Vision par ordinateur : Approche traditionnelle

Le royaume de la créativité

- ***Hans Moravec***: *"it is comparatively easy to make computers exhibit adult level performance on intelligence tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility"*

Plan

- Définir une image
- Traitements de base
- Taches de haut niveau
  - Un point clé : La représentation
  - Classification
  - Détection
  - Segmentation

Qu’est-ce que la vision ?

Percevoir le monde

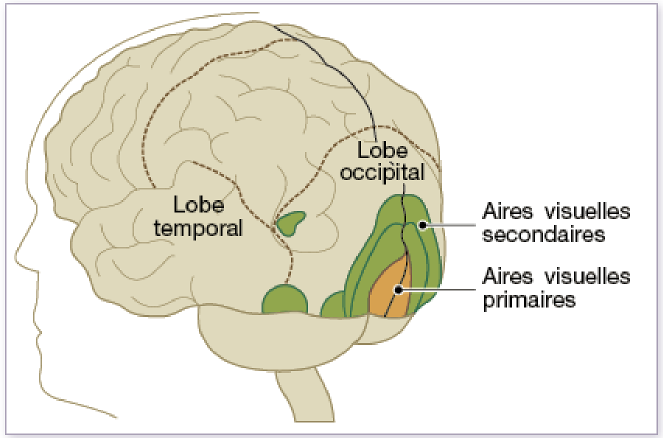
- Structure 3D
- Composé d’objets
- Efficacement interprété par l’être humain

Info recueillie

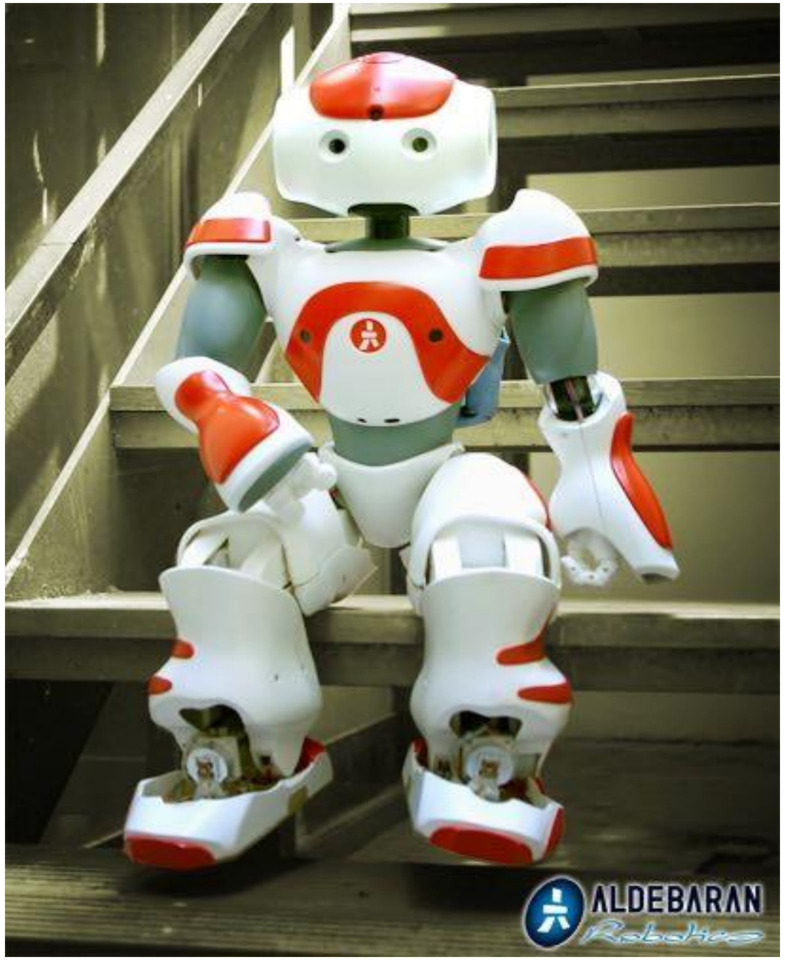
- Ensemble de points
- Pixels -> info sur la lumière
- Quantité et contenu spectral/couleur

Comment « voir » les objets ?

- Les objets n’existent pas sur la rétine
- Interprétation -> processus visuel



Qu’est ce qu’une image pour un ordinateur ?



Représentation des images

79	73	53	46	52	44	41	42	41	42	42	43	43	43	43	42	40	42	44	42	42	44	45	44	43	42	42	42	43	44	45	45	45
65	44	41	38	36	37	38	35	33	34	35	36	37	36	35	35	33	35	36	35	35	36	37	37	37	37	37	38	39	39	39	38	38
77	41	44	40	33	39	37	34	34	34	36	38	38	37	36	37	36	36	37	37	37	36	37	39	38	37	37	38	40	41	40	39	39
79	42	36	31	34	39	35	37	35	34	36	38	39	37	37	38	36	35	35	36	35	34	35	38	40	39	37	37	39	41	41	40	38
82	50	37	33	37	37	34	36	34	33	34	36	36	36	36	37	39	37	36	38	37	35	37	40	40	39	38	38	39	41	41	40	37
64	42	35	36	36	34	38	36	36	36	36	37	37	38	38	38	37	35	34	36	36	35	35	37	39	40	41	41	42	43	41	38	40
47	34	34	35	35	34	37	35	36	37	38	37	37	39	39	37	37	35	35	37	38	37	36	36	38	40	41	41	41	42	40	36	41
44	36	38	35	40	40	34	36	35	38	39	37	37	39	39	36	37	36	37	39	40	40	38	36	40	42	42	41	41	42	41	38	40
44	36	38	36	38	37	36	35	37	36	38	39	40	42	42	37	36	38	38	37	36	37	37	37	41	43	42	42	42	41	39	41	41
43	36	38	37	39	37	36	35	36	36	37	38	38	40	41	38	39	40	39	37	35	36	38	40	40	41	40	41	43	42	40	40	40
44	37	39	37	39	37	36	35	35	35	36	37	37	39	40	39	40	41	40	38	36	37	39	41	39	40	39	41	44	44	41	40	40
46	39	40	37	38	36	35	34	34	36	37	38	37	38	39	40	39	39	40	40	38	38	39	39	41	41	40	41	43	43	42	42	42
46	39	39	37	38	36	36	36	35	37	39	39	39	39	40	41	39	38	39	40	39	40	39	37	42	43	41	40	41	41	41	43	43
46	38	39	36	37	37	38	38	37	39	39	40	41	40	40	42	42	38	38	39	38	39	40	38	41	43	41	40	40	40	40	43	43
46	39	39	37	38	37	38	38	39	39	39	40	41	40	40	42	44	40	40	41	38	38	41	40	41	42	41	41	43	42	42	43	44
47	40	40	37	38	36	36	37	40	40	38	39	41	40	39	42	44	41	43	43	38	37	40	40	41	42	41	43	46	46	44	43	46
47	39	40	41	37	37	38	41	40	40	35	38	41	43	48	45	44	46	44	41	41	41	40	42	41	41	40	40	44	48	47	43	44
46	39	41	41	39	39	38	38	41	40	40	40	42	43	43	47	44	46	44	41	41	41	40	40	41	43	43	42	44	46	46	44	43
47	40	41	42	40	41	40	38	39	38	42	41	43	43	39	48	45	45	43	41	42	41	40	39	41	44	46	44	44	45	45	45	44
49	41	41	40	39	42	41	39	40	38	42	44	43	43	43	45	46	45	43	42	42	42	41	40	42	44	46	45	44	44	45	45	45
50	41	41	41	41	43	42	40	40	41	42	48	46	51	64	58	47	45	43	42	42	43	43	42	42	44	45	45	44	44	45	45	45
52	44	45	47	47	47	45	44	47	46	55	93	120	123	105	61	48	45	43	43	42	43	44	44	42	43	44	44	44	44	45	46	43
53	44	46	50	49	47	46	47	47	55	69	97	111	104	83	52	48	44	43	44	42	42	44	43	42	44	45	44	43	43	45	47	45
51	41	43	47	45	44	44	47	48	49	48	49	52	51	50	50	48	44	44	44	42	42	43	42	42	44	46	45	42	42	45	48	48
48	39	39	41	43	46	45	43	46	46	46	46	46	48	48	48	46	44	44	43	40	44	46	42	45	42	44	44	41	44	46	42	45
47	38	37	38	40	45	48	48	47	47	47	47	47	47	48	48	45	42	44	45	42	42	44	43	43	43	46	46	43	43	45	43	45
51	42	40	39	38	42	45	46	47	46	46	48	47	46	46	47	47	43	45	48	45	43	45	45	43	46	48	47	45	44	45	46	47
48	41	42	42	41	44	46	47	47	46	47	49	48	46	46	47	51	48	48	50	48	47	47	45	47	49	49	47	46	46	47	50	49
44	39	42	45	44	46	46	47	47	46	48	50	50	48	47	49	53	51	49	48	48	49	47	43	48	49	49	48	48	48	48	50	49
45	40	43	44	42	43	43	44	45	45	47	49	49	47	47	49	52	50	47	45	46	48	46	42	47	47	49	51	50	48	48	47	48
46	40	43	43	41	43	44	45	44	45	47	47	47	47	47	48	50	49	47	46	47	47	47	45	47	47	51	53	50	48	48	48	49
47	42	44	44	42	43	43	44	46	48	49	49	48	49	49	50	50	49	48	49	50	48	48	50	48	48	53	55	49	47	50	50	51
48	42	43	43	42	43	46	51	49	50	54	52	50	51	49	50	50	49	49	50	49	48	49	52	53	50	48	49	50	50	51	53	53

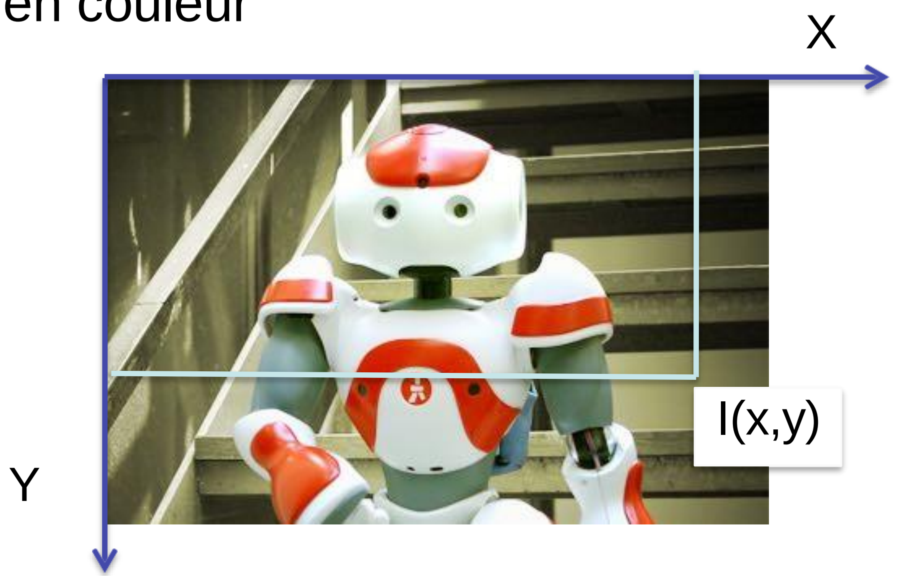
Représentation des images

Fonction

– I(x,y) : valeur d'un pixel

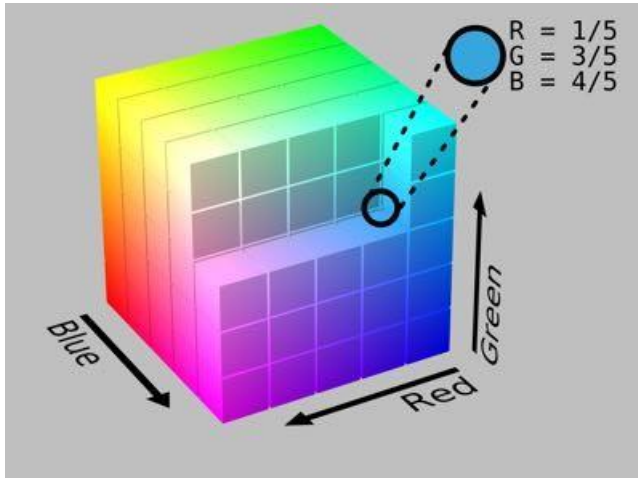
dans R en monochrome

dans R^3 en couleur

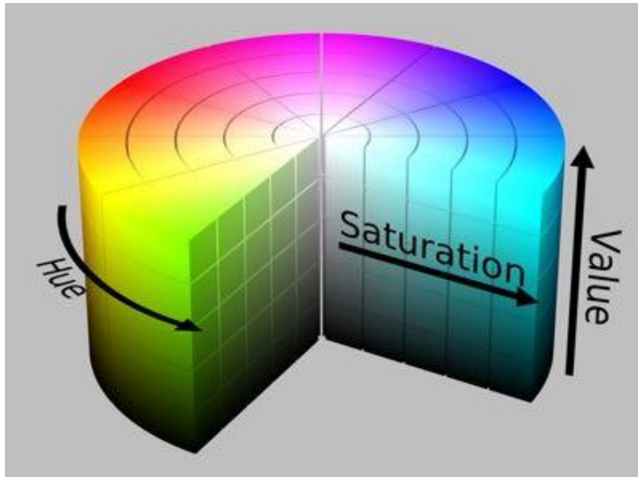


Représentation des images

Espaces colorimétriques

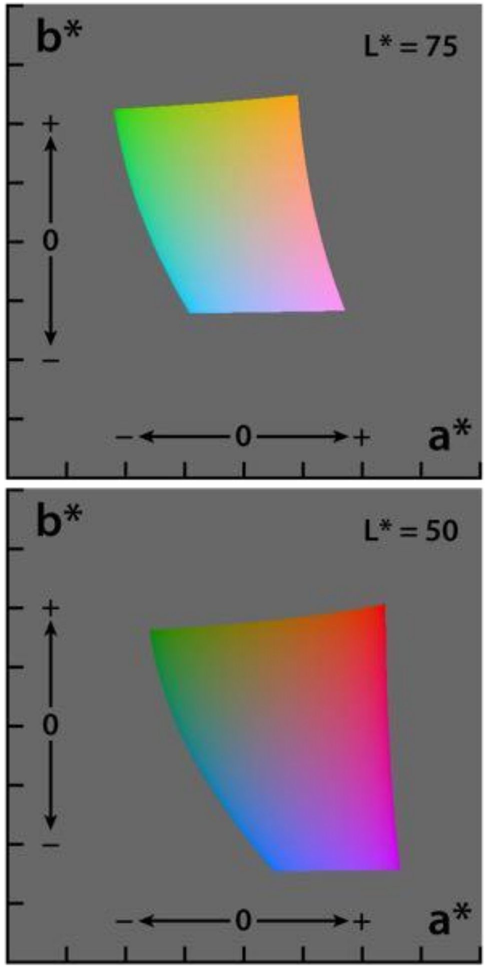


RGB



HSV

Lab



Représentation des images

Variations

- I(x,y) ➔ I(x,y) + α : luminosité globale
- I(x,y) ➔ λI(x,y) : changement de contraste
- I(x,y) ➔ I(x+a,y+b) : translation
- I ➔ A.I+b : transformation affine

➔

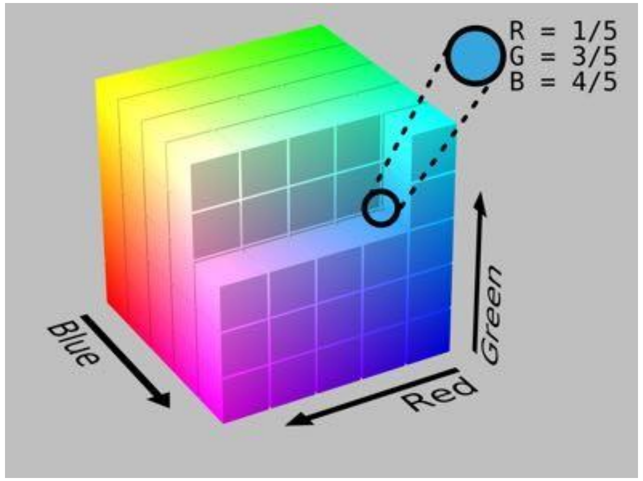
Traitements de base

Filtrage et amélioration

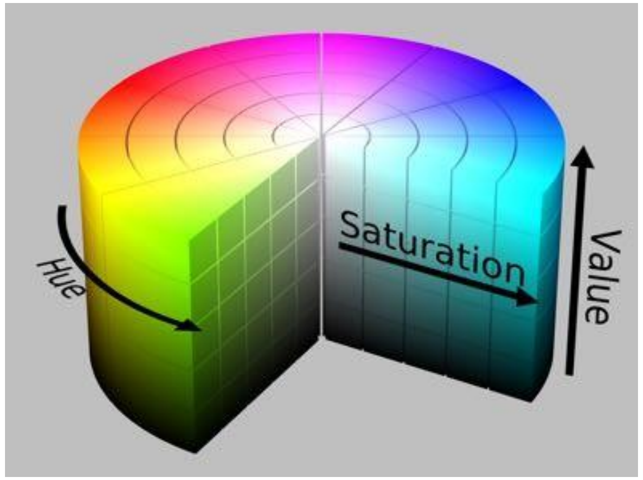
- Changement de color space
- Thresholding
- Egalisation d’histogramme
- Flou
- Morphologie mathématique

Représentation des images

Espaces colorimétriques

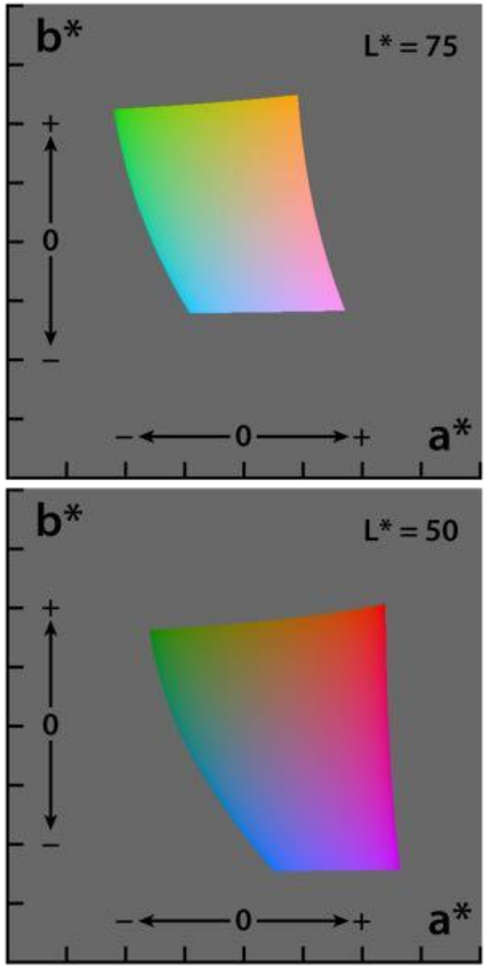


RGB

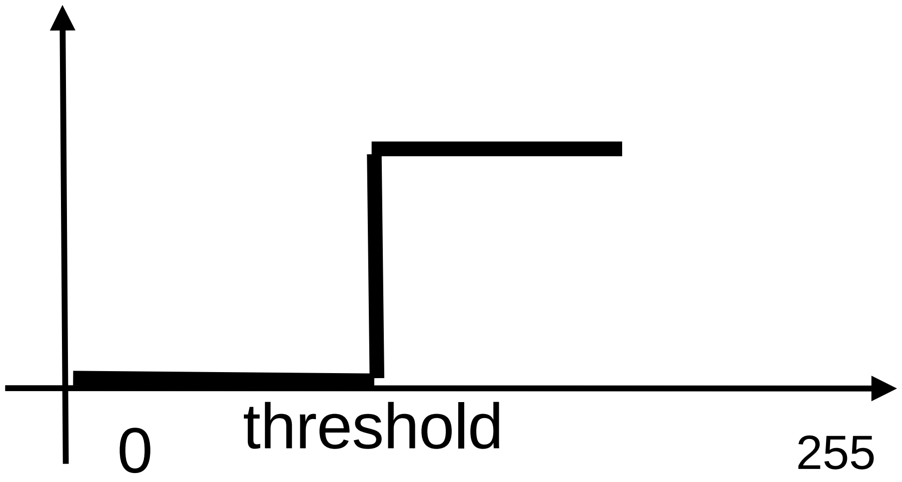


HSV

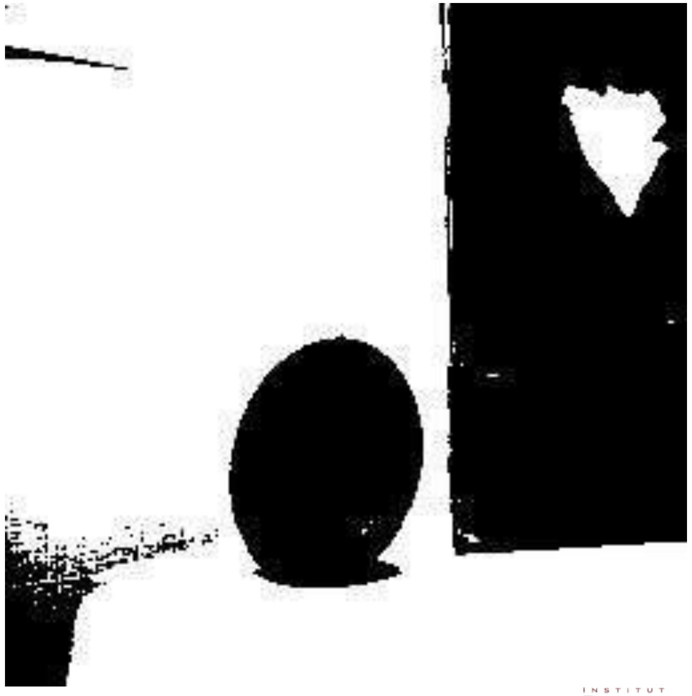
Lab



Binarisation



Input image



T. Chateau

Detections de contours

T.

Detections de contours

En 2D

– Masques de Sobel

Dérivée selon x : I_x

$$H_1 = \frac{1}{8}\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

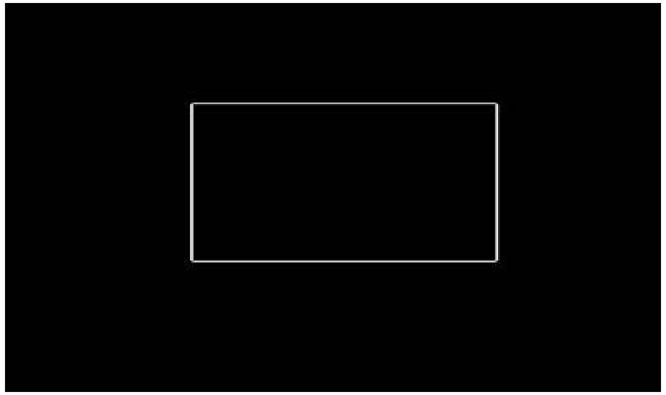
Dérivée selon y : I_y

$$H_2 = \frac{1}{8}\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

Detections de contours

Estimation du gradient

– Norme: $\mathrm{sqrt}(I_x^2+I_y^2)$ Direction: $\mathrm{atan2}(I_y,I_x)$

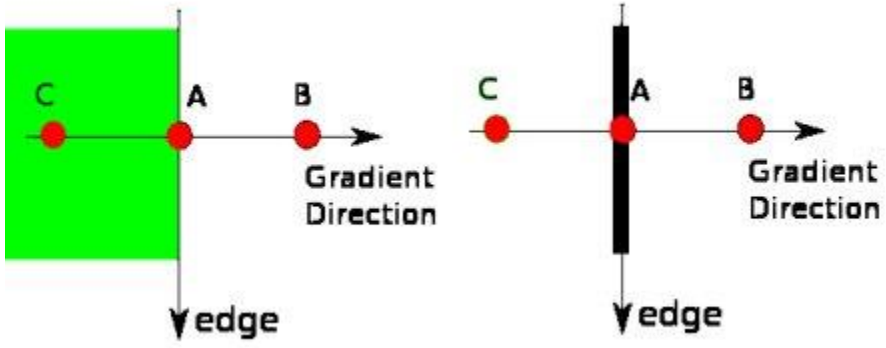


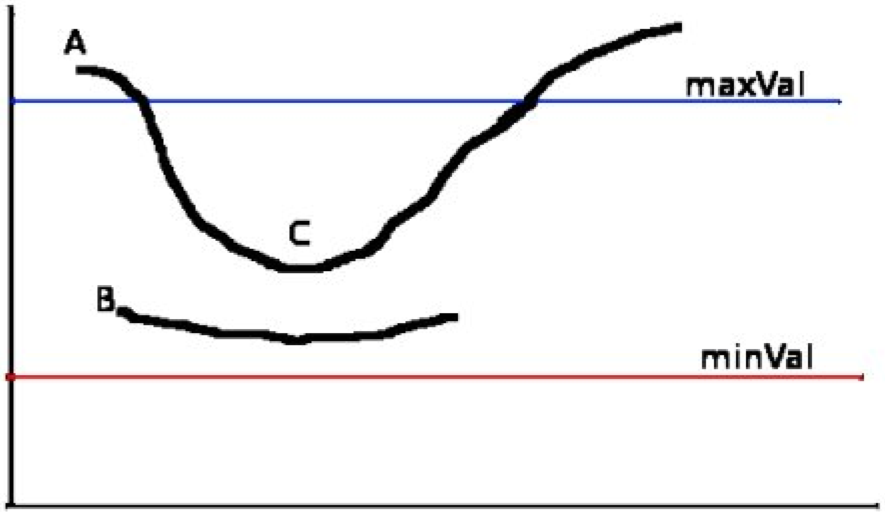
Détections de contours

- Définition
  - Extremas du gradient dans la direction du gradient
  - Zéros du Laplacien
- Filtrage du gradient
  - Suppression des non maxima dans la direction du gradient

Détections de contours : Méthode de Canny

- Suppression du bruit
- Calcul gradient Sobel
- Non maximum suppression
- Seuillage à hysteresis

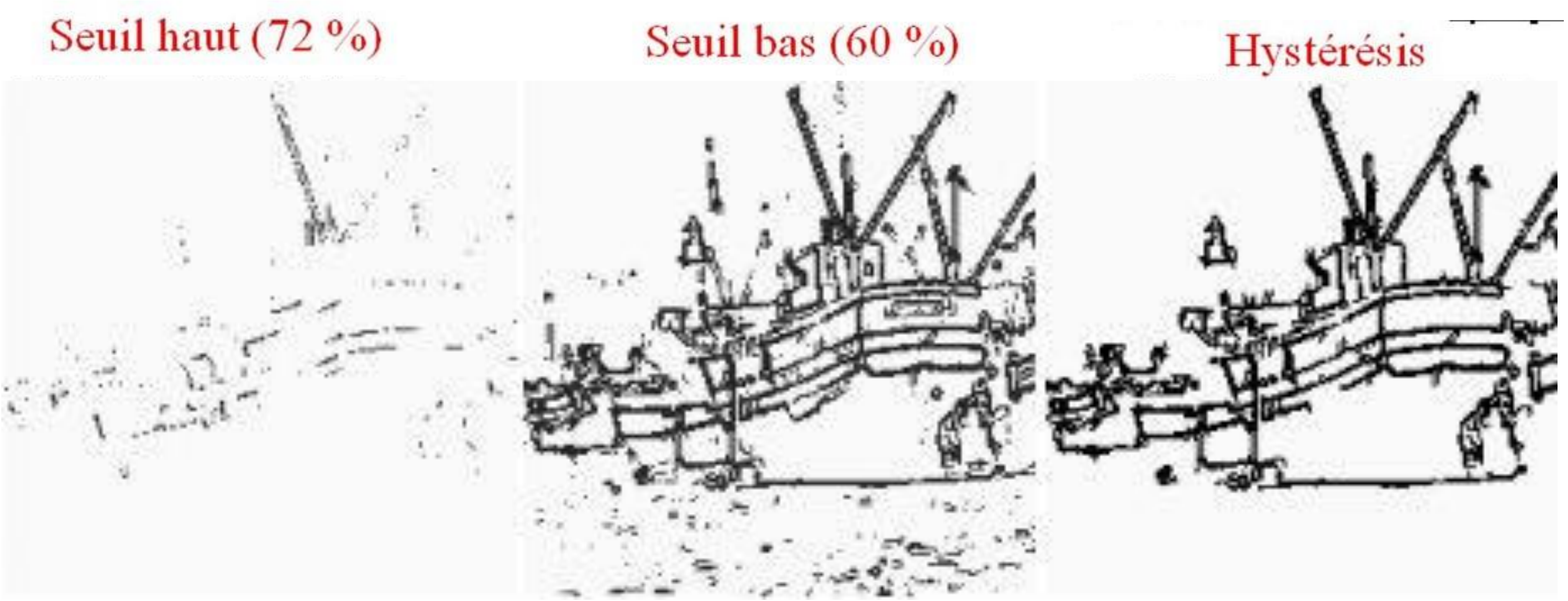




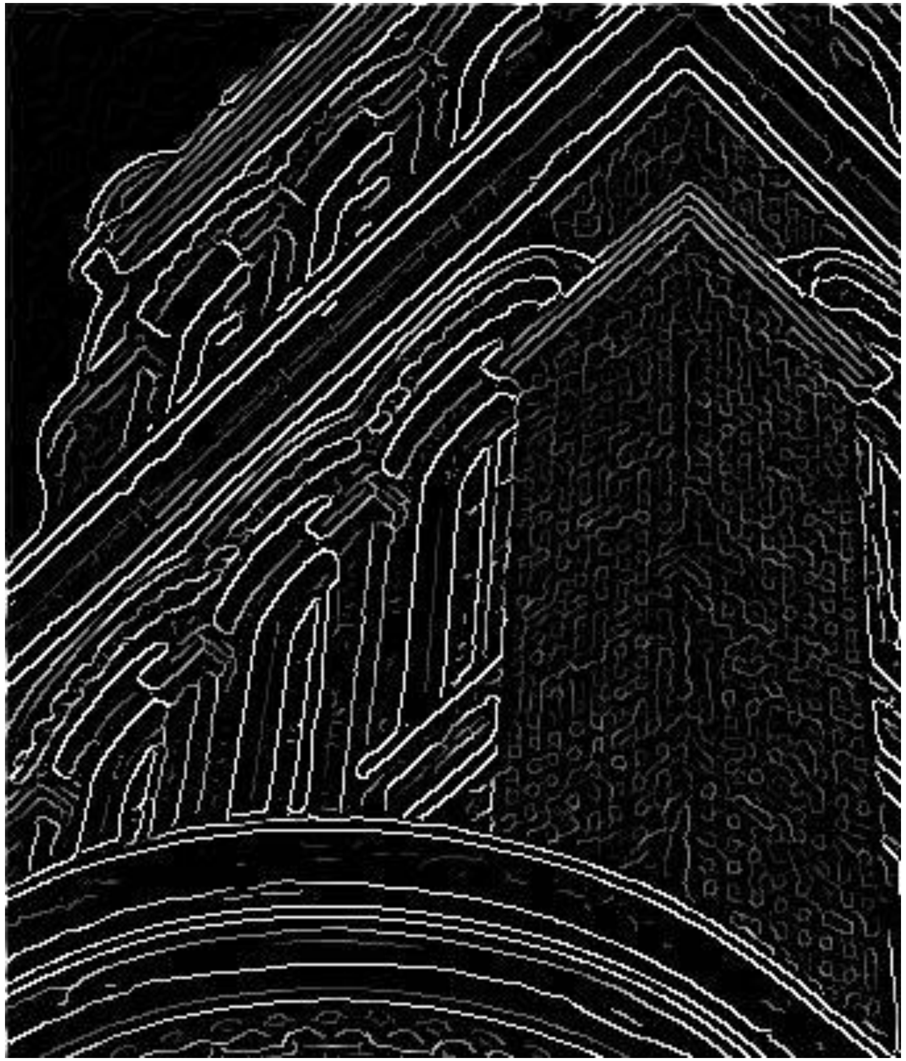
Détections de contours : Méthode de Canny

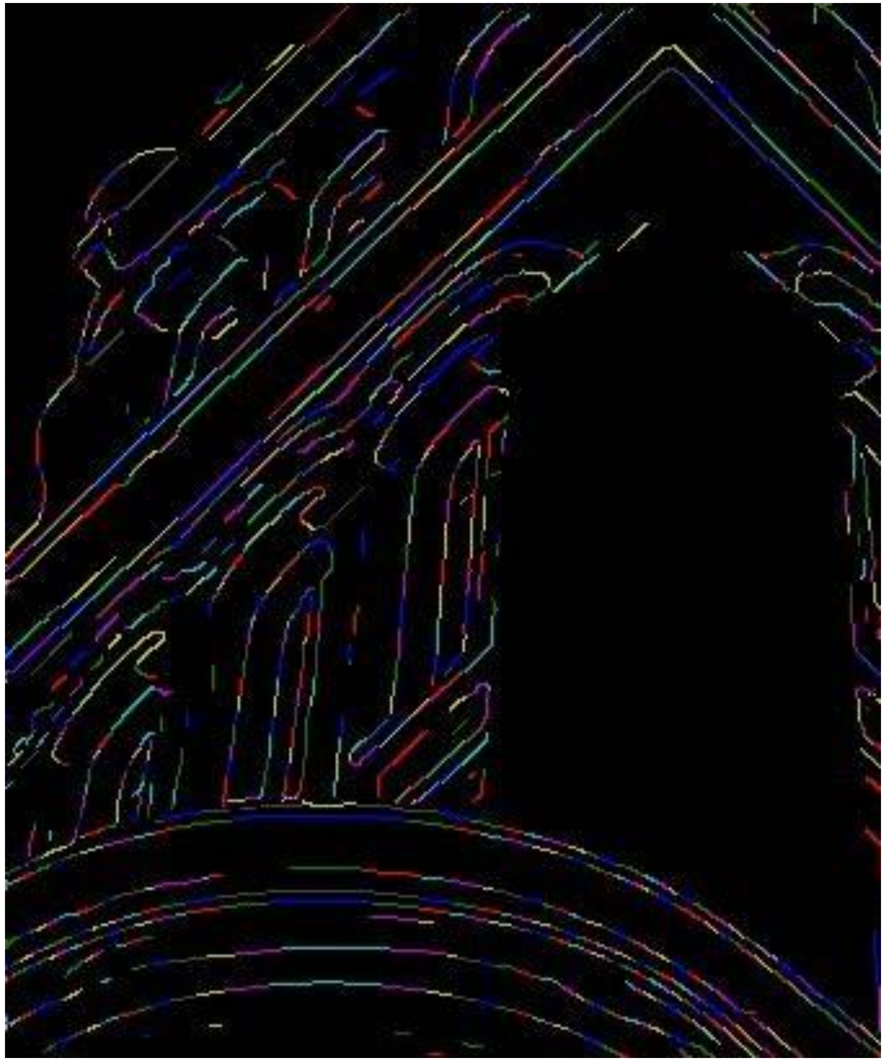
Seuillage par hystérésis

– Gradients > s_1, connectés à des gradients > s_2

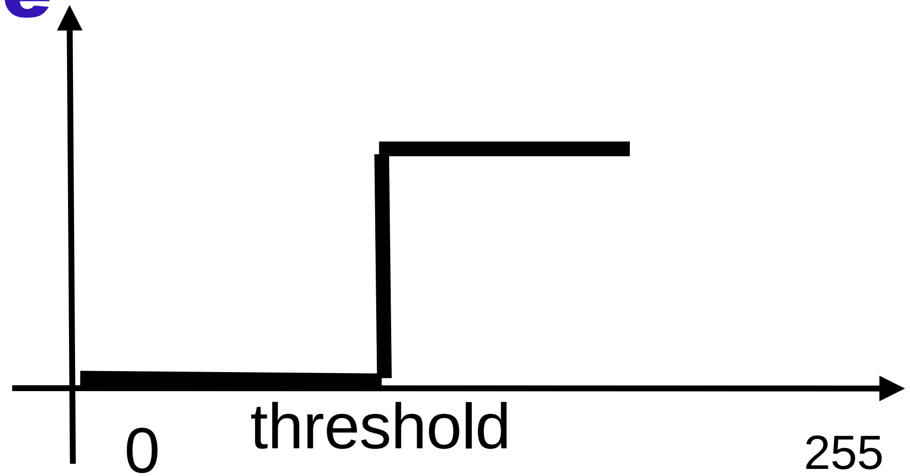


Détections de contours : Méthode de Canny



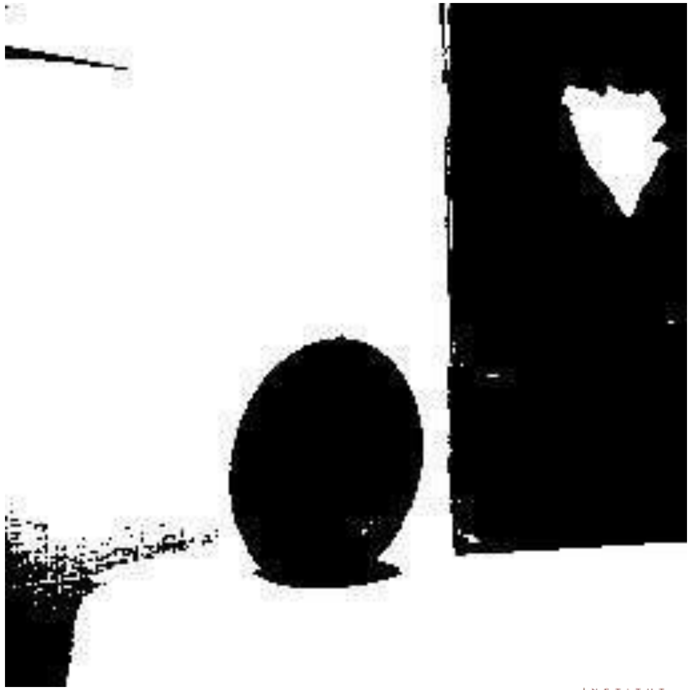


Egalisation d'histogramme



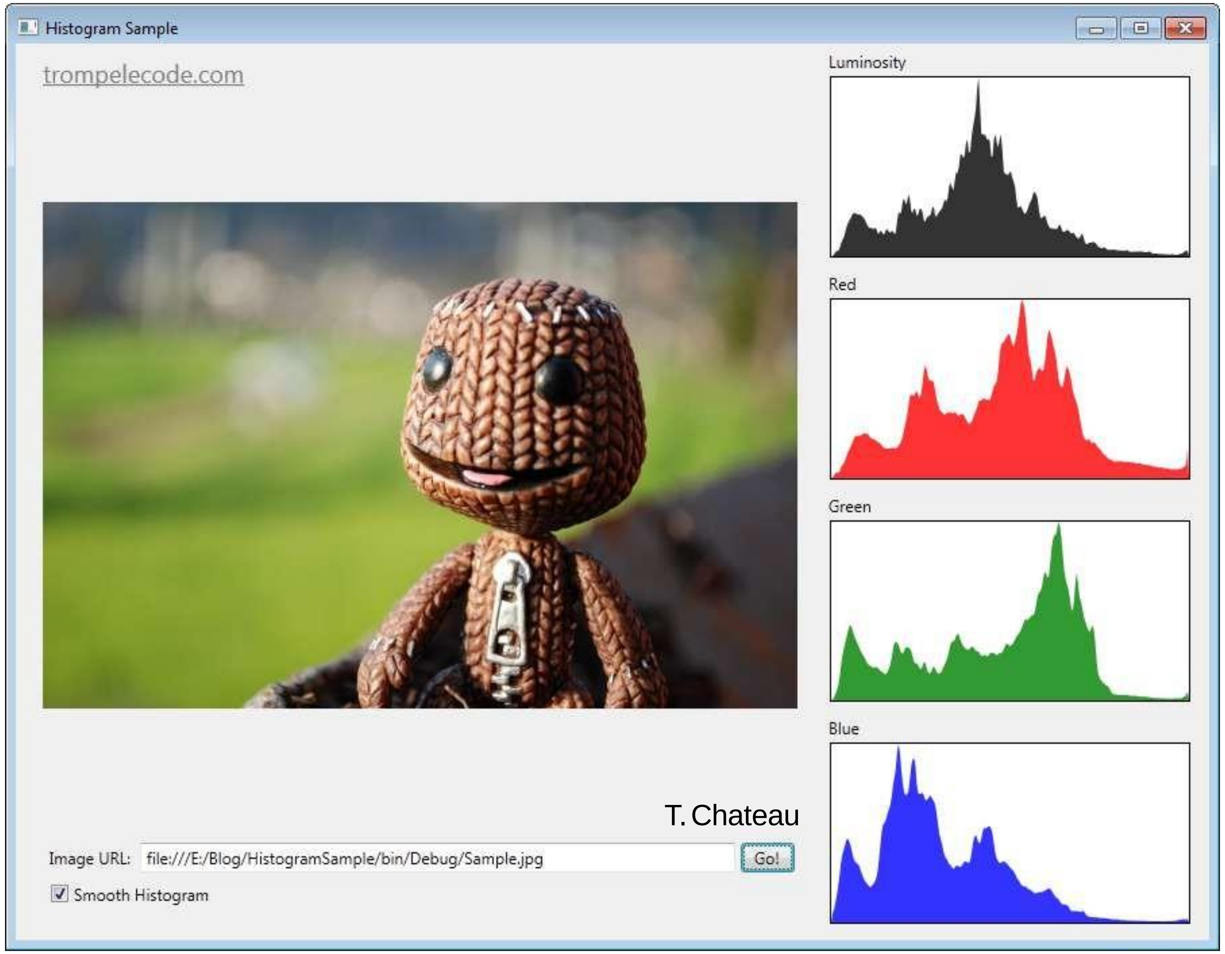
Input image

T. Chateau



Egalisation d'histogramme

$$H(x) = Card\{\mathbf{p} : I(\mathbf{p}) = x\}$$

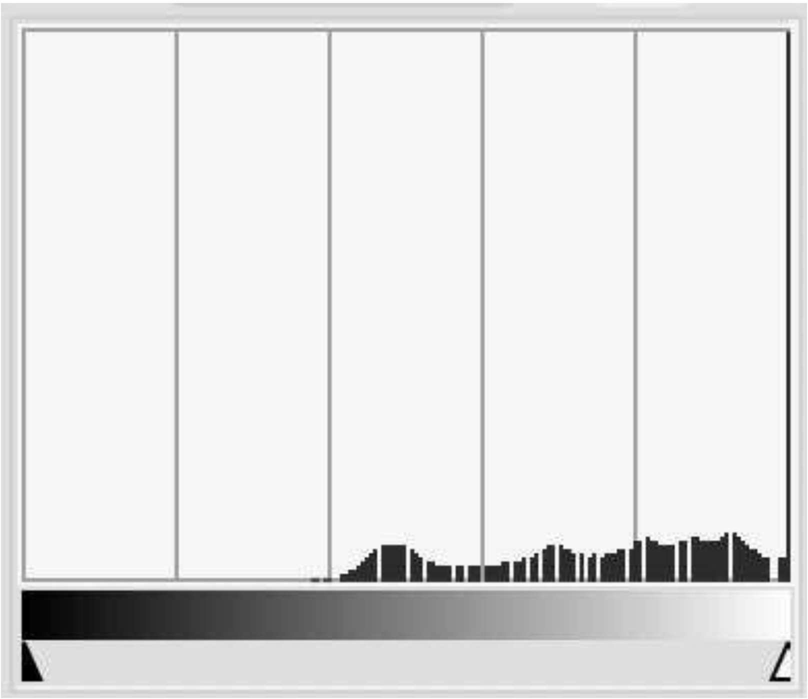


T. Chateau

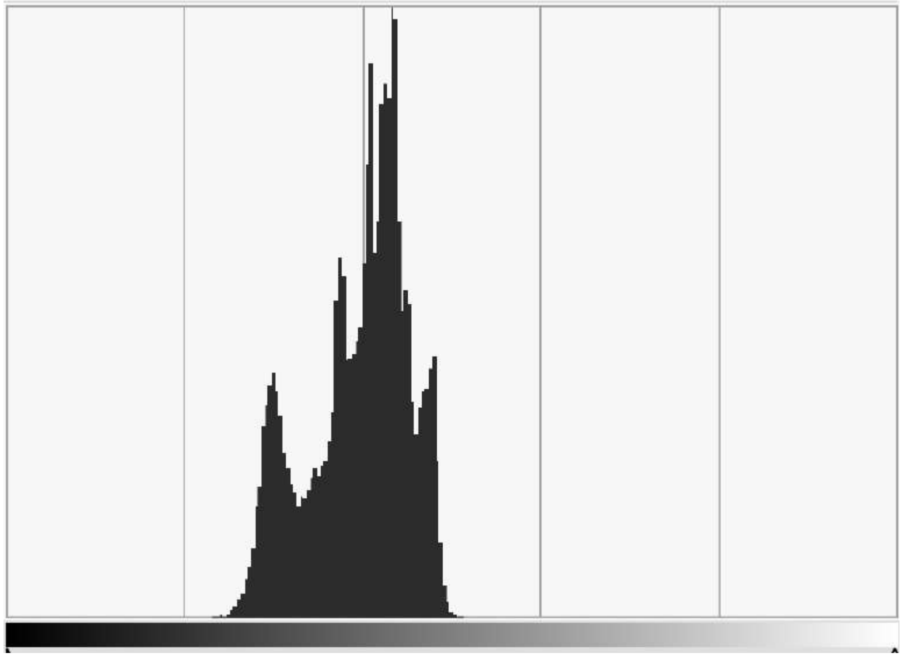
Histogramme

T. Chateau

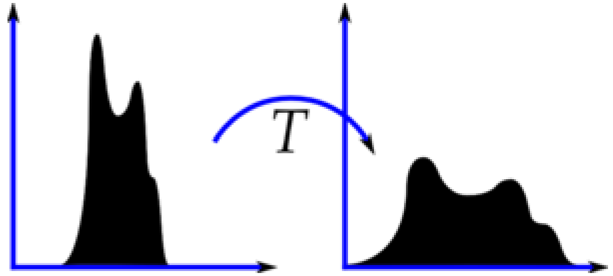
Histogramme

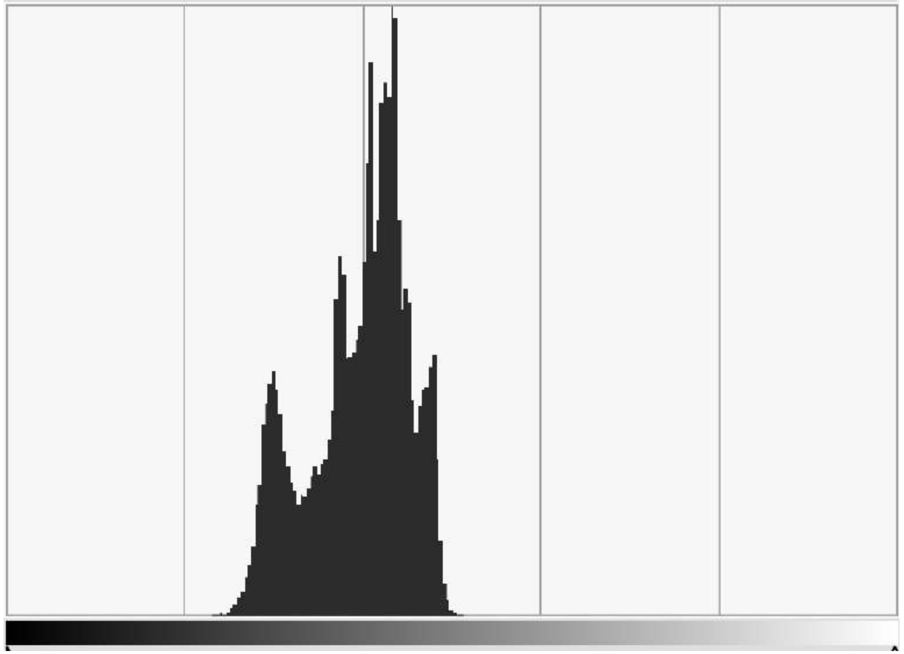


Histogramme

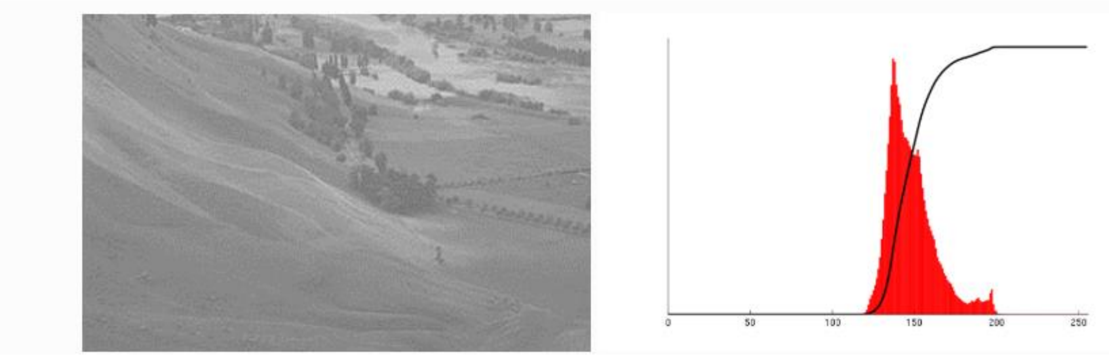


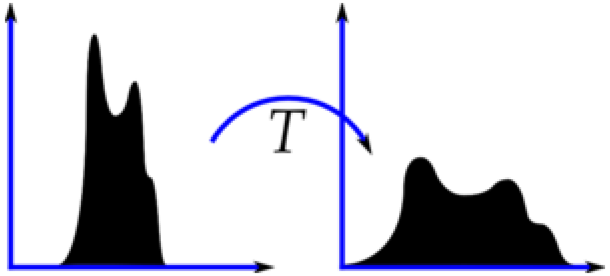
Egalisation d'histogramme



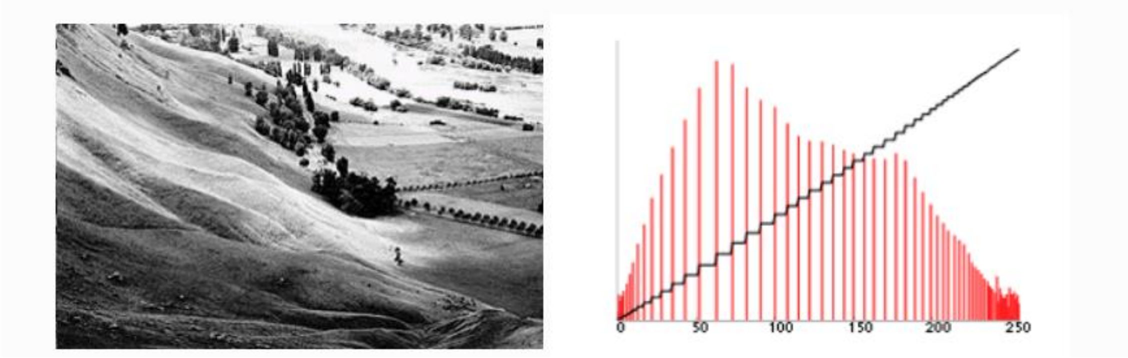


Egalisation d’histogramme

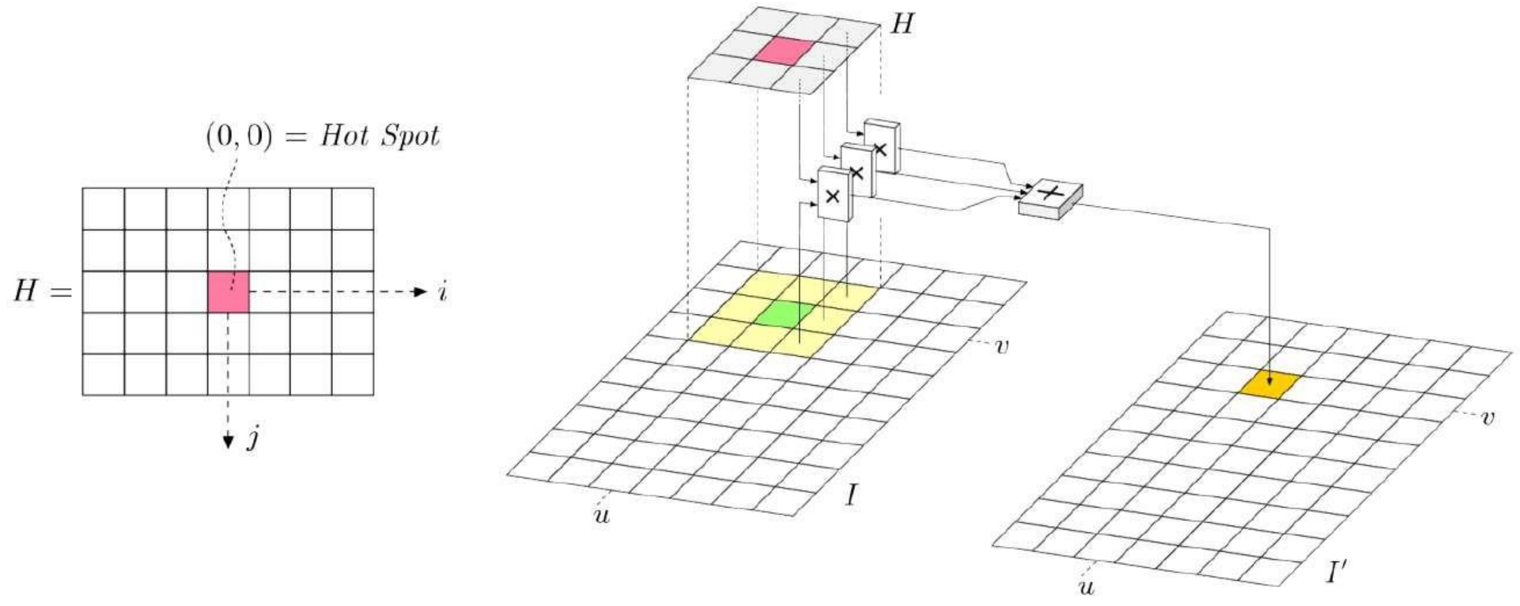




$$\forall v \in [0..n], eg(v) = \frac{V_{max} - V_{min}}{N} C_f(v) + V_{min}$$



Filtrage par convolution



10	5	3
4	5	1
1	1	7

Local image data

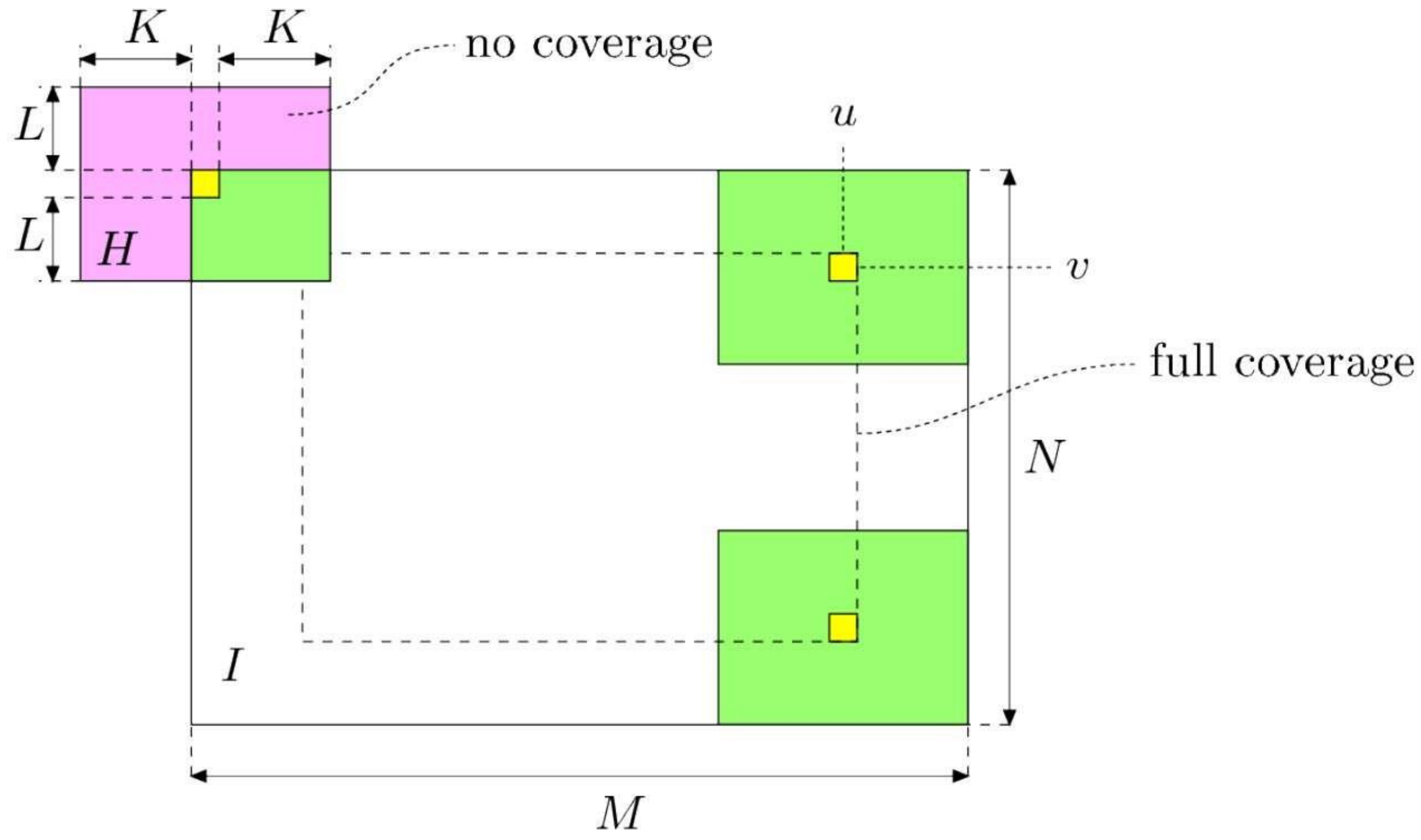
0	0	0
0	0.5	0
0	1	0.5

kernel

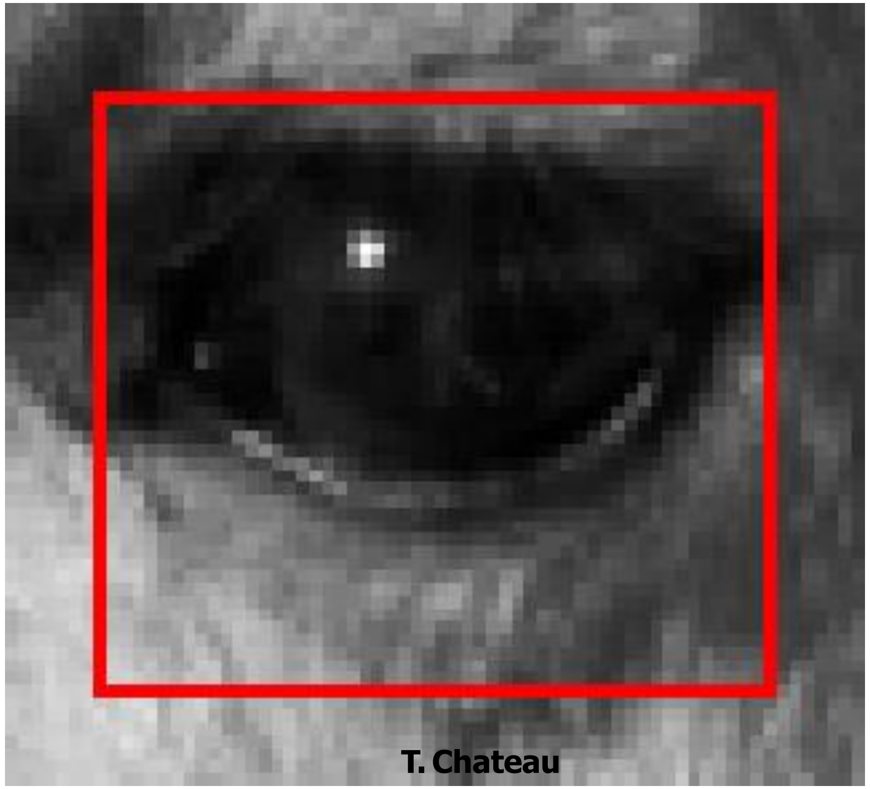
	7	

Modified image data

How to cope with image boundary?



Filtrage par convolution : Crop

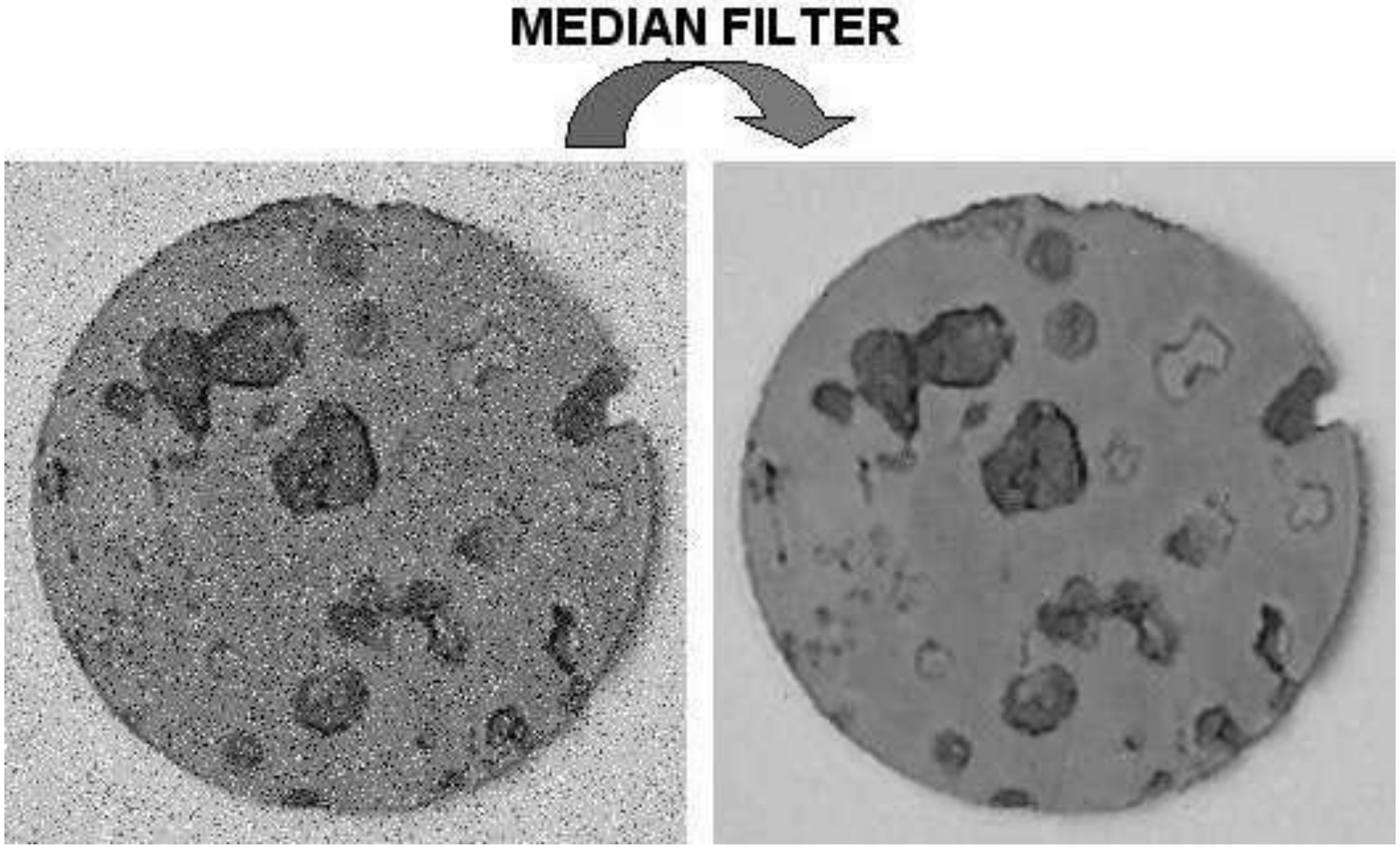


Filtrage par convolution : PAD

Filtrage médian (Bruit poivre et sel)

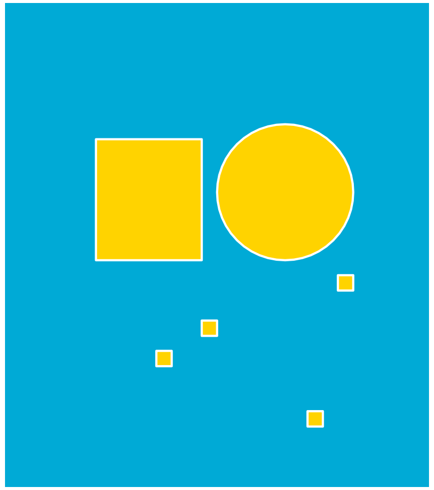
- Robuste
- Non-lineaire
- Chaque pixel est modifié selon la valeur médiane de son voisinage

Filtrage par convolution : Filtre médian

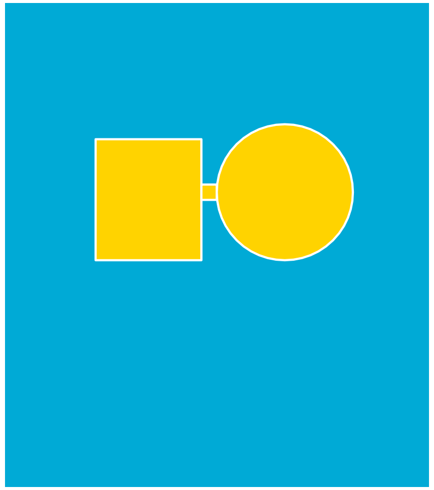


Binary image filtering: Mathematical morphology

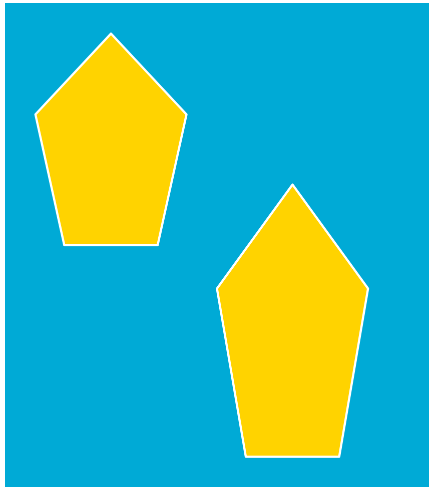
What can we do with MM ?



Remove noise



separate shapes



compare shapes

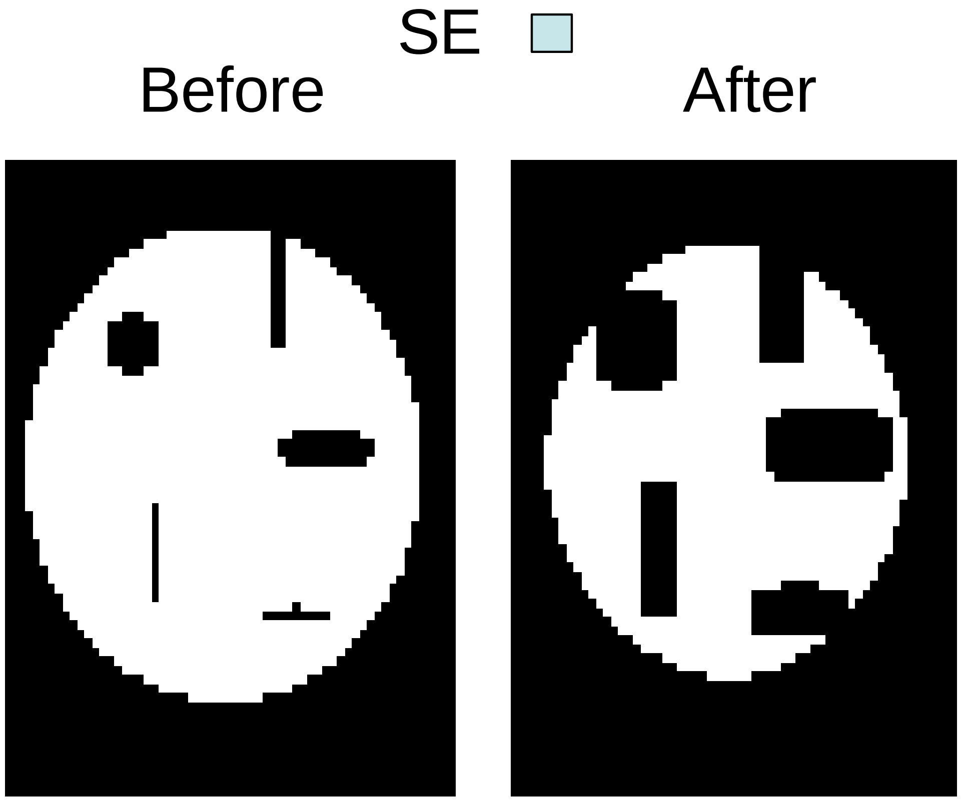
Binary image filtering: Mathematical morphology

- Main idea: probe an image with a simple, pre- defined shape, drawing conclusions on how this shape fits or misses the shapes in the image.

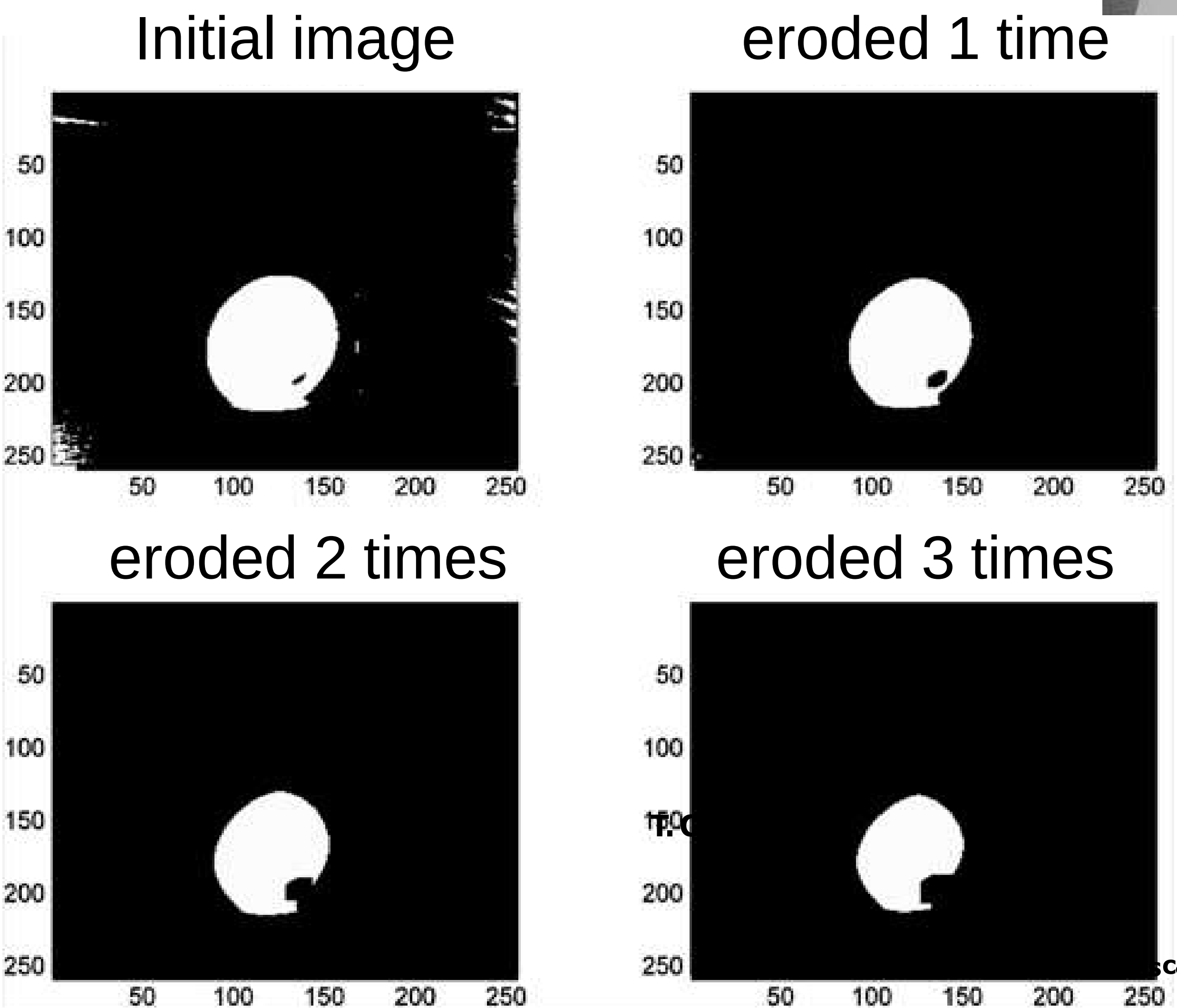
- This simple "probe" is called structuring element, and is itself a binary image (i.e., a subset of the space or grid).

Binary image filtering: Mathematical morphology

Basic operators: erosion

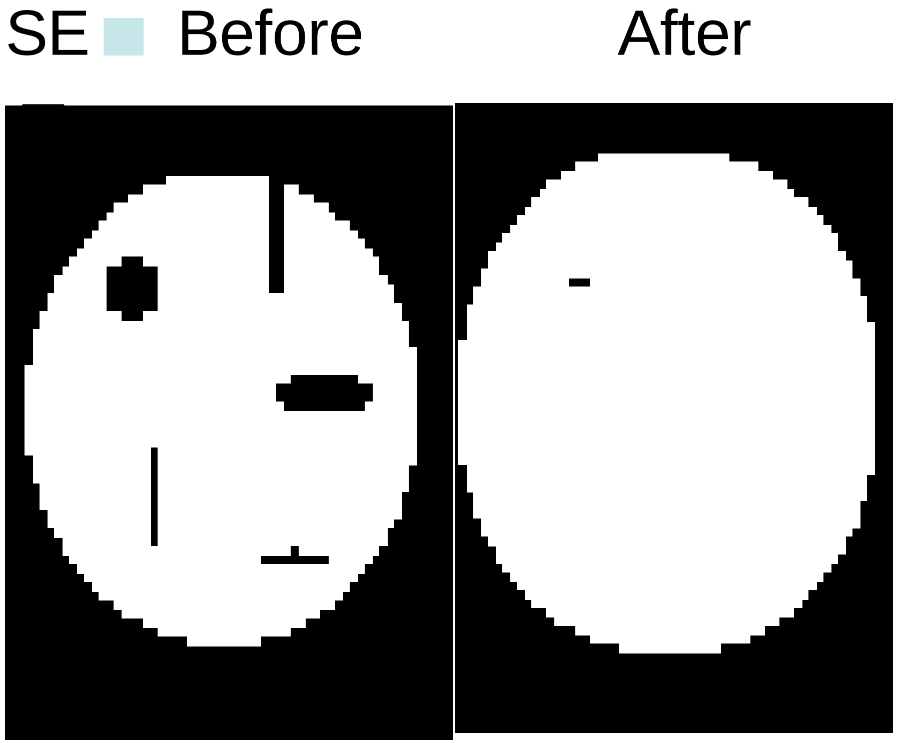


Binary image filtering: Mathematical morphology



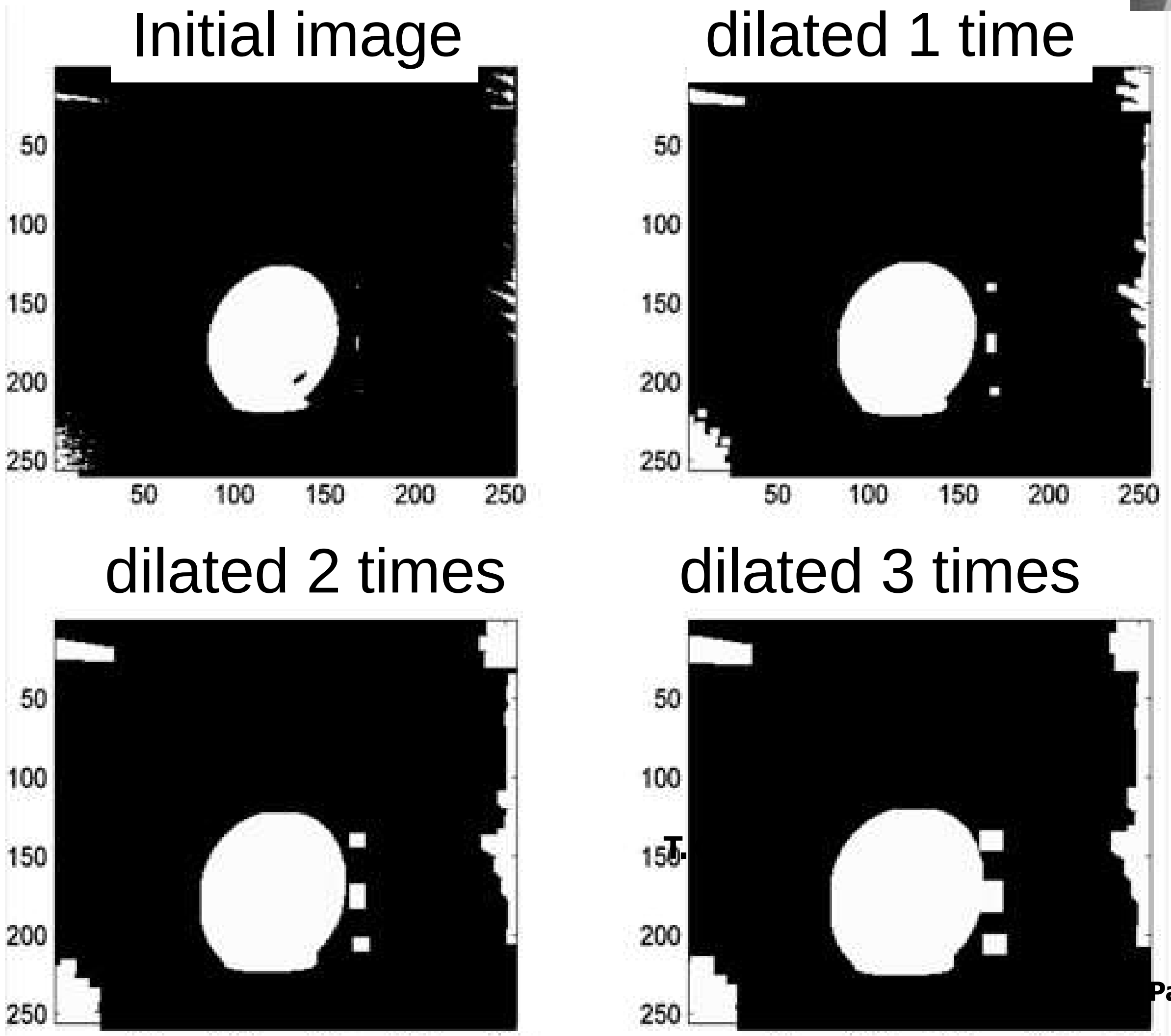
Binary image filtering: Mathematical morphology

Basic operators: dilation



Binary image filtering: Mathematical morphology

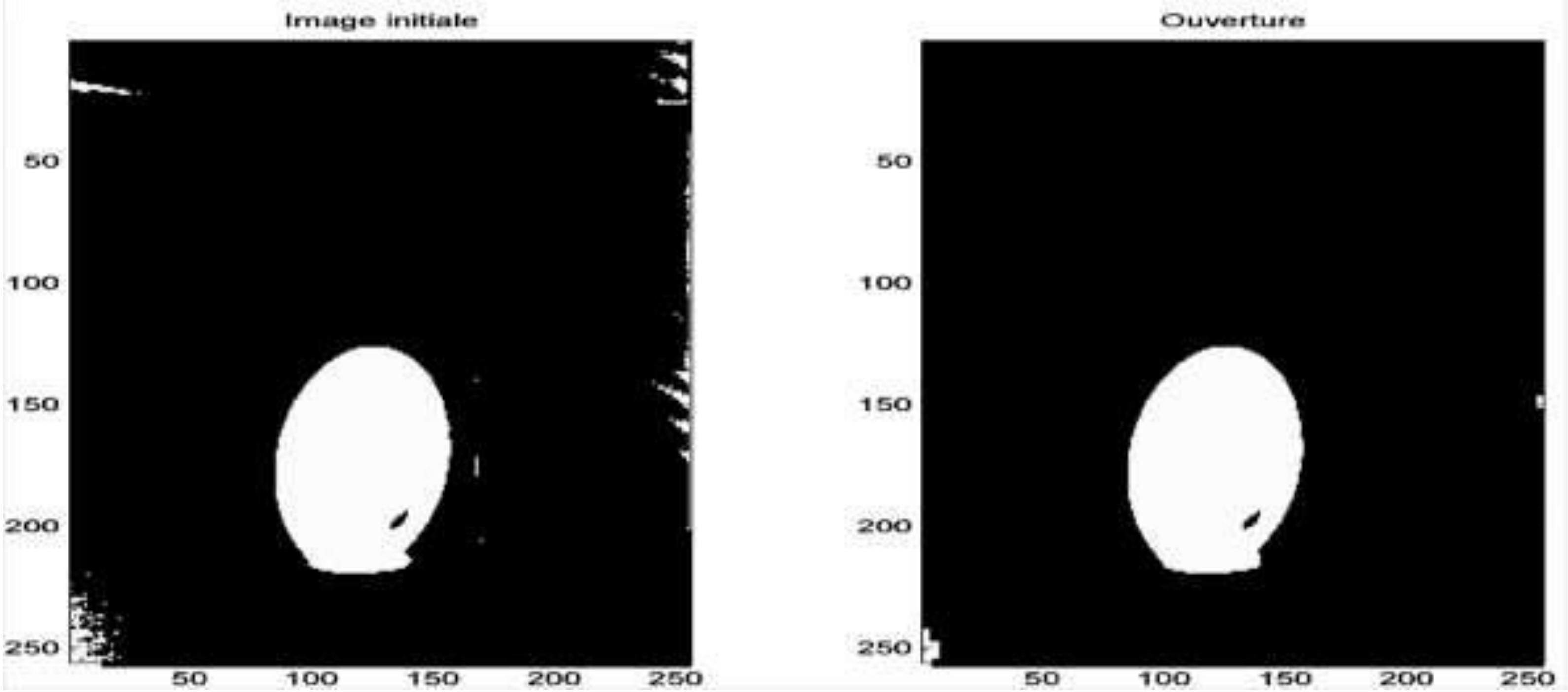
Example



Binary image filtering: Mathematical morphology

Basic operators:

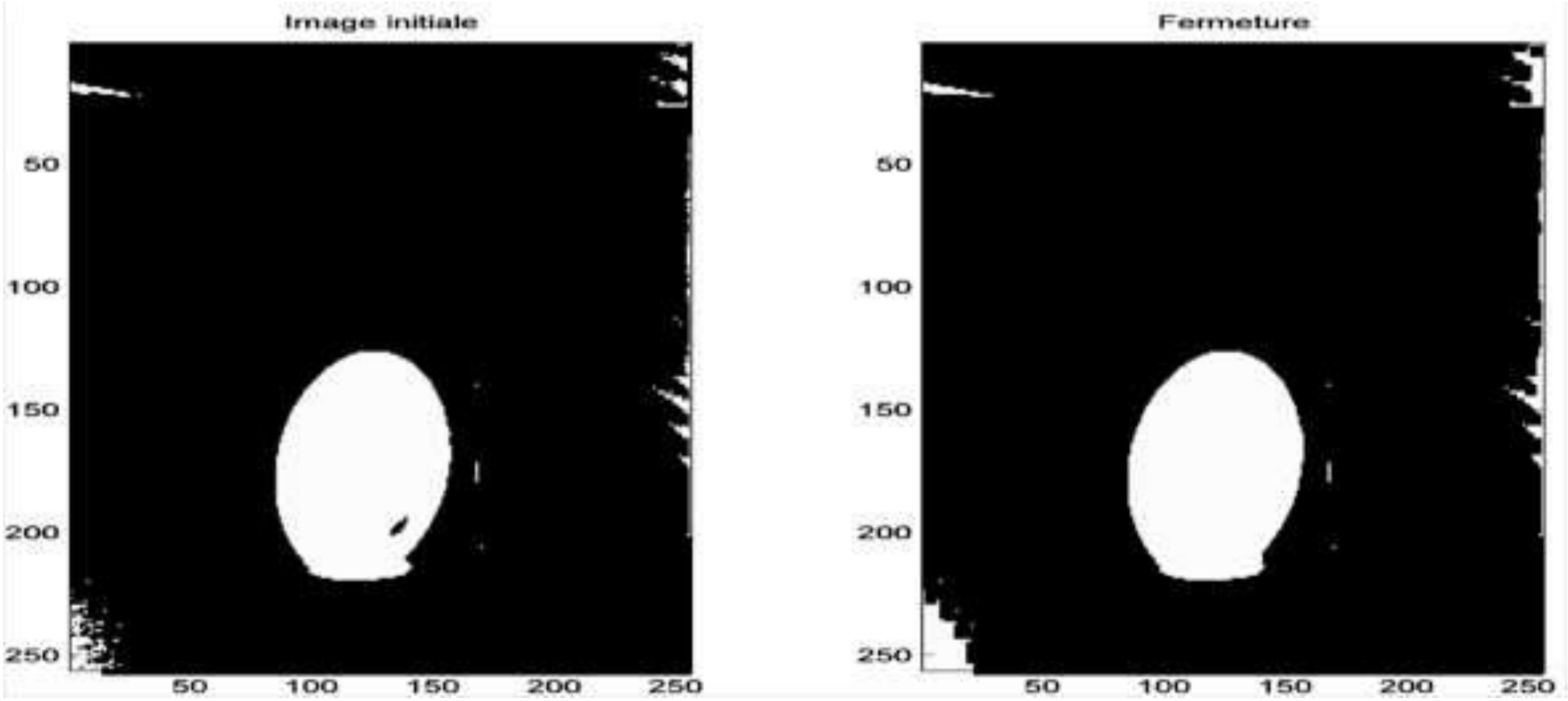
ouverture



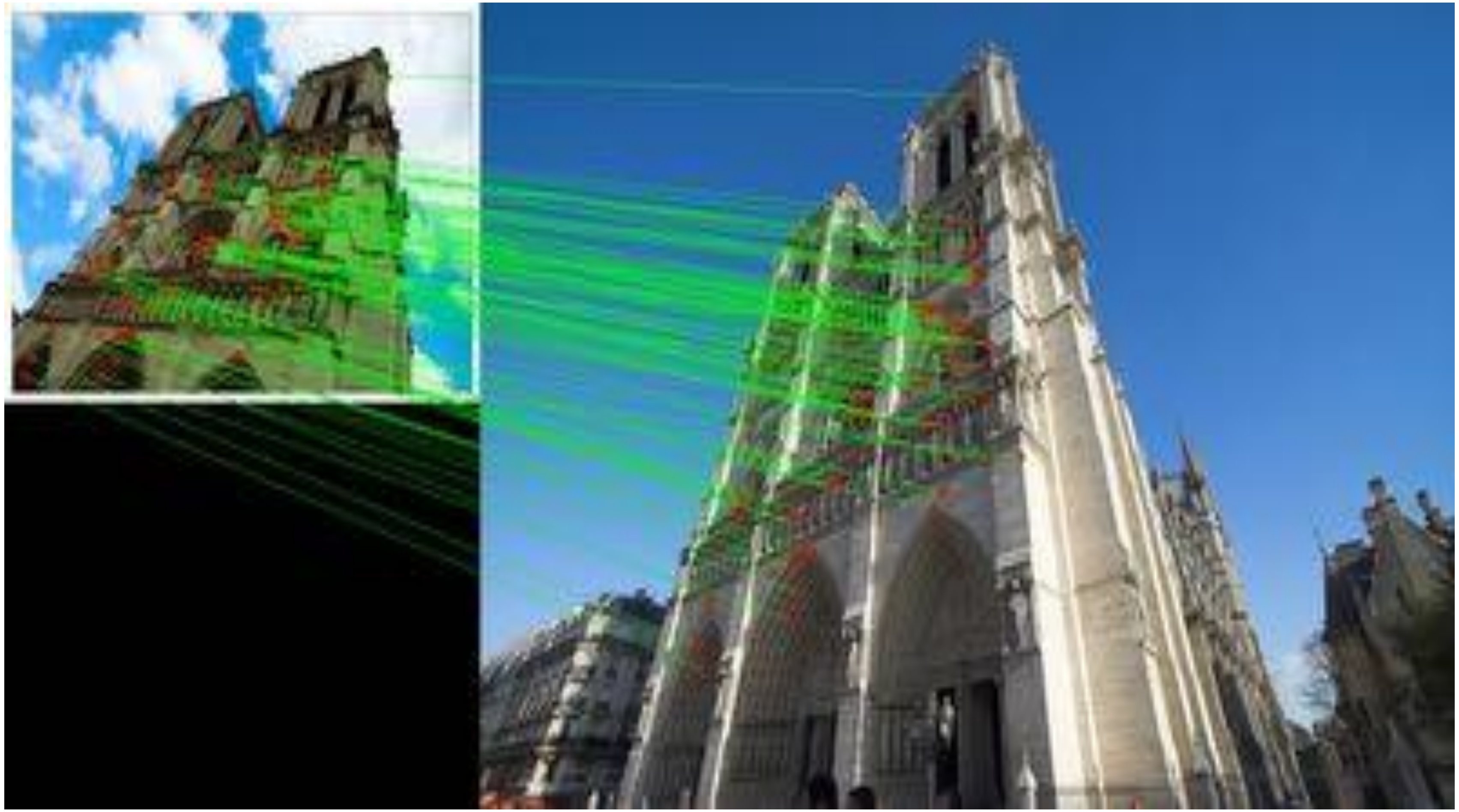
Binary image filtering: Mathematical morphology

Basic operators:

fermeture



Un concept important : La représentation



- On veut raisonner sur une image
- Il nous faut trouver des descripteurs
- Pour avoir une représentation
  - Identifiables dans plusieurs images
  - Reconnaissables / mouvement caméra
  - Robustes aux changements de d’éclairage
  - Robustes aux déformations liées au mouvement

Comparaison de points

Trouver le point le plus similaire

– Hypothèse de décalage en translation

Comment comparer $I_1(x_1,y_1)$ et $I_2(x_2,y_2)$?

Robustesse changement luminosité / contraste ?

Comparaison de points

Sum of Squared Differences

$$SSD(I_1, x_1, y_1, I_2, x_2, y_2) = \sum_{i=-n}^{n} \sum_{j=-p}^{p} (I_1(x_1+i, y_1+j) - I_2(x_2+i, y_2+j))^2$$

Cross corrélation

- Invariance / luminosité et contraste

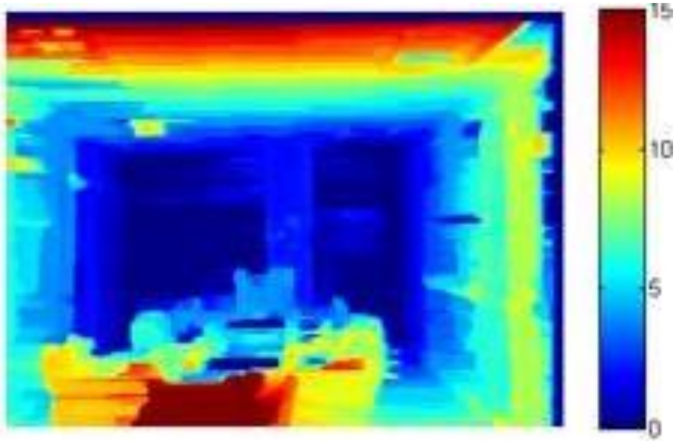
$$NCC(I_1, x_1, y_1, I_2, x_2, y_2) = \frac{\sum_{i,j}(I_1(x_1+i, y_1+j) - \overline{I_1(x_1, y_1)})(I_2(x_2+i, y_2+j) - \overline{I_2(x_2, y_2)})}{\sqrt{\sum_{i,j}(I_1(x_1+i, y_1+j) - \overline{I_1(x_1, y_1)})^2 \sum_{i,j}(I_2(x_2+i, y_2+j) - \overline{I_2(x_2, y_2)})^2}}$$

Application : stéréo-vision

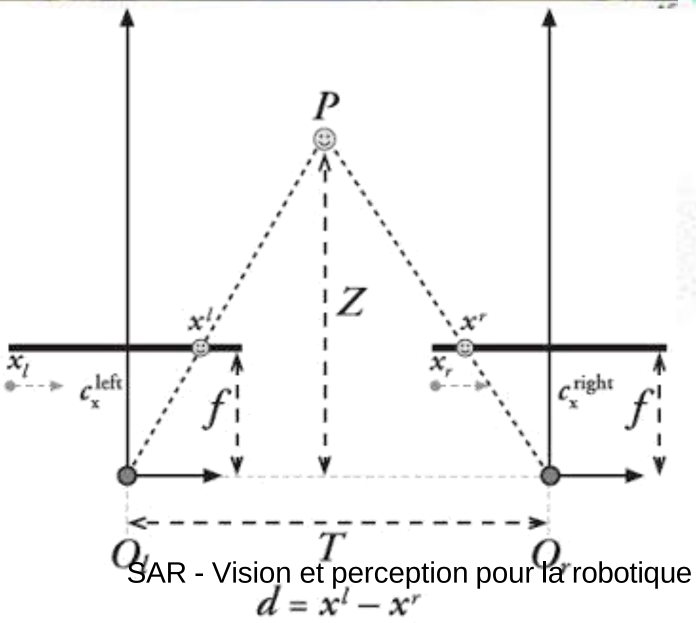
Gauche

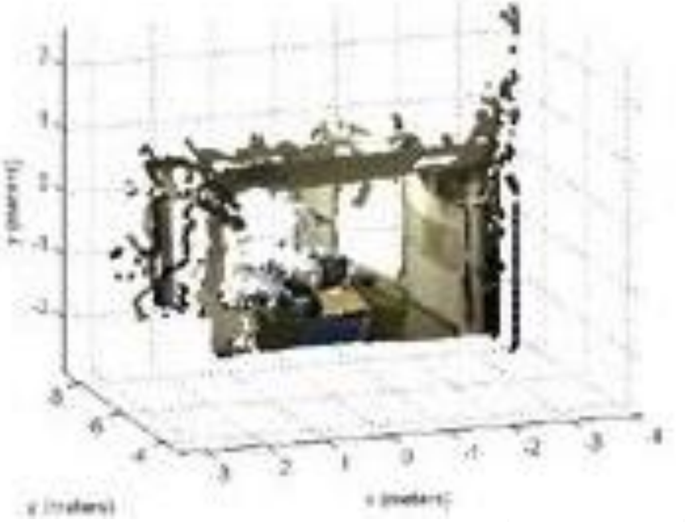
Droite

Disparité



Validité ?





Comparaison de points

Trouver le point le plus similaire

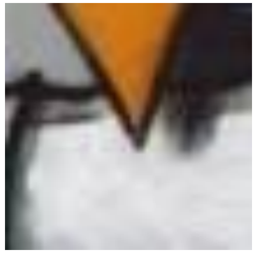
- Translation -> SSD

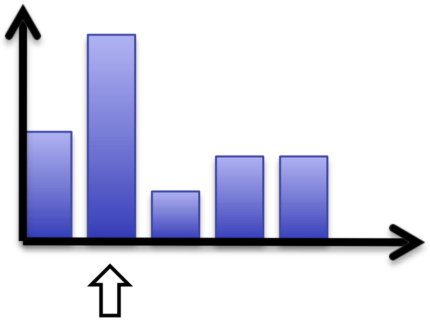
Comment reconnaître un point après rotation ?

Descripteurs SIFT [Lowe99]

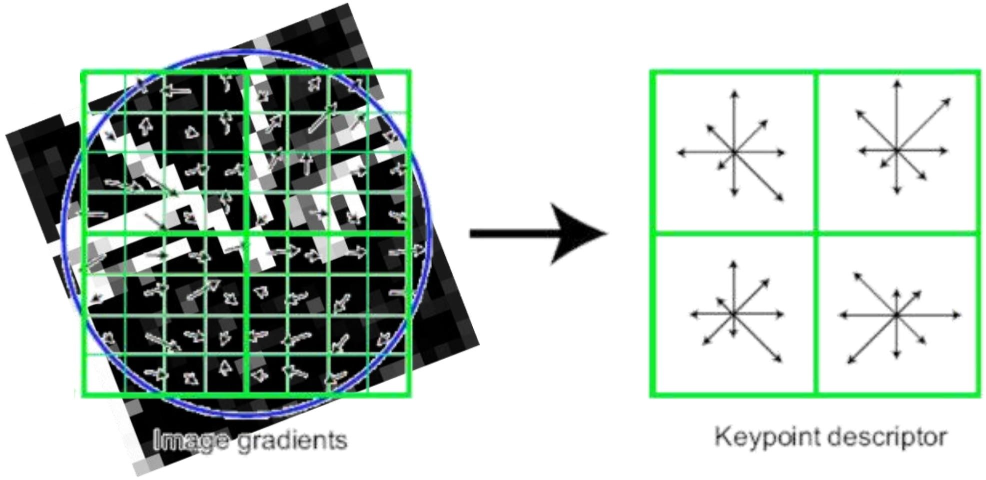
Histogramme d'orientation du gradient

– Orientation de référence





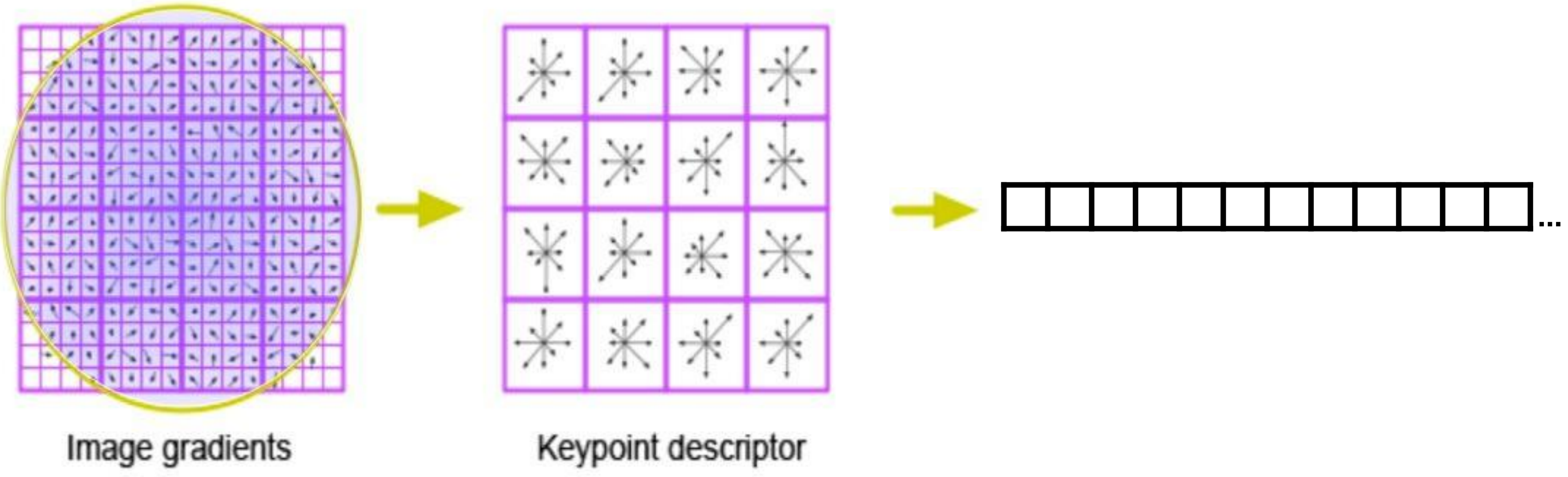
– Histogramme des orientations



Descripteurs SIFT [Lowe99]

Histogramme d’orientation du gradient

- 4x4 fenêtres
- Histogrammes sur 8 directions
- Pondération gaussienne autour du centre
- Dimension 128



Invariance au changement d'échelle

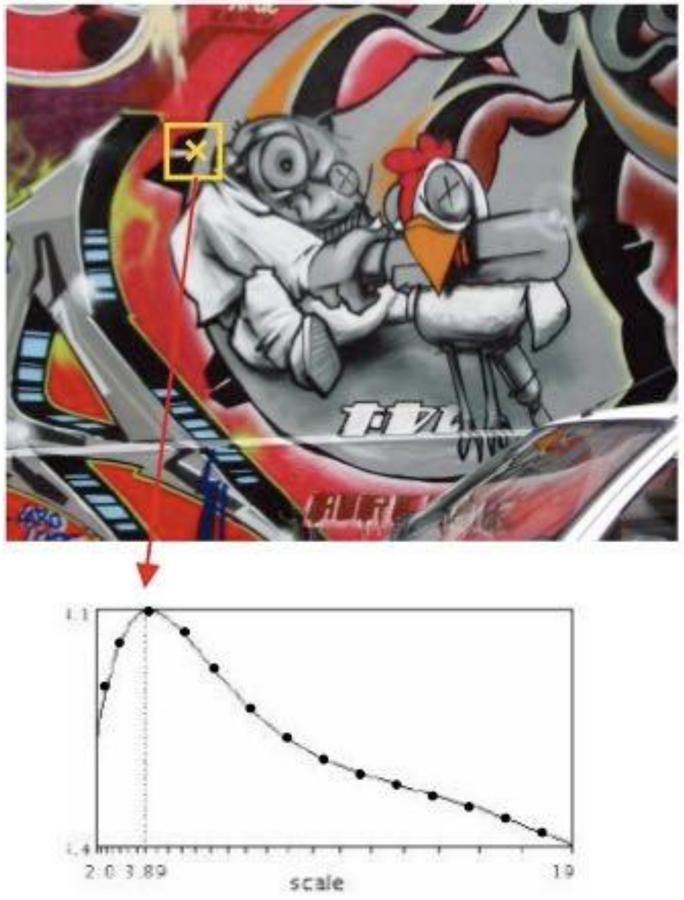
Trouver des point similaires quelque soit l'échelle

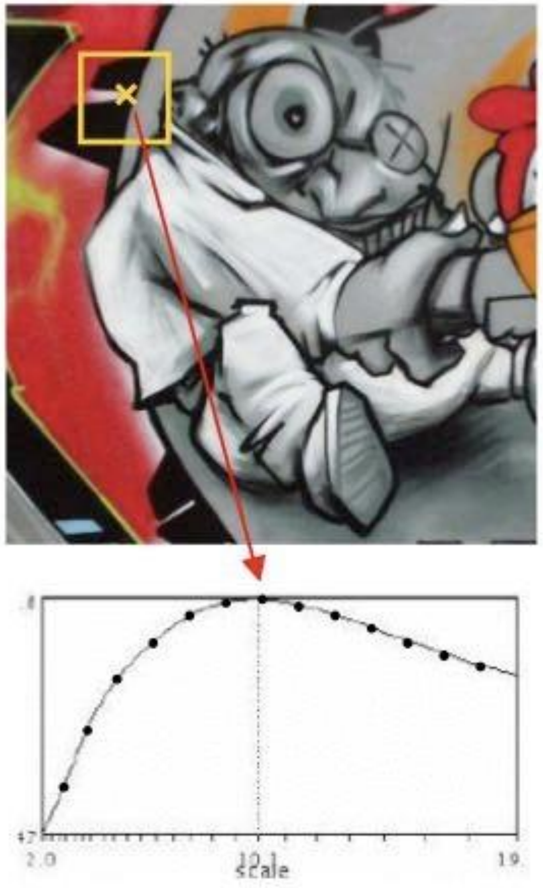
- Translation -> SSD
- Rotation -> orientation de référence

Comment comparer des points à des échelles différentes ?

Détection d’échelle de référence

Analyse par fonction de signature

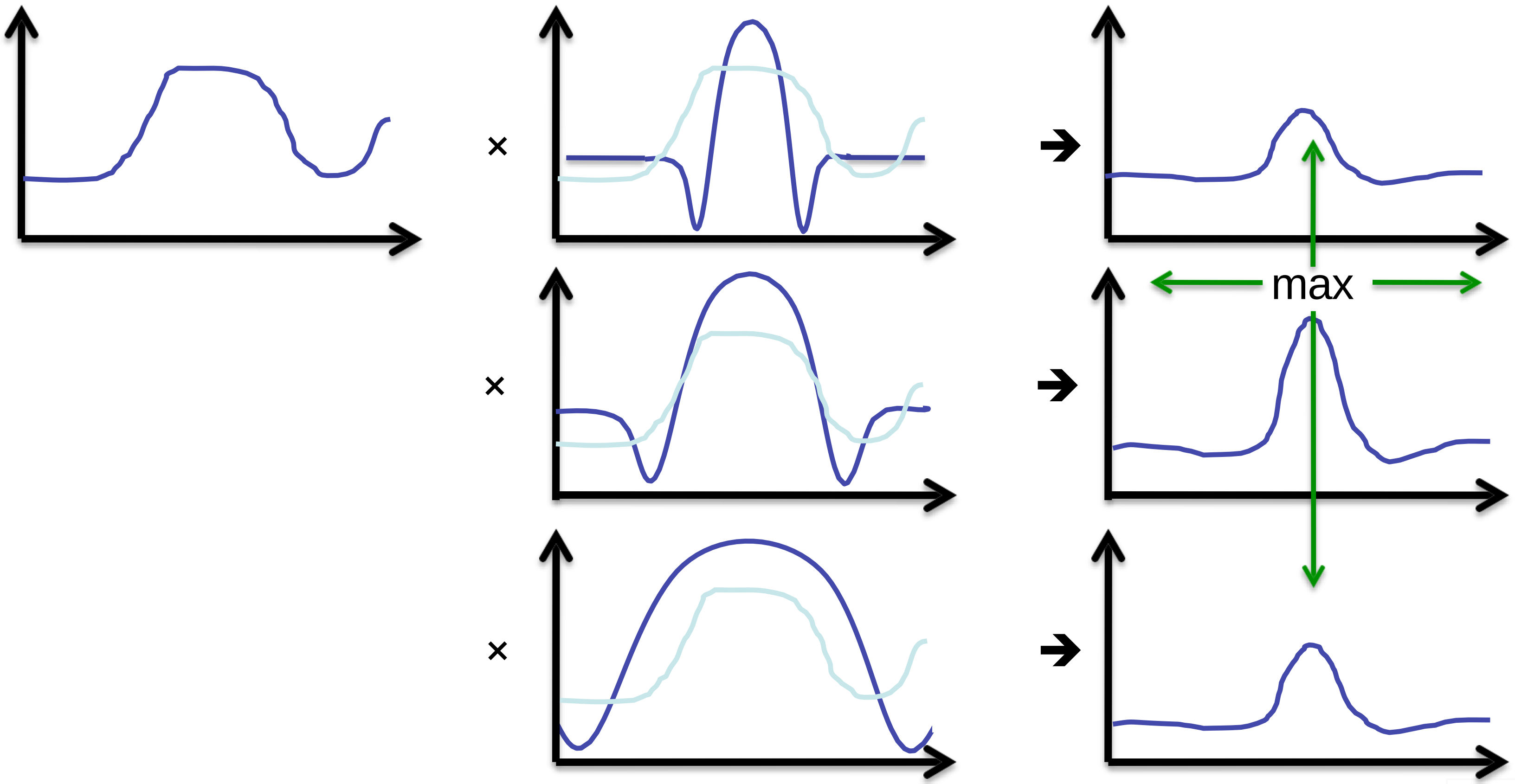




Détection de « blobs »

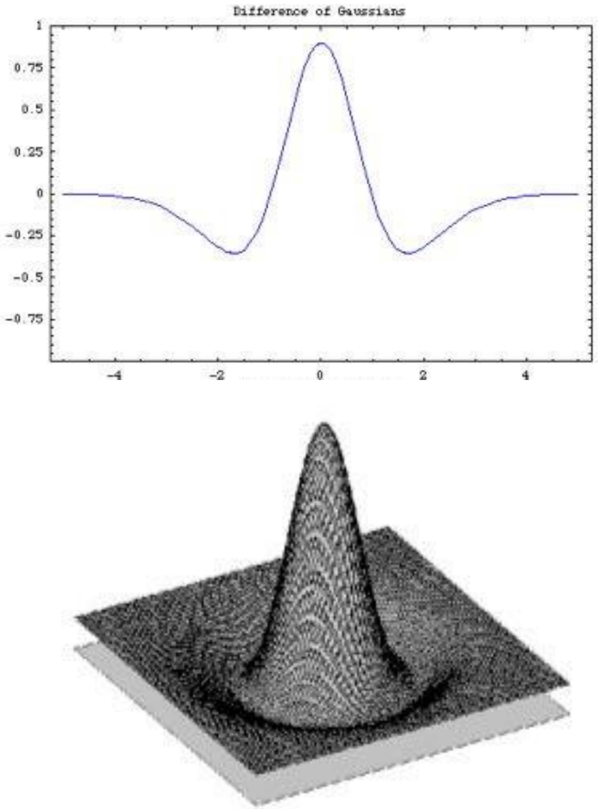
Détection d'échelle avec fonction en « cloche »

– Recherche de points au centre de « blobs »

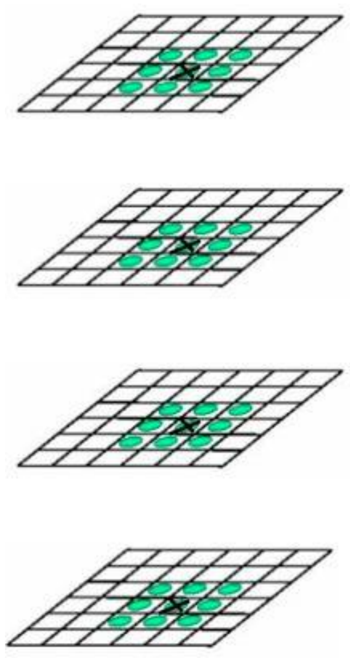


Détecteur SIFT

Détection par différence de gaussiennes



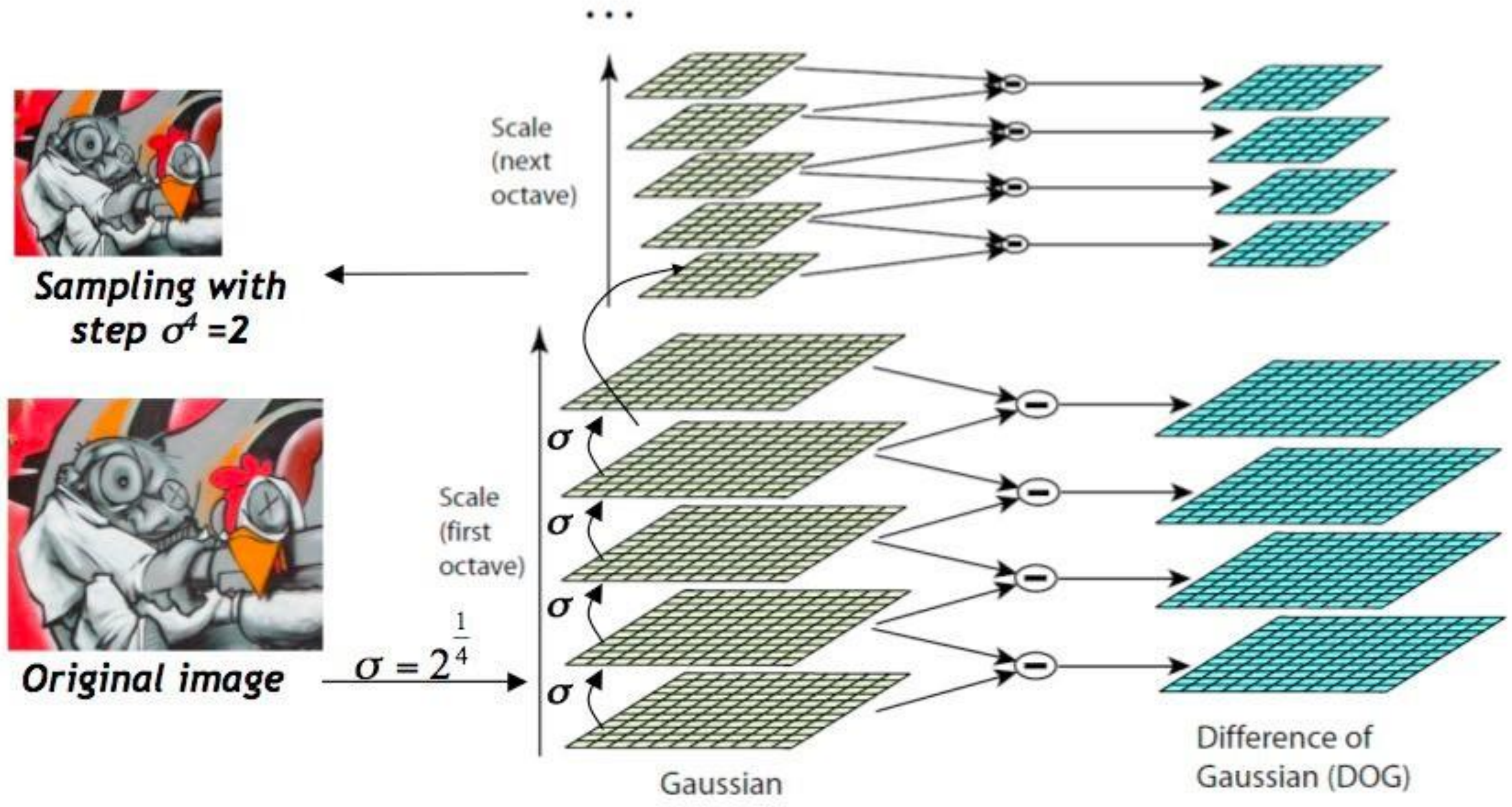
Convolution



Max échelle/espace

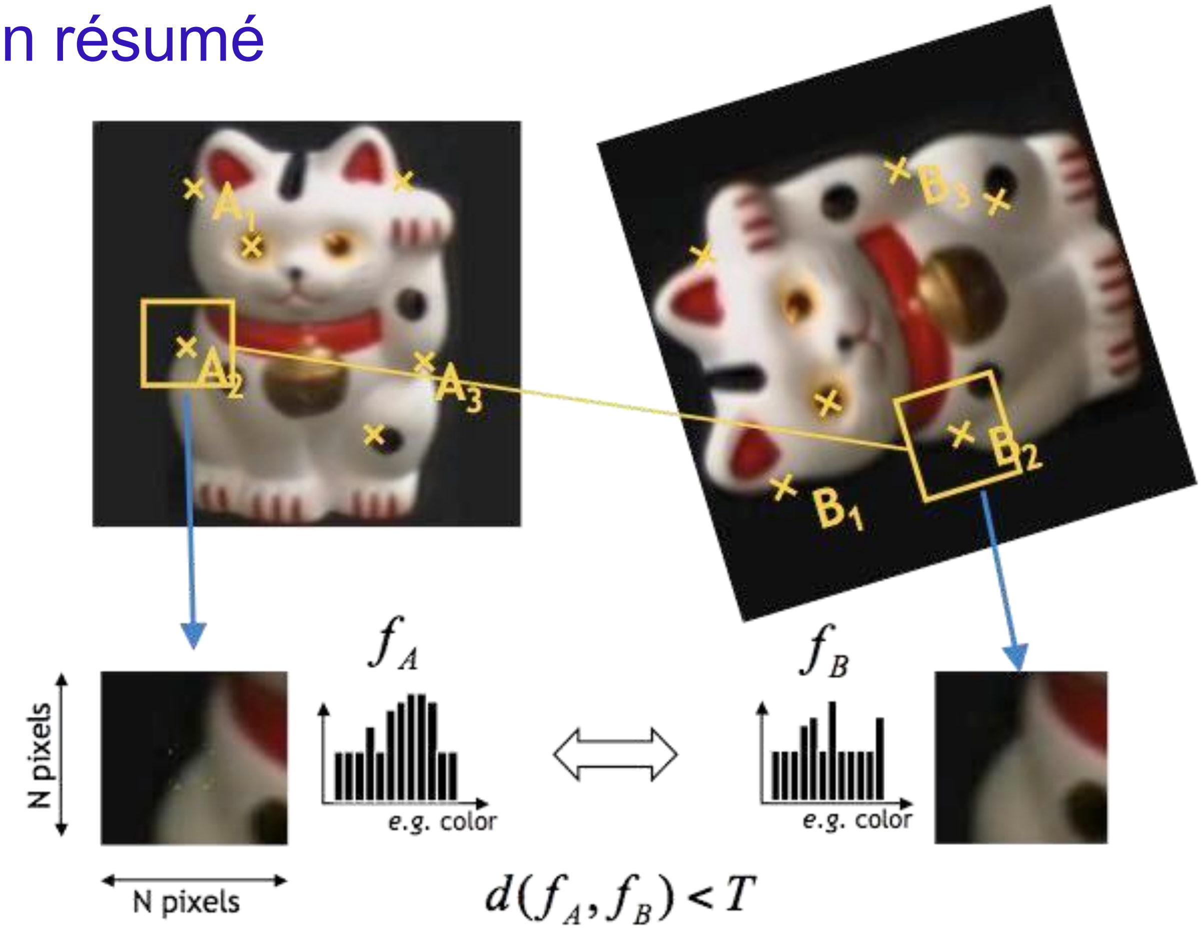
Détecteur SIFT

Implémentation optimisée



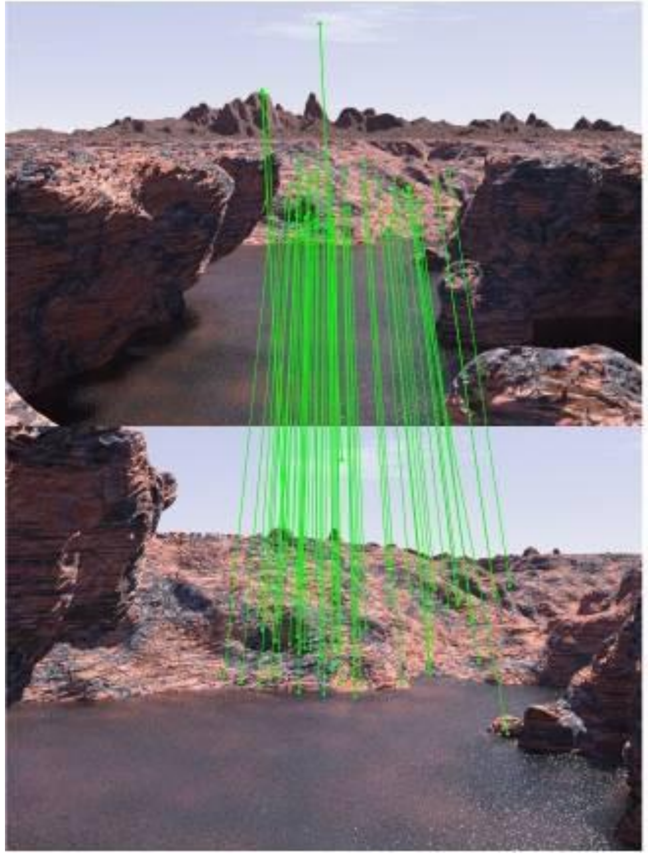
Détecteur SIFT

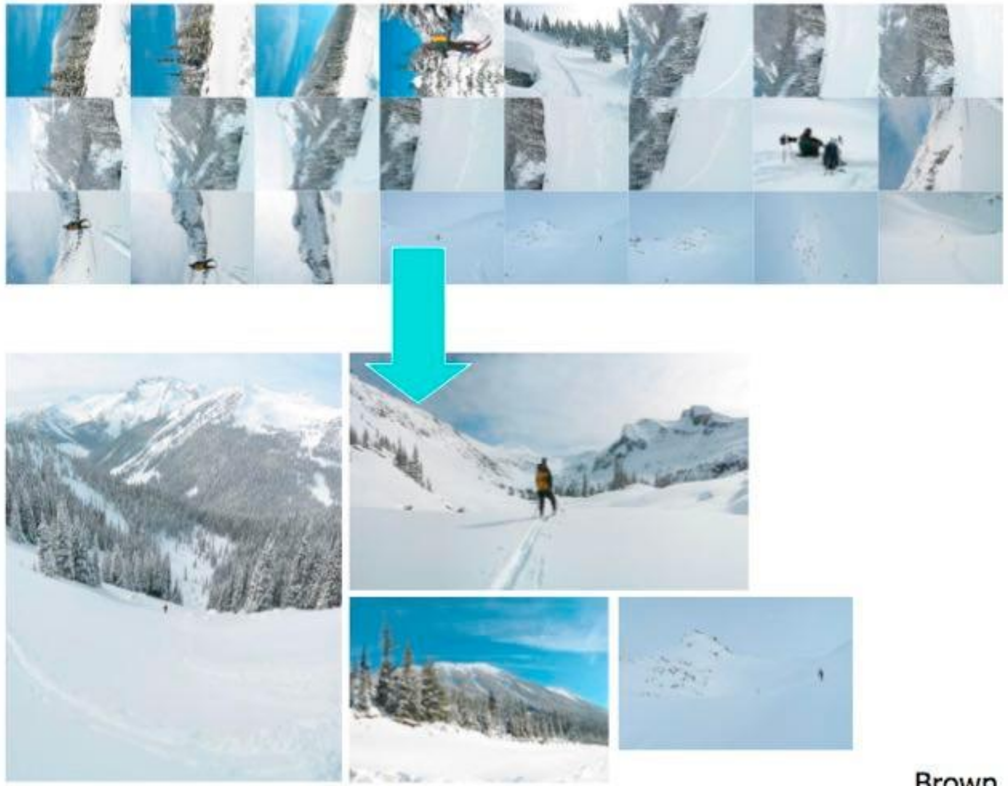
En résumé



Détecteur SIFT

Exemples de correspondances

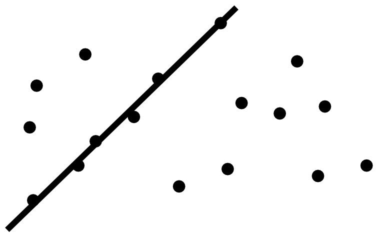




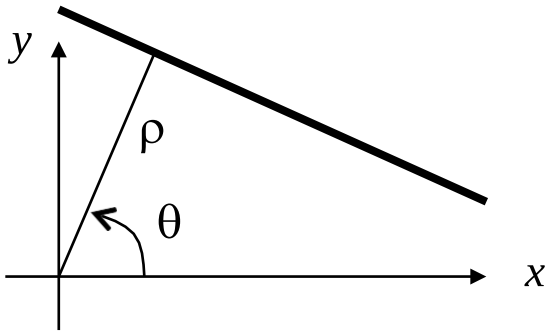
Brown, ICCV 2003

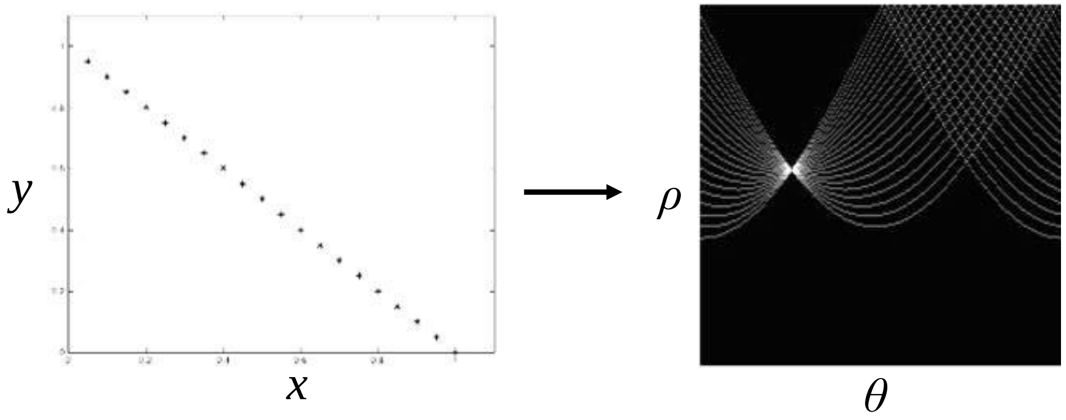
Hough Transform

- Origin: Detection of straight lines in clutter
  - Basic idea: each candidate point votes for all lines that it is consistent with.
  - Votes are accumulated in quantized array
  - Local maxima correspond to candidate lines

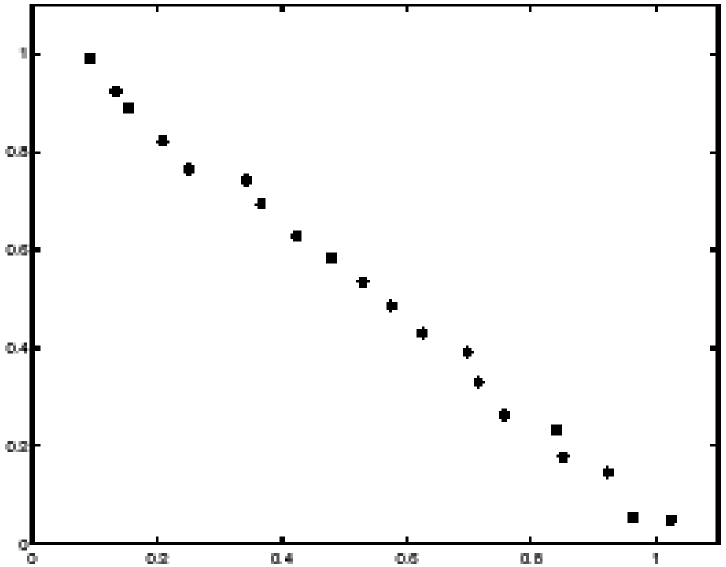


- Representation of a line
  - Usual form $y = a\,x + b$ has a singularity around 90°.
  - Better parameterization: $x \cos(\theta) + y \sin(\theta) = \rho$

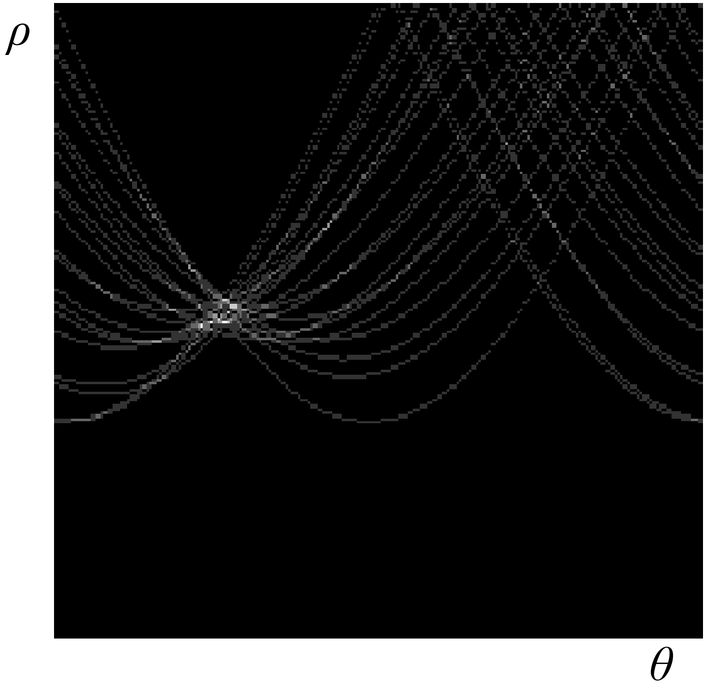




Hough Transform: Noisy Line



Tokens

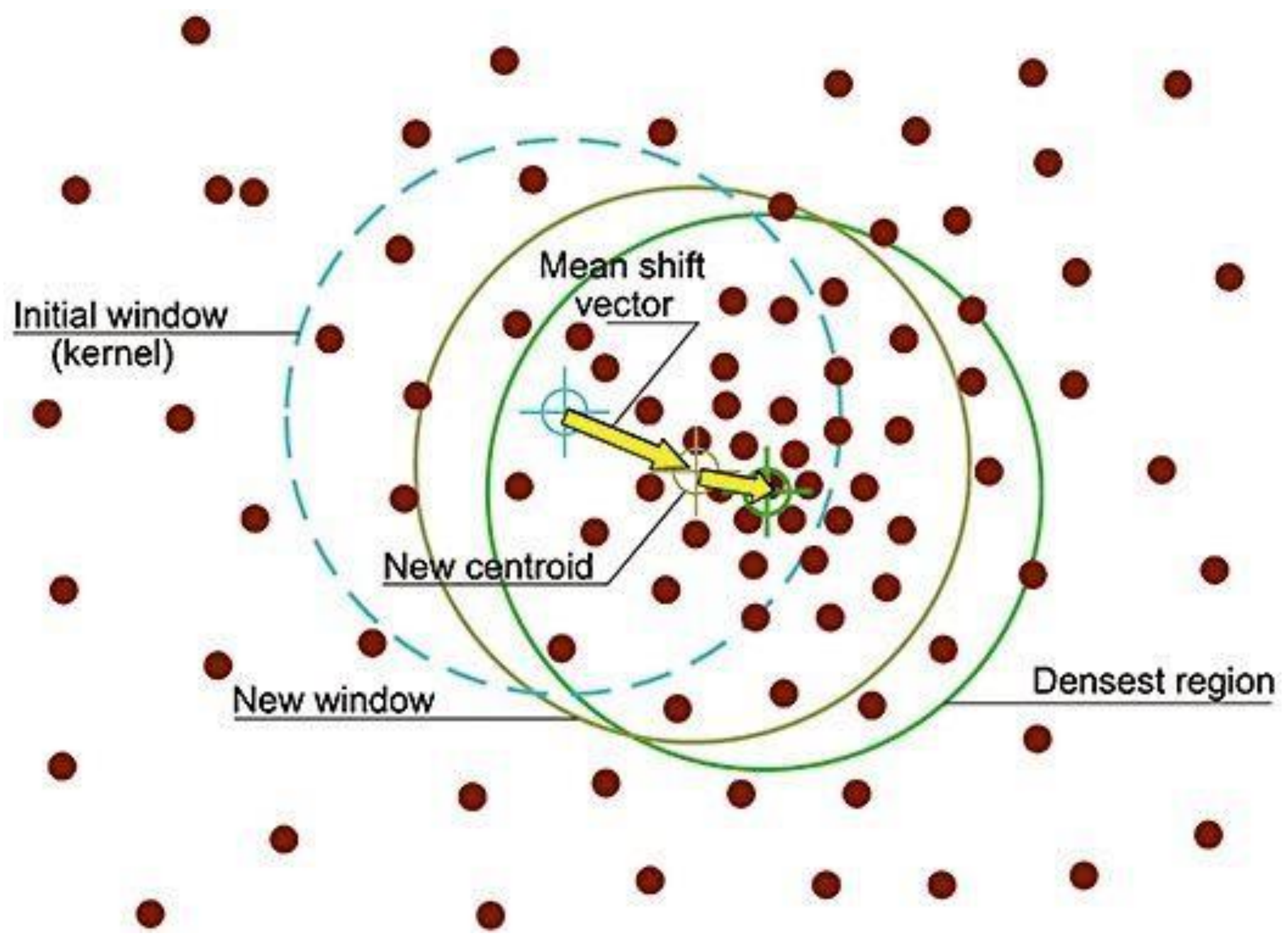


Votes

- Possible problem: Finding the true maximum
  - Mean-shift, Gaussian convolution…

Mean-Shift

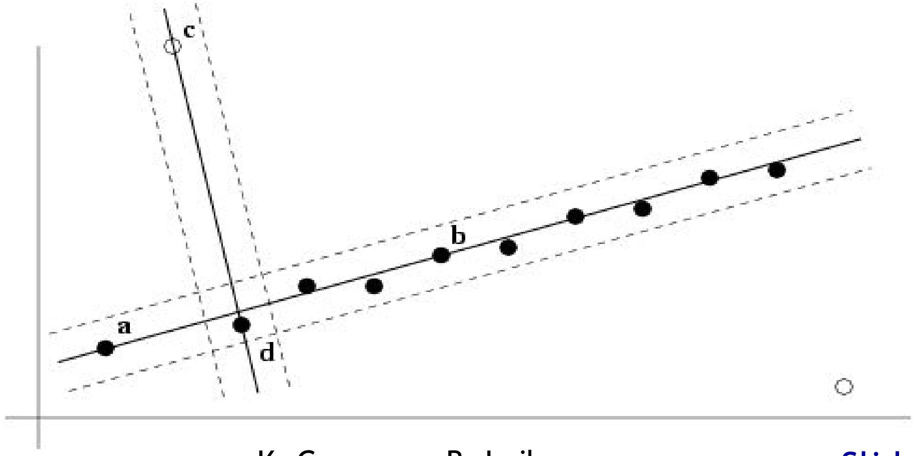
Recherche de max local itératif



RANSAC (RANdom SAmple Consensus)

[Fischler81]

- Randomly choose a minimal subset of data points necessary to fit a model (a *sample*)
- Points within some distance threshold t of model are a *consensus set*. Size of consensus set is model's *support*.
- Repeat for N samples; model with biggest support is most robust fit
  - Points within distance t of best model are inliers
  - Fit final model to all inliers



Slide credit: David Lowe

RANSAC: How many samples?

- How many samples are needed?
  - Suppose w is fraction of inliers (points from line).
  - n points needed to define hypothesis (2 for lines)
  - k samples chosen.
- Prob. that a single sample of n points is correct: w^n
- Prob. that all samples fail is: $(1-w^n)^k$

$\Rightarrow$ Choose k high enough to keep this below desired failure rate.

Slide credit: David Lowe

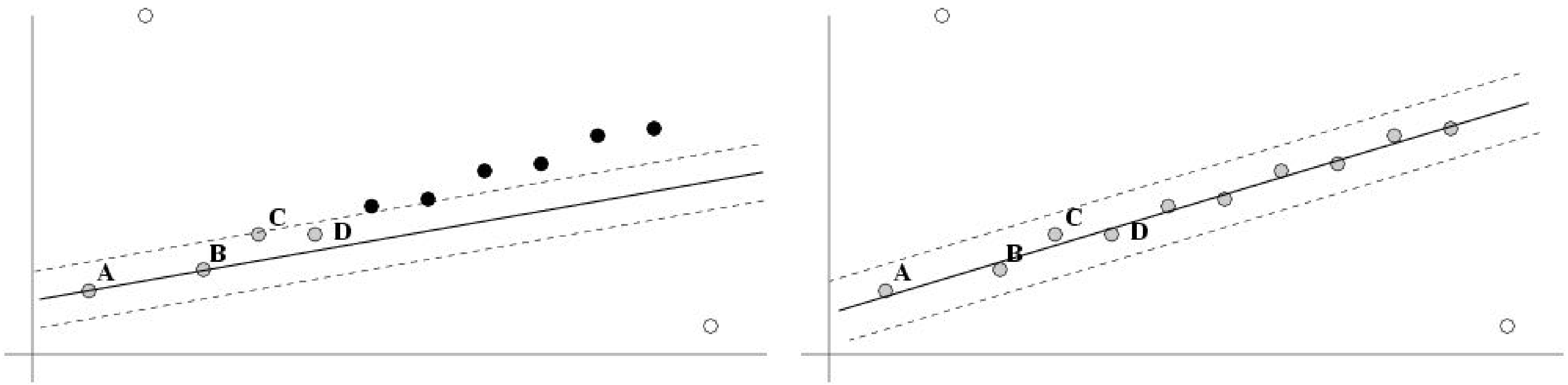
RANSAC: Computed k (p=0.99)

Sample size n	Proportion of outliers						
	5%	10%	20%	25%	30%	40%	50%
2	2	3	5	6	7	11	17
3	3	4	7	9	11	19	35
4	3	5	9	13	17	34	72
5	4	6	12	17	26	57	146
6	4	7	16	24	37	97	293
7	4	8	20	33	54	163	588
8	5	9	26	44	78	272	1177

Slide credit: David Lowe

After RANSAC

- RANSAC divides data into inliers and outliers and yields estimate computed from minimal set of inliers
- Improve this initial estimate with estimation over all inliers (e.g. with standard least-squares minimization)
- But this may change inliers, so alternate fitting with re-classification as inlier/outlier



Slide credit: David Lowe

Comparison

Gen. Hough Transform

- Advantages
  - Very effective for recognizing arbitrary shapes or objects
  - Can handle high percentage of outliers (>95%)
  - Extracts groupings from clutter in linear time
- Disadvantages
  - Quantization issues
  - Only practical for small number of dimensions (up to 4)
- Improvements available
  - Probabilistic Extensions
  - Continuous Voting Space

  [Leibe08]

RANSAC

- Advantages
  - General method suited to large range of problems
  - Easy to implement
  - Independent of number of dimensions
- Disadvantages
  - Only handles moderate number of outliers (<50%)
- Many variants available, e.g.
  - PROSAC: Progressive RANSAC [Chum05]
  - Preemptive RANSAC [Nister05]

Segmentation et detection d’objets

- Ligne de partage des eaux
- Cascade de Haar (Viola Jones detor

En pratique

- Non trivial de faire des opérations basiques de vision
- Problème de robustesse
- Avantage de la compréhension
- Bienvenue dans l’ère du Deep learning (2012)

Quelle heure est t'il ?

Et maintenant ! TP A vous de jouer?