

7. Merkmalsextraktion

Motivation:

- in der Bildverarbeitung und -analyse ist eine präzise Beschreibung auch komplizierter Objekte notwendig
- in vielen Anwendungen ist es wichtig, Klassen von Formen zu erkennen
- die Klassifikation beruht auf Merkmalen der untersuchten Formen

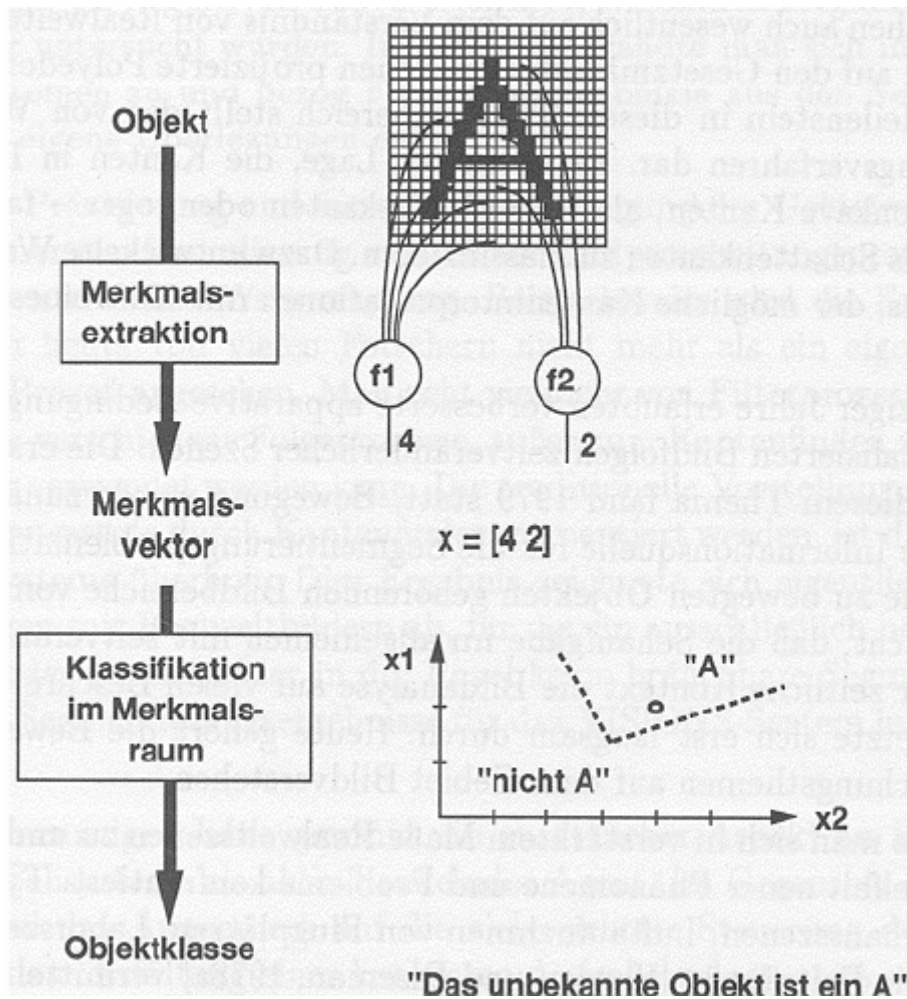
Mögliche Einteilungen von Merkmalen:

- Kontur-, Regionen-Merkmale
 - Art des Merkmals (Skala: nominal, ordinal, metrisch)
 - lokale / globale Merkmale
 - topologische / metrische Merkmale
 - mathematische / heuristische Techniken der Merkmalsextraktion
 - Rekonstruktionsfähigkeit des Objekts aus den Merkmalen
 - Fähigkeit, auch unvollständig dargestellte Objekte korrekt zu beschreiben
 - statistische / syntaktische Merkmale
 - Robustheit in Bezug auf Transformationen des Bildes
-
- geometrische Merkmale
 - densitometrische (photometrische) Merkmale (aus Grauwert-Histogramm)
 - Texturmerkmale
 - Farbmerkmale

Zweck der Merkmalsextraktion:

- Quantitative und qualitative Aussagen über Bildinhalte ermöglichen
- Einordnung von Objekten in Kategorien bzw. Klassen ermöglichen
- Merkmale = komprimierte Form von Bildinformation

Prozess der Merkmalsextraktion, Informationsverarbeitung und Klassifikation:



Lebewesen vs. Maschine:

große Unterschiede in Erkennungsleistung und -geschwindigkeit

<i>Art der Leistung</i>	<i>Lebewesen</i>	<i>Maschine</i>
Quantitative Bestimmung von Merkmalen (z.B. Grauwerte, Größen)	subjektiv, ungenau, relativ	objektiv, präzise
Zählen von Objekten	durch Ermüdung Fehler, langsam	keine Ermüdung, schnell, genau
Wiedererkennen von Objekten in unterschiedlichen Szenen	große Wiedererkennungslleistung, auch bei veräuschten Daten	Objektklassifikation eher problematisch

Geschichte:

Selfridge 1955: Vision von "Augen und Ohren für den Rechner"

- Zeichenerkennung
- Schrift-, Spracherkennung
- Bildauswertung (bes. militärische Anwendungen)

Überblick:

<p>topologische Merkmale</p> <p>Konturmerkmale</p> <ul style="list-style-type: none"> Kettencode Polygonbeschreibung Länge Orientierung Krümmung ... <p>statistische Merkmale</p> <ul style="list-style-type: none"> Momente Anzahl der Pixel Informationsgehalt Autokorrelation ... 	<p>Flächen- (Regionen-) Merkmale</p> <ul style="list-style-type: none"> Intensitätswert / Farbe Lage des Schwerpunkts Umfang, Größe Hauptachsen umschreibendes Rechteck konvexe Hülle ... <p>Texturmerkmale</p> <ul style="list-style-type: none"> statistische Merkmale strukturelle Beschreibung ... <p>fraktale Merkmale</p> <p>morphologische Merkmale</p> <ul style="list-style-type: none"> z.B. granulometrische Kurven
---	--

Topologische Merkmale

Wann ist ein Bildobjekt eine Kurve?

Def. von Menger und Uryson:

Ein Kontinuum C ist *eindimensional* im Punkt p , wenn für jedes $\varepsilon > 0$ eine Umgebung U innerhalb der ε -Umgebung von p existiert, deren Rand das Kontinuum nur in isolierten Punkten schneidet.

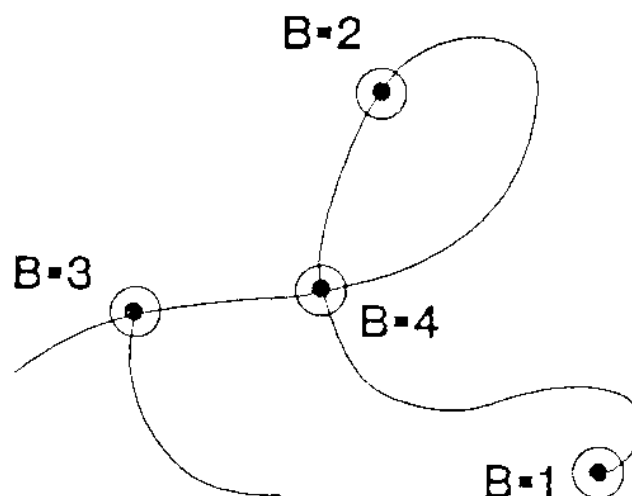
Ein Kontinuum, das in allen Punkten eindimensional ist, heißt *eindimensionales Kontinuum* oder *Kurve*.

(topologische Dimensionsdefinition, zu unterscheiden von algebraischer und fraktaler Dimension!)

Verzweigungsindex:

Es sei $B_\varepsilon(p)$ die Anzahl der isolierten Punkte, aus denen der Durchschnitt der Kurve mit dem lt. Def. gefundenen Rand der Umgebung besteht. Dann wird $B(p) = \lim B_\varepsilon(p)$ für $\varepsilon \rightarrow 0$ (sofern dieser Grenzwert existiert) als Verzweigungsindex der Kurve im Punkt p bezeichnet.

- *Einfache* Kurvenpunkte haben den Verzweigungsindex $B=2$ und Endpunkte den Verzweigungsindex $B=1$.
- Eine *Linie* ist eine unverzweigte Kurve (Verzweigungsindex < 3), die genau 2 Endpunkte besitzt.



Verzweigungsindex für digitale Kurven:

Es sei $U(p)$ die 3×3 -Umgebung eines Bildpunktes p aus einem Binärbild M und $R(p) = U(p) \setminus \{p\}$. Der Punkt p wird dann als Kurvenpunkt bezeichnet, wenn alle Punkte aus $R(p) \cap M$ isoliert sind, d.h. keine 4-Nachbarn in M besitzen. Die Anzahl dieser isolierten Punkte heißt *Verzweigungsindex* $B(p)$.

Wenn alle Punkte von M Kurvenpunkte sind, heißt M *digitale Kurve*.

Beachte: Die Menge der Konturpunkte eines digitalen Objekts bildet nicht notwendig eine Kurve i. Sinne dieser Def.

topologische Charakteristiken von Regionen:

Zusammenhang (Konnektivität)

(vgl. Übung 2)

Sei V die betrachtete Region im Binärbild.

Pixelzusammenhang:

Zwei Pixel p und q ($p, q \in V$) sind zusammenhängend, wenn q benachbart zu p ist, d.h. $q \in N_4(p)$ oder $q \in N_8(p)$, etc. So spricht man vom 4-Zusammenhang, bzw. 8-Zusammenhang etc.

Pfad von p nach q

Folge von Pixeln $(p_0, p_1, \dots, p_{n-1}, p_n) = (p, p_1, \dots, p_{n-1}, q)$, bei dem alle Pixel paarweise zusammenhängend sind, d.h. Es gilt $p_i \in N_4(p_{i-1})$ oder $p_i \in N_8(p_{i-1})$ ($1 \leq i \leq n$) etc. So spricht man von 4-; 8-Pfade, etc. Pfadlänge = n .

Zusammenhängende Komponente K

(maximal zusammenhängendes Gebiet)

I : Bild,

$I = \{ f_{00}, f_{01}, \dots, f_{M-1N-1} \}, S \subseteq I$

Teilmenge S heißt zusammenhängende Komponente K wenn gilt

$$K = \{ q \mid q \in S, \forall p \in S, \exists \text{Pfad}(p, \dots, q) \}$$

Hier:

Komponente: zusammenhängende Komponente aus 1-Pixeln

Löcher: zusammenhängende Komponente aus 0-Pixeln

Zusammenhängende Komponente: 4-Zusammenhang, 8-Zusammenhang
Loch: 8-Zusammenhang, 4-Zusammenhang

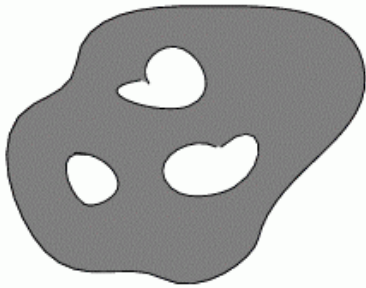
Die Eulerzahl E

Topologischer Deskriptor: $E = N_Z - N_L$

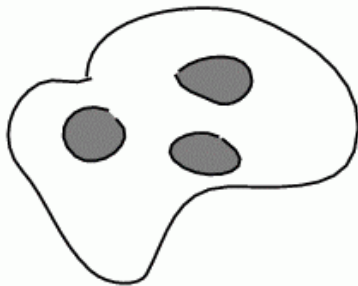
N_Z : Anzahl der zusammenhängenden Komponenten,

N_L : Anzahl der Löcher.

Beispiele



a) „Schweizer Käse“ $E = 1 - 3 = -2$



b) $E = 3 - 0 = 3$



c) $E = 1 - 1 = 0$



d) $E = 1 - 2 = -1$

Einfache geometrische Merkmale

Fläche: Anzahl der Bildpunkte, die zum Objekt gehören

Umfang: Anzahl der Umfangspunkte

Gewinnung der Umfangspunkte durch morphologische Operation ($\mathbf{D}_B(X) \setminus X$) oder durch Konturerkennung, siehe Kap. 5

Umfang genauer aus dem Kettencode:

direkte Nachbarn mit Faktor 1, die restlichen mit Faktor $\sqrt{2}$

Formfaktor:

Verhältnis von Umfangsquadrat zur Fläche

$$V_{U^2F} = \frac{U^2}{4\pi F}$$

ist beim Kreis = 1, größer bei zerklüfteten Objekten

wird manchmal auch *Kompaktheit* genannt

(Achtung: dies hat nichts mit "Kompaktheit" im Sinne der Topologie oder Analysis zu tun!)



Schwerpunkt (bei Binärbild) zur Lagebeschreibung:

Geometrischer Schwerpunkt (x_s, y_s) :

$$x_s = \frac{1}{F} \sum_{i=1}^F x_i \quad \text{bzw.} \quad y_s = \frac{1}{F} \sum_{i=1}^F y_i$$

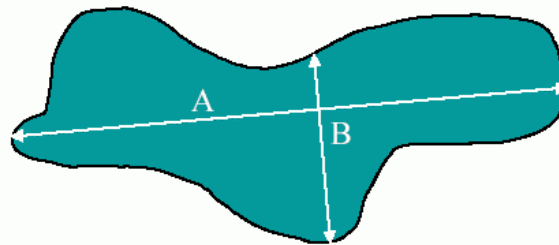
(F = Fläche des Objekts)

Formfaktor zwischen Innen- und Außenkreis um den geometr. Schwerpunkt:

$$V_{RAD} = \frac{R_{\max}}{R_{\min}}$$

Exzentrizität:

- Verhältnis der längsten Sehne innerhalb der Region zur Sehne quer dazu (Hauptachse zu Nebenachse)



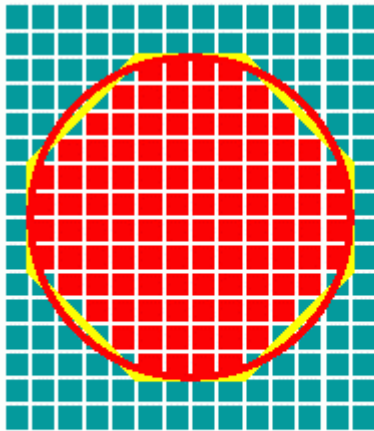
Konturmerkmale

- Beschreibung einer Kontur durch Liste von Pixel-Koordinatenpaaren (kartesisch oder polar)
- oder durch Kettencode (siehe Kap. 3)

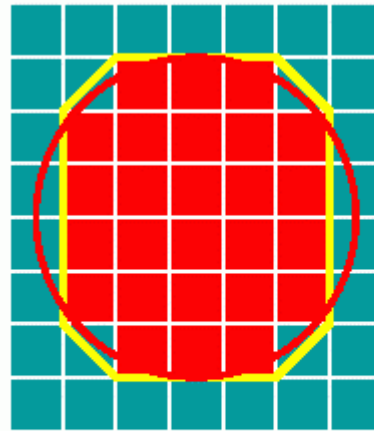
Länge

- **Randlänge:** ermittelt aus dem Chain Code:
 - vertikale und horizontale Elemente: Länge 1
 - diagonale Elemente: Länge $\sqrt{2}$ (8NB)
 - Rand länger in 4NB als 8NB
 - geschlossene Kurven: Randlänge = Umfang (Perimeter)
 - abhängig von der Bildauflösung
 - abhängig von der Randdefinition (innerer Rand, äußerer Rand, extended Border)

Randlänge abhängig von der räumlichen Auflösung:



$$16 + 16 * \sqrt{2} = 38,63$$

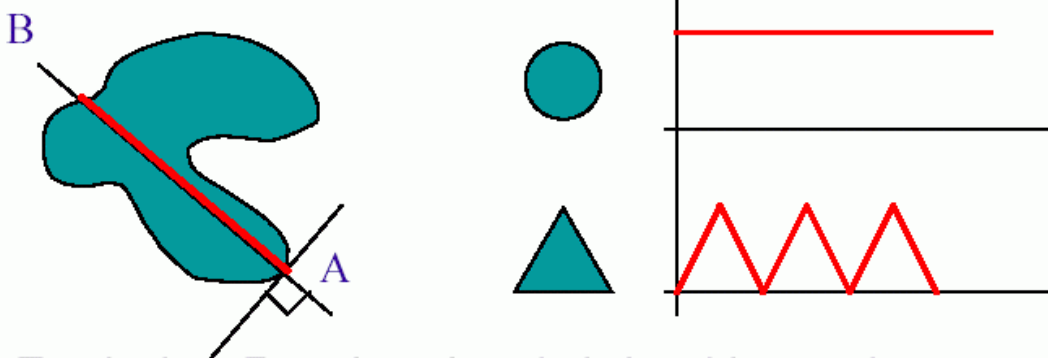


$$2 * (14 + 4 * \sqrt{2}) = 2 * 19,66 = 39,32$$

Krümmung: siehe Übung 7

Signatur:

- Folge von Kontur-Normalabständen



- Für jeden Randpunkt wird der Abstand zum kürzesten gegenüberliegenden Randpunkt, rechtwinklig zur aktuellen Tangente, gesucht

Sehnenverteilung (chord distribution)

- Sehne ist eine Linie, die 2 Randpunkte verbindet
- Konturbeschreibung durch die **Verteilung der Länge und Winkel dieser Sehnen**

– kontinuierlich

$$h(\Delta x, \Delta y) = \iint b(x, y) b(x + \Delta x, y + \Delta y) dx dy$$

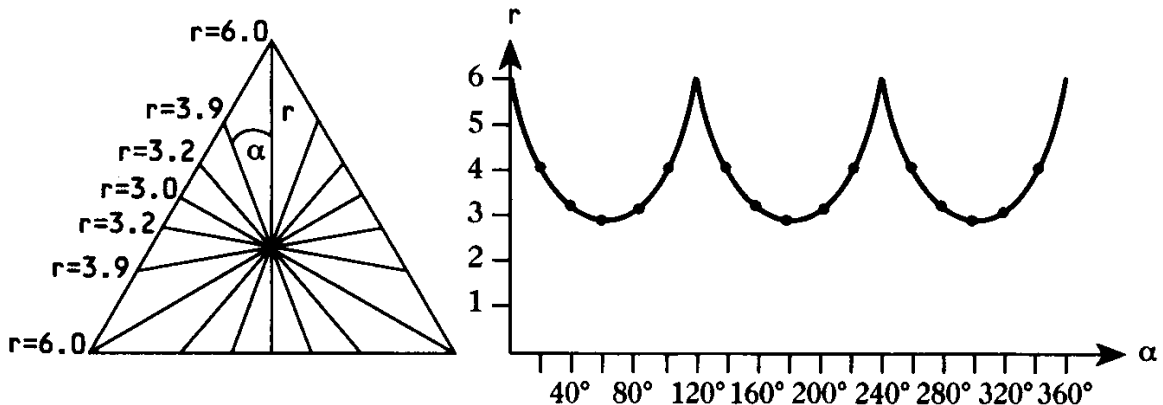
– diskret

$$h(\Delta x, \Delta y) = \sum_i \sum_j b(x_i, y_j) b(x_i + \Delta x, y_j + \Delta y)$$

Polarer Abstand:

Abstand des Bereichsrandes vom Schwerpunkt.

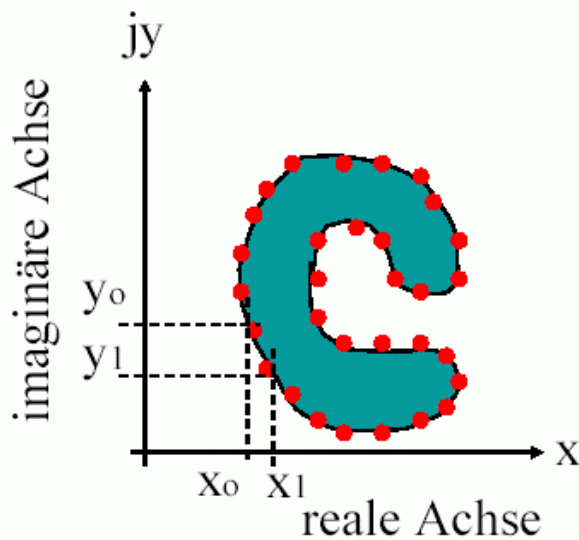
Darstellung als Kurve: aussagekräftig für die Form



Fourier-Deskriptoren:

FT angewandt auf die Kontur als 1D-Funktion

- Geschlossene Kurve in der komplexen Ebene
- Umrundung der Kurve (entgegen dem Uhrzeigersinn) ergibt komplexe Funktion $z(t)$
- Geschwindigkeit wird so gewählt, daß eine Umrundung 2π lange dauert.
- Mehrfache Umrundung ergibt 2π periodische Funktion: Darstellung als Fourier-Folge möglich



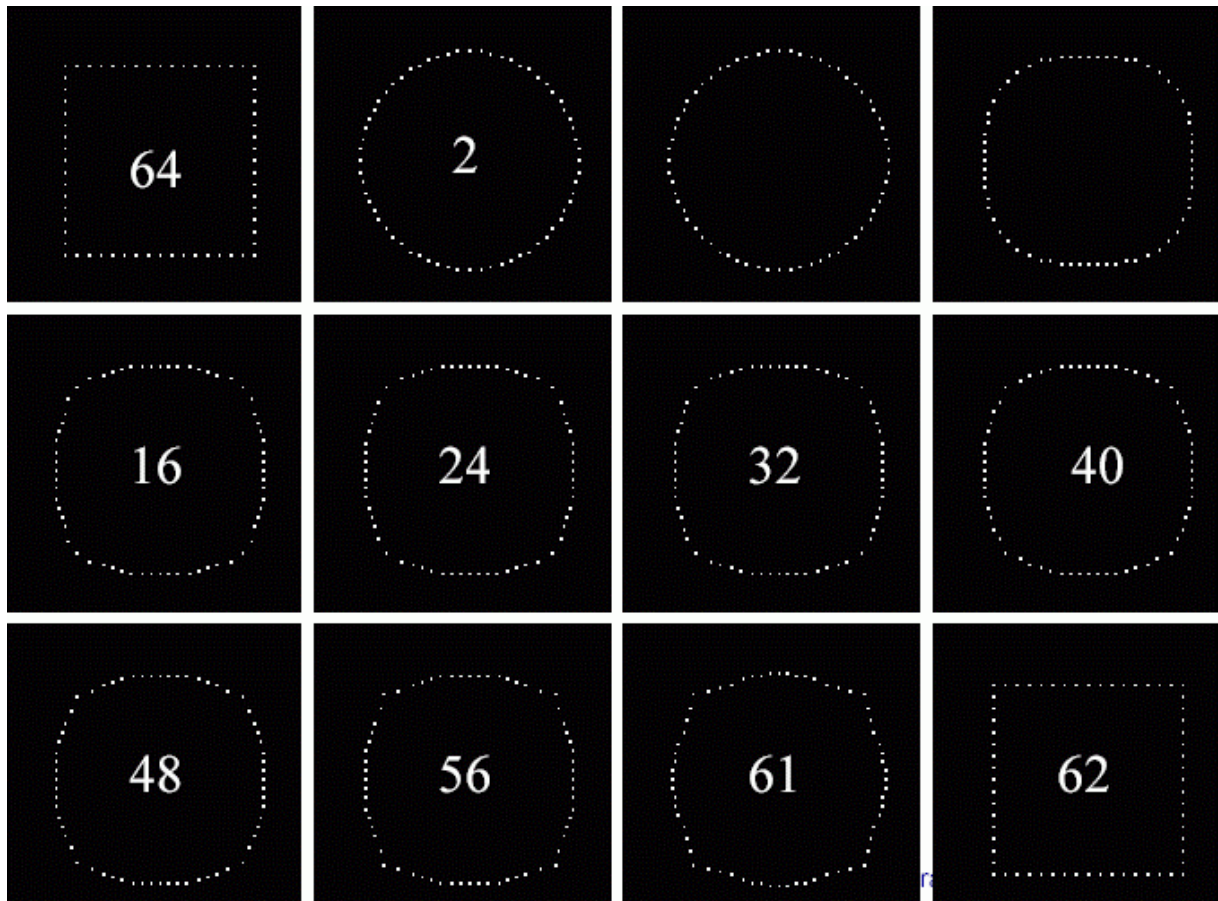
$$z(t) = \sum_n T_n e^{jnt}$$

T_n ... Fourier Deskriptoren

$$t = \frac{2\pi s}{L} \quad s \dots \text{Kurvenlänge}$$

$$T_n = \frac{1}{L} \int_0^L z(s) e^{-j(2\pi/L)ns} ds$$

Beispiel für die Anwendung:
Approximation des Randes durch weniger Koeffizienten in der
Fourier-Darstellung



(aus Beichel 2002)

– analog auch mit Wavelets möglich!

Regionen-Merkmale und Formbeschreibung

einfache Merkmale von Regionen: Fläche, Schwerpunkt, Formfaktor, Exzentrizität (siehe oben)

Ferret box:

kleinstes *achsenparalleles* Rechteck, welches die Region vollständig enthält

- *Aspect ratio* (Breite : Höhe) dieses Rechtecks
- Füllungsgrad dieses Rechtecks

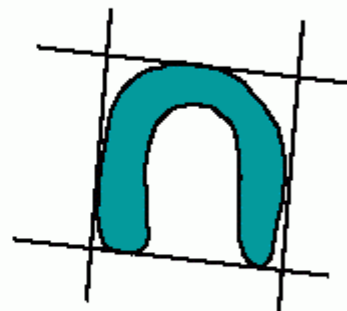
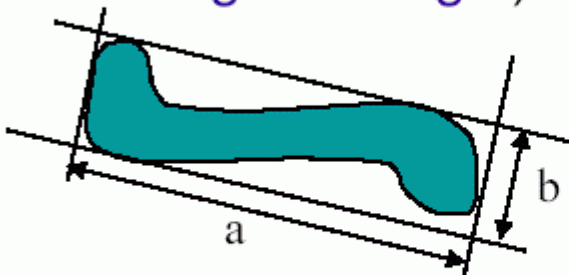
Minimum bounding rectangle (MBR):

kleinstes Rechteck, welches die Region vollständig enthält
– i.allg. nicht achsenparallel

- Angabe als Polygon (4 Eckpunkte)
- oder über Mittelpunkt, Breite, Höhe und Winkel

Ausdehnung (Elongatedness):

- Verhältnis der **Kantenlängen** des kleinsten Umschreibenden Rechtecks (**Minimum Bounding Rectangle**)



- bei **stark kurvigen** Regionen: **Verhältnis** der **Fläche** zum **Quadrat der Dicke**

$$\text{elongatedness} = \frac{\text{area}}{(2d)^2}$$

- *Füllungsgrad* des MBR = *Rectangularity*

- **Rectangularity** (Rechteckigkeit):
Maximales Verhältnis der
Regionenfläche zur Fläche eines
umschreibenden Rechtecks $rect = \max_k(F_k)$
- Auch Dreiecks bzw. Kreisverhältnis

Best ellipse fit:

Das Objekt wird durch eine Ellipse angenähert. Mittelpunkt, Orientierung der Hauptachse und Länge von Haupt- und Nebenachse werden repräsentiert.

Elongiertheit:

wird auch def. als $(1 - \text{Nebenachsenlänge}/\text{Hauptachsenlänge})$ des *best ellipse fit*.

Densitometrische Merkmale

(auch: photometrische, radiometrische Merkmale)

werden aus Grauwertverteilung gewonnen

- einzelnes Objekt
oder
- ganzes Bild
oder
- Fenster

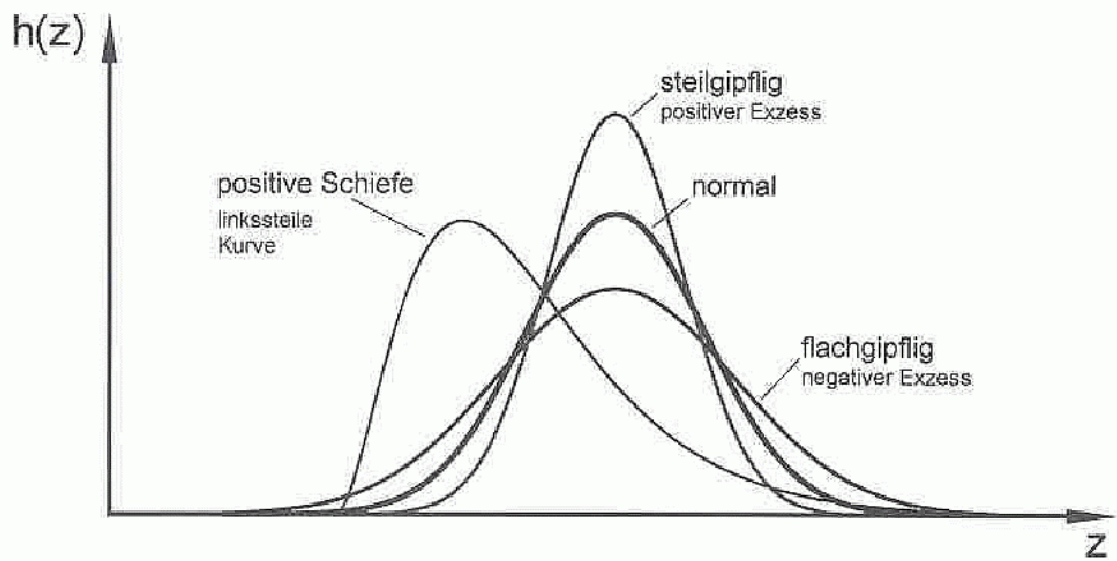
nach Segmentierung Berechnung von Histogrammen für
einzelne Objekte / Regionen

Bildpunkte außerhalb der betrachteten Region werden auf Null
gesetzt (Masken-Operation)

Charakterisierung des Histogramms durch Momente (Mittelwert,
Varianz, Schiefe, Kurtosis; vgl. Übung 1)

➤ Histogrammformen

[Abmayr 94], S. 258



Übersicht über Größen der beschreibenden Statistik
(Achtung: bei Schiefe und Exzess bzw. Kurtosis teilweise unterschiedliche Definitionen in der Literatur!)

Statistik 1. Ordnung

Mittelwert	$y_m = \sum_{y=0}^{L-1} y P(y)$
Varianz	$\sigma_y^2 = \sum_{y=0}^{L-1} (y - y_m)^2 P(y)$
Schiefe	$y_s = \frac{1}{\sigma_y^3} \sum_{y=0}^{L-1} (y - y_m)^3 P(y)$
Exzess	$y_E = \frac{1}{\sigma_y^4} \sum_{y=0}^{L-1} (y - y_m)^4 P(y)$
Energie	$P_y = \sum_{y=0}^{L-1} [P(y)]^2$
Entropie	$H_y = - \sum_{y=0}^{L-1} P(y) \log_2[P(y)]$

Die Schätzung der Wahrscheinlichkeit erfolgt aus einem Grauwert histogramm $N(y)$

$$P(y) = P\{x(m, n) = y\} \approx \frac{N(y)}{M}$$

M = Zahl der Bildpunkte, L =Zahl der Quantisierungsstufen

Statistik 2. Ordnung

Korrelation $\rho = \sum_{y=0}^{L-1} \sum_{z=0}^{L-1} y \cdot z P(y, z)$

Kovarianz $\gamma = \sum_{y=0}^{L-1} \sum_{z=0}^{L-1} (y - y_m)(z - z_m) P(y, z)$

Trägheitsmoment $I = \sum_{y=0}^{L-1} \sum_{z=0}^{L-1} (y - z)^2 P(y, z)$

Absolutwert $y_4 = \sum_{y=0}^{L-1} \sum_{z=0}^{L-1} |y - z| P(y, z)$

Energie $P_y = \sum_{y=0}^{L-1} \sum_{z=0}^{L-1} [P(y, z)]^2$

Entropie $H_{yz} = - \sum_{y=0}^{L-1} \sum_{z=0}^{L-1} P(y, z) \log_2[P(y, z)]$

Die Schätzung der Verbundwahrscheinlichkeit

$$P(y, z) = P\{ x(m, n) = y, x(i, k) = z \}$$

erfolgt aus einem zweidimensionalen Histogramm, der sogenannten Cooccurrence Matrix $N(y, z)$

$$P(y, z) \approx \frac{N(y, z)}{M}$$

M = Zahl der Bildpunkte, L =Zahl der Quantisierungsstufen

Korrelationskoeffizient

(vgl. Kap. 4):

$$r = \frac{s_{xy}}{s_x \cdot s_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

- drückt Stärke des linearen Zusammenhangs zwischen zwei Zufallsgrößen aus
- liegt zwischen -1 und $+1$

wenn 2 Bildpunkte (aufgefasst als Realisierungen von zufälligen Größen) desselben Bildes korreliert werden:

(normierte) Autokorrelationsfunktion

wenn die beiden verglichenen, ortsabhängigen Zufallsvariablen aus verschiedenen Bildern stammen:

(normierte) Kreuzkorrelationsfunktion

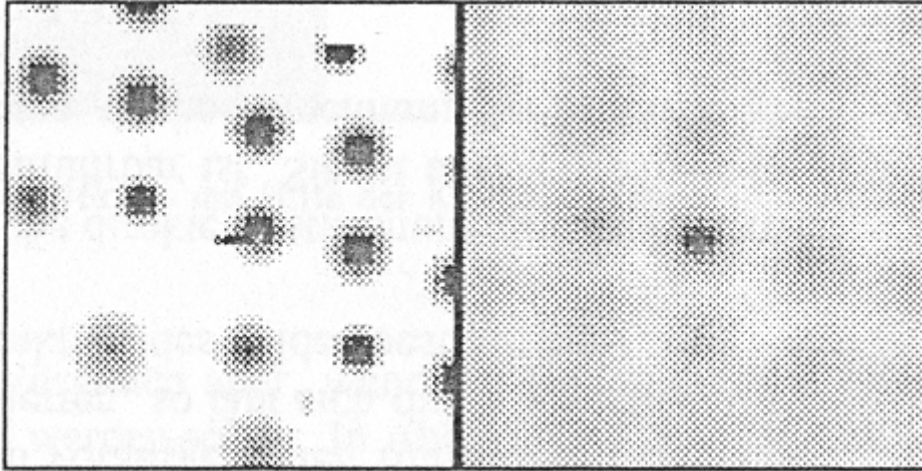
Problem bei der Berechnung:

einzelnes digitales Bild = nur eine einzige Realisierung des angenommenen Zufallsprozesses

⇒ man verwendet die Voraussetzung eines *stationären Prozesses*:

- Mittelwerte und Varianzen sind ortsunabhängig
- die normierte Autokorrelation zweier Bildpunkte ist verschiebungsunabhängig (bei gleicher Verschiebung der beiden Punkte)
- wenn zusätzlich gilt: die Korr. ist nur abhängig von $r^2 = (\Delta x)^2 + (\Delta y)^2$, nicht von Δx und Δy selbst, so spricht man von einem isotropen Prozess.

Beispiel: Autokorrelationsfunktion eines Grauwertbildes



Berechnung der Autokorrelationsfunktion effizienter mit Hilfe der Fourier-Transformierten des Bildes (s. Voss & Süße 1991)

Einsatz der Kreuzkorrelation beim Matching:

Beispiel:
Bestimmung eines Verschiebungsvektors zwischen zwei Satellitenaufnahmen

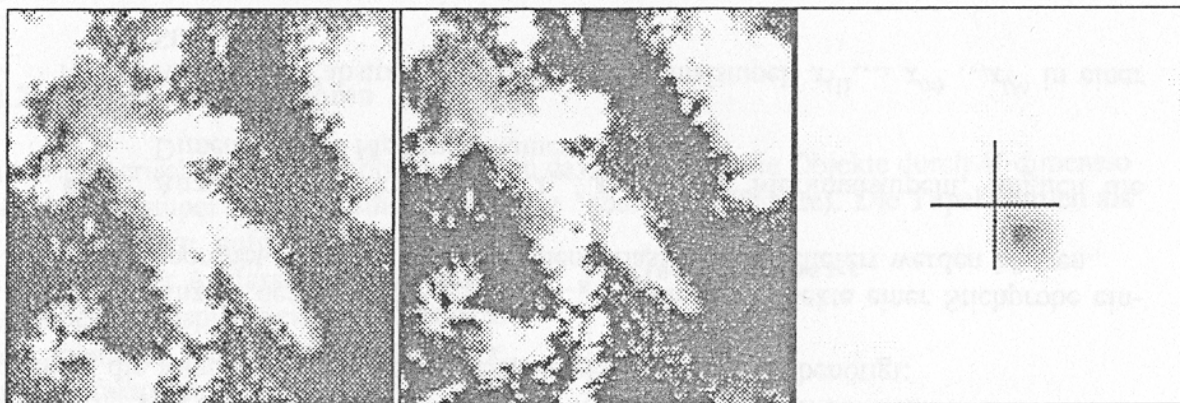


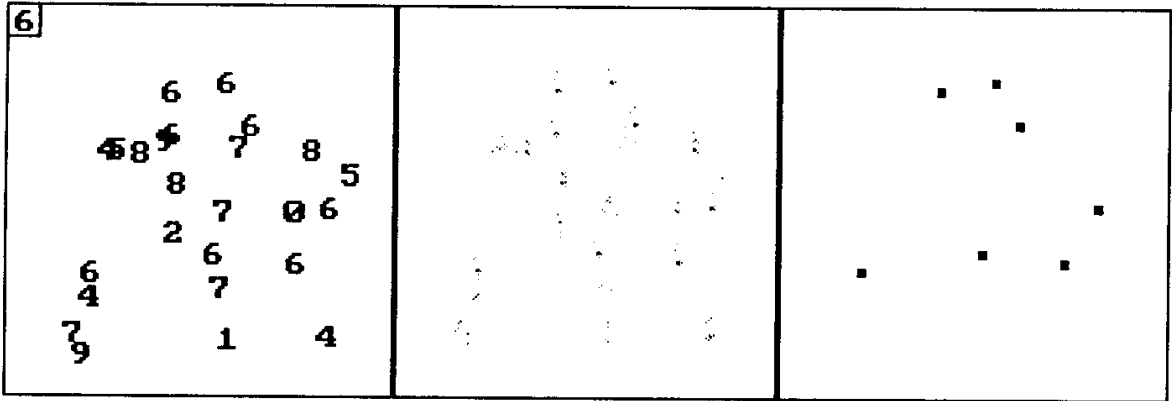
Bild 1

Bild 2

Korrelationsfunktion
(Fadenkreuz = Nullpunkt)

(aus Voss & Süße 1991)

Korrelationsanalyse zur Detektion der Ziffer "6" durch Matching:



Originalbild
mit gesuchtem Muster
(links oben)

Korrelationsfunktion

Cluster der
Korr.funktion