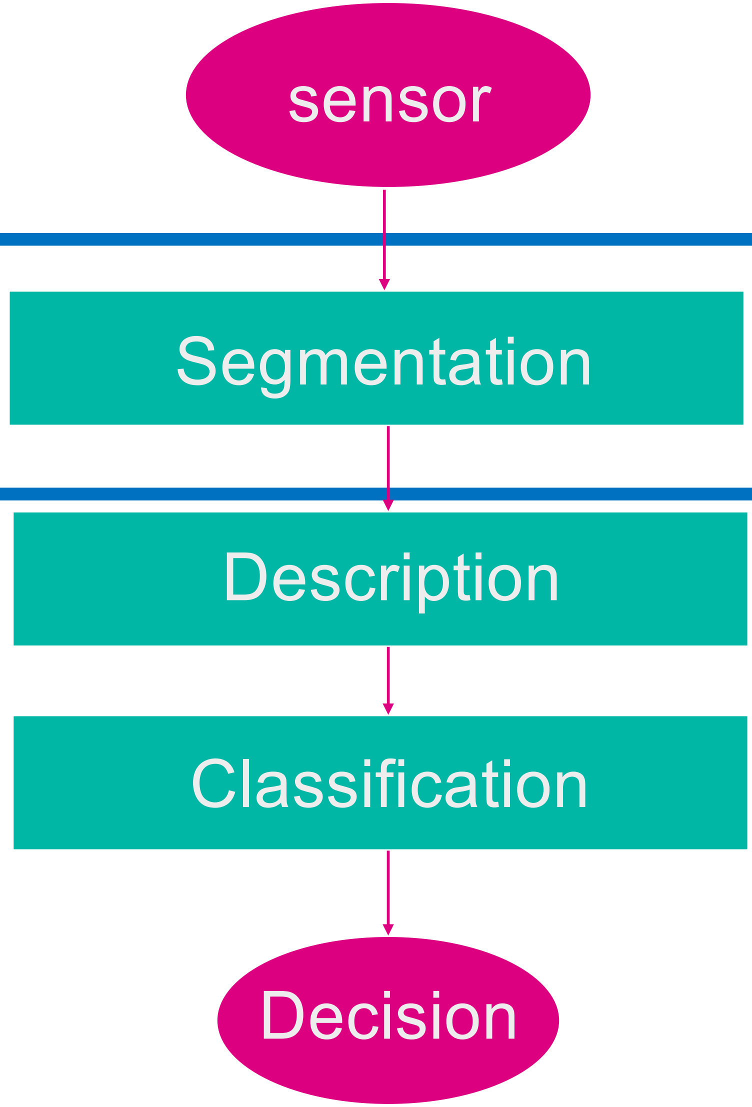
Image analysis and Pattern Recognition

Lecture 2 : image segmentation

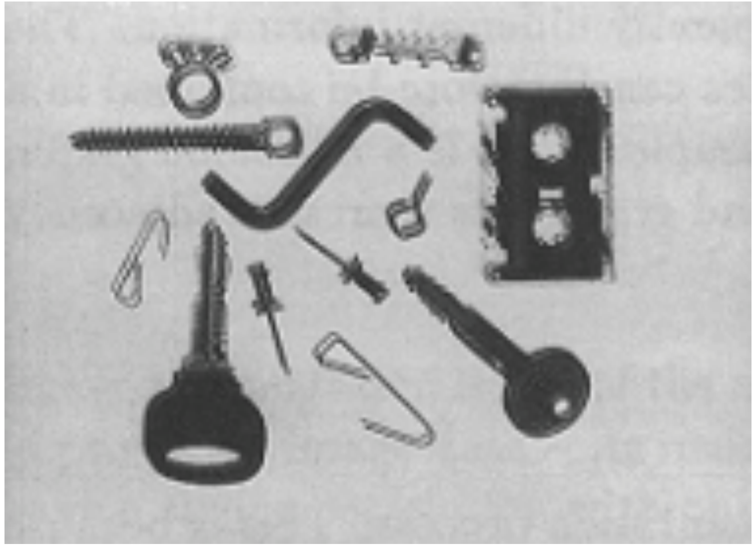
Prof. Jean-Philippe THIRAN
JP.Thiran@epfl.ch

The Pattern Recognition System



- Electricity, electronics, metrology
- Signal Processing
- Image Processing
- Geometry, topology, …
- Statistics
  - AI
- Robotics, …

Image segmentation: introduction



- In image analysis, we are interested in objects present in the images.
- Definition: object: part of an image that is semantically coherent
- In practice: often
  - Connex
  - Coherent color
  - Surrounded by sharp contours
- But also
  - Coherent Texture
  - Prior knowledge

- Definition : Segmentation: partition of an image in a finite number of regions $R_1, \ldots, R_s$ such that

$$R = \bigcup_{i=1}^{s} R_i, \qquad R_i \cap R_j = \varnothing, \quad i \neq j$$

- Segmentation methods aim at defining regions that correspond to the object sin the image
  - Region-based segmentation methods
  - Contour-based segmentation methods

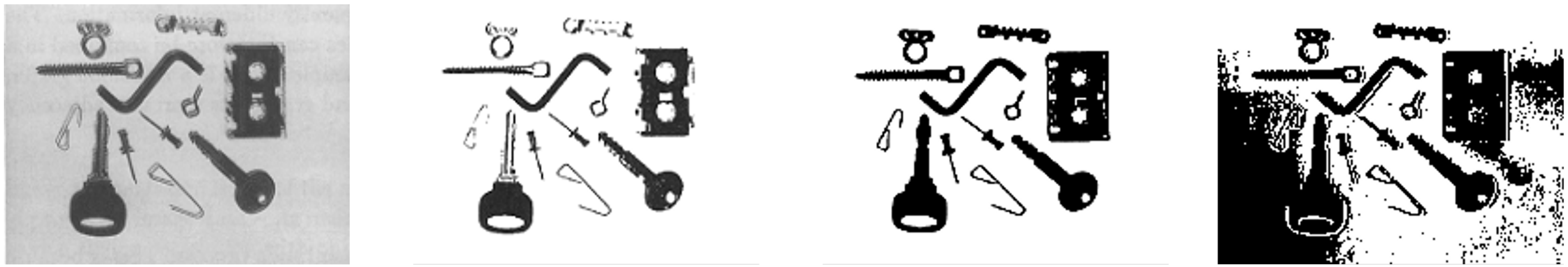
- Separation of the image into regions by setting one or several thresholds on the gray levels:

$$g(i,j) = \begin{cases} 1 & \text{si } f(i,j) \geq T \\ 0 & \text{si } f(i,j) < T \end{cases}$$

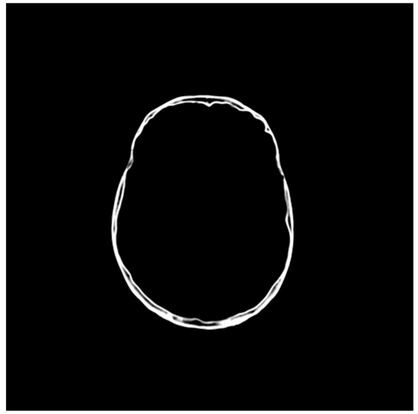
or

$$g(i,j) = \begin{cases} 1 & \text{si } f(i,j) \in D \\ 0 & \text{si } f(i,j) \notin D \end{cases}$$

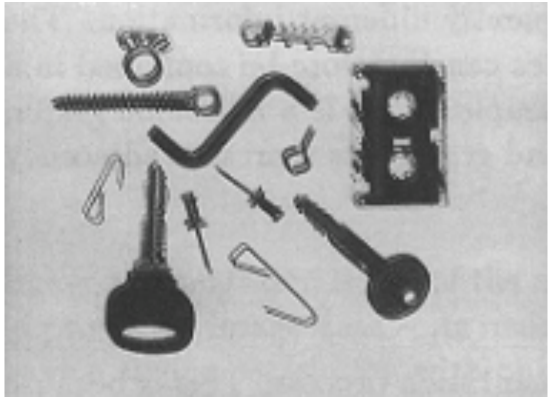
- The problem is how to find the optimal threshold(s)

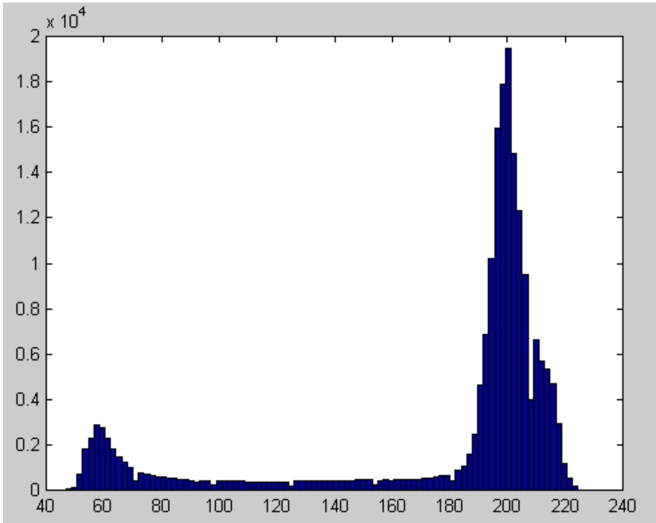


- Threshold detection methods:
  - Sometime we can know the value a priori
    - *Example : medical imaging (CT scan)*
  - If one knows some properties of the object to be segmented, one can use them to set the threshold:
    - *Example : size of the object*
    - *One can then define the threshold such that the object has the correct property*
  - Otherwise, the histogram can be used to set the threshold.

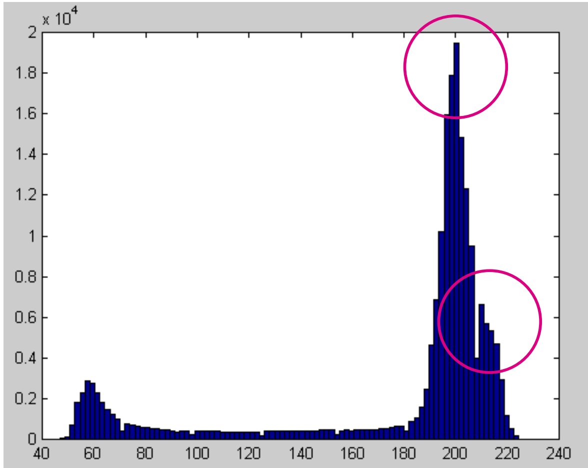


- Definition of the threshold based on the image histogram
  - Often histograms are bimodal
    - *E.g. when we have uniform objects on a uniform background (like back-lightning)*

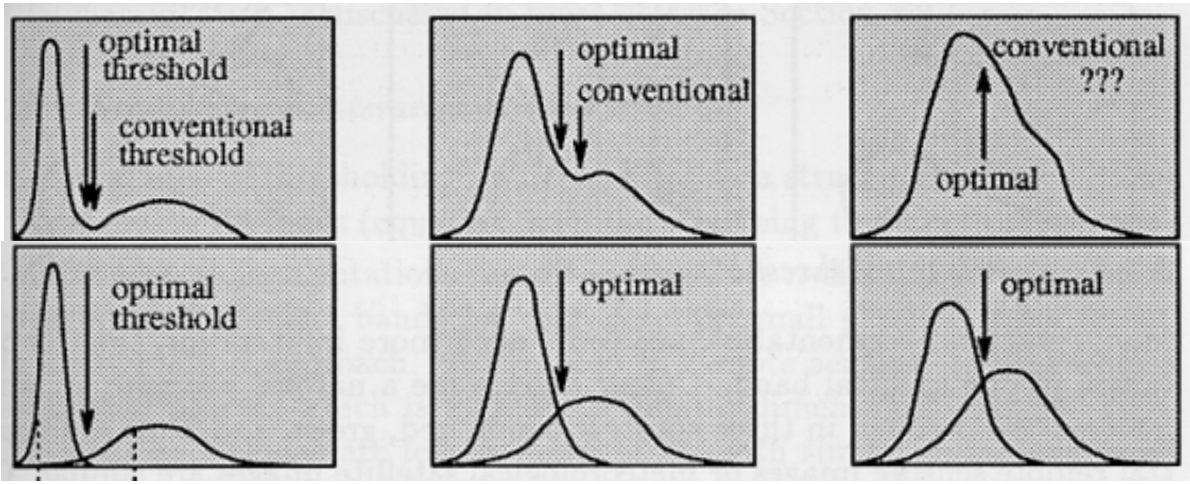




- Definition of the threshold based on the image histogram
  - Choice of the threshold that minimizes the segmentation error
  - Intuitively one could choose the threshold as the minimum between twp picks in the histogram
  - "Mode" method:
    - *Identify the two main maxima*
    - *Take the minimum in between*
    - *Often, we have to avoid to consider two local maxima that belong to the same mode*
  - Improve the « peak-to-valley » ratio
    - *Eliminate high gradient pixels*
    - *Or consider only high gradient pixels*



- Threshold definition based on the histogram: optimal thresholding
  - Minimizes the probability of wrong classification
  - Requires knowledge on the gray level distributions
    - *Often Gaussian, because of the noise*
    - *Not always: Reighley*
  - By identifying the gray level distribution of each class, one can find the optimal threshold
  - This optimal threshold is not (necessarily) the minimum of the histogram



- Threshold definition based on the histogram: optimal thresholding
  - If one knows regions that are obviously in the background, we can easily estimate its distribution
  - Otherwise, we can define a model (e.g. Gaussian) and fit curves to the histogram. Example 2 Gaussians

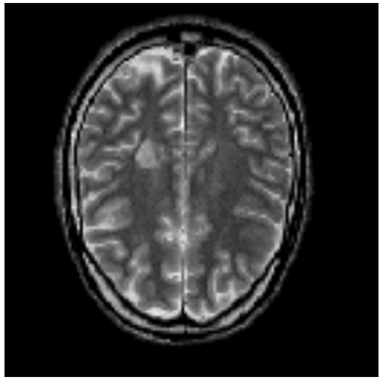
$$d_b(l) = A_b e^{-\frac{(l-c_b)^2}{2\sigma_b^2}} \qquad d_o(l) = A_o e^{-\frac{(l-c_o)^2}{2\sigma_o^2}}$$

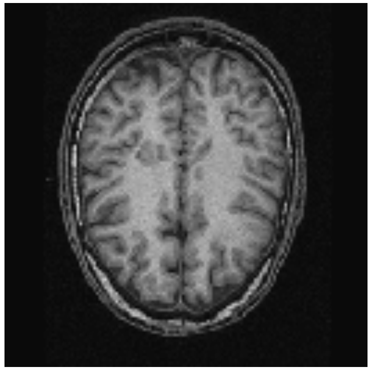
    - *variables :* $A_b, c_b, \sigma_b, A_o, c_o, \sigma_o$
    - *Function to optimize : e.g. minimize the means squared difference:*

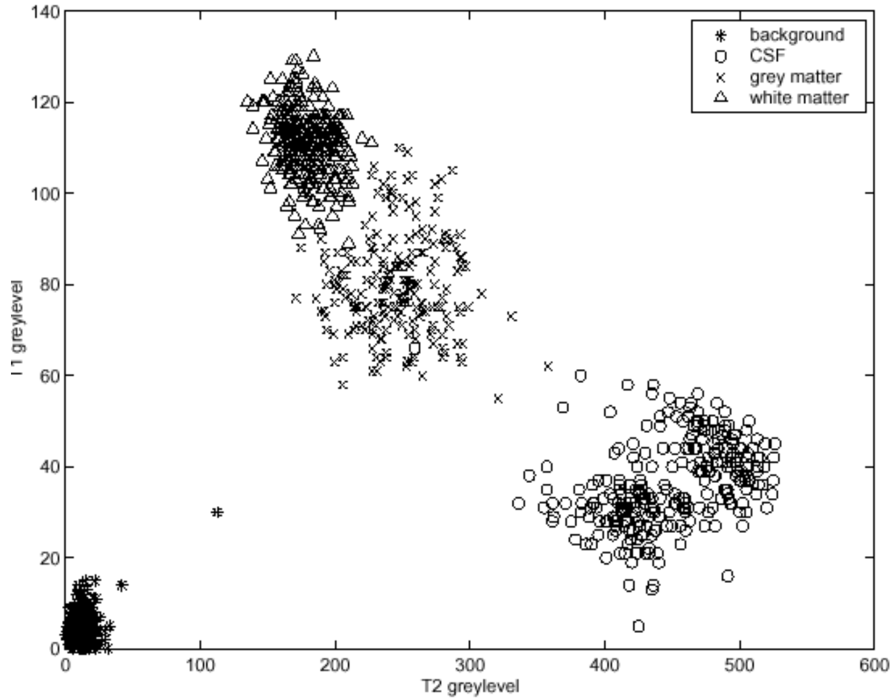
$$\varepsilon = \frac{1}{l_{\max}+1} \sum_{l=0}^{l_{\max}} \left[h(l) - \left(d_b(l) + d_o(l) \right) \right]^2$$

    - *Any optimization method can be used here : Expectation-Maximization (EM), simplex, Newton, …*

- Thresholding: partition in two regions in a 1D space
- Multispectral thresholding: idem, but in a n-D feature space





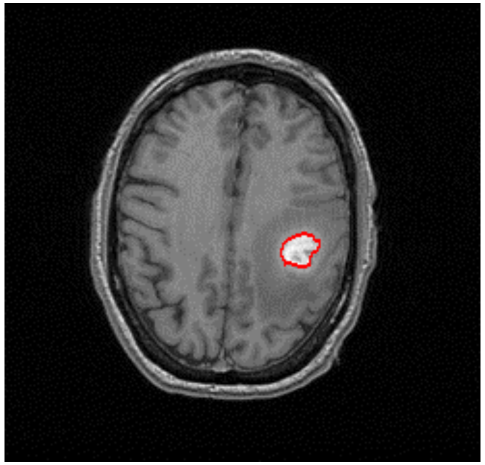


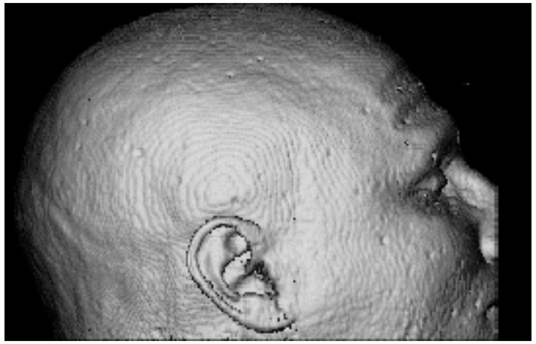
- Classification:
  - Chapter 3

- Segmentation by thresholding is very simple and works very well, but in a limited number of cases.
- Additionally, it does not define objects, but only separates background from foreground. The foreground does not necessarily define the objects (several clusters, not connex, …)
- This is what region growing methods aim at solving

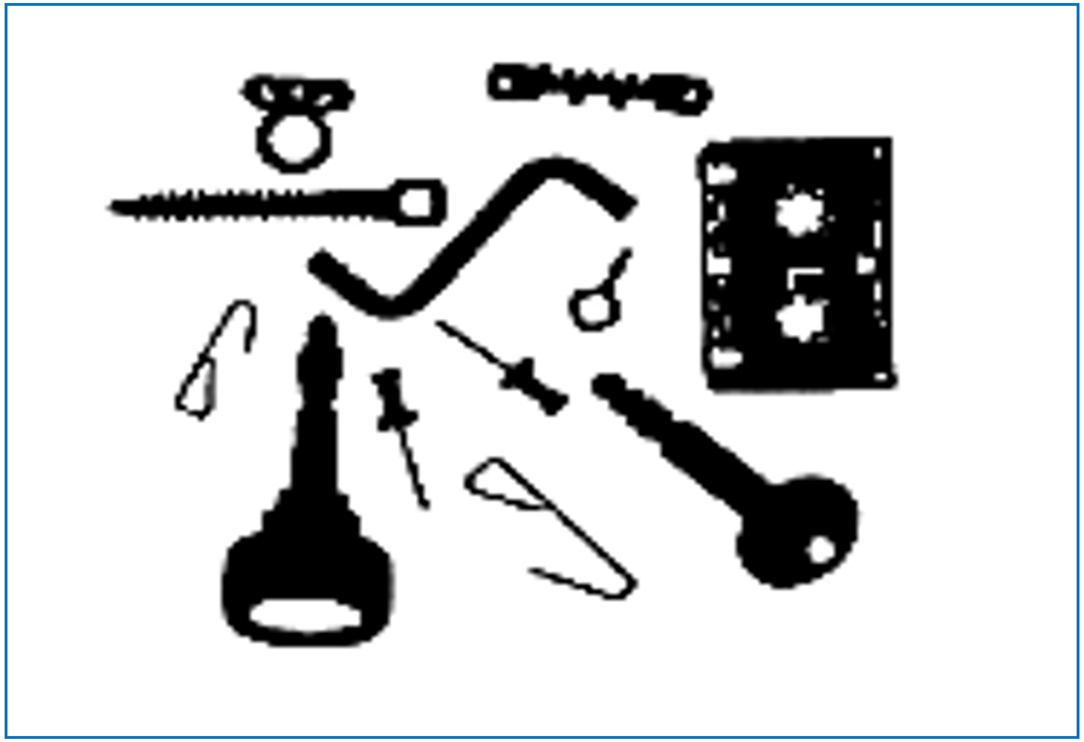
- Principe of region growing:
  - Let us fix a starting point (seed) in the desired region
  - Let us also define the homogeneity criterion used to define the region
    - *e.g. intensity > threshold*
  - By a recursive procedure, (i.e. neighbor to neighbor), let us include in the region all the pixels that are neighbors of the current pixel and that satisfy the homogeneity criterion
    - *By this the region will grow until it contains all the points connected to the seed point*
    - *We obtain a connex region*

- Notice that the homogeneity criterion can be quite sophisticated :
  - examples :
    - *Intensity varies slowly (intensity difference between two neighbors lower than a threshold)*
    - *The local variance is lower than a threshold (homogeneous texture)*
    - *...*
  - Examples of applications:
    - *semi-automatic segmentation of a tumor*
    - *Brain segmentation in MR images*
    - *Definition of the external contours of an object on a dark background*

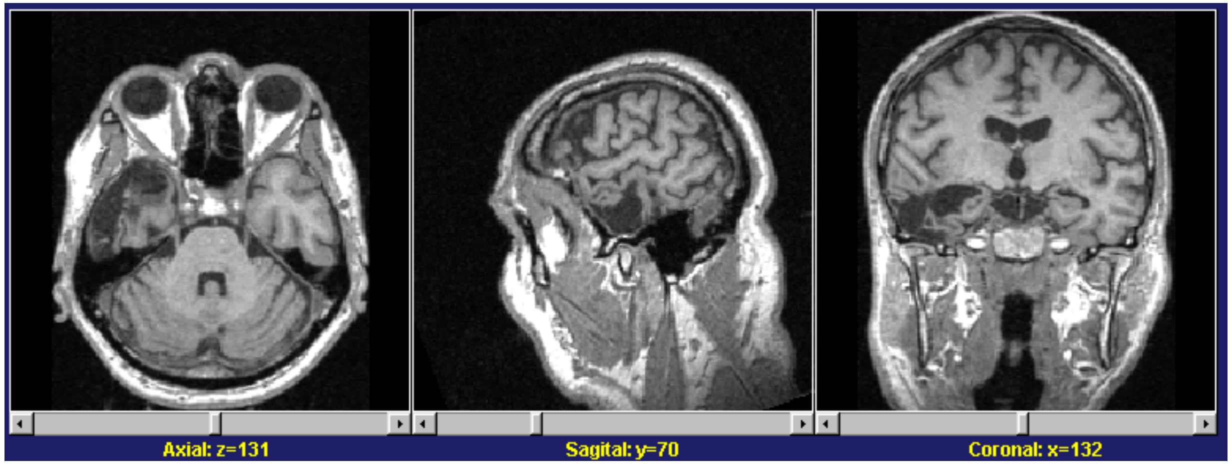


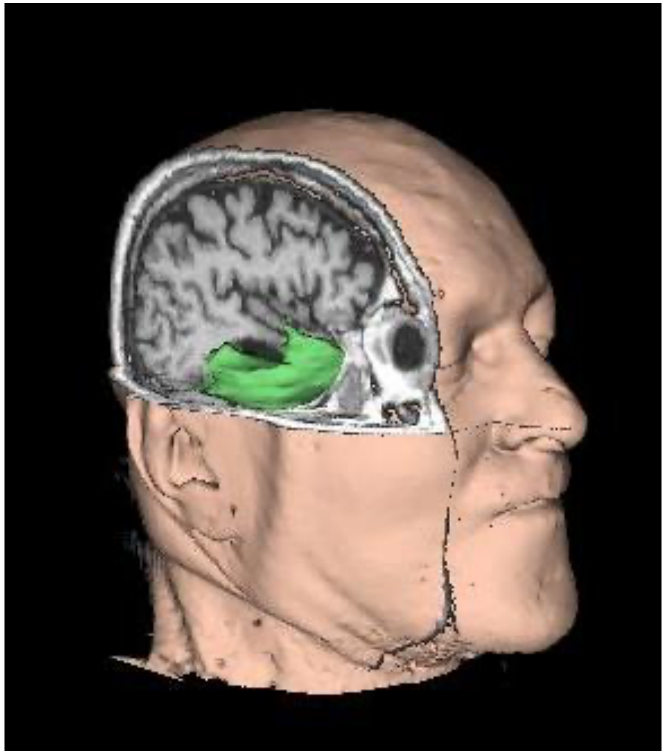


- Object labeling:



- Application example





- Region Merging:
  - On considers each pixel as a region
    - *But they do not all respect the homogeneity criterion*
  - One fuses adjacent regions if the homogeneity criterion of the union is respected
  - But the result depends on the order of the fusion process

  - On can also group blocks of 2x2, 4x4 or 8x8 pixels
  - The result is often a very rough segmentation

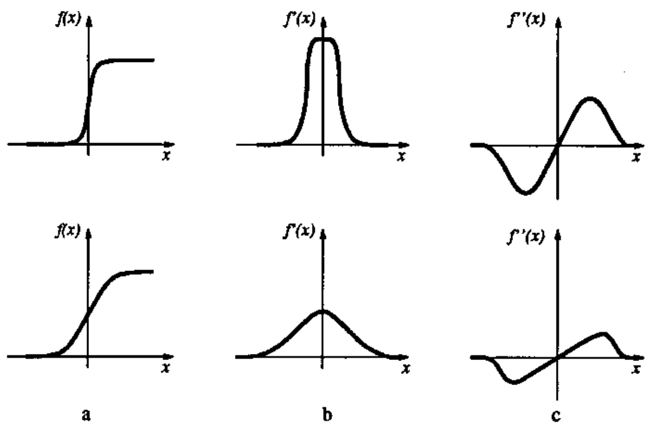
- Region splitting
  - This is the dual approach
  - One starts with the entire image
  - The homogeneity criterion is probably not met
    - *One divides the image, for instance in 4 parts*
  - This is repeated recursively until we reach small regions that can not be separated because they are homogeneous enough
  - The result is often a over-segmentation

- Split and Merge
  - One starts by splitting the image into (too many) small homogeneous regions
    - *Over-segmentation*
  - One regroups them by merging

Image segmentation: edge detection

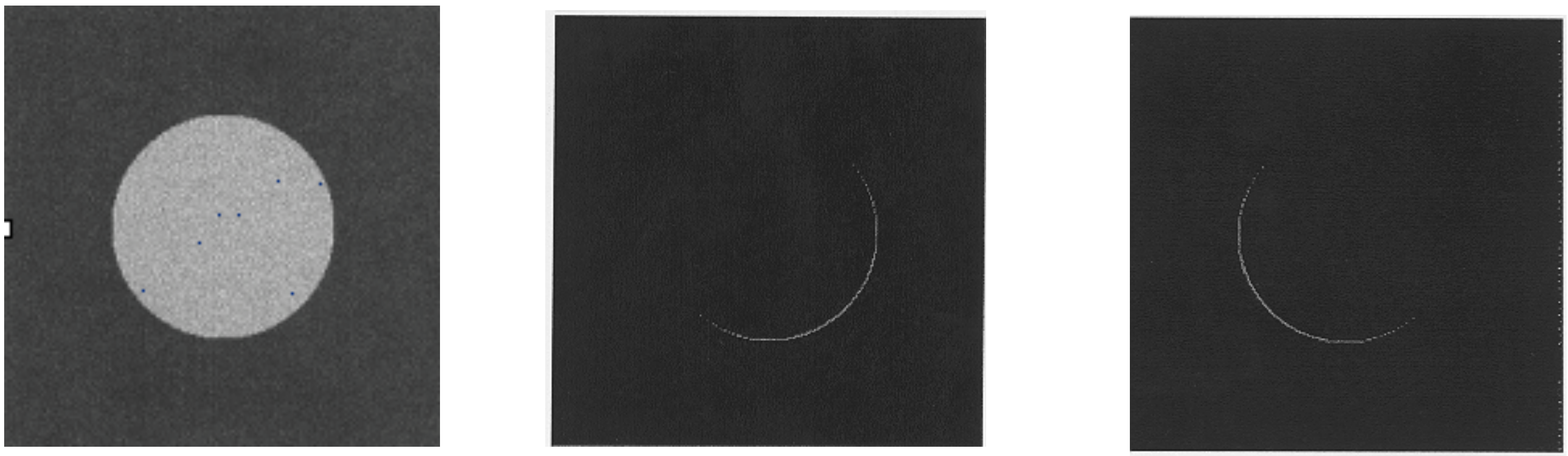
- The other way to define a region is to search for sharp edges: contour-based methods. An edge is
  - A sharp transition of intensity in an image
  - i.e. where the intensity profile is like a step function
  - i.e. where the 1st derivative has a maximum
  - i.e. where the second derivative crosses zero (zero-crossing).

- In the history there has been a lot of methods to detect the maxima of the 1st derivative
  - High-pass filter: convolution with various filters
  - Often direction



- Edge detection: maximum of the 1st derivative
  - Robert Operator:

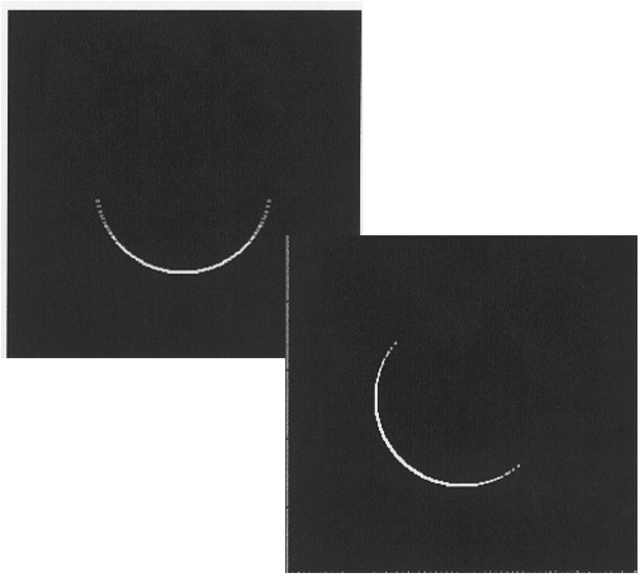
$$h_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \text{ et } h_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$



Edge detection: maximum of the 1st derivative

– Prewitt

$$h_1 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \quad h_2 = \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix} \quad h_3 = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$



…

– Sobel

$$h_1 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad h_2 = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix} \quad h_3 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

…

- Edge detection : zero-crossing of the 2nd derivative
  - Marr-Hildreth : *Laplacian of Gaussian*
  - principle :
    - *Noise will introduce false detections*
    - *Let us filter the image to remove the noise and undesired details: Gaussian filter*

$$G(x,y) = e^{-\frac{x^2+y^2}{2\sigma^2}}$$

    - *Let us calculate the 2nd derivative of the filtered image (Laplacian operator)*

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2}$$

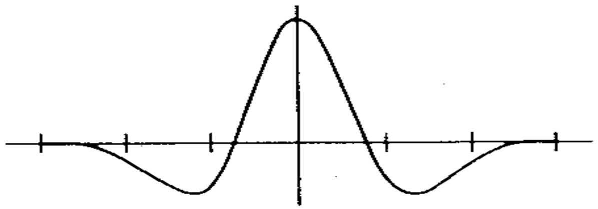
    - *Zero-crossing points of this filtered 2nd derivative are edge points*

Image segmentation: edge detection

- Trick : exploit the property of convolution:

$$\nabla^2\left(G(x,y)**f(x,y)\right)=\left(\nabla^2 G(x,y)\right)**f(x,y)$$

- Thus, it means that we just have to convolve the image by the second derivative of a Gaussian « Laplacian of Gaussian », (LoG) and find the zero crossing of the resulting image



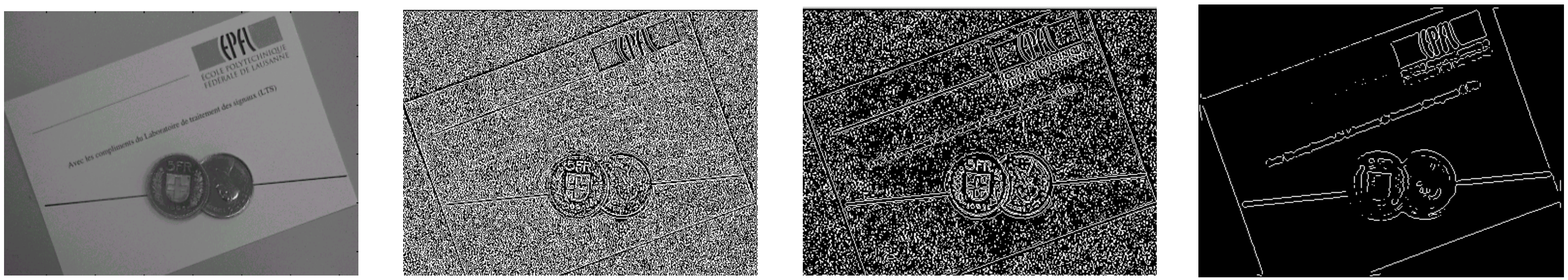
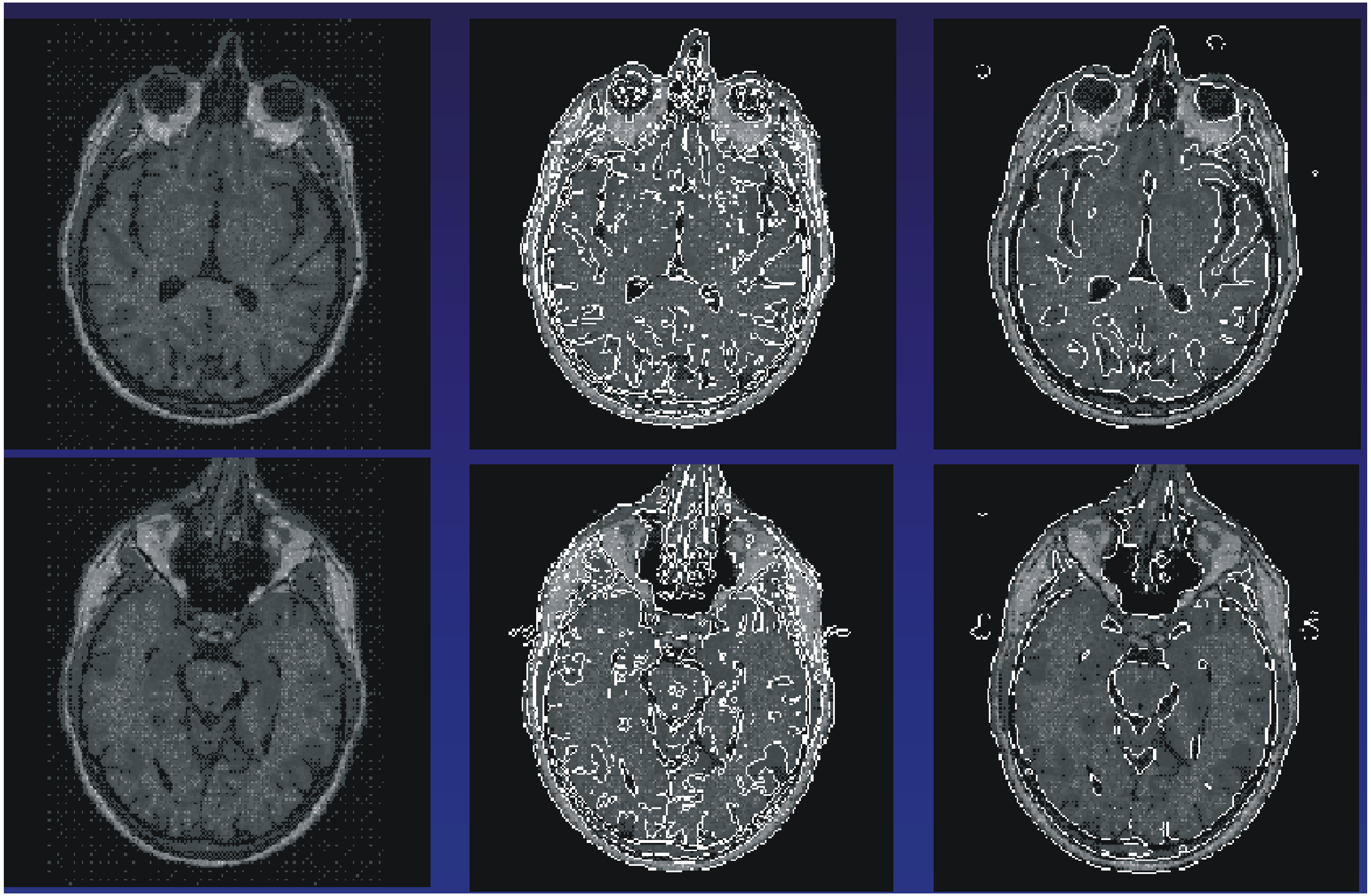
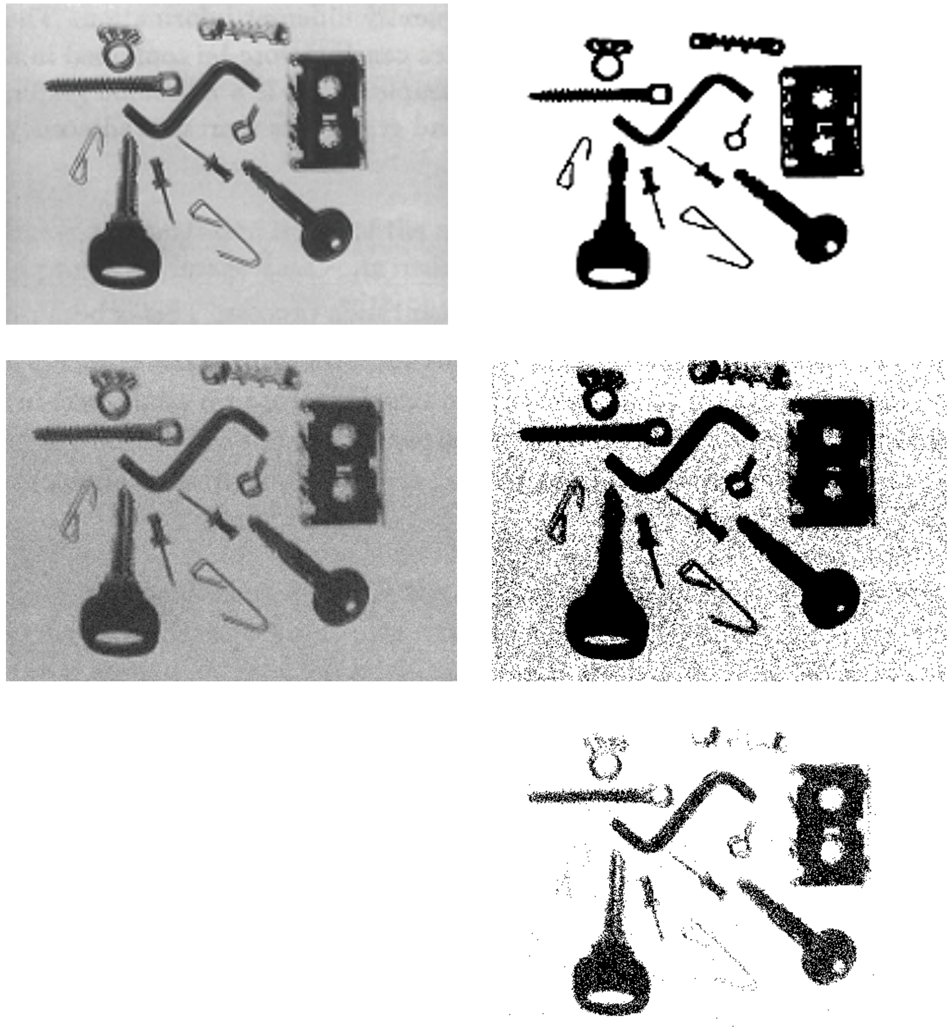


Image segmentation: edge detection



Mathematical morphology: motivations

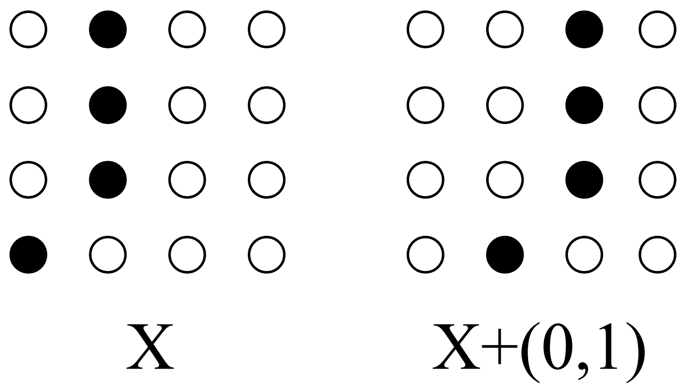
- If the image is not too noisy, no problem
- But most often, noise creates many problems after segmentation:
  - Irregular borders
  - Holes inside
- Solution:
  - Fill the holes
  - Smooth the borders



Mathematical morphology

- Very complete and coherent theory
  - Set of operators in image analysis based on shapes
  - Uses the vocabulary of set theory
- Principle:
  - Compare the objects present in an image with a reference object, of given size and shape, called structuring element
- Basic operators:
  - Dilation, erosion, opening and closing
- Origins: Jean Serra, CMM, École des Mines, Paris
  - *Image Analysis and Mathematical Morphology, London, Academic Press, 1982.*
  - R. Haralick, S. Sternberg et X Zhuang, *Image Analysis using Mathematical Morphology,* IEEE PAMI 9(4), 1987.

- Framework : discrete plan *E*
- Addition:



- Translation:

$$(A)_x = \{ c \in E \ \text{ t.q. } \ c = a + x, \forall a \in A \}$$

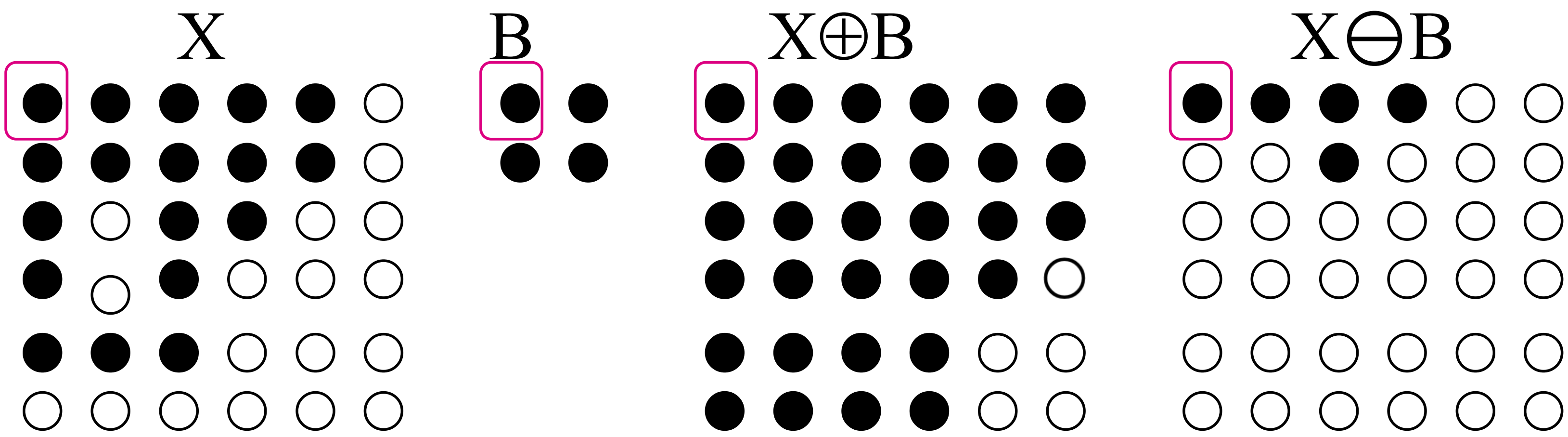
Binary morphology: dilation and erosion

- **Dilation :**

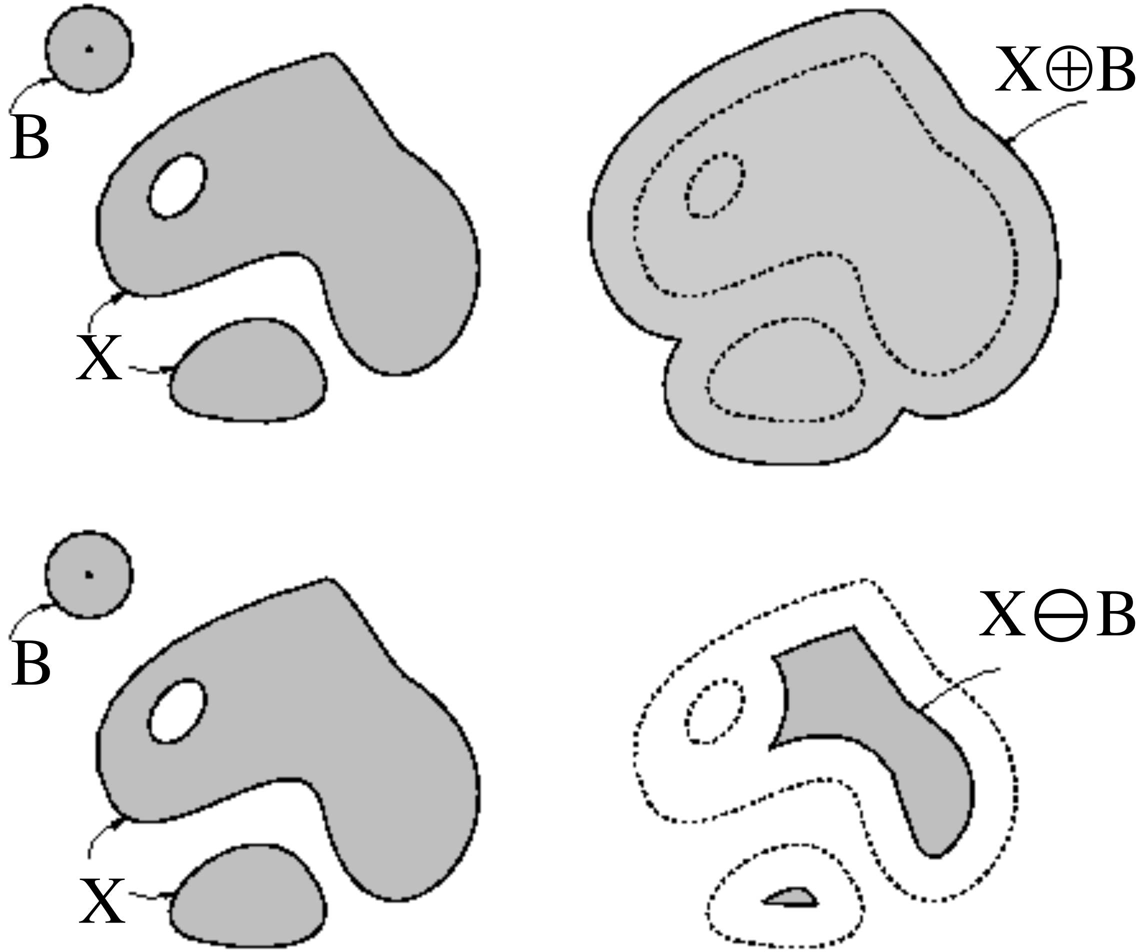
$$X \oplus B = \{c \in E \text{ t.q. } c = a + b \text{ avec } a \in X,\ b \in B\}$$
$$= \bigcup_{b \in B} (X)_b$$

- **Erosion :**

$$X \ominus B = \{c \in E \text{ t.q. } \forall \mathrm{b} \in \mathrm{B},\ c + b \in X\}$$
$$= \bigcap_{b \in B} (X)_{-b}$$

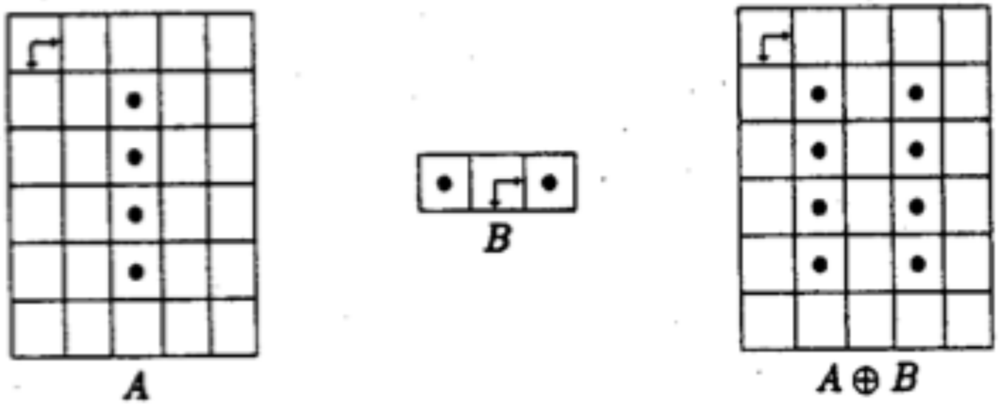


Binary morphology: dilation and erosion



Binary morphology: dilation

- extensivity :
  - if $0 \in B$, $A \subseteq A \oplus B$



- increasing:
  $A \subseteq B \Rightarrow A \oplus C \subseteq B \oplus C$
- $(A \cup B) \oplus C = (A \oplus C) \cup (B \oplus C)$
- $(A \cap B) \oplus C \subseteq (A \oplus C) \cap (B \oplus C)$

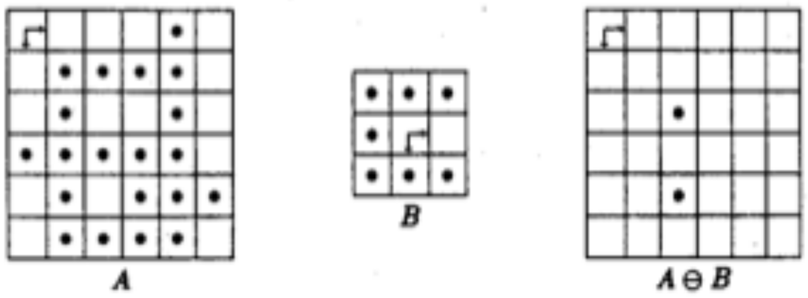
Binary morphology: erosion

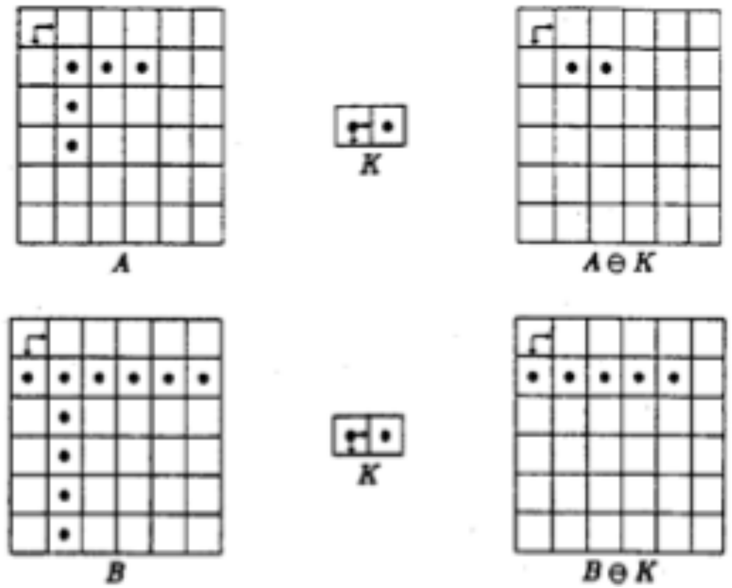
- $A \ominus B = \{x \in E \mid (B)_x \subseteq A\}$
- Anti-extensivity :
  - *si* $0 \in B$, $A \ominus B \subseteq A$
- Invariance by translation

  $(A)_x \ominus B = (A \ominus B)_x$

  $A \ominus (B)_x = (A \ominus B)_{-x}$
- increasing :

$A \subseteq B \Rightarrow A \ominus K \subseteq B \ominus K$





Binary morphology: dilation & erosion

- $A \subseteq B \Rightarrow D \ominus A \subseteq D \ominus B$

- Definitions :
  - Complement of A : $A^c = \{ x \in E \mid x \notin A \}$
  - Reflection of B : $\breve{B} = \{ x \;\; t.q. \;\; -x \in B \}$

- Duality erosion-dilation: $(A \ominus B)^c = A^c \oplus \breve{B}$

- $A \ominus (B \oplus C) = (A \ominus B) \ominus C$
- Invariance by rotation if B is a circle

Binary morphology: opening and closing

- **Opening:**

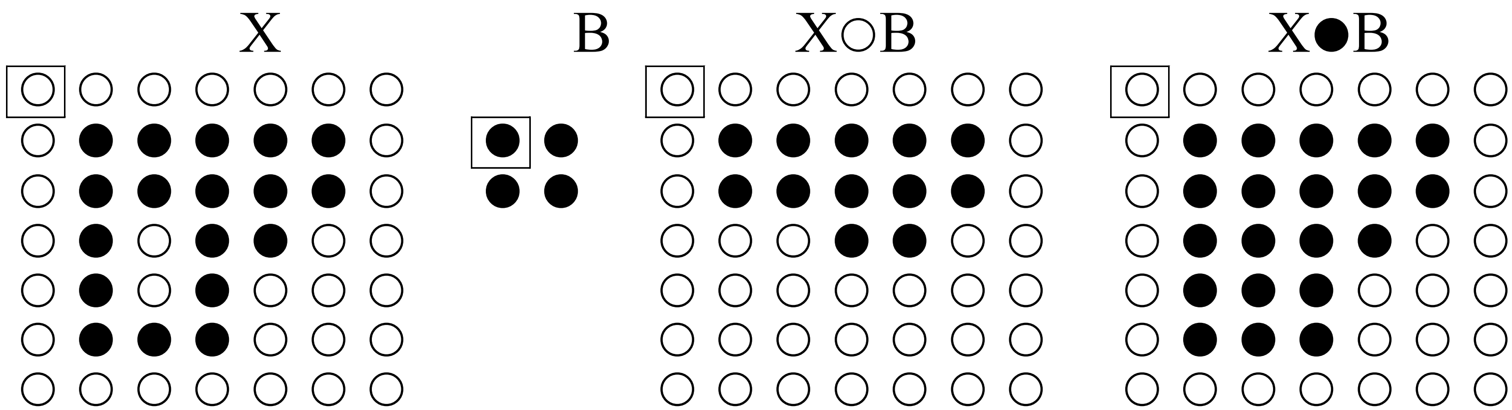
$$X \circ B = (X \ominus B) \oplus B$$

$$= \bigcup_{\{x \text{ t.q. } B_x \subseteq X\}} B_x$$

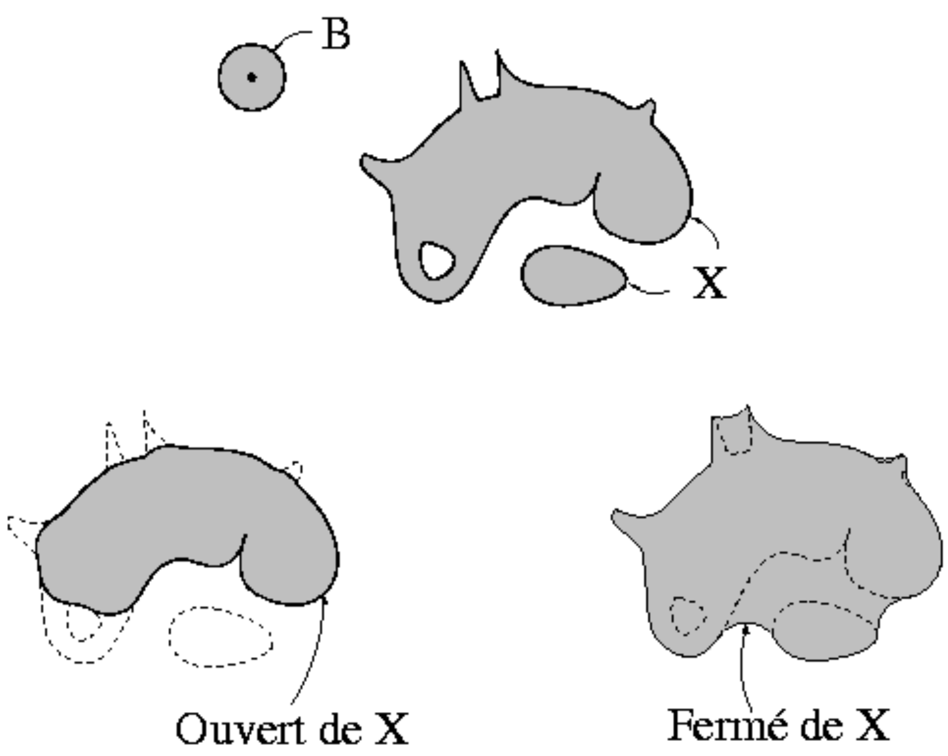
- **Closing:**

$$X \bullet B = (X \oplus B) \ominus B$$

$$= \bigcap_{\{x \text{ t.q. } \breve{B}_x \cap X \neq \emptyset\}} \breve{B}_x^c$$



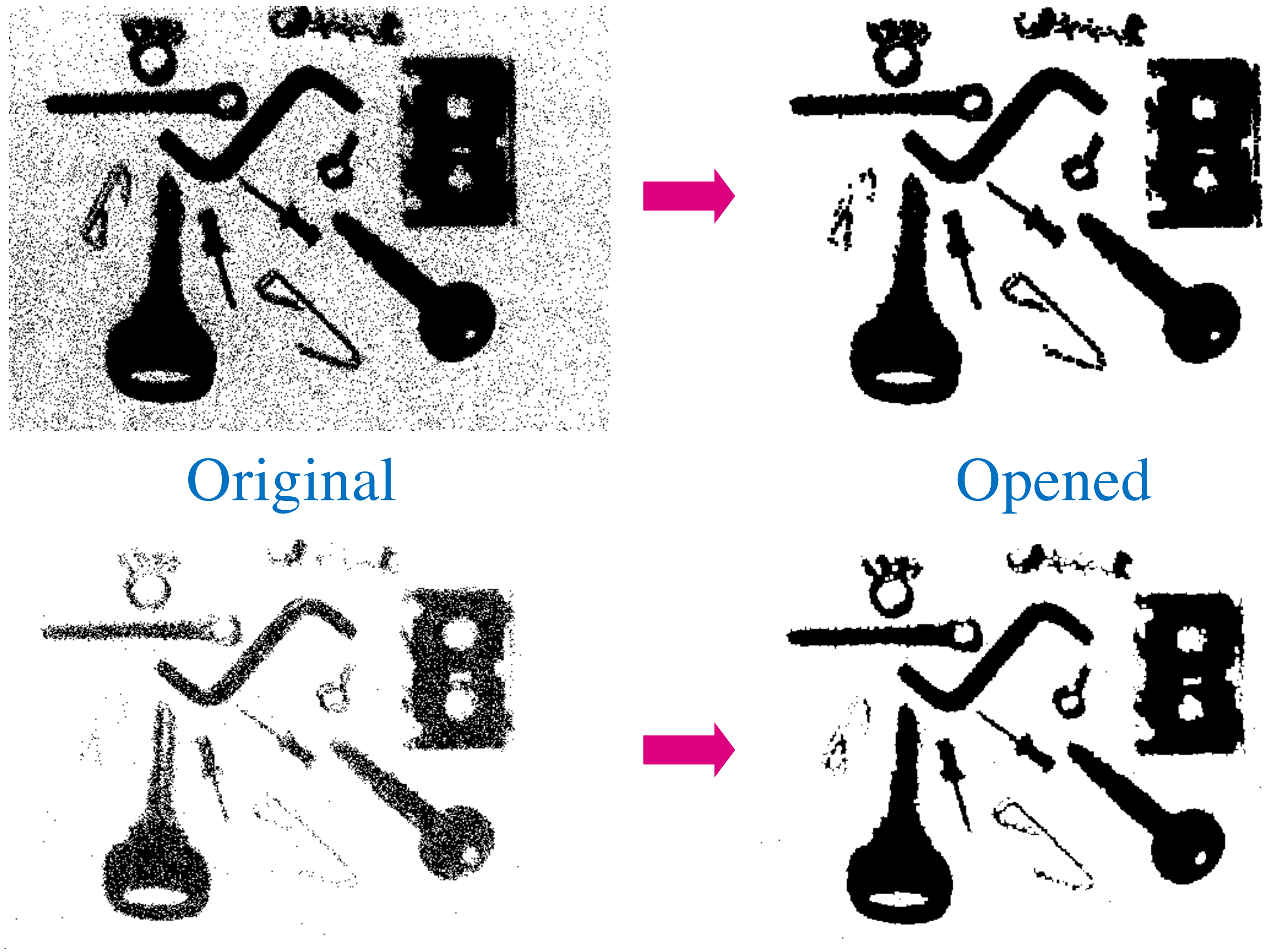
Binary morphology: opening and closing



- Properties:
  - duality : $X \circ B = (X^c \bullet \breve{B})^c$
  - extensivity :
    $X \circ B \subseteq X$ et $X \subseteq X \bullet B$
  - idempotence :
    $(X \circ B) \circ B = X \circ B$
    $(X \bullet B) \bullet B = X \bullet B$

- Principe : comparison with a structuring element
- Erosion & dilation: change the size of the objects
- Opening & closing: fill the holes
  - Inside the objects (closing)
  - External to the objects (opening)
- Use:
  - Post-processing after segmentation
  - cfr. examples

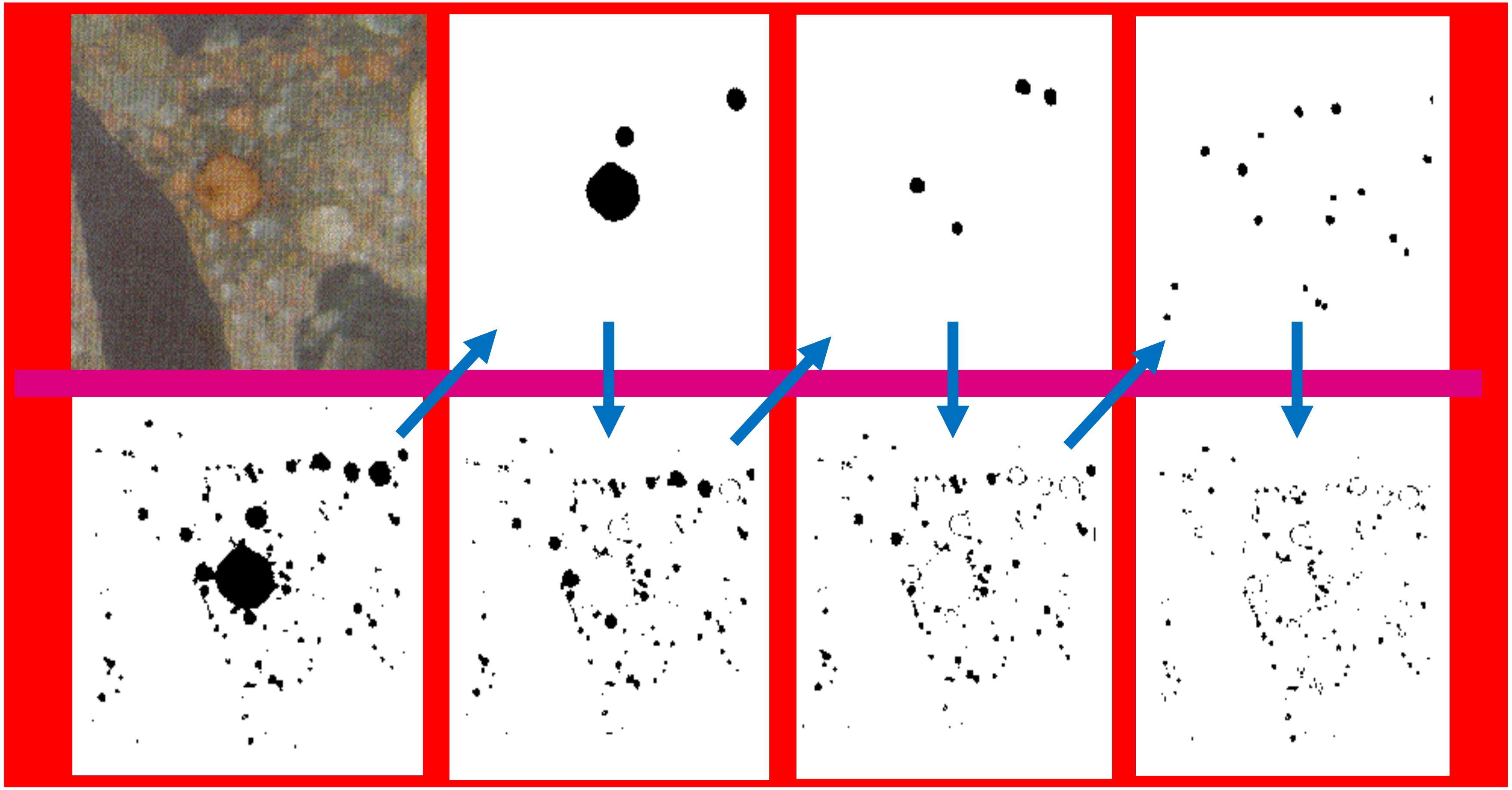
Binary morphology: conclusions



Original — Opened

Original — Closed

- Granulometry in Concrete



- Framework: functions $f : E \rightarrow C : x \rightarrow f(x)$
- Supremum (∨) and infimum (∧)
  - $(f \vee g)(x) = max\ (f(x), g(x))$
  - $(f \wedge g)(x) = min\ (f(x), g(x))$
- Ordering relationship :
  - $f \leq g \Leftrightarrow f(x) \leq g(x) \quad \forall x \in E$
- Translation: $f_b = f(x-b)$
- Correspondence :

Binaire	Niveaux de gris
∪, ∩	∨, ∧
⊆, ⊇	≤, ≥

- Definitions :

$$(f \oplus k)(x) = \bigvee_{z \in K} f_z(x) + k(x)$$

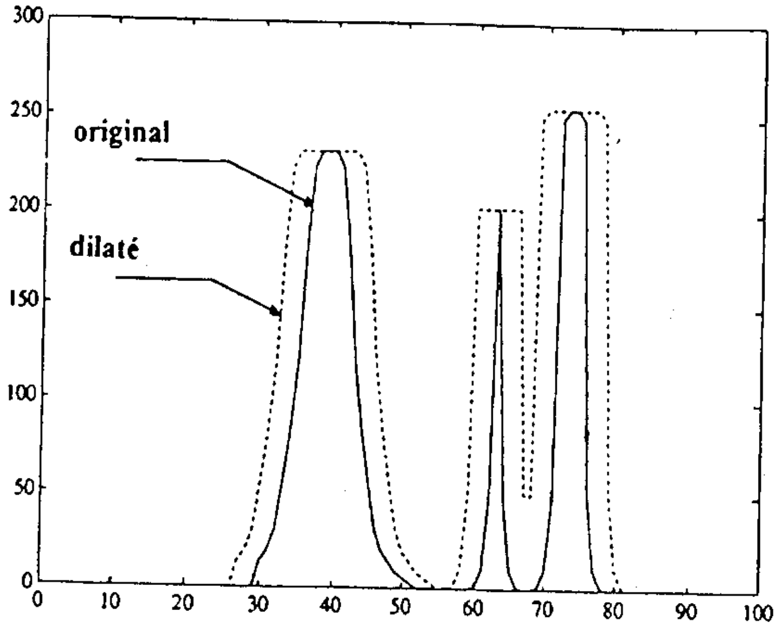
$$(f \ominus k)(x) = \bigwedge_{z \in K} f_{-z}(x) - k(x)$$

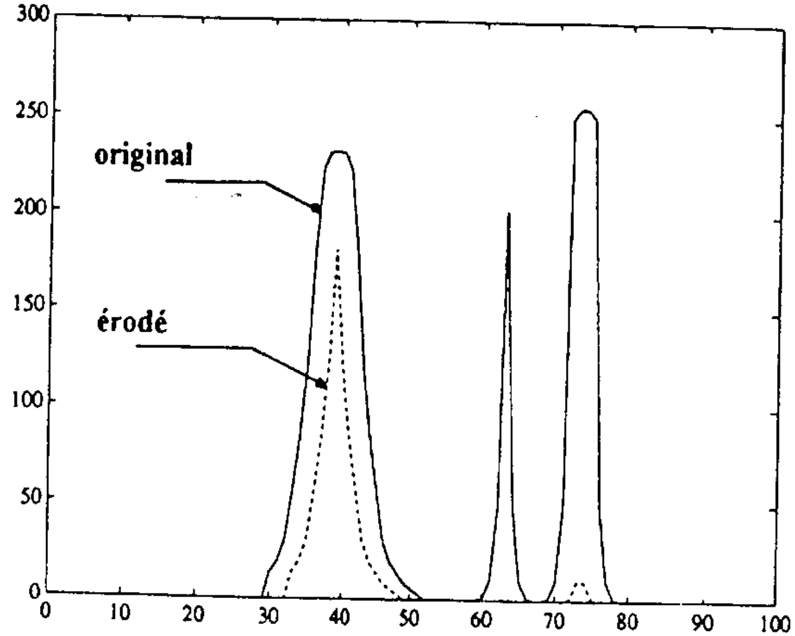
$$(f \circ k)(x) = (f \ominus k) \oplus k$$

$$(f \bullet k)(x) = (f \oplus k) \ominus k$$

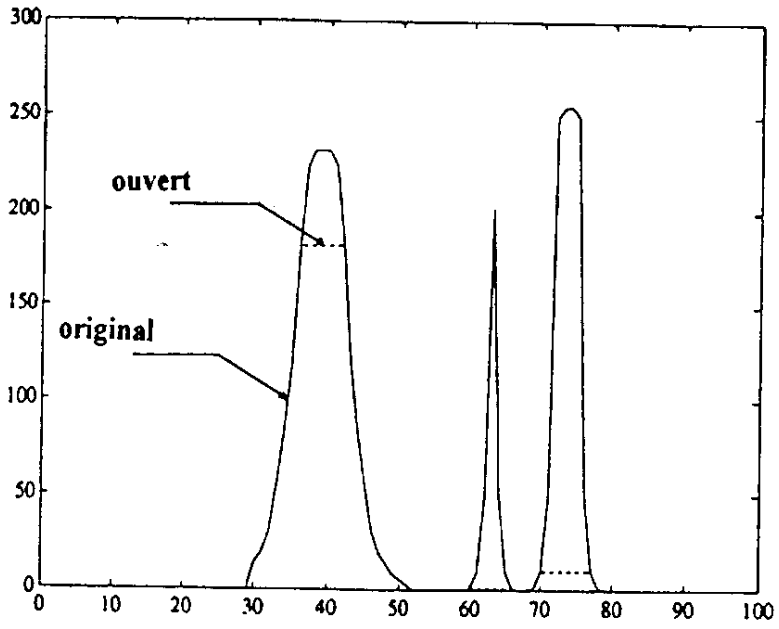
- Properties: cfr binary morphology

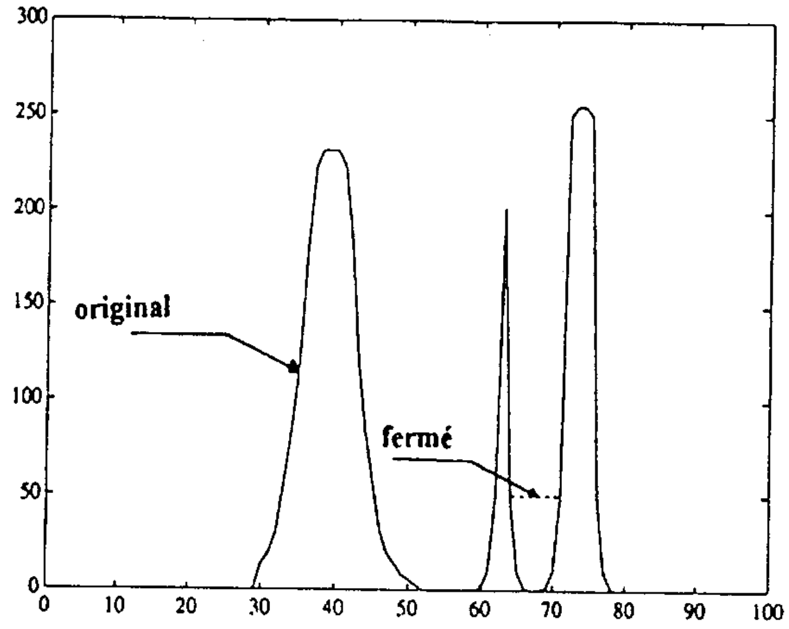
Gray-scale Morphology

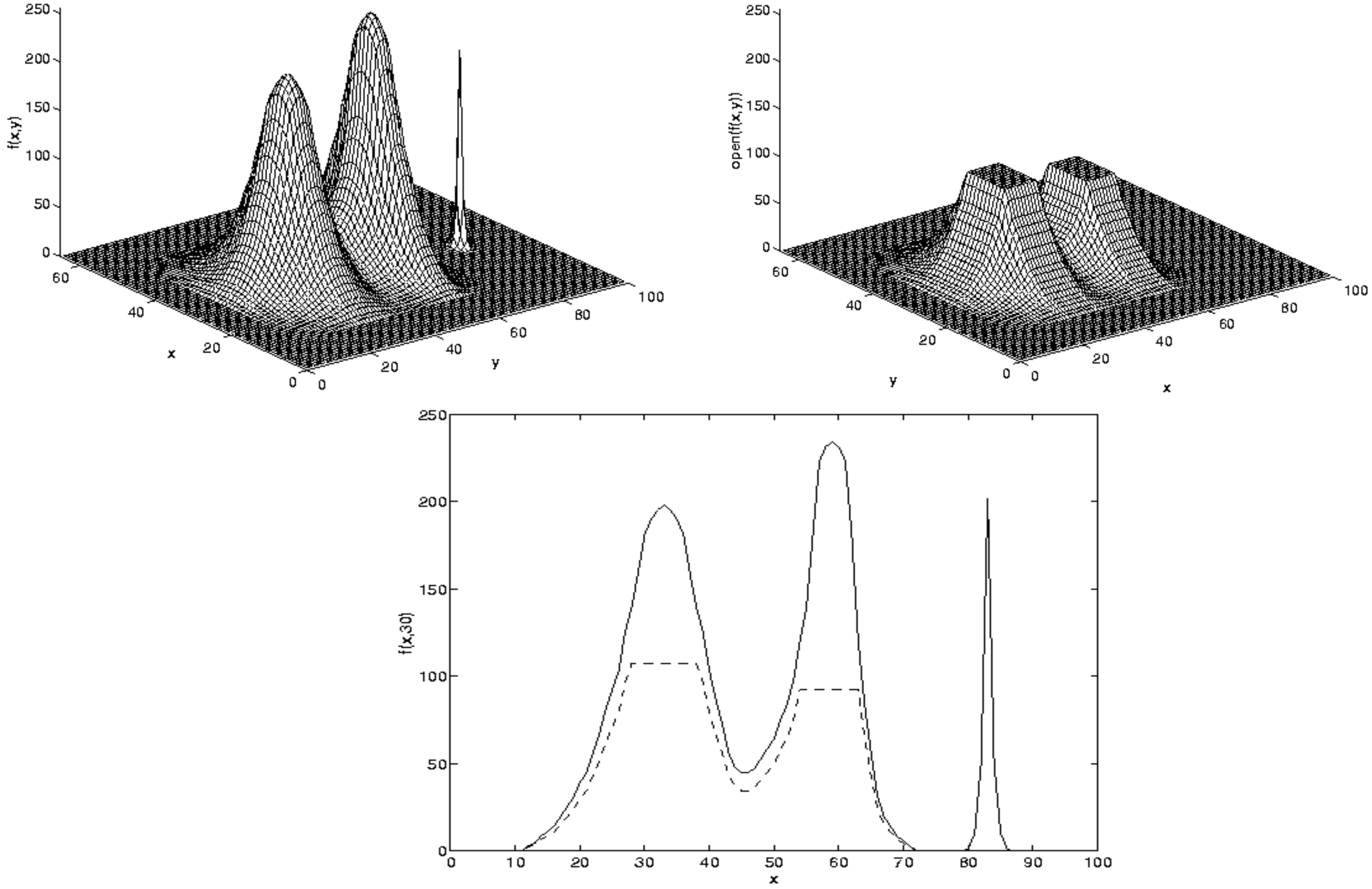




Gray-scale Morphology

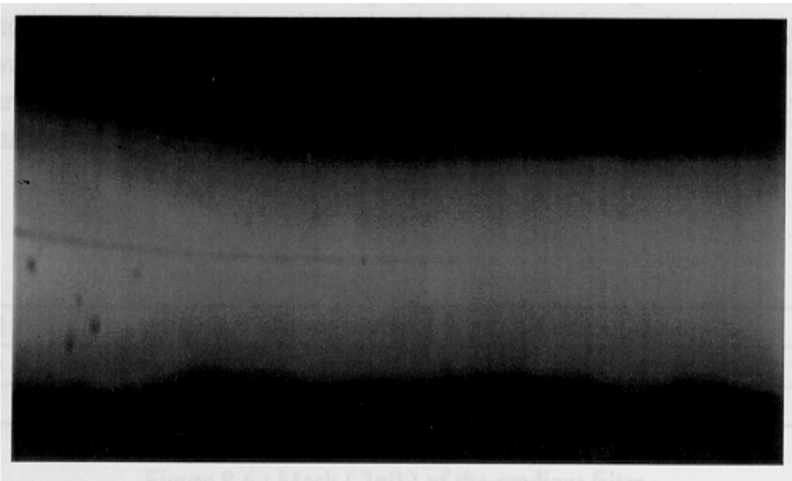


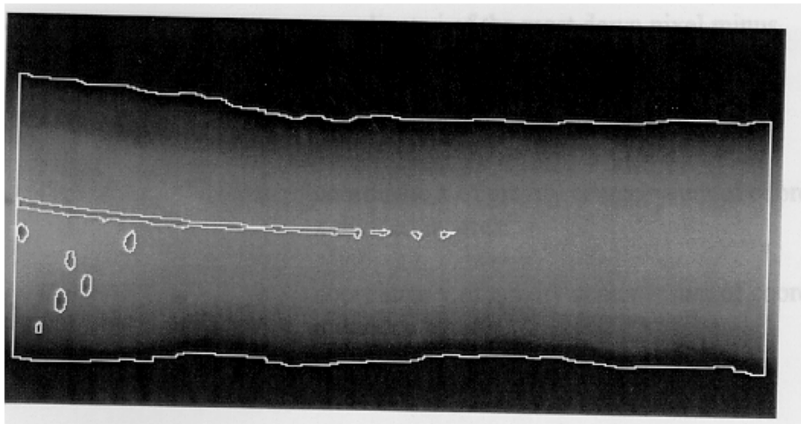


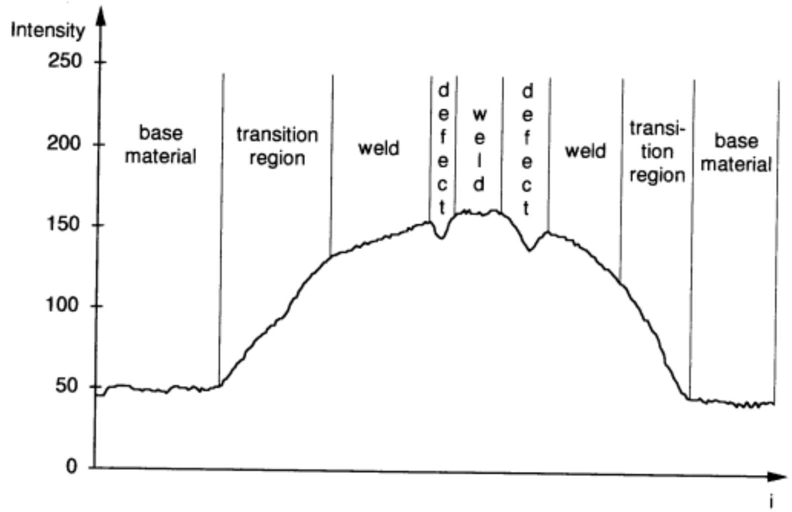

250
200
150
100
50
0
f(x,y)
60
40
x
20
0
0
20
40
60
80
100
y
open(f(x,y))
y
x
f(x,30)
0
10
20
30
40
50
60
70
80
90
100

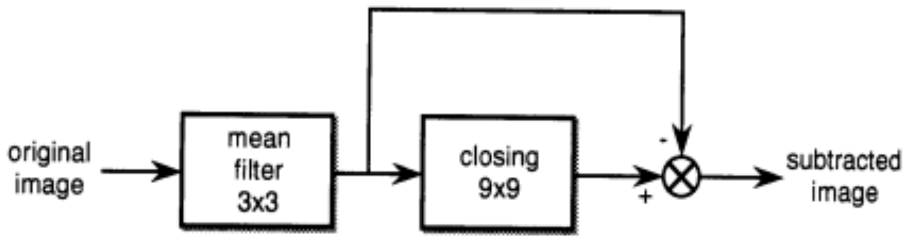
(c)

- Defect detection in welding









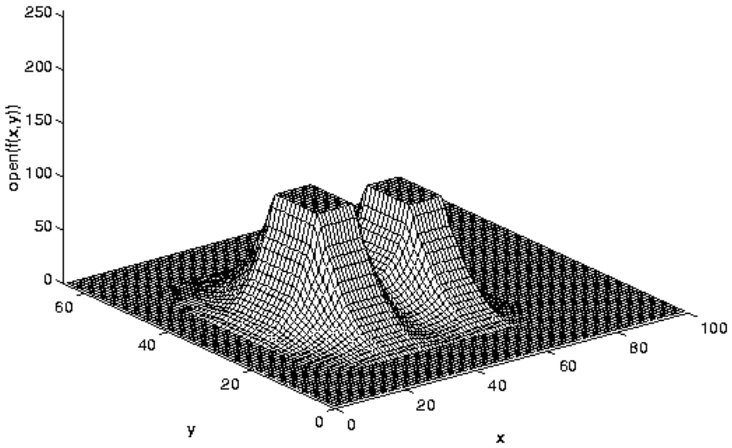
- Geodesic dilation:

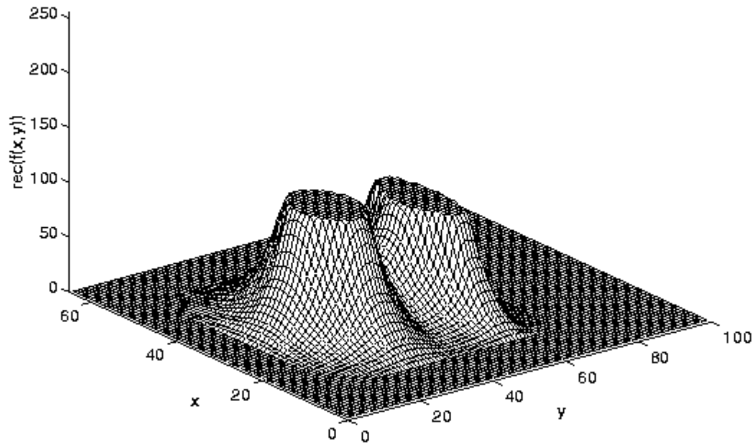
$$(X \oplus B)^{(1)} = (X \oplus B) \cap Y$$

$$(X \oplus B)^{(n)} = \underbrace{(((X \oplus B)^{(1)} \oplus B)^{(1)} \dots)^{(1)}}_{n\ fois}$$

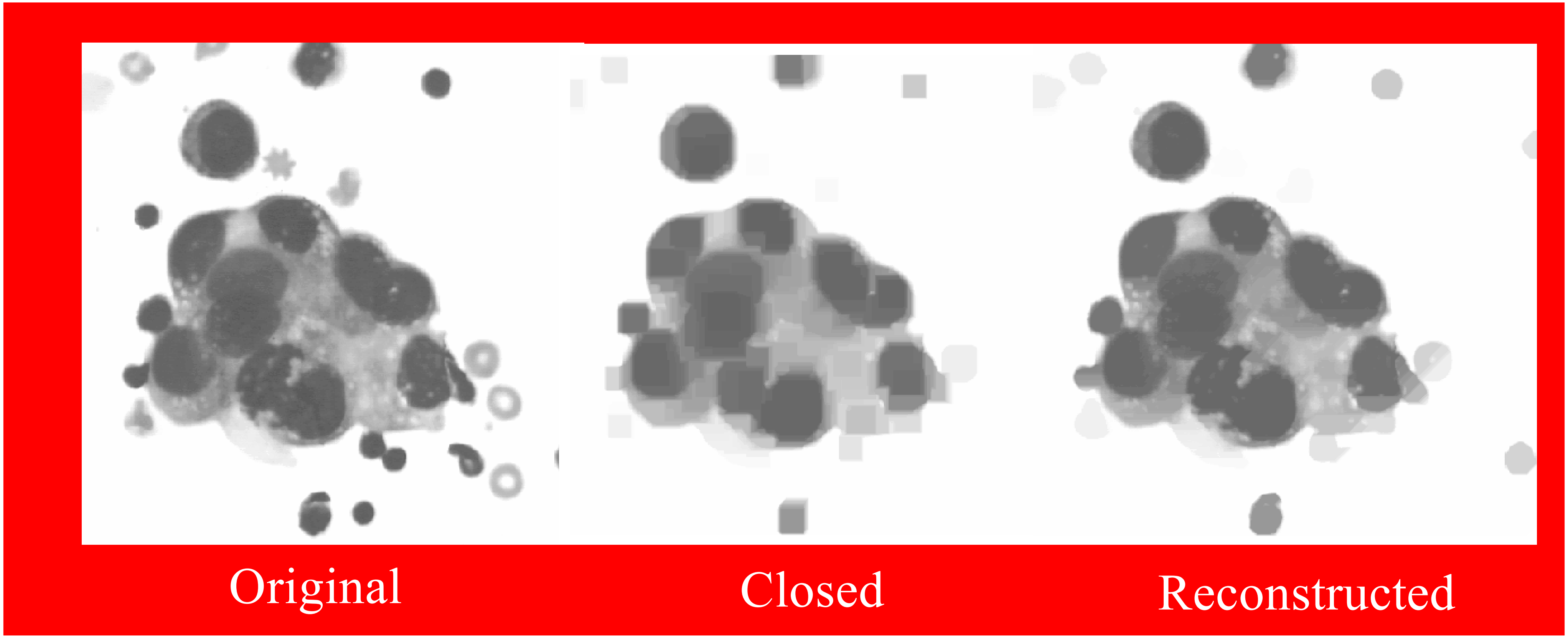
- Reconstruction :

$$R_Y = (X \oplus B)^{(i)}, \text{avec}\,(X \oplus B)^{(i-1)} = (X \oplus B)^{(i)}$$

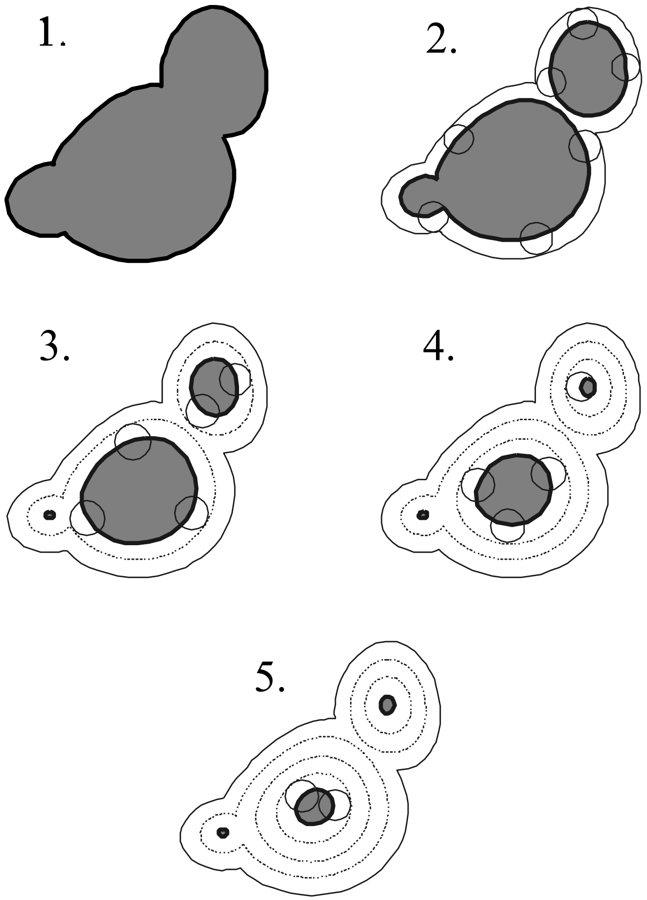


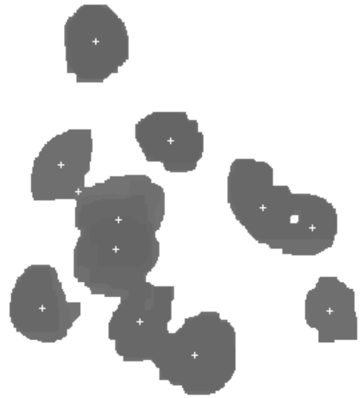


- Morphological segmentation: cancer cell detection and analysis



Application: feature extraction





Thresholded

« Ultimate Eroded »