

Intelligente Sehsysteme

10 Objekterkennung

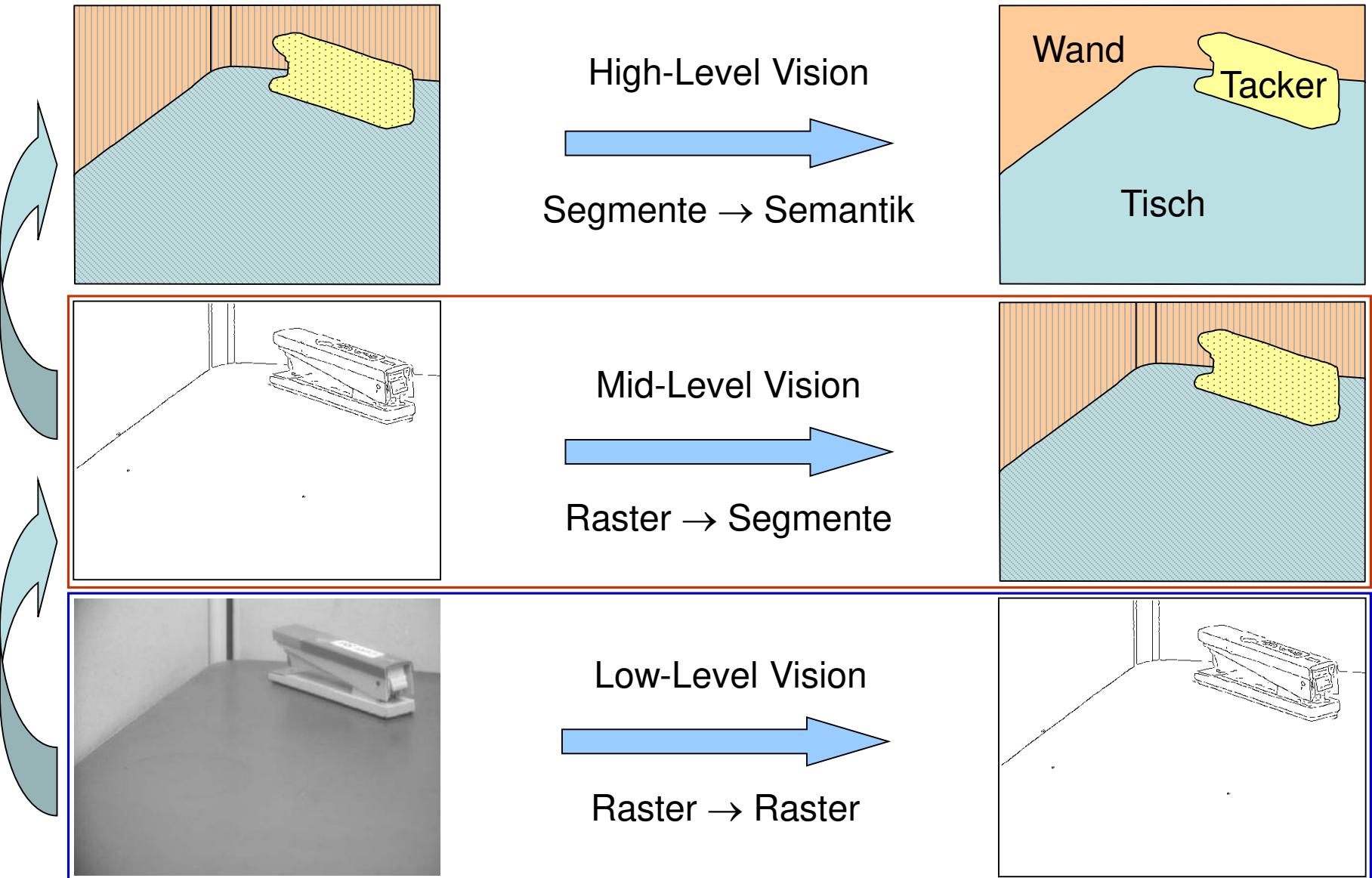
Segmentmerkmale
regionen- und formbasierte Merkmale
MAP-Klassifikation

Volker Steinhage

Inhalt

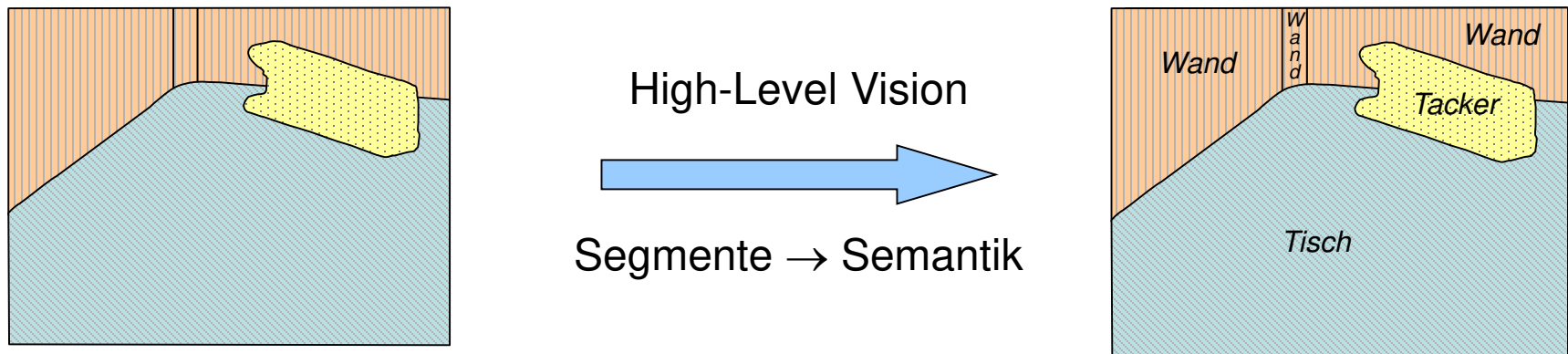
- Segmentmerkmale: regionen- und formbasierte Merkmale
- Within-Class-Scatter & Between-Class-Scatter
- MAP-Klassifikation
 - A-Posteriori-W'keiten, Satz von Bayes, klassenspezifische Merkmalsverteilungen und A-Priori-W'keiten
 - Schätzung von klassenspezifische Merkmalsverteilungen und A-Priori-W'keiten aus Segmentmerkmalen

Phasen des Computersehens (Wiederholung)



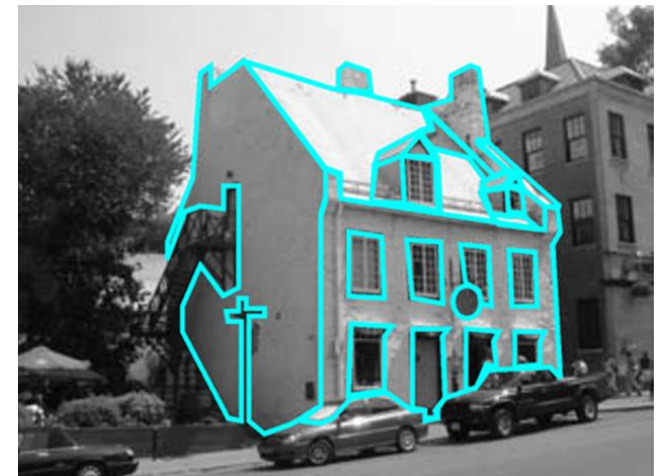
Ausgangspunkt: Segmentierung

- Segmentierte Bilder zeigen Segmente. Bei der **Segmentierung** wurden lediglich **allgemeine Annahmen** über die Ähnlichkeiten von Pixeln oder die Verläufe von Konturen eingesetzt.
- Bei der **Objekterkennung** werden **domänenspezifische Modelle** der zu erkennenden Objekte oder Objektklassen verwendet.
- Damit gehört die Objekterkennung zur 3. Phase der visuellen Verarbeitung, der High-Level-Vision.



Objekterkennung

- Bedeutungszuordnung erfolgt i.A. durch Zuordnung der Segmente zu Objektklassen
- Dazu müssen Beschreibungen der Objektklassen vorliegen: Modelle der Klassen.
- Im einfachsten Fall* basiert ein Klassenmodell alleine auf Eigenschaften der Segmente, den Segmentmerkmalen.
- Die Klassifikationsaufgabe wird dann auf Ähnlichkeitsmaße in einem Merkmalsraum reduziert.

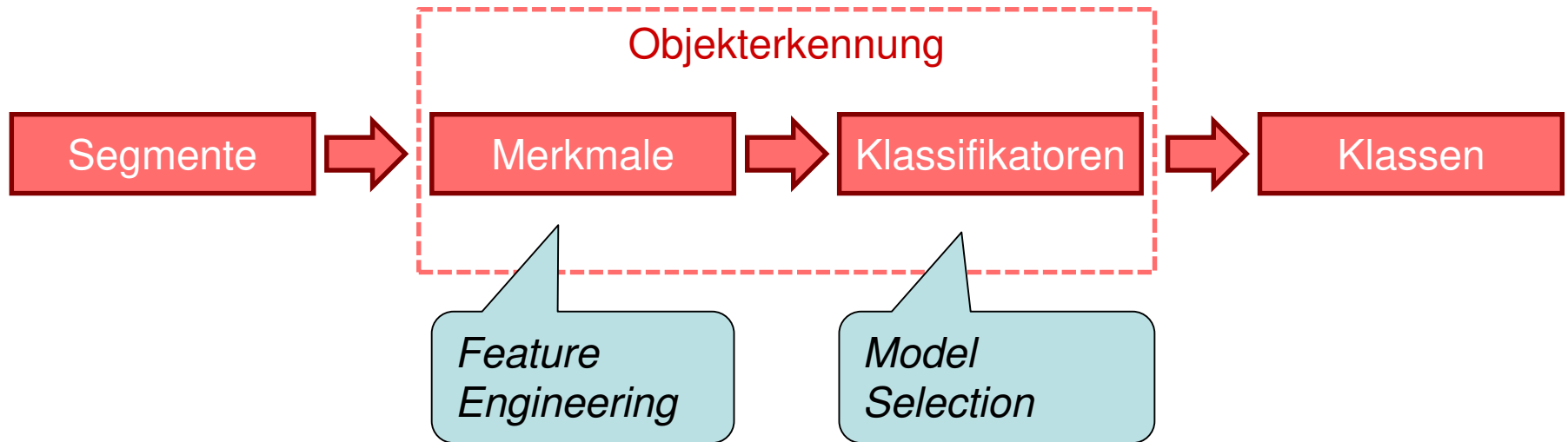


* Komplexere Objekte können durch Segmentgruppen abgebildet werden. Eine Gaube kann aus Dach-, Wand- und Fensterflächen bestehen. Die Klassifikation vergleicht dann Segmentgraphen G_S mit Modellgraphen (Subgraph-Matching).

Entwurfsaspekte der Objekterkennung

Damit zeigt die Objekterkennung zwei Entwurfsaspekte

- (1) **Feature Engineering**: Welche Merkmale (engl. Features) sind geeignet?
- (2) **Model Selection**: Welches Modell des maschinellen Lernens* ist geeignet?



* Siehe Vorlesung „Grundlagen der KI“: Lernverfahren, Hyperparameter, etc. (SVM, LDA, Boosting, KNN, Clustering, ...)

Merkmale und Merkmalsräume

- Gegeben sei ein einzelnes Segment s .
- Das Segment s werde durch M Merkmale m_1, \dots, m_M des Segmentinneren und/oder des Segmentrandes beschrieben.
- Jedes Merkmal m_i sei ein Skalar, das einen bestimmten Aspekt des Objektmodells beschreibt.
- Die M Merkmale m_1, \dots, m_M
 - werden in einem M -dim. Merkmalsvektor $\mathbf{m} = (m_1, \dots, m_M)$ zusammengefasst,
 - spannen so einen M -dimensionalen Merkmalsraum auf.
- Jedes Segment s wird über seinen Merkmalsvektor $\mathbf{m}(s)$ auf einen bestimmten Ort im Merkmalsraum abgebildet.

Merkmale: Beispiel

Zwei Merkmale für Bildsegmente von Schrauben und Muttern:

- Anzahl L der Löcher
- Kreisförmigkeitsmerkmal Kr

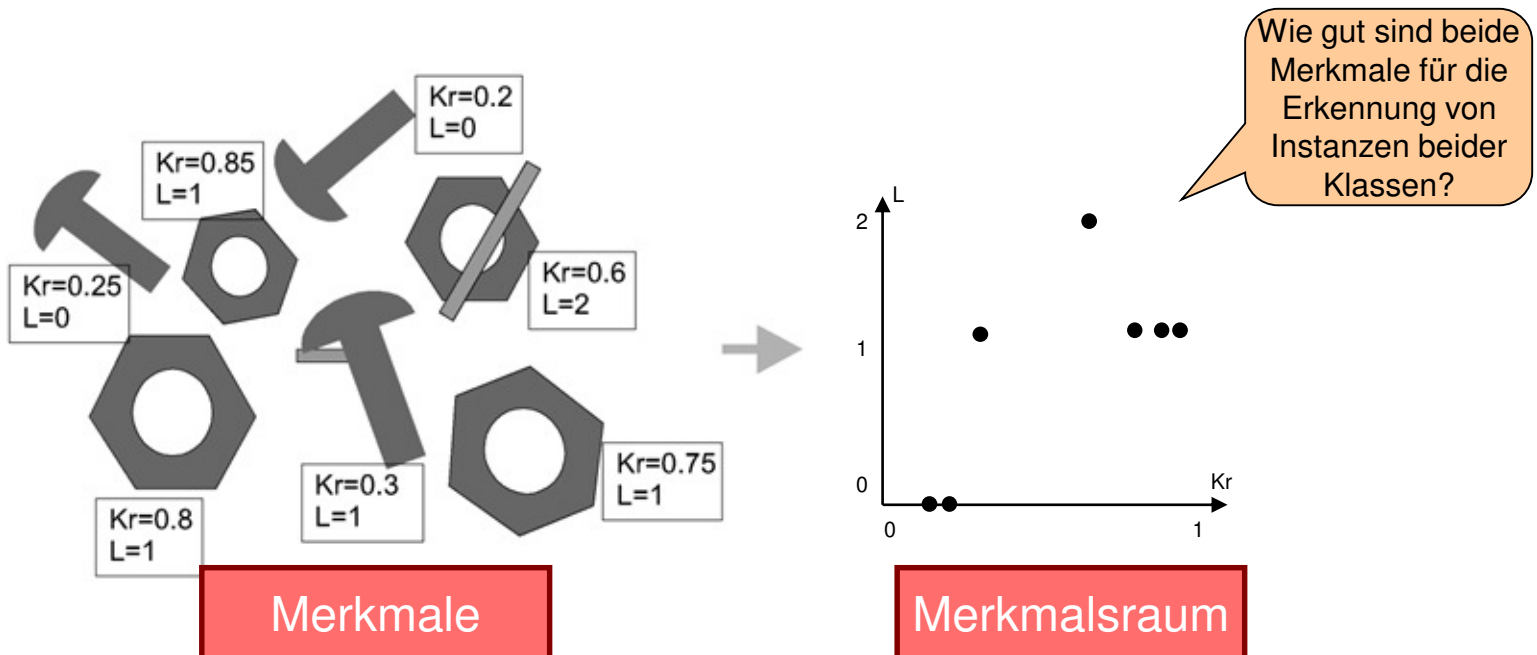
Wie gut sind beide Merkmale für die Erkennung von Instanzen beider Klassen?



Merkmale: Beispiel (Forts.)

Zwei Merkmale für Bildsegmente von Schrauben und Muttern:

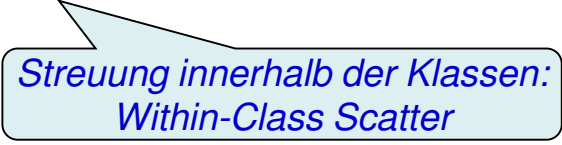
- Anzahl L der Löcher
- Kreisförmigkeitsmerkmal Kr



Within-Class Scatter & Between-Class Scatter (1)

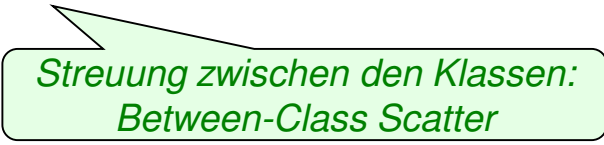
Die M Merkmale m_1, \dots, m_M sind gut gewählt, wenn

- 1) Segmente *einer Klasse* in eine eng begrenzte Region des Merkmalsraums abgebildet werden,



*Streuung innerhalb der Klassen:
Within-Class Scatter*

- 2) Regionen von Segmenten *aus verschiedenen Klassen* gut voneinander getrennt sind.



*Streuung zwischen den Klassen:
Between-Class Scatter*

Within-Class Scatter & Between-Class Scatter (2)

Sei \mathbf{T} eine Menge von klassifizierten Segmenten nach K Klassen c_0, \dots, c_{K-1} , die \mathbf{T} in K disjunkte Mengen $\mathbf{T}_0, \dots, \mathbf{T}_{K-1}$ unterteilen. Dann rechnet sich

- *Within-Class Scatter* S_w (dt. *Intragruppenvarianz*) als durchschnittliche Varianz der Merkmalsvektoren jeder Klasse und
- *Between-Class Scatter* S_b (dt. *Intergruppenvarianz*) als mittlerer quadratischer Abstand zwischen den Mittelwerten der Klassen:

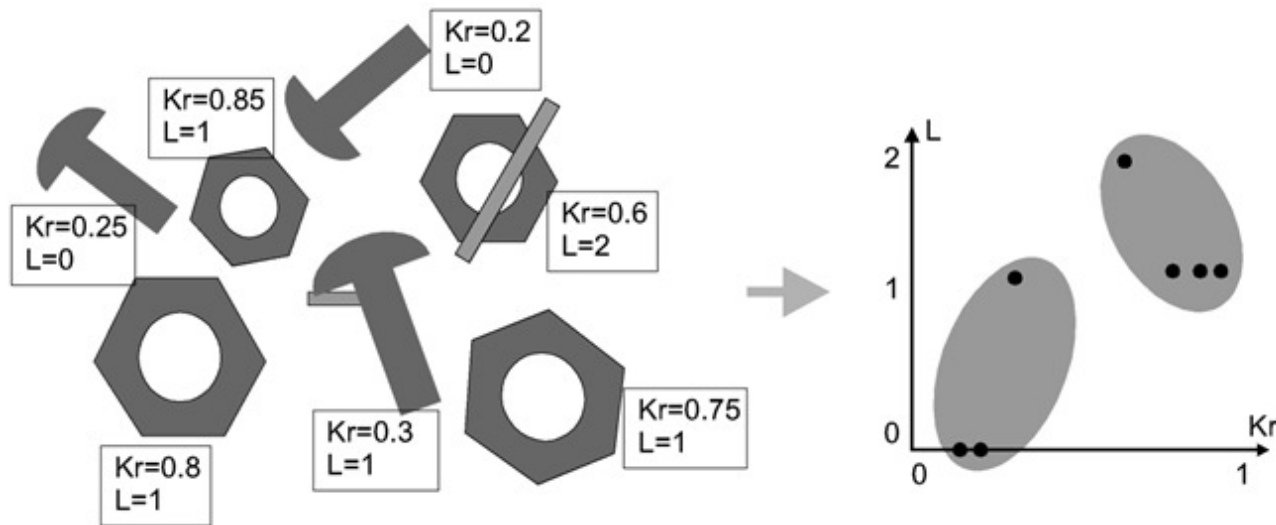
$$S_w = \frac{1}{|\mathbf{T}|} \sum_{k=0}^{K-1} \sum_{s \in T_k} \|(\mathbf{m}(s) - \boldsymbol{\mu}_k)\|^2,$$

$$S_b = \frac{1}{n} \sum_{k_1=0}^{K-1} \sum_{k_2 > k_1}^{K-1} \|(\boldsymbol{\mu}_{k_1} - \boldsymbol{\mu}_{k_2})\|^2$$

$$\text{mit } \boldsymbol{\mu}_k = \frac{1}{|\mathbf{T}_k|} \sum_{s \in T_k} \mathbf{m}(s) \text{ und } n = \binom{K}{2}.$$

Merkmale: Beispiel (Forts.)

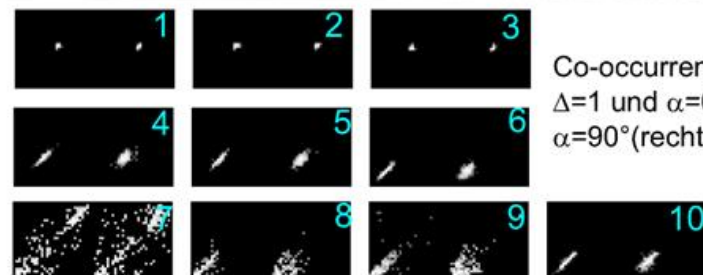
Die Merkmale sind hinreichend gut gewählt, da selbst kleinere Störungen die Trennbarkeit der beiden Klassen nicht beeinflussen.



Regionenbasierte Merkmale

- Werden auch Grauwert oder Farbe eines Segments als Textur bezeichnet, **dann sind alle regionenbasierte Merkmale Texturen.**
- Neben dem durchschnittl. Grauwert eines Segments können dies die Varianz und die in Vorlesung 5 vorgestellten Haralickschen Texturmaße aus den Co-Occurrence-Matrizen sein.

Beispiel: die Co-Occurrence-Matrizen für verschiedene Gebiete in einem Grauwertbild. Man erkennt, dass sie sich für die Klassen „Himmel“ (1,2,3), Gebäude (4,5,6) und Wasser (7,8,9,10) unterscheiden. Eine Klassifikation über Haralickschen Texturmaße in diese drei Klassen sollte möglich sein.



Co-occurrence-Matrizen für $\Delta=1$ und $\alpha=0^\circ$ (links) bzw. $\alpha=90^\circ$ (rechts)

Formbasierte Merkmale: Im Vergleich zu Texturmerkmalen

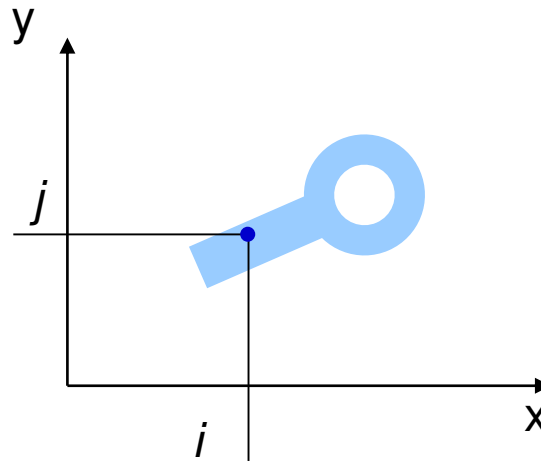
- Formmerkmale definieren sich *ausschließlich über den Objekt- bzw. Segmentrand.*
- Die Folge: *im Ansatz* sind Formmerkmale im Vergleich zu Texturmerkmalen anfälliger gegenüber leichten Segmentierungsfehlern,
 - da diese den Segmentrand leicht korrumpieren können
 - während für Texturmerkmale nur hinreichend viele Stichproben aus dem Segmentinneren nötig sind.
- Einfache Formmerkmale *erhalten ihre Robustheit gegenüber Segmentierungsfehlern durch die Integration des Merkmals* über das gesamte Objekt bzw. Segment.

Formbasierte Merkmale: Flächeninhalt

- Zu den einfachen Formmerkmalen eines Segments s zählen:

1) der Flächeninhalt $F(s) = \text{Zahl aller Pixel von } s$
$$= |\{ (i,j) \mid (i,j) \in s \}|$$

$F(s)$ ist dann translations- und rotationsrobust, wenn die Rasterung vernachlässigbar ist, also bei hinreichend großen, kompakten Bereichen.

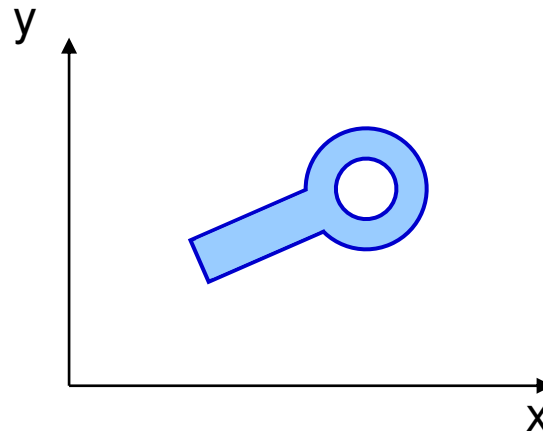


$$F(s) = |\{ (i,j) \mid (i,j) \in s \}|$$

Formbasierte Merkmale: Rand

- Zu den einfachen Formmerkmalen eines Segments s zählen:

2) Die Randlänge $R(s)$ = Zahl der Pixel von s , die mind. ein Nachbarpixel in ihrer 8-Nachbarschaft außerhalb von s haben,
$$= | \{ (i,j) \in s \text{ mit } (k,l) \in N_8(i,j) \wedge (k,l) \notin s \} |$$



$$R(s) = | \{ (i,j) \in s \text{ mit } (k,l) \in N_8(i,j) \wedge (k,l) \notin s \} |$$

Formbasierte Merkmale: Kompaktheit

- Zu den einfachen Formmerkmalen eines Segments s zählen:
 - 3) die Kompaktheit $K(s)$ = Verhältnis von Flächeninhalt von s
zum Quadrat der Randlänge von s
$$= F(s) / R(s)^2$$

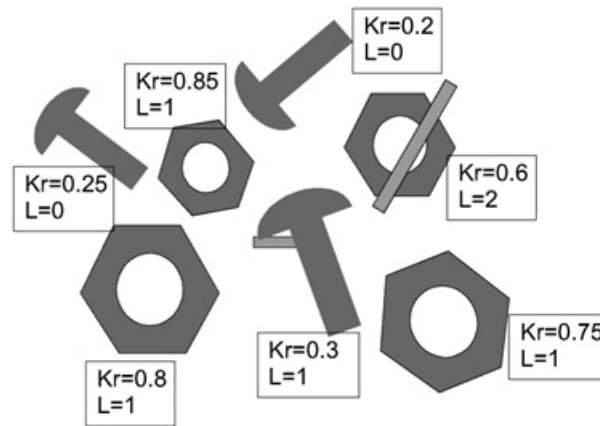
Durch die Quadrierung der Konturlänge wird eine gewisse Skalierungsrobustheit erreicht.

Dennoch zeigen sich Schwankungen bei der Kompaktheit für verschiedene Größen ähnlicher Formen: z.B. für ein diagonal zum Raster orientiertes Quadrat ist $K(s)=0.2$ für $F(s)=13$ und $K(s)=0.15$ für $F(s)=61$.

Formbasierte Merkmale: Kreisähnlichkeit

- Zu den einfachen Formmerkmalen eines Segments s zählen:
 - 4) die Kreisähnlichkeit $Kr(s)$ = Verhältnis vom 4π -fachen Flächeninhalt von s zum Quadrat der Randlänge von s
$$= (F(s) \cdot 4\pi) / R^2(s)$$

Siehe Bspl.:

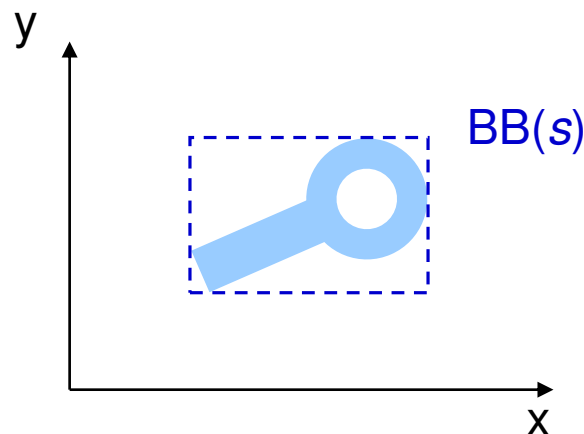


Formbasierte Merkmale: achsenparallele Bounding Box

- Zu den einfachen Formmerkmalen eines Segments s zählen:
 - 5) der Größe des kleinsten, achsenparallelen, umschließenden Rechtecks, der sog. Bounding Box $BB(s)$ von s :

$$F_{BB}(s) = \text{Höhe}_{BB}(s) \cdot \text{Breite}_{BB}(s)$$

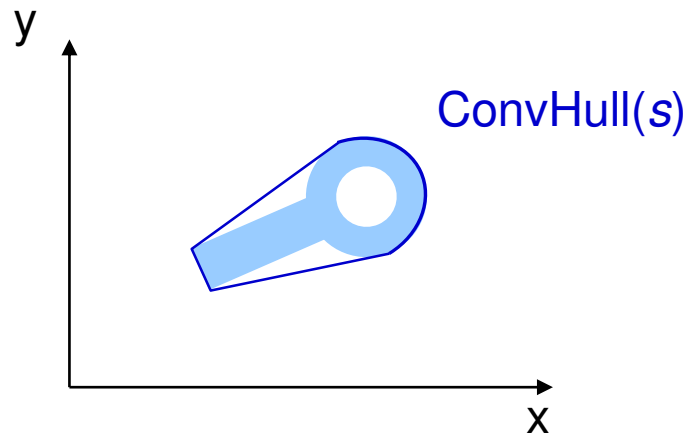
$F_{BB}(s)$ ist in hohem Maße orientierungsabhängig.



Formbasierte Merkmale: Konvexe Hülle

- Zu den einfachen Formmerkmalen eines Segments s zählen:
 - 6) Die Flächendifferenz zwischen der Segmentfläche $F(s)$ und der Fläche der konvexen Hülle von s :

$$F_{\text{Diff_ConvHull}}(s) = F_{\text{ConvHull}}(s) - F(s)$$



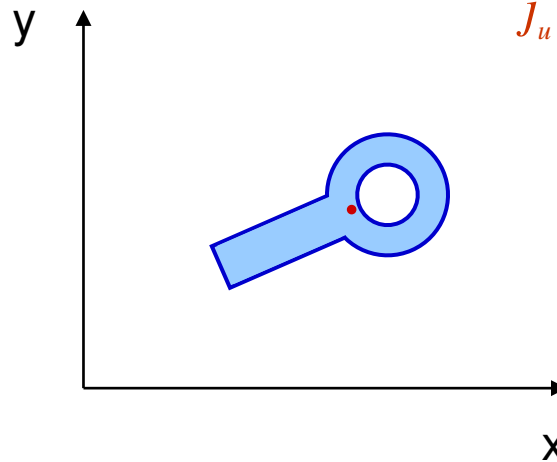
Formbasierte Merkmale: Schwerpunkt

- Die im folgenden dargestellten Flächenträgheitsmomente schätzen die Flächenverteilung um Achsen durch den Schwerpunkt μ

(i_μ, j_μ) = Schwerpunktkoordinaten von s mit

$$i_u = \frac{1}{F(s)} \cdot \sum_{(i,j) \in s} i,$$

$$j_u = \frac{1}{F(s)} \cdot \sum_{(i,j) \in s} j.$$

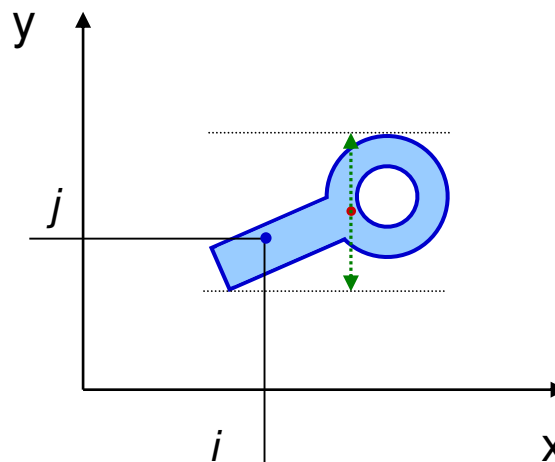


Formbasierte Merkmale: *Trägheitsmoment in y-Richtung*

- Basierend auf dem Schwerpunkt μ :

7) Flächenträgheitsmoment um Achse durch μ in x-Richtung durch
Integration der Abstände aller Flächenelemente in y-Richtung

$$m_x(s) = \sum_{(i,j) \in s} (j - j_\mu)^2$$

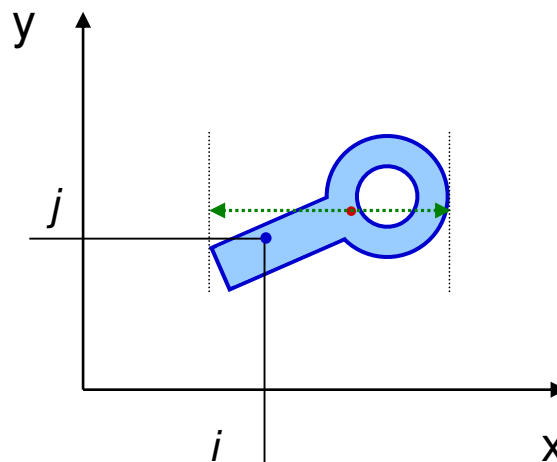


Formbasierte Merkmale: *Trägheitsmoment in x-Richtung*

- Basierend auf dem Schwerpunkt μ :

7) Flächenträgheitsmoment um Achse durch μ in y-Richtung durch
Integration der Abstände aller Flächenelemente in x-Richtung

$$m_y(s) = \sum_{(i,j) \in s} (i - i_\mu)^2$$



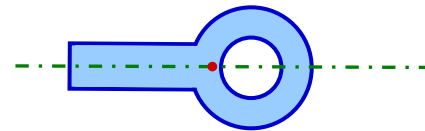
Formbasierte Merkmale: gemischtes *Trägheitsmoment*

- Basierend auf dem Schwerpunkt μ :

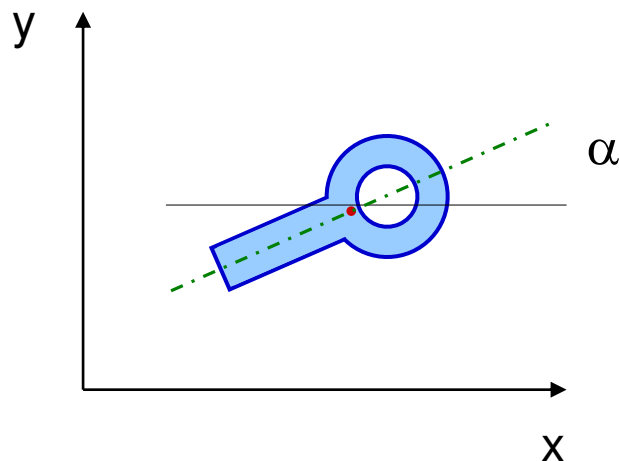
9) Gemischtes Flächenträgheitsmoment

$$m_{xy}(s) = \sum_{(i,j) \in s} (i - i_\mu)(j - j_\mu)$$

$m_{xy}(s) = 0$, wenn eine Bezugsachsenrichtung eine **Symmetrieachse** darstellt



Andernfalls ist **Drehwinkel** α berechenbar nach $\tan(2\alpha) = 2m_{xy}/(m_y - m_x)$.



Anforderungen an Klassifikationsmodelle

- 1) den Segmenten sollen **Bedeutungen** zugeordnet werden
- 2) diesen Entscheidungen sollen **Qualitätsangaben** zugeordnet werden
- 3) die **W'keit einer Fehlzuordnung soll minimal** sein

MAP-Klassifikation



Maximum-a-Posteriori-Klassifikation (MAP-Klassifikation) erfüllt die Anforderungen, indem für ein Segment s mit Merkmalsvektor $\mathbf{m}(s)$ die W'keit $P(s = c_k | \mathbf{m}(s))$ berechnet wird, mit der s zur Klasse c_k gehört.

- Diese bedingte W'keit $P(s = c_k | \mathbf{m}(s))$ heißt **A-Posteriori-Wahrscheinlichkeit**.
- Die W'keit für Fehlzusordnungen wird minimiert, indem für jeden Merkmalsvektor $\mathbf{m}(s)$ **immer die Klasse c_k gewählt wird, welche die A-Posteriori-W'keit $P(s = c_k | \mathbf{m}(s))$ maximiert:**

⇒ Maximum-a-posteriori-Methode (= MAP)

$$P(s = c_k | \mathbf{m}(s)) \geq P(s = c_j | \mathbf{m}(s)) \text{ für alle } j \neq k.$$

MAP-Klassifikation: Likelihood

- Die **A-Posteriori-W'keit** $P(s = c_i \mid \mathbf{m}(s))$ ist aber i.A. schwierig direkt zu bestimmen:

- Für jeden Merkmalsvektor $\mathbf{m}(s)$ – also für *jede Merkmalskombination* – muss die W'keit seiner Zuordnung zu jeder Klasse bekannt sein.
- Leichter ist oft der umgekehrte Sachverhalt zu spezifizieren:
 - die **Wahrscheinlichkeit für einen bestimmten** Merkmalsvektor $\mathbf{m}(s)$ bei geg. Verteilungsparametern einer Objektklasse: $P(\mathbf{m}(s) \mid s = c_i)$.


MAP-Klassifikation: Likelihood & Maximum Likelihood

- Allg. handelt es sich um eine Funktion

$$L: \Theta \rightarrow \mathbb{R}, \theta \rightarrow f(x|\theta),$$

die auf der Basis einer Stichprobenverteilung in Form einer W'keitsdichtefunktion f von x die W'keit $f(x|\theta)$ für x angibt, wenn der (wahre) Parameter der Grundgesamtheit den Wert θ hat.

Eine solche Funktion L wird als **Likelihood**-Funktion* bezeichnet.

- Der Schätzwert desjenigen Parameters $\hat{\theta}_{ML}(x)$, bei dem die beobachtete Stichprobenrealisation x am plausibelsten ist, ist dann

$$\hat{\theta}_{ML}(x) = \arg \max_{\theta \in \Theta} f(x|\theta)$$

und heißt **Maximum-Likelihood-Schätzer** von θ .

* auch *Plausibilitätsfunktion* oder *Mutmaßlichkeitsfunktion*

MAP-Klassifikation: Satz von Bayes

- Nun stehe zusätzlich zur Stichprobenverteilung f eine **A-priori-Verteilung** g von θ zur Verfügung. Dadurch lässt sich θ als eine Zufallsvariable betrachten. Die **A-posteriori-Verteilung** von θ ergibt sich mit dem Satz von Bayes durch

$$\theta \rightarrow f(\theta|x) = \frac{f(x|\theta)g(\theta)}{\int_{\Theta} f(x|\theta')g(\theta')d\theta'} .$$

- Die Maximum-a-Posteriori-Methode verwendet jetzt den Modalwert der A-posteriori-Verteilung als Schätzwert für θ :

$$\hat{\theta}_{MAP}(x) = \arg \max_{\theta \in \Theta} \frac{f(x|\phi)g(\phi)}{\int_{\Theta} f(x|\phi')g(\phi')d\phi'} .$$

Das Integral im Nenner hängt nicht von θ ab, Daher muss es bei der Maximierung nicht berücksichtigt werden.

MAP-Klassifikation: Satz von Bayes

- Im folgenden werden die **A-Posteriori-W'keit** $P(s = c_k | \mathbf{m}(s))$ über den Satz von Bayes aus den **klassenspezifischen Merkmalsverteilungen** $P(\mathbf{m}(s) | s = c_i)$ sowie klassenspezifischen **A-Priori-W'keiten** $P(s = c_i)$ abgeleitet:

$$P(S = C_i | \mathbf{m}(S)) = \frac{P(\mathbf{m}(S) | S = C_i)P(S = C_i)}{\sum_{k=0}^{K-1} P(\mathbf{m}(S) | S = C_k)P(S = C_k)}.$$

Auch hier dient der Nenner der Normalisierung. Auf ihn kann verzichtet werden, sofern nur die Klasse mit *maximaler* A-Posteriori-W'keit gesucht wird, aber die genaue W'keit nicht relevant ist.

MAP-Klassifikation: Trainingsmenge

- Für die Anwendung des Satzes von Bayes sind zu schätzen

- die A-Priori-W'keiten $P(s = c_i)$
- die Merkmalsverteilungen $P(\mathbf{m}(s) \mid s = c_i)$.

⇒ MAP ist überwachtes Lernverfahren

- Beide Schätzungen basieren auf repräsentativer Trainingsmenge $\mathbf{T} = \mathbf{T}_1 \cup \mathbf{T}_2 \cup \dots \cup \mathbf{T}_K$ von bereits in K Klassen zugeordneten Segmenten.

MAP-Klassifikation: Bestimmung der A-Priori-W'keiten

Zwei Möglichkeiten zu Bestimmung der A-Priori-W'keiten $P(s = c_k)$:

- a) die $P(s = c_k)$ sind aus der Anwendungsdomäne bekannt
- b) die $P(s = c_k)$ sind aus der Verteilungsstatistik der Trainingsmenge $\mathbf{T} = \mathbf{T}_1 \cup \mathbf{T}_2 \cup \dots \mathbf{T}_K$ zu schätzen
 - Die Trainingsmenge \mathbf{T} muss repräsentativ für alle Klassen sein

Die A-Priori-W'keiten $P(s = c_k)$ müssen sich natürlich über alle K Klassen zu Eins aufsummieren: $\sum_{k=0, \dots, K-1} P(s = c_k) = 1$

MAP: Schätzung der Merkmalsverteilungen (1)

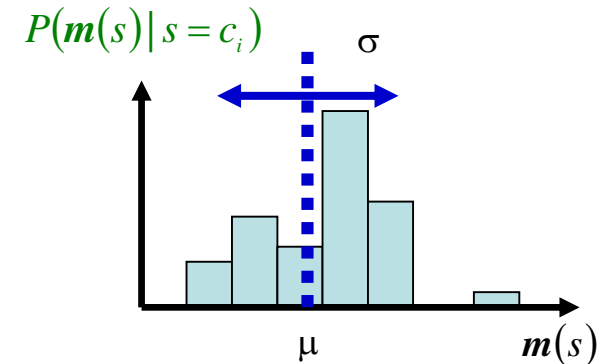
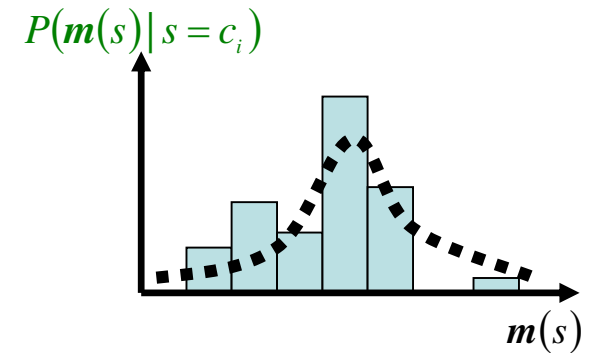
Schätzung der **klassenspezifischen Merkmalsverteilungen**
 $P(\mathbf{m}(s) \mid s = c_k)$:

- Bei M Merkmalen sind die Verteilungen als M -dimens. Funktionen zu schätzen
- Die Schätzung basiert auf aus \mathbf{T} abgeleiteten **Histogrammen** der Merkmalsvektoren der Klassen c_k

MAP: Schätzung der Merkmalsverteilungen (2)

Zur Schätzung aus den **Histogrammen der Merkmalsvektoren der Klassen c_k** :

- Schätzung einer beliebigen Funktion:
Interpolation der Funktion durch die Stichproben
- Schätzung von Parametern einer Funktion
 - Vorgabe einer **Verteilungscharakteristik** (z.B. Normalverteilung (das muss begründbar sein!))
 - Schätzung der Parameter



MAP: Schätzung der Merkmalsverteilungen (3)

- Schätzung der $P(\mathbf{m}(s) \mid s = c_i)$ bei unkorrelierten Merkmalen:
 - es wird die Unabhängigkeit der Merkmale angenommen und für *jedes* Merkmal $m_j(s)$ wird ein separates Histogramm erstellt und daraus die Funktion $P(m_j(s) \mid s = c_k)$ geschätzt
 - die Gesamtwahrscheinlichkeit ergibt sich als Produkt der Einzelw'keiten:

$$P(\mathbf{m}(s) \mid s = c_k) = P(m_1(s) \mid s = c_k) \cdot P(m_2(s) \mid s = c_k) \cdot \dots \cdot P(m_M(s) \mid s = c_k).$$



MAP: Schätzung der Merkmalsverteilungen (4)

- Schätzung der $P(\mathbf{m}(s) \mid s = c_i)$ bei unkorrelierten Merkmalen (Forts.):
 - Jede $P(m_j(s) \mid s = c_k)$ wird durch das normierte Histogramm des Merkmals m_j von allen Elementen der **Trainingsteilmenge** \mathbf{T}_k approximiert.
 - Der *Definitionsbereich* $[m_{j,min}, m_{j,max}]$ jedes Merkmals m_j muss ggf. in B **Bins** bzw. **Intervalle** $[m_{j,min}, m_{j,1}]$, $[m_{j,1}, m_{j,2}]$, ..., $[m_{j,B-1}, m_{j,max}]$ zerlegt werden (engl. *Binning*). Für jedes Intervall wird die Zahl der Stichproben \mathbf{t}_i mit Merkmalswerten von m_j in diesem Intervall ermittelt.
 - Durch B wird die Genauigkeit der Schätzung der Verteilungsfunktion festgelegt. Eine Daumenregel sagt, dass bei $N = |\mathbf{T}|$ Stichproben, $B = N^{1/2}$ eine gute Wahl ist (Bspl. $N = |\mathbf{T}| = 100 \leadsto B = 10$).

MAP: Schätzung der Merkmalsverteilungen (5)

- Schätzung der $P(\mathbf{m}(s) \mid s = c_i)$ bei korrelierten Merkmalen:
 - Die M -dim. Verteilungsfunktion $P(\mathbf{m}(s) \mid s = c_k)$ ergibt sich nicht mehr als Produkt von M 1-dim. Verteilungsfunktionen $P(m_j(s) \mid s = c_k)$, sondern **muss direkt geschätzt werden**.
 - Nach der verallgemeinerten Daumenregel sind bei $N = |\mathbf{T}|$ Stichproben $B = N^{1/(M+1)}$ Intervalle jeder der M Achsen des Merkmalsraumes zu wählen.
 - Eine zuverlässige Schätzung erfordert für ein hinreichendes Binning bereits für kleine M sehr hohe Stichprobenzahlen (z.B. $M = 4$, $N = |\mathbf{T}| = 100.000$, $B = 10$).*

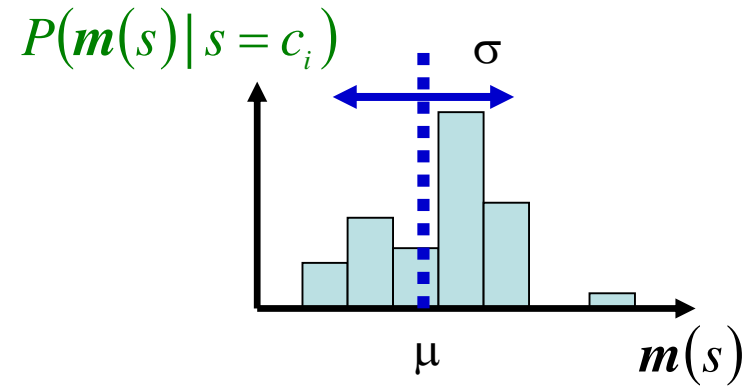
* Aus diesen Überlegungen ergibt sich Folgendes (Zitat aus [BB15, S. 751]): „Die Annahme der Unabhängigkeit trägt wesentlich dazu bei, dass statistische Problemstellungen einfach und lösbar bleiben, obwohl diese Annahme in der Praxis oft (und bekanntermaßen) verletzt ist. Beispielsweise sind die Intensitätswerte benachbarter Bildelemente kaum jemals voneinander unabhängig.“

Gleichwohl bleibt zu überprüfen, ob die Annahme im konkreten Fall wirklich angemessen ist – z.B. durch Trainings- und Testmenge bekannter Daten.

MAP: Schätzung der Merkmalsverteilungen (6)

- Hier kann es vorteilhaft sein, die Charakteristik der Verteilung zu kennen.

- Wird z. B. angenommen, dass die Merkmale unkorreliert und normalverteilt sind, müssen nur Erwartungswert sowie Varianz für jedes Merkmal geschätzt werden.



Bildquelle: Klaus Tönnies: Grundlagen der Bildverarbeitung, Pearson Studium, 2005.

- Als Beispiel für die Schätzung von unkorrelierten und normalverteilten klassen-abhängigen Merkmalsverteilungen $P(m(s) | s = c_i)$ (nach Tönnies) vergleichen wir nun tatsächlich Äpfel mit Birnen bzw. deren Grauwertbilder.

Beispiel für MAP-Klassifikation: Merkmale

- Beispiel für die Schätzung von unkorrelierten und normalverteilten klassen-abhängigen Merkmalsverteilungen (1):
 - Ausgangspunkt seien segmentierte Grauwertbilder von Äpfeln und Birnen.
 - Ausgewählte Merkmale seien:
 - 1) Die Kreisähnlichkeit $Kr(s)$ des Konturrandes: $Kr(s) = (F(s) \cdot 4\pi) / R^2(s)$ mit Fläche $F(s)$ und Randlänge $R(s)$ – vgl. Folien 15,16 und 18.

Annahme: die Kreisähnlichkeit ist für Äpfel größer als für Birnen.
 - 2) Der mittlere Intensitätswert des Segments:

Annahme: Birnen eher gelb bis grün, Äpfel eher rot bis grün, daher hellere Grauwerte bei Birnen.



Bildquelle: Colourbox (15.01.21)

Beispiel für MAP-Klassifikation: Merkmalverteilungen

- Für die MAP-Klassifikation gehen von der Annahme *von unkorrelierten und normalverteilten klassen-abhängigen Merkmalsverteilungen* aus:

- die Merkmalsverteilungen für die Klassen c_1 und c_2 ergeben sich aufgrund der unkorrelierten Merkmale als Produkte

$$P((m_1(s), m_2(s)) \mid s = c_k) = P(m_1(s) \mid s = c_k) \cdot P(m_2(s) \mid s = c_k), \quad k = 1, 2.$$

- Aus je 10 Stichproben werden Erwartungswert E und Varianz σ^2 geschätzt:

$$E_k(m_i) \approx \frac{1}{10} \sum_{j=0}^9 m_i(s_j), \quad \sigma_k^2(m_i) \approx \frac{1}{9} \sum_{j=0}^9 (m_i(s_j) - E(m_i))^2, \quad i = 1, 2, \quad k = 1, 2.$$

Zur Erinnerung:

- Aus den Normalverteilungen folgt (für $k = 1, 2$): $P(S = C_i \mid \mathbf{m}(S)) = \frac{P(\mathbf{m}(S) \mid S = C_i) P(S = C_i)}{\sum_{k=0}^{K-1} P(\mathbf{m}(S) \mid S = C_k) P(S = C_k)}$

$$P(c_k \mid (m_1(s), m_2(s))) \propto \frac{1}{\sigma_k(m_1) \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{m_1(s) - E_k(m_1)}{\sigma_k(m_1)} \right)^2} \cdot \frac{1}{\sigma_k(m_2) \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{m_2(s) - E_k(m_2)}{\sigma_k(m_2)} \right)^2}.$$

Beispiel für MAP-Klassifikation: Trainingsmenge

- Beispiel für die Schätzung von unkorrelierten und normalverteilten klassen-abhängigen Merkmalsverteilungen (3):

- Die 10 Stichproben m_j ($0 \leq j \leq 9$) für Klasse c_1 (Äpfel)
mit Merkmalen $m_{j1} = Kr(s)$ und $m_{j2} = i_{average}(s)$:

$$\vec{m}_j = \begin{pmatrix} m_{j1} \\ m_{j2} \end{pmatrix} = \begin{pmatrix} 0,95 \\ 0,62 \end{pmatrix}, \begin{pmatrix} 0,92 \\ 0,64 \end{pmatrix}, \begin{pmatrix} 0,84 \\ 0,64 \end{pmatrix}, \begin{pmatrix} 0,99 \\ 0,72 \end{pmatrix}, \begin{pmatrix} 0,91 \\ 0,70 \end{pmatrix}, \begin{pmatrix} 0,92 \\ 0,65 \end{pmatrix}, \begin{pmatrix} 0,94 \\ 0,59 \end{pmatrix}, \begin{pmatrix} 0,98 \\ 0,67 \end{pmatrix}, \begin{pmatrix} 0,99 \\ 0,68 \end{pmatrix}, \begin{pmatrix} 0,83 \\ 0,71 \end{pmatrix}$$

ergeben für c_1 : $E_1(m_1) = 0,93$, $\sigma_1^2(m_1) = 0,0032$, $E_1(m_2) = 0,66$, $\sigma_1^2(m_2) = 0,0017$.

- Die 10 Stichproben für c_2 (Birnen):

$$\vec{m}_j = \begin{pmatrix} m_{j1} \\ m_{j2} \end{pmatrix} = \begin{pmatrix} 0,72 \\ 0,80 \end{pmatrix}, \begin{pmatrix} 0,80 \\ 0,83 \end{pmatrix}, \begin{pmatrix} 0,77 \\ 0,72 \end{pmatrix}, \begin{pmatrix} 0,60 \\ 0,91 \end{pmatrix}, \begin{pmatrix} 0,67 \\ 0,83 \end{pmatrix}, \begin{pmatrix} 0,72 \\ 0,85 \end{pmatrix}, \begin{pmatrix} 0,84 \\ 0,74 \end{pmatrix}, \begin{pmatrix} 0,85 \\ 0,77 \end{pmatrix}, \begin{pmatrix} 0,78 \\ 0,79 \end{pmatrix}, \begin{pmatrix} 0,77 \\ 0,88 \end{pmatrix}$$

ergeben für c_2 : $E_2(m_1) = 0,75$, $\sigma_2^2(m_1) = 0,0058$, $E_2(m_2) = 0,81$, $\sigma_2^2(m_2) = 0,0036$.

Beispiel für MAP-Klassifikation: A priori W'keiten

- Beispiel für die Schätzung von unkorrelierten und normalverteilten klassen-abhängigen Merkmalsverteilungen (4):

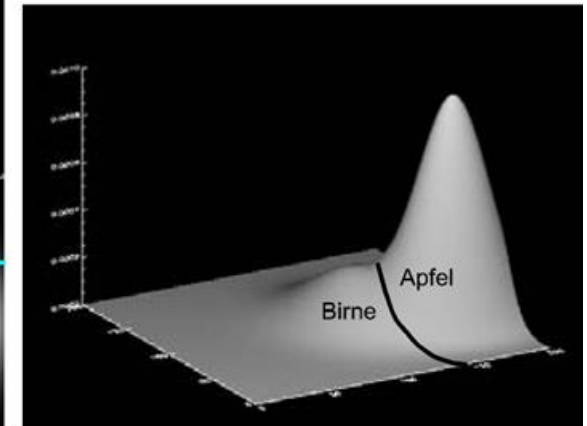
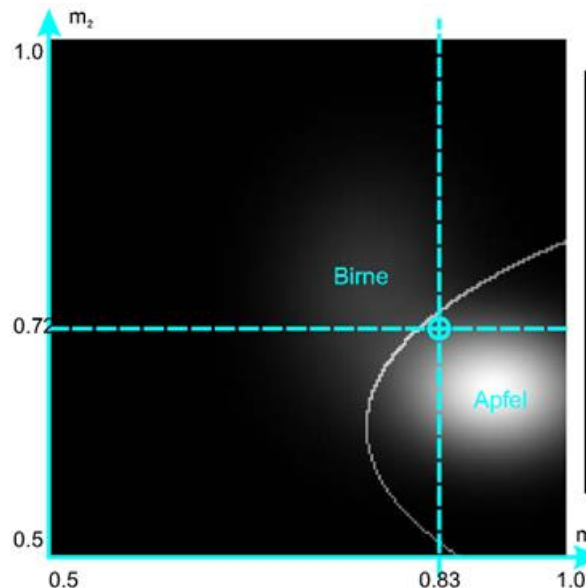
- Die A-Priori-W'keiten werden geschätzt als $P(c_1) = 0,67$ und $P(c_2) = 0,33$.

Zur Erinnerung:

$$P(S = C_i | \mathbf{m}(S)) = \frac{P(\mathbf{m}(S) | S = C_i) P(S = C_i)}{\sum_{k=0}^{K-1} P(\mathbf{m}(S) | S = C_k) P(S = C_k)}$$

- Die beiden Darstellungen zeigen die aus den Stichproben geschätzten, mit den A-Priori-W'keiten gewichteten Merkmalswahrscheinlichkeiten.

Die weiße Linie links bzw. die schwarze Linie rechts gibt die Entscheidungsgrenze an.



Beispiel für MAP-Klassifikation: Inferenz

- Beispiel für die Schätzung von unkorrelierten und normalverteilten klassen-abhängigen Merkmalsverteilungen (5):

- Im Fadenkreuz links wird ein zu klassifizierendes Segment s_{test} mit Merkmalswerten (0,83;0,72) angezeigt.

Zur Erinnerung:

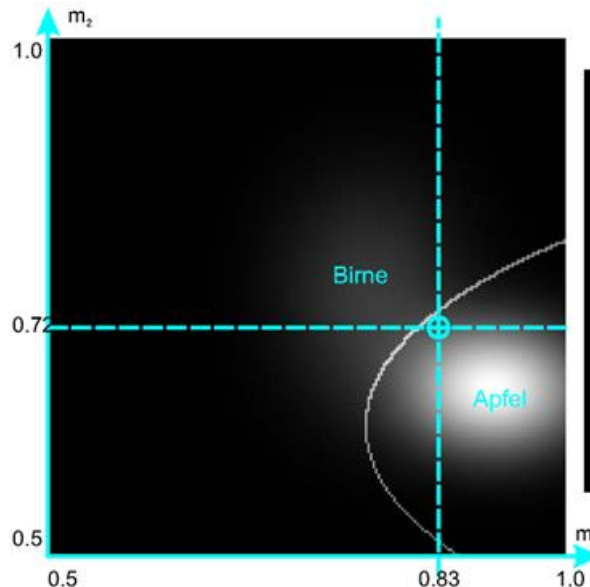
$$P(S = C_i | \mathbf{m}(S)) = \frac{P(\mathbf{m}(S) | S = C_i)P(S = C_i)}{\sum_{k=0}^{K-1} P(\mathbf{m}(S) | S = C_k)P(S = C_k)}$$

- Die Berechnungen der A-Posteriori-W'keiten liefern:

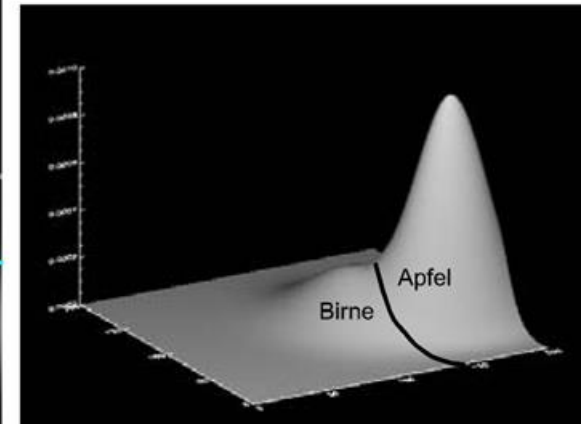
$$P(c_1 | (0,83, 0,72)) = 0,65,$$

$$P(c_2 | (0,83, 0,72)) = 0,35.$$

Also wird Segment s_{test} der Klasse c_1 der Äpfel zugeordnet.



Die Abbildungen zeigen die Position des Merkmalsvektors $\mathbf{f}(S)$ im Raum der mit den A-Priori-W'keiten gewichteten Merkmalswahrscheinlichkeiten.



Beispiel aus: Klaus Tönnies: Grundlagen der Bildverarbeitung, Pearson Studium, 2005. Wegen der kontinuierl. Normalverteilung ist die W'keit über ein Intervall zu rechnen. Vereinfachend wird eine Intervallbreite von 0,01 mit im Intervall konstanten Werten (0,83;0,72). Die Intervallgrößen in Zähler und Nenner kürzen sich gegenseitig.

Weitere Klassifikationsverfahren

- Die MAP-Klassifikation wurde hier stellvertretend behandelt.
- Wichtig sollte bei dieser Behandlung sein, wie die Schnittstelle zwischen Segmentmerkmalen einerseits und Klassifikationsverfahren andererseits exemplarisch ausbuchstabiert wird.
- Generell sind alle anderen überwachten Klassifikationsansätze sowie Methoden des unüberwachten Clusterings für die bildbasierte Objektklassifikation bzw. Objekterkennung anwendbar.

Zusammenfassung (1)

- Objekterkennung heißt häufig: **Klassifikation** der im Bild abgebildeten Objekt in zuvor spezifizierte Objektklassen.
- Ausgangspunkt der Klassifikation sind segmentierte Bilder. Den Segmenten haben Klassifikationsverfahren als Verfahren der High-Level-Vision **Bedeutung** zuzuordnen.
- Die Modellierung der Objektklassen kann unterschiedlich komplex sein.
 - Die Zuordnung der Segmente kann alleine aufgrund ihrer eigenen Segmentmerkmale erfolgen.
 - Die Zuordnung kann zudem Relationen zwischen Segmenten und damit Segmentgruppen berücksichtigen → Subgraphen-Matching.

Zusammenfassung (2)

- **Segmentmerkmale** werden aus dem Segmentinneren und dem Segmentrand abgeleitet.
- Die **Merkmale sind gut gewählt**, wenn die Segmente einer Klasse in eine eng begrenzte Region des Merkmalsraums abgebildet werden und Regionen von Segmenten aus verschiedenen Klassen gut voneinander getrennt sind: niedriger **Within-Class-Scatter** und hoher **Between-Class-Scatter**.
- **Regionenmerkmale**: werden auch Grauwert oder Farbe eines Segments als Textur bezeichnet, dann sind alle regionenbasierte Merkmale Texturen. Sie werden durch hinreichend viele Stichproben aus dem Segmentinneren geschätzt.
- **Formmerkmale** werden aus dem Segmentrand abgeleitet. Einfache Formmerkmale erhalten ihre **Robustheit** gegenüber Segmentierungsfehlern durch **Integration des Merkmals** über das gesamte Objekt bzw. Segment.

Zusammenfassung (3)

- Bei MAP-Klassifikationsverfahren können die A-Posteriori-W'keiten über den Satz von Bayes aus den klassenspezifischen Merkmalsverteilungen und A-Priori-W'keiten erfolgen.
- Die klassenspezifischen Merkmalsverteilungen und A-Priori-W'keiten können aus Trainingsmengen bestimmt werden.
 - Dies ist bzgl. der Merkmalsverteilungen schwierig, wenn die Merkmale korreliert sind, da dann die i.A. hochdimensionale Verteilungsfunktion $P(\mathbf{m}(s) \mid s = c_k)$ direkt aus den Stichproben geschätzt werden muss, aber die Trainingsmenge dafür häufig nicht groß genug ist.
 - Einfacher ist es, wenn die Merkmale unkorreliert sind, da die M -dim. Verteilungsfunktion $P(\mathbf{m}(s) \mid s = c_k)$ dann als Produkt von M eindim. Verteilungsfunktionen $P(m_j(s) \mid s = c_k)$ ableitbar ist.
 - Ist zudem die Charakteristik der Verteilungsfunktion bekannt, müssen lediglich deren Parameter geschätzt werden.

Literatur zu ML und MAP

- [Tö05] K. Tönnies. Grundlagen der Bildverarbeitung. Pearson, 2005.
- [BB15] W. Burger, M. J. Burge. Digitale Bildverarbeitung (3. Aufl.). Springer Vieweg, 2015.
- [Wi23] Wikipedia, <https://de.wikipedia.org/wiki/Maximum-a-posteriori-Schätzung>, (Abr.: 18.12.2023)