

1 1 **Klassifikation**

One of the most interesting aspects of the world is that it can be considered to be made up from patterns.

A pattern is essentially an arrangement. It is characterized by the order of its elements for which it is made rather than by the intrinsic nature of these elements

Norbert Wiener

Einführung

Segmentierung

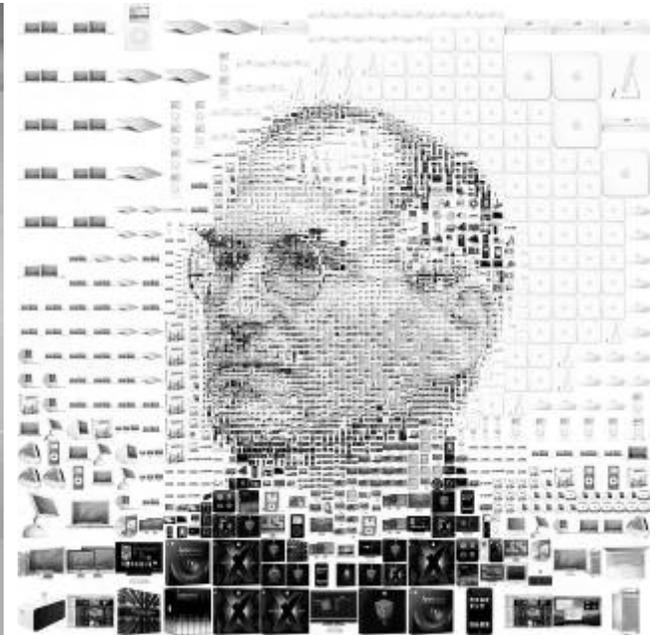
- Erste Schritt der erkennenden Bildverarbeitung
 - Zusammenfassen von Punkten zu größeren Einheiten
 - Unterteilung des Bildes in Bereiche
- ⇒ Bildpunkt orientiert

Klassifikation

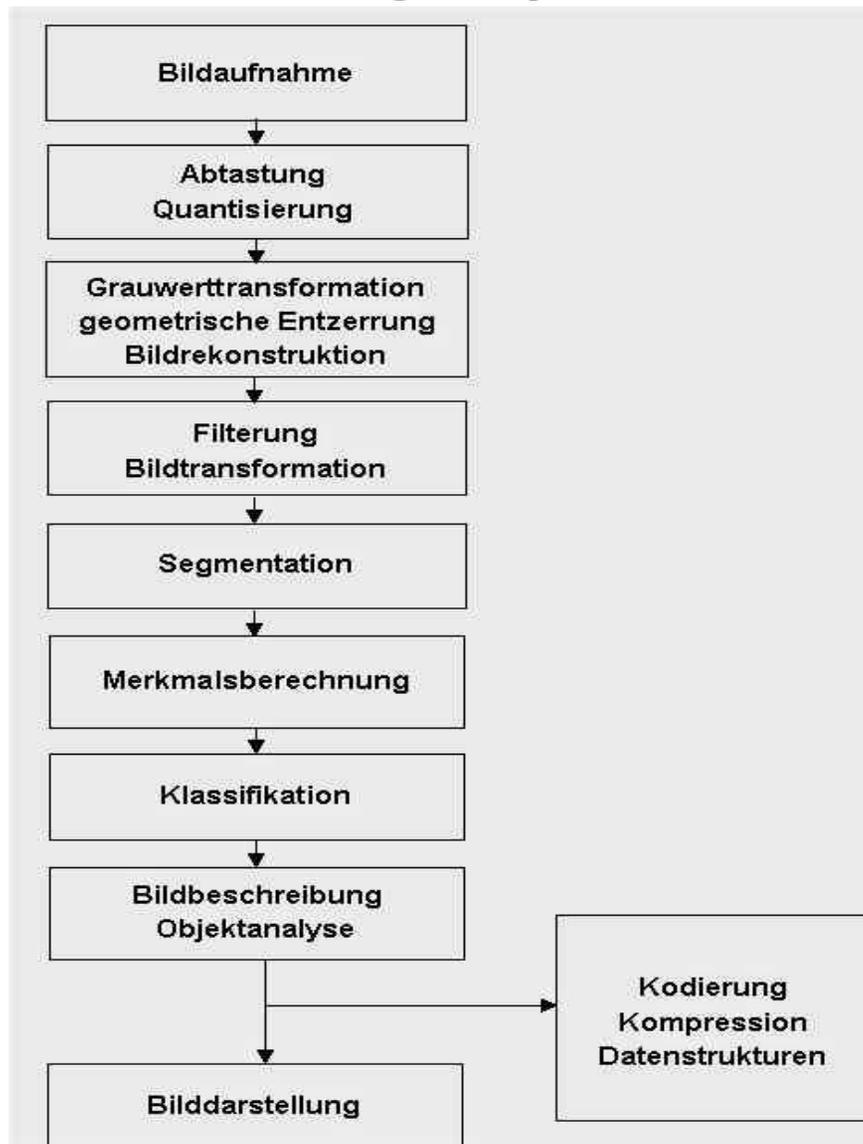
- Zuweisung von Bereichen in „Objektklassen“
- ⇒ Regionen orientiert

Muster

- Muster = Anordnungen von Elementen, welche eine “typische Ordnung” zeigen. Ordnung ist dabei wichtiger als die Elemente per-se



Bildverarbeitungssystem



Schritte bei der Klassifikation

Schritt 1: **Segmentierung**

=> *Definition des Objektbereichs*

Schritt 2: **Merkmalsextraktion.**

=> *Eigenschaften des Objekts*

Schritt 3: **Klassifikation.**

=> *Übereinstimmung mit
vorbestimmten Objektmerkmalen*

Segmentierung

Gegeben Bild B mit Bildpunkten f_1, \dots, f_n

Gesucht Regionen R_i für die gilt:

- $\bigcup_{i=1, \dots, s} R_i = B$
- $R_i \cap R_j = \emptyset$
- $f_s \in R_i, f_t \in R_j : f_s \sim f_t \Leftrightarrow i = j$

Merkmalsextraktion

„Merkmale“ beschreiben Eigenschaften von Regionen/Objekten, nicht mehr von einzelnen Pixeln

Einfache Objektmerkmale

- Umschreibendes Rechteck
- Mittlerer Intensität
- Flächeninhalt
- Schwerpunkt
- Umfang

Abgeleitete Merkmale

- Kompaktheit, Konvexheit
- Lage, Orientierung

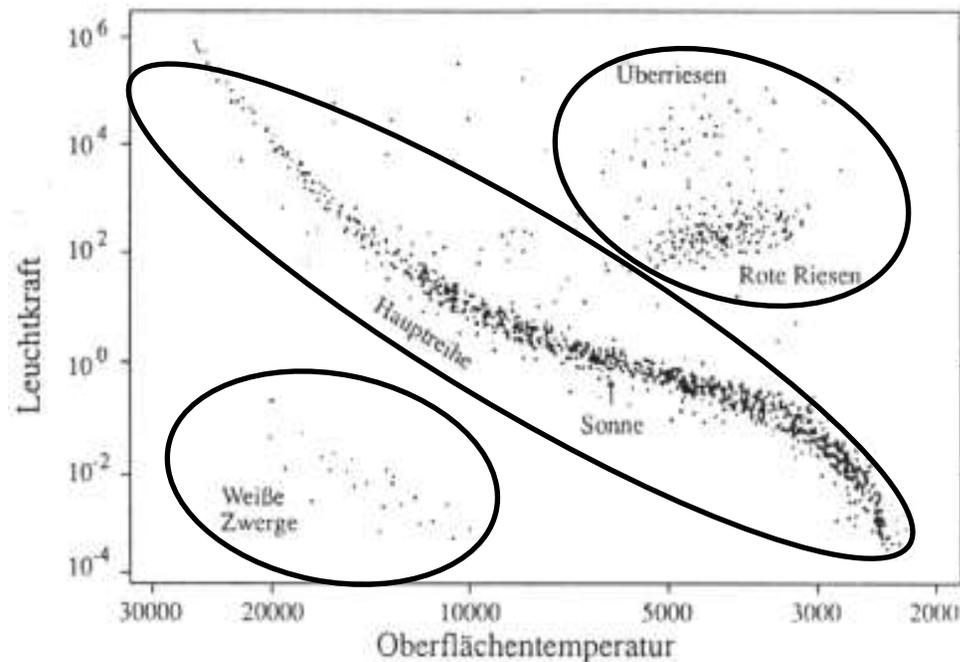
Klassifizierung

Klassifizierung von Objekten anhand ihrer Merkmale

Notwendige Voraussetzung

Merkmalsanalyse / Clustering des Merkmalraums

Beispiel



Klassifikation

Modellierung einer kognitiven Leistung mit dem Ziel einer Interpretation des Bildinhaltes

- Segmentierung ist ein „dummer“ Prozess, der lediglich ähnliche **Pixel** anhand der Signalinformation **zu Regionen** zusammenfasst
- Klassifikation ist ein „intelligenter“ Prozess, der **Regionen** **zu Objekt(klassen)** zuordnet und somit den Inhalt des Bildes erkennen will

=> Grenze fließend, Unterscheid eher quantitativ als qualitativ

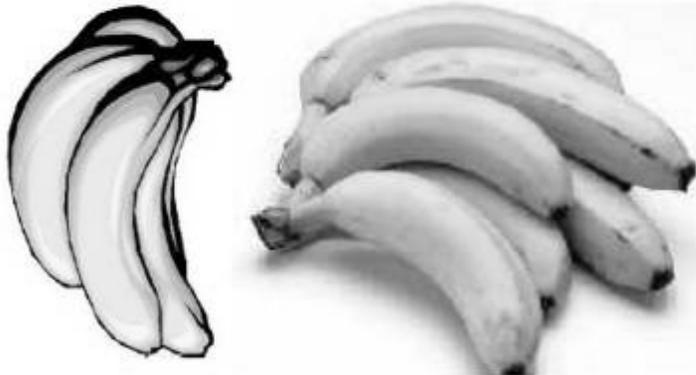
Grundlegende Ansätze

1. **Syntaktisch:** Dinge werden so durch Folgen von Symbolen beschrieben, dass Objekte der gleichen Kategorie die selben Beschreibungen aufweisen. Das Problem der Mustererkennung stellt sich in diesem Fall als Suche nach einer formalen Grammatik dar, also nach einer Menge von Symbolen und Regeln zum Zusammenfügen derselben.
2. **Statistisch:** Ziel ist es hier, ein Objekt in die Kategorie mit der höchsten Wahrscheinlichkeit einzusortieren. Statt Merkmale nach vorgefertigten Regeln auszuwerten, werden sie hier einfach als Zahlenwerte gemessen und in einem Merkmalsvektor zusammengefasst. Eine mathematische Funktion ordnet dann jedem denkbaren Merkmalsvektor eindeutig eine Kategorie zu.
3. **Strukturell:** verbindet verschiedene syntaktische und/oder statistische Verfahren zu einem einzigen neuen Verfahren. Die grundlegende Merkmalserkennung wird dabei allgemeinen statistischen Verfahren überlassen, während übergeordnete Inferenzverfahren Spezialwissen über das Sachgebiet einbringen.

Syntaktisch



rot, rund



gelb, länglich

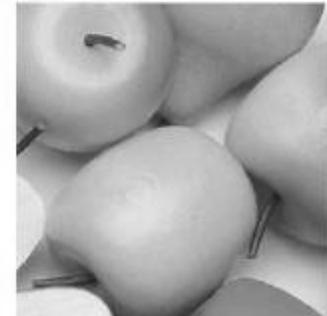
Syntaktischer Mustererkenner:

Regel 1: „Wenn gelb und länglich,
dann Banane“

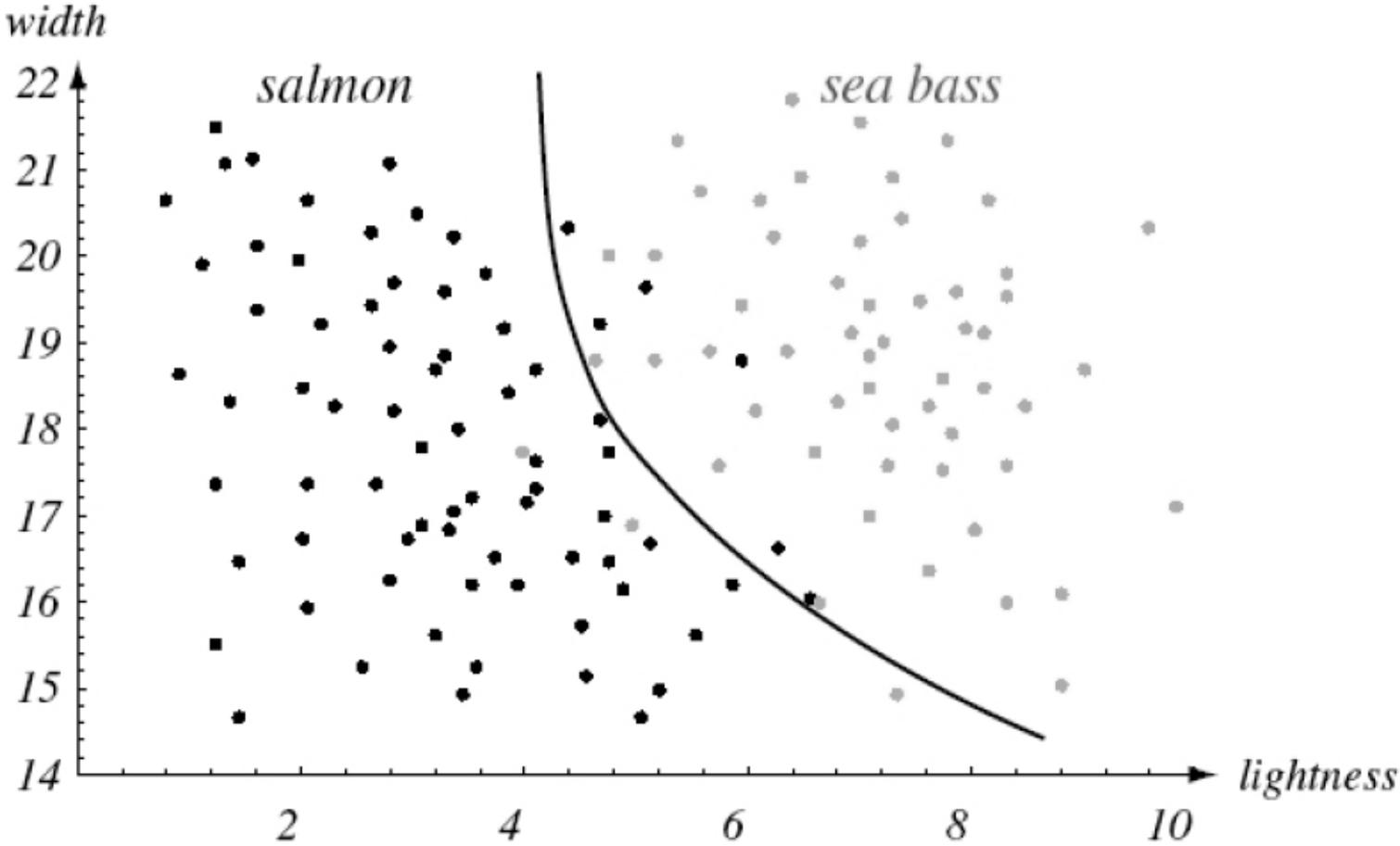
Regel 2: „Wenn rot und rund,
dann Apfel“



Erfordert
eindeutige
Attribute

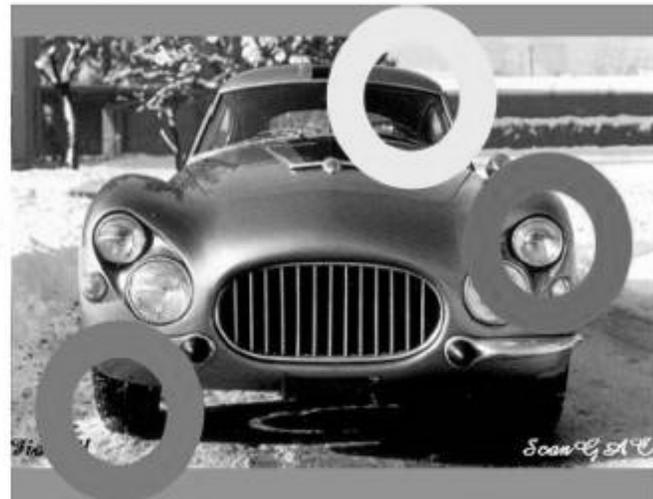
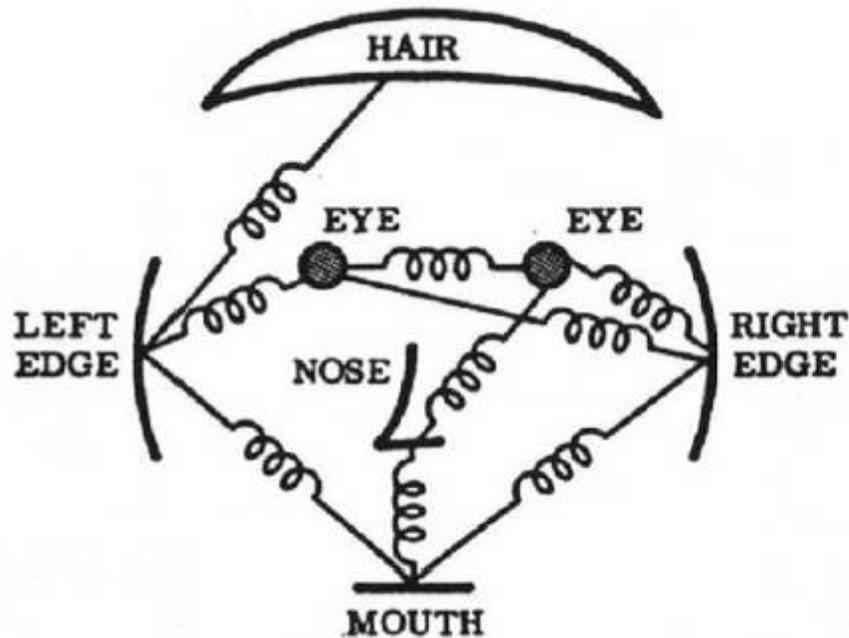


Statistisch



[aus Duda et al., 2001]

Strukturell, Wissens-/Modellbasiert



- Teile des Objekts werden mit statistischen Techniken detektiert.
- Struktur, d.h. Konstellation der Teile wird zusätzlich gelernt.

Relation einfacher Merkmale



Schritte

1. Auswahl Merkmale
 - Kleinste Anzahl, voneinander unabhängig
 - Klassenzuordnung ermöglichen
 - Ggf. normiert/normierbar
2. Definition von Klassen
 - Objektive Kriterien, sinnvolle Auswahl von Klassen
 - Gut trennbar, keine fließende Grenzen
3. Bestimmung von Klassengrenzen
4. Zuordnungsmethode eines unbekanntes Objektes zu einer der Klassen

Postulate

1. Eine repräsentative Stichprobe liegt vor zur Sammlung von Information
2. Ein (einfaches) Muster besitzt charakteristische Merkmale für seine Klassenzugehörigkeit
3. Merkmale bilden kompakte Clusters, verschiedene Klassen sind „an sich“ getrennt
4. Komplexes Muster läßt sich in einfachere Bestandteile zerlegen
5. Komplexe Muster haben bestimmte Struktur (= keine beliebige Anordnung einfacherer Bestandteile)
6. Zwei (komplexe) Muster sind ähnlich, wenn deren einfache Bestandteile sich wenig unterscheiden

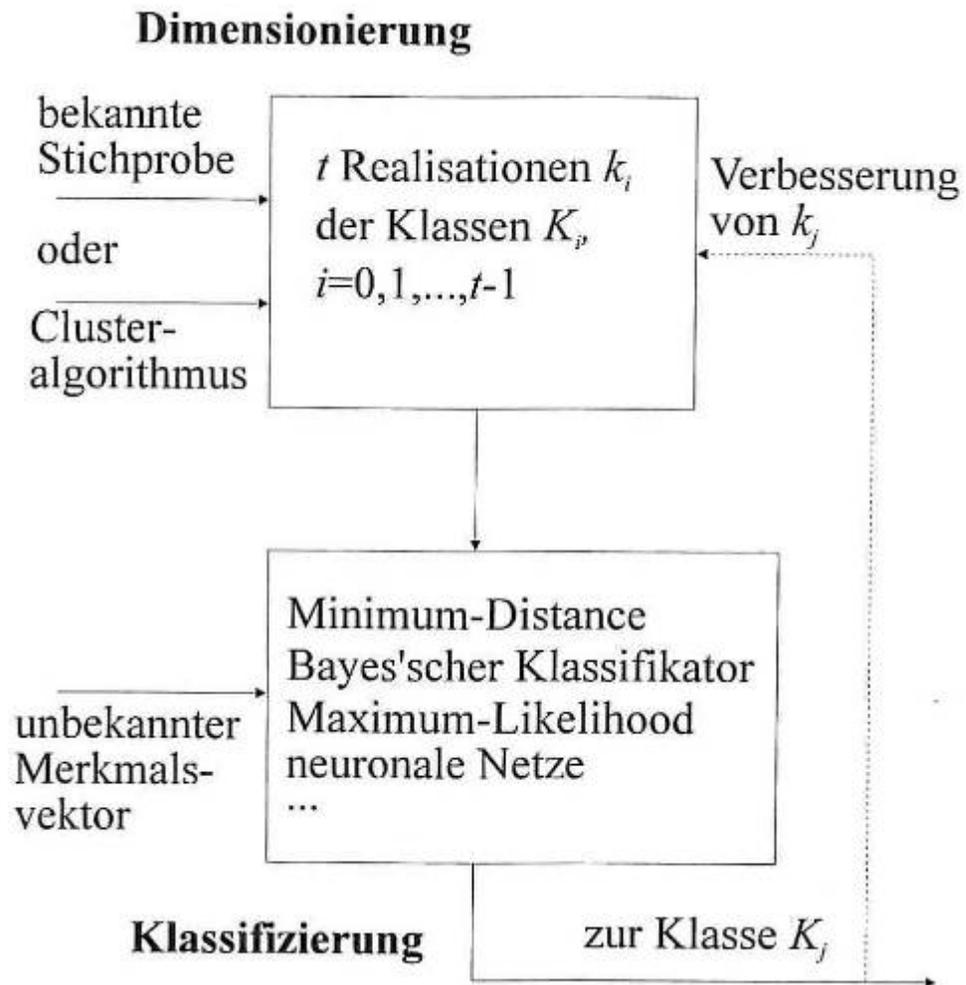
Klassifizierung

1. „Mehrkanaliges“ Bild (Merkmalsvektor) $S = (g_1, \dots, g_n)$
2. Definition von Klassen
Realisation/Stichprobe = Bei n-Beobachtungen „Häufungen“ im Merkmalsraum
3. Bestimmung von Klassengrenzen
Dimensionierung = Segmentierung/Unterteilung des Merkmalsraumes in Klassen
4. Zuordnungsmethode
Klassifizierung = Zuordnung eines Vektors zu einer Klasse

Probleme

1. Anzahl & Lage/Grenzen der Klassen im n-D Merkmalsraum unbekannt
2. Zuordnungsfunktion von Merkmalen zu Klassen

Klassifizierungsvorgang



Schritte

1. Auswahl Merkmale

- Kleinste Anzahl, voneinander unabhängig
- Klassenzuordnung ermöglichen
- Ggf. normiert/normierbar

2. Definition von Klassen

- Objektive Kriterien, sinnvolle Auswahl von Klassen
- Gut trennbar, keine fließende Grenzen

3. Bestimmung von Klassengrenzen

4. Zuordnungsmethode eines unbekanntes Objektes zu einer der Klassen

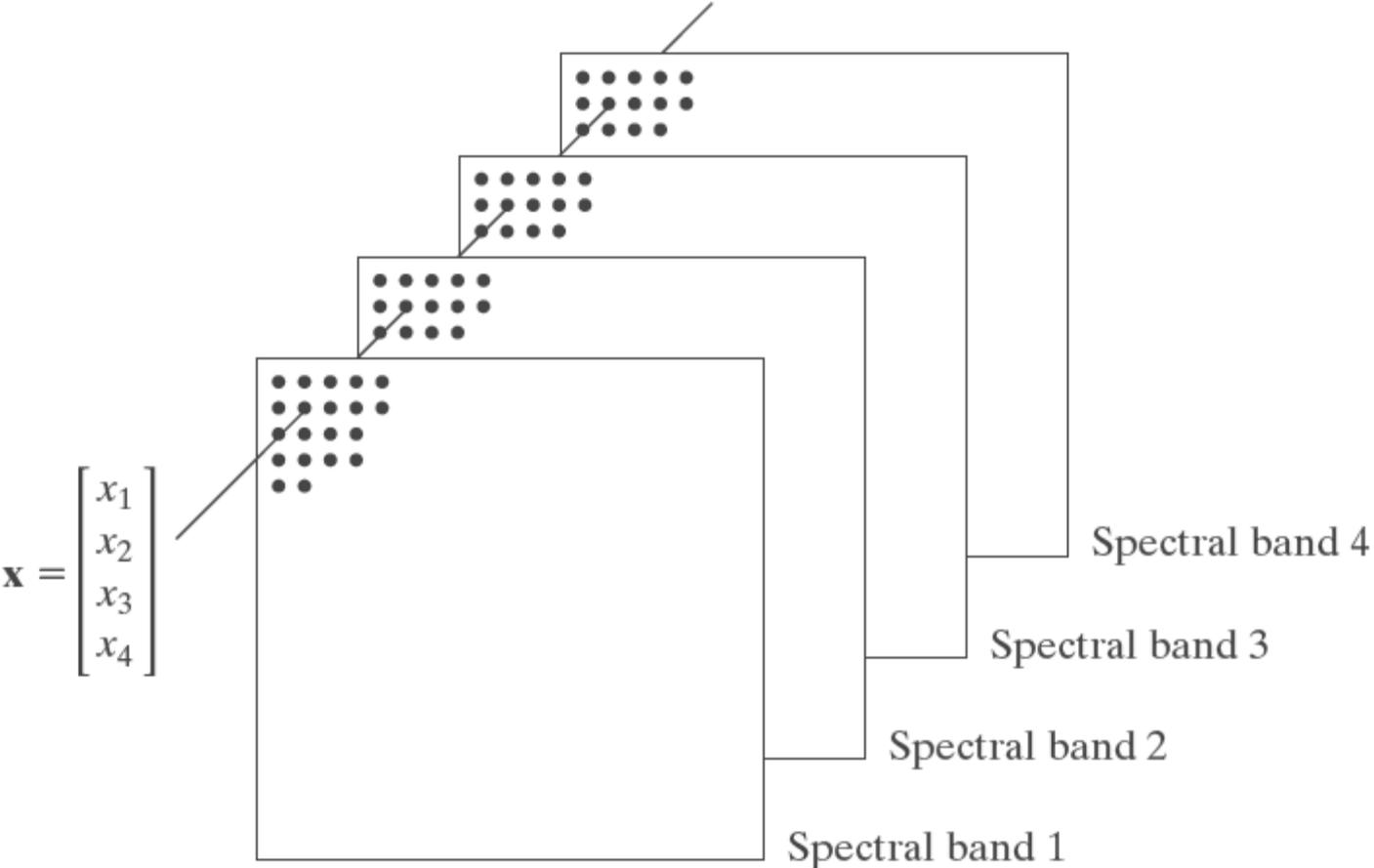
Auswahl der Merkmale

- Symbolische: Mitglieder eines endlichen Alphabets, „Symbolketten“ => syntaktische Klassifikatoren
- Numerische: In (reelle oder komplexe) Zahlen zu erfassende „Vektoren“ => numerische Klassifikatoren

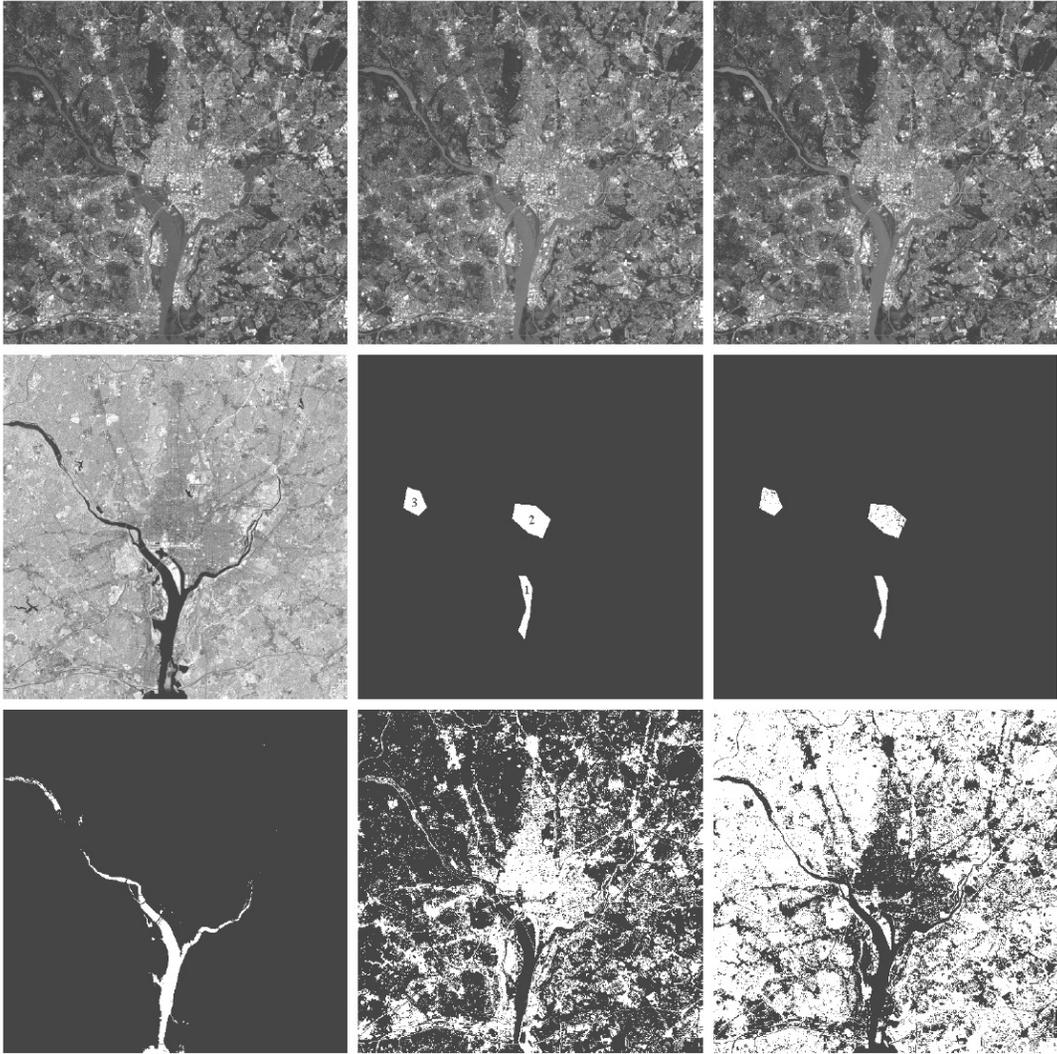
Zur Gewinnung von Merkmalen:

- Heuristisch: Intuition, Erfahrung, Beobachtung. Oft mangelhaft, insignifikant, wenig aussagefähig
- Analytisch: „optimale“ Merkmale bezüglich eines Kriteriums

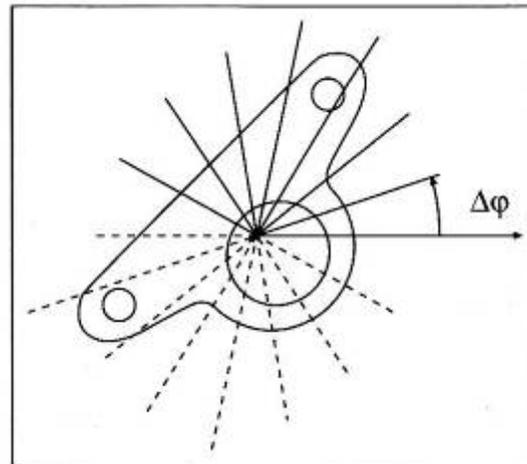
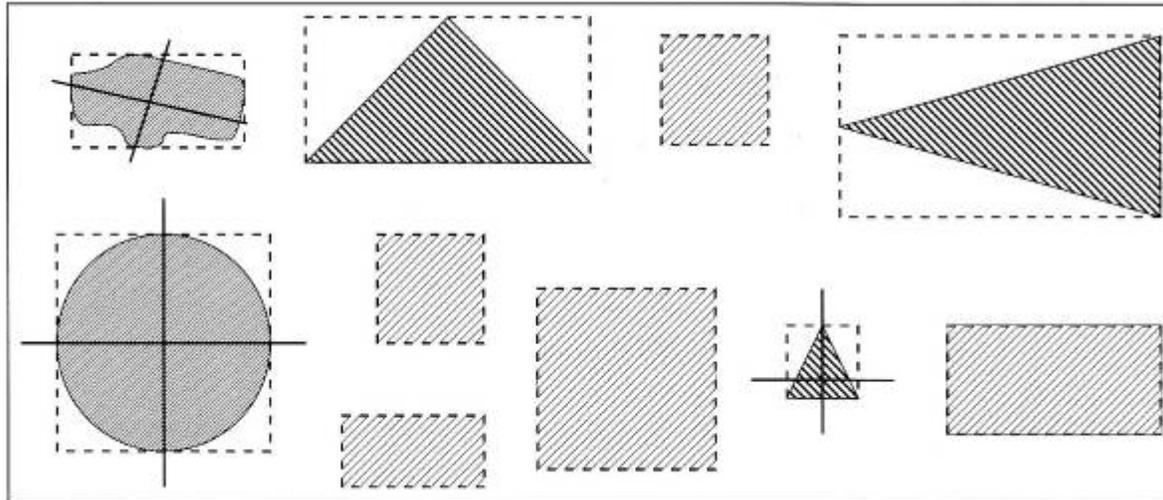
Mehrkanaliges Bild



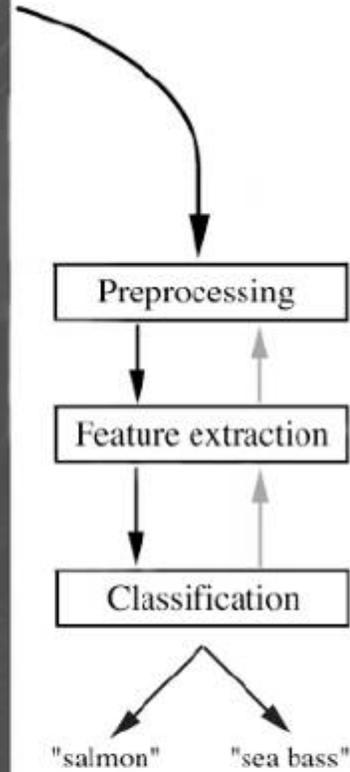
Mehrkanaliges Bild



A: Heuristische Methode

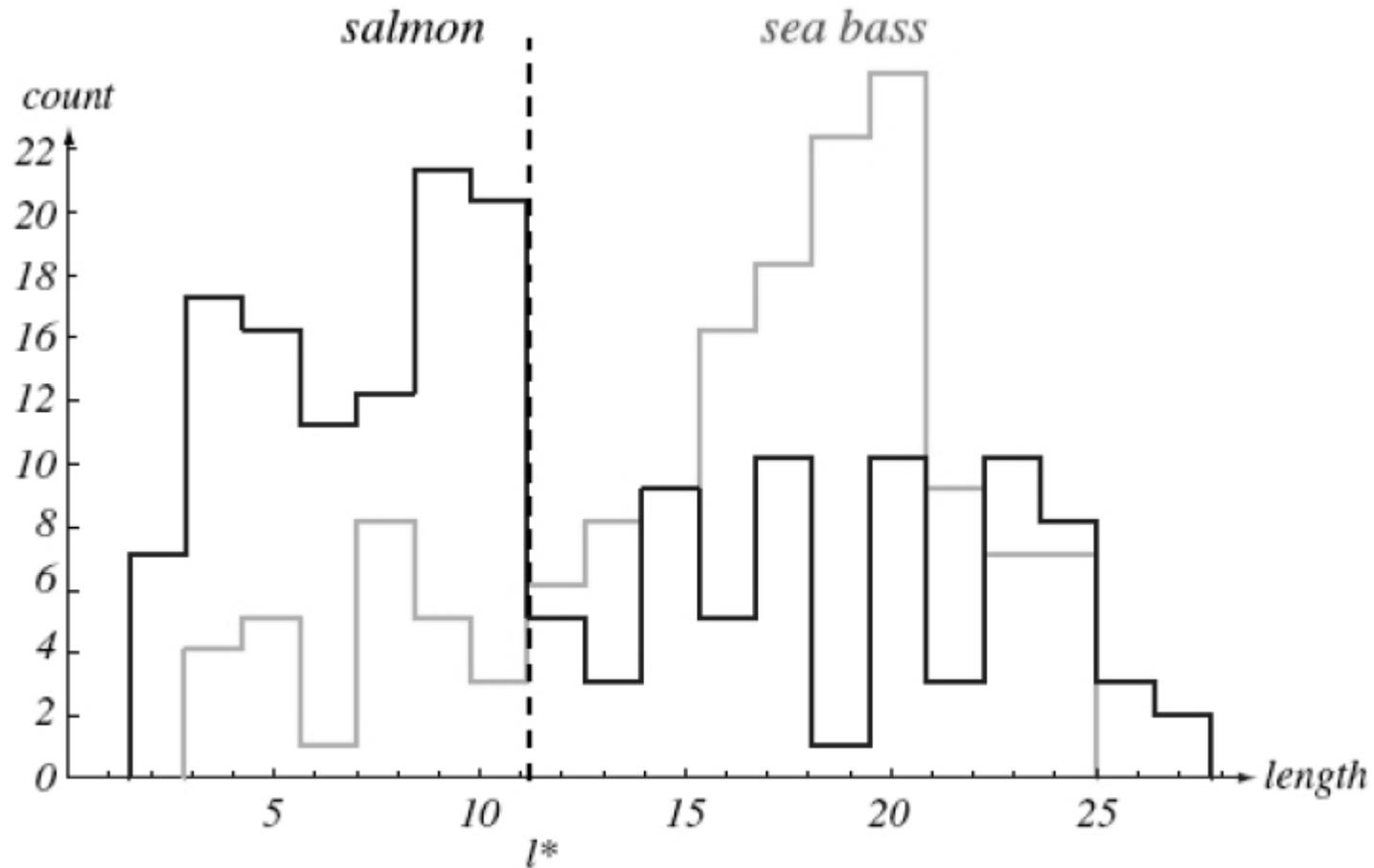


Beispiel



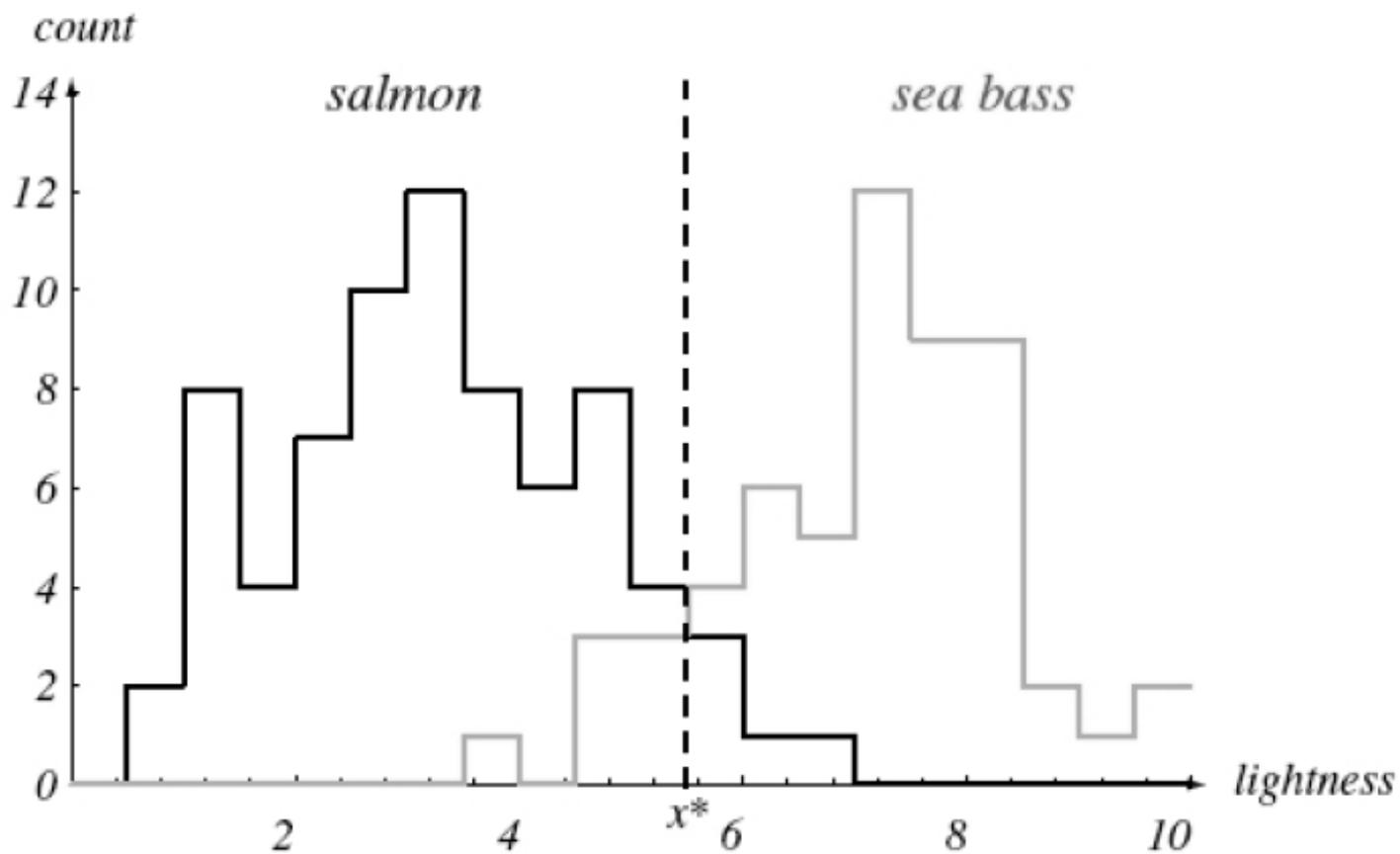
[aus Duda et al., 2001]

Beispiel: Ein Merkmal (Längenhistogramm)



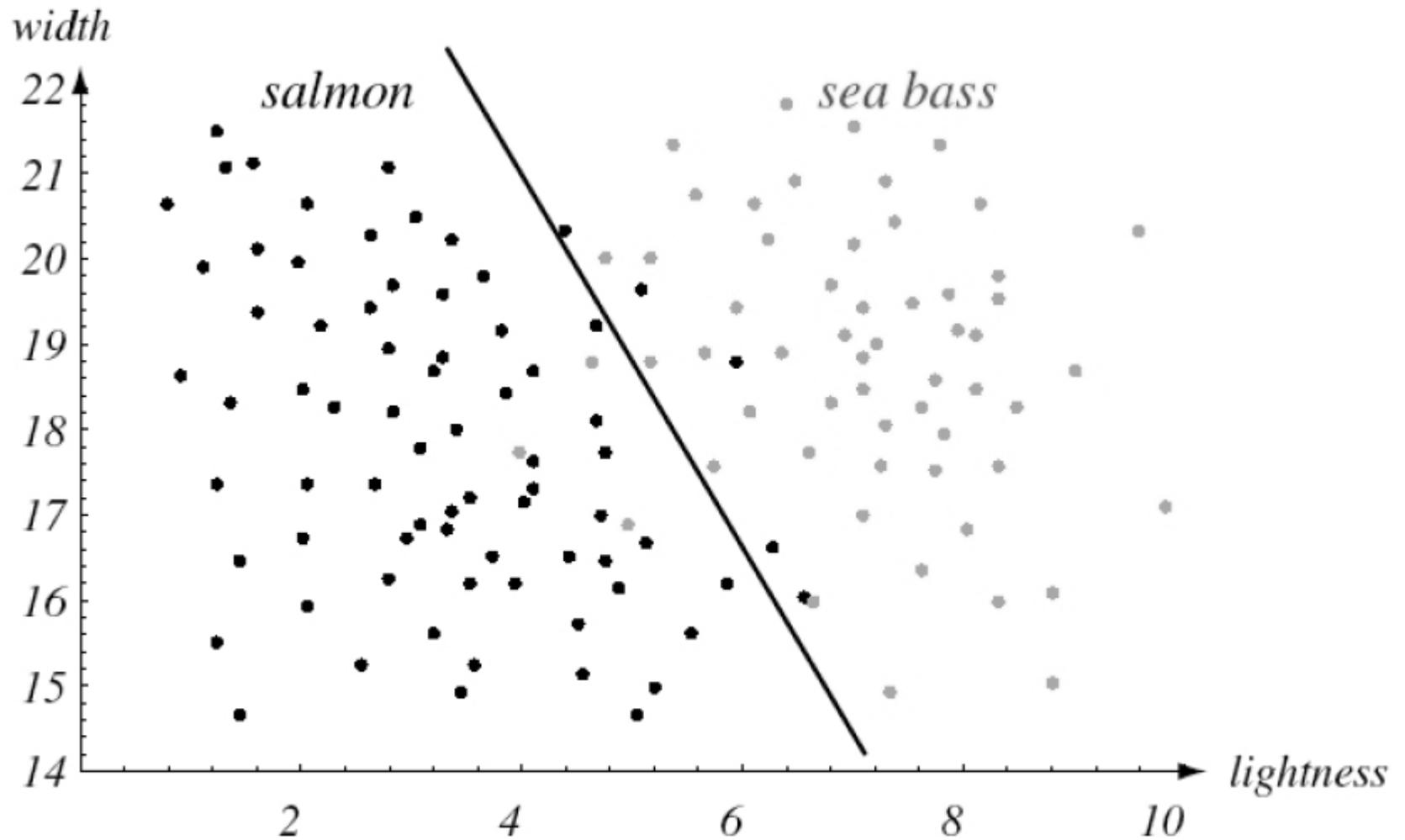
[aus Duda et al., 2001]

Beispiel: Ein Merkmal (Helligkeitshistogramm)



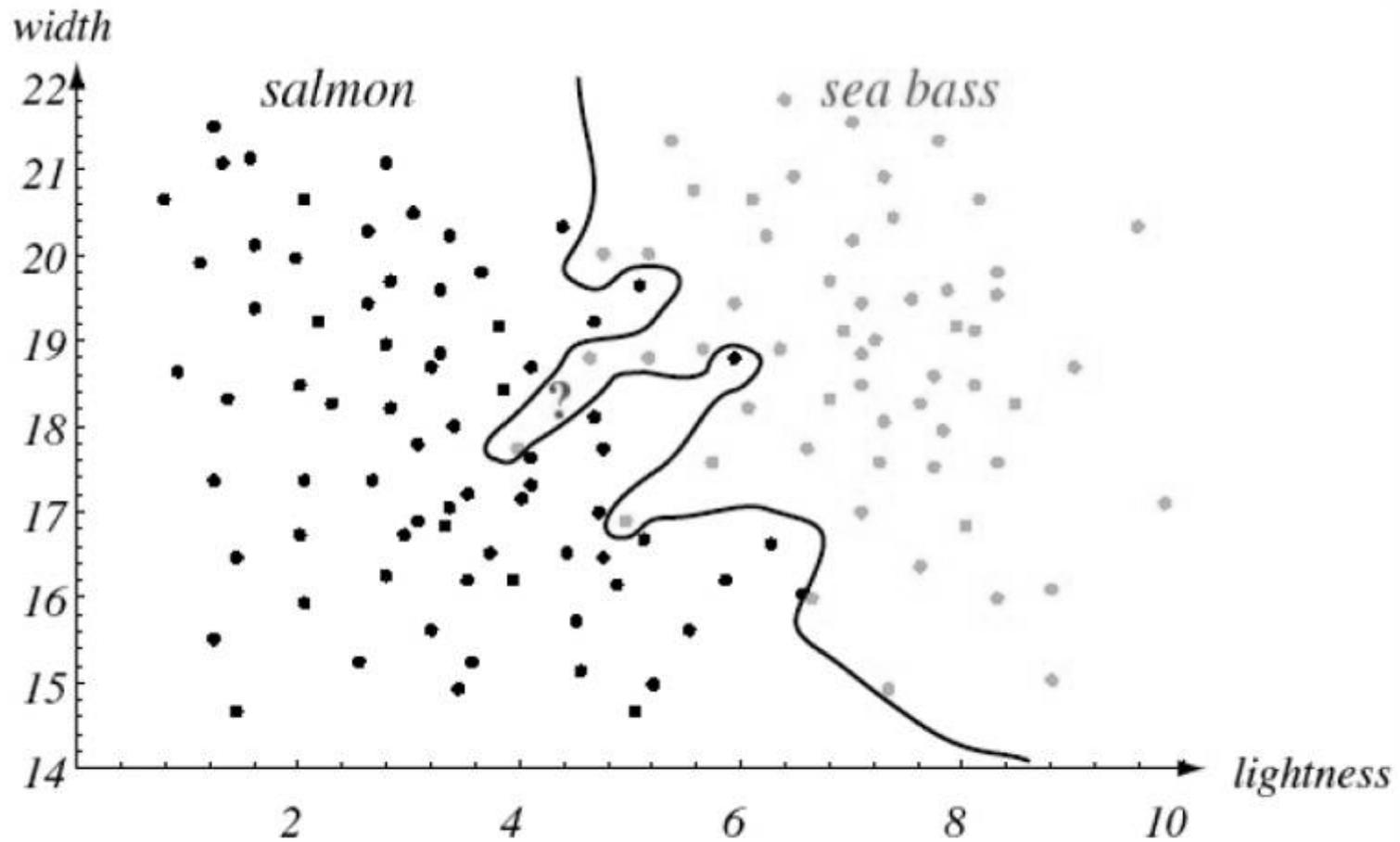
[aus Duda et al., 2001]

Beispiel: BEIDE Merkmale



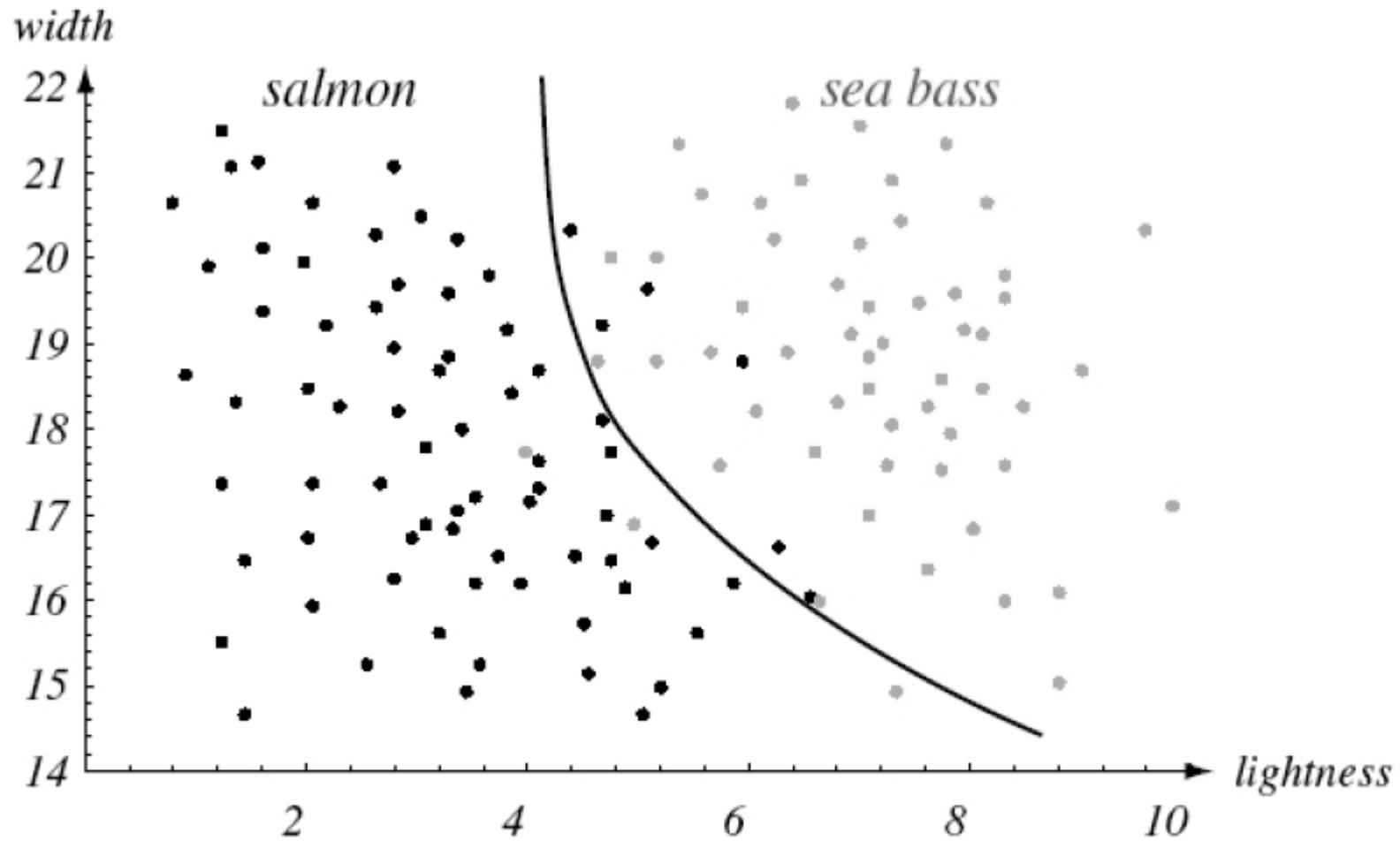
[aus Duda et al., 2001]

Beispiel: Klassentrennung?



[aus Duda et al., 2001]

Beispiel: Klasentrennung?



[aus Duda et al., 2001]

B: Analytische Methode

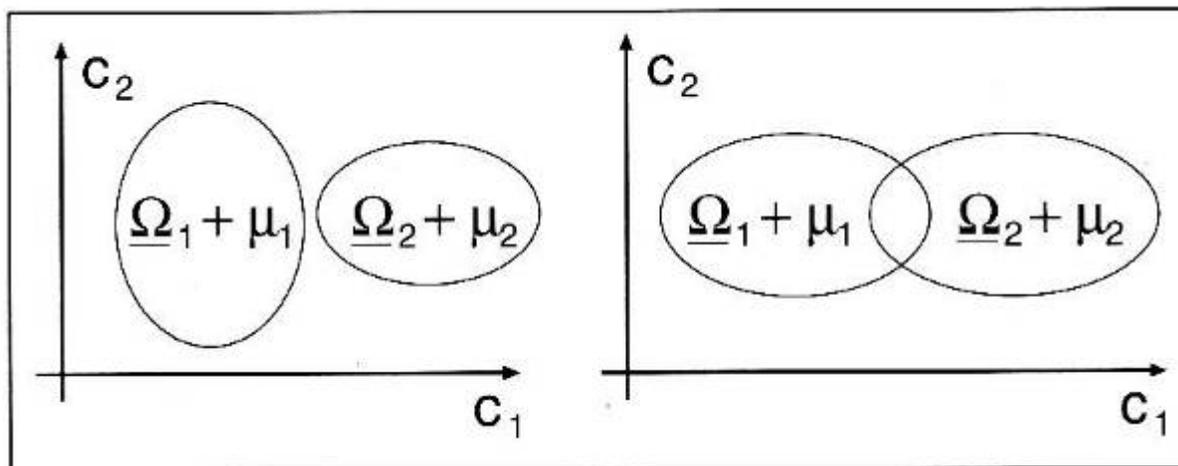
- Auswahl von M Merkmalen ($m = 1 \dots M$), welche ein Kriterium am besten optimieren

Zwei grundsätzliche Arten von Kriterien:

1. Gruppierung von Merkmalen (clustering), Trennbarkeit von Klassen
2. Klassifikationsfehlwahrscheinlichkeit

Auswahl von Merkmalen

- Merkmalsmenge so klein wie möglich
- Abstand der Clustermittelpunkte kein ausreichendes Kriterium



Auswahl von Merkmalen

- Theoretisch sinnvoll: Fehlwahrscheinlichkeit minimieren, Zuordnung zu der Klasse mit höchster a-posteriori Wahrscheinlichkeit
- Optimaler Bayes-Abstand

- Problem: bedingte Dichten $p(c|W)$ i.d.R. unbekannt
- Näherung über Stichprobe

Strategien zur Merkmalsauswahl

- Jedes Merkmal allein bewerten, N beste auswählen
- Erst das best bewertete Merkmal, dann Paare, dann Tripel etc.
- Zuerst zwei schlechteste zu trennende Klassen ermitteln, Merkmale auswählen, die zur Trennung bestens beitragen
- *Leaving-one-out*: erst Klassifikation mit allen K Merkmalen, dann mit K-1 bis beste Kombination, dann K-2 etc. bis optimale Kombination erreicht

=> Auswertung heuristischer Merkmale aufwendig, daher besser Bewertung durch Lagebeurteilung der Vektoren im Merkmalsraum

Schritte

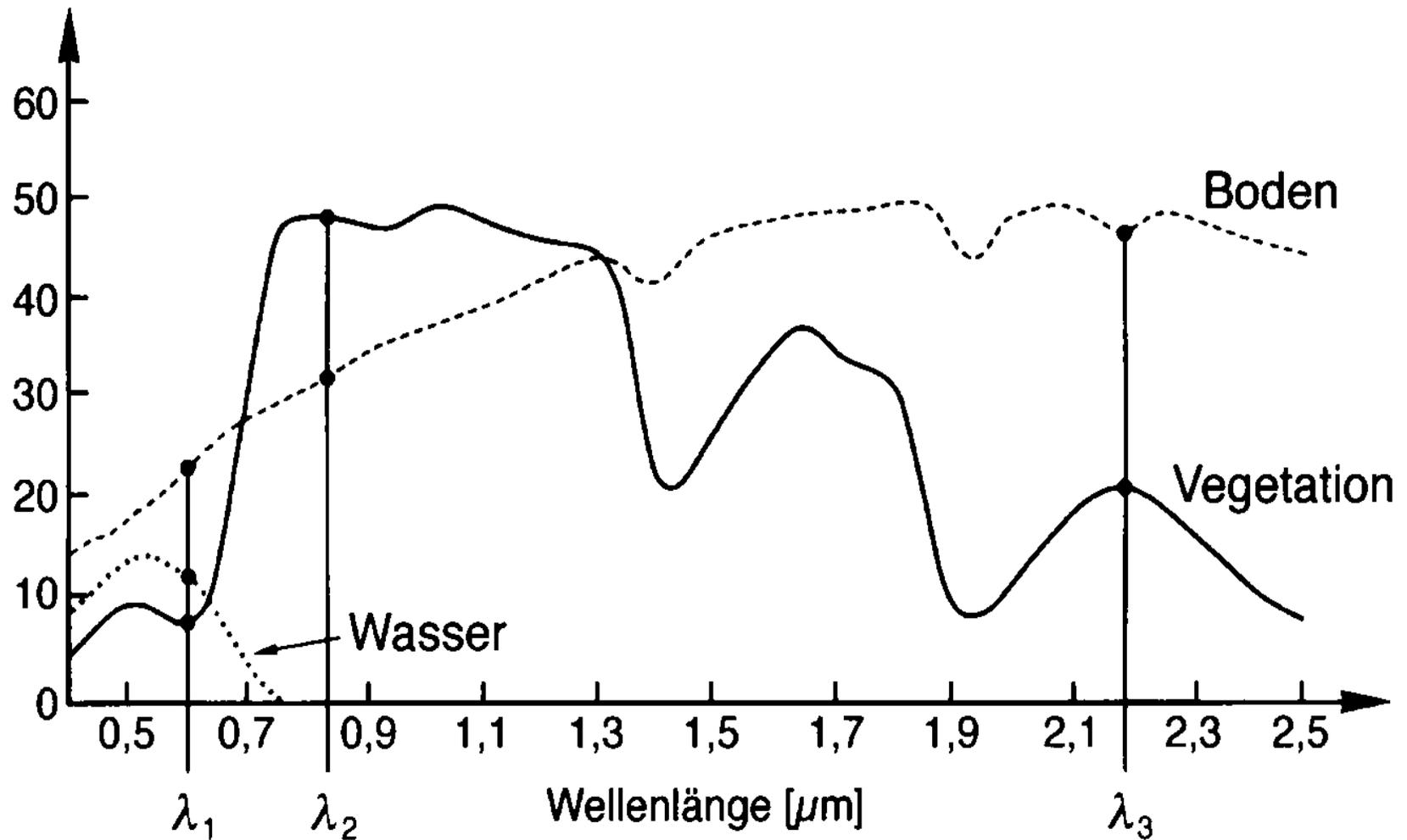
1. Auswahl Merkmale
 - Kleinste Anzahl, voneinander unabhängig
 - Klassenzuordnung ermöglichen
 - Ggf. normiert/normierbar
2. Definition von Klassen
 - Objektive Kriterien, sinnvolle Auswahl von Klassen
 - Gut trennbar, keine fließende Grenzen
3. Bestimmung von Klassengrenzen
4. Zuordnungsmethode eines unbekanntes Objektes zu einer der Klassen

Auswahl der Klassen & -grenzen

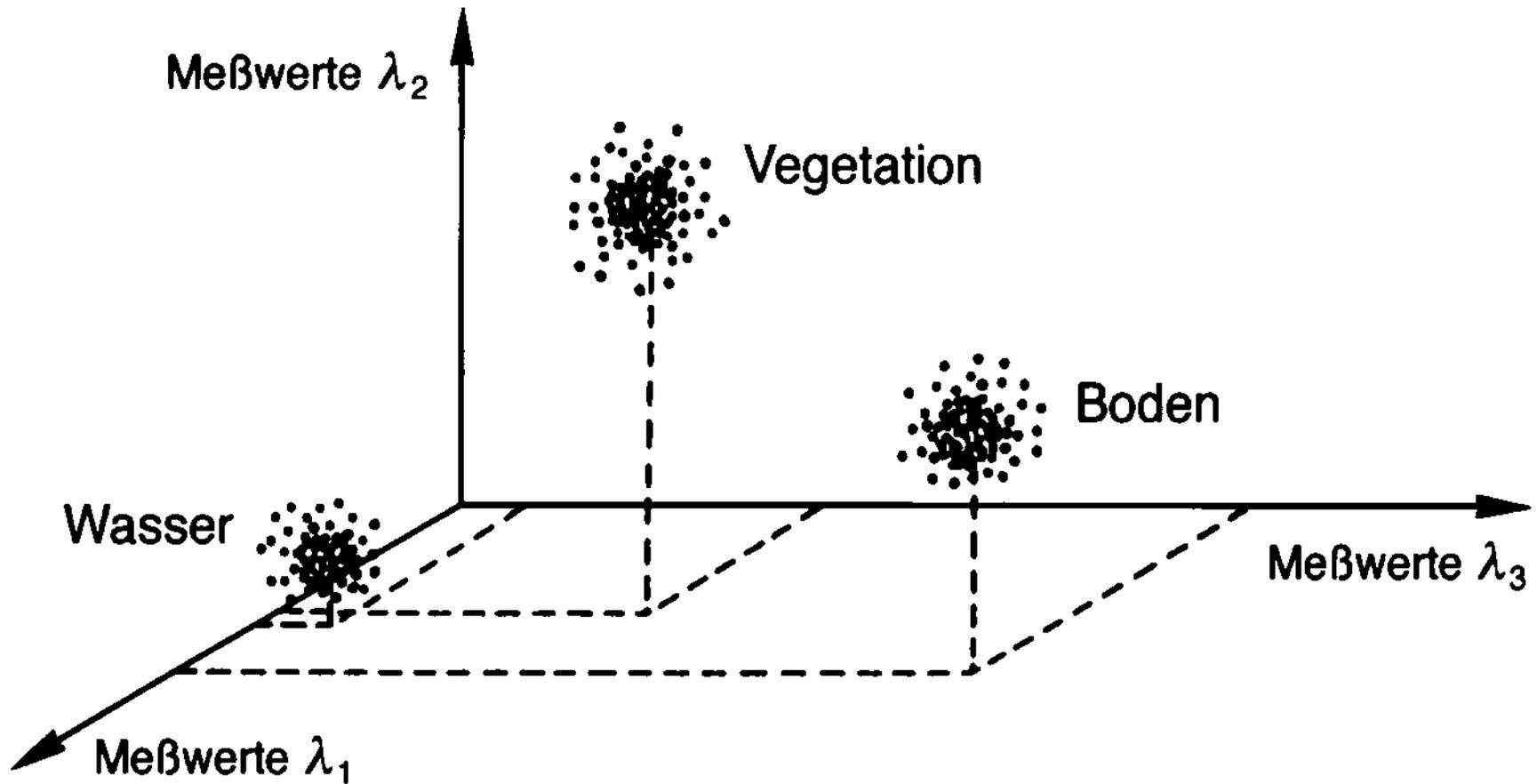
- Ermittlung der Punktwolken (Cluster) im Merkmalsraum
 - Trennung der Cluster
 - Zuordnungsabbildungen für unbekannte Vektoren
- ⇒ Nach Postulat 3 führen ähnliche Merkmale zu (einigermaßen) zusammenhängende Clusters
- ⇒ Wie Cluster bestimmen?

Beispiel: Meßwerte

Reflexionsgrad [%]



Beispiel: Cluster Bildung



Vorgehensweisen

Dimensionierung: Wahl der Stichprobe

- Fest dimensionierte = konstante Lernmenge
- Lernende = sich erweiternde Lernmenge

Überwachung: Wahl der Klassen und deren Grenzen

- Überwachte: Klassenanzahl & -grenzen fest
- Unüberwachte: Klassenanzahl & -grenzen veränderlich

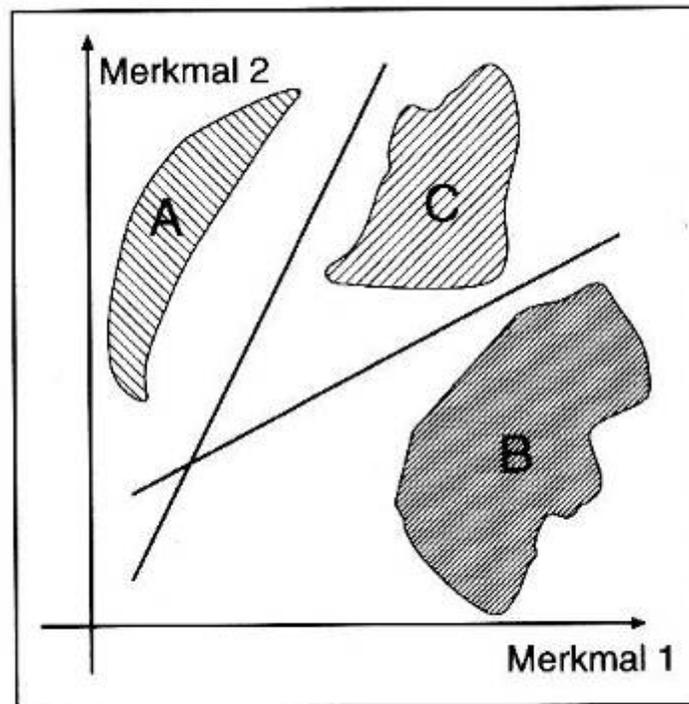
1. Fest dimensionierte überwachte Strategie
2. Fest dimensionierte unüberwachte Strategie
3. Lernende überwachte Strategie
4. Lernende unüberwachte Strategie

1: Fest-Dimensionierte überwachte Strategie

Vorgabe Stichprobe bekannter Objekte (Testbilder)

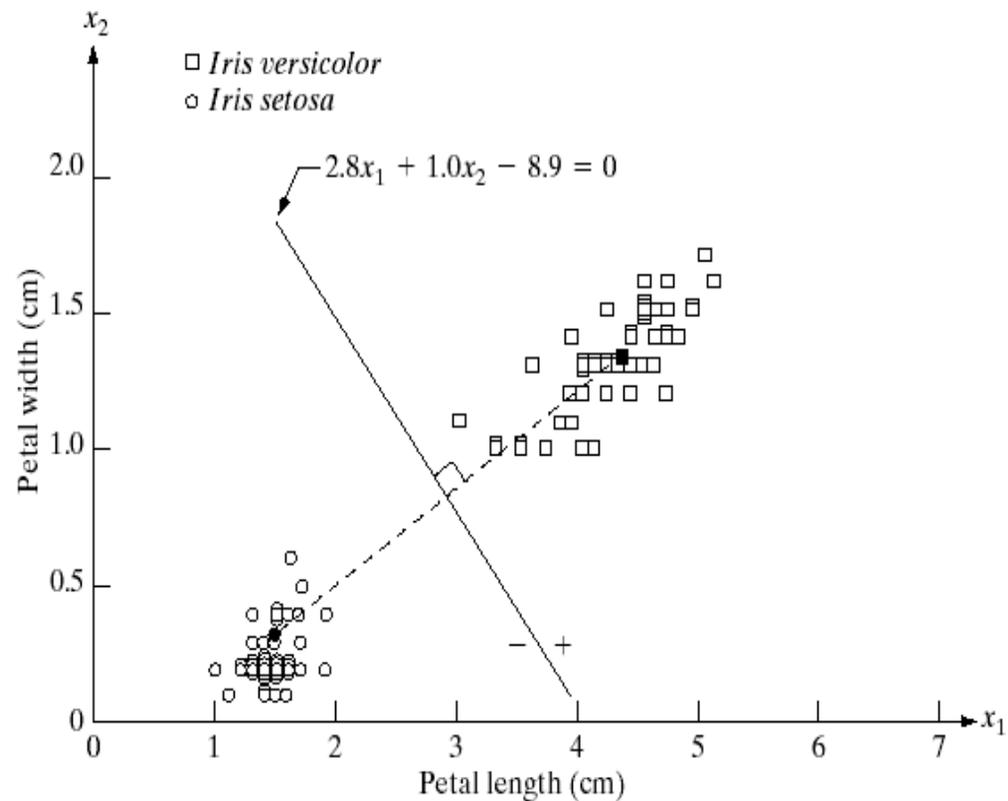
⇒ Musterklassen eindeutig durch die Stichproben festgelegt

⇒ (c_j, h_j) Wahrscheinlichkeiten a-priori bekannt/gemessen



Fest-Dimensionierte überwachte Strategie

- Repräsentative Stichprobe
- Definition der Klassen (z.B. per Hand)



Probleme

Stichprobe (Realisation) ist immer eine Approximation einer Klasse

Stichprobe „repräsentativ“?

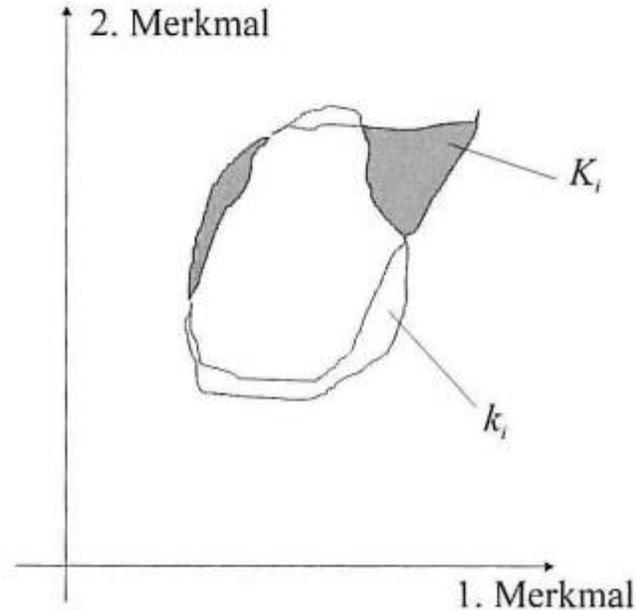
- Nicht verfügbar
- Verfügbar, aber anders als in der aktuellen Beobachtung
- Keine vollständige Abdeckung des Merkmalraumes
- Überabdeckung des Merkmalraumes

Dimensionierung „statisch“:

- Anzahl der Klassen vorab bekannt
- Grenzen unverschiebbar

Fehlerquellen

1. Stichprobe k_i umfasst nicht alle Merkmale der Klasse $K_i \Rightarrow$ Vektoren können der Klasse nicht zugeordnet werden (dunkler Bereich)
2. Stichprobe erfasst Merkmale, welche nicht zur Klasse gehören \Rightarrow „fremde“ Vektoren werden der Klasse zugeordnet (heller Bereich)



K_i = Objekt

k_i = Stichprobe

2: Fest-Dimensionierte unüberwachte Strategie

Feste Menge Trainingsvektoren, aber:

1. Anzahl Klassen unbekannt
2. Klassengrenzen unbekannt, wird während der Klassifikation generiert (clustering)

Vorteile

- Stichprobe einfacher & weniger anfällig
- Grenzen flexibler
- Re-clustering (dynamische Dimensionierung)

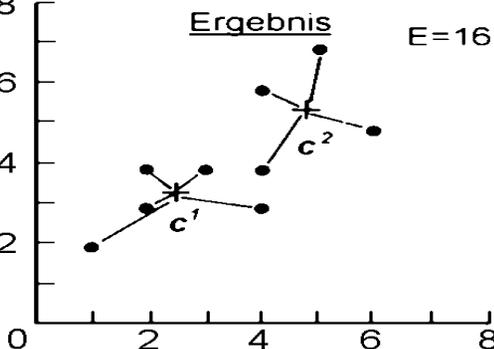
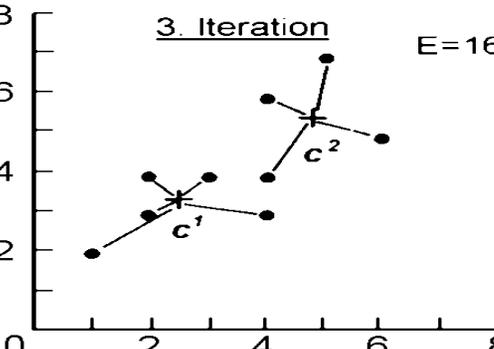
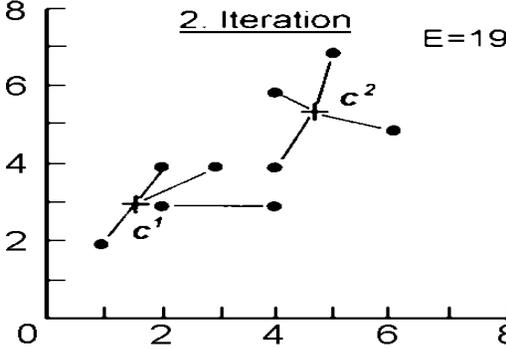
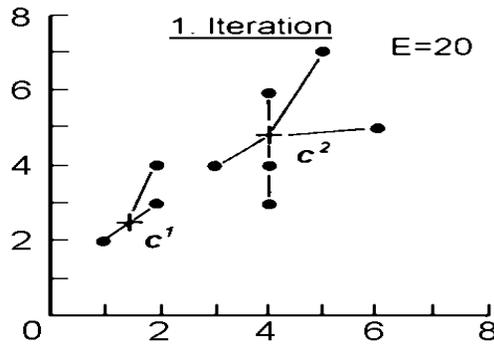
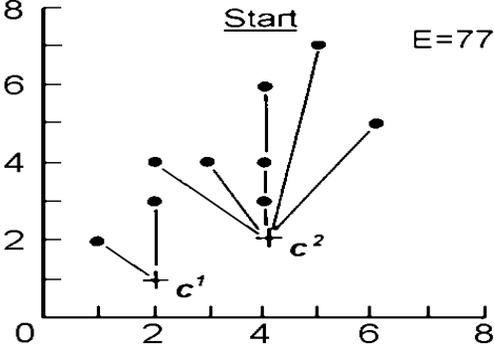
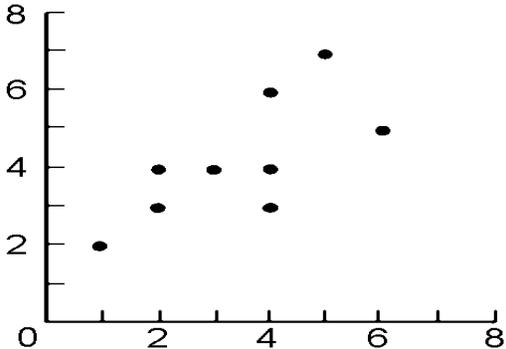
K-Means Vorgehensweise

1. Initiale Generierung (zufälligen, beliebigen, vordefinierten) von Clustern
2. Für jeden Sample: Berechnung der Distanz zu allen Cluster
3. Zuordnung jedes Samples zu Cluster, falls Distanz kleiner Schwelle ist
4. Neubewertung Clusteranzahl, -Grenzen, -Schwerpunkte
 - Wenn Vektor zu keinem Cluster zugehörig, neue Cluster
 - Cluster, die „zu nahe“ kommen, werden vereinigt
 - Cluster, die „zu groß“ werden, können geteilt werden
 - Neuberechnung der (neuen) Cluster-Schwerpunkte
5. Erneute Zuordnung aller Vektoren zu den neuen Cluster
6. Berechne Summe der Distanzmaße für jeden Cluster
7. Wiederhole 3, 4 & 5 bis Summe der Distanzmaße minimiert

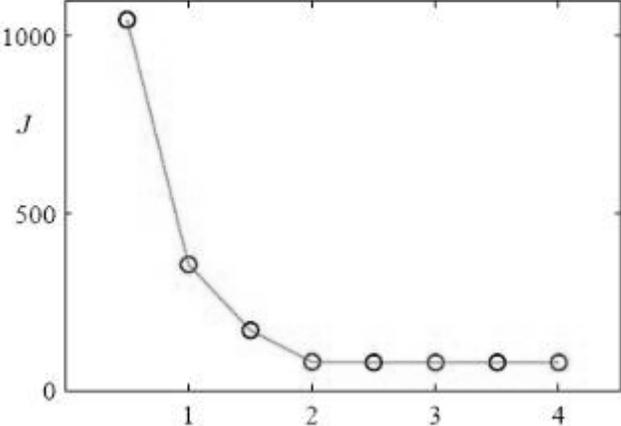
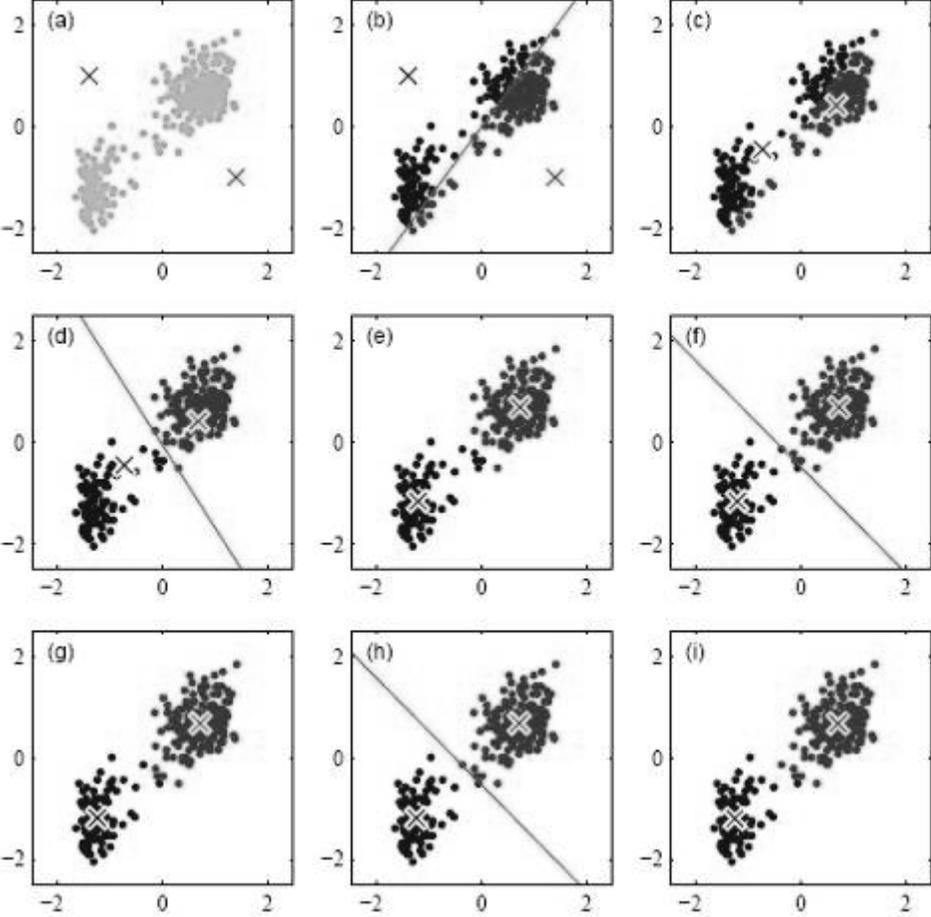
k-Means Clustering

1. Initialisierung der Clusterzentren
Zufällig oder aufgrund von Beurteilung des Bildinhalts
2. Zuordnung aller Daten x_i zum nächsten Clusterzentrum
3. Neuberechnung der Clusterzentren als „Datenschwerpunkte“
4. Schritte 2 und 3 werden so lange wiederholt, bis die Summe der Streuungsquadratsumme über alle Cluster minimiert

k-Means Beispiel



k-Means Beispiel

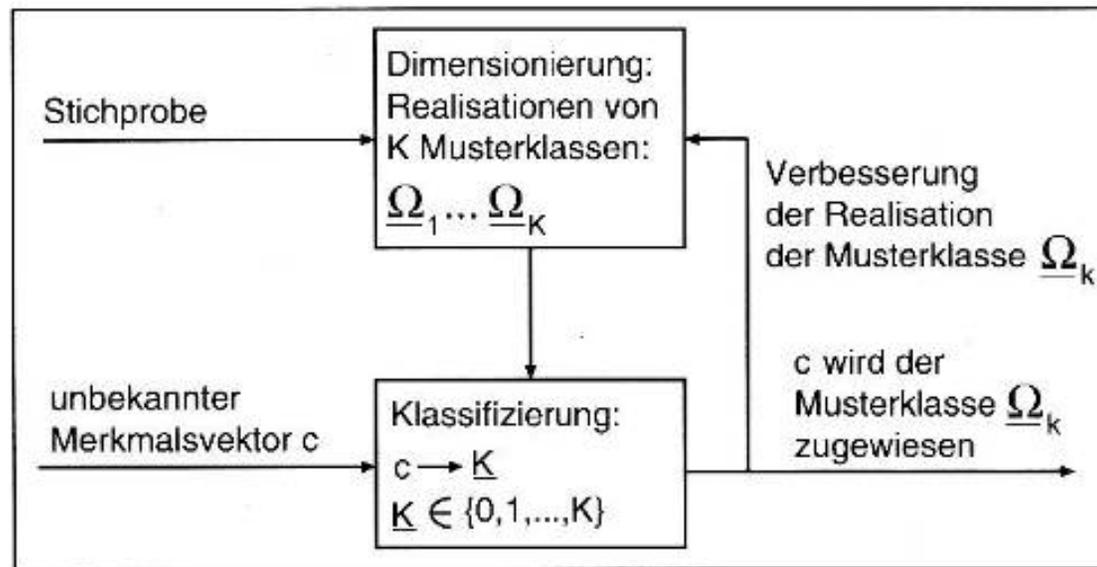


Lernende Strategien

- Stichprobe kann sich mit der Zeit verändern (Geldautomat)
- Trenderkennung durch Rückkopplung
 - Jeden neuen Vektor wird klassifiziert
 - Danach wird er zu der Trainingsmenge zugeführt (= Erweiterung!)
 - Dadurch Neudefinition der Klassengrenzen (überwacht) und – Anzahl (unüberwacht)

3: Lernende Überwacht

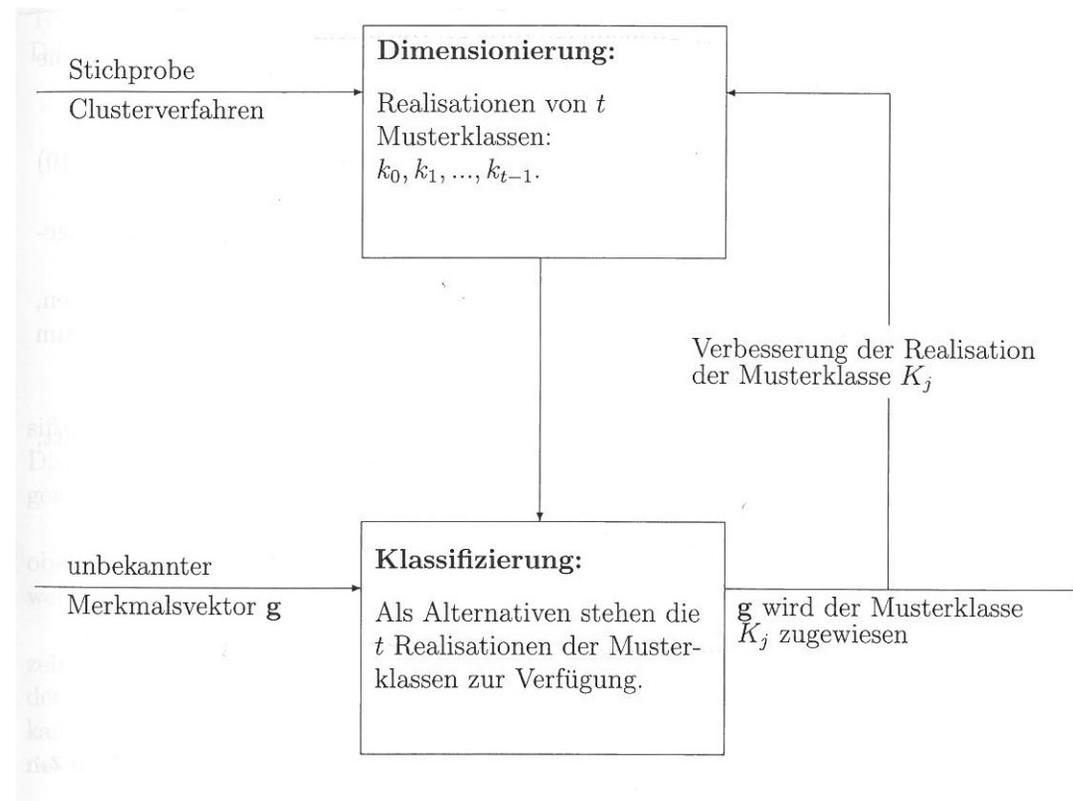
- Erstdimensionierung: bekannte Stichprobe, # Klassen
- Probe wird klassifiziert und zur Stichprobe zugeführt
- Trendanpassung: Nach Klassifikation re-Dimensionierung der Klassengrenzen (nicht Anzahl, da *überwacht*)



4: Lernende unüberwacht

train-on-the-job

- Initiale bzw. keine bekannte Stichprobe
- Anzahl und Grenzen der Klassen dynamisch durch Clustering



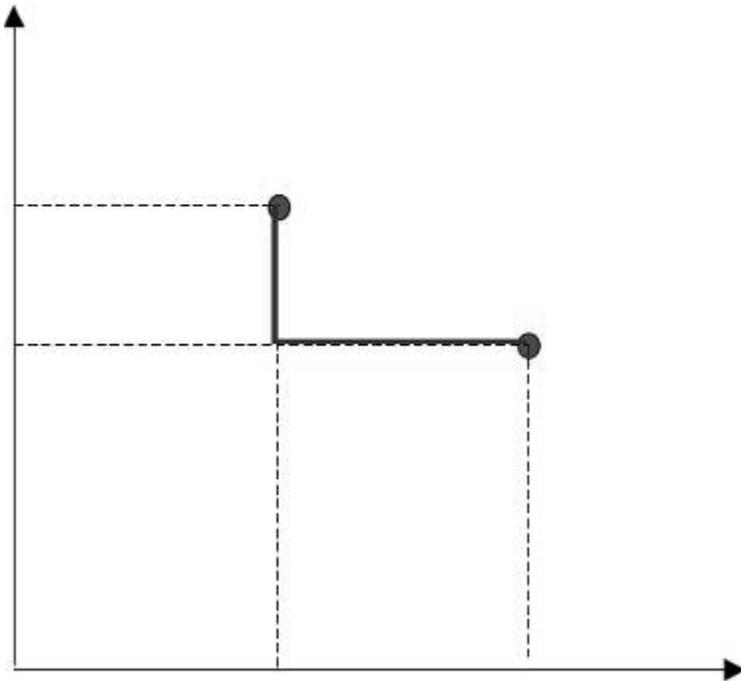
Schritte

1. Auswahl Merkmale
 - Kleinste Anzahl, voneinander unabhängig
 - Klassenzuordnung ermöglichen
 - Ggf. normiert/normierbar
2. Definition von Klassen
 - Objektive Kriterien, sinnvolle Auswahl von Klassen
 - Gut trennbar, keine fließende Grenzen
3. Bestimmung von Klassengrenzen
4. Zuordnungsmethode eines unbekanntes Objektes zu einer der Klassen

Metrik: Manhattan

Manhattan, City block: Summe der absoluten Differenzen
(Minkowski-Metrik mit $p=1$)

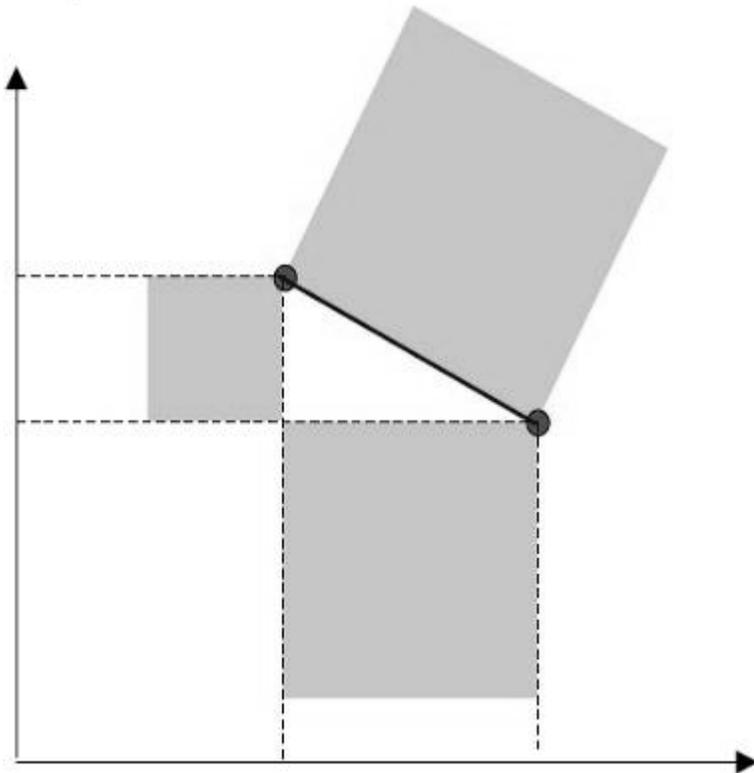
$$d(x, y) = L_1(x, y) = \sum_{i=1}^n |x_i - y_i|$$



Länge des Weges,
wenn man sich auf den
Koordinatenachsen bewegt

Metrik: Euklidisch

- Euklidischer Abstand: Wurzel aus der Summe der quadratischen Differenzen (Minkowski-Metrik mit $p=2$)

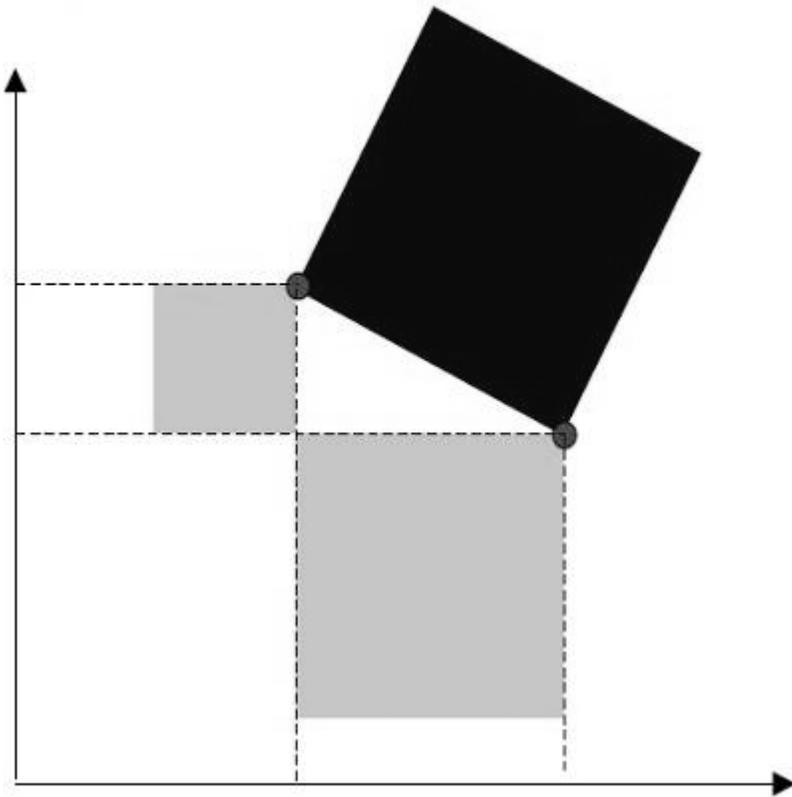


$$d(x, y) = L_2(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Entfernung entspricht der
Luftlinie

Metrik: Quadratische Euklidisch

- Quadratischer euklidischer Abstand: Summe der quadratischen Differenzen



$$d(x, y) = \sum_{i=1}^n (x_i - y_i)^2$$

Große Entfernungen fallen stärker in Gewicht

Minimum-Distance Klassifikator

Geometrischer Ansatz

- Zuordnung zu der Klasse, welche „am nächsten“ zu dem Vektor liegt
- Größte Nähe = kleinster Abstand
 - Verschiedene Metriken
- Abstand im Vergleich...
 1. zu einem Cluster-Vertreter (Mittelwert, Schwerpunkt, Median, Häufigster Wert)
 2. zu mehreren Cluster Elemente
 3. Zuordnung zum nächsten klassifizierten Nachbar (NN)
 4. Zuordnung zur Mehrheit von mehreren klassifizierten Nachbarn (k-NN)

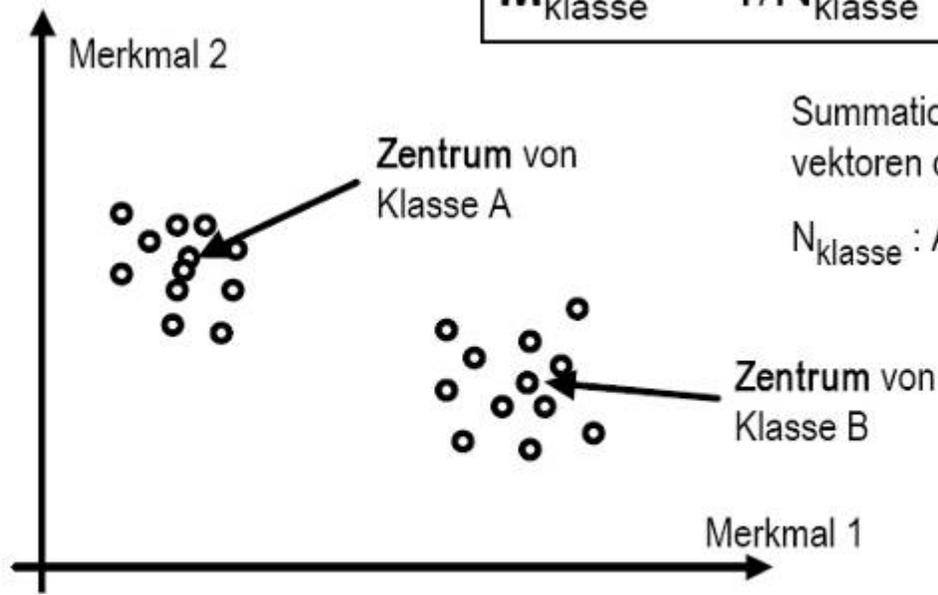
Minimum-Distance zum Clusterzentrum

Cluster-Zentren:

Von jedem Cluster einer Objektklasse wird das *Cluster-Zentrum* bestimmt (*Prototyp der Klasse*).

Das Cluster-Zentrum ist der Mittelwert der Merkmalsvektoren von einem Cluster.

$$\vec{M}_{\text{klasse}} = 1/N_{\text{klasse}} * \sum \vec{M}_{i(\text{klasse})}$$



Summation erfolgt über alle Merkmalsvektoren der Klasse.

N_{klasse} : Anzahl der Objekte einer Klasse

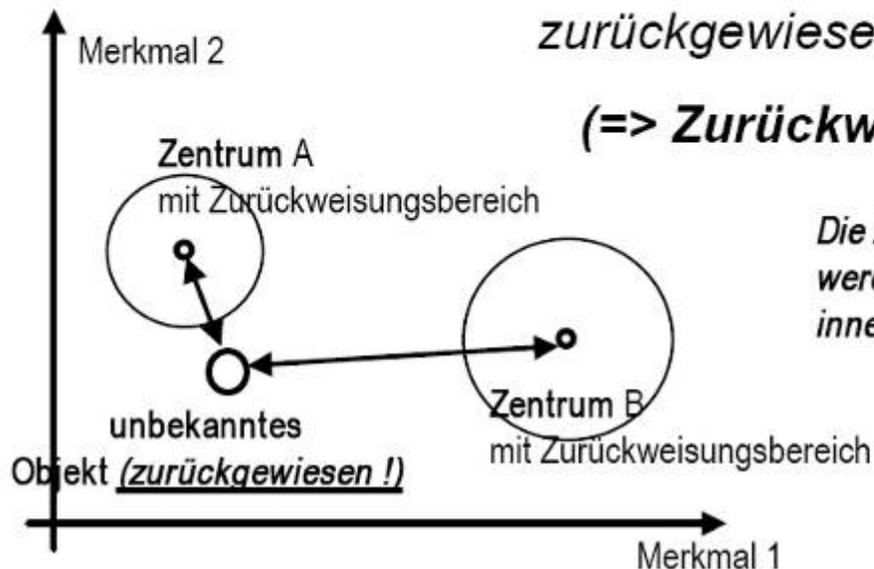
Minimum-Distance zum Clusterzentrum

Geeignet für „kompakte“ Cluster mit geringer Streuung
(gültige Approximation mit einem Zentralelement)

Eine Zuordnung eines unbekannten Objekts zu einer Objektklasse kann nur dann erfolgen, wenn der Abstand d_{Klasse} zum Zentrum der Klasse kleiner als ein gegebener Grenzwert ist.

Ist das bei keiner Klasse der Fall, wird das Objekt zurückgewiesen.

(=> Zurückweisungsklasse)



Die Zurückweisungs-Radien werden proportional zur Streuung innerhalb eines Clusters gewählt.

Nächster Nachbar (NN), festdimensioniert

Geeignet für „irreguläre“ Cluster, z.B. große Streuung, mehrere Ballungsanhäufungen im Merkmalsraum etc.

Für jeden Vektor:

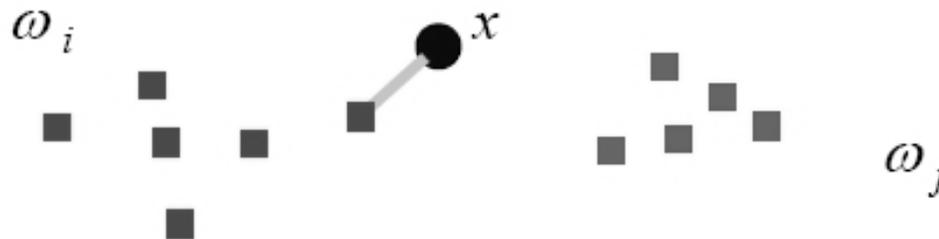
Für alle i Cluster C_i :

Für jeden der j Vertreter jedes der i Cluster:

- berechne j Distanzen $D_{i,j}$ des Vektors zu allen j Vertreter des jeweiligen Clusters C_i
- Distanz D_i für Cluster C_i = Minimum der berechneten Distanzen $D_{i,j}$
- Zuordnung zum Cluster C_i mit kleinster Distanz D_i

Nächster-Nachbar (NN), lernernd

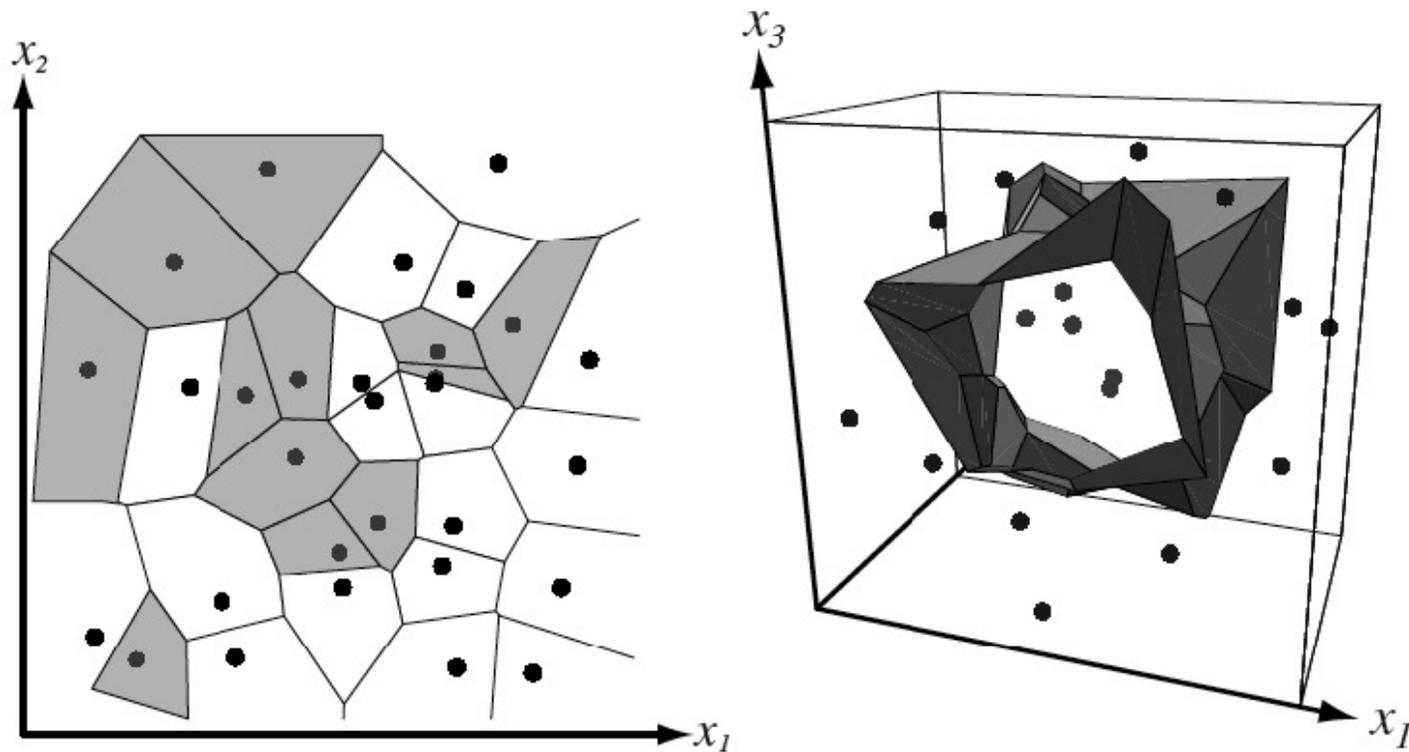
- Stichprobe: Es bestehen bereits j klassifizierten Vektoren $S_{i,j}$, welche zu i bekannten Klassen C_i zugeordnet sind
- Für jeden neuen Vektor X berechne alle Distanzen $D(X, S_{i,j})$, wobei D beliebiges geeignetes Distanzmaß
- Suche kleinstes Distanzmaß $D_m = \min \{D(X, S_{i,j})\}$
- Ordne X zu der Klasse aus C_i zu der der vorklassifizierter Vektor $S_{i,j}$ mit dem kleinsten Distanzmaß D_m zu X bereits gehört
- X wird nach Klassifikation zu $S_{i,j}$, d.h. es erweitert die klassifizierte Stichprobe



⇒ Klassengrenzen werden dynamisch Verschoben (überwacht lernend)

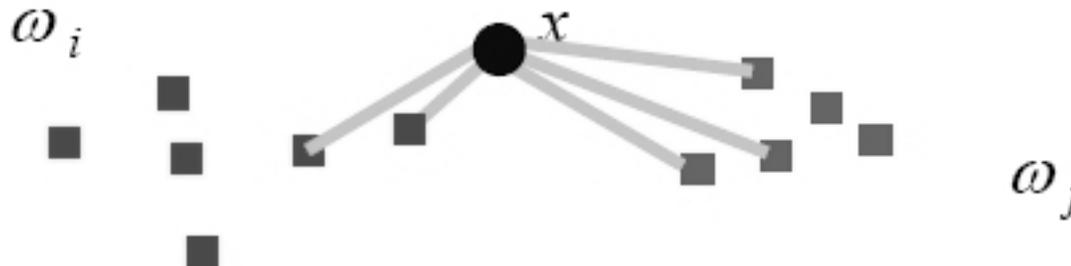
Minimum-Distance Klassifikator

- Verbindungen der Klassenzentren: Delauny Triangulierung
 - Klassengrenzen: Voronoi Diagramm
-

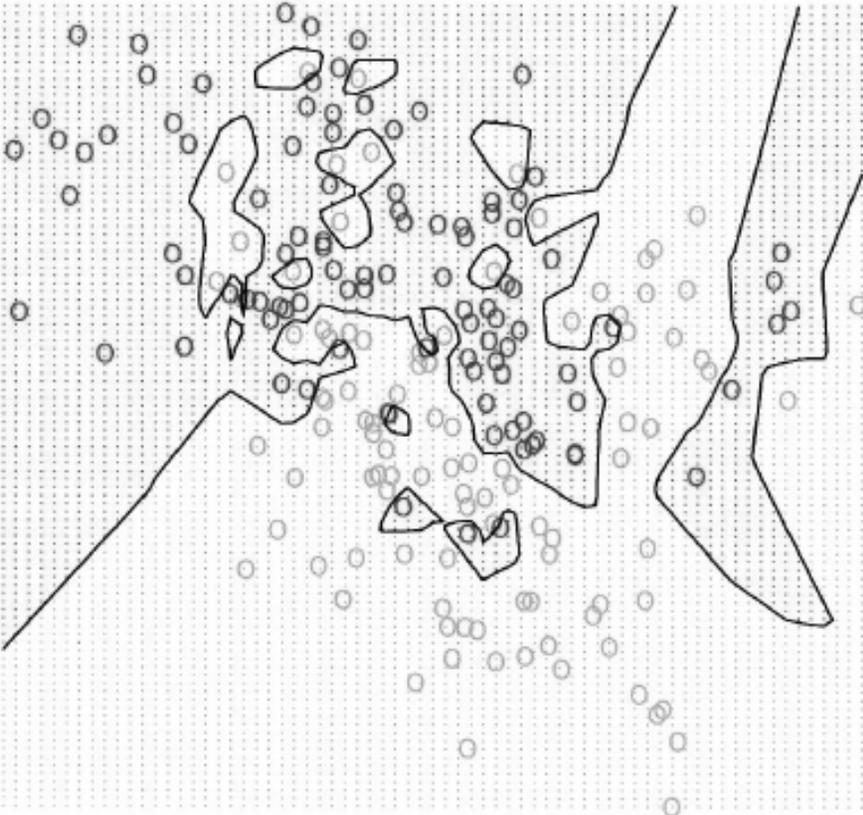


k-Nächster-Nachbar (k-NN)

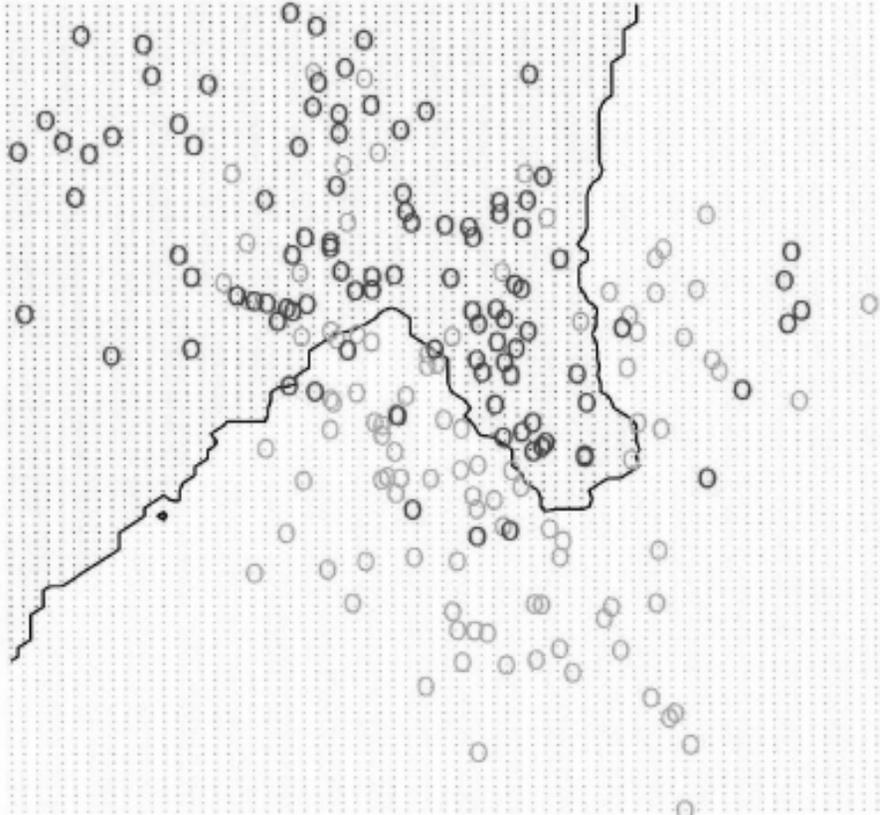
- Berechne alle Distanzen $D(X, S_{i,j})$ wie bei NN
- Bestimme k nächste Nachbarn zu X (k ungerade)
- Mehrheitsentscheidung, welche Klasse C_i „gewonnen“ hat
 - Kleines k = Berücksichtigung nahe Klassen, Lokalverhalten
 - Großes k = Berücksichtigung auch ferner Klassen, robuster
- Erweiterung: höhere Gewichtung näheren Nachbarn (z.B. quadratische Euklidische Metrik)



NN vs k-NN

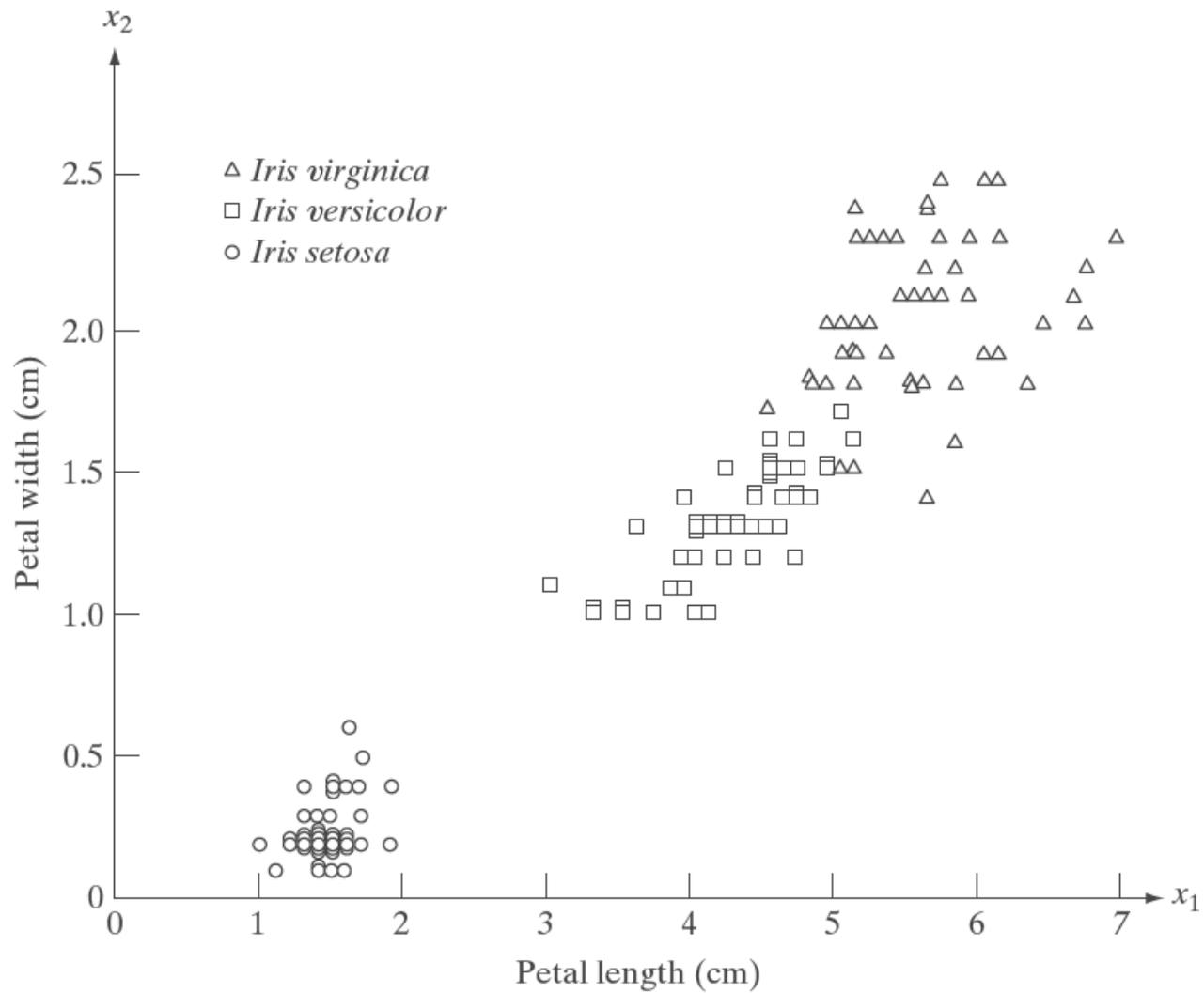


k = 1



k = 15

Maximum-Likelihood-Klassifikator



Maximum-Likelihood-Klassifikator

- Sonderfall des Bayes'schen
- Statistischer Ansatz

$p(x)$: Verteilungsdichte des Vektors x

$P(w_i)$: a-priori Wahrscheinlichkeit der Klasse w_i

$p(x|w_i)$: Verteilungsdichte von x in $w_i =$ Wahrscheinlichkeit, daß der Wert x vorliegt, wenn w_i vorliegt

- $P(w_i) * p(x|w_i)$ ist Maßstab (*Likelihood*) für Zugehörigkeit von x zu w_i

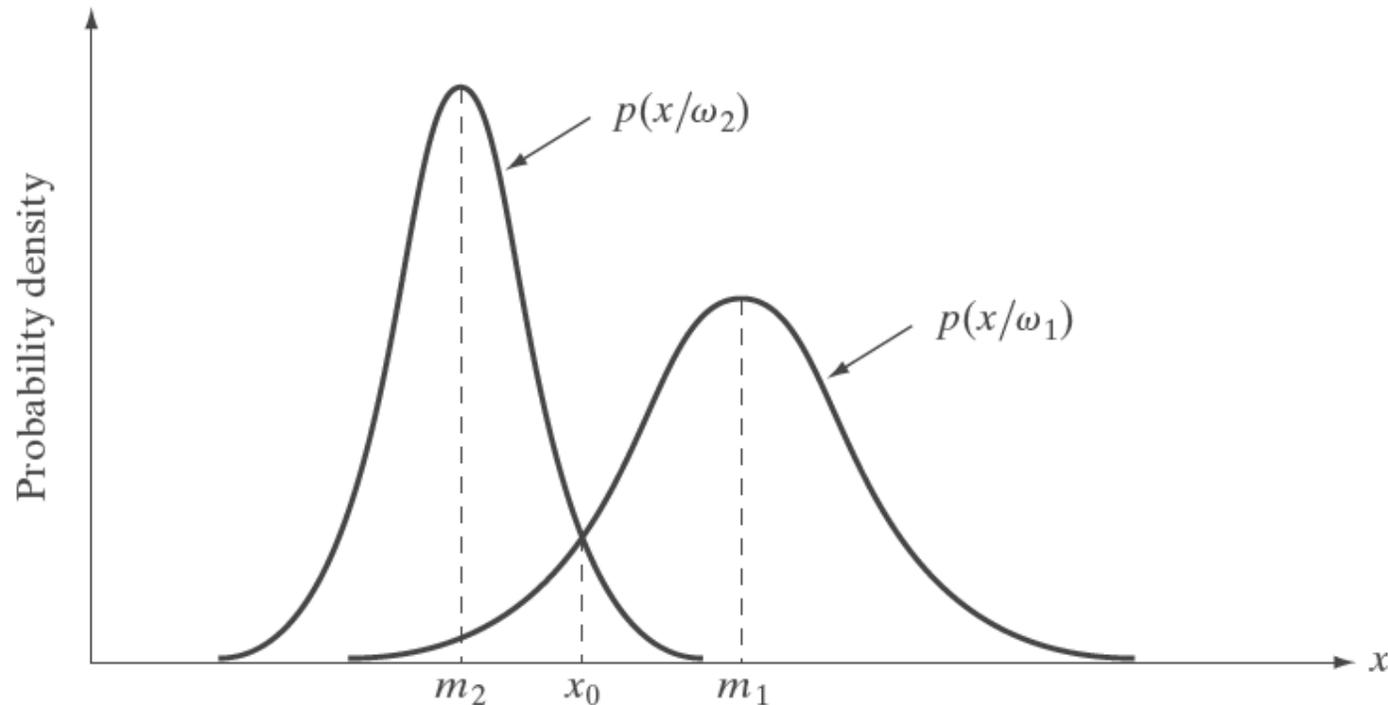
Maximum-Likelihood-Klassifikator

- Ordnet w zu der Klasse w_k so, daß der Erwartungswert der gesamten Fehlklassifikation minimiert wird:

$$p(x) - P(w_i) * p(x|w_i) < p(x) - P(w_j) * p(x|w_j)$$

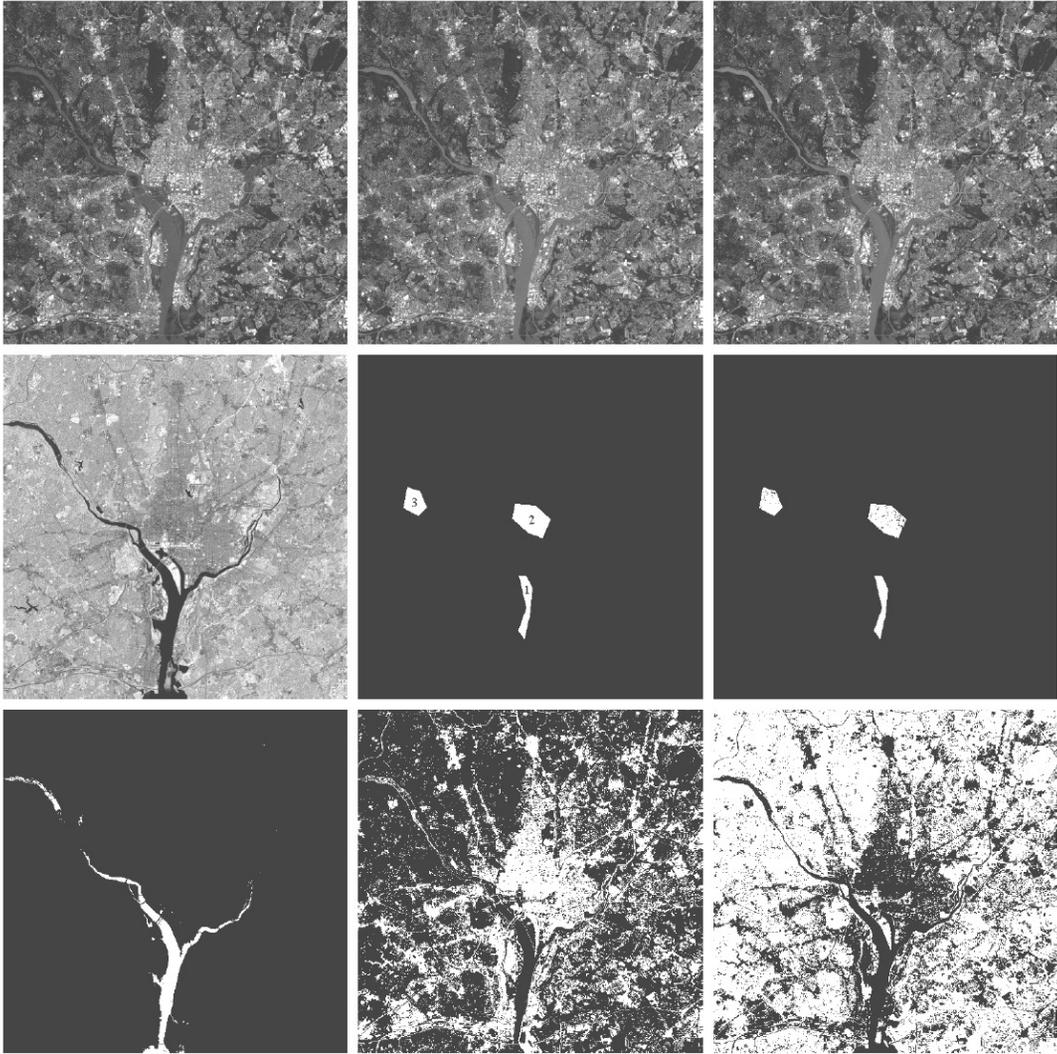
- Nachteil: Wahrscheinlichkeiten $P(w_i)$ und Verteilungsdichten $p(x|w_i)$ müssen a-priori bekannt sein
 - $P(w_i)$: i.d.R. kein Problem, Dimensionierung mit repräsentative Trainingsmenge
 - $p(x|w_i)$: Schwer. Annahme z.B. Gaus'schen Verteilung

Maximum-Likelihood-Klassifikator



- Häufigkeit der Klasse $P(\omega_1) = \text{Fläche unter } p(x|\omega_1) \text{ -- bzw. } 2$
 - Wenn Flächen gleich groß, Häufigkeit beider Klassen identisch
- ⇒ Klassengrenze Punkt x_0

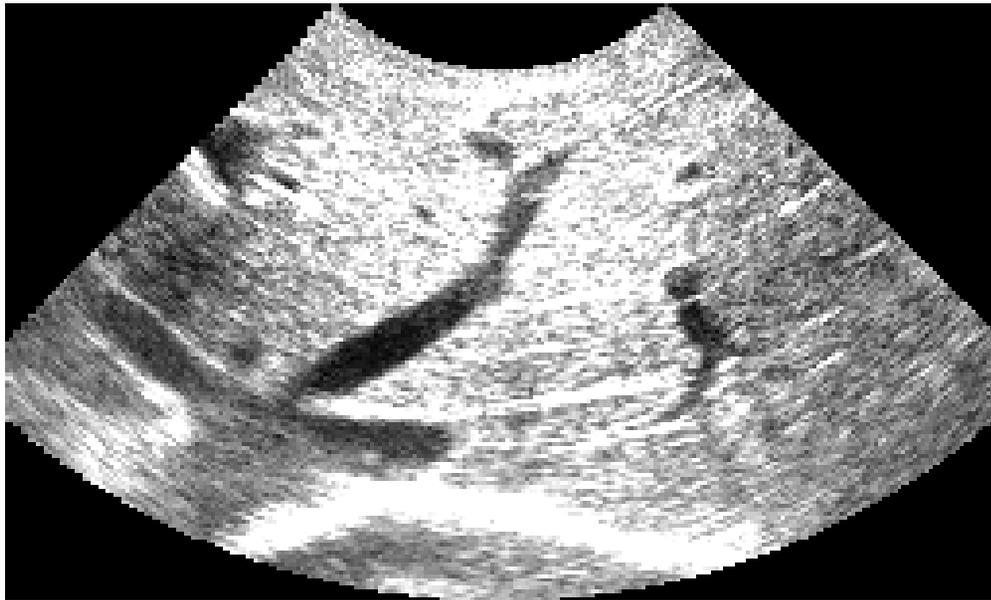
Maximum-Likelihood-Klassifikator



Ursachen für Fehlklassifikation

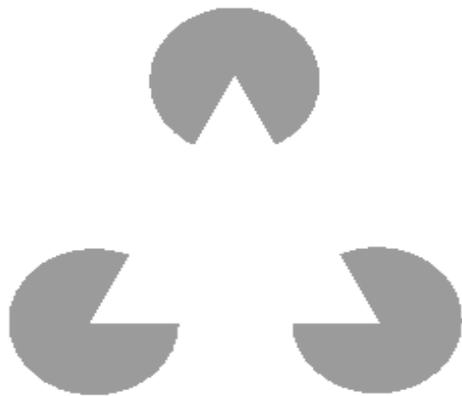
- ▶ Klassen überschneiden sich oder haben fließende Grenzen
 - Z.B. Klassifikation von Farben
 - Hier können u.U. Fuzzy Klassifikationsverfahren helfen
- ▶ Zu wenige oder nicht hinreichend signifikante Merkmale führen zu zufälligen Ergebnissen
 - Beispiel: Unterscheidung von Äpfeln und Birnen anhand der Farbe
- ▶ Messfehler bei der Merkmalsextraktion
- ▶ Unzureichendes Training eines Lernverfahrens
- ▶ Ungeeigneter Trainingsdatensatz
 - Z.B. ungleichmäßige Verteilung der Stichproben auf die Klassen

Wissen Berücksichtigung



Modellbasiertes Segmentieren

- Berücksichtigung von *a priori* Informationen über Art und Aussehen der zu segmentierenden Objekte
- Mensch besitzt ein „Modell“ des Objektes im Kopf, das Segmentieren vereinfacht:



Die Kreisausschnitte werden zu durchgehenden Strecken ergänzt

Danksagung

- Dr. Andrea Miene, Bremen
- Prof. Hillen, Aachen/Jülich
- Prof. Franz, Konstanz
- Prof. Jiang, Münster