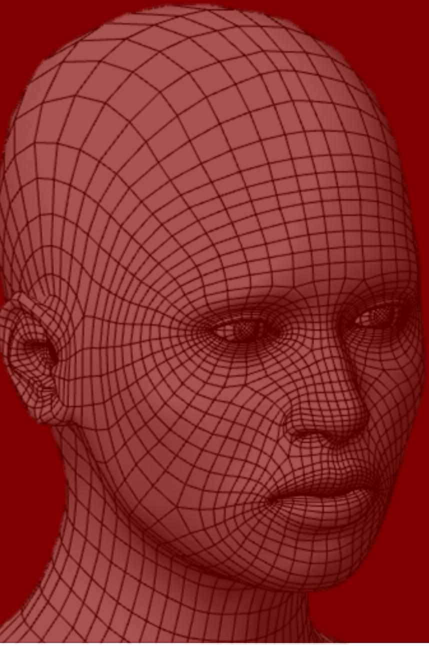


Advanced Image Processing

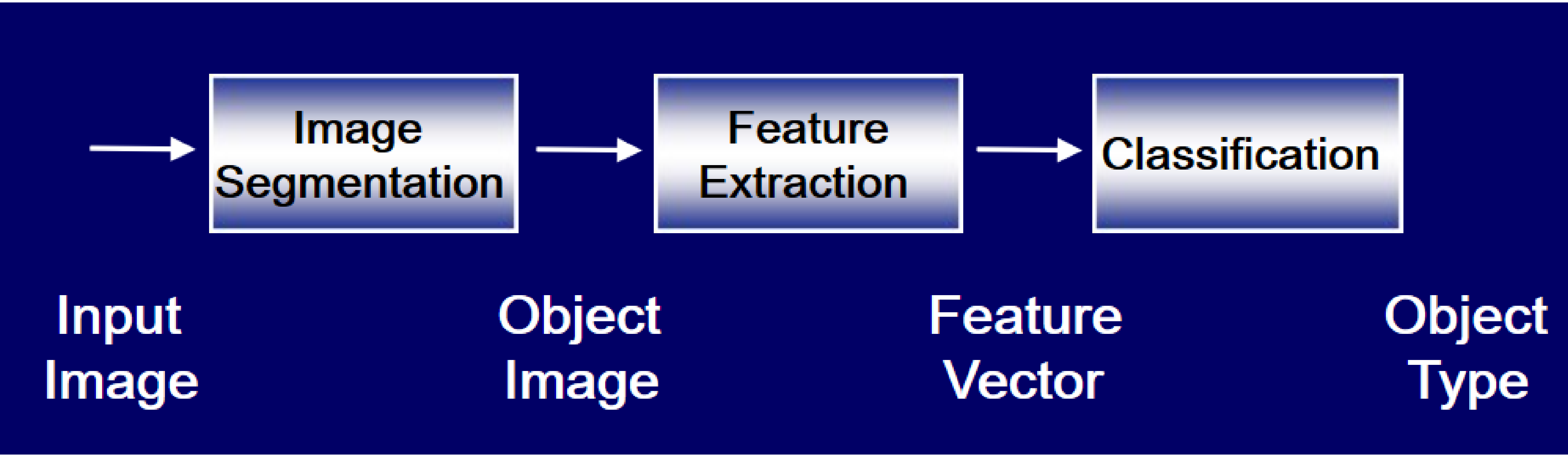
Lesson 2: Feature Detectors and Feature Descriptors: Where We are Now ?

Speaker: Alice OTHMANI, PhD
Associate professor at UPEC

Email: alice.othmani@u-pec.fr

UPEC
UNIVERSITÉ PARIS-EST CRÉTEIL VAL DE MARNE
Connaissance - Action

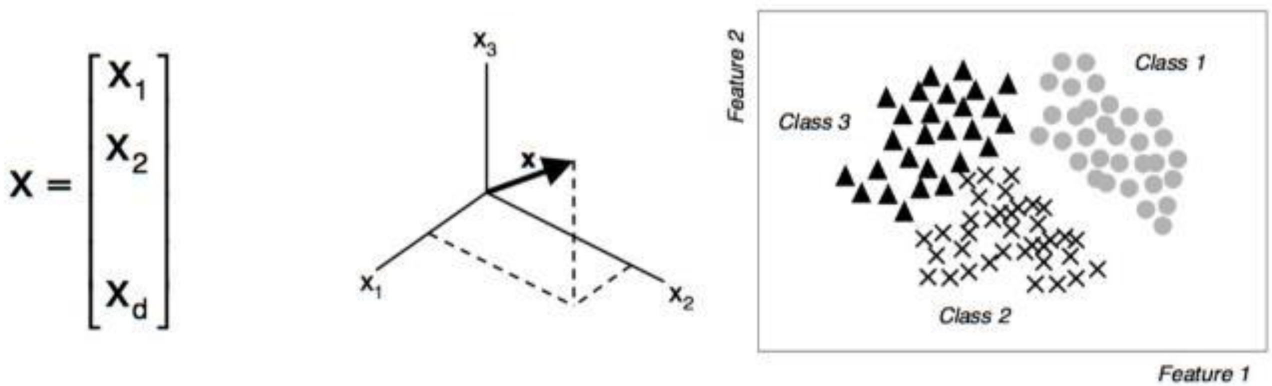
Image processing pipeline based on Shallow learning



- In computer vision and image processing, a feature is a piece of information which is **relevant** for **solving the computational task** related to a **certain application.**

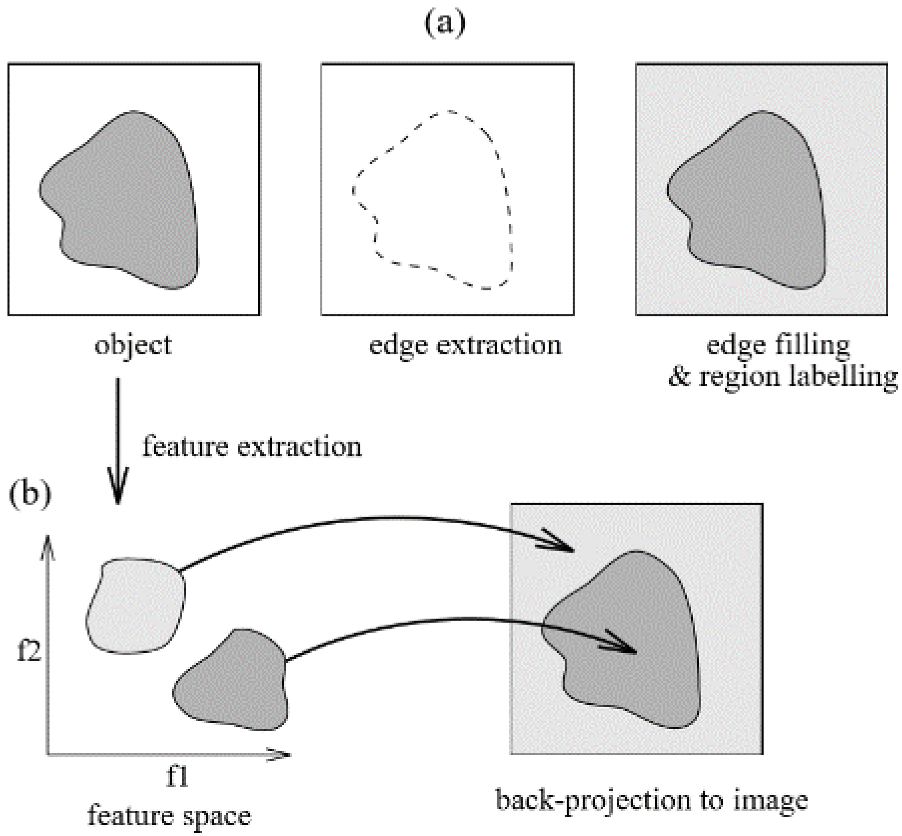
- The feature concept is very general and the choice of features in a particular computer vision system may be **highly dependent on the specific problem at hand.**

- As a result, we will get a **Feature Vector** which can be used in later stages in many ways: it can be compared to a series of Feature Vectors extracted from objects in a database to perform object recognition.



Feature vector | Feature space (3D) | Scatter plot (2D)

Image processing pipeline based on Shallow learning



An example of image segmentation: Observation space (a) and feature space (b) segmentation.

UNIVERSITÉ PARIS-EST CRÉTEIL VAL DE MARNE

Image processing pipeline based on Shallow learning

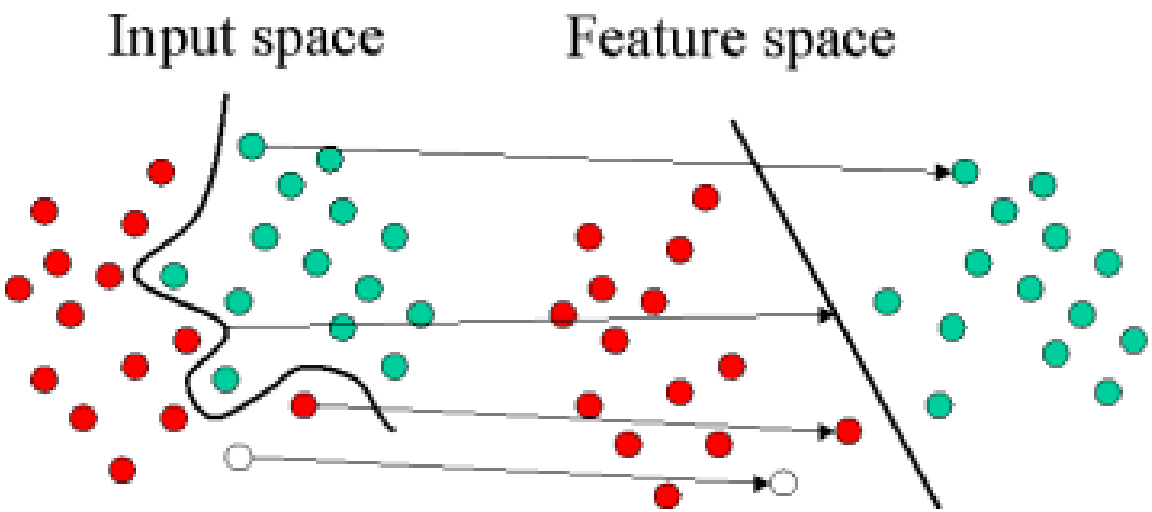
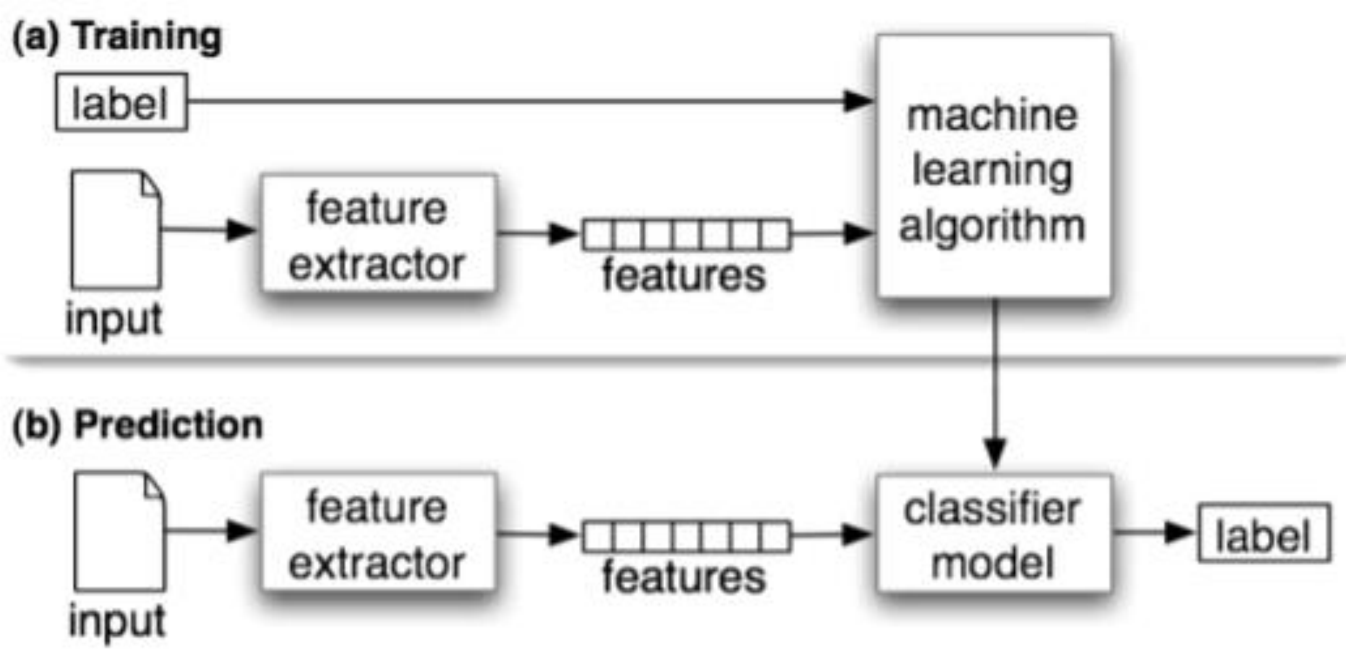
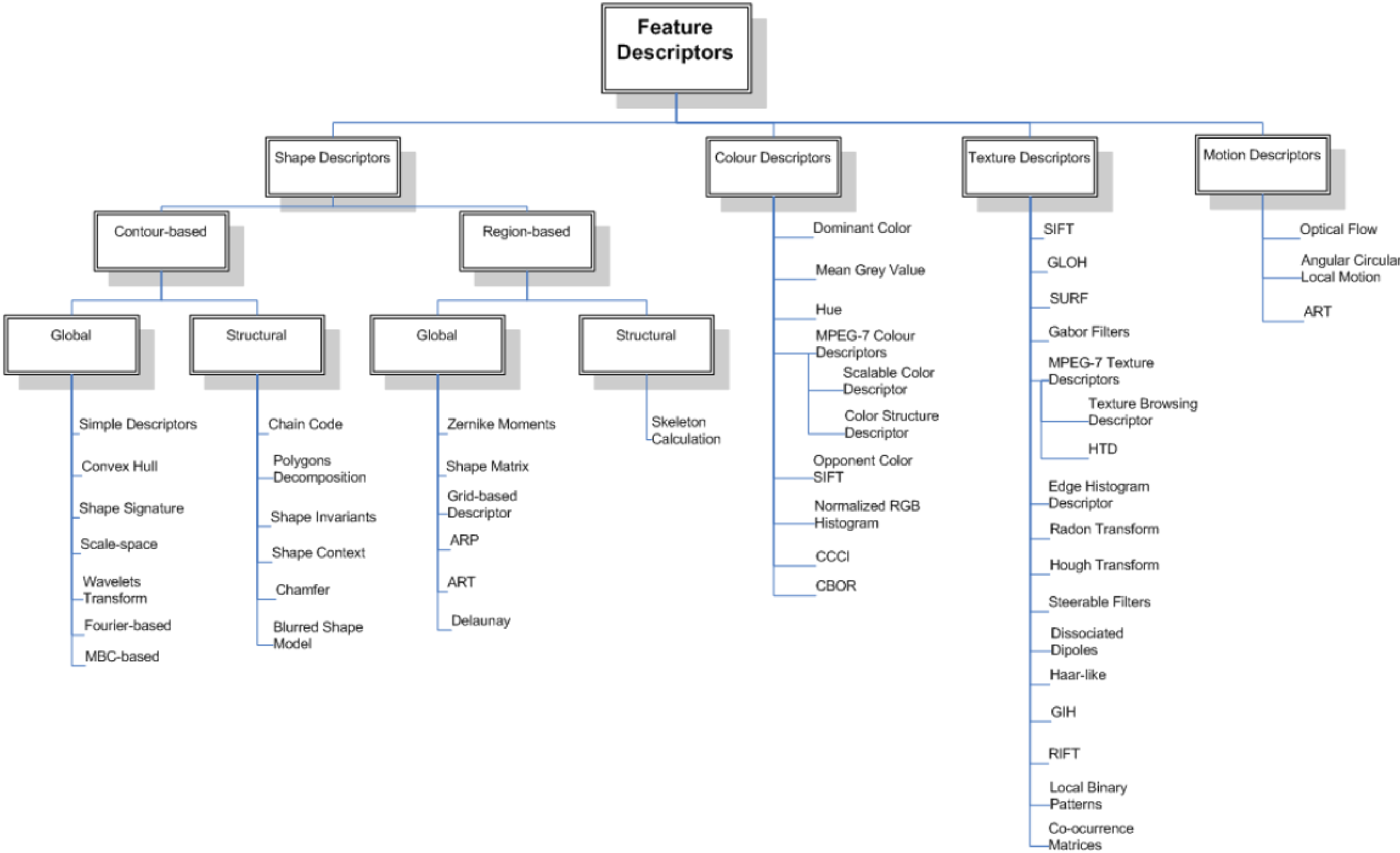


Image processing pipeline based on Shallow learning



- **Shape descriptors**
  - Contour-based
  - Region-based
- **Colour Descriptors**
- **Texture Descriptors**
- **Motion Descriptors**

Feature Descriptors



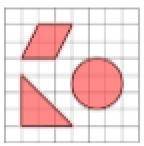
- Shape is an **important visual cue** and it is one of the most used to describe image content.

- It also has its complications because we can not forget that when we project 3-D objects into a 2-D image we are losing one whole dimension of information, so the 2-D representation of an object gives only a partial representation of it.

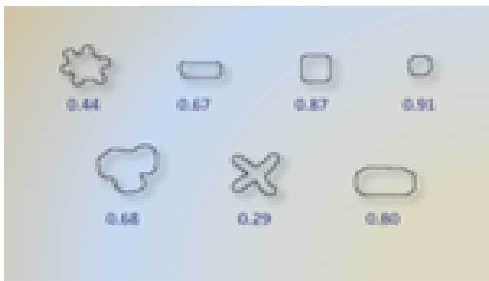
- Even shape is a feature that is **highly affected by noise**, **defects**, **occlusion** or **deformation**, *which can make harder the process of object recognition*.

1) Global:

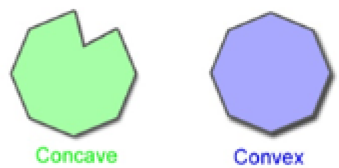
- **Simple Descriptors:** area, perimeter, eccentricity, axis orientation, radius of the principal axis, ..



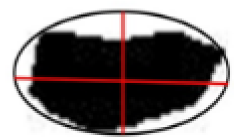
Area and Perimeter



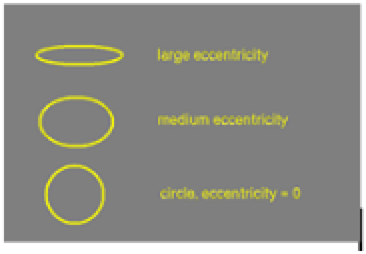
Circularity



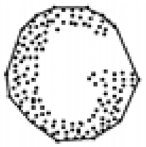
Concavity and convexity



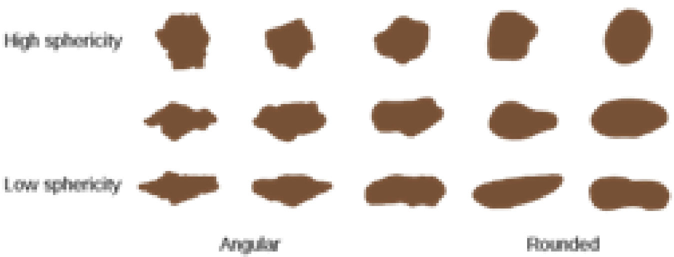
Major and Minor axis length



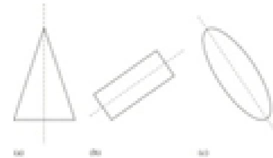
Eccentricity



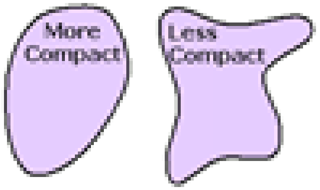
Convexity



Sphericity



Orientation

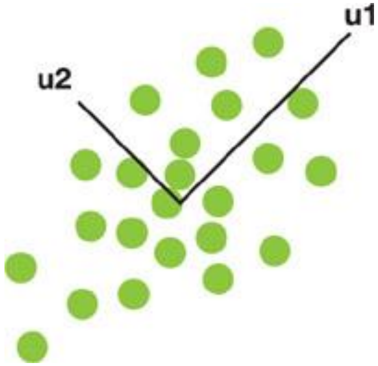


Compactness

1) Global:

- **Principal component analysis features:**

  - The percentage of the total variance explained by each principal component.
  - The longest and shortest diameter (length of the major and the minor axis).
  - The orientation or the direction of the first and the second principal component ….



1) Global:

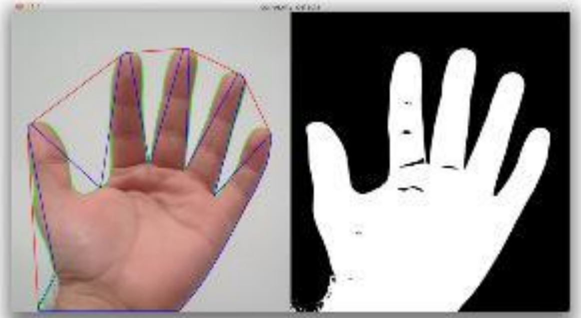
- **Simple Descriptors:** area, perimeter, eccentricity, axis orientation, radius of the principal axis, ..

→ Do not give enough information to describe two similar objects (and point out the differences between them), they can be used to **make a first decision**.

→ They can be used as a first and easy approach or combined with some other descriptors, but they are rarely used as the only descriptors.

Convex Hull:

- One region is considered as **convex** only if by taking any two points of it the segment that binds them is inside the region.
- Convex Hull is defined as the **minimal convex region**.
- Before dividing the contour in segments, it is smoothed to avoid some non-desired effects such as hysterical responses to noise. At last the whole shape will be represented as a chain of concavities.
- **Descriptor:** the convexity which can be defined as the ratio of the perimeters of the Convex Hull of the contour and the original contour.



Shape signature:

- This method represents the shape of an object by means of an **uni-dimensional function** which is extracted from the points belonging to the contour of the shape.

- There are several possible Shape Signatures such as the centroidal profile, shape radii, complex coordinates, distance to the centroid, tangent or accumulative angle, curvature, arc length, etc.

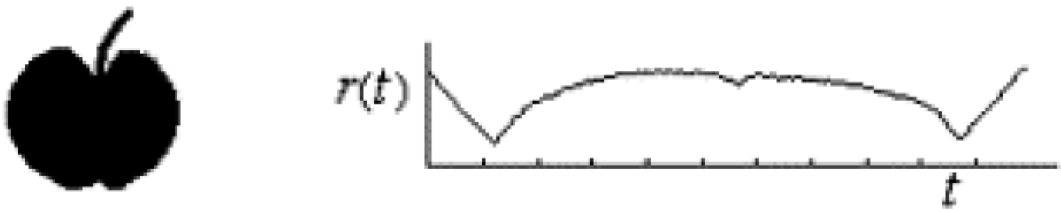


Figure: An apple and its centroidal distance Shape Signature.

Shape signature:

- Shape Signatures are usually **normalized** for translation and scale invariance.
- In order to compensate for orientation changes, shift matching is needed to find the best matching between two shapes.

- Disadvantages:
  - high matching cost,
  - sensitive to noise,
  - slight changes in the boundary can cause large errors in matching

→ Prior and Further processing is necessary to increase its robustness and reduce the matching load.

Frequency Analysis of Texture

Textures are quasi-periodic signals that have energy localized in the Frequency domain.

We focus here on **spatio-frequency** methods.

These methods make it possible to characterize the texture at different scales.

Wavelets

Give a good approximation of the visual system of mammals. [Mallat, 1989]

The input image is decomposed into layers in the retina.

UPEC Connaissance - Action
UNIVERSITÉ PARIS-EST CRÉTEIL VAL DE MARNE

Birth of wavelets

Morlet in 1983 uses the Sliding Window Fourier Transform (SWFT) for the analysis of seismic signals.

$$\widehat{s}(\omega) = \int_{-\infty}^{+\infty} s(t) e^{-j\omega t} dt$$

$$T_{fg} s(t, \omega) = \langle \psi_{\omega,t}, s \rangle = \int_{-\infty}^{+\infty} s(u) g^*(u-t) e^{-j\omega u} du$$

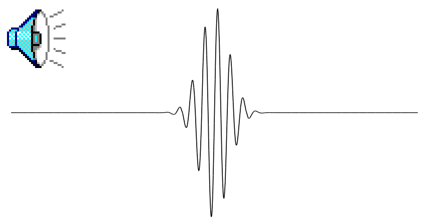
Insufficient method:

- (redundant: no base with good time-frequency localization))
- Using a window whose length is expanded or contracted

→ **Birth of the idea of wavelets**

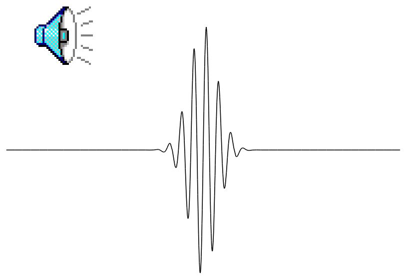
Wavelets Transform:

$\psi(t)$

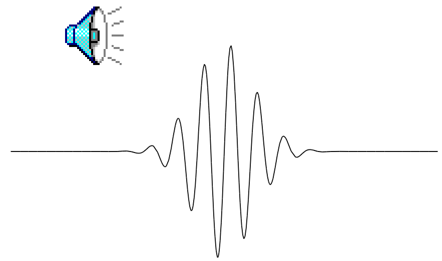


Famille construite par dilatation

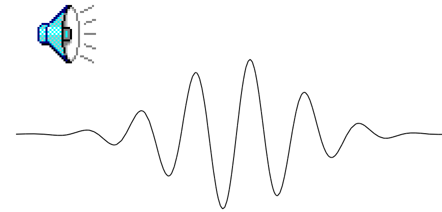
$\psi(t)$



$\psi(t/2)$

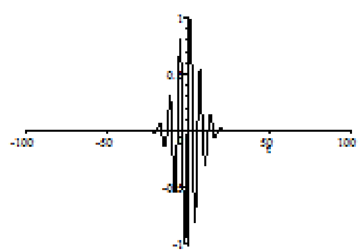


$\psi(t/4)$

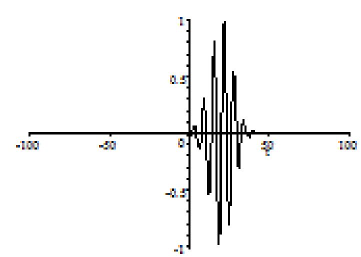


et translation : $\psi(\frac{t-b}{a}$

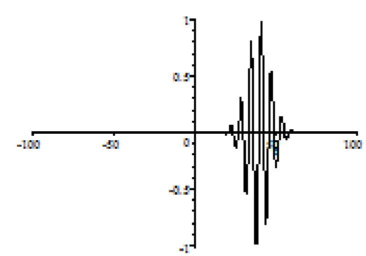
$\psi(t)$



$\psi(t-20)$



$\psi(t-40)$



Wavelets (Ondelettes, fr)

- Dilatation (a)
- Translation (b)
- Norme invariante dans $L^2(R)$
- Expression dans l'espace de départ et dans Fourier:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi(\frac{t-b}{a}) \qquad \widehat{\psi_{a,b}}(\omega) = \sqrt{a}e^{j\omega b}\hat{\psi}(a\omega)$$

- Donc le spectre est contracté d'un facteur a si la « largeur » de la fonction est dilatée du même facteur

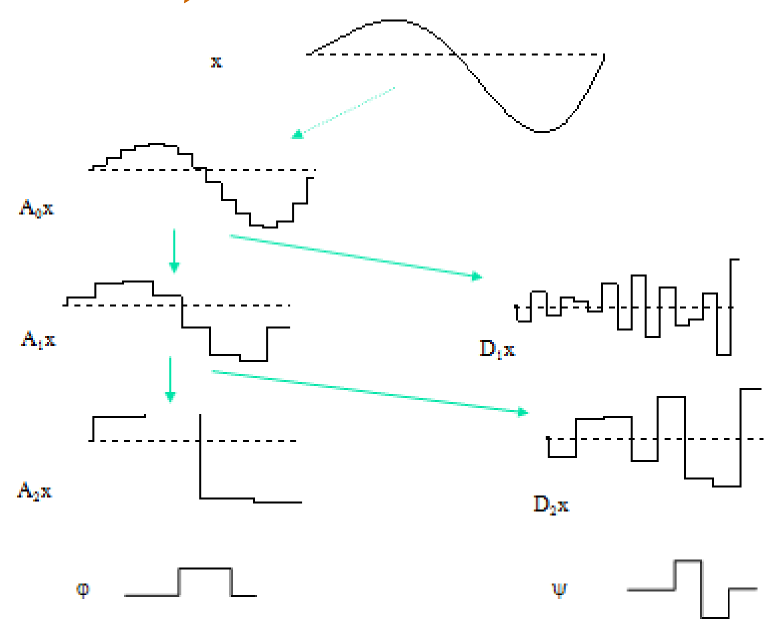
Wavelets Transform

Recursive Algorithm: **Multiresolution Analysis**

$$A_{i-1}x(t) = A_i x(t) + D_i x(t)$$

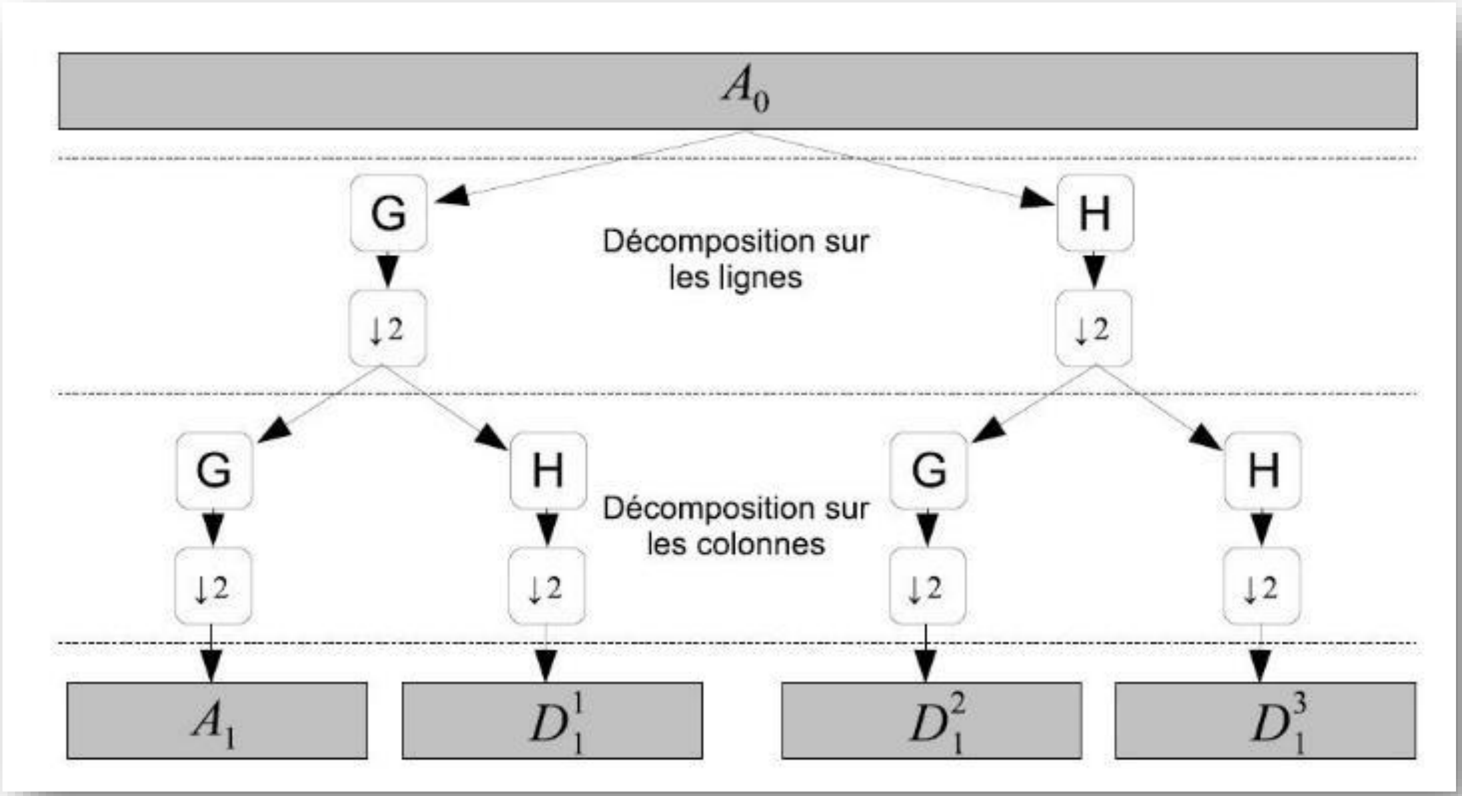
Approximation + ***Details***

(Wavelets coefficients)



UPEC
UNIVERSITÉ PARIS-EST CRÉTEIL VAL DE MARNE
Connaissance - Action

Wavelet Frames



Wavelet decomposition of a 2D image

UPEC Connaissance - Action
UNIVERSITÉ PARIS-EST CRÉTEIL VAL DE MARNE

In the case of an image, the transform is applied in two stages:

- then column by column

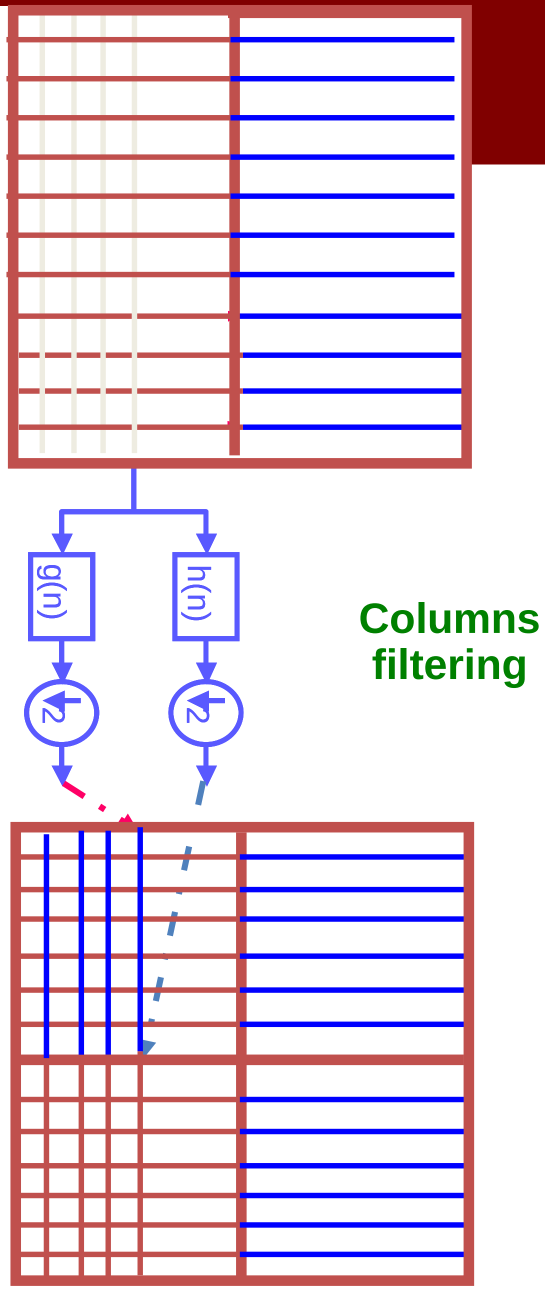
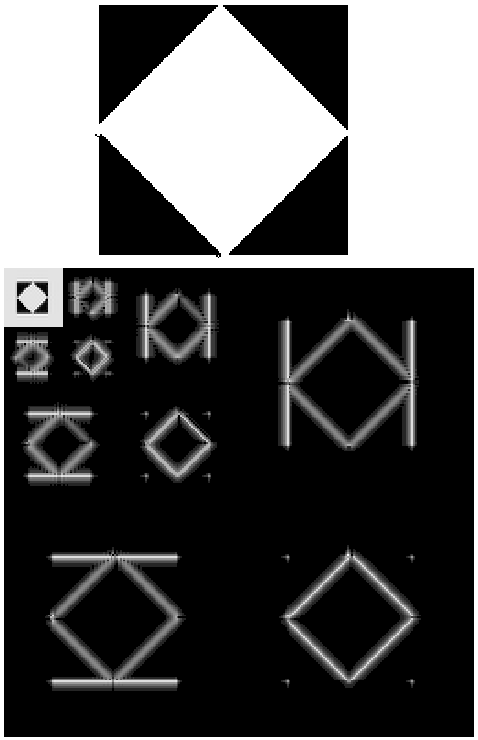
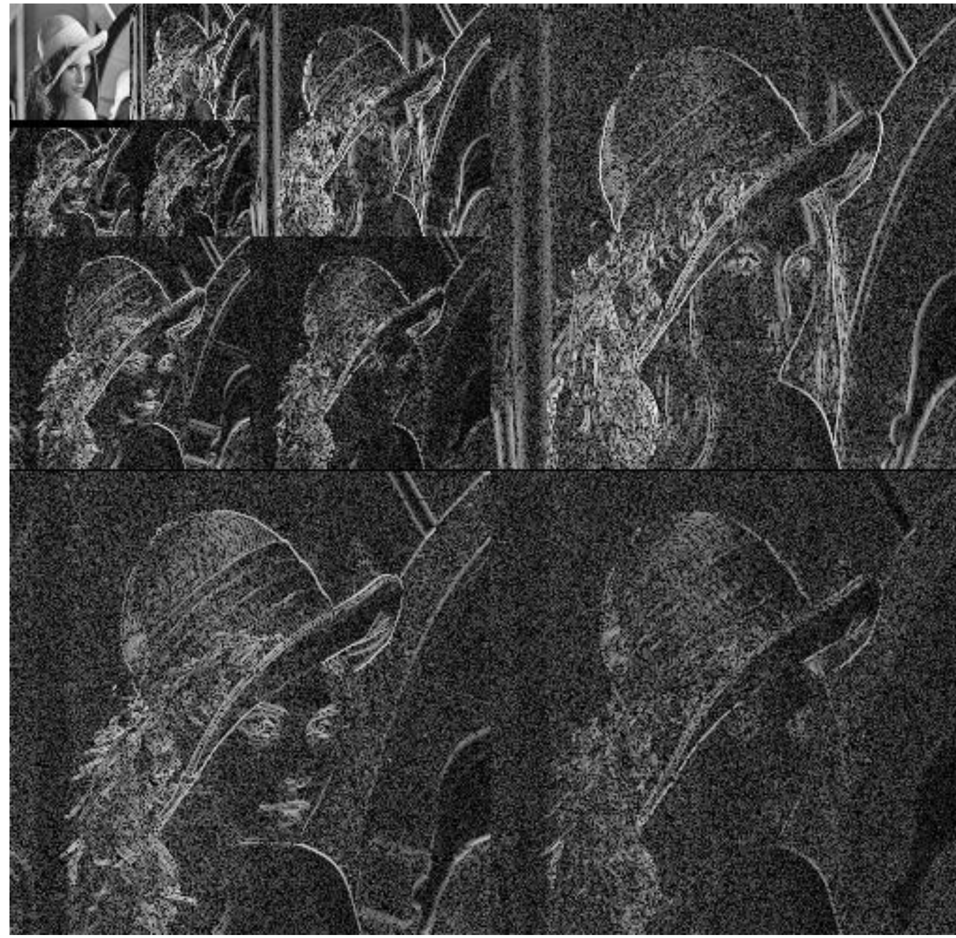


Image decomposition using Wavelet Transform





- It is very common to use the **energy** (E_k), the **mean** (μ_k) and the **standard deviation** (σ_k) of the wavelet coefficients through the multiscale subbands as texture features for classification:

$$E_k = \left(\frac{1}{m \times n}\right) \times \sum_{i=1}^{m}\sum_{j=1}^{n} |im_k(i,j)|^2$$

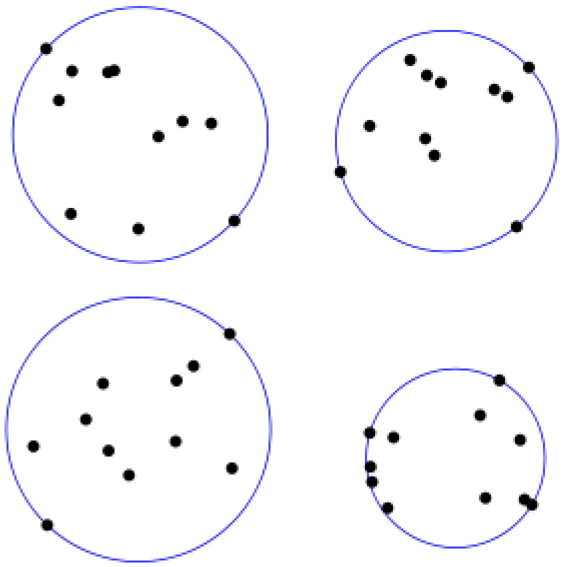
$$\mu_k = \left(\frac{1}{m \times n}\right) \times \sum_{i=1}^{m}\sum_{j=1}^{n} |im_k(i,j)|$$

$$\sigma_k = \sqrt{\left(\frac{1}{m \times n}\right) \times \sum_{i=1}^{m}\sum_{j=1}^{n} (im_k(i,j) - \mu_k)^2}$$

where im_k is a k-level subband image of size m*n.

Minimum Boundary Circle-based: features are extracted from the minimum circle that surrounds the object (that is, the circle that touches its further borders).

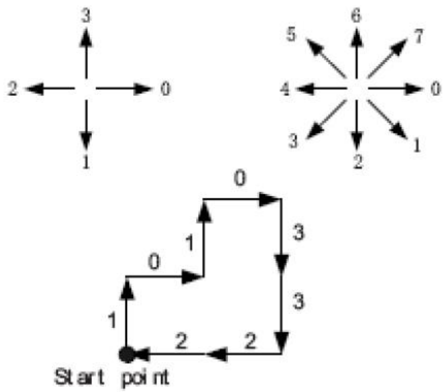
Feature Descriptors: center coordinates, radius, minimum circle crossing points, angle sequence of these crossing points, vertex angle sequence and angle sequence starting point, etc.



2) Structural:

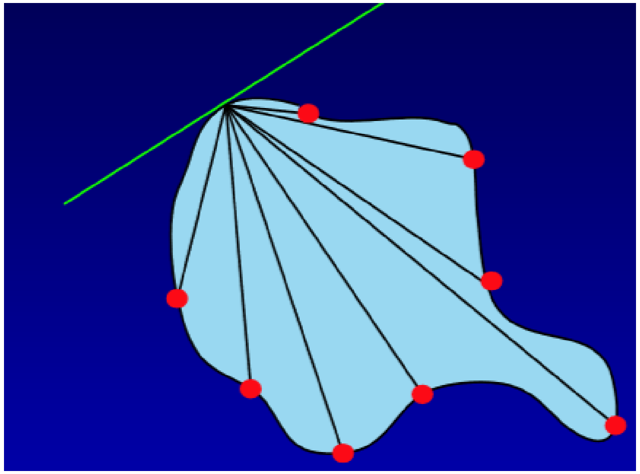
Chain code:

- Codifying lines into a determinate code → ***Descriptor***
- The nodes that surround a certain point (or central node) are enumerated counter-clockwise in ascending order from inside to outside.
- A chain will consist of an ordered link sequence.
- The inverse and the length of the chain can also be calculated.
- **Disadvantage**: depends on the starting point.
- **Advantage**: invariant to translation.



(up-left) One possible chain code (up-right) Another possible chain code (down) Contour described by using a 4 words chain code.

Shape context :



From a point P, measure distance and angle to all other points. Histogram it. That histogram is the shape context for that point.

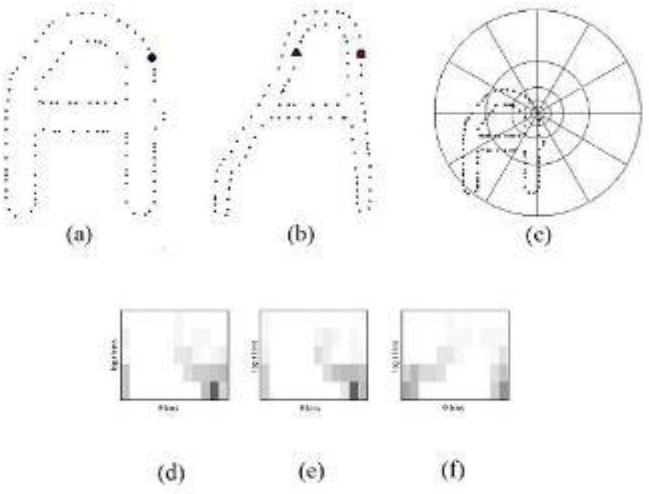
No, of course it's more complicated than that. The angle is relative to the local tangent. And the measurements are logs of distance, but that's the gist of it.

[1] Belongie, S., Malik, J., & Puzicha, J. (2001). Shape context: A new descriptor for shape matching and object recognition. In *Advances in neural information processing systems* (pp. 831-837).

[2] Belongie, S., Malik, J., & Puzicha, J. (2002). Shape matching and object recognition using shape contexts. *IEEE transactions on pattern analysis and machine intelligence*, *24*(4), 509-522. (cited 6500 times)

Shape context :

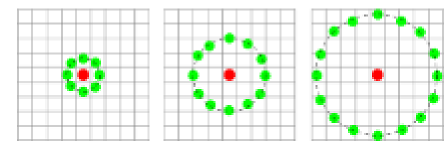
- The basic idea is to pick n points on the contours of a shape.
- For each point pi on the shape, consider the (n−1) vectors obtained by connecting pi to all other points.
- The set of all these vectors is a rich description of the shape localized at that point but is far too detailed.
- For the point pi, the coarse histogram of the relative coordinates of the remaining (n−1) points,


(a) (b) (c)
(d) (e) (f)

$$h_i(k) = \#\{q \neq p_i : (q - p_i) \in \mathrm{bin}(k)\}$$

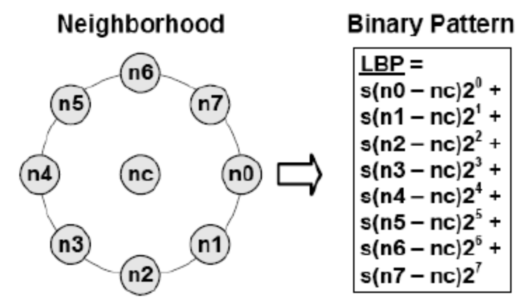
is defined to be the shape context of pi.
The bins are normally taken to be uniform in log-polar space.

LBP – Local Binary Patterns



- Local binary patterns (LBP) is a type of visual descriptor used for classification in computer vision:
  - to recognize textures,
  - for object detection in digital image.

- The operator **describes each pixel by the relative grey levels of its neighboring pixels.**



- The features have proven to be **robust against illumination changes**, they are very fast to compute, and do not require many parameters to be set.

T. Ojala, M. Pietikäinen, and D. Harwood (1994), "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions", Proceedings of the 12th IAPR International Conference on Pattern Recognition (ICPR 1994), vol. 1, pp. 582 - 585.

LBP – Local Binary Patterns

For each PIXEL of an image, a BINARY CODE is produced
→ to make a new matrix with the new value (binary to decimal value).

$$LBP_{p,r}(N_c) = \sum_{p=0}^{P-1} g\left(N_p - N_c\right)2^p$$

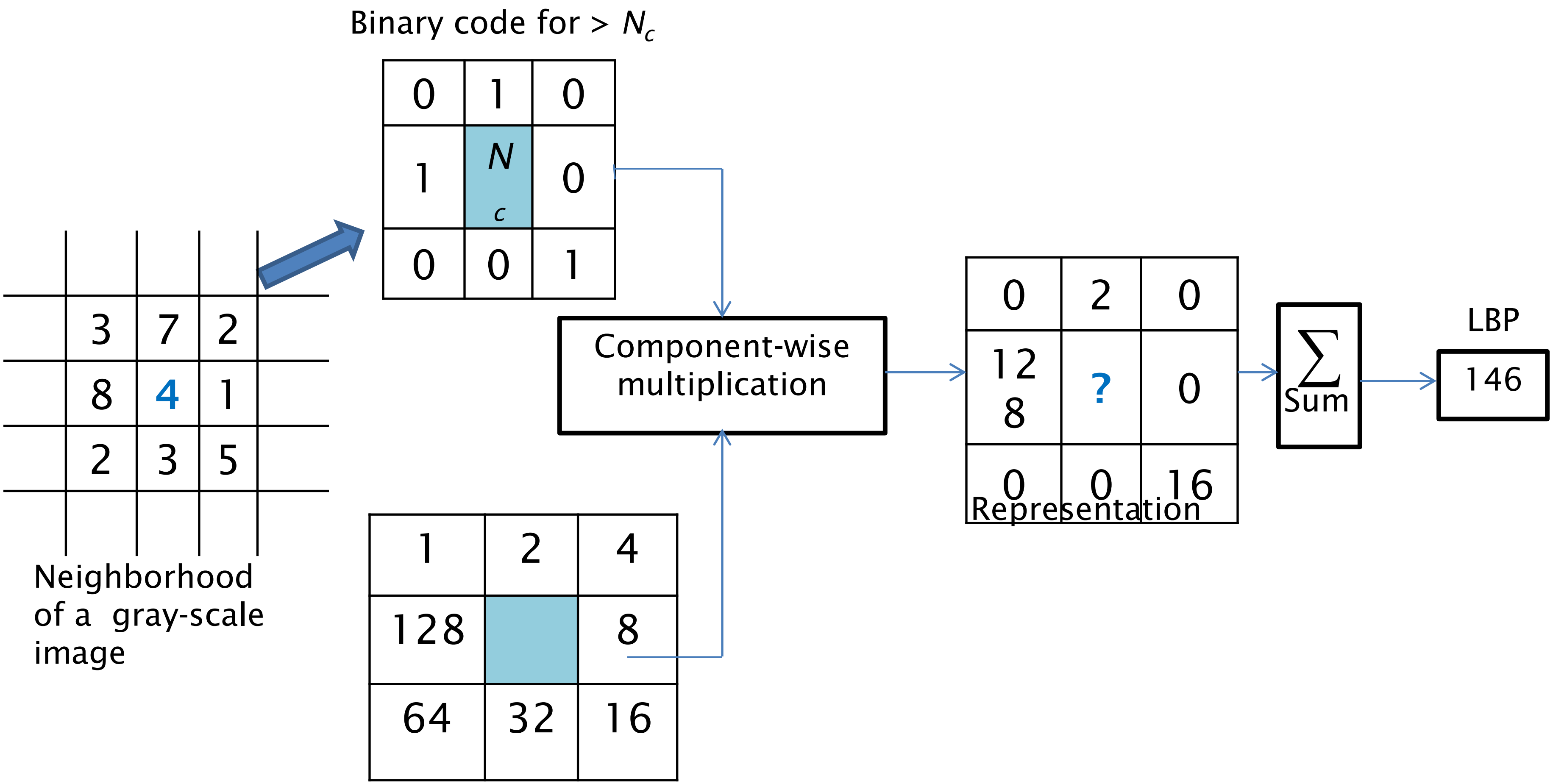
where,
neighborhood pixels $\left(N_p\right)$ in each block →
is thresholded by its center pixel value (N_c)
p→ sampling points (e.g., p = 0, 1, …, 7 for a 3x3 cell, where P = 8)
r → radius (for 3x3 cell, it is 1).

Binary threshold function $g(x)$ is,

$$g(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

UPEC Connaissance - Action
UNIVERSITÉ PARIS-EST CRÉTEIL VAL DE MARNE

Computation of Local Binary Pattern



LBP – Local Binary Patterns

	20	10	5	
	7	10	12	
	9	1	11	

1 Pick a pixel of the image and compute its LBP

Principe

UPEC Connaissance - Action | UNIVERSITÉ PARIS-EST CRÉTEIL VAL DE MARNE

LBP – Local Binary Patterns

	20	10	5	
	7	10	12	
	9	1	11	

2

Determine a mask on the neighboring 3x3

Principe

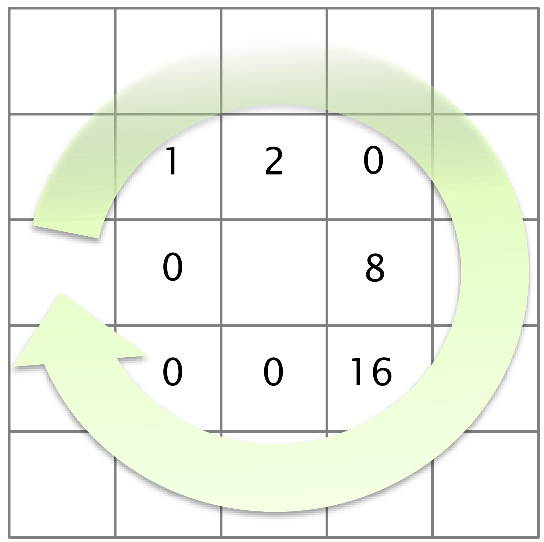
LBP – Local Binary Patterns

	1	1	0	
	0		1	
	0	0	1	

3 Where the center pixel's value is greater than the neighbor's value, write "0". Otherwise, write "1". This gives an 8-digit binary number (which is usually converted to decimal for convenience).

Principe

LBP – Local Binary Patterns



4 Calculate each value of neighboring pixels

LBP – Local Binary Patterns

Principe

		27		

5 Calculate the LBP of the chosen pixel

Gray Level Co-Occurrence Matrix (GLCM)

- A co-occurrence matrix or co-occurrence distribution is a matrix that is defined over an image to be the **distribution of co-occurring pixel values** (grayscale values, or colors) **at a given offset**.

- Represents the **distance and angular spatial relationship** over an image sub-region of specific size.

Robert M Haralick; K Shanmugam; Its'hak Dinstein (1973). "Textural Features for Image Classification" (PDF). IEEE Transactions on Systems, Man, and Cybernetics. SMC-3 (6): 610–621.

Gray Level Co-Occurrence Matrix (GLCM)

- The GLCM is created from a gray-scale image.
- The GLCM calculates how often a pixel with gray-level (grayscale intensity or Tone) value *i* occurs either horizontally, vertically, or diagonally to adjacent pixels with the value *j*.

GLCM directions of Analysis

- 1. Horizontal (0^0)
- 2. Vertical (90^0)
- 3. Diagonal:
  - a.) Bottom left to top right (-45^0)
  - b.) Top left to bottom right (-135^0)
- Denoted P_0, P_{45}, P_{90}, & P_{135} Respectively.
- Ex. $P_0(i, j)$

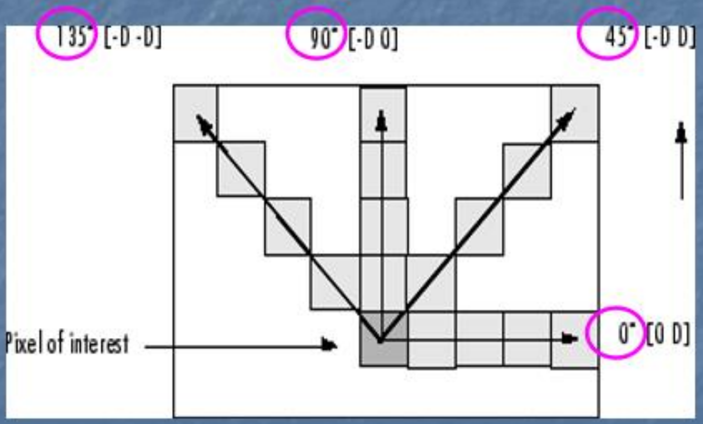
Gray Level Co-Occurrence Matrix (GLCM)

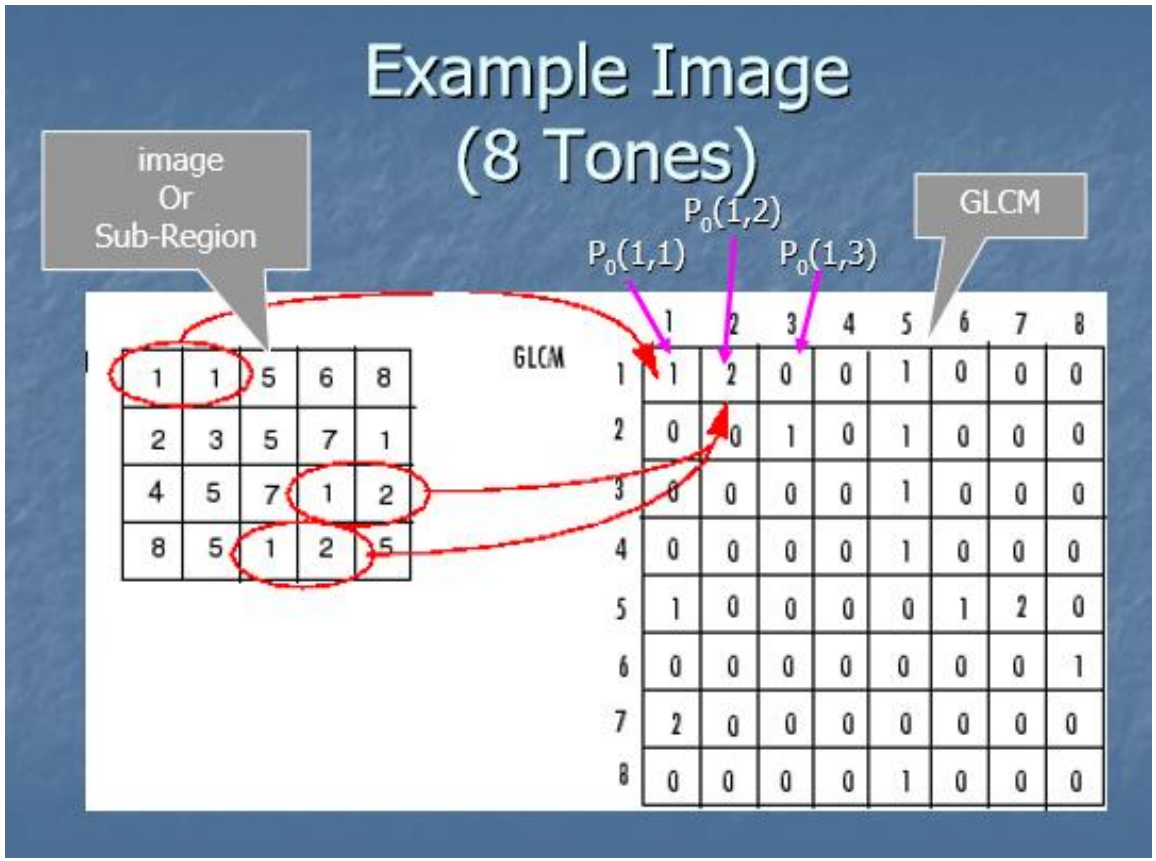
GLCM directions of Analysis

- 1. Horizontal (0^0)
- 2. Vertical (90^0)
- 3. Diagonal:
  - a.) Bottom left to top right (-45^0)
  - b.) Top left to bottom right (-135^0)
- Denoted P_0, P_{45}, P_{90}, & P_{135} Respectively.
- Ex. $P_0(i, j)$

Example of directional Analysis

P_0, P_{45}, P_{90}, & P_{135}



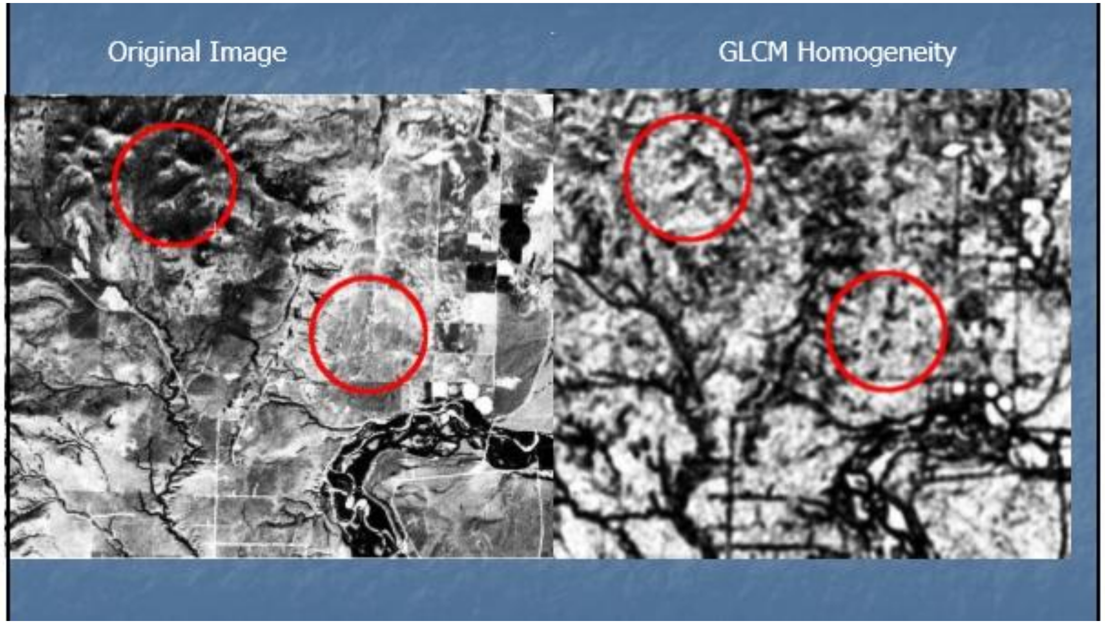

Example Image
(8 Tones)
image
Or
Sub-Region
GLCM
P0(1,1)
P0(1,2)
P0(1,3)
1 1 5 6 8
2 3 5 7 1
4 5 7 1 2
8 5 1 2 5
GLCM
1 2 3 4 5 6 7 8
1 1 2 0 0 1 0 0 0
2 0 0 1 0 1 0 0 0
3 0 0 0 0 1 0 0 0
4 0 0 0 0 1 0 0 0
5 1 0 0 0 0 1 2 0
6 0 0 0 0 0 0 0 1
7 2 0 0 0 0 0 0 0
8 0 0 0 0 1 0 0 0

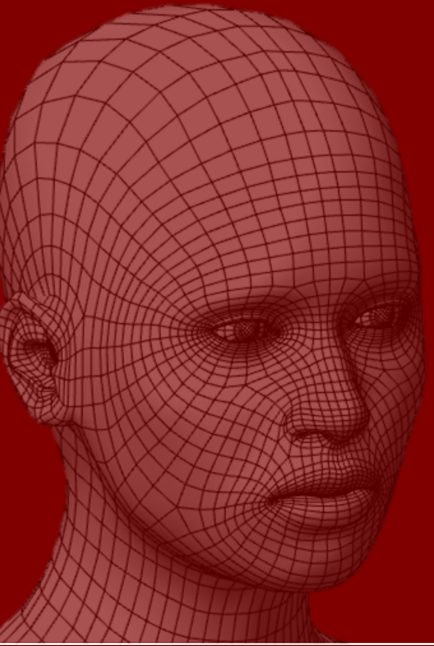
- After you create the GLCMs, you can derive several statistics from them using the different formulas.
- These statistics provide information about the texture of an image :

Statistic	Description
Contrast	Measures the local variations in the gray-level co-occurrence matrix.
Correlation	Measures the joint probability occurrence of the specified pixel pairs.
Energy	Provides the sum of squared elements in the GLCM. Also known as uniformity or the angular second moment.
Homogeneity	Measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal.

- The textures below were run using a 7*7 window.
- All used the invariant direction, which is an average of all four spatial arrangements.
- Pixel offset is 1 in all cases.

- The textures below were run using a 7*7 window.
- All used the invariant direction, which is an average of all four spatial arrangements.
- Pixel offset is 1 in all cases.





Thank you for your attention