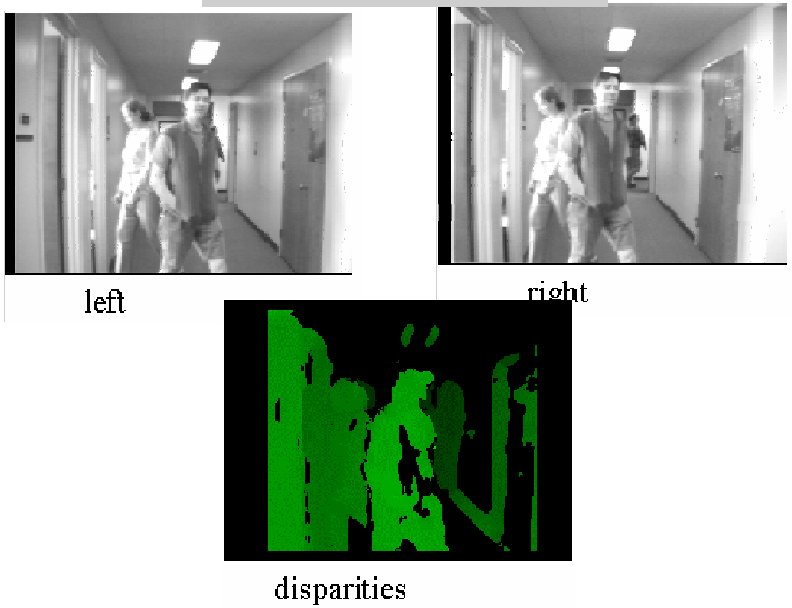
Computer Vision: introduction

- Science that develops the **theoretical** and **algorithmic base** by with **useful information is obtained and analyzed** about the world/environment, from:

An image

A set of images

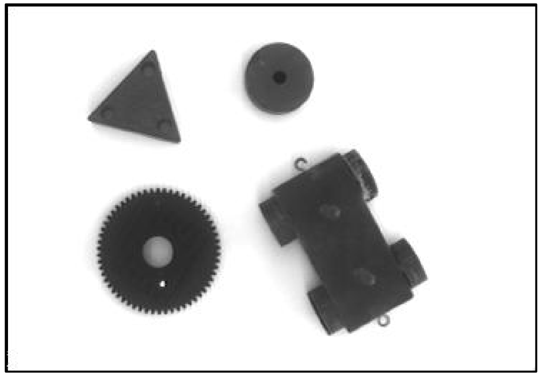


A sequence of images

- Sensors do not offer this information directly, processing is required.

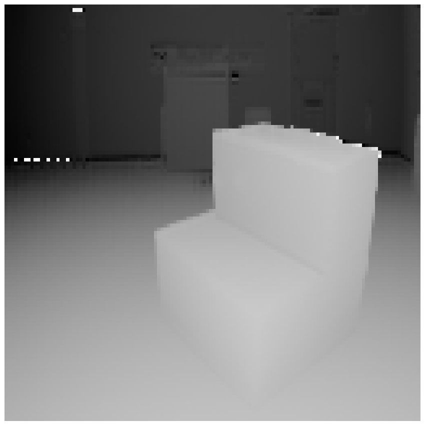
Types of images

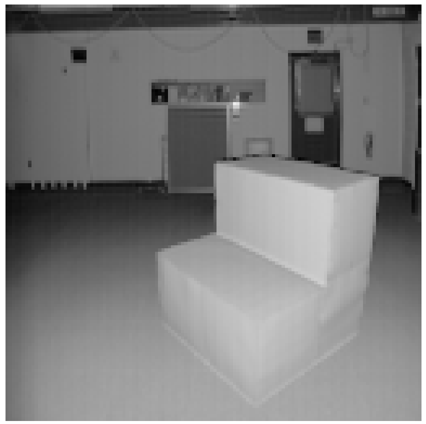
- Optical sensors: **luminosity**



- Range sensors: **distance to objects**

range

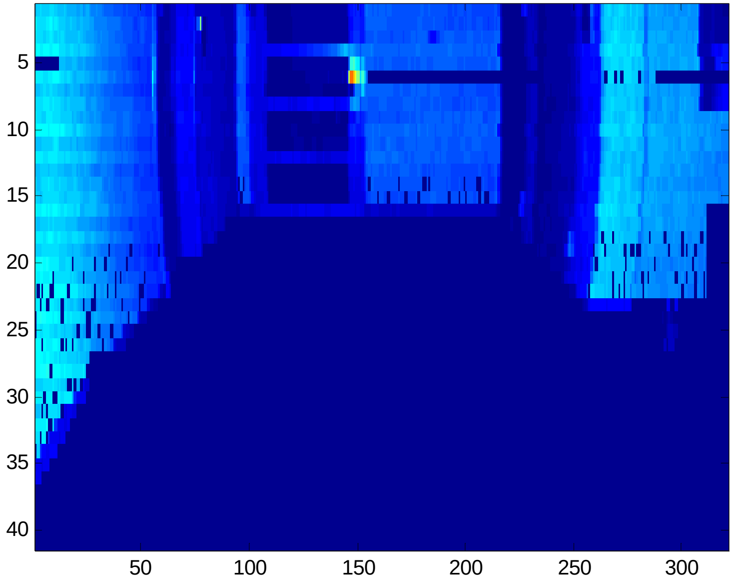


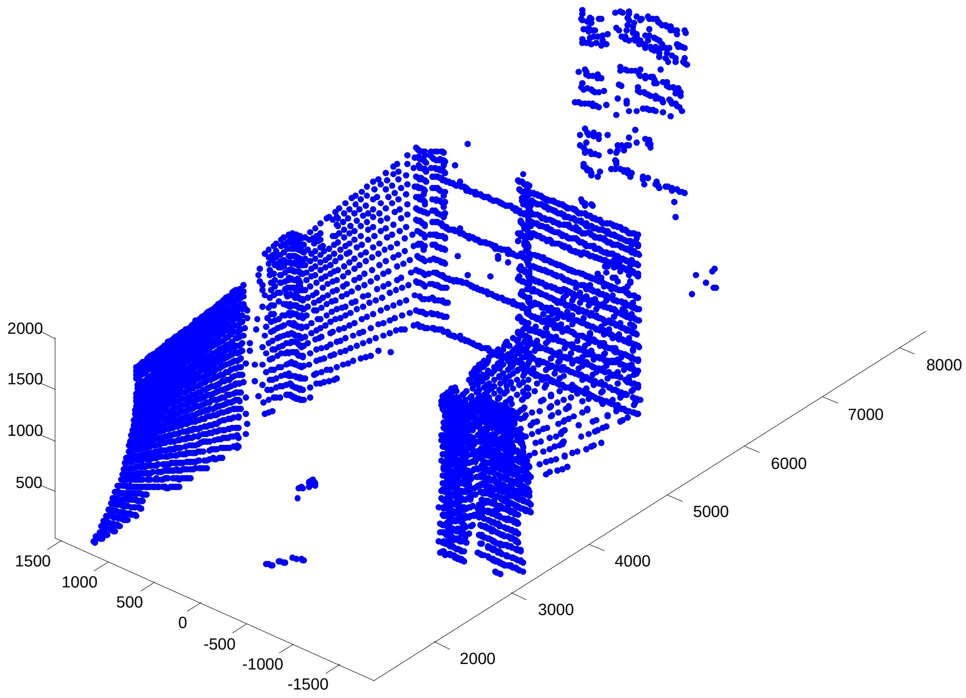


optical

Types of images

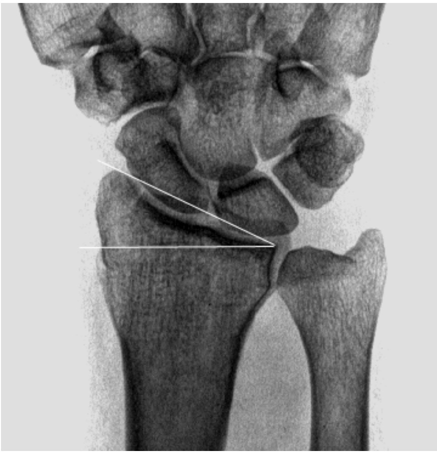
- Range sensors: **distance to objects**



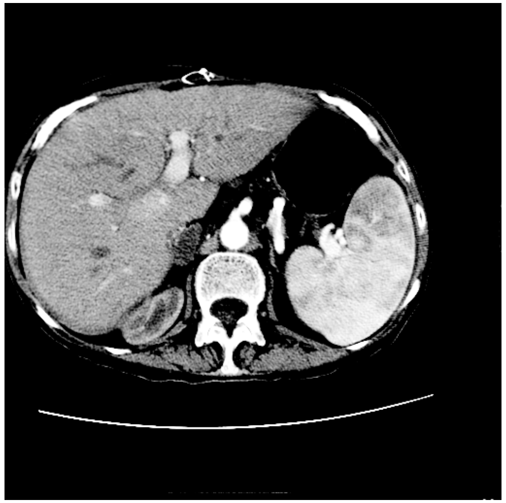


Types of images

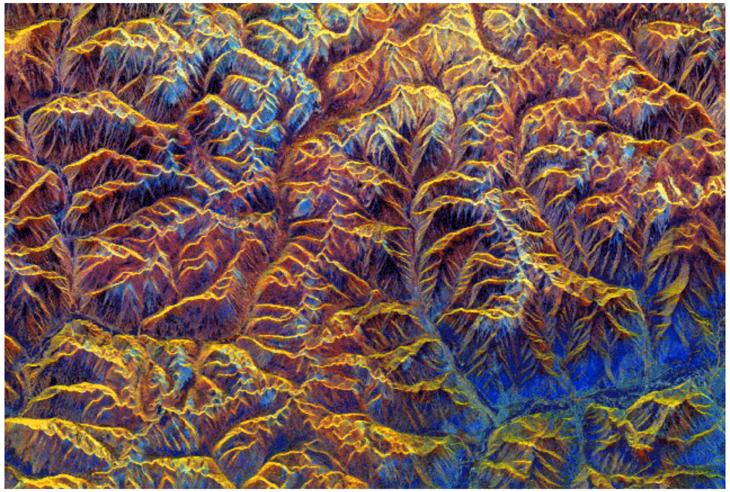
- Radiography: x rays



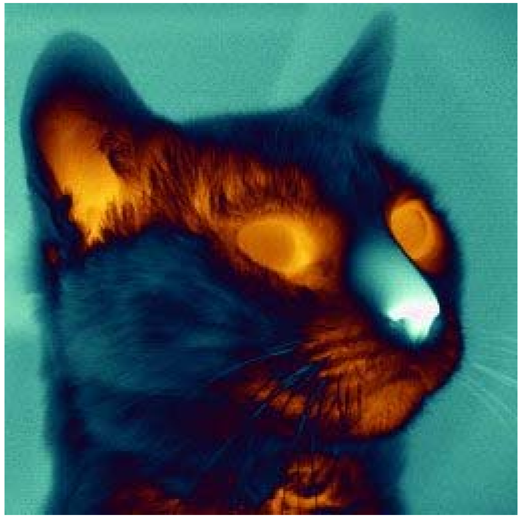
- Tomography



- Sonar: echolocation



- Infrared:



Fundamental problem

- **Observable units:** a sequence of integers

00	00	00	00	00	00	00	00	00
00	01	02	03	04	03	02	01	00
00	02	04	06	08	06	04	02	00
00	03	06	09	12	09	06	03	00
00	04	08	12	16	12	08	04	00
00	03	06	09	12	09	06	03	00
00	02	04	06	08	06	04	02	00
00	01	02	03	04	03	02	01	00
00	00	00	00	00	00	00	00	00

- **Analytical units:** high level descriptions

- What objects are there?
- Where are they?

- What properties do they have?
- How do they relate?

Fundamental problem

- Observable units **are not equal to** analytical units.

Observable units:

PIXEL (PIctures(x?) ELement)

Properties:

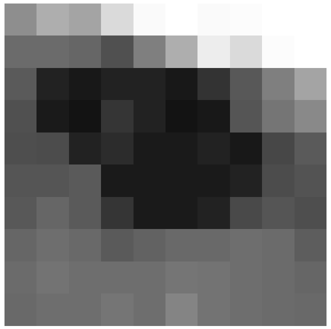
- position
- value

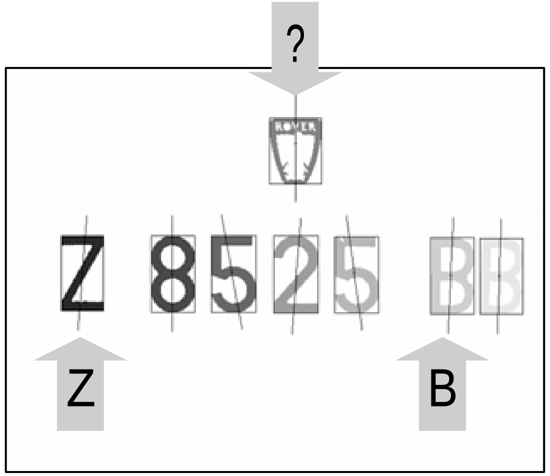
Analytical units:

Objects present:

- identity
- shape
- position
- orientation
- description

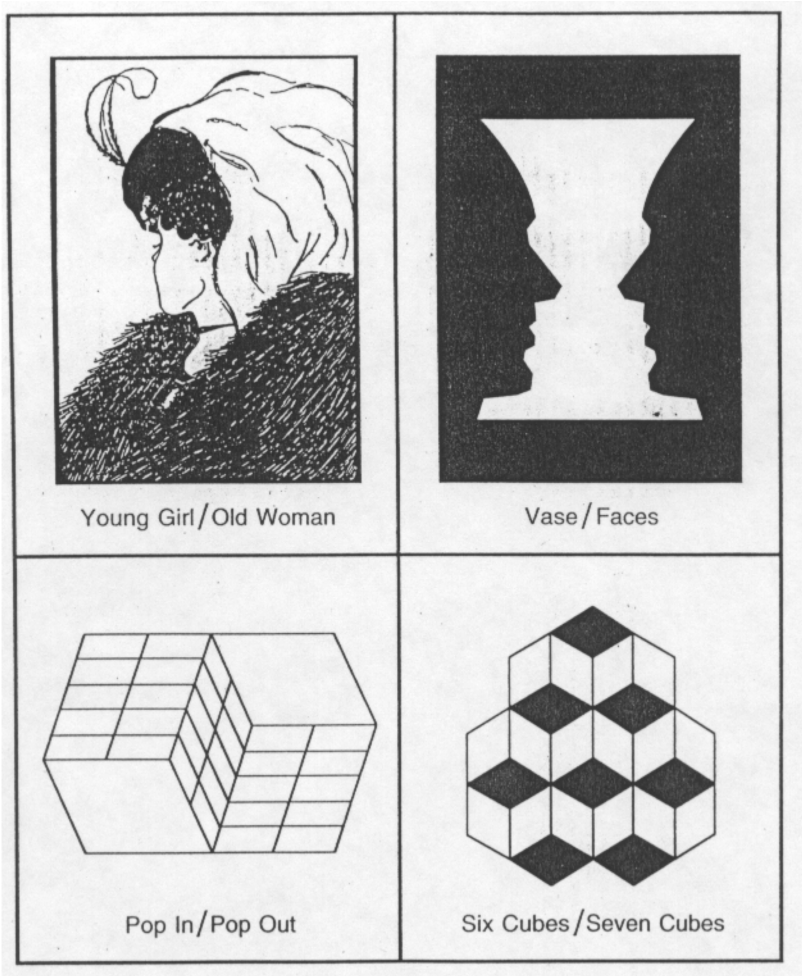
- Each pixel **DOES NOT** in isolation contain information about the object.





Additional problems

- Occlusions
- Shades
- Reflectance
- Noise
- Complex background

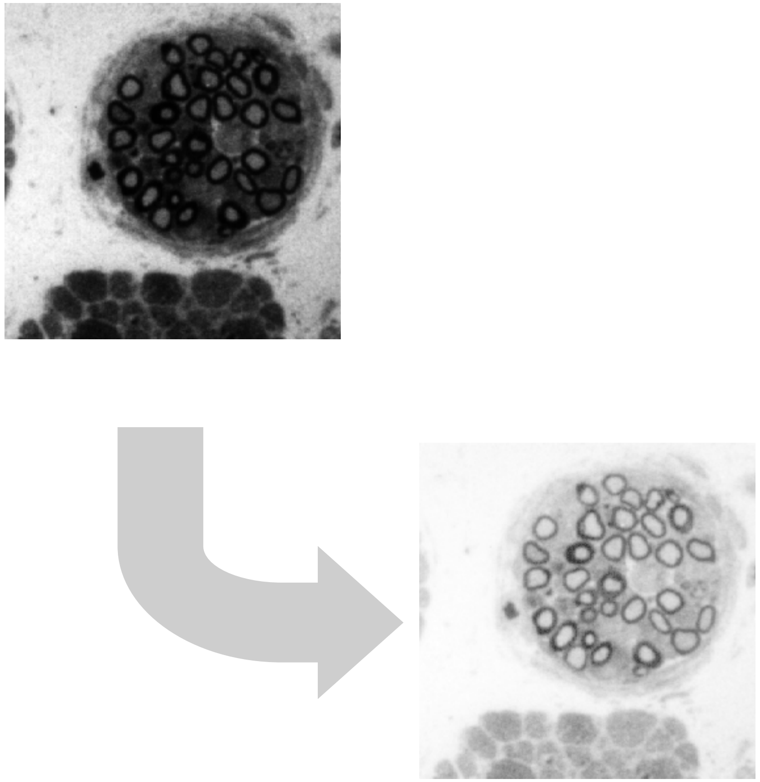


Difficult problems for all vision systems.

These are complex even for human vision

Relation to other fields

- **Digital image processing:** Techniques to transform images with no interpretation (improvement, compression, noise reduction).



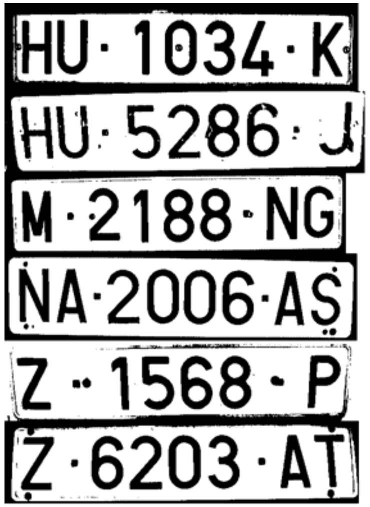
- **Computer graphics:** Generation of synthetic images from scene descriptions (it is the opposite problem!).

- **Pattern recognition:** Numerical information classification (statistical), and also symbolic (syntactical).
- **Artificial intelligence:** It studies computational aspects of intelligence, to design systems that can behave intelligently.
- **Psycophysics:** It studies the effect of physical phenomena or stimuli in the organism. Human vision has been long studied.

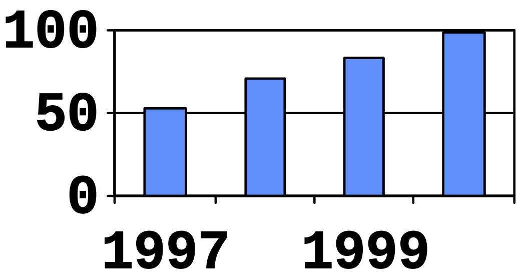
Applications

- Medical diagnosis
- Meteorology
- Robotics
- Agriculture
- Space exploration
- …

Recognition of planar shapes:

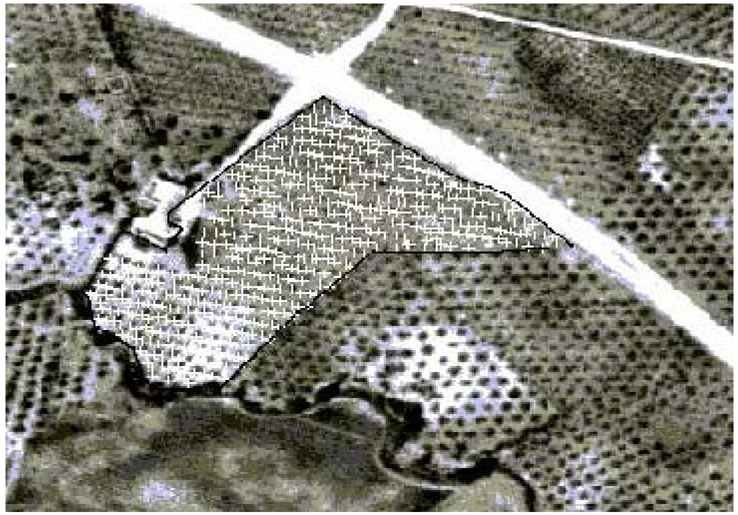


Mercado europeo vision (miles millones ptas)



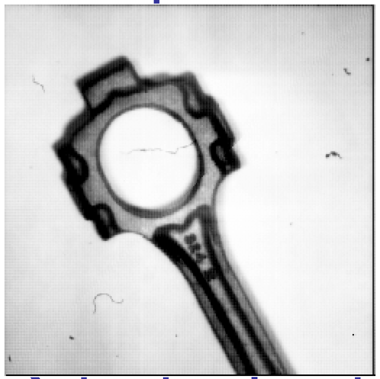
- 17% annual increment

Certification of olive plantation:

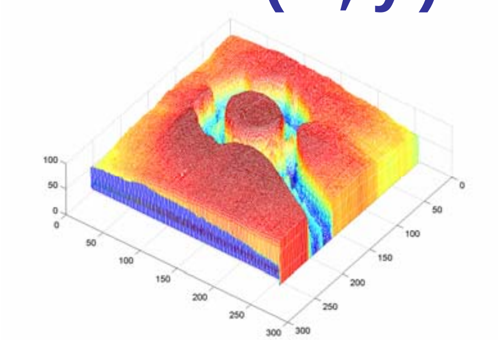


Generalities

- **Image:** continuous function **f** of two variables **x** and **y**, defined in a rectangular region of a plane.



- **f(x,y)** is the luminosity value of the image at coordinates **(x,y)**



- Images that we perceive in our visual activities (the ones we will consider) normally consist of **light reflected** by the objects that we see.

- Light being a form of **energy**, it must have a positive and finite value:

$$0 < f(x,y) < \infty$$

Generalities

Characterization:

(1) ***Illumination:*** amount of light in the scene that we visualize.

(2) Reflectance: amount of light reflected by the objects in the scene.

- These two components $i(x,y)$ and $r(x,y)$ combine in product form $f(x,y)$:

$$f(x,y) = i(x,y)\, r(x,y)$$

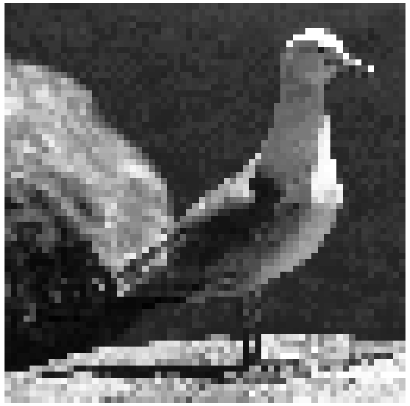
$$0 < i(x,y) < \infty$$
$$0 < r(x,y) < 1$$

- The nature of $i(x,y)$ is determined by the light source, and that of $r(x,y)$ by the characteristics of the object.
- Reflectance is bounded by 0 (total absorption) y 1 (total reflectance).

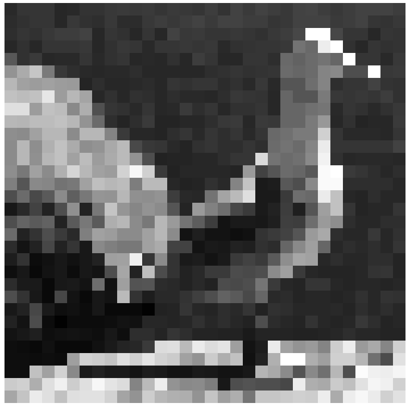
Generalities

- An image, being a continuous function, cannot be represented perfectly in a digital computer.
- It becomes necessary to sample the signal in a finite number of points.
- **Resolution:** image sampling rate; it determines the amount of image elements (pixels).

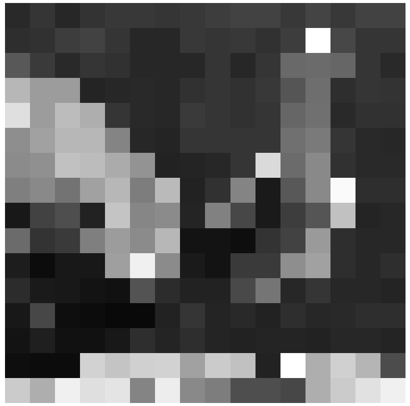
128 x 128 pixels



64 x 64 pixels



32 x 32 pixels



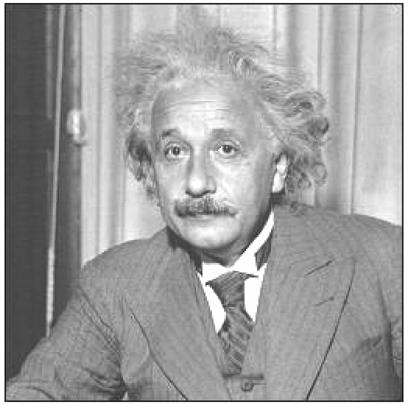
16 x 16 pixels

- Goal: use minimal resolution required for adequate processing

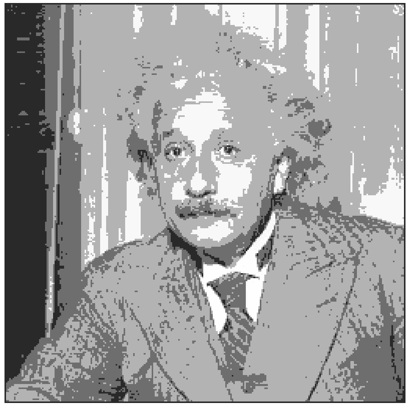
Generalities

- It becomes necessary to quantize each sampled value in the finite size of a memory unit in the computer.

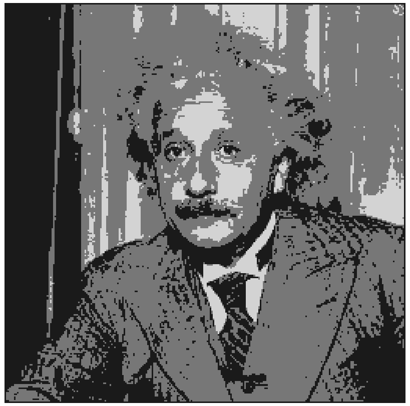
- **Quantization:** intensity levels used to represent the value of a pixel.



256 niveles



50 niveles



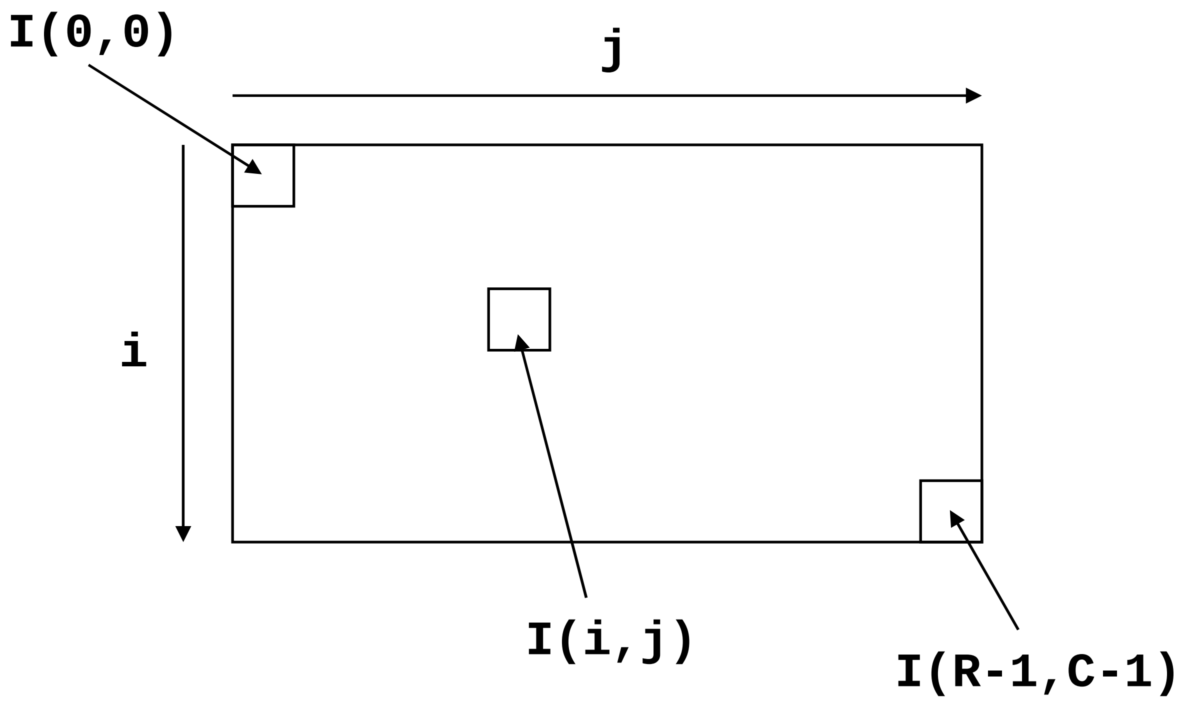
10 niveles

2 niveles

- The justification of binary vision lies in the fact that in many applications two level quantization is sufficient.

Generalities

- **Digital images:** discrete approximation **`I(i,j)`** of an image **`f(x,y)`** in the form of a matrix **`R X C`**.

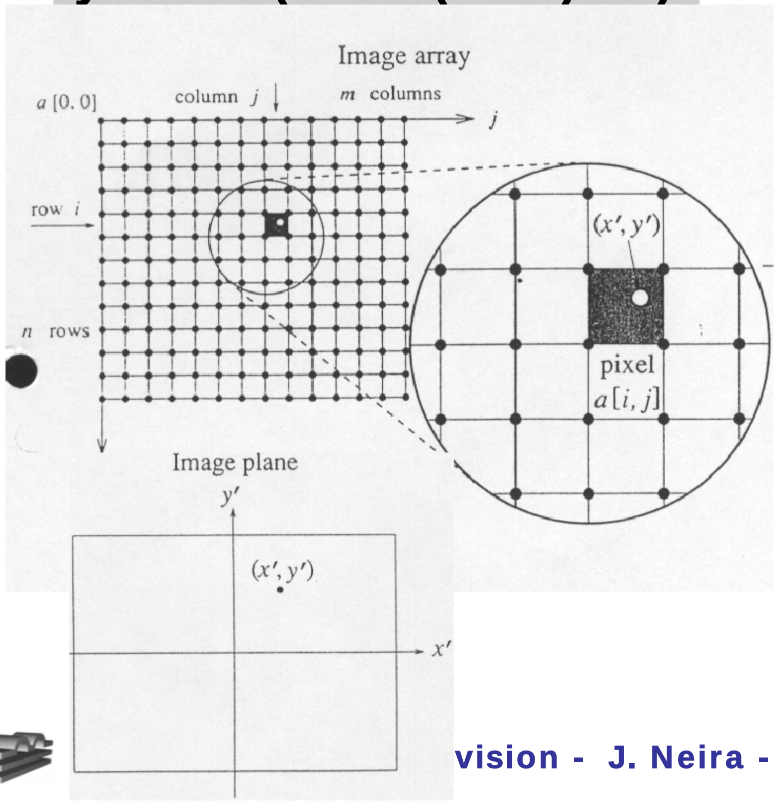


- **Pixel:** a sample of **`f(x,y)`** stored in **`I(i,j);`** usually a positive integer of 8 bits [0:255].

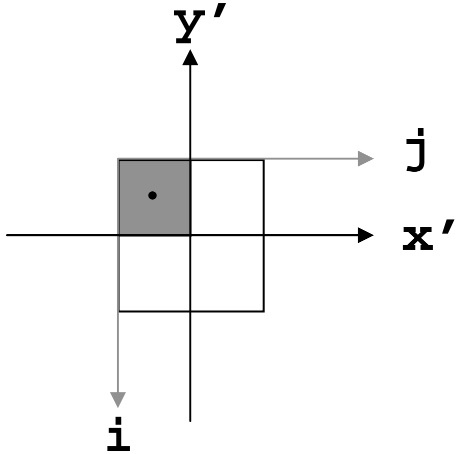
Generalities

- The coordinates (x', y') of the center of the pixel in the image plane can be computed from the coordinates (i, j) of the pixel:

$$x' = j - (R-1)/2$$
$$y' = -(i - (C-1)/2)$$

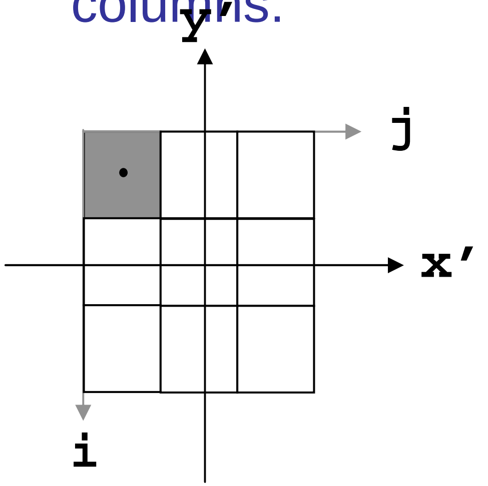


- Even number of rows and columns:




```
   R  = 2
   C  = 2
   i  = 0
   j  = 0
x'  = -1/2
 y'  = 1/2
```

- Odd number of rows and columns:

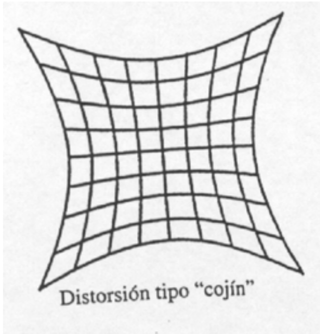



```
 R  = 3
 C  = 3
 i  = 0
 j  = 0
x'  = -1
y'  = 1
```

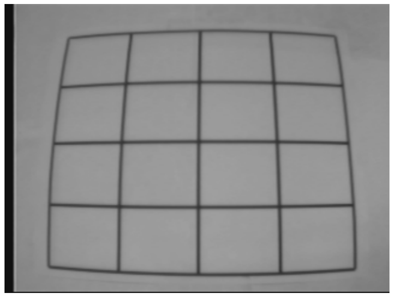
Assumptions

1. The origin of the system of the image plane coordinates corresponds to the center of the digital image matrix.

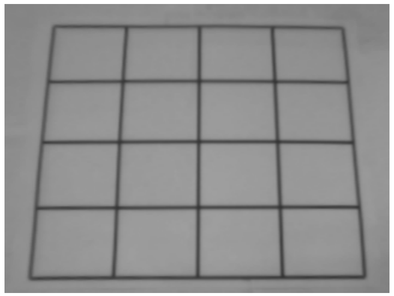
2. Lens imperfections, camera construction errors can be corrected during the *calibration* process.



Uncalibrated image:

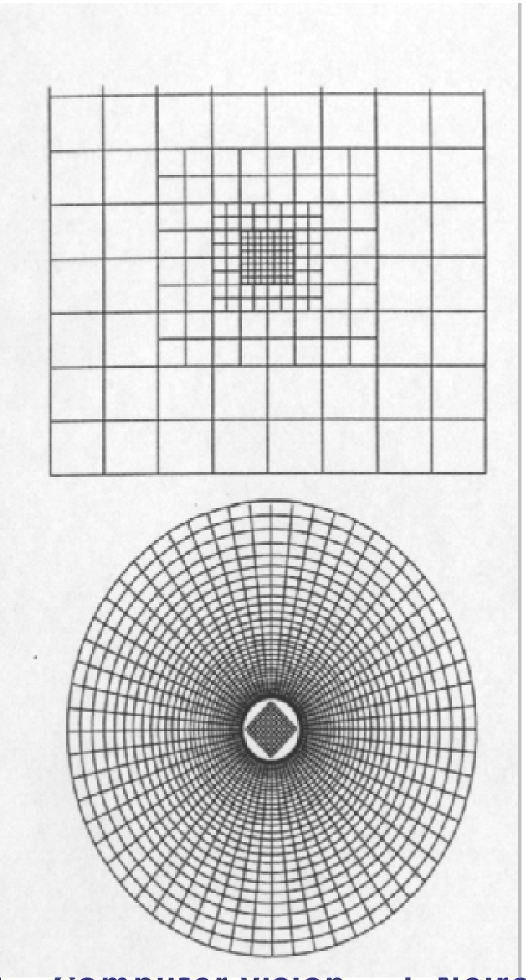


Calibrated image:



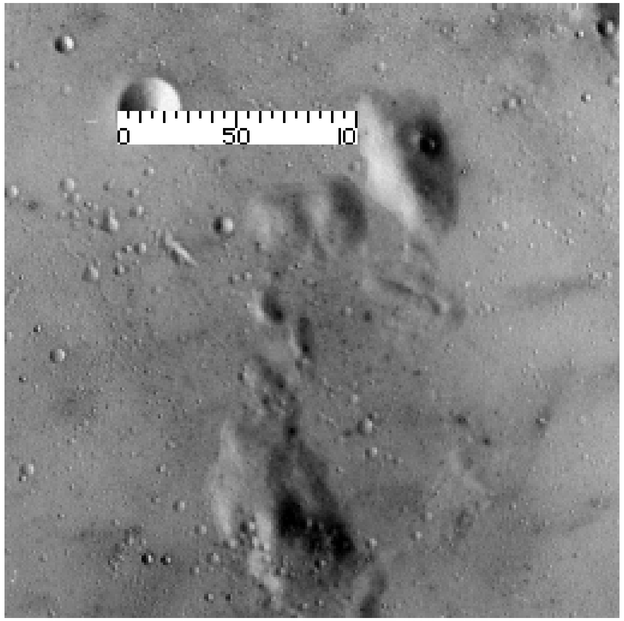
Assumptions

3. Sampling is carried out along a uniform rectangular mesh (many useful applications do not make this assumption).

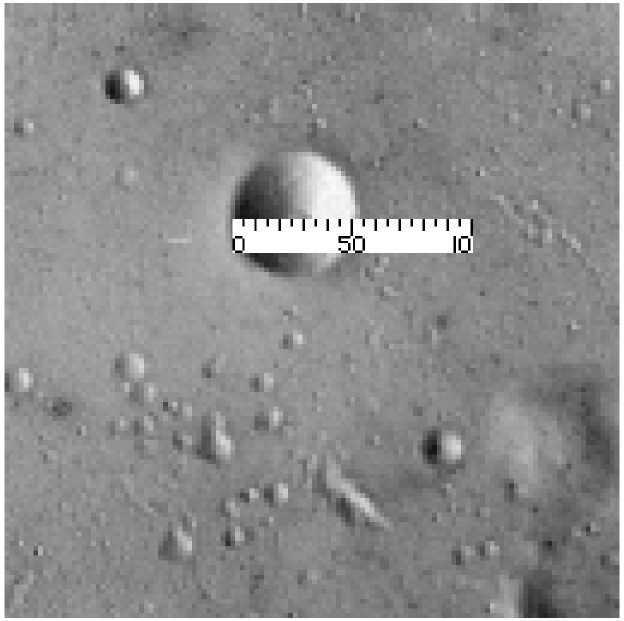


Assumptions

4. During the **calibration process**, spatial resolution is computed.



20 km/pixel

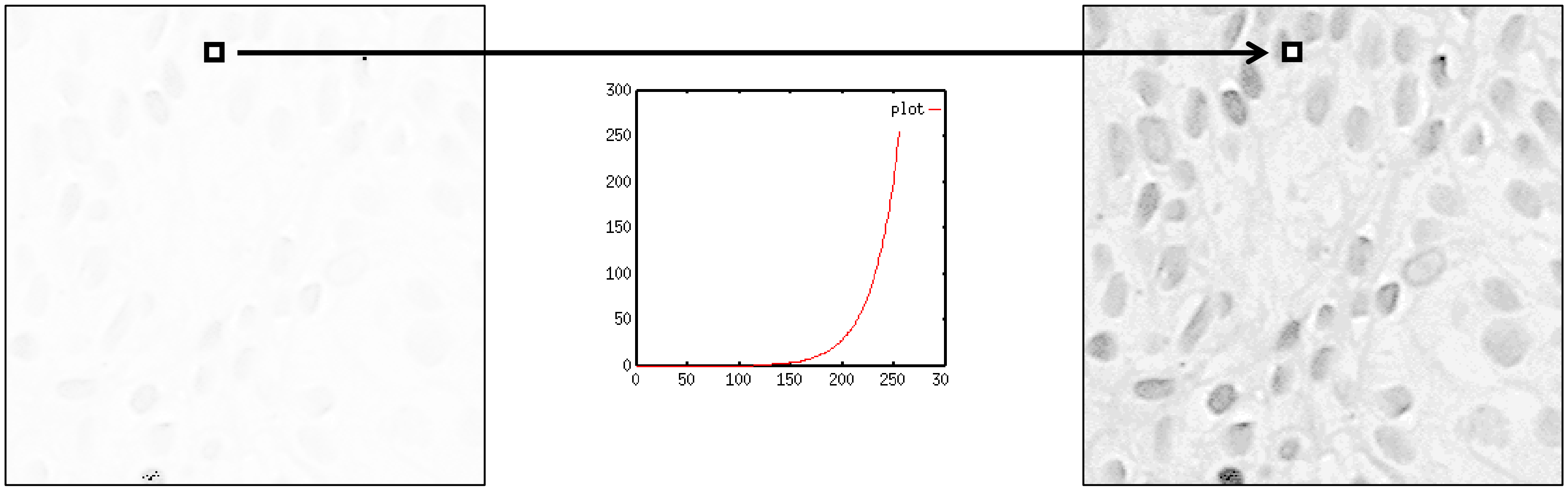


10 km/pixel

Four levels of processing

1. **Pixel level:** the result is computed for each pixel individually

- Example: contrast enhancement



Thresholding:

1.........1	0......0

0 **T** **255**

Inversion (negative):

?......................?

0 **255**

Application of a LUT:

```
for i = 0,..,R-1
   for j = 0,..,C-1
      B[i,j] = LUT[A[i,j]]
   endfor
endfor
```

- Efficiently implemented using LUTs.

Generalities

2. **Local level:** the result is computed using the set of neighbors of the pixel.

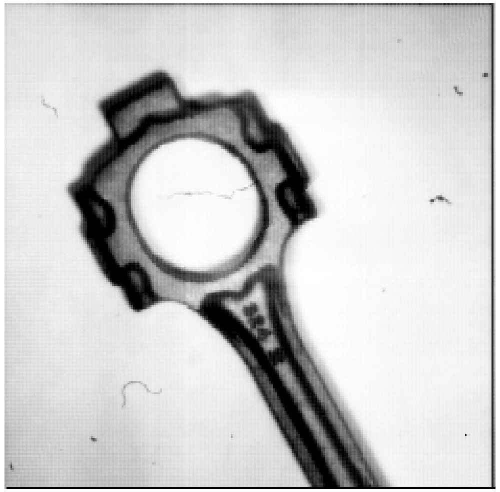
- Example: smoothing

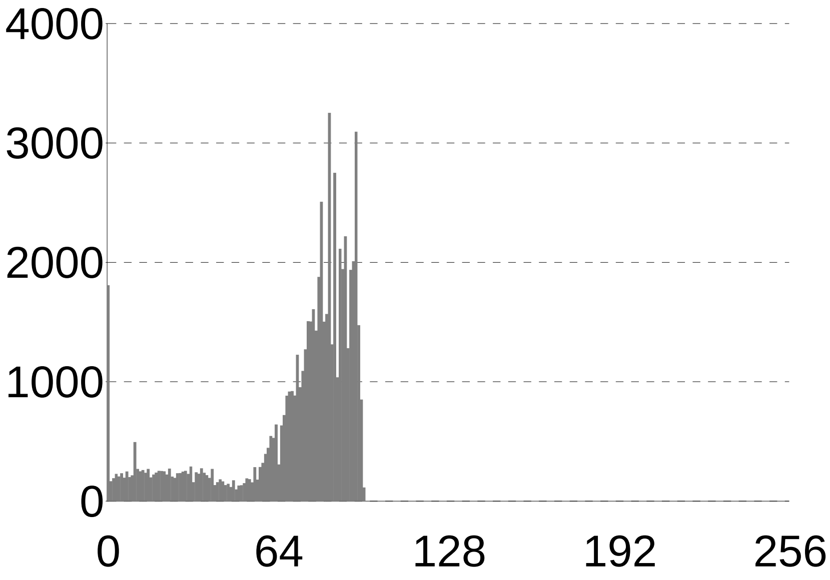
- It can be executed in real time in SIMD architectures.

Generalities

3. **Global level:** the result is computed using ALL the pixels in the image.

- Example: histograms





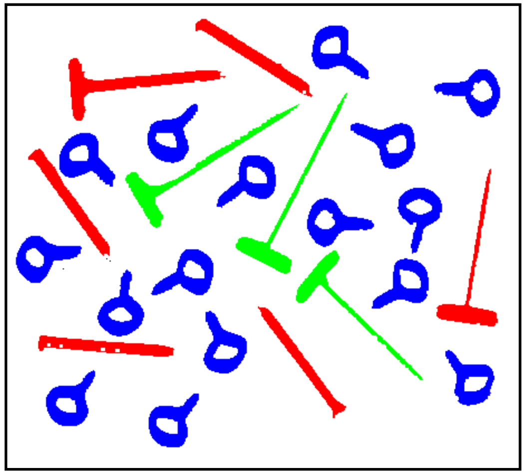
- These are the operations that make vision systems slow.

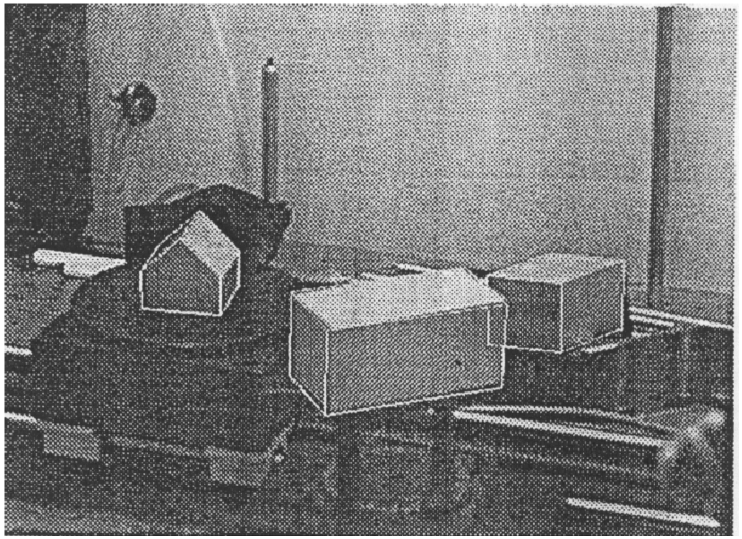
Generalities

4. **Object level:** these operations require to identify which pixels correspond to the same object, and then compute descriptors using those pixels.

- Examples: object identification using perimeter and pixel number.

The 3D problem is more complex





- This is the goal of computer vision.