Robot Vision

TTK4255

Lecture 08 – Shape Analysis

Annette Stahl

(Annette.Stahl@ntnu.no)

Department of Engineering Cybernetics – ITK

NTNU, Trondheim

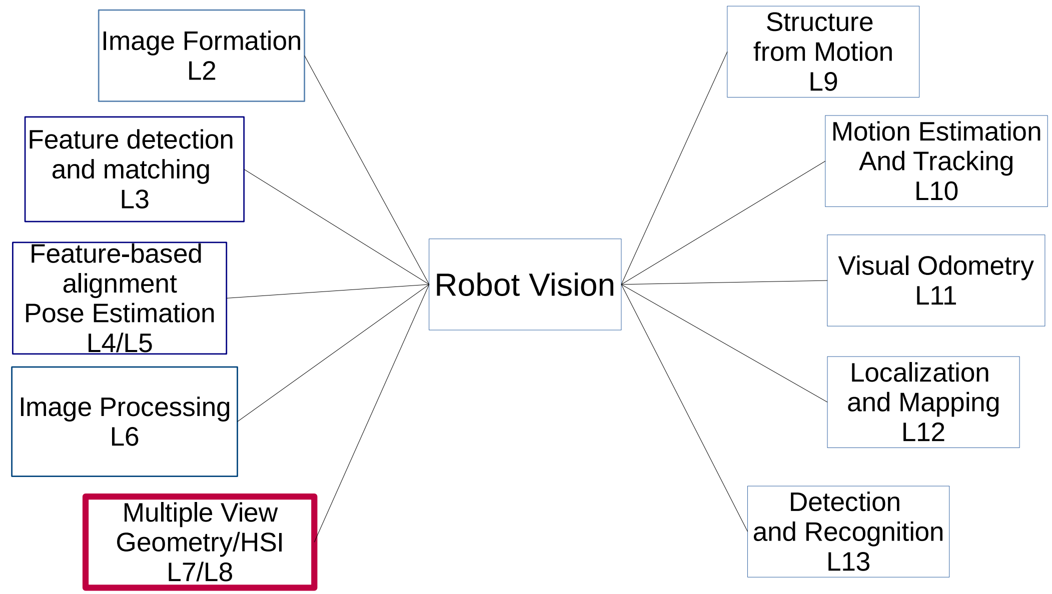
Spring Semester

24. February 2020

Lecture 08 – Shape Analysis

Annette Stahl (Annette.Stahl@ntnu.no)

Simen Haugo (Simen.Haugo@ntnu.no)



Outline of the fifth lecture:

- Shapes
- Some notes about shapes
- Prerequisites (Assumptions we make for the presented methods.)
- Shape Description
- Two classical Shape Representations/Descriptions
- Fourier Descriptors
- Curvature Scale Space

Shapes

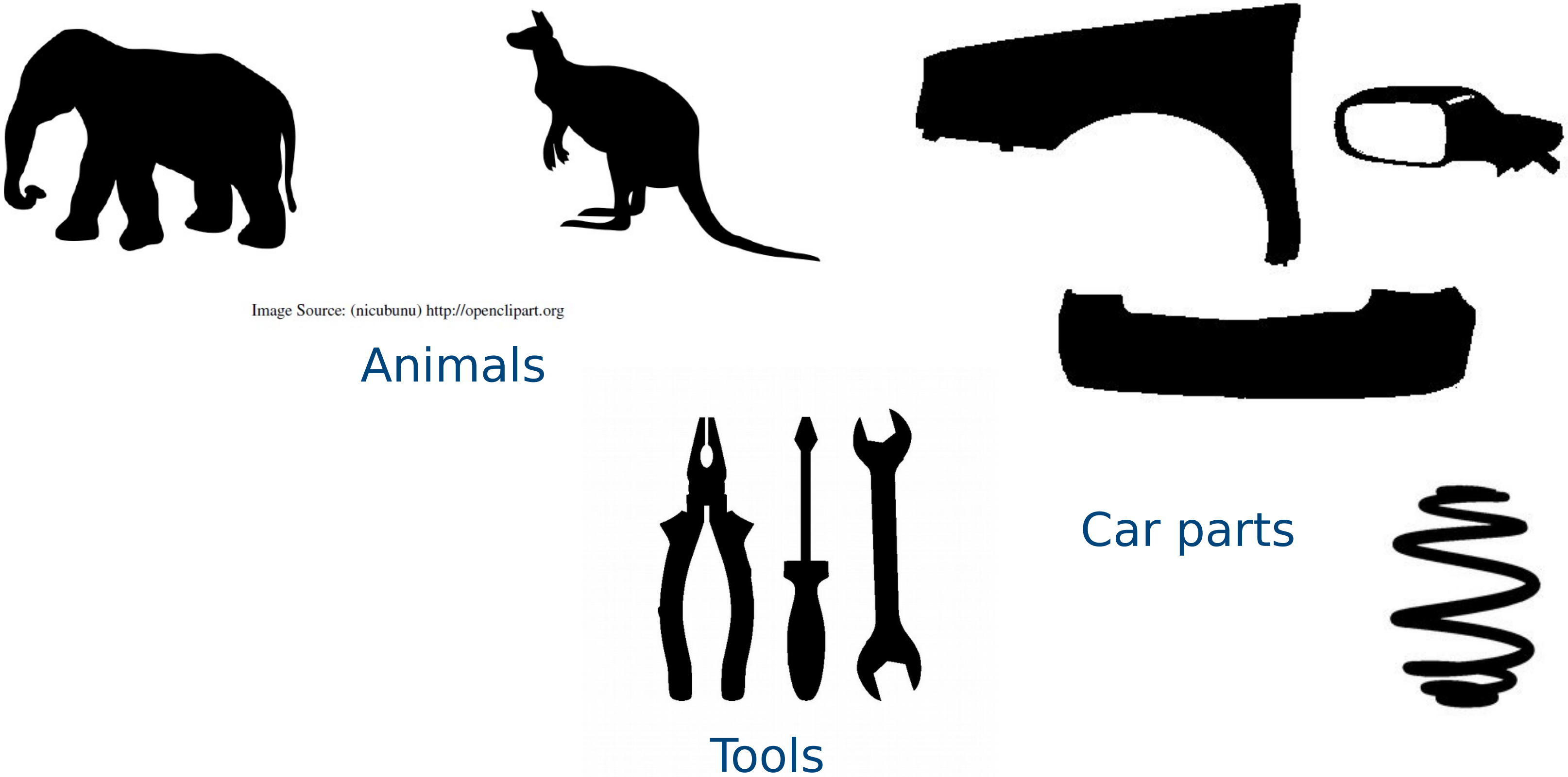


Image Source: (nicubunu) http://openclipart.org

Humans perform very well in recognizing objects from shape!

Shape descriptors: attempts to quantify shape in ways that agree with human intuition (or task-specific requirements)

Applications for Shape-Recognition

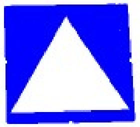
Industrial Manufacturing

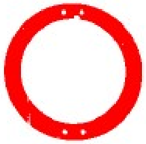
Many tasks in the industry are still performed manually
Non rigid objects

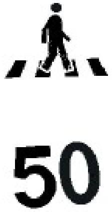
Autonomous Vehicles

Recognition of street-signs

prototype Google car







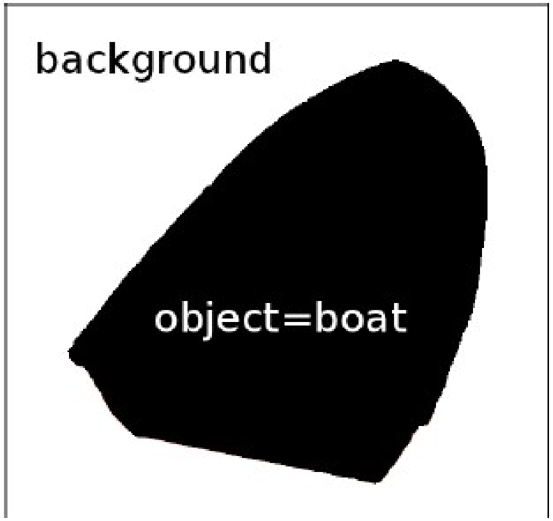
How can we obtain shapes?

Answer: By segmentation approaches. (Note: This is a large field on its own!)

A subdivision of an image into multiple image regions.
(The partitions should correspond to objects within the image.)

Image

Segmentation we would like to get



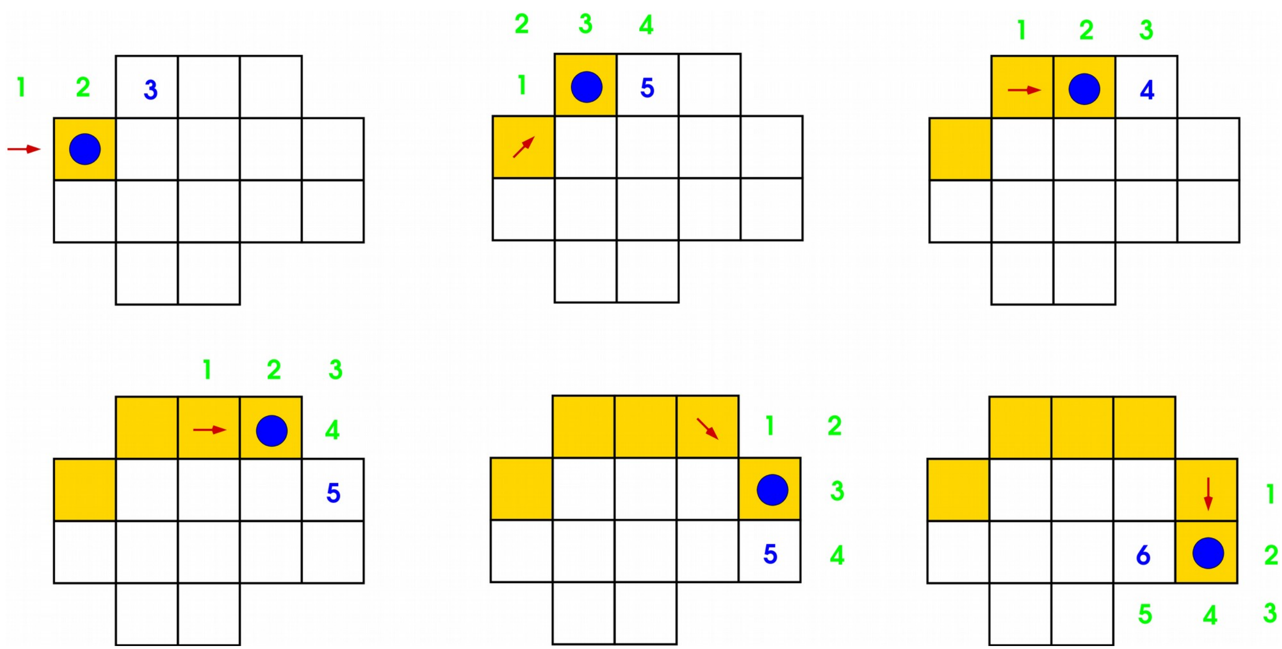
Color-based segmentation

Many challenges of its own. In simple cases we can just use a thresholding.

Prerequisites for Shape Description

We assume that:

- the image has already been segmented → (most simple approach: thresholding)
- the extracted shape is a closed curve.
- the contour tracing around the object is already performed.
  → i.e. Moore contour tracing, sequence of 2D curve points



$$\begin{pmatrix} x_1, & y_1 \\ x_2, & y_2 \\ \vdots & \vdots \\ x_N, & y_N \end{pmatrix}$$

Shape Description

Aim:

Creating numerical representations of the shape that can be further processed by a robot vision system. For example the classification of objects.

Problems:

- **no** generally **“best“ method** can be named **for** a **shape description**
- Unclear what the important parts of a shape are.

→ Useful information available in the variations of the boundaries (i.e. first and second derivatives, location of points with high curvature or curvature change).

Descriptor-Properties of Interest

Properties of shape descriptors that are of particular interest:

- **invariance to** geometric **transformations** (translation, scaling, rotation)
  → how robust is the description to these transformations

- **reconstruction ability**
  → how good can the original shape be reconstructed from the description

- **suitability** of the descriptor **in the presents of occlusions**
  → how robust is the description if only partial shape information is available

Contour Based Descriptor

- Fourier Descriptors
- Curve description by its Curvature Scale Space

Fourier Descriptors

Fourier Descriptors were first suggested by Cosgriff in 1960.

They can be used to describe the closed boundary of a shape in a 2 dimensional space using the Fourier methods.

- N (pixel-) coordinates describing the boundary shape

$$\begin{pmatrix} x_0, & y_0 \\ x_1, & y_1 \\ \vdots & \vdots \\ x_{N-1}, & y_{N-1} \end{pmatrix}$$

$$\Gamma(k) = (x_k, y_k) \quad \text{with} \quad 0 < k \leq N-1$$

- interpret the points as complex numbers

$$s(k) = x_k + iy_k$$

x-direction as real axis and y-direction as imaginary axis

- Now we have N complex numbers instead of $2N$ real numbers (2D → 1D).

Fourier Descriptors

The discrete Fourier transform of $s(k)$ results in N coefficients

$$a(m) = \sum_{k=0}^{N-1} s(k) \cdot e^{-i2\pi \frac{m}{N} k}$$

The absolute values of the Fourier coefficients are used as **Fourier descriptor** of the curve $s(k)$, $0<m<N$.

The **inverse** Fourier transform is able to **recover the original curve** $s(k)$, $0<k<N$,

$$s(k) = \frac{1}{N} \sum_{m=0}^{N-1} a(m) \cdot e^{+i2\pi \frac{m}{N} k}$$

the recovered coordinates are exactly the same as the ones that we started with!
(**Note** that s is the **complex valued** description of the curve)

Fourier Descriptors

To obtain a **real valued** curve

$$\Gamma(k) = (x(k), y(k))$$

from the complex valued signal $s(k)$ again we just have to **extract the real and imaginary parts** of s

$$x(k) = \mathrm{Re}(s(k)) \qquad y(k) = \mathrm{Im}(s(k)$$

Approximate the Shape

As coefficients $a(m)$ that belong to higher frequencies tend to have a small contribution to the signal, one can **drop terms**

$$s(k) = \frac{1}{N} \sum_{m=0}^{f_d - 1} a(m) \cdot e^{+i2\pi \frac{m}{N} k}$$

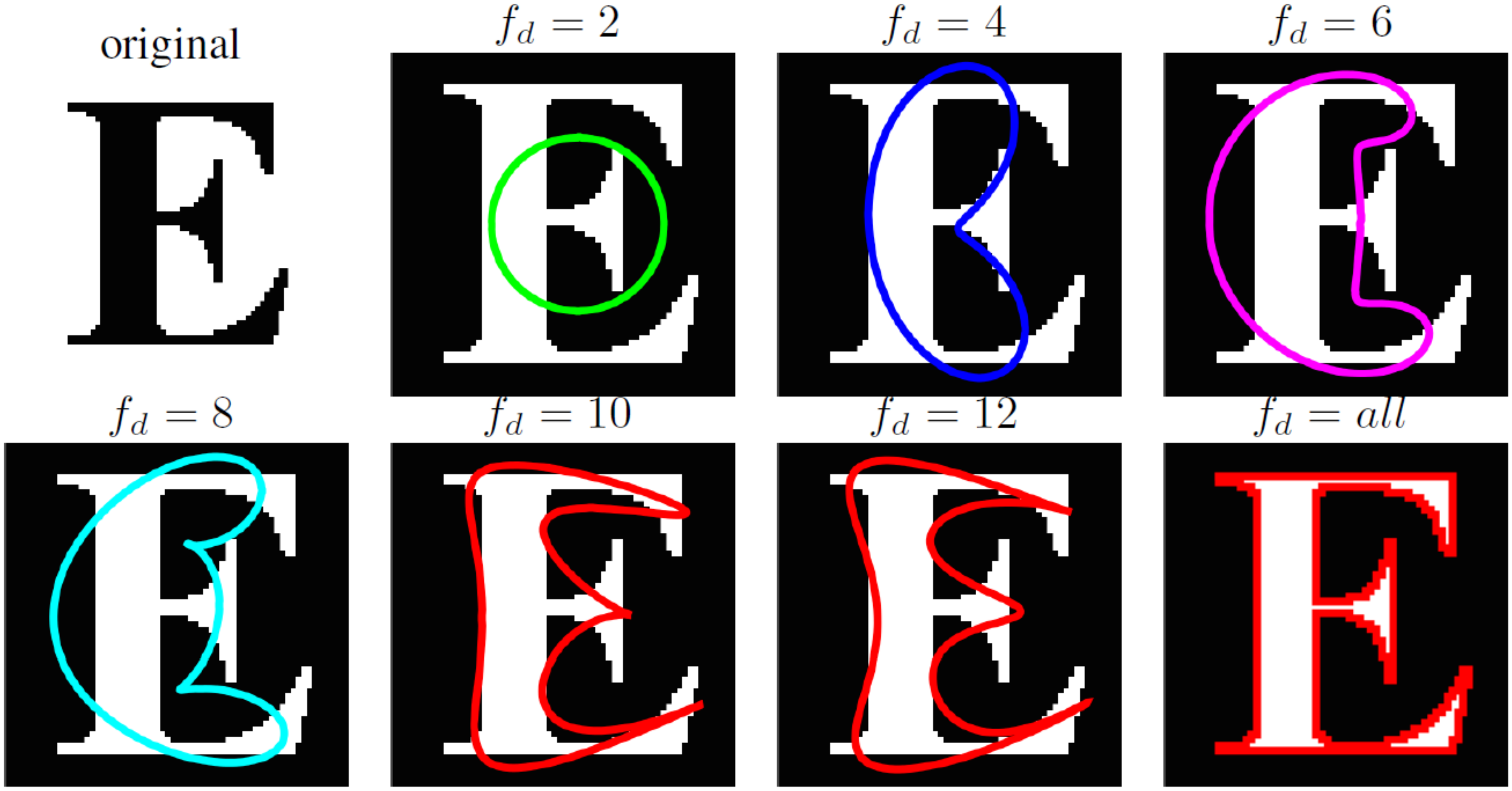
in order **to obtain an approximation** of the shape. This is equivalent to setting all terms

$$a(m) = 0 \qquad \text{for higher frequencies}$$

Note: The more descriptors (coefficients $a(m)$) we use to reconstruct the original curve the closer the result gets to the original curve.

Example: Fourier Descriptor

Boundary reconstruction using an increasing number of Fourier coefficients.



Note: As we have N boundary points each reconstruction has N points too!

Use of Fourier Descriptors

- **Smoothing contours:** removing higher frequency descriptors smooth out fine details of a shape.
- **Find similar shapes:** Ignoring $a(0)$ the other Fourier descriptors can be compared against the Fourier descriptors of unknown objects.

F. Larsson, M. Felsberg, Fourier Descriptors for Traffic Sign Recognition

Fig. 3. Upper left: Query image. Upper right: Extracted contours. Lower left: Contours that matched any of the contours in the pedestrian crossing prototype are shown in a non-yellow color. Lower right: The final result after matching against all sign prototypes.

Fourier Descriptors

Some properties of Fourier Descriptors

- apart from the first coefficient $a(0)$, Fourier descriptors are independent of the location
- scale invariance by normalizing the Fourier descriptors ($||a||^2 = 1$, with $a(0)=0$).
- Fourier descriptors show rotational invariance (phase offset)

$$a(m) \rightarrow e^{(i\phi)} a(m)$$

- index shift s (shifted start point of the curve ->phase offset)

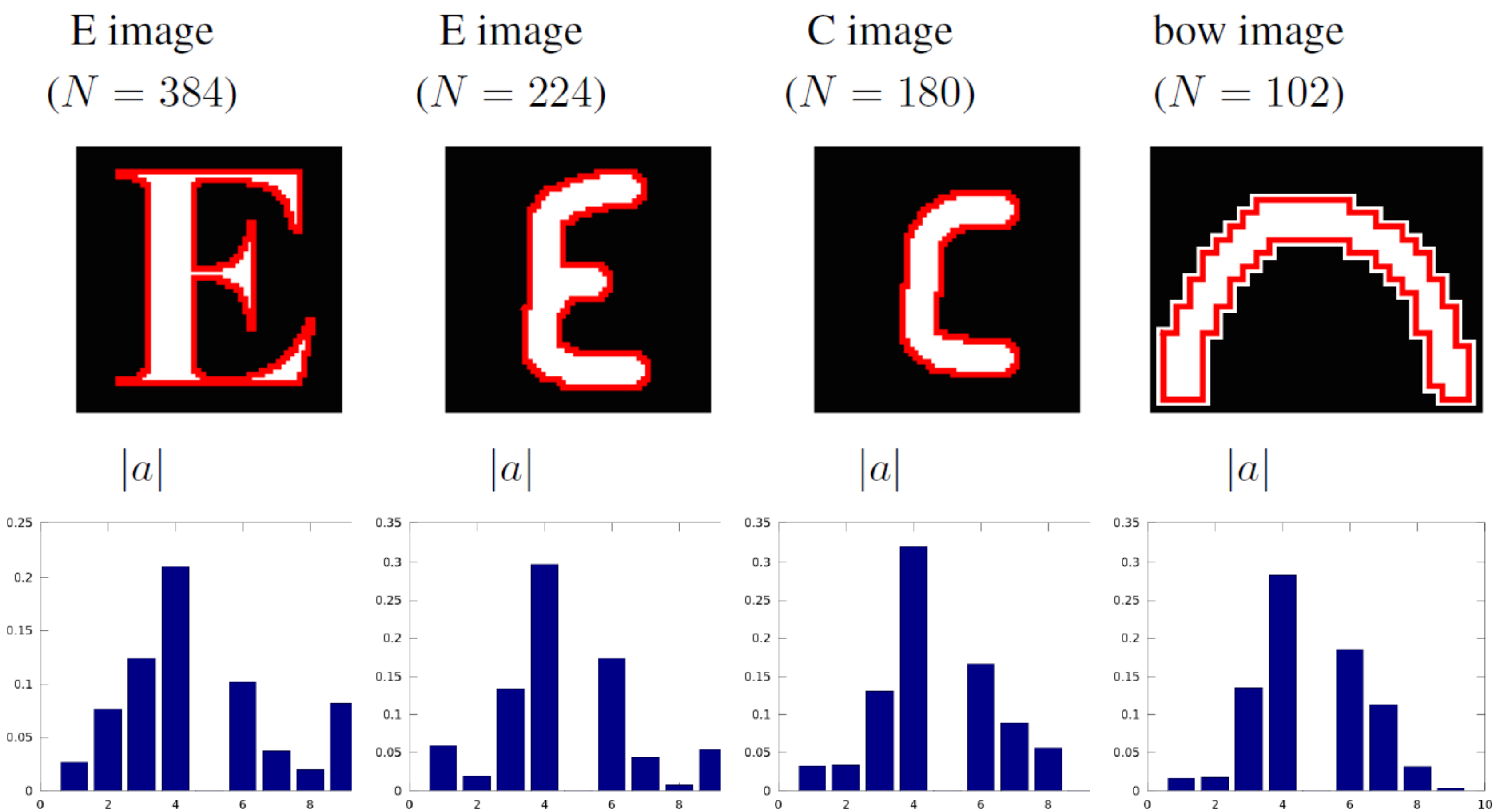
$$a(m) \rightarrow e^{(i 2\pi m \Delta s / N)} a(m)$$

Main Issue:

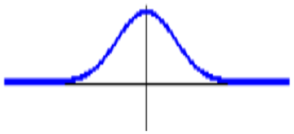
- Fourier descriptors are sensitive to every boundary point
- → occlusion or any distortion affects the full representation

Example: Fourier Descriptor

Visual comparison of a few Descriptors:

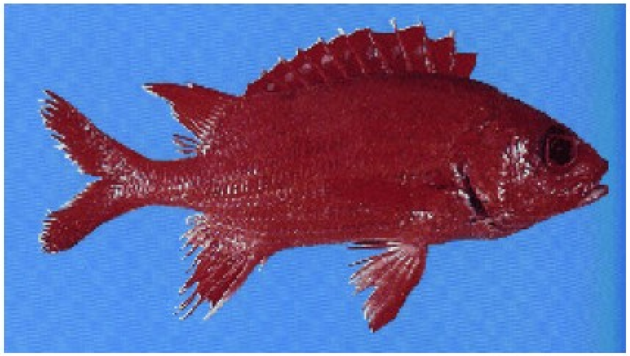


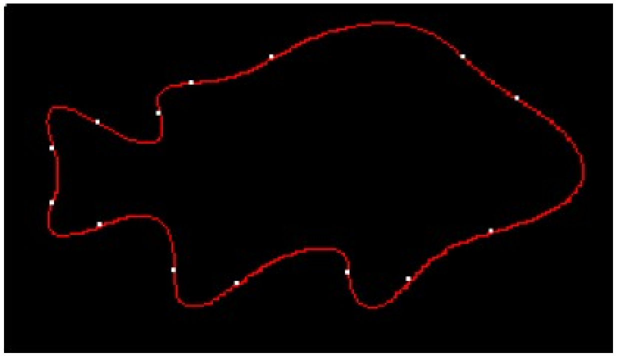
Curvature Scale Space

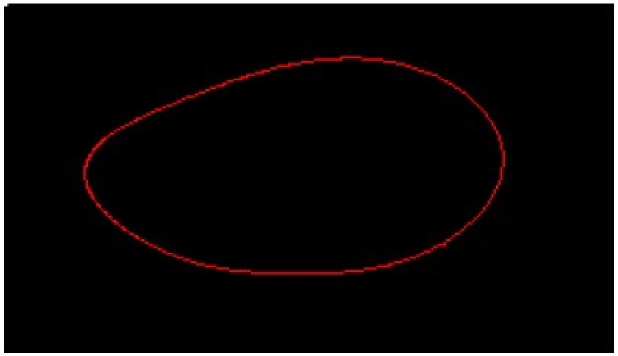


Basic idea:

- **Iteratively smooth** a boundary curve with a Gaussian of increasing size sigma -> results in increasingly smooth shapes
- Find for each iteration **the zero-crossings of the curvature** on the boundary. These points divide the curve into segments.
- Use the run-length **position** (i.e. the length of the path along the border) **of curvature zero-crossings** for all Gaussians to build the ’Curvature Scale Space'.



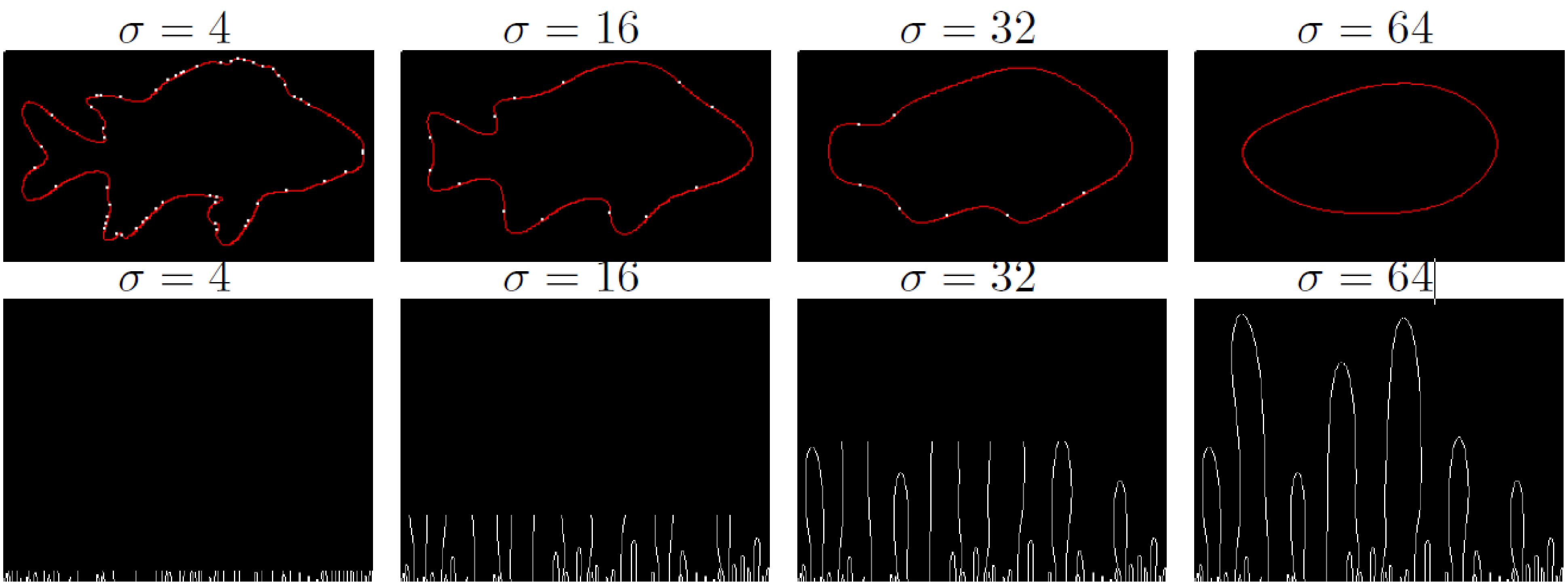




Zero-crossings of the curvature: Points on the contour where the sign of the curvature is changing.

Curvature Scale Space

Construction of the Curvature Scale Space:

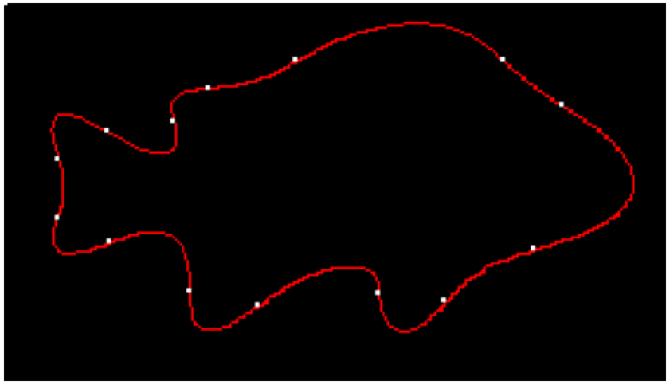
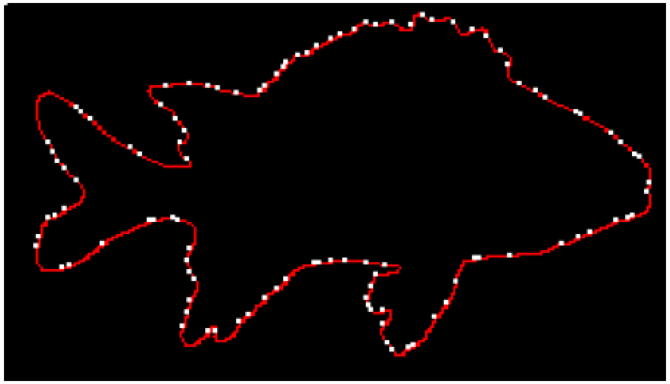
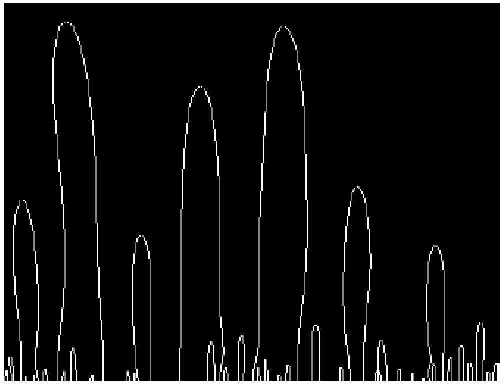


Demo

Curvature Scale Space

Image boundary curvature scale space

Source: http://www.cse.iitk.ac.in/users/amit/courses/768/99/gunjan/

From low resolution (highly smooth) to higher resolution (detailed contour):

- new segmentation points can only appear at higher resolutions
- no existing segmentation point can disappear

→ agrees with our intuition that finer details are detected in higher resolution and that significant structures should not disappear with increasing resolution.

Computational Details

$$\begin{pmatrix} x_1, & y_1 \\ x_2, & y_2 \\ \vdots & \vdots \\ x_N, & y_N \end{pmatrix}$$

Curvature of a Gaussian smoothed curve:

Signed curvature of a parametric curve $\Gamma(t) = (x(t), y(t))$

$$k = \frac{x'y'' - y'x''}{(x'^2 + y'^2)^{3/2}}$$

With $x' = \frac{\partial}{\partial t}x(t)$ and $x'' = \frac{\partial^2}{\partial^2 t}x(t)$, similar for y.

Smooth $\Gamma(t)$ by convolving it with a 1D Gaussian $G(t,\sigma)$ with width σ :

$$\Gamma_\sigma(t) = (x(t) * G(t,\sigma), y(t) * G(t,\sigma))$$

Setting $X(t,\sigma) = x(t) * G(t,\sigma)$ and $Y(t,\sigma) = y(t) * G(t,\sigma)$ we can now compute the **first** and **second derivative** (recall the associativity of the convolution):

$$\frac{\partial}{\partial t}(x * G) = x * \frac{\partial}{\partial t}G = x * G_t$$

(similarly for Y_t and Y_{tt}). With this we can express the curvature of the smoothed curve!

$$X_t(t,\sigma) = x(t) * G_t(t,\sigma) \quad , \quad X_{tt}(t,\sigma) = x(t) * G_{tt}(t,\sigma)$$

Computational Details

Computation:

Convolve $x(t)$, and $y(t)$ with the first and second derivative of the 1D Gaussian $G_t(t,\sigma)$ and $G_{tt}(t,\sigma)$ which gives $X_t(t,\sigma), Y_t(t,\sigma), X_{tt}(t,\sigma)$ and $Y_{tt}(t,\sigma)$

→ these can directly be used in the curvature computation:

Curvature of a Gaussian smoothed curve:

$$k(t,\sigma) = \frac{X_t(t,\sigma)Y_{tt}(t,\sigma) - Y_t(t,\sigma)X_{tt}(t,\sigma)}{(X_t(t,\sigma)^2 + Y_t(t,\sigma)^2)^{3/2}}$$

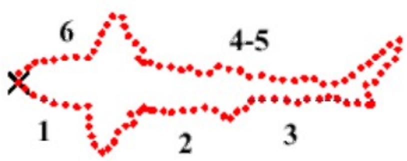
→ **Finally determine zero-crossings in the curvature** k

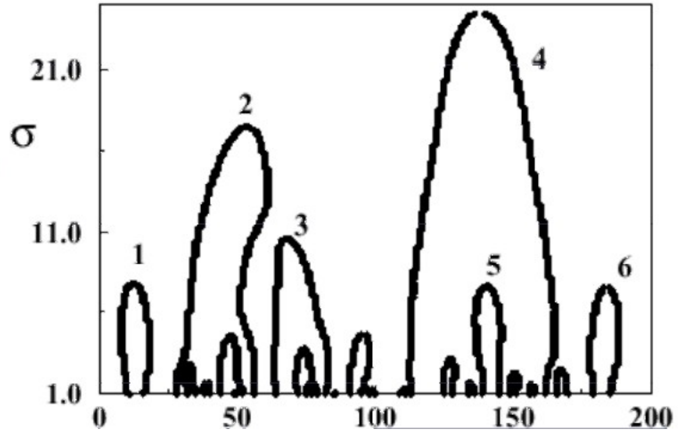
Note: In the discrete case t denotes the index of the border points and the Gaussian can be approximated by a binomial mask allowing to perform discrete convolution.

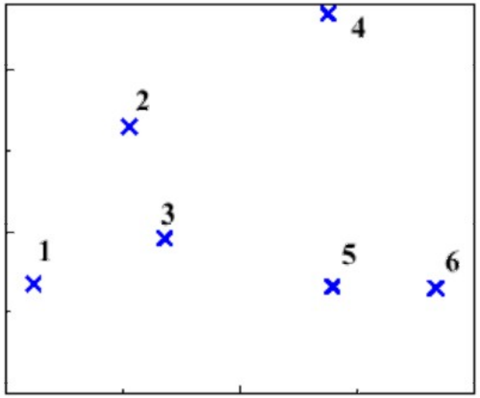
Shape Representation

The Curvature Scale Space can be used for a compact representation of shape that allows to compare different shapes.

Idea: Represent the shape by the **locations of the maxima** of its CSS image

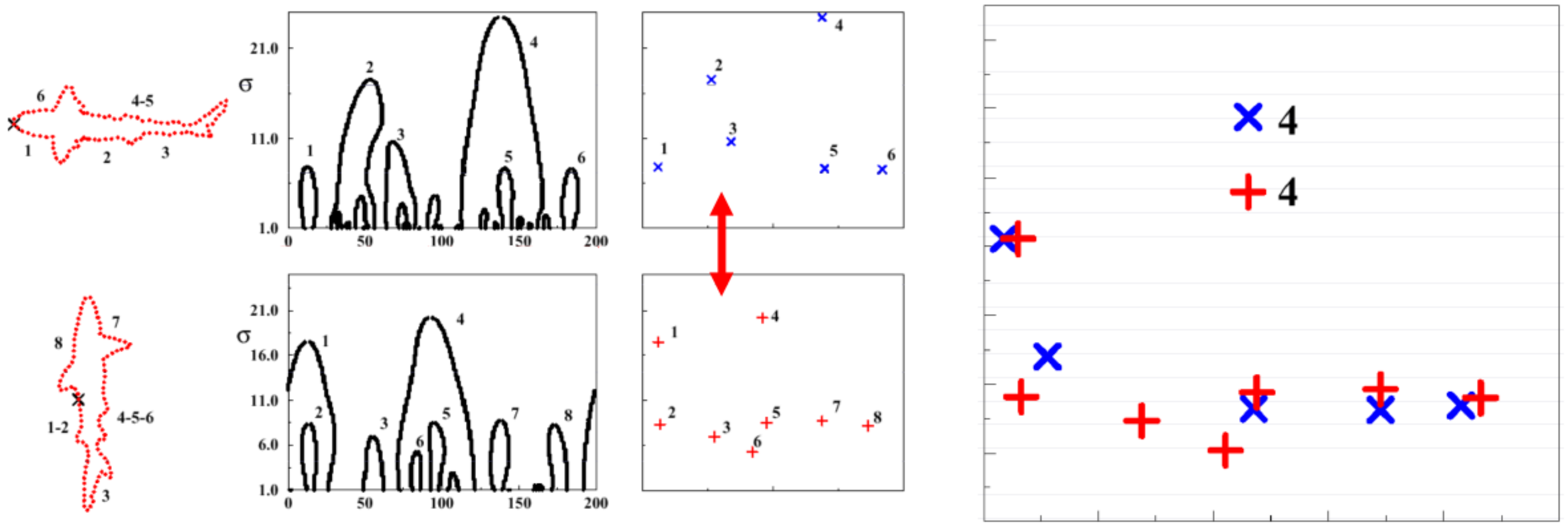




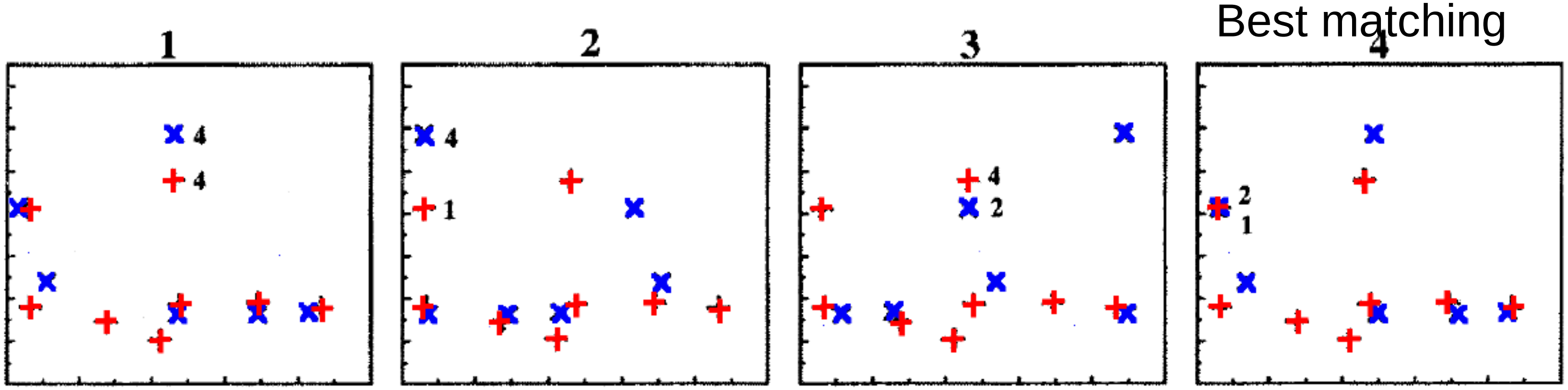


→ the Shape is then represented by a set of 2D points that can be matched against reference points of a “model“ curve.

Matching



Best matching



Four best selected choices for matching the two sets of maxima.

Application: Shape Retrieval

Curvature scale space image in shape similarity retrieval (S.Abbasi,F. Mokhtarian,J. Kittler)

→ compare a given shape with all prototype shapes of the shape classes

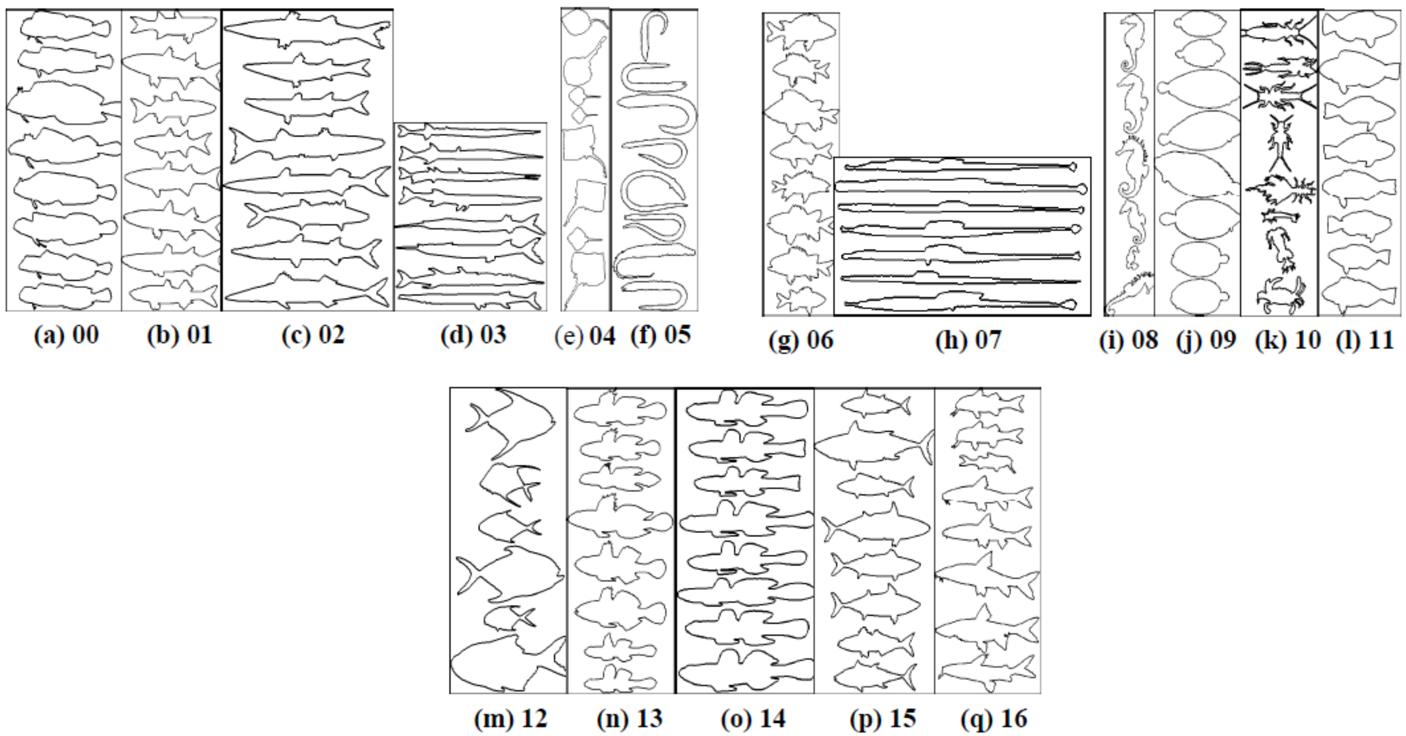


Fig. 6a–q. Classified database used for objective evaluation

The “nearest“ shape determines the class of the shape

Shape Representation/Description

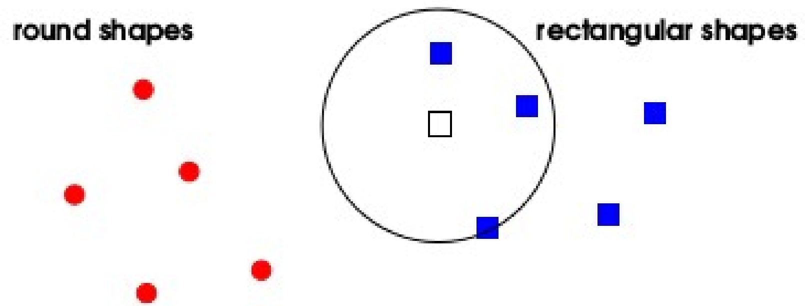
Step towards an image understanding:

Classification/Recognition of image regions.

Requires: A description as **numeric feature vector** which allows distance computations and makes it suitable for the use within some classifiers.

An example is the **k-Nearest Neighbor** classifier: round shapes rectangular shapes.

Other classifiers include **Bayes-Classifier, Support Vector Machines or Neural Networks.**



Summary

Two nice ways to represent Shapes of objects in images → classical Shape Representations/Descriptions:

- Fourier Descriptors
  - Shape description that is invariant against scale, rotation, translation

- Curvature Scale Space
  - Shape description that is invariant against scale, translation and can handle rotation

Applications Robot Vision:
- Street-sign recognition autonomous vehicle (Google)
- Object identification (industrial environments)

Literature

- R.L. Cosgriff, Identification of shape, Ohio State Univ. Res. Foundation, Columbus, OH, Tech. Rep. ASTIA AD, Vol. 254, 1960
- C.T. Zahn and R.Z. Roskies, "Fourier descriptors for plane close curves", IEEE Trans. Computers, Vol C-21, March 1972, pp. 269-281.
- F. Mokhtarian, S. Abbasi, J. Kittler. Efficient and robust retrieval by shape content through curvature scale space, Series on Software Engineering and Knowledge Engineering, 1997