Snakes, level sets and graph-cuts
(Deformable models)

Centro de Visión por Computador,
Departament de Matemàtica Aplicada i Anàlisi,
Universitat de Barcelona

Snakes, level sets and graph-cuts (Deformable models)

Petia Radeva

Centro de Visión por Computador,

Departamento de Matemática Aplicada y Análisis,
Universidad de Barcelona

The problems of Medical image analysis vs. Computer Vision

Segmentation

Object recognition

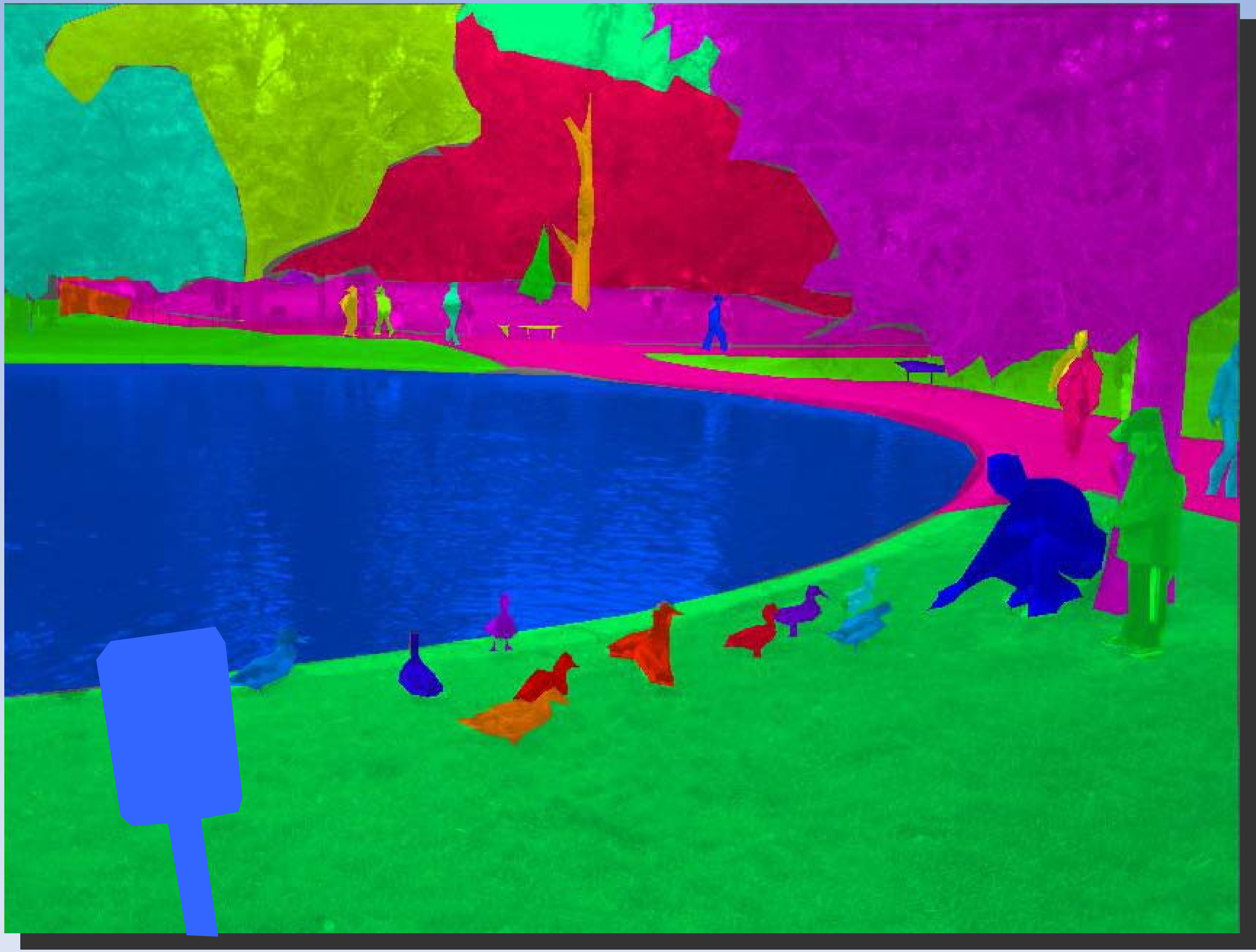
Atlas matching

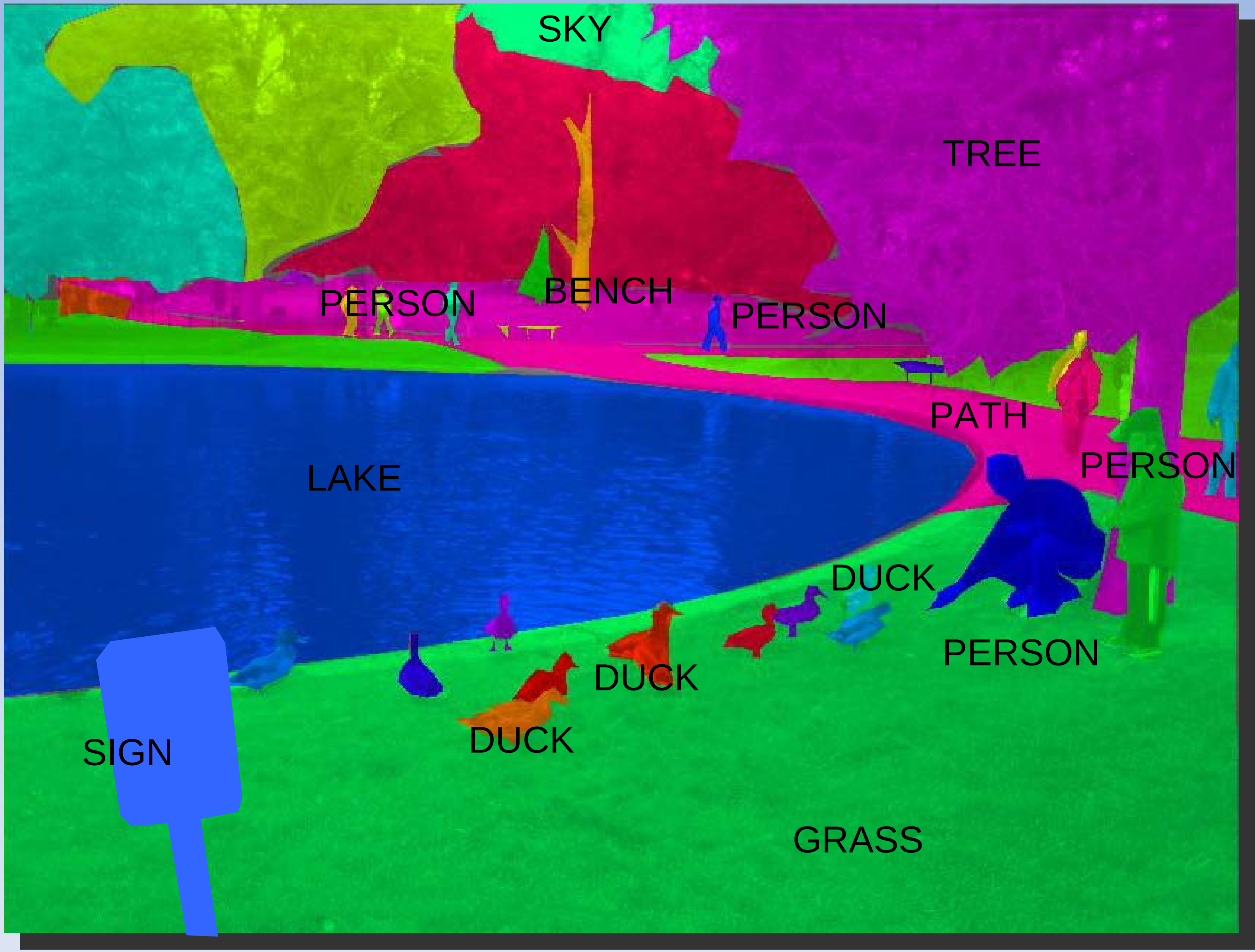
Registration

3D reconstruction

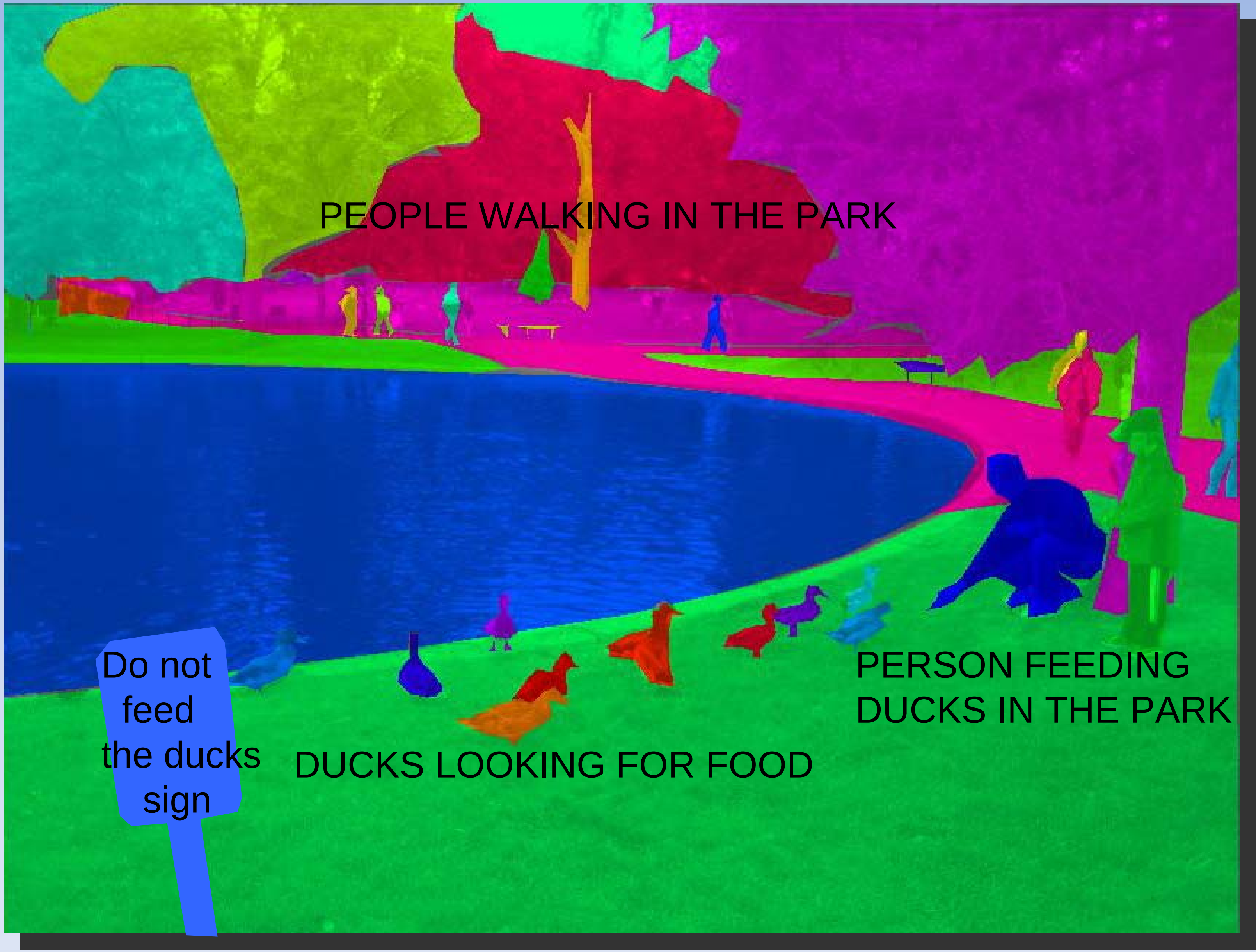
Deformation/Motion analysis

From tutorial de Antonio Torralba




SKY
TREE
PERSON
BENCH
PERSON
PATH
PERSON
LAKE
DUCK
DUCK
PERSON
DUCK
SIGN
GRASS


A VIEW OF A PARK ON A NICE SPRING DAY


PEOPLE WALKING IN THE PARK
Do not feed the ducks sign
DUCKS LOOKING FOR FOOD
PERSON FEEDING DUCKS IN THE PARK


PEOPLE UNDER THE
SHADOW OF THE TREES
PLEASE DO NOT FEED THE DUCKS
DUCKS ON TOP
OF THE GRASS

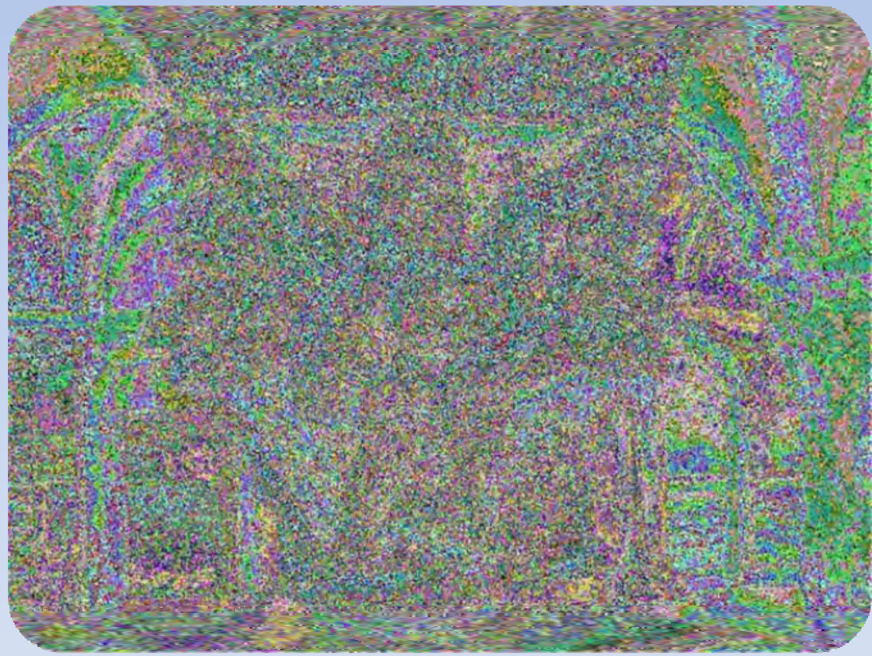
Image Understanding

What you see

+ context, prior information

It's not just issue of quantity!!

What the computer sees



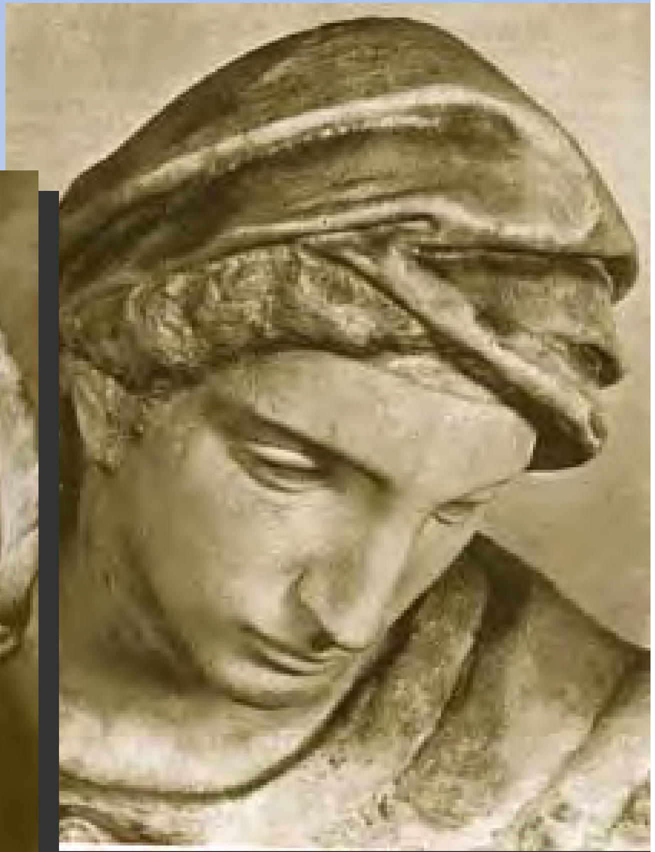
(same image in false colors)
Same information
Presented in a way your brain
is not trained to cope with.

Image Understanding

Automated Image Analysis is a **VERY** difficult problem

- “**Sensory Gap**”
  - The sensor doesn’t measure exactly what we want to observe (eg. 2D projection, noise, occlusions…)
- “**Semantic Gap**”
  - Images are low level collections of numbers…How to extract semantically significant, abstract descriptors? (eg: “a chair”, “forest”). Processes as complex as those happening in the brain are needed.

Challenges 1: view point variation (Sensory gap)



Michelangelo 1475-1564

Challenges 2: illumination

slide credit: S. Ullman

Challenges 3: occlusion

Magritte, 1957

Challenges 4: scale

Challenges 6: background clutter

Klimt, 1913

Challenges 5: deformation

Xu, Beihong 1943

Challenges 7: intra-class variation

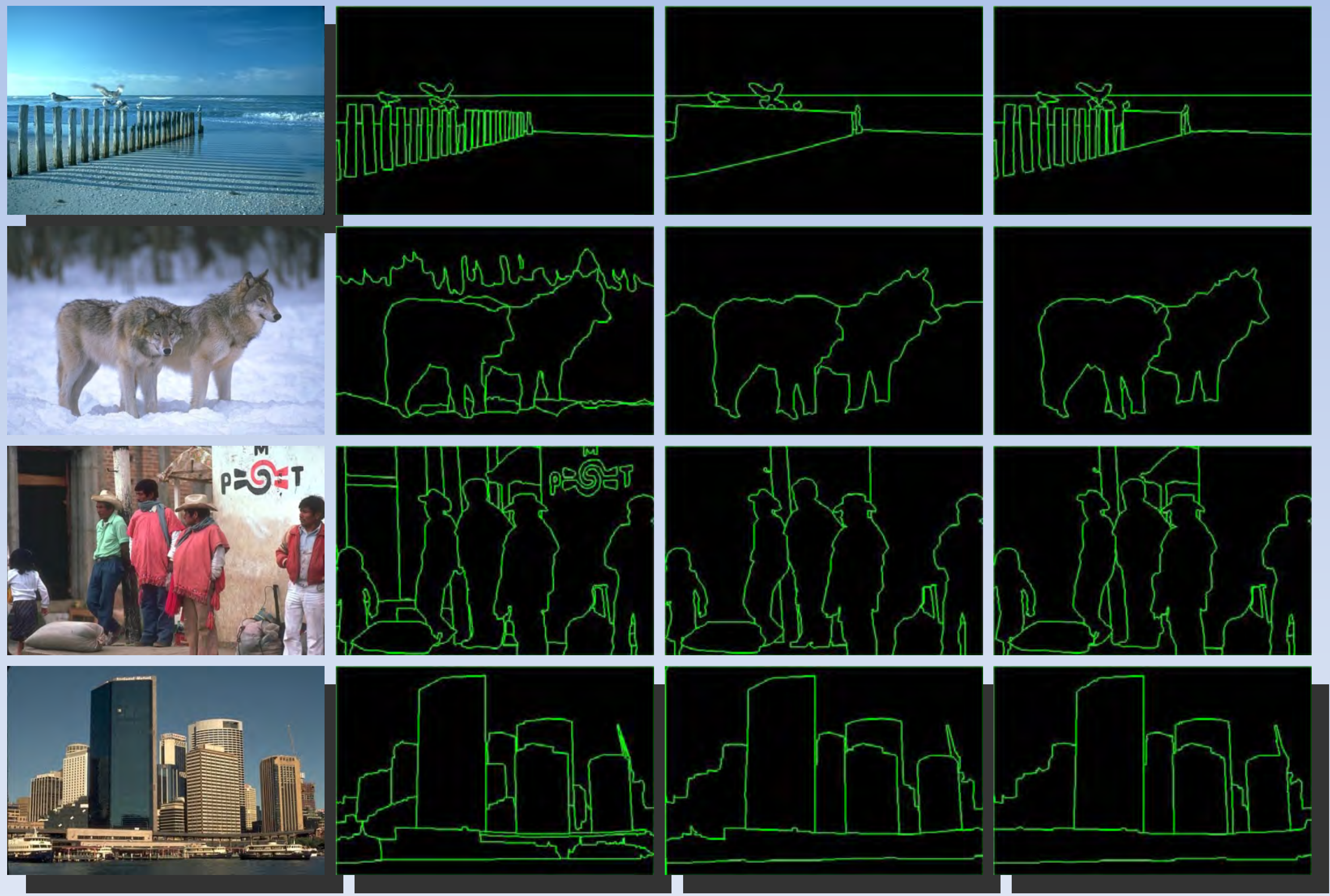





Some chairs

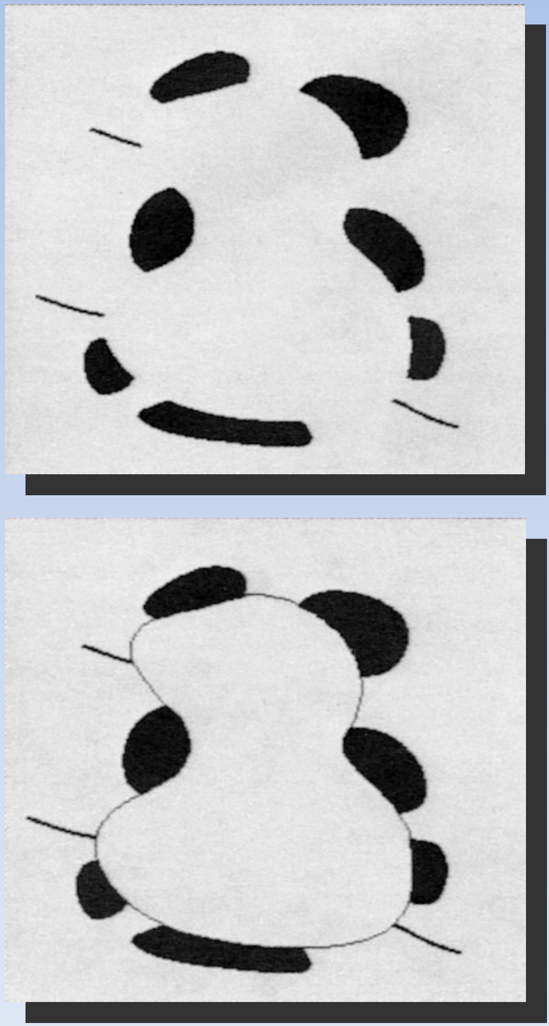
Related by function, not form

Segmentation is a subjective process (semantic gap)

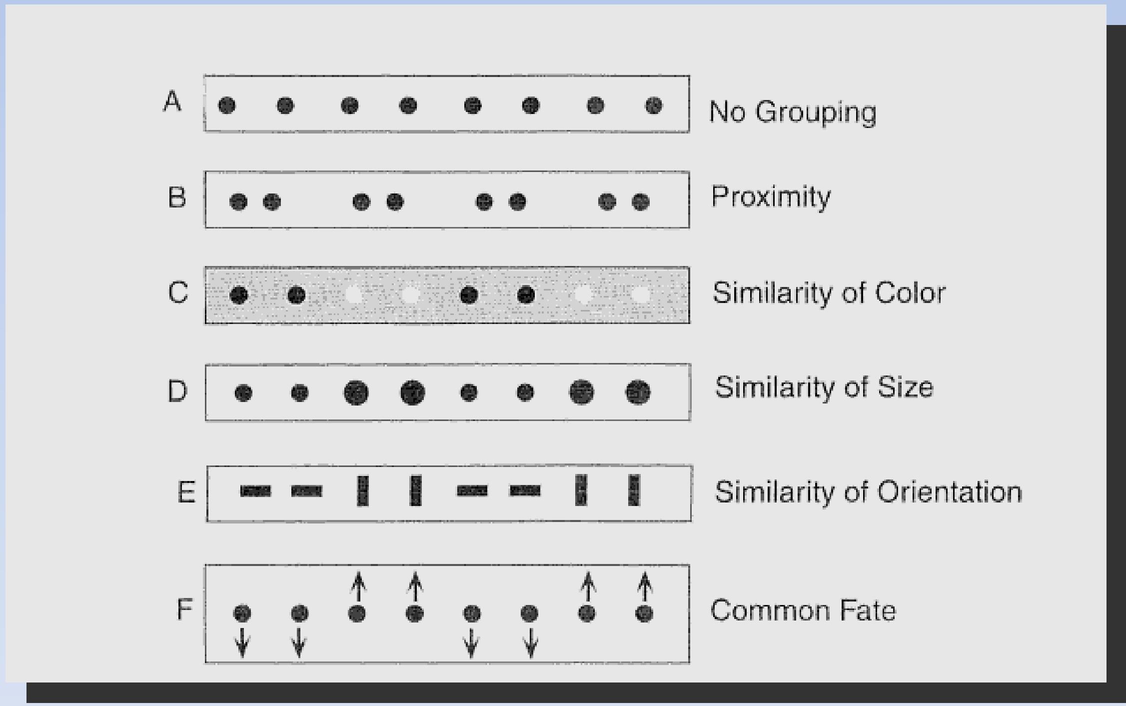


From tutorial Jitendra Malik

Subjective contours and free-form models

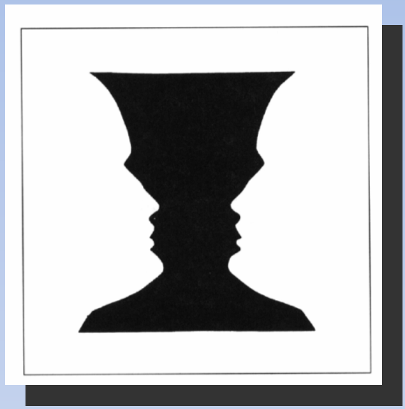


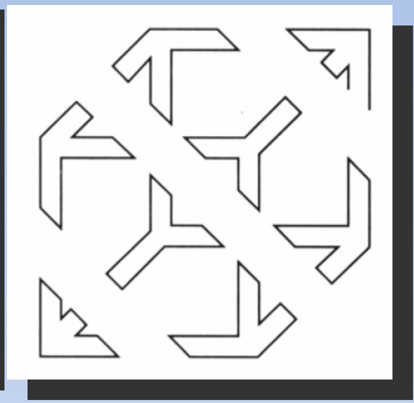
From Kass, Witkin & Terzopoulos

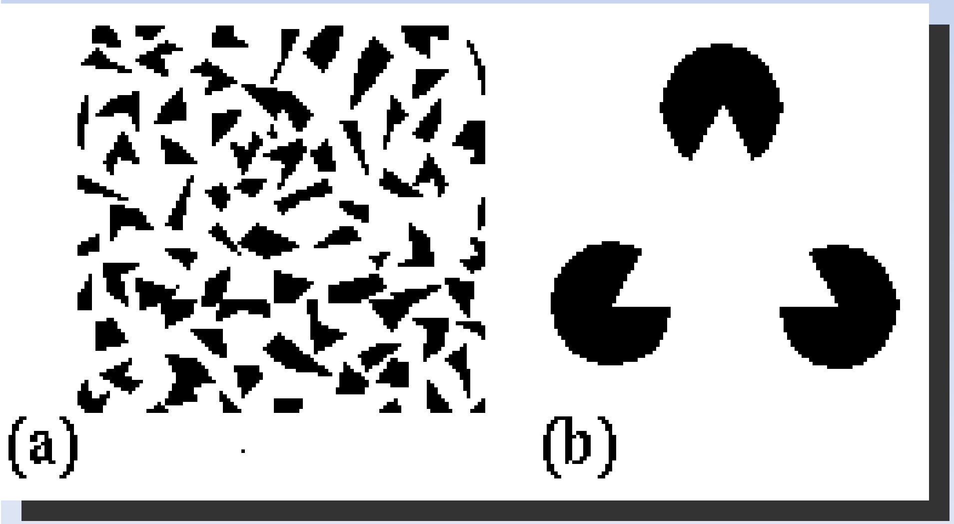


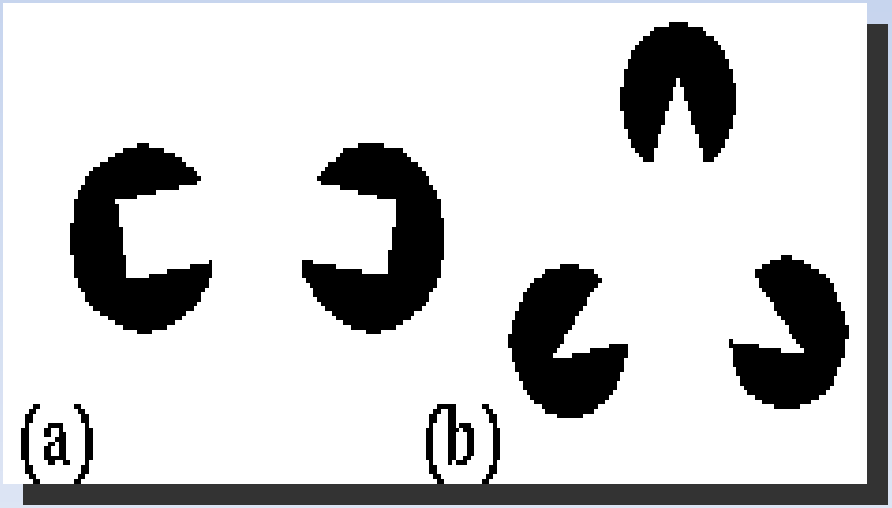
Grouping factors

Background vs. Foreground



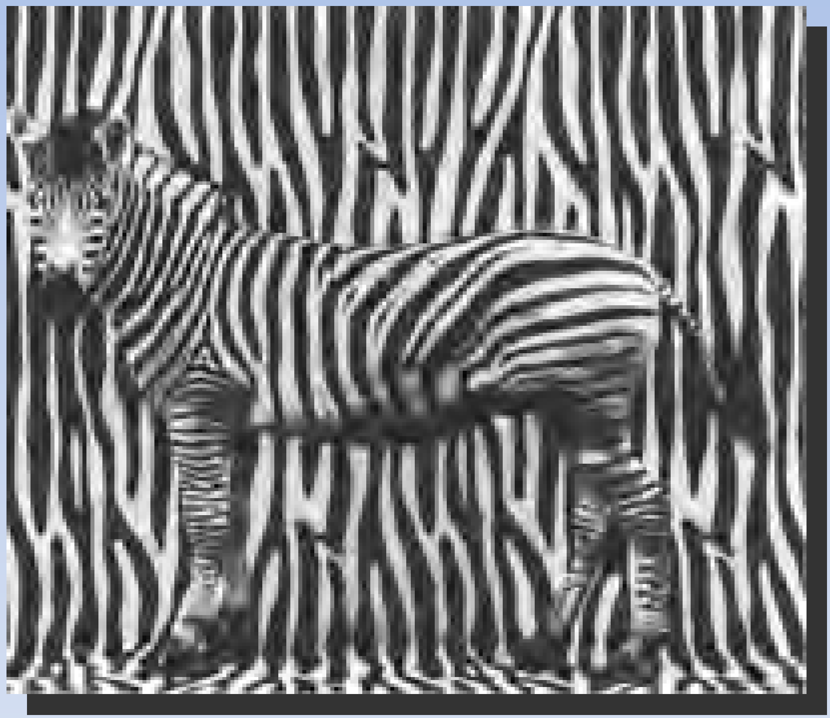




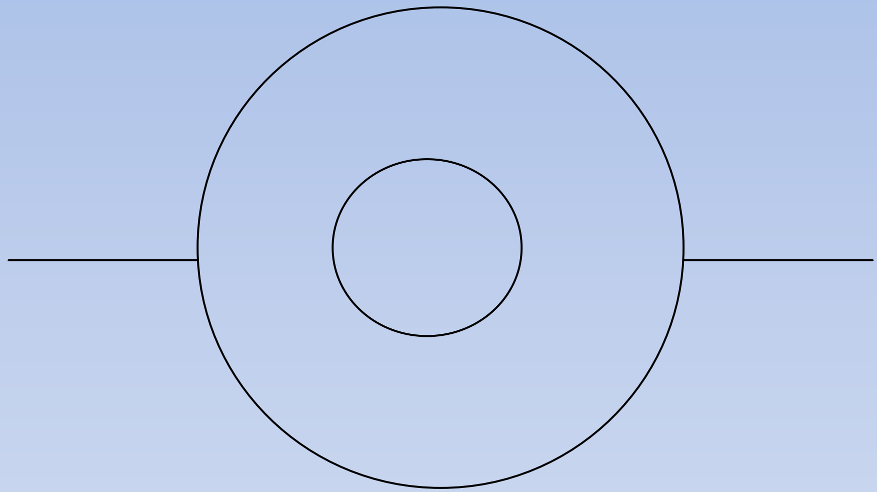


Spatial clustering

The segmentation challenge



Priming with prior knowledge (top-down or bottom-up image processing?!)

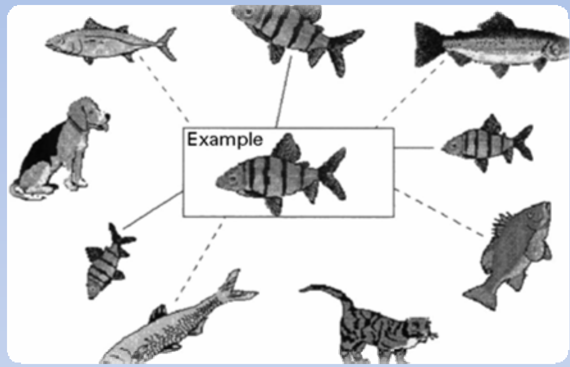


If you have never seen it before, this figure probably means little at first sight?!

Need of high-level knowledge to interpret images
real-time analysis needs selective processing
no need of considering the whole scene (less comp. load).

The problem of image segmentation

Models should allow for the expected variations in size, shape and appearance of the structure



Usually, models are hand-crafted or too general.

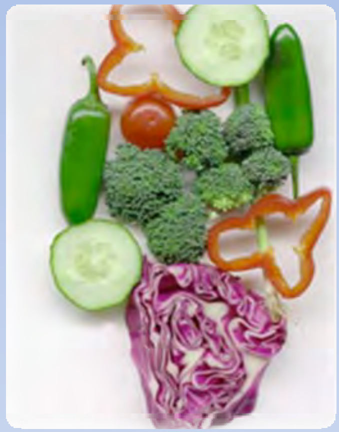
Aim: Statistically based technique for building compact models of the shape and appearance of flexible objects

How to introduce high-level knowledge to regularize the segmentation problem?

- Similar pixels properties
- General high-level constraints
  - location of images
  - boundary smoothness, etc.
- Model-guided segmentation and recognition

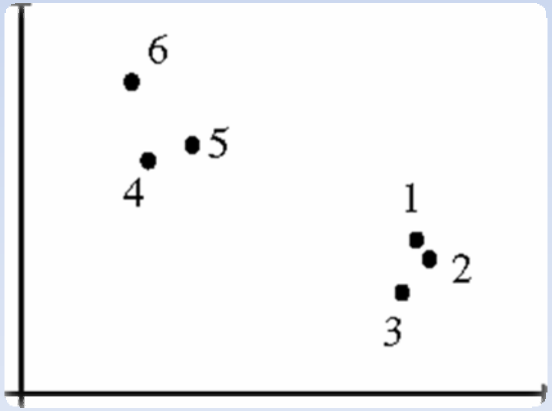
Segmentation as a clustering problem

Clustering (píxels, elements, etc.) with the same properties

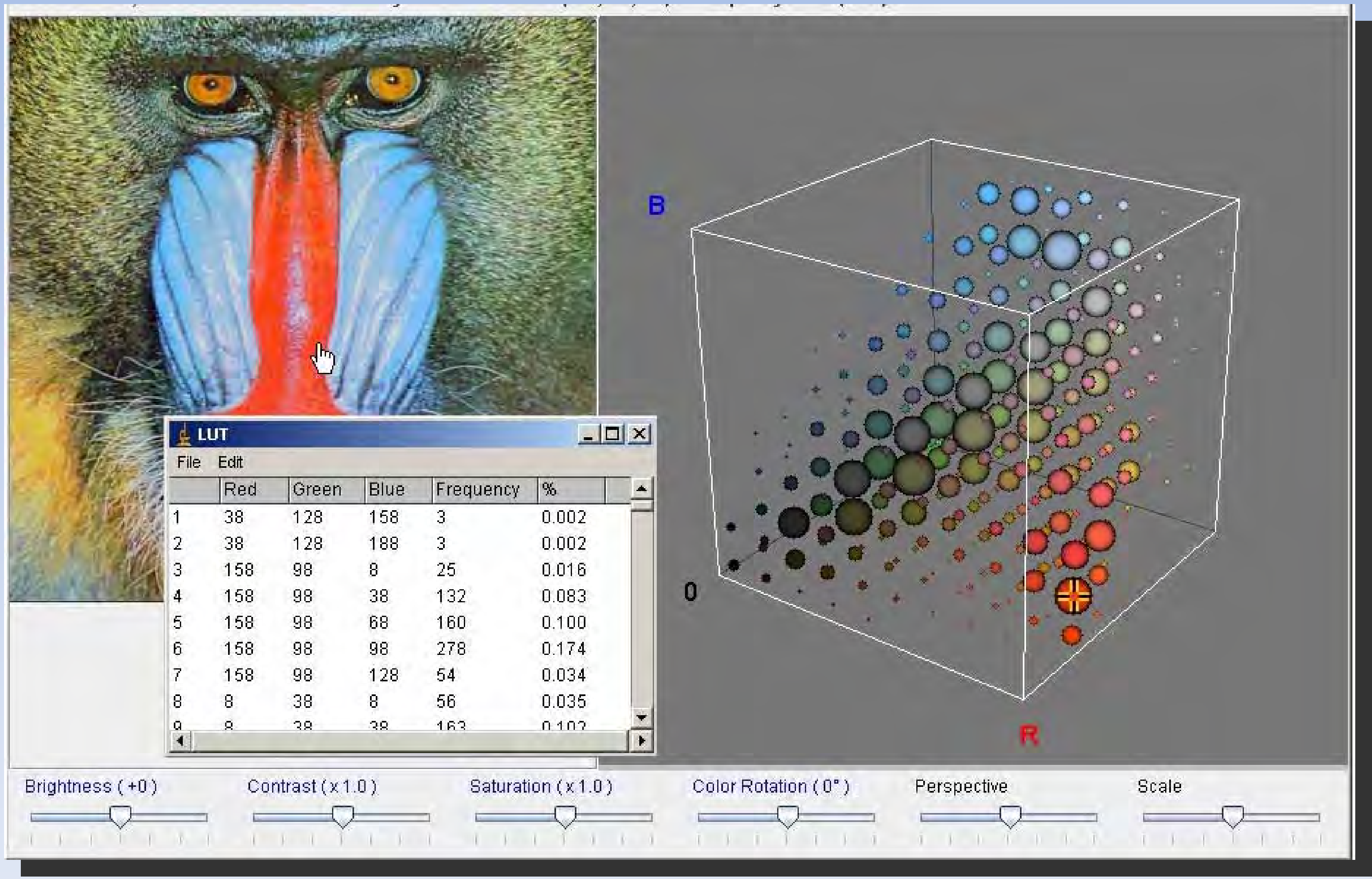


- “Agglomerative clustering”

- “Divisive clustering”

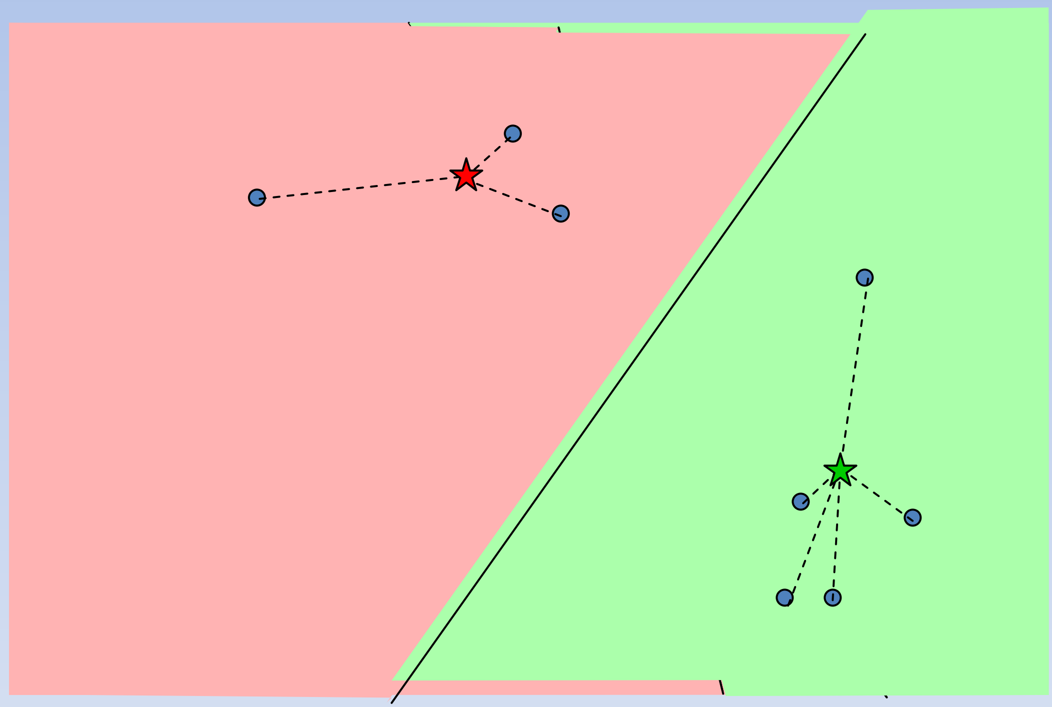


Histogram 3D



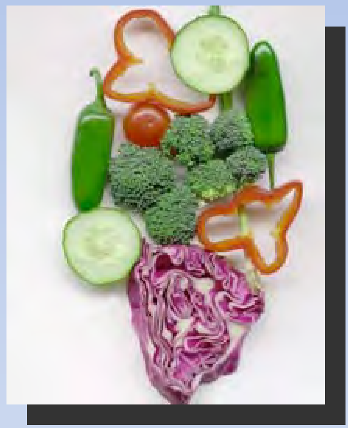
K-Means

- Algorithm
  - Fix cluster centres;
  - Assign points to the most similar clusters
  - Recalculate clusters centres
- x can be any feature as long as features distance can be estimated.



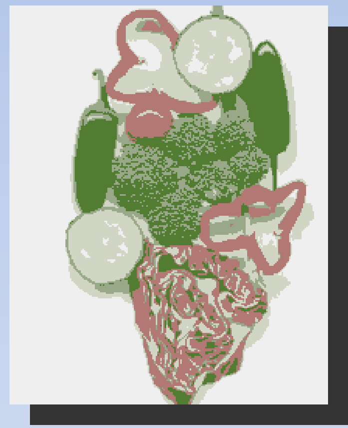
Results of the clusterization by K-Means

Image



Clusters based on intensity

Clusters based on colour



Example



http://www.ece.neu.edu/groups/rpl/kmeans/

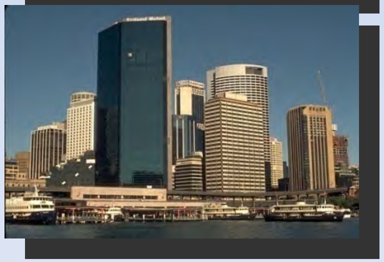
How to introduce high-level knowledge to regularize the segmentation problem?

- Similar pixels properties

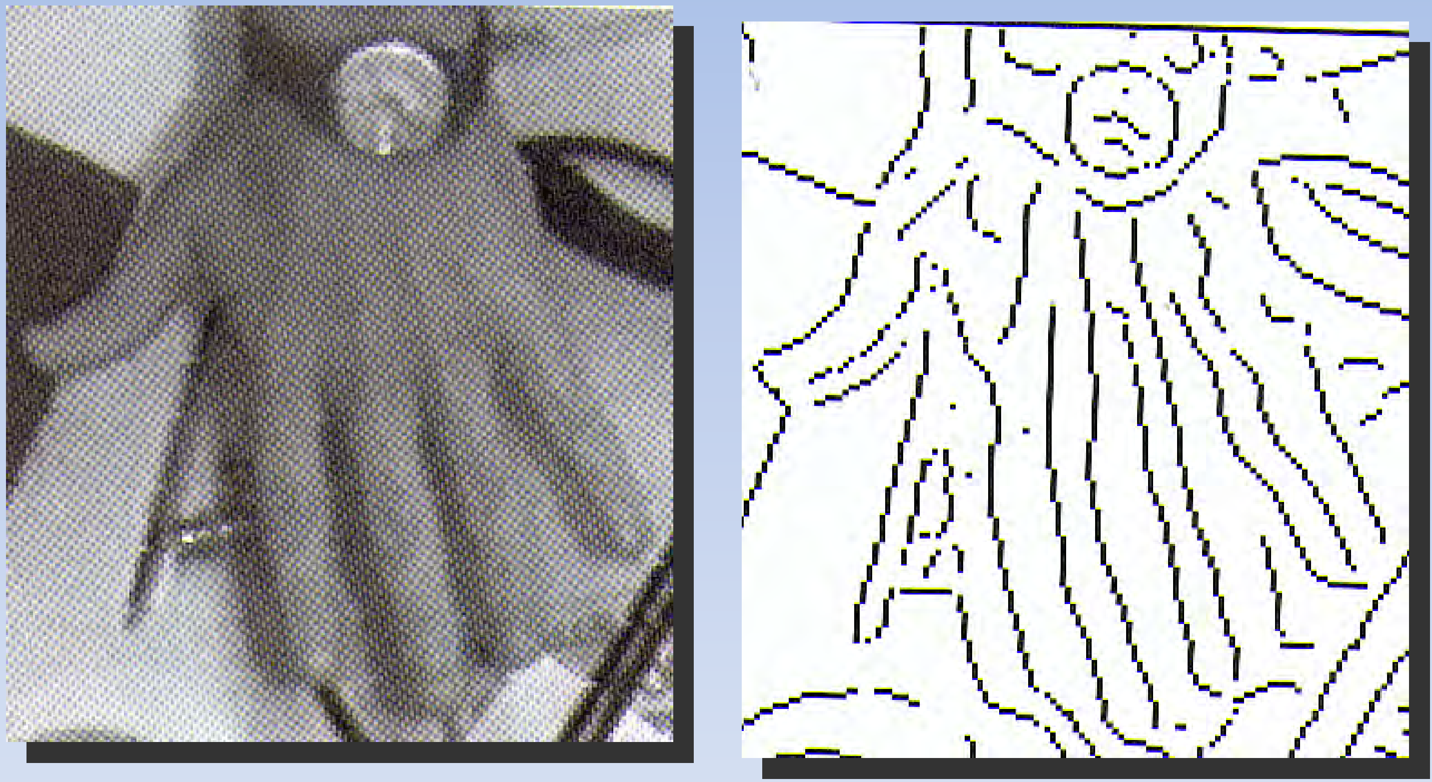
- General high-level constraints
  - Boundary smoothness.
  - Physics-based models, etc.

- Model-guided segmentation and recognition



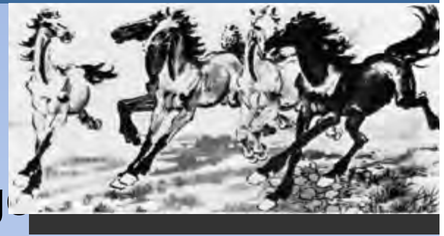
Why snakes



Edge map is ambiguous to be interpreted

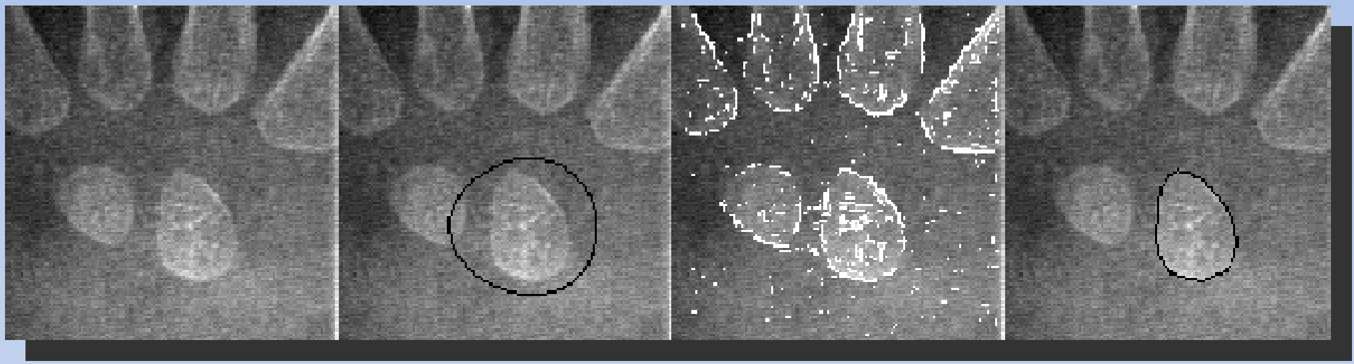
From A. Blake

Motivation



- Challenge – locate and recognize different objects in an image
- How to integrate and interpret the diverse local image cues (intensity, gradient, texture, etc.)
- Bottom-up or top-down approach?!
- Geometrical shape information – local and generic to global and specific (smoothness, elasticity, hand-crafted shapes)
- “There are no 2 leaves of the same shape” – intrinsic intraclass variation
- Object deformation – varying imaging conditions, sensor noise, occlusion, imperfect segmentation
- Can we come up with a versatil and flexible approach for object modeling and representation to deal with a variety of shape deformationbs and variations while maintaining a certain structure?!

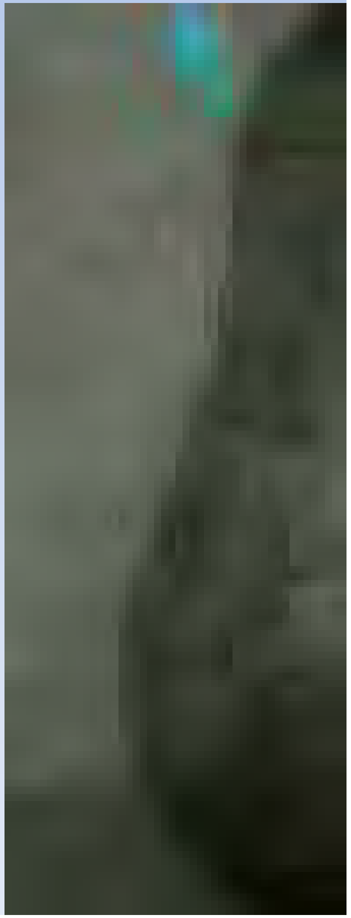
What is a snake in Computer Vision?!



- Snake - elastic continuous curve that from an initial position begins to deform to adjust the object's contour.

- External forces attract the snake towards image features.

- Internal forces avoid discontinuities in the snake shape.



Snake representation

- A snake is an elastic curve $u(s) = (x(s), y(s))$ defined by means of:

a discrete representation

- - point-based snake - elastic curve as a sequence of snaxels :

$$Q(u) = \{x_i(u), y_i(u)\},\ i = 0, \ldots, N \qquad \Omega = [0, 1]$$

a continuous representation

- - a tesselation established over the parametrization set
- - decomposition of the curve in a basis of functions (usually piecewise polynomials)
- - small support of the basis functions

Energy-Minimizing Curve

- Snake - an elastic curve with associated energy:

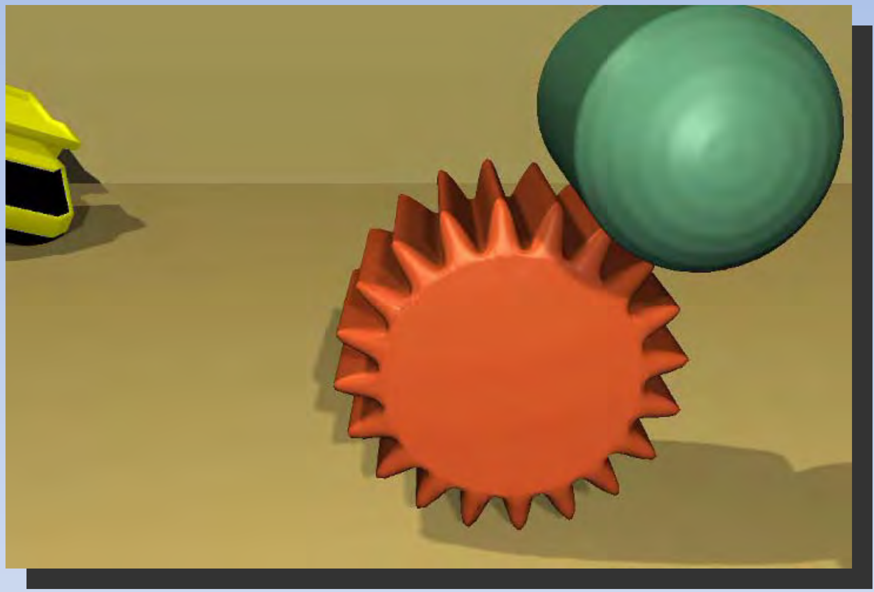
$$E_{snake} = \int_0^1 E_{int}(u(s)) + E_{ext}(u(s))\, ds.$$

- Potential - a surface P (x, y) with valleys corresponding to image features

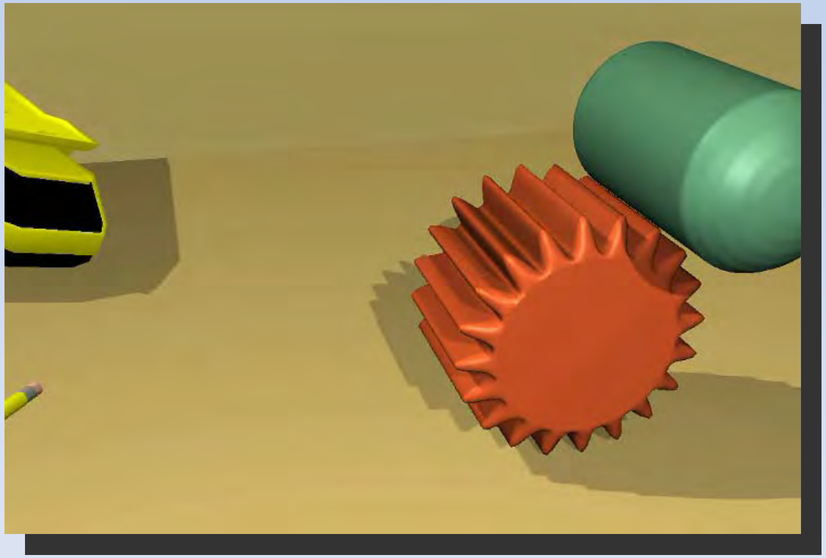
- External (image) forces attract the snake to the potential valleys:

$$E_{ext}(u(s)) = P(x, y)$$

Deformable models are physics-based models



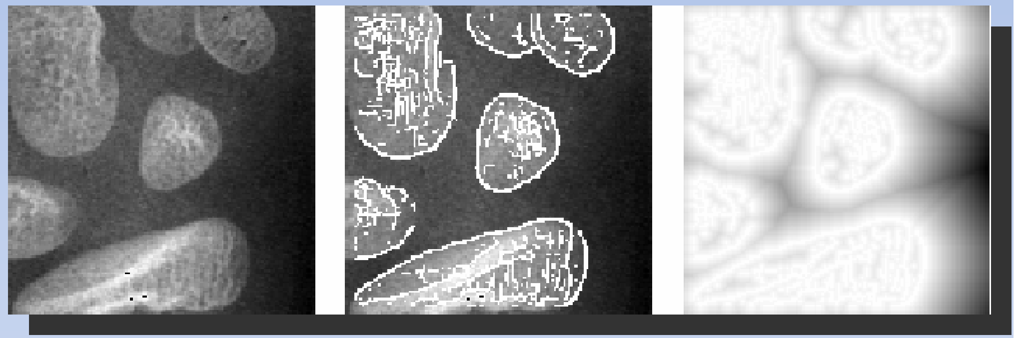
Model the objects as physics-based ones



Energy-Minimizing Curve

External (image) forces attract the snake to the potential valleys

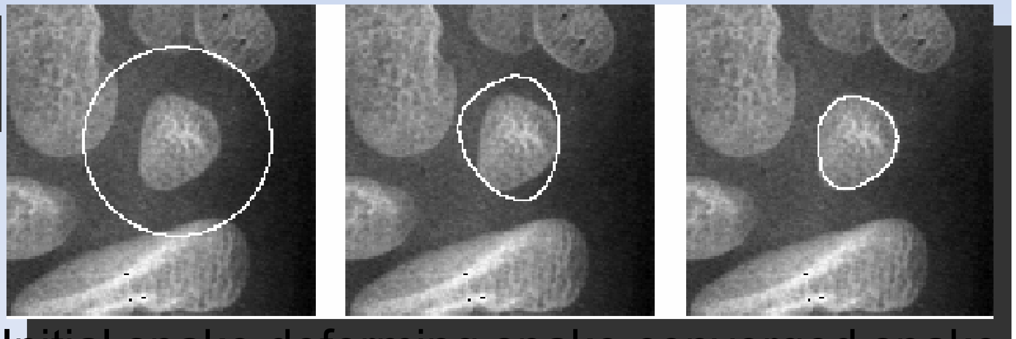
$$E_{ext}(u(s)) = P(x, y)$$



Original image image features potential field

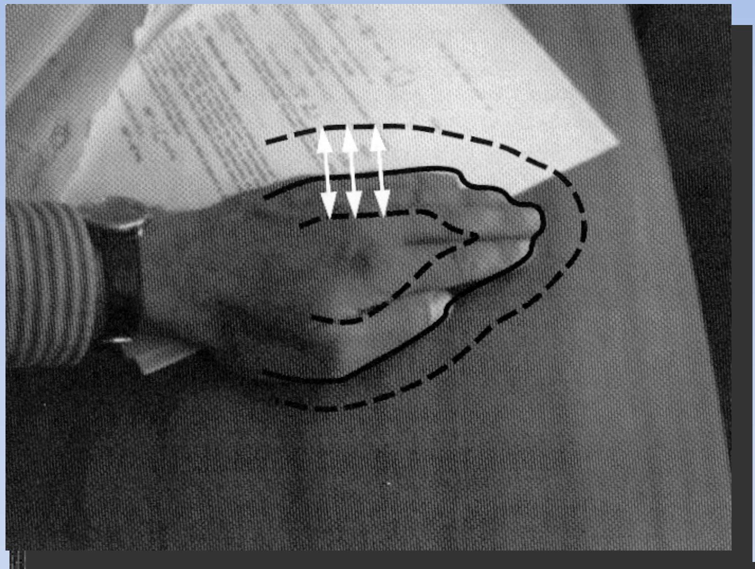
Internal forces penalize stretching and bending:

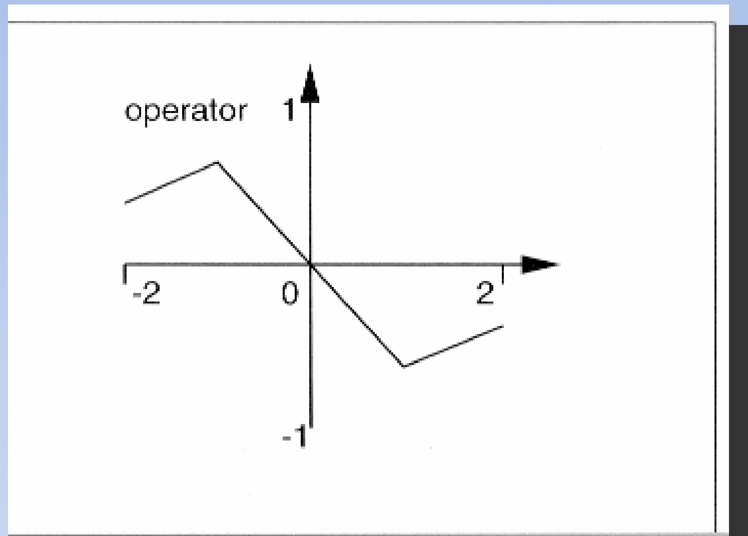
$$E_{int}(u(s)) = \alpha E_{membrane} + \beta E_{thin-plate}$$

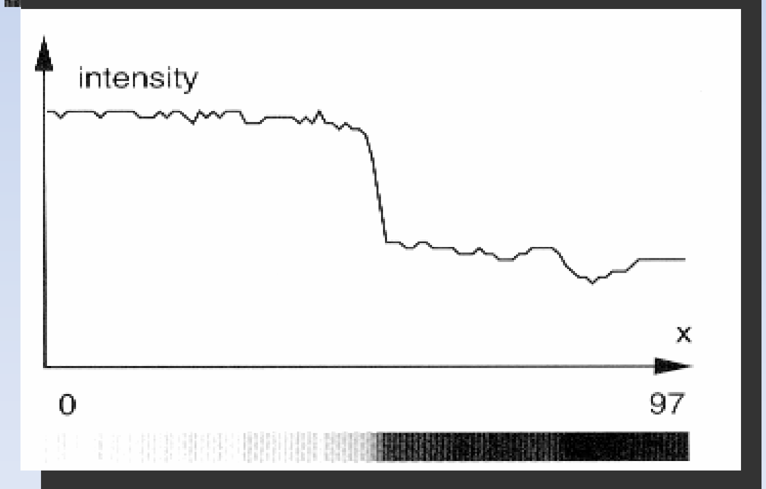


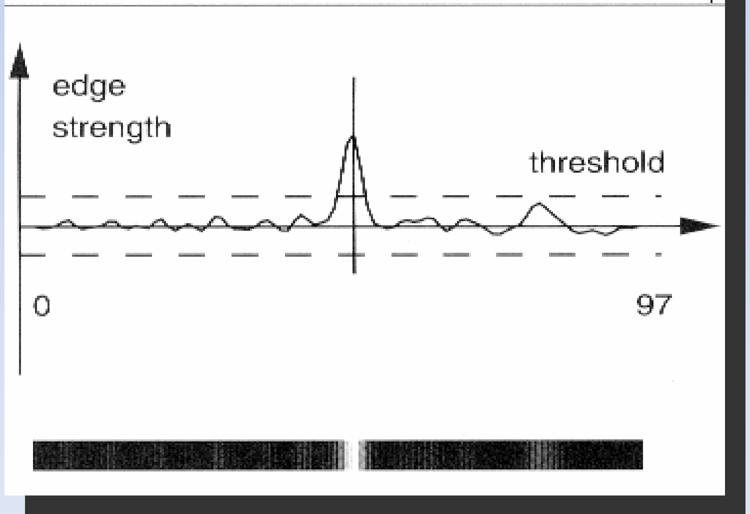
Initial snake deforming snake converged snake

Image feature extraction - edge detection



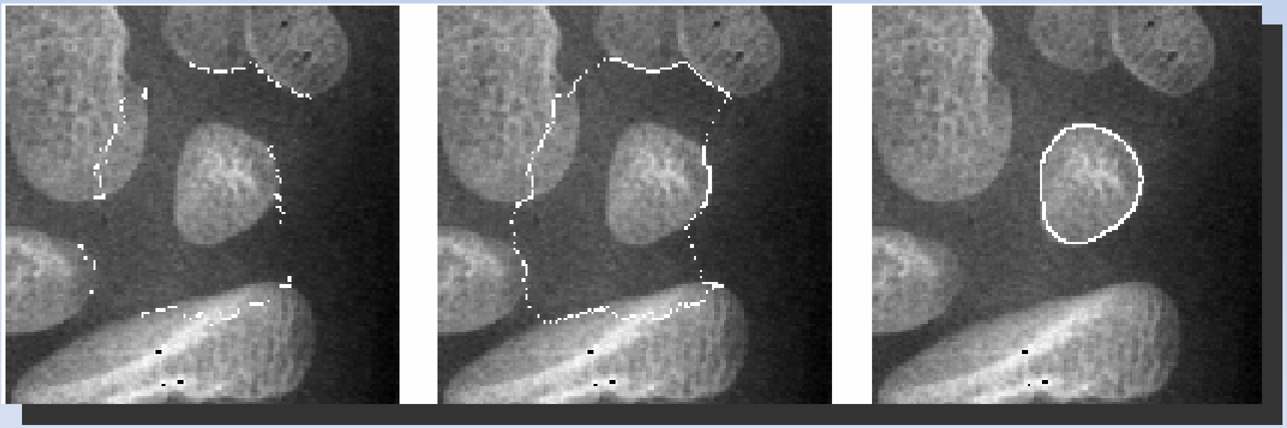






Internal Energy of the Snake

- membrane energy given by the first derivative of the curve that avoids the stretching and discontinuity of the curve: $E_{membrane}(Q) = Q_u{}^2(u)$
- thin-plate energy given by the second derivative of the curve that avoids the bending of the curve: $E_{thin-plate}(Q) = Q_{uu}{}^2(u)$



Snake deformations with different elastic properties

Segmentation by snakes - an energy-minimization procedure.

Energy Minimization Procedure for a Point-Based Snake

$$E_{snake} = \int_0^1 E_{int}(u(s)) + E_{ext}(u(s))\, ds.$$

An energy minimum satisfies the Euler-Lagrange equation:

$$\begin{cases} -\frac{d}{du}(\alpha(u)Q_u) + \frac{d^2}{du^2}(\beta(u)Q_{uu}(u)) + \nabla E_{ext}(Q(u)) = 0 \\ + \text{ boundary conditions.} \end{cases}$$

The equation can be decoupled wrt both spatial parameters:

$$\begin{cases} -\frac{d}{du}(\alpha(u)x_u) + \frac{d^2}{du^2}(\beta(u)x_{uu}) + \frac{d}{dx}E_{ext}(Q(x,y)) = 0 \\ -\frac{d}{du}(\alpha(u)y_u) + \frac{d^2}{du^2}(\beta(u)y_{uu}) + \frac{d}{dy}E_{ext}(Q(x,y)) = 0 \\ + \text{ boundary conditions.} \end{cases}$$

Energy-Minimization Procedure

Discretizing by the Finite Difference Method (DFM):

$$\begin{cases} \alpha_i(x_i - x_{i-1}) - \alpha_{i+1}(x_{i+1} - x_i) + \\ \beta_{i+1}(x_{i+2} - 2x_{i+1} + x_i) - 2\beta_i(x_{i+1} - 2x_i + x_{i-1}) + \\ \qquad \beta_{i-1}(x_i - 2x_{i-1} + x_{i-2}) + \\ \qquad \frac{d}{dx}E_{ext}(u(x_i, y_i)) = 0 \\ \alpha_i(y_i - y_{i-1}) - \alpha_{i+1}(y_{i+1} - y_i) + \\ \beta_{i+1}(y_{i+2} - 2y_{i+1} + y_i) - 2\beta_i(y_{i+1} - 2y_i + y_{i-1}) + \\ \qquad \beta_{i-1}(y_i - 2y_{i-1} + y_{i-2}) + \\ \qquad \frac{d}{dy}E_{ext}(u(x_i, y_i)) = 0 \end{cases}$$

In matrix form we get:

$$\begin{cases} Ax + \frac{d}{dx}E_{ext}(x, y) = 0 \\ Ay + \frac{d}{dy}E_{ext}(x, y) = 0 \end{cases}$$

Stiffness matrix A

$$A = \begin{pmatrix} c_1 & d_1 & e_1 & 0 & 0 & a_1 & b_1 \\ b_2 & c_2 & d_2 & e_2 & 0 & 0 & a_2 \\ a_3 & b_3 & c_3 & d_3 & e_3 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ e_{N-1} & 0 & 0 & a_{N-1} & b_{N-1} & c_{N-1} & d_{N-1} \\ d_N & e_N & 0 & 0 & a_N & b_N & c_N \end{pmatrix}$$

where

$$\begin{aligned} a_i &= \beta_{i-1} \\ b_i &= -2\beta_i - 2\beta_{i-1} - \alpha_i \\ c_i &= \beta_{i+1} + 4\beta_i + \beta_{i-1} + \alpha_{i+1} + \alpha_i \\ d_i &= -2\beta_{i+1} - 2\beta_i - \alpha_{i+1} \\ e_i &= \beta_{i+1} \end{aligned}$$

Introducing in the system of kind $Ax = b$ a snake energy dissipation functional:

$$\begin{cases} -\gamma(x_t - x_{t-1}) = Ax_t + \frac{d}{dx}E_{ext}(x_{t-1}, y_{t-1}) \\ -\gamma(y_t - y_{t-1}) = Ay_t + \frac{d}{dy}E_{ext}(x_{t-1}, y_{t-1}) \end{cases}$$

γ - *damping parameter*, determining the rate of convergence of the minimization process

Snake Energy-Minimization Procedure

- Iterative procedure for snake energy minimization:

$$\begin{cases} x_t = (A+\gamma I)^{-1}(\gamma x_{t-1} + F_{ext}(x_{t-1}, y_{t-1})) \\ y_t = (A+\gamma I)^{-1}(\gamma y_{t-1} + F_{ext}(x_{t-1}, y_{t-1})) \end{cases}$$

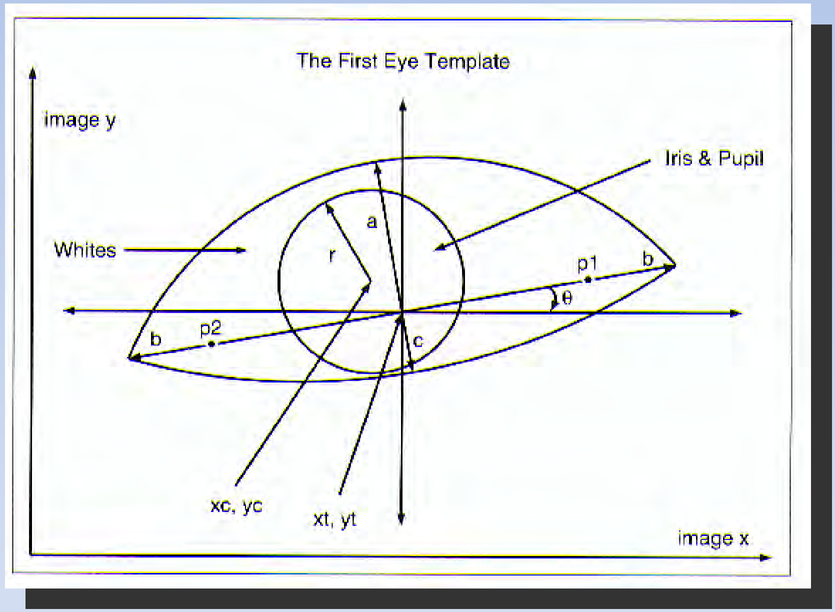
where $F_{ext} := -\nabla E_{ext}(Q(x,y))$ - external forces

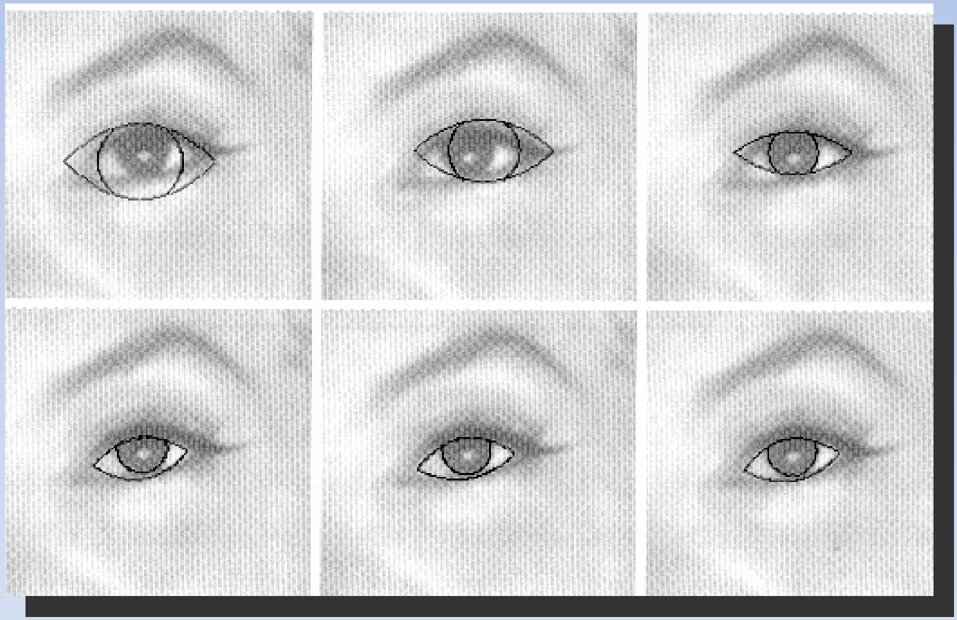
The snake deformation is a composition of:

- snake attraction by external forces
- snake smoothing due to internal forces.

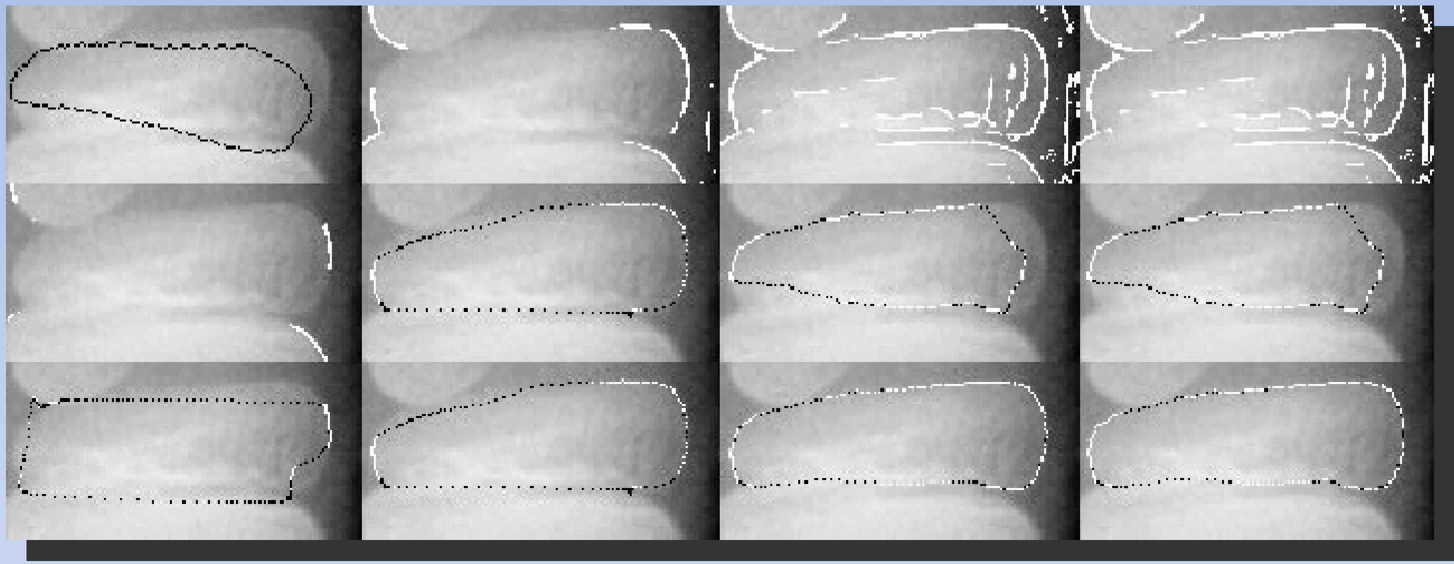
Given α, β and γ positive $\Rightarrow$ $(A+\gamma I)$ - constant, banded, positive definite $\Rightarrow$ **the snake converges!**

Flexible templates (Yuille et al.)





Snake in a MultiScale Deformation Scheme

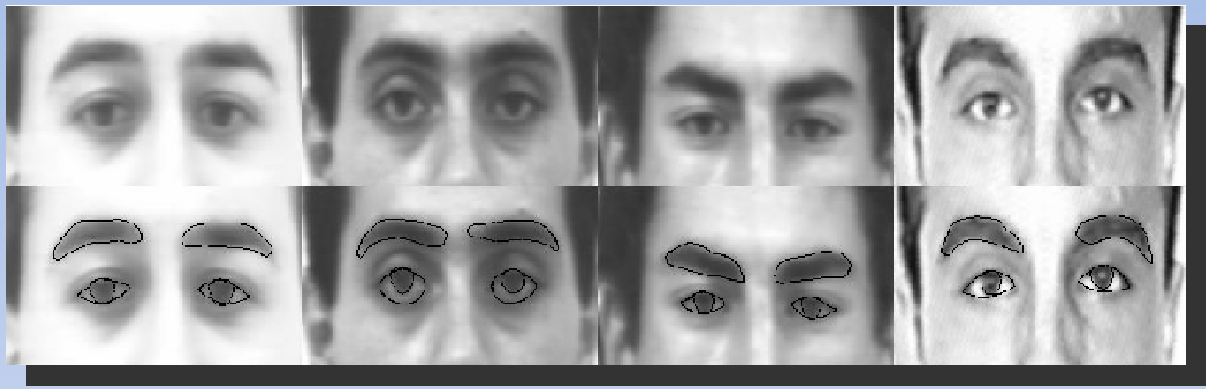


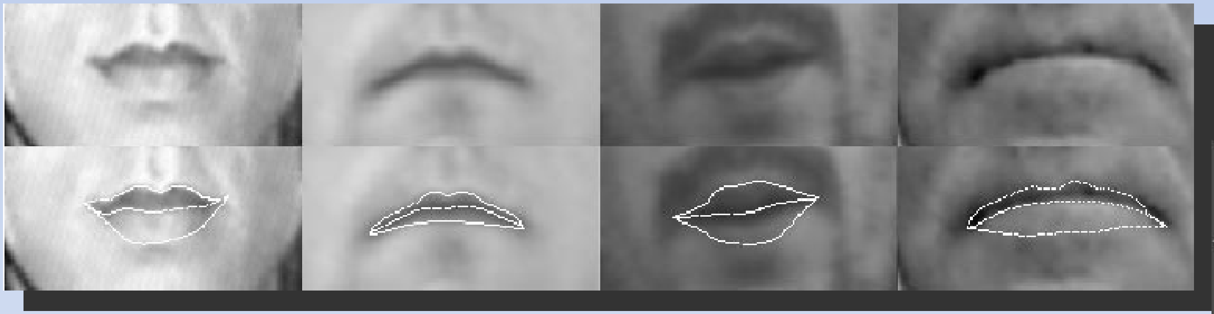
Different behaviour of a snake without (II row) and with (III row) MSDS application

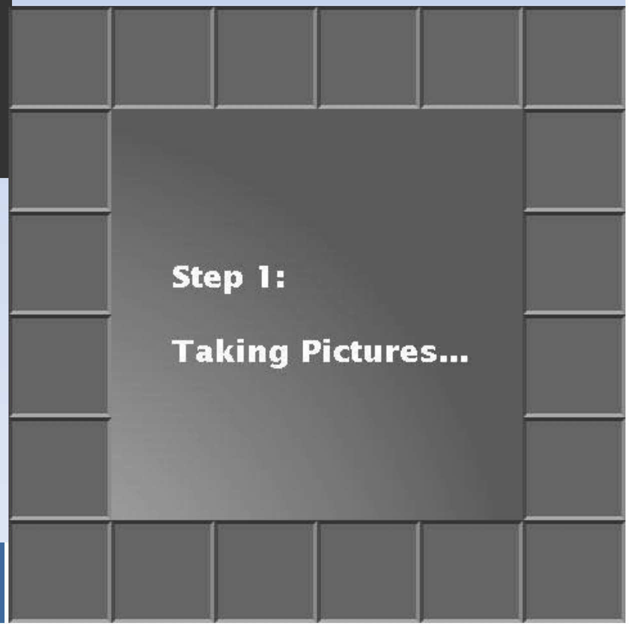
Properties:

- Priority to strong image features (local minima of the snake energy),
- Less dependence of snake initialization.

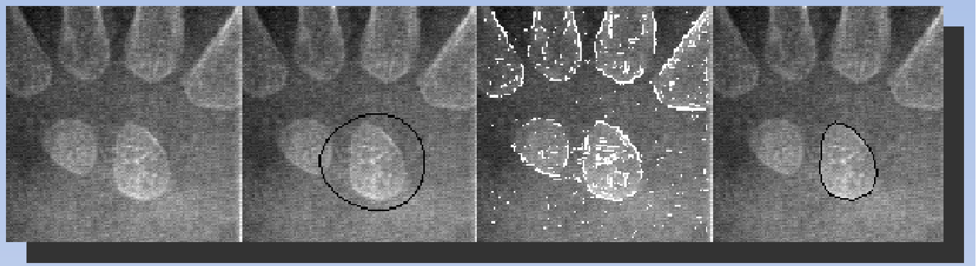
Segmentation of facial features in snapshots







Advantages of deformable shapes



- Physics-based and active models;
- Soft and hard constraints;
- Selective wrt false image features;
- Interpret sparse, incomplete and redundant information;
- Integration of data from multiple cues;
- Well-elaborated mathematical apparatus;
- Local and non-affine deformations;
- Generic restrictions allowed;
- Use of an approximate geometric object's model;
- Snakes regularize in a natural way ill-posed problems of Computer Vision.