, r)$ im Tangentenraum entspricht der Bogenlänge des im Originalraum durch die Rotation beschriebenen Kreissegmentes.

Die Deutung des Stufenausgleiches im Originalraum ist also, die Segmente s_1, s_2 über die kleinste Bogenlänge zu rotieren bis sie parallel ausgerichtet sind, die überstrichene Bogenlänge der dabei beschriebenen Kreissegmente muß gleich sein.

Zu beachten ist, daß die Segmente parallel, aber i.a. nicht kollinear enden. Bedeutung erlangt dieses bei der Suche nach der nächsten Abstraktion mittels Polygonlinearisierung: die nächste Abstraktion entsteht *nicht* durch Rücktransformation der Tangentenraumdarstellung nach dem Stufenausgleich. Der Grund dafür ist einfach: die Transformation eines geschlossenen Polygons in den Tangentenraum mit folgendem Stufenausgleich und Rücktransformation ergibt nicht notwendig ein geschlossenes Polygon. Daher wird der Stufenausgleich nur für die Berechnung der visuellen Relevanz benötigt, die nächste Abstraktion wird anhand des minimalen Kostenfunktionswertes durch direktes Verbinden der Nachbarpunkte q, r gebildet.

Die Abbildung 3.7 zeigt verschiedene Stufen eines Randpolygons, das mit dem Evolutionsverfahren mit Stufenausgleich bearbeitet wurde. Die rechte Spalte zeigt die entsprechenden Tangentenraumdarstellungen. Weitere Beispiele sind in Abbildung 3.23 zu finden.

3.5 Verschiedene Kostenfunktionen

An dieser Stelle soll näher auf die Wahl der Kostenfunktion eingegangen werden, insbesondere soll gezeigt werden, wie das Evolutionsverfahren bei sorgloser Wahl fehlschlagen kann.

Abbildung 3.8 zeigt das Ausgangspolygon 'Kangaroo', bestehend aus 143 Eckpunkten. Die Spalten a) bis d) zeigen die Evolution vier verschiedener Kostenfunktionen, die Polygone haben jeweils eine Punktanzahl von 46 (mittlere Zeile) bzw. 16 (untere Zeile).

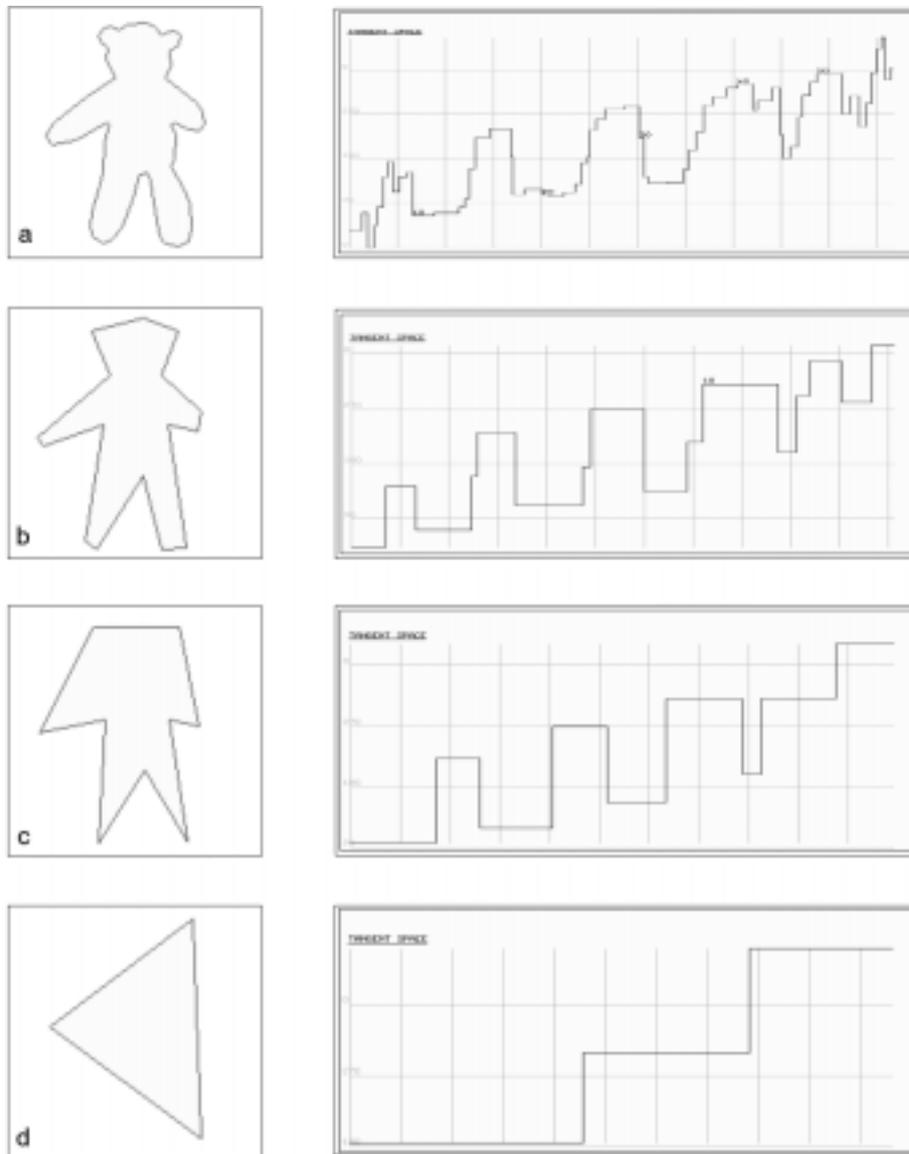


Abbildung 3.7: Evolution einer Randkurve. Die rechte Spalte zeigt die Tangentenraumdarstellung der entsprechenden Evolutionsstufe

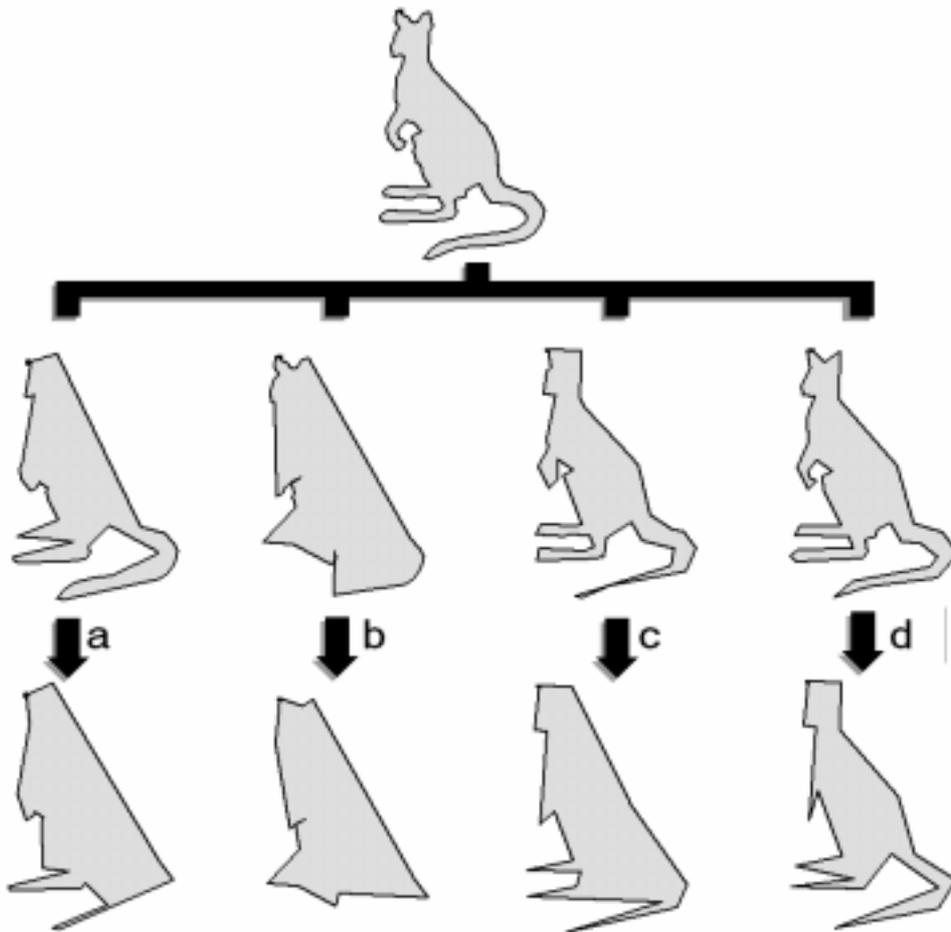


Abbildung 3.8: Vergleich der Evolution mit verschiedenen Kostenfunktionen

Als Kostenfunktion wurde benutzt (die Bezeichnungen beziehen sich auf Abbildung 3.6):

- Spalte a): $K = (\alpha \leftrightarrow \beta)$, es wird nur der Wendewinkel des jeweiligen Punktes beachtet.
- Spalte b): $K = \frac{l_1 l_2}{l_1 + l_2}$, es wird nur der hyperbolische Durchschnitt der Längen der an den jeweiligen Punkt angrenzenden Segmente beachtet.
- Spalte c) $K = |p \leftrightarrow \frac{q+r}{2}|$, der Abstand zum Mittelpunkt der Nachbarpunkte.
- Spalte d) Der in diesem Kapitel hergeleitete Stufenausgleich

Die Spalten a), b) zeigen deutlich die Unzulänglichkeit der *Einzelinformationen* über Länge und Wendewinkel, aus denen der Stufenausgleich zusammengesetzt ist.

Das Maß in Spalte c) ist entnommen aus Fejes u. Rosenfeld [7]. Die dort vorgeschlagene Polygonevolution entfernt jedoch keine Punkte, sondern verschiebt jeden Eckpunkt eines Randpolygons, auf den Mittelpunkt seiner Nachbarn. Ein solches Verfahren hat grundlegend verschiedene Eigenschaften zu dem in dieser Arbeit vorgestellten, wie der Vergleich in Kapitel 3.6.2.2 zeigt. Dieses Maß, direkt übernommen in das Evolutionsverfahren durch Bildung der nächsten Abstraktion, liefert deutlich schlechtere Ergebnisse als der Stufenausgleich in Spalte d). Man beachte, daß beim Stufenausgleich trotz erheblicher Reduktion der Punktzahl die Form erkennbar bleibt.

3.6 Eigenschaften des Evolutionsverfahrens

Das Evolutionsverfahren durch iterative Berechnung der nächsten Abstraktion hat einige nennenswerte Eigenschaften, die in diesem Kapitel gezeigt werden sollen. Die ersten drei sind unabhängig von der Wahl der Kostenfunktion, es sind also Eigenschaften, die direkt aus der Teilmengenbildung resultieren:

Das Evolutionsverfahren...

1. ...endet in einem konvexen Polygon
2. ...reduziert schrittweise die Komplexität des gegebenen Polygons
3. ...erhält die Position von wichtigen Randpunkten, der Rand wird nicht verwischt

Die anderen Eigenschaften sind abhängig von der Art der Kostenfunktion, gelten aber zumindest für den oben eingeführte Stufenausgleich K :

Das Evolutionsverfahren mit Stufenausgleich K ...

1. ...ist invariant gegenüber Translation, Rotation, Skalierung
2. ...besitzt eine adaptive Auflösungsanpassung
3. ...ist robust gegenüber Störungen
4. ...führt sowohl zur Störungselimination als auch auch zur Formabstraktion
5. ...führt parameterfrei zu einer universellen Abstraktionsstufe
6. ...extrahiert Liniensegmente

7. ...stimmt überein mit menschlicher Intuition

Die Aufzählung der Eigenschaften soll nun mit Leben gefüllt werden. Die evolutionssimmanenten Eigenschaften 1—3 resultieren aus der einfachen Vorgehensweise, die Evolution über Teilmengen zu bilden, also gewisse Punkte zu entfernen. In dieser Vorgehensweise liegt der große Unterschied zu klassischen Evolutionsverfahren, welche zumeist Randpunkte versetzen, z.B. [5], [10].

3.6.1 Eigenschaften im Zusammenhang mit der Teilmen- genrelation

Die folgenden drei Beweise zeigen die einfache mathematische Handhabbarkeit des Evolutionsverfahrens bedingt durch iterative Bildung der nächsten Abstraktion.

3.6.1.1 Konvexität

Gegeben sei ein Polygon $\mathcal{P} = (p_0, \dots, p_n)$ mit $n + 1 = \#\mathcal{P} \geq 3$

Satz:

Das Evolutionsverfahren $\mathcal{E}(\mathcal{P}, \cdot)$ endet in einem konvexen Polygon \mathcal{P}^m

Beweis:

Der Beweis ist trivial: die Evolutionssequenz $\mathcal{E} = \mathcal{P}^0, \dots, \mathcal{P}^m$ endet mit einem Polygon \mathcal{P}^m mit $\#\mathcal{P}^m = 3 \Rightarrow \mathcal{P}^m$ ist konvex. ■

Man beachte, daß bis auf die Geschlossenheit (die natürlicherweise für Konvexitätsaussagen Voraussetzung ist) keine weiteren Anforderungen an das Polygon gestellt werden, so gilt diese Aussage z.B. auch für selbstüberschneidende Polygone. Die Endstufe \mathcal{P}^m der Evolution ist natürlich i.a. nicht die konvexe *Hülle* des Polygons \mathcal{P} .

Aussagen bezüglich der Konvexität sind bei anderen Verfahren der Scale-Space Theorie keineswegs einfach. Man findet eine tiefgehende Konvexitätsaussage über Evolutionsverfahren mit Diffusionsgleichungen z.B. bei Grayson [10]. Die dortige Aussage gilt nur für einfache, geschlossene und glatte Kurven. Ein polygonales Analogon zur Lösung des Evolutionsproblems mit Diffusionsgleichungen findet sich bei Bruckstein et. al. [5]. Die Frage der Konvexität ist dort noch offen, es werden Experimente angegeben, die auf eine Entwicklung der Evolution auf ein konvexes Polygon hin deuten.

3.6.1.2 Komplexität

Die nächste Aussage ist stark visuell motiviert. Es ist deutlich, daß sich bei höheren Evolutionsstufen die ursprünglich gegebene Form (stark) abstrahiert, intuitiv nimmt ihre *Komplexität* ab. Die Komplexität eines Polygons läßt sich wie folgt modellieren:

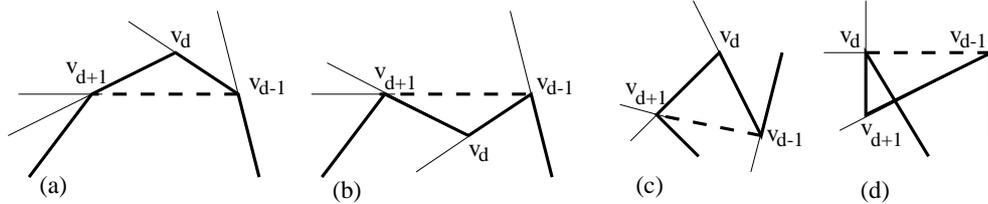


Abbildung 3.9: Entnahme eines Randpunktes: die Komplexität bleibt unverändert (a) bzw. wird reduziert (b,c,d)

Definition: Komplexität eines Polygons

Die *Komplexität* C eines Polygons $\gamma = (p_0, \dots, p_n)$ ist die Summe der absoluten Wendewinkel:

$$C(\gamma) = \sum_{i=0}^n |\text{turn}(p_i)|$$

wobei $\text{turn}(p_i)$ den Wendewinkel im Punkt p_i angibt, siehe Abbildung 3.9. Einfach zu sehen ist, daß die Komplexität eines geschlossenen konvexen Polygons genau 2π beträgt, die Komplexität eines geschlossenen nichtkonvexen Polygons größer als 2π ist.

Mit dieser Definition gilt nun für das Evolutionsverfahren der folgende

Satz:

Die Komplexität $C(\gamma_i)$ eines geschlossenen Polygons nimmt in der Evolutionssequenz $\mathcal{E} = \gamma_0, \dots, \gamma_m$ monoton ab, d.h. $\forall i = 0, \dots, m \Leftrightarrow 1 : C(\gamma_i) \geq C(\gamma_{i+1})$

Beweis:

Der Beweis soll hier nur kurz angedeutet werden. Es seien γ^k und γ^{k+1} zwei aufeinanderfolgende Polygone der Evolutionssequenz $\mathcal{E}(\gamma)$, v_d sei der aus γ^k entfernte Punkt, d.h. $\gamma^k \setminus \gamma^{k+1} = \{v_d\}$, v_{d-1}, v_{d+1} seien die Nachbarpunkte von v_d . Es sei $A \subset \gamma^k$ das viersegmentige Teilpolygon aus γ^k , das die Punkte v_{d-1}, v_d, v_{d+1} enthält. Ist A konvex (konkav), so ist $C(\gamma^k) = C(\gamma^{k+1})$, siehe Abbildung 3.9 (a).

Ist A nicht konvex (konkav), so ist $C(,^k) > C(,^{k+1})$, siehe Abbildung 3.9(b,c,d). Der Beweis kann durch einfache Geometrie erbracht werden. ■

3.6.1.3 Positionstreue

Die folgende Eigenschaft ist besonders interessant im Zusammenhang mit Verfahren, welche auf verschiedenen Stufen $,^k$ der Evolution $\mathcal{E}(,)$ operieren. In solchen Verfahren muß oftmals ein Bezug zwischen diesen Stufen hergestellt werden, so können z.B. grobe Formeigenheiten in hoher Abstraktionsstufe und deren Untergliederung in weniger stark evolvierten Stufen ermittelt werden. Dazu müssen Formteile der Stufe $,^k$ einen Verweis auf ihre Lage in der Stufe $,^l$, $l > k$ besitzen. Die diskrete Polygonevolution liefert diesen Zusammenhang automatisch, da die Stufen $,^k$ in Teilmengenrelation $,^0 \supset ,^1 \supset \dots \supset ,^m$ stehen, d.h. Randpunkte hoher Abstraktionsstufe $,^l$ sind in niedrigerer Abstraktionsstufe $,^k$, $l > k$ enthalten und damit leicht aufzufinden, siehe dazu Abbildung 3.10

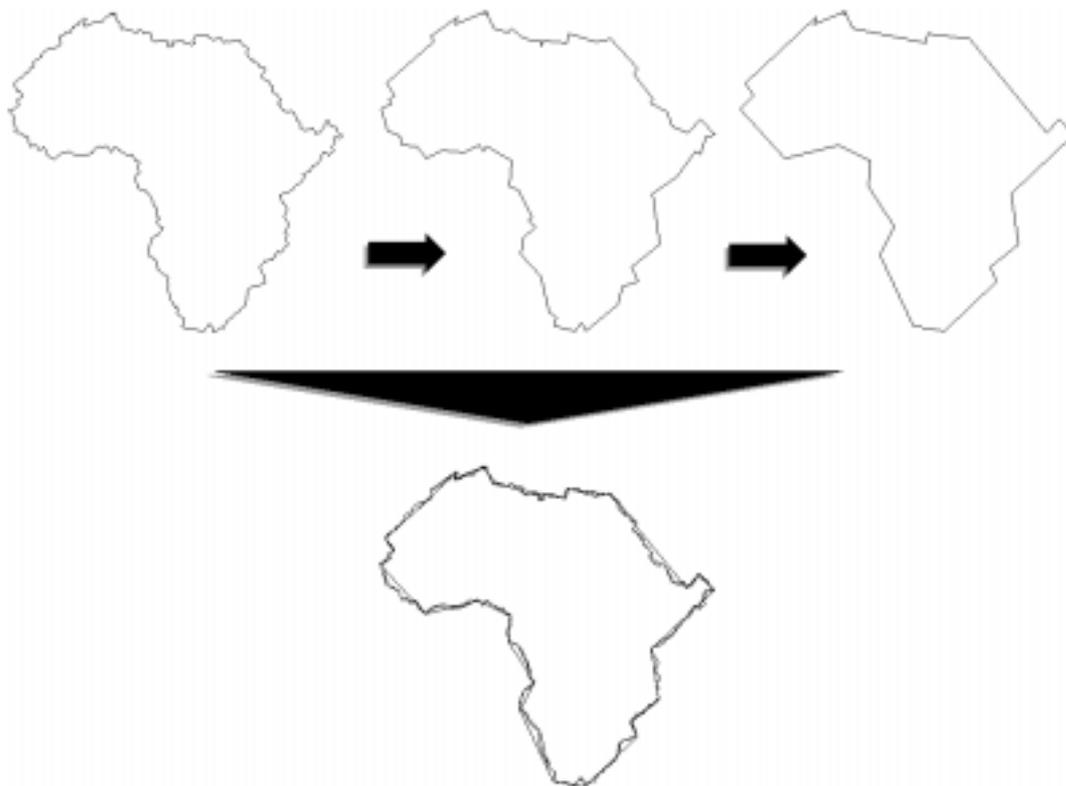


Abbildung 3.10: Verschiedene Evolutionsstufen und ihre Überlagerung: die Eckpunkte behalten ihre Position

Satz:

Die diskrete Polygonevolution $\mathcal{E}(\cdot, \cdot)$ erhält die Position wichtiger Randpunkte von \mathcal{P} ,

Beweis:

Der Beweis ist trivial und folgt direkt aus dem Algorithmus: das Polygon \mathcal{P} wird nur um einige Eckpunkte reduziert, die übrigen bleiben unverändert. Die 'Wichtigkeit' der Punkte bezieht sich auf die Kostenfunktion K , die Aussage ist also per Definition richtig. ■

Der Erhalt der Position ist ein wichtiger Vorteil der Polygonevolution gegenüber klassischen Evolutionsverfahren durch Diffusionsgleichungen. Kapitel 5 wird hierauf näher eingehen.

3.6.2 Eigenschaften im Zusammenhang mit dem Stufen- ausgleich

Wie schon in Kapitel 3.4 erwähnt, ist die Kostenfunktion K_Γ der Kern des Evolutionsverfahrens, ihre Wahl ist keineswegs unkritisch, wie der Vergleich mit anderen möglichen Kostenfunktionen in Kapitel 3.5 gezeigt hat. Die im folgenden aufgezeigten Eigenschaften des Evolutionsverfahrens stehen in engem Zusammenhang mit dem Stufenausgleich K und können als Argumentation für die Wahl eben dieser Kostenfunktion angesehen werden. Falls nicht explizit erwähnt, beziehen sich die folgenden Eigenschaften, Sätze und Beweise auf den Stufenausgleich.

3.6.2.1 Invarianz

Satz:

Das Evolutionsverfahren $\mathcal{E}(\cdot, \cdot)$ mit Stufenausgleich K ist invariant gegenüber Translation, Rotation und Skalierung

Beweis:

- Translation: die Transformation von \mathcal{P} in den Tangentenraum ist translationsinvariant. Der Stufenausgleich K berechnet sich im Tangentenraum \Rightarrow die Aussage ist richtig.
- Rotation: die Rotation des Polygons \mathcal{P} entspricht einer Translation der Stufenfunktion in Richtung der y -Achse im Tangentenraum. Der Stufenausgleich benutzt als Winkelinformation nur y -Differenzen im Tangentenraum,

diese sind invariant gegenüber der Translation in y -Richtung \Rightarrow die Aussage ist richtig.

- Skalierung: zur Berechnung des Stufenausgleichs wird die Länge des Randpolygons , normiert \Rightarrow die Aussage ist richtig

■

Bemerkungen:

- Diese Eigenschaft gilt für alle Kostenfunktionen, welche keine weitergehenden Lageinformationen benutzen. Die genaue Berechnung, z.B. die Bildung des hyperbolischen Durchschnitts, spielt hier keine Rolle.
- Die Aussage über Rotationsinvarianz und Skalierungsinvarianz ist mit Vorsicht zu genießen: der Stufenausgleich bezieht sich auf eine Menge an Eckpunkten in \mathbb{Z}^2 . Unter Rotationen um beliebige Winkel und Skalierungen um beliebige Faktoren versteht man die Einbettung des Polygons in \mathbb{R}^2 , die gängige Rotations- oder Skalierungsoperation, gefolgt von einer Digitalisierung \mathcal{D} , also einer Rücktransformation von \mathbb{R}^2 nach \mathbb{Z}^2 . Diese Digitalisierung \mathcal{D} ist nicht formerhaltend, d.h. für ein Polygon $P \subset \mathbb{R}^2$ gilt i.a. $P \setminus \mathcal{D}(P) \neq \emptyset$. Die Veränderung der Forminformation kann zu einer veränderten Gewichtung der Kosten der Eckpunkte führen. Dieses ist jedoch in der Praxis nicht relevant. Die obige Aussage ist also praktisch, nicht mathematisch korrekt zu verstehen.

3.6.2.2 Adaptivität

Die im folgenden beschriebene Eigenschaft resultiert aus dem Zusammenspiel des Evolutionsalgorithmus' mit dem Stufenausgleich K . Es ist zunächst zu beachten, daß die Kosten $K_\Gamma(p)$ im Punkt bezüglich der Darstellung des Polygons , durch seine Eckpunkte p_0, \dots, p_n lokal berechnet werden, diese lokale Betrachtung im Kantenzug $\bar{} \subset \mathbb{R}^2$ – also der in \mathbb{R}^2 eingebetteten Menge der Verbindungssegmente zwischen den Punkten $p_i \in \bar{}$, – jedoch keine lokale Entsprechung hat. Die Berechnung der Kosten $K_\Gamma(p_i)$ in $\bar{}$ bezieht sich auf ein Gebiet $\mathcal{G} \subset \bar{}$, $\mathcal{G} = \overline{p_{i-1}p_i} \cup \overline{p_{i-1}p_i}$, welches ein zweisegmentiges Teilpolygon um p_i darstellt. Dieses Zusammenspiel zwischen lokaler und gebietsweiser Betrachtung sowie die Auswirkung auf die Polygonevolution \mathcal{E} soll nun näher beleuchtet werden.

Definition: Umgebung

Gegeben sei eine Menge \mathcal{M} . Unter einer *Umgebung* \mathcal{N} versteht man eine Abbildung $\mathcal{N} : \mathcal{M} \rightarrow \mathcal{P}(\mathcal{M})$, $p \rightarrow \mathcal{N}(p)$ falls gilt: $\forall p \in \mathcal{M} \ p \in \mathcal{N}(p), \mathcal{N}(p) \setminus p \neq \emptyset$

Definition: Auflösung einer Menge

Gegeben sei eine Menge \mathcal{M} mit Metrik d und einer Umgebung \mathcal{N} . Unter der Auflösung $R_{\mathcal{N}}(\mathcal{M})$ von \mathcal{M} versteht man den minimalen Abstand eines Punktes $p \in \mathcal{M}$ zu Punkten seiner Umgebung:

$$R_{\mathcal{N}}(\mathcal{M}) = \min_{p \in \mathcal{M}} \{d(p, q) \mid q \in \mathcal{N}(p), q \neq p\}$$

Eine Menge \mathcal{M} heißt *feiner* aufgelöst als eine Menge \mathcal{M}' , falls

$$R_{\mathcal{N}}(\mathcal{M}) < R_{\mathcal{N}'}(\mathcal{M}')$$

Als Beispiel sei die Auflösung eines geschlossenen Polygons $\mathcal{P} = p_0, \dots, p_n = p_0$ mit (kanonischer) Nachbarschaftsabbildung

$$\mathcal{N} : \mathcal{P} \rightarrow \mathcal{P}(\mathcal{P}), \mathcal{N}(p_i) = \{p_{i-1}, p_i, p_{i+1}\}$$

und euklidischer Metrik d_e genannt:

$$R_{\mathcal{N}}(\mathcal{P}) = \min_i \{d_e(p_i, p_{i+1})\}$$

Die Auflösung eines Randpolygons ist entscheidend für die Darstellung von Feinheiten und für die Erkennbarkeit der Form. Ist sie zu gering, können wichtige formgebende Randpunkte nicht erreicht und abgebildet werden. Andererseits ist auch eine hohe Auflösung nicht immer von Vorteil: durch die größere Informationsdichte erhöht sich der Bearbeitungsaufwand, zusätzlich – und dieses ist der entscheidende Nachteil zu hoher Auflösung – kann für eine ganze Klasse von Verfahren mit dem Gewinn lokaler Informationen ein Verlust globaler Informationen einhergehen. Genauer:

Definition: Umgebungsvorschrift

Eine Abbildung $\mathcal{U}(\mathcal{M})$, die einer Menge \mathcal{M} die Folgen $M = \mathcal{M}^0, \mathcal{M}^1, \dots, \mathcal{M}^n$ von nichtleeren Mengen sowie eine Folge $N = \mathcal{N}^0, \mathcal{N}^1, \dots, \mathcal{N}^n$ von Umgebungen zuweist, heißt *Umgebungsvorschrift*.

Definition: Auflösungsabhängigkeit

Gegeben sei eine Menge \mathcal{M} sowie die durch eine Umgebungsvorschrift $\mathcal{U}(\mathcal{M})$ definierten Folgen M und N . N heißt *auflösungsabhängig*, falls gilt: $\exists p \in \mathcal{M}^j, j \in 1..k : \exists i < j : \mathcal{N}^i(p) \neq \mathcal{N}^j(p)$

Die Umgebungsvorschrift $\mathcal{U}(\mathcal{M})$ heißt auflösungsabhängig, falls N auflösungsabhängig ist.

Definition: Lokale Information

Gegeben sei eine Menge \mathcal{M} mit Umgebung \mathcal{N} . Eine *lokale Information* $K_{\mathcal{M}}$ ist eine Abbildung $K_{\mathcal{M}} : \mathcal{M} \times \mathcal{N}(\mathcal{M}) \rightarrow \mathbb{R}_0^+$, $(p, \mathcal{N}(p)) \rightarrow k$

Es sei nun \mathcal{M} eine Menge mit Metrik d und Umgebungsvorschrift $\mathcal{U}(\mathcal{M})$, $(\mathcal{M}^i, \mathcal{N}^i), (\mathcal{M}^j, \mathcal{N}^j)$ $i < j$ seien zwei Elemente aus $\mathcal{U}(\mathcal{M})$. \mathcal{M}^i sei feiner aufgelöst als \mathcal{M}^j .

Es seien K^i, K^j lokale Informationen auf $\mathcal{M}^i, \mathcal{M}^j$. Ist $\mathcal{U}(\mathcal{M})$ auflösungsabhängig, so gilt i.a. für einen Punkt $p \in \mathcal{M}^j$: $K^i(p, \mathcal{N}^i(p)) \neq K^j(p, \mathcal{N}^j(p))$, insbesondere kann also auch für einen Punkt $p \in \mathcal{M}^j$ gelten: $\mathcal{R}_{\mathcal{N}^i}(\mathcal{M}^i) < \mathcal{R}_{\mathcal{N}^j}(\mathcal{M}^j)$ und $K^i(p, \mathcal{N}^i(p)) < K^j(p, \mathcal{N}^j(p))$, d.h bei höherer Auflösung ist die lokale Information in p geringer.

Als Beispiel sei eine polygonale Randkurve verschiedener Abstraktionsstufen gegeben, siehe Abbildung 3.11

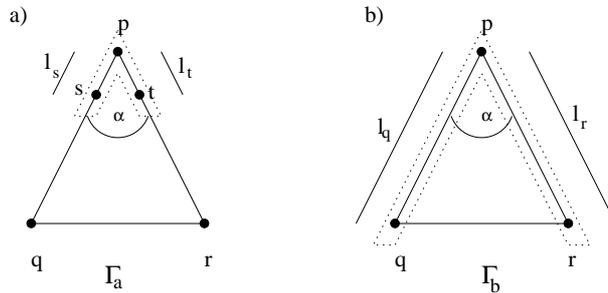


Abbildung 3.11: Visuell identisches Polygon, dargestellt durch Mengen von Eckpunkten verschiedener Auflösung

Die Abbildung zeigt zwei visuell identische Polygone, d.h. als Kurvenzüge in \mathbb{R}^2 aufgefaßt sind (a) und (b) identische Punktmengen, aufgefaßt als Menge von Eckpunkten jedoch ist die linke Figur eine feiner aufgelöste Obermenge der rechten.

Die Mengen $\mathcal{M}^0, \mathcal{M}^1$ sind gegeben durch

$$\mathcal{M}^0 = \{p, s, q, r, t\}, \quad \mathcal{M}^1 = \{p, q, r\}$$

die Umgebungen $\mathcal{N}^0, \mathcal{N}^1$ sind kanonisch gegeben. Als lokale Information soll der Stufenausgleich K angenommen werden. Die Umgebungsvorschrift ist somit:

$$\mathcal{U}(\mathcal{M} = \mathcal{M}^0) = ((\mathcal{M}^0, \mathcal{M}^1), (\mathcal{N}^0, \mathcal{N}^1))$$

Die gestrichelte Umrandung in Abbildung 3.11 zeigt die Nachbarschaft \mathcal{N} des Punktes p , in Abbildung (a) ist $\mathcal{N}_a(p) = \{p, s, t\}$, in Abbildung (b) ist $\mathcal{N}_b(p) = \{p, q, r\}$, d.h. $\mathcal{U}(\mathcal{M})$ ist auflösungsabhängig.

Die Information $K(p)$ des Punktes p , ist in der feiner aufgelösten Figur geringer:

$$l_s < l_q \wedge l_t < l_r \Rightarrow K_{\Gamma_a}(p) = \alpha \frac{l_s l_t}{l_s + l_t} < \alpha \frac{l_q l_r}{l_q + l_r} = K_{\Gamma_b}(p)$$

Man beachte, daß die *visuelle* Relevanz von p , welche durch den Stufenausgleich K modelliert werden soll, dagegen in beiden Formen gleich ist, und sich die Figuren nur um Punkte s, t mit Informationsgehalt $K(s) = K(t) = 0$ unterscheiden, also nur um (bezogen auf K) *nicht relevante* Punkte.

Derartige Beispiele finden sich auch bei Verfahren der klassischen Bildverarbeitung, insbesondere auch bei solchen, die in der Praxis weit verbreitet sind. Es gehören dazu lineare Filteroperationen auf digitalen Grauwertbildern, die auf einer z.B. durch $\mathcal{N}_{4,8}$ Nachbarschaftsrelation festgelegten und damit auflösungsabhängigen Umgebung operieren.

Ein Beispiel sei hier (eindimensional) anhand des diskreten *Laplace-Operators* \mathcal{L} gezeigt (siehe Abbildung 3.12):

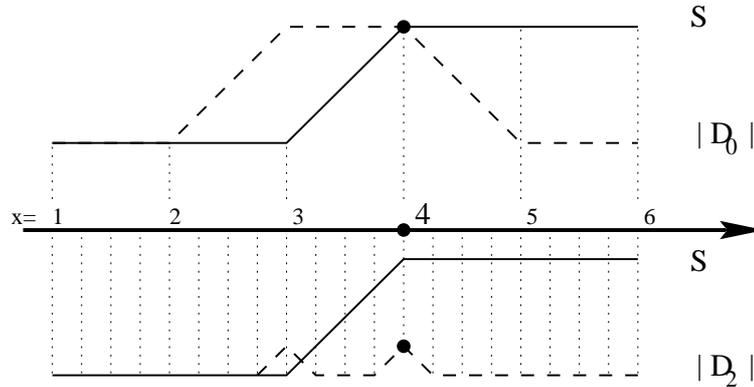


Abbildung 3.12: Informationsverlust durch feinere Auflösung beim Laplace-Operator

Es sei $S : \mathbb{R} \rightarrow \mathbb{R}$ ein eindimensionales kontinuierliches Signal:

$$x < 0 \Rightarrow S(x) = 0, \quad 0 \leq x < 4 \Rightarrow S(x) = x, \quad x \geq 4 \Rightarrow S(x) = 4$$

S werde mit verschiedenen Auflösungen $a_i \in \mathbb{R}$, $a_i = \frac{1}{2^i}$, $i \in 1..r_{max}$ in diskrete Signale D_i gewandelt:

$$D_i : \mathbb{Z} \rightarrow \mathbb{R} \times \mathbb{R}, n \rightarrow (na_i, S(na_i)) =: (d_i^x(n), d_i^y(n))$$

Der diskrete Laplace-Operator \mathcal{L} auf den Signalen D_i ist definiert durch:

$$\mathcal{L} : D_i \rightarrow \mathbb{R} \times \mathbb{R}$$

$$(d_i^x(n), d_i^y(n)) \rightarrow (d_i^x(n), \Leftrightarrow d_i^y(n \Leftrightarrow 1) + 2d_i^y(n) \Leftrightarrow d_i^y(n+1)) =: (l_i^x(n), l_i^y(n))$$

Es sei nun als Informationsgehalt des Punktes $(d_i^x(n), d_i^y(n))$ der durch \mathcal{L} gegebene Wert $|l_i^y(n)|$ angenommen.

Die Umgebungsvorschrift $\mathcal{U}(D = D_{r_{max}})$ ordne der Folge von Mengen

$$D_{r_{max}} \supset D_{r_{max}-1} \supset \dots \supset D_0$$

die (kanonischen) Umgebungen

$$\mathcal{N}_i : \mathcal{N}_i(D_i(n)) = \{D_i(n \Leftrightarrow 1), D_i(n), D_i(n+1)\}$$

zu, $\mathcal{U}(D)$ ist damit auflösungsabhängig.

Für den Informationsgehalt der digitalen Versionen des Punktes $S(4)$ gilt in den verschiedenen Auflösungsstufen $D_{i,j}$:

$$i > j \Rightarrow l_i^y\left(\frac{4}{a_i}\right) < l_j^y\left(\frac{4}{a_j}\right)$$

d.h. bei feinerer Auflösung wird auch hier der Informationsgehalt geringer.

Bemerkung: Es gilt in beiden obigen Beispielen, daß sich die Darstellungen der Formen in verschiedenen Auflösungen nur durch Punkte p_i mit Informationsgehalt $K(p_i) = 0$ unterscheiden, d.h. der Informationsgehalt der in den Beispielen betrachteten Punkte wird bei feinerer Auflösung geringer, obwohl keine (bezogen auf K) relevanten Punkte eingefügt wurden. Auch gilt: durch die feinere Auflösung wurde nicht nur die Information in einzelnen Punkten, sondern auch die Gesamtinformation $\sum K$ reduziert.

Auflösungsabhängigkeit bedeutet, daß die Ergebnisse abhängig sind von der Feinheit der Diskretisierung, also in der Praxis der Scanner- oder Kameraauflösung.

Auflösungsabhängige Verfahren sind damit bei gleicher ursprünglicher Bildinformation abhängig von der Repräsentation dieser Information und können damit nicht in dem Sinne stetig sein, daß ähnliche Bilder zu ähnlichen Ergebnissen führen (da schon identische Bilder, verschieden repräsentiert, unterschiedliche Ergebnisse liefern).

Man kauft sich zumeist durch einen zusätzlich eingeführten Parameter von diesem Nachteil frei: die Bestimmung der Umgebung \mathcal{N} wird abhängig von der Auflösung gewählt, d.h. die Auflösung geht als *metrischer* Parameter in das Verfahren ein (die obigen Umgebungen \mathcal{N} waren *topologisch* definiert).

Die Bestimmung dieses Parameters ist ein Grundproblem lokaler Verfahren, d.h. in der Praxis wird der Parameter entweder als konstant vorausgesetzt (=die Auflösung der Kamera und die Skalierung der Formen werden als konstant vorausgesetzt) oder durch optische Kontrolle des Anwenders justiert.

Die Unabhängigkeit von der Auflösung erreicht die Polygonevolution über einen anderen, parameterfreien Weg: die Auflösung von Teilpolygonen wird auf die minimal notwendige Auflösung reduziert, indem Punkte ohne Informationsgehalt entfernt werden. Die Auflösung paßt sich also der Form an. Die kanonische Topologie erzeugt automatisch Umgebungen, welche nur Punkte mit relevanter Information enthalten. Durch diese *Adaption* wird das Gesamtverfahren unabhängig von der Auflösung der Anfangsform, obgleich die benutzte lokale Informationsfunktion, z.B. der Stufenausgleich, auflösungsabhängig ist.

Dieses kann wie folgt formuliert werden:

Definition: Auflösungsadaption

Gegeben sei ein Polygon $\mathcal{P} = p_0, \dots, p_n \subset \mathbb{R}^2$ durch seine Eckpunkte, sowie eine zugehörige lokale Information K_{Γ} . \mathcal{P} enthalte keine irrelevanten Punkte, d.h.

$\forall p \in \mathcal{P} : K_{\Gamma}(p) > 0$. $\mathcal{P}' \supset \mathcal{P}$, sei ein Polygon mit zugehöriger lokaler Information $K_{\Gamma'}$ derart, daß

$\mathcal{P}' = \mathcal{P} \cup \{p_{i=0..k}\}, p_i \in \mathbb{R}^2, K_{\Gamma'}(p_i) = 0$.

Das Evolutionsverfahren $\mathcal{E} = (\mathcal{P}' = \mathcal{P}'_0, \dots, \mathcal{P}'_m)$ heißt auflösungsadaptiv, falls eine Umgebungsvorschrift $\mathcal{U}(\mathcal{P}, \mathcal{P}')$ derart existiert, daß die Mengenfolge $\mathcal{P}'_0, \dots, \mathcal{P}'_n$ der Evolution $\mathcal{E}(\mathcal{P}, \mathcal{P}')$ entspricht und in \mathcal{E} ein Element $\mathcal{P}'_i = \mathcal{P}$ existiert.

Satz:

Das Evolutionsverfahren \mathcal{E} mit Stufenausgleich K ist auflösungsadaptiv

Beweis:

Es sei $\Delta = \mathcal{P}' \setminus \mathcal{P}$, d.h. $\Delta = \{p \in \mathcal{P}' | K_{\Gamma'}(p) = 0\}$. Zu zeigen ist:

1) $\forall p \in \Gamma, : K_{\Gamma'}(p) > 0$

2) $\forall i \in 0..m : p \in \Gamma'_i \cap \Delta \Rightarrow K_{\Gamma'_i}(p) = 0$ sowie $p \in \Gamma'_i \setminus \Delta \Rightarrow K_{\Gamma'_i}(p) > 0$.

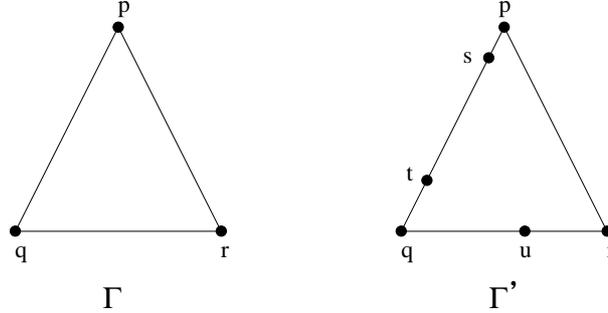


Abbildung 3.13: Das Evolutionsverfahren mit Stufenausgleich K ist auflösungsadaptiv

Für den Stufenausgleich gilt mit $p_j \in \Gamma'_i : K_{\Gamma'_i}(p_j) = 0 \Leftrightarrow p_{j-1}, p_j, p_{j+1}$ sind kollinear.

Zu 1):

Es gilt: $\forall p_i \in \Gamma, : p_{i-1}, p_i, p_{i+1}$ nicht kollinear.

$\forall p_j \in \Gamma' \setminus \Delta, : K_{\Gamma'} = 0 \Rightarrow \exists p_i, p_{i+1} \in \Gamma, : p_j \in \overline{p_i p_{i+1}} \Rightarrow \forall, \ni p_i = p'_k \in \Gamma' : p_{i-1}, p_i, p_{i+1} \text{ kollinear} \Leftrightarrow p'_{k-1}, p'_k, p'_{k+1} \text{ kollinear}$
 \Rightarrow Aussage 1

Zu 2):

Es sei $\Gamma'_i \setminus \Gamma_{i+1} = \{p_{i+1}\} \subset \Delta$ sowie $K_{\Gamma'_i}(p_i) = 0$.

$\Rightarrow p_{i-1}, p_i, p_{i+1}, p_{i+2}$ kollinear

$\Rightarrow p_{i-1}, p_i, p_{i+2}$ kollinear

$\Rightarrow K_{\Gamma'_{i+1}}(p_i) = 0$

Es sei $\Gamma'_i \setminus \Gamma_{i+1} = \{p_{i+1}\} \subset \Delta$ sowie $K_{\Gamma'_i}(p_i) > 0$.

$\Rightarrow p_{i-1}, p_i, p_{i+1}$ nicht kollinear, p_i, p_{i+1}, p_{i+2} kollinear,

$\Rightarrow p_{i-1}, p_i, p_{i+2}$ nicht kollinear

$\Rightarrow K_{\Gamma'_{i+1}}(p_i) > 0$

\Rightarrow Aussage 2. ■

Die Adaption, erzwungen durch die Entfernung auflösungsbedingter Punkte irrelevanter Information, ist der entscheidende Punkt im Zusammenspiel der lokalen, auflösungsabhängigen Informationsfunktion mit dem Gesamtverfahren.

Als Gegenbeispiel zu der adaptiven Polygonevolution durch Stufenausgleich sei die in [7] von Fejes u. Rosenfeld beschriebene Polygonevolution genannt, welche in den einzelnen Stufen keine Punkte *entfernt*, sondern verschiebt:

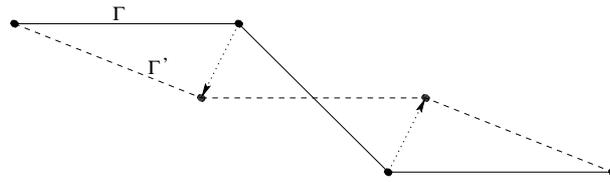


Abbildung 3.14: Polygonevolution nach Fejes und Rosenfeld [7]

Das Polygon $\Gamma^{i+1} = (p'_0, p'_1, \dots, p'_n = p'_0)$ entsteht aus dem Polygon $\Gamma^i = (p_0, p_1, \dots, p_n = p_0)$ durch die Vorschrift (siehe Abbildung 3.14):

$$p'_i = \frac{p_{i-1} + p_{i+1}}{2}$$

Die Information $K_{\Gamma}^f(p_i)$ ist dabei gegeben durch den Betrag der Verschiebung:

$$K_{\Gamma}^f(p_i) = |p_i \Leftrightarrow p'_i|$$

Klar: K_{Γ}^f ist auflösungsabhängig. Diese Abhängigkeit bleibt im Gesamtverfahren erhalten (dabei sei die Auflösungsabhängigkeit der Evolution nicht formal definiert, sondern intuitiv verstanden: die Bezeichnung gelte, falls die Ergebnisse der Evolution zweier Mengen, die sich nur um Punkte mit Informationsgehalt 0 unterscheiden, 'erheblich differieren'). Dieses zeigt Abbildung 3.15:

Die linke und rechte Spalte zeigen Evolutionsen eines als Kantenzug aufgefaßten (und visuell) identischen Polygons, als Punktmenge Γ , Γ' aufgefaßt gilt: $\Gamma \subset \Gamma'$, Γ' feiner aufgelöst als Γ , und $\forall p \in \Gamma' \setminus \Gamma : K_{\Gamma'}^f(p) = 0$. Die Evolution von Γ ergibt ein (bis auf Skalierung) in zwei Stufen oszillierendes System, die Evolution von Γ' dagegen zeigt eine Entwicklung zum Kreis: die Evolutionsen 'differieren erheblich'.

Bemerkungen:

- Zu beachten ist, daß für ein auflösungsadaptives Verfahren nur gefordert ist, daß Punkte p eines Polygons $\Gamma^0 = \Gamma^0$, die auch in einer späteren Evolutionsstufe Γ^i enthalten sind, in allen vorhergehenden Stufen Γ^j , $0 \leq j < i$ eine lokale Information $K_{\Gamma^j}(p)$ besitzen, die größer als 0 ist. Nicht gefordert dagegen ist, daß die Informationsordnung beibehalten werden muß, also mit $j < i$ und $p, q \in \Gamma^i$ gelten muß: $K_{\Gamma^j}(p) < K_{\Gamma^j}(q) \Rightarrow K_{\Gamma^i}(p) < K_{\Gamma^i}(q)$.

Eine solche Forderung würde nicht zu einer *iterativen* Evolution, sondern zu einem *rein lokalen* Verfahren führen, welches die Evolutionsschritte aus

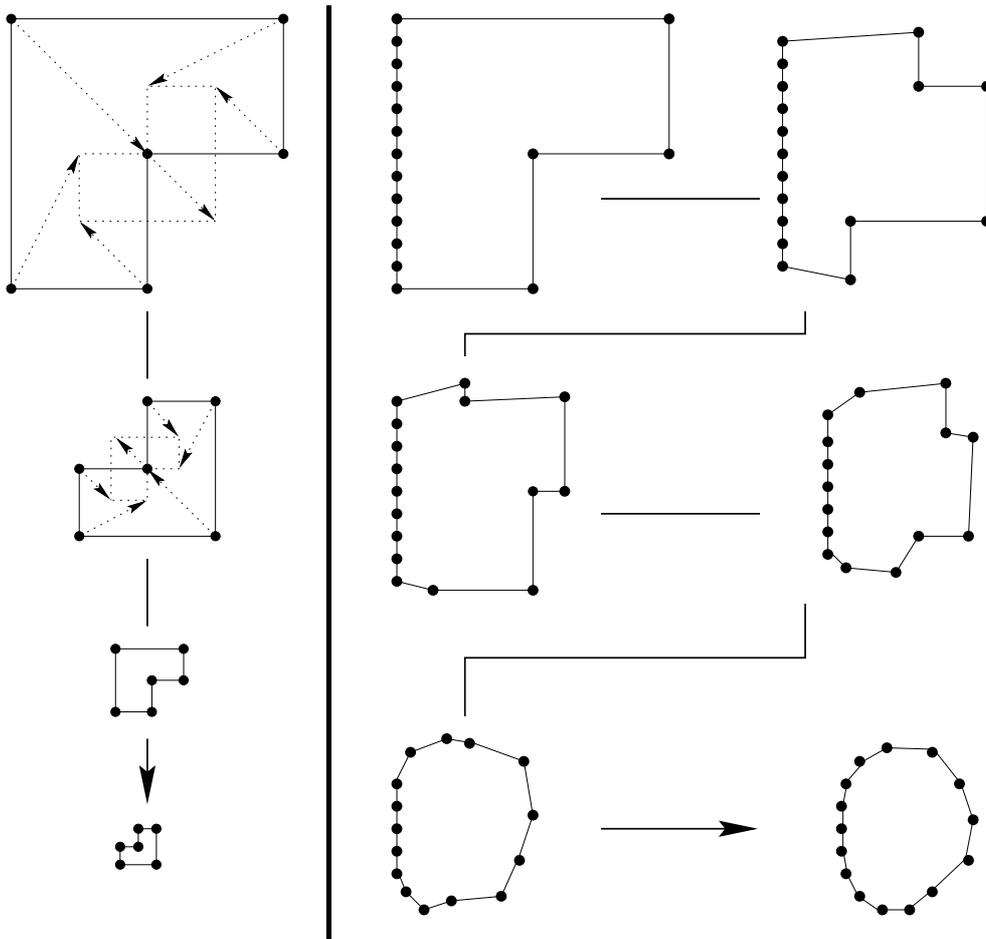


Abbildung 3.15: Polygonevolution nach Fejes und Rosenfeld [7]

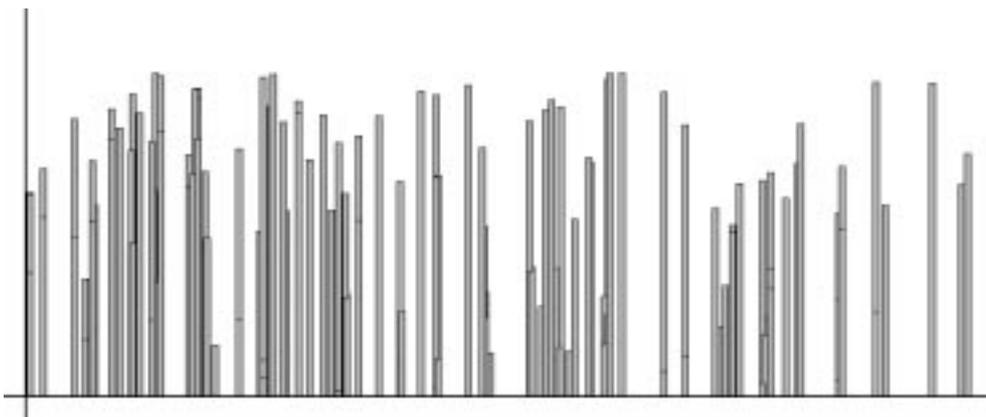


Abbildung 3.16: Informationsgehalt von Punkten eines Polygons (x -Achse) und ihre Lebensdauer (y -Achse)

lokalen Informationen des Ausgangspolygons gewinnt. Ein Beispiel dazu zeigte die 'falsche' Evolution in Abbildung 3.4. Die Polygonevolution mit Stufenausgleich K ist dagegen *kein lokales Verfahren*, obgleich die Kostenberechnung K eine *lokale* Information ist. Abbildung 3.16 zeigt eine Aufstellung des Informationsgehaltes der Eckpunkte in der Form ρ , aus Abbildung 3.17 links gegenüber ihrer 'Lebensdauer' in der Evolution mit Stufenausgleich K_Γ . Die Lebensdauer eines Punktes p ist dabei der maximale Index von Evolutionselementen $i \in \mathcal{E}$, in welchen der Punkt enthalten ist. Deutlich ist, daß die Verteilung von Punkten hoher und niedriger Lebensdauer keine Abhängigkeit von der lokalen Relevanz in der Ausgangsform zeigt, d.h. entscheidend für den Gesamtalgorithmus ist weder die Teilmengenbildung durch Berechnung der nächsten Abstraktion, noch der Stufenausgleich: entscheidend ist das *Zusammenspiel* zwischen Stufenausgleich und dem Evolutionsalgorithmus.

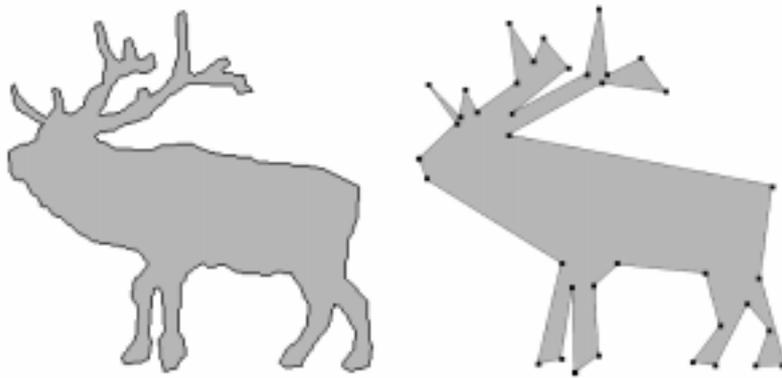


Abbildung 3.17: Adaptive Auflösungsanpassung

- Die Definition der Auflösung macht nur intuitiven Sinn bei der Betrachtung von *Teilpolygone* einer gegebenen Form ρ , man betrachte dazu Abbildung 3.17: die linke Form zeigt die Originalrandkurve, jedes Pixel ist hier ein Eckpunkt des Randpolygons, d.h. bei Betrachtung von Teilpolygone zeigt sich eine nahezu einheitliche Auflösung. Die rechte Seite zeigt eine Abstraktionsstufe der Evolution mit Stufenausgleich K . Nur die markierten Punkte sind Eckpunkte des Randpolygons. Bei Betrachtung von Teilpolygone zeigt sich, daß in Gebieten mit hoher Dichte an visuell relevanten Eckpunkten (hier dem Geweih) die Auflösung der Teilpolygone feiner ist als die Auflösung von Teilpolygone, die Gebiete mit relativ wenigen relevanten Punkten beschreiben (z.B. dem Rücken).

Die visuell irrelevanten Randpunkte p_i der linken Form besitzen nicht notwendigerweise eine lokale Information vom Wert $K_\Gamma(p_i) = 0$, wie in der Definition der Auflösungsadaption gefordert. Jedoch können sie als *gestörte* Version von Randpunkten mit Information $K_\Gamma(p_i) = 0$ angesehen werden. Die Evolution geht hier also weit über die Adaptionseigenschaft hinaus: sie bemerkt und entfernt Störungen.

3.6.2.3 Robustheit

Der Ausgangspunkt des vorherigen Kapitels war eine ideale Bildvorlage \mathcal{I} , also ein Randpolygon, nur bestehend aus visuell relevanten Punkten (mit Informationsgehalt > 0). Diese Randkurve wurde durch Einfügen visuell irrelevanter Teile (Informationsgehalt $= 0$) gestört. Die Evolution \mathcal{E} , angewendet auf die gestörte Version \mathcal{I}' , extrahierte das ursprüngliche Bild. Diese Überlegung ist theoretischer Natur, sie vertieft die Veranschaulichung und ist notwendig für das Verständnis der Vorgehensweise des Evolutionsverfahrens.

Für die Praxis aber ist die Erkenntnis von wenig Wert: visuell irrelevante Punkte können mit Einschrittverfahren erkannt und entfernt werden, ein iteratives Verfahren ist nicht notwendig. Zusätzlich bietet die Praxis keine Entsprechung: wie in der Einleitung zu diesem Kapitel erwähnt und in Abbildung 3.1 sowie Abbildung 3.17 gezeigt, sind die Bildstörungen in der Praxis nicht vom Informationsgehalt 0, d.h. für den Stufenausgleich K gilt meist, daß Störungspunkte echte Wendewinkel $\omega \neq 0$ besitzen. Insbesondere zeigt die Abbildung 3.16, daß visuell irrelevante Punkte im gestörten Ausgangspolygon \mathcal{I}' der Evolution \mathcal{E} einen *höheren* Informationsgehalt besitzen können als Punkte, welche in der Evolution erhalten bleiben und als visuell relevant angesehen werden müssen. In einem solchen Fall ist klar, daß das ursprüngliche Bild nicht *exakt* genähert werden kann, Abbildung 3.18 zeigt dazu ein Beispiel.

Zahlreiche Experimente zeigen, daß dieses Verhalten in der Praxis keine negative Auswirkung auf den visuellen Eindruck hat. In den Abbildungen 3.19 und 3.20 sind dazu zwei Beispiele dargestellt.

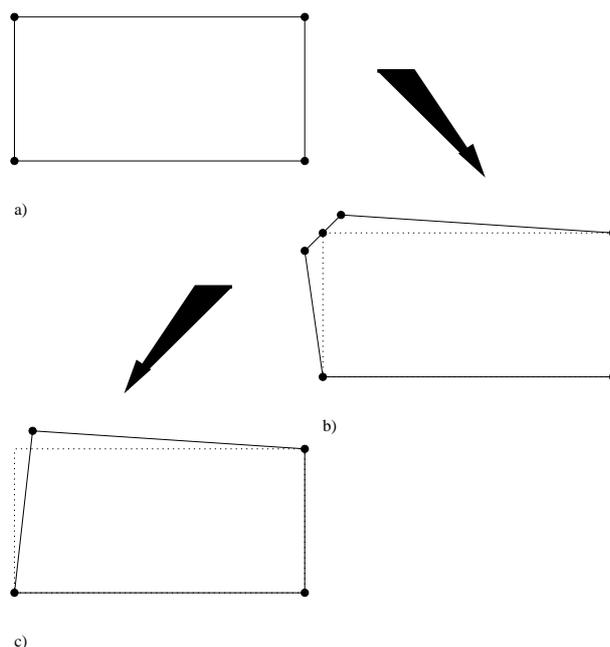


Abbildung 3.18: Beispiel für eine Störung, in welcher die ursprüngliche Form nicht exakt wiederhergestellt werden kann: a) Ausgangsbild b) gestörtes Bild c) Rekonstruktion

Man beachte, daß die Störungen zum Teil erheblich sind (3.19), und daß sowohl global verteilte Störungen (3.20, Spalte b) als auch gebietsbeschränkte Störungen (3.20, Spalte c) erkannt werden.

Deutlich ist, daß die grobe Form erhalten und erkennbar bleibt, d.h. der Hausdorffabstand der Evolutionsergebnisse der verschiedenen gestörten Versionen (intuitiv) gering ist. *Ein Beweis dazu steht jedoch noch aus*.

Ein Beweis für den Fall der *exakten* Rekonstruktion wird jedoch in Kapitel 3.7 gegeben. Dort wird gezeigt, daß das Evolutionsverfahren *stetig* ist, d.h. daß sich für jede gegebene Form eine Klasse von gestörten Versionen findet, deren Evolution zur ursprünglichen Form zurückführt.

3.6.2.4 Störungselimination und Formabstraktion

Eine weitere Eigenschaft, die im Zusammenhang mit der Störungselimination steht, ist die intuitiv richtige *Reihenfolge* der Entfernung von Randpunkten. Die Evolution umfaßt dabei zwei Stufen:

1. Elimination leichter Störungen, z.B. Digitalisierungsrauschen

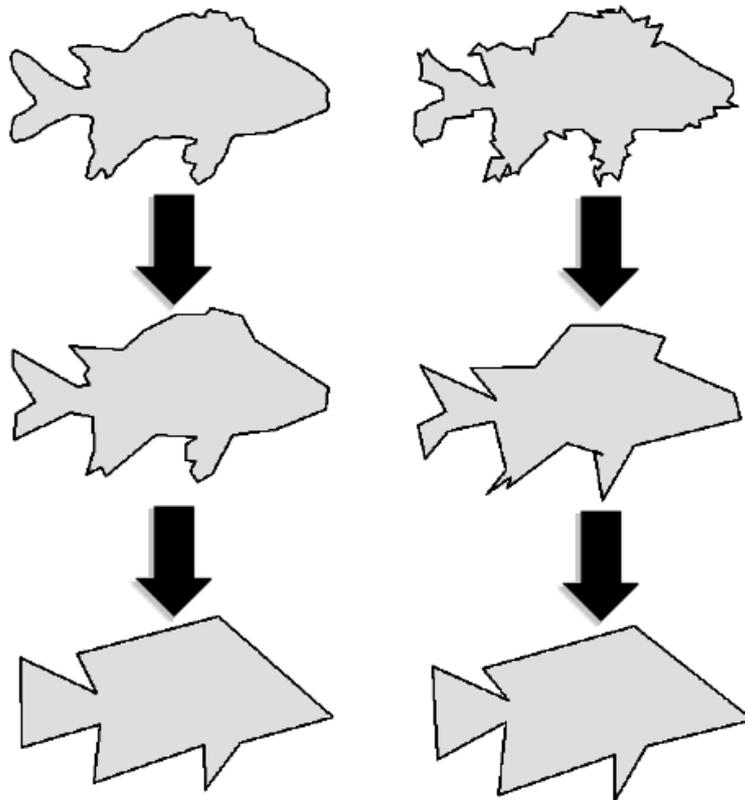


Abbildung 3.19: Drei Evolutionstufen der Vorlage 'Fisch' (linke Spalte), Evolution der gestörten Version (rechte Spalte).

2. Formabstraktion

Wichtig ist diese Eigenschaft in der Praxis: das Verfahren kann als *Rauschfilter* eingesetzt werden, indem die Evolution nach geeigneter Schrittzahl i_{noise} abgebrochen wird. Das bedeutet, daß die Stufen $, 0, \dots, i_{noise}$ als nahezu identisch angesehen werden können und in der Praxis nicht beachtet werden müssen.

Abbildung 3.21 zeigt den typischen Verlauf der Kostenentwicklung der Evolution einer digitalisierten Bildvorlage. Derartig gewonnene Randkurven enthalten zunächst viele kurze Segmente, die während der Phase der Rauschfilterung zusammengefasst werden. Die Kostenkurve ist während dieser Phase nahezu konstant. Sie zeigt daraufhin steigende Tendenz, die Form wird stärker verändert, die Phase der Formabstraktion setzt ein.

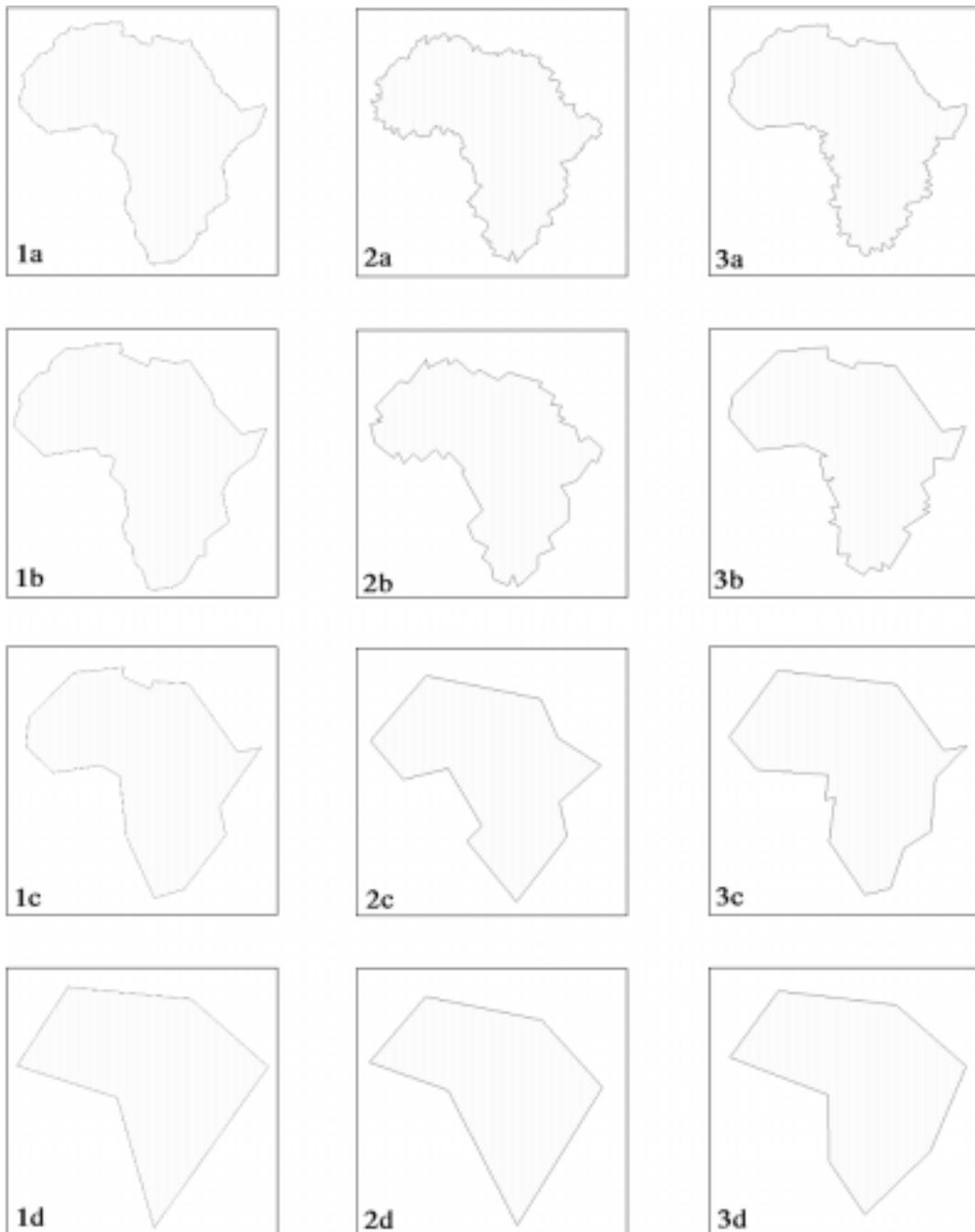


Abbildung 3.20: Evolution dreier verschieden gestörter Versionen der Form 'Afrika'.

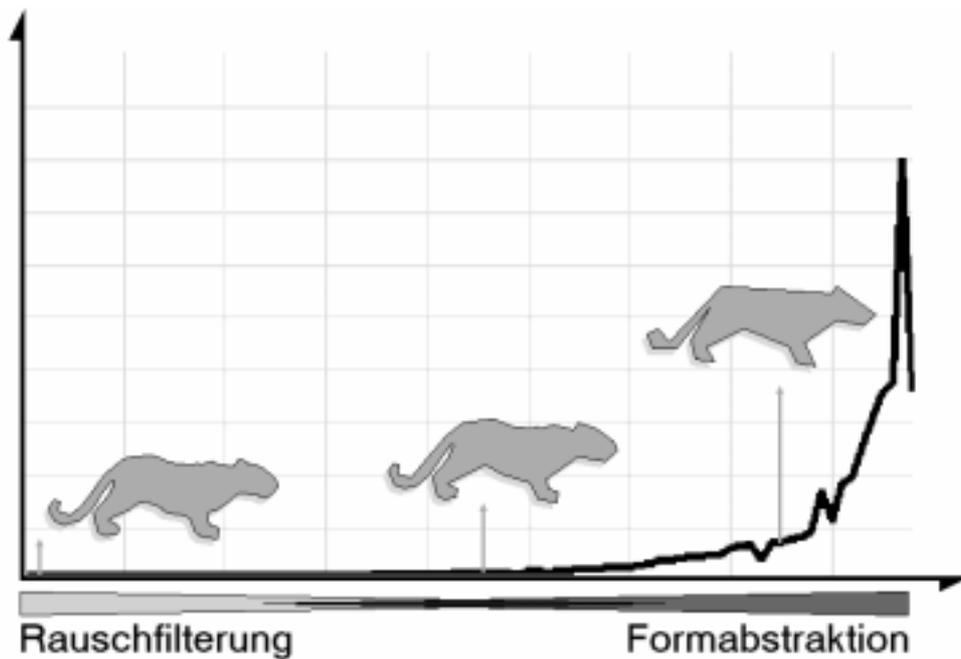


Abbildung 3.21: Kostenentwicklung der Evolution: die x -Achse zeigt die Evolutionsstufe, die y -Achse die Kosten des zuletzt entfernten Randpunktes.

3.6.2.5 Abstraktionsstufen

Entscheidend ist das automatische Erkennen des Übergangs von der Rauschelimination zur Formabstraktion, siehe Abbildung 3.21. Das zugehörige Verfahren basiert auf dem *Formvergleich* der einzelnen *Teile* der Ausgangsform mit den korrespondierenden Teilen jeder Abstraktionsstufe (die Definition und das Verfahren des Formvergleiches folgen in Kapitel 5):

Es seien $\mathcal{P} = \mathcal{P}^0, \mathcal{P}^1, \dots, \mathcal{P}^m$ die Stufen der Evolution \mathcal{E} , \mathcal{P}^m sei das erste konvexe Polygon.

Ausgehend von $i = m$ wird wie folgt verfahren:

- \mathcal{P}^i wird in konvexe/konkave Teilpolygone zerlegt. Jedes dieser Polygone stellt ein visuell relevantes Teil dar, siehe Kap. 4
- Jedes Teilpolygon $\gamma_j^i \subset \mathcal{P}^i$ wird mit dem zu ihm korrespondierenden Teilpolygon $\gamma_j^0 \subset \mathcal{P}^0$ verglichen³. Das zu $\gamma_j^i = (p_s, \dots, p_e)$ korrespondierende Teilpolygon ist diejenige Teilfolge in \mathcal{P}^0 mit $\gamma_j^0 = (p_s, \dots, p_e)$ und $\gamma_j^i \subseteq \gamma_j^0$.

³Der Vergleich der Teilpolygone greift dem Vergleich von Formen durch polygonale Randkurven voraus. Dieser Vergleich ist Inhalt des späteren Kapitels 5, das Vergleichsmaß wird hier nicht vorgestellt.

- Falls der Vergleich mind. eines Teilpolygons zu einem Wert führt, der oberhalb eines gegebenen Schwellwertes T liegt, wird davon ausgegangen, daß das Polygon eine zu stark abstrahierte Evolutionsstufe darstellt, es wird mit der vorhergehenden Stufe, i^{-1} fortgefahren.

Diese Prozedur wird fortgeführt, bis der Vergleich aller Teilpolygone γ_j^i mit ihren korrespondierenden Teilen γ_j^0 einen Wert unterhalb T erreicht. Die dritte Spalte der Abbildung 3.23 zeigt Ergebnisse der automatischen Abstraktion.

Bemerkungen:

- Der Vergleich der Teilpolygone hoher Abstraktionsstufen mit ihren korrespondierenden Teilen nutzt die Eigenschaft der Evolution, daß eine Stufe i , h Teilmenge aller vorhergehenden Stufen i , $0 \leq i < h$ ist, d.h. Randpunkte ihre Position behalten; die Korrespondenz existiert somit und kann durch einfachen Positionsvergleich festgestellt werden.
- Die Benutzung des festen Schwellwertes T führt zu intuitiv richtigen Abstraktionsstufen in verschiedensten Experimenten, kann jedoch nicht eine endgültige Lösung darstellen. Denkbar wäre eine komplexere Berechnung unter Einbeziehung statistischer Daten (z.B. mittlere Segmentlänge).
- In Kapitel 3.6.2.2 wurde erwähnt, daß das Evolutionsverfahren mit Stufenausgleich auflösungsadaptiv ist, obgleich der Stufenausgleich K von der Auflösung abhängt. Als Vorteil zu einstufigen Lösungen wurde insbesondere die Parameterfreiheit erwähnt, die Evolution wurde der Einführung eines metrischen Parameters bei einstufigen Verfahren entgegengestellt. Die in diesem Kapitel vorgestellte automatische Abstraktion ist der Ersatz für diesen Parameter.

3.6.2.6 Liniensegmente

Das Evolutionsverfahren ist motiviert durch die Zielsetzung, beliebig geformte Objekte bearbeiten zu können, und dabei die Eigenheiten der Vorgaben, nämlich digitaler Bilder, zu berücksichtigen.

Ein spezielles Interesse der Bildverarbeitung gilt der Erkennung von beliebig geformten, geradlinig begrenzten Objekten, z.B. Gebäude und Straßen in Luftaufnahmen. Die Schwierigkeit dabei besteht insbesondere in den technisch bedingten Störungen der aufzufindenden Kanten (Aufnahmerauschen, digitale Darstellung).

Bisherige Verfahren zur Erkennung derartiger Objekte arbeiten modellbasiert, z.B. Brunn et al. [6]. Dieses Verfahren arbeitet mit Informationen über lokale Eigenschaften von Eckpunkten anhand von Modellen typischer Punktkonstellationen, es ist also speziell auf die oben genannte Objektklasse ausgerichtet.

Das Evolutionsverfahren erkennt und extrahiert Liniensegmente ohne solches Vorwissen. Dabei ist zu beachten, daß das Ausgangspolygon durch ein digitales Randpolygon, also den Chaincode gegeben ist. Wie in den vorigen Kapiteln gezeigt, enthält diese Darstellung bzgl. des Stufenausgleichs keine relevante lokale Information, wie die Tangentenraumdarstellung des Beispiels in Abbildung 3.22 verdeutlicht. Besonders hier zeigt sich die Stärke der robusten Auflösungsadaptation.

Ein Beispiel zeigte Abbildung 3.2, die Ausgangsform (links) stammt aus einer Luftaufnahme aus [6], die evolvierte Form (rechts) zeigt deutlich die Extraktion der Kanten.

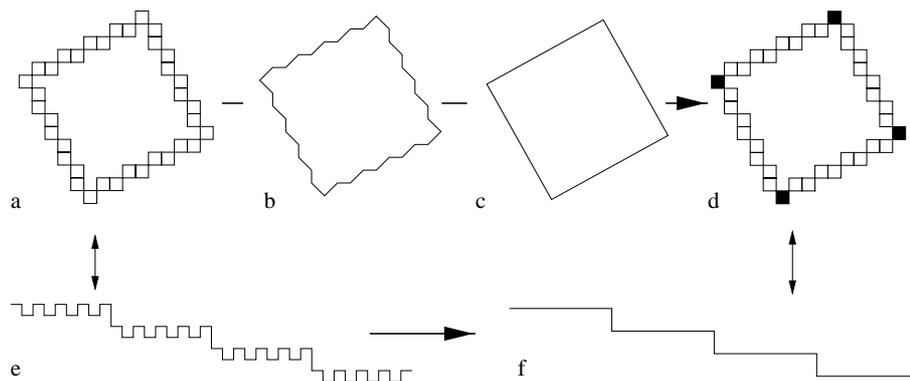


Abbildung 3.22: Extraktion von Strecken: a) Quadrat aus 4 digitalen Strecken; b) zugehöriges Randpolygon; c) Evolution von (b); d) von (c) extrahierte Ecken in (a); e) und f) Tangentenraumdarstellungen zu a) und d)

Die Extraktion von Liniensegmenten führt zu einer starken Beziehung zwischen der Polygonevolution und digitalen Strecken, siehe Abbildung 3.22, diese wird in Kapitel 3.9 näher untersucht.

3.6.2.7 Übereinstimmung mit menschlicher Intuition

Die Evolution wurde im Rahmen dieser Arbeit als Vorfilter der eigentlichen Formerkennung entwickelt, d.h. die Abstraktion darf, nach menschlicher Auffassung, die Form nicht verändern. Die Evolution stimmt überein mit menschlicher Intuition: zur Untermauerung dieser Aussage können nur Ergebnisse von Experimenten

gezeigt werden. Abbildung 3.23 zeigt einige Evolutionsstufen von Formen verschiedener Komplexität. Die Spalten zeigen (von links) die Ausgangsform, eine Stufe der Phase der Rauschverminderung, die automatisch erkannte Abstraktion, eine weiterführende, starke Abstraktion. Zu beachten ist, daß auch in der rechten Spalte zwar wichtige Teile entfernt wurden, die übrigen Punkte jedoch weiterhin wichtige Unterteilungen der Ausgangsform darstellen. Diese Beobachtung wird in späteren Arbeitsschritten (und Kapiteln dieser Arbeit) genutzt.

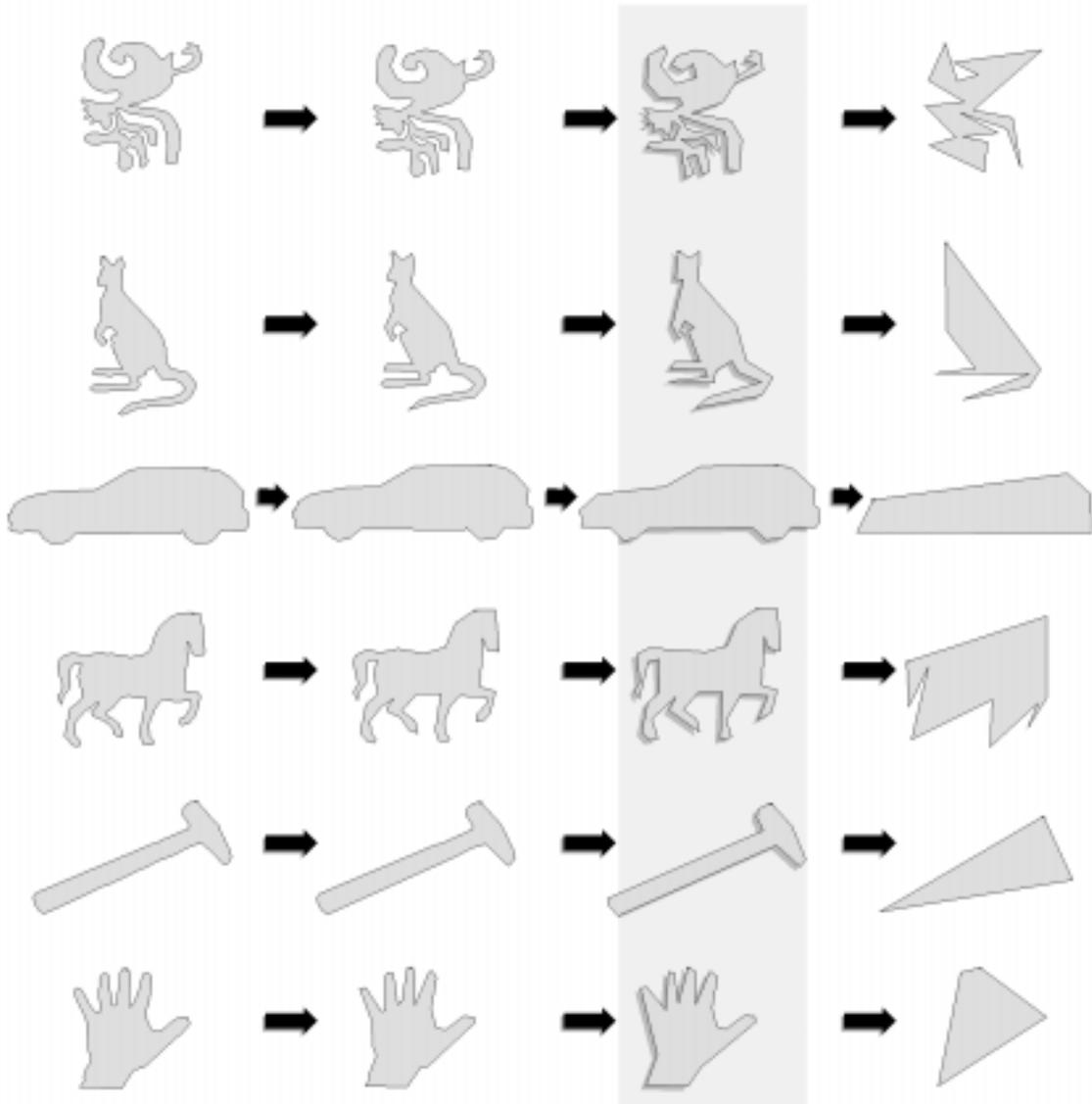


Abbildung 3.23: Ergebnisse der Evolution. Die unterlegte Spalte zeigt die automatisch erkannte Abbruchstufe.

3.7 Die Stetigkeit des Evolutionsverfahrens

Das folgende Kapitel wird zeigen, daß der Verlauf der Evolution zweier (im wesentlichen im Hausdorffschen Sinn) hinreichend ähnlicher Polygone \mathcal{P} und \mathcal{Q} 'gleichartig' verläuft. An die benutzte Kostenfunktion K werden dabei nur Forderungen gestellt, die visuell einfach motiviert werden können.

Neben der theoretischen Ergänzung durch diese Eigenschaft bekommt die im vorigen Kapitel beschriebene Adaptivität praktischen Sinn. Die dort beschriebene Auflösungsadaption war über den Vergleich zweier Mengen \mathcal{P} , \mathcal{Q} , \mathcal{Q}' definiert, deren Differenzmenge $\mathcal{P} \setminus \mathcal{Q}' =: \Delta_\Gamma$ nur Punkte p_i mit Informationsgehalt $K_{\Gamma'}(p) = 0$ enthalten. Für visuell sinnvolle Informationsfunktionen K bedeutet dieses, daß diese Punkte keine visuelle Ergänzung des Kantenzuges $\overline{\mathcal{Q}'}$ darstellen, d.h. für $p_i \in \Delta_\Gamma$ wird gefordert: $p_i \in \overline{\mathcal{Q}'}$. Die Kurvenevolution arbeitet in der Praxis mit Polygonen $\mathcal{P} \subset \mathbb{Z}^2$, d.h. mit diskretisierten Versionen kontinuierlicher Vorlagen. Die obige Forderung kann somit aufgrund des Diskretisierungsfehlers nicht erfüllt werden. Ist jedoch die diskrete Version des kontinuierlichen Polygons 'hinreichend ähnlich', so verläuft die Evolution wie von der Adaptiondefinition gefordert, die Punkte $p_i \in \Delta_\Gamma$ werden als erstes eliminiert.

Es seien \mathcal{P}, \mathcal{Q} zwei Polygone ⁴ für die o.B.d.A. gelte: $\#\mathcal{P} \leq \#\mathcal{Q}$.

Definition: Zugeordnete Eckpunkte

$\mathcal{V}_\mathcal{Q}^\mathcal{P} \subset \mathcal{Q}$ heißt die Menge der \mathcal{P} zugeordneten Eckpunkte aus \mathcal{Q} , falls gilt:

1. $\mathcal{V}_\mathcal{Q}^\mathcal{P} = (q_i | q_i \in \mathcal{Q}, \exists p \in \mathcal{P} : d(p, q_i) < d(p, q_j) \forall q_j \in \mathcal{Q} \setminus q_i)$
2. Die Eckenreihenfolge in $\mathcal{V}_\mathcal{Q}^\mathcal{P}$ entspricht derjenigen in \mathcal{Q} :
mit $\mathcal{Q} = (q_{i_0}, q_{i_1}, \dots, q_{i_n})$ und $\mathcal{V}_\mathcal{Q}^\mathcal{P} = (q_{j_0}, q_{j_1}, \dots, q_{j_m})$ gilt:
 $j_a < j_b \Leftrightarrow i_a < i_b, a, b \in 0, \dots, m = \#\mathcal{V}_\mathcal{Q}^\mathcal{P} \Leftrightarrow 1$

$q_i \in \mathcal{V}_\mathcal{Q}^\mathcal{P}$ heißt *nahester Eckpunkt* zu $p_i \in \mathcal{P}$, geschrieben $q_i = NV(p_i)$,
 $\Leftrightarrow d(q_i, p_i) < d(q_j, p_i) \forall q_j \in \mathcal{V}_\mathcal{Q}^\mathcal{P} \setminus q_i$.

Definition: Eckverbindung

Ein Teilpolygon $\mathcal{Q}^{p_s, s+1} = (q_0, \dots, q_n) \subset \mathcal{Q}$ heißt *Eckverbindung*, falls zwei auf-

⁴Zur besseren Lesbarkeit wird in diesem Kapitel die Notation der Polygone von \mathcal{P} und \mathcal{Q} zu \mathcal{P} und \mathcal{Q} gewandelt

einanderfolgende Eckpunkte $p_s, p_{s+1} \in \mathcal{P}$ existieren, so daß $q_0 = NV(p_s)$, $q_n = NV(p_{s+1})$.

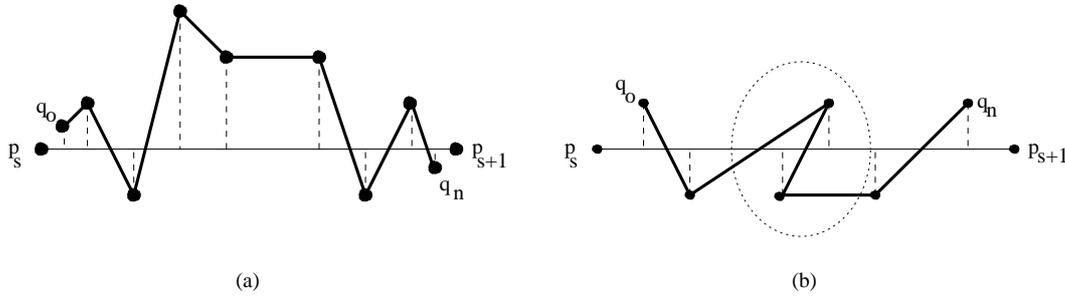


Abbildung 3.24: links projektiv geordnet, rechts nicht

Eine Eckverbindung $Q^{p_s, p_{s+1}} = (q_0, \dots, q_n) \subset \mathcal{Q}$ heißt *projektiv geordnet*, falls gilt:
 $\forall q_j \in Q^{p_s, p_{s+1}} \setminus q_n: d(p_s, P_{\mathcal{P}}(q_j)) < d(p_s, P_{\mathcal{P}}(q_{j+1}))$, wobei $P_{\mathcal{P}}(q)$ die metrische Projektion von q auf $\overline{p_s p_{s+1}}$ bezeichnet.

\mathcal{Q} heißt *vollständig projektiv geordnet*, falls alle Eckverbindungen projektiv geordnet sind.

Definition:

$Strip_H(\overline{p_s p_{s+1}}) = B(\overline{p_s p_{s+1}}, \frac{H}{2})$, $p_s, p_{s+1} \in \mathcal{P}$ bezeichnet einen *Hausdorff-Streifen* mit Abstand $\frac{H}{2}$ um die Kante $\overline{p_s p_{s+1}}$.

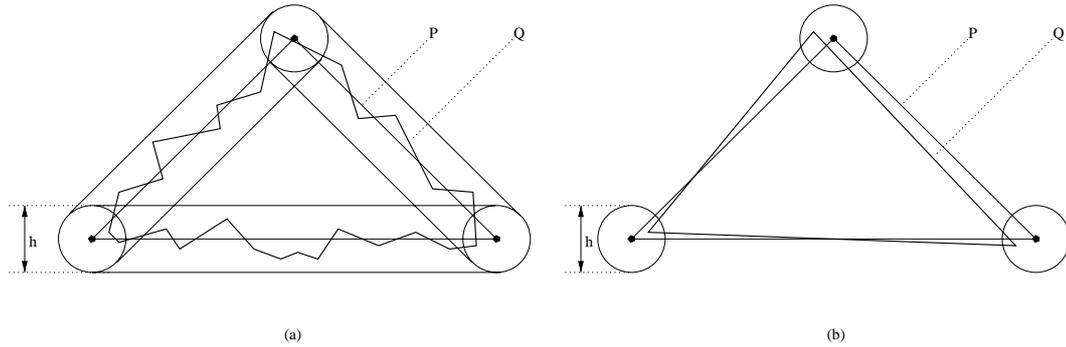
Definition: (H, d) -Approximation

\mathcal{Q} heißt *schwache (H, d) -Approximation* an \mathcal{P} , falls gilt:

1. \mathcal{Q} ist vollständig projektiv geordnet
2. Für alle Eckverbindungen $Q^{p_{s_i}, p_{s_{i+1}}}$ gilt: $Q^{p_{s_i}, p_{s_{i+1}}} \subset Strip_H(\overline{p_{s_i} p_{s_{i+1}}})$
3. $\forall p_i \in \mathcal{P} : NV(p_i) \in B(p_i, \frac{H}{2})$
4. $d > 0$ ist die minimale Kantenlänge in \mathcal{Q}

\mathcal{Q} heißt *starke (H, d) -Approximation* an \mathcal{P} , falls zusätzlich gilt:

5. $\#\mathcal{Q} = \#\mathcal{P}$

Abbildung 3.25: (a) schwache und (b) starke (H, d) -Approximation

Bemerkung:

Durch Punkt 3. ist gesichert, daß $\#\mathcal{V}_Q^{\mathcal{P}} = \#\mathcal{P}$, d.h. ist \mathcal{Q} schwache (H, d) -Approximation, so ist $\mathcal{V}_Q^{\mathcal{P}}$ starke (H, d) -Approximation.

Forderungen an die Kostenfunktion:

Es seien $p, q, r \in \mathbb{R}^2$. Die für das Evolutionsverfahren verwendete Kostenfunktion C besitze folgende Eigenschaften (siehe Abbildung 3.26):

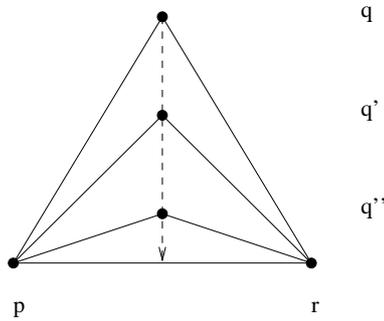


Abbildung 3.26: Zur Kostenfunktion

1. C ist stetig (bzgl. Hausdorff-Norm in \mathbb{R}^2)
2. $C(p, q, r) = 0 \Leftrightarrow p, q, r$ kollinear
3. $\forall p, q, r: d(q, \overline{pr}) \rightarrow 0 \Leftrightarrow C(p, q, r) \rightarrow 0$

Bei den folgenden Aussagen über das Evolutionsverfahren wird vorausgesetzt, daß die verwendete Kostenfunktion die obigen Forderungen erfüllt. Zusätzlich soll gelten, daß \mathcal{P} nur 'echte' Ecken besitzt, d.h. $C(p_{i-1}, p_i, p_{i+1}) > 0 \forall p_i \in \mathcal{P}$.

Die Bezeichnung $C(p_i)$ steht im folgenden vereinfachend für $C(p_{i-1}, p_i, p_{i+1})$, falls p_i aus einem geschlossenen Polygon ist.

Satz: Stetigkeit des Evolutionsverfahrens

Es sei $\mathcal{P} = \mathcal{P}^0, \dots, \mathcal{P}^m$ eine vom Evolutionsverfahren produzierte Folge von Polygonen.

Dann existiert zu jedem $d > 0$ ein $\delta > 0$ derart, daß zu jeder schwachen (δ, d) -Approximation \mathcal{Q} an \mathcal{P} eine Evolutionsstufe \mathcal{Q}^k existiert, für die gilt:

1. \mathcal{Q}^k ist starke (δ, d) -Approximation an \mathcal{P}^0
2. $\forall i = 0..k : \mathcal{Q}^i$ ist schwache (δ, d) -Approximation an \mathcal{P}^0

Gilt zusätzlich:

$\forall \mathcal{P}^0, \dots, \mathcal{P}^{m-1} : \#(\mathcal{P}^i \setminus \mathcal{P}^{i+1}) = 1$ sowie $\forall \mathcal{Q}^k, \dots, \mathcal{Q}^{k+m-1} : \#(\mathcal{Q}^i \setminus \mathcal{Q}^{i+1}) = 1$, d.h. wird in jedem Evolutionsschritt von \mathcal{P} und \mathcal{Q}^k nur genau ein Punkt entfernt, dann:

3. $\forall i = 0, \dots, m : \mathcal{Q}^{i+k}$ ist starke (δ, d) -Approximation an \mathcal{P}^i

(Anmerkung: Die Relation 'starke (δ, d) -Approximation' ist symmetrisch.)

Beweis:

Der Beweis ist eine direkte Folge aus den nächsten zwei Sätzen. ■

Die folgenden zwei Sätze kennzeichnen zwei Stufen des Beweises.

In der ersten Stufe wird gezeigt, daß aus einer schwachen (δ, d) -Approximation \mathcal{Q} mit geeignetem δ – die nach Voraussetzung eine höhere Anzahl von Eckpunkten als \mathcal{P} besitzt – vom Evolutionsverfahren zunächst alle Ecken aus \mathcal{Q} entfernt werden, die nicht Ecken aus \mathcal{P} zugeordnet sind.

Die zweite Stufe zeigt, daß das Evolutionsverfahren für die in der ersten Stufe entstandene starke (δ, d) -Approximation \mathcal{Q} an \mathcal{P} mit geeignetem δ in \mathcal{P} und \mathcal{Q} gleich verläuft, d.h. wird in einem Evolutionsschritt \mathcal{P} um p_i reduziert, so wird in diesem Evolutionsschritt \mathcal{Q} um $NV(p_i)$ reduziert.

Satz:

Gegeben sei ein Polygon \mathcal{P} . Dann existiert ein δ derart, daß zu jeder schwachen

(δ, d) -Approximation \mathcal{Q} eine Evolutionsstufe \mathcal{Q}^k existiert, die eine starke (δ, d) -Approximation an \mathcal{P} ist.

Beweis:

Man betrachte eine Folge (δ_i) mit $\mathbb{R}^+ \ni \delta_i \rightarrow 0$ sowie alle zugehörigen (δ_i, d) -Approximationen \mathcal{Q}_{δ_i} .

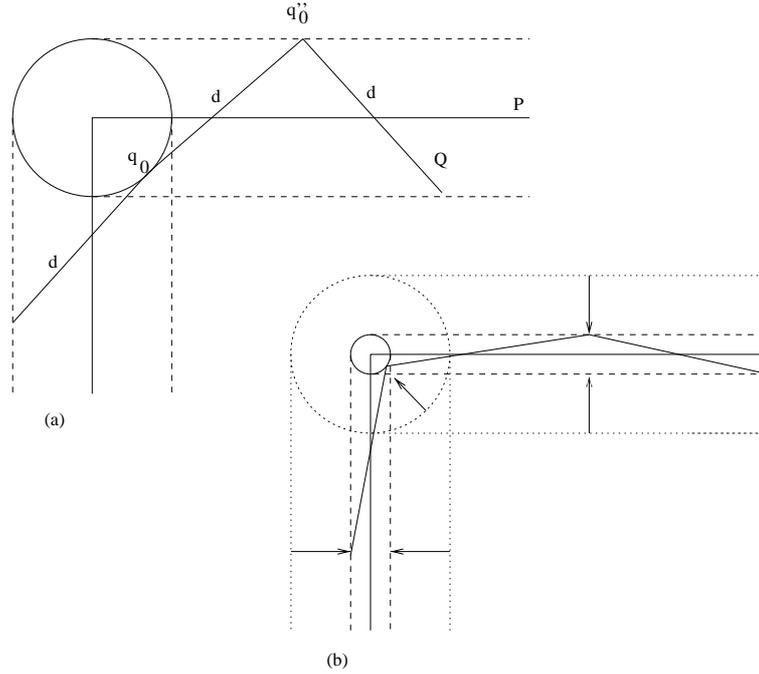


Abbildung 3.27: Alle Kanten in a) und b) haben die minimal Länge d . Gezeigt wird die Konstellation, die $C(q_0)$ minimiert und $C(q''_0)$ maximiert. a) $C(q_0) < C(q''_0)$, b) $C(q_0) > C(q''_0)$

$\mathcal{T}_{\delta_i}^{p_{s,s+1}}$ bezeichne die Menge aller geordneten Tripel von Eckpunkten einer Eckverbindung $\mathcal{Q}_{\delta_i}^{p_{s,s+1}} = (q_0, \dots, q_n)$, $p_s, p_{s+1} \in \mathcal{P}$, $\#\mathcal{Q}_{\delta_i}^{p_{s,s+1}} > 2$, die die Reihenfolge in $\mathcal{Q}_{\delta_i}^{p_{s,s+1}}$ respektieren:

$$\mathcal{T}_{\delta_i}^{p_{s,s+1}} = \{(q_{i_0}, q_{i_1}, q_{i_2}) \mid q_{i_0}, q_{i_1}, q_{i_2} \in \mathcal{Q}_{\delta_i}^{p_{s,s+1}}, i_0 < i_1 < i_2\}$$

Es gilt $\forall q_{i_1} \in \mathcal{Q}_{\delta_i}^{p_{s,s+1}} \setminus \{q_0, q_n\}$ aufgrund der Eigenschaften der Kostenfunktion mit $\delta_i \rightarrow 0$ (siehe Abbildung 3.27):

$$\forall (q_{i_0}, q_{i_1}, q_{i_2}) \in \mathcal{T}_{\delta_i}^{p_{s,s+1}} : C(q_{i_0}, q_{i_1}, q_{i_2}) \rightarrow 0$$

Man betrachte nun die Punkte $\mathcal{Q}_{\delta_i}^{p_{s,s+1}} \cap \mathcal{V}_{\mathcal{Q}_{\delta_i}}^{\mathcal{P}} = \{q_0, q_n\}$:

Es seien q'_0 und $q''_0 \in \mathcal{Q}_{\delta_i}$ die linke und rechte Nachbarecke von q_0 (zu beachten: $q'_0 \notin \mathcal{Q}_{\delta_i}^{p_{s,s+1}}$). Dann gilt (wiederum mit $\delta_i \rightarrow 0$) für (q'_0, q_0, q''_0) : $C(q_0) \rightarrow \epsilon > 0$,

denn:

$$\delta_i \rightarrow 0 \Rightarrow d(q'_0, \overline{p_{s-1}p_s}) \rightarrow 0 \text{ und } d(q''_0, \overline{p_s p_{s+1}}) \rightarrow 0 \text{ sowie } d(q_0, p_s) \rightarrow 0$$

$C(p_s) > 0 \Rightarrow p_{s-1}, p_s, p_{s+1}$ nicht kollinear $\Rightarrow q'_0, q_0, q''_0$ nicht kollinear und zusätzlich: $\min(d(q'_0, q_0), d(q_0, q''_0)) \rightarrow d' \geq d \Rightarrow d(q'_0, q_0, q''_0) \not\rightarrow 0 \Rightarrow C(q_0) \not\rightarrow 0$.

$$\Leftrightarrow \exists \epsilon > 0: C(q_0) > \epsilon.$$

(analog für q_n).

$$\Rightarrow \exists \delta^{p_s, s+1} \in (\delta_i) > 0: C(q_{i_1}) < \min\{C(q_0), C(q_n)\} \forall q_{i_1} \in (q_{i_0}, q_{i_1}, q_{i_2}) \in \mathcal{T}_{\delta_i}^{p_s, s+1}$$

Setzt man $\delta_{min} = \min_{p_s} \{\delta^{p_s, s+1}\}$, so gilt somit für jede schwache

(δ_{min}, d) -Approximation $\mathcal{Q}_{\delta_{min}}$ mit $\mathcal{Q}_{\delta_{min}} \neq \mathcal{V}_{\mathcal{Q}_{\delta_{min}}}^{\mathcal{P}}$:

$$C(q_{min}) \leq C(q_j) \forall q_j, q_{min} \in \mathcal{Q}_{\delta_{min}} \Rightarrow q_{min} \notin \mathcal{V}_{\mathcal{Q}_{\delta_{min}}}^{\mathcal{P}}$$

d.h. das Evolutionsverfahren entfernt keine Punkte aus $\mathcal{V}_{\mathcal{Q}_{\delta_{min}}}^{\mathcal{P}}$, solange $\mathcal{Q}_{\delta_{min}} \setminus \mathcal{V}_{\mathcal{Q}_{\delta_{min}}}^{\mathcal{P}} \neq \emptyset$. Da das Verfahren terminiert, existiert eine Stufe $\mathcal{Q}_{\delta_{min}}^k$ mit

$$\#\mathcal{Q}_{\delta_{min}}^k = \#\mathcal{V}_{\mathcal{Q}_{\delta_{min}}^k}^{\mathcal{P}} = \#\mathcal{P}$$

$\Rightarrow \mathcal{Q}_{\delta_{min}}^k$ ist starke (δ_{min}, d) -Approximation ■

Lemma:

Gegeben sei ein Polygon \mathcal{P} sowie eine schwache (δ, d) -Approximation \mathcal{Q} an \mathcal{P} . Jede Evolutionsstufe \mathcal{Q}^i von \mathcal{Q} , für die gilt: $\mathcal{V}_{\mathcal{Q}^i}^{\mathcal{P}} = \mathcal{V}_{\mathcal{Q}}^{\mathcal{P}}$ ist schwache (δ, d) -Approximation an \mathcal{P} .

Beweis:

Die Stufen \mathcal{Q}^i entstanden durch Entfernung von Punkten, deren Nachbarpunkte Elemente derselben Eckverbindung war. ■

Der obige Satz und das Lemma beweisen die Aussagen (1) und (2) des Stetigkeitssatzes. Zu Aussage (3):

Satz:

Es sei $\mathcal{P} = \mathcal{P}^0, \dots, \mathcal{P}^m$ eine vom Evolutionsverfahren produzierte Folge von Polygonen für welche gelte:

$$\forall \mathcal{P}^0, \dots, \mathcal{P}^{m-1}: \#(\mathcal{P}^i \setminus \mathcal{P}^{i+1}) = 1$$

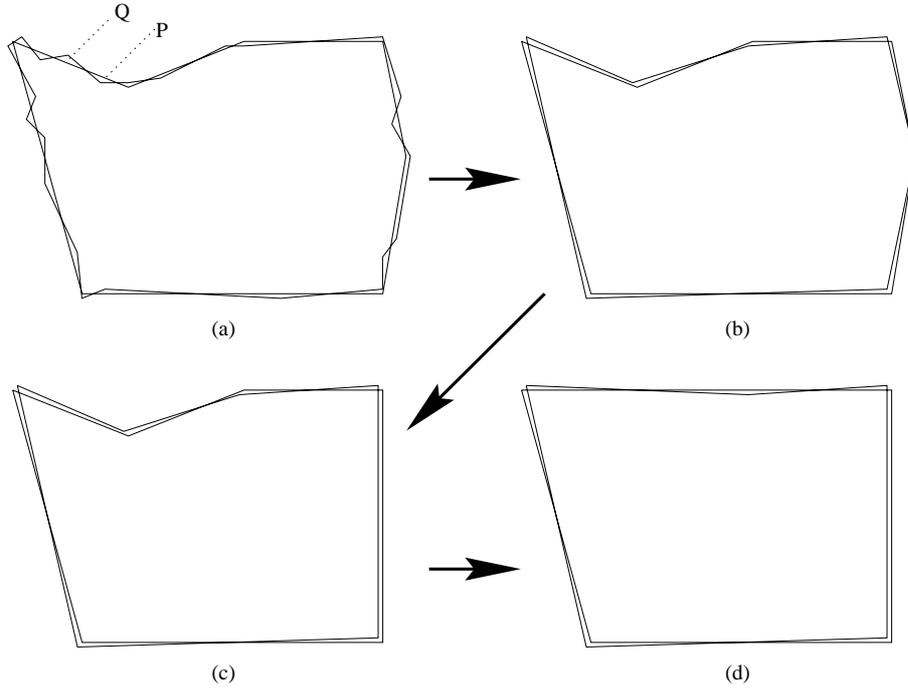


Abbildung 3.28: Zum Stetigkeitsatz: (a)→(b) zeigt die Elimination von nicht \mathcal{P} zugeordneten Punkten, (b)→(c)→(d) zwei Evolutionsschritte in \mathcal{P} , sowie die gleichartigen Evolutionsschritte der starken (δ, d) -Approximation

Dann existiert ein δ derart, daß für jede starke (δ, d) -Approximation \mathcal{Q} und für alle Evolutionsstufen $\mathcal{P}^k, \mathcal{Q}^k$ gilt:

$$\mathcal{P}^k \setminus \mathcal{P}^{k+1} = \{p_{i_0}\} \Rightarrow \mathcal{Q}^k \setminus \mathcal{Q}^{k+1} = \{NV(p_{i_0})\}$$

Beweis:

Es sei $\mathcal{T}_{\mathcal{P}}$ die Menge aller Tripel von Eckpunkten in \mathcal{P} :

$$\mathcal{T}_{\mathcal{P}} = \{(p_{i_0}, p_{i_1}, p_{i_2}) \mid p_{i_0}, p_{i_1}, p_{i_2} \in \mathcal{P}\}$$

$C(p_{i_1}), p_{i_1} \in (p_{i_0}, p_{i_1}, p_{i_2}) \in \mathcal{T}_{\mathcal{P}}$ seien die zugehörigen Kosten.

$\mathcal{M} \subset \mathcal{P}^0$ bezeichne alle Punkte mit minimalen Kosten bezüglich aller Tripel in $\mathcal{T}_{\mathcal{P}}$:

$$\mathcal{M} = \{p_{i_1} \in (p_{i_0}, p_{i_1}, p_{i_2}) \in \mathcal{T}_{\mathcal{P}} \mid C(p_{i_1}) \leq C(p'_{i_1}) \forall p'_{i_1} \in (p'_{i_0}, p'_{i_1}, p'_{i_2}) \in \mathcal{T}_{\mathcal{P}}\}$$

Es sei:

$$c_{sec} = \min(C(p_{j_1}), p_{j_1} \in (p_{j_0}, p_{j_1}, p_{j_2}), p_{j_1} \notin \mathcal{M},$$

$$c_{\Delta} = c_{sec} \Leftrightarrow c_{min}$$

Aufgrund der Eigenschaften der Kostenfunktion gilt mit $\delta \rightarrow 0$:

$$\forall p_j \in \mathcal{P}: |C(NV(p_j)) \Leftrightarrow C(p_j)| \rightarrow 0 \Rightarrow \exists \delta_{min} : |C(NV(p_j)) \Leftrightarrow C(p_j)| < \frac{c_{\Delta}}{2}$$

Es sei p_i der Punkt minimaler Kosten in Stufe \mathcal{P}^i . Dann folgt $\forall p_j \in \mathcal{P}^i \setminus p_i$:

$$\underbrace{C(p_i) + \frac{c\Delta}{2}}_{>C(NV(p_i))} < C(p_j) \Leftrightarrow \underbrace{C(p_j)}_{<C(NV(p_j))} < \frac{c\Delta}{2}$$

$$\Leftrightarrow C(NV(p_i)) < C(NV(p_j))$$

und damit: $\mathcal{Q}^{i+1} = \mathcal{Q}^i \setminus NV(p_i)$ ■

Bemerkung:

Die zusätzliche Annahme zu Aussage 3 des Stetigkeitssatzes, nämlich $\forall \mathcal{P}^0, \dots, \mathcal{P}^{m-1}: \#(\mathcal{P}^i \setminus \mathcal{P}^{i+1}) = 1$ ist notwendig, wie das Beispiel in Abbildung 3.29 zeigt:

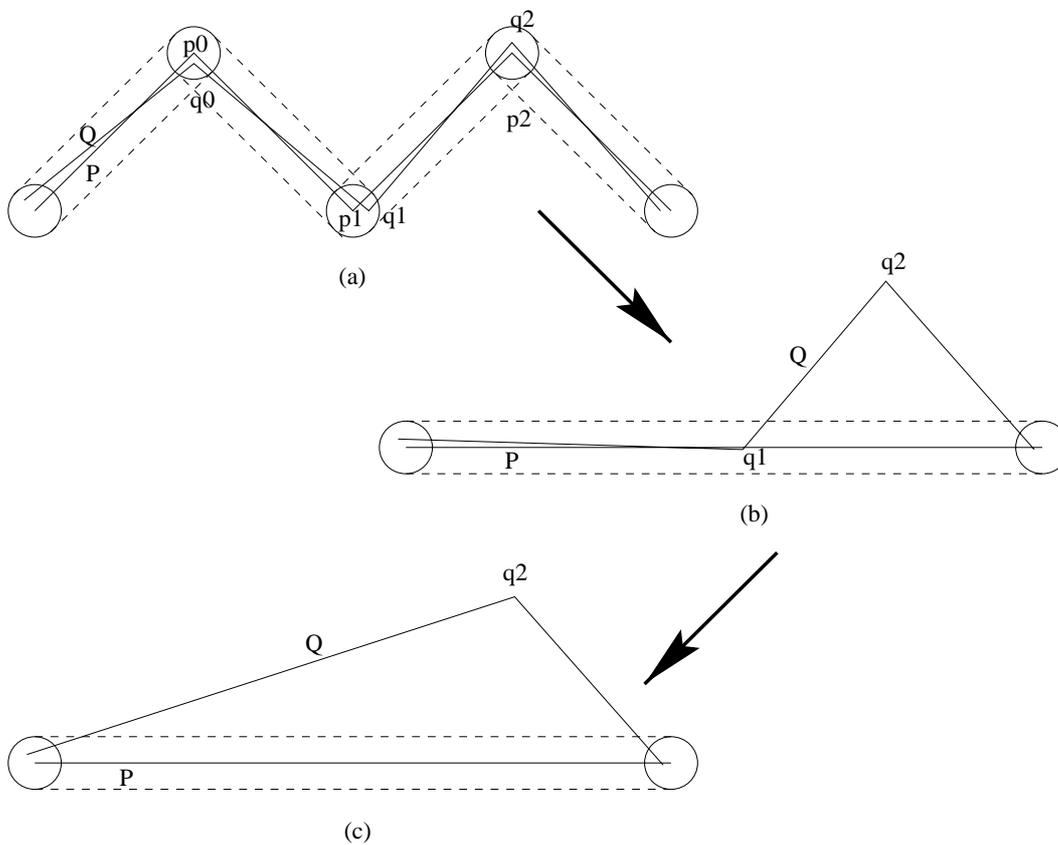


Abbildung 3.29: Gegenbeispiel

- a) Mit $C(p_0) = C(p_1) = C(p_2) = c_{min}$ werden die Punkte p_1, p_2, p_3 gleichzeitig aus \mathcal{P} entfernt.
- b), c) Die Evolution in \mathcal{Q} dagegen verläuft anders, da $C(p_0), C(p_1), C(p_2)$ paarweise verschieden sind. Die Evolutionsschritte in \mathcal{Q} sind keine (H, d) -Approxima-

tionen.

3.8 Die topologierhaltende Evolution

Kapitel 3 zeigte, wie die naheste Abstraktion durch Polygonlinearisierung einfach gebildet werden kann. Jedoch kann diese simple Vorgehensweise in der Praxis zu einer erheblichen Schwierigkeit führen, es kann z.B., so das Ausgangspolygon $\overline{\gamma} = \overline{\gamma}^0$ eine Jordankurve darstellt, die Jordaneigenschaft nicht für alle Polygone $\overline{\gamma}^i, i=1, \dots, m$ gesichert werden, siehe Abbildung 3.30.

Ein für die spätere Weiterverarbeitung des Polygons (die Gliederung in Formteile, Kapitel 4) wichtiges Problem ist das folgende: Die Gliederung des Polygons γ in Teilpolygone γ_i bezieht sich auf maximale konvexe/konkave Bögen. Diese werden über die Drehrichtung der Wendewinkel in den Punkten $p_i \in \gamma$ bestimmt.

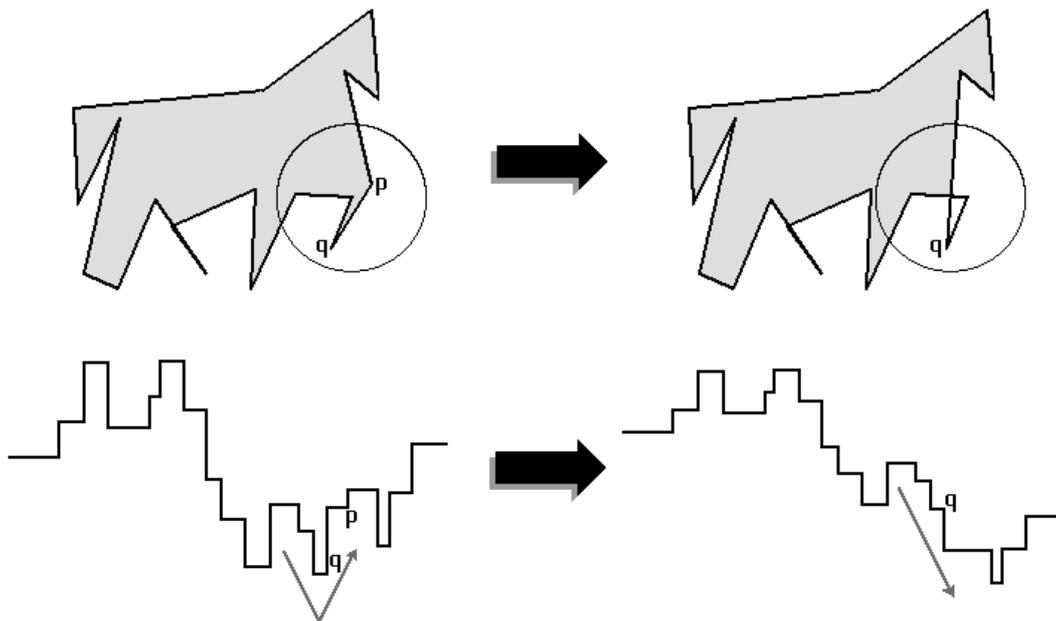


Abbildung 3.30: Selbstüberschneidung und Änderung der Drehrichtung

Das einfache Evolutionsverfahren erhält die Drehrichtung nicht notwendigerweise, wie die Abbildung 3.30 zeigt: durch Entfernung des Punktes p beschreibt γ' in q eine Linkskurve, und nicht, wie ursprünglich in γ , eine Rechtskurve. Die Tangentenraumdarstellung verdeutlicht den drastischen Unterschied der Kurvenverläufe: Der Wechsel von einem konkaven zu einem konvexen Bogen in q entfällt,

es ergibt sich ein einzelner Bogen (mit großem Wendewinkel). Diese nichtintuitive Verschmelzung führt zu einer Fehlunterteilung der Randkurve. (Bem.: wie Abbildung 3.33 zeigt, ist die Änderung des Drehwinkels unabhängig von der Jordaneigenschaft).

Das topologieerhaltende Evolutionsverfahren sichert, daß auch γ' ein Jordan-Polygon ist, indem Eckpunkte, deren Entfernung zu Selbstüberschneidung führen würden, zur Reduzierung nicht zugelassen werden. Die Bedingung, die dabei zur *Sperrung* der entsprechenden Punkte führt, erhält dabei automatisch die Drehrichtung des Wendewinkels.

Definition: Sperrung

Eine Ecke $p_i \in \gamma$ heißt *gesperrt* von einer Ecke $p_j \in \gamma$, (Notation: $p_i \leftarrow p_j$), falls $p_j \neq p_{i-1}, p_i, p_{i+1}$ und p_j innerhalb des Dreiecks $\Delta(p_{i-1}, p_i, p_{i+1})$ liegt, siehe Abbildung 3.31.

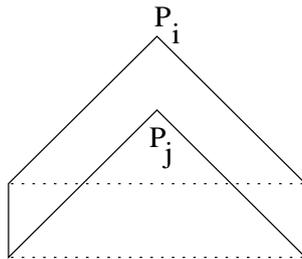


Abbildung 3.31: $p_i \leftarrow p_j$, p_j ist nicht gesperrt.

Existiert in γ mindestens eine Ecke, welche nicht gesperrt ist, so heißt γ *reduzierbar*.

Es stellt sich die Frage, ob jedes Jordanpolygon $\overline{\gamma} \subset \mathbb{R}^2$ reduzierbar ist, d.h. ob weiterhin die Evolution bis hin zum Dreieck vollziehbar ist, oder ob Polygone entstehen können, in welchen alle Ecken gesperrt sind. Im folgenden wird gezeigt, daß derartig gesperrte Polygone nicht existieren, d.h. das Evolutionsverfahren endet weiterhin in einem konvexen Polygon.

$I(\gamma)$ bezeichne das (durch den Jordanschen Kurvensatz gesicherte) Innere von γ , vereinigt mit γ .

Lemma 1:

Es sei $p_i \in \gamma$, eine echt konvexe und nicht gesperrte Ecke, $\gamma' := \gamma \setminus \{p_i\}$ das aus γ

reduziert um p_i entstehende Polygon. Dann gilt:

p_i echt konvex und reduzierbar $\Leftrightarrow p_i \notin I(, ')$.

Beweis:

p_i echt konvex und reduzierbar

$\Leftrightarrow \Delta(p_{i-1}, p_i, p_{i+1}) \subset I(,)$

$\Leftrightarrow I(, ') = I(,) \setminus \Delta(p_{i-1}, p_i, p_{i+1}) \cup \overline{p_{i-1}p_{i+1}}$

$\Leftrightarrow p_i \notin I(, ')$ ■

Satz 1:

Es sei wiederum $p_i \in ,$ eine nicht gesperrte Ecke, $,'$ das aus $,$ reduziert um p_i entstehende Polygon. Dann gilt:

1. $,'$ ist Jordan-Polygon, d.h. $\overline{p_{i-1}p_{i+1}}$ schneidet keine Segmente von $,$
2. Die Umlaufrichtung von $,'$ entspricht der Umlaufrichtung von $,$, d.h. $I(,)$ rechts von $\overline{p_{i-1}p_i} \Leftrightarrow I(, ')$ rechts von $\overline{p_{i-1}p_{i+1}}$.

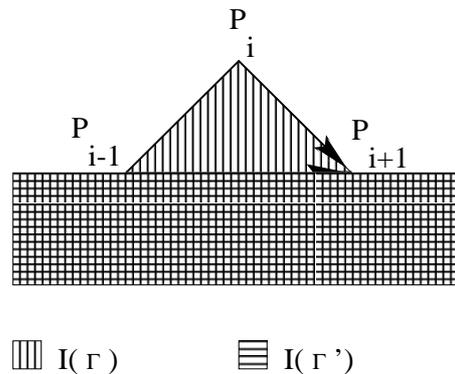


Abbildung 3.32: Reduzierung ohne Änderung der Umlaufrichtung

Beweis:

1. (Beweis klar)
2. p_i ist echt konvexe bzw. echt konkave Ecke, denn mit $\overline{p_{i-1}p_i} \parallel \overline{p_i p_{i+1}}$ und p_i gesperrt wäre $,$ kein Jordan-Polygon.

Es sei p_i eine echt konvexe Ecke, die Umlaufrichtung o.B.d.A so gewählt, daß $I(,)$ rechts von $\overline{p_{i-1}p_i}$ liegt. Dann:

$$\overline{p_{i-1}p_{i+1}} \setminus \{p_{i-1}, p_{i+1}\} \subset I(\cdot), I(\cdot') = (I(\cdot) \setminus \Delta(p_{i-1}, p_i, p_{i+1})) \cup \overline{p_{i-1}p_{i+1}}$$

$$\Rightarrow I(\cdot') \text{ ist rechts von } \overline{p_{i-1}p_{i+1}} \text{ (siehe Abbildung 3.32)}$$

(analog verläuft der Beweis für eine echt konkave Ecke)

■

Corollar:

In der topologieerhaltenden Evolution $\gamma = \gamma^0, \dots, \gamma^n$ haben alle γ^i die gleiche Umlaufrichtung.

Bemerkung:

Die Sperrungsbedingung durch den Dreieckstest ist durch den Erhalt der Umlaufrichtung stärker als ein reiner Test auf Selbstüberschneidung. Das Beispiel in Abbildung 3.33 zeigt ein Polygon, welches ohne Selbstüberschneidung reduzierbar ist, jedoch seine Umlaufrichtung ändert.

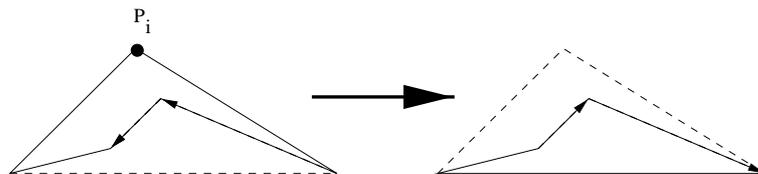


Abbildung 3.33: Zu Satz 1

Satz 2:

Jedes geschlossene Jordan-Polygon γ in \mathbb{R}^2 ist durch die topologieerhaltende Evolutionsvorschrift reduzierbar.

Beweis: Beweis durch Widerspruch, es wird angenommen, γ sei nicht reduzierbar.

Nach Voraussetzung ist γ Jordankurve in \mathbb{R}^2 .

Es gilt:

\exists konvexe, nicht reduzierbare Ecke $p_{i_0} \in \gamma$,

$\Rightarrow \exists$ Ecke $p_{i_1} \in \gamma$, für welche gilt:

- Sperrung:

$$p_{i_0} \leftarrow p_{i_1}$$

- Abstandsforderung:

p_{i_1} ist p_{i_0} bezüglich der metrischen Projektion auf $\overline{p_{i_0-1}p_{i_0+1}}$ am nächsten:

$$d(p_i, \overline{p_{i_0-1}p_{i_0+1}}) \leq d(p_{i_1}, \overline{p_{i_0-1}p_{i_0+1}}) \quad \forall p_i \neq p_{i_0}, \text{ (siehe Abbildung 3.34).}$$

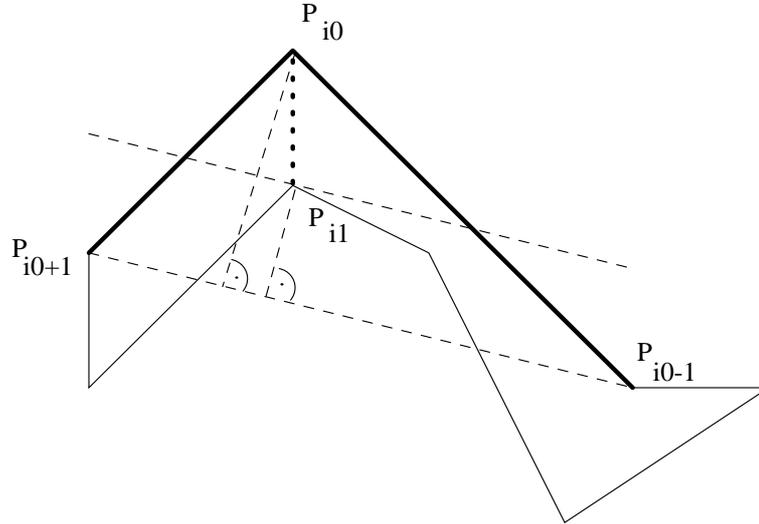


Abbildung 3.34: Zur Abstandsforderung

Durch die Abstandsforderung gelten für die Ecke p_{i_1} folgende Aussagen:

- Die Verbindungsstrecke $\overline{p_{i_0}p_{i_1}}$ liegt vollständig in $I(\cdot, \cdot)$, denn:
 1. p_{i_0} konvexe Ecke $\wedge p_{i_1} \in \Delta(p_{i_0-1}, p_{i_0}, p_{i_0+1})$
 $\Rightarrow (\overline{p_{i_0}p_{i_1}} \setminus \{p_{i_0}, p_{i_1}\}) \cap I(\cdot, \cdot) \neq \emptyset$
 2. $(\overline{p_{i_0}p_{i_1}} \setminus \{p_{i_0}, p_{i_1}\}) \cap \overline{p_i p_{i+1}} = \emptyset \quad \forall p_i \in \cdot, \cdot$, denn
 $(\overline{p_{i_0}p_{i_1}} \setminus \{p_{i_0}, p_{i_1}\}) \cap \overline{p_i p_{i+1}} \neq \emptyset$
 $\Rightarrow \exists p_i$ mit $d(p_i, \overline{p_{i_0-1}p_{i_0+1}}) > d(p_{i_1}, \overline{p_{i_0-1}p_{i_0+1}})$ im Widerspruch zur Abstandsforderung.
- 1, 2 $\Rightarrow \overline{p_{i_0}p_{i_1}} \subseteq I(\cdot, \cdot)$
- p_{i_1} ist konkav (Beweis analog zu 1,2)

$\Rightarrow (p_{i_0}, p_{i_0+1}, \dots, p_{i_1}, p_{i_0}) =: \cdot', \cdot'$ ist wiederum ein geschlossenes Jordan-Polygon mit $\# \cdot' < \# \cdot$, weil $p_{i_0-1} \neq p_{i_1} \Rightarrow p_{i_0-1} \notin \cdot'$

\cdot', \cdot' Jordan-Polygon $\Rightarrow \exists$ konvexe Ecke $p_{i_2} \neq p_{i_0}, p_{i_1}$, da jedes geschlossene Jordan-Polygon mind. 3 konvexe Ecken besitzt.

Es sei $p_{i_2} \neq p_{i_0}$ die von p_{i_0} aus in Umlaufrichtung *erste* konvexe Ecke. Dann gilt:

p_{i_2} ist nicht reduzierbar, da $(p_{i_2-1}, p_{i_2}, p_{i_2+1})$ ein Teilweg von γ ist (Anm.: dieses gilt nicht für p_{i_0})

Es folgt wiederum: $\exists p_{i_3}$ mit:

- $p_{i_2} \leftarrow p_{i_3}$
- $d(p_i, \overline{p_{i_0-1}p_{i_0+1}}) \leq d(p_{i_3}, \overline{p_{i_0-1}p_{i_0+1}}) \forall p_i \neq p_{i_0}, p_{i_1}$

Der naehste sperrende Punkt p_{i_3} kann nicht aus dem 'abgeschnittenen' Teil $\gamma \setminus \gamma'$ von γ stammen, da:

$\Delta(p_{i_2-1}, p_{i_2}, p_{i_2+1}) \subset I(\gamma, \gamma')$ und

$\forall p_j \in \gamma \setminus \gamma' : p_j \notin I(\gamma, \gamma')$ (da keine Seite von $\gamma \setminus \gamma'$ die Strecke $\overline{p_{i_0-1}p_{i_0+1}}$ schneidet), aber $\overline{p_{i_3}p_{i_2}} \subset I(\gamma, \gamma')$ (siehe (2))

$\Rightarrow p_{i_3} \in \gamma'$

Wie schon oben γ' aus einer echten Teilmenge der Ecken aus γ entstanden ist, kann nun wiederum das Jordan-Polygon

$\gamma'' := (p_{i_2}, p_{i_2+1}, \dots, p_{i_3}, p_{i_2})$ betrachtet werden

Es gilt: $p_{i_2-1} \neq p_{i_3} \Rightarrow p_{i_2-1} \notin \gamma'' \Rightarrow \# \gamma'' < \# \gamma'$.

Durch Induktion gelten somit für γ^k die gleichen Aussagen wie für γ^{k+1}

Es muß also erfüllt sein: $\# \gamma^k > \# \gamma^{k-1} > \# \gamma^{k-2} > \dots$, wobei jedes der γ^i ein geschlossenes Jordan-Polygon mit zumindest einer irreduziblen konvexen Ecke aus γ enthält. Da nach Voraussetzung $\# \gamma = n$ (endlich), folgt:

$\exists N \leq n \Leftrightarrow \exists : \# \gamma^N = 3$, d.h. nach spätestens $n \Leftrightarrow 3$ Schritten entsteht ein konvexes Jordan-Polygon, welches keine irreduziblen Ecken enthält.

\Rightarrow Widerspruch ■

Bemerkung: Die Gültigkeit des Jordanschen Kurvensatzes ist notwendig, wie das Beispiel in Abbildung 3.35 zeigt:

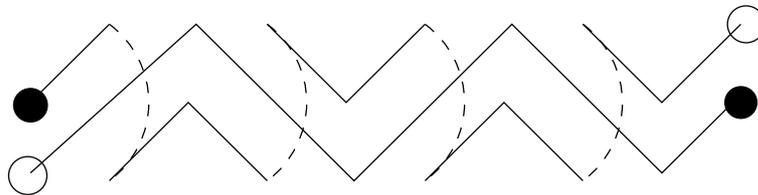


Abbildung 3.35: Geschlossenes, nicht-reduzierbares Jordan-Polygon auf einem Torus (die mit gleichen Marken versehenen Punkte sollen als identisch angesehen werden)

3.9 Polygonevolution und digitale Strecken

Die Polygonevolution wurde ursprünglich nur zur Glättung der Randkurve benutzt. Die Ausgangsrandkurve, also der Chaincode, wurde dazu zunächst in digitale Strecken zerlegt, diese (nicht eindeutigen !) digitalen Strecken lieferten die Eckpunkte des zu glättenden Polygons.

Im Laufe der Entwicklung des Systems stellte sich jedoch heraus, daß dieser Vorverarbeitungsschritt nicht notwendig ist, da die Polygonevolution, direkt auf den Chaincode angewendet, ähnliche, sogar zum Teil intuitivere Aufteilungen lieferte.

Bei näherer Betrachtung ergibt sich ein enger Zusammenhang zwischen der Polygonevolution und digitalen Strecken, d.h. die Polygonevolution liefert den digitalen Strecken 'ähnliche' Aufteilungen. Diese Idee soll im folgenden präzisiert werden.

Vorbetrachtungen:

Man betrachte die Kostenfunktion $C : \mathbb{R} \rightarrow \mathbb{R}$, $C(x) = w(x) \frac{l_1(x)l_2(x)}{l_1(x)+l_2(x)}$ für folgende Konstellationen:

1. $l_1 = \sqrt{h^2 + b_1^2}$, $l_2(x) = \sqrt{(x + b_2)^2 + h^2}$, $w(x) = \arctan \frac{h}{b_1} + \arctan \frac{h}{x+b_2}$,
 $x \in (0, \infty]$, $h \in \mathbb{R}^+$ fest, $b_1, b_2 \in \mathbb{R}^+$ fest.

Dieses entspricht der Kostenfunktion im Punkt p in Abbildung 3.36.

Es gilt:

$$b_1 \geq b_2: \max_x \{C(x)\} = C(0)$$

$$b_1 < b_2: \max_x \{C(x)\} = \lim_{x \rightarrow \infty} C(x)$$

2. $l_1(x) = \sqrt{(x + b)^2 + h^2}$, $l_2(x) = \sqrt{(L \Leftrightarrow b \Leftrightarrow x)^2 + h^2}$,
 $w(x) = \arctan \frac{h}{x+b} + \arctan \frac{h}{L-b-x}$,
 $x \in (0, L \Leftrightarrow 2b)$, $L \in \mathbb{R} \geq 2b$ fest, $h, b \in \mathbb{R}^+$ fest.

Dieses entspricht der Kostenfunktion in $p(x)$ in Abbildung 3.37.

Es gilt: $\max_x \{C(x)\} = \{C(0), C(L \Leftrightarrow 2b)\}$

3. Die Kostenfunktion ist monoton in w, l_1, l_2 , d.h.:

$$l_1 < l'_1 \Leftrightarrow w \frac{l_1 l_2}{l_1 + l_2} < w \frac{l'_1 l_2}{l'_1 + l_2}, \quad l_2 < l'_2 \Leftrightarrow w \frac{l_1 l_2}{l_1 + l_2} < w \frac{l_1 l'_2}{l_1 + l'_2}, \quad w < w' \Leftrightarrow w \frac{l_1 l_2}{l_1 + l_2} < w' \frac{l_1 l_2}{l_1 + l_2}$$

Satz: Beschränktheit der Kostenfunktion in einem Rechteck

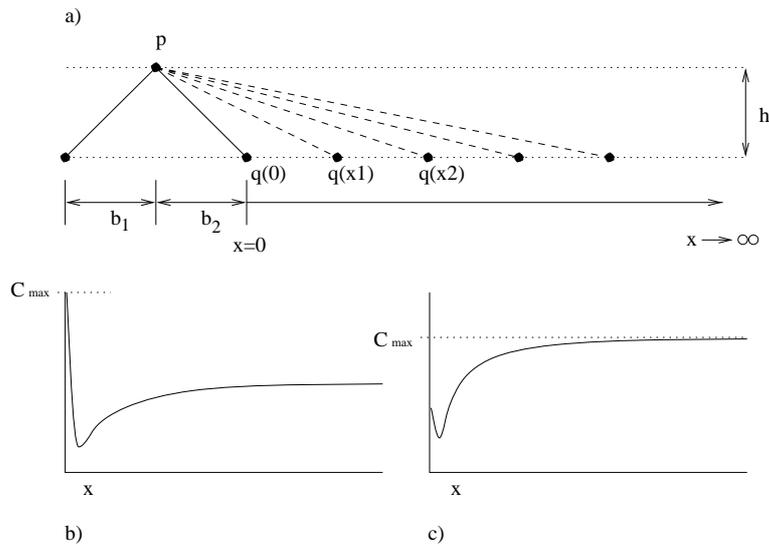


Abbildung 3.36: a) Wanderndes q und Verlauf der Kostenfunktion für b) $b_1 \geq b_2$, c) $b_1 < b_2$

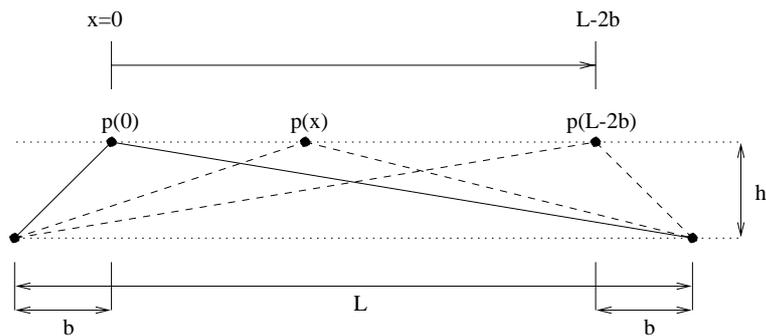


Abbildung 3.37: Wanderndes p

Gegeben sei ein Rechteck $\mathcal{R} \subset \mathbb{R}^2$ der Breite b und Höhe h mit $b > h > 0$ sowie ein offener Polygonzug $(q_0, p, q_1) \subset \mathcal{R}$ mit minimaler Kantenlänge $d \geq h$. Zusätzlich gelte für q_0, p, q_1 : die Projektion auf eine Seite mit Länge b des Rechtecks erhält die Ordnung (q_0, p, q_1) .

Dann gilt:

Die Kostenfunktion $C(p)$ ist nach oben beschränkt durch $c_{max} = d \arcsin \frac{h}{d}$

Beweis:

Die Seiten des Rechtecks seien bezeichnet wie in Abbildung 3.38, o.B.d.A liege p_0 auf Seite D . $P_a(q_0), P_a(q_1)$ seien die Projektionen von q_0, q_1 auf A , $P_c(p)$ die Projektion von p auf C . p' sei der Punkt auf C mit $d(P_a(q_0), p') = d$, q'_1 der Punkt

auf A mit $d(P(q_0), p'_1) = 2\sqrt{d^2 \Leftrightarrow h^2}$.

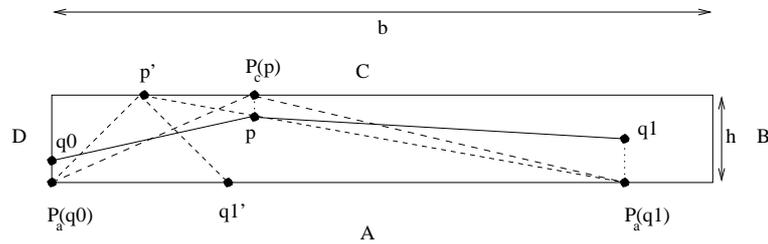


Abbildung 3.38: Die Kostenfunktion in p ist beschränkt

Aus dem Monotonieverhalten der Kostenfunktion folgt:

$$C(p) = C(q_0, p, q_1) \leq C(P_a(q_0), P_c(p), P_a(q_1)).$$

Aus den Eigenschaften 1. und 2. der Kostenfunktion folgt:

$$C((P_a(q_0), P_c(p), P_a(q_1)) \leq C((P_a(q_0), p', P_a(q_1)) \leq C((P_a(q_0), p', q'_1) = c_{max} \quad \blacksquare$$

Die folgenden Betrachtungen sind durch die Frage motiviert, welche Eigenschaften die Stufen des Evolutionsverfahrens aufweisen, wenn die Eckpunkte des Ausgangspolygons eine digitale 8-Kurve bzw. ein digitales 8-Segment beschreiben.

Gegeben sei ein digitales 8-Segment $\mathcal{D} \subset \mathbb{Z}^2$

Definition:

Ein Punkt $P \in \mathcal{D}$ heißt *Stufenwechsel*, falls er entweder Anfangs- bzw. Endpunkt des Segments ist, oder einen Nachbarpunkt besitzt, der nicht 4-Nachbar ist.

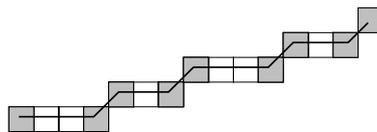


Abbildung 3.39: Digitales Segment mit Stufenwechseln und zugehörigem Stufenpolygon

$\mathcal{W} = (p_0, \dots, p_n)$ ist diejenige geordnete Menge aller Stufenwechsel, die die Reihenfolge der Stufenwechsel in \mathcal{D} beachtet.

Ein Polygon $\mathcal{Q} = (q_0, \dots, q_n) \subset \mathbb{R}^2$ heißt *Stufenpolygon* zu einem digitalen Segment \mathcal{D} , falls gilt: $\forall i \in (0, \dots, n): q_i \in \mathcal{Q} \Leftrightarrow q_i = p_i \in \mathcal{W}$.

Bemerkung:

Ist ein Polygon \mathcal{P} durch alle Punkte eines digitalen Segments \mathcal{D} definiert, so existiert immer eine Evolutionsstufe \mathcal{P}^k , die genau das Stufenpolygon \mathcal{Q} zu \mathcal{D} darstellt, da die Kostenfunktion genau in $\mathcal{D} \setminus \mathcal{W}$ den Wert 0 annimmt, diese Punkte von der Evolution also als erste entfernt werden.

Satz: Beschränktheit der Kostenfunktion eines Stufenpolygons

Gegeben sei ein zu einem digitalen Segment zugehöriges Stufenpolygon \mathcal{Q} .

Dann gilt:

Die Kosten der in allen Evolutionsstufen $\mathcal{Q}^0, \dots, \mathcal{Q}^m$ entfernten Punkte sind nach oben durch $\frac{\pi}{2\sqrt{2}}$ beschränkt.

Beweis:

O.B.d.A verlaufe \mathcal{Q} in Oktant 0.

Unterscheidung in 3 Fälle:

1. $\#\mathcal{Q} = 2$
 \mathcal{Q} besteht nur aus Endpunkten, es gibt keine weiteren Evolutionsstufen
 \Rightarrow Aussage gilt.
2. $\#\mathcal{Q} > 2$, minimale Kantenlänge $d > 1$
 Da \mathcal{Q} Stufenpolygon, gilt: $d > 1 \Rightarrow d \geq \sqrt{2}$.
 \exists Rechteck $\mathcal{R} \subset \mathbb{R}^2$ mit Höhe h , Breite b , $b \geq 1 = h$ welches \mathcal{Q} umschließt, also $\mathcal{Q} \subset \mathcal{R}$. Für alle Tripel (q_0, p, q_1) von Eckpunkten aus \mathcal{Q} , welche die Ordnung von \mathcal{Q} beachten (und nur solche Tripel sind während der Evolution möglich), folgt aus dem Satz über die Beschränktheit der Kostenfunktion im Rechteck:
 $\forall d \geq \sqrt{2} : C(p) \leq d \arcsin \frac{h}{d} \leq \frac{\pi}{2\sqrt{2}}$.
3. $\#\mathcal{Q} > 2$, minimale Kantenlänge $d = 1$
 Es existiert ein Rechteck wie in Fall 2.
 Betrachtet werden wiederum die Kosten aller Tripel (q_0, p, q_1) , welche die Ordnung von \mathcal{Q} beachten. Gilt $\min\{|\overline{q_0p}|, |\overline{pq_1}|\} > 1$, so gilt der Beweis aus Fall 2. Es sei also $\min\{|\overline{q_0p}|, |\overline{pq_1}|\} = 1$, o.B.d.A $|\overline{q_0p}| = 1$. Der maximale Wendewinkel in p beträgt in diesem Fall $\frac{\pi}{4}$.

Aus dem Monotonieverhalten der Kostenfunktion folgt, daß die maximalen Kosten $C_{max}(p)$ unter dem maximalen Wendewinkel $\frac{\pi}{4}$ und der maximalen

Länge $|\overline{pq_1}| \rightarrow \infty$ gegeben sind, also durch:

$$C_{max}(p) = \lim_{|\overline{pq_1}| \rightarrow \infty} \frac{\pi |\overline{pq_1}|}{4(|\overline{pq_1}|+1)} < \frac{\pi}{2\sqrt{2}}.$$

■

Im folgenden wird das Verhalten des Evolutionsverfahrens in Polygonzügen untersucht, die zu zwei sich schneidenden digitalen 8-Segmenten gehören.

Definition: digitale Schnittkurve

Eine *digitale Schnittkurve* \mathcal{S} ist eine digitale 8-Kurve, die durch Δ_4 -Diskretisierung zweier Strecken $S_1 = \overline{s_1s}$, $S_2 = \overline{s_2s} \subset \mathbb{R}^2$ mit gemeinsamem Endpunkt s entstanden ist.

$\mathcal{S}_1, \mathcal{S}_2 \subset \mathcal{Q}$ seien die zwei digitalen Segmente maximaler Länge, die die Endpunkte $\Delta_4(s_1), \Delta_4(s_2) \in \mathcal{S}$ enthalten (siehe Abbildung 3.40 a).

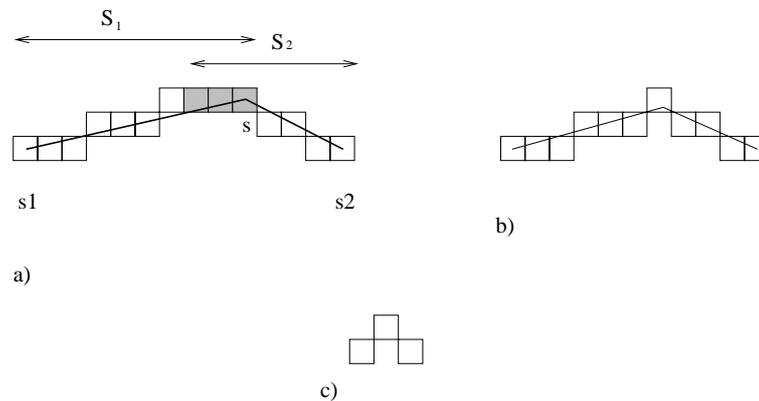


Abbildung 3.40: Schnitt von 2 digitalen Segmenten

Die Schnittmenge zweier solcher digitaler 8-Segmente ist i.a. kein einzelner Punkt, sondern wiederum ein digitales 8-Segment, welches mehr als einen Punkt enthält. Die Einteilung von Schnittkurven in digitale Segmente ist daher nicht notwendig eindeutig, sondern startpunktabhängig. Um den daraus entstehenden Schwierigkeiten zu entgehen, seien zunächst nur Schnittkurven \mathcal{S} betrachtet, für welche $\#(\mathcal{S}_1 \cap \mathcal{S}_2) = 1$, die also einen eindeutigen *Schnittpunkt* p besitzen (Abbildung 3.40 b). 3.40 c) zeigt die einzige Punktkombination (modulo Rotation um $n\frac{\pi}{2}$), in welcher p auftreten kann.

Anmerkung: p und seine Nachbarn sind Stufenwechsel.

In allen folgenden Betrachtungen liege p im Ursprung des Koordinatensystemes, die Oktanten seien wie üblich (gg. den Uhrzeigersinn, Oktant 0 beschreibt Winkel aus $[0, \frac{\pi}{4})$) numeriert.

Satz: Gegeben sei eine Schnittkurve \mathcal{S} mit eindeutigem Schnittpunkt p im Ursprung des Koordinatensystemes. $\mathcal{Q} = (q_0, \dots, q_{i1}, p, q_{i2}, \dots, q_n)$ sei das zu \mathcal{S} zugehörige Stufenpolygon.

Dann gilt:

Es gibt eine Evolutionsstufe \mathcal{Q}^m mit $\mathcal{Q}^m = (q_0, p, q_n)$, d.h. die Evolution erhält exakt den Schnittpunkt p .

Beweis:

Fallunterscheidung:

1. \mathcal{Q} verläuft in einem oder zwei benachbarten Oktanten.

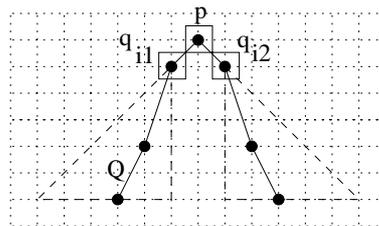


Abbildung 3.41: Die Dreiecke unter q_{i1} und q_{i2} begrenzen die möglichen Fortführungen von \mathcal{Q} in den Oktanten 5,6

Es sei $\mathcal{T} = \{(q_i, p, q_j) | q_i, q_j \in \mathcal{Q}, i \leq i1, j \geq i2\}$ die Menge aller Tripel um p , die die Ordnung in \mathcal{Q} erhalten. Es gilt:

$$\begin{aligned} \min_i(|\overline{q_i p}|) &= |\overline{q_{i1} p}|, \\ \min_j(|\overline{p q_j}|) &= |\overline{p q_{i2}}| \text{ und} \\ \min_{i,j}(\angle(\overline{q_i p}, \overline{p q_j})) &= \angle(\overline{q_{i1} p}, \overline{p q_{i2}}) \\ \Rightarrow \min_{i,j} C(q_i, p, q_j) &= C(q_{i1}, p, q_{i2}) = \frac{\pi}{2\sqrt{2}} \end{aligned}$$

Es seien $\mathcal{Q}_1 = (q_0, \dots, p) \subset \mathcal{Q}, \mathcal{Q}_2 = (p, \dots, q_n) \subset \mathcal{Q}, \mathcal{T}_i = (t_1, t, t_2) \subset \mathcal{Q}_i, i = 1, 2$ die Menge aller geordneten Tripel aus \mathcal{Q}_i , welche die Ordnung in \mathcal{Q}_i erhalten.

\mathcal{Q}_1 sowie \mathcal{Q}_2 lassen sich von einem Rechteck mit Höhe h , Breite $b, b > h \geq 1$ umschließen, somit gilt durch den Satz über die Beschränktheit der Kostenfunktion eines Stufenpolygons:

$$\forall (t_1, t, t_2) \in \mathcal{T}_{1,2}: C(t_1, t, t_2) < \frac{\pi}{2\sqrt{2}} = C(q_{i1}, p, q_{i2}) \Rightarrow \text{die Evolution erhält } p.$$

2. \mathcal{Q} verläuft in 2 Oktanten die einen Winkel von $\frac{\pi}{2} < w \leq \frac{3\pi}{4}$ bilden, o.B.d.A Oktant 5 und 7 (siehe Abbildung 3.42).

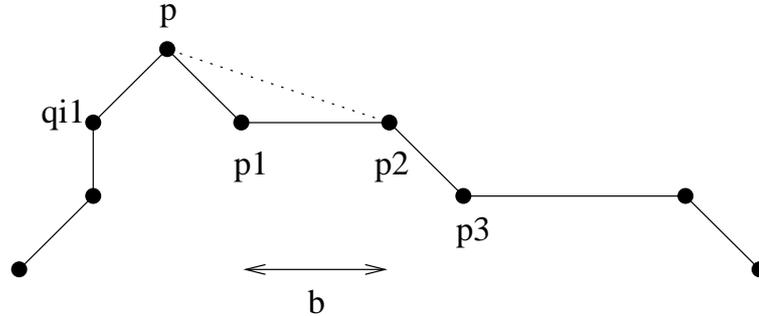


Abbildung 3.42:

Es sei $\mathcal{T} = \{(q_i, p, q_j) | q_i, q_j \in \mathcal{Q}, i \leq i_1, j \geq i_2\}$ die Menge aller Tripel um p , die die Ordnung in \mathcal{Q} erhalten. Es gilt:

$\forall t = (t, p, q), t' = (t', p, q) \in \mathcal{T}: |\overline{t_1 p}| = \sqrt{2} \Rightarrow C(t, p, q) \leq C(t', p, q)$, d.h. zur Untersuchung einer unteren Kostenschranke in p kann ein Nachbar q_{i1} von p auf $(\Leftrightarrow 1, \Leftrightarrow 1)$ festgelegt werden.

Die Punkte seien wie in Abbildung 3.42 bezeichnet, $\mathcal{Q}_2 = (p, \dots, q_n) \subset \mathcal{Q}$ sei das Stufenpolygon in Oktant 7.

Solange nicht p_1 entfernt wird, bleibt p erhalten, da die Kosten $C(q_{i1}, p, p_1)$ bzgl. der Kosten in einem Stufenpolygon maximal sind.

Wird p_1 entfernt und ist p_2 noch vorhanden, so wird p_2 im nächsten Schritt entfernt, da für alle Breiten b (siehe Abbildung 3.42) und für alle Punkte $q \in \mathcal{Q}_2 \setminus \{p, p_1, p_2\}$ gilt: $C(q_{i1}, p, p_1) > C(p, p_2, q)$ sowie $C(q_{i1}, p, p_2) > C(p, p_2, q)$.

Sind p_1 und p_2 nicht mehr vorhanden, bleibt p erhalten, da für alle Punkte $q \in \mathcal{Q}_2 \setminus \{p, p_1, p_2\}$ gilt: $C(q_{i1}, p, q) > \frac{\pi}{2\sqrt{2}}$.

3. \mathcal{Q} verläuft in 2 Oktanten, welche einen Winkelbereich von $\frac{\pi}{2} \leq w \leq \pi$ überstreichen, o.B.d.A Oktanten 4 und 7.

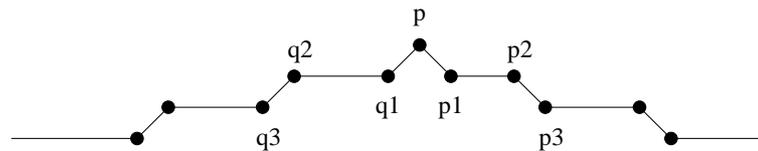


Abbildung 3.43:

Es seien $\mathcal{Q}_1 = (q_0, \dots, p) \subset \mathcal{Q}, \mathcal{Q}_2 = (p, \dots, q_n) \subset \mathcal{Q}$, o.B.d.A. verlaufe \mathcal{Q}_2 in Oktant 7, und es gelte mit den Bezeichnungen aus Abbildung 3.43 $|\overline{q_1 q_2}| \geq |\overline{p_1 p_2}|$.

Solange q_1 nicht entfernt wurde, bleibt p analog zu dem Beweis aus Fall 2 erhalten. Dieser Beweis zeigte ebenfalls, daß mit p_1 auch immer p_2 entfernt wird. Zusätzlich gilt, daß p_1 (und damit p_2) vor q_1 gelöscht werden, da $|\overline{q_1 q_2}| \geq |\overline{p_1 p_2}|$. Betrachtet werden muß also die Konstellation, in welcher q_1 entfernt wird. In dieser Konstellation wird p immer benachbart von q_1 sowie von einem Punkt $p' \in \mathcal{Q}_2 \setminus \{p, p_1, p_2\}$ (siehe Abbildung 3.44).

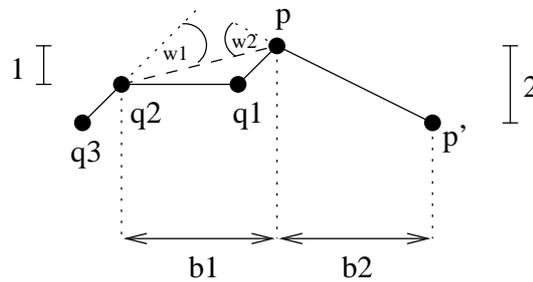


Abbildung 3.44:

Analog zum vorherigen Fall kann gezeigt werden, daß im Schritt nach der Entfernung von q_1 auch q_2 (so noch vorhanden) entfernt wird. Durch Entfernung von q_1 , dessen Kosten minimal waren, ändern sich die Kosten in p und q_2 . Es gilt (Eigenschaft 1 der Kostenfunktion):

$\forall q' \in \mathcal{Q}_1 \setminus \{q_3, q_2, q_1, p\}: C(q', q_2, p) < C(q_3, q_2, p) < C(q_3, q_2, q_1)$, d.h. falls $C(q_3 q_2, p) < C(q_2, p, p')$, wird im Folgeschritt q_2 entfernt. Die Kostenungleichung lautet:

$$w_1 \frac{l_1 l_2}{l_1 + l_2} < w_2 \frac{l_2 l_3}{l_2 + l_3} \text{ mit}$$

$$l_1 = \sqrt{2}, l_2 = \sqrt{1 + b_1^2}, l_3 = \sqrt{4 + b_2^2}, w_1 = \frac{\pi}{4} \Leftrightarrow \arctan \frac{1}{b_1}, w_2 = \arctan \frac{1}{b_1} + \arctan \frac{2}{b_2}$$

Aus der Voraussetzung $|\overline{q_1 q_2}| \geq |\overline{p_1 p_2}|$ folgt $b_1 \geq \frac{b_2}{2}$, mit dieser Voraussetzung ist die Ungleichung erfüllt.

\Rightarrow Mit q_1 wird auch q_2 entfernt, die entstehende Evolutionsstufe enthält das Teilpolygon (q', p, p') , wobei $q' \in \mathcal{Q}_1 \setminus \{q_2, q_1, p\}, p' \in \mathcal{Q}_2 \in \{p, p_1, p_2\}$. Die minimale Höhe eines (q', p, p') umfassenden Rechtecks ist 2, somit gilt für die Kosten in p : $C(p) > 2 > \frac{\pi}{2\sqrt{2}}$

$\Rightarrow p$ bleibt erhalten. ■

Der obige Beweis hat also gezeigt, daß, so eine *eindeutige* Zerlegung eines digitalen Polygons in digitale Strecken existiert auch eine Evolutionsstufe existiert, die genau aus den eindeutigen Schnittpunkten dieser digitalen Strecken besteht.

Eine derartige Formulierung für den nicht eindeutigen Fall kann nicht gefunden werden, hier müssen Näherungsaussagen gemacht werden, die aufzeigen, daß sich das Evolutionsverfahren nicht 'beliebig weit' von der Darstellung durch digitale Strecken entfernen kann. Hilfreich dabei sind die Betrachtungen über die Steigkeit des Verfahrens, eine genaue Formulierung sowie ein zugehöriger Beweis stehen noch aus. Die zahlreichen Experimente zeigen jedoch, daß auch im nichteindeutigen Fall ein starker Zusammenhang zwischen dem Ergebnis des Evolutionsverfahrens und der Extraktion von digitalen Strecken besteht.

3.10 Vergleich mit anderen Verfahren

Die digitale Polygonevolution ist ein Verfahren zur Vereinfachung einer gegebenen Randkurve: sie ist ein Ansatz zur Reduzierung der in der Randkurve enthaltenen Information, erwachsen aus der Forderung nach

- Reduzierung des Einflusses von Störungen und
- Erkennung relevanter Formeigenschaften.

Damit läßt sie sich in drei verschiedene Kategorien von Techniken der Bildverarbeitung einordnen:

1. Polygonale Approximation
2. Suche nach wesentlichen Merkmale (critical point detection)
3. Scale Space Evolution

Ziel der polygonalen Approximation ist es, zu einem gegebenen Polygon P , ein Teilpolygon $P' \subset P$, zu finden, welches ein Näherungskriterium (z.B. Hausdorffabstand) erfüllt. Es existieren zahllose Ansätze zur Polygonapproximation, genannt seien Ramer [33], Wall and Danielsson [38], Kurozumi und Davis [17].

Auf die Polygonapproximation soll hier nicht näher eingegangen werden, das Evolutionsverfahren soll aber in Vergleich zu typischen Vertretern der beiden anderen Kategorien, besonders der Scale Space Evolution gestellt werden.

3.10.1 Critical Point Detection

Als '*critical points*' werden formbestimmende oder dominante Punkte (z.B. Eckpunkte, Wendepunkte, Punkte maximal negativer Krümmung) bezeichnet. Sie können als Grundlage der Formbeschreibung dienen. Im Unterschied zur polygonalen Approximation liegt die Aufgabenstellung bei der Suche nach dominanten Punkten in der möglichst exakten Lokalisierung dieser Punkte, die Approximation dagegen zielt auf visuelle Ähnlichkeit.

Die meisten Ansätze dieser Kategorie basieren auf Krümmungsabschätzungen, begleitet von den damit zusammenhängenden Nachteilen:

- Die digitale Geometrie kennt keinen praktisch einsetzbaren Begriff der Krümmung, d.h. Krümmung in digitalen Kurven ist eine lokale Eigenschaft, die aufgrund ihrer Störungsanfälligkeit praktisch unbrauchbar ist. Daher wird zur Berechnung der Krümmung die digitale Natur der Vorlagen zugunsten mathematisch bekannterer Handhabbarkeit verlassen: es werden kontinuierliche Modelle erstellt.
- Auch im kontinuierlichen Fall ist die Bestimmung der Krümmung eine lokale Operation. Zur Dämpfung der typischen Störungseinflüsse werden die Randkurven daher zunächst mit (linearen) Glättungsfilttern bearbeitet, benutzt werden zumeist Gauß-Filter (z.B. Asada und Brady [11]).
- Die entscheidenden Nachteile solcher Glättungsfiltter:
 1. Sie benötigen einen festen Parameter (z.B. Gaußfilter: die Breite der Gaußregion zur Faltung), die Bestimmung dieses Parameters ist auflösungsabhängig, das Verfahren ist nicht adaptiv (oder anders ausgedrückt: bei festem Parameter ist die Filterung nicht invariant gegenüber Skalierungen).
 2. Die Punkte der Randkurve werden durch die Filterung umpositioniert.
 3. Die Filterung bedeutet erheblichen Rechenaufwand.

(Auf diese Eigenschaften wird im nächsten Kapitel (Scale Space Evolution) näher eingegangen).

Wie die vorhergehenden Kapitel gezeigt haben, existieren diese Nachteile bei der Polygonevolution nicht.

Zhu und Chirlian stellen in [42] einen nichtlinearen Ansatz vor, der ohne Abschätzung der Krümmung arbeitet und der Kurvenevolution ähnlich ist. Es wird jedem

Randpunkt ein Relevanzmaß zugeordnet, welches vom Punkt selbst und seinen zwei Nachbarn abhängig ist: benutzt wird die Höhe des aufgespannten Dreiecks. Als kritische Punkte gelten alle Punkte, deren Relevanzmaß oberhalb eines Schwellwertes \mathcal{T} liegt.

Die Unterschiede zur Polygonevolution mit Stufenausgleich sind:

- Die Polygonevolution mit Stufenausgleich ist stetig, wie Kapitel 3.7 zeigt. Die notwendigen Voraussetzungen an die Kostenfunktion (siehe Kapitel 3.7, Abbildung 3.26) sind bei den von Zhu und Chirlian verwendeten Kosten, nämlich der Dreieckshöhe, nicht gegeben (verletzt ist Forderung 3).
- Die Randpunkte werden bei Zhu und Chirlian einmalig in der Ausgangskurve mit dem Relevanzmaß belegt. Die dadurch festgelegte Information enthält nicht notwendig Aussagen über die visuelle Relevanz, wie die Abbildungen 3.4 und 3.16 gezeigt haben. Die Polygonevolution dagegen erlaubt eine Neubelegung der Punkte mit ihrem Relevanzmaß bei veränderter Nachbarschaftskonstellation.
- Bei Zhu und Chirlian ist der Schwellwert \mathcal{T} festgelegt, d.h. er muß als zusätzlicher Parameter angegeben werden. Die Polygonevolution dagegen benötigt kein solches Abbruchkriterium.

3.10.2 Scale Space Evolution

Eine klassische, kontinuierliche Variante der Polygonevolution findet sich bei Mokhtarian [31]. Die Familie der Evolutionskurven γ^t einer gegebenen kontinuierlichen Jordankurve γ^0 wird durch Faltung von γ^0 mit Gaußkernen verschiedener Größe $0 < t < \infty$ errechnet, siehe Abbildung 1.6. Gekoppelt an diese Evolution ist eine Formbeschreibung, welche wiederum zu einer Critical-Point-Detection führt: die Form wird durch die Veränderung der Lage der Wendepunkte auf der Randkurve beschrieben. Die daraus resultierenden Eigenschaften werden in Kapitel 5 näher beleuchtet, hier sollen zunächst die Eigenschaften der kontinuierlichen Evolution in Vergleich zur Evolution mittels Stufenausgleich gesetzt werden. Dabei sollen Unterschiede im praktischen Einsatz nicht ausgeschlossen werden, da beide Verfahren im Rahmen der MPEG-7 Standardisierung in einem sehr praxisbezogenen Rahmen im Vergleich stehen.

- Der Rechenaufwand des Verfahrens von Mokhtarian ist erheblich höher, da die kontinuierliche Modellierung nicht der digitalen Natur der Vorlage entspricht. Die theoretisch bewiesenen Ergebnisse müssen in der Praxis durch

beliebig feine Diskretisierung der kontinuierlich angenommenen Randkurve ermittelt werden. Die Berechnungen der Gaußfaltung sind aufwendiger (abhängig von der betrachteten Umgebung) als der Stufenausgleich (pro Evolutionsstufe zwei (Nachbar-)Punkte).

- Das kontinuierliche Verfahren endet in einem Kreis, d.h. alle Punkte erhalten äquivalente Information. Der Stufenausgleich dagegen erzeugt ein Dreieck, welches bedeutende Punkte der Randkurve auszeichnet. Der praktische Vorteil: die Evolution erzeugt direkt eine Lageinformation über das Objekt, welche für den anschließenden Formvergleich von Bedeutung ist. Diese Information ergibt sich direkt auf dem Polygon, ohne eine Transformation des Polygons in eine andere Darstellung. Dazu Weickert [39], S.6:

Gaussian smoothing does not only reduce noise, but also blurs important features such as edges and, thus, makes them harder to identify. Since Gaussian scale-space is designed to be completely uncommitted, it cannot take into account any a-priori information on structures which are worth being preserved (or even enhanced).

- Das kontinuierliche Verfahren erhält nicht die Position von Randpunkten, der Rand wird global geglättet, indem Punkte versetzt werden. Eine Korrespondenz der Punkte zwischen Darstellungen t_1, t_2 ist schwierig. Diese 'Verwischung' ist ein grundsätzliches Problem dieser Art von Evolutionsverfahren. Dazu wiederum Weickert [39]:

Diffusion dislocates features when moving from finer to coarser scales. So features identified at a coarse scale do not give the right location and have to be traced back to the original image [40]. In practice, relating dislocated information obtained at different scales is difficult and bifurcations may give rise to instabilities. These coarse-to-fine tracking difficulties are generally denoted as the correspondence problem.

- Das kontinuierliche Verfahren glättet global, d.h. auch falls nur eine Teilkurve gestört ist, führt die Glättung zur Abrundung der gesamten Kurve. Die Adaptivität verhindert solches Verhalten bei der Polygonevolution mit Stufenausgleich.
- Das kontinuierliche Verfahren benötigt einen festgelegten Abbruchparameter, d.h. eine durch die Breite t festgelegten Schwellwert.

3.11 Weitere Einsatzgebiete der diskreten Kurvenevolution

Die Polygonevolution wurde im Rahmen dieser Arbeit als Vorfilter für den Formvergleich entwickelt. Die Eigenschaften der Evolution jedoch machten dieses Verfahren für verschiedenste Einsätze interessant, von denen hier kurz drei Anwendungen vorgestellt werden sollen.

Die folgenden Anwendungen der Evolution werden nicht vom Autor dieser Arbeit bearbeitet, es soll lediglich ein Ausblick auf die übergreifenden Möglichkeiten der Evolution gegeben werden.

3.11.1 Signalverarbeitung

Es sei $x(n)$ ein diskretes eindimensionales Signal, also eine Folge $x_0, x_1, \dots \subset \mathbb{R}$. Diese Folge kann als Treppenfunktion

$$S : \mathbb{R} \rightarrow \mathbb{R}, \quad x \rightarrow x_i, \quad \Leftrightarrow x \in [i, i + 1)$$

dargestellt werden. Auf diese Treppenfunktion kann nun, analog zur Evolution im Tangentenraum, der Stufenausgleich angewendet werden. Die Evolution kann hier als nichtlineares eindimensionales Filter $\mathcal{E}(S)$ angesehen werden. Experimente zeigen, daß typischerweise wesentliche Signalinformationen von Rauscheffekten getrennt werden können, siehe Abbildung 3.45.

Dabei hat das Filter \mathcal{E} zwei wesentliche Eigenschaften:

1. Es ist konservativ, d.h. $\int |S(x)|dx = \int |\mathcal{E}(S(x))|dx$
2. Es erhält steile Flanken im Signal, d.h. es zeigt sich nicht die für klassische lineare Filter typische Abrundung solcher Signaleigenschaften.

Die Evolution wurde erfolgreich (mit kleinen Anpassungen an die Anwendungsumgebung) für die Filterung von Signalen aus der Berechnung von Stoßkurven eingesetzt.

3.11.2 Vereinfachung von graphenähnlichen Strukturen: U-Bahnpläne

Ein breites Anwendungsgebiet automatischer Datenverarbeitung ist die Erfassung und Vereinfachung geographischer Daten. Typischer Einsatz sind vereinfachte

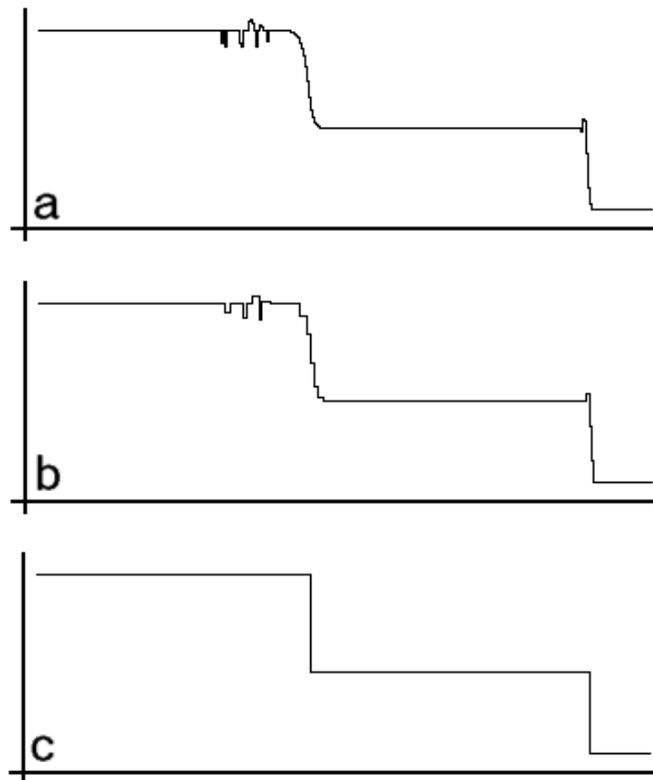


Abbildung 3.45: Evolution als nichtlineares Signalfilter: a) Ausgangssignal b) und c) verschiedene Evolutionsstufen. Wesentliche Signalinformationen, insbesondere steile Kanten bleiben erhalten.

Darstellungen von Verkehrsnetzen. Die Evolution wurde hier bei der Abstraktion von U-Bahnplänen eingesetzt.

Die Ausgangskurven sind dabei nicht wie bisher geschlossene Randkurven, sondern Netze polygonaler Strecken, deren Eckpunkte die Haltestellen darstellen.

Die Vereinfachung verläuft wiederum durch Polygonlinearisierung, allerdings werden die Punkte nicht aus dem Polygonnetz, sondern nur aus der Nachbarschaftsrelation entfernt (die Haltestellen sollen in der Ergebnisgrafik erhalten bleiben). Weiterhin wird die Anwendung durch topologische Restriktionen erweitert. Die Ergebnisse zeigt Abbildung 3.46.

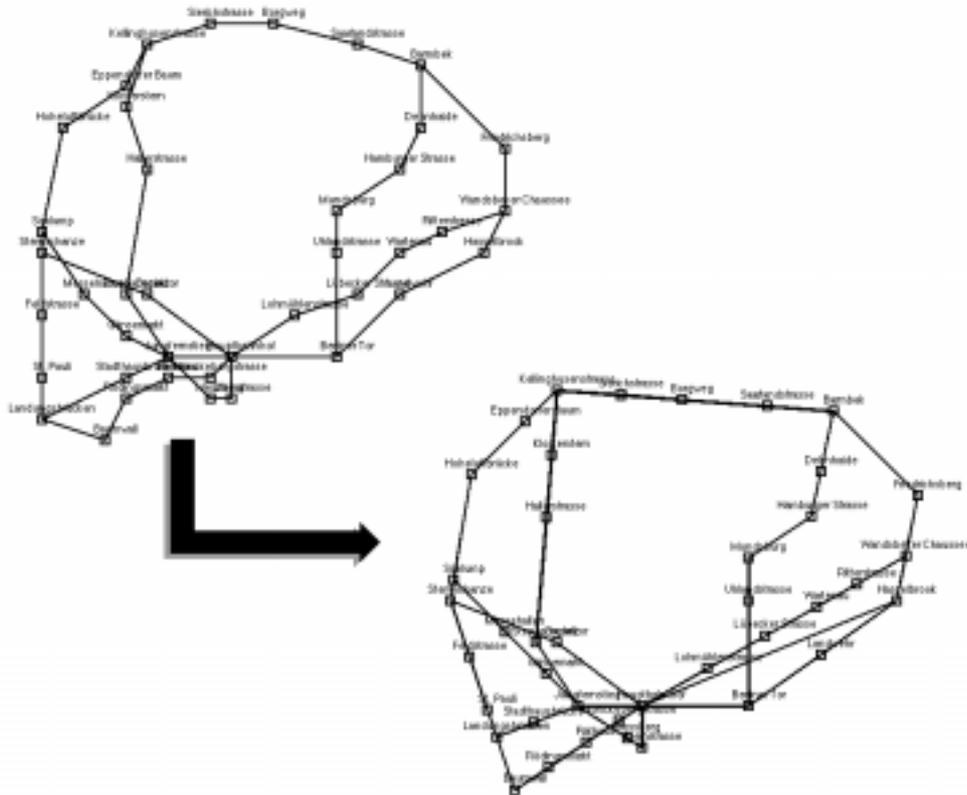


Abbildung 3.46: Abstraktion des Hamburger U-Bahnnetzes

3.11.3 Segmentierung zeitdiskreter mehrdimensionaler Signale

Die Polygonevolution wurde bisher, da sie nur im Rahmen der Verarbeitung 2-dimensionaler Bilder vorgestellt wurde, lediglich für polygonale Kurven $\gamma : \mathbb{Z} \rightarrow \mathbb{R}^2$ definiert. Die benutzten Parameter für den Stufenausgleich, nämlich Winkel und Länge, erlauben jedoch den Einsatz in beliebigdimensionalen polygonalen Kurven $\gamma : \mathbb{Z} \rightarrow \mathbb{R}^n$, $n \in \mathcal{N}$.

Hierdurch erschließt sich ein weites Feld im Gebiet der Verarbeitung zeitdiskreter Signale, z.B. bei Bewegtbildern (TV): jedes Einzelbild kann als Punkt in einem mehrdimensionalen Raum gedeutet werden, der zeitliche Ablauf bestimmt die Reihenfolge dieser Punkte und damit ein Polygon γ , wie oben definiert. Bei geeigneter Wahl der Bildparametrisierung können starke Änderungen zwischen aufeinanderfolgenden Einzelbildern zu markantem Verlauf des Polygons γ führen, so daß die durch diese Bilder bestimmten Punkte auf γ bei Abstraktion mittels der

Kurvenevolution erhalten bleiben. Das bedeutet: ähnliche ⁵ Filmszenen werden zusammengefaßt, indem die Evolution die zu den Bildern dieser Szene zugehörigen Punkte eliminiert, die Anfangs- und Endpunkte der Szenen aber bleiben erhalten. So ist es z.B. möglich, Videomaterial szenenbasiert zu indizieren.

Versuche mit einem solchen System wurden sehr erfolgreich in Zusammenarbeit mit dem *Center for Automatic Research, University of Maryland, USA* bearbeitet.

3.12 Ausblick

Kapitel 3.11.3 zeigte den erweiterten Einsatz der Evolution eindimensionaler Kurven in mehrdimensionalen Räumen, ebenso denkbar ist natürlich eine Erweiterung des vorgestellten Abstraktionsprinzips auf mehrdimensionale Kurven, z.B. Oberflächen in \mathbb{R}^3

Die Abstraktion verläuft hier analog zu dem beschriebenen 2-dimensionalen Fall: das Objekt liegt als 3d-Gitter vor, die Knotenpunkte werden bewertet und entfernt. Allerdings stellen sich zwei Probleme:

- Die Bestimmung der Kostenfunktion muß neu überdacht werden. Das Hauptproblem ist dabei, eine Entsprechung zur Bedeutung des Wendewinkels zu finden.
- Die durch die Entfernung eines Punktes entstehende Fläche ist i.a. nicht planar, eine Linearisierung wie im zweidimensionalen Fall ist also nicht möglich. Stattdessen muß die neu entstehende Fläche triangularisiert werden, dieser Vorgang ist jedoch nicht eindeutig. Die Bestimmung einer visuell sinnvollen Triangulation steht noch offen.

Erste Ergebnisse zeigt Abbildung 3.47. Interessant ist die gewaltige Datenkompression zwischen Bild a) und b): obgleich kein deutlicher Bildqualitätsverlust bemerkbar ist, ist die Datenmenge in Bild b) auf ca. 12 % reduziert (a): 34835 Punkte, b): 4196 Punkte).

⁵die Bedeutung der 'Ähnlichkeit' wird durch die Parametrisierung bestimmt

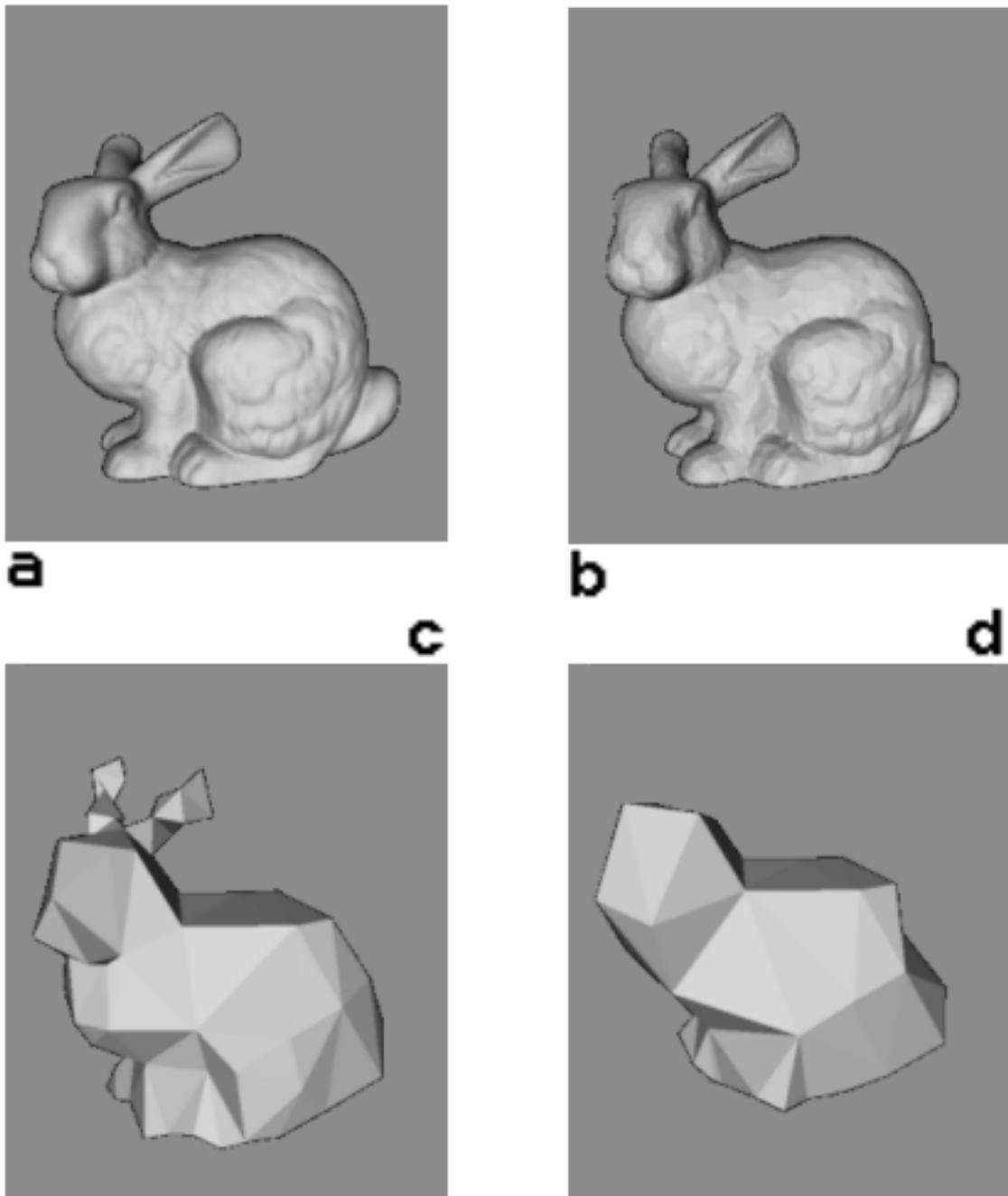


Abbildung 3.47: Evolution einer 3d-Gitterstruktur. Anzahl der Gitterpunkte: a) 34835 b) 4196 c) 290 d) 244

Kapitel 4

Formteile

4.1 Einleitung und Motivation

Ein wichtiger Schritt zur Erkennung von Formen ist die Unterteilung in *visuell bedeutsame Teile*. Diese Vorgehensweise ist vor allem in natürlichem Kontext sinnvoll, wie Hoffman und Singh [14] durch zwei grundlegende Beobachtungen zeigen:

1. Objekte sind auch bei *teilweiser* Überdeckung erkennbar
2. Objekte sind i.a. nicht starr und können auch unter *teilweiser* Verformung erkannt werden (z.B. Beinstellung bei laufendem Tier)

Der Nachteil einer Behandlung einer Form als Ganzes, ohne Unterteilung, läßt sich ebenso einfach als Folge der Umkehrung der obigen Beobachtungen beschreiben:

Ist die Gesamtinformation mindestens zum Teil überdeckt / verzerrt / gestört, ist das Objekt nicht erkennbar. Ein interessantes Beispiel für ein solches Verfahren, das eben bei nur teilweiser Störung versagt, ist in Arkin [1] gegeben. Es wird dort ein globales Vergleichsmaß zur Erkennung der Übereinstimmung von polygonalen Randkurven vorgestellt. Dieses Maß versagt insbesondere dann, wenn die Randinformation *nicht gleichverteilt* gestört ist, d.h. eine über den gesamten Rand verteilte Störung wirkt sich auf das Verfahren weniger negativ aus als eine (optisch geringere) teilbeschränkte Störung. Kapitel 5 wird auf dieses Verfahren und eine auf Formunterteilung basierte Anpassung näher eingehen.

Eine Formunterteilung sollte einigen Mindestanforderungen genügen. Dazu Hoffman [14]:

'The Parts must be computable (else they cannot be obtained), defined on any shape (else they will not help us recognize some shapes), and invariant under generic perturbations of viewpoint (else we shall see new parts each time we move, which would defeat indexing). These principles suggest that to define the part boundaries of a shape we should use its intrinsic geometry.'

Insbesondere die Forderung nach Allgemeingültigkeit der Teilbeschreibung, die auch Ziel der vorliegenden Arbeit ist, legt die Benutzung solcher 'intrinsisch geometrischer' Eigenschaften nahe. Es scheidet damit die Benutzung von Verfahren aus, die darauf zielen, die Form vorgegebenen (einfachen geometrischen) Grundkörpern anzupassen, da durch die Einschränkungen an diese Körper keine allgemeingültige Formbeschreibung erzielt werden kann. Literaturangaben modellbasierter Formbeschreibungen sowie eine weitergehende Diskussion finden sich in [14], S.31.

Die Modellierung objekteigener Merkmale verfolgt zwei Ansätze, die Unterteilung nach Flächen- oder nach Randeigenschaften ¹.

Vertreter der Flächenunterteilung arbeiten typischerweise über Skelettierung, z.B. klassische Skelettierung bei Held und Abe [12], oder weitergehend 'Shock Graphs' bei Siddiqi et al. [35]. Dabei werden im wesentlichen Engstellen (oder physikalisch: Sollbruchstellen) der Form gesucht. Eine Begründung wird Siddiqi und Kimia mit der Untersuchung von 'Necks' in [34], [36], [37] gegeben. Abbildung 4.1 zeigt ein Beispiel.

Die vorliegende Arbeit unterteilt die durch ein Polygon gegebene Form nach Randmerkmalen. Diese Vorgehensweise wurde insbesondere motiviert durch die Darstellung von 'Limbs' in [34], [37] sowie durch psychologische Experimente von Hoffmann und Richards in [13] und die Abhandlung über Formteile in Hoffman und Singh [14]. Die dort vorgestellten Modelle unterteilen Randkurven anhand ihres Krümmungsverhaltens. Die Motivation soll hier kurz dargestellt werden, da die aus ihr entspringende Formunterteilung für das weitere Vorgehen dieser Arbeit, insbesondere den Formvergleich (Kapitel 5), entscheidend ist:

Die Benutzung starker Krümmungen, genauer starker *konkaver* Krümmungen, wird in [13] empfohlen. Dort wird zunächst eine Beobachtung zur Krümmungsstärke aus den Eigenschaften der Projektion dreidimensionaler Objekte in die

¹Diese Unterteilung ist praxisbasiert: theoretisch enthalten Rand- und Flächendarstellung dieselbe Information, da sie ineinander transformierbar sind. Wesentlich für die Unterscheidung ist nur, welche Informationen in den verschiedenen Darstellungen implizit und welche explizit enthalten sind.

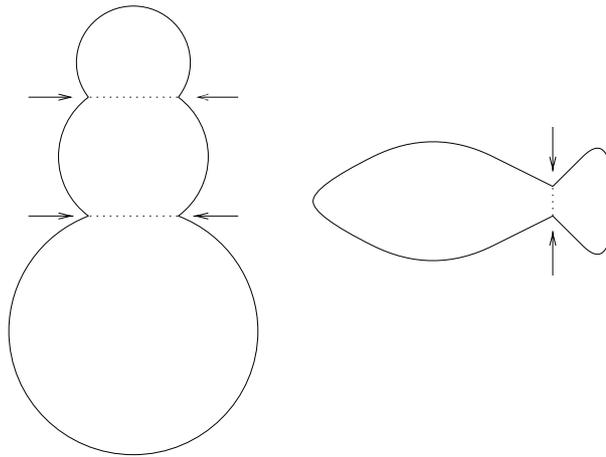


Abbildung 4.1: Aufteilung von Formen durch *necks* nach Siddiqi et. al. [34]

zweidimensionale Darstellung hergeleitet, siehe Abbildung 4.2: die Projektion erzeugt bei Überschneidung der Teile eine starke konkave Krümmung.

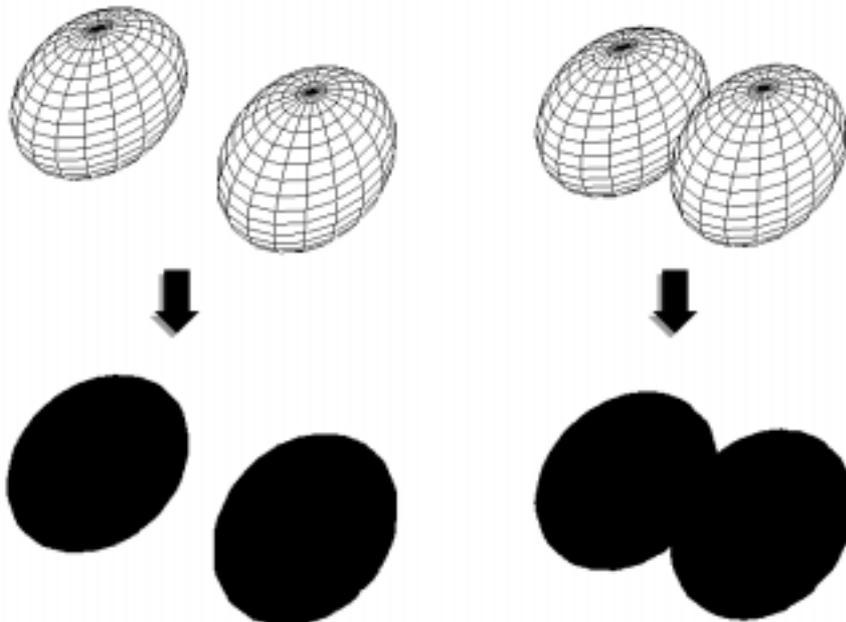


Abbildung 4.2: Starke konkave Krümmung des Randes, erzeugt durch Projektion

Die besondere Rolle *konkaver* Krümmungen wird in [13] anhand zahlreicher Bildexperimente hervorgehoben, für die stellvertretend Abbildung 4.3 stehen soll. Sie zeigt ein für derartige Experimente typisches 'Kippbild': die Treppe ist von

oben, als normal begehbar oder in einer nach gängigen Gravitationsrichtlinien nicht begeharen Form gekippt zu sehen. Die visuelle Unterteilung in die Treppe ist abhängig von der jeweiligen Deutung, es wird jedoch immer an (bezogen auf die Deutung) *konkaven* Ecken unterteilt. Die Unterteilung in 'stoffliche' Elemente (die Stufen) scheint natürlicher als diejenige in 'nichtstoffliche' (die Aussparungen).

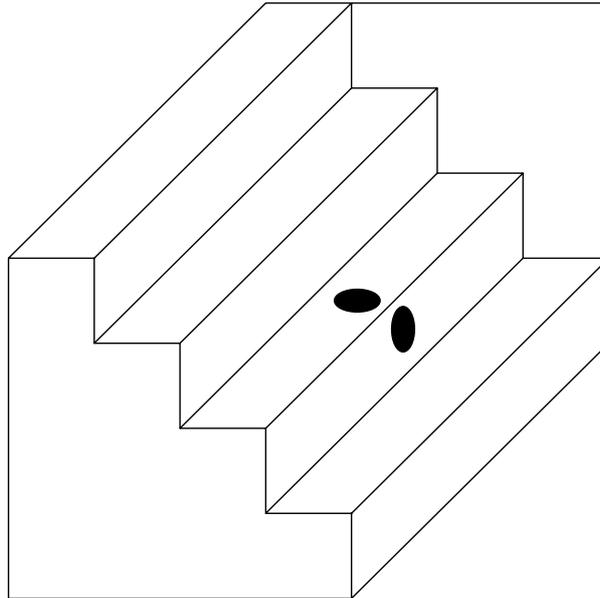


Abbildung 4.3: Die Schröder-Treppe: die Unterteilung in Stufen erfolgt in beiden Sichtarten an konkaven Ecken

Diese Beobachtungen führen zur *Minima rule for silhouettes* [13]:

'For any silhouette, all negative minima of curvature of its bounding curve are boundaries between parts.'

Ein Beispiel dazu zeigt Abbildung 4.4.

Eine weitergehende Eigenschaft des Krümmungsverhalten ist in Siddiqi et al. [34] beobachtet, sie stellt einen Zusammenhang zwischen der Form eines Objektes und der Funktion seiner enthaltenen Teilobjekte dar:

'An object's two-dimensional form is influenced by both the nature of its interaction with other objects and the nature of its projection onto the retinal image. First, as a result of this interaction, biological and man-made entities, whether by evolution or by design, give rise

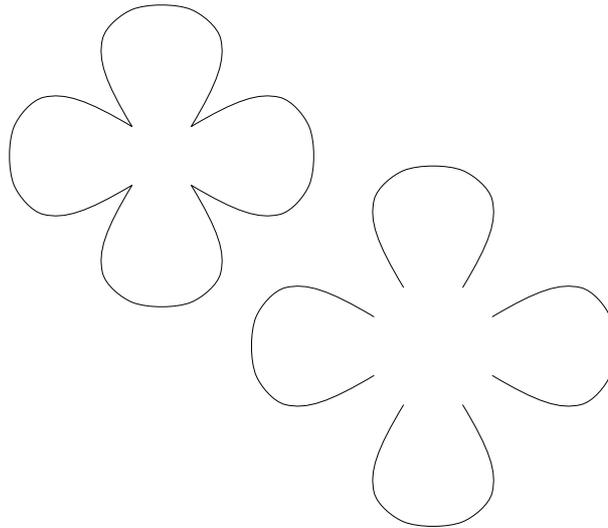


Abbildung 4.4: Aufteilung der Randkurve an konkaven Minima nach der *minima rule*

to parts which specialize in their function. When an object specializes function independently in two different, but connected regions, the result is often a sharp change in the three-dimensional surface of the object, e.g., the join between the beak of a bird and its head. The projection of these sharp changes yield a pair of high curvature points, where the tangent at one point smoothly continues to the tangent at the other point. This is a restatement of the Gestalt principle of 'good continuation',... '

Ein Beispiel dazu zeigt Abbildung 4.5: es werden zunächst Randpunkte minimaler Krümmung extrahiert. Die Unterteilung der Form geschieht durch Verbindung von Punkten mit 'natürlicher' Richtungsfortsetzung.

Zu beachten ist, daß hier, gefordert durch die Eigenschaft der natürlichen Fortsetzung, die gesamte Form, also die Formfläche anhand von Randeigenschaften zerschnitten wird, die Randpunkte werden durch sogenannte *part cuts* (Beusmans et al. [4]), also Schnittkanten verbunden.

Reine Randunterteilung – als Voraussetzung für die *part-cuts* oder auch als eigenständige Formbeschreibung – erhält man durch lokales Auffinden der Krümmungsminima.

Die Objekte, mit denen sich die Bildverarbeitung beschäftigt, sind meistens digitaler Natur, auch die Randpolygone, auf denen die vorliegende Arbeit operiert, entstammen digitalen Liniensegmenten. Die Unterteilung dieser Randpoly-

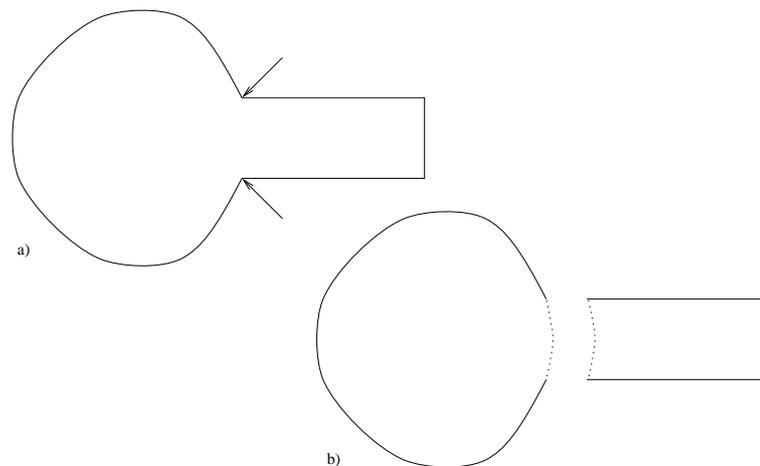


Abbildung 4.5: Aufteilung von Formen durch *limbs*, gegeben durch a) Randkonkavitäten und b) natürliche Richtungsfortsetzung

gone sollte robust sein, so das Ziel dieser Arbeit. Die Bestimmung von Punkten minimaler Krümmung ist, wie die Bestimmung aller Extrempunkte, eine lokale Berechnung und daher aufgrund der digitalen Natur der Vorlage keinesfalls robust. Dieses Problem wird auch in [13] deutlich:

'...the minima rule states in principle what these boundaries are, leaving open how in practice they may be computed from images despite noise and other resources.'

Die Lösung liegt darin, das lokale Konzept der *Punkte* minimaler Krümmung aufzugeben zugunsten einer Aufteilung in konvexe und konkave *Teilkurven*, und die Formunterteilung anhand der Hierarchie, die durch die Kurvenevolution vorgegeben ist, vorzunehmen.

4.2 Formunterteilung durch Kurvenevolution

Gegeben sei ein Polygon $\mathcal{P} = (p_0, p_1, \dots, p_n) \subset \mathbb{Z}^2$, $\# \mathcal{P} \geq 3$ derart, daß der zugeordnete Kantenzug $\bar{\mathcal{P}} \subset \mathbb{R}^2$ eine geschlossene Jordankurve darstellt. $\mathcal{L} \subseteq \mathcal{P}$, sei die Menge der konvexen Ecken, $\mathcal{R} \subset \mathcal{P}$, sei die Menge der konkaven Ecken.

Definition: Maximaler konvexer (konkaver) Bogen

Ein Teilpolygon $c = (p_s, p_{s+1}, \dots, p_e) \subseteq \mathcal{P}$, heißt *maximaler konvexer Bogen*, falls gilt:

$c = \gamma \Leftrightarrow \gamma = \mathcal{L}$ oder

$c \subset \gamma \Leftrightarrow p_s, p_e \in \mathcal{R}, p_{s+1}, \dots, p_{e-1} \in \mathcal{L}$

Ein Teilpolygon $c = (p_s, p_{s+1}, \dots, p_e) \subset \gamma$, heißt *maximaler konkaver Bogen* falls gilt:

$p_s, p_e \in \mathcal{L}, p_{s+1}, \dots, p_{e-1} \in \mathcal{R}$

Klar: Jeder maximale konvexe(konkave) Bogen besteht aus mindestens drei Punkten.

Die Menge aller maximalen konvexen / konkaven Bögen von γ , sei mit \mathcal{C} bezeichnet. Über die Ordnung der Punkte in γ , wird eine Ordnung in \mathcal{C} induziert:

Mit $c_i = (p_{s_i}, p_{s_i+1}, \dots, p_{e_i}), c_j = (p_{s_j}, p_{s_j+1}, \dots, p_{e_j})$ folgt: $i < j \Leftrightarrow s_i < s_j$

Eine allgemeine Theorie über Randpolygone und Bögen findet sich in Latecki und Rosenfeld [29].

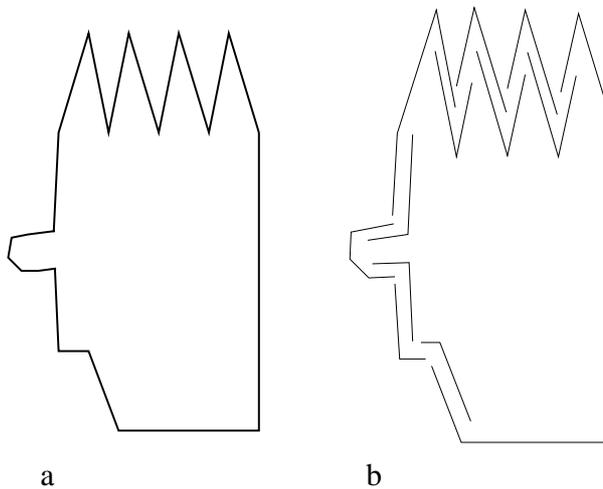


Abbildung 4.6: Aufteilung eines Randpolygons in maximale konvexe / konkave Bögen

Die durch die Kurvenevolution erhaltene Formhierarchie ist die Grundlage der Unterteilung einer Form in visuell relevante Teile. Dabei wird die Form auf einer gegebenen Hierarchiestufe in maximale konvexe und konkave Bögen unterteilt, siehe Abbildung 4.6.

Zu beachten bei einer solchen Unterteilung:

1. Die konvexen und konkaven Bögen schneiden sich jeweils in genau einem Segment.

2. Die konvexen Bögen entsprechen nicht notwendig exakt der im vorhergehenden Kapitel angeregten Unterteilung an Punkten maximaler negativer Krümmung. Dagegen existiert zu jedem Punkt p maximaler negativer Krümmung genau ein maximaler konkaver Bogen c mit $p \in \gamma$.
3. Als Teile einer Form werden die maximalen konvexen Bögen gedeutet. Die max. konkaven Bögen beinhalten Information über die (Winkel-) Lage dieser Teile zueinander.

Die ersten zwei Beobachtungen beschreiben den Übergang vom lokalen Konzept der Extremalpunkte zum robusten Konzept der Extremalsegmente: Die vorliegende Unterteilung bezieht sich auf polygonale Randkurven, d.h. Extremaleigenschaften können nicht mit Mitteln der Differentialgeometrie angegeben werden. Das in Punkt 1 angegebene Überlappungssegment kann jedoch analog zum Wendepunkt als 'Wendesegment' angesehen werden. Die Stärke von Krümmungen analog der Krümmung in kontinuierlichen Kurven kann in Polygonzügen nur über den Wendewinkel der jeweiligen Ecke angegeben werden. Anstatt des Wendewinkels in einer Ecke wird nun der Gesamtwinkel des maximalen konkaven Bogens als Krümmungsinformation angesehen. Dabei ist der Gesamtwinkel als Summe der einzelnen Wendewinkel definiert, siehe Abbildung 4.7

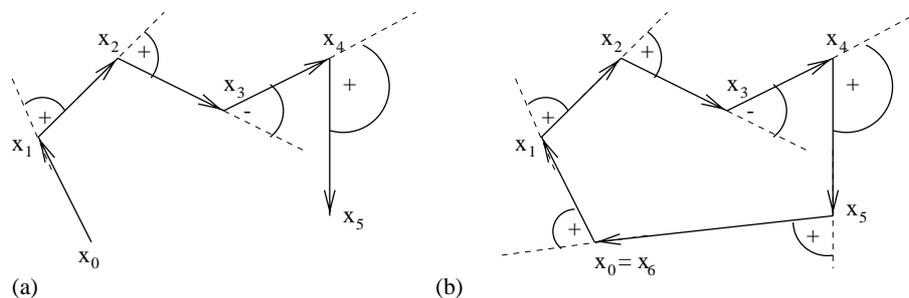


Abbildung 4.7: (a) Der **Gesamtwinkel** eines polygonalen Kantenzuges ist die Summe der Wendewinkel seiner Eckpunkte $\tau(x_1) + \dots + \tau(x_4)$. (b) Für ein geschlossenes Polygon gilt: $\tau(x_0) + \dots + \tau(x_5)$.

Die Tatsache, daß eine Beziehung zwischen visuell relevanten Teilen und Konvexität besteht, folgt direkt aus der obigen Motivation und wird z.B. in der Literatur bei Basri et al. [2] wie folgt beschrieben:

'Parts generally are defined to be convex or nearly convex shapes separated from the rest of the object at concavity extrema, as in Hoffmann and Richards [13], or at inflections, as in Koenderink and Doorn [16]'

Das Problem besteht dabei in der Definition des *'nearly convex'*. Dieses Problem kann durch Verwendung der Formhierarchie gelöst werden, die durch den Einsatz der Kurvenevolution gegeben ist: je höher die Stufe der Abstraktion, desto wichtiger sind die verbleibenden Eckpunkte für die Gesamtform. Auf hohen Abstraktionsstufen verbleibende maximale konvexe / konkave Bögen zeigen also wichtige Unterteilungen der Ausgangsform. Unwichtige, 'kleine' Konvexitäten oder Konkavitäten werden in früheren Abstraktionsstufen eliminiert (und können als Definition des *'nearly convex'* dienen).

Die hierarchische Formunterteilung liefert damit automatisch die Signifikanz der Teile:

- Je höher die Abstraktionstufe, desto visuell signifikanter die Teile
- Die Unterteilung auf hoher Abstraktionsstufe wird auf die Struktur in schwächerer Abstraktion vererbt

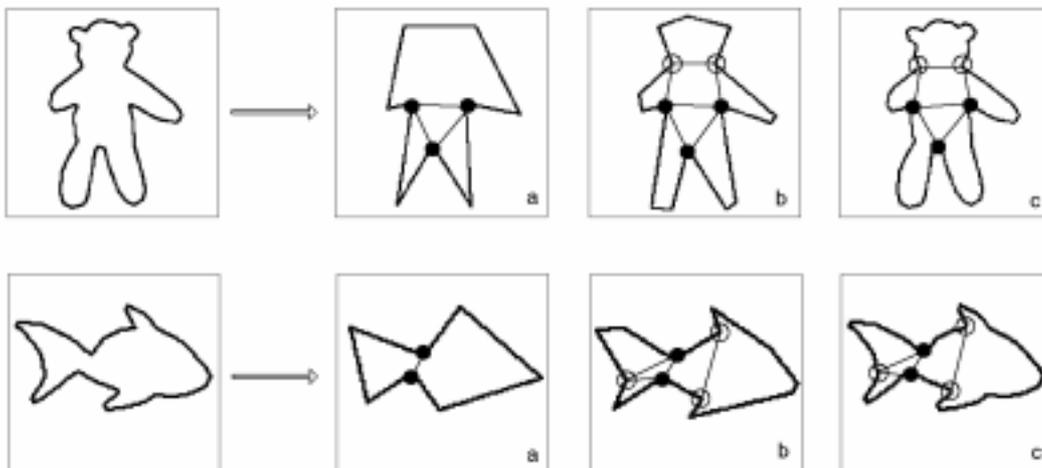


Abbildung 4.8: Unterteilung von Randkurven durch hierarchische Aufspaltung in konvexe Bögen

Wichtig ist wiederum die Eigenschaft der Positionstreue der Evolution: die auf hoher Abstraktionsstufe gefundenen stark abstrahierten Bögen geben nur die Unterteilung an, d.h. nicht die abstrakten Bögen selbst sind die Unterteilung der Form, sondern die Endpunkte p_s, p_e dieser Bögen. p_s, p_e sind aufgrund der Positionstreue der Evolution in der Ausgangsform enthalten. Die dadurch definierten Bögen (der Ausgangsform), die zu den abstrakten Bögen korrespondieren, geben die signifikanten Formteile an. Diese sind nicht notwendig konvex, sondern i.a. eben nur *'nahezu konvex'*.

4.3 Ergebnisse

Abbildung 4.8 zeigt zwei Beispiele für die Unterteilung von Randkurven durch die hierarchische Aufspaltung in konvexe Bögen. Auf der höchsten Abstraktionsstufe (a) wird die von der Evolution abstrahierte Form 'Bär' in drei konvexe Bögen, also visuell relevante Teile, die Form 'Fisch' in zwei Teile zerlegt. Die visuellen Teile der Ausgangsform, die zu diesen Bögen korrespondieren, repräsentieren die wichtigsten Formteile, für den Fisch die Schwanzflosse und den Fischkörper, für den Bären die Beine sowie den Oberkörper (auf dieser Stufe noch in Zusammenfassung der Arme und des Kopfes). Bei weiterer Unterteilung in (b) erhält man feinere visuelle Teile. So wird z.B. der den Oberkörper des Bären repräsentierende Bogen weiter unterteilt in Kopf und Arme. Die rechte Spalte (c) zeigt die korrespondierenden Unterteilungen auf der Originalform.

Die Robustheit des Verfahrens, die aus der Verwerfung der Konzepte lokaler Eigenschaften in Verbindung mit den Eigenschaften der Kurvenevolution folgt, zeigt Abbildung 4.9.

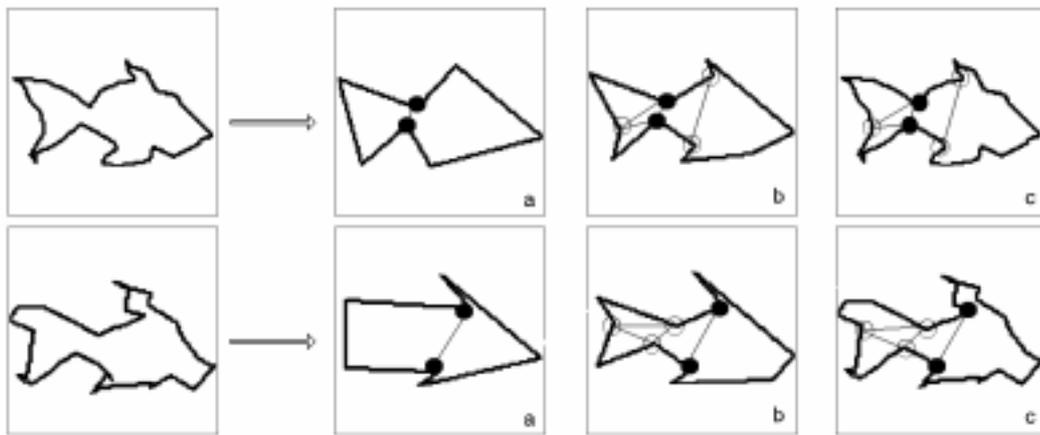


Abbildung 4.9: Robustheit des Verfahrens: Die Abbildung zeigt zwei gestörte Versionen der Form 'Fisch' aus 4.8. Die Aufteilung beider Versionen entspricht derjenigen der Originalform

Ogleich das nichtlokale Konzept der Unterteilung in maximale Bögen nicht exakt der Unterteilung an Punkten minimaler Krümmung entspricht, erzeugt es doch in vielen Experimenten ein ähnliches Ergebnis.

Dieses zeigt Abbildung 4.10, welche alle möglichen 'codon quadrupel' [13] zeigt. Die Originalobjekte entstammen Figur 7 in [13]. Die abstrahierten Versionen befinden sich jeweils links der Pfeile, auf dieser Abstraktionsstufe wurde die Formunterteilung bestimmt. Die Endpunkte maximaler konvexer Bögen sind durch

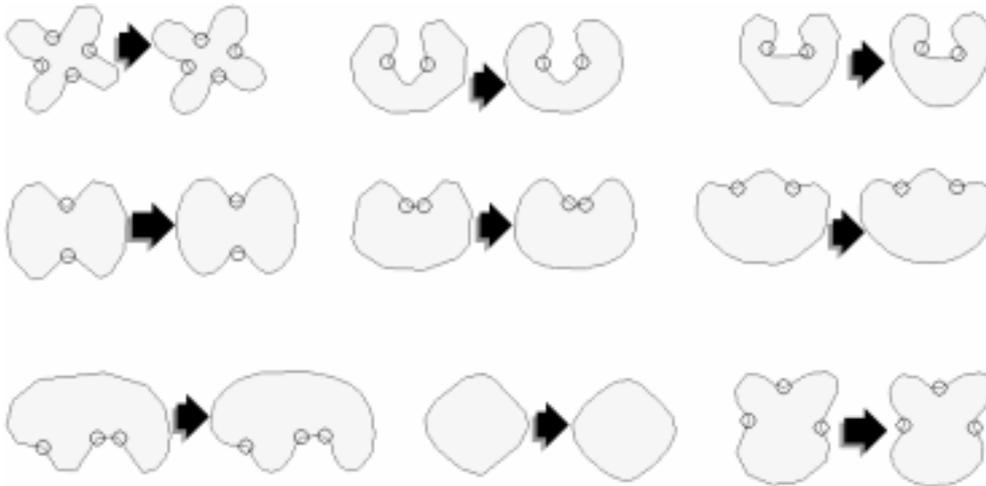


Abbildung 4.10: Die Endpunkte maximaler konvexer Bögen entsprechen Punkten minimaler Krümmung. Die Ausgangsobjekte (rechts der Pfeile) entstammen [13]

Kreise sowohl in der abstrahierten, als auch in der Originalform gekennzeichnet. Es zeigt sich, daß diese Endpunkte nahe den Krümmungsminima liegen. Dieses gilt ebenfalls für die Randkurven in Figur 4.8.

Zu beachten ist, daß die Abbildungen 4.8 bzw. 4.9 die *Fläche* unterteilen, 4.10 dagegen den Rand. Da zur Teilbestimmung ausschließlich Randinformationen benutzt wurden, erscheint eine reine Randunterteilung angemessener. Die durch Verbindung von Randpunkten erzeugten Teilflächen, bei Beusmans et al. [4] als *part-cuts* bezeichnet, benötigen weitergehende Theorien, wie in Hoffman und Singh [14] angemerkt.

Ein Beispiel dafür ist die schon oben erwähnte Theorie der *limbs and necks* von Siddiqi und Kimia [34]. Die Verbindung der auf dem Rand gefundenen Punkte zur Erzeugung von *limbs* benutzt die Information der Richtungsfortsetzung.

Obgleich eine solche Zusatzinformation bei der hierarchischen Zerlegung durch die Kurvenevolution nicht gegeben ist, werden für viele Objekte intuitive *part-cuts* allein durch die Verbindung der Endpunkte in den gegebenen Hierarchiestufen erzeugt, siehe Abbildung 4.8. Die Zusatzinformation geht hier über die Vererbung der Unterteilung in verschiedenen Abstraktionsstufen ein.

Die Unterteilung mittels der hierarchischen Zerlegung zeigt in vielen Experimenten eine starke Übereinstimmung mit den Ergebnissen von Siddiqi und Kimia, siehe Abbildung 4.11. Die dort gezeigten Formen entstammen aus [36], sie wurden dort für psychologische Experimente zur Verifizierung der Theorie der *limbs and necks* benutzt. Abbildung 4.11 zeigt pro Form drei Stufen der Evolution,

sowie die in ihnen erkannte Flächenunterteilung. Die rechte Spalte zeigt die unterteilte Originalform. Verschiedene Graustufen zeigen dabei die Hierarchie bzw. die Vererbung der Unterteilung. Eine dunklere Graustufe ist dabei die spätere Unterteilung einer (sie in einer abstrakteren Unterteilung enthaltenden) hell gefärbten Fläche.

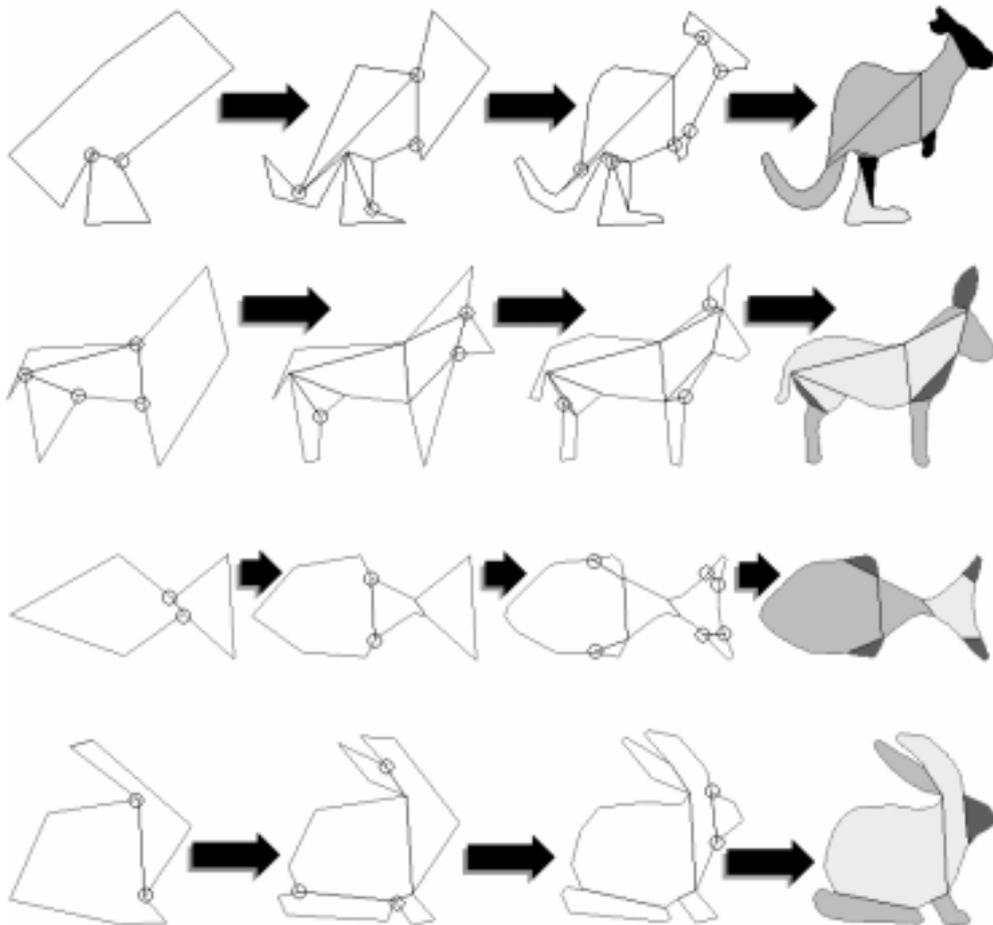


Abbildung 4.11: Flächenunterteilung: die Teile entsprechen im Wesentlichen denen durch *limbs and necks* aufgefundenen Teile in [36].

Ogleich die hierarchische Randunterteilung für eine Vielzahl von Objekten zu einer intuitiven Flächenunterteilung führt, gilt diese Aussage nicht für allgemeine Formen. Im folgenden wird daher nicht weiter auf die Flächenunterteilung eingegangen, benutzt wird ausschließlich die Unterteilung in maximale konvexe / konkave Bögen und ihre Vererbung bzw. weitere Unterteilung auf verschiedenen Abstraktionsstufen.

Kapitel 5

Formvergleich

5.1 Einleitung

Die formbezogene Einordnung eines Objektes in eine Bilddatenbank verlangt nach einem Vergleichsmaß, welches die Zuordnung des Eingabeobjektes zu einer Formklasse ermöglicht. Dieses Vergleichsmaß sollte kognitiv motiviert sein, d.h. mit der menschlichen Wahrnehmung übereinstimmen. Aus dieser einfachen Forderung lassen sich fünf grundlegende Eigenschaften formulieren:

1. Das Vergleichsmaß sollte die Erkennung von intuitiv ähnlichen Objekten erlauben, auch wenn sie nicht mathematisch identisch sind.
2. Das Vergleichsmaß sollte robust gegenüber Störungen, z.B. Digitalisierungsfehlern, arbeiten.
3. Das Vergleichsmaß sollte die Aufteilung der Gesamtform in Teilformen respektieren.
4. Das Vergleichsmaß sollte invariant gegenüber grundlegenden Transformationen sein.
5. Das Vergleichsmaß sollte universell sein, d.h. es sollte keine a priori Information benötigen, die nicht spezifizierte Objekte von der Klassifizierung ausschließt.

Schon die Kurvenevolution und die Unterteilung der Randpolygone wurden unter Berücksichtigung dieser Forderungen entwickelt. Die Evolution zielt auf (1)

und (2), die hierarchische Formunterteilung auf (3), die Forderungen (4) und (5) werden sowohl von der Evolution als auch von der Unterteilung erfüllt.

Das im folgenden vorgestellte Vergleichsmaß nutzt Evolution und Unterteilung dadurch, daß zwei als Randpolygone ρ_0, ρ_1 gegebene Formen als Vorverarbeitung mittels der Evolution zunächst abstrahiert, daraufhin auf einer (automatisch erkannten) geeigneten Abstraktionsstufe unterteilt werden. Das Vergleichsmaß operiert auf dieser Zerlegung.

Da die Unterteilung der Formen ρ'_0, ρ'_1 nur durch eine Folge von maximalen Bögen gegeben ist, welche keine Lage- bzw. Rotationsinformation enthält, muß nach Forderung (3) zusätzlich sichergestellt werden, daß nur visuell korrespondierende Teile miteinander verglichen werden:

- Die Teilpolygone von ρ'_0 und ρ'_1 müssen einander zugeordnet werden, eine *Korrespondenz* zwischen ihnen muß erstellt werden.

Die Korrespondenz wird im wesentlichen durch den Vergleich verschiedener Kombinationen maximaler Bögen errechnet. Es wird eine Abbildung zwischen den Gruppen definiert, welche einen maximalen Bogen aus ρ'_0 einem oder mehreren aufeinanderfolgenden maximalen Bögen aus ρ'_1 , oder einen maximalen Bogen aus ρ'_1 einem oder mehreren aufeinanderfolgenden maximalen Bögen aus ρ'_0 zugeordnet, niemals aber mehrere maximale Bögen aus ρ'_0 mehreren aus ρ'_1 (bzw. umgekehrt). Diese Bedingung ist darin begründet, daß

- ein maximaler Bogen einer stark abstrahierten Evolutionsstufe zu möglicherweise mehreren maximalen Bögen auf einer weniger abstrahierten Stufe korrespondiert
- ein maximaler Bogen aus $\rho'_{0(1)}$ einer gestörten und daher aus mehreren maximalen Bögen bestehenden Version des korrespondierenden visuellen Teils in $\rho'_{1(0)}$ besteht.

Da durch die Verwendung einer rauschfreien Abstraktionsstufe von der Annahme ausgegangen werden kann, daß jeder maximale konvexe Bogen genau ein visuelles Teil darstellt, erhält das Verbot der Zuordnung von mehreren–zu–mehreren Teilpolygonen die visuelle Struktur der Formen. Es ist i.a. nicht möglich, eine reine eins–zu–eins Zuordnung zwischen Teilpolygonen zu finden, da nach Forderung (5) nicht von mathematisch identischen Objekten ausgegangen werden kann.

Die möglichen Teilpolygonkombinationen werden durch *dynamisches Programmieren* erzeugt. Die Korrespondenz wird in Kapitel 5.2.1 definiert, die praktische Umsetzung der Modelle in den Kapiteln 5.2.3, 5.2.4 vorgestellt.

Aus Forderung (1) folgt:

- Der Korrespondenz muß ein Vergleichsmaß von Polygonen zugrunde liegen, das der Auffassung intuitiver Ähnlichkeiten entspricht.

Der vorliegende Ansatz lehnt sich an ein Maß an, daß von Arkin et al. in [1] vorgestellt wurde. Es basiert auf der L_2 -Metrik der Tangentenraumdarstellung der Randpolygone. Dabei ist der wesentliche Unterschied zu [1] dadurch gegeben, daß der vorliegende Ansatz das Maß auf *Teilpolygone* anwendet, während in [1] die *Gesamtpolygone* verglichen werden. Genaueres dazu findet sich in Kapitel 5.3.

Aus der Korrespondenz der Teilbögen kann daraufhin ein Maß für den Vergleich der Gesamtform hergeleitet werden, Kapitel 5.5 behandelt dieses Thema.

5.2 Korrespondenz der visuellen Teile

5.2.1 Definition der Korrespondenz

Gegeben sei ein Polygon P . Die Menge \mathcal{C} bezeichne die geordnete Menge aller maximalen konvexen / konkaven Bögen (c_0, c_1, \dots, c_n) , $c_i \subseteq P$. Die Ordnung in \mathcal{C} wird dabei über die Ordnung der Punkte p_i induziert (siehe Kapitel 4.2).

Definition: Gruppe

Eine *Gruppe* $g = (p_{i_0}, p_{i_0+1}, \dots) \subseteq P$, ist ein Teilpolygon aus P , welches durch Vereinigung $c_s \cup c_{s+1} \cup \dots \cup c_{s+n-1}$, $n > 0$ n aufeinanderfolgender maximaler Bögen $c_i \in \mathcal{C}$ gebildet wird. Die Menge aller Gruppen von P wird mit $\mathcal{G}(P)$ bezeichnet.

Definition: Gruppierung

Eine Folge $G = (g_0, g_1, \dots, g_{n-1})$ von Gruppen g_i heißt *Gruppierung*, falls gilt:

$\#(g_i \cap g_{i+1 \pmod n}) = 2$, d.h. zwei aufeinanderfolgende Gruppen überschneiden sich in genau einem Liniensegment in \overline{P} .

Die Bedingung beinhaltet, daß P vollständig von G überdeckt wird, da insbesondere $\#(g_0 \cap g_{n-1}) = 2$ gelten muß.

Die Menge aller Gruppierung von P wird mit $\mathcal{G}(P)$ bezeichnet.

Definition: Korrespondenz

Gegeben seien zwei Polygone P_0, P_1 . Die Gruppen $G_0 \in \mathcal{G}(P_0)$ und $G_1 \in \mathcal{G}(P_1)$ *korrespondieren*, falls eine Abbildung $k : G_0 \rightarrow G_1$ existiert, für die gilt:

1. k ist bijektiv
2. $\forall g \in G_0 : g \in \mathcal{C}_0 \vee k(g) \in \mathcal{C}_1$

k heißt *Korrespondenz* zwischen G_0 und G_1 . Die Menge aller Gruppierungen $(G_0, G_1) \in \mathcal{G}(\cdot, \cdot) \times \mathcal{G}(\cdot, \cdot)$, zu denen eine Korrespondenz k existiert, wird mit $\mathcal{K}_{\Gamma_0, \Gamma_1}$ bezeichnet.

Falls zu zwei Gruppen $g_0 \in G_0, g_1 \in G_1$ eine Korrespondenz k zwischen G_0 und G_1 derart existiert, daß $k(g_0) = g_1$, so heißt g_1 Teilkorrespondenz zu g_0 .

Bemerkung:

Bedingung (1) sichert, daß beide Formen in eine gleiche Anzahl von Gruppen zerlegt werden. Bedingung (2) sichert, daß zumindest einer der jeweils zugeordneten Bögen ein maximaler konvexer / konkaver Bogen ist, die Zuordnung von mehreren zu mehreren Bögen wird also ausgeschlossen.

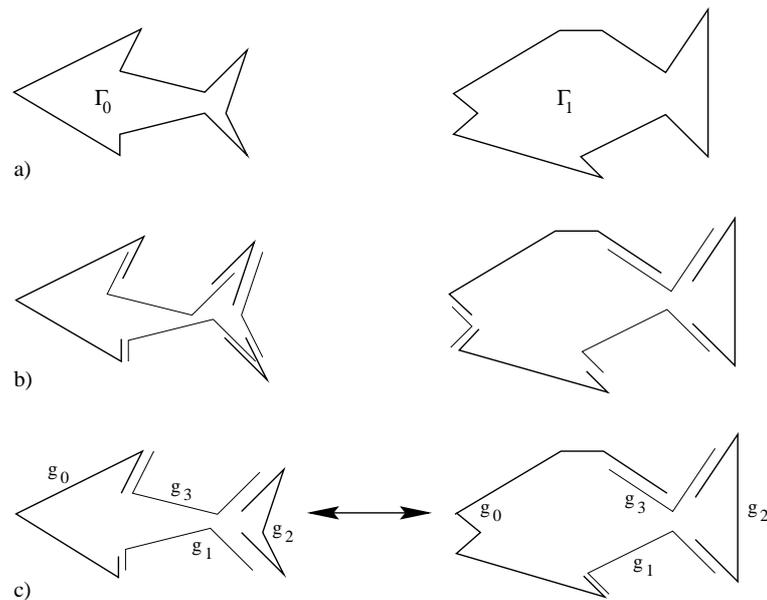


Abbildung 5.1: Zeile a) Ausgangsform b) Unterteilung in maximale Bögen c) mögliche Gruppierung, die Korrespondenz wird durch gleiche Indizierung gezeigt.

Abbildung 5.1 zeigt ein Beispiel: die zu vergleichenden Polygone Γ_0, Γ_1 der Zeile (a) werden in maximale konvexe / konkave Bögen unterteilt (b). Zeile (c) zeigt jeweils eine Gruppierung G_0 bzw. G_1 aus $\mathcal{G}(\cdot, \cdot)$ bzw. $\mathcal{G}(\cdot, \cdot)$. Die Gruppen $g_0(\cdot, \cdot)$ und $g_2(\cdot, \cdot)$ bestehen aus zusammengefügt maximalen Bögen.

5.2.2 Der Vergleich zweier Objekte

Im folgenden werden maximale konvexe/konkave Bögen der Einfachheit halber nur mit 'Bogen' bezeichnet.

Die Korrespondenzen k aus $\mathcal{K}_{\Gamma_0, \Gamma_1}$ werden durch ein Maß, die *Fehlstellung*, bewertet, welches die Güte der intuitiven Zuordnung modelliert. Dieses Maß wird definiert über das Vergleichsmaß S der einzelnen Teile, welches in Kapitel 5.3 hergeleitet wird.

Definition: Fehlstellung einer Korrespondenz

Es sei $\mathcal{S} : g(, 0) \times g(, 1) \rightarrow \mathbb{R}_0^+$ ein Polygonvergleich (siehe Kapitel 5.3).

Die Abbildung $\mathcal{M} : \mathcal{K}_{\Gamma_0, \Gamma_1} \rightarrow \mathbb{R}_0^+, (G_0, G_1) \rightarrow \sum_{g \in G_0} \mathcal{S}(g, k(g))$

heißt *Fehlstellung* der Korrespondenz.

Definition: Optimale Korrespondenz

Eine Korrespondenz $k_0 \in \mathcal{K}_{\Gamma_0, \Gamma_1}$ heißt optimale Korrespondenz, falls gilt:

$$\forall k \in \mathcal{K}_{\Gamma_0, \Gamma_1} : \mathcal{M}(k_0) \leq \mathcal{M}(k)$$

Definition: Formunterschied

Ist durch $k_0 \in \mathcal{K}$ die optimale Korrespondenz gegeben, so heißt $\mathcal{M}(k_0)$ *Formunterschied* zweier durch $, 0, 1$ gegebener Objekte.

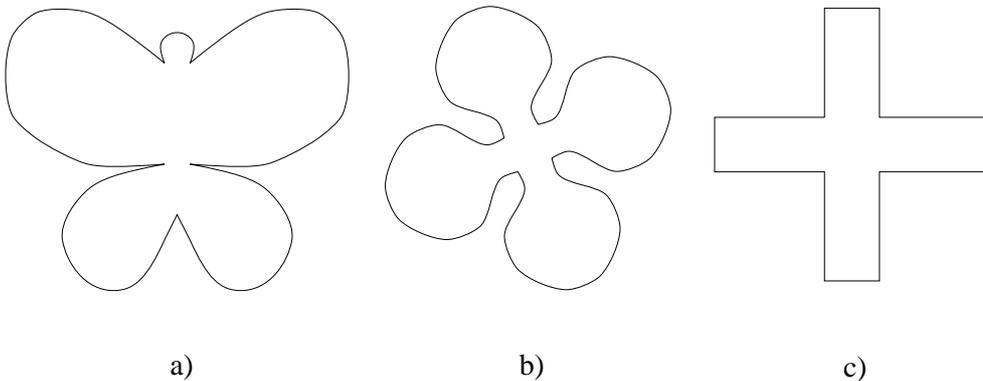


Abbildung 5.2: Das intuitive Maß der Ähnlichkeit erfüllt nicht die Dreiecksungleichung

Bemerkung:

- Durch den Formunterschied wird der Ähnlichkeitswert zweier Objekte angegeben. Je geringer er ist, desto ähnlicher sind die Objekte.

- Der Formunterschied ist keine Metrik, er erfüllt nicht die Dreiecksungleichung. Die Dreiecksungleichung allerdings entspricht ebenfalls nicht der menschlich intuitiven Ähnlichkeitsbewertung, siehe Abbildung 5.2: Die 'intuitive Summe' der Unterschiede zwischen a),b) sowie b),c) scheint geringer als diejenige zwischen a),c).
- Die Metrikaxiome der Definitheit sowie der Symmetrie sind erfüllt, diese sind auch für das intuitive Maß sinnvoll.

Um die Berechnung aller möglichen Kombinationen der Bögen zu Gruppierungen zu umgehen, wird dieses Problem durch *dynamisches Programmieren* gelöst. Zum Verständnis des Einsatzes dieses Verfahrens soll jedoch zunächst ein Beispiel mit vollständiger Kombination durchgespielt werden:

5.2.3 Mögliche Gruppierungen: ein Beispiel

Gegeben seien zwei fischähnliche Formen wie in Abbildung 5.3. Beide Formen lassen sich in vier Bögen unterteilen. Die Indizes der Abbildung suggerieren die intuitiv richtige Korrespondenz, diese soll aus allen Korrespondenzen in $\mathcal{K}_{\Gamma_0, \Gamma_1}$ gefunden werden.

Die dabei verwendete Technik ist startpunktabhängig: die maximalen Bögen a_0, a_3 bzw. b_0, b_3 sind ausgezeichnet, es gibt keine Gruppe mit Vereinigung $a_0 \cup a_3$ bzw. $b_0 \cup b_3$. Diese Einschränkung der Allgemeinheit wird in der Praxis durch Rotation der Startbögen ausgeglichen.

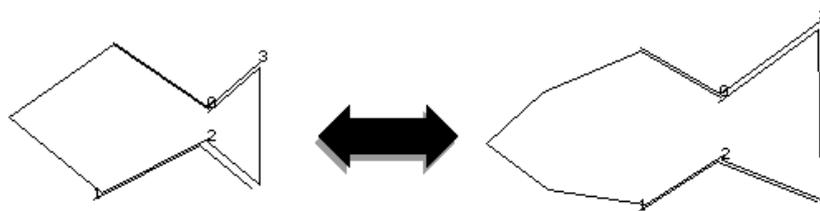
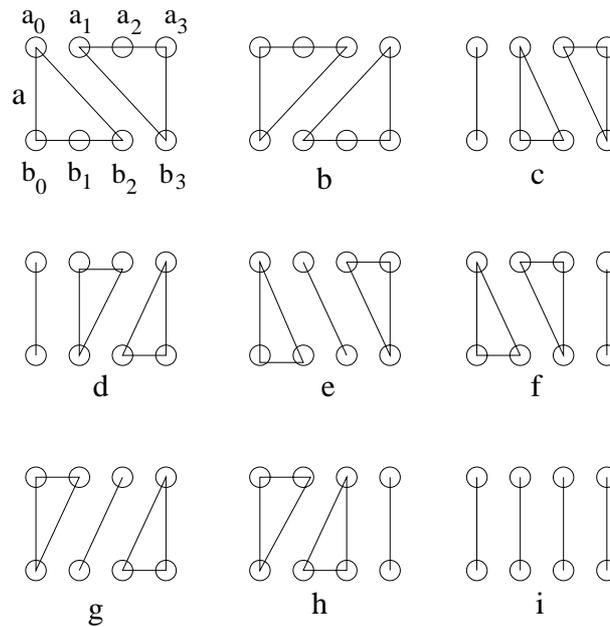


Abbildung 5.3: Die Ausgangspolygone des Beispiels, links P_0 , rechts P_1

Abbildung 5.4 zeigt die Kombinationen. Die Kreise der jeweils oberen Zeile der Abbildungen a)–i) stehen für die Bögen des Polygons P_0 , die Kreise der jeweils unteren Zeile für Bögen des Polygons P_1 aus Abbildung 5.3. Die Verbindungslinien innerhalb einer Zeile kennzeichnen die zu einer Gruppe zusammengefaßten Bögen, die Verbindungen zwischen den Zeilen die Teilkorrespondenzen. Mit $g_i \in \mathcal{G}(P_0)$ und $h_i \in \mathcal{G}(P_1)$ ist beispielsweise Abbildung 5.4 c) wie folgt zu deuten:

Abbildung 5.4: Die Korrespondenzen in $\mathcal{K}_{\Gamma_1, \Gamma_2}$

- Γ_0 ist aufgeteilt in drei Gruppen, bestehend aus den Bögen $g_0 = a_0, g_1 = a_1, g_2 = a_2 \cup a_3, a_i \in \mathcal{C}(\Gamma_0)$
- Γ_1 ist aufgeteilt in drei Gruppen, bestehend aus den Bögen $h_0 = b_0, h_1 = b_1 \cup b_2, h_2 = b_3, b_i \in \mathcal{C}(\Gamma_1)$
- Die Teilkorrespondenzen:
 - $h_0 = b_0$ ist Teilkorrespondenz zu $g_0 = a_0$
 - $h_1 = b_0 \cup b_1$ ist Teilkorrespondenz zu $g_1 = a_1$
 - $h_2 = b_3$ zugeordnet ist Teilkorrespondenz zu $g_2 = a_2 \cup a_3$

Die Suche nach der optimalen Korrespondenz in den möglichen Kombinationen a)–i) kann durch den gerichteten Graphen in Abbildung 5.5 verdeutlicht werden: Jede Kante zeigt eine Teilkorrespondenz, so entspricht z.B. die Kante 1012,0 derjenigen Teilkorrespondenz, die die Bögen a_0, a_1, a_2 dem Bogen b_0 zuordnet. Diese Teilkorrespondenz ist Teil der in Abbildung 5.4 b) gezeigten Korrespondenz. Die Indizes der Knoten zeigen die Indizes der Bögen, die genau an den Teilkorrespondenzen der Kanten beteiligt sind, welche vom Knoten *start* aus bis zu dem jeweils indizierten Knoten durchschritten wurden. So bedeutet der Knotenindex

¹Die Kantenbezeichnungen sind ziffernweise zu lesen

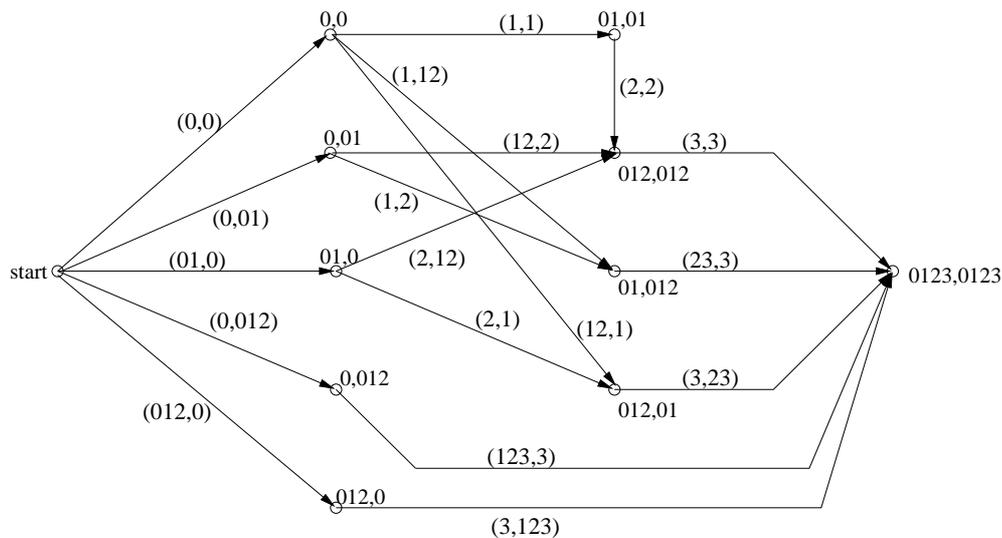


Abbildung 5.5: Graph zur Erzeugung aller Korrespondenzen in $\mathcal{K}_{\Gamma_0, \Gamma_1}$. Der Graph ist gerichtet, d.h. die Kanten sind nur in Pfeilrichtung zu durchschreiten.

012,01, daß alle Wege, die zu diesem Knoten führen, genau die Bögen a_0, a_1, a_2 den Bögen b_0, b_1 so zuordnen, daß sie an einer Korrespondenz beteiligt sind.

Als Beispiel diene wiederum Korrespondenz c) aus Abbildung 5.4:

- Die Teilkorrespondenz g_0 zu h_0 entspricht Kante (0, 0) und führt zu Knoten 0, 0
- Die Teilkorrespondenz g_1 zu h_1 entspricht Kante (1, 12) und führt zu Knoten 01, 012
- Die Teilkorrespondenz g_2 zu h_2 entspricht Kante (23, 3) und führt zu Knoten 0123, 0123

Die Erstellung aller gültigen Korrespondenzen $\mathcal{K}_{\Gamma_0, \Gamma_1}$ entspricht dem Durchschreiten aller möglichen Wege in dem gerichteten Graphen vom Knoten *start* zum Knoten 0123, 0123.

Werden den Kanten nicht nur die Teilkorrespondenzen, sondern auch das Maß S , also der Polygonvergleich der Teilkorrespondenz zugeordnet, so ergibt sich durch Aufsummierung der Polygonvergleichswerte der durchschrittenen Kanten die Fehlstellung der erzeugten Korrespondenz.

Die optimale Korrespondenz entspricht also dem *billigsten* Weg, d.h. dem Weg minimaler Summe der Kantenmaße.

5.2.4 Auffinden der optimalen Korrespondenz durch dynamisches Programmieren

Um die kombinatorische Explosion zu umgehen, wird die Erzeugung von Korrespondenzen durch dynamisches Programmieren gelöst. Die Vorgehensweise soll hier nicht allgemein beschrieben werden, sondern direkt anhand der Anpassung an das vorliegende Problem.

Notation:

$G = (E, V)$ bezeichne den Graphen, bestehend aus der Menge der Kanten $E = e_0, \dots, e_n$ sowie der Menge der Knoten $V = v_0, \dots, v_m$.

Der Wert einer Kante e_i sei durch $\mathcal{S}(e)$ gegeben, wobei \mathcal{S} der Polygonvergleich der Teilkorrespondenz ist, die durch die Kante dargestellt wird.

Die Menge aller Wege w_i zwischen zwei Knoten v_a und v_b sei durch $W_{v_a}^{v_b}$ bezeichnet.

Definition: Verbindungswert

Der Wert $M(w)$ eines Weges $w = (e_{i_1}, e_{i_2}, \dots, e_{i_l})$ sei definiert durch die Summe $\sum_{e_i \in w} \mathcal{S}(e_i)$ der Werte seiner Kanten.

Der Verbindungswert $C(v_{i_1}, v_{i_2})$ zweier Knoten v_{i_1}, v_{i_2} ist der *minimale* Wert aller Wege zwischen v_{i_1} und v_{i_2} : $C(v_{i_1}, v_{i_2}) = \min_{w \in W_{v_{i_1}}^{v_{i_2}}} (M(w))$

Definition: Vorgänger

Ein Knoten v_p heißt *Vorgänger* des Knotens v_b bezüglich des Knoten v_a , falls gilt:

- $\exists w = (e_{i_1}, e_{i_2}, \dots) \in W_{v_a}^{v_b}$ mit $v_p \in e_{i_k} \in w$
- $\exists e \in E$ mit $e = (v_p, v_b)$

Die Menge aller Vorgänger von v_b bzgl. v_a wird mit $P_{v_a}(v_b)$ bezeichnet.

Die Grundidee des Verfahrens läßt sich durch folgenden Satz formulieren:

Satz:

Gegeben seien zwei Knoten v_a, v_b . Dann gilt:

$$C(v_a, v_b) = \min_{p \in P_{v_a}(v_b)} (C(v_a, p) + M((p, v_b)))$$

Beweis: Naheliegend. ■

Gesucht ist der Weg minimaler Kosten innerhalb eines gerichteten Graphen zwischen zwei ausgezeichneten Knoten, im obigen Beispiel die Knoten $v_a = start$ und

$v_b = 0123, 0123$. Der Satz legt nahe, jedem Knoten $v \in V$ den Verbindungswert $C(v_a, v)$ anhand der Verbindungswerte $C(v_a, v_p)$, $v_p \in P_{v_a}(v)$ zuzuordnen.

Die Berechnung der Verbindungskosten zwischen zwei Punkten läßt sich algorithmisch rekursiv formulieren:

Rekursive Berechnung der Verbindungskosten

Start: $v = v_b$
BERECHNE $C_{v_a}(v)$
 End.

BERECHNE $C_{v_a}(v)$:

if $P_{v_a}(v) = \emptyset$:
 $C_{v_a}(v) = 0$

else
 $\forall p_i \in P_{v_a}(v)$: **BERECHNE** $C_{v_a}(p_i)$
 $C_{v_a}(v) = \min_{p_i \in P_{v_a}(v)} (C_{v_a}(p_i) + M((p_i, v)))$

Bemerkungen:

- Die maximale Weglänge in G entspricht $\min(\#\mathcal{C}_{\Gamma_0}, \#\mathcal{C}_{\Gamma_1})$, wobei \mathcal{C}_{Γ_i} wie bisher die Menge aller maximalen Bögen in Γ_i bezeichnet
- Der zu den Verbindungskosten zugehörige Weg zwischen $start$ und $0123, 0123$ entspricht der Korrespondenz minimaler Fehlstellung.
- Geschwindigkeitsoptimierte Ansätze zur Berechnung der optimalen Korrespondenz sollen hier nicht vorgestellt werden.

Das Beispiel soll beendet werden mit einer Tabelle realer Daten zu den Formen aus Abbildung 5.3, erstellt durch das Datenbankprogramm.²

Erklärung der Tabellen:

Die linke Spalte zeigt die Indizes der Knoten, die folgenden vier Spalten zeigen

²Natürlich liefert ein JAVA-Programm der späten 90er Jahre keine Tabellen, sondern Mengen an Fenstern und Farbgrafiken. Die gezeigten Tabellen sind Auszüge der Berechnungen zwischen den Fenstern !

Knoten	Weglänge			
	1	2	3	4
0,0	0.101	-	-	-
0,01	0.410	-	-	-
0,012	0.3.458	-	-	-
0,0123	4.145	-	-	-
01,0	0.488	-	-	-
01,01	-	0.153	-	-
01,012	-	3.427	-	-
01,0123	-	3.483	-	-
012,0	4.088	-	-	-
012,01	-	2.804	-	-
012,012	-	1.255	0.261	-
012,0123	-	3.346	1.154	-
0123,0	6.189	-	-	-
0123,01	-	4.125	-	-
0123,012	-	3.417	1.180	-
0123,0123	-	16.150	1.320	0.326

Tabelle 1: Verbindungskosten

Knoten	Weglänge			
	1	2	3	4
0,0	start	-	-	-
0,01	start	-	-	-
0,012	start	-	-	-
0,0123	start	-	-	-
01,0	start	-	-	-
01,01	-	0,0	-	-
01,012	-	0,01	-	-
01,0123	-	0,012	-	-
012,0	start	-	-	-
012,01	-	01,0	-	-
012,012	-	01,0	01,01	-
012,0123	-	01,0	01,01	-
0123,0	start	-	-	-
0123,01	-	012,0	-	-
0123,012	-	0,01	01,01	-
0123,0123	-	012,0	012,012	012,012

Tabelle 2: Indizes der Vorgänger

Daten, die von der Weglänge³ abhängen. Der Eintrag '-' bedeutet, daß keine Daten berechnet werden können, da der Knoten nicht durch einen Weg der entsprechenden Länge zu erreichen ist.

Tabelle 1 zeigt die Verbindungskosten $C(start, v)$ zu den jeweiligen Knoten v , wobei pro Spalte nur Wege einer Länge, die durch die Spalte bestimmt ist, zugelassen sind. So bedeutet der Eintrag 1.320 in Zeile 0123, 0123, Spalte 3, daß die Verbindungskosten zwischen $start$ und 0123, 0123 unter ausschließlicher Berücksichtigung von Wegen der Länge=3 1.320 betragen. Dieses entspricht der minimalen Fehlstellung aller Korrespondenzen mit 3 Teilkorrespondenzen, d.h. der Korrespondenzen $c - h$ in Abbildung 5.4. Der Minimalwert der gesamten Zeile 0123, 0123 entspricht den Verbindungskosten (unter Berücksichtigung *aller* möglichen Wege, wie in der Definition gefordert), im Beispielfall der unterlegte Wert 0.326.

Tabelle 1 zeigt die Verbindungskosten sowie implizit die Anzahl der an der optimalen Korrespondenz beteiligten Teilkorrespondenzen. Tabelle 2 dokumentiert die Entstehung der Tabelle 1, über sie läßt sich die genaue optimale Korrespondenz ermitteln.

Tabelle 2 zeigt zu jedem Knoteneintrag den Vorgänger auf dem Weg, der zu minimalen Verbindungskosten führt. Der (nichtunterlegte) Eintrag 012, 012 in Zeile 0123, 0123, Spalte 3 bedeutet z.B., daß der Vorgänger des Knotens 0123, 0123 unter Berücksichtigung aller Wege der Länge=3 der Knoten 012, 012 ist. Der Weg läßt sich nun spaltenweise zurückverfolgen: der Vorgänger des Knotens 012, 012 ist Knoten 01, 0, gegeben durch den Eintrag in Zeile 012, 012, Spalte 2. Weitere Rückverfolgung ergibt den Gesamtweg $start - 01, 0 - 012, 012 - 0123, 0123$. Dieser Weg entspricht der Korrespondenz 'h' in Abbildung 5.4, da für diesen Weg die Kanten (\rightarrow Teilkorrespondenzen) (01, 0), (2, 12), (3, 3) durchschritten werden.

Die unterlegten Eintragsfelder zeigen die optimale Korrespondenz

$$start - 0, 0 - 01, 01 - 012, 012 - 0123, 0123$$

die zugehörigen Kanten sind:

$$(0, 0), (1, 1), (2, 2), (3, 3)$$

Die optimale Korrespondenz ist also Korrespondenz 'i' aus Abbildung 5.4, wie erwartet die eins-zu-eins Zuordnung der Formteile aus Abbildung 5.3.

³Zur Erinnerung: die Weglänge entspricht der Anzahl der Teilkorrespondenzen

5.3 Ähnlichkeitsmaß für Polygone

Dieses Kapitel definiert ein Maß, anhand dessen die visuellen Formteile verglichen werden können, die durch Teilpolygone der Randkurven γ_0, γ_1 gegeben sind. Dieses Maß ist die Grundlage u.a. zur Berechnung der Verbindungskosten und damit der optimalen Korrespondenz.

Gegeben seien zwei Randkurven γ_0, γ_1 sowie zwei Gruppen $g_{i=0,1} \in g(\gamma_i)$. Die Länge von g_i sei zu 1 normiert.

Zur Definition des Maßes werden die Gruppen g_i in die Tangentenraumdarstellung (siehe Kapitel 2.3) $\mathcal{T}(g_i)$ transformiert.

Dann ist das Vergleichsmaß, genannt Polygonvergleich, gegeben durch:

Definition: Polygonvergleich S

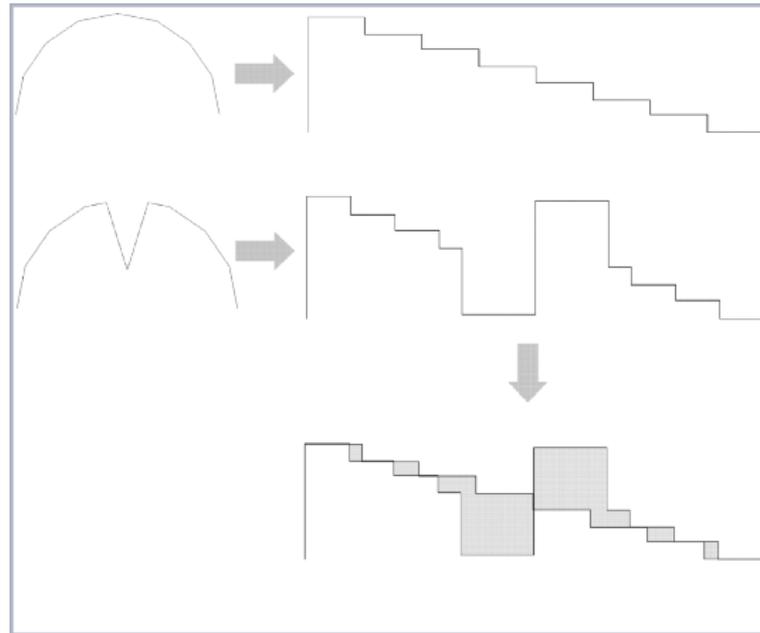
$$S(g_0, g_1) = \left(\int_0^1 (T(g_0)(s) \Leftrightarrow T(g_1)(s) + \theta_0)^2 ds \right)^{\frac{1}{2}} \max(l(g_0), l(g_1)) \max\left(\frac{l(g_0)}{l(g_1)}, \frac{l(g_1)}{l(g_0)}\right)$$

Dabei bezeichnet $l(g_i)$ die relative Länge der Gruppe g_i bezüglich der Länge der Randkurve γ_i , θ_0 ist eine Translationskonstante, die das Integral minimiert. Existenz und Berechnung dieser Konstante werden in Arkin et al. [1] aufgezeigt.

Die Definition des Polygonvergleichs ist wie folgt motiviert, siehe Abbildung 5.6:

- Das L_2 -Maß, berechnet durch das Integral, detektiert Formunterschiede der Polygone, siehe als Beispiel Abbildung 5.6
- Der Formunterschied wird mit dem Straffaktor $\max\left(\frac{l(g_0)}{l(g_1)}, \frac{l(g_1)}{l(g_0)}\right)$ multipliziert, der die unterschiedliche relative Länge der Gruppen zu den Ausgangspolygonen bewertet, d.h. dieser Faktor ist die Strafe für die Normierung beider Gruppen auf gleiche Länge
- Schließlich wird die Wichtigkeit der Gruppe innerhalb des Ausgangspolygons bewertet: der Faktor $\max(l(g_0), l(g_1))$ benutzt dazu die 'wichtigere' Gruppe, d.h. die Gruppe größerer relativer Länge.

⁴Die Definition des Maßes behandelt allgemeine Polygone $P \subset \mathbb{R}^2$. Im Rahmen des Formvergleichs werden allerdings nur Gruppen verglichen.

Abbildung 5.6: zu $D_{L_2}(T(g_0), T(g_1))$

5.4 Ergebnisse der Korrespondenz

Zunächst sollen nur Ergebnisse der Korrespondenz, also die Ergebnisse der Zuordnungen und nicht das Vergleichsmaß gezeigt werden. Die Möglichkeit, Korrespondenzen zu ermitteln, ist eine wichtige Eigenschaft dieses Verfahrens, da Formteile innerhalb eines Bildes lokalisiert werden können. Dieses ist bei Verfahren, welche die Form zum Vergleich in eine andere Darstellung transformieren, oftmals nicht bzw. nur erschwert möglich.

Die Ergebnisse der Korrespondenzberechnung sollen unter 6 verschiedenen Testbedingungen gezeigt werden, siehe Abbildung 5.7. Die Bezifferung der Teilbögen der Abbildungen entspricht korrespondierenden Teilen.

- a) Einfacher Test, ähnliche Formen.
- b) Ähnliche Form, verschiedene Anzahl von Teilen. Der untere Fisch besitzt eine zusätzliche Flosse. Die Flosse wird als Teil des Fischkörpers korrekt zugeordnet. Die Zuordnung der restlichen Teile bleibt bestehen.
- c) Nichtstatische Deformation der Teile. Die Gesamtformen sind ähnlich, die einzelnen Teile jedoch unterscheiden sich u.a. durch Rotation und Biegung. Auch hier werden alle Teile korrekt zugeordnet.

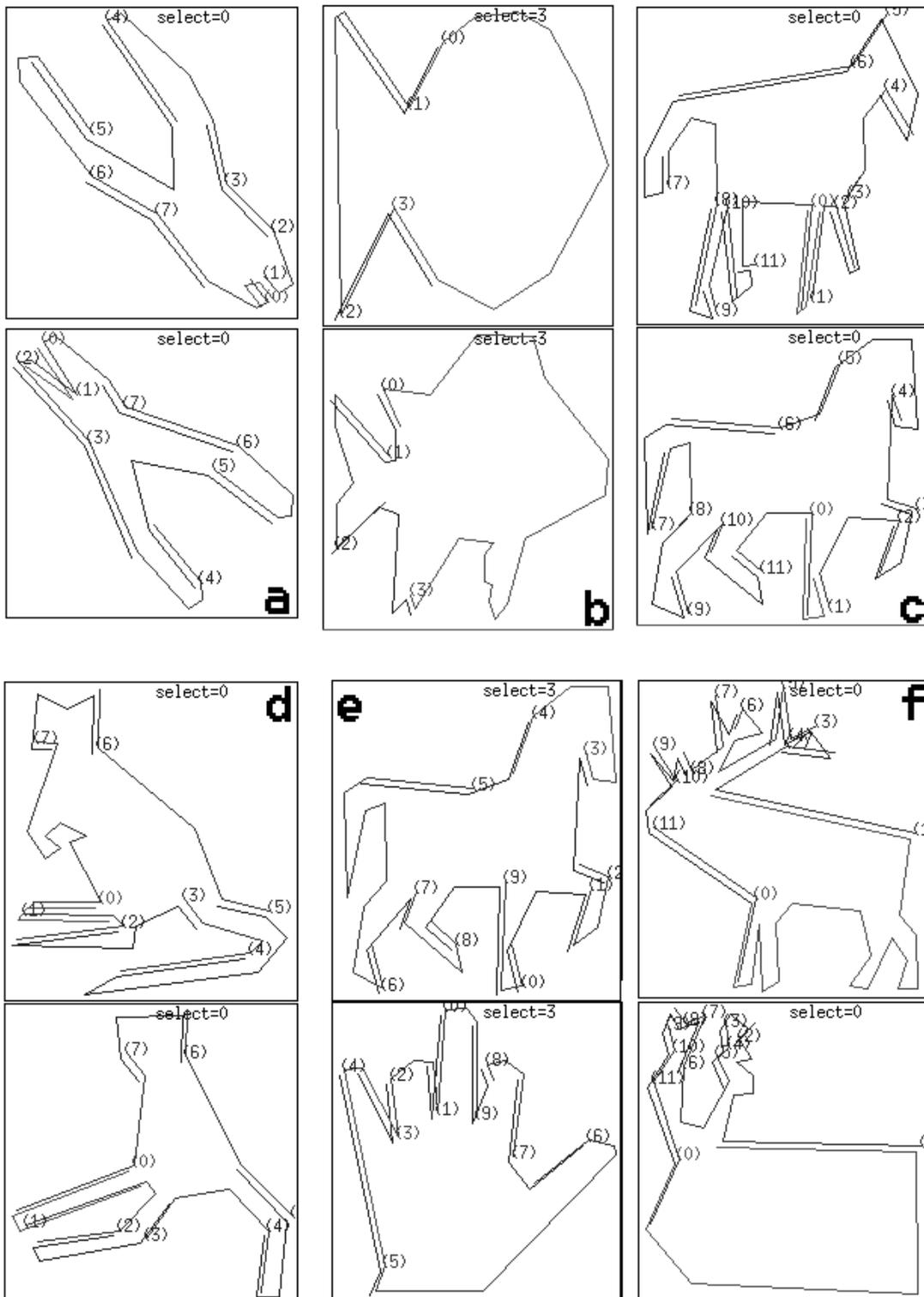


Abbildung 5.7: Ergebnisse der Korrespondenz

- d) Starke Abstraktion. Die Formen besitzen kaum Ähnlichkeit, das untere Objekt ist nur noch mit gutem Willen als Känguruh zu erkennen. Der Algorithmus erstellt sinnvolle Korrespondenzen.
- e) Verschiedene Formen. Der Zusammenhang zwischen diesen Formen kann nur noch teilbasiert erstellt werden, die Gesamtform unterscheidet sich erheblich. Die Korrespondenz ordnet die Teile intuitiv richtig zu.
- f) Überdeckung: Die Formen sind nur in einem Teil ähnlich, nämlich dem Geweih. Auch diese Teilkorrespondenz wird gefunden.

Insgesamt ergeben diese (sowie zahlreiche andere) Tests sehr gute Übereinstimmung mit der intuitiven Zuordnungsweise menschlicher Betrachter.

5.5 Ergebnisse des Formunterschieds

Der Formunterschied wurde definiert als das Maß $\mathcal{M}(k_0)$ der optimalen Korrespondenz $k_0 \in \mathcal{K}$. Es soll nun in Vergleich zu bekannten Ähnlichkeitsmaßen gestellt werden.

Dabei soll sich der Vergleich auf universelle (d.h. ohne explizite Vormodellierung arbeitende) Ähnlichkeitsmaße beschränken, die invariant gegenüber den 'gängigen' Transformationen (Translation, Rotation, Spiegelung, Skalierung) sind. Dieses schließt z.B. solche Verfahren aus, die die Ähnlichkeit zwischen Objekten über den Hausdorffabstand berechnen (dieses Maß ist zwar universell, aber besitzt nicht alle der gewünschten Invarianzeigenschaften)

Zum Vergleich werden die Verfahren von Siddiqi et al. [35], Mokhtarian et al. [31], Basri et al. [2] sowie Arkin et al. [1] herangezogen.

5.5.1 Vergleich mit Siddiqi et al.

Das in [35] vorgestellte Verfahren von Siddiqi et al. basiert auf einer hierarchischen Darstellung von 'shocks' in Skelettierungen zweidimensionaler Objekte, d.h. es ist kein Objektrand- sondern ein Objektflächenbezogenes Verfahren. Obgleich diese Verfahren sich grundsätzlich unterscheiden, werden ähnliche Ergebnisse erzielt. Die in Abbildung 5.8 gezeigten Objekte wurden direkt aus [35] entnommen und können mit der dortigen Tabelle verglichen werden.

Die intuitiv richtige Auswahl ähnlicher Objekte durch das in dieser Arbeit vorgestellte Verfahren zeigt Abbildung 5.13 (rechts). Gezeigt werden die Objekte geringsten Formunterschieds in Bezug auf Tabelle 5.8.

5.5.2 Vergleich mit Mokhtarian et al.

Ein Teil des Verfahren von Mokhtarian et al. [31] wurde schon in Kapitel 3.10.2 vorgestellt. Die dort angesprochene kontinuierliche Kurvenevolution mittels Gaußfaltung wird für den Formvergleich benutzt, indem die Lage der Wendepunkte im Laufe der Evolution als Formbeschreibung herangezogen wird.

Dieses Verfahren ist, wie auch das in dieser Arbeit vorgestellte, als Vorschlag zum MPEG-7 Standard aufgenommen, Kapitel 6 zeigt Ergebnisse beider Verfahren anhand des MPEG-7 Testmaterials. Hier sollen nur Ergebnisse des vorliegenden Formunterschieds auf Basis von Testmaterial aus [31] gezeigt werden, siehe Abbildung 5.9. Dieses Testmaterial besteht aus einer Randkurvdatenbank von 1100 Meerestieren. Die Anfrageobjekte, markiert mit der '1' in Abbildung 5.9 entsprechen in [31] den Objekten 4(a), 4(d), 5(a), 7(a). Die Ergebnisse ähneln sich wiederum stark und sind intuitiv richtig.

5.5.3 Vergleich mit Basri et al.

Ein weiteres Vergleichsmaß wird von Basri et al. in [3] gegeben. Zusätzlich wird dort ein Forderungskatalog wünschenswerter Eigenschaften eines Ähnlichkeitsmaßes erstellt. Basri et al. vergleichen in [3] drei Ähnlichkeitsmaße: 'spring-model', 'linear-model' sowie 'continous deformation model'. Diese Maße beruhen auf verschiedenen Berechnungen von Deformationsenergie bei der Überführung zweier Randkurven ineinander ('morphing'). Dabei werden zur Ermittlung der bestmöglichen Korrespondenzen von ausgezeichneten Punkten der Randkurve lokale Krümmungsberechnungen benötigt. Die daraus resultierenden Schwierigkeiten wurden schon an verschiedenen Stellen dieser Arbeit erwähnt.

Es soll sich nun das Verfahren der vorliegenden Arbeit dem Forderungskatalog aus [3] stellen. Die benutzten Testobjekte stammen aus [3].

1. Forderung an das Ähnlichkeitsmaß bei regulären Polygonen P : je mehr Ecken P besitzt, desto ähnlicher ist es einem Kreis. Abbildung 5.11 zeigt die Ergebnisse des Tests mit dem Formunterschied dieser Arbeit, die Forderung wird erfüllt. Dabei sind die relativen Differenzen des vorliegenden Verfahrens deutlicher als die in [3] erreichten.

Instance		Class Prototype								
										
	->	245	2772	3436	3078	1274	1950	4840	8189	1631
	->	2515	898	4564	4347	2348	1864	5933	7638	3408
	->	3205	5721	3607	3982	2702	3921	3812	7005	3407
	->	3348	5346	1480	806	4191	5242	3896	6424	3161
	->	4868	5376	1537	1906	3890	5097	5402	7403	4137
	->	5279	5433	1789	526	3902	4790	3945	5921	3389
	->	5627	5429	1777	1545	3783	5801	4844	6095	4043
	->	774	2998	2724	4198	626	1465	4979	8862	2569
	->	1587	2814	2861	4749	652	1685	5633	8951	2336
	->	2779	5075	3737	4893	1567	258	4631	8057	1995
	->	6231	9451	4418	4332	4627	5759	519	4611	4145
	->	5007	8250	4860	4202	4218	4815	1998	5938	3522
	->	8668	12057	6201	5854	6295	7175	3186	2914	5720
	->	2406	4935	5385	3868	1772	2251	5320	6124	798
	->	1855	4296	3205	3222	1045	2528	4761	6054	877

Abbildung 5.8: Ergebnisse des Formunterschieds anhand von Testobjekten aus Tabelle 1, S. 27, in [35].

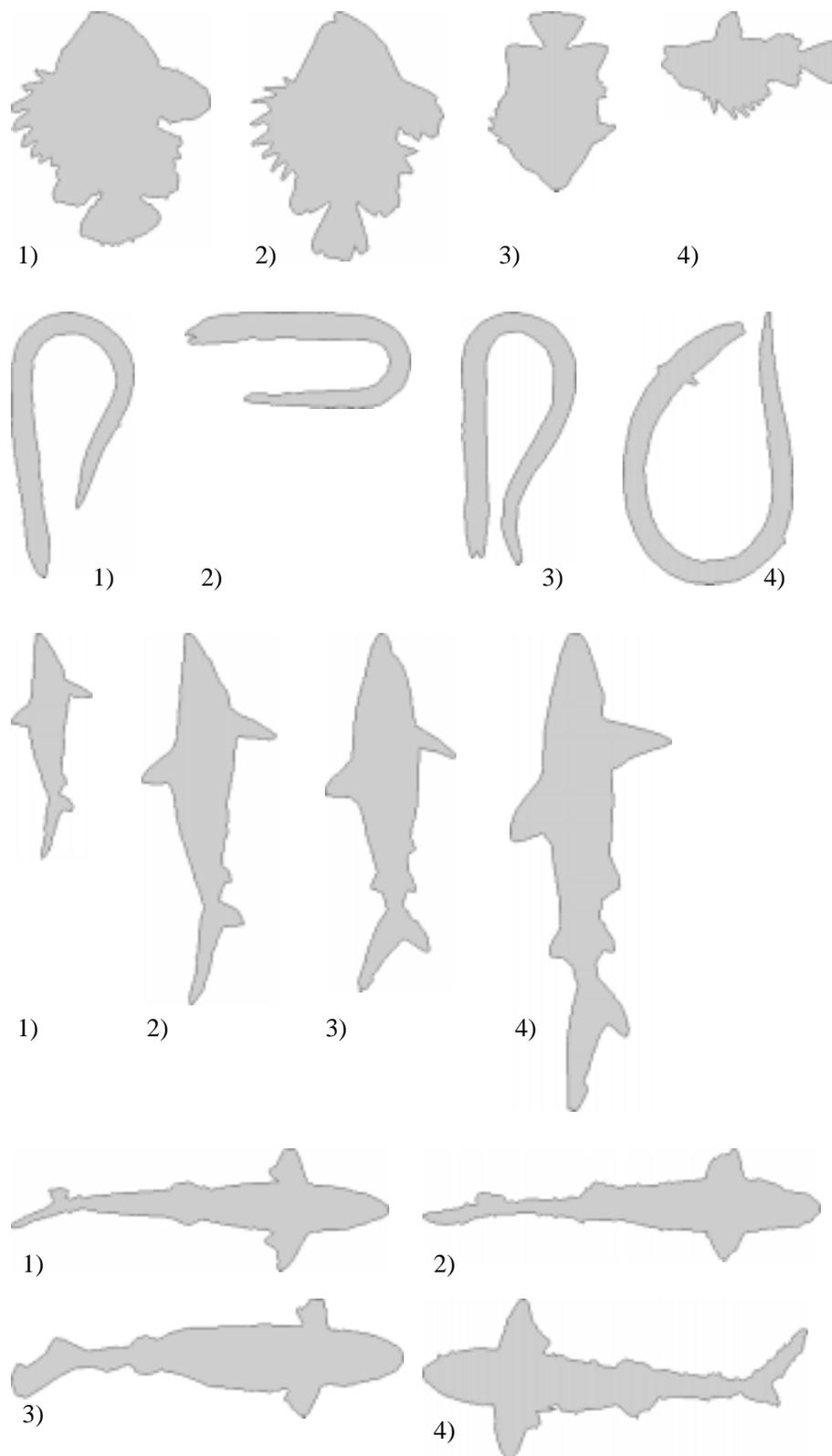


Abbildung 5.9: Ergebnisse des Formunterschieds auf Fischkonturen aus [31]. Die Anfrageobjekte sind durch '1)' gekennzeichnet.

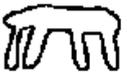
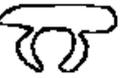
	Table1	Table2	Table3	Table4	Table5	Table6
						
Table1	0	15	57	39	34	61
Table2	15	0	57	44	42	47
Table3	57	57	0	41	42	58
Table4	39	44	41	0	17	44
Table5	34	42	42	17	0	44
Table6	61	47	58	44	44	0

Abbildung 5.10: 'Tischähnliche Objekte' aus [3]

.	▽	□	⬡	⬢
○ →	1220	662	313	193

Abbildung 5.11: Ergebnisse des Formunterschieds anhand von Testobjekten aus Tabelle 1 in [3].

2. Forderung an das Ähnlichkeitsmaß bei strukturellen Veränderungen von Formteilen: eine Verbiegung inmitten eines Formteils muß zu stärkerer Abweichung der Ähnlichkeit zum ursprünglichen Objekt führen als ein Verbiegung des gesamten Teils gegenüber der Gesamtform. Abbildung 5.12 zeigt die Ergebnisse, die wiederum in der relativen Differenz deutlicher ausfallen als die in [3] erreichten.

.		
	216	974

Abbildung 5.12: Ergebnisse des Formunterschieds anhand von Testobjekten aus Tabelle 2 in [3].

3. Das Maß sollte invariant unter euklidischen Transformationen sein. Diese Forderung erfüllt der Formunterschied dieser Arbeit, zusätzlich ist er skalierungsinvariant.
4. Das Maß sollte visuelle Teile der Objekte respektieren. Diese Forderung war eine der Grundlagen dieser Arbeit, sie wird erfüllt. In Zusammenhang mit dieser Forderung kann auch Abbildung 5.12 gesehen werden.

Weitere Forderungen aus [3] sind fragwürdig, z.B. wird von dem Ähnlichkeitsmaß gefordert, die Axiome einer Metrik zu erfüllen. Diese Forderungen werden hier nicht weiter behandelt.

An dieser Stelle sollen zum weiteren Vergleich Ergebnisse des Formunterschieds anhand von Objekten aus [3] gegeben werden: Abbildung 5.10 zeigt Objekte

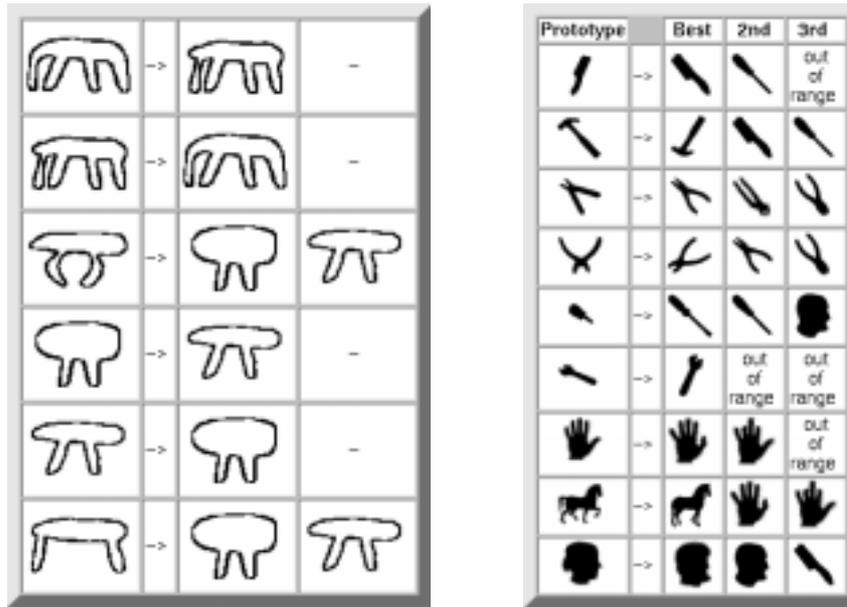


Abbildung 5.13: Objekte geringsten Formunterschieds zu Objekten aus 5.10 und 5.8.

aus [3], Tabelle 4. Die Objekte sollen nach [3] 'tischähnliche Formen' darstellen. In Übereinstimmung mit [3] werden Tisch 4 und 5 als 'sehr ähnlich' klassifiziert, ebenfalls Tisch 1 und 2, d.h. die Tische 1,2 sowie 4,5 Formen zwei Klassen von Objekten. Eine solch deutliche Aufteilung, die in Übereinstimmung mit der Intuition steht, erreicht keines der in [3] vorgestellten Maße. Die besten Übereinstimmung zeigt Abbildung 5.13.

5.5.4 Vergleich mit Arkin et al.

Das von Arkin et al. [1] vorgestellte Vergleichsmaß ist die Grundlage des in dieser Arbeit vorgestellten Polygonvergleiches 5. Der wesentliche Unterschied zu [1] besteht darin, daß Arkin et al. den Polygonvergleich auf die gesamte Form anwenden, das Verfahren dieser Arbeit jedoch auf einzelne Teile. Dieser Unterschied eliminiert einen wesentlichen Nachteil des Verfahrens aus [1], der darin besteht, daß nicht gleichverteilte Störungen auf der Randkurve zu nichtrobustem Verhalten führen.

Ein Beispiel dazu zeigt Abbildung 5.14: der Buchstabe 'H' wurde an nur einer Seite stark gestört (linke Spalte). Die rechte Spalte zeigt die zum Original und zur gestörten Version zugehörigen Tangentenraumdarstellungen. Es zeigt sich,

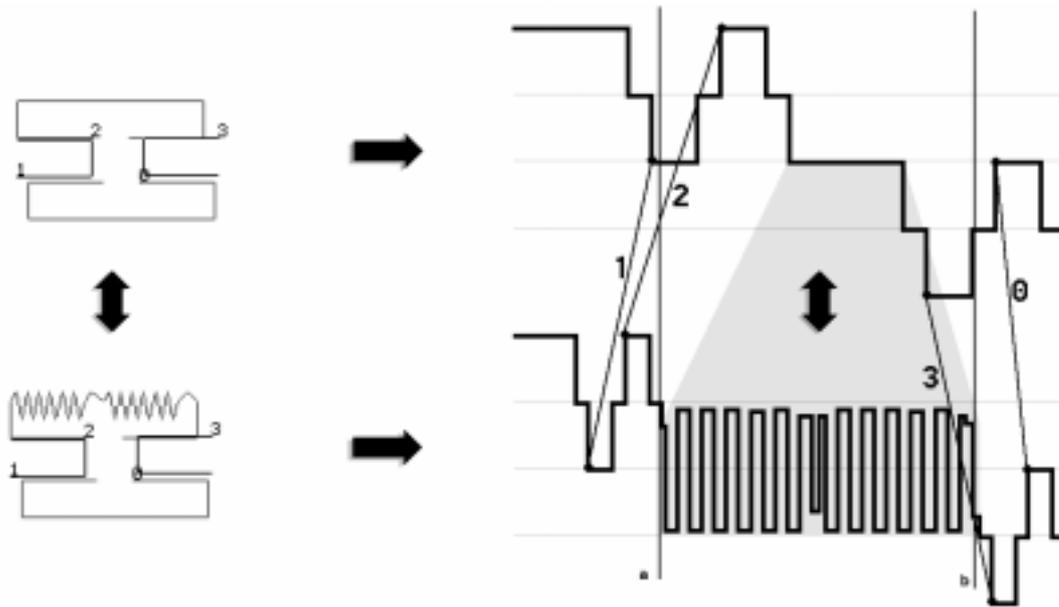


Abbildung 5.14: Ergebnis des Formunterschieds im Vergleich zu Arkin et al. [1]

daß die intuitiv korrespondierenden Teile beider Versionen nicht zur Deckung gebracht werden können, falls nur über die gesamte Kurve gleichmäßig skaliert wird, wie in [1] vorgesehen. Das dortige Verfahren versagt somit im vorliegenden Fall.

Die Indizes der linken Spalte zeigen die ermittelte Korrespondenz des Verfahrens aus dieser Arbeit. Durch die Unterteilung des Buchstabens in vier Bögen wird nur eine Fehlerstrafe für den gestörten Bogen verhängt. Da auch die Tangentenraumdarstellung unterteilt wird (die Indizes zeigen die Startpunkte der mit gleichem Index versehenen Bögen der Formen der linken Spalte), wird durch die Skalierung der Einzelabschnitte vor dem Vergleich die gesamte Kurve nicht gleichmäßig, sondern den Abschnitten angepaßt skaliert.

5.6 Weitere Ergebnisse

Zusätzlich zu den Vergleichen mit anderen Verfahren wurden Tests des Verhaltens des Formunterschieds bei verschieden transformierten Objekten vorgenommen. Dazu wurden die Objekte perspektivisch rotiert, perspektivisch verzerrt, planar verzerrt und mit Störungen versehen.

Erklärung zu den Abbildungen: die erkannten Teile sind gegeneinander Versetzt gezeichnet, um besser identifiziert werden zu können. Die gegen den Uhrzeigersinn

verlaufenden Indizes zeigen die jeweilige Korrespondenz.

5.6.1 Perspektivische Rotation

Das Experiment 'Hand 1' in Abbildung 5.15 zeigt in der rechten Spalte eine Rotation der Hand A um die x -Achse, gefolgt von einer Projektion in die xy -Ebene, d.h. die Objekte werden in vertikaler Richtung abhängig vom Cosinus der Rotationswinkel gestaucht. Es wurde um 45, 70 und 80 Grad rotiert.

5.6.2 Perspektivische Verzerrung

Abbildung 5.16 zeigt Ergebnisse des Vergleiches dreier Hände, die sich voneinander durch Verzerrung einzelner visueller Teile unterscheiden, d.h. die Finger werden verschieden abgespreizt bzw. in Richtung der randdefinierenden Projektion geknickt.

5.6.3 Planare Verzerrung

Das dritte Experiment (Abbildung 5.17) zeigt ein Pferd, dessen visuelle Teile planar verzerrt wurden, d.h. die einzelnen Gliedmaße sind in verschiedenen Bewegungsstellungen gezeigt.⁵ Die Rösser unterscheiden sich durch planare Verzerrung von Kopf, Fesseln und Schweif.

5.6.4 Rauschen

Im Puma-Experiment in Abbildung 5.18 wurden drei verschieden verrauschte Pumasilhouetten verglichen. Die Formen ergeben sich durch: A) keine Störung, B) viele aber geringe Störungen, C) wenige aber große Störungen. Die Störungen wurden durch zufälliges Versetzen von Randpunkten erzeugt.

⁵Ein geübter Reiter wird beruhigt feststellen, daß die Bewegungsstudien nicht am lebenden Tier, sondern mit einer mangelhaften Pferdesimulation erstellt wurden.

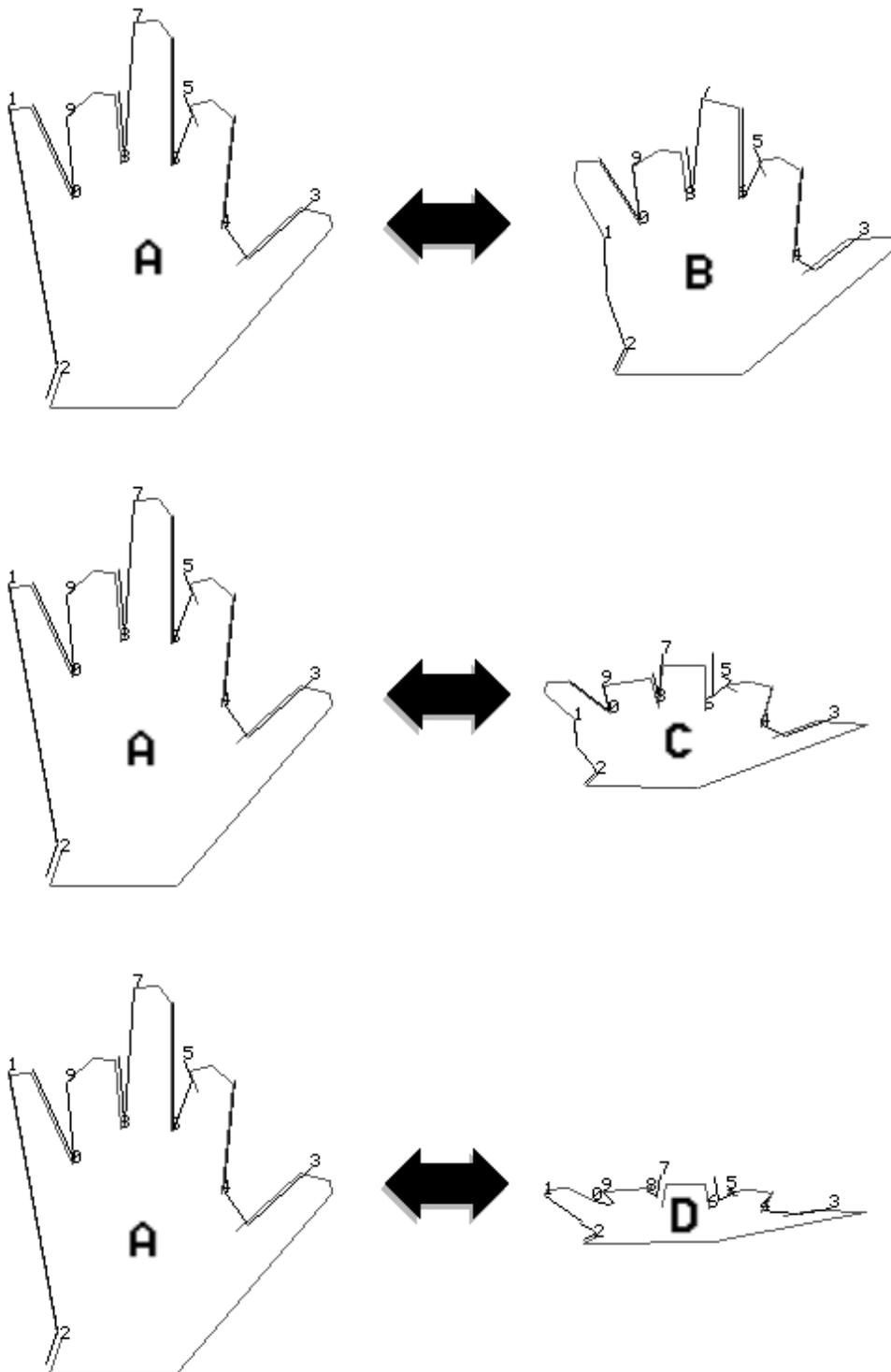


Abbildung 5.15: Experiment Hand1: Vergleichsmaße: $d(a,b)=381$, $d(a,c)=2309$, $d(a,d)=4237$

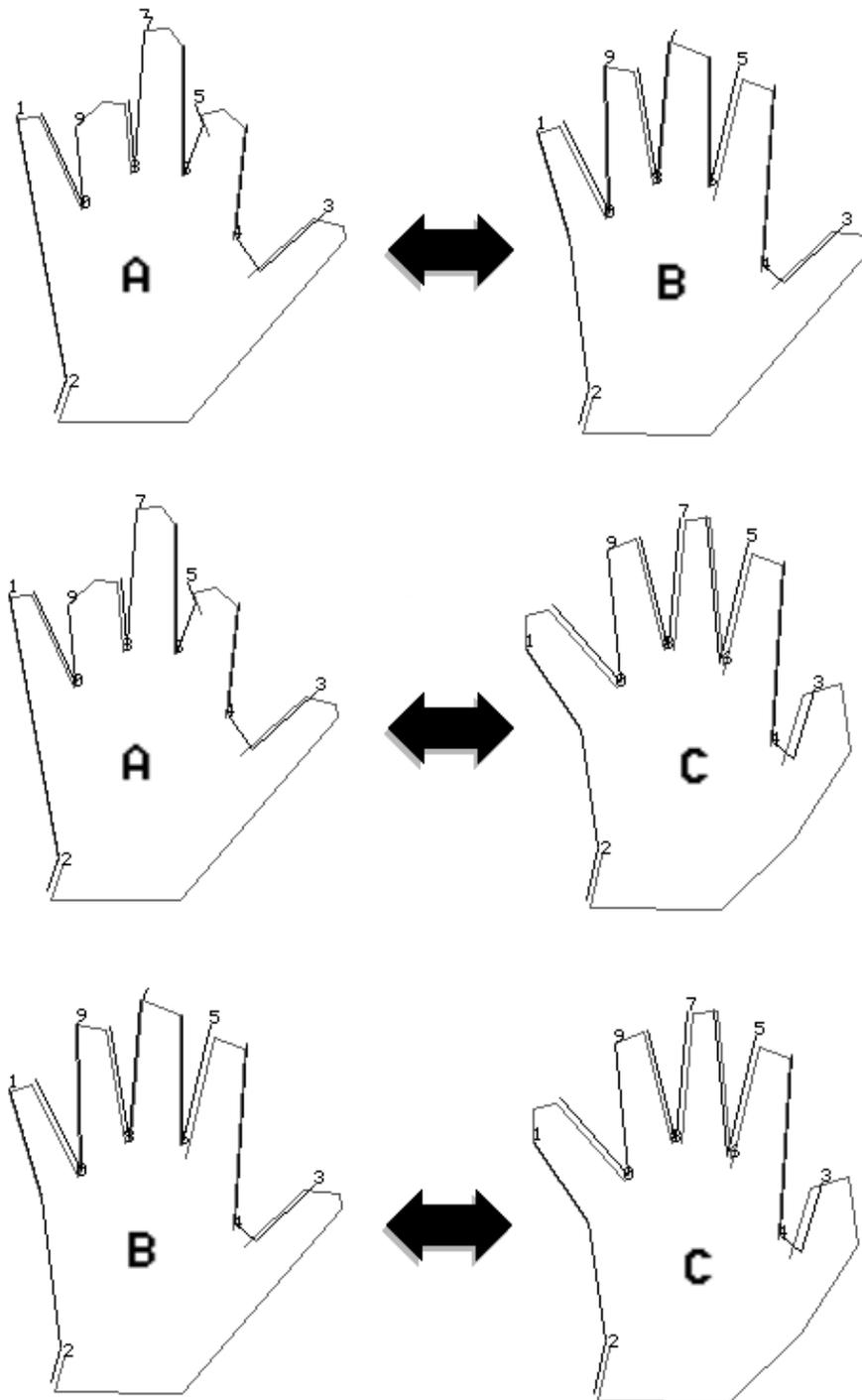


Abbildung 5.16: Experiment Hand2: Vergleichsmaße: $d(a,b)=1998$, $d(a,c)=2080$, $d(b,c)=519$

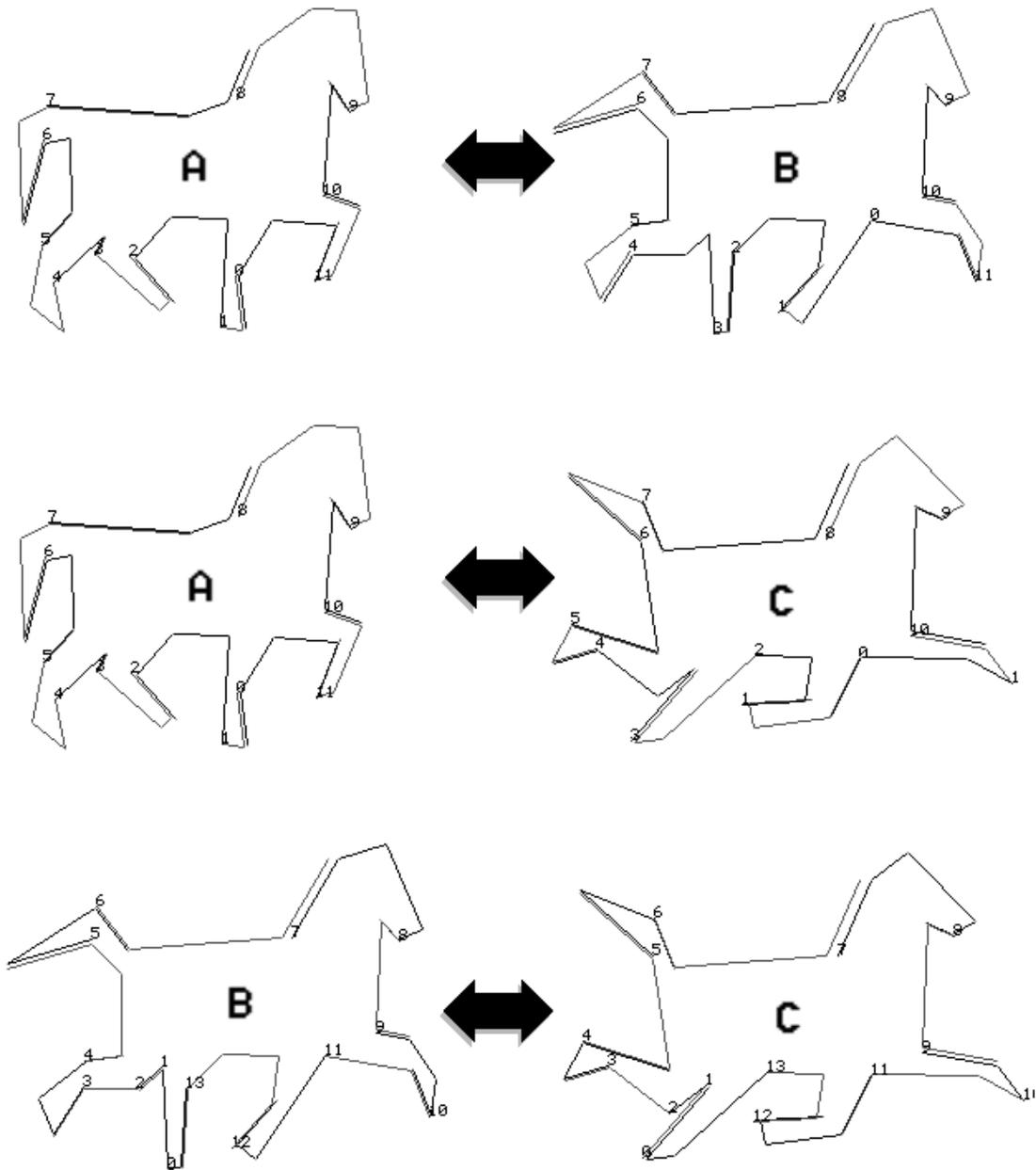


Abbildung 5.17: Experiment Pferd: Vergleichsmaße: $d(a,b)=2188$, $d(a,c)=3471$, $d(b,c)=1156$

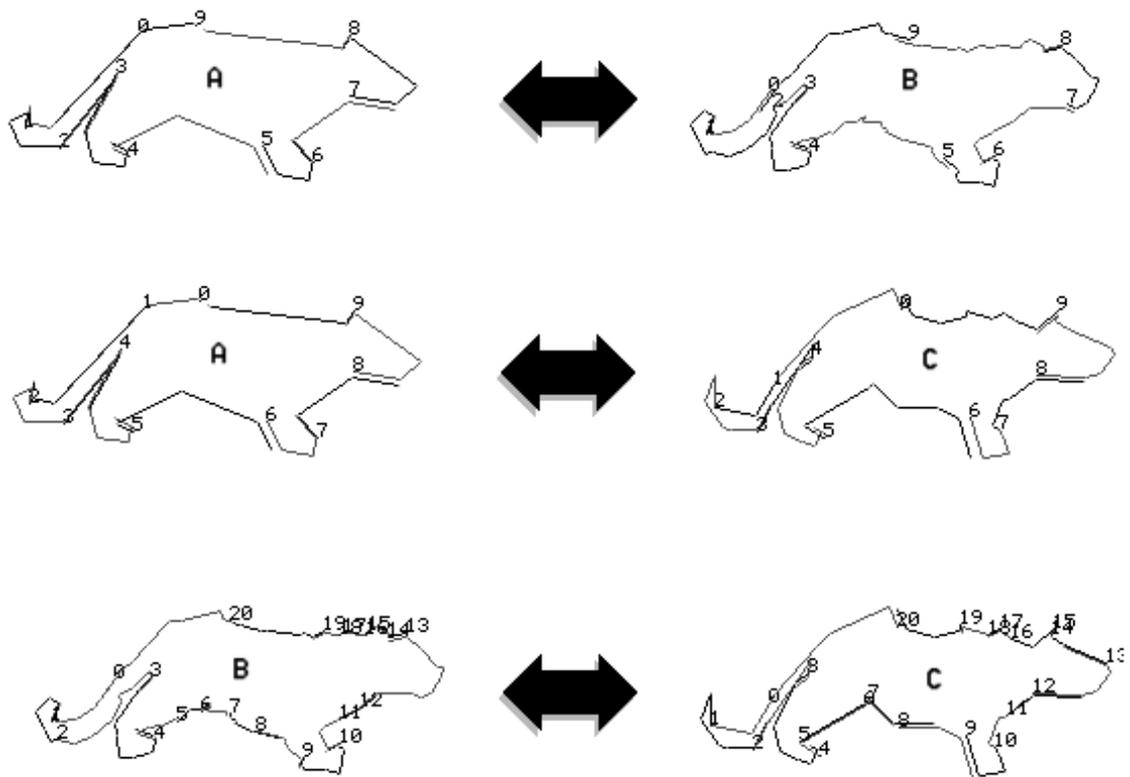


Abbildung 5.18: Experiment Puma: Vergleichsmaße: $d(a,b)=1458$, $d(a,c)=1698$, $d(c,b)=1650$

Kapitel 6

Die Datenbank, MPEG-7

6.1 Die Datenbank

Die Datenbank ist die praktische Umsetzung der Ergebnisse der vorherigen Kapitel. Ihre Beschreibung ist die Zusammenfügung der bisher beschriebenen Module.

Ziel dieser Arbeit ist es, wie schon einleitend erwähnt, einen Beitrag zur Definition des MPEG-7 Standards zu liefern. Hierzu müssen aufwendige Tests, welche durch die MPEG-Kommission festgelegt werden, in normaler Personalcomputerumgebung durchgeführt werden. Die Erstellung eines Datenbanksystems, das unter praktischen Bedingungen robust lauffähig sein muß, gehört somit zum Umfang dieser Arbeit.

Das gesamte System wurde in der Entwicklungsphase in JAVA implementiert. Dem Vorteil der stark objektorientierten Entwicklungsumgebung stand beim Praxiseinsatz die fehlende Geschwindigkeit störend gegenüber, das gesamte System mußte in C und C++ nachimplementiert werden. Vom JAVA-Ursprungscode blieb dabei nur die grafische Oberfläche übrig. Um das System offen für Erweiterungen zu halten, wurde die Schnittstelle zum Kerndatenbanksystem in CORBA implementiert. Denkbare Erweiterungen sind Ergänzungen z.B. von Textursuchmaschinen.¹

Die Module zeigt Abbildung 6.1

Die Eingabe kann entweder anhand einer Skizze (Zeichnung per Maus auf dem Bildschirm) oder durch einen vorliegenden Randkurvendatensatz erfolgen. Die

¹eine solche Kombination des vorliegenden Systems mit einer Texturdatenbank wurde in Zusammenarbeit mit der Siemens-AG erstellt. Die Beschreibung sprengt den Rahmen dieser Arbeit.

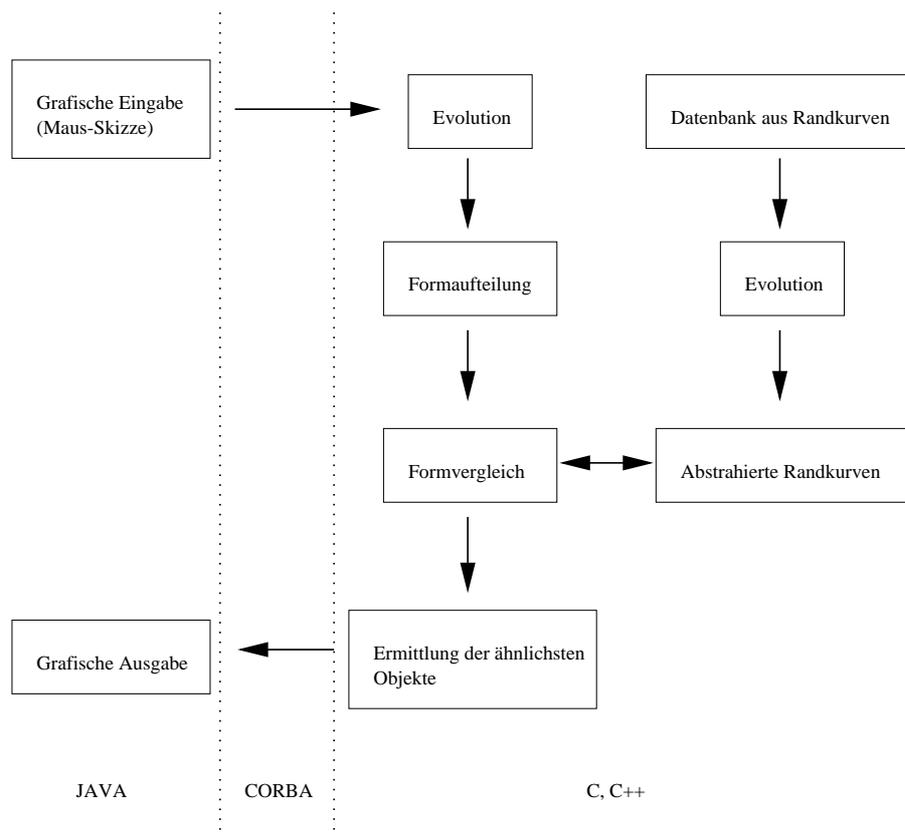


Abbildung 6.1: Aufbau der Datenbank

Bearbeitung erfolgt durch die Arbeitsschritte Abstraktion (Evolution), Formaufteilung und Formvergleich. In der momentanen Version wird die Eingabe mit jedem Objekt der Datenbank vorgenommen. Eine geplante Ergänzung wird die Objekte der Datenbank in Klassen aufteilen, ein Vorauswahl durch Vergleich mit jeweils nur einem Vertreter pro Klasse erhöht die Arbeitsgeschwindigkeit erheblich. Abbildung 6.2 zeigt die grafische Oberfläche der Datenbank.

6.2 MPEG-7

Der Video und Audiostandard MPEG-7 sieht vor, Daten inhaltsbasiert zu codieren und Anfragen an diese Inhalte zu stellen (\rightarrow <http://drogo.csel.stet.it/mpeg>). Der Standard besteht u.a. aus verschiedenen *Deskriptoren*, durch welche die Eigenschaften der Inhalte repräsentiert werden, dazu zählen Textur-, Bewegungs- und Formdeskriptoren. Das System der vorliegenden Arbeit wurde als Vorschlag im Bereich Formdeskriptoren angenommen und steht zur Zeit der Drucklegung

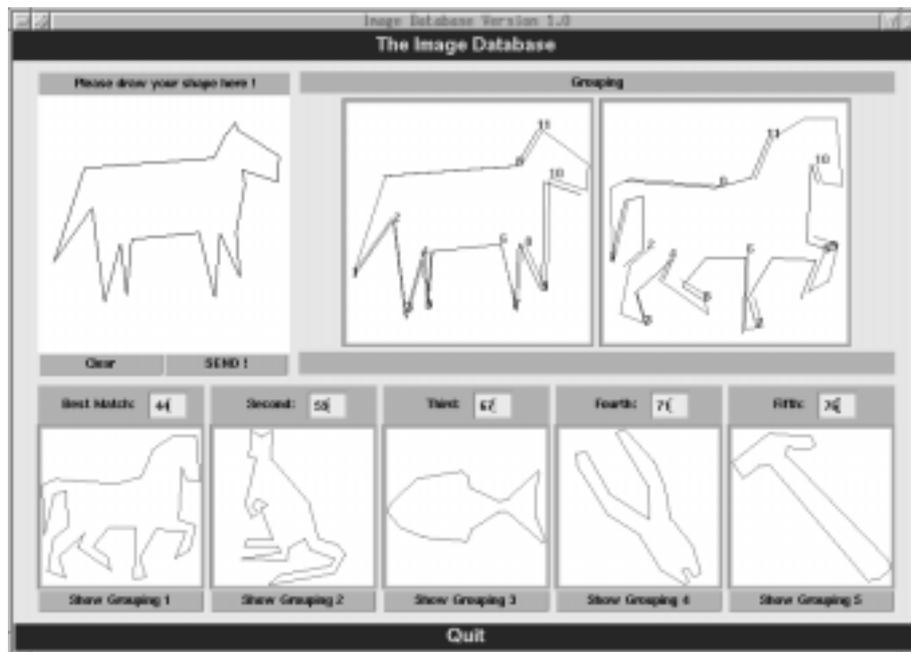


Abbildung 6.2: Die Oberfläche der Datenbank

(Dezember 1999) in Konkurrenz zu fünf weiteren Verfahren. Die Auswahl wird anhand von festgelegten Tests vorgenommen. Die Ergebnisse der Tests werden in diesem Kapitel beschrieben, sie zeigen das Verhalten des Verfahrens in praktischer Umgebung.

6.2.1 Überblick

Das Kernexperiment 'CE-Shape-1' für Formdeskriptoren ist definiert unter N2690 des MPEG-7 Standards. Ziel dieses Kernexperimentes ist die Ermittlung der Qualität der Formdeskriptoren anhand von geschlossenen Konturen unter verschiedenen Transformationen. Der in dieser Arbeit vorgestellte Deskriptor hat innerhalb von MPEG-7 die Arbeitsbezeichnung P298. Der folgende Vergleich bezieht sich auf die Ergebnisse der MPEG-Tagung in Vancouver im Juli 1999. Es waren dort als Formdeskriptoren zugelassen:

P162-Turning angle shapes	no results	IBM
P320-Curvature scale-space	M4731	Mitsubishi ITE-VIL
P567-Wavelet contour descriptor	M4740	Heinrich Hertz Institute Berlin
P687-Zernike moments	M4810	Hanyang University
P517-Multilayer eigenvector	M4793	Hyundai
P144-Geometric hashing	no results	IBM
P579-Normalised Contour	no results	Tektronix inc.
P298-Correspondence of visual parts	(dieses Kapitel)	Hamburg University and Siemens
P??-DAG Ordered Trees	M4751	Mitsubishi and Princeton University

6.2.2 Test A: Robustheit gegenüber Skalierung und Rotation

6.2.2.1 A1: Skalierung

Die Datenbank enthält 420 Formen, gebildet aus 70 Grundformen in 6 verschiedenen Skalierungsstufen. Die Skalierungsfaktoren betragen 2, 1, 0.3, 0.25, 0.2, 0.1. Jedes der 420 Objekte wird als Eingabeobjekt in die Datenbank eingespeist, die besten 6 Vergleiche jedes Eingabeobjektes werden bewertet, d.h. das bestmögliche Ergebnis besteht aus 2520 Treffern.

Bemerkung:

Etwa 17 Formen der Skalierungsstufe 0.1 sind durch die Verkleinerung derart verändert, daß ihre Form keinen intuitiven Bezug mehr zur Ausgangsform hat bzw. ein größerer Formbezug zu Objekten anderer Klassen gegeben ist. Ein Beispiel dazu zeigt Abbildung 6.3. Es ist daher nur eine maximale Trefferrate von $2520 \Leftrightarrow 170 = 2350 = 93\%$ erreichbar.

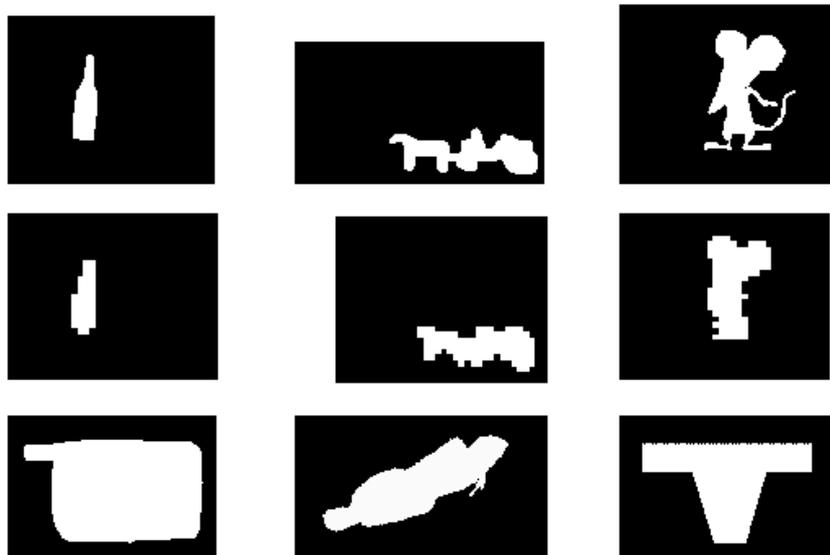


Abbildung 6.3: Deformation durch digitale Skalierung

6.2.2.2 A2: Rotation

Die Datenbank enthält 420 Formen, gebildet aus 70 Grundformen in 6 verschiedenen Rotationsstufen. Die Rotationswinkel betragen (in Grad): 0,9,36,45,90,150.

Jedes der 420 Objekte wird als Eingabeobjekt in die Datenbank eingespeist, die besten 6 Vergleiche jedes Eingabeobjektes werden bewertet, d.h. das bestmögliche Ergebnis besteht aus 2520 Treffern.

6.2.2.3 Ergebnisse aus A1 und A2

	P687	P567	P517	P320	<u>P298</u>	DAG
A1	92.54	88.04	92.42	89.76	88.65	?
A2	99.60	97.46	100.0	99.37	100.0	?
Durchschnitt	96.07	92.75	96.211	94.57	94.33	85

6.2.3 Test B: Ähnlichkeit

6.2.3.1 Beschreibung

Die Datenbank enthält 1400 Formen, gebildet aus 70 Grundformen in 20 verschiedenen Verzerrungen bzw. beliebigen Veränderungen. Jedes der 1400 Objekte wird als Eingabeobjekt in die Datenbank eingespeist, die Anzahl von Objekten der gleichen Klasse innerhalb der ersten 40 besten Treffer wird gezählt (bull's-eye-test). Das bestmögliche Ergebnis besteht somit aus 28000 Treffern.

Bemerkungen:

Wiederum läßt das Testmaterial eine Trefferquote von 100% nicht zu, da die Objekte innerhalb einer Klasse z.T. derart differierten, daß eine größere Ähnlichkeit zu Objekten anderer Klassen gegeben ist. Der Zusammenhang innerhalb der Klassen besteht aus semantischem Wissen, welches jedoch nicht geprüft werden soll. Beispiele zeigt Abbildung 6.4.

Bis auf wenige Ausnahmen sind die Ergebnisse des vorliegenden Verfahrens intuitiv richtig, auch wenn keine Klassenzugehörigkeit erkannt werden kann (da sie im Testmaterial nicht vorhanden ist).

Der Ähnlichkeitstest spiegelt am ehesten die Umgebung des praktischen Einsatzes eines Systems innerhalb einer MPEG-7 Anwendung. Er wird jedoch im Rahmen der Testauswertung nicht stärker gewertet.

6.2.3.2 Ergebnisse

	P687	P567	P517	P320	<u>P298</u>	DAG
B	70.221	67.76	70.33	75.44	76.45	60



Abbildung 6.4: Die Objekte haben nur semantischen Zusammenhang, ihre Randkurven sind nicht formähnlich.

6.2.4 Test C: Bewegung und Deformation

6.2.4.1 Beschreibung

Die Datenbank besteht aus zwei Klassen von Formen: 1100 verschiedenen Fischen sowie 200 Formen, welche einen Fisch in verschiedenen Bewegungen zeigt. Eine Seitenansicht dieses Fisches wurde als Eingabe in die Datenbank eingespeist. Die Anzahl der Objekte aus dieser Fischklasse innerhalb der ersten Treffer wurde gezählt.

Bemerkung:

Auch dieser Test enthält Datenmaterial, welches nur mit semantischen Vorwissen eine Trefferquote von 100% zuläßt. In 14 der 200 Fälle wird der Fisch in Bewegungsstufen gezeigt, welche keinerlei Formähnlichkeit mit dem Eingabeobjekt aufweisen. Andererseits enthält die Klasse der anderen 1100 Fische Objekte, die dem Anfrageobjekt stark ähneln. Es konnte somit eine maximale Trefferquote von 93% (=186/200) erreicht werden. Abbildung 6.5 enthält Beispiele beider Klassen. Es ist deutlich zu sehen, daß die Fische der unteren Reihe dem Eingabeobjekt (obere Reihe links) stärker ähneln als die Bewegungsstudien dieses Fisches, gezeigt in der oberen Reihe, Bilder 2,3,4. Die Fische der unteren Reihe entstammen dem Ergebnis der vorliegenden Arbeit.

6.2.4.2 Ergebnis

	P687	P567	P517	P320	<u>P298</u>	DAG
C	94.5	93.0	88.0	96.0	92.0	83.0

6.2.5 Gesamtergebnis

Die vorgeschriebene Auswertung sieht bei der Endauswertung eine gleiche Gewichtung aller Tests vor. Das führt zu folgendem Gesamtergebnis:

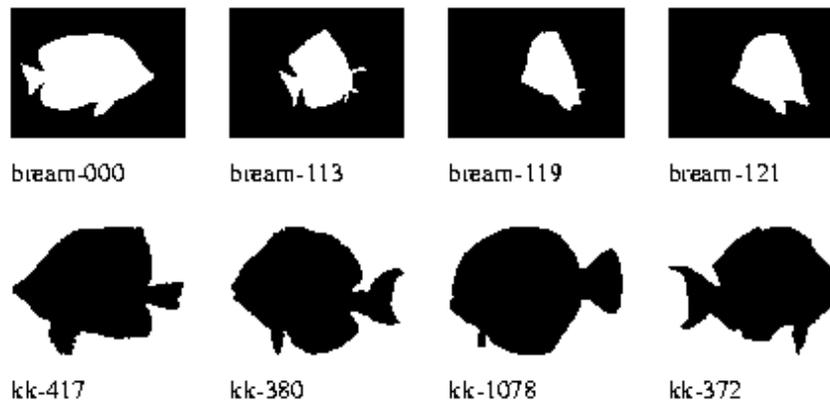


Abbildung 6.5: Die Objekte der unteren Reihe sind zum Anfrageobjekt 'bream-000' ähnlicher als die im Testmaterial vorgesehenen Objekte der oberen Reihe

	P687	P567	P517	P320	<u>P298</u>	DAG
A,B,C	86.93	84.50	84.85	88.67	87.59	76

Wird die Auswertung dagegen nach Anfragen gewichtet, d.h. bekommt insbesondere Test B stärkeres Gewicht, so ergibt sich durch:

$$840/2241A + 1400/2241B + 1/2241C$$

	P687	P567	P517	P320	<u>P298</u>	DAG
A,B,C	79.92	77.14	80.04	82.62	83.16	69.38

Es zeigt sich, daß durch stärkere Bewertung der Tests auf Übereinstimmung mit menschlicher Intuition das Verfahren der vorliegenden Arbeit die beste Trefferquote erzielt.

Die Entscheidung, ob das hier vorgestellte Verfahren als Standarddeskriptor innerhalb von MPEG-7 angenommen wird, fällt auf einer Tagung innerhalb der nächsten Monate nach der Drucklegung.

Kapitel 7

Zusammenfassung

Die zunehmende Bedeutung der Automatisierung visueller Kontrollprozesse verlangt Verfahren zur formbasierten Identifikation in digitalen Bildern. Die vorliegende Arbeit zeigt dazu einen neuen Ansatz, der im Gegensatz zu klassischen Algorithmen der Bildverarbeitung die digitale Natur der Objektvorlagen respektiert und im Formvergleich mit der menschlichen Intuition übereinstimmt. Die Arbeit ist ein Beitrag zum internationalen Videostandard MPEG-7, Zielsetzung neben theoretischer Entwicklung und Analyse ist damit die Implementierung robuster und schneller Verfahren.

Die untersuchten Objekte werden als digitales Randpolygon dargestellt, alle Operationen verbleiben in dieser Darstellung. Es werden zwei unabhängige Module entwickelt, die *Kurvenevolution* sowie der *Formvergleich*.

Die Kurvenevolution ist ein Beitrag zur Scale Space Theorie: einer gegebenen Funktion f^0 , hier dem digitalen Randpolygon, wird eine Folge $f^{i=0..n}$ von Funktionen zugeordnet, in welcher mit aufsteigendem Index i gewünschte Eigenschaften der Ursprungskurve f^0 gefiltert werden. In der vorliegenden Arbeit wird die Evolution zur Extraktion bedeutender Randpunkte benutzt, es wird eine automatische Abstraktion der ursprünglichen Form erstellt. Durch Aufspaltung in konvexe und konkave Teilbögen dieser Abstraktion kann daraufhin die Form unterteilt werden, ein grundlegender Schritt in der Formerkennung. Das Evolutionsverfahren ist Kern dieser Arbeit, seine Eigenschaften und damit verbundene Vorteile gegenüber klassischen Verfahren werden aufgezeigt.

Im zweiten Modul, dem Formvergleich, wird ein Maß von Randpolygonen entwickelt, anhand dessen eine Korrespondenz zwischen den Formteilen zweier zu

vergleichender Objekte erstellt und eine Aussage über die Ähnlichkeit dieser Objekte getroffen werden kann.

Die Module werden in einer Beispielanwendung, der Bilddatenbank, zusammengefügt, durch experimentelle Ergebnisse wird die Übereinstimmung mit der menschlichen Intuition gezeigt.

Das in dieser Arbeit entwickelte Verfahren wurde als Vorschlag im Bereich Formdeskriptoren des MPEG-7 Standards anerkannt. Es steht damit in Konkurrenz zu sechs weiteren Verfahren. Die Ergebnisse der Vergleiche zu diesen Verfahren schließen die Arbeit ab.

Literaturverzeichnis

- [1] M. Arkin, L. P. Chew, D. P. Huttenlocher, K. Kedem, and J. S. B. Mitchell. An efficiently computable metric for comparing polygonal shapes. *IEEE Trans. PAMI*, 13:209–206, 1991.
- [2] R. Basri, L. Costa, D. Geiger, and D. Jacobs. Determining the similarity of deformable shapes. In *Proc. IEEE Workshop on Physics-Based Modeling in Computer Vision*, pages 135–143, 1995.
- [3] R. Basri, L. Costa, D. Geiger, and D. Jacobs. Determining the similarity of deformable shapes. *Vision Research*, 38:2365–2385, 1998.
- [4] J. Beusmans, D.D. Hoffman, and B.M. Bennett. Description of solid shape and its inference from occluding contours. *Journal of the Optical Society of America*, A. 4:1155–1167, 1987.
- [5] A.M. Bruckstein, G. Shapiro, and C. Shaked. Evolutions of planer polygons. *Int. J. of of Pattern Recognition and AI*, 9:991–1014, 1995.
- [6] A. Brunn, U. Weidner, and W. Förstner. Model-based 2d-shape recovery. In *Proc. of 17. DAGM Conf. on Pattern Recognition (Mustererkennung)*, pages 260–268, Bielefeld, Springer-Verlag, Berlin, 1995.
- [7] S. Fejes and A. Rosenfeld. Discrete active models and applications. *Pattern Recognition*, 30:817–835, 1997.
- [8] G. Frederix and E. J. Pauwels. Automatic interpretation based on robust segmentation and shape-extraction. In *Proc. of 3rd Int. Conf. on Visual Infrmation Systems*, Amsterdam, June 1999, submitted.
- [9] H. Freeman. Shape description via the use of critical points. *Pattern Recognition*, 10:159–166, 1978.
- [10] M.A. Grayson. The heat equation shrinks embedded plane curves to round points. *Journal of Differential Geometry*, 26:285–314, 1987.

- [11] M. Brady H. Asada. The vurvature primal sketch. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8:2–14, 1986.
- [12] A. Held and K. Abe. On the decomposition of binary shapes into meaningful parts. *Pattern Recognition*, 27:637–647, 1994.
- [13] D. D. Hoffman and W. A. Richards. Parts of recognition. *Cognition*, 18:65–96, 1984.
- [14] D. D. Hoffman and M. Singh. Saliency of visual parts. *Cognition*, 63:29–78, 1997.
- [15] B. B. Kimia, A. R. Tannenbaum, and S. W. Zucker. Shapes, shocks, and deformations. i: The components of shape and the reaction-diffusion space. *Int. J. Computer Vision*, 15:189–224, 1995.
- [16] J. J. Koenderink and A. J. Doorn. The shape of smooth objects and the way contours end. *Perception*, 11:129–137, 1981.
- [17] Y. Kurozumi and W.A. Davis. Polygonal approximation by minimax method. *Computer Graphics and Image Processing*, 19:248–264, 1982.
- [18] R. Lakämper, L. J. Latecki, and U. Eckhardt. Contour-based shape similarity. In *Proc. SPIE's Vision Geometry VII*, volume 3454, pages 202–213, San Diego, July 1998.
- [19] R. Lakämper and F. Seytter. Manipulation objektbasiert codierter Bilder als Anwendungsbeispiel neuer Videostandards. In *Proc. DAGM Conf. on Pattern Recognition (Mustererkennung)*, pages 427–434, Springer-Verlag 1997.
- [20] L. J. Latecki, R.-R. Ghadially, R. Lakämper, and U. Eckhardt. Continuity of the discrete curve evolution. In *SPIE and SIAM Conf. on Vision Geometry VIII*, July 1999, to appear.
- [21] L. J. Latecki and R. Lakämper. Convexity rule for shape decomposition based on discrete contour evolution. *Computer Vision and Image Understanding*, 73:441–454, 1999.
- [22] L. J. Latecki and R. Lakämper. Contour-based shape similarity. In D. P. Huijsmans and A. W. M. Smeulders, editors, *Proc. of Int. Conf. on Visual Information Systems*, volume 1614, pages 617–624, Amsterdam, June 1999.

- [23] L. J. Latecki and R. Lakämper. Polygon evolution by vertex deletion. In *Proc. of Int. Conf. on Scale-Space Theories in Computer Vision*, Corfu, Greece, September 1999.
- [24] L. J. Latecki and R. Lakämper. Discrete approach to curve evolution. In *Proc. DAGM Conf. on Pattern Recognition (Mustererkennung)*, pages 85–92, Stuttgart, September 1998.
- [25] L. J. Latecki and R. Lakämper. Shape decomposition and shape similarity measure. In *Proc. DAGM Conf. on Pattern Recognition (Mustererkennung)*, pages 367–376, Stuttgart, September 1998.
- [26] L. J. Latecki and R. Lakämper. Shape similarity measure based on correspondence of visual parts. *IEEE Trans. Pattern Analysis and Machine Intelligence*, submitted.
- [27] L. J. Latecki and R. Lakämper. Comparing planar shapes with application to object retrieval in image databases. *Pattern Recognition*, to appear.
- [28] L. J. Latecki, R. Lakämper, and U. Eckhardt. <http://www.math.uni-hamburg.de/home/lakaemper/shape>. In (*internet page*), 1998.
- [29] L. J. Latecki and A. Rosenfeld. Supportedness and tameness: Differentialless geometry of plane curves. *Pattern Recognition*, 31:607–622, 1998.
- [30] F. Mokhtarian. Search for similar shapes in the squid system. *Internetpage*, <http://www.ee.surrey.ac.uk/Research/VSSP/imagedb/demo.html>, 1998.
- [31] F. Mokhtarian, S. Abbasi, and J. Kittler. Efficient and robust retrieval by shape content through curvature scale space. In A. W. M. Smeulders and R. Jain, editors, *Image Databases and Multi-Media Search*, pages 51–58. World Scientific Publishing, Singapore, 1997.
- [32] M. Spiel. Seminar 1996. <http://www.informatik.uni-mannheim.de>, internet-page, 1996.
- [33] U. Ramer. An iterative procedure for the polygonal approximation of plane curves. *Computer Graphics and Image Processing*, 1:244–256, 1972.
- [34] K. Siddiqi and B. B. Kimia. Parts of visual form: Computational aspects. *IEEE Trans. PAMI*, 17:239–251, 1995.

- [35] K. Siddiqi, A. Shokoufandeh, S. J. Dickinson, and S. W. Zucker. Shock graphs and shape matching. *Int. J. of Computer Vision*, to appear; <http://www.cim.mcgill.ca/siddiqi/journal.html>.
- [36] K. Siddiqi, K. Tresness, and B. B. Kimia. Parts of visual form: Ecological and psychophysical aspects. In *Proc. IAPR's Int. W. on Visual Form*, Capri, 1994.
- [37] K. Siddiqi, K. J. Tresness, and B. B. Kimia. Parts of visual form: Psychophysical aspects. *Perception*, 25:399–424, 1996.
- [38] K. Wall and P.E. Danielsson. A fast sequential method for polygonal approximations of digitized curves. *Computer Graphics and Image Processing*, 28:220–227, 1984.
- [39] J. Weickert. A review of nonlinear diffusion filtering. In B. M. ter Haar Romeny, L. Florack, J. Koenderink, and M. Viergever, editors, *Scale-Space Theory in Computer Vision*, pages 3–28. Springer, Berlin, 1997.
- [40] A.P. Witkin. Scale-space filtering. In *Proc. IJCAI*, volume 2, pages 1019–1022, 1983.
- [41] C. T. Zahn and R. Z. Roskies. Fourier descriptors for plane closed curves. *IEEE Trans. on Computers*, 21:269–281, 1972.
- [42] P. Zhu and P.M. Chirlian. On critical point detection of digital shapes. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 17:737–748, 1995.

Lebenslauf

Name: Rolf Lakämper
Geburtstag: 6.12.1965
Geburtsort: Gütersloh, NRW
Familienstand: ledig

Schulbildung

1972-1976 Grundschule Gütersloh
1976-1985 Städt. Gymnasium Gütersloh, Abschluß Abitur

Berufliche Tätigkeiten

1985-1987 Zivildienst
1983 freiberuflicher Programmierer der ARA-Soft Gütersloh
1984 Gründung der Software GbR micro-partner/Magic-Bytes
(Bereich Computerspiele), Leitung der Softwareentwicklung
1987 Software-Entwicklungsleiter der Rainbow-Arts GmbH Düsseldorf
Leitung der GoldenGoblin-GmbH
ab 1989 weitere freiberufliche Tätigkeit als Programmierer und in der
Projektentwicklung bei Rainbow-Arts/GoldenGoblins
ab 1997 Wissenschaftlicher Mitarbeiter am FB Mathematik, Universität Hamburg

Ausbildung

1989 Studium der angewandten Mathematik an der Universität Hamburg,
Nebenfächer Informatik, Physik
1996 Viermonatiger Aufenthalt bei der Siemens AG München,
Abteilung Zentrale Forschung und Entwicklung
1996 Diplom in Zusammenarbeit mit der Siemens AG. Thema:
Untersuchung und Implementation morphologischer
Verfahren in der Bildverarbeitung
seit 1997 Promotionsstudium am FB Mathematik Uni Hamburg

Veröffentlichungen

1. R. Lakämper and F. Seytter, Manipulation objektbasiert codierter Bilder
als Anwendungsbeispiel neuer Videostandards, Proc. DAGM Conf. on Pat-

- tern Recognition (Mustererkennung), S.427–434, Springer-Verlag 1997
2. L. J. Latecki and R. Lakämper, Discrete Approach to Curve Evolution, Proc. DAGM Conf. on Pattern Recognition (Mustererkennung), S.85-92, Stuttgart, September 1998
 3. L. J. Latecki and R. Lakämper, Shape Decomposition and Shape Similarity Measure, Proc. DAGM Conf. on Pattern Recognition (Mustererkennung), S.367-376, Stuttgart, September 1998
 4. R. Lakämper and L. J. Latecki and U. Eckhardt, Contour-based shape similarity, Proc. SPIE's Vision Geometry VII 3454, S. 202–213, San Diego, July 1998
 5. L. J. Latecki and R. Lakämper, Convexity Rule for Shape Decomposition Based on Discrete Contour Evolution, Computer Vision and Image Understanding 73, S.441–454, 1999
 6. L. J. Latecki and R. Lakämper, Contour-based shape similarity Proc. of Int. Conf. on Visual Information Systems, Springer-Verlag, Lecture Notes in Computer Science 1614, S. 617–624, Amsterdam, June 1999
 7. L. J. Latecki and R. Lakämper Polygon Evolution by Vertex Deletion, Proc. of Int. Conf. on Scale-Space Theories in Computer Vision, Corfu, Greece, September 1999
 8. L. J. Latecki and R. Lakämper, Shape Similarity Measure Based on Correspondence of Visual Parts, IEEE Trans. Pattern Analysis and Machine Intelligence, (submitted)
 9. L. J. Latecki and R.-R. Ghadially and R. Lakämper and U. Eckhardt, Continuity of the discrete curve evolution, SPIE and SIAM Conf. on Vision Geometry VIII, July 1999, to appear
 10. L. J. Latecki and R. Lakämper, Comparing Planar Shapes with Application to Object Retrieval in Image Databases, Pattern Recognition, to appear