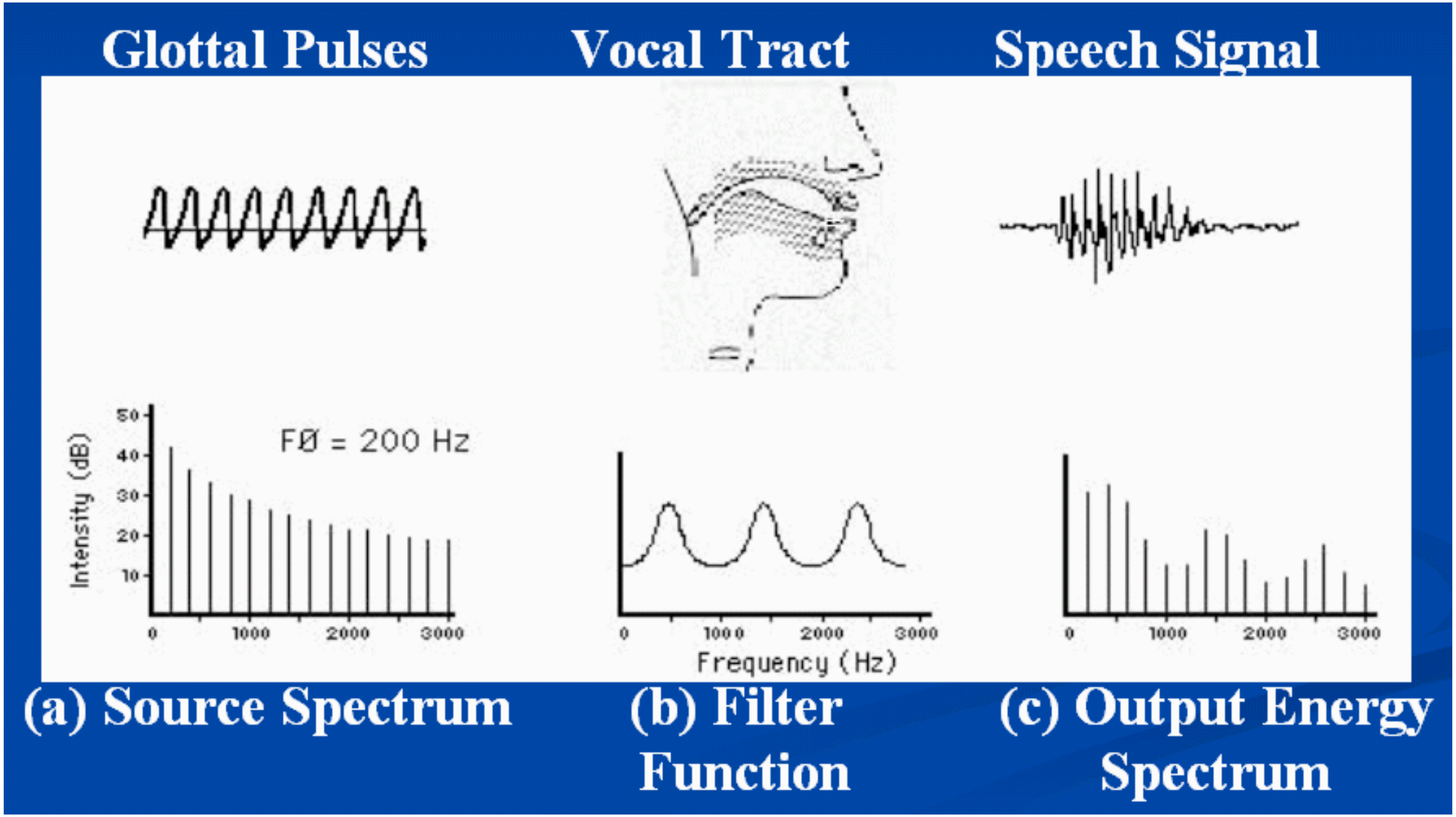
M2RI, Parcours Images et Données
Module Acquisition et Représentation des Données
Extraction de caractéristiques

Features Extraction - 2

Bertrand Coüasnon
Irisa/Insa
Imadoc
Bureau C314
http://www.irisa.fr/imadoc
couasnon@irisa.fr

Global Descriptor: Speech

- **Human speech production**



Global Descriptor: Speech

- **Source + filter model**
  - Time model $x(t) = (h \star s)(t)$
  - Frequency model

$$X(f) = H(f)S(f)$$

$$\log|X(f)| = \log|H(f)| + \log|S(f)|$$

varies slowly varies rapidly

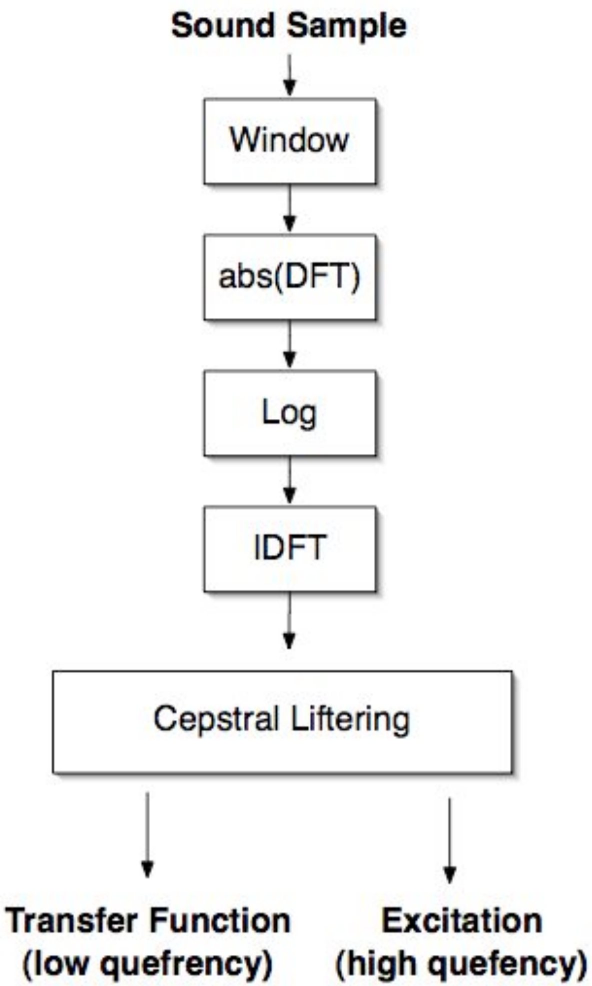
  - "Quefrency" model

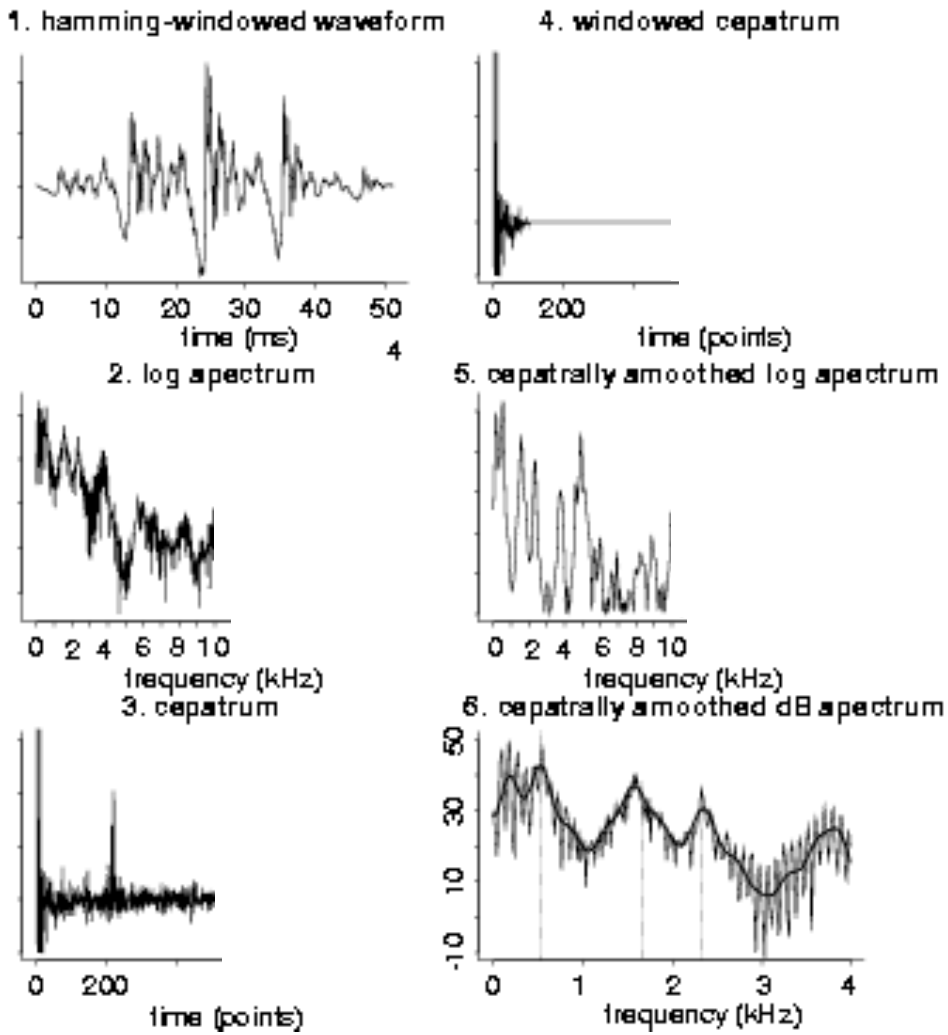
$$F^{-1}\{\log|X(f)|\} = F^{-1}\{\log|H(f)|\} + F^{-1}\{\log|S(f)|\}$$

low quefrency high quefrency

Global Descriptor: Speech

▪ Cepstrum





Global Descriptor: Speech

- **Use of the Cepstrum in speech analysis**
  - Compute the cepstrum on 20 msec. window frames, every 10 msec.
  - N=12 lowest “quefrency” coefficients
  - +1 energy on each frame
  - + first & second time derivatives

⇒ 3(N+1) dimensional feature vector every 10 msec.

Outline

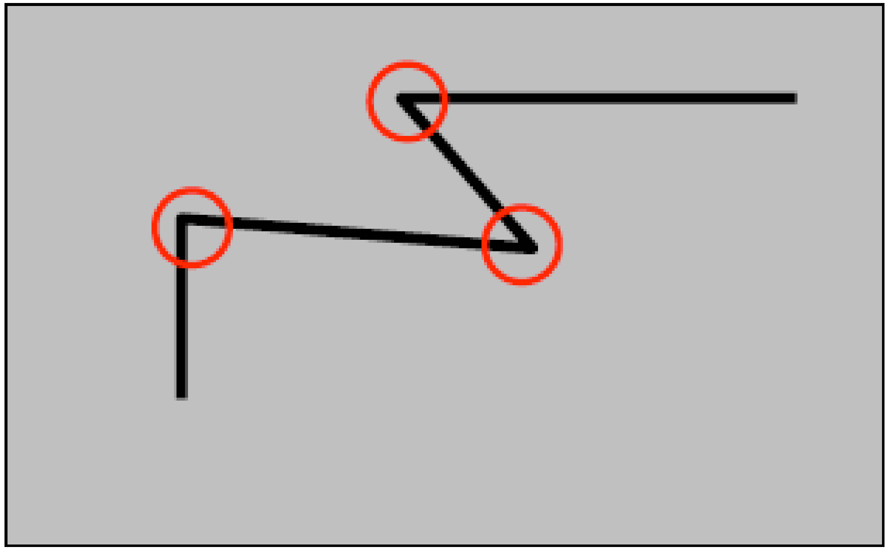
- Difficulties of Feature Extraction
  - Segmentation
  - Recognition in Images
  - Recognition in Document Images
- Feature Extraction/Detection on the whole Signal
  - Global Descriptor
  - Local Descriptor
- Feature Extraction/Detection on Objects
  - Object Localization / Extraction
  - Object Characterization
- Interest of Multi-Resolution
- Using Features

Local Descriptor: Introduction

- Stable across images
  - Invariant to small camera transformations
  - Robust to changes in illumination
- Descriptive
  - Different image patches will be represented differently
  - Descriptions can be reliably compared
- Interest Point Detector
  - Harris Corner Detector (very popular)
- Feature Descriptor
  - SIFT (best performer)

Local Descriptor: Harris

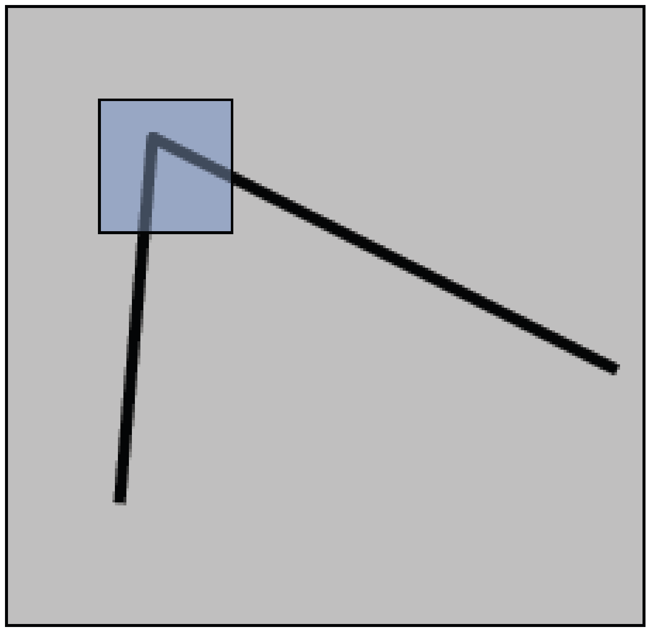
- **Harris corner detector**



  - **C.Harris, M.Stephens. "A Combined Corner and Edge Detector". 1988**

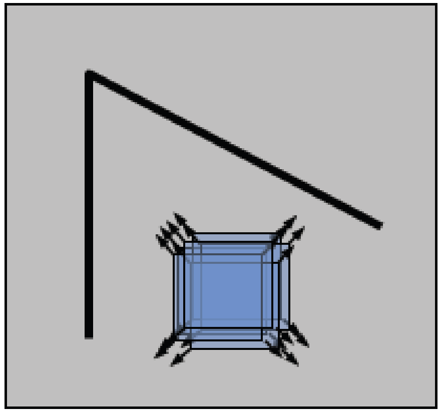
Local Descriptor: Harris

- We should easily recognize the point by looking through a small window
- Shifting a window in any direction should give a large change in intensity

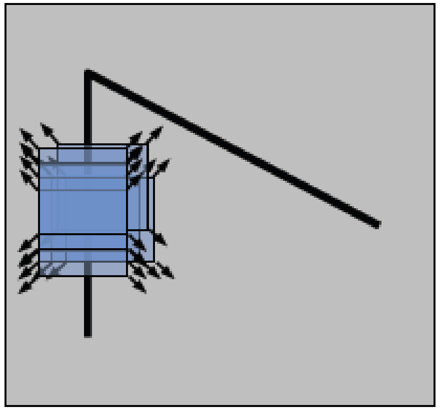


Local Descriptor: Harris

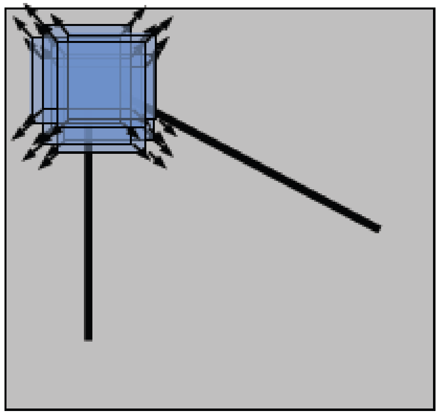
- “flat” region: no change in all directions



- “edge”: no change along the edge direction



- “corner”: significant change in all directions



Local Descriptor: Harris

Change of intensity for the shift $[u,v]$:

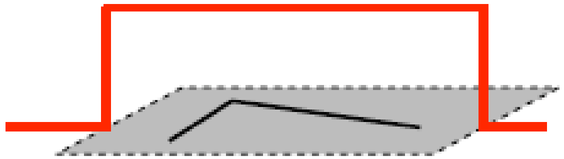
$$E(u,v)=\sum_{x,y} w(x,y)\left[I(x+u,y+v)-I(x,y)\right]^2$$

Window function

Shifted intensity

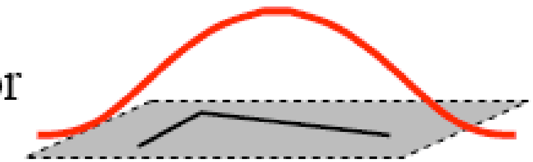
Intensity

Window function $w(x,y)$ =



1 in window, 0 outside

or



Gaussian

$$=\sum_{x,y} w(x,y)\left[I_x u+I_y v+O(u^2,v^2)\right]^2$$

Local Descriptor: Harris

For small shifts $[u,v]$ we have a *bilinear* approximation:

$$E(u,v) \cong [u,v] \quad M \quad \begin{bmatrix} u \\ v \end{bmatrix}$$

where M is a 2×2 matrix computed from image derivatives:

$$M = \sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$

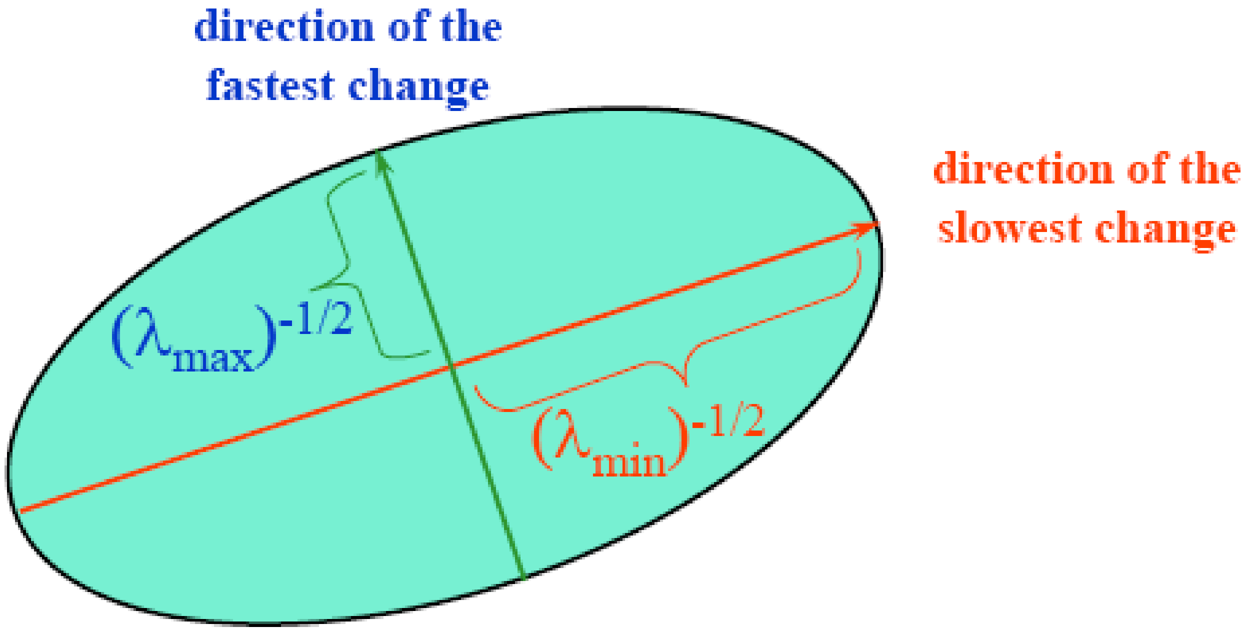
Local Descriptor: Harris

Intensity change in shifting window: eigenvalue analysis

$$E(u,v) \cong [u, v] \;\; M \begin{bmatrix} u \\ v \end{bmatrix}$$

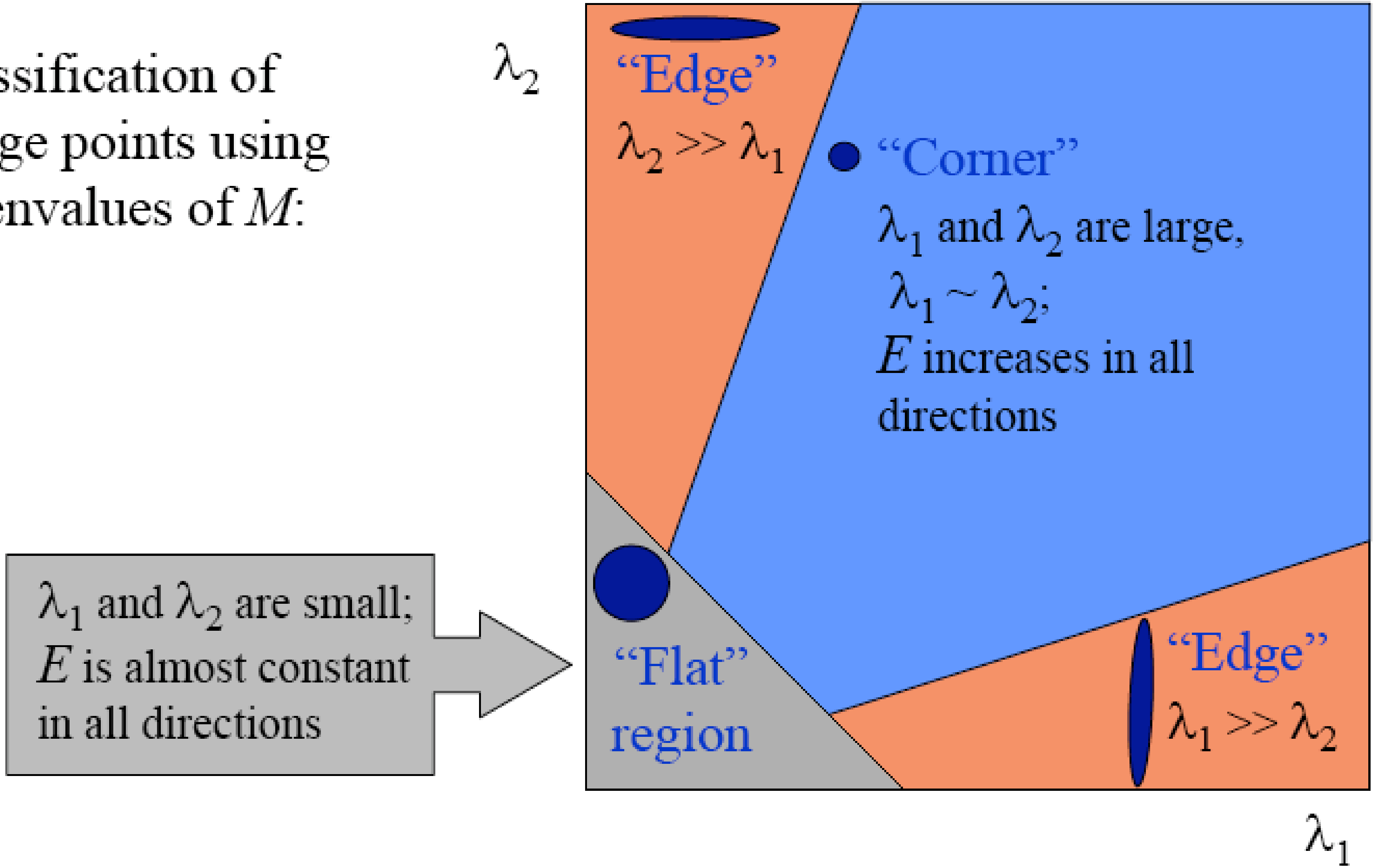
λ_1, λ_2 – eigenvalues of M

Ellipse $E(u,v) = 2$



Local Descriptor: Harris

Classification of image points using eigenvalues of M:



Local Descriptor: Harris

Measure of corner response:

$$R = \det M - k\left(\operatorname{trace} M\right)^2$$

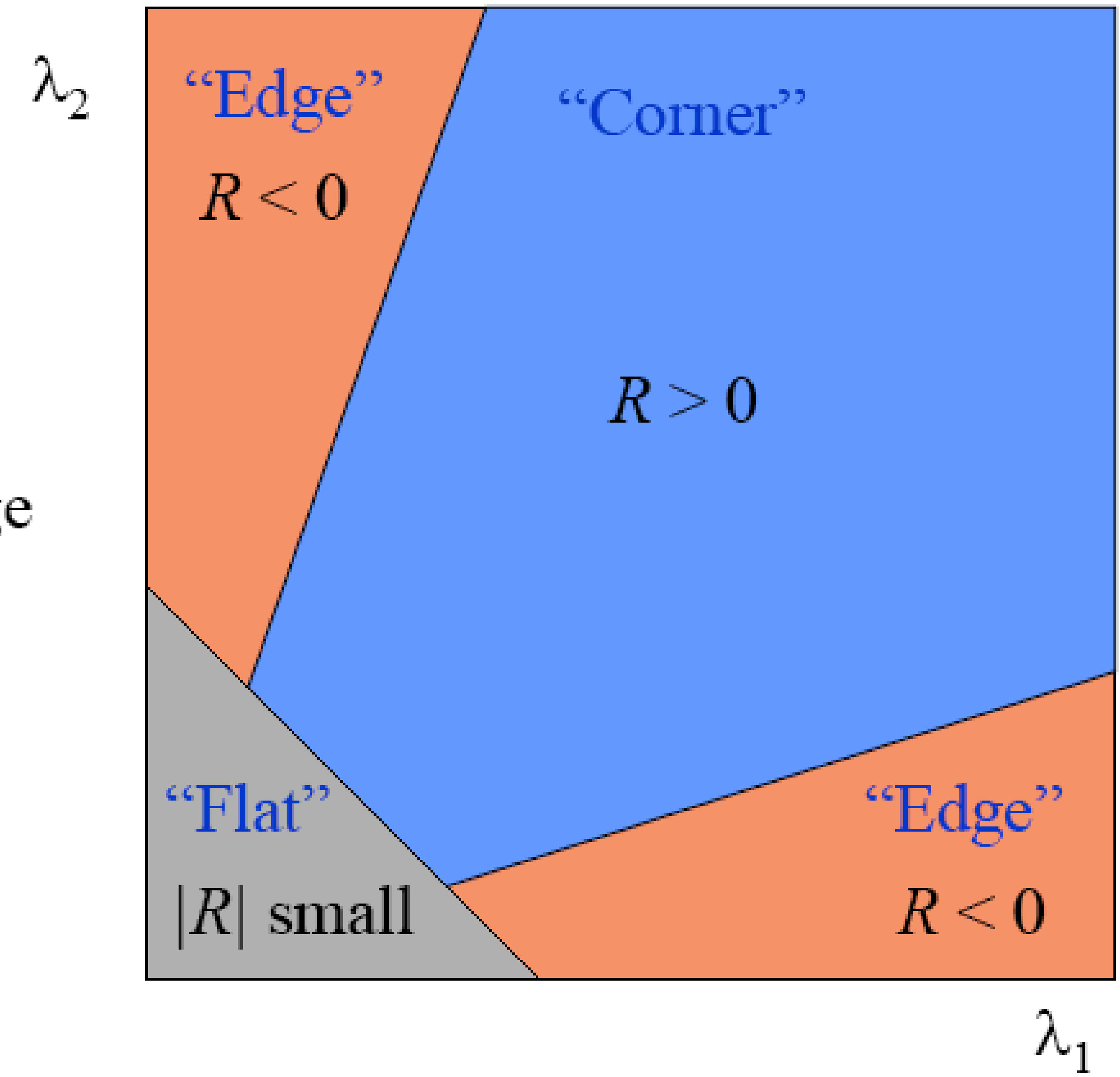
$$\det M = \lambda_1 \lambda_2$$
$$\operatorname{trace} M = \lambda_1 + \lambda_2$$

(k – empirical constant, k = 0.04-0.06)

Motivation: computation avoids explicit eignevalue decomposition

Local Descriptor: Harris

- R depends only on eigenvalues of M
- R is large for a corner
- R is negative with large magnitude for an edge
- $|R|$ is small for a flat region



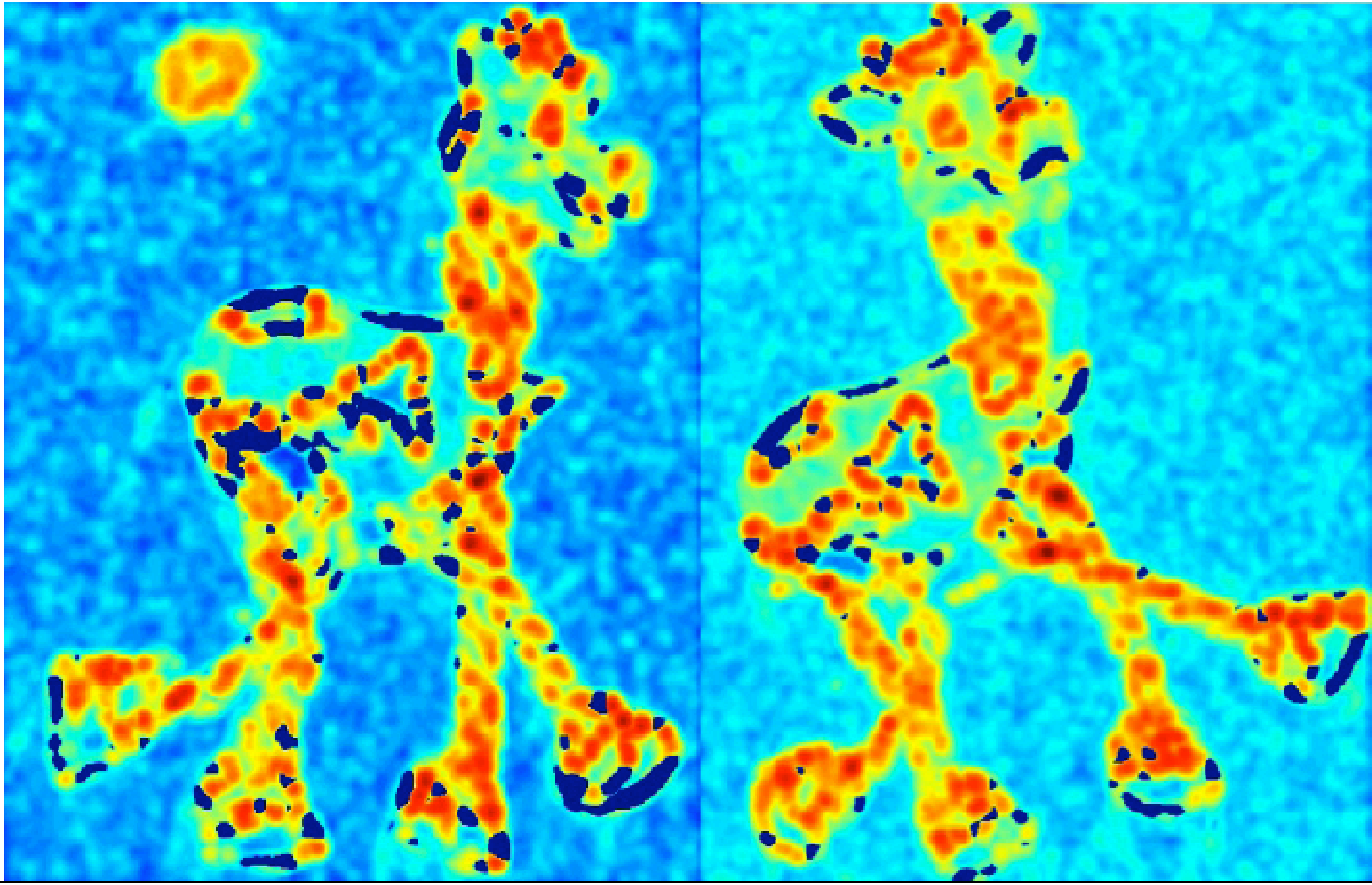
Local Descriptor: Harris

- **The Algorithm**
  - **Find points with large corner response function *R* (*R* > threshold)**
  - **Take the points of local maxima of *R***

Local Descriptor: Harris - Example

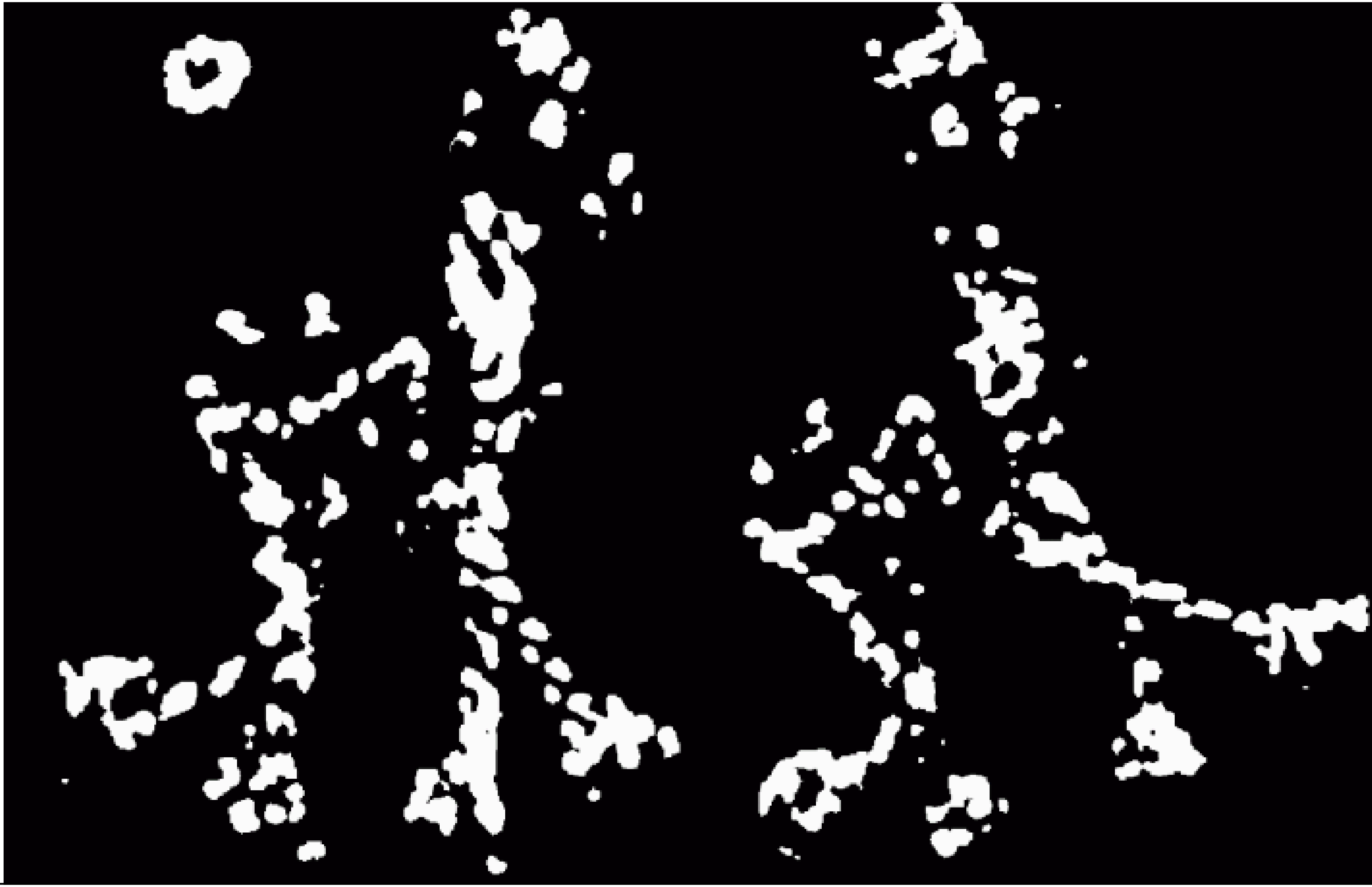
Local Descriptor: Harris - Example

- **Compute corner response R**



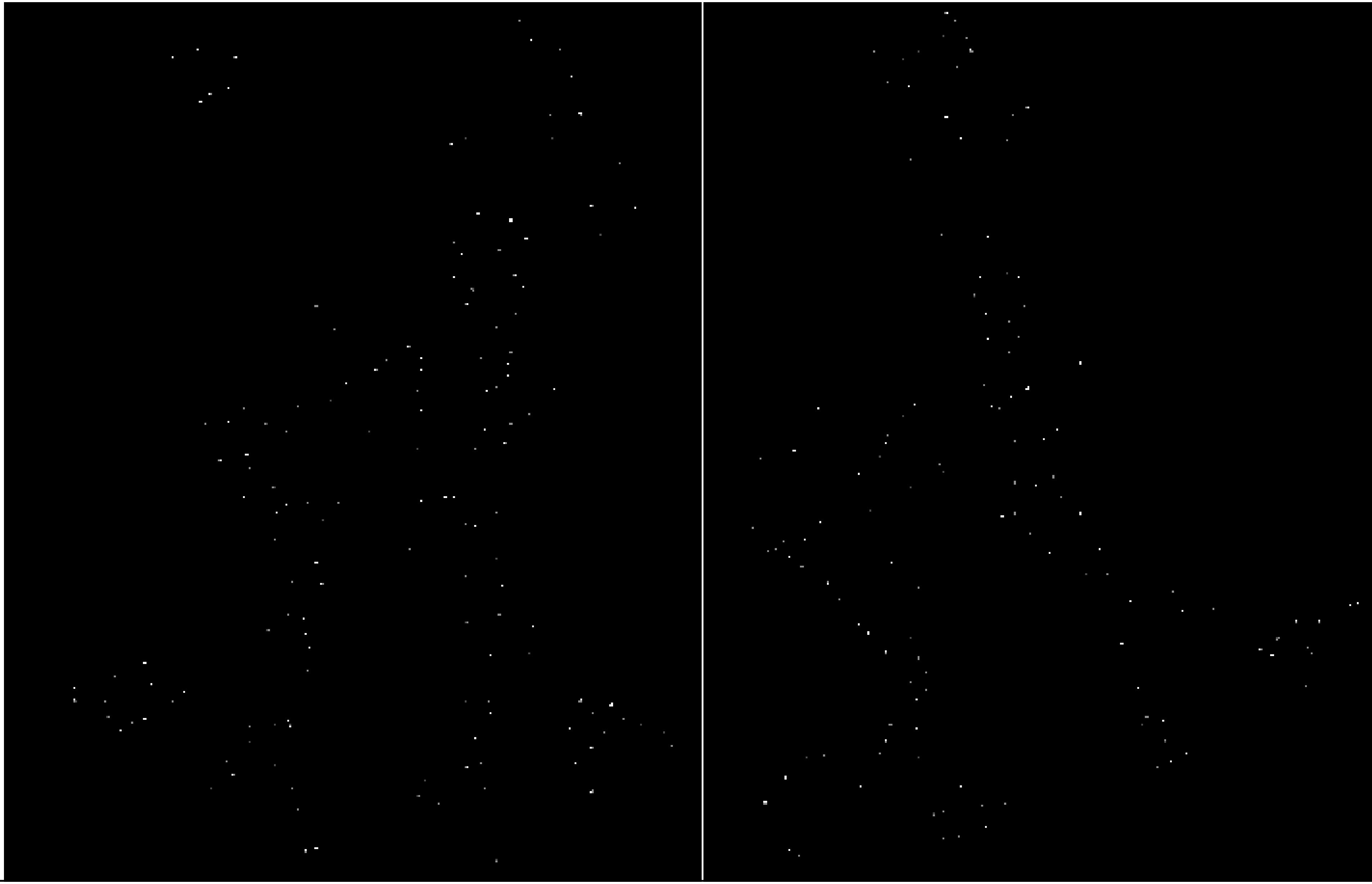
Local Descriptor: Harris - Example

- Find points with large corner response: R>threshold



Local Descriptor: Harris - Example

- Take only the points of local maxima of R



Local Descriptor: Harris - Example

Local Descriptor: Harris - Summary

- Average intensity change in direction $[u,v]$ can be expressed as a bilinear form:

$$E(u,v) \cong [u,v] \; M \begin{bmatrix} u \\ v \end{bmatrix}$$

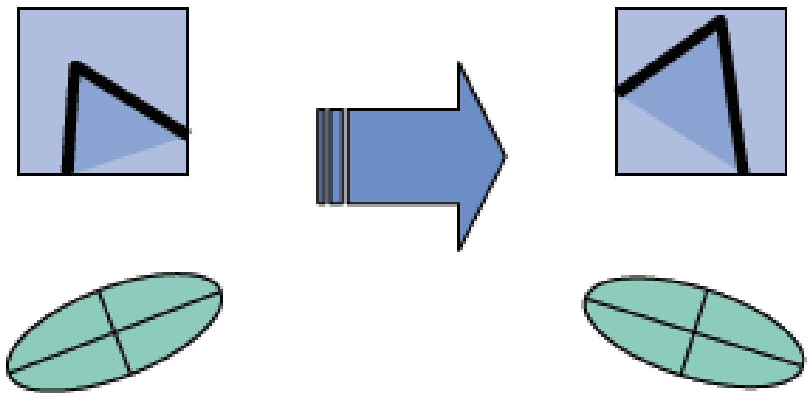
- Describe a point in terms of eigenvalues of M: *measure of corner response*

$$R = \lambda_1 \lambda_2 - k\left(\lambda_1 + \lambda_2\right)^2$$

- A good (corner) point should have a *large intensity change* in *all directions*, i.e. R should be large and positive

Local Descriptor: Harris - Some Properties

- Rotation invariance

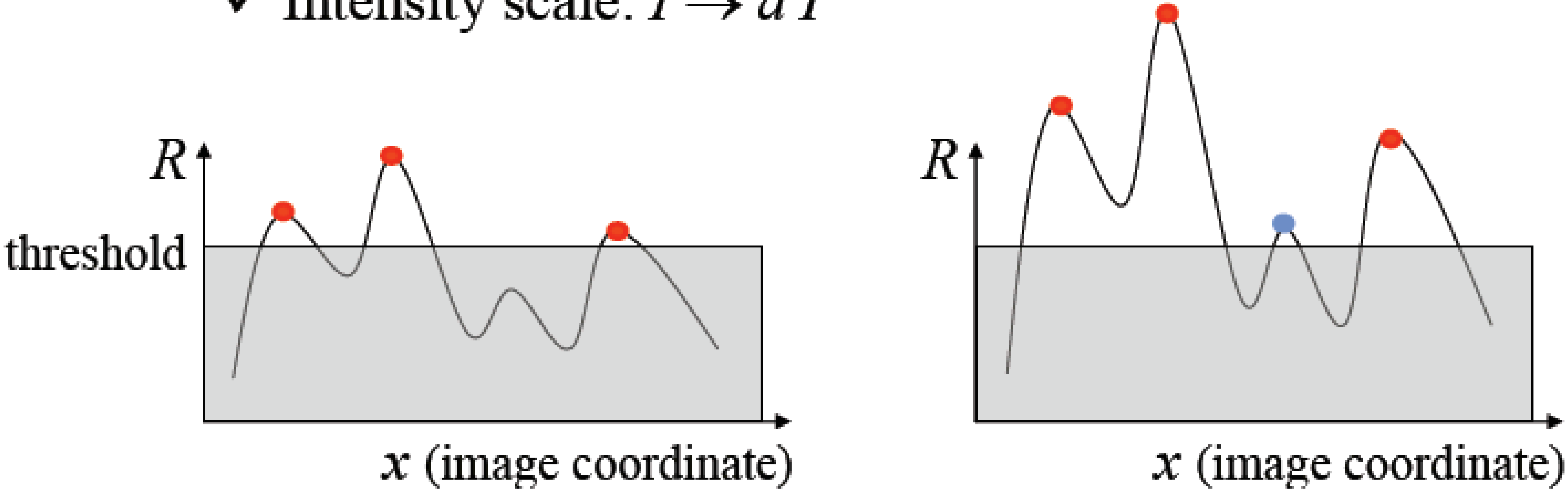


Ellipse rotates but its shape (i.e. eigenvalues) remains the same

Corner response R is invariant to image rotation

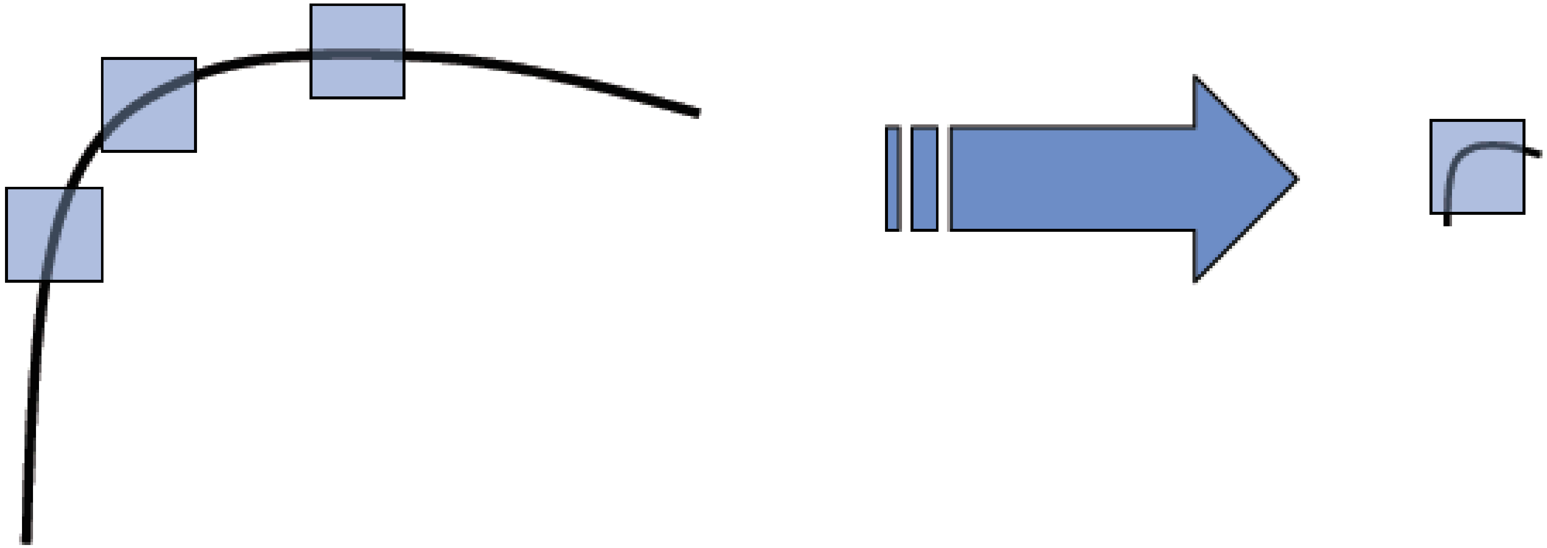
Local Descriptor: Harris - Some Properties

- Partial invariance to *affine intensity* change
  - ✓ Only derivatives are used => invariance to intensity shift $I \rightarrow I + b$
  - ✓ Intensity scale: $I \rightarrow a\,I$



Local Descriptor: Harris - Some Properties

- But: non-invariant to *image scale*!

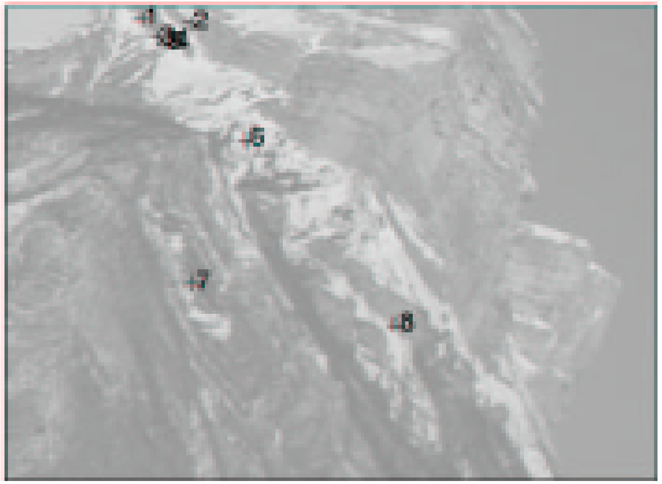


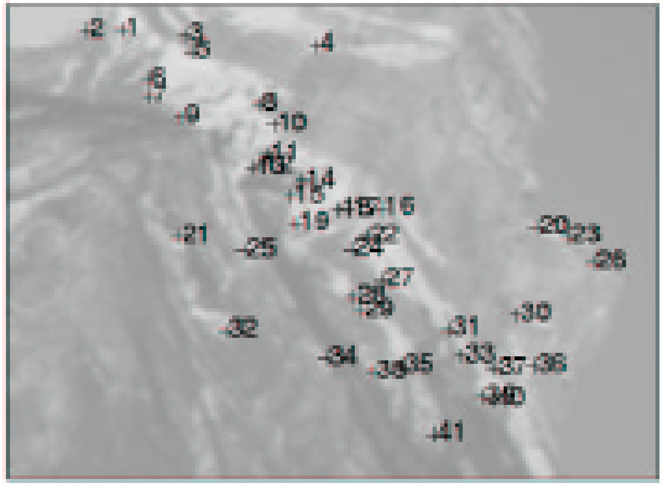
All points will be classified as edges

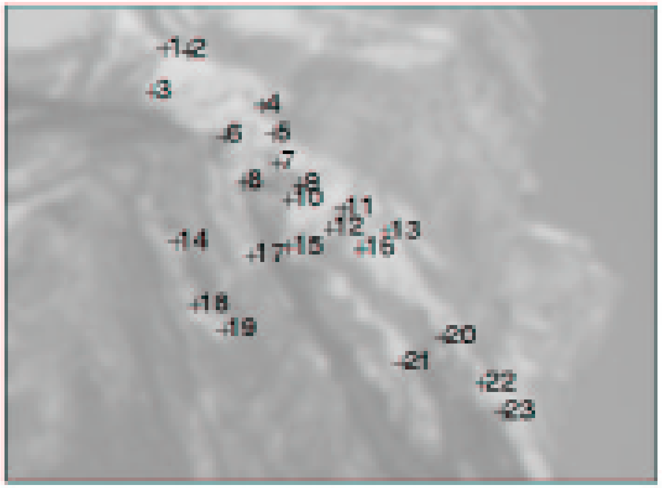
Corner !

Local Descriptor: Harris

- **Example of Harris Corner Detector at different scales**

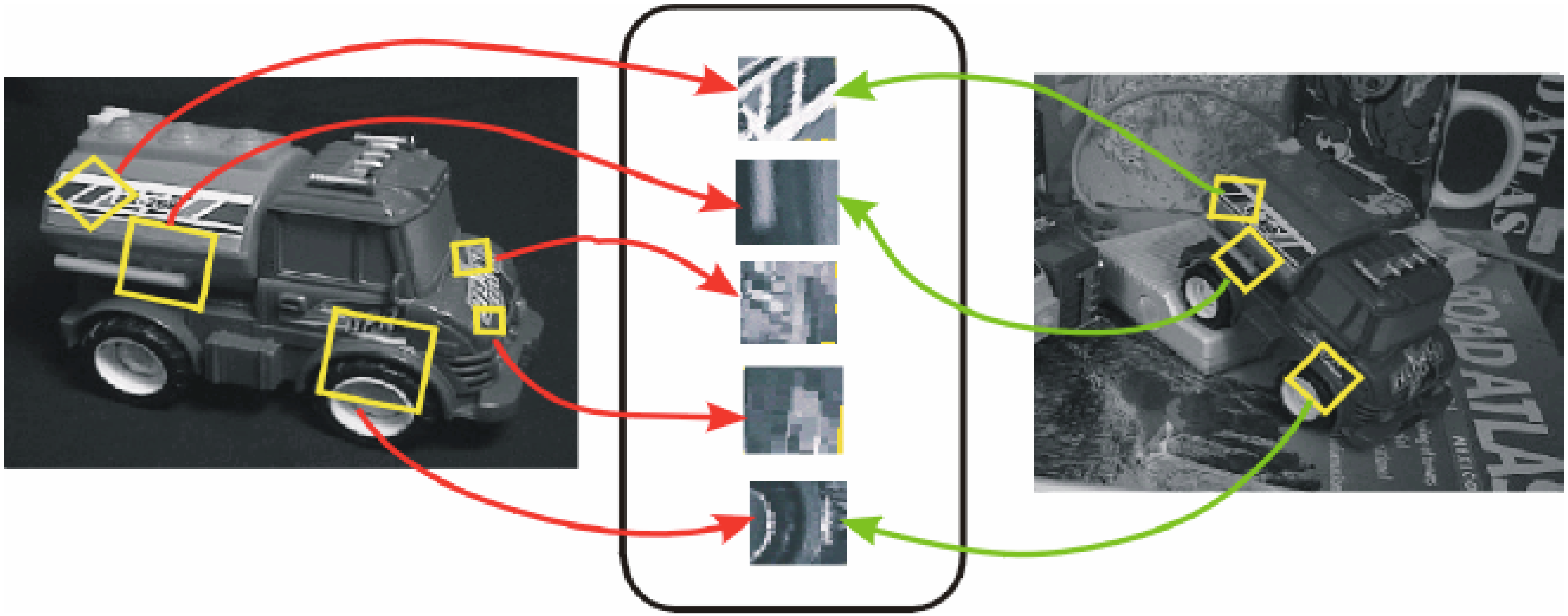






Local Detector: SIFT

- SIFT: Scale Invariant Feature Transform
- Image content is transformed into local feature coordinates that are invariant to translation, rotation, scale, and other imaging parameters



SIFT Features

- D.G. Lowe, 1999, 2004

Local Detector: SIFT

- **1) Scale-Space Extrema Detection**
  - Interest Point
    - Local extrema of difference-of-Gaussian filters at different scales
- **2) Keypoint localization**
  - Removes extrema with low contrast
- **3) Orientation assignment**
  - Gradient Orientation
- **4) Generation of keypoint descriptors**

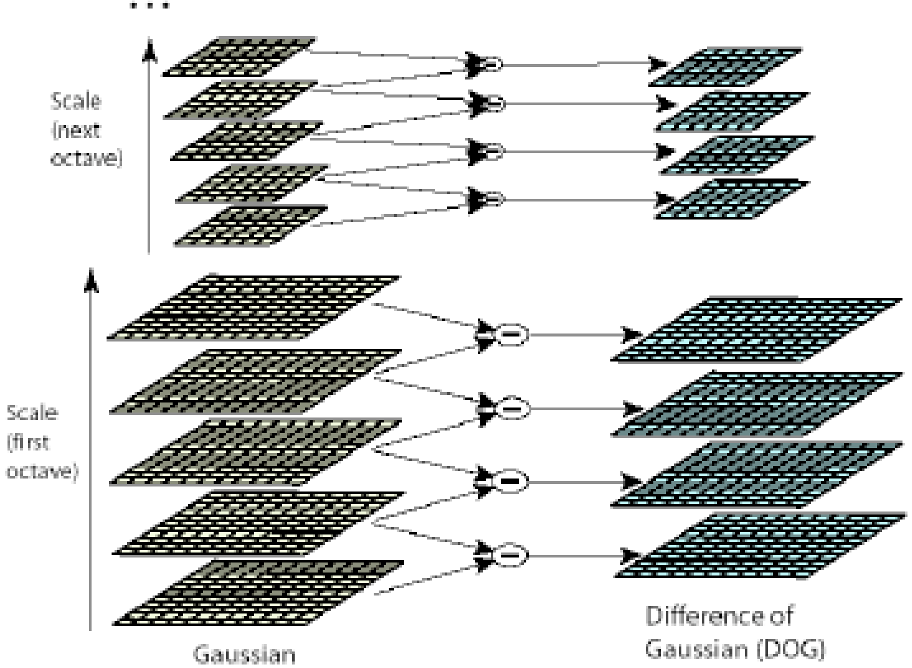
SIFT: 1) Scale-Space Extrema Detection

From Gaussian scale pyramid -- create Difference of Gaussian (DOG) images

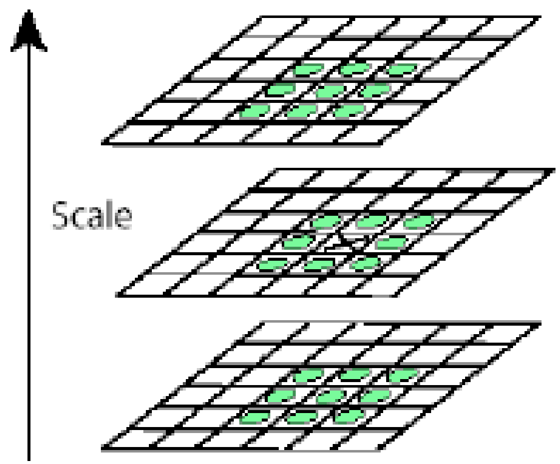
$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$

$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y)$

$x, y, k\sigma) - L(x, y, \sigma).$

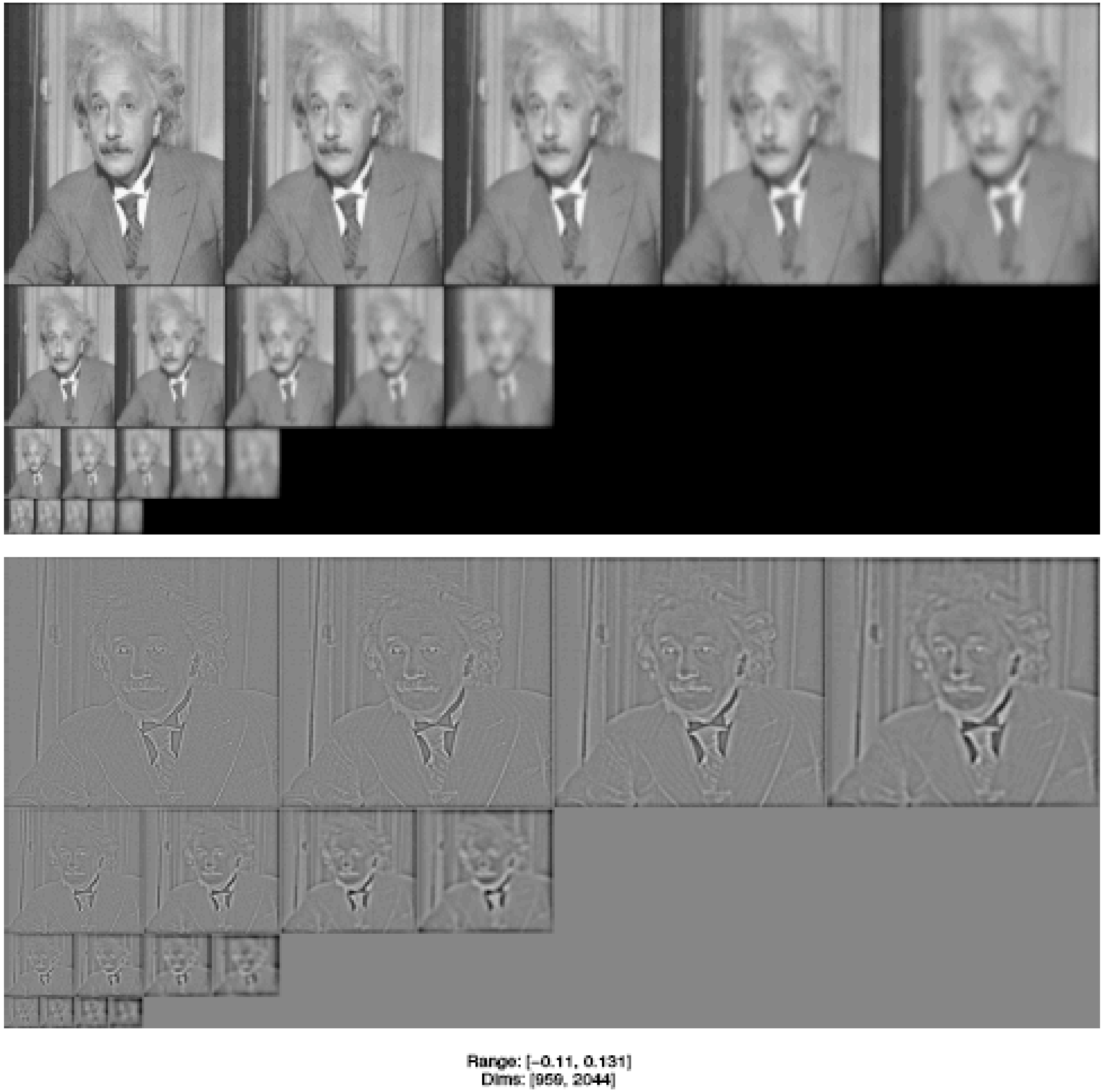


- **Find maximun response over space and scale**

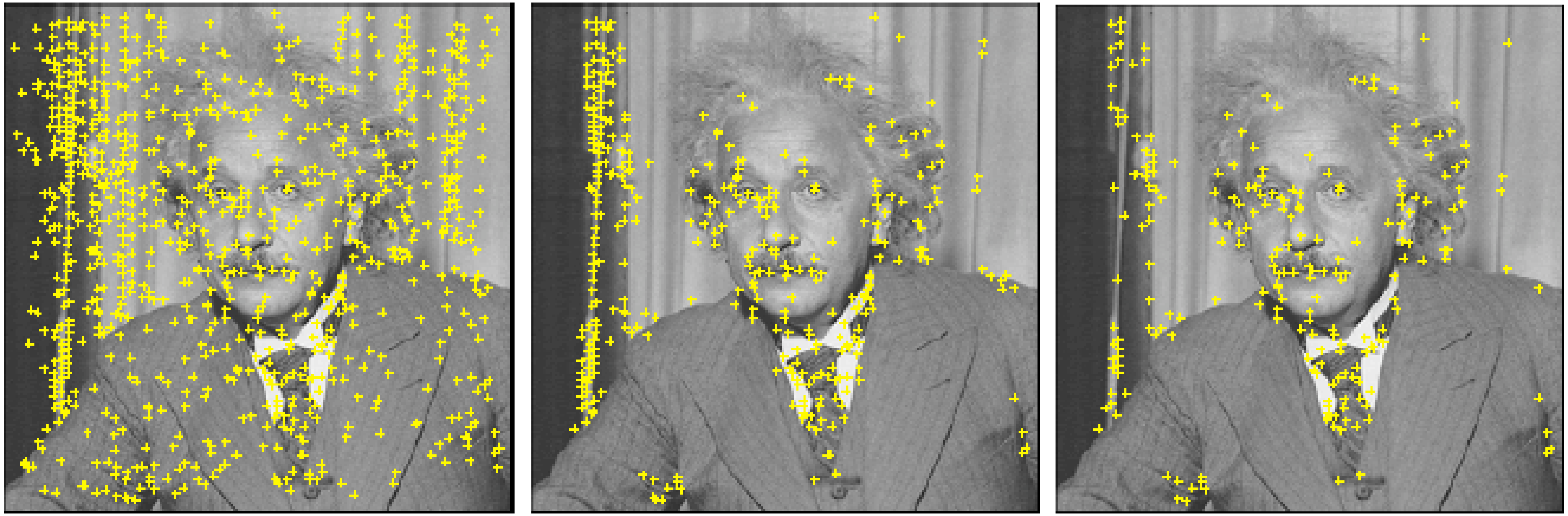


SIFT: 1) Scale-Space Extrema Detection

- Gaussian blurred images and DoG images grouped by octave

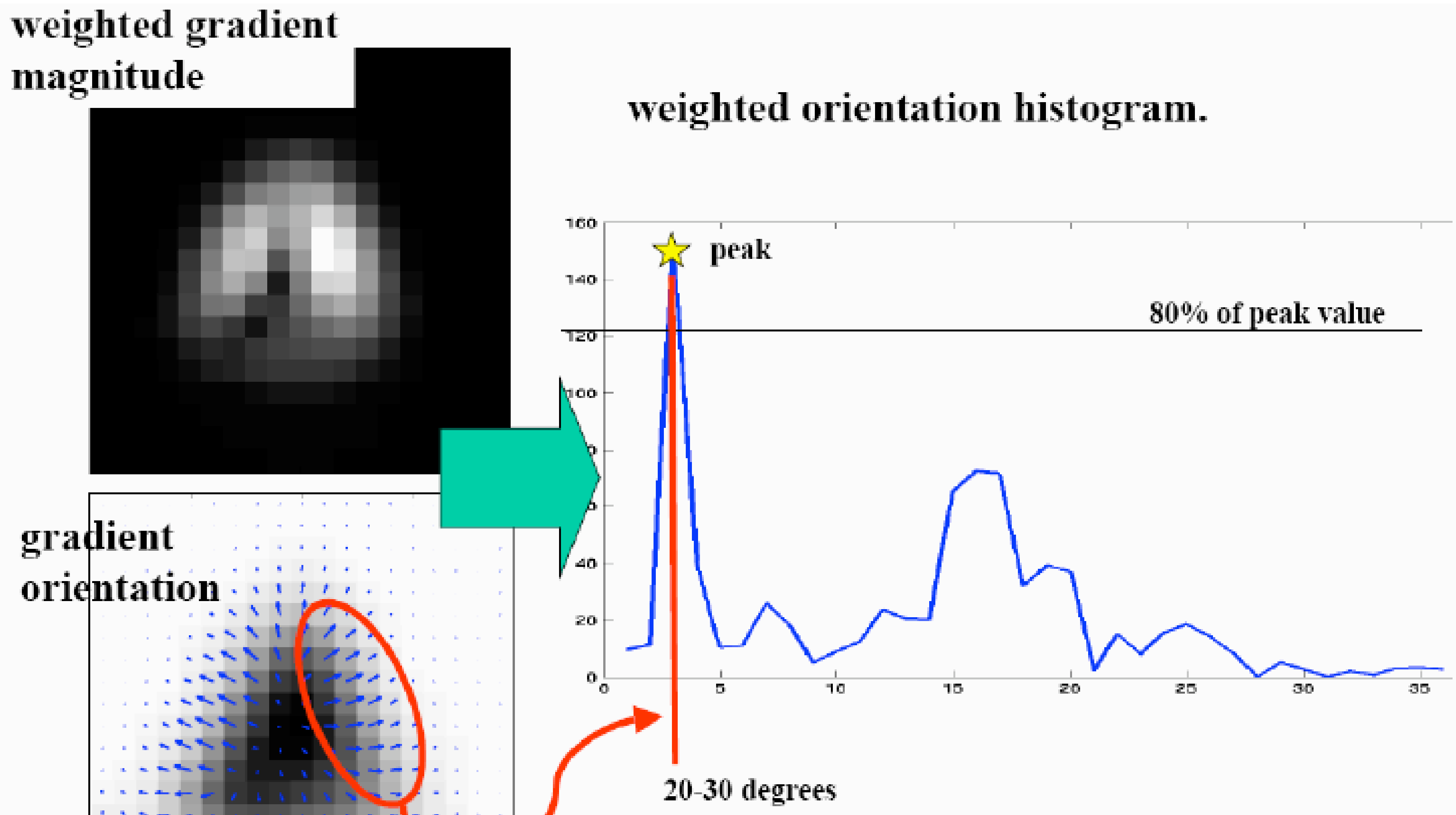


SIFT: 1) and 2)



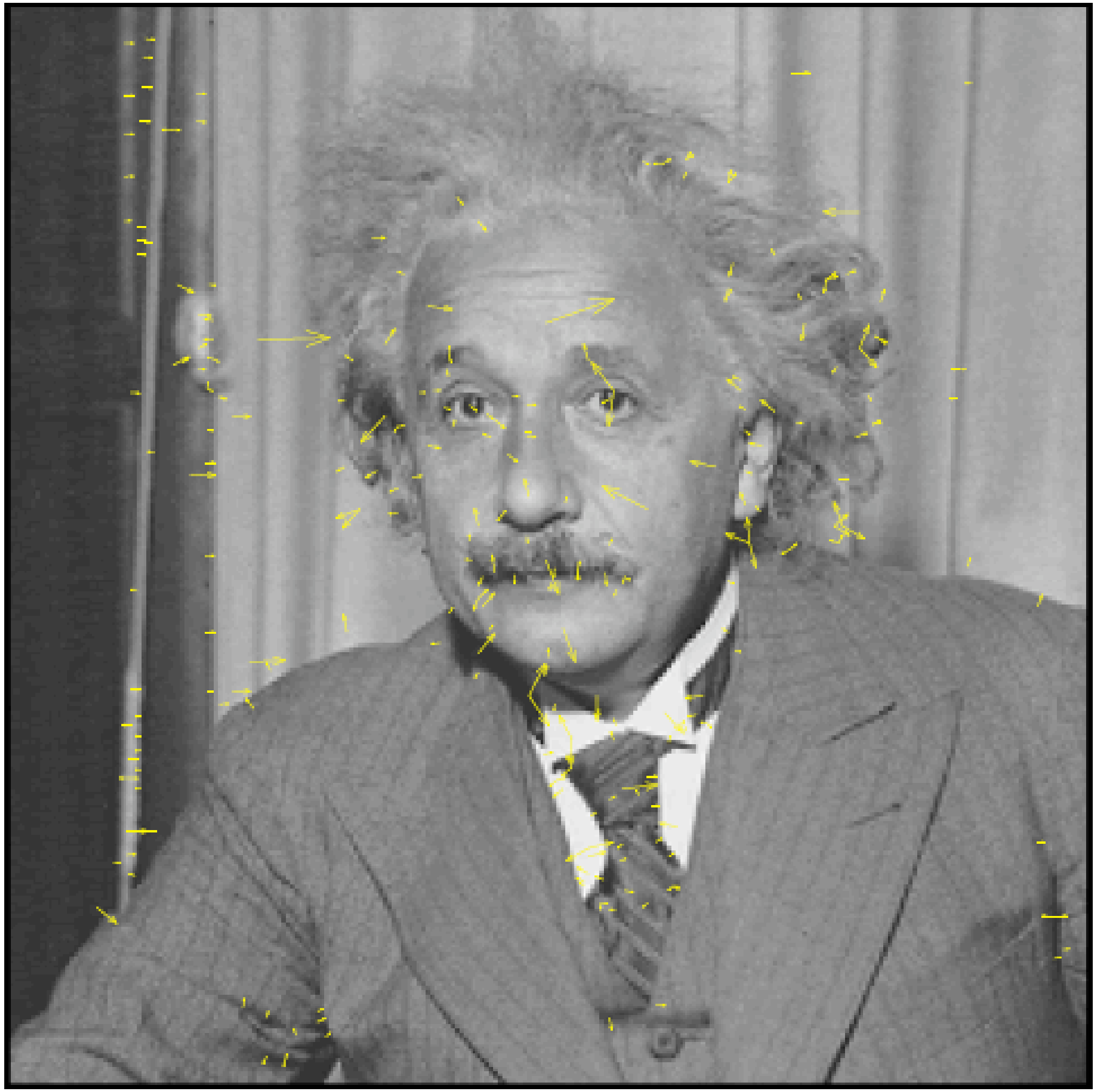
- 1) Scale-Space Extrema Detection
  - Left: Maxima of DoG across scales
- 2) Keypoint Localization
  - Middle: Remaining Keypoint after removal of low contrast point
  - Right: Remaining Keypoint after removal of edge responses

SIFT: 3) Orientation Assignment



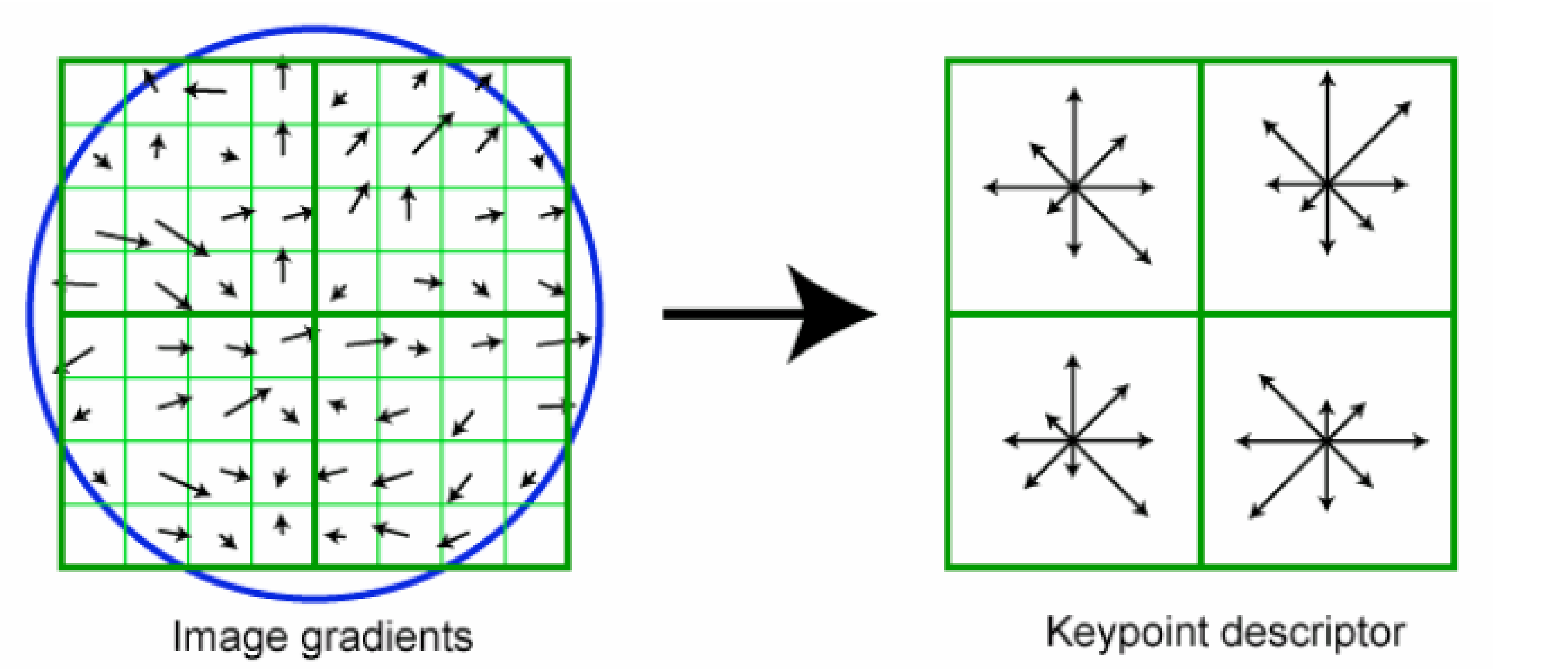
Orientation of keypoint is approximately 25 degrees

SIFT: 3) Orientation Assignment



SIFT: 4) Keypoint Descriptor

- Thresholded image gradients are sampled over 16x16 array of locations in scale space
- Create array of weighted orientation histograms

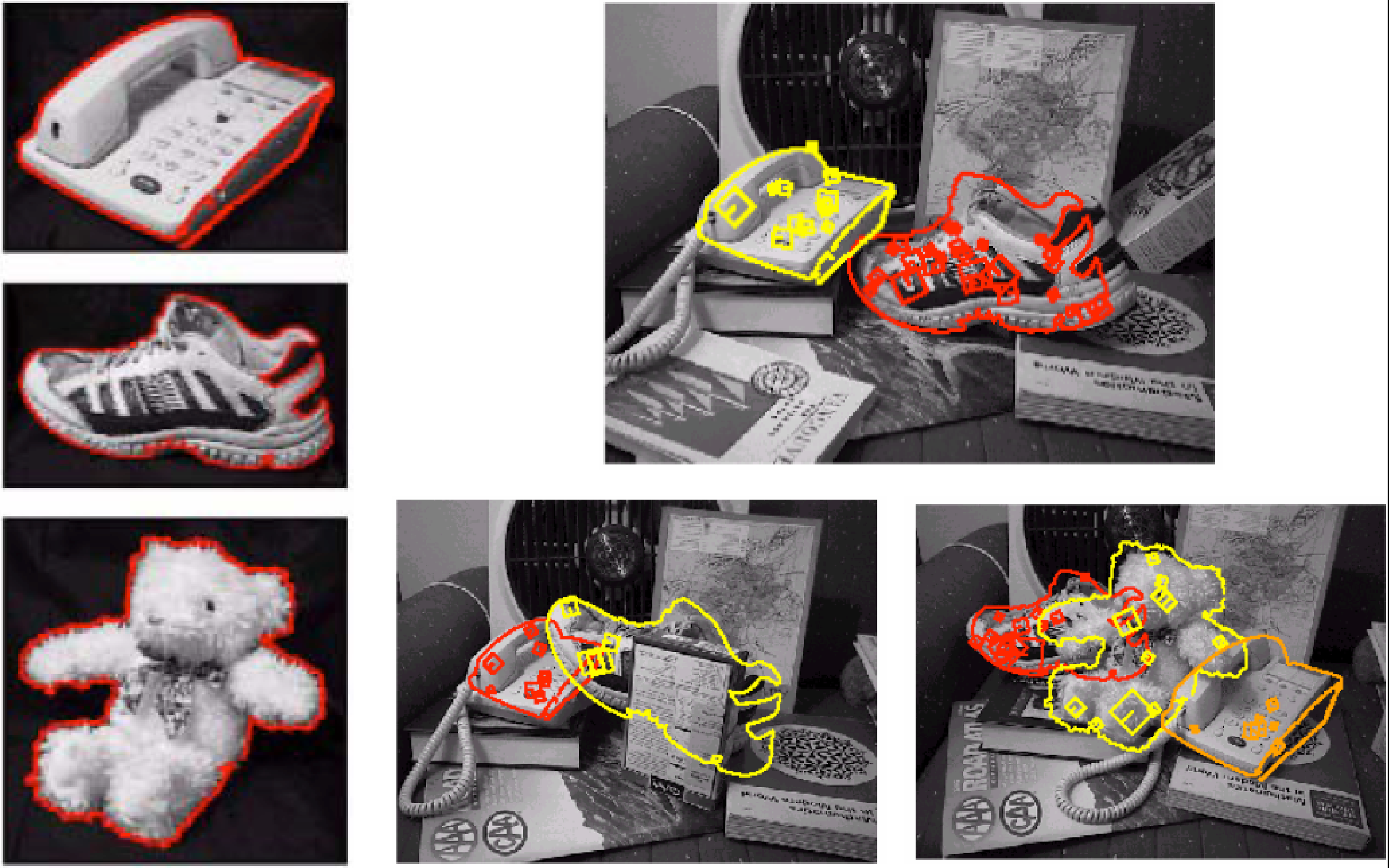


- 8 orientations x 4x4 histogram array = 128 dimensions

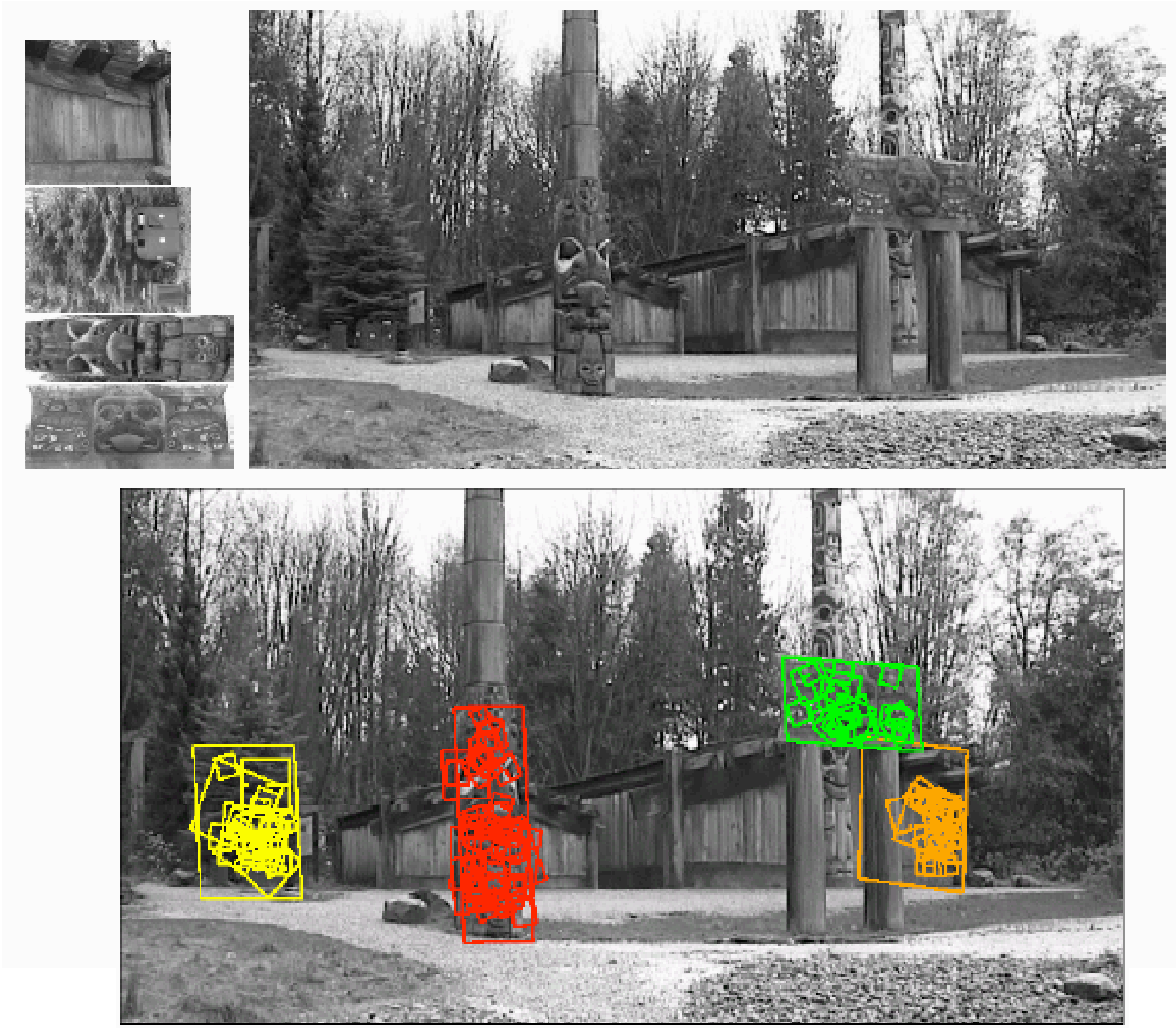
SIFT: Application

SIFT: Application

SIFT: Application



SIFT: Application



SIFT: Properties

- Invariances:
  - Scaling — Yes
  - Rotation — Yes
  - Illumination — Yes
  - Perspective Projection — Maybe
- Provides
  - Good localization — Yes

Outline

- Difficulties of Feature Extraction
  - Segmentation
  - Recognition in Images
  - Recognition in Document Images
- Feature Extraction/Detection on the whole Signal
  - Global Descriptor
  - Local Descriptor
- Feature Extraction/Detection on Objects
  - Object Localization / Extraction
  - Object Characterization
- Interest of Multi-Resolution
- Using Features

Feature Extraction/Detection on Objects

- **Features on Objects**
  - **Need to detect Objects in Images**
    - Segmentation
    - Difficult task
    - Recognize => Segment but Segment => Recognize
    - See section Segmentation difficulties
  - **Object characterization**
    - Features extracted on Segmented Objects

Outline

- **Difficulties of Feature Extraction**
  - Segmentation
  - Recognition in Images
  - Recognition in Document Images
- **Feature Extraction/Detection on the whole Signal**
  - Global Descriptor
  - Local Descriptor
- **Feature Extraction/Detection on Objects**
  - Object Localization / Extraction
  - Object Characterization
- **Interest of Multi-Resolution**
- **Using Features**

Features: Object Localization / Extraction

- Difficulties of Feature Extraction
- Feature Extraction/Detection on the whole Signal
- Feature Extraction/Detection on Objects
  - Object Localization / Extraction
    - Region Labeling
    - Edge Detection
    - Connected Components Labeling
    - Skeletonization
    - Handwriting extraction
    - Line-Segment Detection
  - Object Characterization
- Interest of Multi-Resolution
- Using Features

Object Extraction: Region

- **Criteria: Luminosity, Color, Texture…**
  - **Homogeneous Regions / Criteria**
  - **Significant Criteria Difference**
- **Simple Inside, no small holes**
- **Regular Edges**

Region: Approaches

- Measure Space Partition

- Region Growing
  - Find a Germ
  - Aggregation

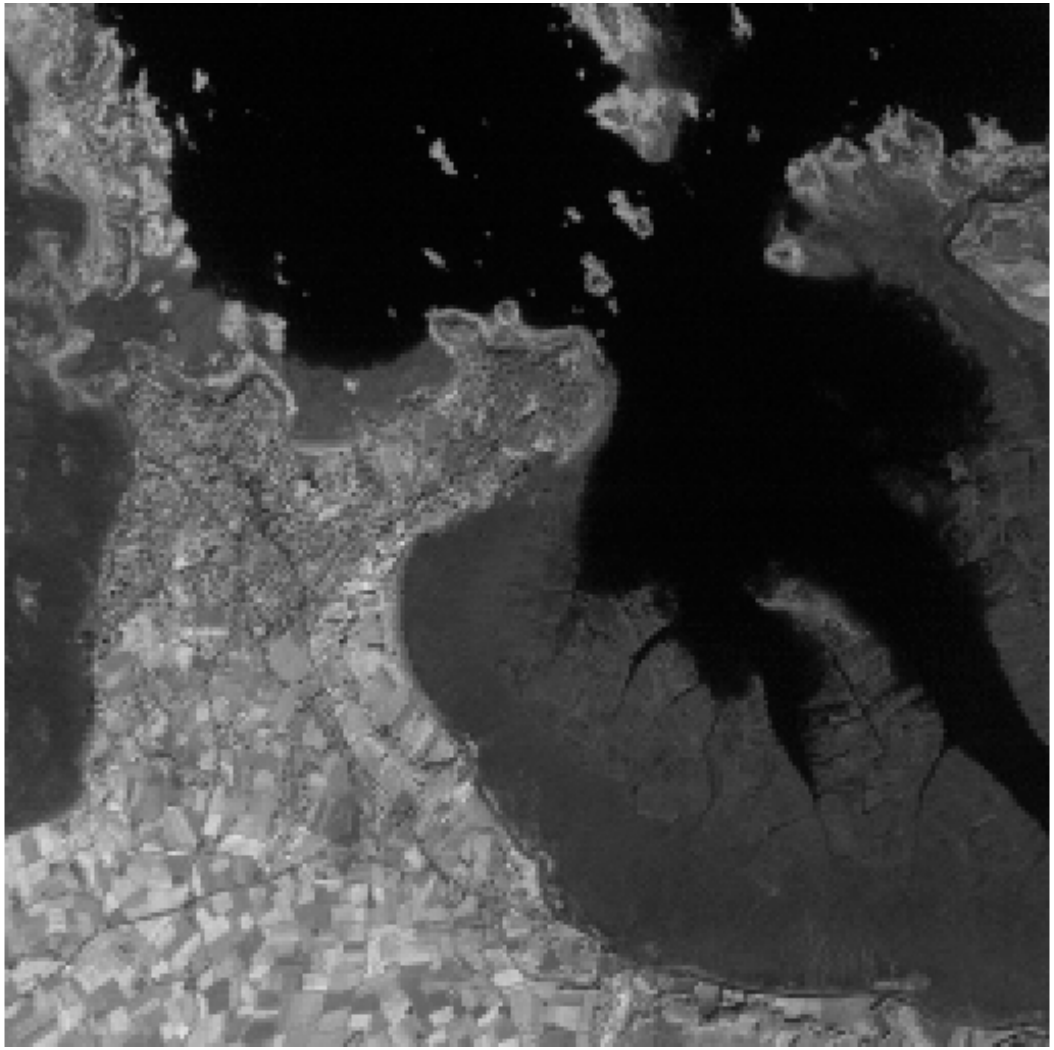
- Edge-Region Duality
  - Edge Detection
    - See next section
  - Fill inside

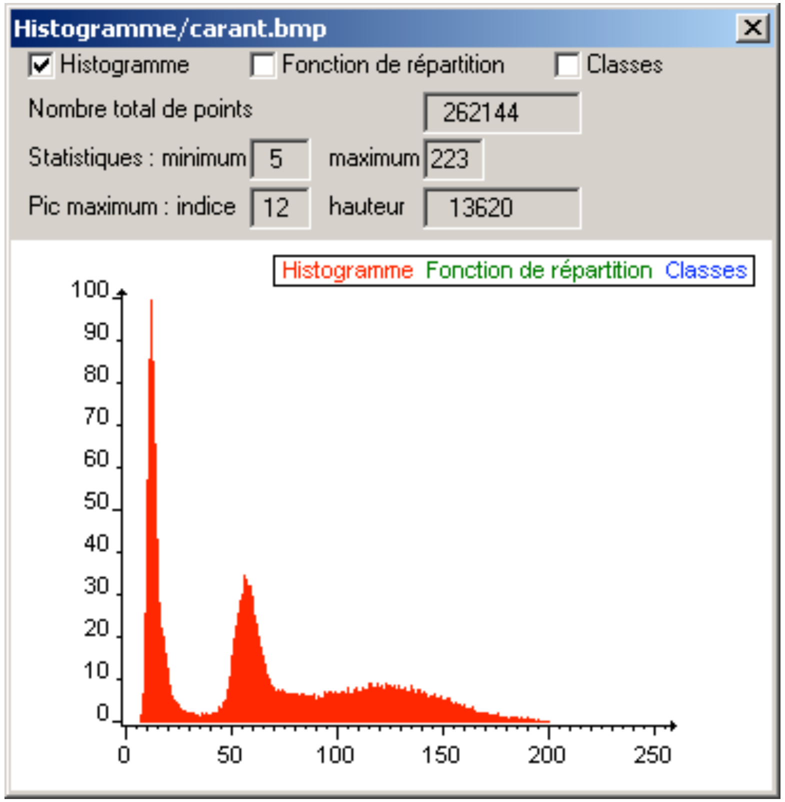
- Using Time in Image Sequence
  - Object Movement can be detected
  - What Changed from one image to another
  - Follow Object

Region: Measure Space Partition

- **Measure Space**
  - Luminosity
  - Color
  - Texture
  - Gradient
- **Measure Space Partition**
  - For example: black/white
- **Pixel Labeling by a class name**
  - Connected Components of the same name

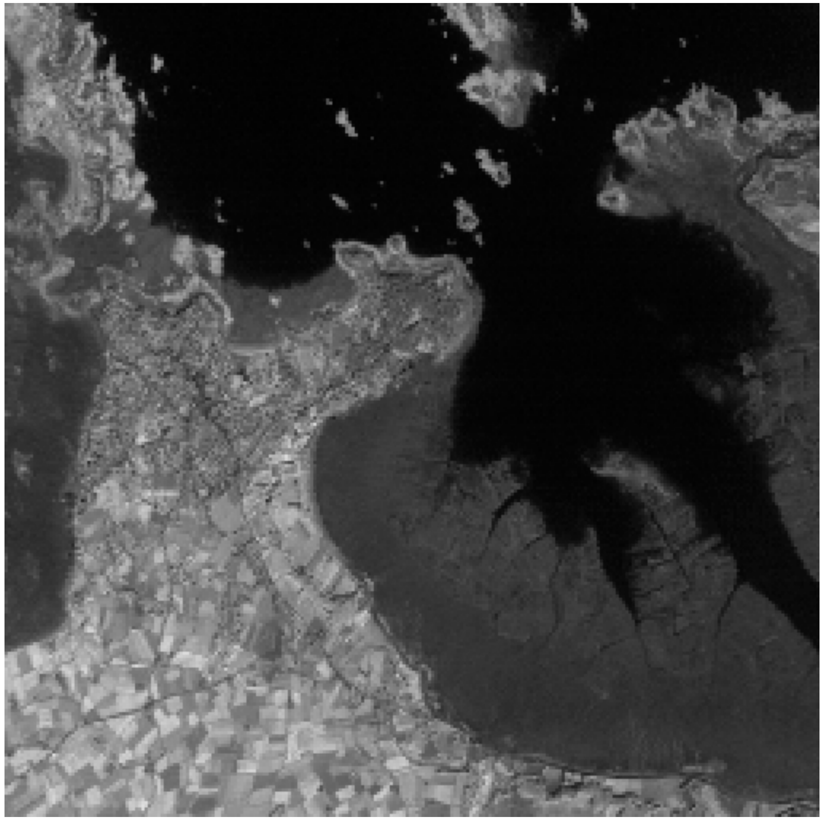
Luminosity Partition

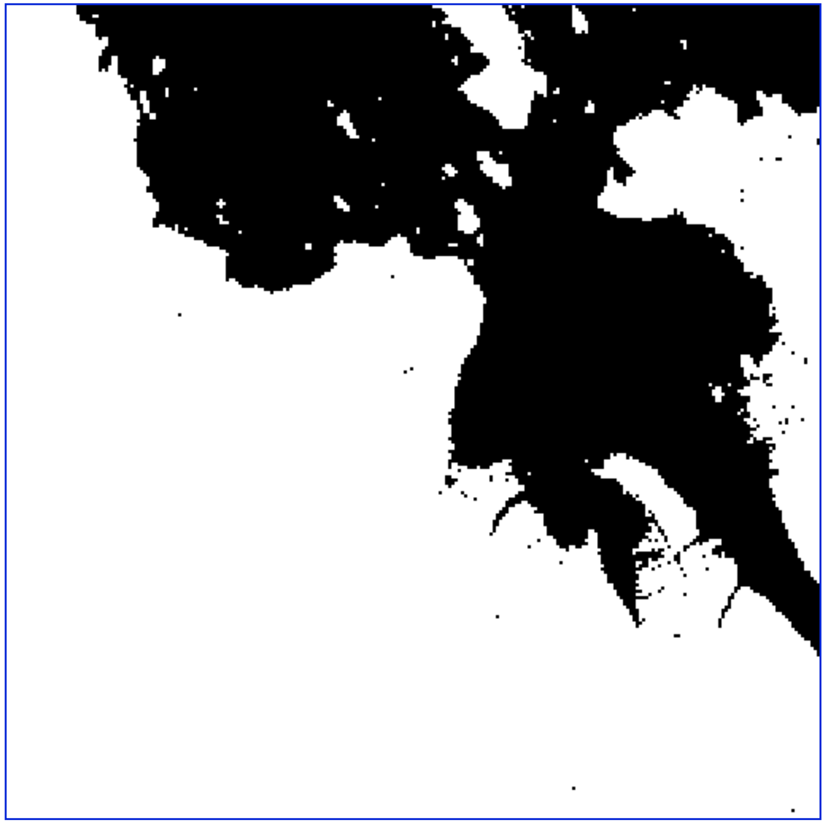




Seuil manuel
0-40 eau
41-98 estran
>98 terre

Luminosity Partition

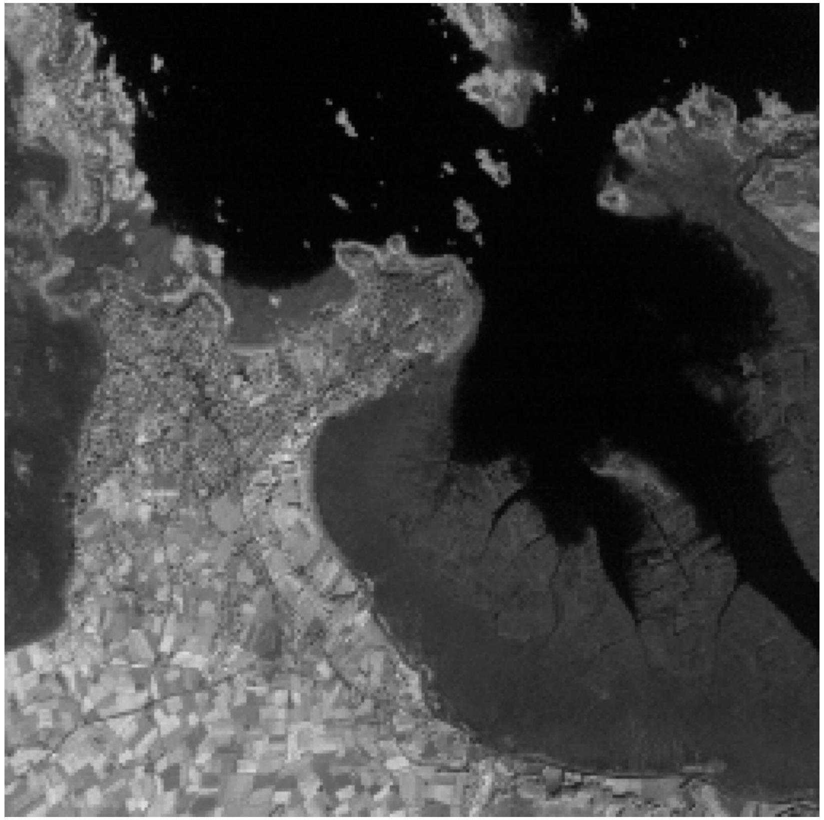


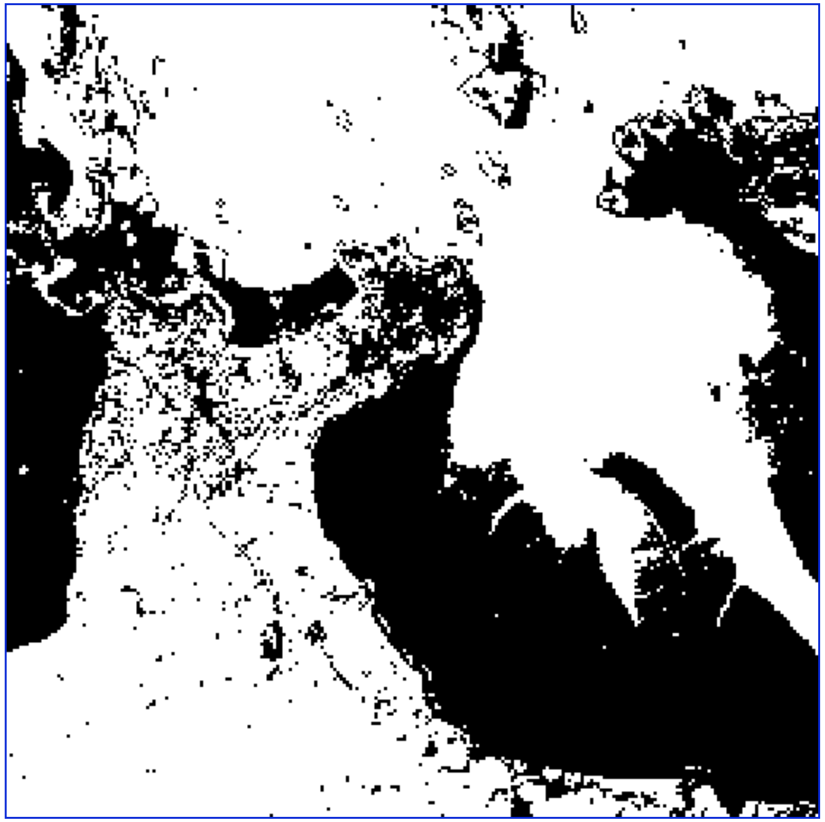


eau < 40

Noter les pixels isolés

Luminosity Partition

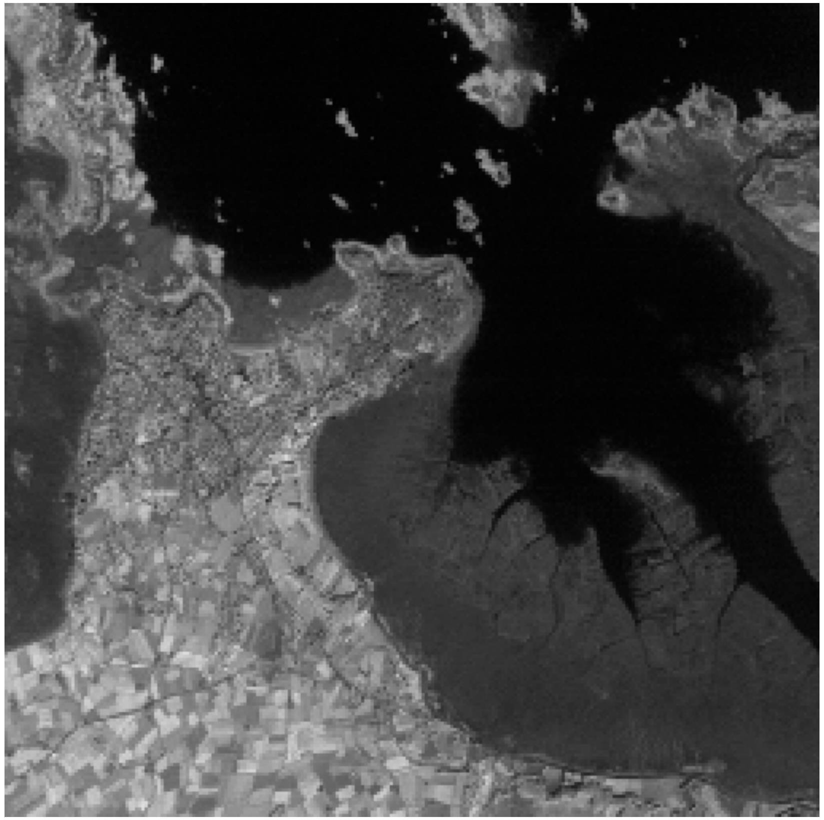


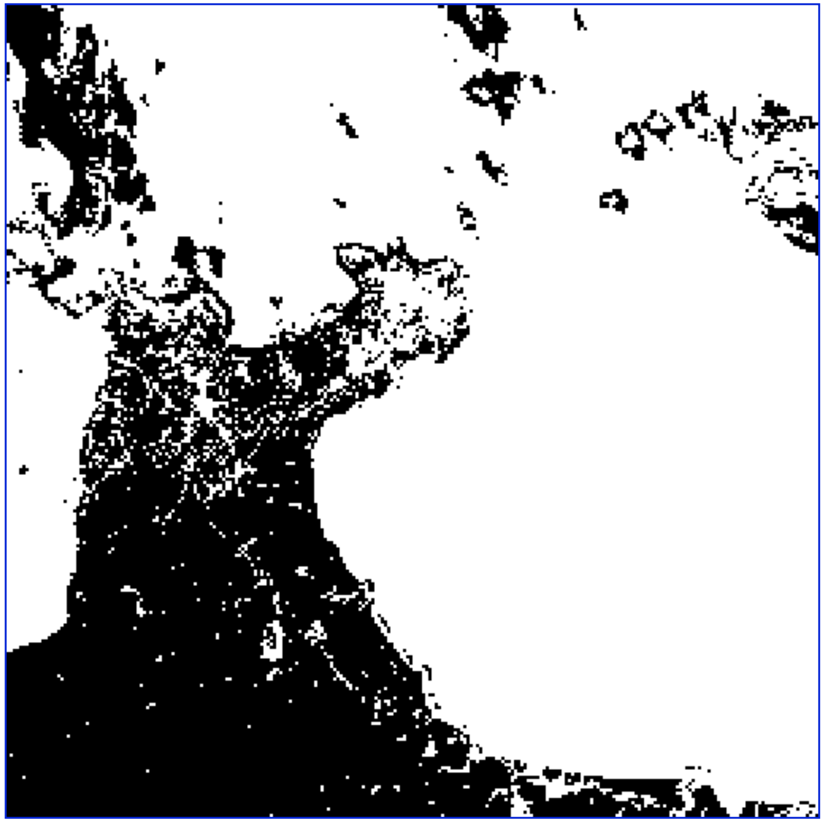


41< estran < 98

Besoin d'un critère en plus

Luminosity Partition

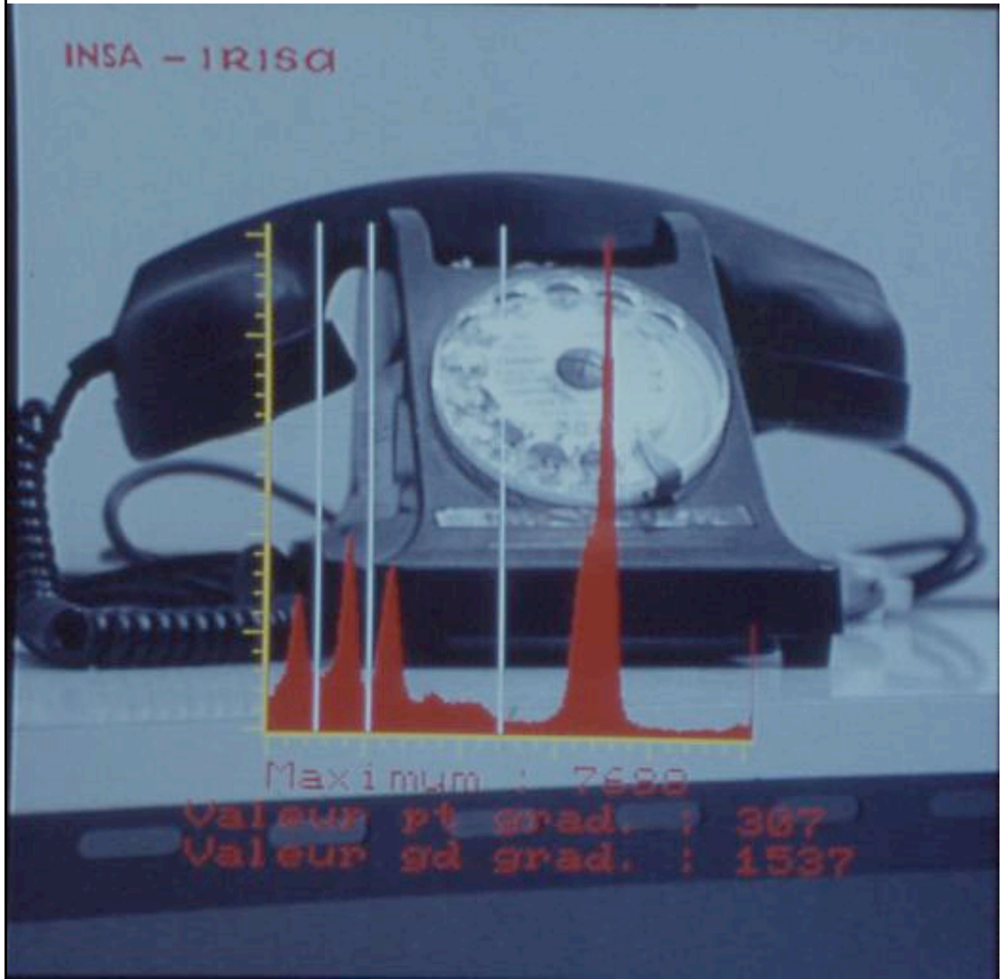


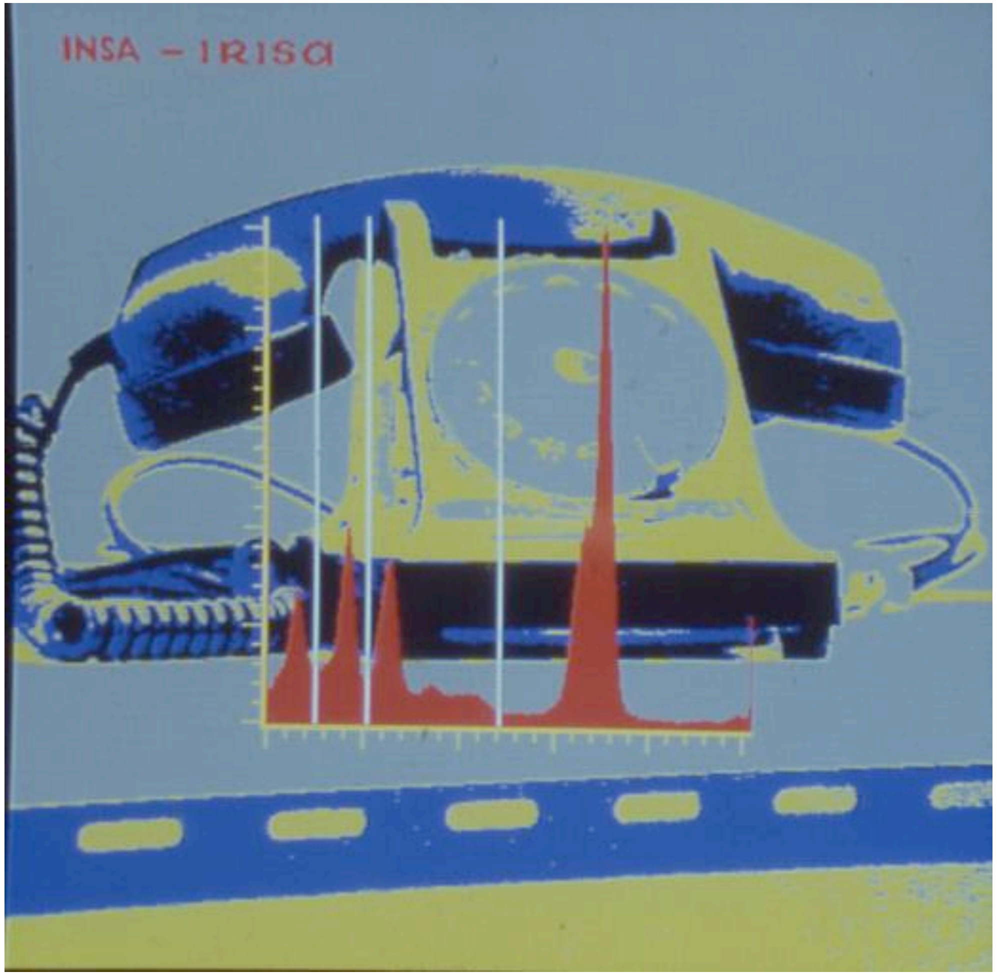


terre > 98

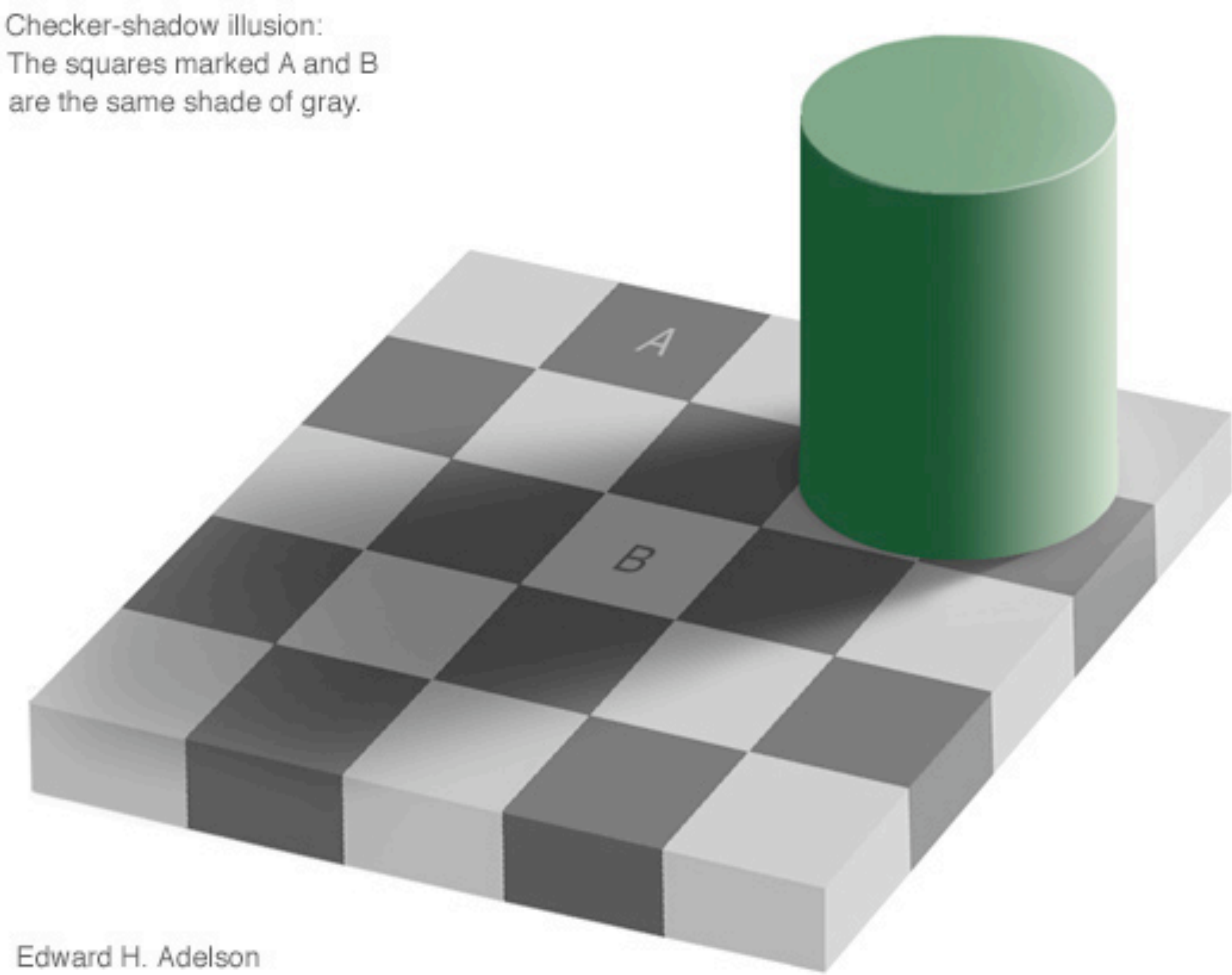
Besoin d'un critère en plus

Multi-Class Luminosity Partition

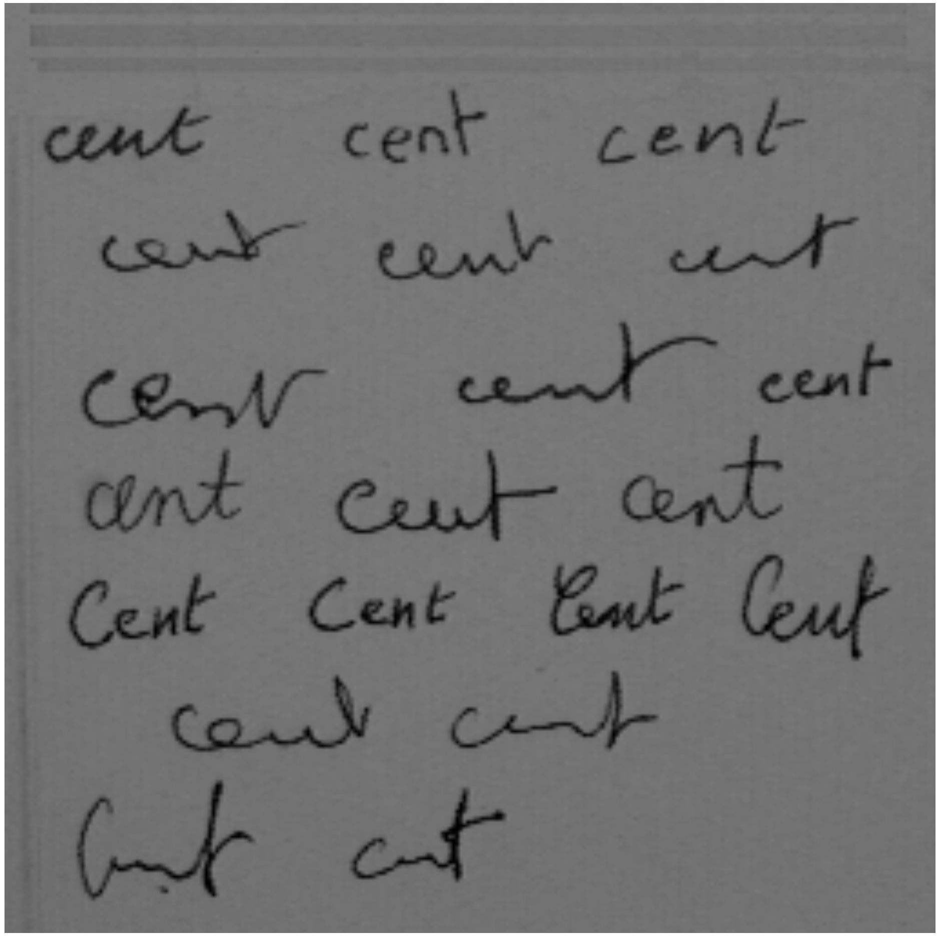


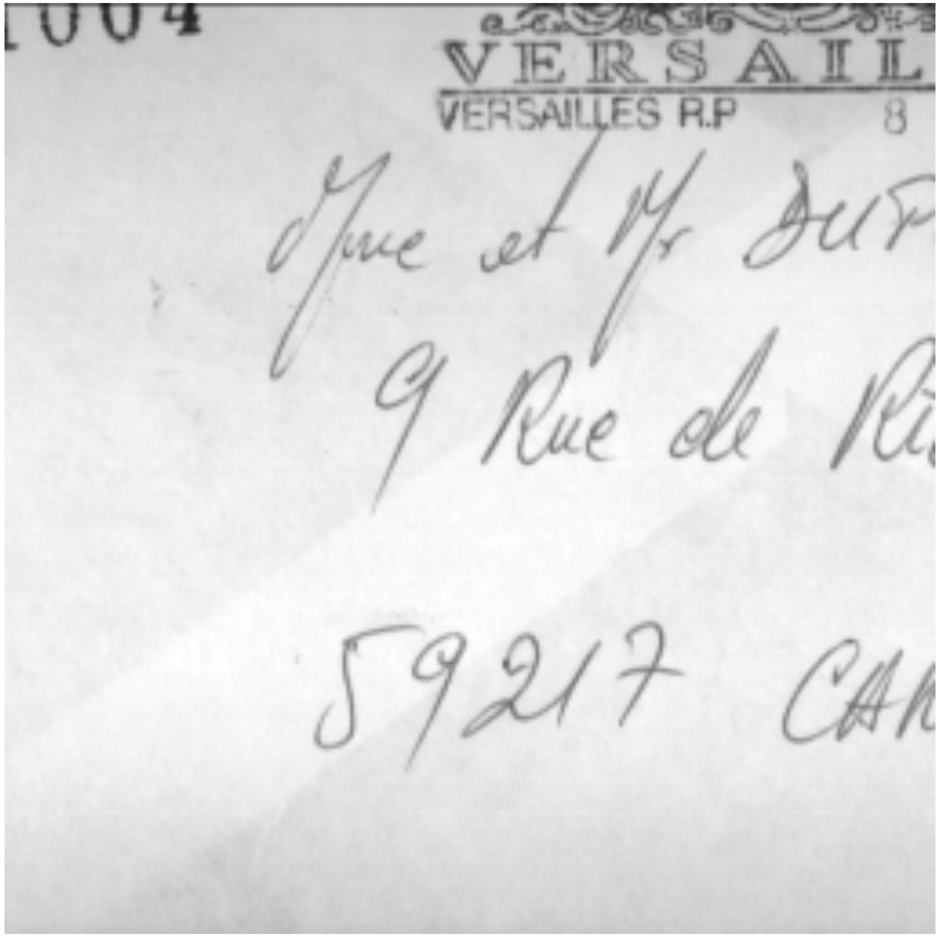


Difficult Task

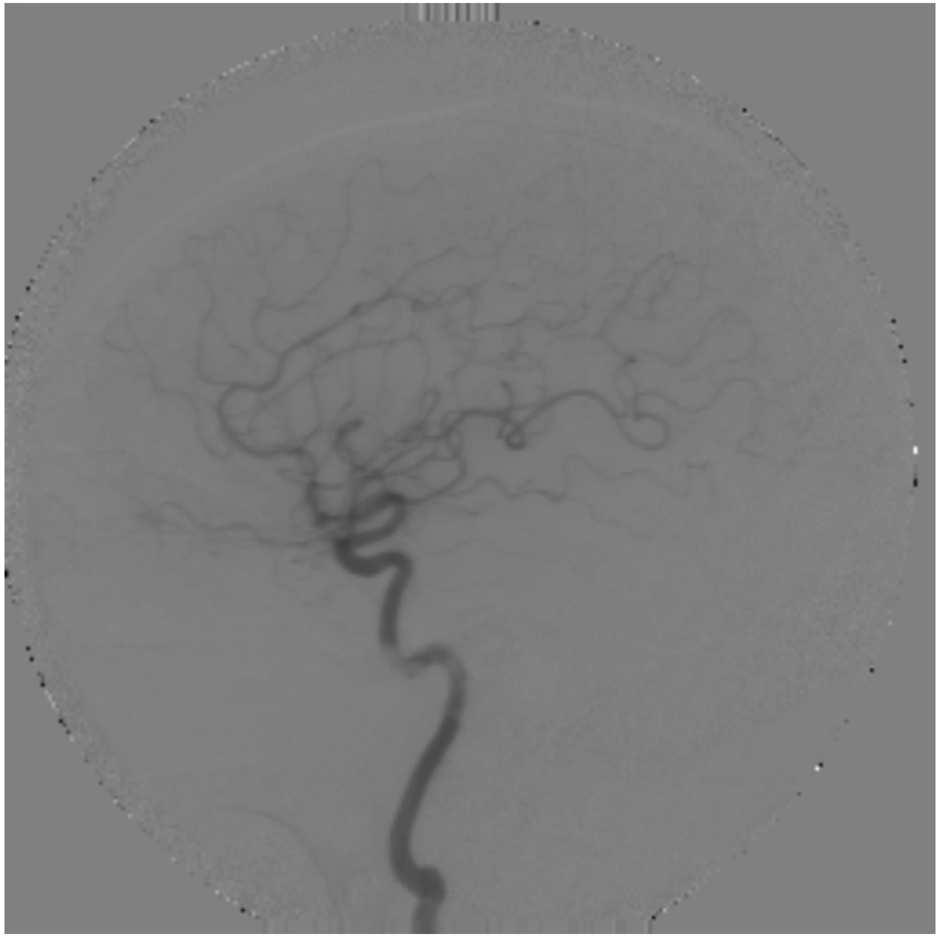


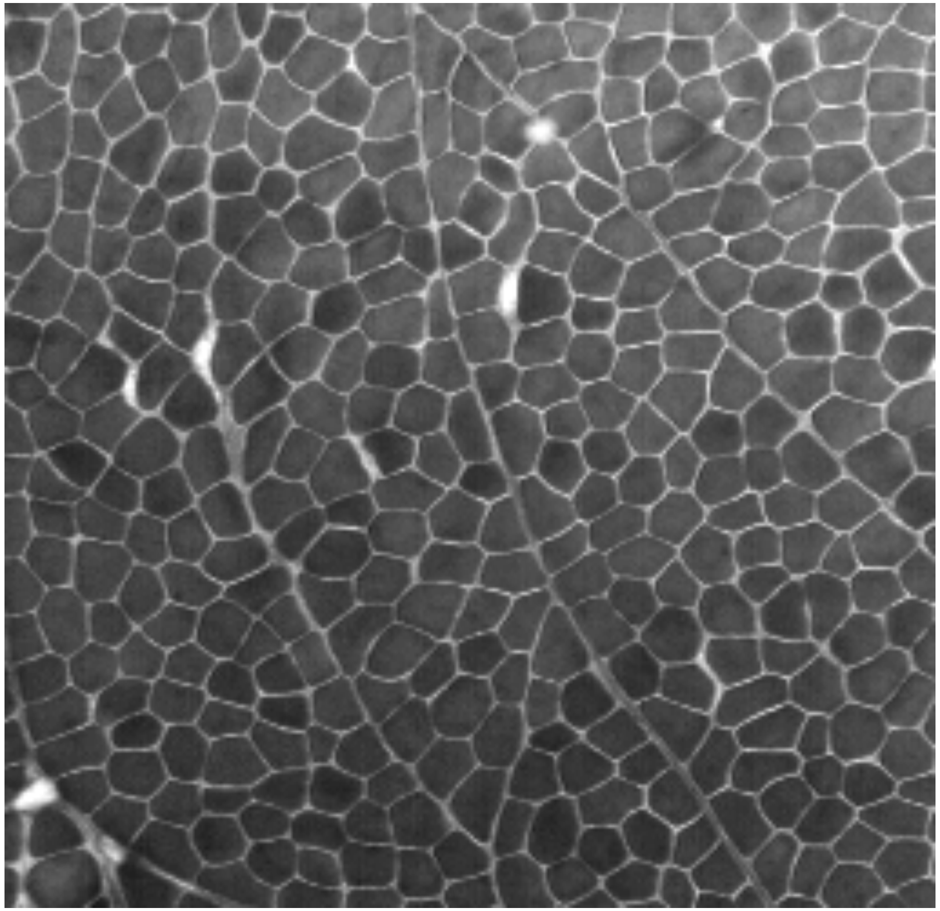
Binarization Possible on Text Images





Binarization Possible on Bio-Medical Images

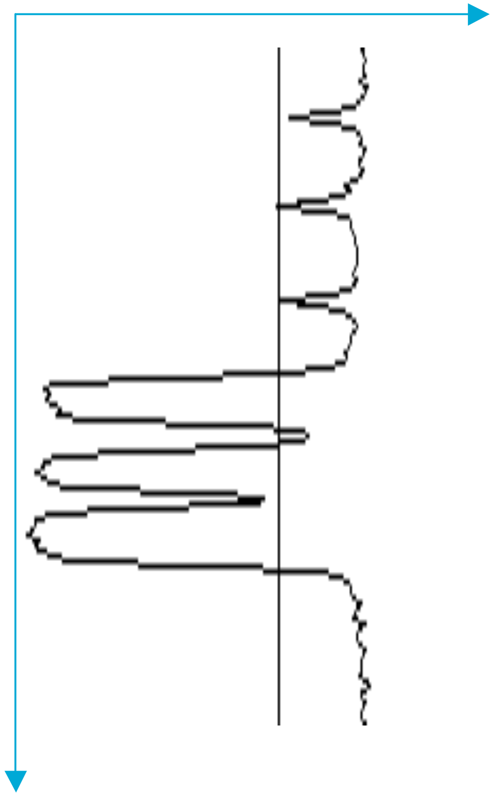


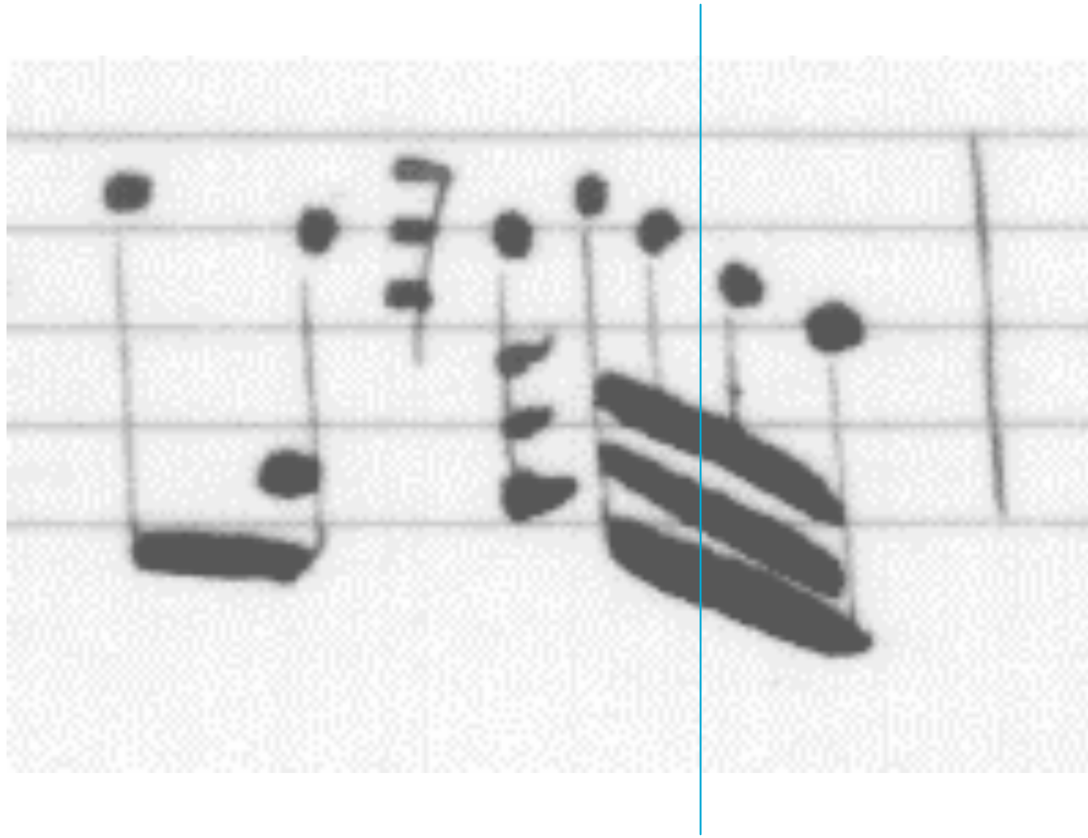


Binarization?

Threshold

Luminosité





Threshold

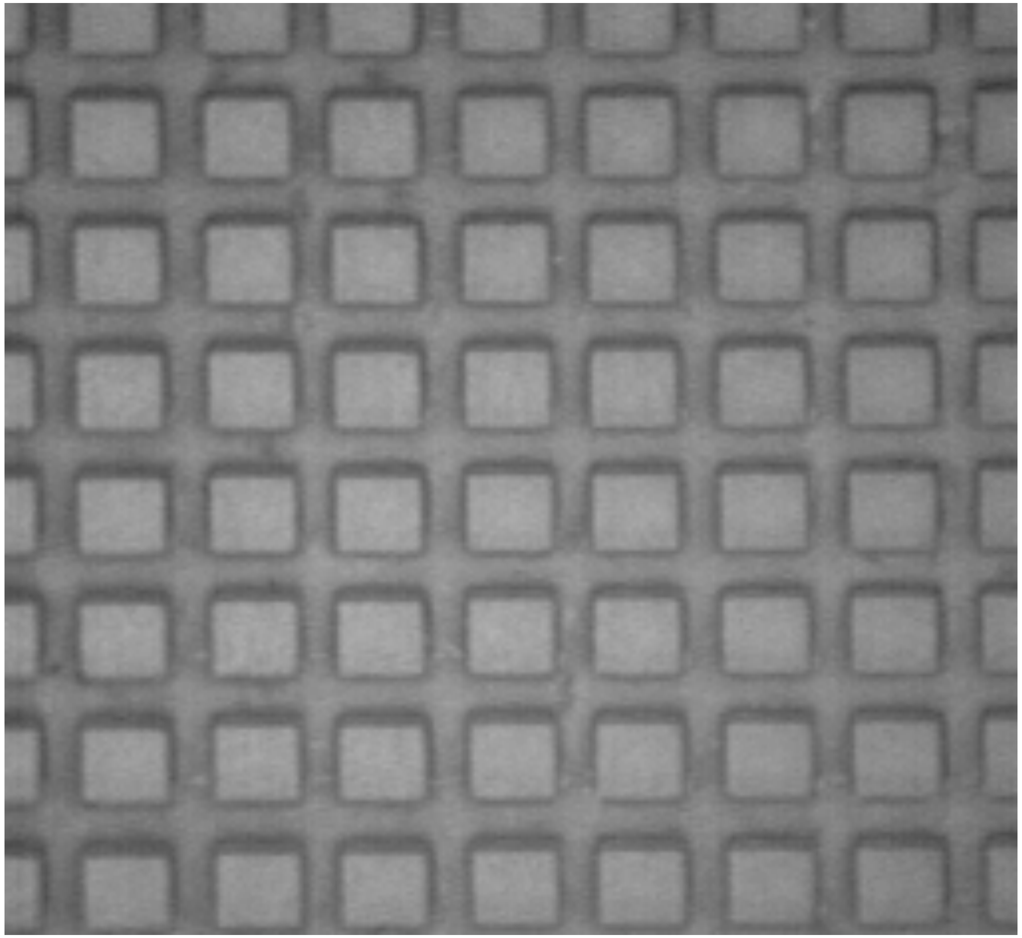
210

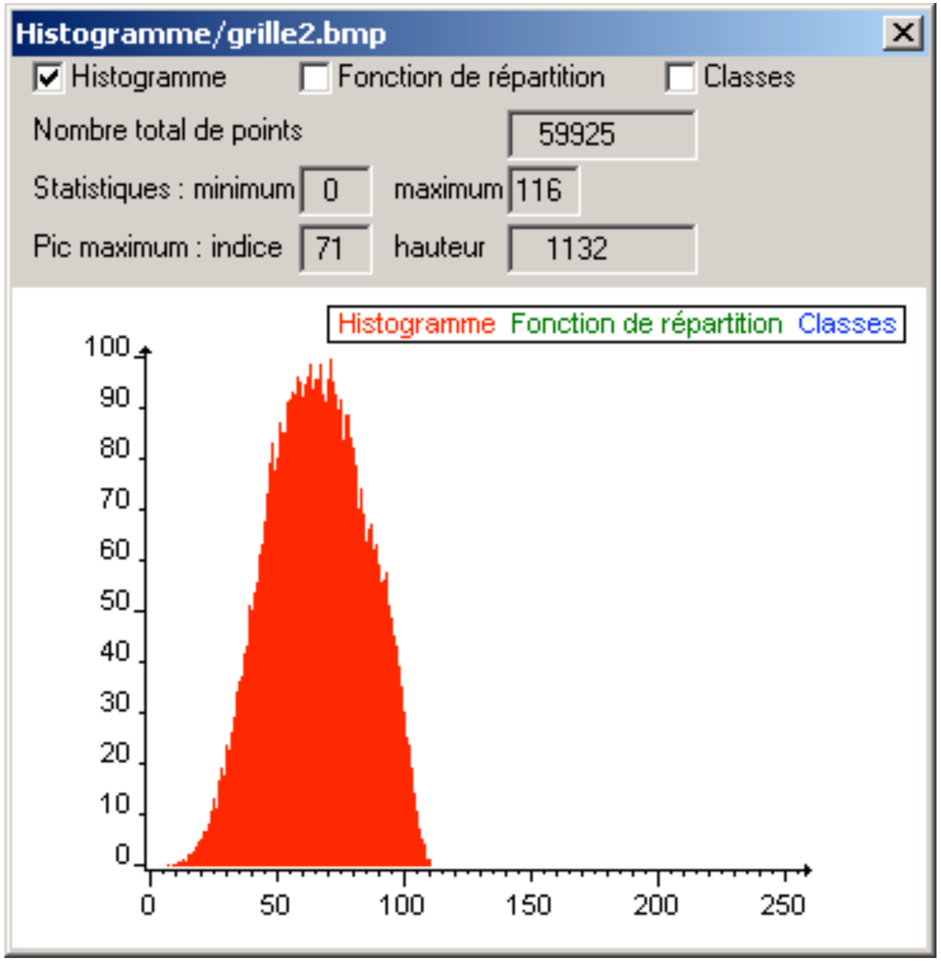
195

110

Seuillage adaptatif

Mono-modal Histogram

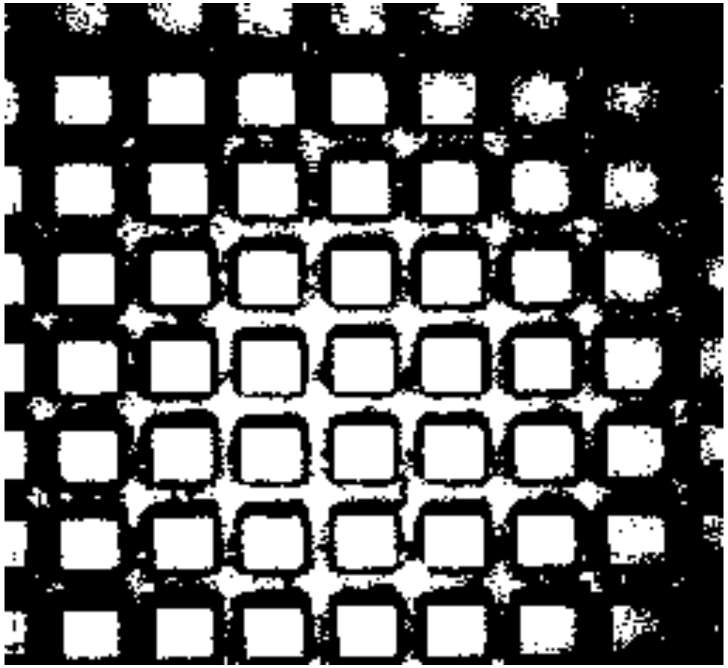




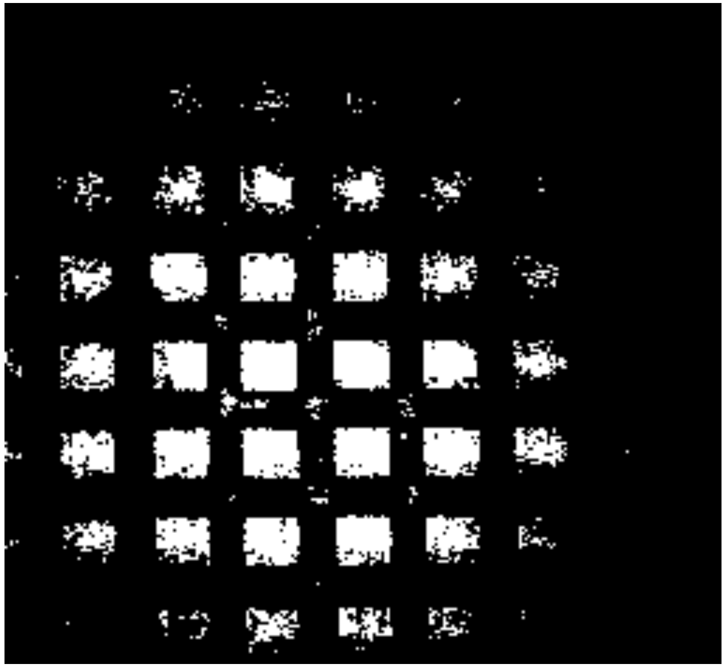
Bi-Modal Image

Mono-Modal Histogram

Mono-modal Histogram



Threshold at 71, max peak



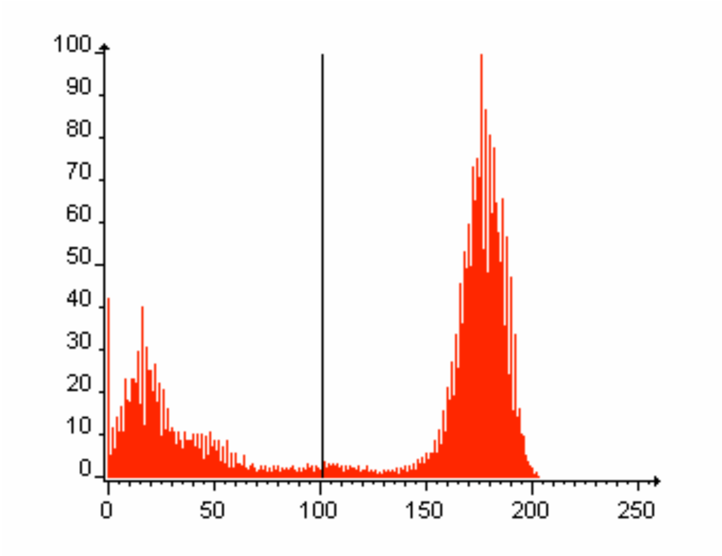
Too high threshold

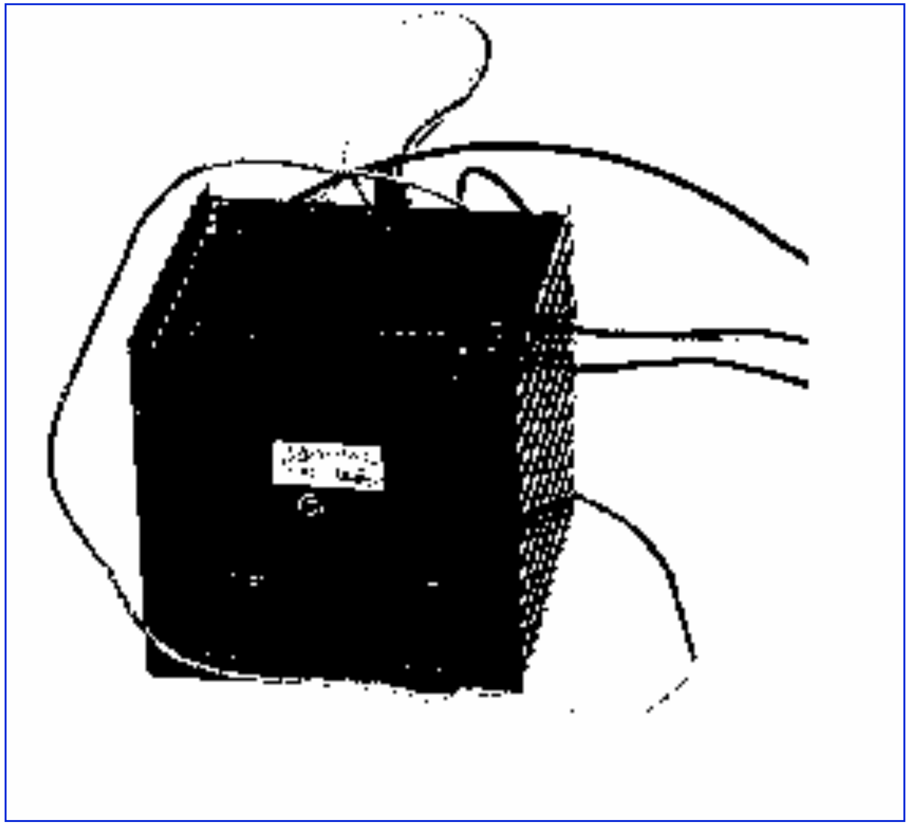
Conclusion: non uniform lightning

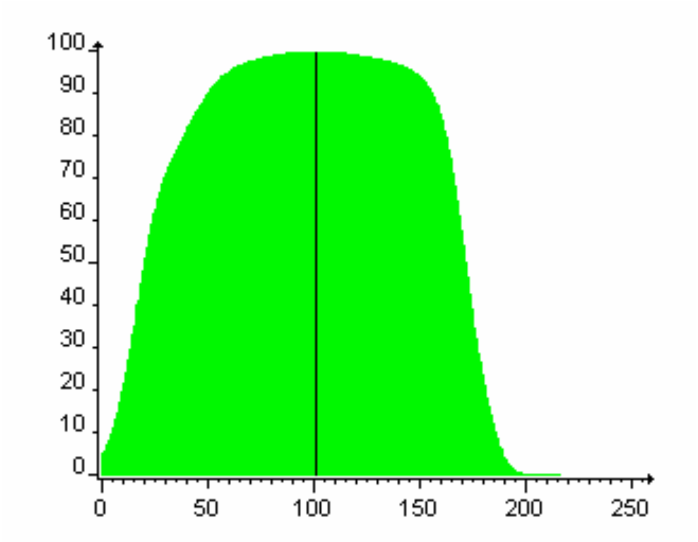
Choosing Threshold

- **Global Methods**
  - Interclasse Variance : Otsu
  - Histogram Concavity Analysis
  - Entropy : Pun
  - Moments
  - Histogram Modification: Mason
  - Transition Matrix: Deravi
  - …
- **Local Methods**

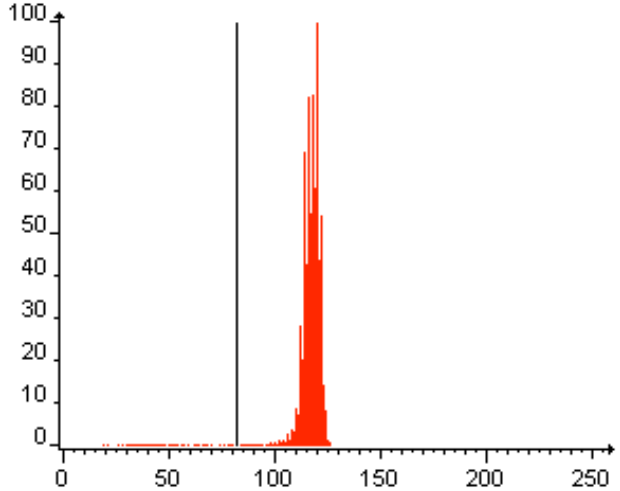
Binarization: Otsu Method

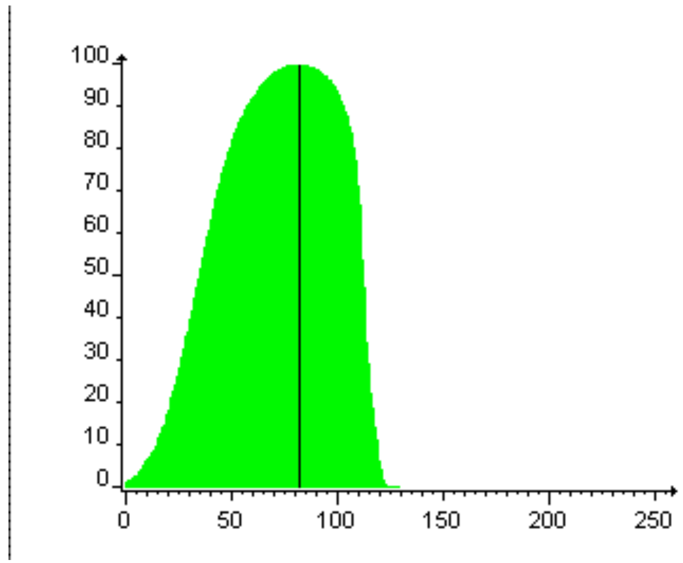


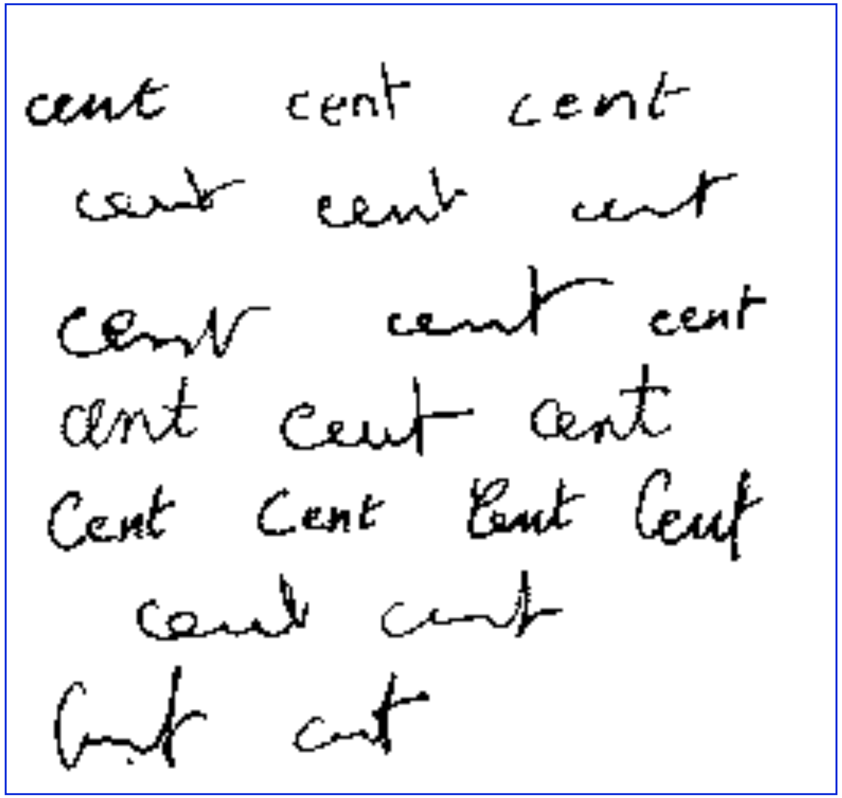




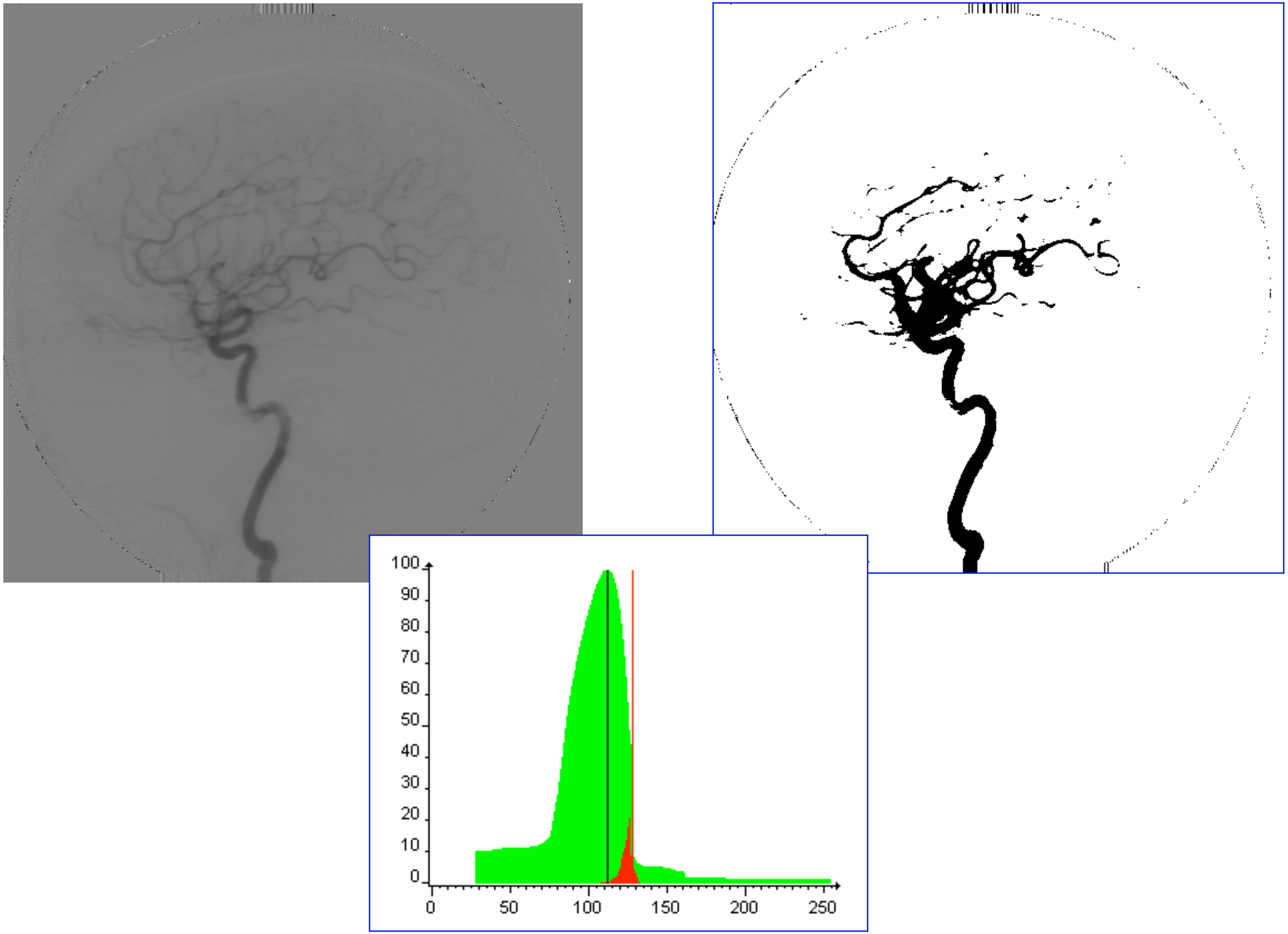
Binarization: Otsu Method





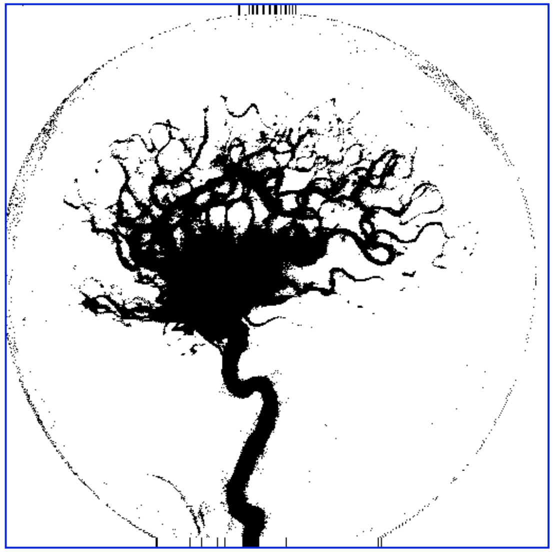


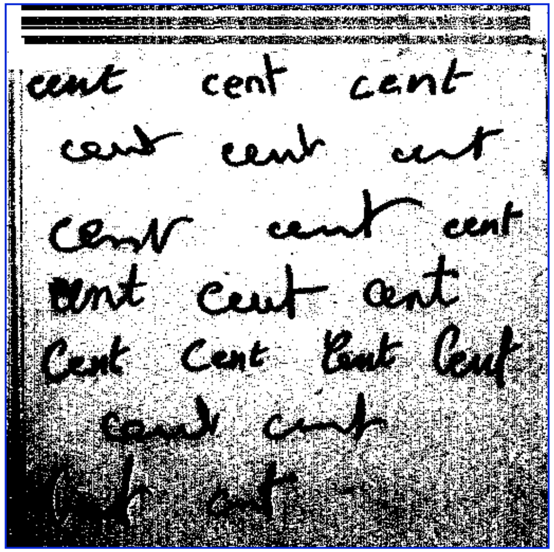
Binarization: Otsu Method



Binarization: PUN Method

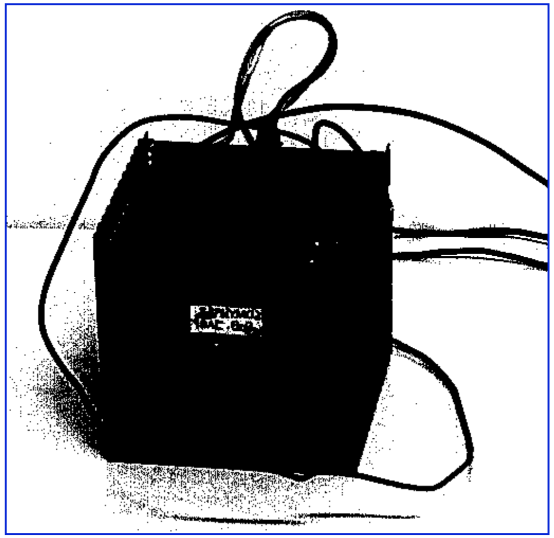
121

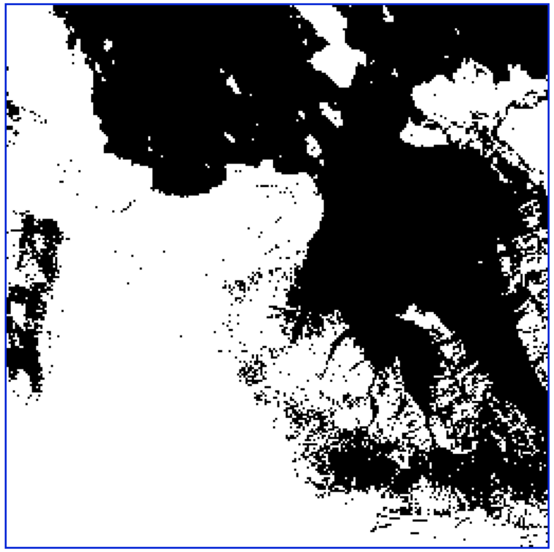




116

157

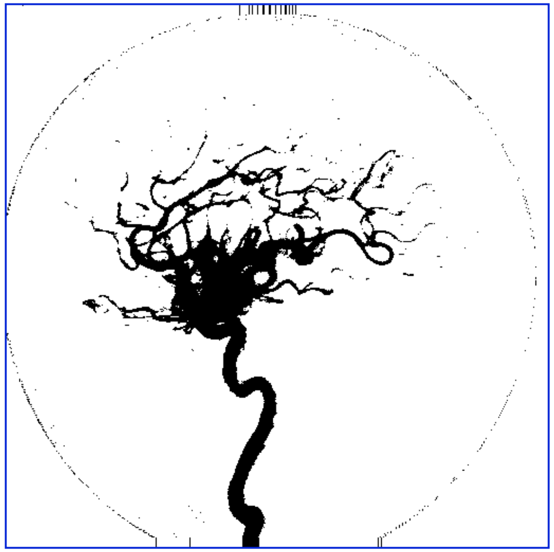


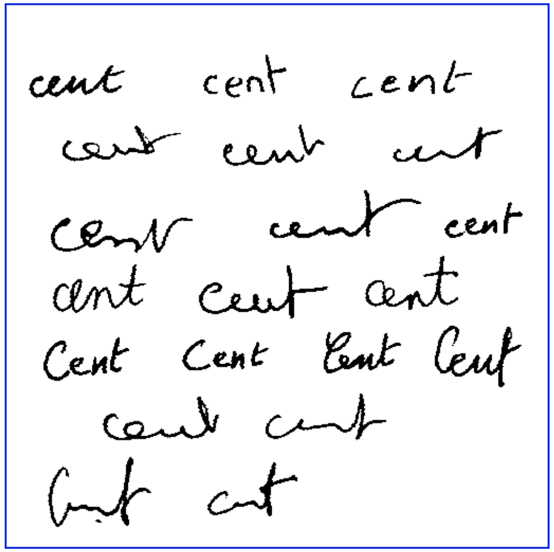


57

Binarization: Moments Method

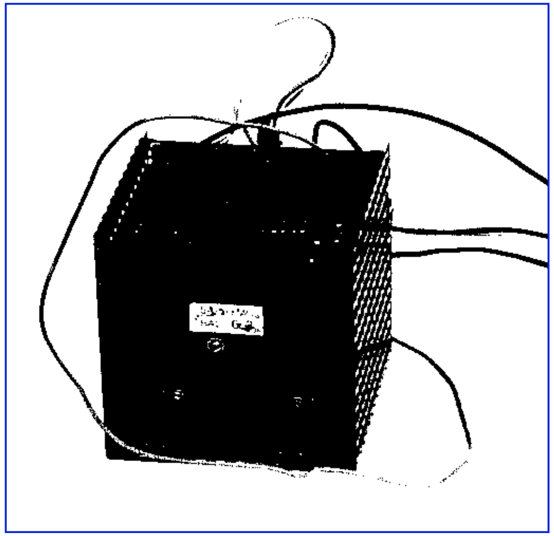
117

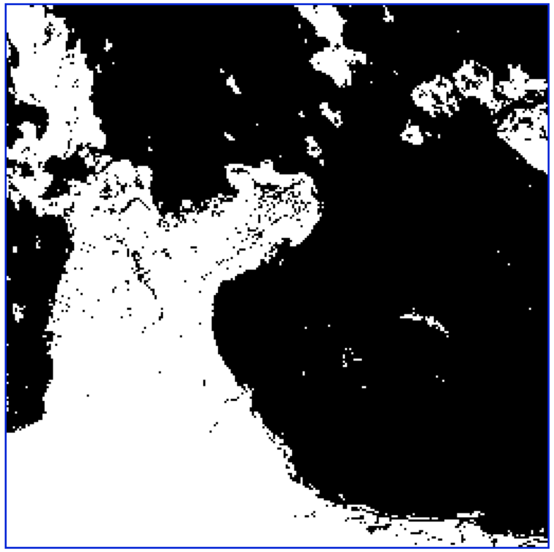




85

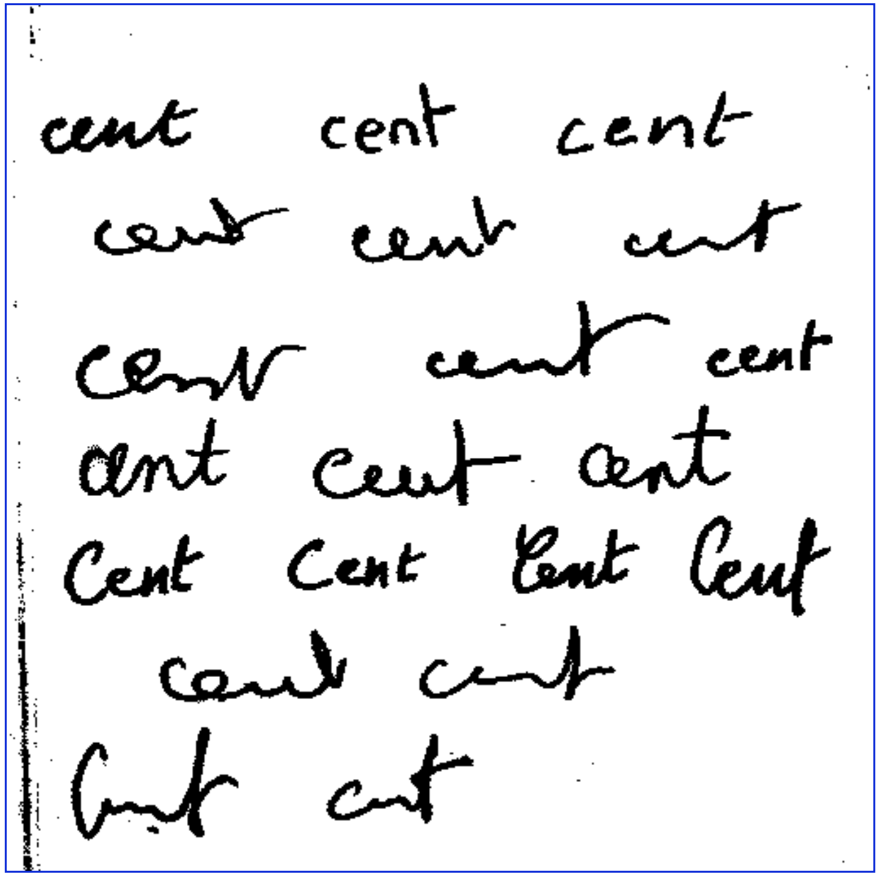
100



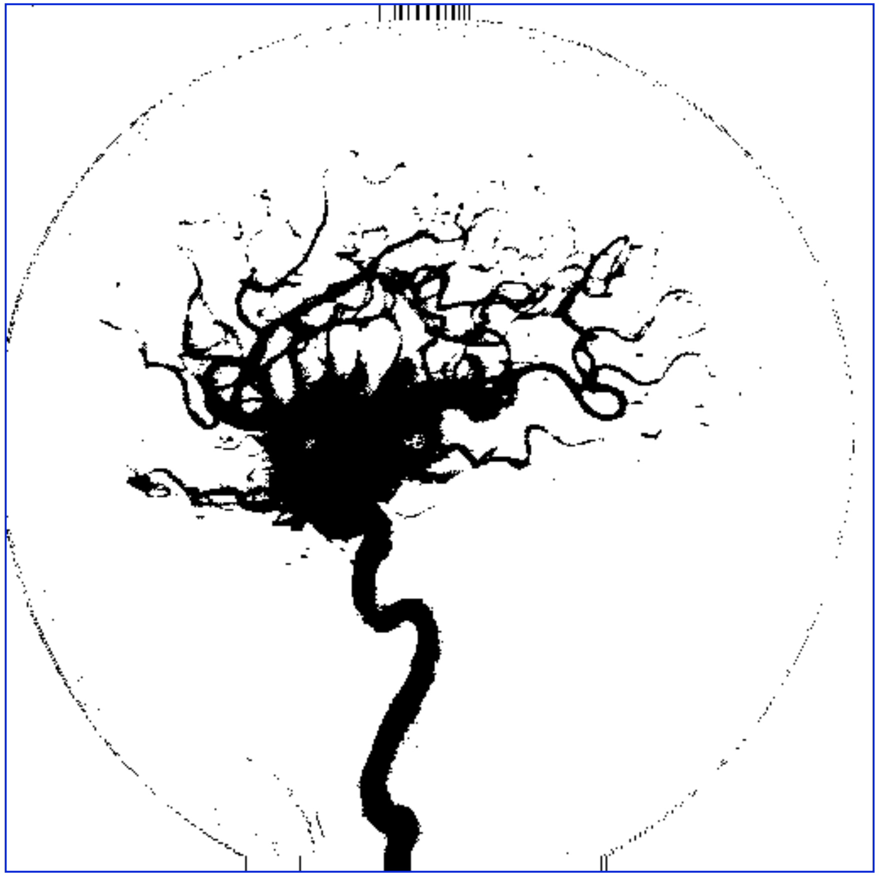


77

Binarization: Deravi



106

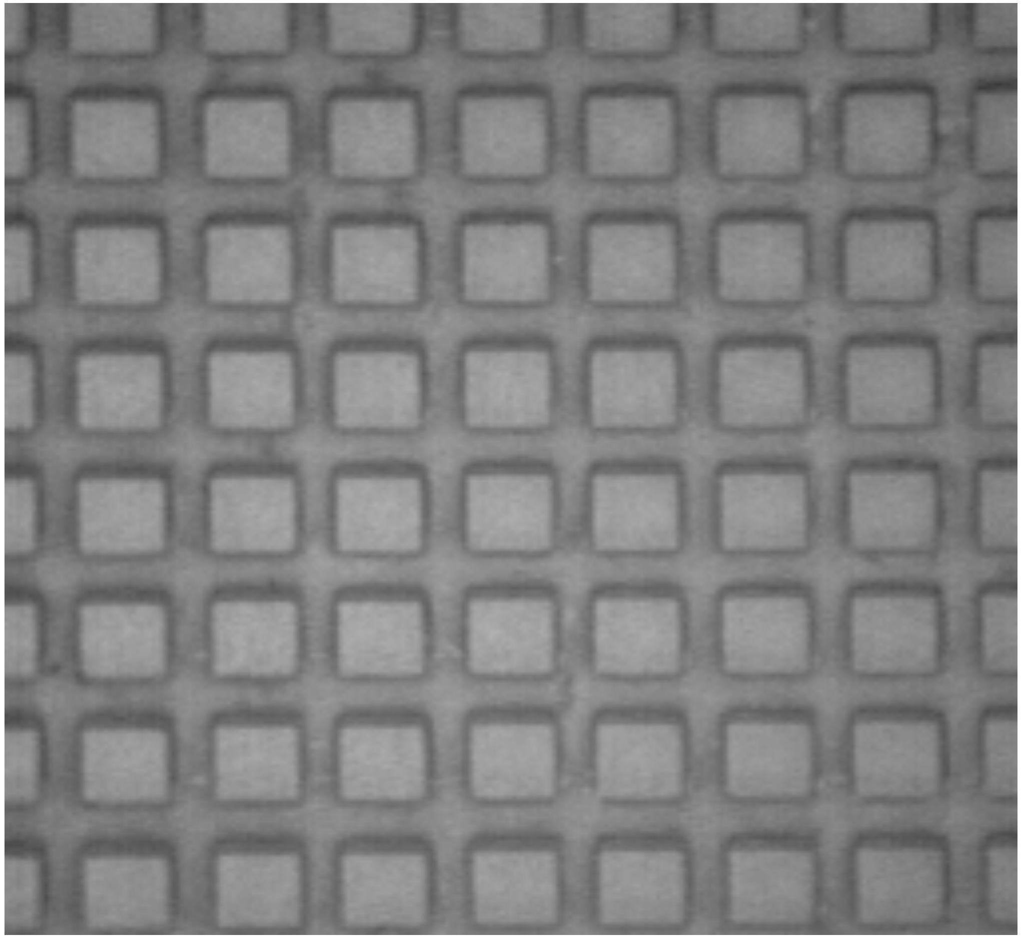


119

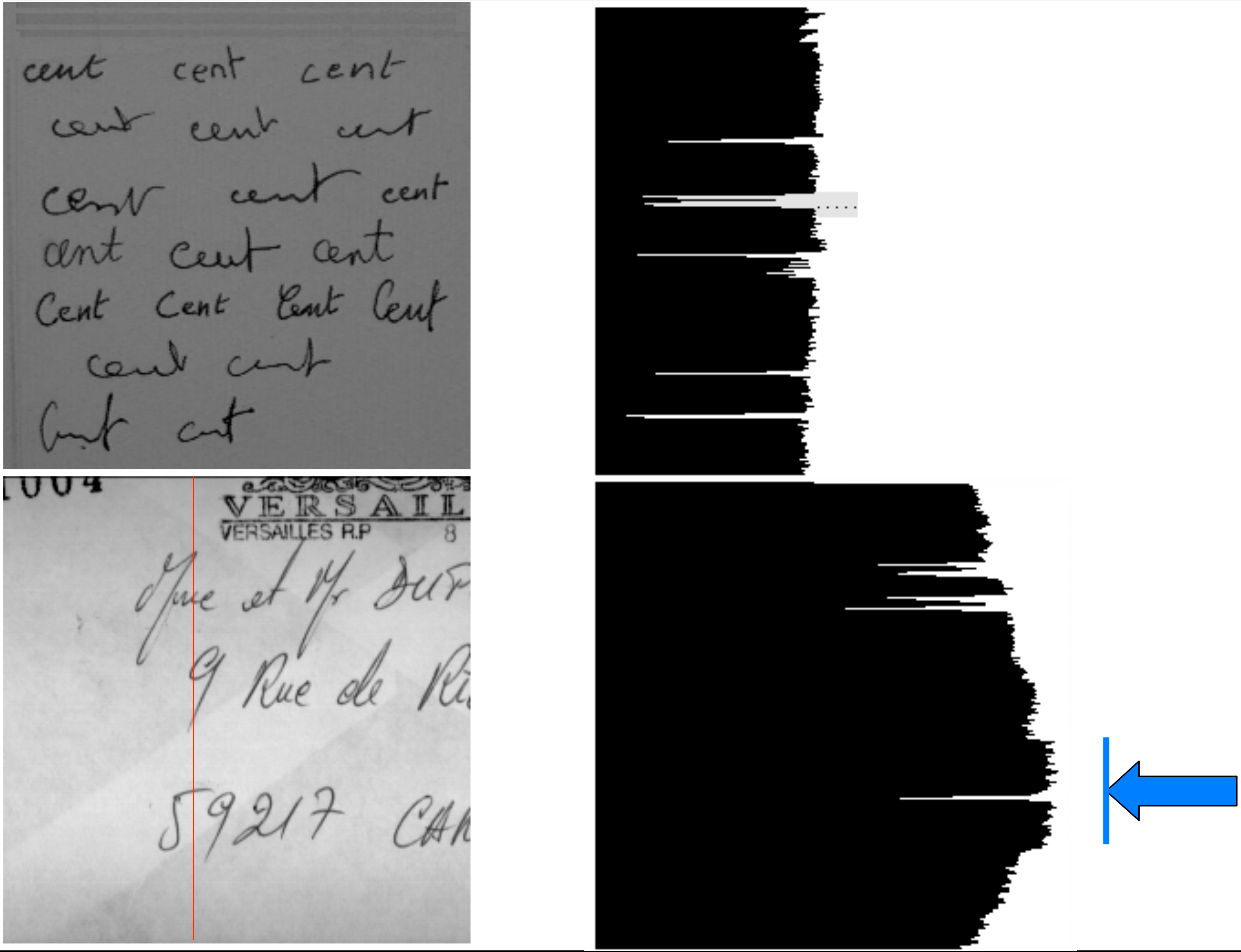
Binarization: Global Methods

- **Lot of Different Methods**
- **None really usable**
- **Problem**
  - **Global Threshold**

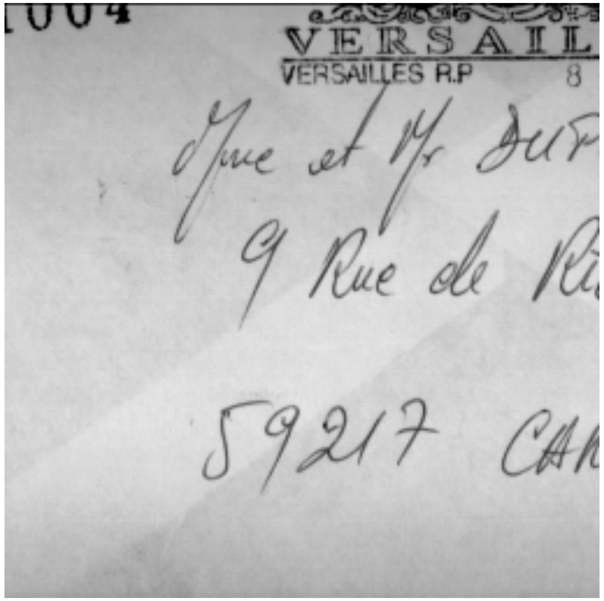
Non-uniform Lightning

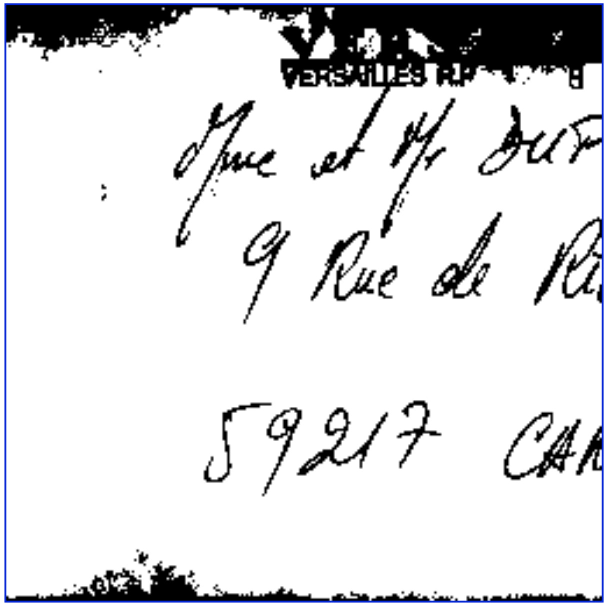


Non-uniform Lightning: Column



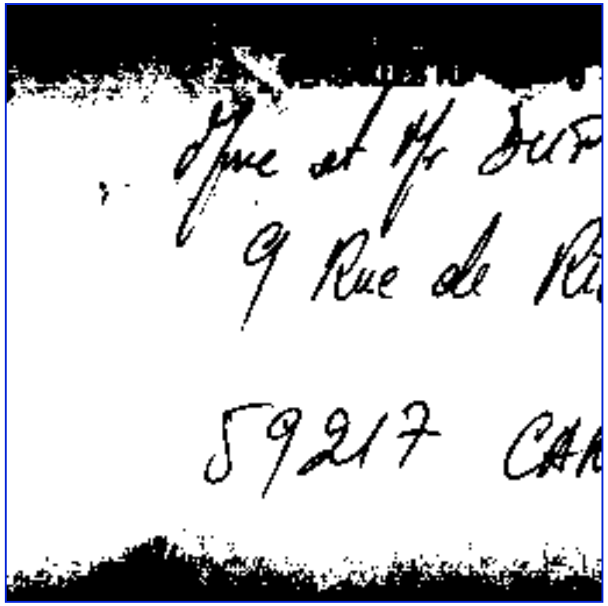
Non-uniform Lightning: Column

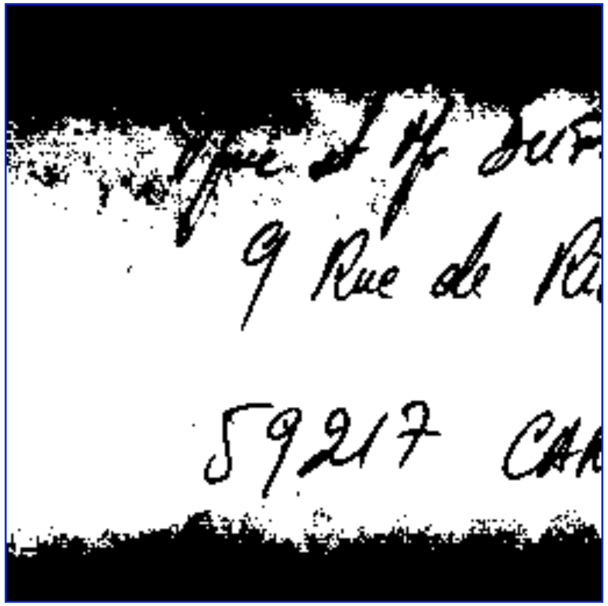




170

180



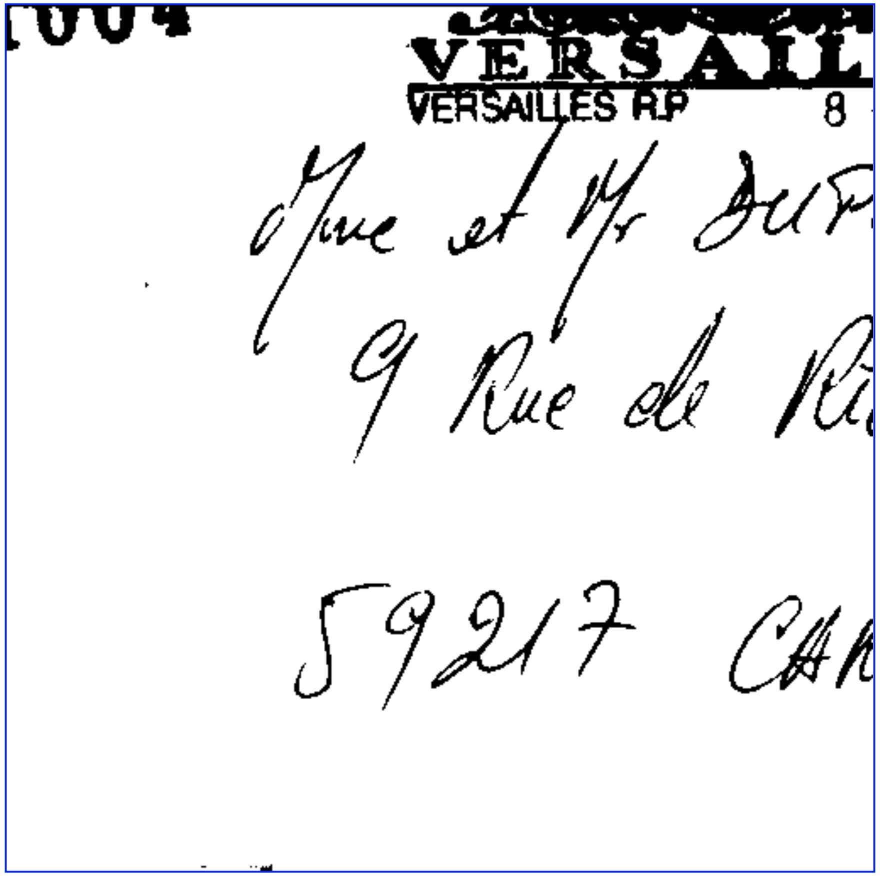


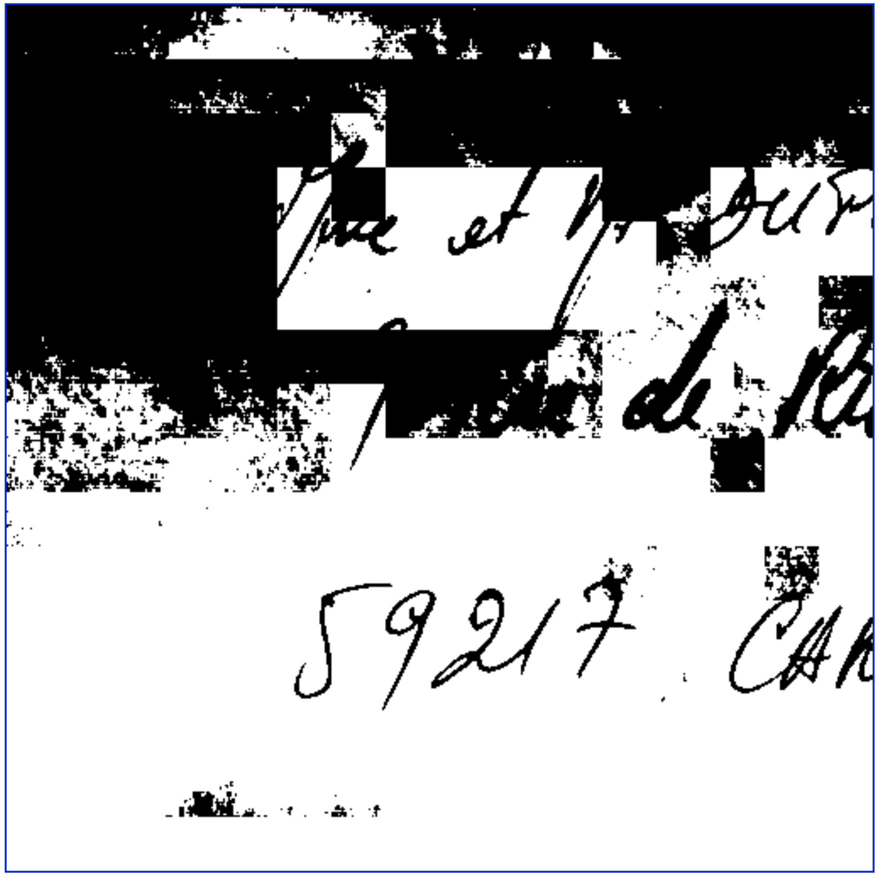
190

Binarization: Local Method

- **Split Image in Squares**
  - Square Size
  - Threshold for each Square
- **Using Threshold**
  - By Square
  - By Interpolation
  - Overlaping Squares
- **Variation Lightning Evaluation**

Binarization: Local Method



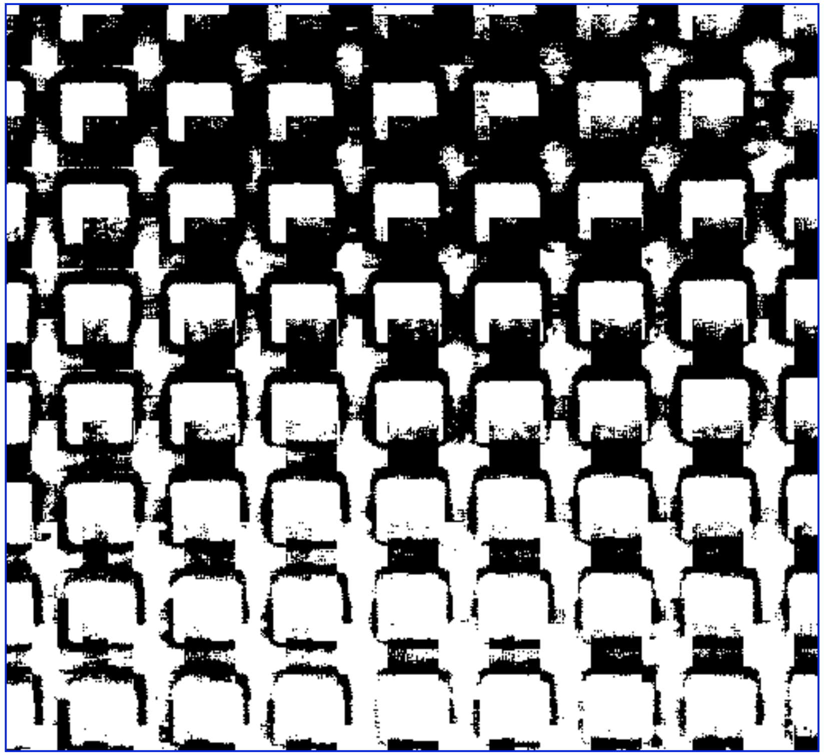


Otsu: Threshold at 188

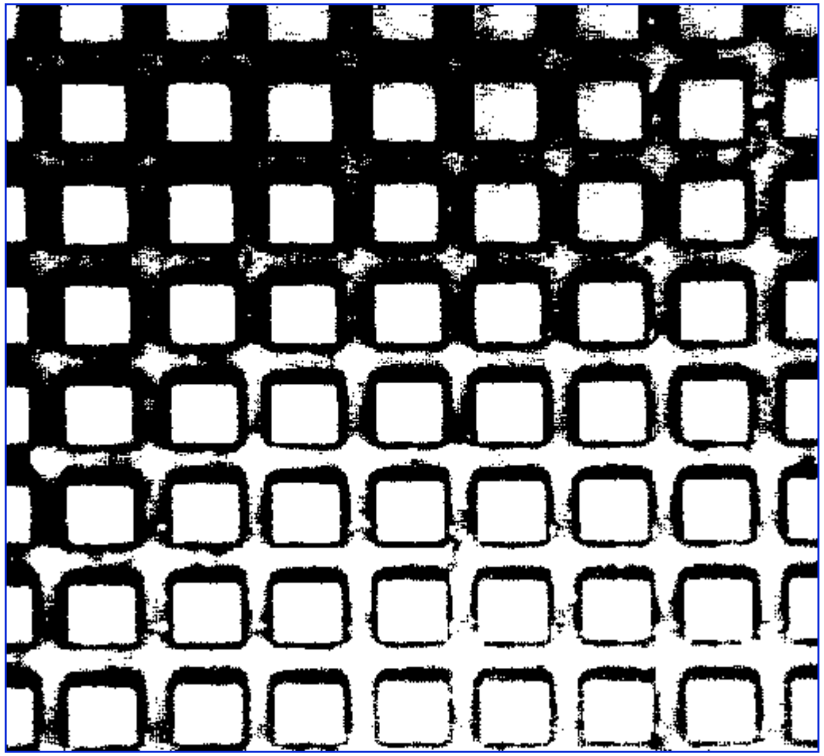
Otsu: 32x32 Squares

Binarization: Local Method, Square Size

Otsu Method in each Square



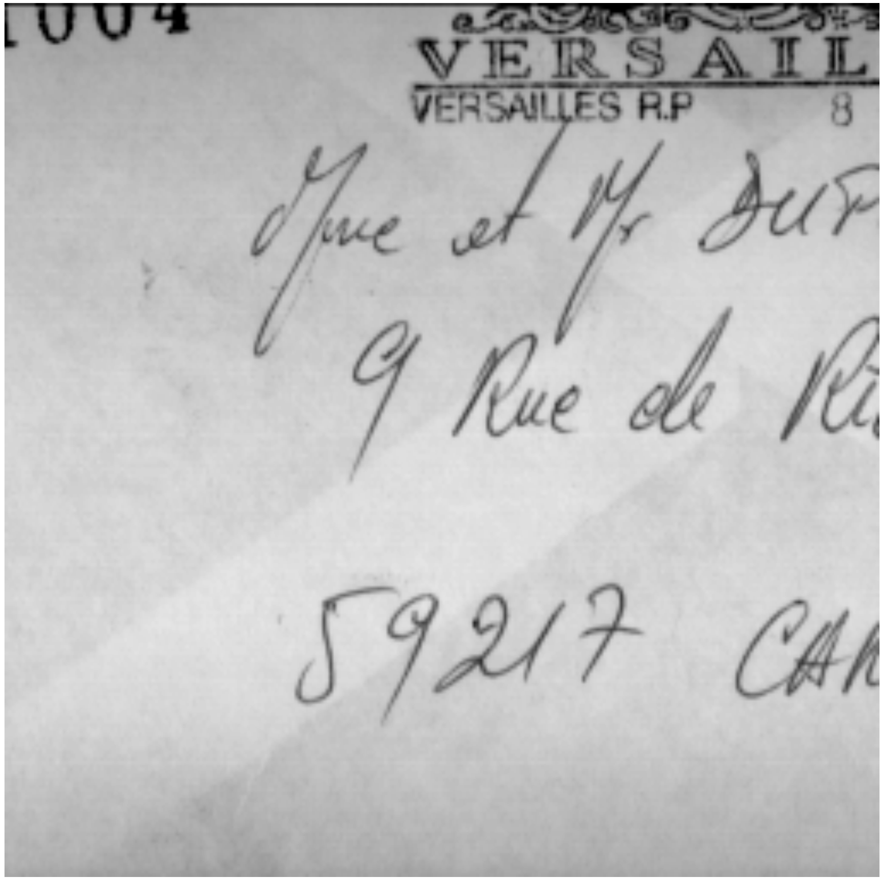
16x16 Square, too small

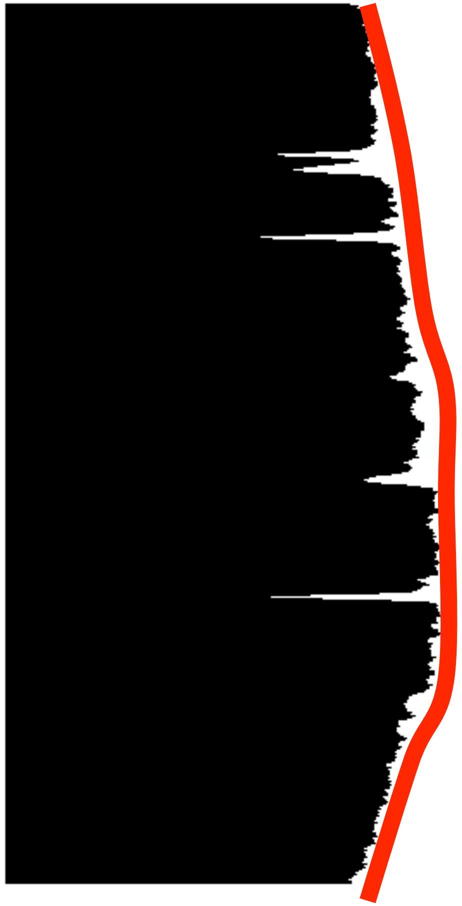


64x64 Square, good fit

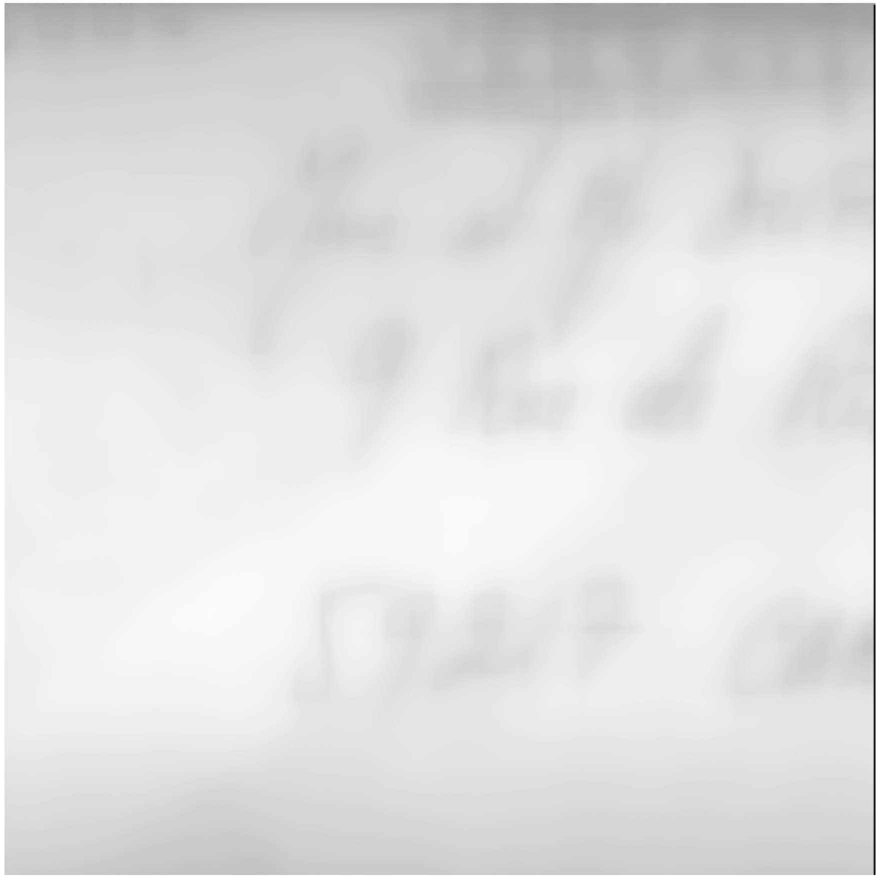
Binarization: Low-Pass Filtering

- **Variation Lightning Evaluation**

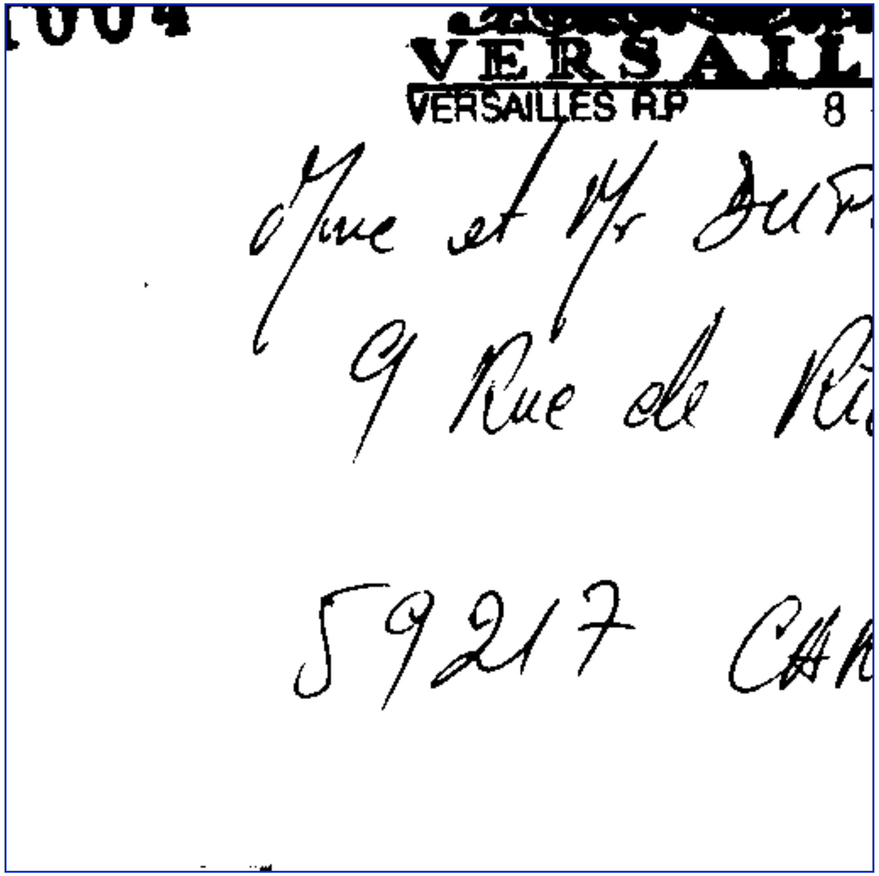


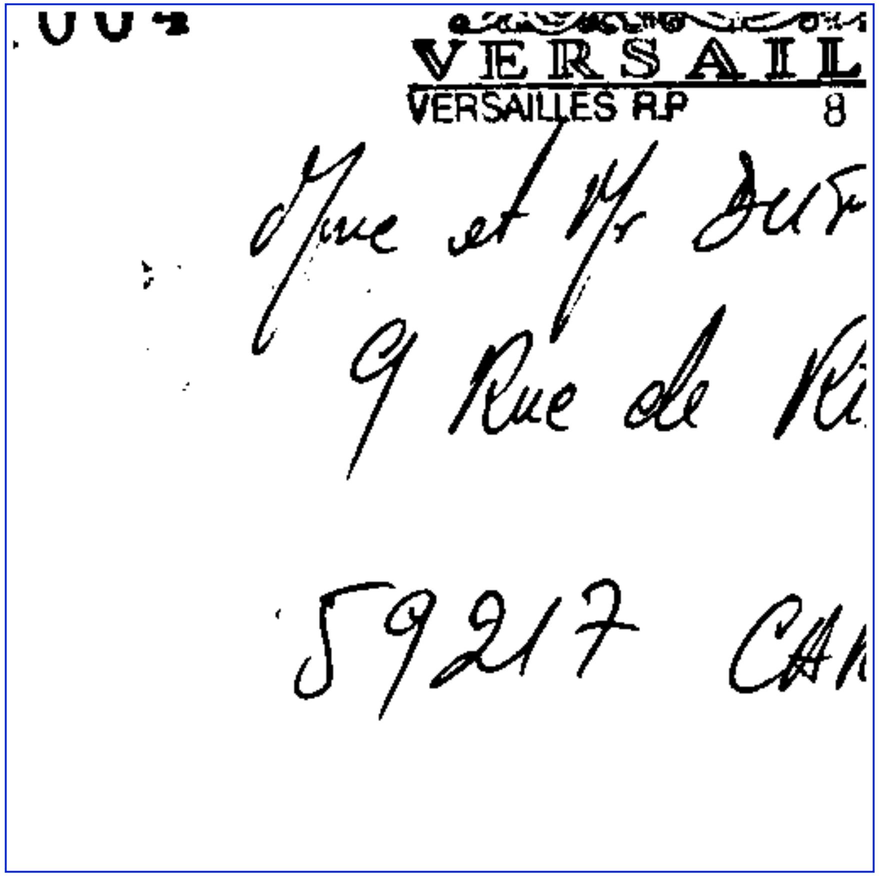


Binarization: Low-Pass Filtering



Binarization: Low-Pass Filtering

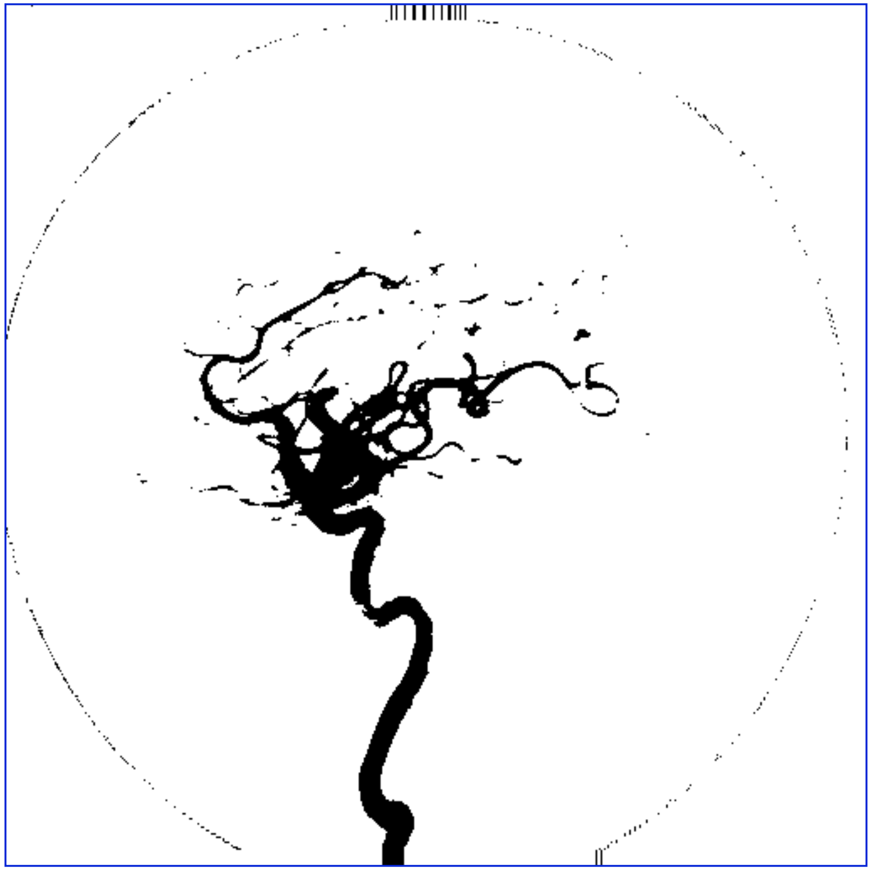




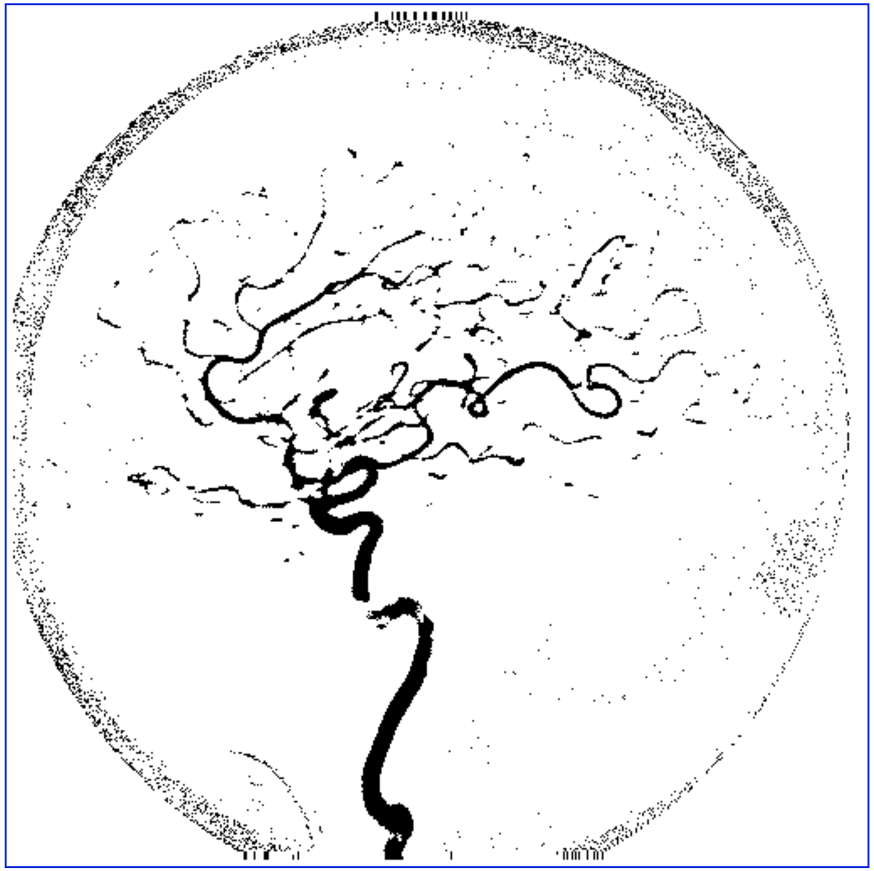
Otsu : seuil à 188

Filtrage passe-bas, contraste de 10

Binarization: Low-Pass Filtering



Otsu : seuil à 109



Filtrage passe-bas, contraste de 3