Visuelle Perzeption für Mensch-Maschine Schnittstellen

Vorlesung, WS 2012 / 2013

Prof. Dr. Rainer Stiefelhagen

Institut für Anthropomatik
Karlsruher Institut für Technologie - KIT

http://cvhci.anthropomatik.kit.edu/

rainer.stiefelhagen@kit.edu

Computer Vision for Human-Computer Interaction
Research Group, Universität Karlsruhe (TH)

cv:hci

Basics:
Computer Vision &
Image Understanding

WS 2012/13

Dr. Saquib Sarfraz

saquib.sarfraz@kit.edu

- The purpose of neo-cortex is three fold

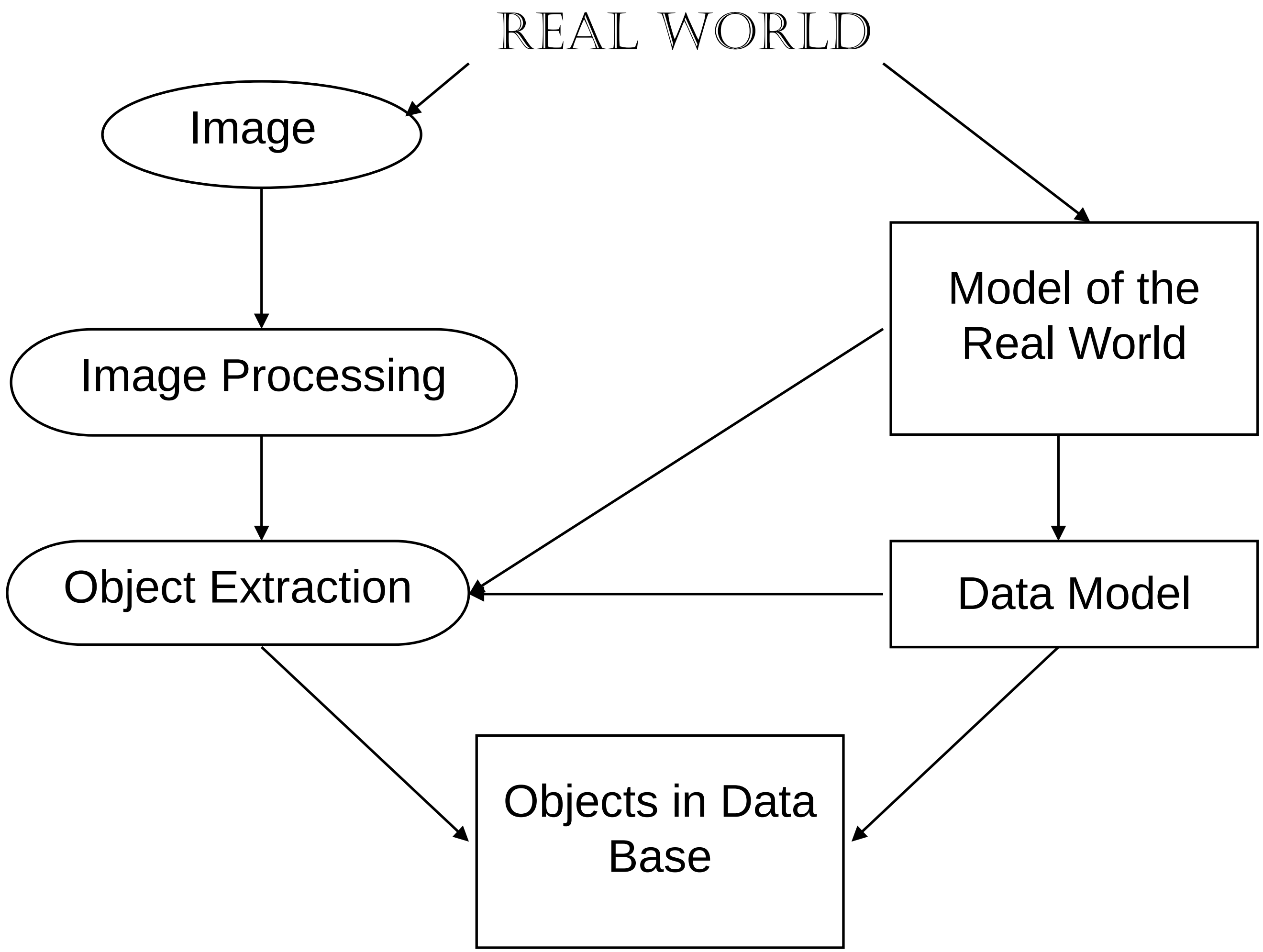
to RECOGNIZE

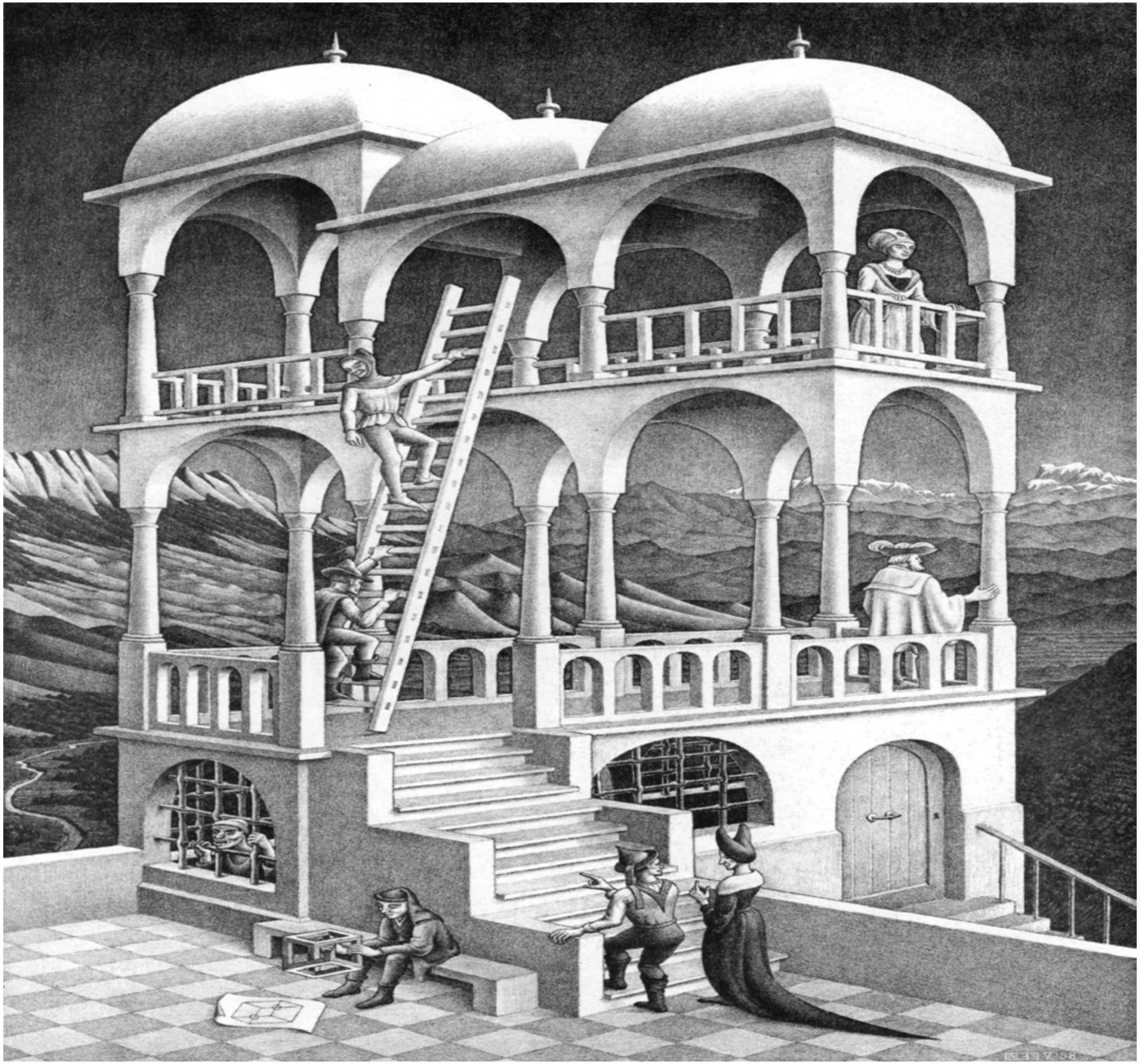
OBJECTS

OBJECTS

OBJECTS

Objects from Digital Images





Poseidon and the Mermaid, by David M. Weimer, 1990. This India-ink drawing is ambiguous: It may be perceived either as Poseidon, the Greek god of the sea, or as a mermaid. The mermaid's tail fin is Poseidon's moustache. (Courtesy, David M. Weimer.)

Introduction
Image Understanding: Definition

- Generation of an
  explicit, meaningful description
  of **physical objects**
  in one or several **images**
  at a certain point in **time** or period of time
  of the imaged **scene**
  with the help of computers.

  - Physical objects are in the scene.
  - Images of the scene are taken as input.
  - Explicit description of the scene is the result.

- Aim: complete automation

Introduction
Image Understanding: Object

- **Main questions:**
  - Why objects?
  - What are objects?
- **Holistic view:**
  - Everything is a whole.
- **Discrimination of objects:**
  - Creatures differentiate among important things to improve their survival capabilities.
- **Object:**
  - Homogeneous area of the scene

Introduction
Image Understanding: Concepts

- **Scene:**

  Part of time and space in the **real world.**
  - Static: 3D (e.g. interior, building and environment, landscape, country, planet, universe, inside of humans)
  - Dynamic: 3D + time
  - Also known as **object space**
- **Image:**

  2D projection of the scene
  - Intensity image (active: radar, microwave, passive: mono-/multi-spectral)
  - Depth image

Introduction
Knowledge-based object extraction

- **Object extraction:**

  Is based on what we **know** about an object

  - What kind of characteristics/details are there?
  - What is the difference with respect to other objects?
  - What are the relations to other objects?

    Object model

- Object model depends on human cognition of the world;
  i.e., the result of the automatic object extraction should be identical to the interactive (human) interpretation.
  This does not mean that the methods used have to be the same!

  Formal representation of knowledge

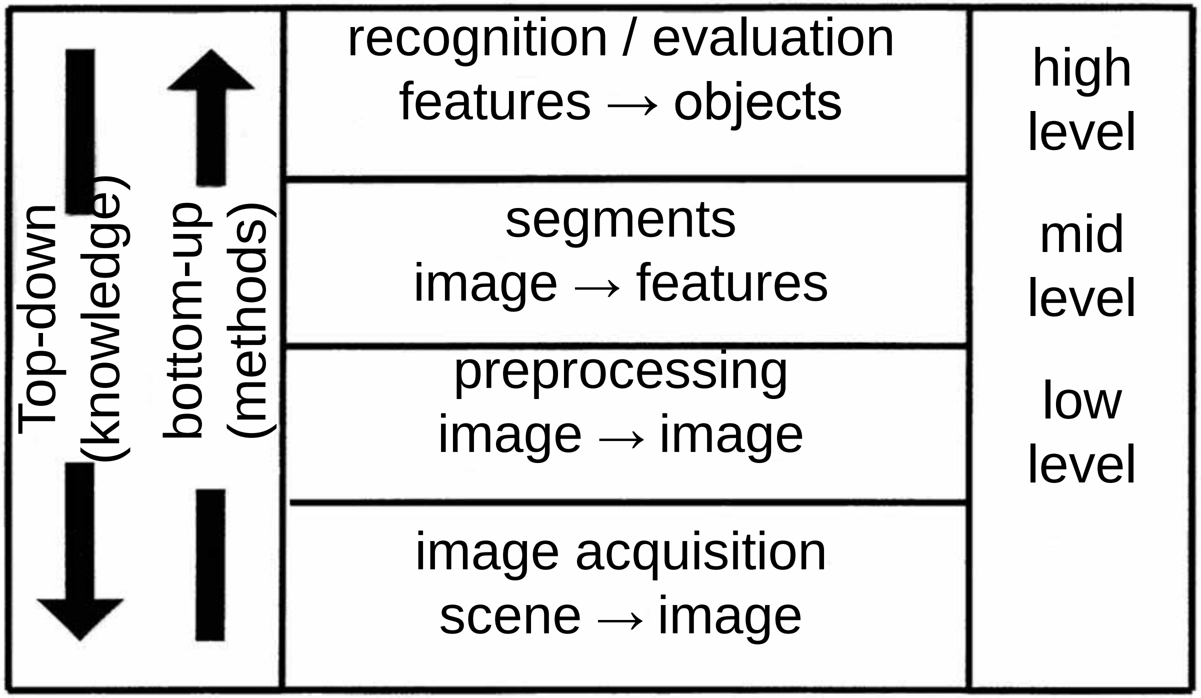
    Techniques of artificial intelligence

Utilization of Model: Standard Approach to Image Understanding

- Problem: Given are
  - Formal representation of the model
  - Image(s) of the scene

  ⇒ Compute instances from (instantiate) concepts of the representation.
- Operations:
  - Extract characteristics (invariants) from images (description)
  - Project concepts on the characteristics (prediction)
  - Compare characteristics with the representation (matching)

Utilization of the Model: Standard Approach: Levels



- Image formation

  (Input: scene, result: image)

  - Digital data (e.g. CCD-camera, Landsat TM, MRI,CT, SAR)
  - **Scanner**

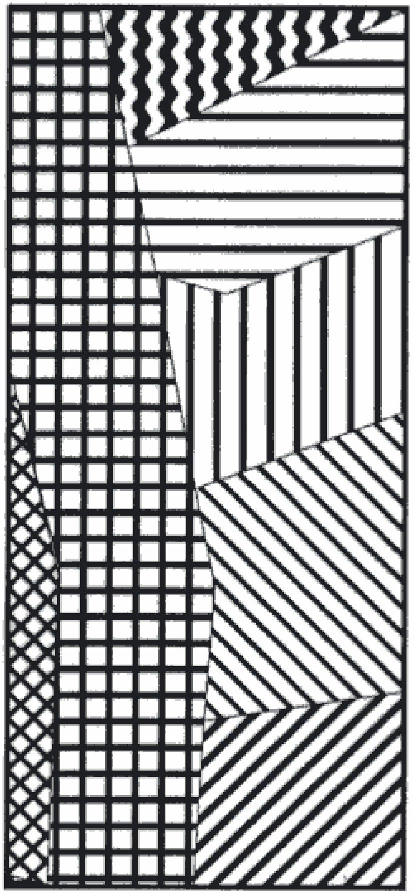
Utilization of Model
Standard Approach: Levels

Preprocessing (low-level; iconic)
(Input: image, result: image)

- Image enhancement (visual image quality; e.g. contrast improvement, change of LUT)
- Image restoration (physical image quality: e.g. removal of noise/disturbances, blur)

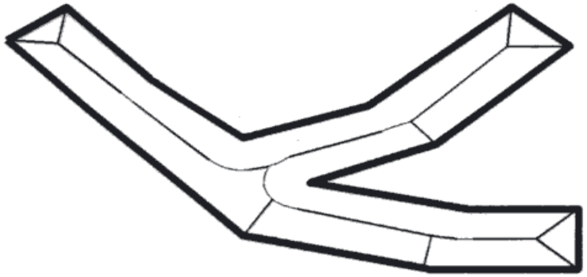
Segmentation (mid-level)
(Input: image, result: features)

- For area (2D) objects: Subdividing pixels into a set of non-overlapping, non-empty **regions** filling the image plain without any gaps (features: contours)

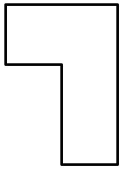


Utilization of Model
Standard Approach: Levels

- Segmentation (contd.)
  - Conditions for creation of **regions**:
    e.g. same color value, same texture or hatching.
  - In case of linear objects the **centerlines** are the **features**.



- Recognition/evaluation (high-level)
  (input: attributes, results: objects)
  - Explicit knowledge about attributes and relations is used to match features and objects.
  - Example: Contour of a region gains meaning “house”,
    if it does not exceed a certain area size and is composed of
    straight lines oriented perpendicularly to each other.



Utilization of Model: Standard Approach: Procedure

Bottom-up (data driven; based on methods)

- Recognize objects starting with processing the image.

Top-down (model driven; based on knowledge)

- Project model on image and search for components.

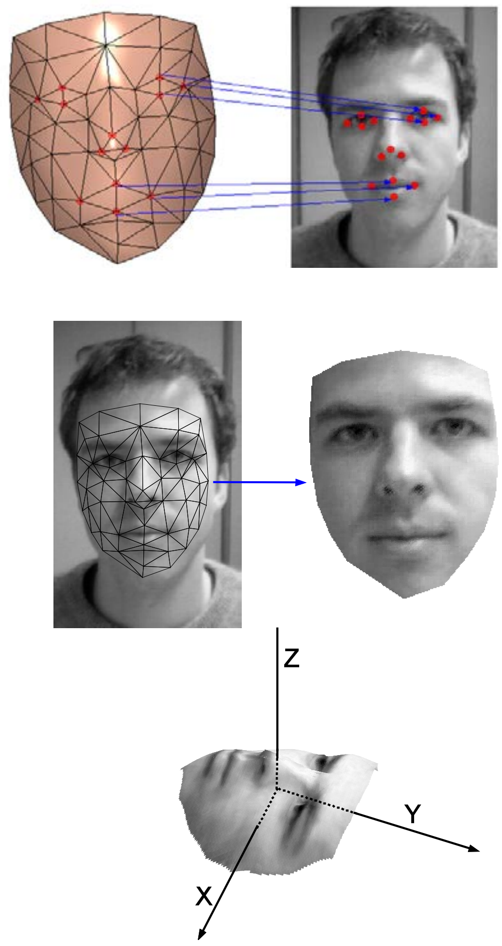
Mixed mode

- Switch between bottom-up and top-down procedure.

- **Problem**: When should we switch?

Example- Top-down

Head Pose Estimation



Example- Bottom-Up

Road Extraction

Module 1

Initial road segments:

- hypotheses for road axes:

  bright lines in low resolution
- hypotheses for road sides:

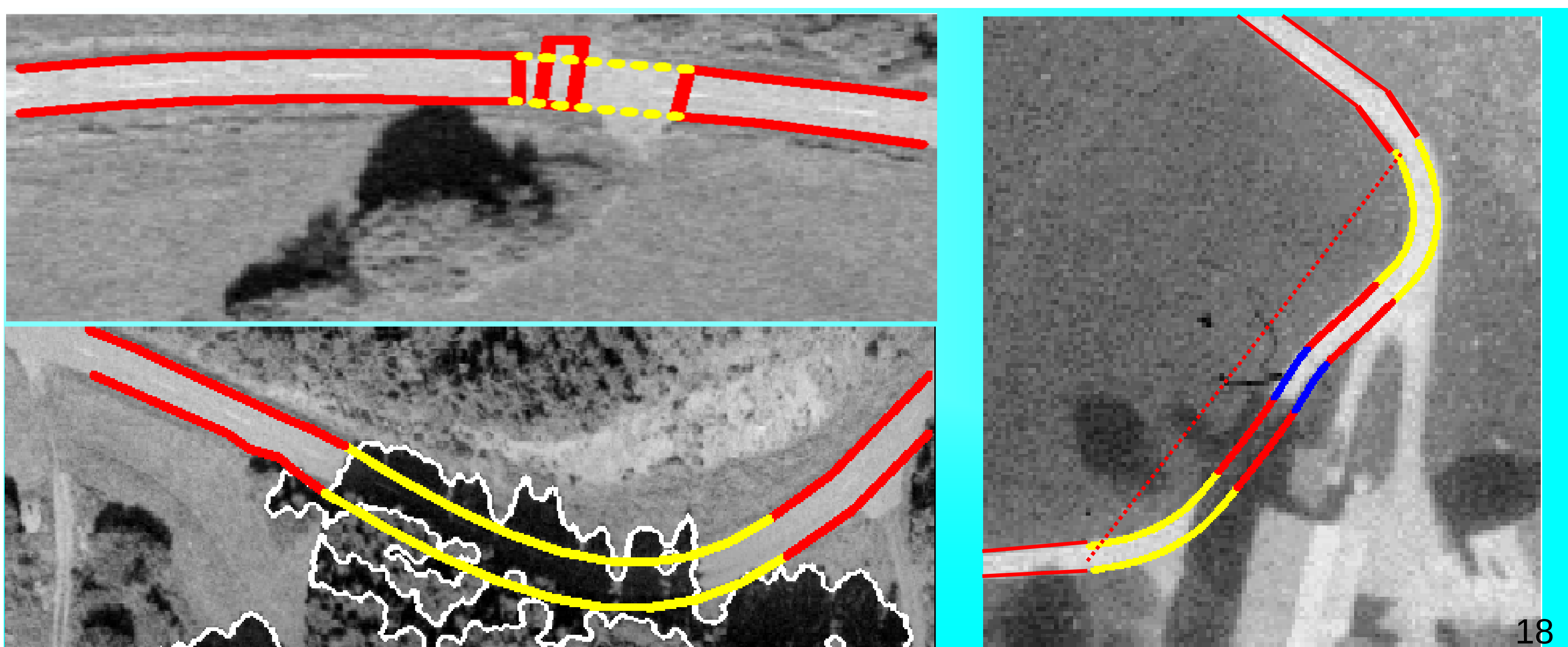
  parallel edges enclosing homogeneous region in high

Extraction

Module 1

Connections between road segments:

1) perceptually evident connections (short, similar gray value statistics)
2) geometrically possible connections (verification by zip-lock ribbon snakes)
3) connections that can be explained by road context

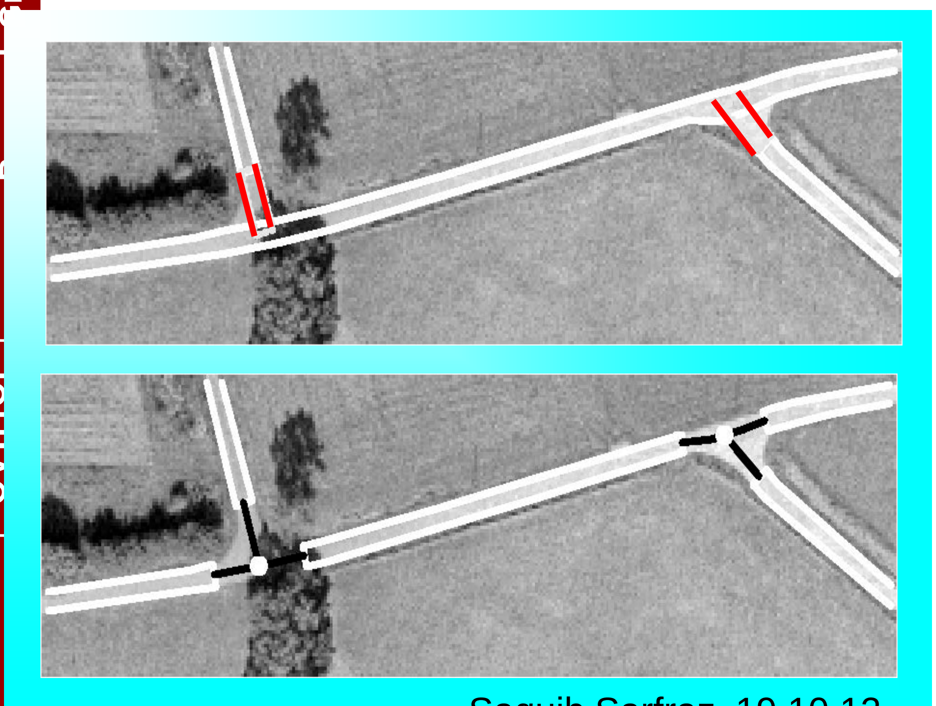


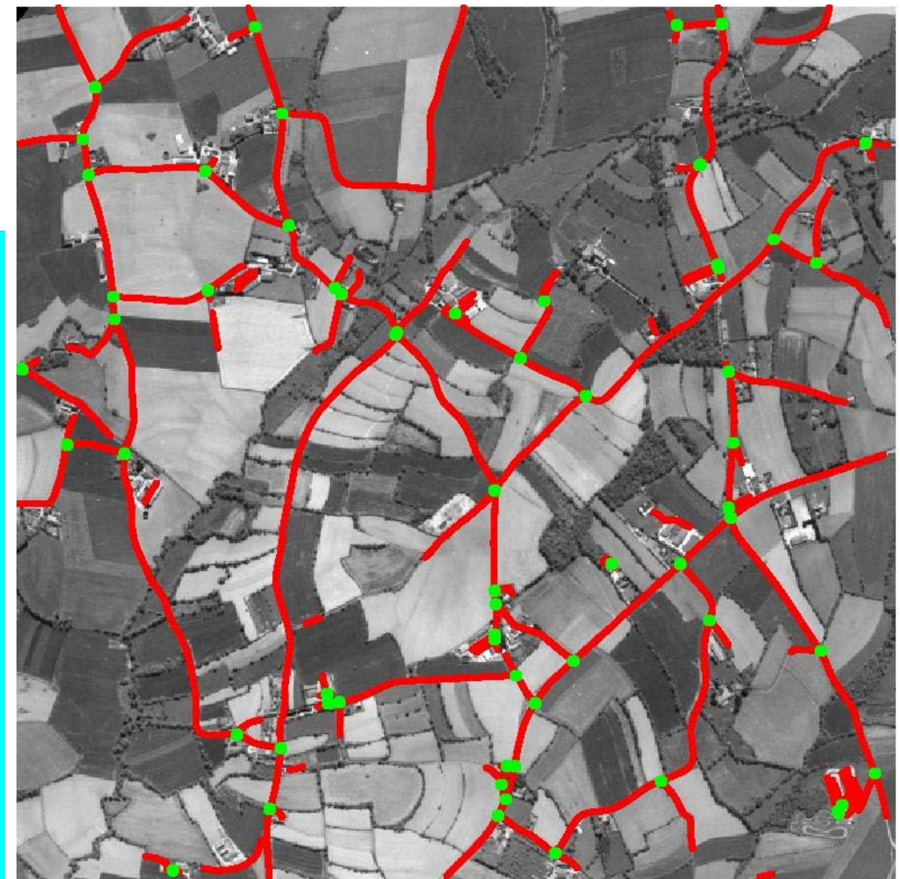
Extraction

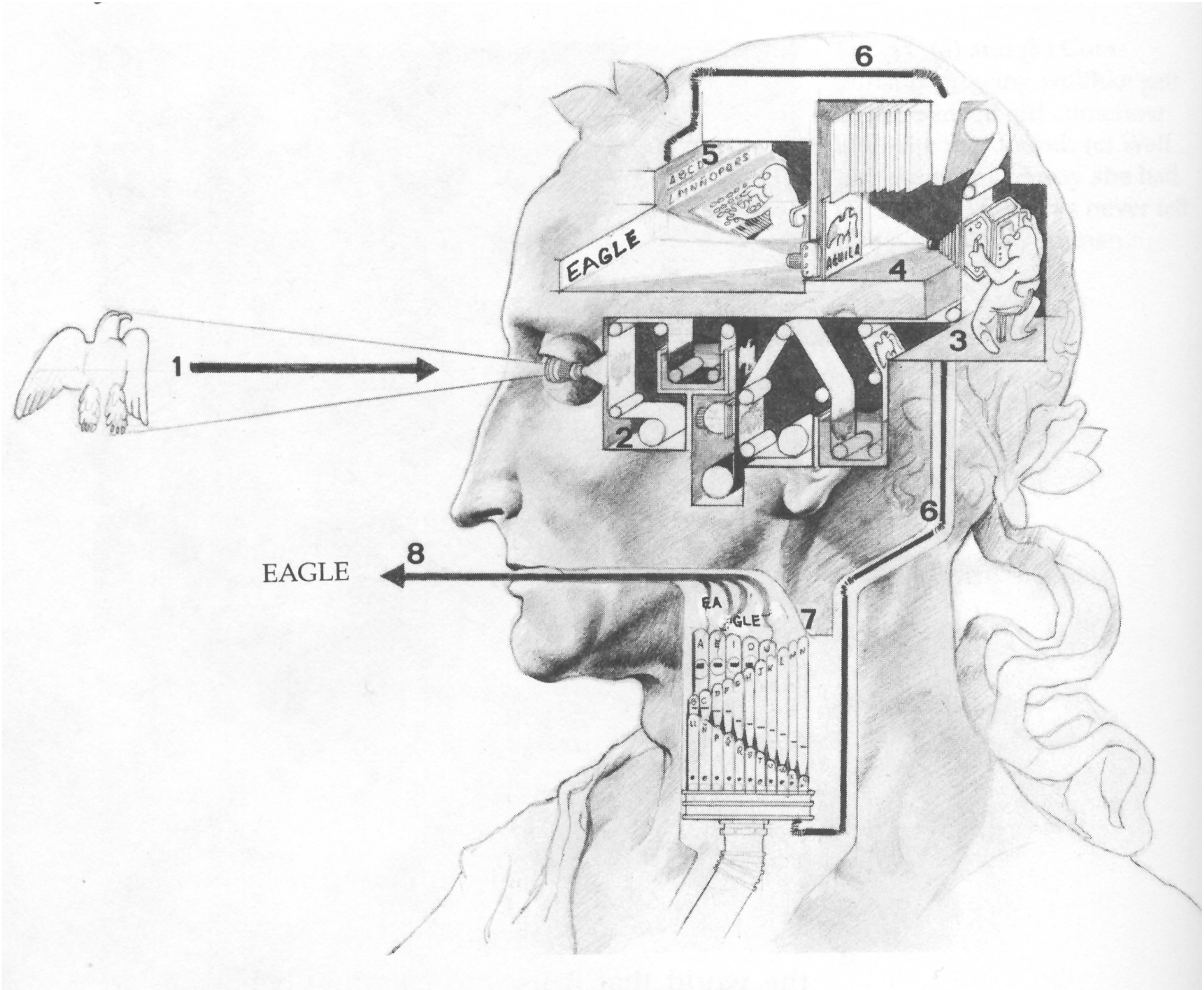
Module 1

Junctions:

1) extend ends of road segments
2) check gray value statistics
3) check context






6
5
EAGLE
AQUILA
4
3
1
2
6
8
EAGLE
7

Math is the Doctrine of the Patterns

W. W. Sawyers (1955). Prelude to Mathematics. London: Penguin Books:

...we may say, “Mathematics is the classification and study of all possible patterns.” Pattern is here used in a way that not everybody may agree with. It is to be understood in a very wide sense, to cover almost *any kind of regularity that can be recognized by the mind*. Life, and certainly intellectual life, is only possible because there are certain regularities in the world. A bird recognizes the black and yellow bands of a wasp; man recognizes that the growth of a plant follows the sowing of seed. In each case, a mind is aware of pattern. (p. 12; highlighted in the original)

Mathematics:
The Science of Patterns

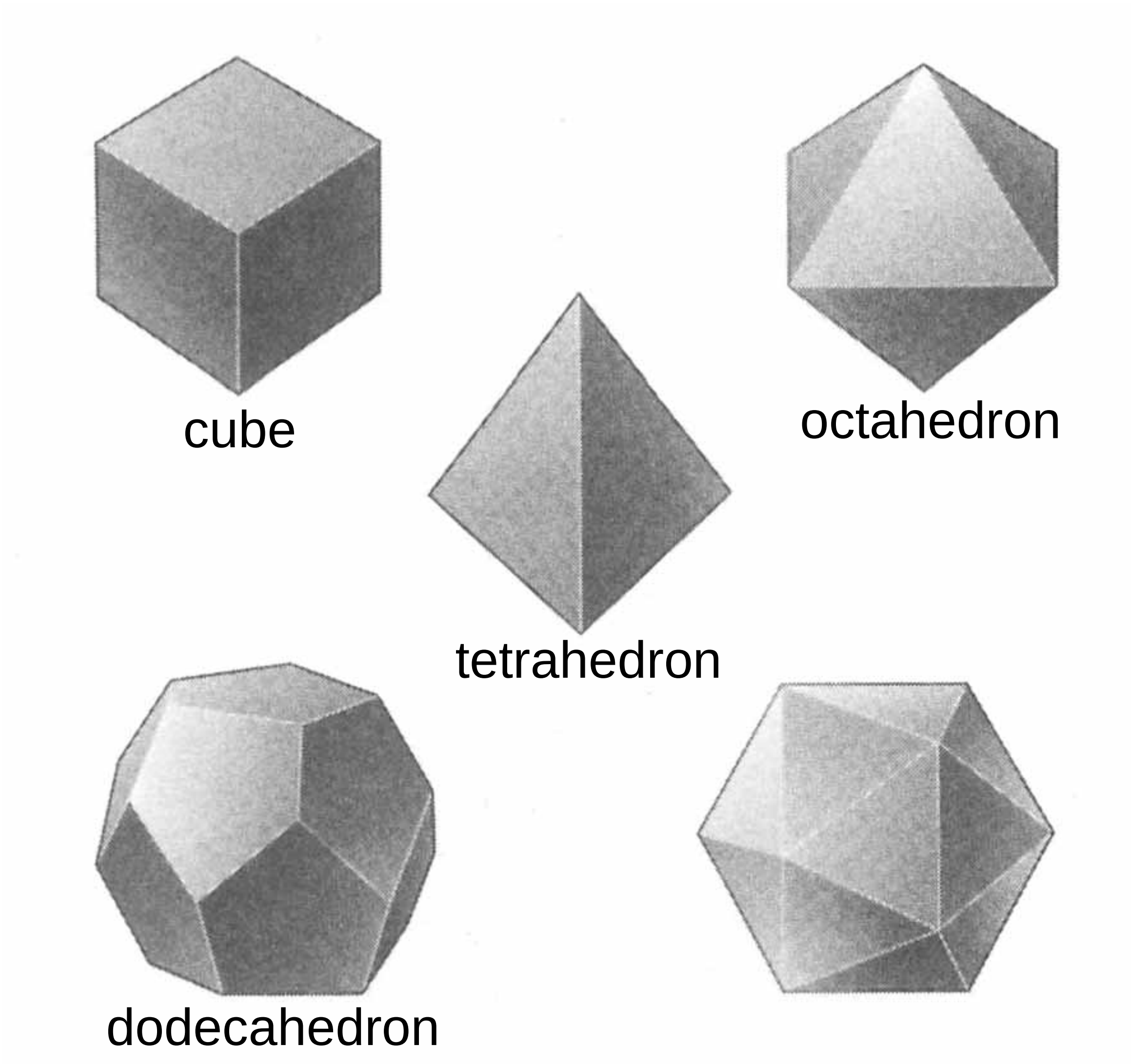
- Andrew Gleason in the 1984 October issue of the Bulletin of the American Academy of Arts and Sciences:

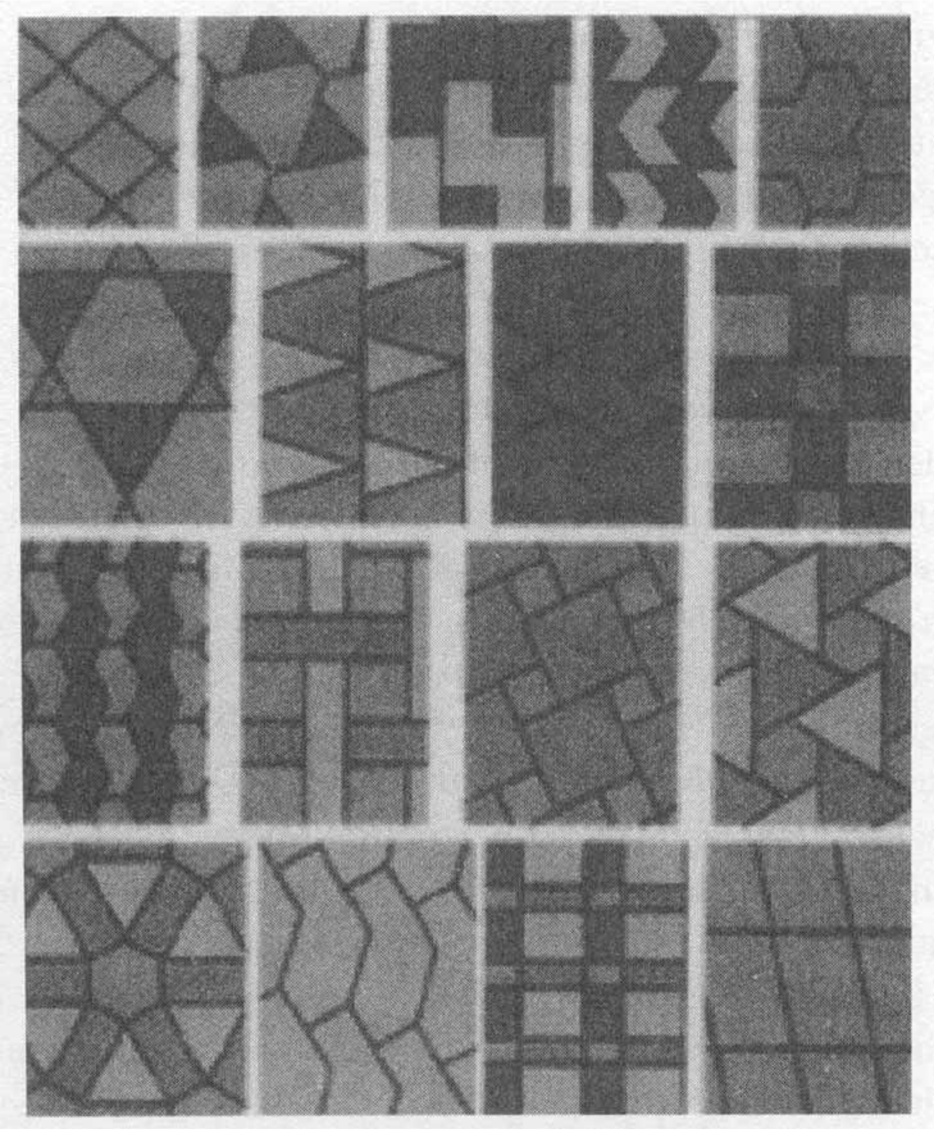
  Mathematics is a science of order in terms of regularities and patterns. Aim of mathematics is the identification and description of such orders, kinds of orders and coherences, which exist between those various kinds of order.

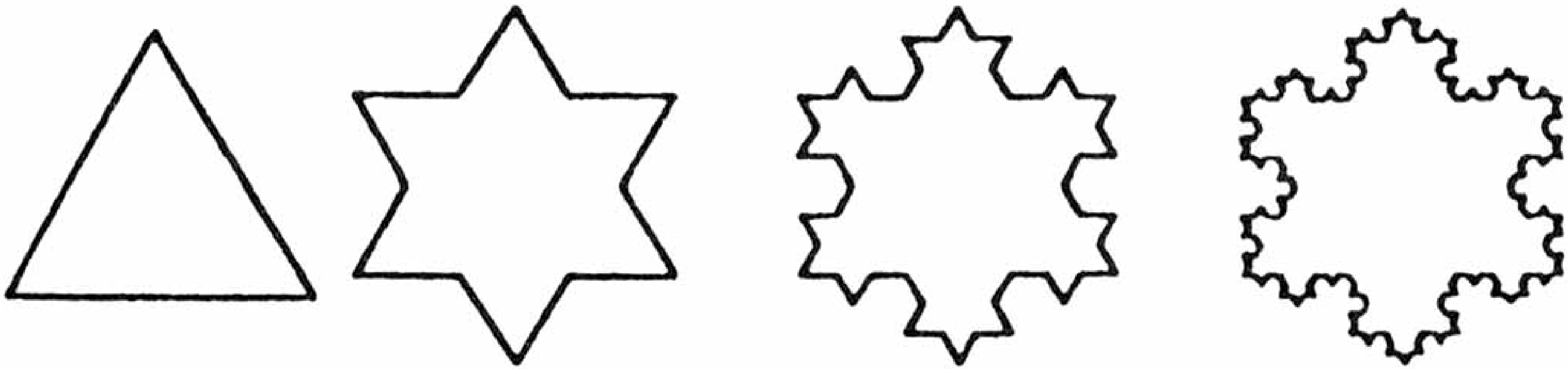
- In: Keith Devlin: The Math Gene. J.G. Cotta, Stuttgart 2001.

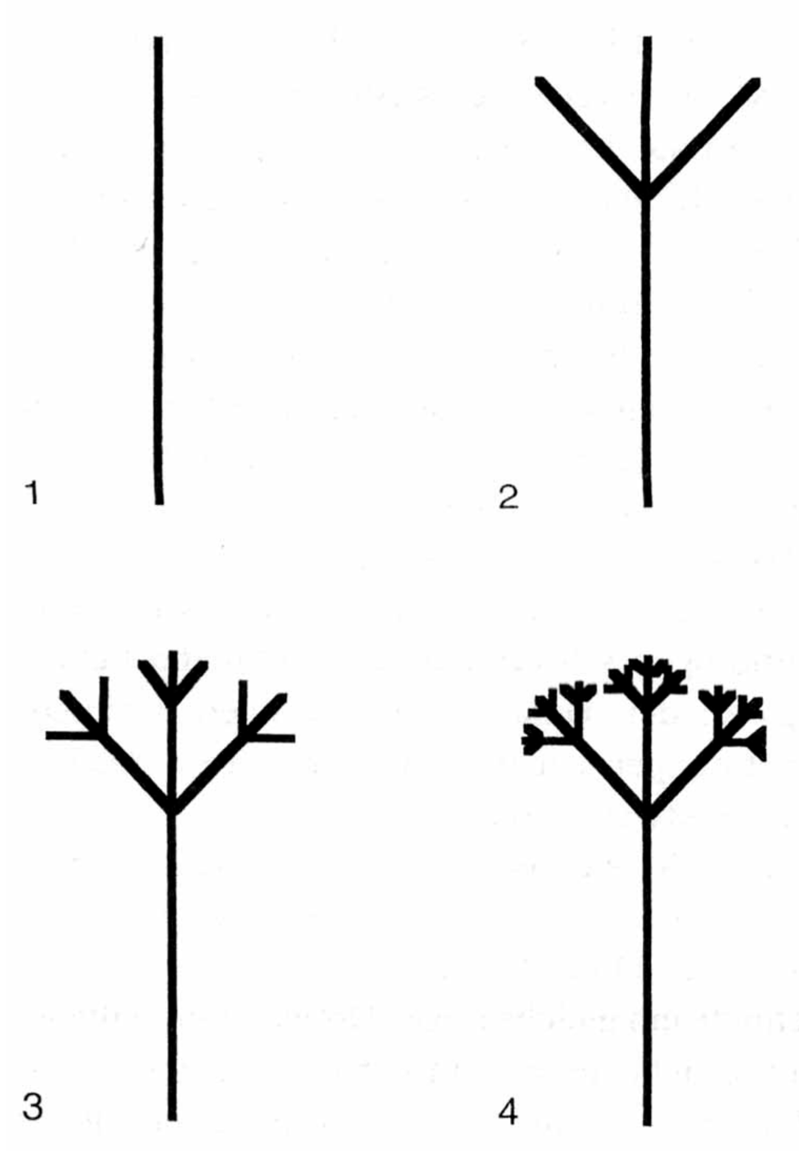
Computer Vision for Human-Computer Interaction
Research Group, Universität Karlsruhe (TH)

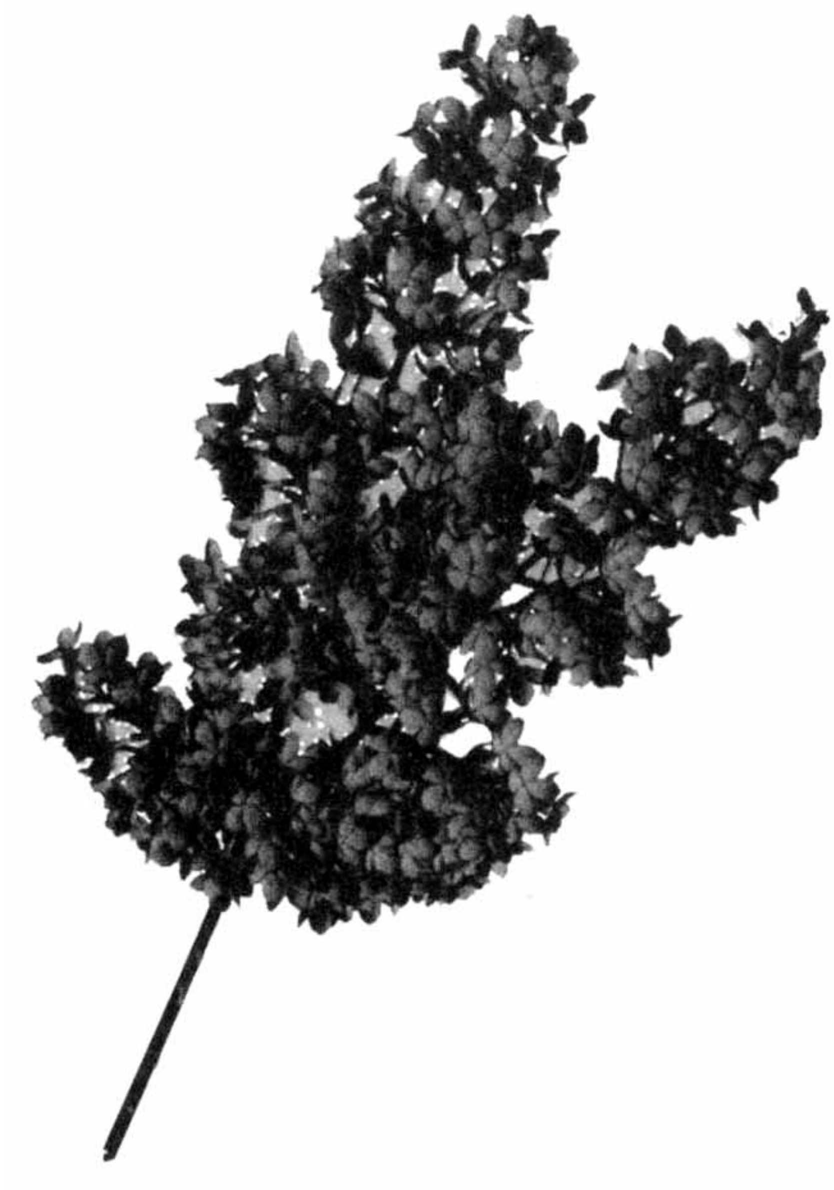
cv:hci



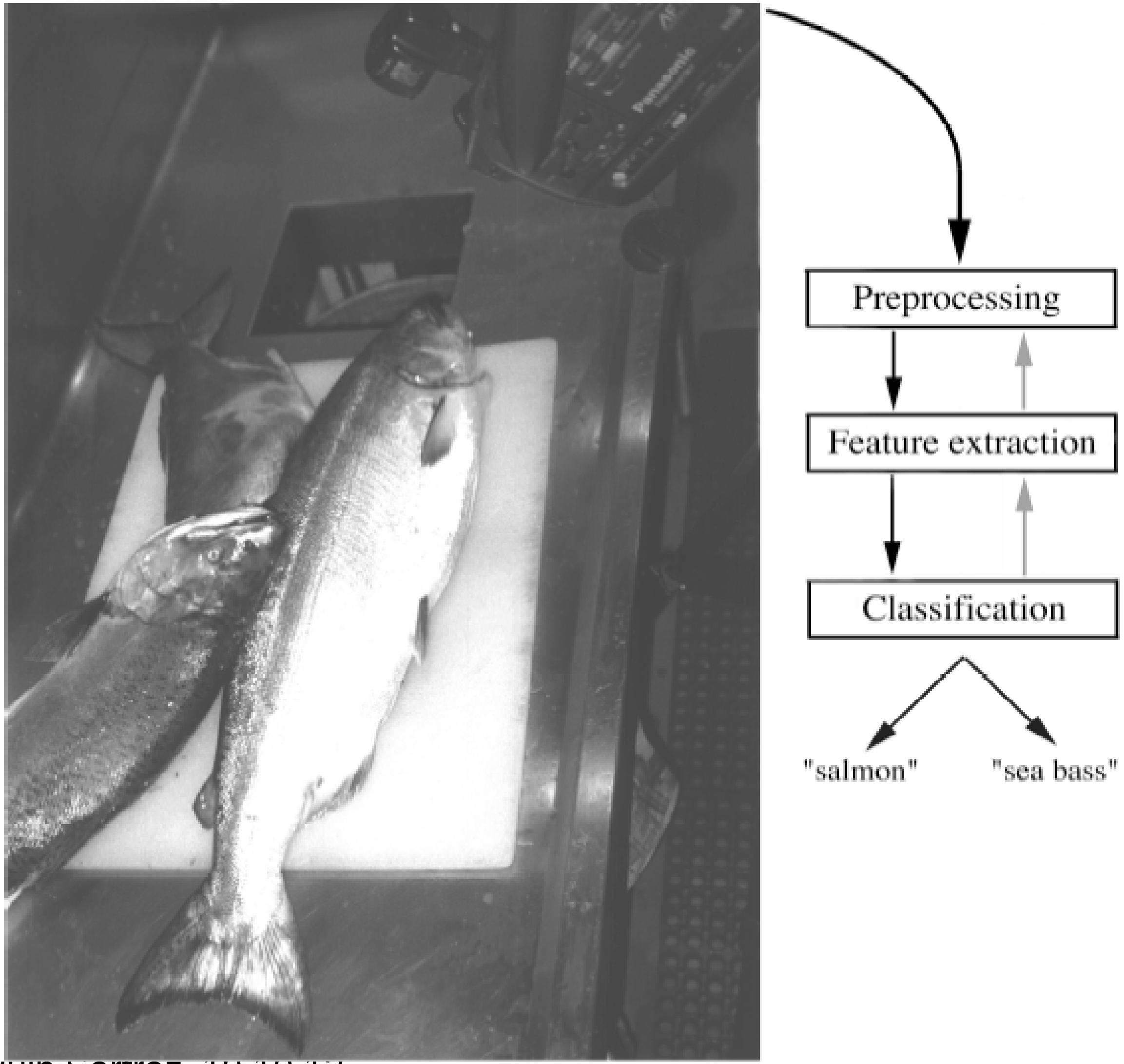


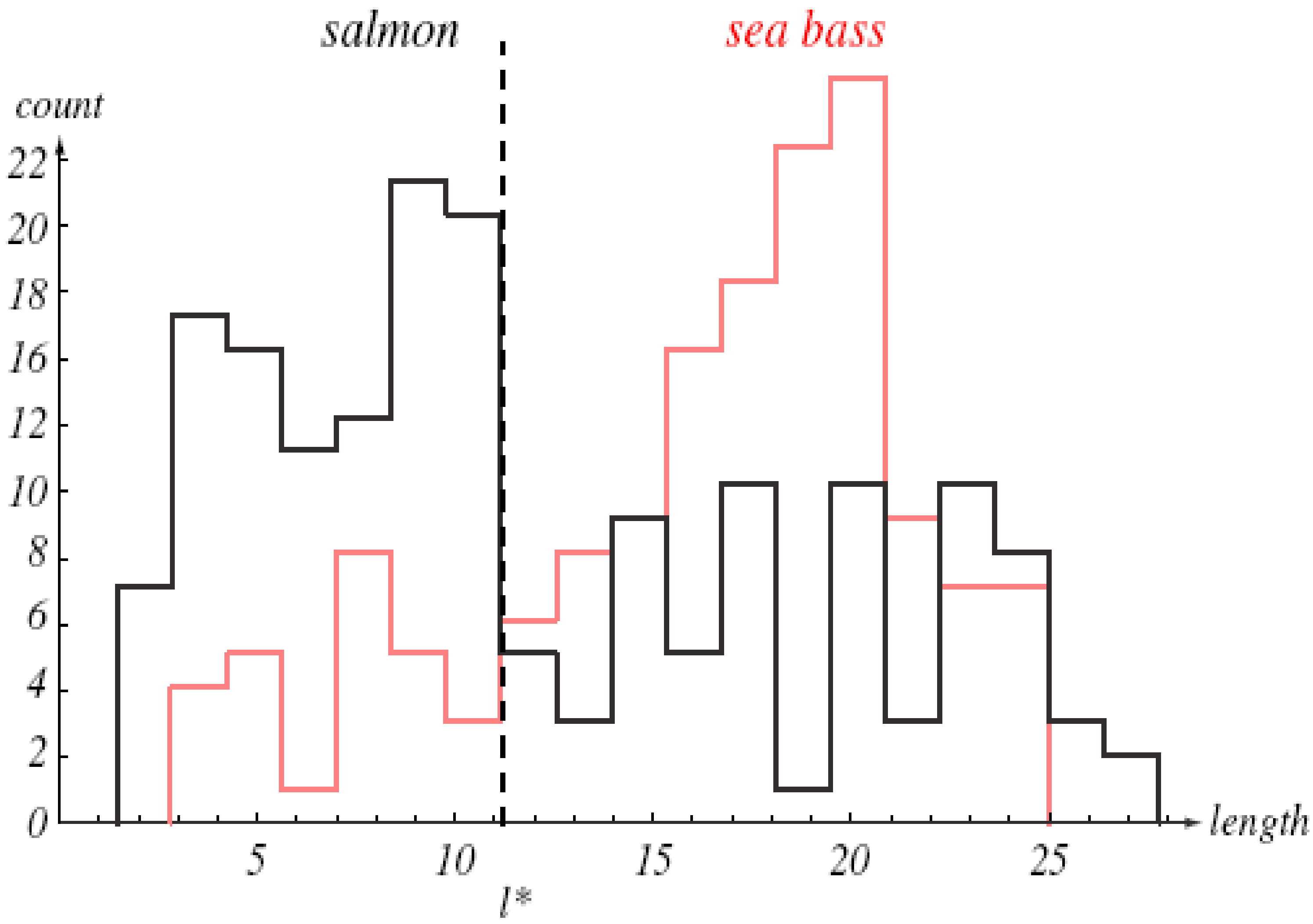


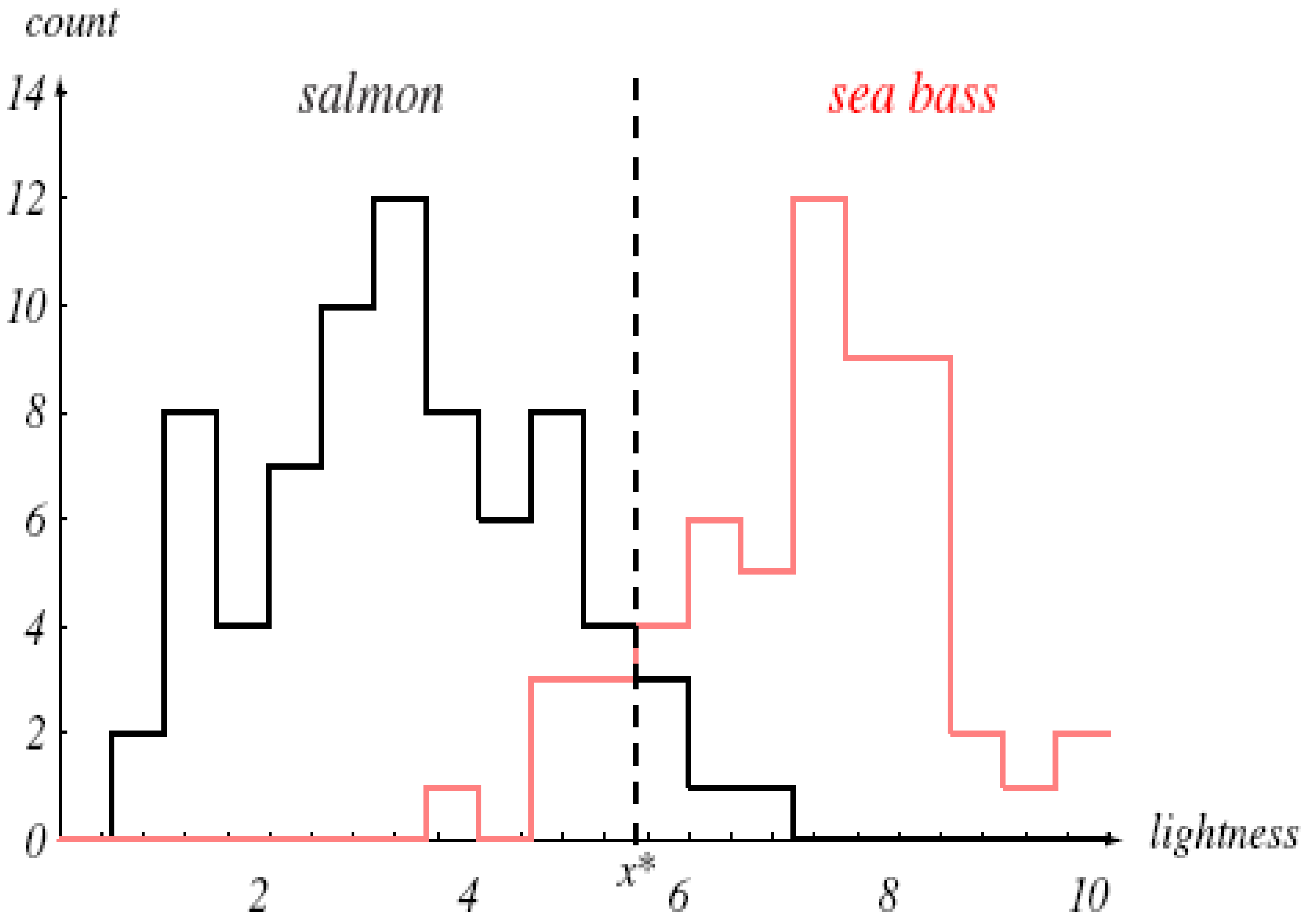

1
2
3
4

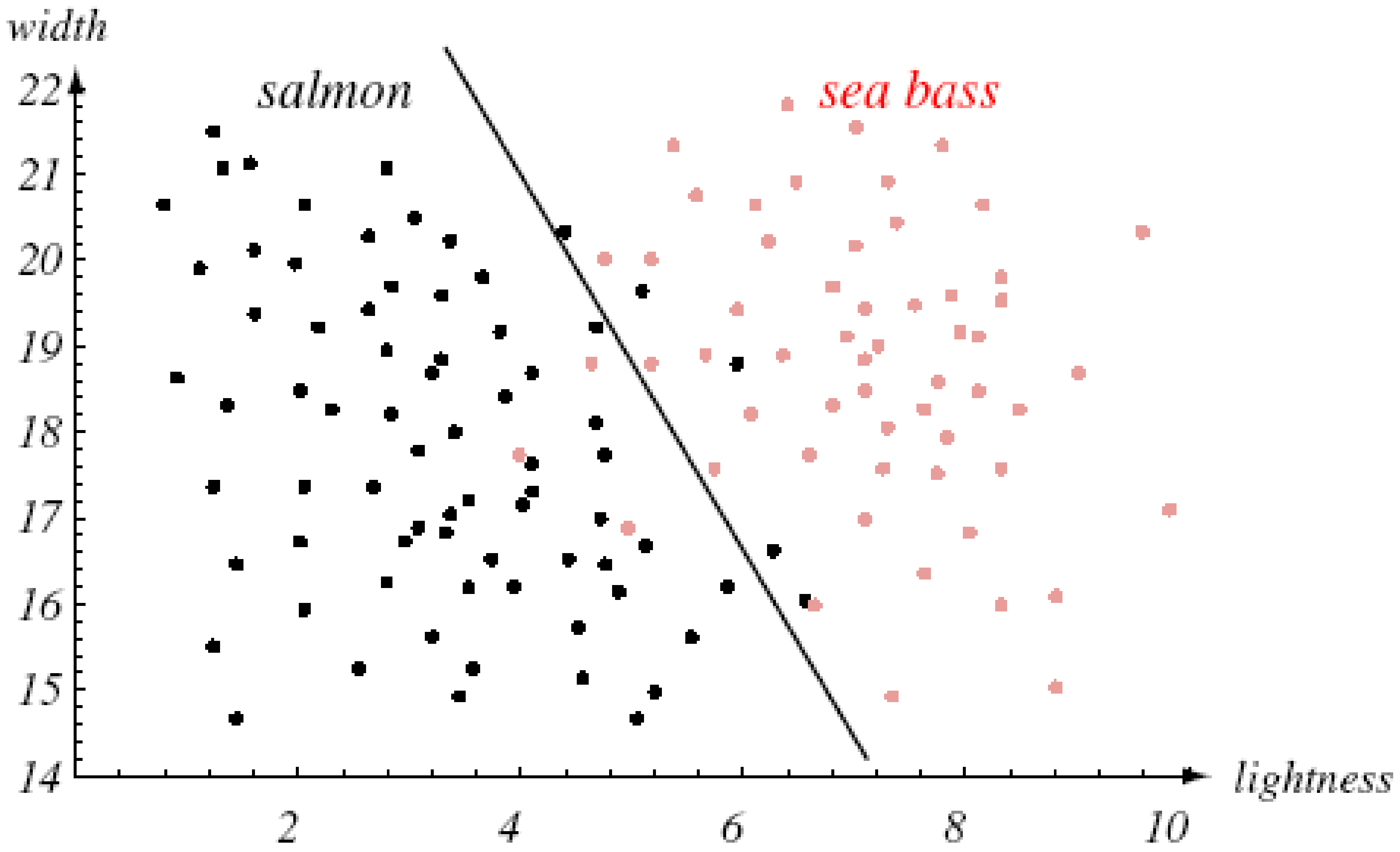


Pattern Recognition




salmon
sea bass
count
22
20
18
16
12
10
8
6
4
2
0
length
5
10
15
20
25
l*


count
14
12
10
8
6
4
2
0
salmon
sea bass
x^*
2
4
6
8
10
lightness


width
salmon
sea bass
22
21
20
19
18
17
16
15
14
2
4
6
8
10
lightness

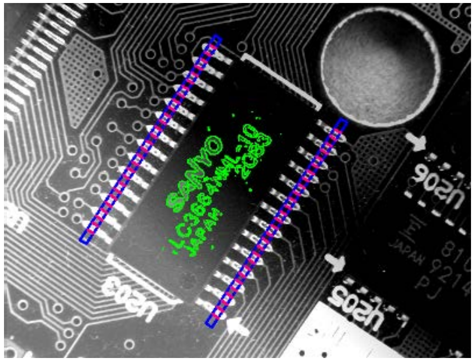
Real-Time Object Recognition: Definition and Motivation

Definition of object recognition

- Determination of the location of a 2D-object in the image
- Location of the object is described by:
  - Object position
  - Object orientation

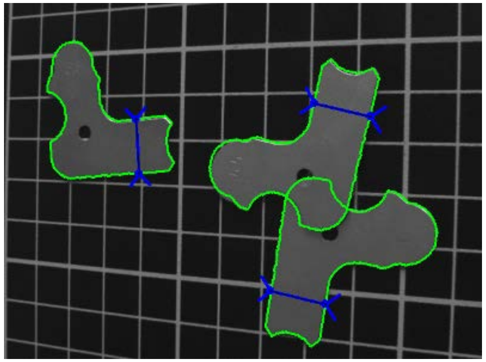
Prerequisite for many applications in industries, e.g.:

semiconductor inspection



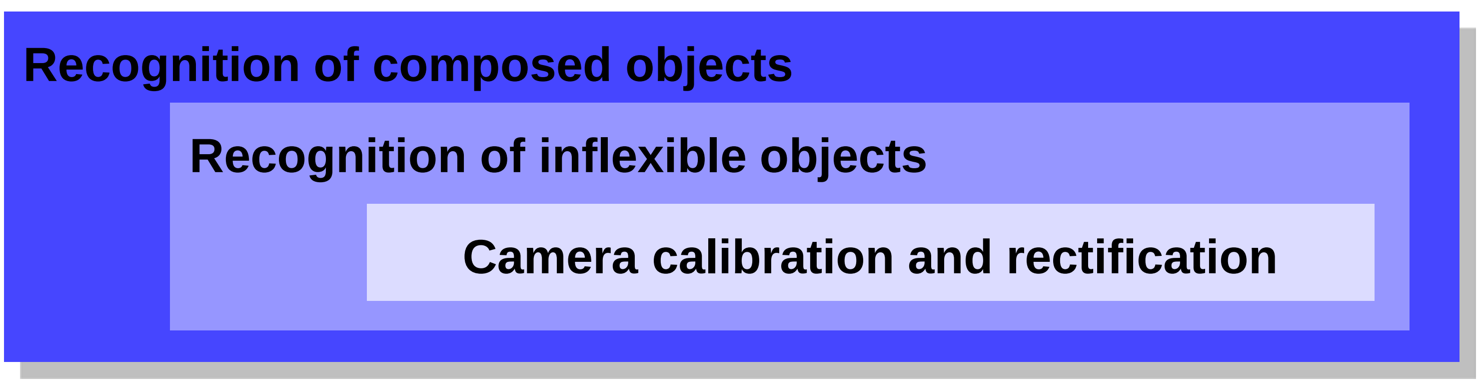
print image quality control

„Pick-and-Place"



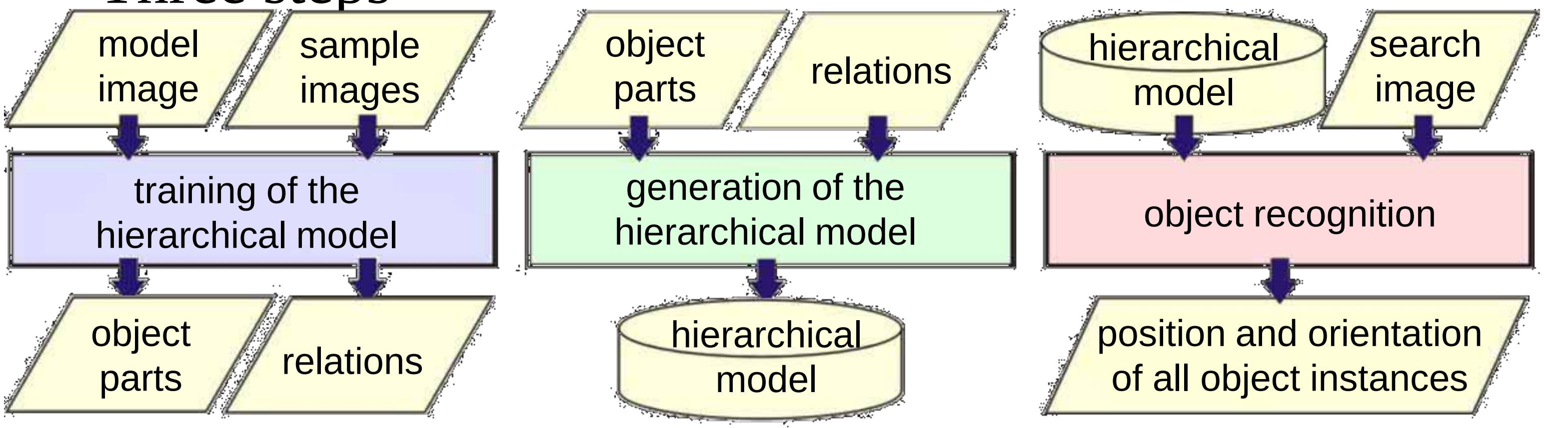
Object Recognition Procedure

- Procedure is composed of three components
  1. Camera calibration and rectification
     ⇒ Robust with respect to perspective distortion
  2. Recognition of inflexible objects
     ⇒ Robust with respect to occlusion, disturbances, brightness fluctuation and noise
     ⇒ Accurate
     ⇒ Real-time compliant
  3. Recognition of composed objects



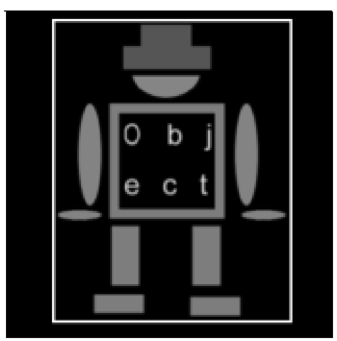
Recognition of Composed Objects

- Three steps

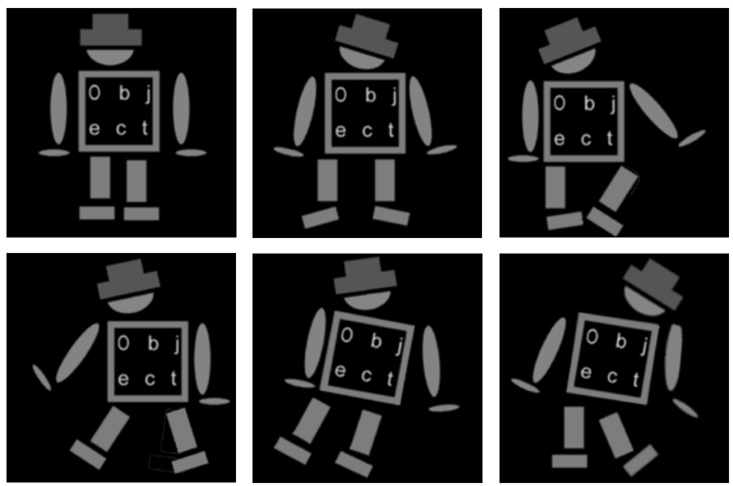


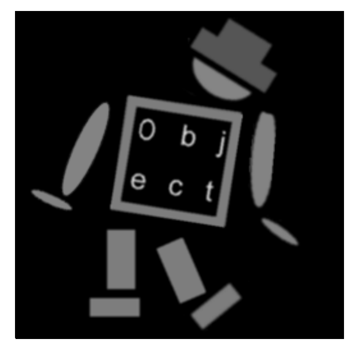
- Example

Image of model



Images of examples

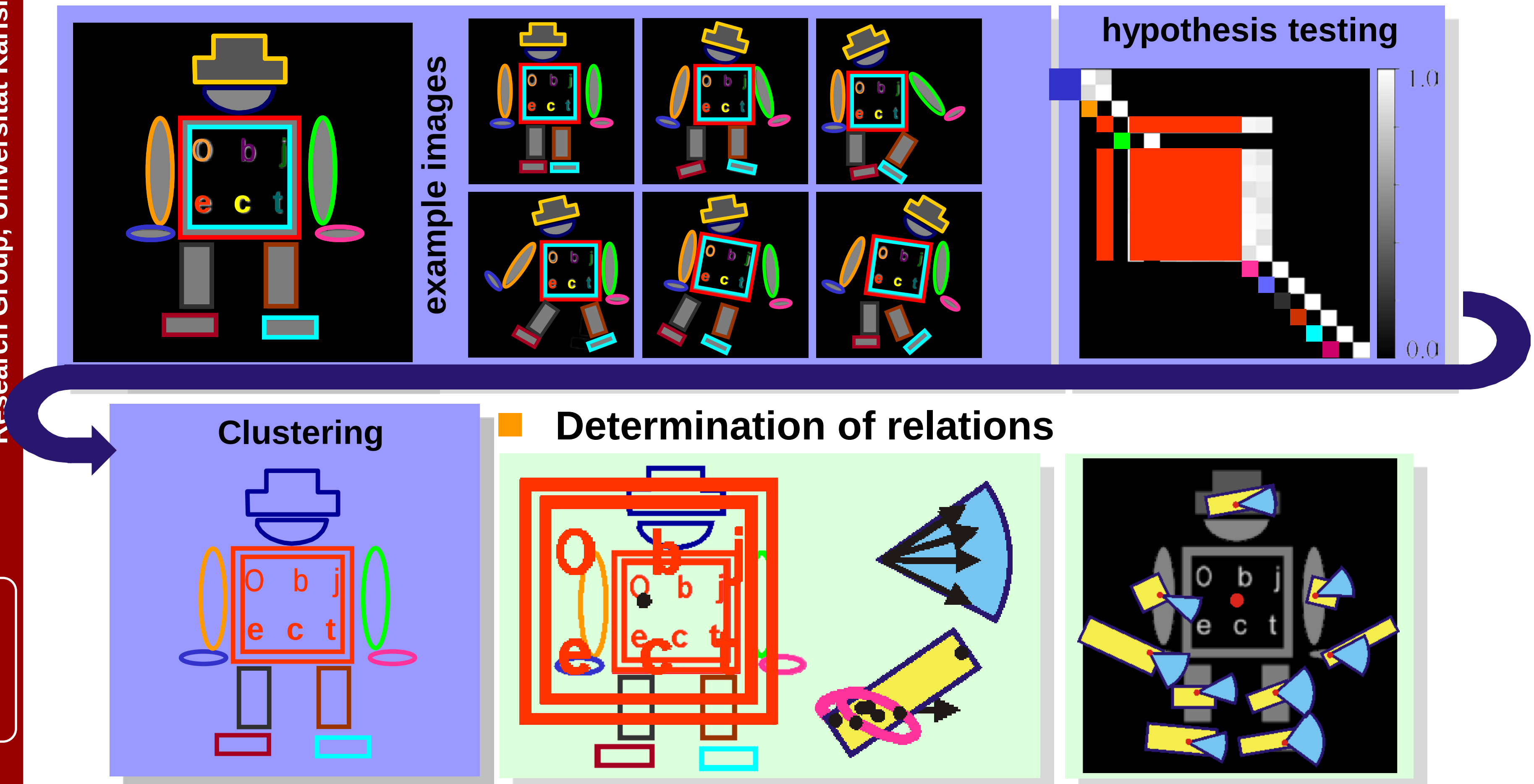




Computer Vision for Human-Computer Interaction
Research Group, Universität Karlsruhe (TH)
cv:hci

Step 1: Training of a Hierarchical Model

- Determination of un-deformable object parts



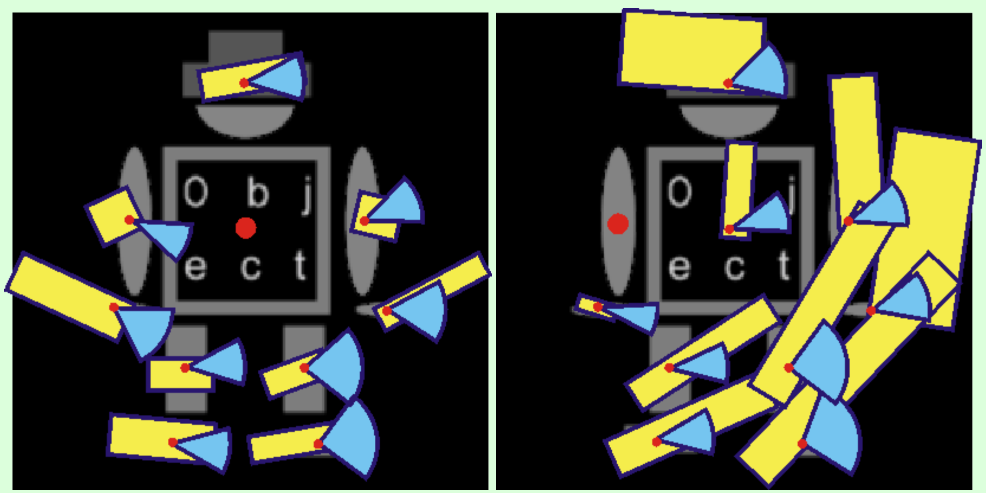
- **Determination of relations**

Step 2: Generation of the Hierarchical Model

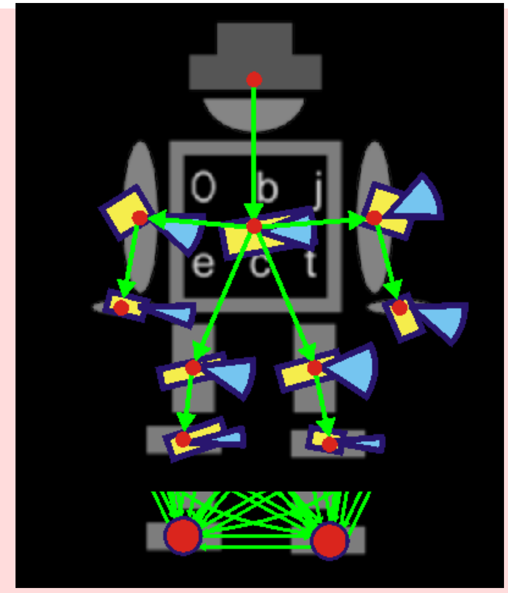
- Goal: Minimization of computation time during object recognition
- Idea:
  - Only one object part is extracted from the image
  - Remaining object parts are extracted recursively

- **Quantization of the search efforts $\Omega_{i,j}$**

$\Omega_{i,j}$ = f(relations,…)



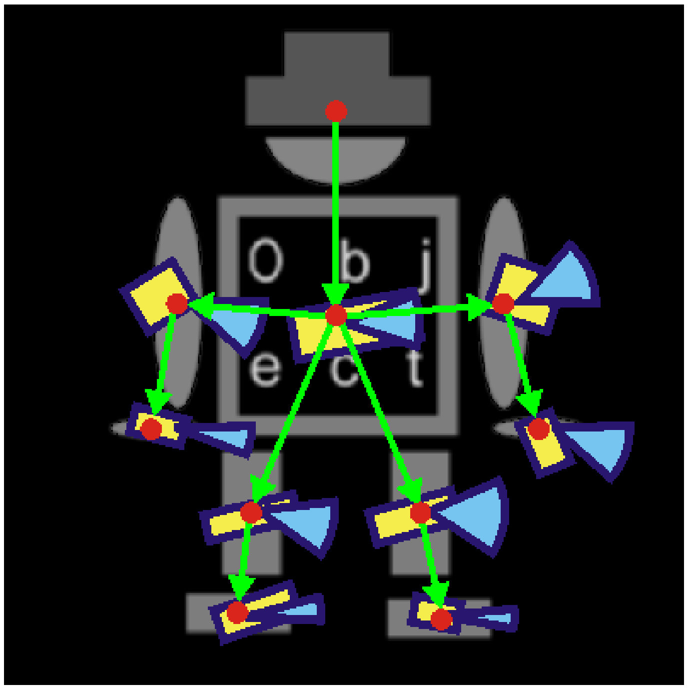
- **Optimization of search**
  $\Sigma\, \Omega_{i,j} \rightarrow \min$



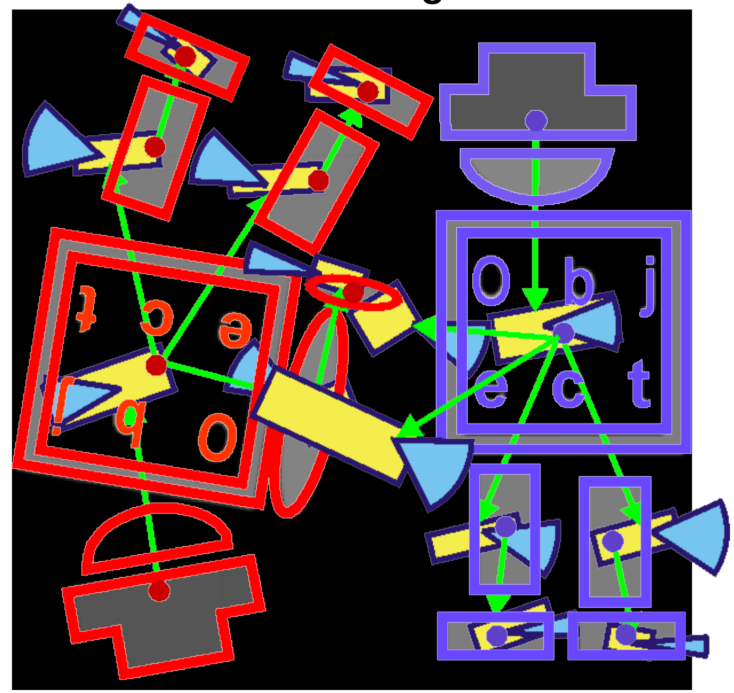
Step 3: Object Recognition

- Hierarchical search

hierarchical model



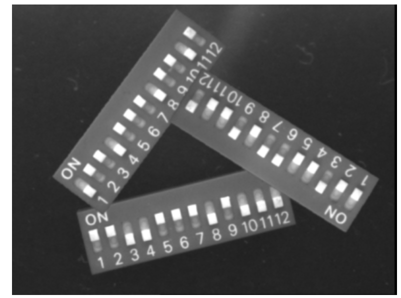
search image

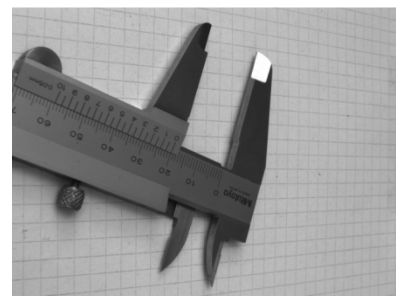


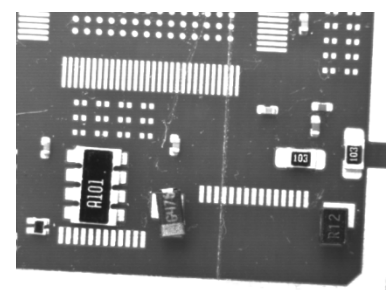
- **Advantages in comparison with independent extraction of object parts**
  - **Real-time feasibility (z.B. here: 20 ms instead of 310 ms)**
  - **Inherent correspondence analysis**

Examples

- Gas Pipe Pliers
- DIP-Switch
- Minimum Durability
- Sliding Caliper
- Comparison:

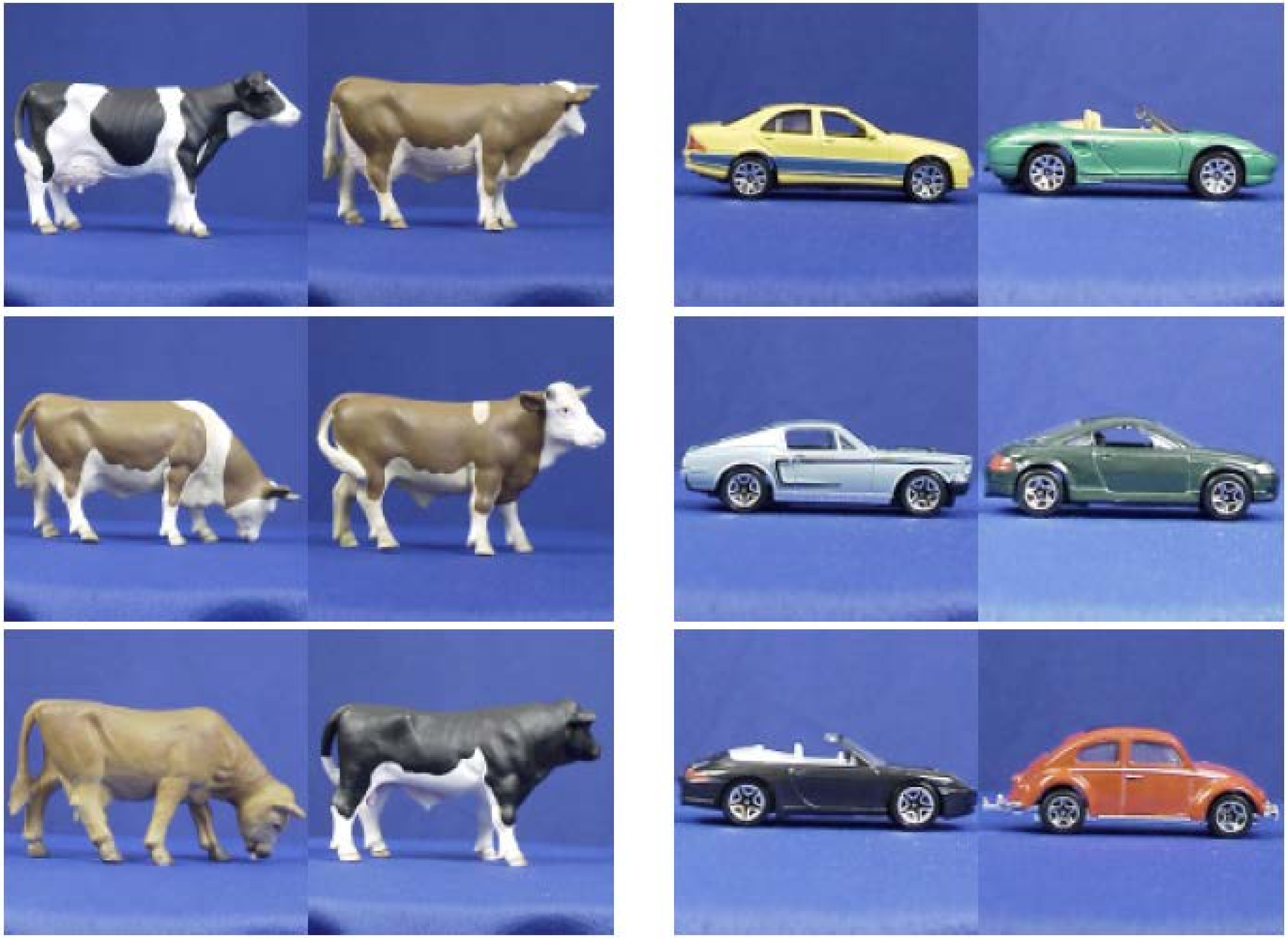






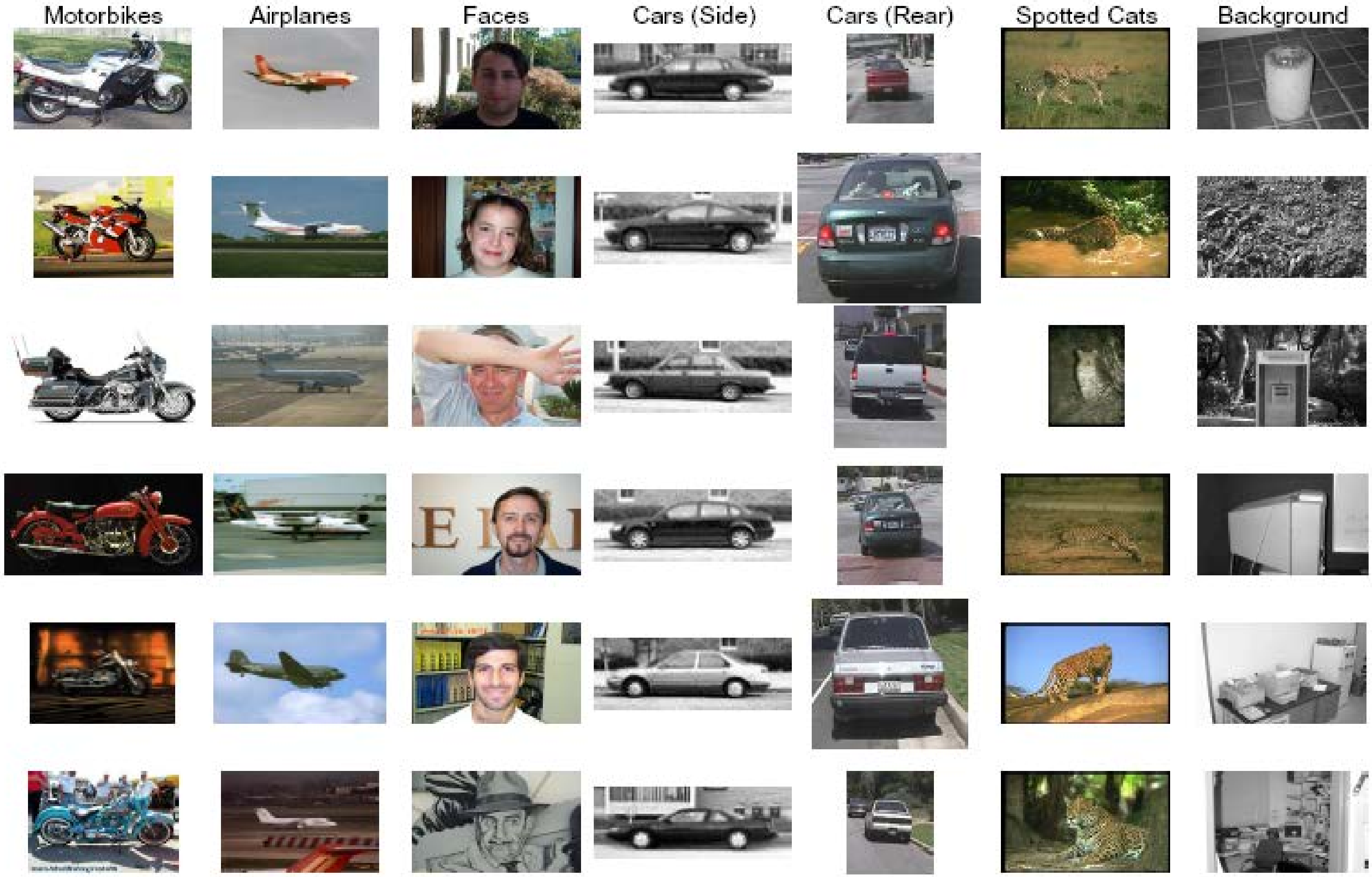
Recognition of fixed objects ↔ **Recognition of composed objects**

Object Categorization:Two Object Categories



Objects and Clutter/Context

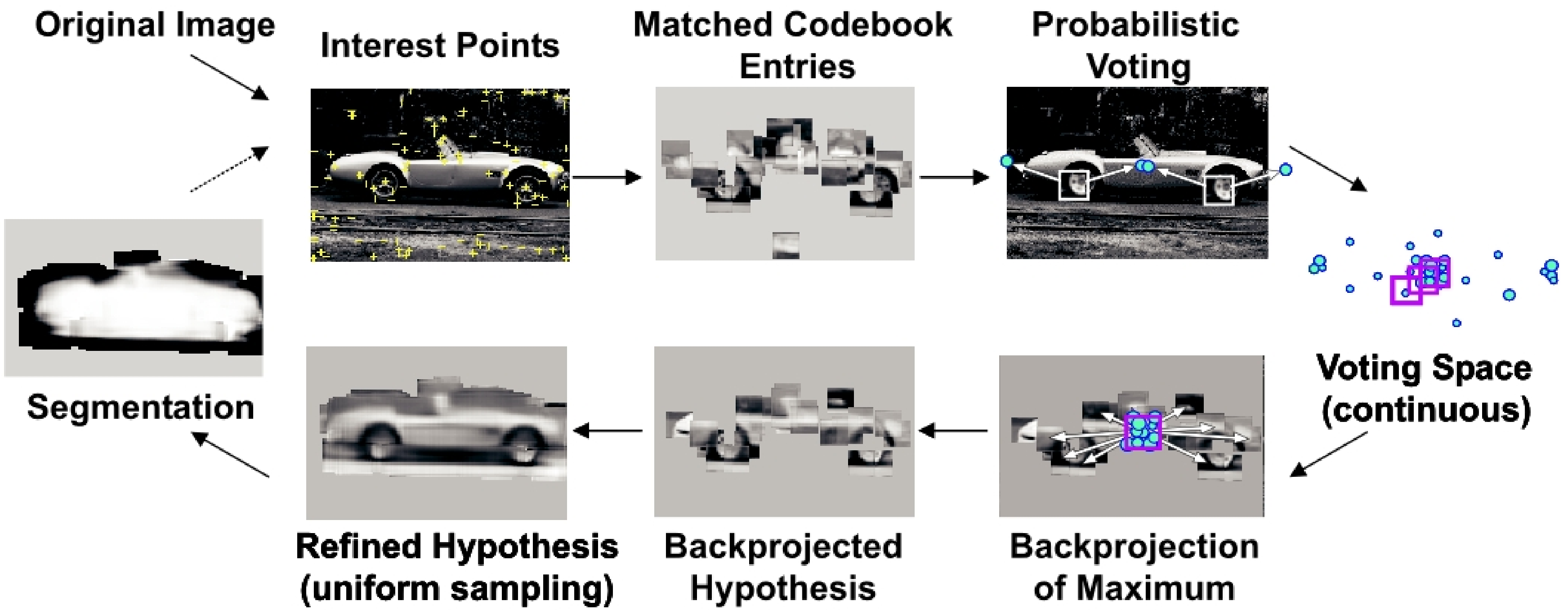
Different Object Categories and "In-Category" Variations



Problems/Challenges/Topics

- Object parts
- Geometry / mutual position of parts
- Automatic identification of (new) object categories
- Geometric normalization / identification of control pts.
- Invariants: gray value, scale, affine/projective transf.
- Unsupervised Training under clutter
- Size of training sets
- Variability of appearance: category, parts
- Variability of shape: deformations, also: position of parts
- Handling of occlusions
- Extraction of salient features: points, edges, regions
- “Transformation of image to numbers”: feature vector
- Grouping of features
- Classification: Bayesian decision theory
- Model and estimation of model parameters
- Which features to use for which object categories?
- Learning: improvement of learned model parameters
- Automatic procedure: generic approaches

An Approach to Object Categorization (Leibe, 2004)



It’s a miracle that curiosity survives formal education

(Albert Einstein)