Inhalt

1. Einleitung

Die Vorlesung beschäftigt sich mit Teilgebiet der:

Mustererkennung : mathematisch-technischer Aspekt von Perzeption

Perzeption : Erfassen, Auswerten und Interpretieren von Sinneseindrücken

Perzeption ist für die Aktivitäten aller Lebewesen wichtig, aber auch für technische Systeme, die (immer) mit der Umwelt/Umgebung interagieren.

Bei der Realisierung solcher perzpetiver Systeme steht die Spannweite von Nachbildung

- eines biologischen Vorbildes bis zur
- rein technisch-ingenieurwissenschaftlicher Sichtweise.

1. Einleitung

In vielen Bereichen zeigen biologische visuelle Systeme erstaunliche Fähigkeiten, die automatisch i.d.R. nur sehr unzulänglich erreicht werden können.

Bei quantitativen Auswertungen und repititiven Aufgaben sind technische Systeme typischerweise überlegen.

In der Vorlesung: im wesentlichen technisch-ingenieurwissenschaftlicher Standpunkt

1.1 Beispiele

1. gesprochener Text:
   - Identifizieren/Authentifizieren der Person
   - möglichst exakte Transspriktion (Diktiersystem)
   - Auskunftsdialog: Erkennung aller Wörter nicht nötig, die Bedeutung muß erkannt werden
2. handschriftlicher Name:
   - Name des Adressaten (Postanschriftenleser)
   - Charaktereigenschaften (Graphologie)
   - gültig/ungültig (Zugangskontrolle)
3. gescannte Formulare: Geldbeträge, Empfänger
4. Bildsequenz (Vidoesequenz) der Umgebung:
   - Navigation einer autonomen Systems (z.B. Hindernisvermeidung, Verfolgung)
   - Erkennen von Teilen und deren Lage (Montage)
5. - CT/MR-Bilder: erkennen von Organen, deren Größe, Form
   - EKG: Diagnose

1.1 Beispiele

6. Satellitenbilder:
   - Waldschäden erfassen
   - Umweltverschmutzung
   - Kartographie
   - militärische Aufklärung
7. chemische Analyse des Gehalts an verschiedenen Fettsäuren in Olivenöl: Herkunftsregion

1.2 Definitionen, Begriffe

Perzpetion hat zum Gegenstand die Eindrücke der Umwelt

Umwelt : $U = \{{}^{\rho}\vec{b}(\vec{x}) | \rho = 1, 2, \ldots\}$
(Menge aller meßbaren Größen; $\vec{x}$ ist Zeit, Raum und $\vec{b}$ sind Messungen physikalischer Eigenschaften, z.B. Lautstärke)

Problemkreis : $\Omega = \{{}^{\rho}\vec{f}(\vec{x}) | \rho = 1, 2, \ldots\} \subset U$
(Menge von Mustern des gleichen Typs)

Muster : Elemente eines Problemkreises

$$\vec{f}(\vec{x}) = \begin{bmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ f_m(x_1, \ldots, x_n) \end{bmatrix}$$

1.2 Definitionen, Begriffe

Beispiele für Muster haben wir oben schon kennengelernt.

- Farbfernsehbild
- Sprache
- S/W – Foto

Muster werden in zwei Typen unterschieden:

- einfache Muster (direkt klassifizierbar)
- komplexe Muster

1.2 Definitionen, Begriffe

Klassifikation von (einfachen) Mustern, ist die Zuordnung eines Musters zu einer von k erlaubten Klassen Ω_λ
Beispiel: Buchstabe, isoliert gesprochenes Wort

Musterklasse od. Klasse : Zerlegung des Problemkreises in disjunkte Klassen:

$$\begin{aligned} \Omega_\lambda &\neq 0 \\ \Omega_\lambda \cap \Omega_\kappa &= \emptyset \quad \text{für } \lambda \neq \kappa \\ \bigcup_{\lambda=1}^{k} \Omega_\lambda &= \Omega \end{aligned}$$

Erweiterung um Rückweisungsklasse $\Omega_0 \Rightarrow \bigcup_{\lambda=0}^{k} \Omega_\lambda = \Omega$

Einfache Muster :
– dem Anwender genügt der Klassenname
– die Klassifikation ist als Ganzes möglich

1.2 Definitionen, Begriffe

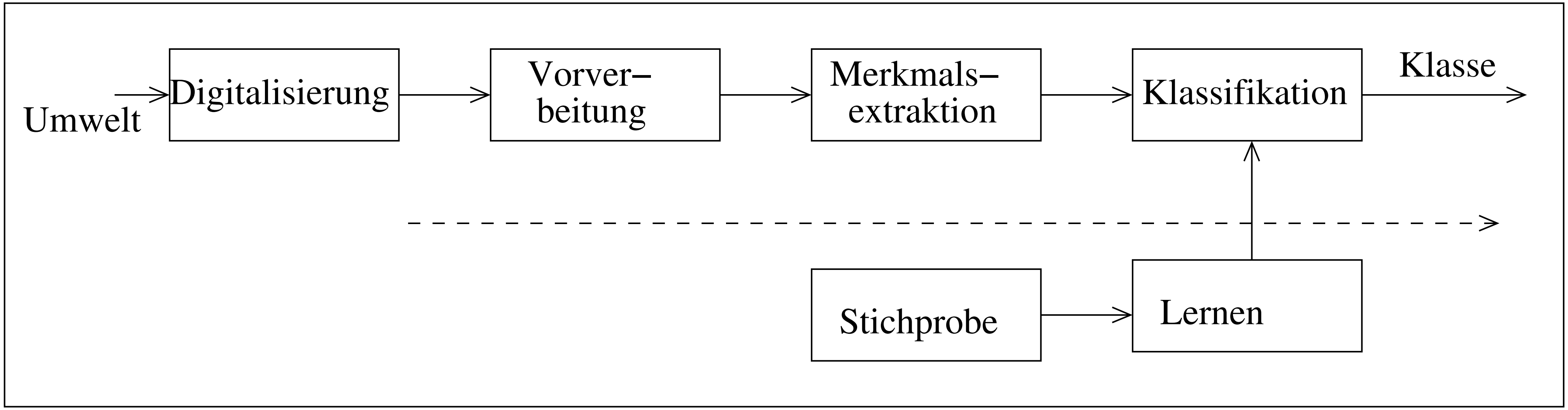
Analyse von (komplexen) Mustern erzeugt eine symbolische Beschreibung

Typen von symbolischen Beschreibungen:
1) ausführliche Auflistung der Teile und Relationen
2) interessante Teile
3) Beschreibung von Veränderungen
4) Klassennamen

Komplexes Muster : liegt vor, wenn der Anwender mehr als einen Klassennamen benötigt oder die Klassifikation als Ganzes nicht möglich ist

1.3 Aufbau von Klassifikationssystemen

In der Vorlesung geht es um Klassifikation von (einfachen) Mustern (bemerke nochmals: die Unterscheidung ist fließend!)



1.4 Grundsätzliche Vorgehensweise

Bei allen Unterschieden zwischen verschiedenen Ansätzen gibt es einige gemeinsame Prinzipien (Niemann 83):

Postulat 1: es steht eine repräsentative Stichprobe ω zur Verfügung

$$\omega = \left\{ {}^{1}\vec{f}(\vec{x}), \ldots, {}^{N}\vec{f}(\vec{x}) \right\} \subset \Omega$$

- $\omega \subset \Omega$: da alles andere irrelevant
- in der Regel klassifizierte Stichprobe
- repräsentativ: wichtig, aber schwer zu erreichen, sogar schwer zu beurteilen
- Frage nach N (Größe der Stichprobe)
  - abhängig von statistischen Eigenschaften des Problemkreises
  - oft N eher zu klein, als groß (wegen Aufwand)

Für die Klassifikation:

Postulat 2: Ein einfaches Muster besitzt Merkmale, die für die Zugehörigkeit zu seiner Klasse charakteristisch sind.

- numerische Merkmale: Vektor reeler Zahlen
- symbolische Merkmale: Kette von Symbolen

Merkmale werden aus dem Muster durch die **Merkmalsextraktion** gewonnen, für numerische Merkmale:

$$\varphi : \varphi\left(\vec{f}\,(\vec{x})\right) = \vec{c} = \begin{pmatrix} c_1 \\ \vdots \\ c_M \end{pmatrix}$$

Dabei sollte die Dimension der Merkmalsvektoren M möglichst klein sein, u.a. da von dieser Größe der mathematische Aufwand bei der Klassifikation abhängt.

Postulat 3: Die Merkmale für Muster einer Klasse bilden einen einigermaßen kompakten Bereich. Die Bereiche für verschiedene Klassen sind getrennt.

- geeignete Merkmale finden ist (oft) schwierig
- die Klassifikation wird relativ gut beherrscht
- lokale Optimierung gewährleistet i.A. nicht eine Optimierung des gesamten Systems

1.4 Grundsätzliche Vorgehensweise

Für die Analyse:

Postulat 4: Ein komplexes Muster besitzt einfachere Bestandteile, die untereinander eine bestimmte Beziehung haben. Das Muster läßt sich in diese Bestandteile zerlegen.

Postulat 5: Ein komplexes Muster eines Problemkreises hat eine bestimmte Struktur:

1. nicht jede Anordnung von Bestandteilen kommt vor
2. es genügen relativ wenige einfachere Bestandteile zur Darstellung eines Musters

Fragen:

- welche einfacheren Bestandteile ?
- Technik zur Repräsentation und Analyse der strukturellen Eigenschaften
- Systemstruktur (z.B. Black-Board- oder Datenbankarchitektur)

1.4 Grundsätzliche Vorgehensweise

Klassifikation und Analyse haben einiges gemeinsam, Trennlinie nicht fest

Postulat 6: Zwei Muster sind ähnlich, wenn ihre Merkmale oder einfacheren Bestandteile sich wenig unterscheiden