



Mustererkennung

— Prinzipien, Methoden, Anwendungen —

Gernot A. Fink

Technische Universität Dortmund
Fakultät für Informatik / Institut für Roboterforschung

28. November 2007

- ▶ Einleitung *Was ist Mustererkennung?*
- ▶ Prinzipien
- ▶ Methoden *... und Verarbeitungsschritte*
(Vorverarbeitung, Merkmalsextraktion, Segmentierung & Klassifikation)
- ▶ Ausblick *Multimodalität, mehr Mustererkennung?*



Mustererkennung - Was ist das?

Hauptsächliches Ziel: Wahrnehmungsleistungen des Menschen nachbilden

Achtung: Werden vom Menschen scheinbar mühelos erbracht, z.B.:

- ▶ Sprachbeherrschung
(z.B. spontansprachliche Kommunikation im Alltag)

Maschinen: Kontrollierte Aufnahmebedingungen

[d.h. kein Hörsaal!, Mikrofon am *Sprecher*]



- ▶ Visuelle Wahrnehmung
(z.B. Erkennung von Personen auf einer Party)

Maschinen: Kontrollierte Umgebung

[Beleuchtung, Objekte, eingeschränkte Dynamik]



Auch: Automatisierung von Entscheidungsprozessen
i.d.R. maschinelle Klassifikation von Meßdaten

Einleitung: Anwendungsgebiete von Mustererkennung

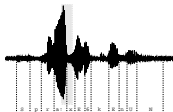
Schriftzeichen (maschinell oder handschriftlich):

- ▶ Digitale Erfassung von Dokumenten (OCR)
- ▶ Verarbeitung von Formularen (z.B. Überweisungsträgern)
- ▶ Automatische Sortierung von Post (= Lesen von Adressen)
- ▶ Texteingabe (insbes. auf PDAs, SmartPhones)



Sprachverarbeitung:

- ▶ Gesprochene Sprache: Mensch-Maschine-Interaktion, Diktiersysteme, automat. Telefondienste, ...
- ▶ Text: Übersetzung, Indizierung/Suche, ...



Medizinische Diagnostik (bzw. deren Unterstützung):

- ▶ Interpretation von EKG, EEG, ... (Zeitsignale)
- ▶ Segmentierung/Interpretation von Röntgen-, Ultraschall-, Magnetresonanz- und Computertomographiebildern



[NASAexplores]

Einleitung: Anwendungsgebiete II

Industrielle Anwendungen:

- ▶ Qualitätskontrolle (z.B. Inspektion von Bauteilen, Überwachung von Herstellungsprozessen)
- ▶ Fertigungsautomatisierung (z.B. Produktionsstraßen, Industrieroboter)



[WOLLIN]

Autonome Roboter / Serviceroboter

- ▶ Autonome Fahrzeuge (z.B. "Stanley", auch: militärisch)
- ▶ Dienstleistungen im häuslichen Bereich (insbes. Betreuung von Senioren)
- ▶ "Search and Rescue" im Katastrophenfall



[Stanford]

Automotive:

- ▶ Sicherheitssysteme (z.B. Kollisionsdetektion [Airbag])
- ▶ Fahrerassistenzsysteme (z.B. Einparkassistent, Spurassistent, Verkehrszeichenerkennung, automatisches Kolonnenfahren für LKW)

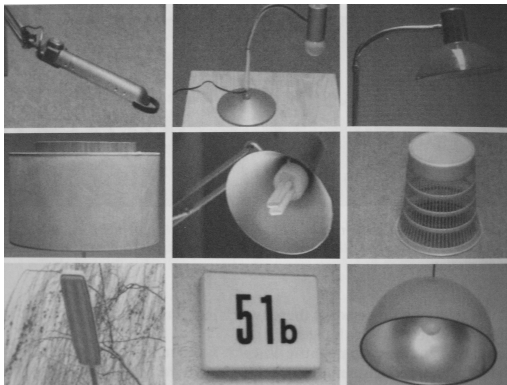


[Siemens VDO]



Einleitung: Warum ist Mustererkennung schwierig?

Kategorisierung: Was haben diese Objekte gemeinsam? ... bis auf eines!



[Jähne 2002]

- ⚡ Kategorien nicht notwendigerweise durch äußere Erscheinung definiert (obwohl i.d.R. von Menschen aufgrund Erscheinung *unterscheidbar* durch Einbeziehung von Kontext-/Weltwissen!)

... Warum ist Mustererkennung schwierig?

Segmentierung: Welche Bestandteile sind die "richtigen"?

Original Image



Alternative segmentations



Potential elementary segments, strokes, ...



⚡ Segmentgrenzen im Signal nicht markiert (z.B. Wörter in gesprochener Sprache)



Mustererkennung: Begriffsbestimmung

[nach Niemann 2003]

Problemkreis: Menge Ω aller (physikalisch) messbaren Größen/Funktionen eines (begrenzten) Anwendungsgebiets, z.B.

- ▶ Sprachsignal [einer gesprochenen Ziffernfolge] = Schalldruckpegel, 1-dim.
- ▶ RGB-Bild [eines Gesichts] = Intensität in 3 Frequenzbändern des elektromag. Spektrums

Muster: Ein Element des Problemkreises

Musterklassen ergeben sich durch (disjunkte) Zerlegung des Problemkreises in Mengen Ω_{κ}

z.B. $\Omega =$ Gesichtsbilder $\Rightarrow \Omega_{\kappa} =$ Bilder einer der κ -ten Person

Mustererkennung: Mathematisch-technische Aspekte der *automatischen* Verarbeitung und Auswertung von Mustern

- ▶ Musterklassifikation: (einfaches) Muster $\rightarrow \Omega_{\kappa}$
- ▶ Musteranalyse: (komplexes) Muster \rightarrow symbolische Beschreibung



Mustererkennung: Anforderungen

[nach Niemann 2003]

Postulat 1: Zur Gewinnung von Informationen über einen Problembereich liegt eine *repräsentative Stichprobe* ω vor.

(ω = Menge von Mustern + ggf. Annotation)

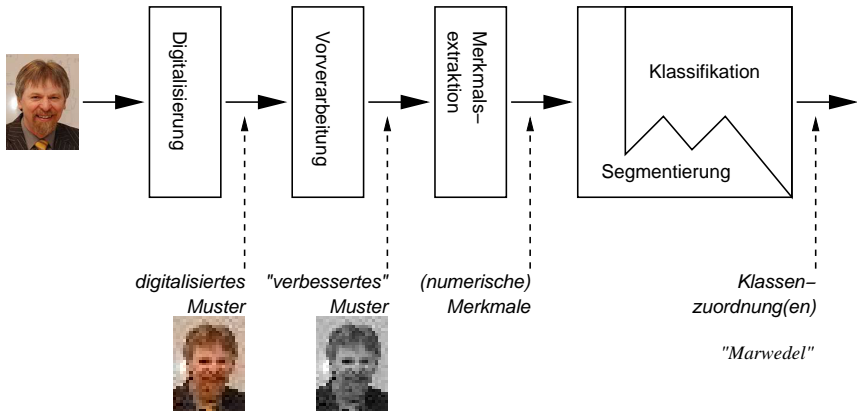
Postulat 2: Ein (einfaches) Muster besitzt **Merkmale**, die charakteristisch für seine Klassenzugehörigkeit sind.

Postulat 3 (*Kompaktheitshypothese*): Die Merkmale von Mustern einer Klasse bilden einen (eingermaßen) kompakten Bereich im Merkmalsraum (auch: Bereiche verschiedener Klassen sind getrennt).

Postulat 4: Muster sind *ähnlich*, wenn sich ihre Merkmale nur wenig unterscheiden.

Hinweis: Bei komplexen Mustern gelten Aussagen analog zu 2 und 4 für einfache Bestandteile der Muster.

Aufbau eines ME-Systems



Digitalisierung

Automatische Verarbeitung von Mustern ausschließlich mit *digitalen* Rechnern
⇒ digitale Repräsentation erforderlich

Digitalisierung = Abtastung + Quantisierung

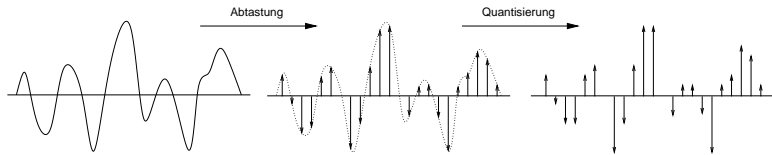
Abtastung: Messung der *analogen* Funktionswerte an

diskreten (Zeit-)Punkten (äquidistant!) ⇒ Folge von *analogen* Meßwerten

Abtasttheorem: Kein Informationsverlust, wenn Abtastrate höher als doppelte Grenzfrequenz (des *bandbegrenzten* Musters)

Quantisierung: Abbildung der analogen Meßwerte auf eine *endliche* Menge *diskreter* Werte

Approximation erzeugt Fehler (*Quantisierungsrauschen*)





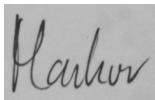
Vorverarbeitung

Ziel: "Verbesserung" eines Musters für nachfolgende Verarbeitungsschritte
⇒ Reduktion unerwünschter Signalanteile / Variabilität

Verfahrensklassen:

- ▶ *Schwelwertoperationen* (z.B. Binarisierung)

Original



Binarized image (Otsu's method)



- ▶ *Filterung* (z.B. Rauschunterdrückung)
- ▶ *Normierung* (z.B. Lage, Größe, ...)

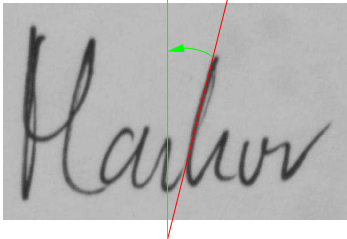
Problem: Wie bewerten? Beeinflußt *gesamte* weitere Verarbeitung!

Achtung: Kaum Standardmethoden! Höchst anwendungsabhängig!

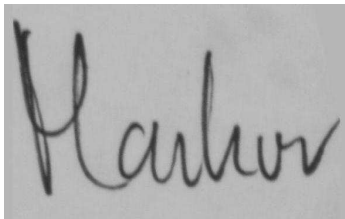
Vorverarbeitung II

Beispiel: Normierung der Schriftnéigung

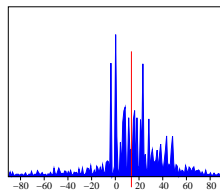
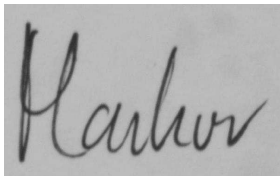
Original



Corrected Slant



Grundlage: Schätzung des Neigungswinkels im Gradientenhistogramm





Merkmalsextraktion

Ziel: Repräsentation charakteristischer Mustereigenschaften zum Zweck der *Unterscheidung* von Musterklassen

Auch: Unterscheidungs*irrelevante* Information nicht repräsentiert

⇒ (i.d.R.) Reduktion der Datenmenge

Vorherrschendes Prinzip: Transformation eines Musters $\underline{f} \in \mathbb{R}^m$ in einen Merkmalsvektor $\underline{c} \in \mathbb{R}^n$, $n \ll m$:

$$\underline{c} = T\{\underline{f}\}$$

Bewertung: Optimale Merkmale minimieren Fehlerrate des Klassifikators

Problem: Im Allgemeinen keine *systematischen* Verfahren zur Bestimmung *optimaler* Merkmale bekannt!

⇒ Methoden sind – mehr oder weniger – heuristisch!

(d.h. beruhen auf Erfahrung / Kreativität des Entwicklers)

Verfahrensklassen:

- ▶ Reihenentwicklungen (“technisch bewährt”)
- ▶ Analytische Methoden (vereinfachtes Optimierungskriterium)
- ▶ Anwendungsspezifische Techniken



Merkmalsextraktion: Reihenentwicklungen

Prinzip: Entwicklung eines Musters $\underline{f} \in \mathbb{R}^m$ nach einer orthogonalen Basis $\Phi = \{\underline{\phi}_\nu\}$ (Basis Φ spannt \mathbb{R}^m auf)

$$c_\nu = \langle \underline{f}, \underline{\phi}_\nu \rangle = \underline{\phi}_\nu^T \underline{f} = \sum_{j=0}^{m-1} \phi_{\nu j} f_j \quad \Rightarrow \quad \underline{c} = (c_1, c_2, \dots, c_n)^T$$

- ▶ Transformation ist *linear*: $\underline{c} = \Phi \underline{f}$
- ▶ Rekonstruktion $\hat{\underline{f}} = \Phi^T \underline{c}$ minimiert (mittleren) Approximationsfehler

Beispiele:

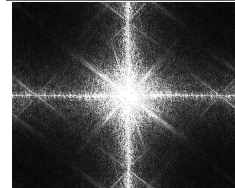
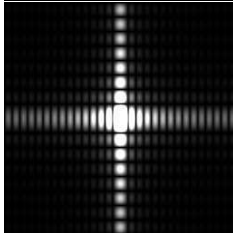
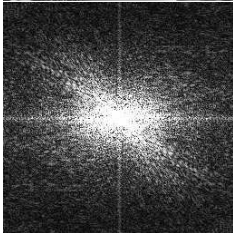
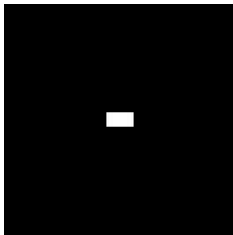
- ▶ *Diskrete Fourier-Transformation (DFT)*
Frequenzbereichsrepräsentation (Amplitude + Phase)

$$F(k) = \sum_{m=0}^{M-1} f(m) e^{-\frac{i2\pi km}{M}} \quad \text{mit} \quad e^{ix} = \sin x + i \cos x$$

⇒ Basisfunktionen: sin/cos mit wachsender Frequenz

- ▶ *Diskrete Cosinus-Transformation (DCT)* → Bildkompression (JPEG)

Merkmalsextraktion: DFT



Beispielbilder (oben) und zugehöriges Fourier-Betrags-Spektrum (unten)



Merkmalsextraktion: Analytische Methoden

Idee: Optimierung eines *vereinfachten* Gütekriteriums a.d. Basis der Verteilung der Merkmalsvektoren im Merkmalsraum

Beschreibung durch *Streuungsmatrizen*, z.B. Gesamtstreuungsmatrix:

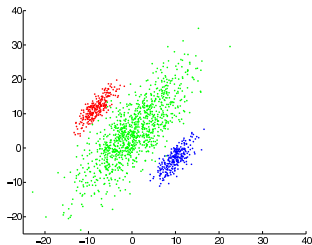
$$S_T = \sum_{j=1}^N (\underline{f}^j - \underline{\mu})(\underline{f}^j - \underline{\mu})^T$$

Auch: Intra- und Interklassenstreuung S_w und S_b

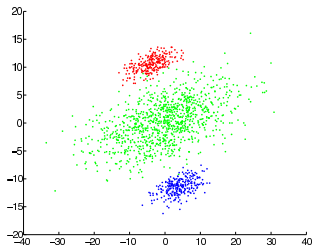
PCA (Hauptachsentransformation): Optimale Beschreibung der *Gesamtvarianz* der Daten durch die (reduzierte) Merkmalsrepräsentation
⇒ Transformationsmatrix = Eigenvektoren von S_T zu größten Eigenwerten

LDA (lineare Diskriminanzanalyse): Transformation zur (näherungsweise) Verbesserung der *Klassentrennbarkeit* (große Streuung zwischen Klassen bei gleichzeitig konstanter Intraklassenstreuung)
⇒ Transformationsmatrix = Eigenvektoren von $S_w^{-1} S_b$ (zu größten Eigenwerten)

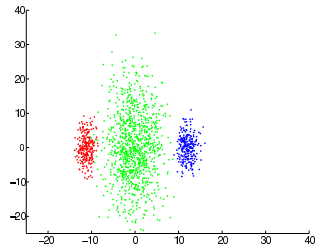
Merkmalsextraktion: PCA & LDA



PCA



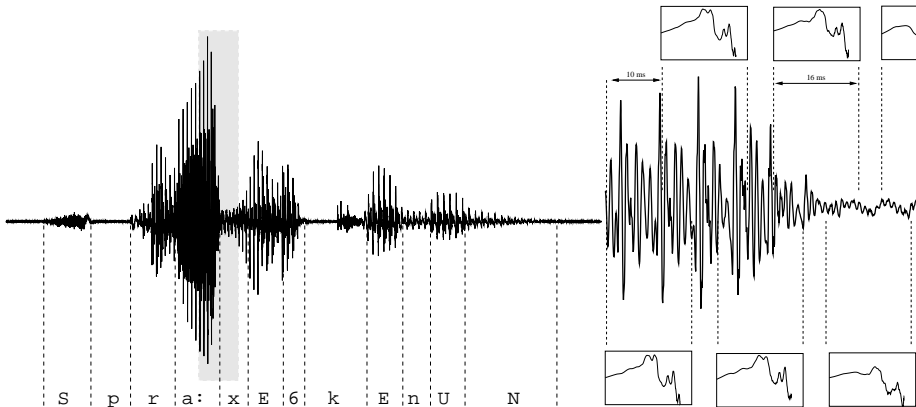
LDA



Merkmalsextraktion V

Beispiel: Kurzzeitanalyse eines Sprachsignals \approx Berechnung eines Modellspektrums

Signal \rightarrow DFT \rightarrow $|\cdot|$ \rightarrow Mel \rightarrow \log \rightarrow DCT \rightarrow Mel-Cepstrum



Segmentierung

Ziel: Unterteilung eines komplexen Musters in kleinere, *bedeutungstragende* Bestandteile

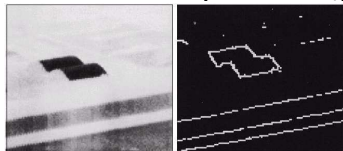
Problem: Ohne eingebettete Klassifikation (\rightarrow *später*) keine allgemeinen Grundlagen!

Bekannte Domäne: *Bildverarbeitung*

- ▶ **Regionen**segmentierung
a.d. Basis eines
Homogenitätskriteriums
(Region = homogener Bildbereich)



- ▶ **Kontur**segmentierung
a.d. Basis von Kanten bzw.
Diskontinuitäten



[Tel Aviv University]

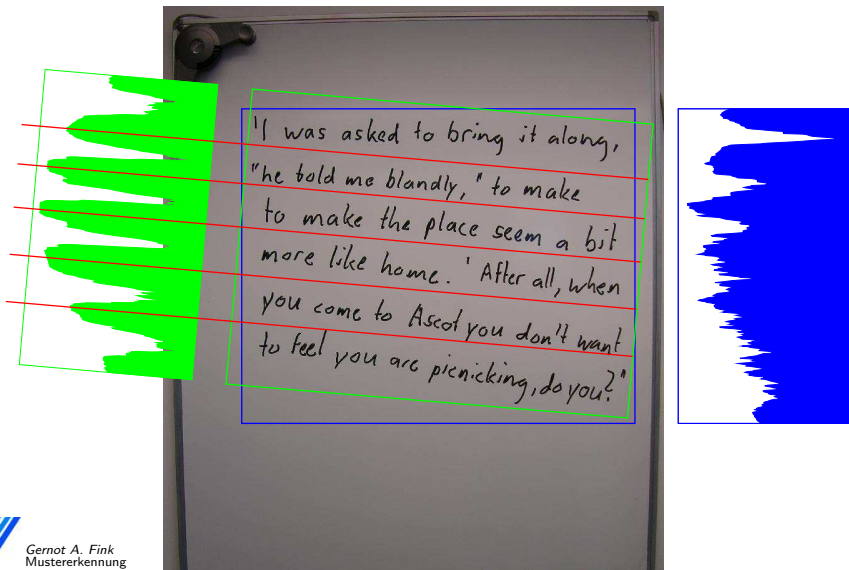
[Gonzalez & Woods, 2002]

Generell: Verfahren überwiegend heuristisch, anwendungsspezifisch



Segmentierung II

Beispiel: Zeilenextraktion aus aus Dokumentabbildern (hier: Handschrift)





Klassifikation

Ziel: Zuordnung eines Musters *als Ganzes* zu einer Klasse (\approx Kategorie)

Vorgehen: Klassifikation definiert eine Abbildung der Merkmalsvektoren \underline{c} auf einen Klassenindex κ :

$$\underline{c} \rightarrow \kappa \in \{1, \dots, K\} \quad \text{oder} \quad \underline{c} \rightarrow \kappa \in \{0, \dots, K\}$$

Zuordnung zur Klasse Ω_0 bedeutet *Rückweisung* des Musters
(z.B. Postanschriftenleser: Adresse manuell auswerten; biometrische
Zugangskontrolle: ungültiger Versuch, wiederholen)

Methoden existieren in großer Vielfalt

- ▶ Einfachstes mögliches Verfahren: Nächster-Nachbar-Klassifikator
- ▶ Viele (im weitesten Sinne) probabilistische Verfahren
(Bayes-Klassifikator, Neuronale Netze, Markov-Modelle, ...)
- ▶ Allen modernen Methoden gemeinsam: Parameter werden auf Beispieldaten optimiert ($\hat{=}$ "gelernt")



Klassifikation: Nächster Nachbar

Voraussetzung: Annotierte Stichprobe

(d.h. Menge von Mustern \underline{f} [repräsentiert als Merkmalsvektoren \underline{c}] mit bekannter Klassenzugehörigkeit)

Methode:

1. Für alle Merkmalsvektoren \underline{c}
Bestimme Abstand $d(\underline{c}, \underline{c})$ zum neuen Muster \underline{c}
Abstand sei minimal für index j
2. Ordne neues Muster der Klasse ω_λ zu, in der auch \underline{f} (d.h. Muster mit min. Abstand) liegt

Grundlage: Abstands- oder Ähnlichkeitsmaß zwischen Merkmalsvektoren

Gebräuchliche Lösung: Euklidischer Abstand $d(\underline{x}, \underline{y}) = \|\underline{x} - \underline{y}\|$
(ex. Vielzahl weiterer Abstandsmaße!)

Problem: Aufwand des Verfahrens (Speicherung, Suche)

Aber: Fehlerrate kann *mit beliebigem Verfahren* höchstens noch *halbiert* werden!



Bayes-Klassifikator

Idee: Bestimmung des *optimalen* Klassifikators

- ▶ *Hier:* Minimierung des Klassifikationsrisikos (d.h. der Kosten)
- ▶ Vereinfachte *Annahme:* (konstante) Kosten für Fehler, erzwungene Entscheidung (keine Rückweisung)

Optimaler Klassifikator entscheidet sich für die Klasse Ω_λ mit maximaler *a posteriori* Wahrscheinlichkeit (Bayes-Klassifikator):

$$\lambda = \operatorname{argmax}_{\kappa} P(\Omega_{\kappa} | \underline{c}) = \operatorname{argmax}_{\kappa} \frac{P(\Omega_{\kappa}) p(\underline{c} | \Omega_{\kappa})}{p(\underline{c})} = \operatorname{argmax}_{\kappa} P(\Omega_{\kappa}) P(\underline{c} | \Omega_{\kappa})$$

Erforderlich:

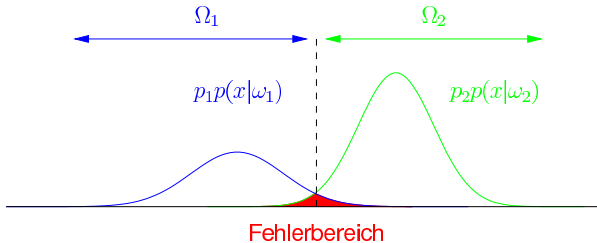
1. *A priori* Wahrscheinlichkeiten $P(\Omega_{\kappa})$ der Klassen
⇒ auf Stichprobe auszählen
2. Repräsentation der klassenbedingten Dichten $p(\underline{c} | \Omega_{\kappa})$ der Merkmalsvektoren



Bayes-Klassifikator II

Wie klassenbedingte Dichten $p(\underline{c}|\Omega_{\kappa})$ modellieren?

Einfach: Normalverteilungen, hier: Zwei-Klassen-Problem



⇒ setzt *unimodale* Dichten voraus!

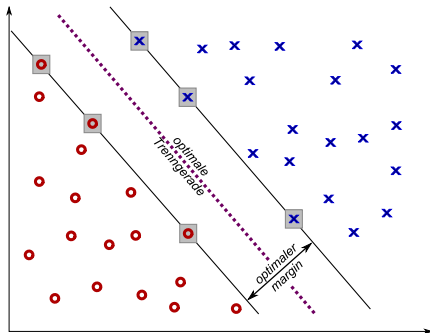
Allgemein: Mischverteilungen (a.d. Basis von Gauss-Dichten)

$$p(\mathbf{x}|\omega_{\kappa}) = \sum_{i=1}^{M_{\kappa}} c_{\kappa,i} \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_{\kappa,i}, \mathbf{K}_{\kappa,i})$$

⇒ beliebige Dichten können *approximiert* werden!

Support-Vektor-Maschinen (SVM)

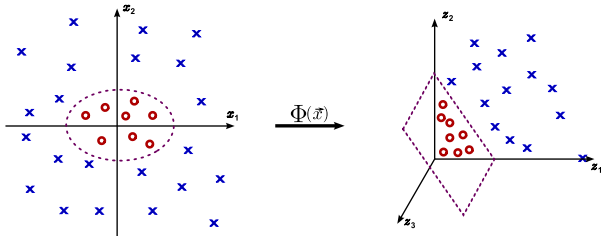
Grundprinzip: Klassentrennung durch Hyperebene im Merkmalsraum



- ▶ Maximierung der *margin* (= Abstand der [zwei] Klassen) von Trennebene
- ▶ Vektoren mit minimalem Abstand = Support-Vektoren

Support-Vektor-Maschinen II

Idee: Sofern eine *lineare* Trennung der Klassen (Hyperebene) nicht möglich, transformiere Merkmalsrepräsentation in höherdimensionalen Raum



Vereinfachung durch Wahl eines geeigneten *Kernels*

⇒ explizite Berechnung der hochdim. Repräsentation vermieden!

Eigenschaften:

- ✓ Sehr leistungsfähiges Verfahren
- ✓ Kann auf Mehrklassenprobleme erweitert werden
- ⚡ Aufweichung des Prinzips i.d. Praxis erforderlich (*slack Variable*)
- ⚡ Parameteroptimierung extrem aufwendig

Segmentierung & Klassifikation

Problem: Abfolge *Segmentierung* \Rightarrow *Klassifikation* vielfach problematisch

Original Image



Alternative segmentations



⋮



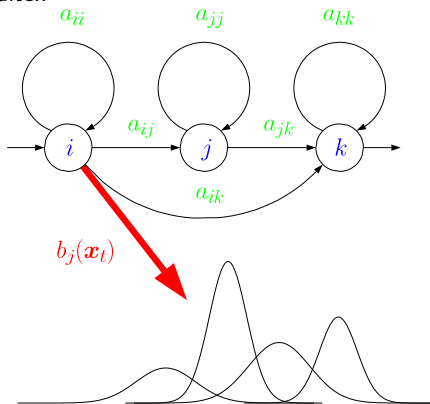
Potential elementary segments, strokes, ...



Besser: Segmentierung mit integrierter Klassifikation \rightarrow *segmentierungsfrei*
Bekanntestes Verfahren für *sequenzhafte* Muster: Hidden-Markov-Modelle

Segmentierung & Klassifikation: Hidden-Markov-Modelle

Prinzip: Hidden-Markov-Modell [HMM] \approx generierender endlicher Automat mit statistischem Verhalten



- ▶ Zustände $i, j \in \{1, 2, \dots\}$
- ▶ Übergangswahrscheinlichkeiten a_{ij}
- ▶ Ausgabewahrscheinlichkeiten $P(O_t|j)$ bzw. -dichten $b_j(\mathbf{x})$



Segmentierung & Klassifikation: HMMs II

Problem: Wie *generatives* Modell zur *Erkennung* von Mustern einsetzen?

Annahme: Muster (z.B. Sprache, Handschrift, biolog. Sequenzen) werden von *prinzipiell vergleichbarem* Modell erzeugt
⇒ dessen Verhalten mit HMM nachbilden
⇒ Dann Rückschlüsse anhand des HMMs möglich

Fragestellungen:

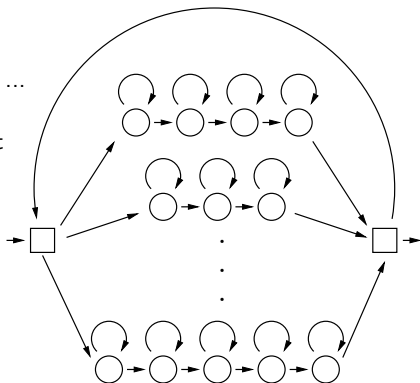
- ▶ **Bewertung:** *Wie gut beschreibt Modell Daten?* (\approx Klassifikation)
⇒ Generierungswahrscheinlichkeit berechnen
- ▶ **Parameterschätzung:** *Wie optimales Modell für best. Daten finden?*
⇒ ausgehend von initialem Modell interaktiv verbessern
- ▶ **Dekodierung:** ...

Problem 2: *Wie HMM zur Segmentierung verwenden?*

Segmentierung & Klassifikation: HMMs III

Ziel: Segmentierung

- ▶ Elementare Einheiten: Zeichen, Laute, ...
- ▶ Wörter o.Ä. durch Verkettung gebildet
- ▶ Lexikon/Inventar = Parallelschaltung [Pseudo-Zustände zur Kantenbündelung]
- ▶ Modell für Folge beliebiger Einheiten aus dem Inventar: Rückkopplung



⇒ Dekodierung des Modells erzeugt Segmentierung
(d.h.. Bestimmung der optimalen
Zustands/Einheiten-Sequenz)



Multimodalität

Multimodale Mustererkennung: Daten *unterschiedlicher* Sensoren werden verarbeitet

[Im Gegensatz zu z.B. Stereo-Bildverarbeitung (2 Bilder → Tiefe)]

Problem: Wie Erkennungsergebnisse kombinieren bzw. Ergebnisse unter Einbeziehung *aller* verfügbaren Modalitäten erzeugen?

⇒ keine *allgemeinen* Verfahren bekannt!

Beispiele:

- ▶ Kombination von Sprache und Gestik bei der MMI
- ▶ Multimodale Sprecherlokalisierung (Audio + Video)



Mehr Mustererkennung?

Vorlesung Mustererkennung

- ▶ MA-Basismodul bzw. Hauptstudium [4V+2Ü, jeweils im WS]
- ▶ Inhalt: *Was Sie heute gehört haben in ordentlich.*

Vorlesung Computer Vision

- ▶ MA-Vertiefungsmodul bzw. Hauptstudium [2V+2Ü, jeweils im WS]
- ▶ Inhalt: Fokus auf speziellen Bildverarbeitungstechniken, keine Klassifikation

Vorlesung Spracherkennung

- ▶ MA-Vertiefungsmodul bzw. Hauptstudium [3V+1Ü, SS 2008]
- ▶ Inhalt: Schwerpunkt auf HMMs für Segmentierung+Klassifikation

Diplom / Studienarbeiten / Hilfskrafttätigkeiten

www.is.irf.de → Career