flusser@utia.cas.cz

www.utia.cas.cz/people/flusser

Prof. Ing. Jan Flusser, DrSc.

# Lecture 1 – 2D Features

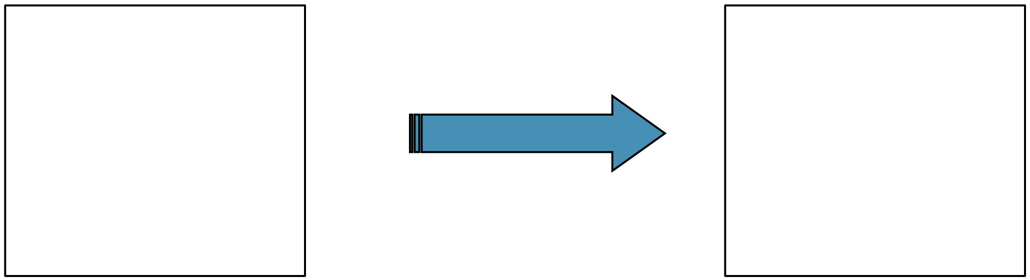
# Digital Image Processing

- Digitization (sampling + quantizing)
- Preprocessing (contrast and brightness changes, denoising, sharpening, … )
- Image analysis (object detection and recognition, scene understanding)
- Image coding (compression)

# Topics of ROZ2

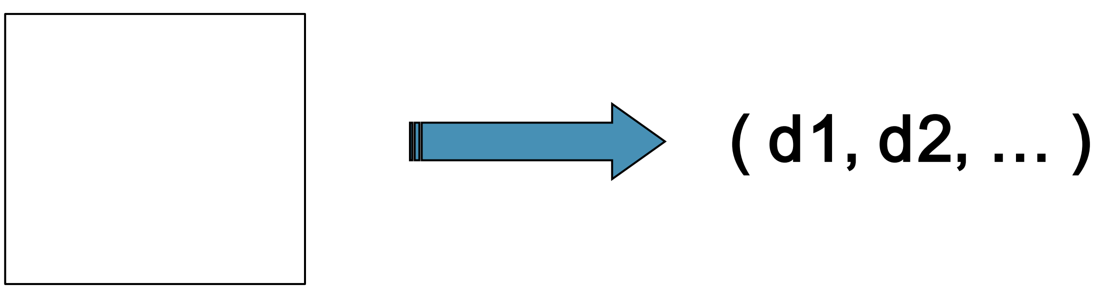
- Digitization (sampling + quantizing)
- Preprocessing (contrast and brightness changes, denoising, sharpening, … )
- Image analysis (object detection and recognition, scene understanding)
- Image coding (compression)

## Image (pre)processing



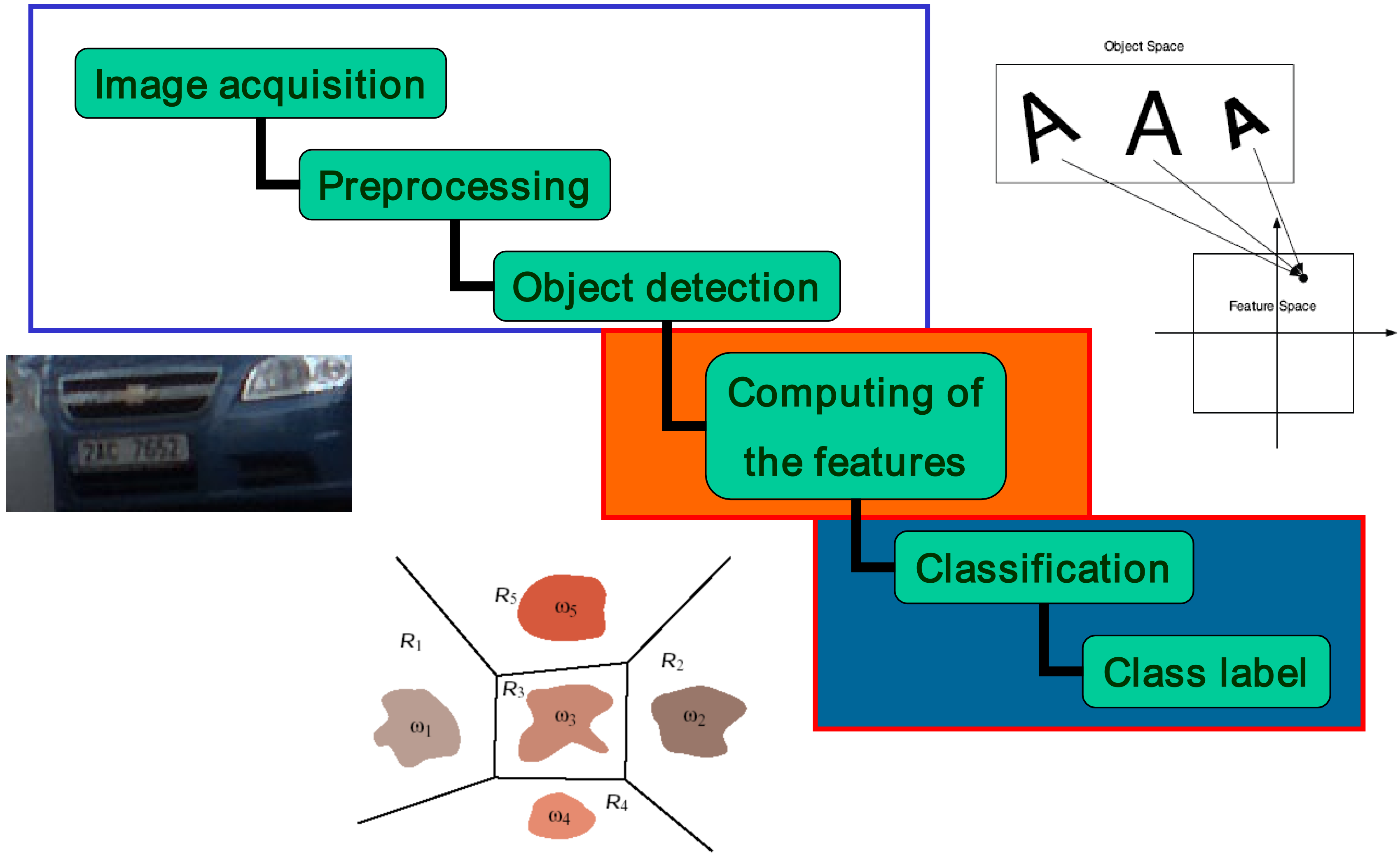
Image

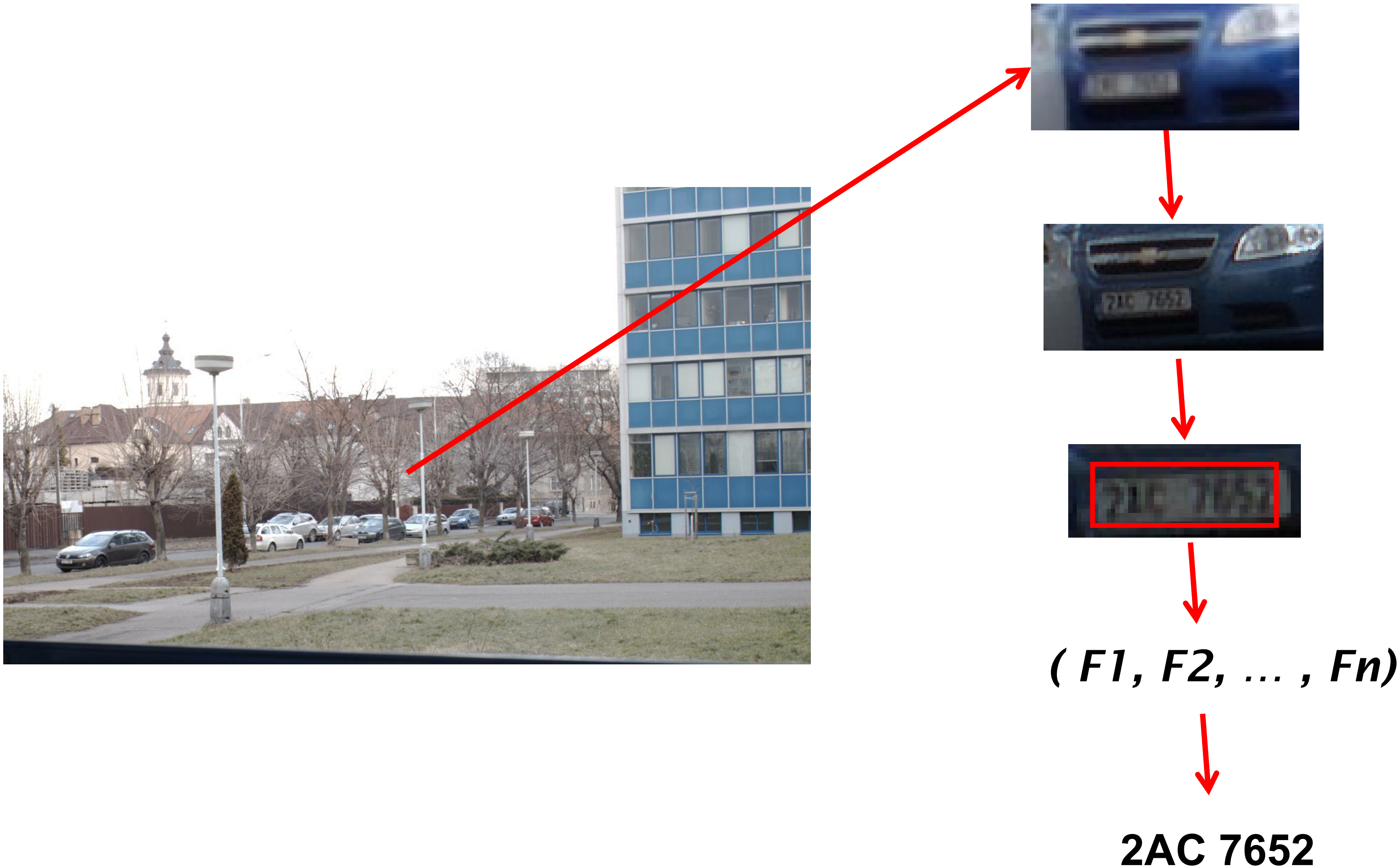
## Image analysis




Non-image

# Object Recognition System




( F1, F2, ... , Fn)
2AC 7652

# Why is visual object recognition so difficult for machines?

## Human beings

- use their lifetime experience as a prior

# Why is visual object recognition so difficult for machines?

## Human beings

- use supplementary information (sound, touch)

# Why is visual object recognition so difficult for machines?

## Human beings

- are very robust to object degradations


# Why is visual object recognition so difficult for machines?

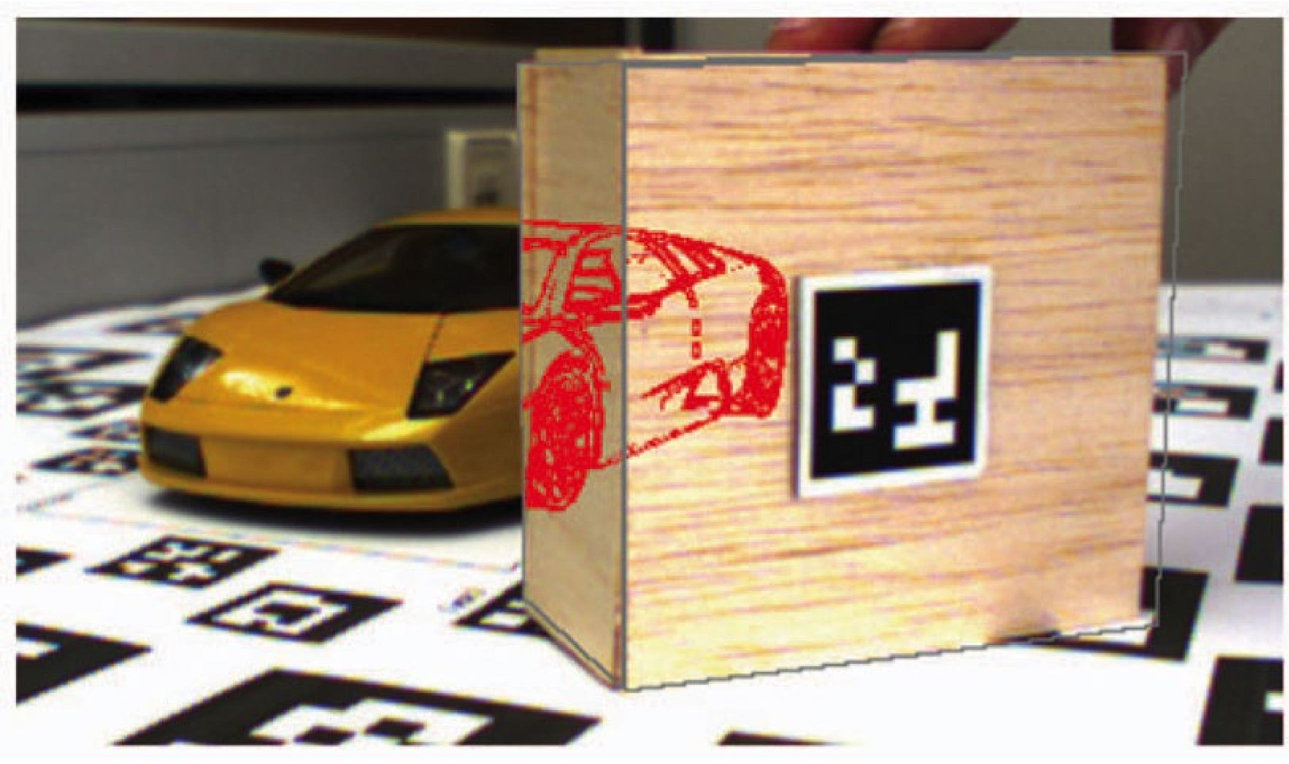
## Human beings

- learn not only from the previous results but also from the recognition process itself

# Why is visual object recognition so difficult for machines?

## Human beings

- can efficiently work with incomplete information



# Why is visual object recognition so difficult for machines?

## Human beings

- can use a broad context

# Features for description

# and recognition of 2D objects

Feature = a point in a metric space (usually in *n*-D Euclidean space) that describes the object

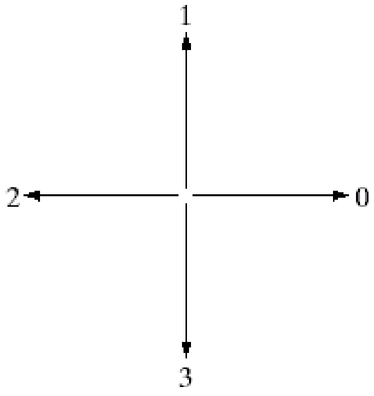
# What is a 2-D object?

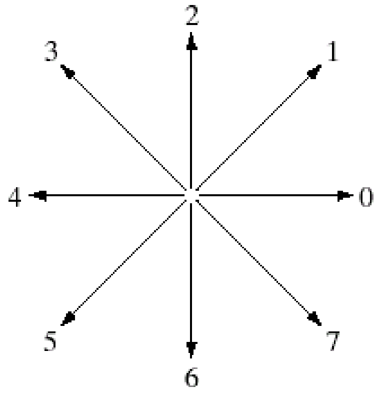
- Binary
- Finite
- Boundary – a simple closed curve or a finite set of them

# What is a (discrete) boundary?

- Boundary pixel – an object pixel having a background neighbor

- What is a neighbor? (Definition of discrete topology)

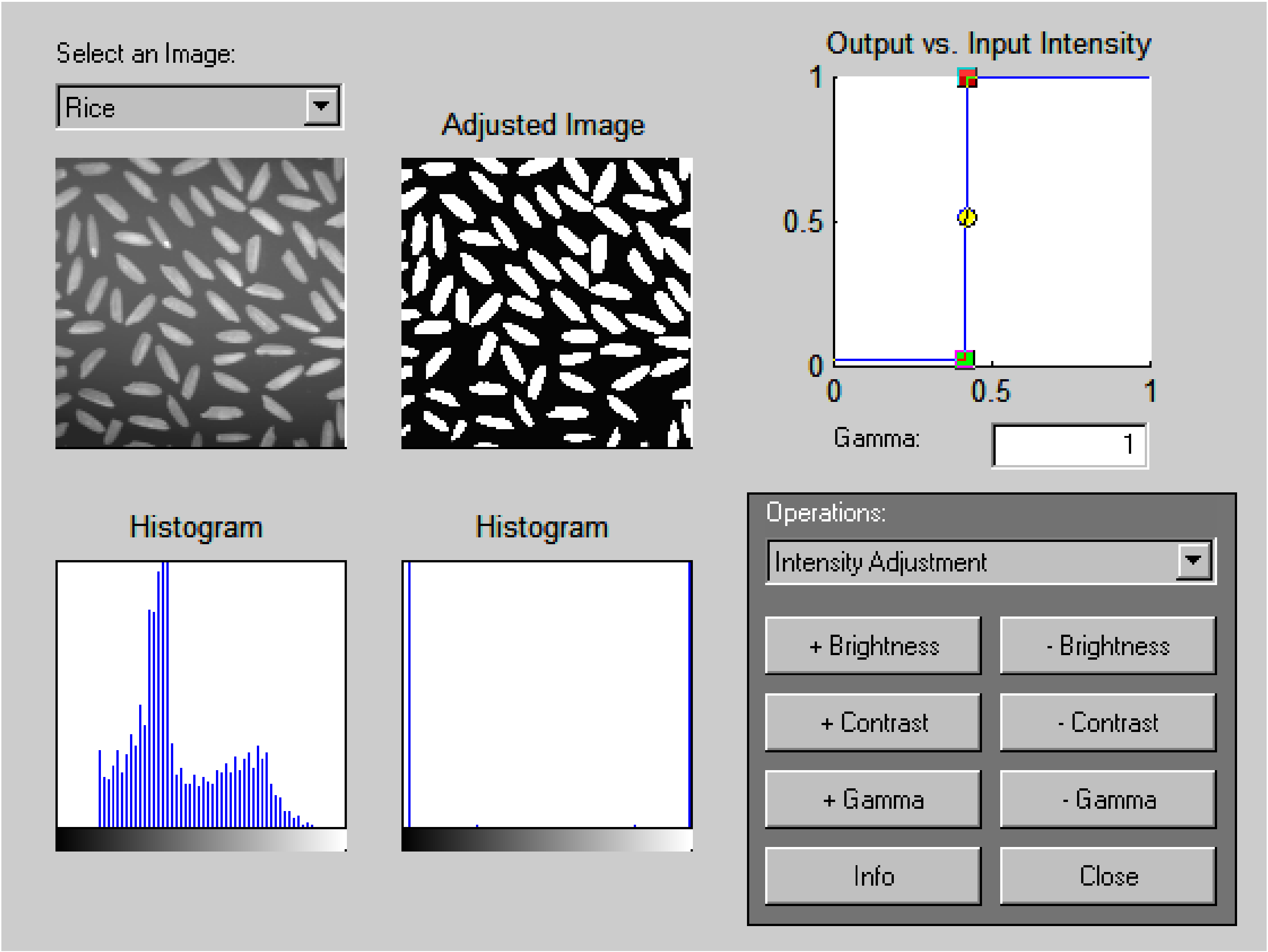


# How to detect objects in the image? (Image segmentation)

- Thresholding
- Edge linking
- Region growing

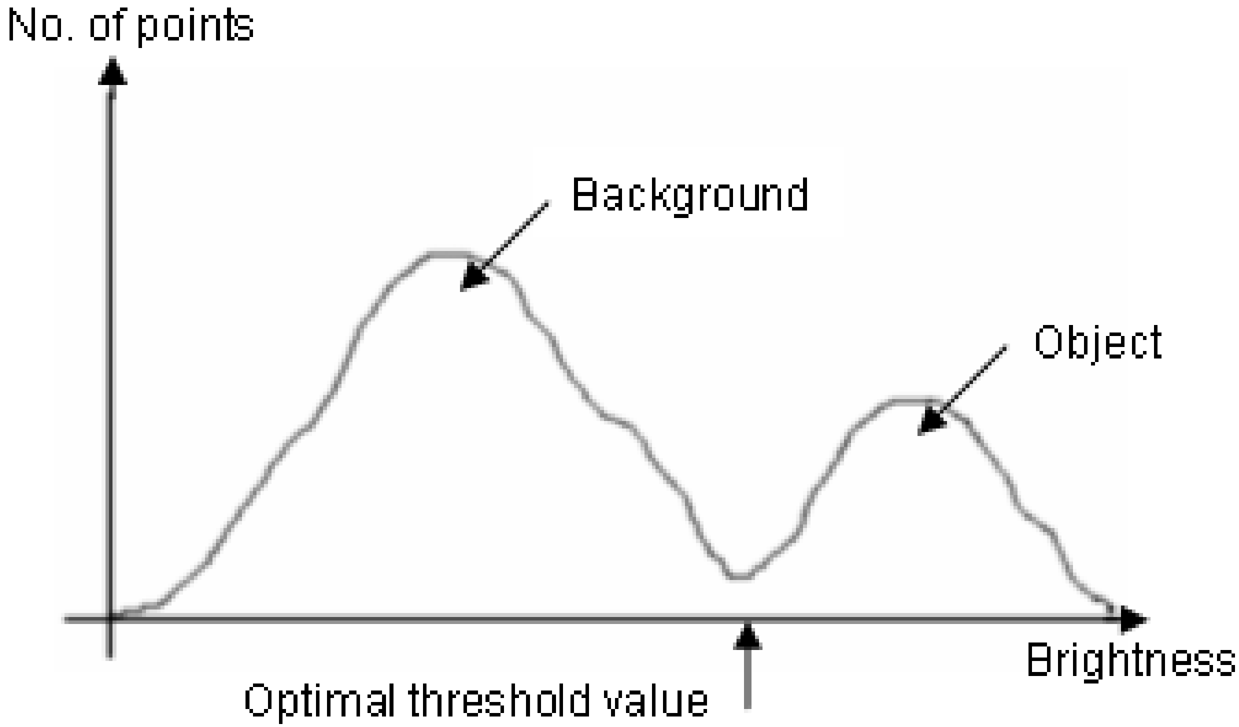
# Thresholding



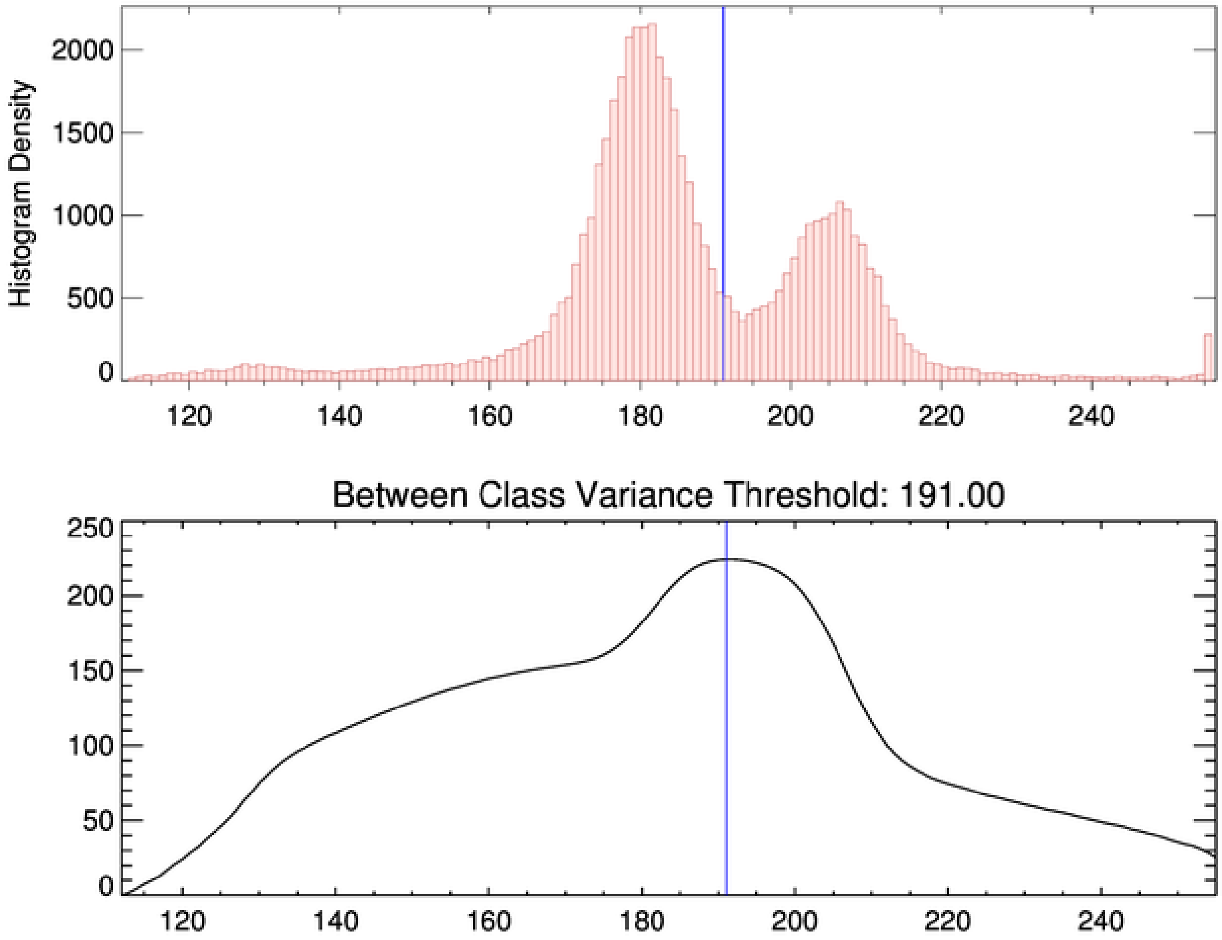
# The Otsu's threshold

$$\sigma_w^2(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_2^2(t)$$

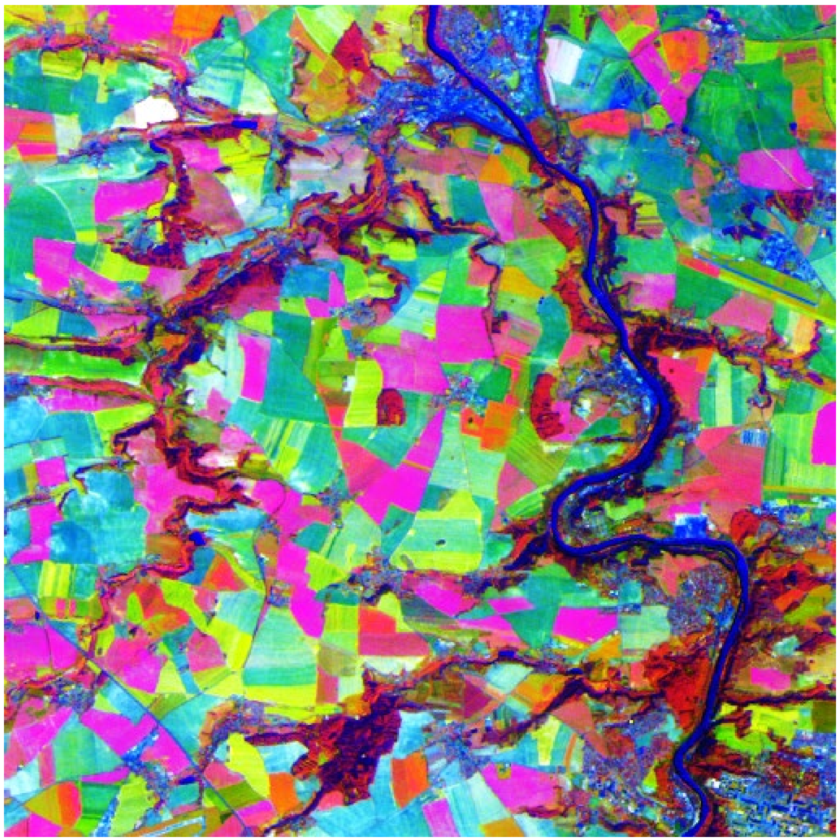
$$\sigma_b^2(t) = \sigma^2 - \sigma_w^2(t) = \omega_1(t)\omega_2(t)\left[\mu_1(t) - \mu_2(t)\right]^2$$



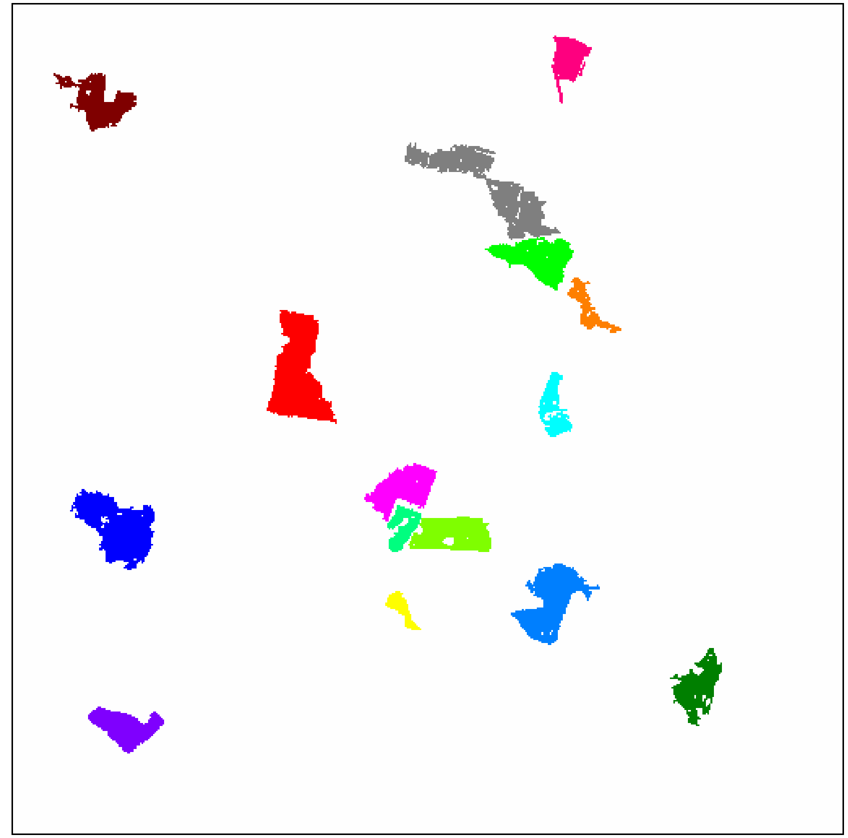
# The Otsu's threshold



# Image segmentation

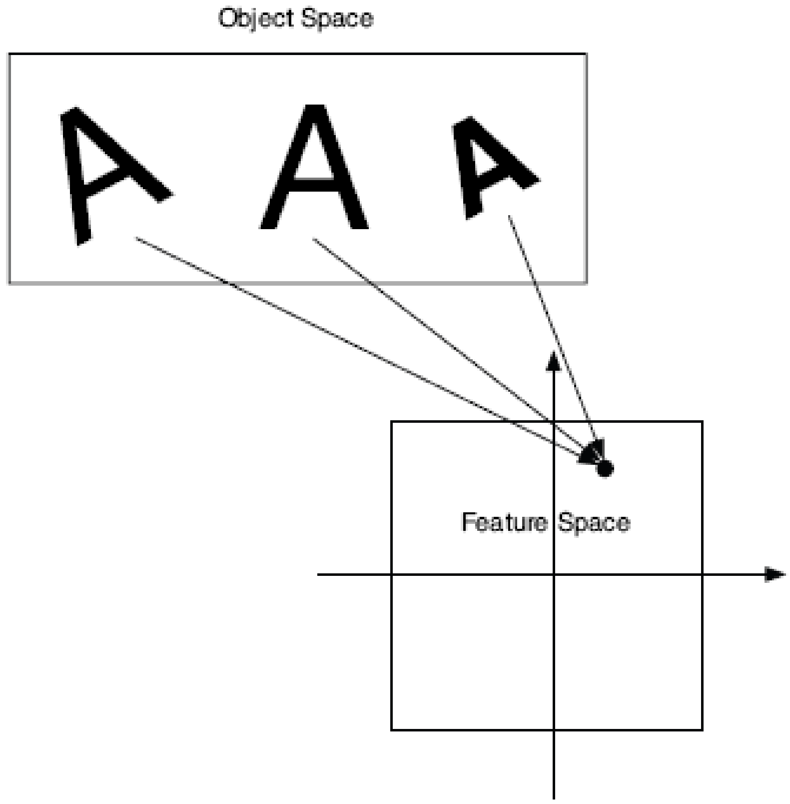


Original

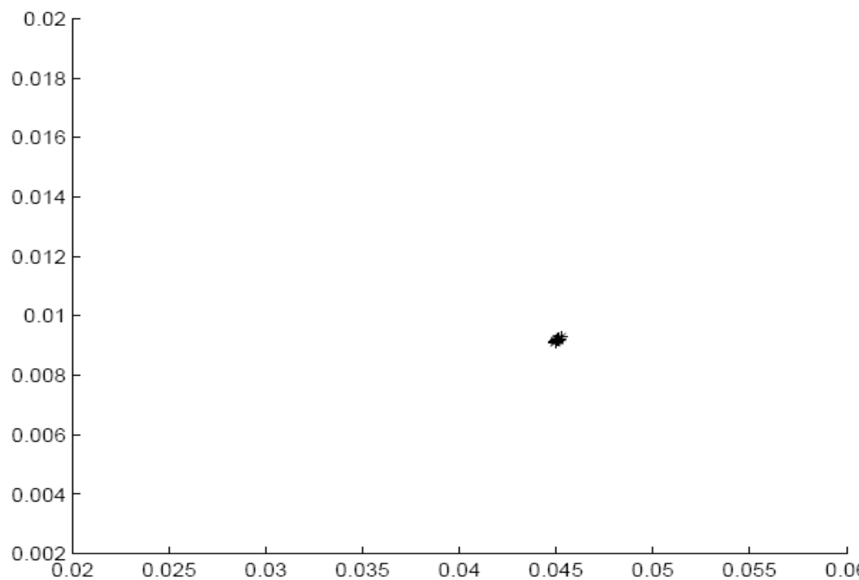


Partial segmentation
Full segmentation

# What are features?



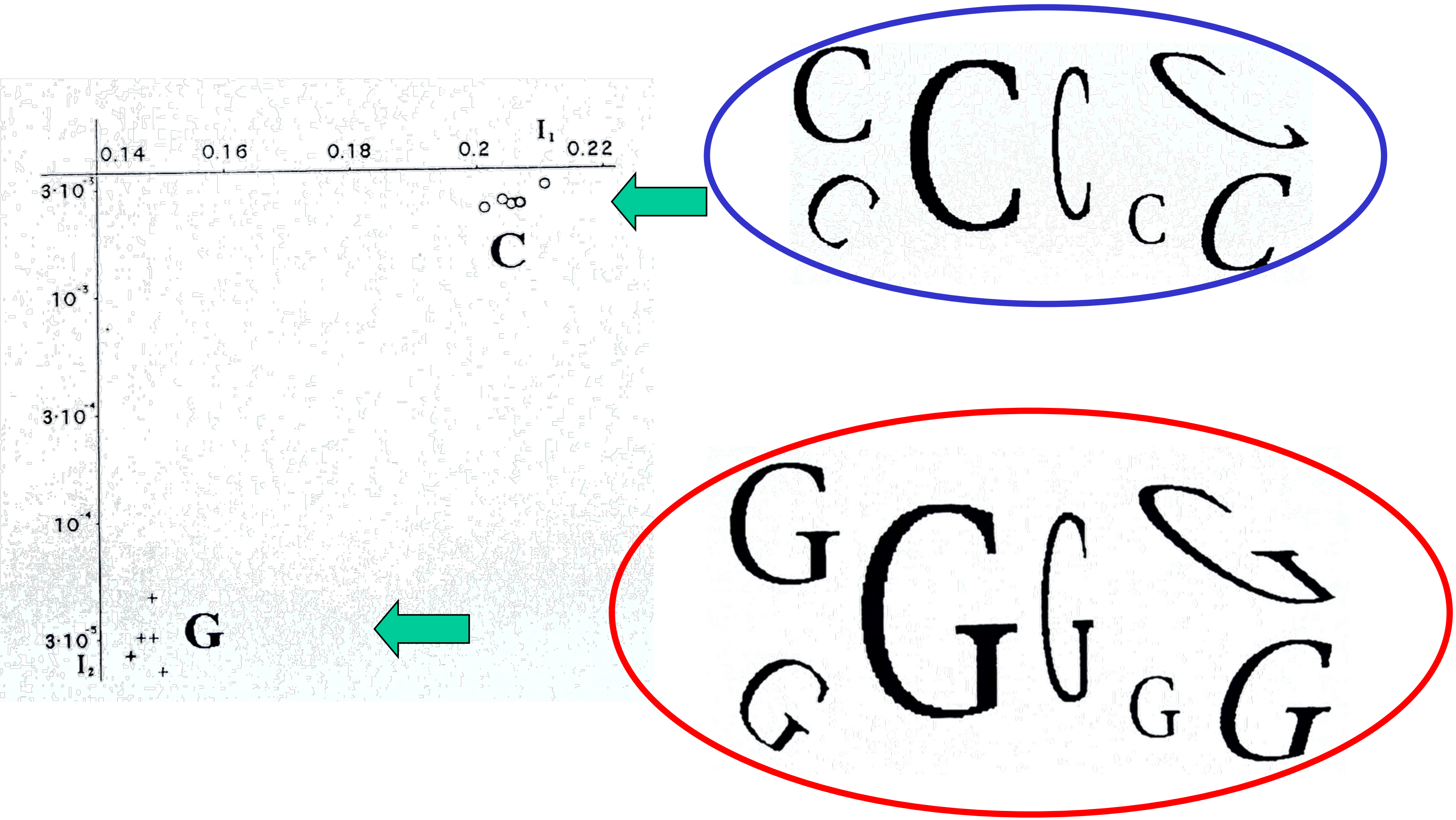
# Example: TRS


0.02
0.018
0.016
0.014
0.012
0.01
0.008
0.006
0.004
0.002
0.02
0.025
0.03
0.035
0.04
0.045
0.05
0.055
0.06

# Desirable properties of the features

- Invariance
- Discriminability
- Robustness
- Efficiency, independence, completeness

# Discrimination power



# Major categories of invariants

**Simple "visual" shape descriptors**

- compactness, convexity, elongation, ...

**Transform coefficient invariants**

- Fourier descriptors, wavelet features, ...

**Point set invariants**

- positions of dominant points

**Differential invariants**

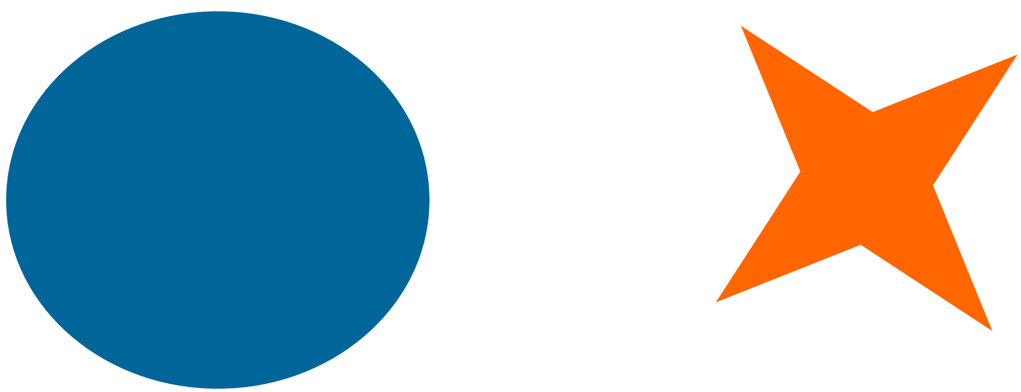
- derivatives of the boundary

**Moment invariants**

# Visual features for binary objects

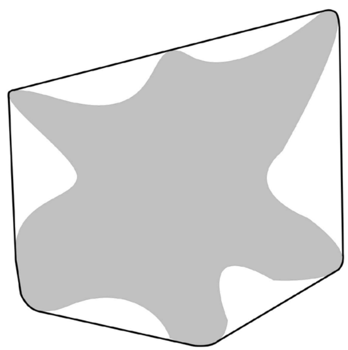
**Simple features**

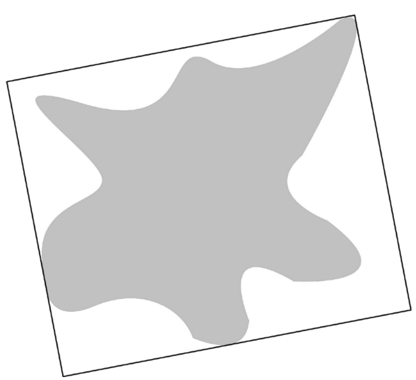
- Compactness $\frac{4\pi P}{O^2}$
- Convexity $\frac{P(A)}{P(C_A)}$



# Visual features for binary objects

## Simple features





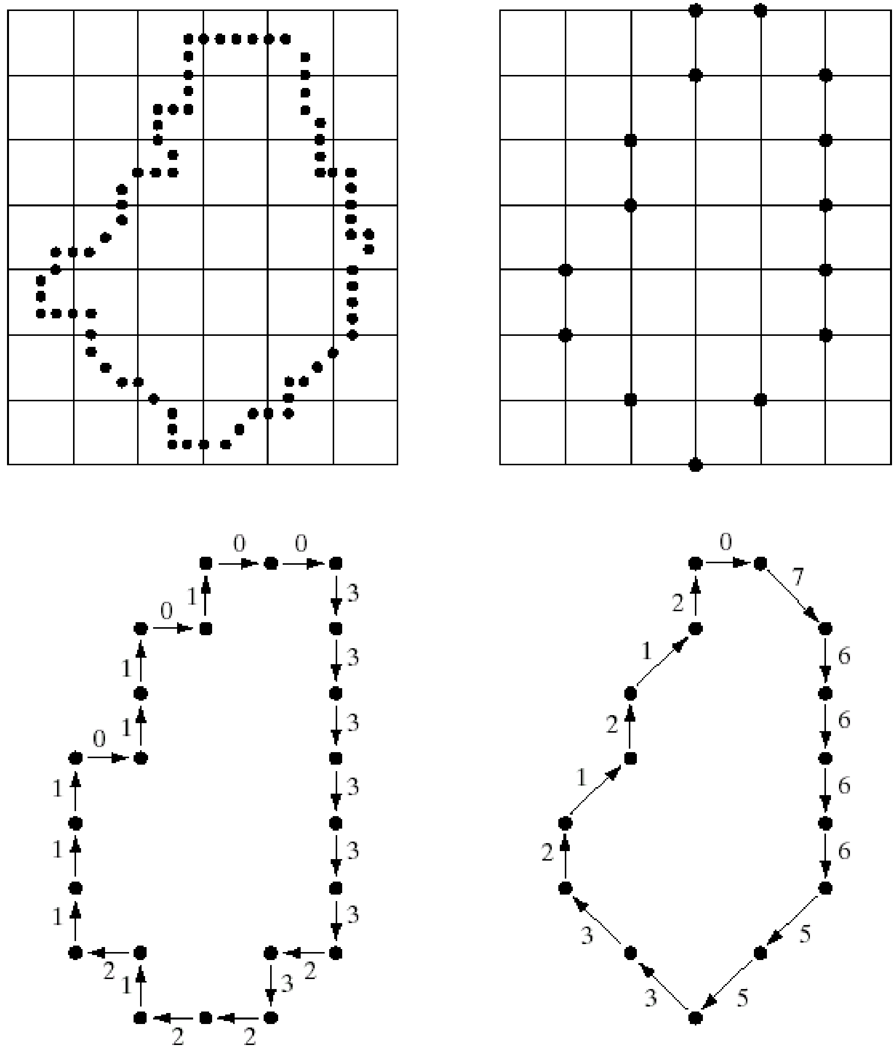
- Compactness $\frac{4\pi P}{O^2}$
- Convexity $\frac{P(A)}{P(C_A)}$
- Elongation
- Rectangularity
- Euler number

# Visual features for binary objects

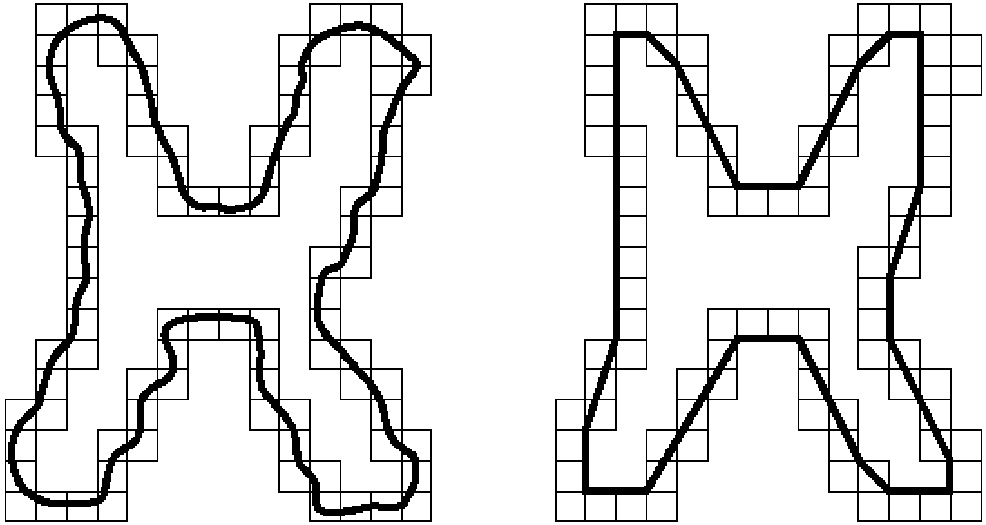
## “Complete” features

- Chain code
- Polygonal approximation
- Shape vector
- Shape matrix
- Other encodings of the radial function

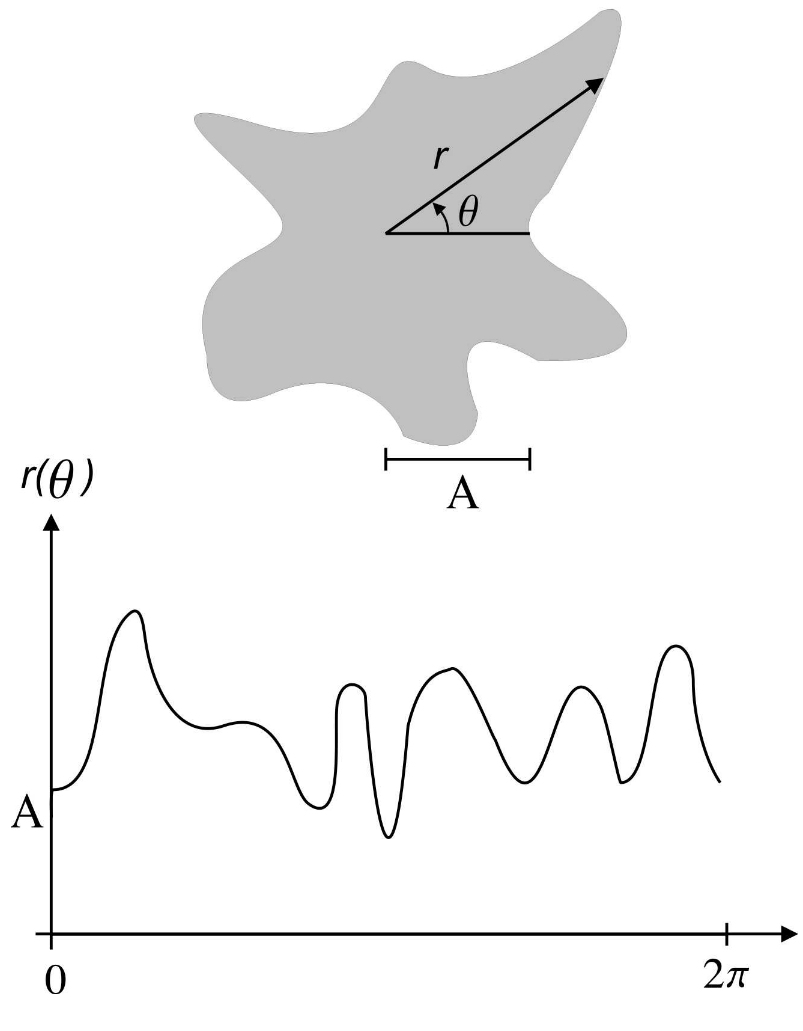
# Chain code



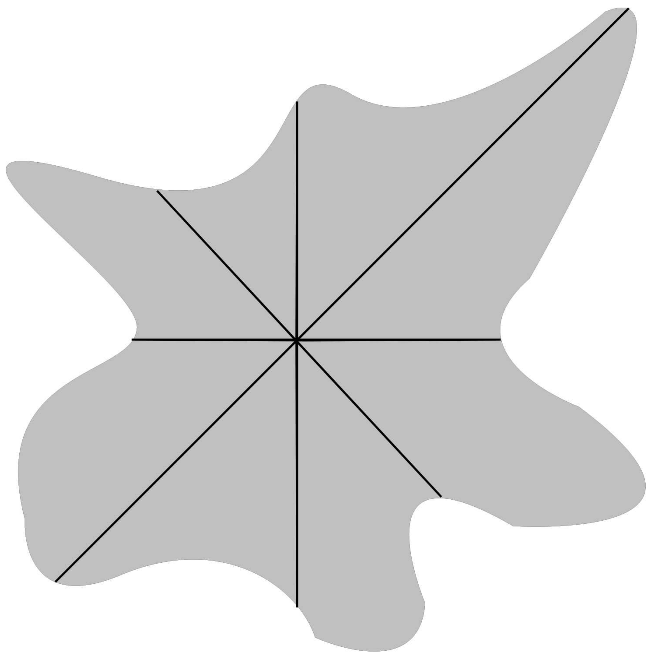
# Polygonal approximation



# Radial function



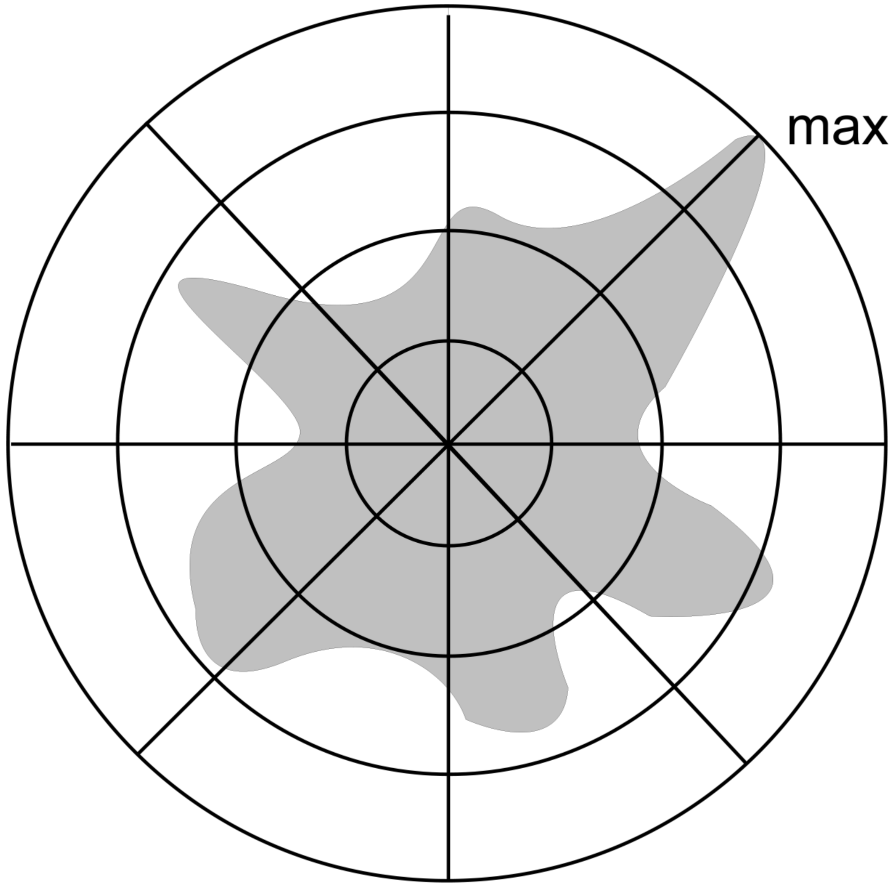
# Shape vector



$$v = (d1, d2, \ldots, dn)$$

$n = 8$

# Shape matrix




$m = 4$

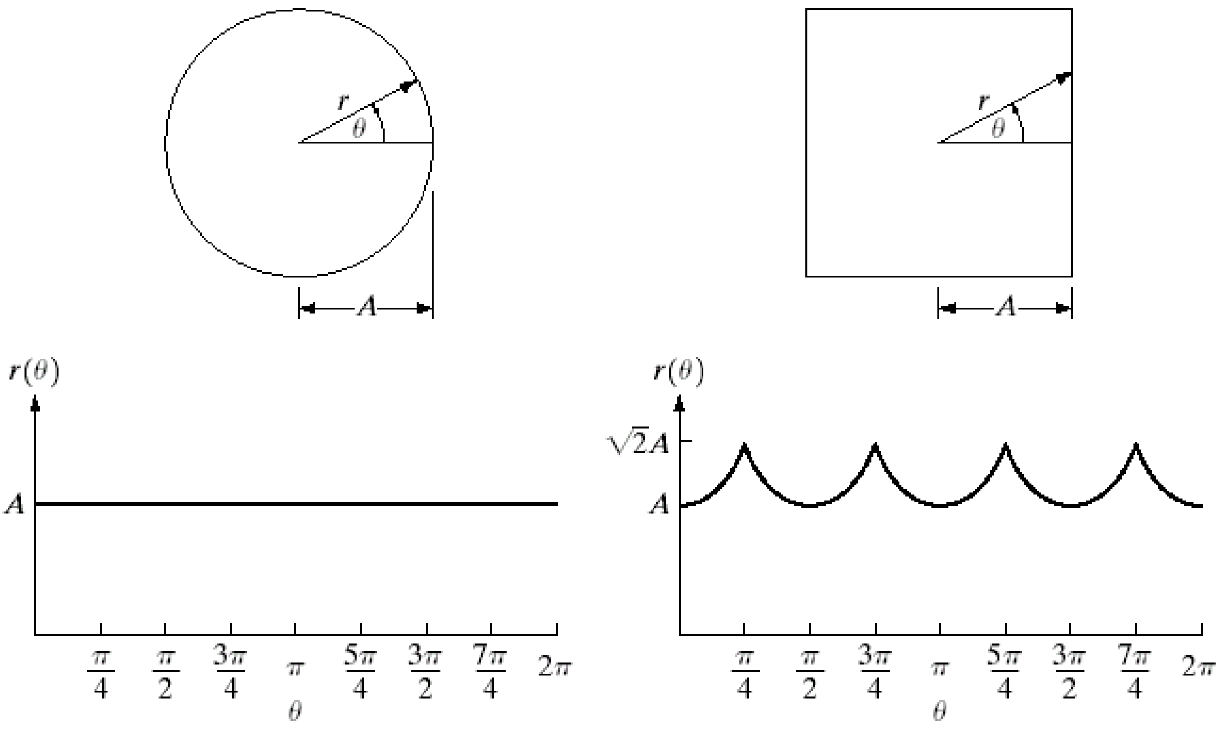
$n = 8$

$$B = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{vmatrix}$$

# Transform coefficient features

- Fourier descriptors
- Wavelet-based features
- Other transform coefficients

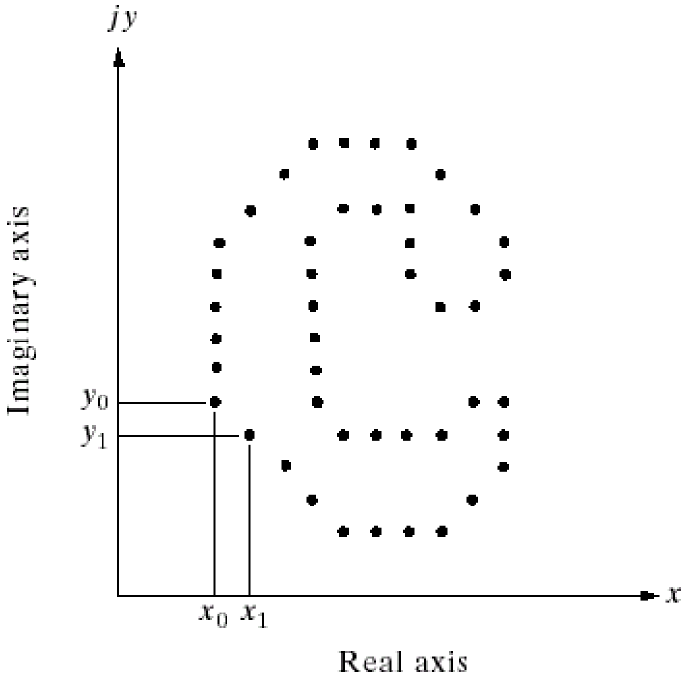
# Fourier descriptors




$$f(t) = ([x(t) - x_c]^2 + [y(t) - y_c]^2)^{1/2}$$

# Fourier descriptors

$$z_n = x_n + i \cdot y_n$$




$$Z_k = \sum_{n=0}^{N-1} z_n e^{-2\pi i k n / N}$$

$$C_k = |Z_k| / |Z_1|, \; k = 2, 3, \ldots$$

# Shift invariance

$$\sum_{n=0}^{N-1} (z_n - z) e^{-2\pi ikn/N} = \sum_{n=0}^{N-1} z_n e^{-2\pi ikn/N} + z \sum_{n=0}^{N-1} e^{-2\pi ikn/N}$$

$$\sum_{n=0}^{N-1} e^{-2\pi ikn/N} = 0, \; k \neq 0$$

$$\sum_{n=0}^{N-1} e^{-2\pi ikn/N} = N, \quad k = 0$$

# Rotation invariance

$$\sum_{n=0}^{N-1} (z_n e^{\varphi i}) e^{-2\pi i k n/N} = e^{\varphi i} \sum_{n=0}^{N-1} z_n e^{-2\pi i k n/N}$$

# Scaling invariance

$$\sum_{n=0}^{N-1} (cz_n) e^{-2\pi ikn/N} = c \sum_{n=0}^{N-1} z_n e^{-2\pi ikn/N}$$

# Invariance to the starting point

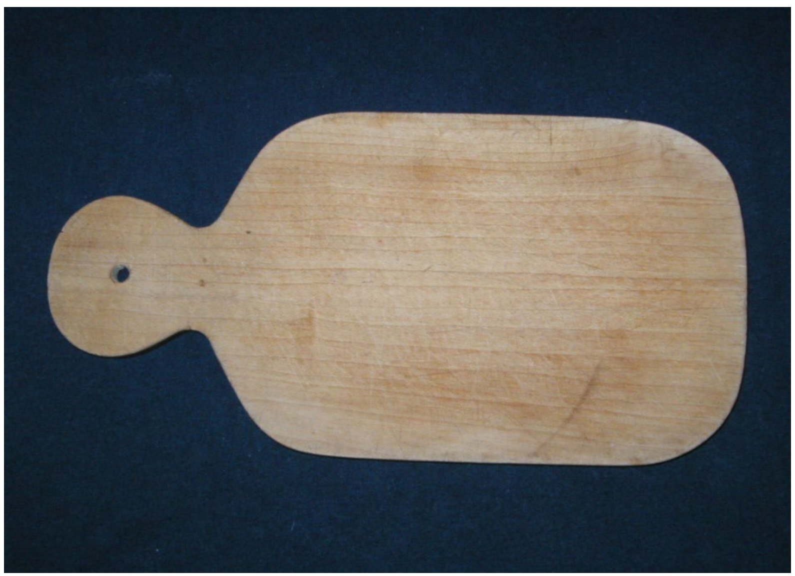
- Fourier Shift Theorem

# Wavelet features

# Differential invariants

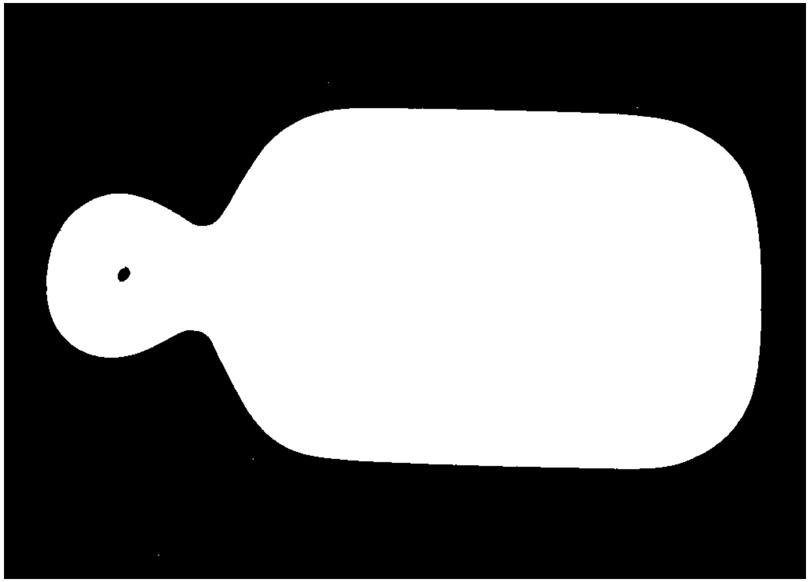
Motivation: recognition of occluded objects

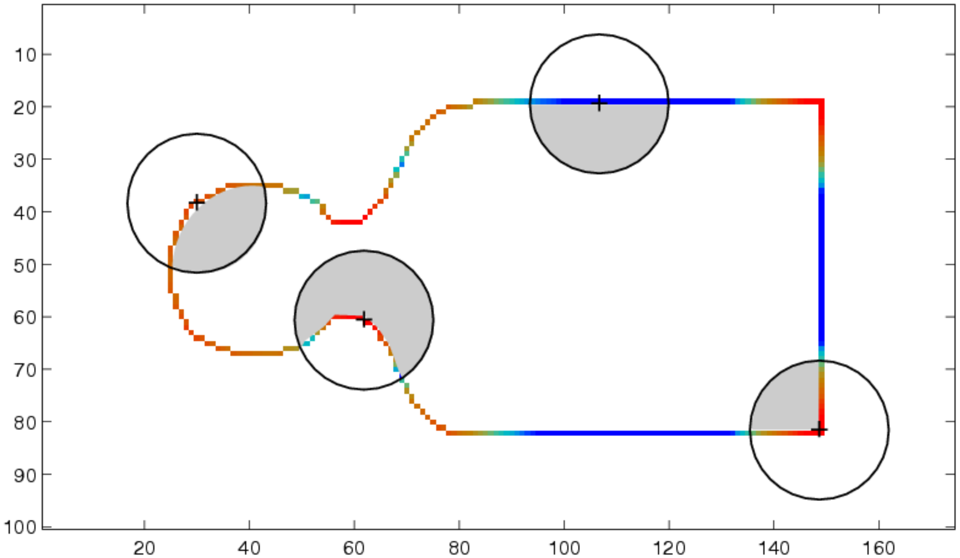
→ Features must be local



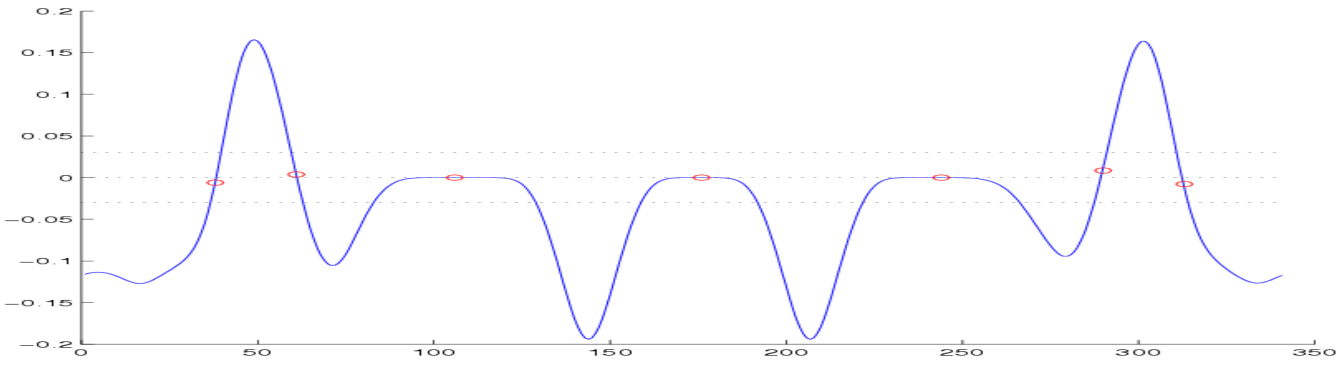
# Differential invariants – an example

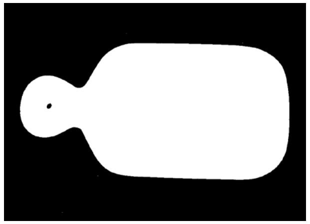
$$c(t) = \frac{\dot{x}\ddot{y} - \ddot{x}\dot{y}}{(\dot{x}^2 + \dot{y}^2)^{3/2}}$$

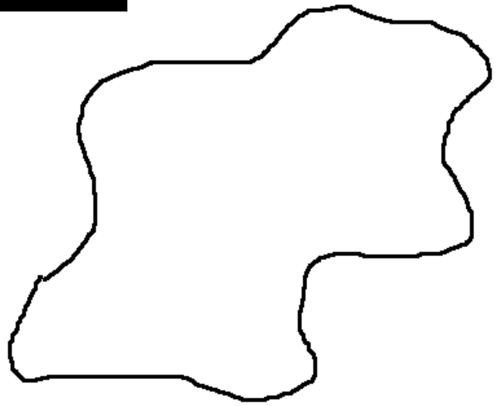


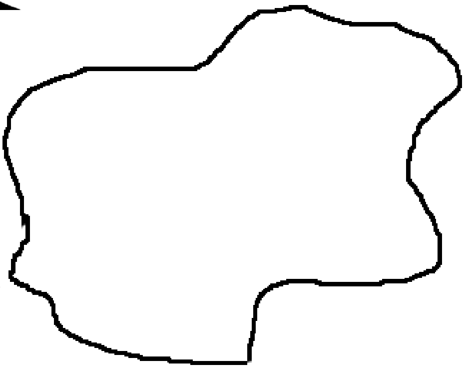


# Differential invariants – an example







# Differential invariants

DI’s are composed from higher-order derivatives of the boundary

DI’s are invariant to affine and even to projective transform but are extremely unstable

# Semi-differential invariants

**Motivation: avoiding high-order derivatives while preserving the locality**

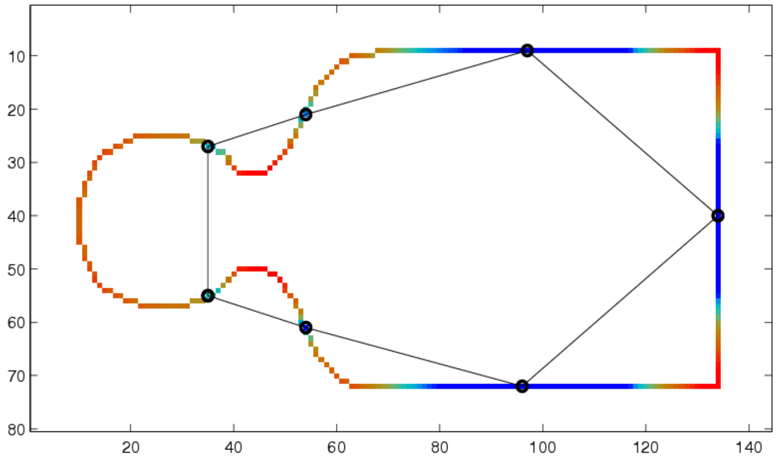
Decomposition of the object into parts, each part is then described by global invariants

The decomposition is often based on inflection points (they are affine- and projective-invariant)

$$\dot{x}\ddot{y} - \ddot{x}\dot{y} = 0$$

# Dividing the object into invariant parts

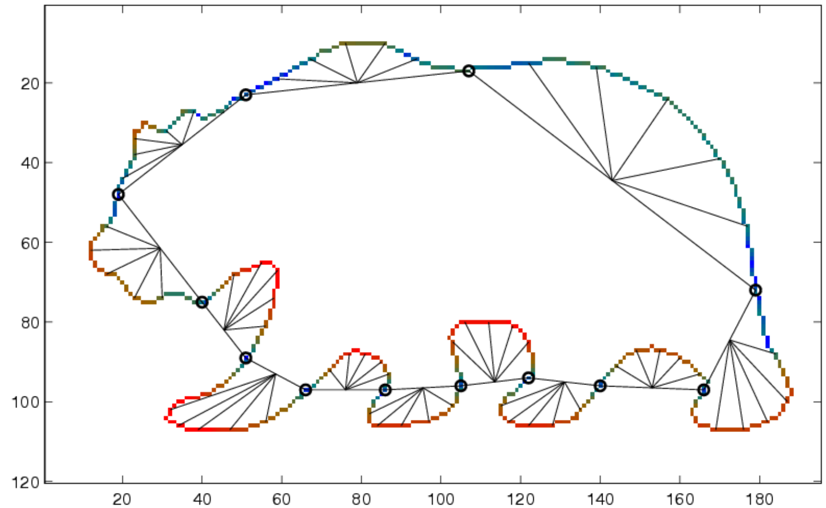
- Inflection points and centers of straight lines

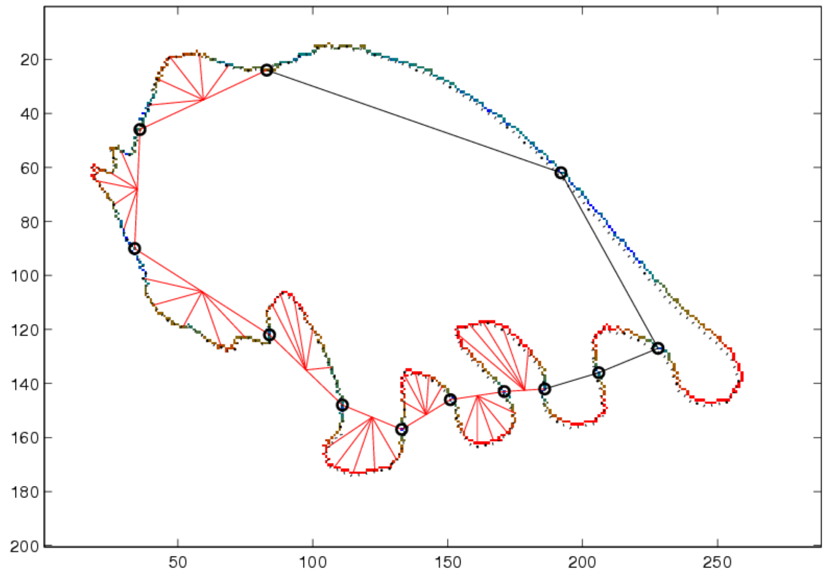


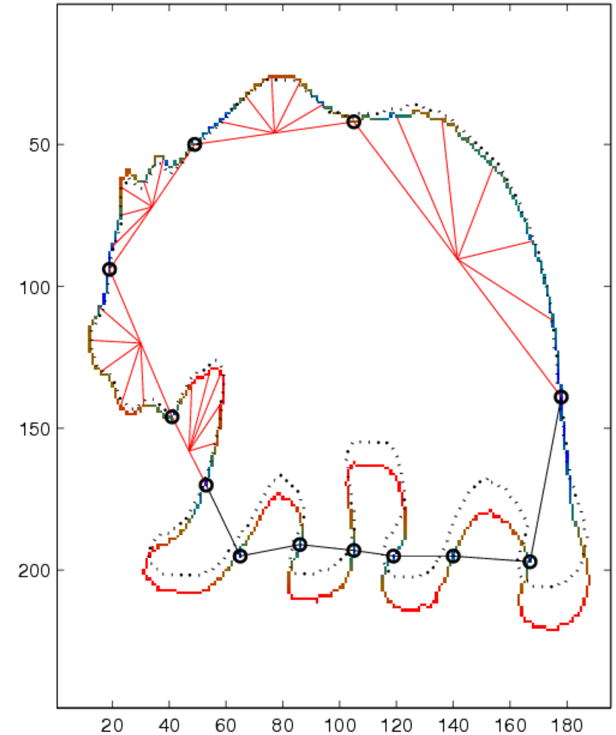

$$\dot{x}\ddot{y} - \ddot{x}\dot{y} = 0$$

- Computing invariants of each part

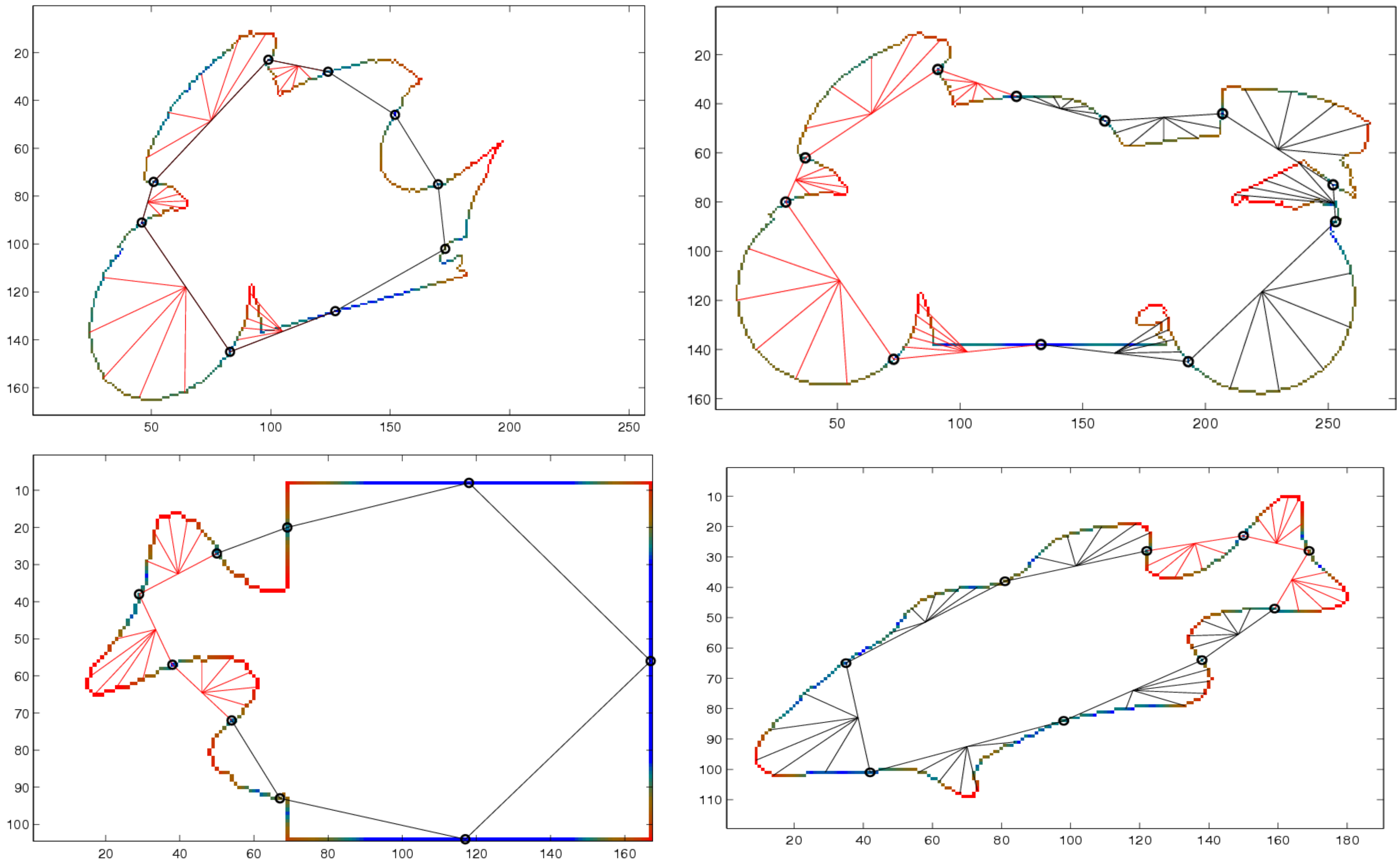
# Affine-invariant radial vectors



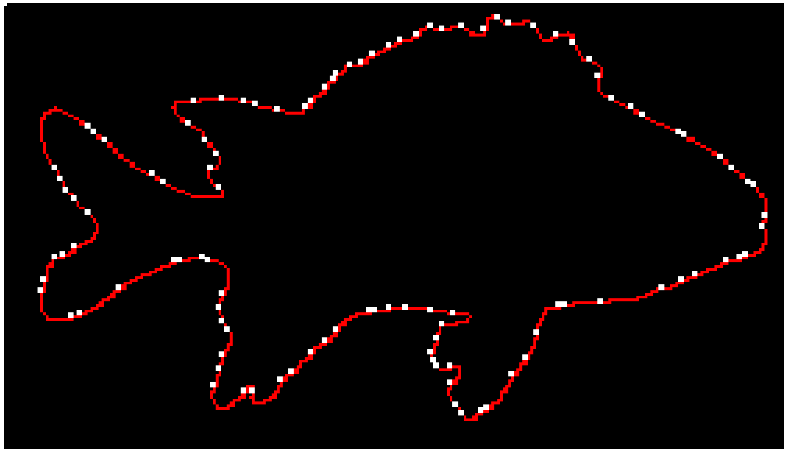



# Recognition of occluded objects

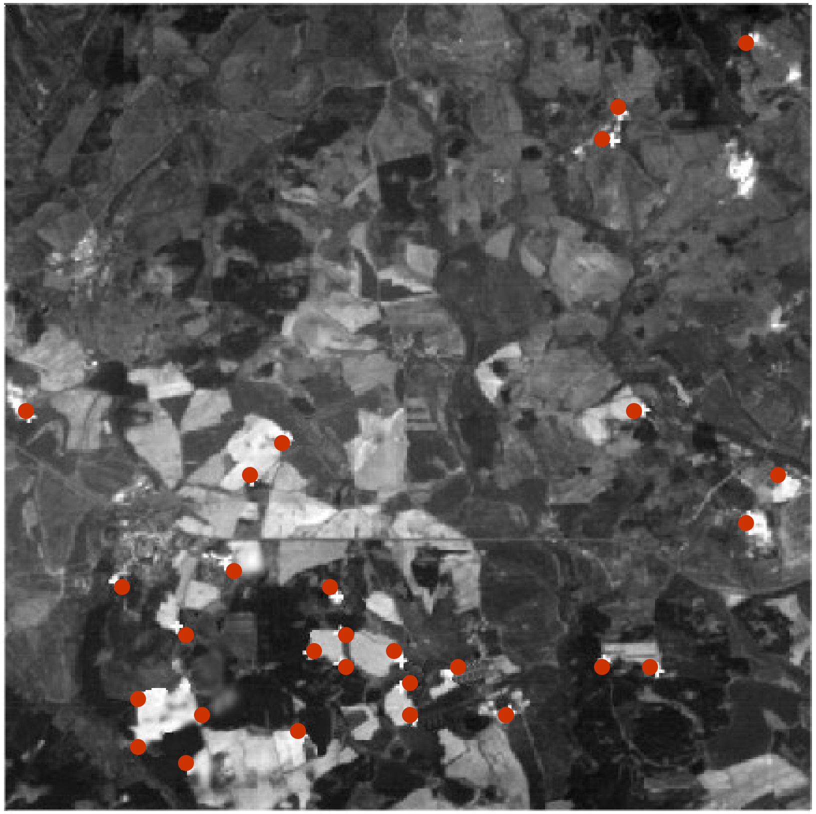


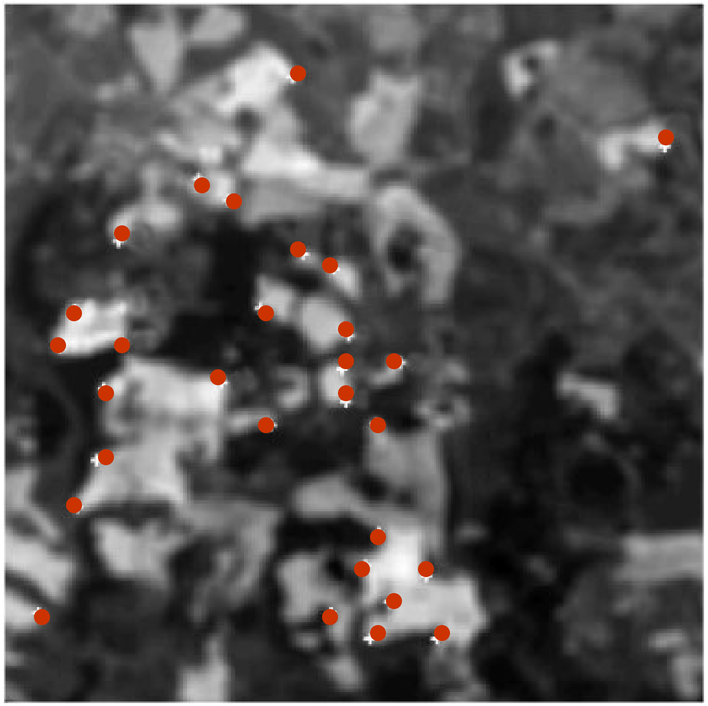
# CSS – Curvature Scale Space

Evolution of inflection points at different scales



# Local invariants of graylevel images



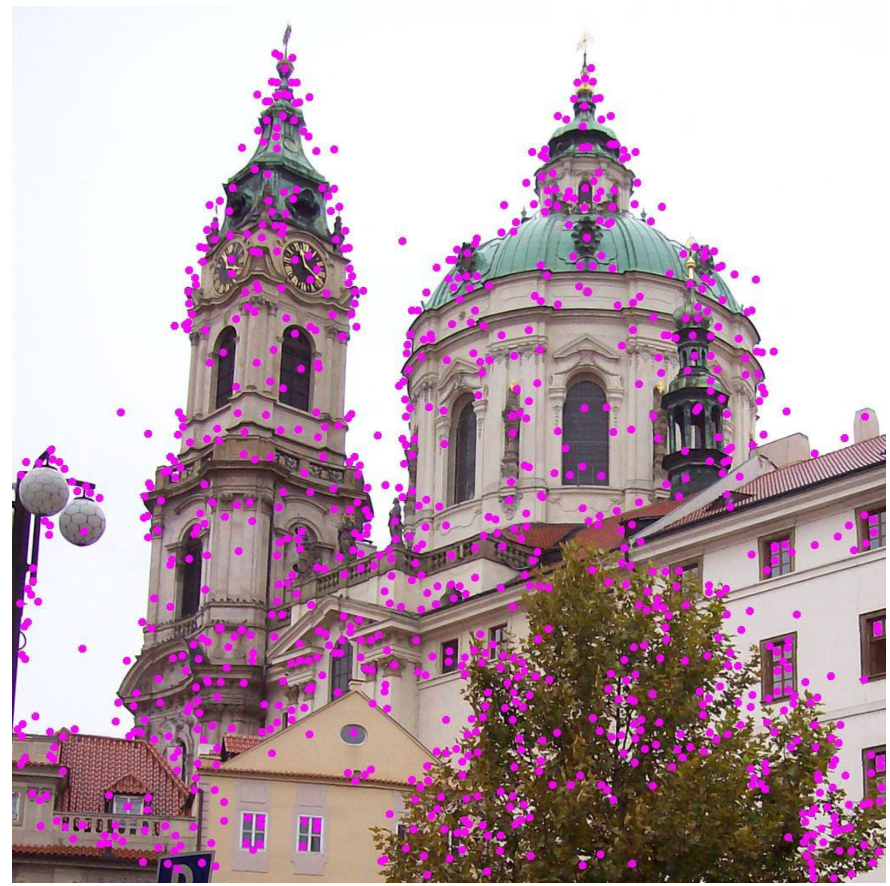


$$\min_{k,m} \text{distance}((v1_k, v2_k, v3_k, \ldots), \overline{(v1_m, v2_m, v3_m, \ldots)})$$

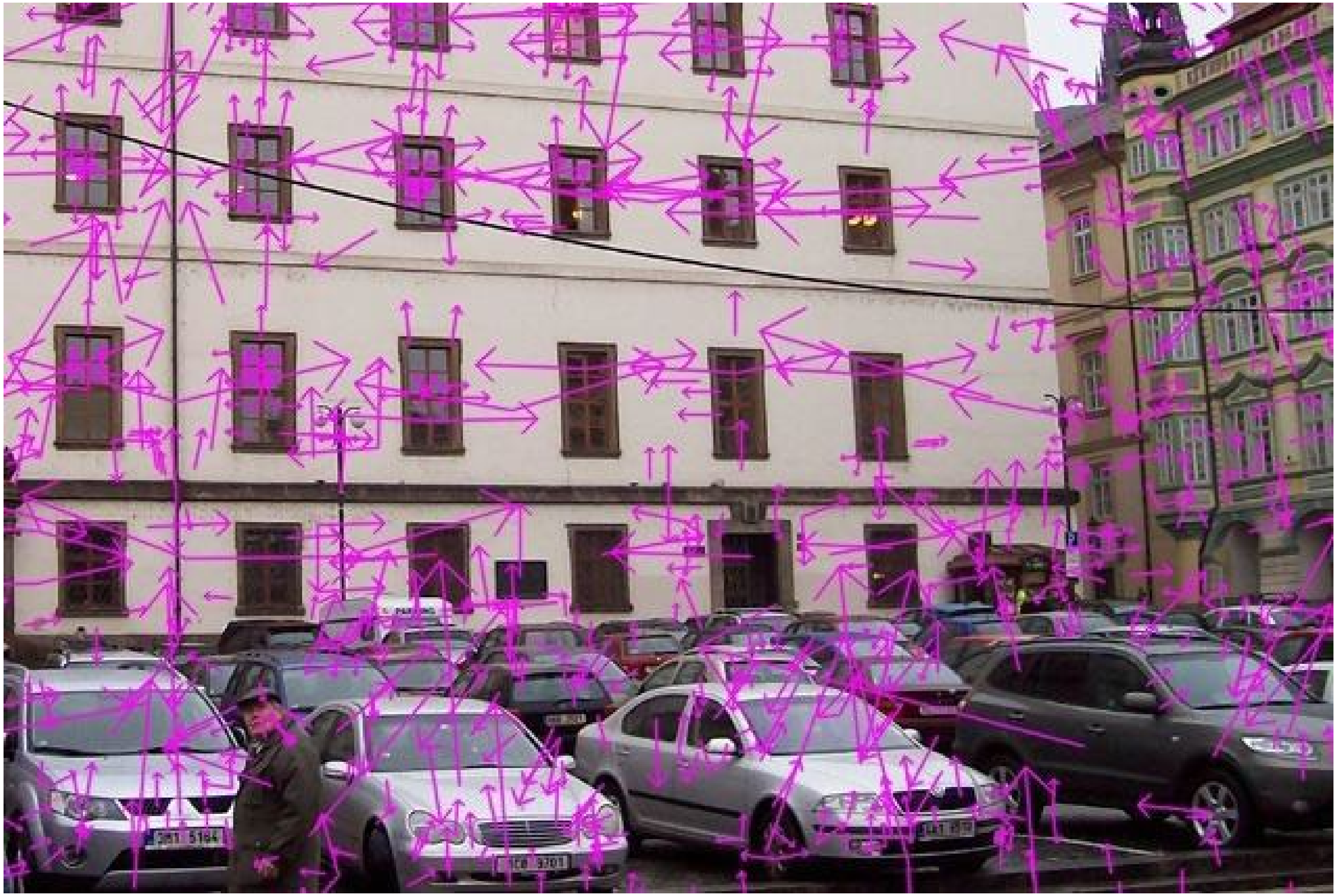
# SIFT (D.G. Lowe, 1999)

# Find local extrema in the difference image

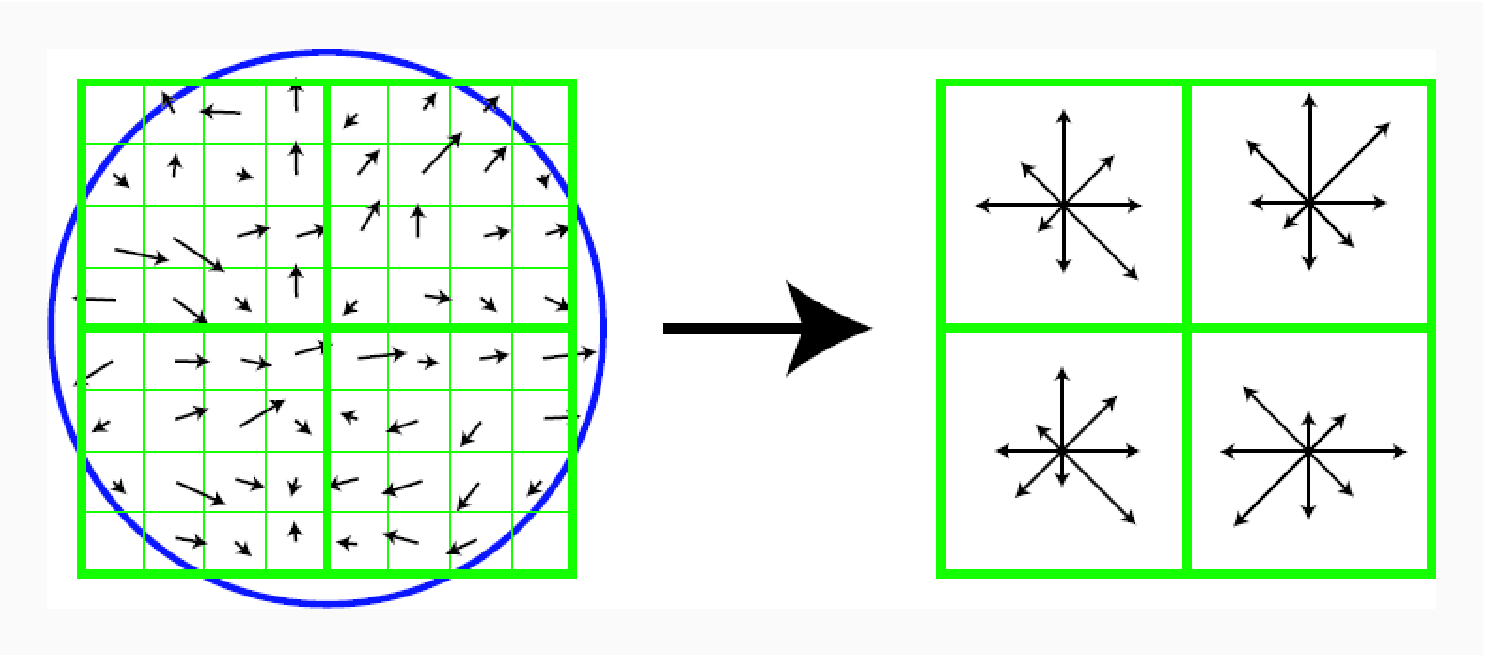
# Delete local extrema with low contrast and those on edges



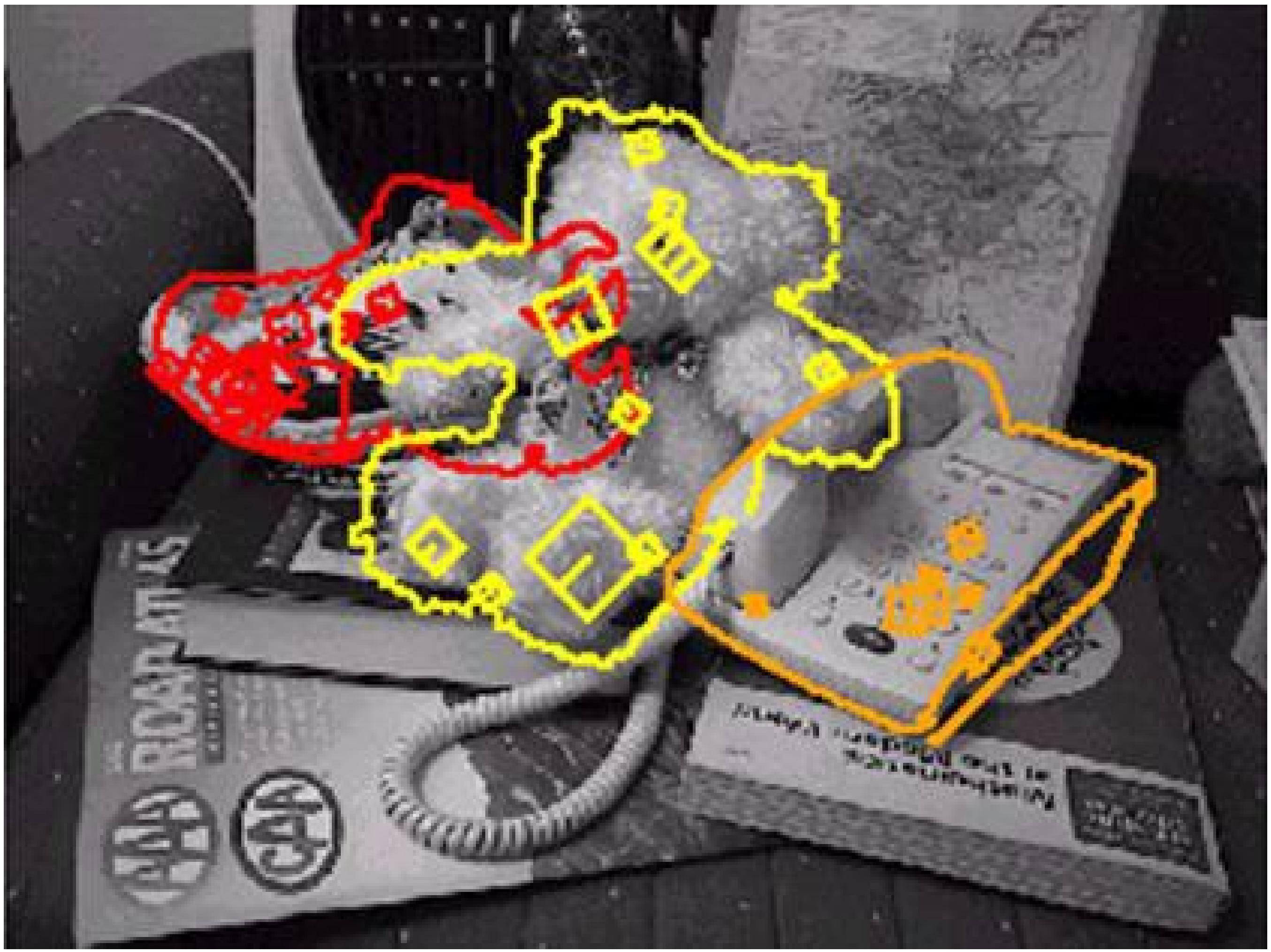
# Gradient-based neighborhood



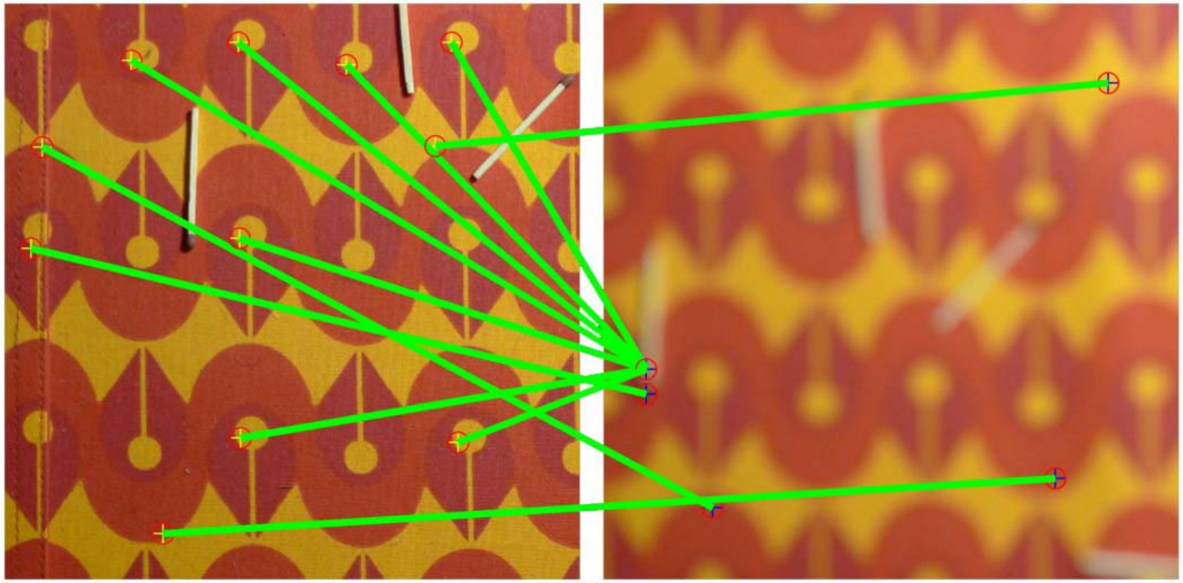
# Feature vector: histogram of gradients
# In the neighborhood

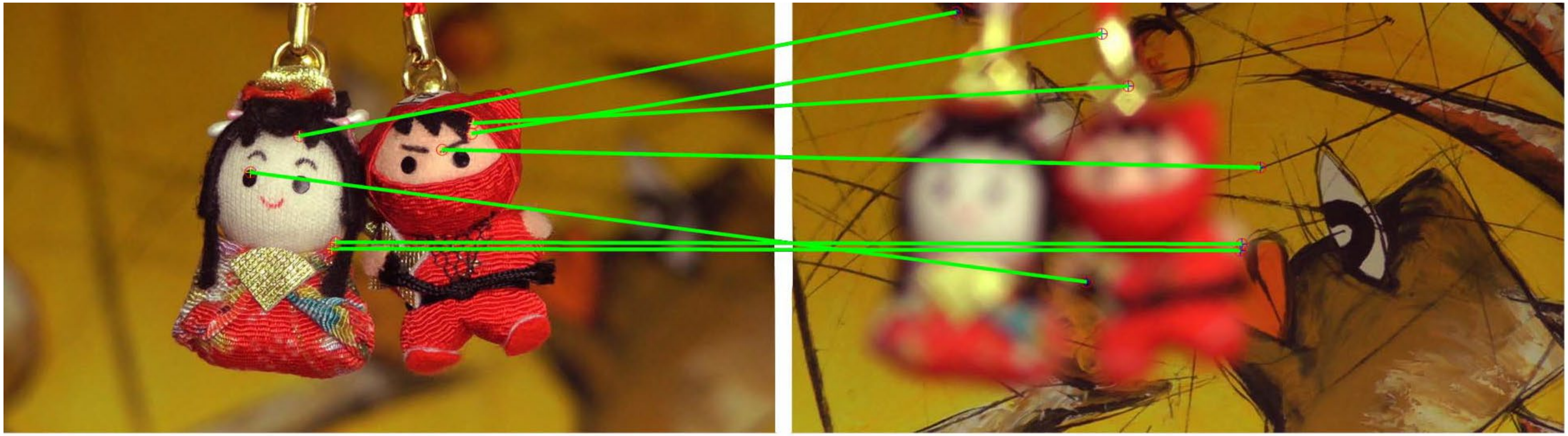


# Recognition of occluded objects



# Problems





# Moment invariants

**Moments are "projections" of the image function into a polynomial basis**

$f(x,y)$ – piecewise continuous image function defined on bounded $\Omega \subset \mathcal{R} \times \mathcal{R}$

$\{\mathcal{P}_{pq}(x,y)\}$ – set of polynomials defined on $\Omega$

$$M_{pq} = \int\int \mathcal{P}_{pq}(x,y) f(x,y) dx dy$$

# Common types of moments

## Geometric moments

$$m_{pq}^{(f)} = \int_{R^2} x^p y^q f(x,y) dx dy$$

## Central moments

$$\mu_{pq}^{(f)} = \int_{R^2} (x - x_t)^p (y - y_t)^q f(x,y) dx dy$$

# What are moment invariants?

Functions of moments, invariant to certain class of image degradations

- Rotation, translation, scaling
- Affine transform
- Convolution/blurring
- Combined invariants

# Invariants to translation and scaling

## Normalized central moments

$$\nu_{pq} = \frac{\mu_{pq}}{\mu_{00}^{w}} \qquad w = \frac{p+q}{2} + 1$$

# Invariants to rotation

$$\phi_1 = \mu_{20} + \mu_{02}$$

$$\phi_2 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2$$

$$\phi_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2$$

$$\phi_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2$$

# Thank you !

## Any questions ?