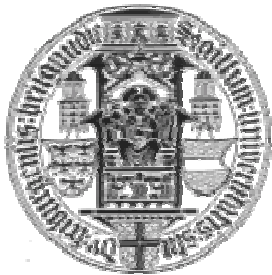


# Seminar

## Inhaltsbasierte Bildsuche

### (2+2 SWS)



Albert-Ludwigs Universität Freiburg  
Lehrstuhl für Mustererkennung und Bildverarbeitung  
Prof. Dr.-Ing. Hans Burkhardt

- Definition und Begriffe
  - Inhaltsbasierte Bildsuche
  - Kategorien
  - Semantic Gap
- Zielgruppen u. Suchparadigmen
- Entwicklung der inhaltsbasierten Bildsuche
- CBIR-Systeme
- Taxonomie Inhaltsbasierte Bildsuche
  - Anfragemethoden
  - Merkmale
  - Ähnlichkeit
  - Ergebnispräsentation
- Standardisierungsbemühungen
- Herausforderungen
- Verwandte Gebiete
- Beispiele für realisierte Systeme

## Definition:

„A technique for retrieving images from a large database on the basis of *automatically derived features*“.

=> Im Gegensatz zu Bildsuche mit Hilfe manuell erstellter Annotationen

## Probleme damit:

- „Ein Bild sagt mehr als 1000 Worte.“
- Ungenauigkeit / Subjektivität
  - => Schlagwortsysteme/Klassifikationscodes (anwendungsbezogen):
    - Getty Art and Architecture Thesaurus (AAT)
    - OPITZ Code
    - ICONCLASS
- Mancher Inhalt lässt sich schwer verschlagworten
- Sprachabhängigkeit
- (Schreib-) Fehler
- großer Arbeitsaufwand

=> Für sehr große, allgemeine Datenbanken nicht mehr machbar!

# ICONCLASS



Printer 712	_____
Boutesteyn, Cornelis	_____
Alias:	_____
Residence:	_____
Leiden from 1679 until 1710	_____
Device 186	_____
ICONCLASS:	_____
25H1123	_____
41A12	_____
47D31	_____
73C7455	_____

'House built upon a rock, h  
built upon sand': landscap  
castle on a rock; windmill i  
background

<u>Notation (Code)</u>	<u>Textual Correlate (Meaning)</u>
25H1123	rock-formations
41A12	castle
47D31	windmill
73C7455	'house built upon a rock; house built upon sand' « doctrine of Christ on love, etc. (Matthew 7:24-27; Luke 6:47-49)



## Einteilung:

Level 1: Suche mit Hilfe von primitiven Merkmalen (Farbe, Textur, Form)

Level 2: Suche mit Hilfe von logischen (abgeleiteten) Merkmalen:  
Suche nach Objekten (Haus, Hund, Katze) oder speziellen  
Objekten/Personen (z.B. „Freiburger Münster“, „G. W. Bush“)

Level 3: Suche nach abstrakten Attributen  
Suche nach benannten Ereignissen („Mauerfall“), Typen  
(Fußballspiel), Konzepten (Hass, Liebe, Freundschaft)

## Begriffe:

- ⇒ Content Based Image Retrieval
- ⇒ Visual Image Retrieval
- ⇒ Appearance Based Image Retrieval

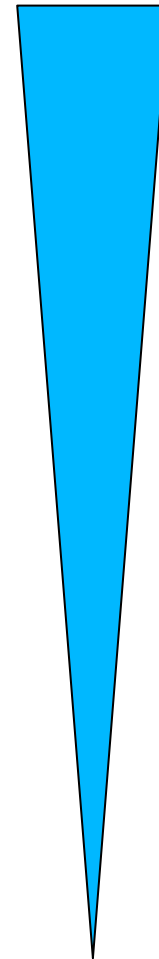
# Semantic Gap



- „Semantic Gap“  
„Lücke“, die zwischen Bildsuche den Stufen 1 und 2,3 entsteht.
- Auch:
  - ⇒ „sensory gap“
  - ⇒ „numerical gap“



- Internetsuche (Privatpersonen)
- (Bild-)Verlage (Verwaltung allg. Bilddatenbanken)
- Journalisten (Bild zu best. Thema)
- ...
- Designer (Material mit best. Farbe/Textur)
- Mode (ähnliche Stoffe)
- Ingenieure (ähnliche Bauteile, CAD)
- Medizin (Gewebe, Knochen, Zellen, Krebs)
- Architekten
- Kunsthistoriker
- ...
- Kriminalitätsprävention
  - Fingerabdrücke, Gesichtserkennung



allgemein

speziell



- Search by association
  - Benutzer hat nur sehr vage Vorstellung, was er will
  - Ziel: „interessante“ Bilder zu finden
  - „Durchstöbern“ der Bilddatenbank => hoch interaktiver Vorgang
- Target search
  - Benutzer weiß genau, welches Bild er sucht
- Category search
  - Man sucht irgendein Bild, das zu einer bestimmten Klasse gehört



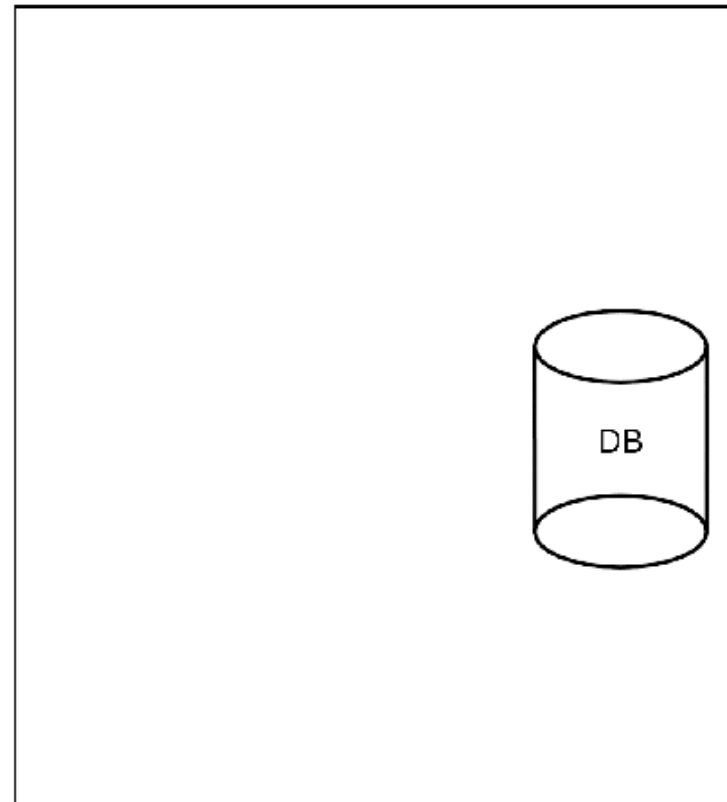
Wichtige Eigenschaft von CBIRS: Interdisziplinarität

Betroffene Gebiete:

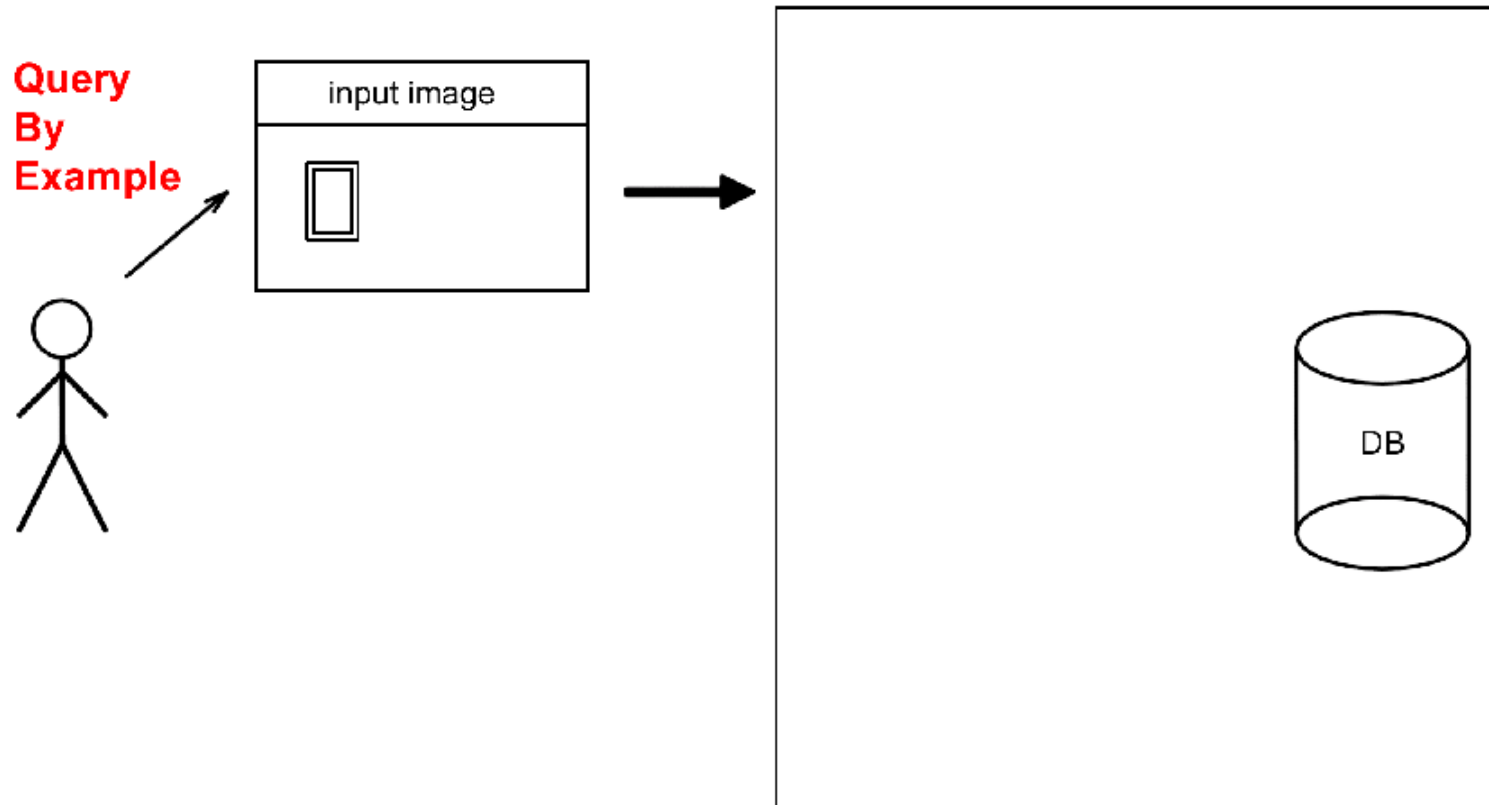
- Digitale Bildverarbeitung
  - Computer Vision
  - Mustererkennung
  - Information Retrieval
  - Künstliche Intelligenz
  - Datenbanktechniken
  - Psychologie
- 
- Beachtung aller Gebiete wichtig für ein erfolgreiches Bildsuchsystem!

- Zunächst Fokus auf textbasierten Suchtechniken (ab ca. 1979)  
Retrieval durch Standard-Datenbank/Textsuchtechniken
- Ab 1990
  - ⇒ Es werden neue Lösungen für die Verwaltung großer Bilddatenbanken gesucht (erster großer Workshop 1992)  
Ansatz: Verwendung der Bildinformationen direkt
- seit 1995  
Erste kommerzielle Systeme, z.B. QBIC von IBM

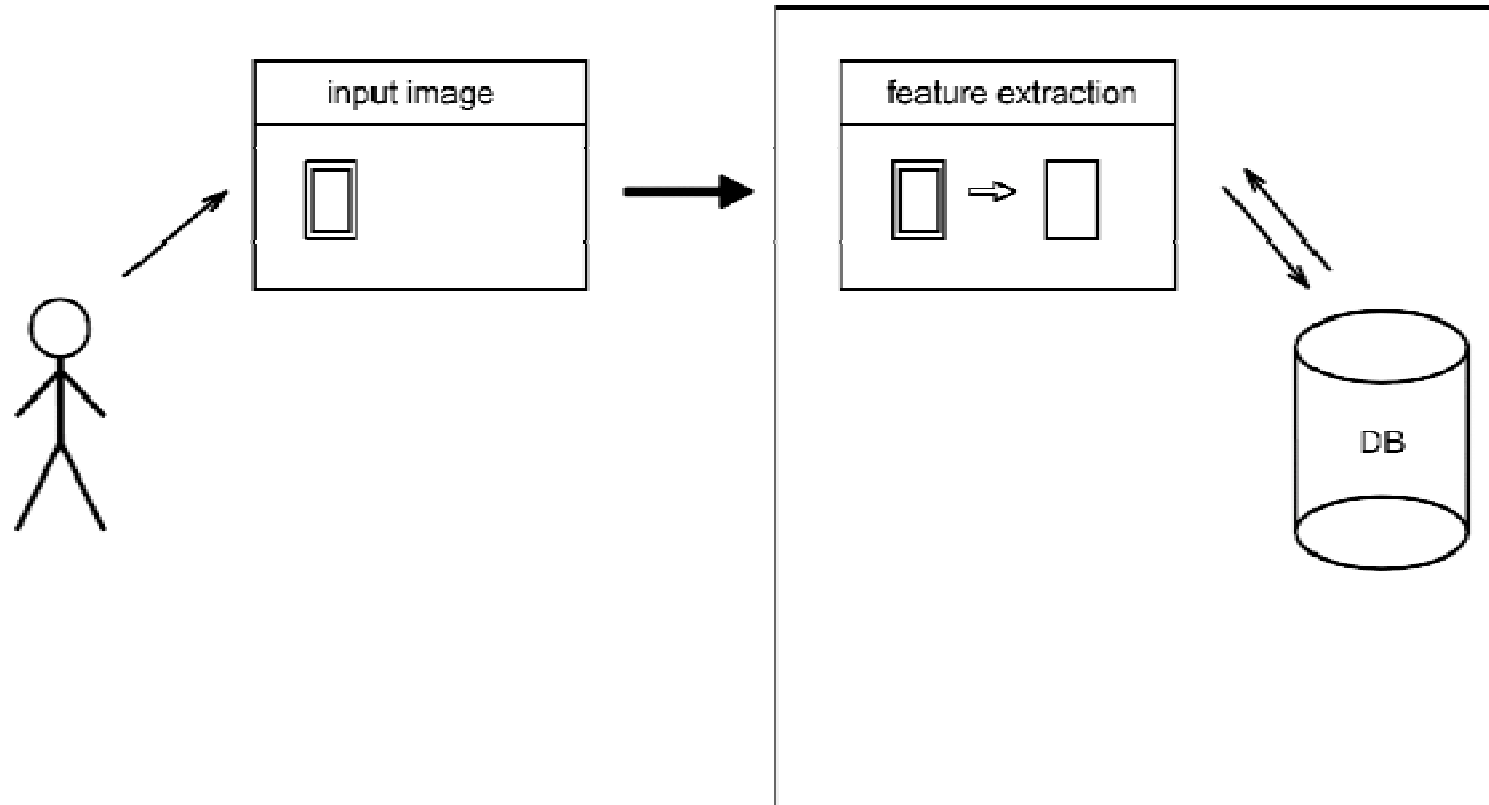
Ab dann: Explosion der Anzahl an Publikationen und Systemen!



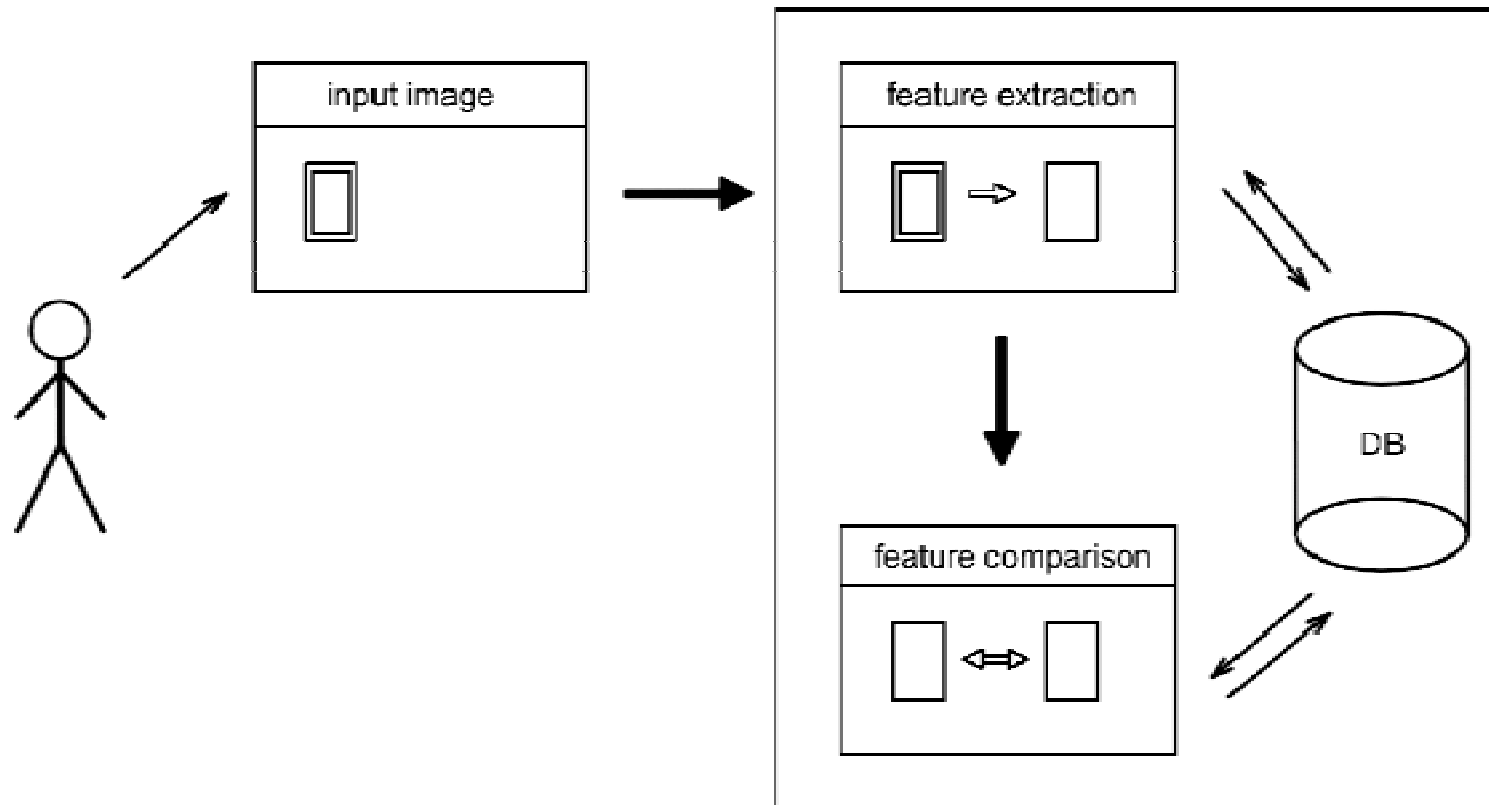
# CBIR - Systeme



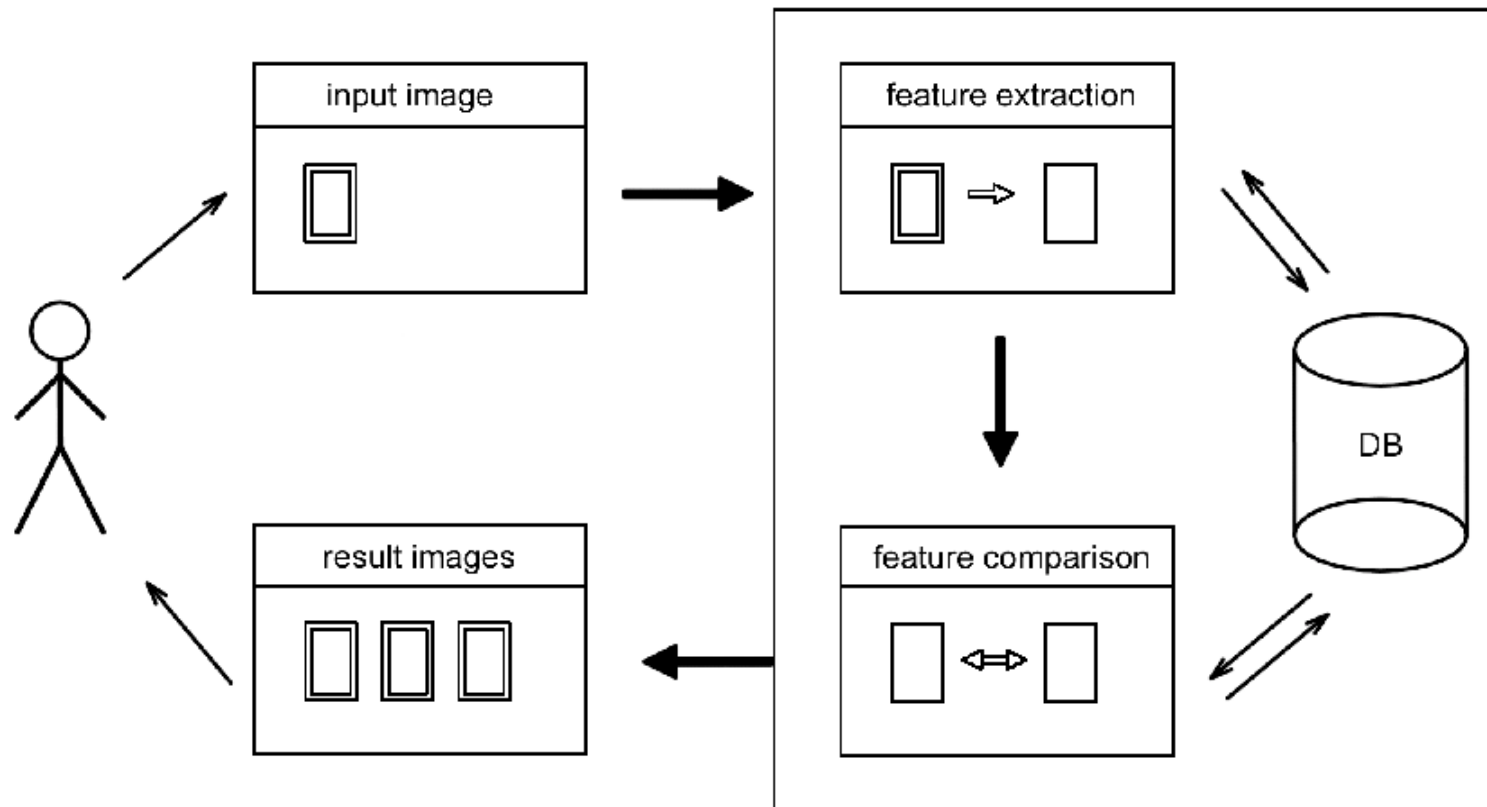
# CBIR - Systeme



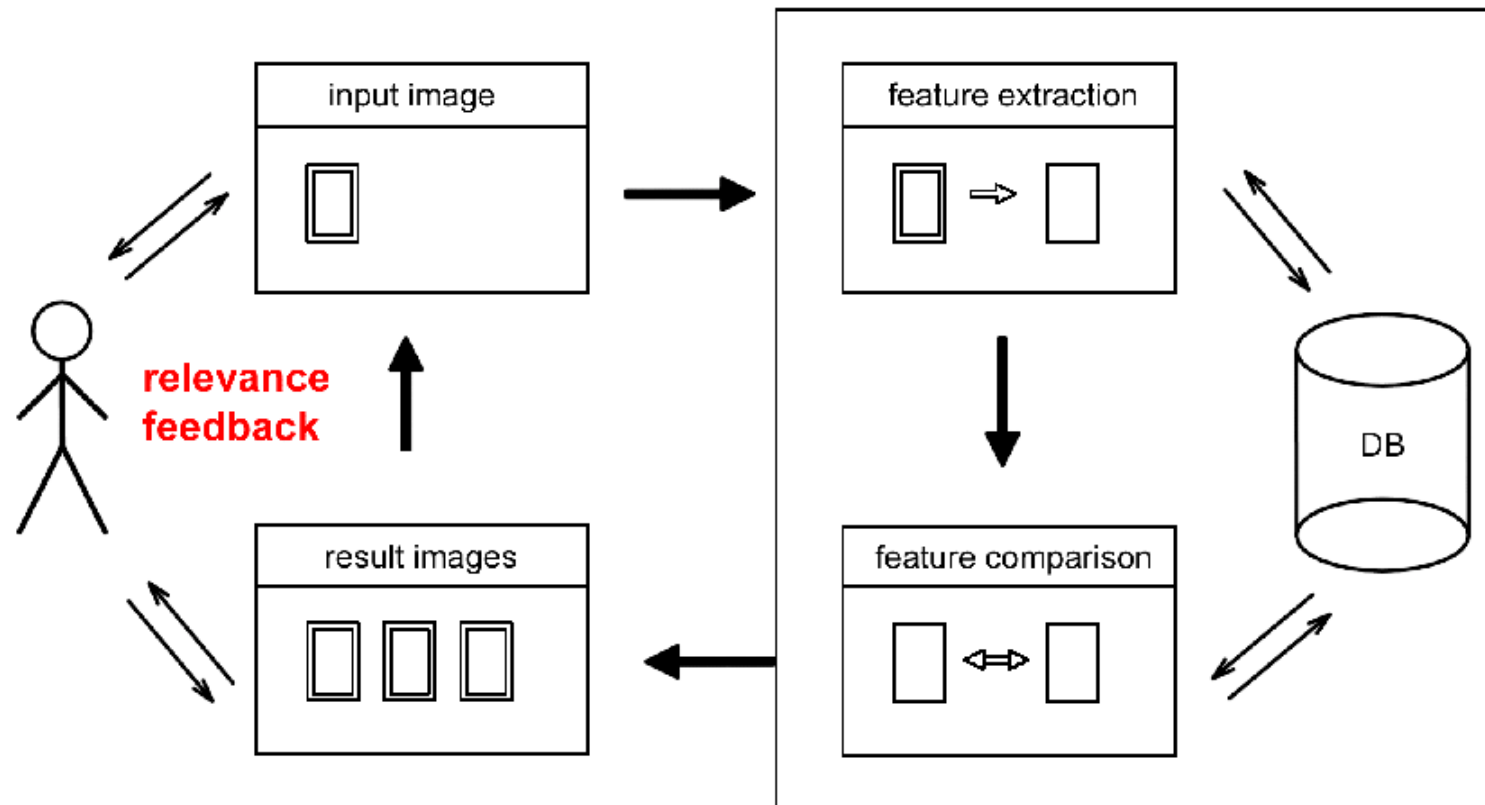
# CBIR - Systeme



# CBIR - Systeme



# CBIR - Systeme





- Beispiel für ein CBIR-System: [SIMBA](#)



Search template



## Anfrage- formulierung

- query by example
- query by sketch
- query by color
- query by color layout
- query by concept
- visual Thesaurus

## Merkmale

- Arten:
  - primitiv /fortgeschritten
  - global/lokal
- Farbe      Form      Segmentierung
- Textur     Kanten      salient points
- Anordnung      salient regions
- Invarianz
- Merkmalsauswahl und –fusion
- Speicherung und Zugriff:
  - Indexstrukturen
  - Aktualisierung
- Merkmalsdarstellung

## Ähnlichkeits- bestimmung

- Ähnlichkeitsmaße
- Distanzmaße
- Klassifikatoren

## Ergebnis- präsentation

- Listen
- Flächen, Räume
- Userinteraktion
  - Relevance Feedback
  - Lernfähigkeit

# Benutzerschnittstellen - Anfrage



- Query by example  
„image example“ oder „group example“



(VIPER)

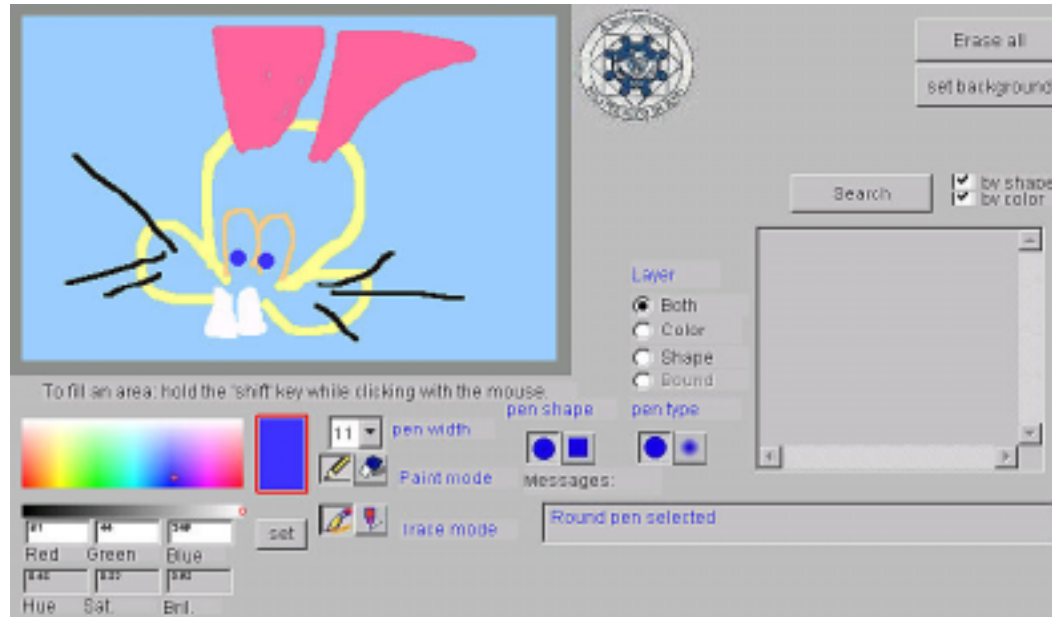
„page zero“ Problem!

- ⇒ erste(s) Bild(er) zufällig
- ⇒ Angabe von Farbverhältnissen
- ⇒ Aussuchen von Farben aus einer Palette
- ⇒ Kategorienangabe (vorherige Einteilung nötig)
- ⇒ Schlüsselwörter



(QBIC)

- Query by sketch

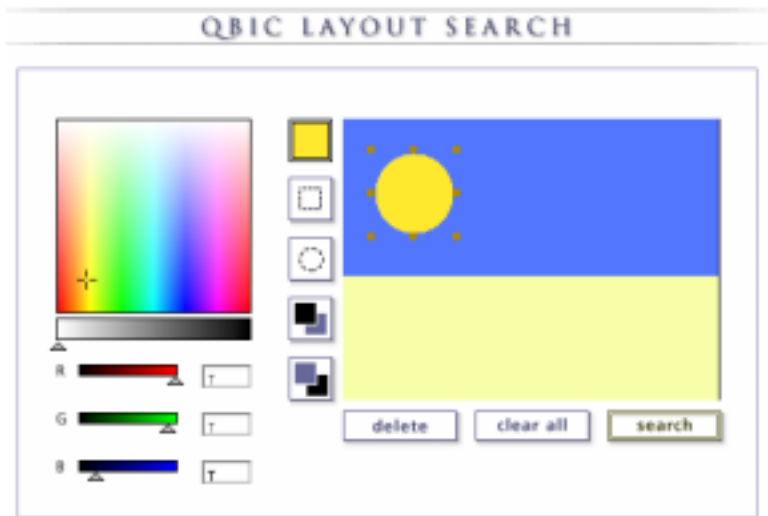


(DrawSearch)

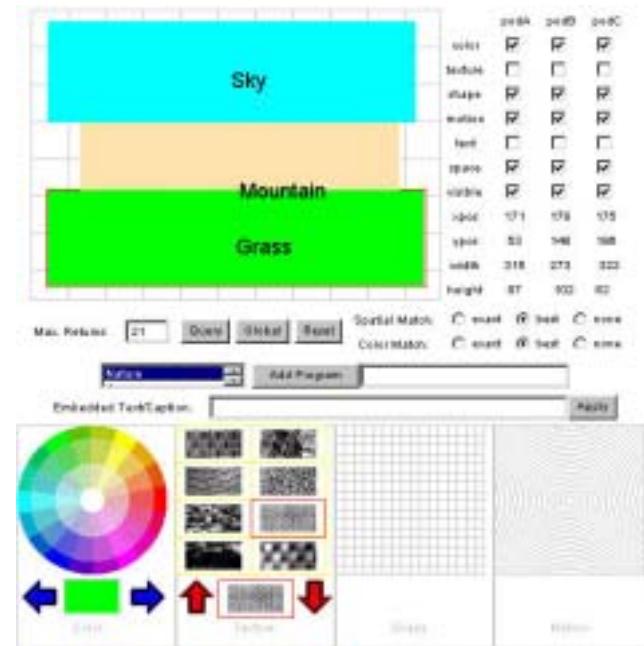
Probleme:

- ⇒ Künstlerische Begabung der Anwender?
- ⇒ Skizzen oft mehrdeutig
- ⇒ Texturen, Farbverläufe etc. schwer darstellbar
- ⇒ Merkmale i.A. nicht direkt mit denen der Bilder vergleichbar

- Query by color layout



(QBIC)



(VIRage)

- „Vereinfachte“ Form des Query by sketch
- Im wesentlichen die selben Probleme

# Benutzerschnittstellen - Anfrage



- Query using a Visual Thesaurus
  - Suche nach einem „mental image“



(N. Boujema, INRIA)

## Anfrage- formulierung

- query by example
- query by sketch
- query by color
- query by concept
- query by concept
- visual Thesaurus

## Merkmale

- Arten:
  - primitiv /fortgeschritten
  - global/lokal
- Farbe      Form
- Textur     Kanten
- Anordnung
- Invarianz
- Segmentierung
- salient points
- salient regions
- Speicherung und Zugriff:
  - Indexstrukturen
  - Aktualisierung

## Ähnlichkeits- bestimmung

- Ähnlichkeitsmaße
- Distanzmaße
- Klassifikatoren

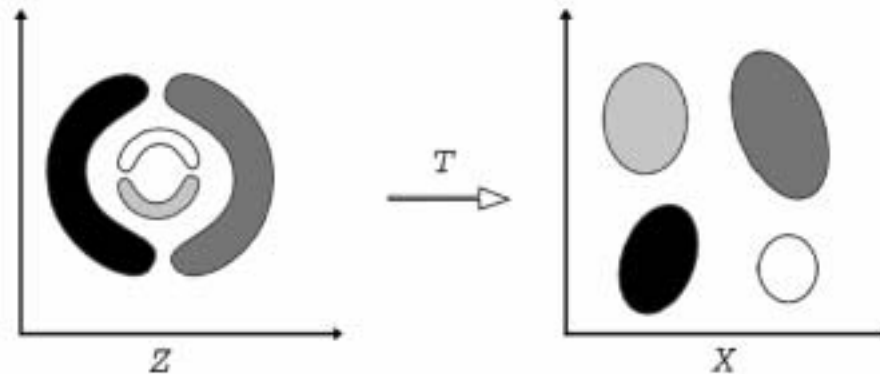
## Ergebnis- präsentation

- Listen
- Flächen, Räume
- Userinteraktion
  - Relevance Feedback
  - Lernfähigkeit

- Erfassung von Eigenschaften des Bildes
- Arten:
  - Einfache, allgemeine Merkmale: („Level 1“), z.B.
    - Farbe
    - Textur
    - Form
    - Kanten
  - ⇒ Jeweils viele verschiedene Ausprägungen
  - ⇒ Viele Merkmale lassen sich nicht strikt einer Klasse zuordnen  
z.B. Integralinvarianten enthalten sowohl Textur- als auch Farbinfo
  - „Spezialmerkmale“
    - Gesichtserkennung
    - Fingerabdruckerkennung

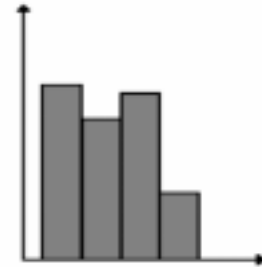


- Normalerweise werden die Pixel nicht direkt verglichen (Gegensatz zu OCR in best. Fällen)
- Transformation des Bildes in eine andere Repräsentation
  - Hoffnungen:
    - Bessere Diskriminierungsfähigkeit



- Kompaktere Darstellung (z.B. mit PCA)

- Wahrscheinlichkeiten, Dichtefunktionen
  - Histogramm
  - Parametrische Darstellung  
(z.B. multidimensionale Gaußverteilungen)
- Merkmalsvektoren
  - Punkte in hochdimensionalen Räumen





# Merkmale - lokal



- Segmentierungsmethoden

- Starke Segmentierung
  - Modellbasierte Ansätze (Vorwissen!)
- Schwache Segmentierung => salient regions
  - Clustering im Featureraum
    - K-means,
    - Expectation Maximization (Blobworld)
  - Graphentheoretische Ansätze
    - Normalized Cuts
  - Edge Flow
  - ...

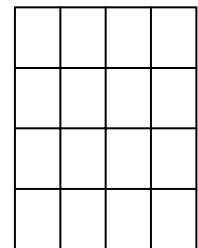
$T = O$

$T \ll O$



- Partitioning
  - Blöcke festgelegter Größe
  - Gleiche oder verschiedene Ausschnitte

$T \neq O$



- Salient points (besonders informationstragende Stellen im Bild)



- Es gibt eine Vielzahl von
  - verschiedenen Merkmalen
  - verschiedenen Merkmalsdarstellungen
  - Methoden, wie Merkmale verglichen werden
- Fragestellung
  - ⇒ Welche Merkmale/Vergleichsmethoden sollen benutzt werden?
  - ⇒ Sollen immer die selben Merkmale verwendet werden, oder kann der Benutzer entscheiden bzw. das System lernen, was „gut“ ist?

- Standard-Datenbankentechniken aus der Textsuche sind i.a. nicht mehr anwendbar
    - Von Suchbegriffen (Wörtern, Wortteilen) kann auf Dokumente verlinkt werden
    - Die Merkmale der Bilder sind jedoch meist floating-point Vektoren und haben an sich keine Bedeutung!
  - ⇒ Standardtechnik: Ähnlichkeitsberechnung zu allen Bildern in der Datenbank
  - ⇒ für sehr große Datenbanken (z.B. Internet) nicht machbar, Suchdauer wächst linear mit der Datenbankgröße
- Ansätze:
- Clustering
  - Baumstrukturen

## Anfrage- formulierung

- query by example
- query by sketch
- query by color
- query by concept
- query by concept
- visual Thesaurus

## Merkmale

- Arten:
  - primitiv /fortgeschritten
  - global/lokal
- Farbe      Form
- Textur     Kanten
- Anordnung
- Invarianz
- Segmentierung
- salient points
- salient regions
- Speicherung und Zugriff:
  - Indexstrukturen
  - Aktualisierung

## Ähnlichkeits- bestimmung

- Ähnlichkeitsmaße
- Distanzmaße
- Klassifikatoren

## Ergebnis- präsentation

- Listen
- Flächen, Räume
- Userinteraktion
  - Relevance Feedback
  - Lernfähigkeit

- Vielzahl an Möglichkeiten, die Ähnlichkeit von Bildern zu bestimmen
- Ähnlichkeit ist bei Bildern höchst subjektiv!



- Bei CBIR-Systemen muss klar sein, welche Art der Ähnlichkeit zugrunde liegt (Ähnlichkeit nach Farbe, Textur, Kombination, Semantik...)

⇒ Nichtbeachtung führt zu enttäuschten Benutzern!



## Anfrage- formulierung

- query by example
- query by sketch
- query by color
- query by color layout
- query by concept
- visual Thesaurus

## Merkmale

- Arten:
  - primitiv /fortgeschritten
  - global/lokal
- Farbe      Form      Segmentierung
- Textur      Kanten      salient points
- Anordnung      salient regions
- Invarianz
- Merkmalsauswahl und –fusion
- Speicherung und Zugriff:
  - Indexstrukturen
  - Aktualisierung
- Merkmalsdarstellung

## Ähnlichkeits- bestimmung

- Ähnlichkeitsmaße
- Distanzmaße
- Klassifikatoren

## Ergebnis- präsentation

- Listen
- Flächen, Räume
- Userinteraktion
  - Relevance Feedback
  - Lernfähigkeit

# Ergebnisrepräsentation



- Listen

The screenshot displays the MICELscope search interface. On the left, the 'Query image' section shows a German postage stamp with two figures in traditional costumes. Below it, the 'Search method' section includes a 'Color' dropdown menu, a 'Textual' checkbox, and three checked checkboxes: 'Motive', 'Size', and 'Aspect Ratio'. The 'Database' dropdown is set to 'BRD'. A 'Start Query' button is at the bottom of this panel. The 'Results' section on the right shows a grid of 15 search results, each a postage stamp with a numerical label below it. The stamps vary in design, color, and value, including stamps with figures, symbols, and text like 'DEUTSCHE BUNDESPOST' and 'WOLFFELT'S MARKE'.

MICELscope

# Ergebnisrepräsentation



- Flächen
  - „Durchstöbern“ des Fine Arts Museums San Francisco ([link](#))



Perceptual Image Similarity  
Experiments Rogowitz et. al.

- Räume



Perceptual Image Similarity  
Experiments Rogowitz et. al.

⇒ Für Benutzer vermutlich schwer handhabbar

- Interaktionsmöglichkeiten für den Benutzer
  - Auswahl bestimmter Merkmale
    - ⇒ Ein Spezialmerkmal ist meist nicht ausreichend!
    - ⇒ (Minka & Pikard: „Society of Models“)



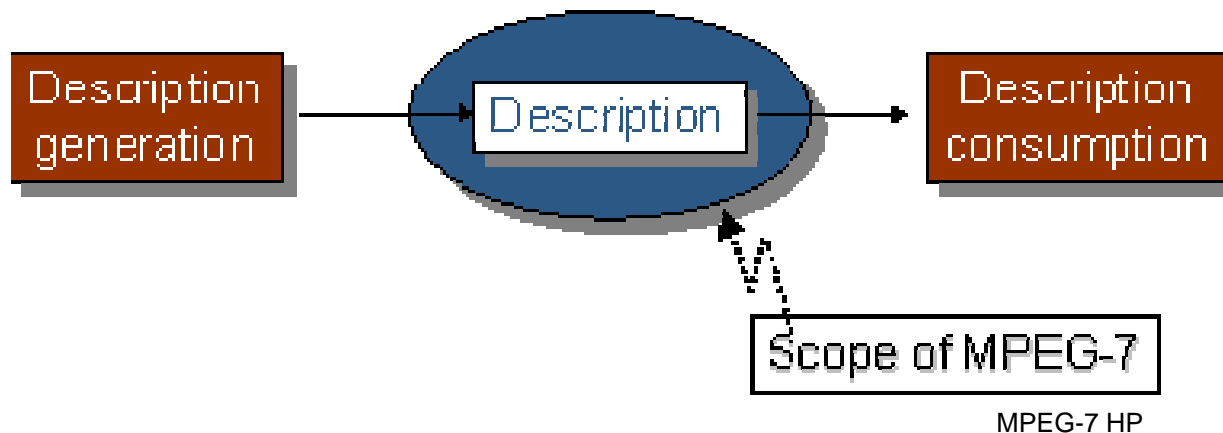
MICHELscope



JP2FF

- Einstellung von Gewichten
- Transparenz der Parameterauswahl
- Lernen aus Benutzereingaben (kurzfristig, langfristig)

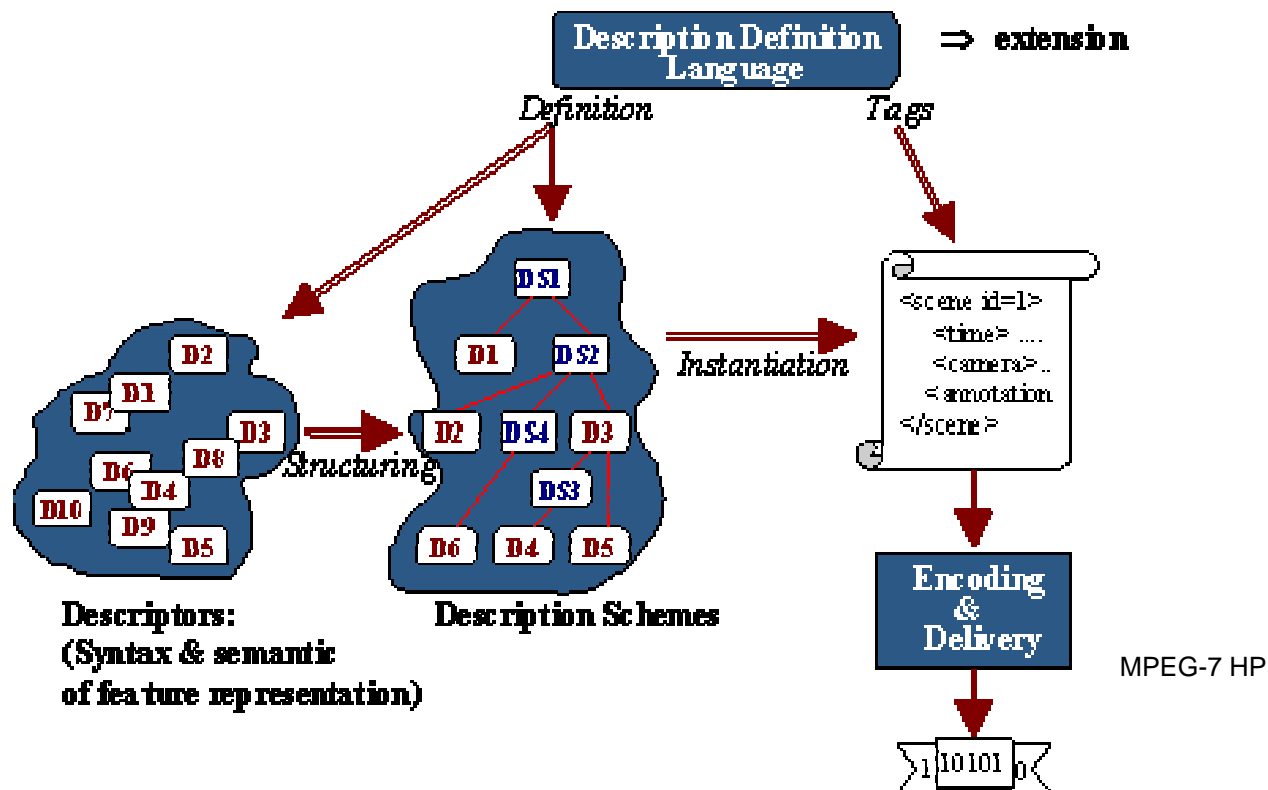
- MPEG-7 “Multimedia Content Description Interface“  
⇒ Ziel: Beschreibung des Inhalts von Multimediaobjekten aller Art (nicht ausschließlich für Videos)



# Standardisierungsbemühungen



- MPEG-7 “Multimedia Content Description Interface“
  - ⇒ Ziel: Beschreibung des Inhalts von Multimediaobjekten aller Art (nicht ausschließlich für Videos)



- MPEG-7 “Multimedia Content Description Interface“
  - ⇒ Beschreibung von:
    - Low-level Beschreibungen jedes Objekts in der Szene (Form, Größe, Farbe, Position und Bewegung)
    - High-level Beschreibungen der Szene (enthaltene Objekte und Vorgänge)
    - Andere Metadaten: Autor, Erstellungsdatum, etc.
- MPEG-7 Visual
  - Toolbox für die visuelle Beschreibung von Daten (Descriptors und Description Schemes):
    - Farbe: Farbraum, Quantisierung, dominante Farbe(n), skalierbares Farblayout, GoF/GoP Farbe
    - Textur: Homogene Texturbeschreibung (Gaborfeatures), Kantenhistogramm
    - Form: Lokale Form, Form der Kontur, 3D-Form
    - Bewegungsbeschreibung
    - Spezialmerkmale für Gesichtserkennung



- MRML (Multimedia Retrieval Markup Language) [Müller et. al.]
  - Betrifft die Kommunikation zwischen verschiedenen Retrievalkomponenten
  - ⇒ Vorschlag einer strukturierten Methode, Anfragen an einen Server zu schicken (Bilder, Relevanzwerte, Algorithmen) und Ergebnisse zurück zu bekommen.

```
<mrml
  session-id = "1"
  transaction-id = "44" >
  <query-step
    session-id = "1"
    resultsize = "30"
    algorithm-id = "algorithm-default" >
    <user-relevance-list >
      <user-relevance-element
        image-location = "http://viper.unige.ch/1.jpg"
        user-relevance = "1" />
      <user-relevance-element
        image-location = "http://viper.unige.ch/2.jpg"
        user-relevance = "-1" />
    </user-relevance-list>
  </query-step>
</mrml>
```

⇒ Semantische Bildsuche (Level2)

Verbindung von low-level Merkmalen zu high-level Konzepten

Wie weit kann man unter Benutzung von primitiven Merkmalen kommen?

Zutaten:

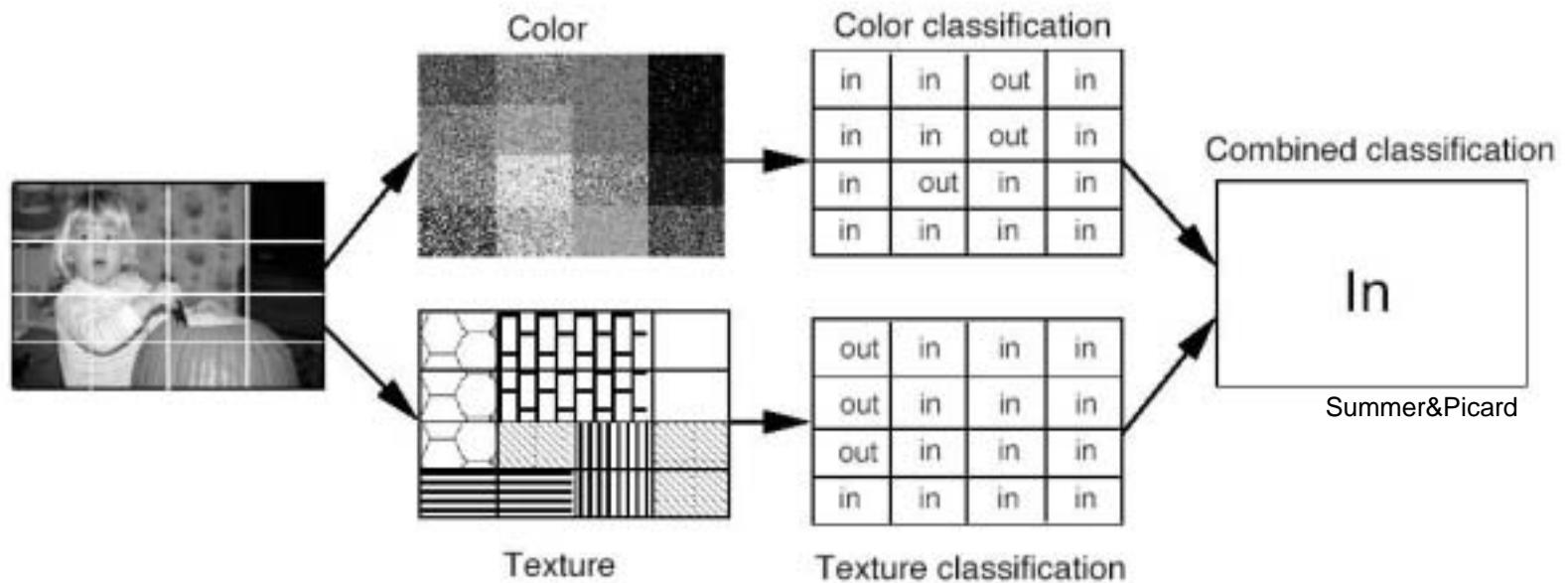
- Leistungsfähige Merkmale
- Einbringen von Vorwissen („lernen“)

# Herausforderungen



Beispiele für semantische Erkennung:

- Szenenerkennung (drinnen/draußen)

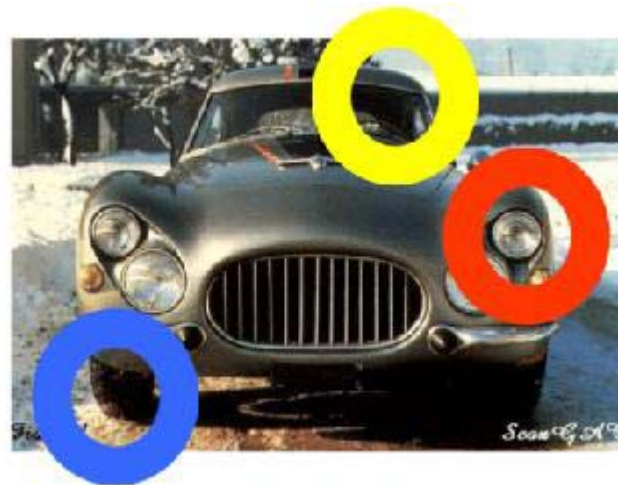


# Herausforderungen



Beispiele für semantische Erkennung:

- Automatisches Erkennen von Objektkategorien bzw. Objekten (Auto, Motorrad, etc.)

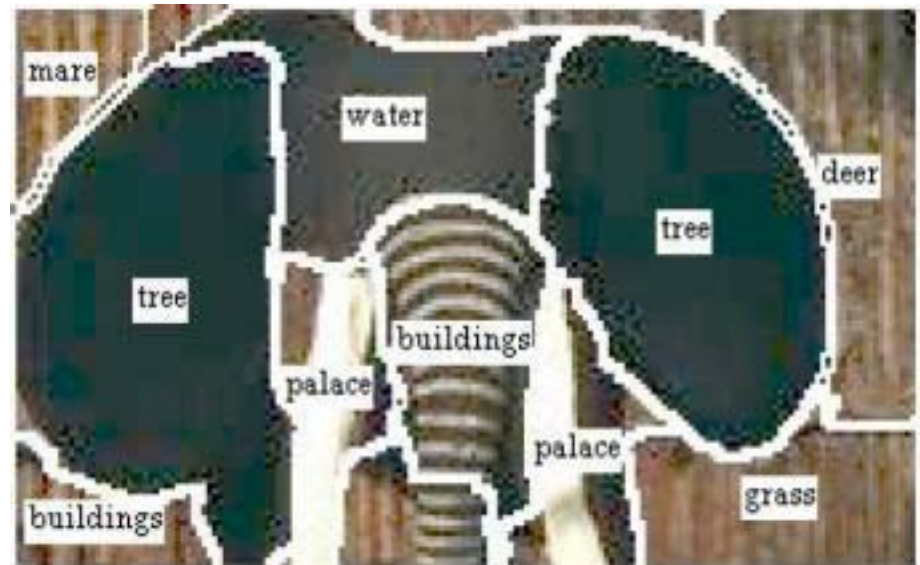
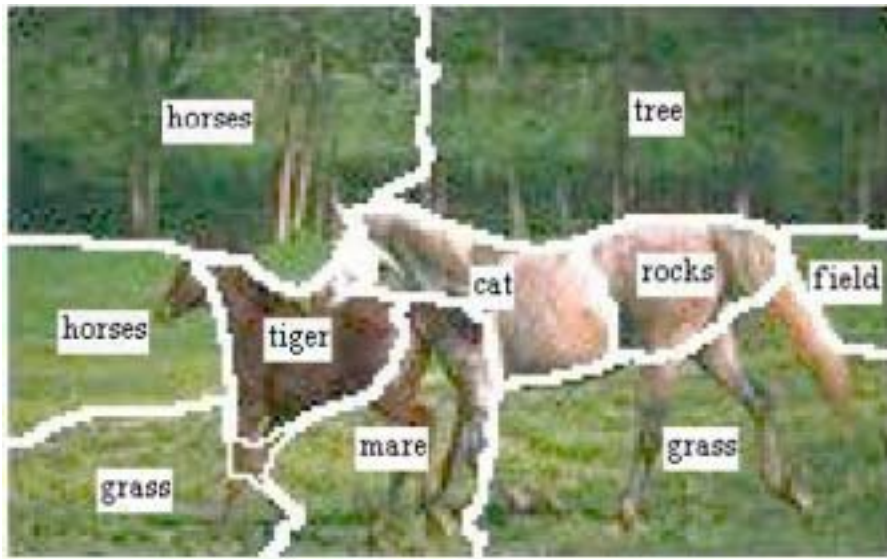


Perona et. al.

# Herausforderungen



- Automatische Verschlagwortung von Bildern/Bildregionen
  - James Wang (ALIp => ganzes Bild)
  - Kobus Barnard (Bildregionen)



[K. Barnard]

# Systeme (Beispiele)

---



- QBIC
- Simba
- Blobworld
  
- Viper/Gift
- SIMPLicity
- WebSeek
- Ikona (INRIA)
- FIRE

- R. Veltkamp, H. Burkhardt, H. Kriegel (Eds.); State-of-the-Art in Content Based Image and Video Retrieval; Kluwer Academic Publishers, 2001
- J. Eakins, M. Graham; Content-based image retrieval; Technical Report, University of Northumbria at Newcastle, 1999
- W. Smeulders et al.; Content-Based Image Retrieval at the End of the Early Years, TPAMI Vol. 22, No. 12, Dec. 2000
- Y. Rui, T. Huang, S. Chang; Image retrieval: Current techniques, promising directions and open issues; Journal of Visual Communication and Image Representation, 1999
- R. Duda, P. Hart and G. Stork; Pattern Classification, Wiley Interscience, 01.