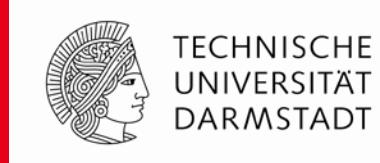


# **Computer Vision – Erkennung**

**Visual Computing**  
**Winter Semester 2018-2019**



**Prof. Dr. A. Kuijper**

Mathematical and Applied Visual Computing (MAVC)  
Graphisch-Interaktive Systeme (GRIS)  
Fraunhofer IGD  
Fraunhoferstrasse 5  
D - 64283 Darmstadt

E-Mail: [office@gris.tu-darmstadt.de](mailto:office@gris.tu-darmstadt.de)  
<http://www.gris.tu-darmstadt.de>  
<https://www.mavc.tu-darmstadt.de>

# Semesterplan



Datum	
26. Okt	Einführung + <u>Visual Computing</u>
02. Nov	<u>Wahrnehmung</u>
09. Nov	Objekterkennung und Bayes
16. Nov	Fourier Theorie
23. Nov	Bilder
30. Nov	Bildverarbeitung
07. Dez	Grafikpipeline & Eingabemodalitäten & VR+AR
14. Dez	Transformationen & 2D/3D Ausgabe
18. Jan	3D-Visualisierung
25. Jan	X3D – 3D in HTML
01. Feb	Informationsvisualisierung
08. Feb	Farbe
15. Feb	User Interfaces + Multimedia Retrieval
07. Mrz	Klausur





# Warum Computer Vision?

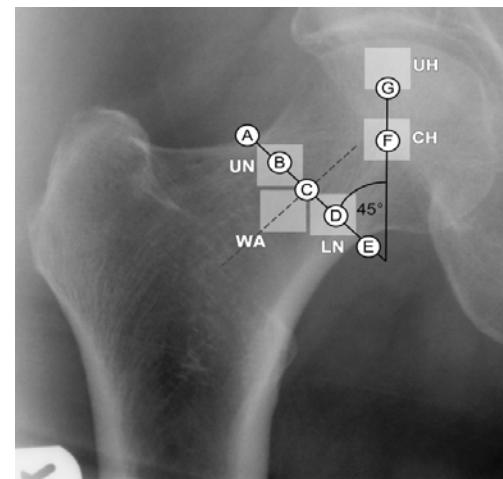
- Wissenschaft
  - Grundlagen der Wahrnehmung: Wie sehen WIR (Menschen)?
  - Erforschen eines computergestützen Modells des menschlichen Sehens
- Ingenieurwesen
  - Wie baut man Systeme, die die Welt erkennen?
  - Lösen realer Anwendungsfälle: Autos, die Fußgänger erkennen
- Anwendungsgebiete
  - Medizinische Bildverarbeitung (Diagnosen unterstützen, Visualisierung)
  - Überwachung (an Flughäfen, Bahnhöfen, ...)
  - Unterhaltung (kamerabasierte Spiele [EyeToy, Kinect, ...])
  - Computergrafik (bildbasiertes Rendering, Computer Vision zur Unterstützung realistischer Grafik)
  - ...

# Einige Anwendungen

- Erkennung von Nummernschildern
  - London Congestion Charge („Staugebühr“)
  - [http://www.cclondon.com/  
imagingandcameras.html](http://www.cclondon.com/imagingandcameras.html)
  - [http://en.wikipedia.org/wiki/  
London\\_congestion\\_charge](http://en.wikipedia.org/wiki/London_congestion_charge)



- Überwachung
  - Gesichtserkennung
  - Flughafensicherheit
    - Tracking von Menschen
    - Scannen von Gepäck



- Medizinische Bildverarbeitung
  - (Halb) automatische Segmentieren und Auswerten von Bildern



DARMSTADTER ECHO • Freitag, 27. April 2012

11

# DARMSTADT

ECHO

## Gezähmte Bestien

Von Stefan Benz

Nicht alles, was wir zum Kauen gern haben, verpesten wir auch. Das gilt für Mensch und Tier. Wir können das aus Mythen und Märchen. Eine Wölfe hat Romulus und Remus gesagt. Zuhörer und Zuhörer waren es, die im Vergleich zur Mauer-WG gründeten. Aus allen Schnitten lernen wird der Radfahrer und beweisen sozusagen, als stammt.

Dannnoch ist Verhältniszucker immer wieder erstaunlich, wenn die Kreatur ihrer Natur zuwiderhandelt. Wie bei den legendären Löwen, die seit 2002 mehrere kleine Antilopen adoptiert. Jetzt staunt die Nachwelt über eine Jung-Dame, die in Großsegelpark

Saaching vier Überwälzen ausbrachte und sie zu füttern, statt sie zu fressen.

Was verleiht den Räubern daran den Appetit? Ist das zu stecken? Und warum springt dieses Verhältnis so ungern vom Tier auf den Menschen über? Frauen sollten aufpassen, ob ihr Mann plötzlich mit einer Grille schwärmt. Wenn der ein „Wille“ Jäger nur noch handelt auf den Sofakissen hin und her, ist nicht dann magisch, dass er darunter seine Wandvorlese gebracht. Statt sie zu essen. Weil er von dem Elternteil voller Kündnis ist, der nie den Verdacht hatte, dass die gezähmte Hausbestie tatsächlich mit der Pizza kuschelt.



Stummer Späher: Überwachungskameras sieht man vielerorts in der Stadt, etwa auf der Mathildenhöhe, auf dem Luisenplatz und dem Willy-Brandt-Platz (von links)

FOTOS: DALES VOLK

**ACHTUNG!!**  
**freetime**  
**Aktionswochenende**  
Heute 27.4.2012 und Samstag 28.4.2012  
(nicht wie angekündigt: Samstag/Sonntag,  
Sonntag geschlossen)  
**freetime Fahrräder und Zubehör**  
Tel. 06151/871507

## Piraten fragen: Wer filmt hier eigentlich?

**VIDEOÜBERWACHUNG** Verzeichnis der Kameras an öffentlichen Plätzen entsteht im Internet – Aufruf zur Mitwirkung

# Wetzlar: Optikparcours

<http://de.wikipedia.org/wiki/Optikparcours>



Überwachung? ☺



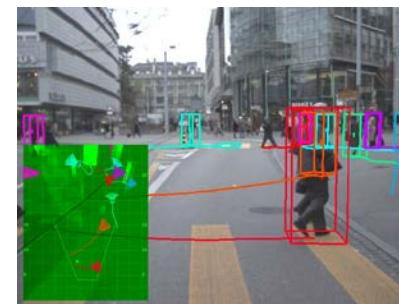
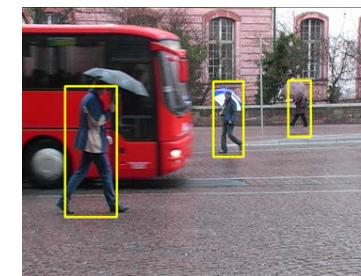
# Einige Anwendungen

## ■ Automobilindustrie

- Heute schon Produkt (bzw. produktreif):
  - Warnung bei Spurwechsel
  - Warnungen bezüglich des toten Winkels
  - Einparkhilfe
  - Warnungen kurz vor einem Zusammenstoß sowie automatisches Bremsen
  - ...



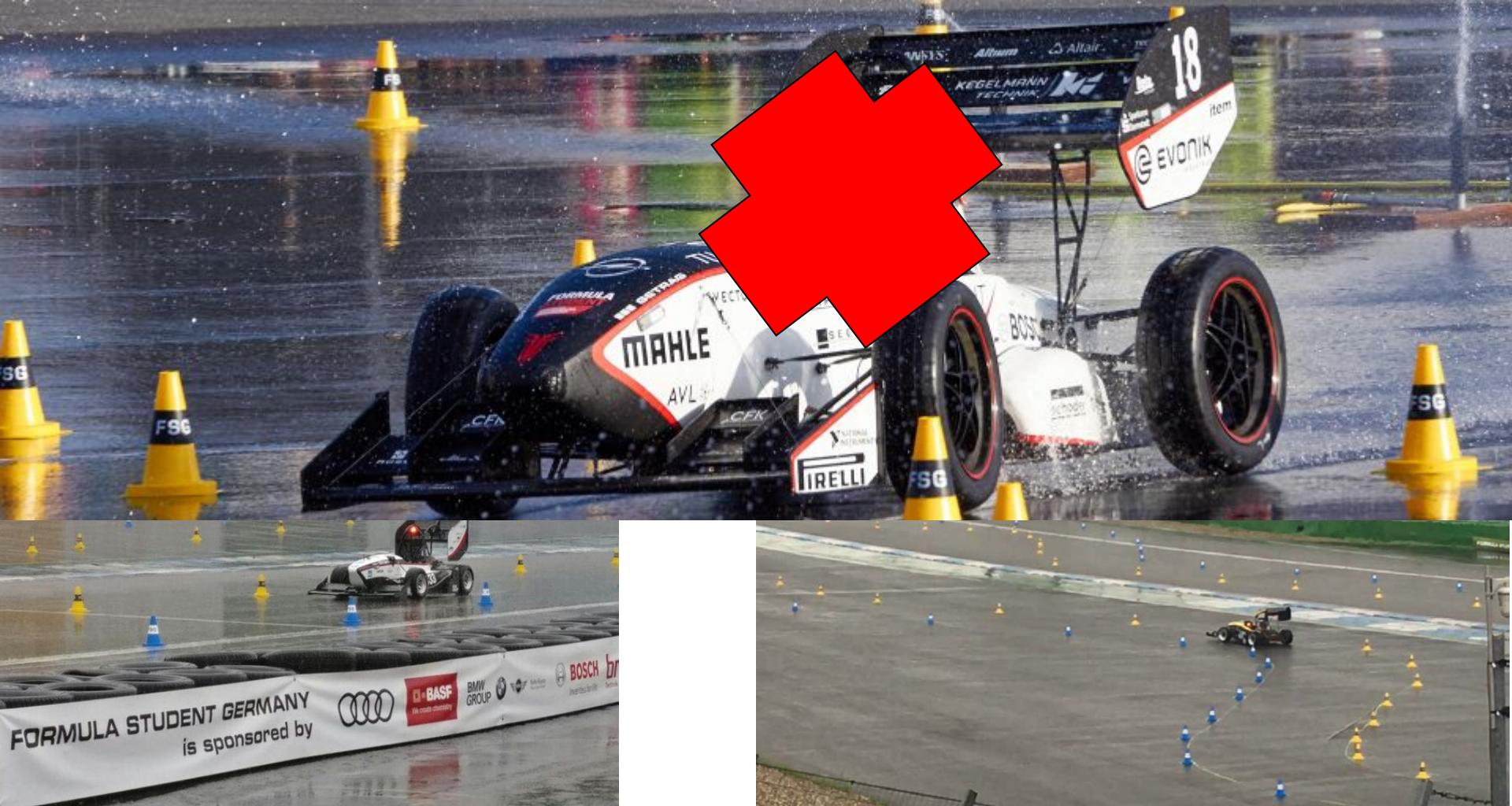
- In naher Zukunft:
  - Bessere Detektion von Fußgängern
  - Wetterkonditionen
  - ...



# Autonomes Fahren: <https://dart-racing.de/>



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT





# Überblick

## ■ Computer Vision

- **Objekterkennung und Objektdetektion als Fallbeispiel**
- **Motivation heutiger Ansätze**

## ■ Bayes Decision Theory

- Verwendung von Merkmalen zur Objekterkennung
- Naive Bayes Klassifikator

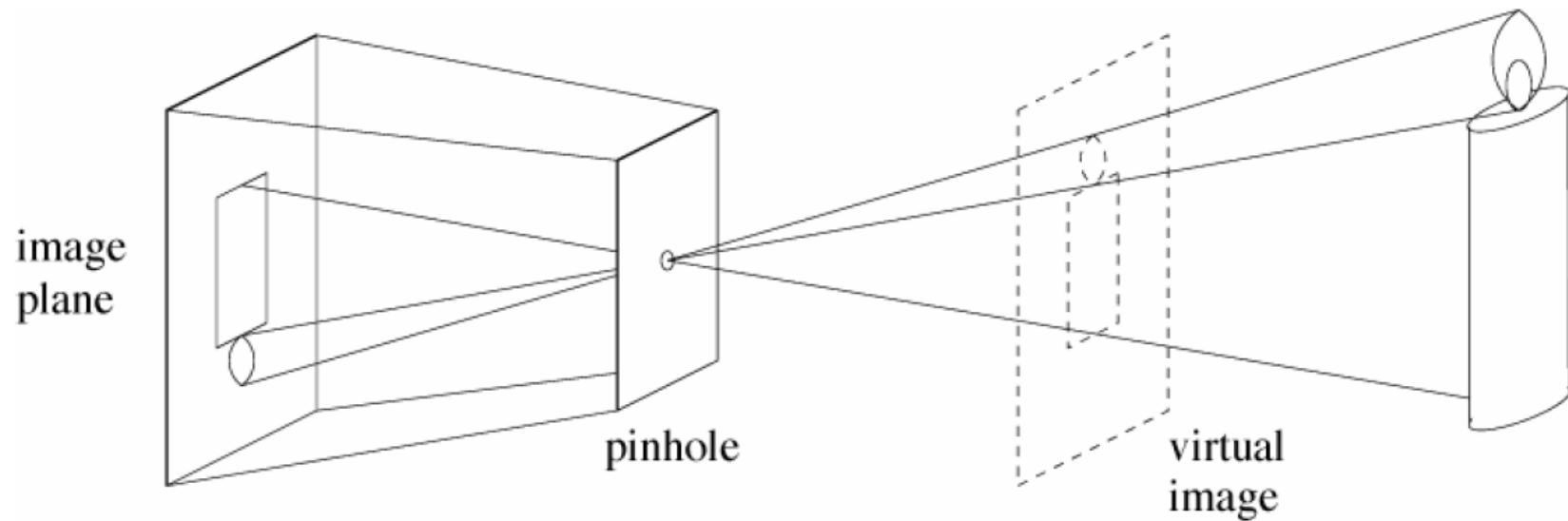
## ■ Gesichtsdetektion

- Schneiderman & Kanade: Naive Bayes Klassifikator  
de facto einer der erfolgreichsten Verfahren

## ■ Diskussion und Ausblick

# Lochkamera (Modell)

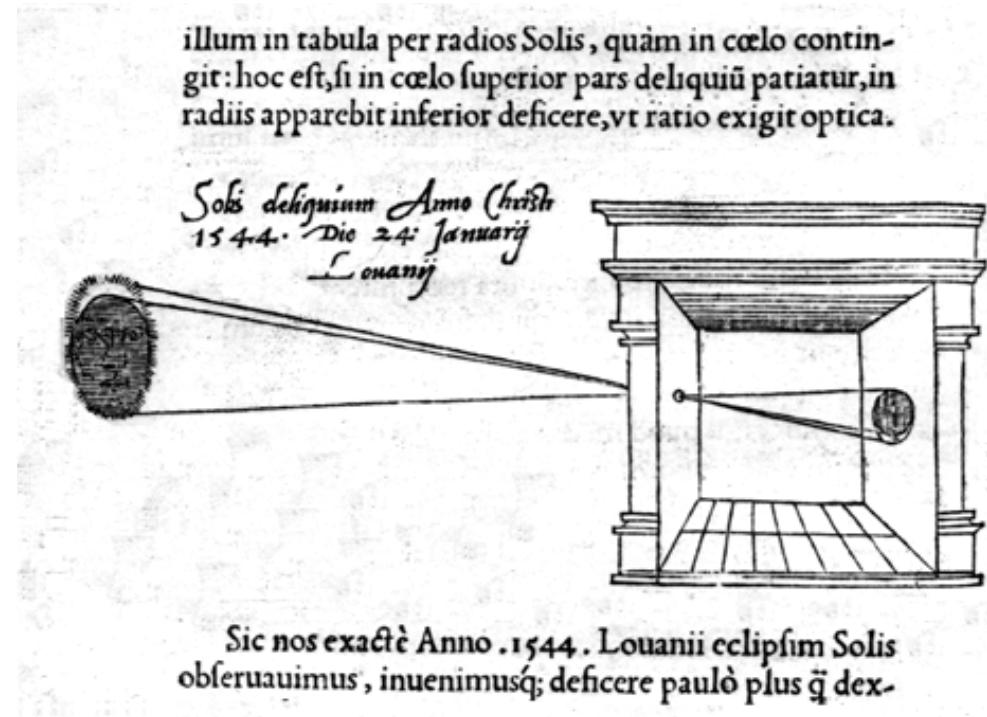
- (Einfaches) Standardmodell und abstraktes Modell
  - Kasten mit einem kleinen Loch



# Camera Obscura

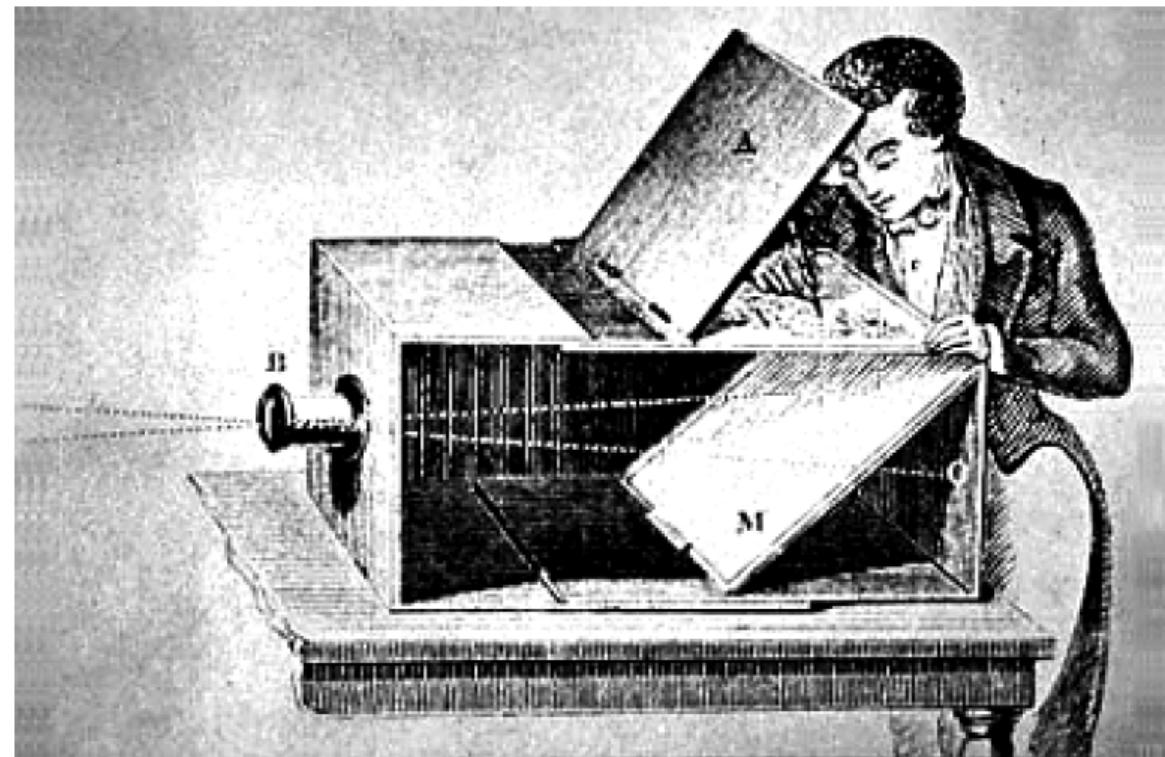
- Um 1519, Leonardo da Vinci (1452 - 1519)
  - [http://de.wikipedia.org/wiki/Camera\\_obscura](http://de.wikipedia.org/wiki/Camera_obscura)

► “when images of illuminated objects ... penetrate through a small hole into a very dark room ... you will see [on the opposite wall] these objects in their proper form and color, reduced in size ... in a reversed position owing to the intersection of the rays”



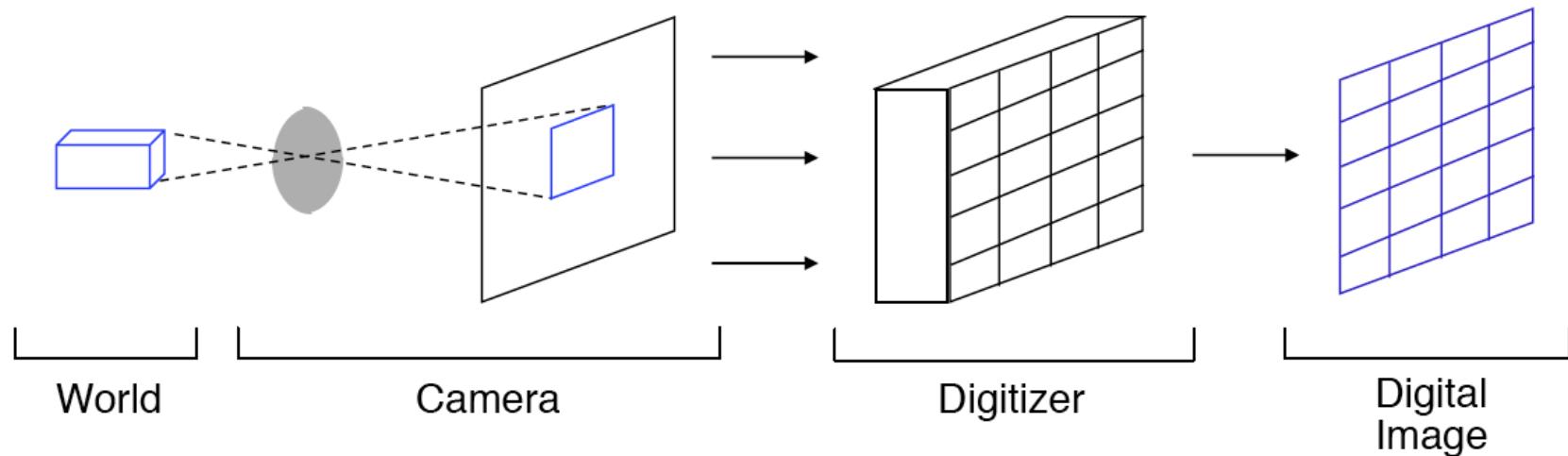
# Das Prinzip der Lochkamera...

- ...wurde von Künstlern...
  - (z.B. Vermeer  
17. Jahrhundert,  
Niederländer)
- ...und Wissenschaftlern  
verwendet



# Digitale Bilder

- Bildverarbeitungsprozess:
  - (Loch-) Kameramodell
  - Rasterisierung zur Erzeugung eines digitalen Bildes



# (Graustufen-) Bild

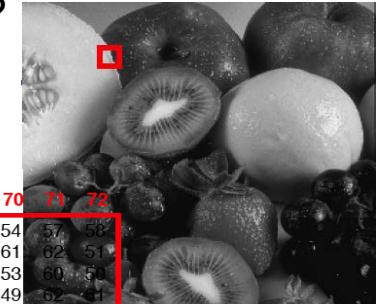
## ■ Ziele der Computer Vision

- Wie kann man Obst in einem Raster von (graustufen) Zahlen erkennen?
- Wie kann man Tiefen-informationen aus einem Raster von (graustufen) Zahlen herleiten?
- ...

## ■ computer vision = das Problem der „umgekehrten Grafik“?

## ■ Ziele der Computergrafik

- Wie kann man ein Raster von (graustufen) Zahlen erzeugen, das wie Obst aussieht?
- Wie kann man ein Raster von (graustufen) Zahlen erzeugen, sodass der menschliche Betrachter Tiefe wahrnimmt?
- ...



$x =$	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	
$y =$	41	210	209	204	202	197	247	143	71	64	80	84	54	54	57	58
42	206	196	203	197	195	210	207	56	63	58	53	53	53	61	62	51
43	201	207	192	201	198	213	156	69	65	57	55	52	53	60	59	50
44	216	206	211	193	202	207	208	57	69	60	55	77	49	56	51	41
45	221	206	211	194	196	197	220	56	63	60	55	46	97	58	106	111
46	209	214	224	199	194	193	204	173	64	60	59	51	62	56	48	55
47	204	212	213	208	191	190	191	214	60	62	66	76	51	49	55	55
48	214	215	215	207	208	180	172	188	69	72	55	49	56	52	56	56
49	209	205	214	205	204	196	187	196	86	62	66	87	57	60	48	48
50	208	209	205	203	202	186	174	185	149	71	63	55	55	45	56	56
51	207	210	211	199	217	194	183	177	209	90	62	64	52	93	52	52
52	208	205	209	209	197	194	183	187	187	239	58	68	61	51	56	56
53	204	206	203	209	195	203	188	185	183	221	75	61	58	60	60	60
54	200	203	199	236	188	197	183	190	183	196	122	63	58	64	66	66
55	205	210	202	203	199	197	196	181	173	186	105	62	57	64	63	63

# Konfluenz

Konfluenz: wo Vision und Graphics zusammenfließen



# Vor dem digitalen Zeitalter: Menschen und Kunst Wiedergewinnung von 3D-Strukturen in 2D



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT



# Menschen und Kunst

## Wiedergewinnung von 3D-Strukturen



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT



“Yves Saint Laurent – the film”

125 Jahre Eiffelturm

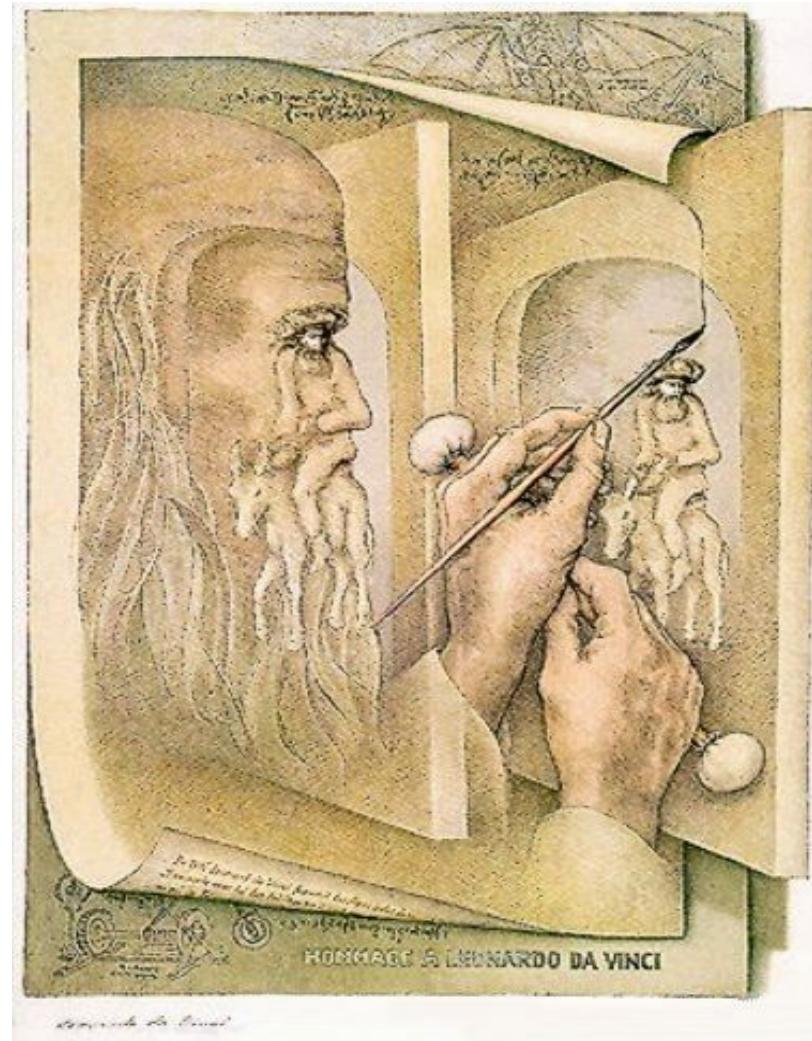
# Fallstudie

## Computer Vision und Objekterkennung

- Ist es mehr als umgekehrte Grafik?
- Wie erkennt man (als Mensch)
  - die Banane?
  - das Glas?
  - das Handtuch?
- Wie kann man Computer dazu bringen, dies ebenso zu tun?



# Schwierigkeit des Erkennens



# Schwierigkeit des Erkennens



# Schwierigkeit des Erkennens

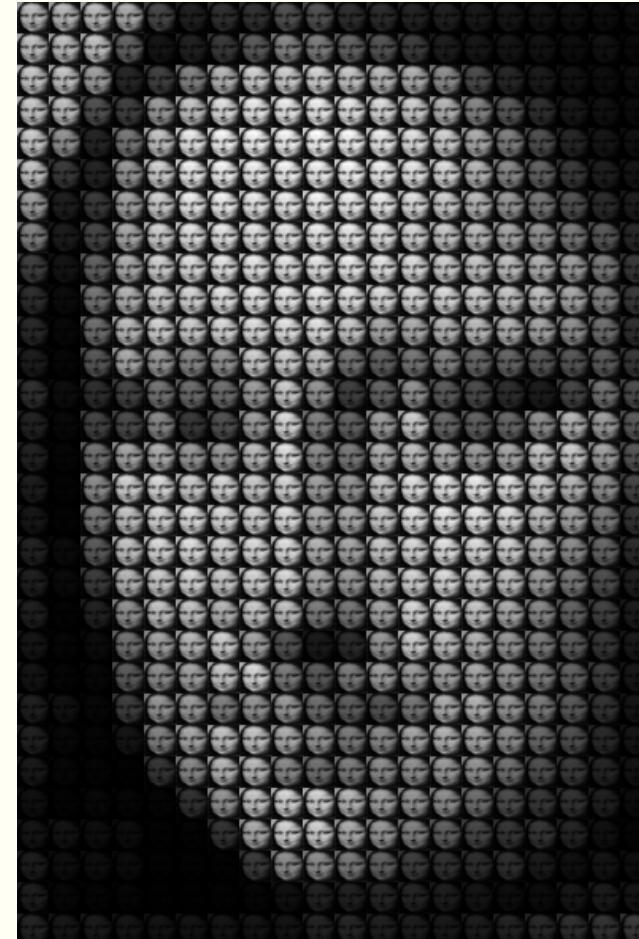
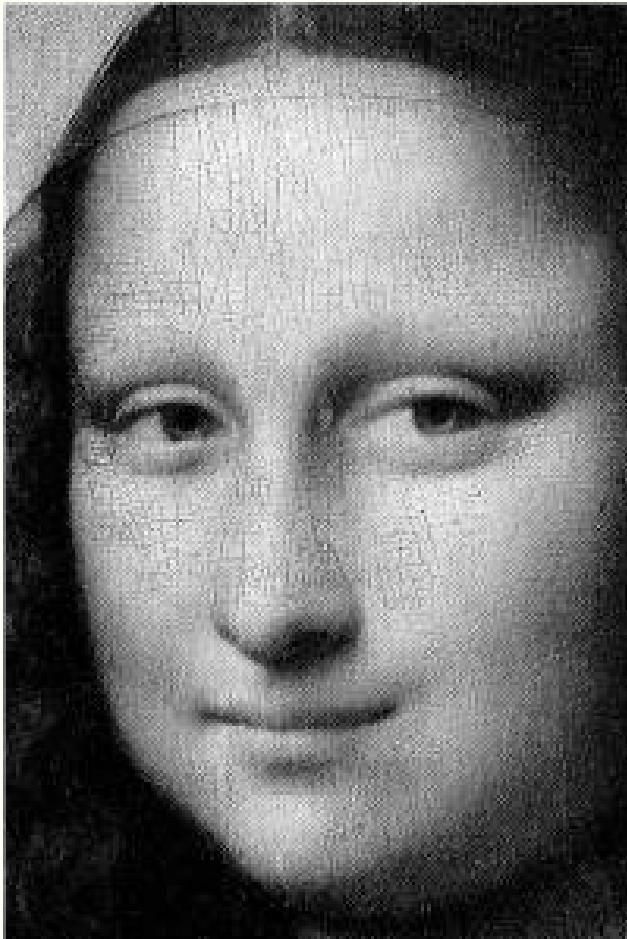


<http://www.victormolev.com>

# Schwierigkeit des Erkennens



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT



# Frau in einem gelben Bikini

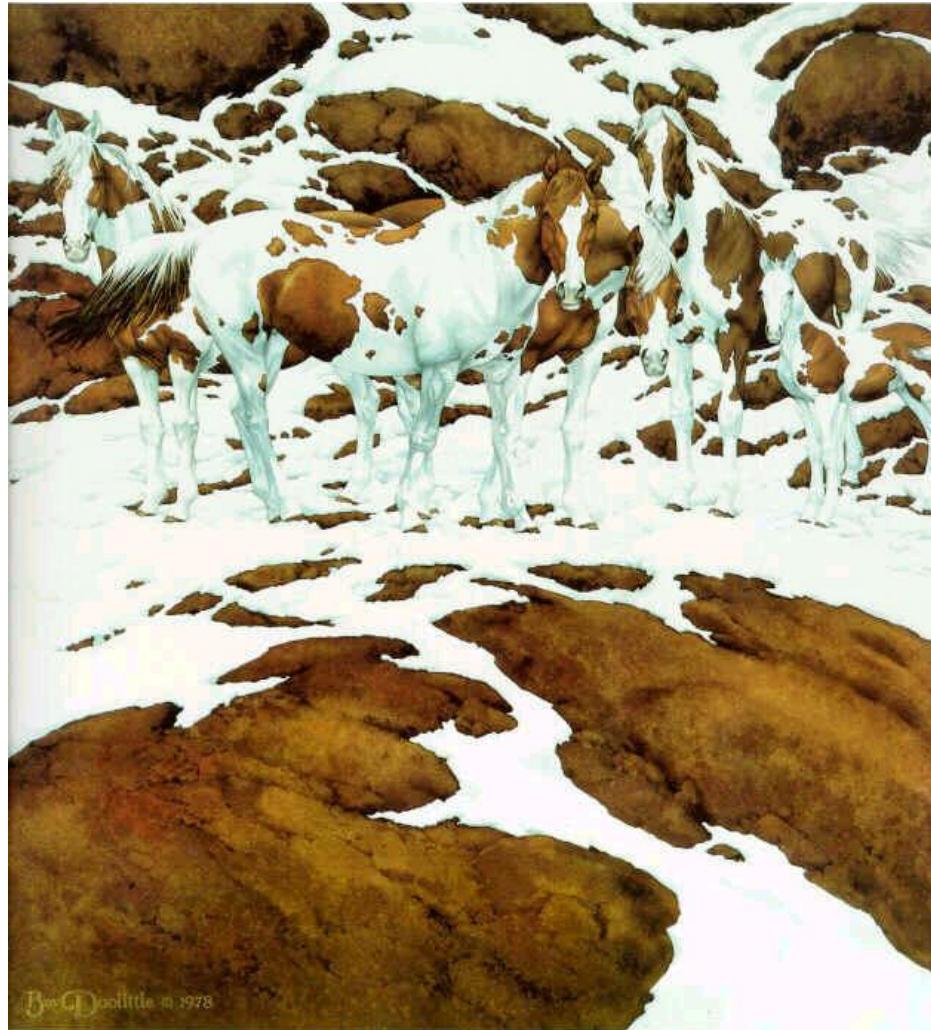


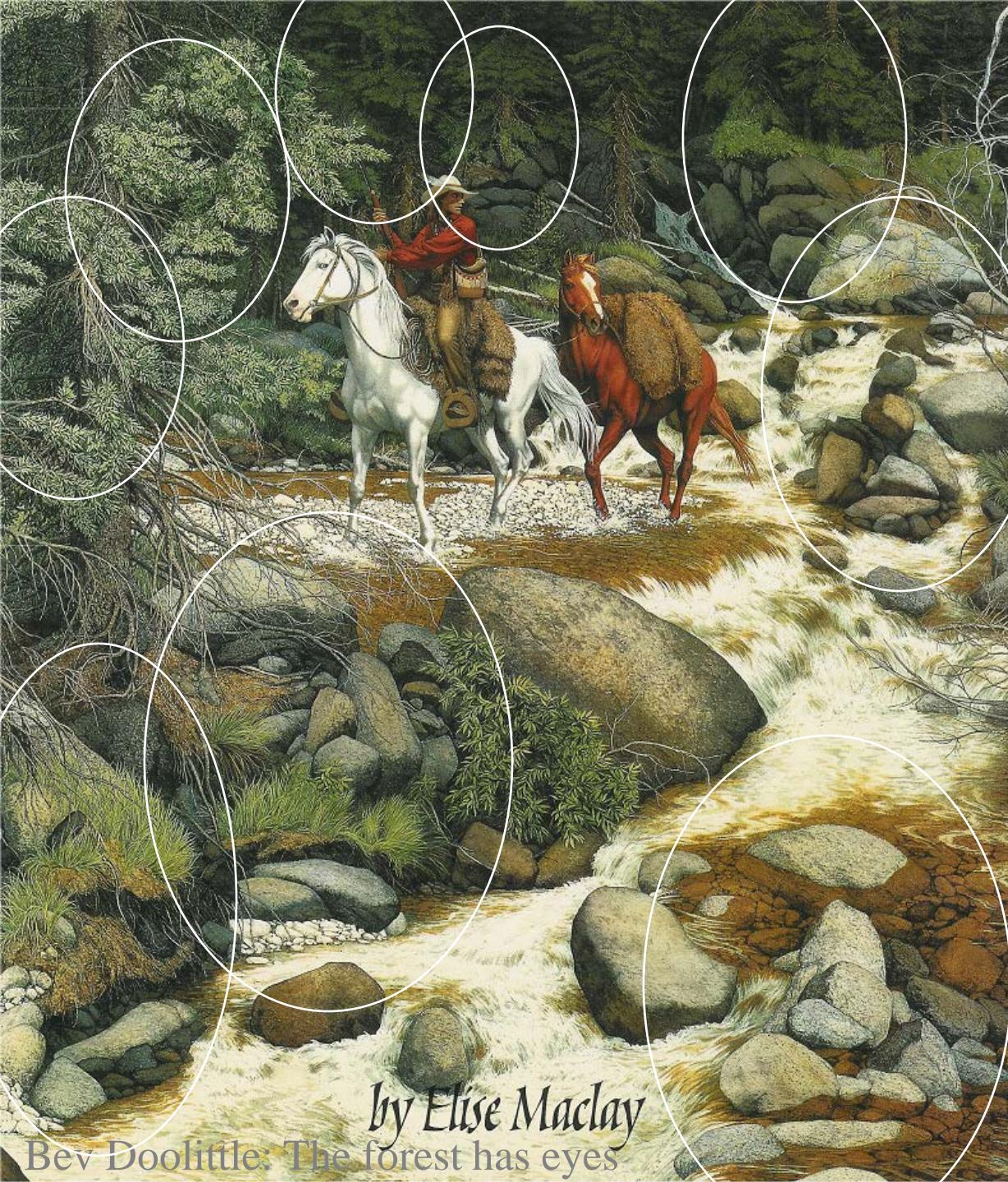
TECHNISCHE  
UNIVERSITÄT  
DARMSTADT



- David Mach: <http://www.dailyrecord.co.uk/news/scottish-news/2012/04/21/stunning-works-by-scots-artist-set-to-go-under-the-hammer-86908-23831695/>

# Schwierigkeit des Erkennens

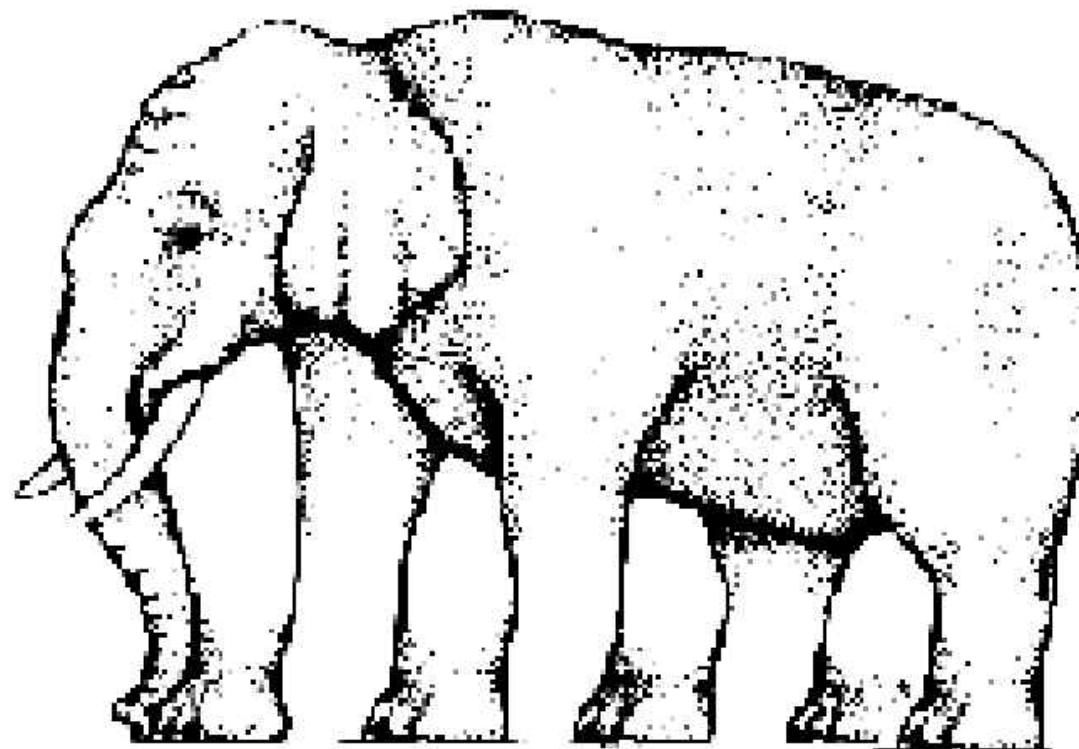




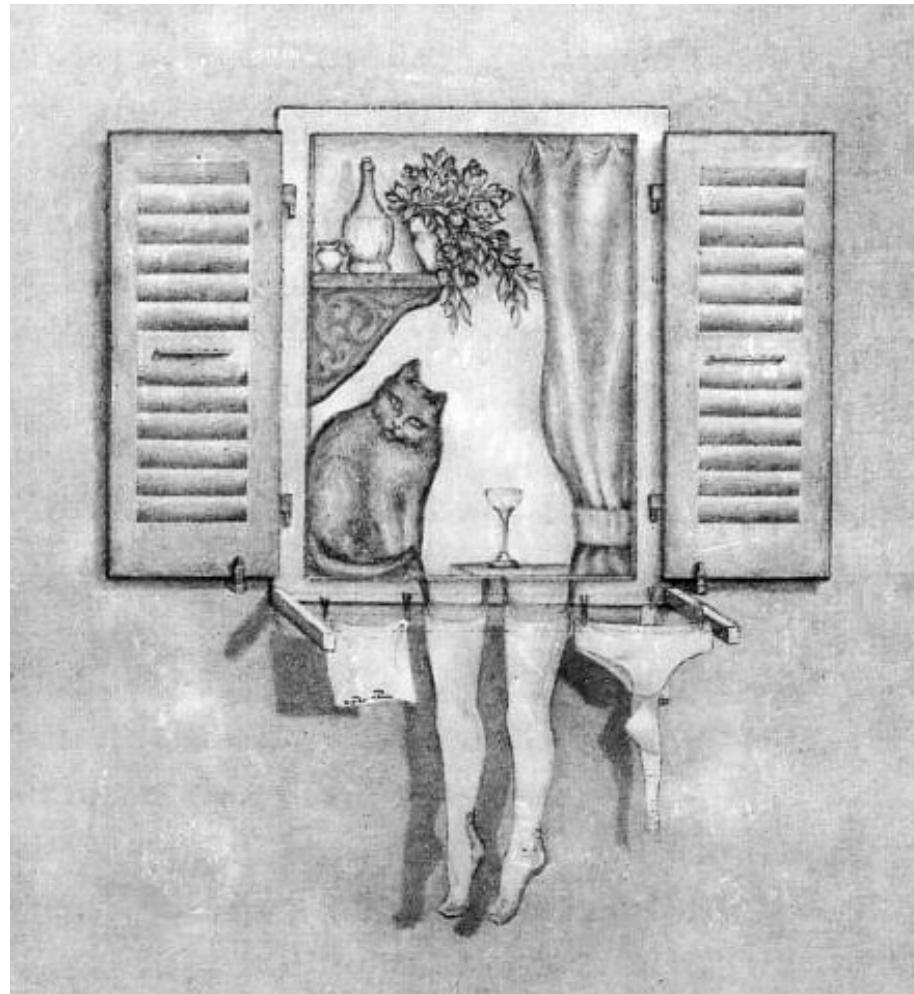
TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

by Elise Maclay  
Bev Doolittle: The forest has eyes

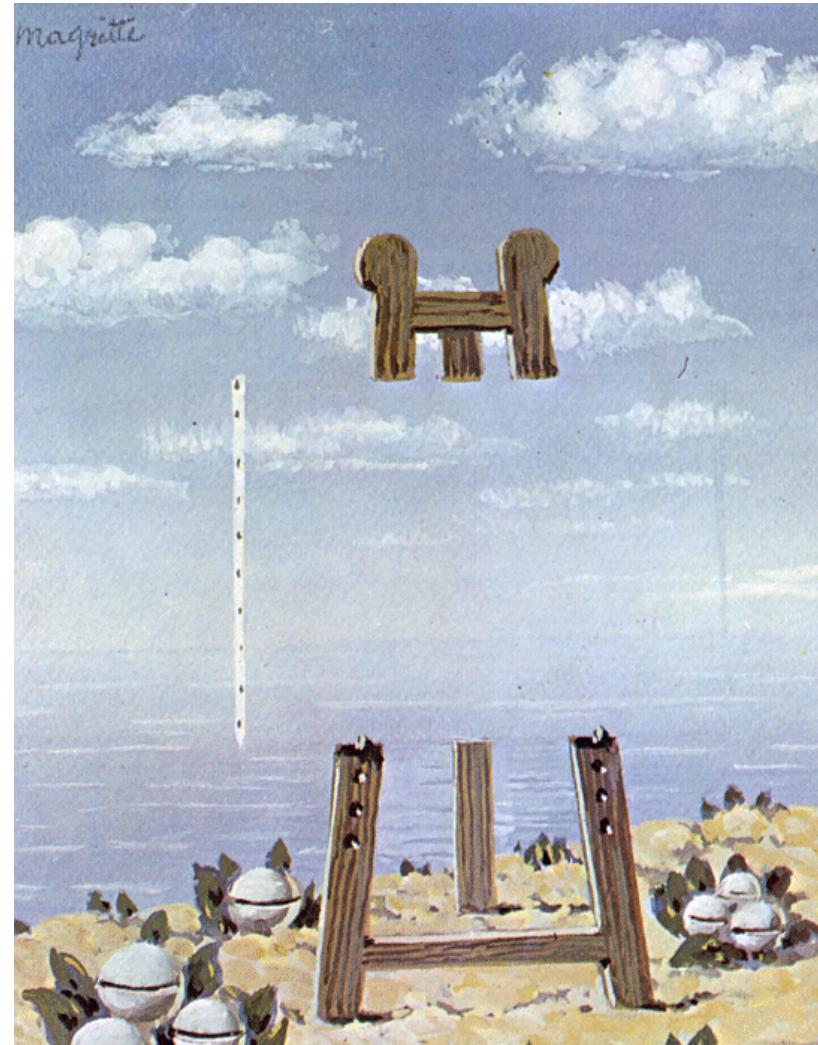
# Schwierigkeit des Erkennens



# Schwierigkeit des Erkennens



# Schwierigkeit des Erkennens





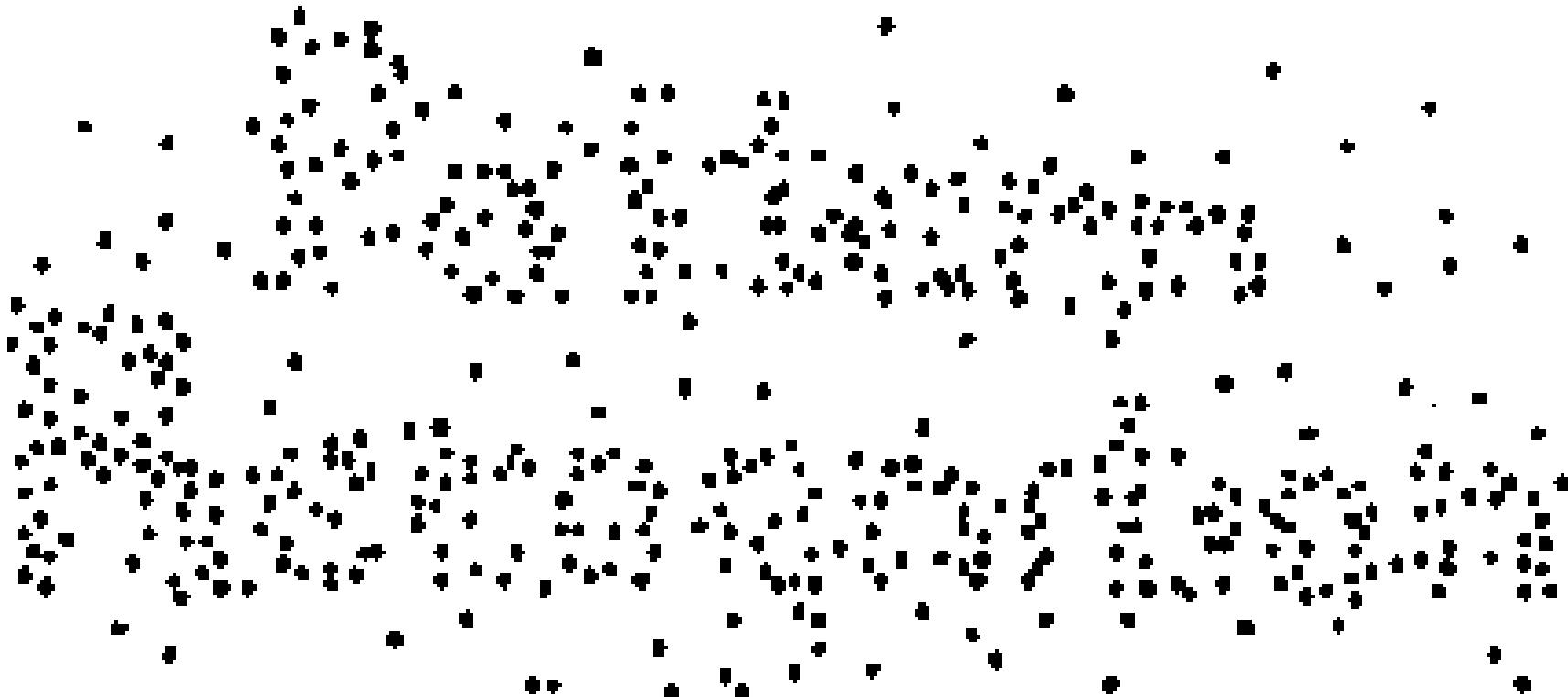
# Erkennung: Die Rolle des Zusammenhangs

- Antonio Torralba



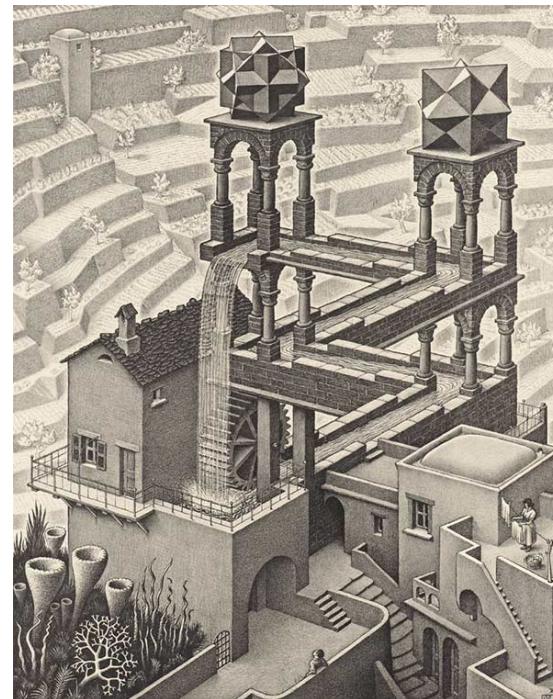


# Erkennung: Die Rolle des Zusammenhangs



# Fehler im Zusammenhang

- <http://www.mcescher.com/gallery/impossible-constructions/>





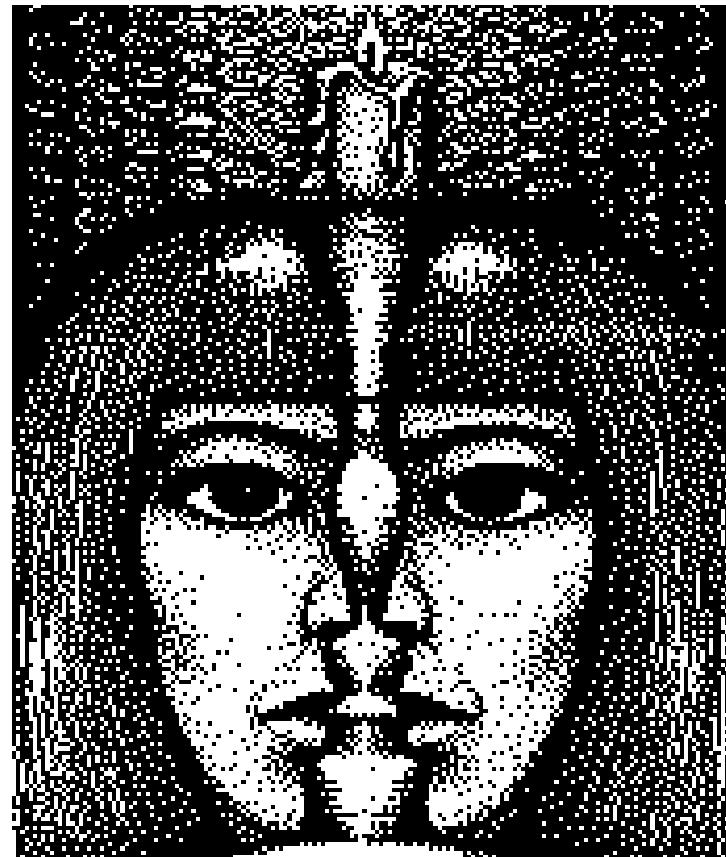
# Objekterkennung - Intuitionen...

- Wichtige Komponenten der Beschreibung
  - lokale Beschreibung / Merkmale
    - z.B. Augen, Mund, Nase
  - globale Anordnung der lokalen Merkmale
    - z.B. relative Positionen, relative Grösse
- Aber auch:
  - schnelle Generierung von guten Hypothesen
  - Segmentierung der Bildbereiche
  - Szenenkontext
  - ...



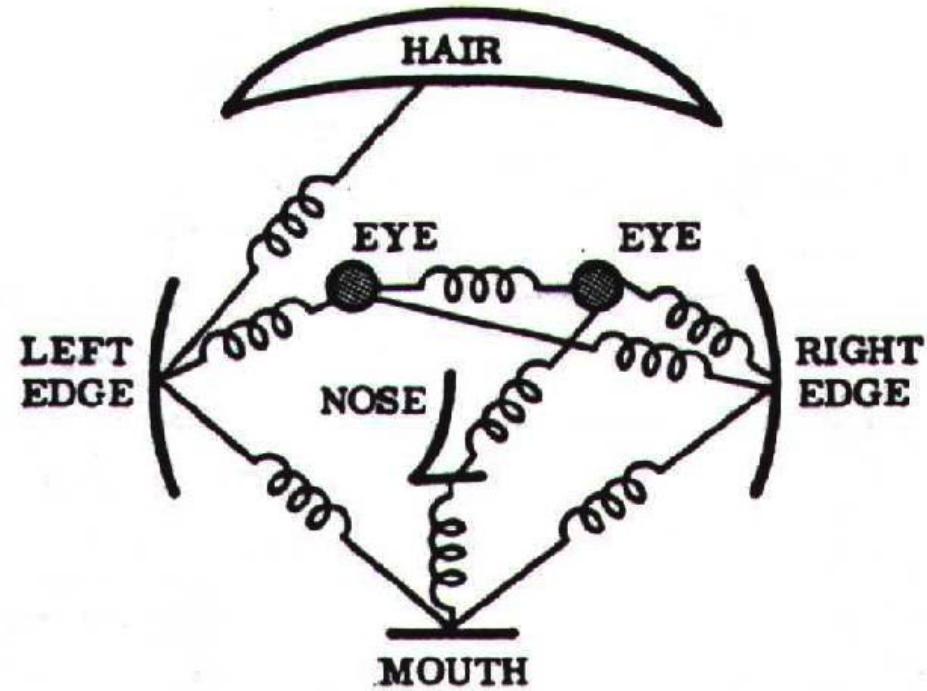
TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

# Ein oder zwei Gesichter?



# Bildaufbau (Pictorial Structure)

- Fischler und Elschlager 1973
- Das Modell hat zwei Komponenten:
  - Teile (parts)  
(2D Bildfragmente)
  - Aufbau (structure)  
(Anordnung der Teile)



# Deformationen



**A**



**B**



**C**



**D**

# Durcheinander (Clutter)



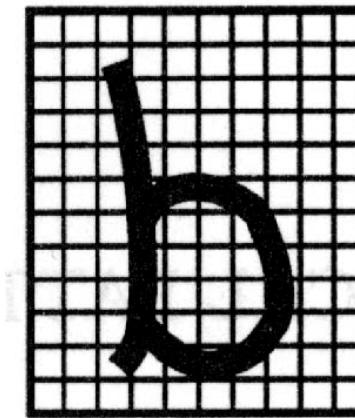
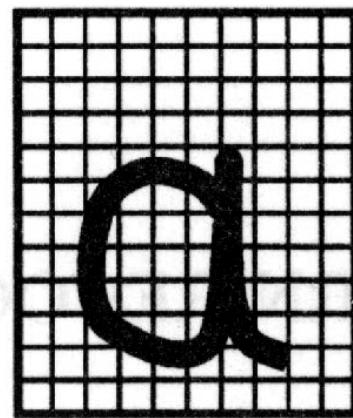


# Überblick

- Computer Vision
  - Objekterkennung und Objektdetektion als Fallbeispiel
  - Motivation heutiger Ansätze
- **Bayes Decision Theory**
  - **Verwendung von Merkmalen zur Objekterkennung**
  - **Naive Bayes Klassifikator**
- Gesichtsdetektion
  - Schneiderman & Kanade: Naive Bayes Klassifikator  
de facto einer der erfolgreichsten Verfahren
- Diskussion und Ausblick

# Bayes Decision Theory

- Beispiel: Buchstabenerkennung



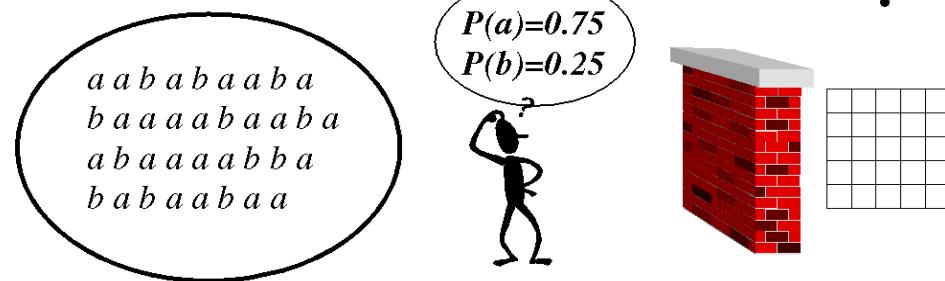
- Ziel: klassifiziere neuen Buchstaben, so dass die Wahrscheinlichkeit der Fehlklassifikation minimiert wird

# Bayes Decision Theory

- 1. Konzept: **Priors** (a priori Wahrscheinlichkeit)

$$P(C_k)$$

- Beispiel:



$$C_1 = a \quad P(C_1) = 0.75$$

$$C_2 = b \quad P(C_2) = 0.25$$

- generell:

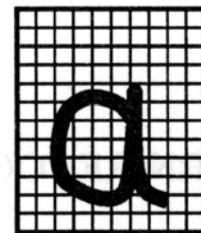
$$\sum_k P(C_k) = 1$$

# Bayes Decision Theory

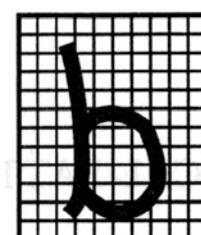
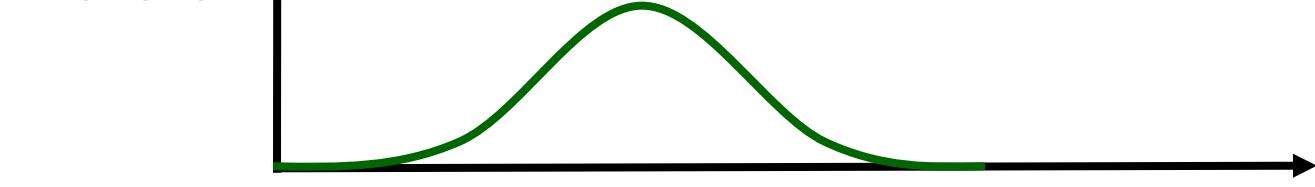
## ■ 2. Konzept: **bedingte Wahrscheinlichkeit**

- mit 'x' Merkmalsvektor (feature)
- 'x' misst/beschreibt Eigenschaft(en)
  - Beispiel: # der schwarzen Pixel, Höhen-Breiten-Verhältnis...

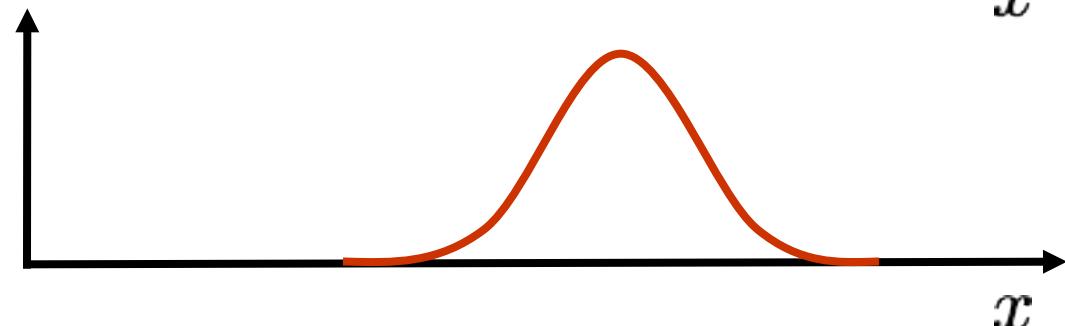
$$P(x|C_k)$$



$$p(x|a)$$



$$p(x|b)$$





## Bemerkung bzgl. ‘pdf’ vs. Probability...

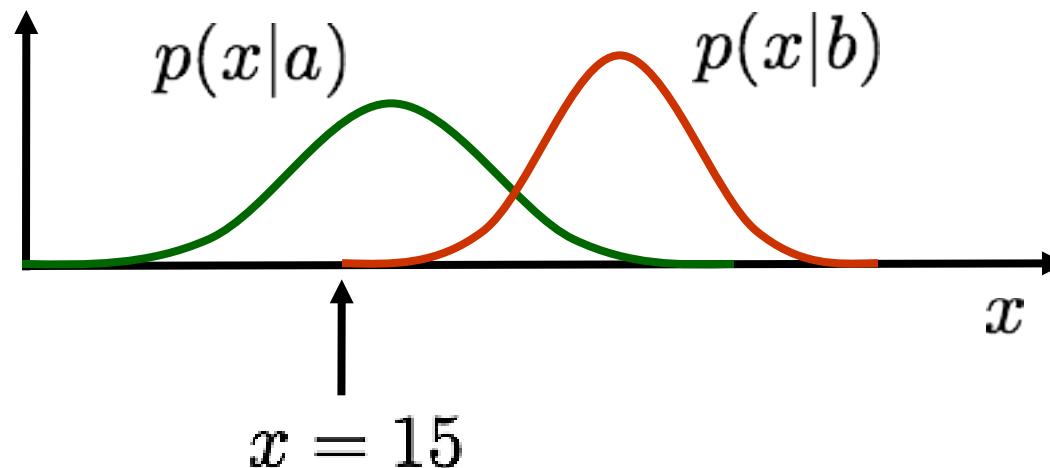
- Zusammenhang zwischen Wahrscheinlichkeitsdichte und Wahrscheinlichkeit für
  - Wahrscheinlichkeit, daß  $x$  ist in dem Intervall  $(x_0, x_1)$ , ist definiert durch:
$$Pr(x_0 < x < x_1) = \int_{x_0}^{x_1} p(t)dt$$
- Wahrscheinlichkeit  $P(x)$  ist somit nicht direkt definiert
- Im folgenden verwenden wir  $P(x)$  in der folgenden Weise:
  - wenn  $dx$  unendlich klein wird, kann die Wahrscheinlichkeit, dass  $x$  ist in dem Intervall  $(x, x+dx)$  ist, wie folgt berechnet werden:

$$P(x) = Pr(x < t < x + dx) = p(x)dx$$

(Riemann....)

# Bayes Decision Theory

- Beispiel:

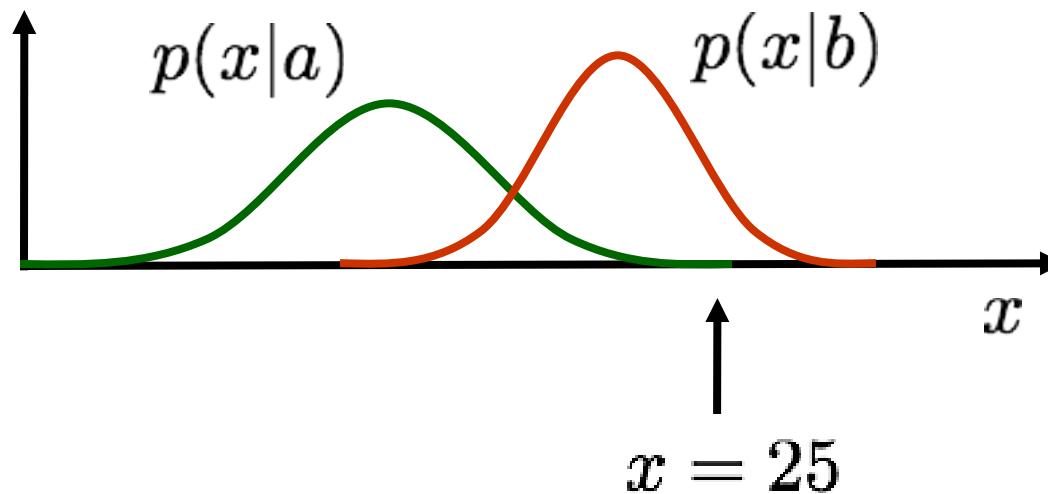


- Frage:

- welche Klasse?
- die Entscheidung sollte hier ‘a’ sein da  $p(x|b) = 0$

# Bayes Decision Theory

- Beispiel:

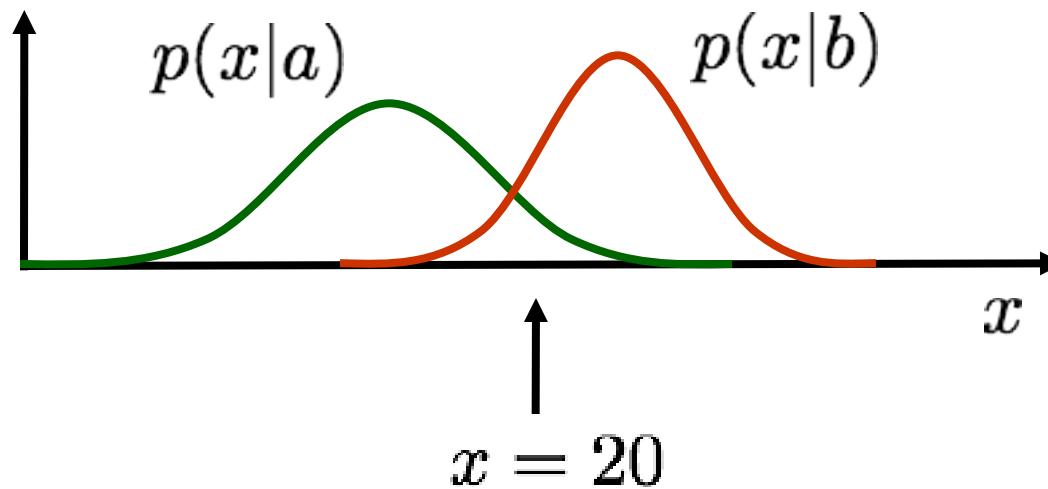


- Frage:

- welche Klasse?
- da  $p(x|a)$  viel kleiner als  $p(x|b)$ , oder besser:  $p(x|a) = 0$ , sollte die Entscheidung hier ‘b’ sein

# Bayes Decision Theory

- Beispiel:



- Frage:

- welche Klasse?
- beachte, dass  $P(a) = 0.75$  und  $P(b) = 0.25\dots$
- $P(x|a) P(a) > P(x|b) P(b)$
- das heisst, die Entscheidung sollte hier wieder ‘a’ sein !!



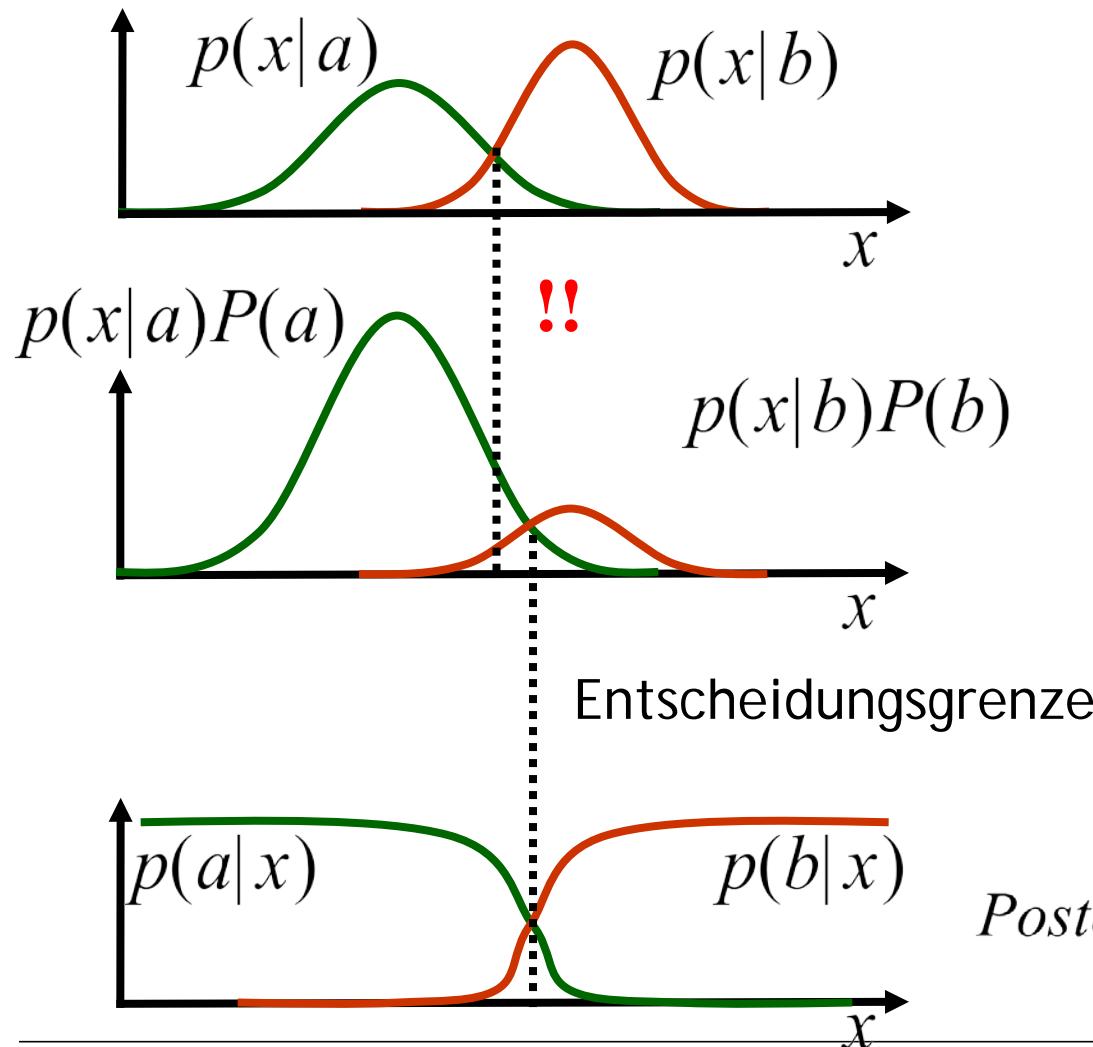
# Bayes Decision Theory

- 3. Konzept: a posteriori Wahrscheinlichkeit (**posterior**)
  - **Wahrscheinlichkeit** der Klasse  $C_k$ ,  
**gegeben** ein bestimmter Merkmalsvektor ‘x’
  - Bayes’ Theorem:

$$P(C_k|x) = \frac{P(x|C_k)P(C_k)}{P(x)} = \frac{P(x|C_k)P(C_k)}{\sum_j P(x|C_j)P(C_j)}$$

$$\textit{Posterior} = \frac{\textit{Likelihood} \times \textit{Prior}}{\textit{Normalization Factor}}$$

# Bayes Decision Theory



$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Normalization factor}}$$

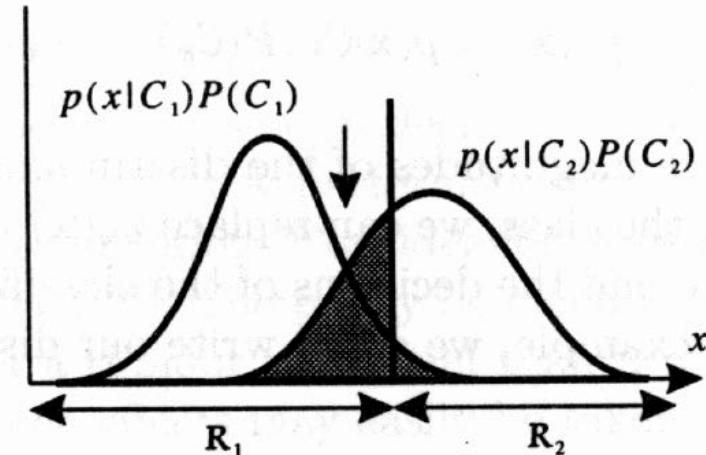
# Bayes Decision Theory

- Ziel:

**Wahrscheinlichkeit der Fehlklassifikation minimieren**

Oder:

Entscheidungsgrenze bestimmen



$$\begin{aligned}
 P(error) &= P(x \in R_2, C_1) + P(x \in R_1, C_2) \\
 &= P(x \in R_2 | C_1)P(C_1) + P(x \in R_1 | C_2)P(C_2) \\
 &= \int_{R_2} p(x \in R_2 | C_1)P(C_1)dx + \int_{R_1} p(x \in R_1 | C_2)P(C_2)dx
 \end{aligned}$$

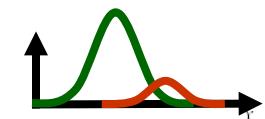
# Bayes Decision Theory

- Entscheidungsregel:
  - entscheide  $C_1$  wenn

$$P(C_1|x) > P(C_2|x)$$

- das ist äquivalent zu:

$$P(x|C_1)P(C_1) > P(x|C_2)P(C_2)$$



- das ist äquivalent zu (genannt Likelihood Ratio Test):

$$\frac{P(x|C_1)}{P(x|C_2)} > \frac{P(C_2)}{P(C_1)} = \lambda$$



# Bayes Decision Theory

- Verallgemeinerung zu mehr als 2 Klassen:
  - entscheide Klasse ‘k’ immer dann, wenn diese die grösste a posteriori Wahrscheinlichkeit hat:

$$P(C_k|x) > P(C_j|x) \text{ for all } j \neq k$$

- oder äquivalent:

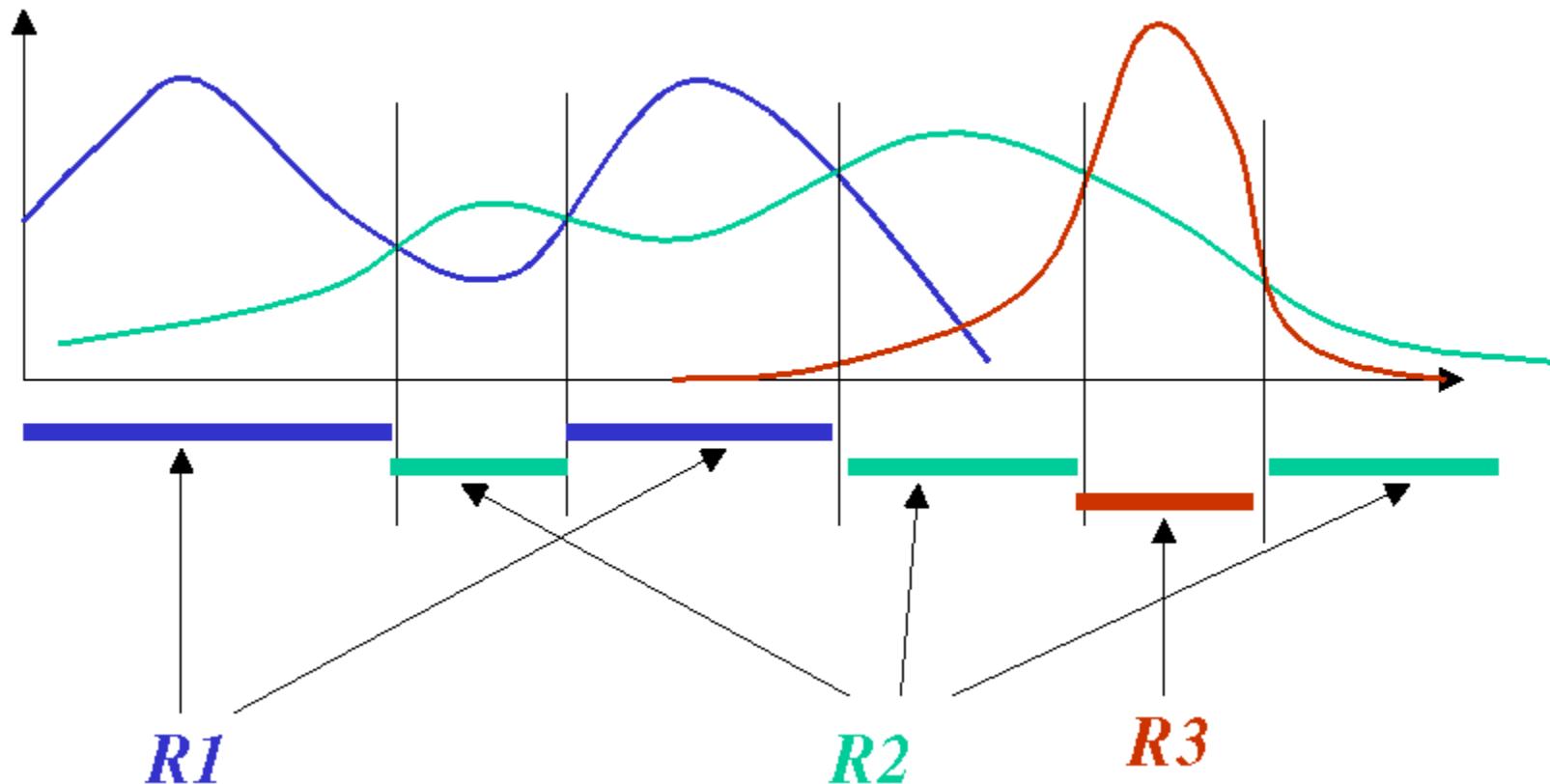
$$P(x|C_k)P(C_k) > P(x|C_j)P(C_j) \text{ for all } j \neq k$$

$$\frac{P(x|C_k)}{P(x|C_j)} > \frac{P(C_j)}{P(C_k)} \text{ for all } j \neq k$$

# Bayes Decision Theory

- Entscheidungsregionen:

$$R_1, R_2, \dots$$





# Bayes Decision Theory

- Verwendung von zwei Merkmalen:  $x_1, x_2$ 
  - z.B. Höhe & Breite

$$P(C_k|x_1, x_2) = \frac{P(x_1, x_2|C_k)P(C_k)}{P(x_1, x_2)}$$

- Problematisch: Schätzung der 2-dimensionalen Wahrscheinlichkeitsdichten nicht immer möglich
  - davon haben wir zwei:

$$P(x_1, x_2|C_k)$$

$$P(x_1, x_2)$$



# Naive Bayes Classifier

- Annahme des naive Bayes Klassifikators:
  - Die beiden Merkmale  $x_1, x_2$  sind (statistisch) *unabhängig*
  - Damit wird:

$$P(x_1, x_2 | C_k) = P(x_1 | C_k)P(x_2 | C_k)$$

$$P(x_1, x_2) = P(x_1)P(x_2)$$

- Naive Bayes Klassifikator (für 2 Merkmale)

$$P(C_k | x_1, x_2) = \frac{P(x_1 | C_k)}{P(x_1)} \frac{P(x_2 | C_k)}{P(x_2)} P(C_k)$$



# Naive Bayes Classifier

- Verwendung von d Merkmalen:  $x_1, x_2, \dots, x_d$

$$P(C_k|x_1, \dots, x_d) = \frac{P(x_1, \dots, x_d|C_k)P(C_k)}{P(x_1, \dots, x_d)}$$

- Naive Bayes Klassifikator:

- Annahme, dass alle Merkmale  $x_i$  statistisch *unabhängig* sind.
- Die Annahme der Unabhängigkeit ist oft **nicht** richtig, ergibt aber häufig gute Ergebnisse und ist somit ein guter Vergleich bzw. Basis zum Vergleich.

$$P(C_k|x_1, \dots, x_d) = \prod_{i=1}^d \frac{P(x_i|C_k)}{P(x_i)} P(C_k)$$



# Naive Bayes Classifier

$$P(C_k|x_1, \dots, x_d) = \prod_{i=1}^d \frac{P(x_i|C_k)}{P(x_i)} P(C_k)$$

- Spezialfall 2-Klassenproblem:

- Entscheide Klasse  $C_1$  wenn:

$$P(C_1|x_1, \dots, x_d) > P(C_2|x_1, \dots, x_d)$$

$$\prod_{i=1}^d \frac{P(x_i|C_1)}{P(x_i)} P(C_1) > \prod_{i=1}^d \frac{P(x_i|C_2)}{P(x_i)} P(C_2)$$

- vereinfacht zum “Likelihood Ratio Test”:

$$\prod_{i=1}^d \frac{P(x_i|C_1)}{P(x_i|C_2)} > \frac{P(C_2)}{P(C_1)} = \lambda$$

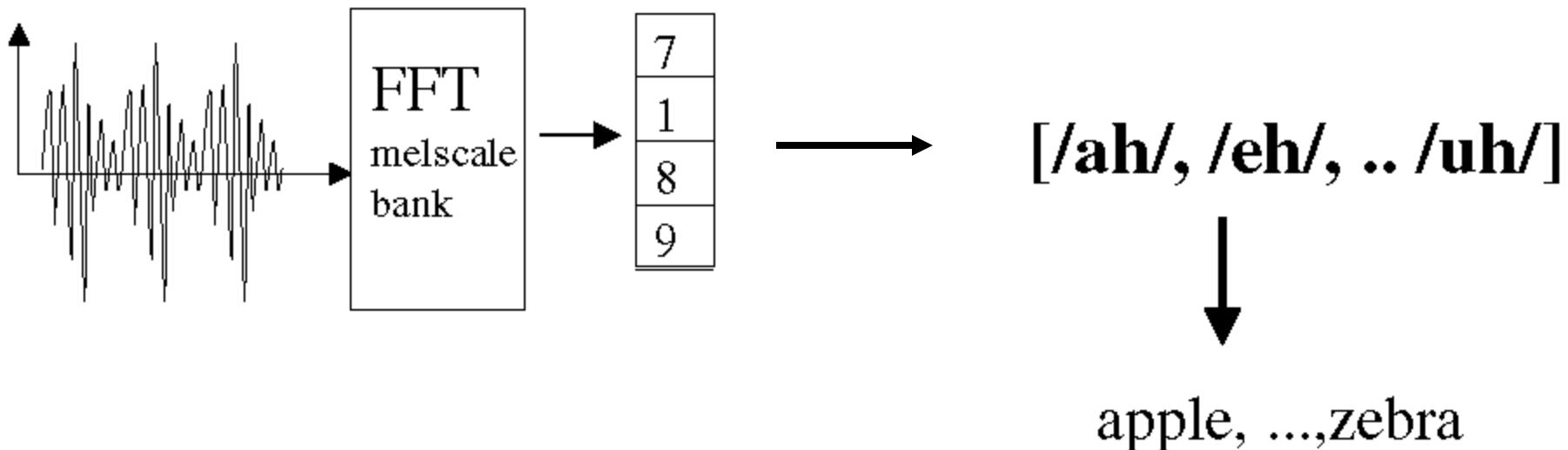


# Probability Density Estimation

- bisher:
  - Bayes optimale Klassifikation
  - Verwendung der Wahrscheinlichkeiten:
$$P(x|C_k)P(C_k)$$
- Schätzung der Wahrscheinlichkeitsdichten (density estimation):
  - Schätzung der Wahrscheinlichkeitsdichte einer Klasse  $C_k$ 
$$p(x|C_k)$$

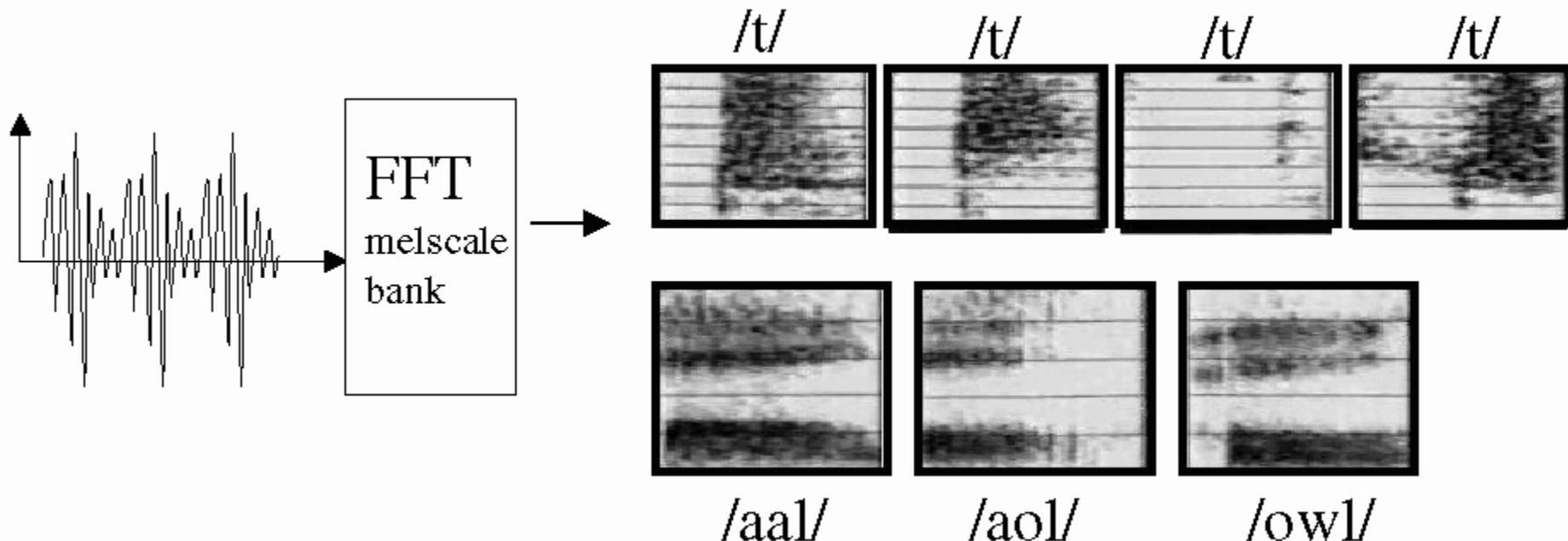
# Bayes Decision Theory

- Was ist an  $P(x|C_k)P(C_k)$  so interessant?
  - Beispiel #1: Spracherkennung:



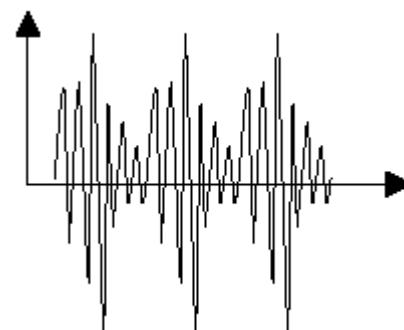
# Bayes Decision Theory

- Was ist an  $P(x|C_k)P(C_k)$  so interessant?
  - Beispiel #1: Spracherkennung:



# Bayes Decision Theory

- Was ist an  $P(x|C_k)P(C_k)$  so interessant?
  - Beispiel #1: Spracherkennung:



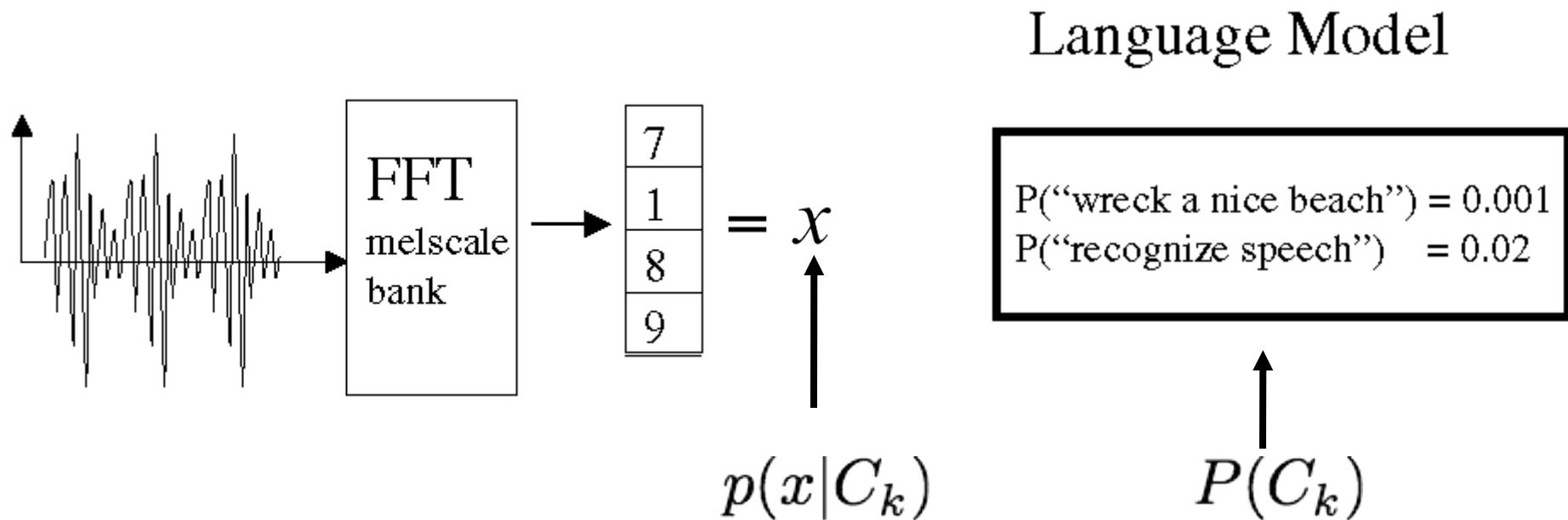
“This machine can recognize speech”

“This machine can wreck a nice beach”

- ▶ wie können wir den richtigen Satz erkennen ??

# Bayes Decision Theory

- Was ist an  $P(x|C_k)P(C_k)$  so interessant?
  - Beispiel #1: Spracherkennung:





# Bayes Decision Theory

- Was ist an  $P(x|C_k)P(C_k)$  so interessant?
  - Beispiel #2: Bildverarbeitung:

Low-Level  
Image  
Measurements

High-Level  
Model  
Knowledge

$$p(x | C_k)$$

$$P(C_k)$$



# Überblick

- Computer Vision
  - Objekterkennung und Objektdetektion als Fallbeispiel
  - Motivation heutiger Ansätze
- Bayes Decision Theory
  - Verwendung von Merkmalen zur Objekterkennung
  - Naive Bayes Klassifikator
- Gesichtsdetektion
  - **Schneiderman & Kanade: Naive Bayes Klassifikator de facto einer der erfolgreichsten Verfahren**
- Diskussion und Ausblick



# Appearance-Based Methods

## ■ Prinzip:

- Erscheinungsmodell (appearance model) aus (üblicherweise) großen Sammlungen von Bildern lernen
- Am häufigsten: Sliding Window Ansatz (d.h. über alle Positionen und Skalierungen!)
- Jedes Fenster (d.h. Teilbereich des Bildes) als „Gesicht“ oder „kein Gesicht“ klassifizieren

## ■ Aspekte:

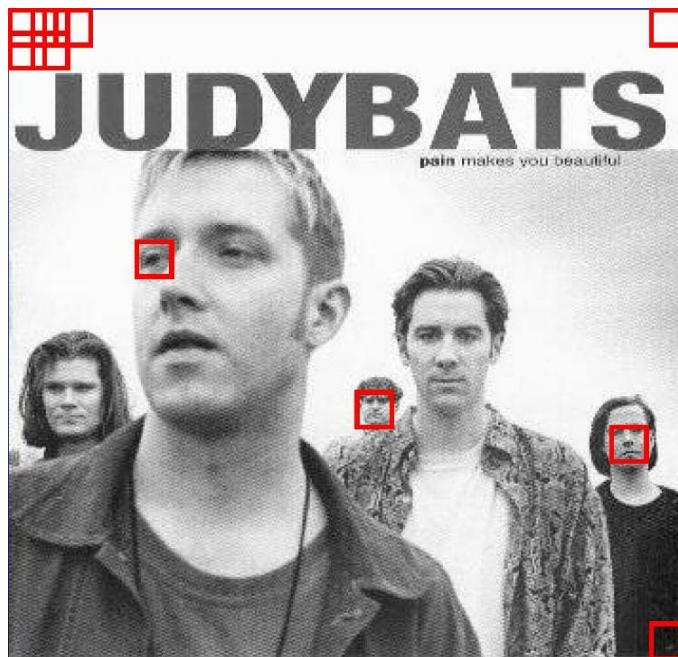
1. Repräsentation des Objektes
  - lokale Merkmale (z.B. Auge, Mund, Nase, ...)
  - globale Anordnung der Merkmale
2. Trainingsdaten
  - sowohl positive Beispiele (z.B. Gesichter)
  - als auch negative Beispiele (z.B. keine Gesichter)
3. Klassifikator und Lernmethode

# Appearance-Based Methods

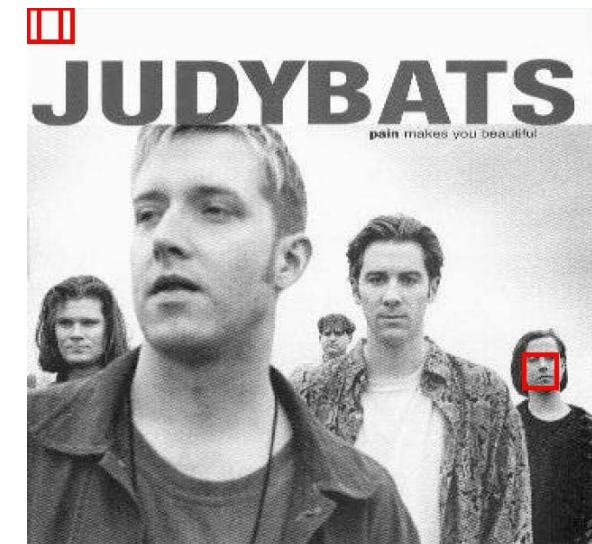
## Suchstrategie



- Suchen über Raum und Skalierung (Sliding Window Approach)



...



Ein Eingabebild wird in Ein-Pixel-Schritten horizontal und vertikal gescannt

Das Bild wird um den Faktor 1,2 verkleinert, die Suche wiederholt

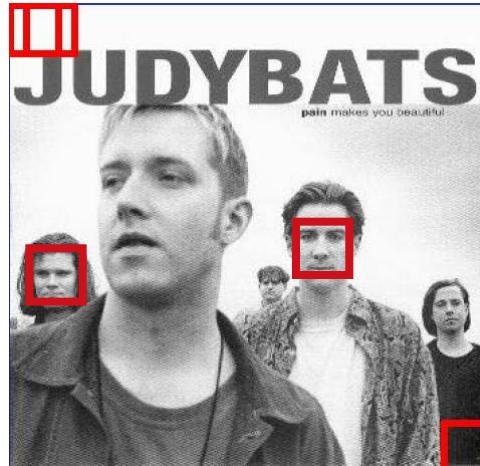
# Appearance-Based Methods

## Suchstrategie

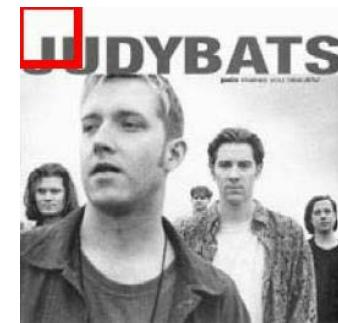


TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

- Suchen über Raum und Skalierung (Sliding Window Approach)



• • •



• • •



Das Bild wird immer weiter um 1,2 verkleinert und es wird immer weiter gesucht, bis das Bild zu klein ist

# Beispiel: Gesichtsdetektion

- Schneiderman & Kanade - 1998
  - einer der performantesten Gesichtdetektoren

## 1. Aspekt: Repräsentation des Objektes

- Verwendung der Waveletzerlegung von Bildern:  
Darstellung der Gesichtsmerkmale mit Frequenzen und deren Ort und Orientierung



Level 2 HL	Level 2 HH	Level 3 HL
Level 2 LH	Level 2 HH	
Level 3 LH		Level 3 HH

- lokale Merkmale
  - wavelet Koeffizienten: Frequenzen der Gesichtsmerkmale ('Kanten') wie z.B. Auge, Mund
- globale Anordnung der Merkmale
  - absolute Position der Frequenzen im Bild

# Beispiel: Gesichtsdetektion

## 2. Aspekt: Trainingsdaten

### ▪ Positive Beispiele

- Sollten möglichst vielfältig sein
- Jedes Bild eines Gesichts wird manuell an den Rändern abgeschnitten und auf eine Standardgröße normalisiert (z.B.  $19 \times 19$  Pixel)
- Es werden virtuelle Beispiele erstellt

### ▪ Negative Beispiele

- Beliebige Bilder, die kein Gesicht enthalten
- Teilbilder von großen Bildern



# Beispiel: Gesichtsdetektion: [Schneiderman and Kanade 98]

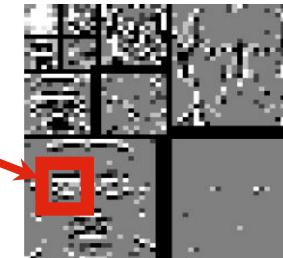
## 3. Aspekt: Klassifikator und Lernmethode

- Naïve Bayes Classifiers:

$$P(C_k|x_1, \dots, x_d) = \prod_{i=1}^d \frac{P(x_i|C_k)}{P(x_i)} P(C_k)$$

- Merkmale:  $x_i$  : Wavelet Koeffizienten an bestimmter Position (u, v)

$$x_i = (f_i, u_i, v_i)$$



- 2 Klassenproblem:

- $C_1$  : Gesichter
- $C_2$  : alles andere = nicht-Gesichter

# Beispiel: Gesichtsdetektion: [Schneiderman and Kanade 98]

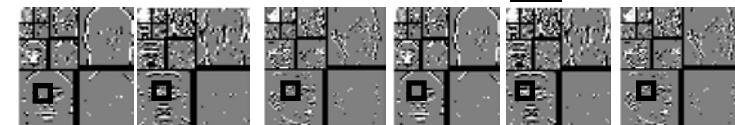


- “Lernen” = Schätze Wahrscheinlichkeit der Wavelet Koeffizienten für

- $C_1$  = Gesichter

$$P(x_i|C_1) = P(f_i, u_i, v_i|C_1)$$

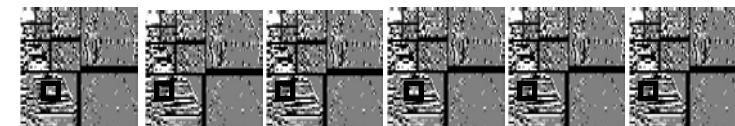
Wavelet Koeffizienten von Bildern MIT Gesichtern



- $C_2$  = nicht-Gesichter

$$P(x_i|C_2) = P(f_i, u_i, v_i|C_2)$$

Wavelet Koeffizienten von Bildern OHNE Gesichter



- Hier:

- Diskretisiere Koeffizienten und Positionen -> diskrete und endliche Anzahl von  $x_i$
  - Zähle, wie häufig jedes  $x_i$  in Bildern mit Gesichtern und Bildern ohne Gesichter vorkommt
  - verwende Likelihood Ratio Test:

$$\prod_{i=1}^d \frac{P(x_i|C_1)}{P(x_i|C_2)} > \frac{P(C_2)}{P(C_1)} = \lambda$$

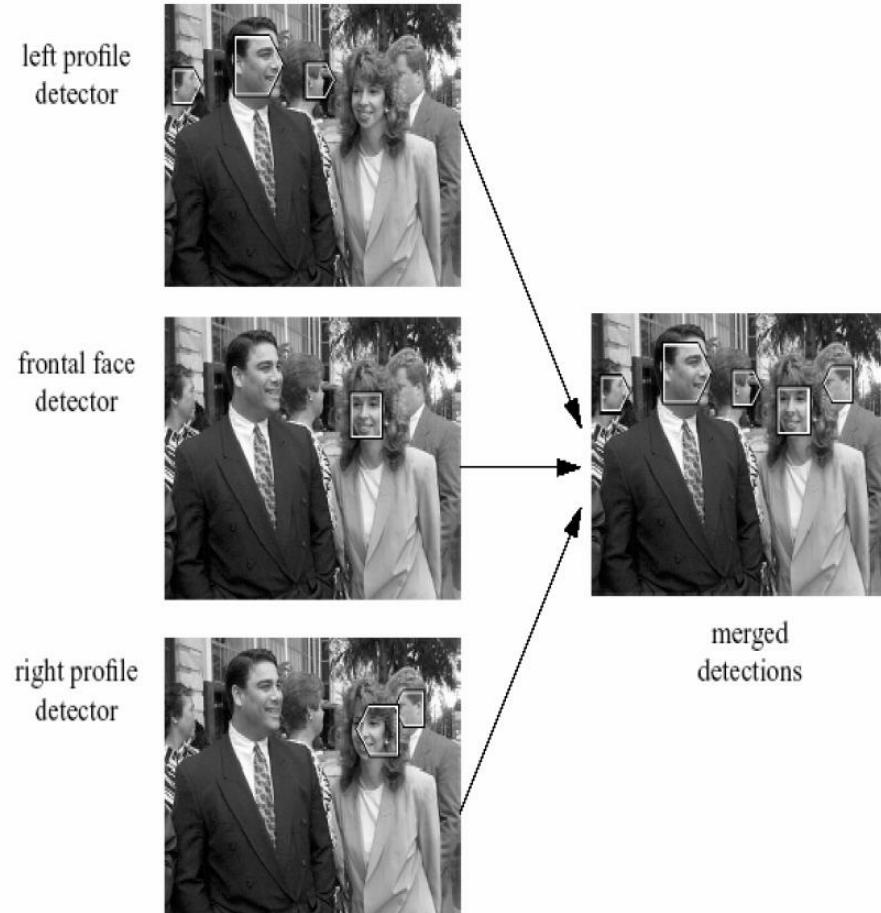
# Appearance-Based Methods

## Naïve Bayes Classifiers [Schneiderman and Kanade 98]



- Bilder aus verschiedenen Ansichten erkennen

- Mehrere Detektoren
- Jeder Detektor ist auf eine Ansicht spezialisiert: frontal, von links, von rechts



# Appearance-Based Methods

## Naïve Bayes Classifiers [Schneiderman and Kanade 98]



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

### ▪ Testergebnisse

- Möglich, Gesichter aus verschiedenen Ansichten zu erkennen
- Erweiterbar, um Autos und andere Objekte zu erkennen



# Koninginnedag 2012



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT



# Erkennung & Identifizierung



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT



# Niederländer!



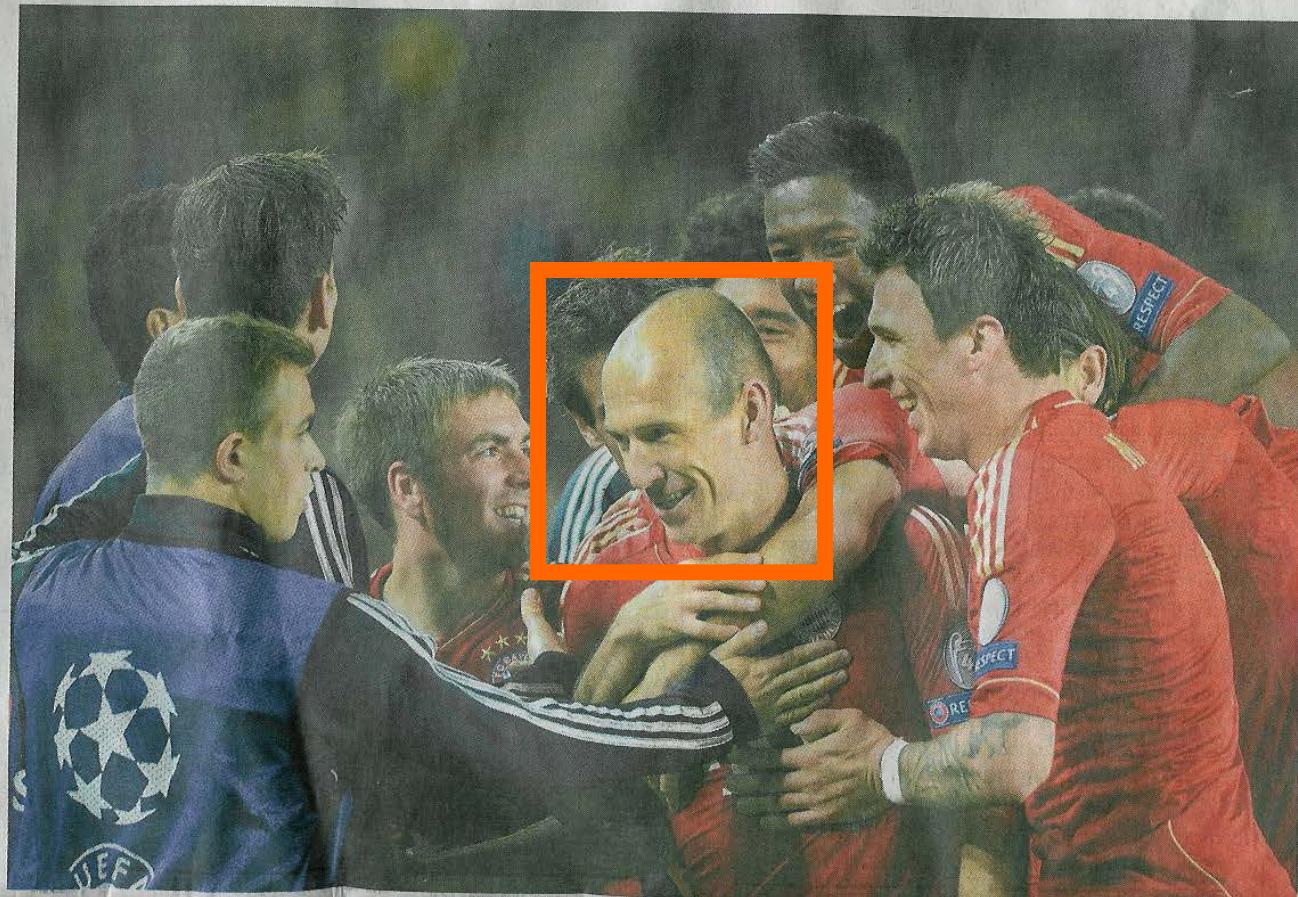
TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

Donnerstag, 2. Mai 2013 · 1,50 Euro

## Bayern und Dortmund im Champions-League-Finale

Der Pott geht in diesem Jahr auf jeden Fall nach Deutschland: Bayern München und Borussia Dortmund stehen im Finale der Champions League am 25. Mai im Londoner Wembley-Stadion. Im Halbfinal-Rückspiel beim FC Barcelona siegten die Münchener am Mittwochabend mit 3:0, nachdem sie bereits das Hinspiel mit 4:0 für sich entschieden hatten. Arjen Robben (auf dem Foto umringt von Mannschaftskameraden) und Thomas Müller erzielten die Tore, hinzu kam ein Eigentor von Gerard Piqué. Am Vortag hatte Dortmund bei Real Madrid in einem nervenaufreibenden Spiel zwar mit 0:2 verloren, doch reichte der 4:1-Sieg im Hinspiel für den Final-einzug. Mehr im Sportteil.

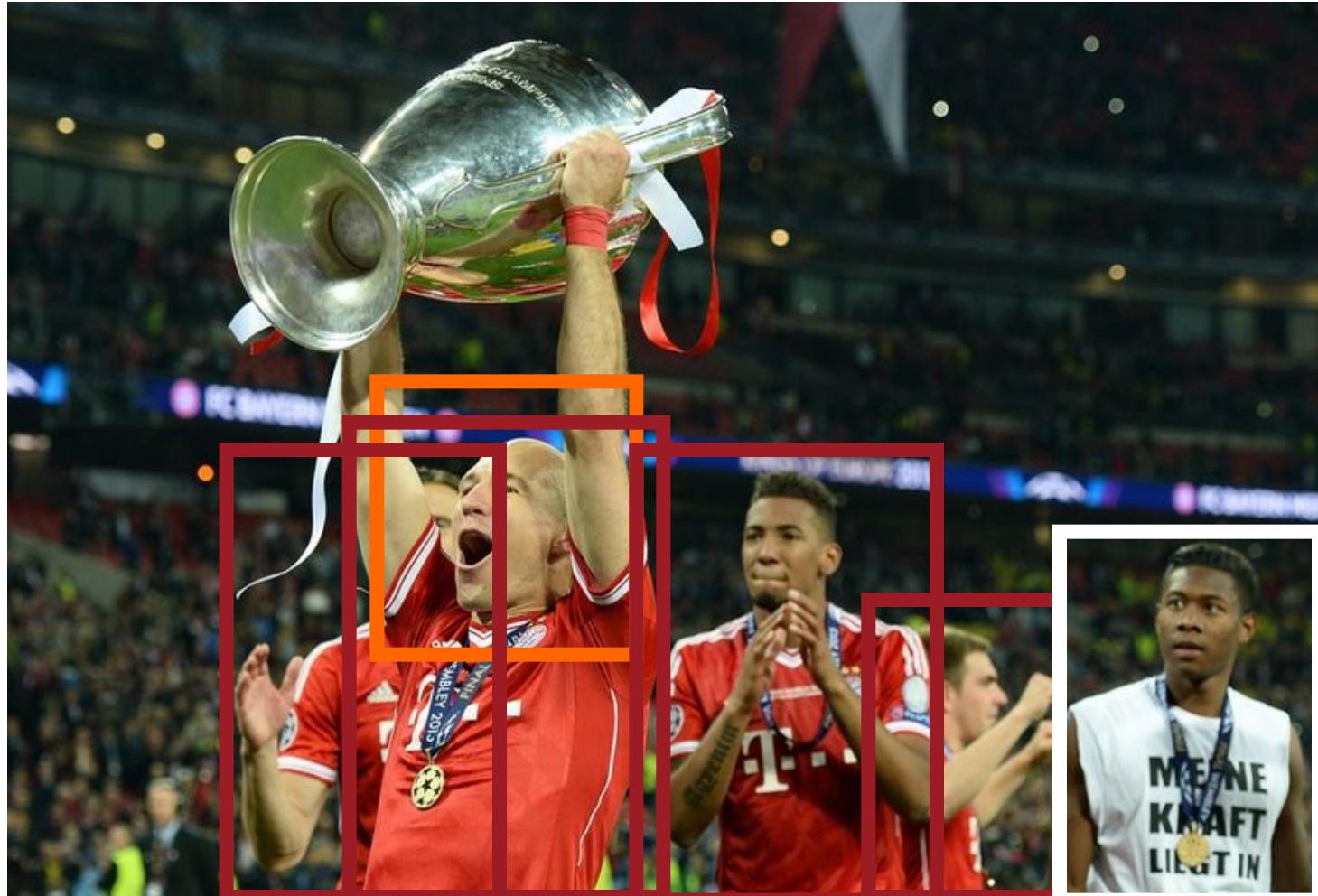
FOTO: DPA



Niederländer!

Bayern?!

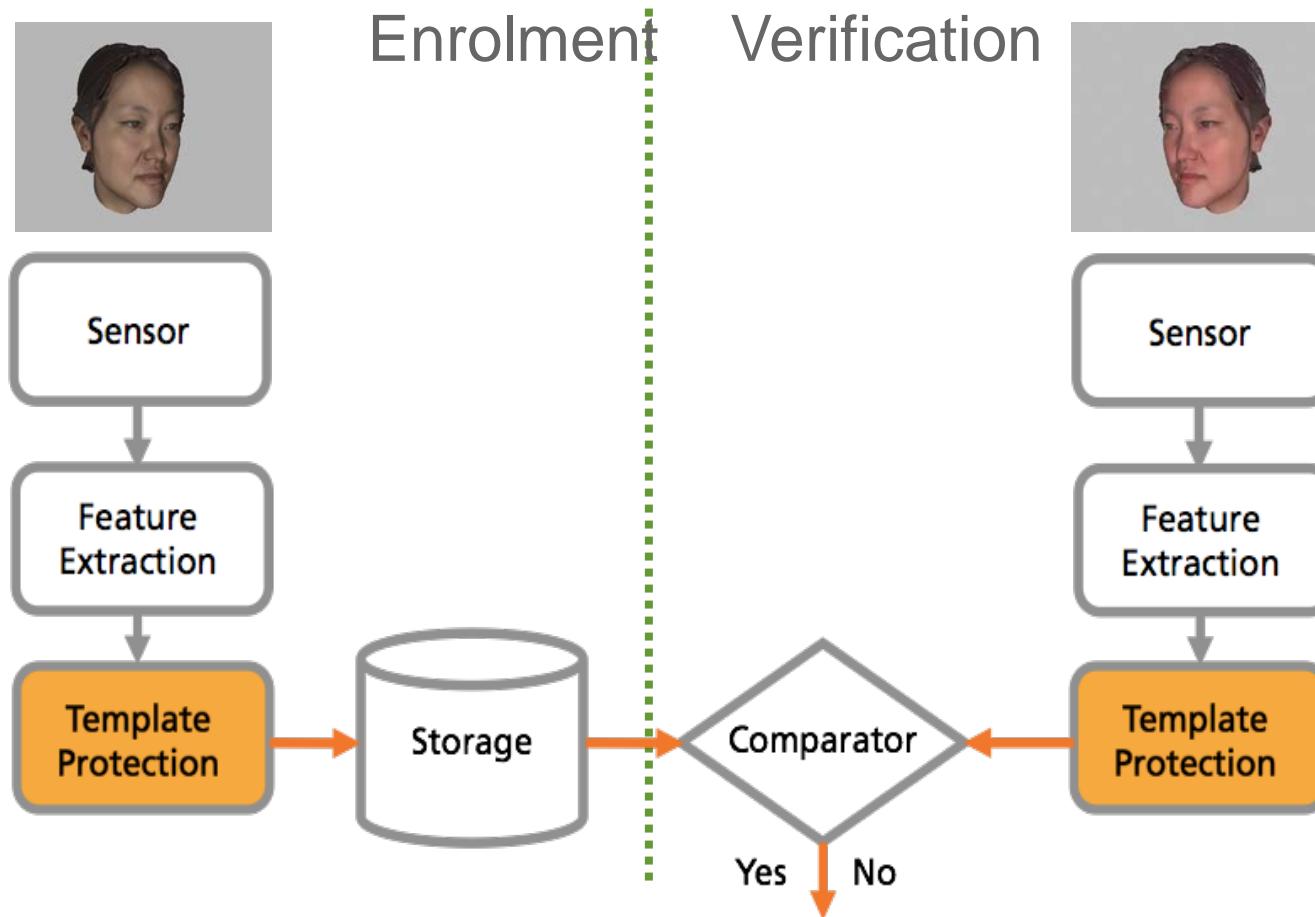
?!





- Gesichtserkennung zählt zu den biometrischen Verfahren.
  - Einsatzgebiete: sicherheitstechnisch, kriminalistisch und forensisch
  - Zweck: Identifikation oder Verifikation natürlicher Personen.
- Verifikation:
  - Die zu verifizierende Person muss dem System ihren Namen oder ihre User-ID mitteilen.
  - Danach entscheidet das [biometrische] System, ob die Person zum zugehörigen Referenzmerkmalsdatensatz gehört oder nicht.
- Identifikation:
  - Die zu erkennende Person offenbart ausschließlich ihr [biometrisches] Charakteristikum.
  - Das System ermittelt daraus durch Vergleich mit den Referenzmerkmalsdatensätzen aller Nutzer den zugehörigen Namen bzw. die User-ID.

# Biometric Systems (with Template Protection)





# Überblick

- Computer Vision
  - Objekterkennung und Objektdetektion als Fallbeispiel
  - Motivation heutiger Ansätze
- Bayes Decision Theory
  - Verwendung von Merkmalen zur Objekterkennung
  - Naive Bayes Klassifikator
- Gesichtsdetektion
  - Schneiderman & Kanade: Naive Bayes Klassifikator  
de facto einer der erfolgreichsten Verfahren
- **Diskussion und Ausblick**



# Diskussion und Ausblick

- Objekterkennung
  - lokale Merkmale UND deren Konstellation ist wichtig

# Beispiel



# Diskussion und Ausblick

## ■ Objekterkennung

- lokale Merkmale UND deren Konstellation ist wichtig
- für die allgemeine Objekterkennung (nicht nur Gesichter) benötigt man häufig aufwändiger Modelle
  - Bestandteile sind aber im wesentlichen auch Merkmale und deren Konstellation
  - Beispiel: Detektion von Personen in Sequenzen von Bildern:



## ■ Weiterführende Veranstaltungen

- Computer Vision (jeweils WS) und Computer Vision 2 (jeweils SS)
- Maschinelles Lernen: Statistische Verfahren (jeweils SS) und Maschinelles Lernen 2: Statistische Verfahren (jeweils WS)

# Demos & Links



- <https://www.youtube.com/watch?v=FuqFYssAE9sEin>  
Einfaches Bayes Klassifizierungsmodel mittels des Tool  
Rapidminer: <https://rapidminer.co>. Das Tool kann alles mögliche bezüglich  
Machine Learning hat den Focus aber nicht unbedingt auf Bilderkennung
- <https://de.mathworks.com/videos/computer-vision-system-toolbox-overview-61219.html>  
Die Computer Vision Toolbox für Matlab.
- Unser Favorit: <http://www.codeproject.com/Articles/3907/Creating-Optical-Character-Recognition-OCR-applica> ist schon etwas in die Jahre gekommen und sehr einfach gehalten. Das Tool an sich ist nur eine kleine Demo für ein Framework, dass mit Neuronalen Netzen arbeitet:
  - <https://github.com/AlexCherkasov/Neuro.NET> Der Sourcecode ist hier verfügbar und lässt sich auch mit neuen Visual Studioversionen problemlos compilieren und auch auf aktuellen Windows Versionen ausführen
  - Die Funktionsweise des Tools ist sehr einfach:
    - Erzeugung eines neuen Trainingspatterns
    - Erstellung eines neuen Neuronalen Netzes
    - Trainieren des Netzes (Hier besser keinen Noise hinzugebe, das kann dazu führen, dass das Lernen nicht terminiert bzw mehrere Stunden dauert. Auch schon bei sehr wenig Noise)
    - Eingabe von Buchstaben -> Netz erkennt sie (hoffentlich) zuverlässig. Hier lässt sich ebenfalls Noise einstellen



---

# Vielen Dank für die Aufmerksamkeit

(Und gleich zu den Übungen!)