

# Was bisher geschah

- ▶ Definition digitaler Bilder  $B : \text{pos} \rightarrow \text{col}$
- ▶ Bildanalyse, statistische Merkmale
- ▶ Signale im Orts- und Frequenzraum
- ▶ Bildbearbeitung durch
  - ▶ Punktoperationen (Farbtransformation)  $f : \text{col}_1 \rightarrow \text{col}_2$   
(punktweise Fortsetzung auf Gesamtbild)
  - ▶ geometrische Transformationen  
(Koordinatentransformation)  $f : \text{pos}_1 \rightarrow \text{pos}_2$
  - ▶ lokale Operationen (abhängig von Nachbarschaft):  
Filter, morphologische Operationen
- ▶ Segmentierung: Erkennung von Regionen im Bild
  - ▶ punktbasiert, z.B. Schwellwertverfahren
  - ▶ regionenbasiert, z.B. region growing
  - ▶ modellbasiert, z.B. Hough-Transformation

# Kantenbasierte Segmentierung

Kanten / Konturen:

meist starke Unterschiede in Intensität, Farbe oder Textur innerhalb kleiner Nachbarschaften

Vorsicht:

Unterschiede haben mitunter andere Ursachen, z.B. Schatten

Schritte der kantenbasierten Segmentierung:

1. Kantenextraktion (Finden potentieller Kantenpunkte)  
z.B. durch Filter, morphologische Operationen  
Ergebnis: Menge von Positionen (evtl. gewichtet)
2. Kantenverfolgung:  
Zusammenfügen von Kanten zu geschlossenen Konturen von Regionen (dabei evtl. Lücken schließen)  
z.B. Edge linking, Canny-Edge-Detector
3. Bestimmung / Markierung der Regionen innerhalb der Konturen

# Edge-Linking

- gegeben: Bild  $B : \text{pos} \rightarrow \text{col}$ ,  
Menge  $P \subseteq \text{pos}$  von **potentiellen** Kantenpositionen  
(z.B. durch Filter, morphologische Operationen)
- gesucht: zusammenhängende Menge  $K \subseteq \text{pos}$   
von Positionen (Konturen)

Idee:

- ▶ systematische Untersuchung aller potentiellen Kantenpositionen in  $P$
- ▶ jeweils Suche nach Kantenpositionen in einer Nachbarschaft in bisheriger Kantenrichtung (parametrisiert durch Richtungsabweichung und Abstand)
- ▶ dabei Konstruktion einer zusammenhängenden Menge  $K$  der tatsächlichen Kantenpositionen  
evtl. auch durch Einfügen kurzer Verbindungen zwischen Kantenpositionen

# Edge-Linking-Algorithmus

- ▶ Beginn an beliebiger (potentieller) Kantenposition  $p \in P$
- ▶ Wiederholung, solange  $P \neq \emptyset$ 
  1. Auswahl eines beliebigen  $p \in P$ ,
  2.  $p$  zu (neuem) Konturabschnitt  $K$  hinzufügen,
  3. zu  $p$  (in bisheriger Kantenrichtung) benachbarte Positionen  $q \in \text{pos}$  mit ähnlichen Kanteneigenschaften (Richtung, Intensität) zu  $P$  und  $K$  hinzufügen,
  4.  $p$  aus  $P$  entfernen,
- ▶ Ende, wenn alle Kantenpositionen in  $P$  verarbeitet

Ergebnis abhängig von

- ▶ (richtungsabhängiger) Nachbarschaftrelation
- ▶ Ähnlichkeitbedingung

Nachbearbeitung möglich, z.B.

- ▶ isolierte Kantenpositionen entfernen
- ▶ Kantenverdünnung / Skelettierung

# Canny-Edge-Detector

1. Glättung mit Gauß-Filter (Tiefpass, Entrauschen)

z.B.  $\frac{1}{16}$ 

1	2	1
2	4	2
1	2	1

 oder  $\frac{1}{24}$ 

1	3	1
3	8	3
1	3	1

2. Kantendetektion (Sobel-Filter)

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

3. Bestimmung der Richtung (aus den in 2. berechneten partiellen Ableitungen) und der Stärke der Kante (Intensität)
4. Reduktion der Kanten auf Breite 1 (Pixel)  
Entfernen der Nachbarpixel senkrecht zur Kantenrichtung mit geringerer Intensität (non-maximal suppression)
5. Kantenverfolgung mit zwei Schwellwerten (Hysterese)  
zur Trennung zwischen sicherem / unsicherem / kein Kantenpixel
6. Zuordnung der unsicheren Kantenpunkte durch (Versuch der) Verbindung mit benachbarten sicheren Kantenpunkten  
Ausgangspunkt: sicherer Kantenpunkt  
unsichere 8-Nachbarn in Kante aufnehmen, wenn Verbindung nur durch sichere Nachbarn nicht möglich ist

# Klassen-Einteilung

**Merkmal** (Attribut): definiert durch Name und Wertebereich  
z.B. Alter  $\in \mathbb{N}$ , Geschlecht  $\in \{m, w\}$

Merkmalsraum: kartesisches Produkt der Bereiche

**Objekt** (Individuum): charakterisiert durch Ausprägungen  
(Werte) der Merkmale (Merkmalsvektor)

z.B. Alter = 23, Geschlecht =  $m$

Punkt (Vektor) im Merkmalsraum, Punktmenge bei  
unbekannten Merkmalsausprägungen

**Klasse** Menge von Objekten mit denselben (ähnlichen)  
Merkmalsausprägungen  
definiert durch

- ▶ Angabe der Merkmale und (Bereichen von)  
Ausprägungen  
z.B. volljährig: Alter  $\geq 18$ , Geschlecht  $\in \{m, w\}$
- ▶ oder Referenzmodell (typischer Vertreter der  
Klasse) und Distanzmaß

Klassen müssen Zerlegung des Merkmalsraumes  
bilden

**Klassifizierung** (eindeutige) Zuordnung der Objektes zu Klassen

# Merkmale von Regionen im Bild

Region  $R \subseteq \text{pos}$  im Bild  $B : \text{pos} \rightarrow \{0, 1\}$   
(bzgl. einer Nachbarschaft zusammenhängend)

Merkmale (Features):

- ▶ topologische Merkmale
- ▶ quantitative Merkmale
- ▶ Formmerkmale
- ▶ geometrische Merkmale

ideal: Merkmale invariant gegenüber Verschiebung, Drehung, Skalierung, Verzerrung

# Topologische Merkmale

**Region:** Menge von Pixeln

**Kontur:** Menge der Randpixel  
(verschiedene Codierungen)

**Skelett:** topologieerhaltendes Bild aus  
Linien von 1 Pixel Dicke

- ▶ Zusammenhangseigenschaften
- ▶ Anzahl von Löchern (Euler-Zahl =  $|\text{Regionen}| - |\text{Löcher}|$ )
- ▶ Eigenschaften von (Ausschnitten des)  
Regionen-Adjazenz-Graphen
- ▶ Symmetrien
- ▶ Kontur der konvexen Hülle

Ziel: invariant gegenüber Verschiebung, Drehung, Skalierung,  
Verzerrung, Helligkeit, Beleuchtung  
z.B. für Zeichen, Gesten

# Quantitative Merkmale

Länge, Breite

Flächeninhalt  $A(R)$ : Anzahl der Positionen in der Region

Umfang der Kontur:  $U(R)$ : Länge, Anzahl der Randpositionen  
(evtl. verschieden gewichtet)

Durchmesser: größter (minimaler Umkreis), kleinster  
(maximaler Inkreis)

Masse: Summe der Intensitäten aller Pixel der Region

# Geometrische Merkmale

Verhältnisse

**Seitenverhältnis:** des kleinsten umschließenden  
(achsenparallelen) Rechtecks (bounding box)

**Kompaktheit:** Verhältnis  $A(R) : U(R)$

**Kreisförmigkeit:**  $4\pi A(R) / U(R)^2$

**fill factor:** Verhältnis  $A(R)$  : Fläche der Bounding-Box

**Dichte:** Verhältnis  $A(R)$  : Fläche der konvexen Hülle der  
Region

**Konvexität:** Verhältnis  $U(R)$  : Umfang der konvexen Hülle der  
Region

Symmetrien

Projektionen (z.B. Linescans)

# Statistische Merkmale

## Schwerpunkt

$$s_x(R) = \frac{1}{A(R)} \sum_{r \in R} r_x$$

$$s_y(R) = \frac{1}{A(R)} \sum_{r \in R} r_y$$

**Intensitäts-Schwerpunkt** von Regionen in Bildern  $B : \text{pos} \rightarrow \text{col}$

$$s'_x(R) = \frac{1}{\sum_{r \in R} B(r)} \sum_{r \in R} r_x B(r)$$

$$s'_y(R) = \frac{1}{\sum_{r \in R} B(r)} \sum_{r \in R} r_y B(r)$$

# Momente

(Region als 2d-Verteilung)

Momente der Region  $R$  im Bild  $B$  : pos  $\rightarrow$  col

$$m_{pq}(R) = \sum_{r \in R} B(r) r_x^p r_y^q$$

Momente der Region  $R \subseteq B$  im Binärbild  $B$  : pos  $\rightarrow$   $\{0, 1\}$

$$m_{pq}(R) = \sum_{r \in R} r_x^p r_y^q$$

Beispiele (für Binärbilder):

- ▶  $m_{00}(R) = \sum_{r \in R} r_x^0 r_y^0 = \sum_{r \in R} 1 = A(R)$
- ▶  $s_x(R) = \frac{1}{A(R)} \sum_{r \in R} r_x^1 r_y^0 = m_{10}(R),$   
 $s_y(R) = \frac{1}{A(R)} \sum_{r \in R} r_x^0 r_y^1 = m_{01}(R)$

## Zentrale Momente

Idee: Verschiebungs-Invarianz um Schwerpunkt

zentrale Momente der Region  $R \subseteq B$ :

$$\mu_{pq}(R) = \sum_{r \in R} B(r) r_{x-s_x(R)}^p r_{y-s_y(R)}^q$$

Verschiebungs-, aber nicht Größen-invariant

Größen-invariante zentrale Momente der Region  $R \subseteq B$ :

$$\mu'_{pq}(R) = \frac{\mu_{pq}(R)}{\mu_{00}(R)^{(p+q+2)/2}}$$

Daraus lassen sich weitere häufig verwendete Merkmale berechnen, z.B. Orientierung, Exzentrizität, Hu-Momente.

# Histograms of Oriented Gradients (HOG)

Idee:

- ▶ lokale Gradientenstruktur ist charakteristisches Merkmal für Regionen und Objekte
- ▶ invariant gegenüber affinen Transformationen und Beleuchtung

Berechnung der HOG (gewichtete)

1. evtl. Glättung
2. Zerlegung des Bildes (bzw. Bildausschnittes) in quadratische Teilbilder (Zellen)  $Z_1, \dots, Z_l$
3. Prewittfilter  $Z_h = (-1, 0, 1)Z$ ,  $Z_v = Z(-1, 0, 1)^T$
4. Bestimmung der in der Zelle vorkommenden Gradienten
  - ▶ lokaler Betrag  $b(p) = \sqrt{Z_h(p)^2 + Z_v(p)^2}$
  - ▶ lokale Richtung  $\theta(p) = \tan^{-1}\left(\frac{Z_v(p)}{Z_h(p)}\right)$  (diskretisiert)
5. Histogramme jeder Zelle  $H_i : \theta \rightarrow \mathbb{R}_{\geq 0}$ ,  $\theta \mapsto |\{p \in Z \mid \theta(p) = \theta\}|$   
(evtl. gewichtet durch  $b(p)$  und Abstand zum Zellmittelpunkt)
6. Normierung der Histogramme über Blöcken von Zellen (ernstfernt Abhängigkeit von Beleuchtung)

Merkmalsvektor des Bildes (Bildausschnittes):

Verkettung der Histogramme aller Blöcke

# Beispiel Verkehrszeichen-Erkennung

**Farb-basiert** statistische Werte verschiedener Farbkanäle, Farb-Verhältnisse und -Differenzen  
Binarisierung und Regionen-Erkennung in verschiedenen Farbkanälen

- + Störungen, verdeckte Schilder  
Extraktion relevanter Bildbereiche
- abhängig von Beleuchtung, Umgebung, Alter und Farbzustand der Schilder

**Form-basiert** Anwendung Klassifikatoren für die fünf typischen Schilderformen auf Bildbereiche verschiedener Größe und Position

- + unabhängig von Beleuchtung, Alter und Farbzustand der Schilder
- Verzerrung, Störungen, verdeckte, gekippte Schilder

Kombination beider Ansätze

ggf. Normierung der farbbasiert gefundenen Bereiche

# Modellbasierte Merkmalsextraktion

Idee: Verwendung von Kontextinformation über zusammenhängende Regionen

Repräsentation von Kandidaten für den Bildinhalt (Modelle), z.B.

- ▶ Mensch aus (bevorzugt) zusammenhängenden Regionen  
Kopf, Rumpf, Arme, Beine
- ▶ Haus mit Fenstern, Dach, Türen
- ▶ Landschaft mit Strand, Meer, Himmel
- ▶ Achtelnote mit Kopf, Hals, Fähnchen

Modell:

- ▶ Menge von Regionen mit jeweils typischen Eigenschaften (Klassen)
- ▶ Beziehungen (Zusammenhang) zwischen den Regionen

Modell = Graph

Klassifikation von (Regionen in) Bildern =

Vergleich des Modellgraphen mit (Teilgraphen des)

Regionen-Adjazenz-Graphen des Bildes

# Klassifikation von Regionen

Ziel: Zuordnung von (Regionen in) Bildern zu Klassen  
(Bedeutung)

Vorbereitung:

- ▶ Segmentierung des Bildes
- ▶ Extraktion der (signifikanten) Merkmale der Bildregionen
- ▶ evtl. Normierung der Merkmale

Klassifikation bei Klassen-Definition durch

- ▶ Merkmalsausprägungen:  
entprechend den extrahierten Bild-Merkmalen
- ▶ Referenzmodelle:  
Zuordnung zur Klasse mit dem ähnlichsten Referenzmodell  
(geringster Abstand im gegebenen Distanzmaß)

Klassifikations-Methoden (vorgestellt in anderen Modulen):  
statistisch, Mustererkennung, Clustering, maschinelles Lernen,  
neuronale Netze, Support Vector Machines

# Bildfolgen

Bild  $B : \text{pos} \rightarrow \text{col}$

Folge von Bildern  $B_0, B_1, B_2, \dots$

$$B : \mathbb{N} \rightarrow (\text{pos} \rightarrow \text{col})$$

z.B. Videosequenz

# Operationen auf Bildfolgen

- ▶ Mittelung mehrerer Bilder

$$\forall p \in \text{pos} : B(p) = \frac{1}{n} \sum_{i=1}^n B_i(p)$$

zum Entrauschen, z.B. in Medizintechnik

- ▶ Differenz zwischen aufeinanderfolgenden Bildern

$$\forall p \in \text{pos} : B(p) = |B_{i+1}(p) - B_i(p)|$$

zur Hervorhebung / Erfassung zeitlicher Veränderungen

- ▶ Detektion von Bewegung (z.B. zur Überwachung)
- ▶ Detektion von Intensitätsänderung (z.B. Wirkung von Kontrastmitteln)
- ▶ Videokompression (MPEG)
- ▶ Feststellung korrespondierender Punkte in mehreren Bildern
- ▶ Mosaicing: Zusammensetzen mehrerer Teilbilder zu einem Gesamtbild (z.B. Panoramabild, Karte aus Satellitenaufnahmen)

# Korrespondenzen in Bildern

Ziel: Bestimmung von Punkten mit derselben Bedeutung in mehreren Bildern

Anwendung z.B. bei

- ▶ Kombination der Informationen mehrerer mit verschiedenen Kameras aufgenommener Bildkanäle
- ▶ Stereobildverarbeitung
- ▶ Suche in Bilddatenbanken
- ▶ Zusammensetzen von Bildern aus überlappenden Teilbildern

z.B. durch descriptives Matching

# Hervorhebung von Bewegungen

Annahme:

- ▶ statischer Hintergrund
- ▶ bewegte Objekte
  
- ▶ Mittelung mehrerer aufeinanderfolgender Bilder:  
bewegtes Objekt „verschwimmt“ mit Hintergrund
- ▶ (Betrag der) Differenz aufeinanderfolgender Bilder ist bei  
bewegten Objekten hoch,  
gering beim statischen Hintergrund