

# Übersicht der Vorlesung

1. Einführung
2. Bildverarbeitung
3. Morphologische Operationen
4. Bildsegmentierung
5. Merkmale von Objekten
6. Klassifikation
7. Dreidimensionale Bildinterpretation
8. Bewegungsanalyse aus Bildfolgen
9. PCA (Hauptkomponentenanalyse)
10. ICA (Independent Component Analysis – Unabhängigkeitsanalyse)

# 5 Merkmale von Objekten

# 5 Merkmale von Objekten

5.1 Einführung

5.2 Geometrische Merkmale

5.3 Momente

5.4 Lauflängenkodierung

5.5 Eulerzahl

5.6 Fourierdarstellung von Segmentkonturen

5.7 Relationalstrukturen

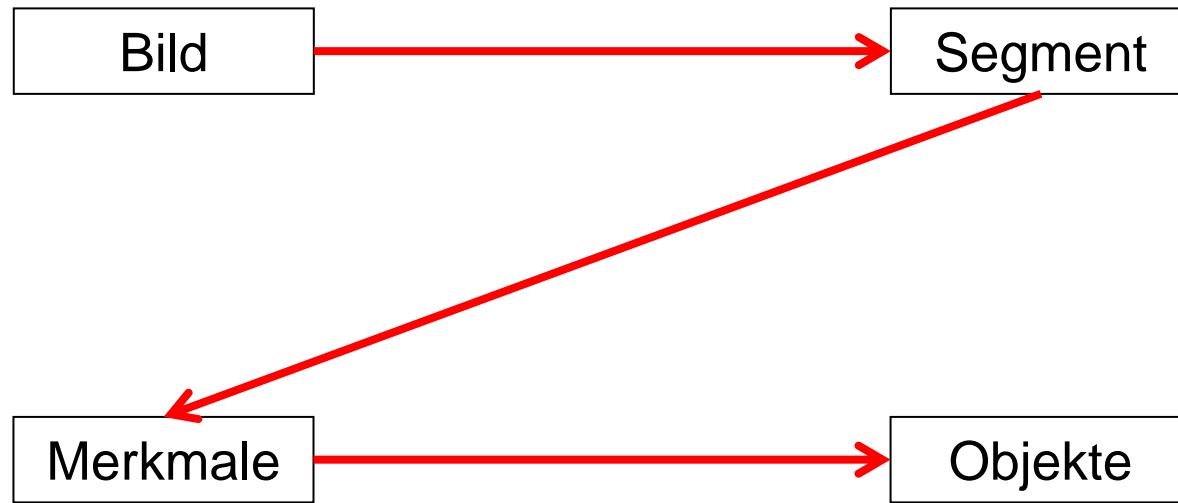
5.8 Statistische Merkmale Textur

# 5.1 Einführung

# Einführung

- Verfahren, die Merkmale von Segmenten oder Objekten aus Bildern extrahieren
- Merkmale dienen zur:
  - charakteristischen Beschreibung von Objekten
  - wesentlichen Unterstützung bei der Objekterkennung
- Ausgangspunkt ist die Segmentierung des Bildes in einzelne Segmente, für die dann Merkmale berechnet werden
- Mit Hilfe von Merkmalen sollen die Segmente bestimmten Objekten zugeordnet werden

# Schritte der Objekterkennung



# Einführung

- Es werden Merkmale benötigt, die zwei beinahe unvereinbare Kriterien erfüllen sollten:
- Sie müssen zum einen aus Bildern berechnet werden können, also auf konkreten Abbildungen von Objekten basieren.



- Zum anderen müssen sie von den Besonderheiten einer konkreten Abbildung unabhängig sein, d.h. invariant bezüglich:
  - Verschiebung
  - Drehung
  - Maßstabsänderung
  - Betrachterperspektive
  - Beleuchtung
  - teilweiser Verdeckung

## 5.2 Geometrische Merkmale

# Fläche eines Segmentes

Anzahl der Pixel, die zum Segment  $B$  gehören

$$F_1 = \left| \{(i, j) : (i, j) \in B\} \right|$$

Falls ein Segment ein Polygon ohne Löcher ist und man die Koordinaten  $(u_i, v_i)$ ,  $i = 0, \dots, N-1$  der Randkontur kennt, kann man die Gauß'sche Flächenformel für Polygone benutzen

$$F_2 = \frac{1}{2} \left| \sum_{i=0}^{N-1} (u_i \cdot v_{[(i+1) \bmod N]} - u_{[(i+1) \bmod N]} \cdot v_i) \right|$$

# Fläche eines Segmentes

- Es kann zu unterschiedlichen Ergebnissen kommen
- Eine weitere Möglichkeit ist die Benutzung der Momente

# Umfang eines Segmentes

Anzahl der Randpixel des Segments  $B$

$$U_1 = \left| \{ (i, j) : (i, j) \in R(B) \} \right|$$

Eine geschlossene Randkontur kann man auch mit Hilfe des Freemancodes beschreiben.

$$R(B) = \{p_1, p_2, \dots, p_M\}$$

$$C_{R(B)} = \{c_0, c_1, \dots, c_{M-2}, c_{M-1}\}$$

$$c_0 = r(p_1, p_2)$$

$$c_i = r(p_{i+1}, p_{i+2}), i = 0, 1, \dots, M-2$$

$$c_{M-1} = r(p_M, p_1)$$

$$c_1 = r(p_2, p_3)$$

$$U_2 = \sum_{i=0}^{M-1} l(c_i) \quad l(c) = \begin{cases} 1 & \text{falls } c = 0, 2, 4, 6 \\ \sqrt{2} & \text{falls } c = 1, 3, 5, 7 \end{cases}$$

$$U_{korr} = 0,95 \cdot U_2$$

# Kompaktheit (Rundheit) eines Segmentes

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

$K$  ist gleich 1 für einen Kreis und wird umso größer, je länger der Umfang eines Segments im Verhältnis zu seinem Flächeninhalt ist

Beispiele:

$$K = \frac{16 \cdot a^2}{4\pi \cdot a^2} = \frac{16}{4\pi} > 1$$

Quadrat mit Seitenlänge  $a$

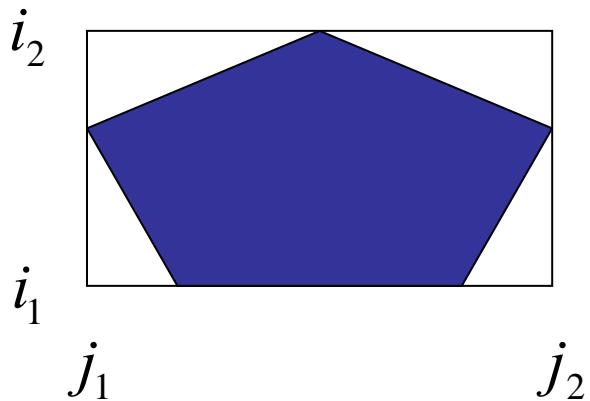
$$K = \frac{100 \cdot a^2}{4\pi \cdot 4a^2} = \frac{25}{4\pi} > \frac{16}{4\pi}$$

Rechteck mit Seitenlängen  $a$  und  $4a$

$$K = \frac{9a^2}{4\pi \frac{a^2}{4}\sqrt{3}} \approx \frac{20.78}{4\pi}$$

gleichseitiges Dreieck mit der Seitenlänge  $a$

# Kleinstes achsenparalleles umschreibendes Rechteck



$$i_1 = \min_{(i,j) \in B} i$$

$$i_2 = \max_{(i,j) \in B} i$$

$$j_1 = \min_{(i,j) \in B} j$$

$$j_2 = \max_{(i,j) \in B} j$$

Höhe, Breite:

$$H = i_2 - i_1$$

$$B = j_2 - j_1$$

Ratio:

$$H/B$$

$$|\log(H/B)|$$

$$|\log(H/B)| = |\log(B/H)|$$

# Umschreibende Rechtecke

- umschreibendes Rechteck minimaler Fläche
- Füllungsgrad (wie viel Prozent des umschreibenden Rechtecks werden ausgefüllt)

# Weitere geometrische Merkmale

- konvexe Hülle
- umschreibende n – Ecke (z.B. n=8)
- kleinste umschreibende Ellipse

## 5.3 Momente

# Moment

dienen zum Finden von Merkmalen, die invariant sind gegenüber Translation, Rotation, Skalierung oder auch allgemeinen affinen Transformationen

$$P = \{(i, j) : i = 0, \dots, I - 1, j = 0, \dots, J - 1\}$$

$$B = \{(i, j) \in P : (i, j) \in \text{Segment}\}$$

Moment  $(p, q)$  – ter Ordnung:

$$m_{pq}(B) = \sum_B i^p j^q$$

In dieser Form hängen die Momente von der Position des Segmentes  $B$  im Bildraster  $P$  ab, so dass eine Translationsinvarianz nicht gegeben ist.

# Beispiele – Schwerpunkt

$$m_{00}(B) = \sum_B 1 \quad \xrightarrow{\text{Fläche von B}}$$

$$m_{10}(B) = \sum_B i \quad m_{01}(B) = \sum_B j \quad m_{11}(B) = \sum_B i \cdot j$$

$$i_c(B) = \frac{m_{10}(B)}{m_{00}(B)}$$

$$j_c(B) = \frac{m_{01}(B)}{m_{00}(B)}$$

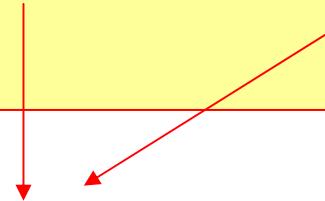
Schwerpunkt  $i$  – Koordinate

Schwerpunkt  $j$  – Koordinate

# Zentriertes Moment

zentriertes Moment  $(p,q)$  – ter Ordnung:

$$\mu_{pq}(B) = \sum_B (i - i_c(B))^p (j - j_c(B))^q$$



Schwerpunkt

zentrierte Momente sind invariant gegenüber Translation

# Normiertes Moment

normiertes Moment  $(p,q)$  – ter Ordnung:

$$\eta_{pq}(B) = \frac{\mu_{pq}(B)}{\mu_{00}(B)^\gamma} \quad \gamma = \frac{p+q}{2} + 1$$

sind invariant gegenüber Ähnlichkeitstransformationen

Beispiele:

$$\eta_{00}(B) = \frac{\mu_{00}(B)}{\mu_{00}(B)^1} = 1$$

$$\eta_{20}(B) = \frac{\mu_{20}(B)}{\mu_{00}^2(B)} = \frac{\sum_B (i - i_c(B))^2}{m_{00}^2(B)}$$

# Rotationsinvariante Merkmale

$$\Phi_1(B) = \eta_{20}(B) + \eta_{02}(B)$$

$$\Phi_2(B) = (\eta_{20}(B) - \eta_{02}(B))^2 + 4(\eta_{11}(B))^2$$

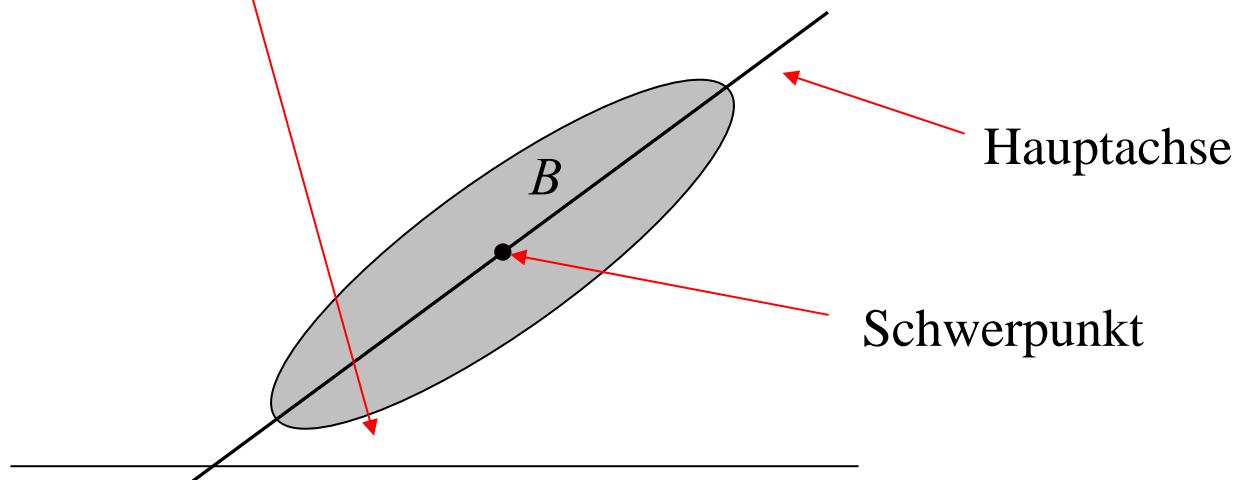
$$\Phi_3(B) = (\eta_{30}(B) - 3\eta_{12}(B))^2 + (3\eta_{21}(B) - \eta_{03}(B))^2$$

$$\Phi_4(B) = (\eta_{30}(B) + \eta_{12}(B))^2 + (\eta_{21}(B) + \eta_{03}(B))^2$$

Es gibt noch 3 weitere solche Merkmale.

# Orientierung eines Segmentes (Richtung der Hauptachse)

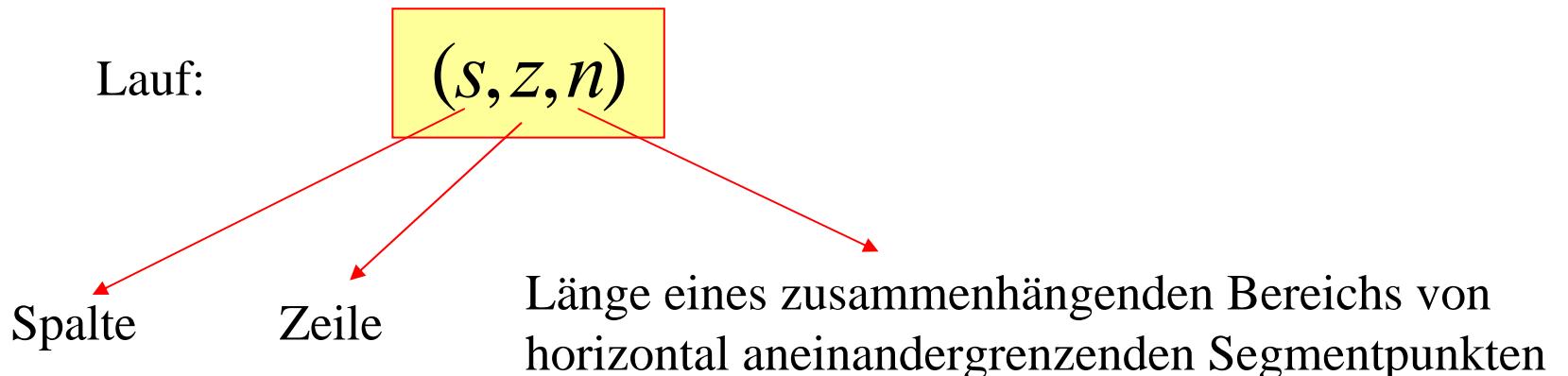
$$\varphi(B) = \frac{1}{2} \arctan \frac{2\eta_{11}(B)}{\eta_{20}(B) - \eta_{02}(B)}$$



## 5.4 Lauflängenkodierung

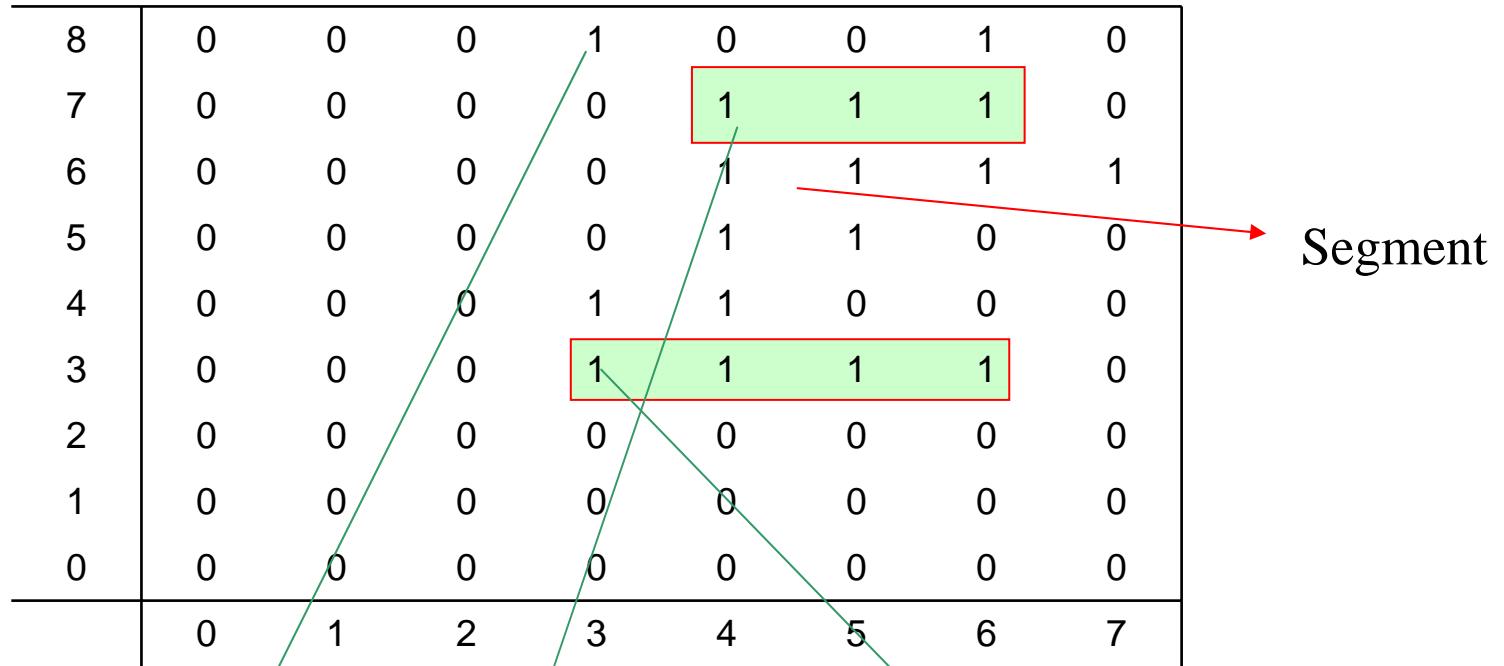
# Lauflängenkodierung

Merkmal ist keine reelle Zahl, sondern eine Menge von Tripeln



Lauflängenkodierung:      Menge von Läufen

# Lauflängenkodierung – Beispiel



Lauflängenkodierung:

$$\{(3,8,1), (6,8,1), (4,7,3), (4,6,4), (4,5,2), (3,4,2), (3,3,4)\}$$

# Lauflängenkodierung

Die Lauflängenkodierung erlaubt eine einfache Berechnung der Fläche und des Schwerpunktes eines Segmentes.

```
type
lauftype = record
    s : integer;
    z : integer;
    n : integer;
    end;
var
    lauf : array[1..num_laeufe] of lauftype;
begin
    for r:=1 to num_laeufe do begin
        z := lauf[r].z;
        s := lauf[r].s;
        n := lauf[r].n;
        flaeche := flaeche + n;
        summe_z := summe_z + z*n;
        summe_s := summe_s + s*n + (n-1)*n/2;
    end;
    scherpunkt_x := summe_s / flaeche;
    scherpunkt_y := summe_z / flaeche;
end.
```

# 5.5 Eulerzahl

# Eulerzahl

Eulerzahl:

$$E = B - L$$

Anzahl der Segmente

Anzahl der Löcher (Hintergrund, welcher von einem Segment umgeben wird).

Es ist zu beachten, dass die Eulerzahl von der Nachbarschaft  $N_4$  oder  $N_8$  abhängig ist.

# Berechnung der Eulerzahl

Wir schieben ein  $2 \times 2$  Fenster über das Bild.

$n(Q_1)$  Anzahl der Situationen, wo im Fenster genau eine Eins zu sehen ist

$$\begin{matrix} 0 & 0 \\ 0 & 1 \end{matrix} \text{ oder } \begin{matrix} 0 & 0 \\ 1 & 0 \end{matrix} \text{ oder } \begin{matrix} 1 & 0 \\ 0 & 0 \end{matrix} \text{ oder } \begin{matrix} 0 & 1 \\ 0 & 0 \end{matrix}$$

$n(Q_3)$  Anzahl der Situationen, wo im Fenster genau drei Einsen zu sehen sind

$$\begin{matrix} 1 & 1 \\ 1 & 0 \end{matrix} \text{ oder } \begin{matrix} 1 & 1 \\ 0 & 1 \end{matrix} \text{ oder } \begin{matrix} 0 & 1 \\ 1 & 1 \end{matrix} \text{ oder } \begin{matrix} 1 & 0 \\ 1 & 1 \end{matrix}$$

$n(Q_D)$  Anzahl der Situationen, wo im Fenster zwei diagonal angeordnete Einsen zu sehen sind

$$\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} \text{ oder } \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}$$

# Eulerzahl – Berechnung

8 – Nachbarschaft:

$$E = E_8 = \frac{n(Q_1) - n(Q_3) - 2 \cdot n(Q_D)}{4}$$

4 – Nachbarschaft:

$$E = E_4 = \frac{n(Q_1) - n(Q_3) + 2 \cdot n(Q_D)}{4}$$

# Beispiel

|   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 |   |
| 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

$$n(Q_1) = 9$$

$$n(Q_3) = 5$$

$$n(Q_D) = 2$$

$$E_8 = \frac{9-5-4}{4} = 0 = 1-1 \quad \xrightarrow{\hspace{1cm}} \text{1 Segment und 1 Loch}$$

$$E_4 = \frac{9-5+4}{4} = 2 = 3-1 \quad \xrightarrow{\hspace{1cm}} \text{3 Segmente und 1 Loch}$$

# Beispiel

|   |   |   |   |   |  |
|---|---|---|---|---|--|
|   |   |   |   |   |  |
| 1 | 1 | 1 | 1 |   |  |
| 1 |   | 1 | 1 |   |  |
| 1 | 1 |   | 1 |   |  |
|   |   | 1 | 1 | 1 |  |
|   |   |   |   |   |  |

$$n(Q_1) = 5$$

$$n(Q_3) = 7$$

$$n(Q_D) = 1$$

$$E_8 = \frac{5-7-2}{4} = -1 = 1-2 \rightarrow 1 \text{ Segment und 2 Löcher}$$

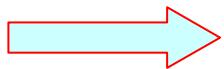
$$E_4 = \frac{5-7+2}{4} = 0 = 1-1 \rightarrow 1 \text{ Segment und 1 Loch}$$

# 5.6 Fourierdarstellung von Segmentkonturen

# Segmentkonturen

Rand eines Segmentes

$N$  Konturpunkte:  $\{(i(n), j(n)): 0 \leq n \leq N-1\}$



$$R(n) = i(n) + i \cdot j(n), \quad i = \sqrt{-1}$$

$$F(k) = \frac{1}{N} \sum_{n=0}^{N-1} R(n) \cdot e^{-2\pi i \frac{nk}{N}} \quad 0 \leq k \leq N-1$$

$$R(n) = \sum_{k=0}^{N-1} F(k) \cdot e^{2\pi i \frac{nk}{N}}$$

Fouriermerkmale für den Rand des Segmentes

Beschreiben die Form des Randes, z.B.:  
gezackte Konturverläufe (große Beträge bei hohen  $k$ )

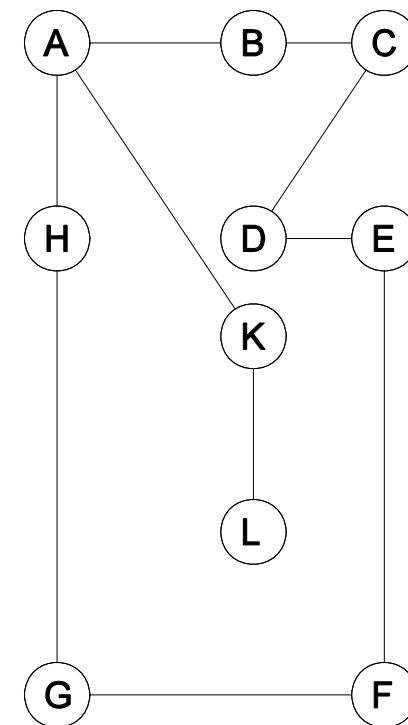
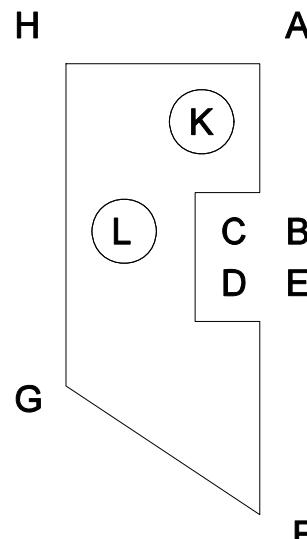
# 5.7 Relationalstrukturen

# Relationalstrukturen

- erlauben eine Segmentbeschreibung auf der Basis lokaler Beschreibungselemente und ihrer Beziehungen untereinander
- Segment wird also durch einen Graph beschrieben
- Beschreibung ist rotationsinvariant
- bleibt auch bei partieller Überdeckung so aussagekräftig, dass eine Klassifizierung durch Vergleich mit gegebenen Mustergraphen möglich ist
- Dieser Vergleich ist allerdings sehr aufwendig

# Beispiel

Relationale Beschreibung eines Stanzteiles mit Hilfe eines Graphen



# 5.8 Statistische Merkmale – Textur

# Textur

- Texturen spielen bei der Unterscheidung und Erkennung von Objekten auf Bildern eine wesentliche Rolle
- Jede Textur besitzt eine Grundtextur, die die lokale Verteilung und Variation der Grauwerte in einem Teilbereich des Bildes bestimmt
- Die Textur entsteht dann durch wiederholte Anordnung dieser Grundtextur
- Texturmodelle:
  - strukturelles Texturmodell (Grundtexturen exakt aneinandergereiht)
  - statistisches Texturmodell (man bekommt bei verschiedener Realisierung keine deckungsgleichen Texturen)
  - Mischformen

# Grauwertübergangsmatrizen (Co-Occurrence-Matrizen)

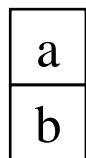
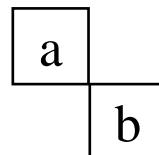
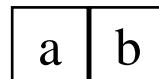
Zeilen- bzw. Spaltenanzahl ( $Z, S$ ) ist gleich der Anzahl der möglichen Grauwerte

Elemente  $f(z, s)$  dieser Matrizen geben die Häufigkeit der in der Textur auftretenden Grauwertkombinationen benachbarter Pixel an  
eine eigene Matrix für jede mögliche Nachbarschaftsbeziehung

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 3 & 3 & 3 & 3 \\ 0 & 0 & 0 & 0 & 3 & 3 & 3 & 3 \\ 0 & 0 & 0 & 0 & 3 & 3 & 3 & 3 \\ 0 & 0 & 0 & 0 & 3 & 3 & 3 & 3 \\ 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Eingabebild (4 Grauwerte)

Nachbarschaft



$$\begin{pmatrix} 0 & 1 & 2 & 3 \\ 12 & 0 & 0 & 0 \\ 0 & 12 & 4 & 0 \\ 0 & 0 & 12 & 0 \\ 4 & 0 & 0 & 12 \end{pmatrix} \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \end{matrix}$$

$$\begin{pmatrix} 9 & 0 & 0 & 0 \\ 1 & 9 & 3 & 3 \\ 3 & 0 & 9 & 0 \\ 3 & 0 & 0 & 9 \end{pmatrix}$$

$$\begin{pmatrix} 12 & 0 & 0 & 0 \\ 0 & 12 & 0 & 4 \\ 4 & 0 & 12 & 0 \\ 0 & 0 & 0 & 12 \end{pmatrix}$$

# Statistische Merkmale

Energie:

$$M_1 = \sum_{z=1}^Z \sum_{s=1}^S f(z,s)^2$$

Kontrast:

$$M_2 = \sum_{z=1}^Z \sum_{s=1}^S (z-s)^2 \cdot f(z,s)$$

Entropie:

$$M_3 = \sum_{z=1}^Z \sum_{s=1}^S f(z,s) \cdot \log(f(z,s))$$

Homogenität:

$$M_4 = \sum_{z=1}^Z \sum_{s=1}^S \frac{f(z,s)}{1+|z-s|}$$