

# Digitale Bildverarbeitung

## Einheit 11

### Klassifikation

---

Lehrauftrag SS 2006

Fachbereich M+I der FH-Offenburg



Dr. Bernard Haasdonk

Albert-Ludwigs-Universität Freiburg

## Ziele der Einheit

---

- Verstehen, dass basierend auf einer Repräsentation von Objekten durch **Merkmalsvektoren** mächtige Operationen möglich sind, z.B. die **Klassifikation**
- **Grundlegende Begriffe** wie Klassifikator, Klassenlabel, Trainingsdaten, Klassifikationsgebiete oder Entscheidungsgrenze werden eingeführt
- Die wichtigsten einfachen **Klassifikatortypen** und einige **Anwendungen** werden präsentiert: Grenzen und Möglichkeiten
- Allgemein wird lediglich ein **Eindruck des Gebiets** der "Klassifikation" vermittelt, da dieses sehr schnell sehr technisch wird.
- Darstellung ist mehr anschaulich, enthält (fast) keine Formeln, soll das Gebiet "schmackhaft" machen

## Motivation

- Ein wichtiger Verarbeitungsschritt ist die **Erkennung** von einem Gesamtobjekt, was der Mensch prima beherrscht.



„Hund“



„Katze“



„Maus“

- Die Frage der **maschinellen Klassifikation** wird auf dem Gebiet des Klassifikatorentwurfs behandelt
- Klassifikatorentwurf ist wie die Merkmalsextraktion ein wichtiges Teilgebiet der Mustererkennung
- Erkennung kann man als eine Art „**Intelligente**“ Leistung oder Handlung sehen.
- Entsprechend machen auch die Felder der **Künstlichen Intelligenz**, des **Maschinellen Lernens** und des **Data-mining** und **Information Retrieval** sich diese Techniken zunutze

8.4.2006

B. Haasdonk, Digitale Bildverarbeitung, FH Offenburg SS 2006, Einheit 11

3

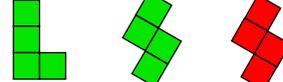
## Grundbegriffe

- Objektvergleich im Merkmalsraum**
  - Nach der **Merkmalsextraktion** ist ein Objekt nur noch ein Punkt in einem abstrakten Raum, dem **Merkmalsraum**
  - Seine **Koordinaten** sind genau die Einträge seines Merkmalvektors
  - Man kann verschiedene Objekte **gleichzeitig** in diesem Raum repräsentieren
  - Man kann verschiedene Objekte **vergleichen**, indem man den **Abstand** zwischen ihren Punkten im Merkmalsraum berechnet (Satz des Pythagoras) und den Abstand als Maß für den Unterschied nimmt.

Beispiel: Abstand 1. & 3. Objekt

$$\sqrt{2^2 + 1^2} = 2.236$$

Objekte



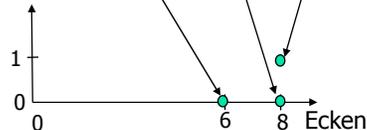
Merkmalsvektoren

(Ecken und Rot-Mittelwert)

$$\begin{pmatrix} 6 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 8 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 8 \\ 1 \end{pmatrix}$$

Merkmalsraum

Rot-Mittelwert



8.4.2006

B. Haasdonk, Digitale Bildverarbeitung, FH Offenburg SS 2006, Einheit 11

4

## Grundbegriffe

- Klassifikationsproblem **anschaulich**:

- Gegeben eine Menge von Objekten, die zu bekannten Kategorien gehören, den sogenannten **Klassen**
- D.h. zu jedem Objekt ist eine **Klassenzugehörigkeit**, oder ein sogenanntes **Label** bekannt

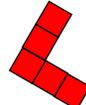
Objekte: Buchstaben



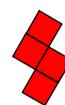
Label: „L“



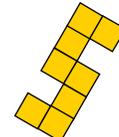
„S“



„L“

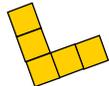


„S“



„S“

- Aufgabe: Finde einen Weg oder Verfahren, zukünftige Objekte ohne Label möglichst gut einer Klasse zuzuordnen.



=?



=?

- Dies heißt **Klassifikation** der neuen Objekte
- Die Zuordnungsvorschrift heißt **Klassifikator**

8.4.2006

B. Haasdonk, Digitale Bildverarbeitung, FH Offenburg SS 2006, Einheit 11

5

## Grundbegriffe

- **Klassifikation im Merkmalsraum**:

- Die Wirkung eines Klassifikators läßt sich schön im Merkmalsraum darstellen
- Der Klassifikator entscheidet für jedes Objekt, das heißt für jeden Merkmalsvektor eine Klasse.

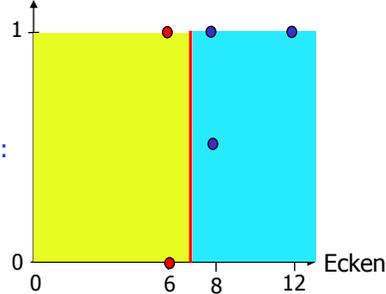
Klassifikator: Klasse „L“ falls Ecken  $\leq 7$ , sonst „S“:

- Diese Entscheidungen kann man grafisch durch **Klassengebiete** und **Entscheidungsgrenzen** zusammenfassen



Merkmalsraum

Rot-Mittelwert



- Das Aufstellen der Klassifikationsvorschrift basiert oft auf die gegebenen Daten. Daher heißt dies auch **Training** des Klassifikators und die Daten heißen **Trainingsdaten**.
- Der **Trainingsfehler** ist ein Gütemaß eines Klassifikators ist (es gibt bessere!) • „L“ • „S“ → Hier Trainingsfehler = 0

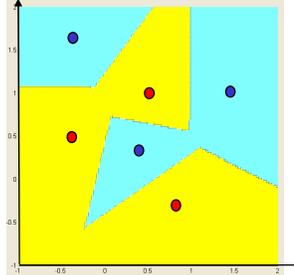
8.4.2006

B. Haasdonk, Digitale Bildverarbeitung, FH Offenburg SS 2006, Einheit 11

6

## Nearest-Neighbour-Klassifikation

- Der Nächste-Nachbar (NN) Klassifikator
- Training: nicht notwendig  
einfach speichern der Trainingspunkte
- Klassifikation:
  - Der nächste Nachbar des zu klassifizierenden Punktes unter den Trainingspunkten wird gesucht
  - Dessen Klasse wird als Klassenzuordnung gewählt.
- Klassifikationsgebiete:
  - Sind begrenzt durch Mittelsenkrechten zwischen Trainingspunkten
- Vorteile:
  - Durch die **nichtlineare** Entscheidungsgrenze können sehr komplexe Probleme gelöst werden
  - **Einfachheit**: das Prinzip ist auf dieser einzelnen Folie erklärbar!
- Nachteile:
  - Leidet unter Überanpassung, sogenanntes „Overfitting“, d.h. sogar Rausch-Punkte werden korrekt klassifiziert.
  - **Klassifikationszeit** und **Speicheraufwand** sind sehr hoch



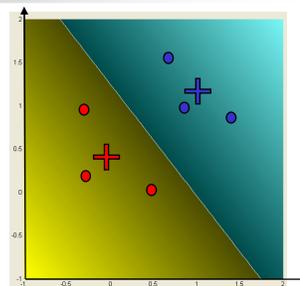
8.4.2006

B. Haasdonk, Digitale Bildverarbeitung, FH Offenburg SS 2006, Einheit 11

7

## Nearest Mean Klassifikation

- Training:
  - Bestimmen der beiden **Klassenmitten** (Schwerpunkte)
- Klassifikation:
  - Abstand des neuen Punktes zu den Klassenmitten berechnen
  - Das Label der nächsten Klassenmitte wird als Klassenlabel gewählt.
- Klassifikationsgebiete
  - Sind Halbebenen, die durch die Mittelsenkrechte der beiden Klassenmitten getrennt sind
- Nachteil:
  - Ist ein **linearer** Klassifikator (bei einem **Zweiklassenproblem**): Komplizierte verschachtelte Probleme, können nicht gelöst werden.
- Vorteil:
  - **Kein Overfitting**
  - Sehr **schnelle Klassifikation**: Egal wieviele Trainingspunkte vorhanden sind, die Klassifikation besteht aus Berechnung von zwei Abständen.



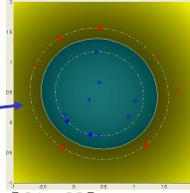
8.4.2006

B. Haasdonk, Digitale Bildverarbeitung, FH Offenburg SS 2006, Einheit 11

8

## Fortgeschrittene Klassifikator-Arten

- K-Nächste-Nachbar Klassifikator
  - Weiterentwicklung des NN-Klassifikators
- Polynomklassifikator [64]:
  - Statt einer linearen Grenzlinie gibt es quadratische Kurven: Kreise, Ellipsen, Hyperbeln
- Sehr wirkungsvoll in der Praxis:
  - Neuronale Netze [63], Support-Vektor-Maschinen [61, 62]
- Typisch für „fortgeschrittene“ Klassifikatoren
  - Nichtlinearität: Voraussetzung für komplexe Probleme
  - Anpassbarkeit der Komplexität der Grenzlinie durch einfache Parameter, d.h. Abwägen von Overfitting und Underfitting



- Kombination von Klassifikatoren für verschiedene Zwecke:
  - Verbesserung der Klassifikation, Objekterkennung durch Teilerkennung, Multiklassen-Klassifikation

8.4.2006

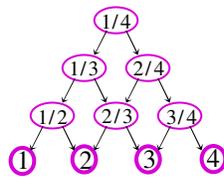
B. Haasdonk, Digitale Bildverarbeitung, FH Offenburg SS 2006, Einheit 11

9

## Multiklassen-Klassifikation

- Bisher: Zwei Klassen erfordern einen sogenannten binären Klassifikator
- Bei einer Aufgabe mit mehr als zwei Klassen braucht man einen Multiklassen-Klassifikator
- Einige Ansätze funktionieren „natürlich“ in diesem Fall, z.B. Nächste Nachbar
- Binäre Klassifikatoren können durch Kombination für ein Multiklassen-Problem verwendet werden

Gerichtete Graphen:



- Für jede Zweiklassen-Kombination wird ein binärer Klassifikator trainiert
- Diese werden in einem Graphen angeordnet
- Bei der Klassifikation wird ein Pfad im Graph abgelaufen
- Die unterste Ebene liefert die endgültige Klassenentscheidung

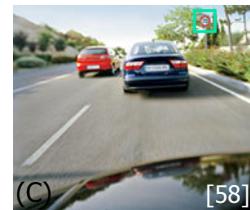
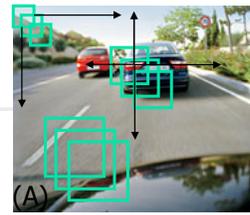
8.4.2006

B. Haasdonk, Digitale Bildverarbeitung, FH Offenburg SS 2006, Einheit 11

10

## Objektdetektion

- Abgrenzung von Objektklassifikation
  - Bei **Objektdetektion** lautet die Aufgabe, dass ein bestimmter Objekttyp lokalisiert werden soll, z.B. Verkehrsschilder
- Es können **mehrere Objekte** in einem Bild vorhanden sein.
- Typische (teure) Vorgehensweise:
  - Es wird ein sehr **schneller Klassifikator** verwendet, der ein Objekt vom Hintergrund unterscheiden kann.
  - Ein **Suchfenster** wird in allen Positionen und in allen Größen über das Suchbild geschoben (A).
  - In jeder dieser zahlreichen überlappenden Positionen wird der Inhalt auf „Hintergrund“ oder „Objekt“ klassifiziert, Ergebnis ist (B).
  - Eventuelle überlappende erkannte Bereiche werden **verschmolzen** (C)



8.4.2006

B. Haasdonk, Digitale Bildverarbeitung, FH Offenburg SS 2006, Einheit 11

11

## Anwendungen

- Webcam-Objekterkennung
  - Durch Subtraktion des Hintergrundes, Schwellwertbildung, Rausch-Filterung werden die Bilder segmentiert.
  - Jedes Segment wird als Merkmalsvektor repräsentiert, der aus zusammenfügen von **rot, grün und blau-Histogramm** entsteht (jeweils 8 Stufen)
  - Als Trainingsmenge werden Bilder von Objekten mit fester Kategorie gespeichert
  - Im Live-Bild wird der **Nächste-Nachbar Klassifikator** angewandt, um die beste Kategorie zu bestimmen (siehe Webseite für MATLAB-Programm-Code)

Livebild



Nächste-Nachbar Klassifikator



Vorhersage

Hand

8.4.2006

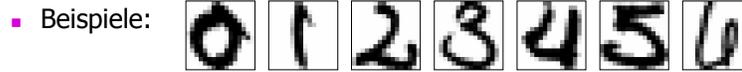
B. Haasdonk, Digitale Bildverarbeitung, FH Offenburg SS 2006, Einheit 11

12

## Maschinelle Schrifterkennung

- **Benchmarkdatensatz der USPS-Ziffern:**

- 7191 Training-, 2007 Testbeispiele von handgeschriebenen Ziffern in Form von 16x16 Graubildern.



- Wird verwendet um neue Klassifikatoren zu testen
- Klassifikationsfehlerraten aus der Literatur [59] zeigen, dass die heute erwähnten Verfahren tatsächlich sehr gut sind.

Neuronales Netz

Support-Vektor-Maschine

k-Nächster-Nachbar Klassifikator

Method	Error rate [%]
Human Performance [13]	2.5
Neural Net (LeNet1) [14]	4.2
SVM, no invariance [11]	4.0
SVM, VSV-method [12]	3.2
k-Nearest Neighbour [14]	*5.7
k-NN + TD [14]	*2.5
TD + kernel densities [7]	2.4
* := extended training set.	

## Maschinelle Schrifterkennung

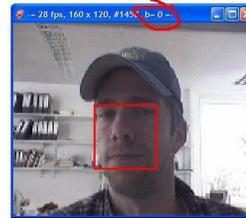
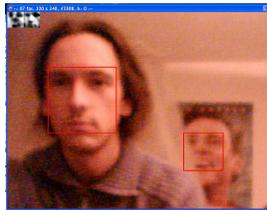
- Blockschrift, Handschrift kann erkannt werden, ist bereits vielseitig im Einsatz: PDA mit Stifteingabe, Scanner mit Texterkennungsfunktion
- Beispiel dafür was (noch) nicht möglich ist:  
*Verzerrte überlagerte Schrift*
- Anwendung bei Internet Zugangskontrollen: **Captchas**
  - Der Benutzer bekommt auf einer Webseite ein Bild vorgesetzt, muss die enthaltene Wörter erkennen und in Textform eingeben
  - Ziel: Zugriff/Einloggen von automatisierten Skripten verhindern



- <http://www.captcha.net/>

## Maschinelle Gesichtsdetektion

- Beispiel dafür, was möglich ist:
- **Gesichtsdetektion**: Personenunspezifische Ortung von Gesichtern (im Gegensatz zu Gesichtserkennung, d.h. Identifikation von Individuen)
- **Demo** basierend auf **Support-Vektor-Maschinen (SVM)**
  - Echtzeit, verschiedene Skalierungen, personenunabhängig, robust
  - Hardware-Voraussetzung: Windows-PC und Webcam
  - Einstellung der Entscheidungsschwelle durch Mausrad



- Download und weitere Informationen in [60] bzw. unter [www.kyb.mpg.de/bs/people/kienzle/facedemo/facedemo.htm](http://www.kyb.mpg.de/bs/people/kienzle/facedemo/facedemo.htm)

## Zusammenfassung

- Der **Merkmalsraum** erlaubt eine gemeinsame Darstellung verschiedener Objekte, indem **jedes Objekt** durch **einen Punkt** gegeben ist, dessen Koordinaten die Merkmale sind.
- Basierend auf der Merkmalsextraktion sind **komplexere Operationen** möglich, z.B. können Objekte klassifiziert werden
- **Klassifikatoren** unterteilen den Merkmalsraum in Gebiete, die den verschiedenen Klassen gehören. Die **Klassifikationsgebiete** sind durch eine **Entscheidungsgrenze** getrennt
- Klassifikatoren unterscheiden sich in verschiedenen Aspekten wie **Linearität**, **Nichtlinearität**, **Aufwand für Training und Test**, etc. Einige anschauliche Vertreter sind der **nearest neighbour** und der **nearest mean Klassifikator**
- Aus Zweiklassen-Klassifikatoren können durch **Kombination** größere Aufgaben gelöst werden, z.B. Multiklassen-Probleme oder Detektion von Objekten in Bildern
- Wichtige **Anwendungsfelder** für Klassifikatoren sind Zeichenerkennung, Objekterkennung, Gesichtsdetektion, etc.

## Referenzen

---

- Weiterführende Literatur
  - [53] R.O. Duda, P.E. Hart und D.G. Stork: „Pattern Classification“. Wiley Interscience, 2nd edition, 2001.
  - [54] S. Theodoridis und K. Koutroumbas: „Pattern Recognition“, Academic Press Inc.(London) Ltd, 2nd edition, 2003.
  - [61] B. Schölkopf and A. J. Smola. Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond. MIT Press, 2002.
  - [63] C. Bishop. „Neural Networks for Pattern Recognition“, Oxford University Press, 1996.
  - [62] J. Shawe-Taylor and N. Cristianini. Kernel Methods for Pattern Analysis. Cambridge University Press, 2004.
  - [64] J. Schürmann. „Pattern classification: a unified view of statistical and neural approaches“, John Wiley & Sons, Inc, 1996.
  - [60] W. Kienzle, G.H. Bakir, M.O. Franz and B. Schölkopf: „Face Detection - Efficient and Rank Deficient“. Advances in Neural Information Processing Systems 17, 673-680. MIT Press, 2005.
  - [56] H. Burkhardt und B. Haasdonk: „Mustererkennung WS 02/03, ein multimedialer Grundlagenkurs im Hauptstudium Informatik“. Institut für Informatik, Universität Freiburg, 2003. (CDs werden verkauft), Online unter <http://lmb.informatik.uni-freiburg.de/lectures/mustererkennung/WS0304/material.html>
- Bildreferenzen
  - [58] Siemens VDO, <http://www.siemensvdo.com/topics/adas/traffic-sign-recognition/>
  - [59] Haasdonk, B., Keysers, D., „Tangent Distance Kernels for Support Vector Machines.“ ICPR 2002, International Conference on Pattern Recognition, 2002.