

Elementos de Reconocimiento Visual

4to bimestre 2023

Docentes

Maria Elena Buemi
Daniel Acevedo
Pablo Negri



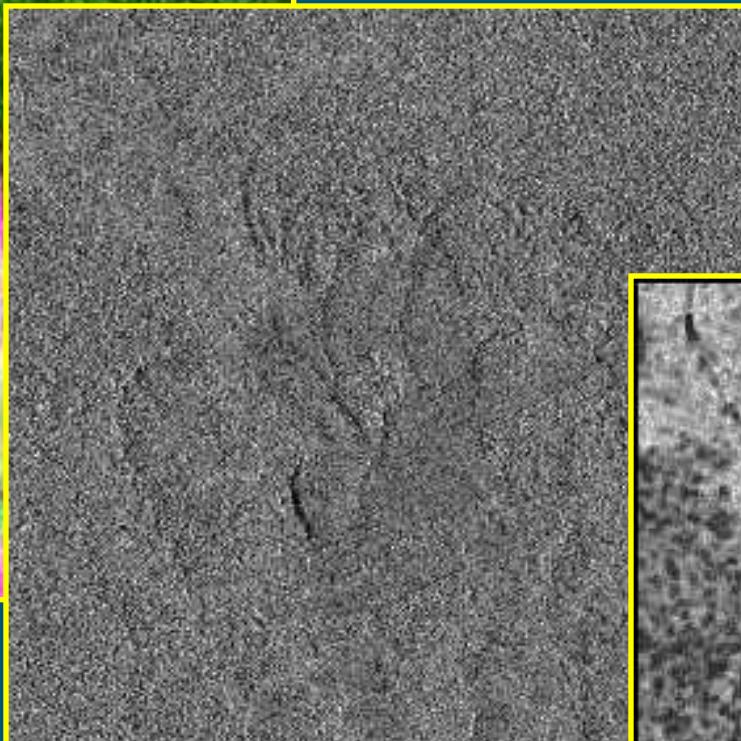
DEPARTAMENTO
DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

¿De qué se trata?

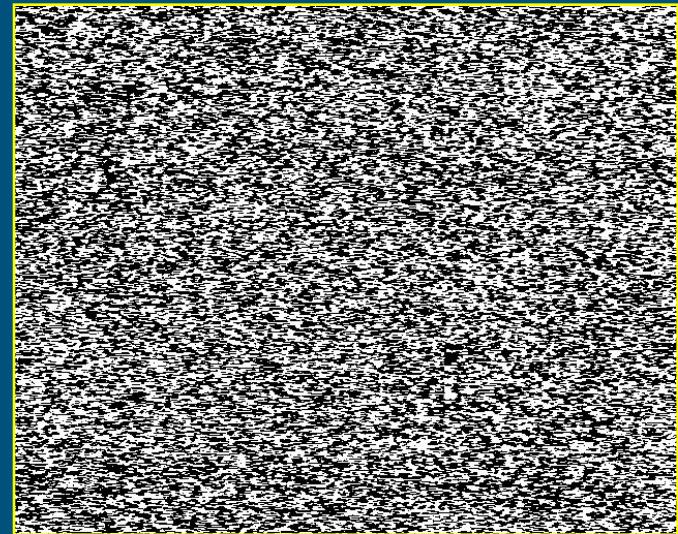
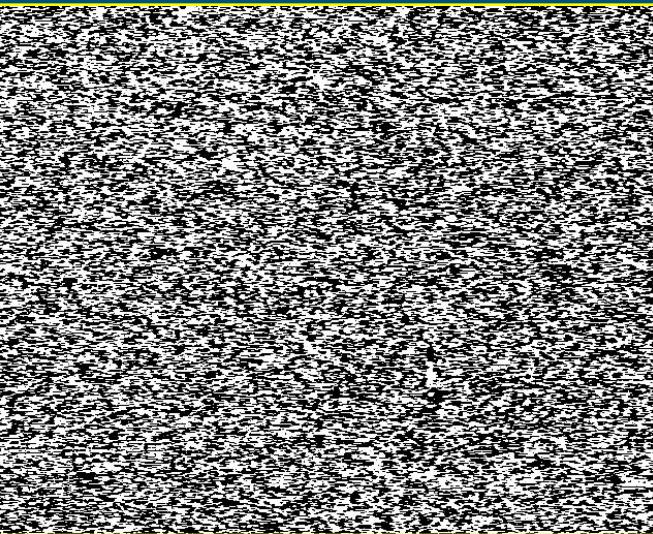
- Extraer información
 - Formular preguntas
 - Buscar respuestas

Ver lo invisible

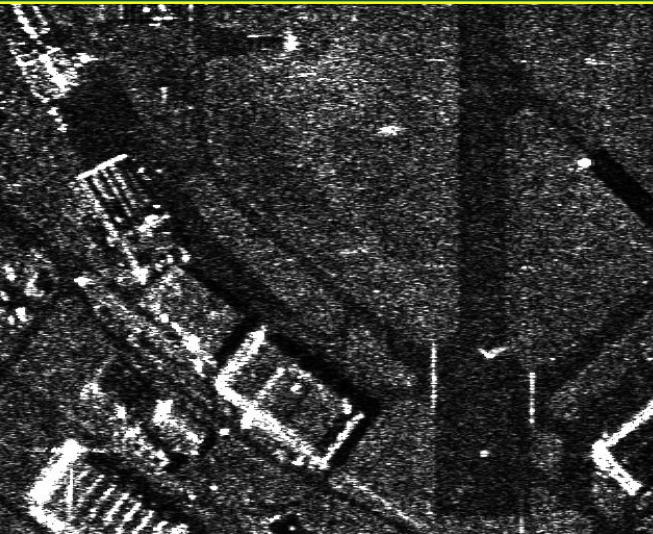


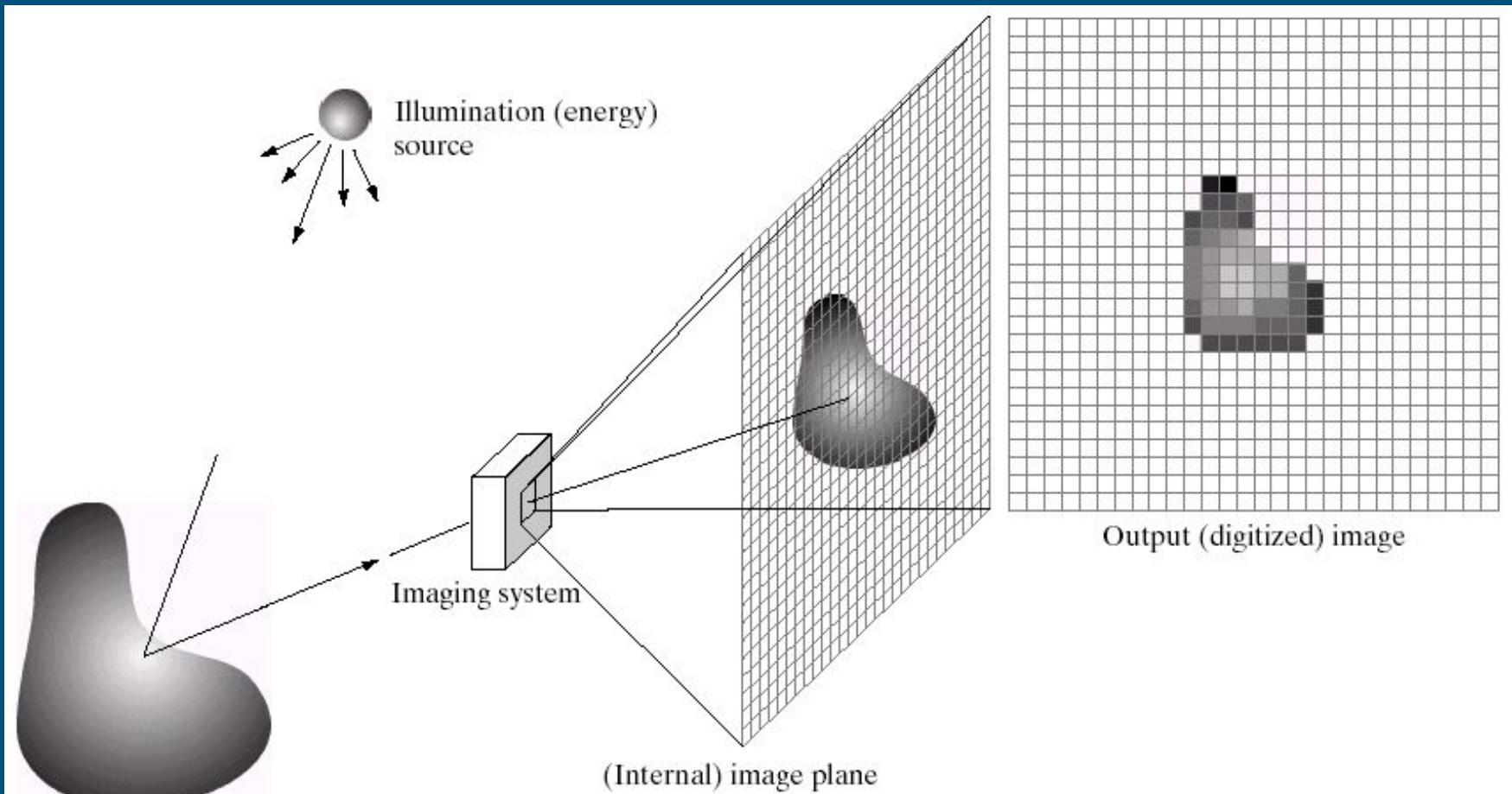
Datos e imágenes de radar de apertura sintética

Complejos



Intensidad y
Amplitud





a
b c d e

FIGURE 2.15 An example of the digital image acquisition process. (a) Energy (“illumination”) source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

Qué es una imagen?

- Cualquier función bi-dimensional

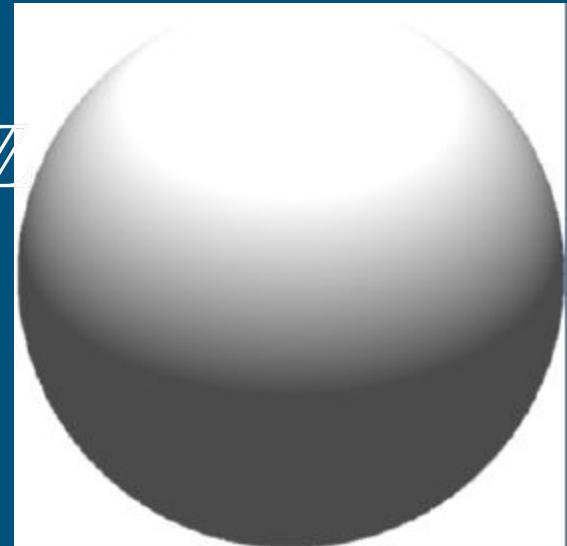
$f : D \rightarrow V$

D : **subconjunto de** $\mathbb{R} \times \mathbb{R}, \mathbb{Z} \times \mathbb{Z}$

coordenadas espaciales.

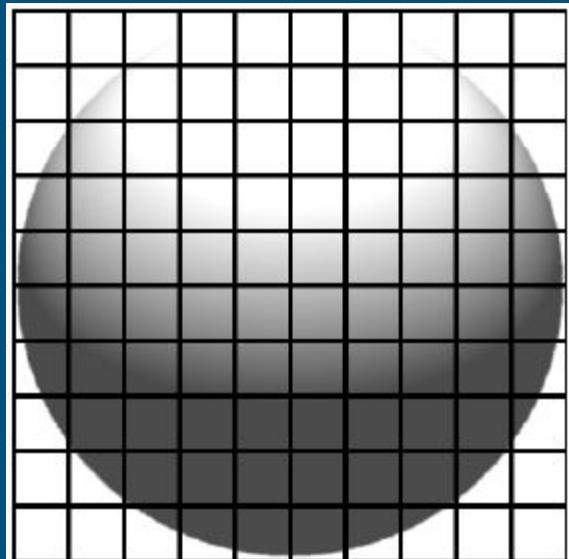
V : **subconjunto de** $\mathbb{C}, \mathbb{R}, \mathbb{Z}$

amplitud



Qué es una imagen?

$$f : D \rightarrow V$$



D: subconjunto de $Z \times Z$
V: subconjunto de C, R, Z.

$$f =$$

$$f(0,0) \dots \dots \dots f(0, M-1)$$

$$f(N-1,0) \dots \dots \dots f(N-1, M-1)$$

Una imagen digital es un muestreo discreto de la señal **f** continua.

Una imagen digital es una matriz, o array bi-dimensional, de números (enteros, reales o complejos)

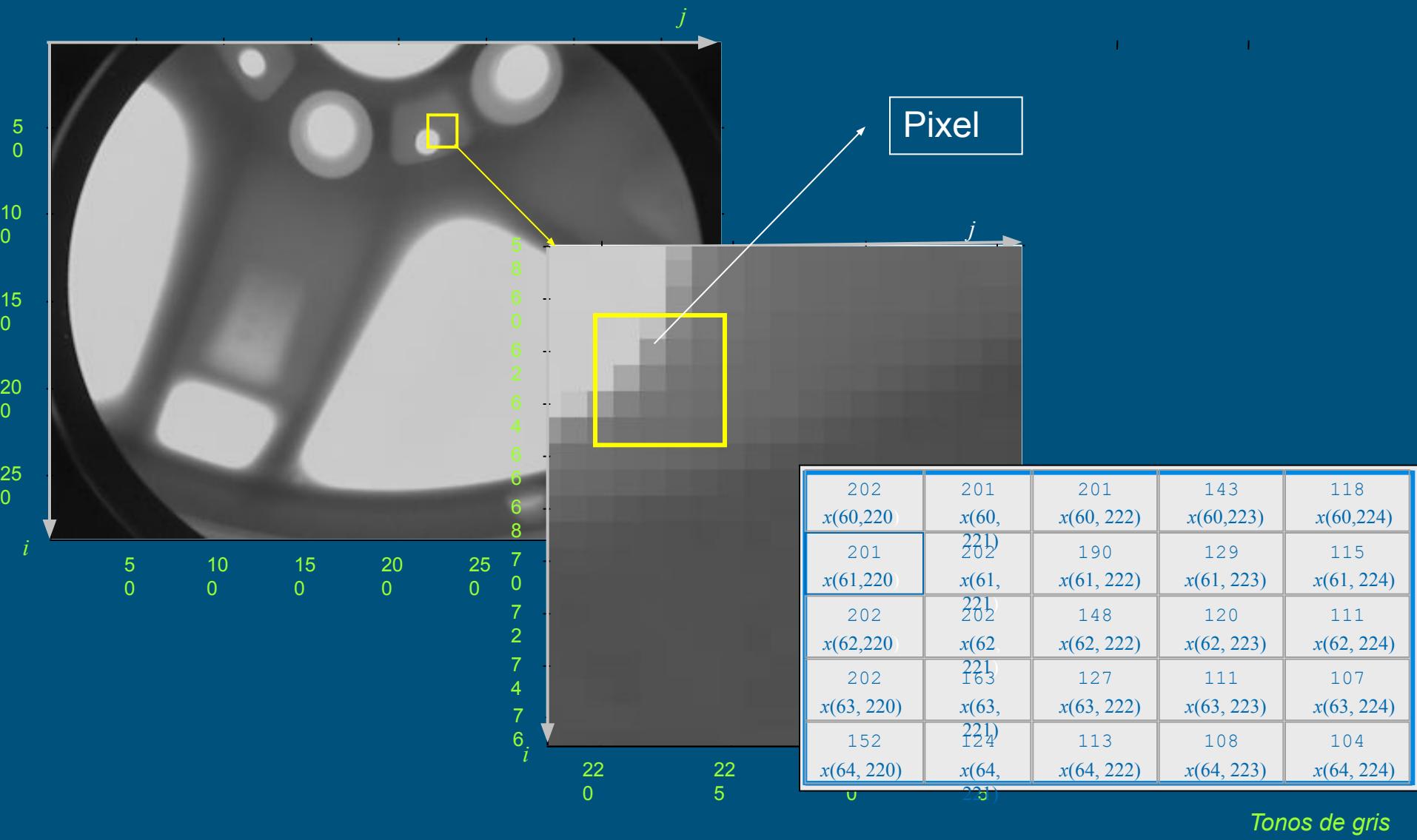
Cada celda de la matriz es un pixel

Un píxel

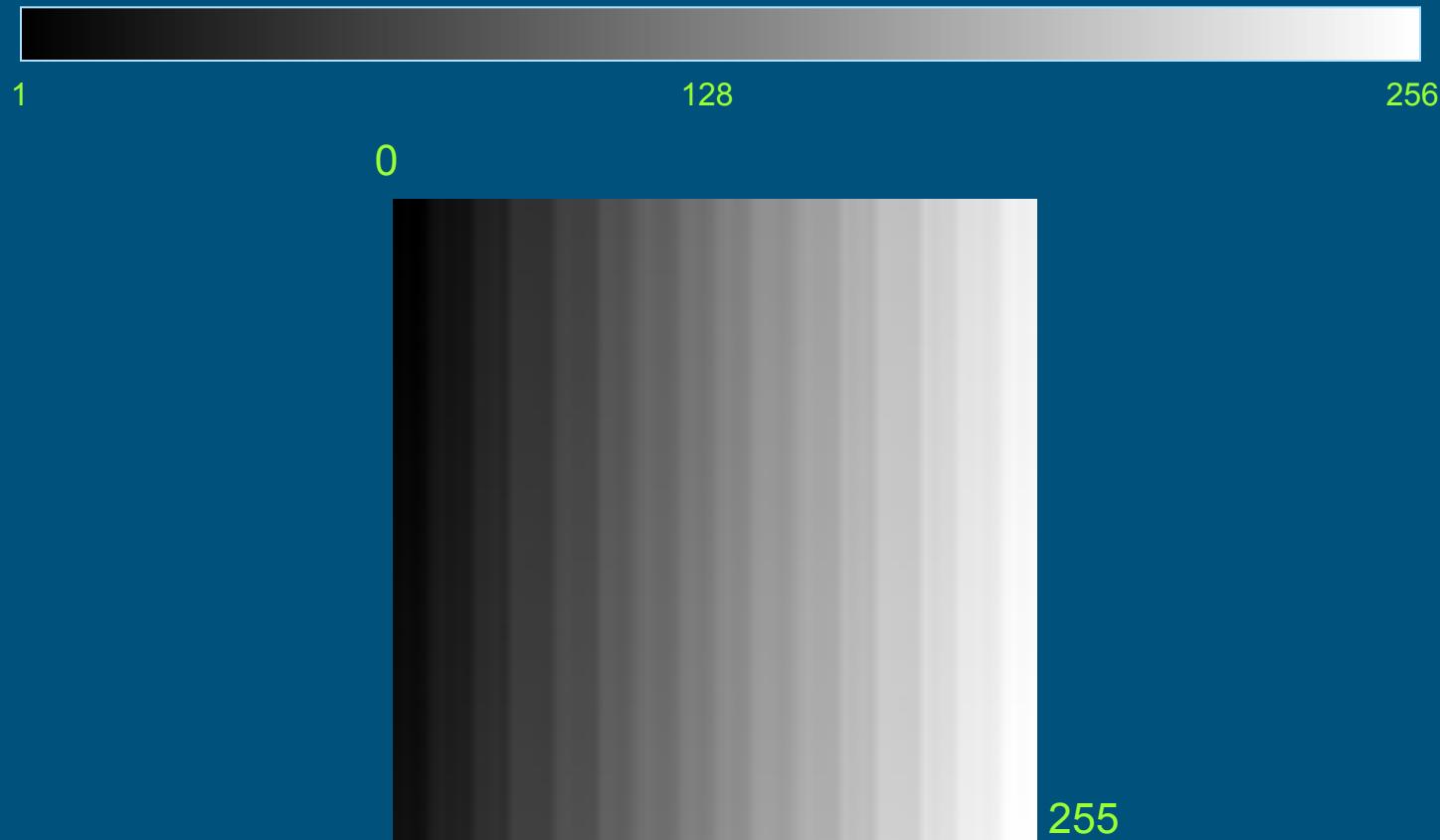
Imagen de 20x15

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|----|----|----|
| 90 | 67 | 68 | 75 | 78 | 98 | 185 | 180 | 153 | 139 | 132 | 106 | 70 | 80 | 81 | 69 | 69 | 67 | 35 | 34 |
| 92 | 87 | 73 | 78 | 82 | 132 | 180 | 152 | 134 | 120 | 102 | 106 | 95 | 75 | 72 | 63 | 75 | 42 | 19 | 29 |
| 63 | 102 | 89 | 76 | 98 | 163 | 166 | 164 | 175 | 159 | 120 | 103 | 132 | 96 | 68 | 42 | 49 | 46 | 17 | 22 |
| 45 | 83 | 109 | 80 | 130 | 158 | 166 | 174 | 158 | 134 | 105 | 71 | 82 | 121 | 80 | 51 | 12 | 50 | 31 | 17 |
| 39 | 69 | 92 | 115 | 154 | 122 | 144 | 173 | 155 | 105 | 98 | 86 | 82 | 106 | 83 | 76 | 17 | 29 | 41 | 19 |
| 34 | 80 | 73 | 132 | 144 | 110 | 142 | 181 | 173 | 122 | 100 | 88 | 141 | 142 | 111 | 87 | 33 | 18 | 46 | 36 |
| 37 | 93 | 88 | 136 | 171 | 164 | 137 | 171 | 190 | 149 | 110 | 137 | 168 | 161 | 132 | 96 | 56 | 23 | 48 | 49 |
| 66 | 117 | 106 | 147 | 188 | 202 | 198 | 187 | 187 | 159 | 124 | 151 | 167 | 158 | 138 | 105 | 80 | 55 | 59 | 54 |
| 127 | 136 | 107 | 144 | 188 | 197 | 188 | 184 | 192 | 172 | 124 | 151 | 138 | 108 | 116 | 114 | 84 | 46 | 67 | 54 |
| 143 | 134 | 99 | 143 | 188 | 172 | 129 | 127 | 179 | 167 | 106 | 118 | 111 | 54 | 70 | 95 | 90 | 46 | 69 | 52 |
| 141 | 137 | 96 | 146 | 167 | 123 | 91 | 90 | 151 | 156 | 121 | 93 | 78 | 82 | 97 | 91 | 87 | 45 | 66 | 39 |
| 139 | 137 | 80 | 131 | 162 | 145 | 131 | 129 | 154 | 161 | 158 | 149 | 134 | 122 | 115 | 99 | 84 | 35 | 52 | 30 |
| 137 | 133 | 56 | 104 | 165 | 167 | 174 | 181 | 175 | 169 | 165 | 162 | 158 | 142 | 124 | 103 | 67 | 19 | 31 | 23 |
| 135 | 132 | 65 | 86 | 173 | 186 | 200 | 198 | 181 | 171 | 162 | 153 | 145 | 135 | 121 | 104 | 53 | 14 | 15 | 33 |
| 132 | 132 | 88 | 50 | 149 | 182 | 189 | 191 | 186 | 178 | 166 | 157 | 148 | 131 | 106 | 78 | 28 | 10 | 15 | 44 |

Digitalización y Representación

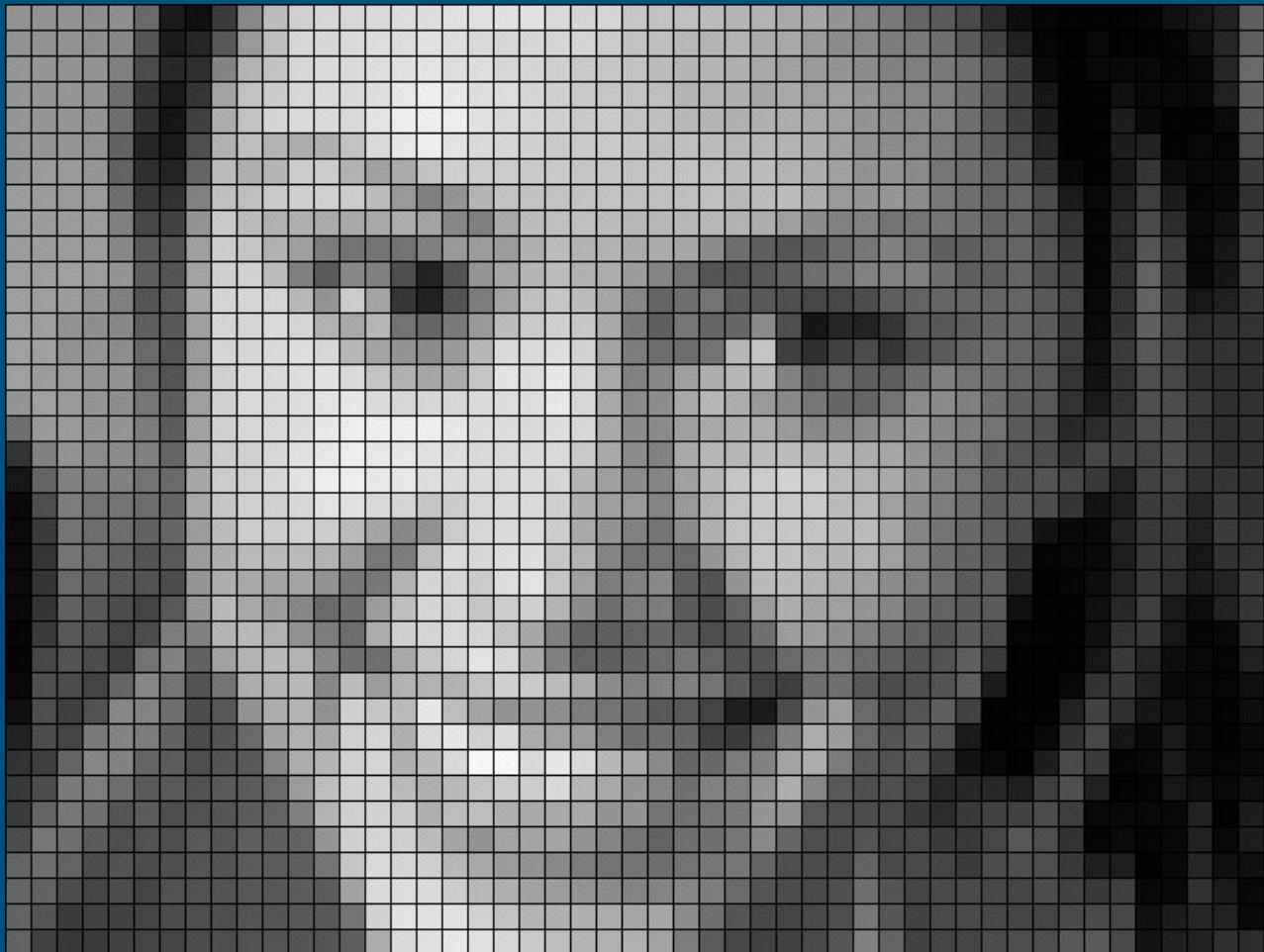
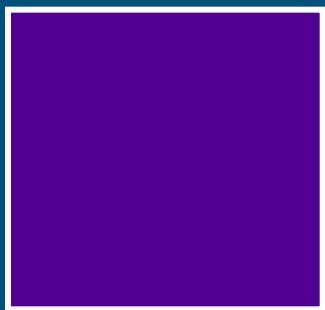


Niveles de gris



Una forma de visualizar una imagen...

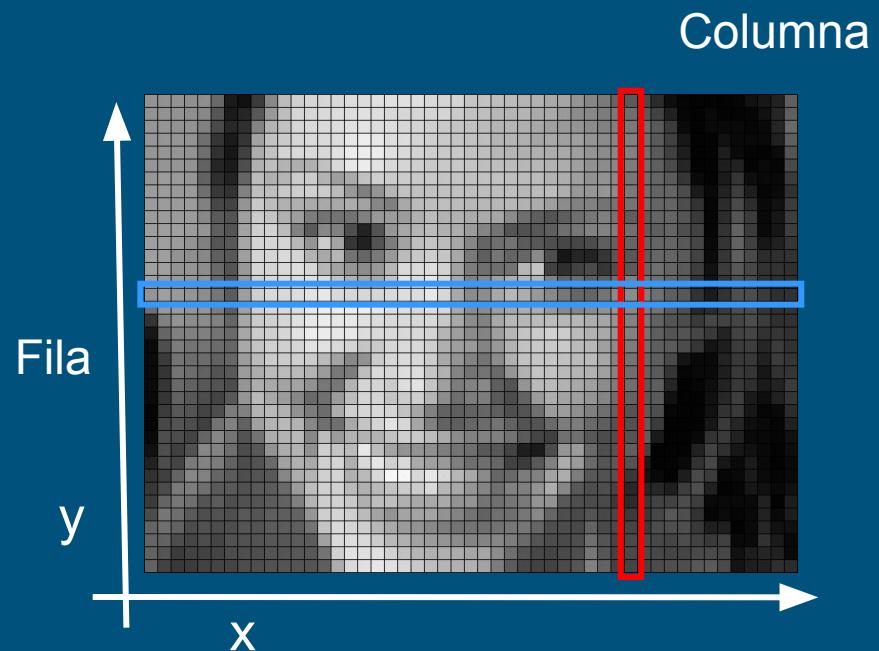
Un píxel



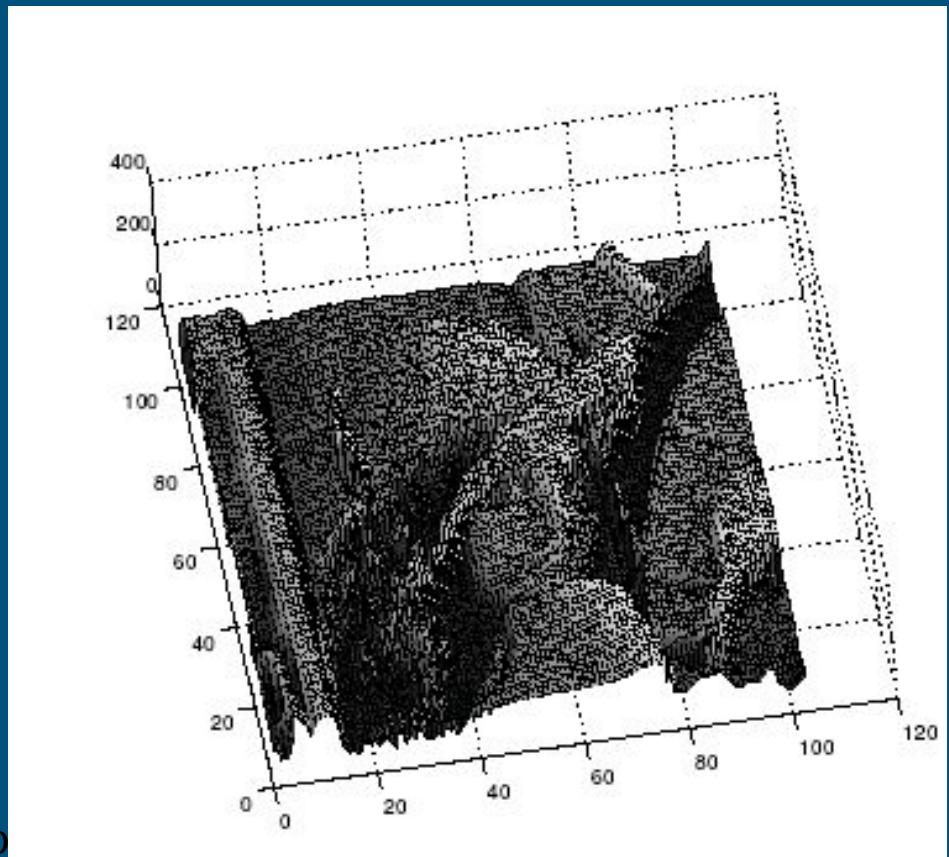
- N° de columnas de la matriz: ancho de la imagen (*width*).
 - N° de filas de la matriz: alto de la imagen (*height*).
 - Eje horizontal: eje x.
 - Eje vertical: eje y.
 - Normalmente el tamaño de la imagen se expresa como: ancho x alto
-
- Tamaños usuales:

320x240, 640x480,

800x600, 1024x768, ...



Una imagen la puedo representar como una superficie



¿Qué tipo de datos tiene cada celda de la matriz?

Imagen binaria:

1 píxel = 1 bit

- 0 = negro; 1= blanco

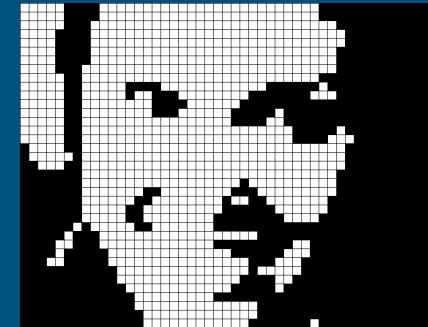


Imagen en escala de grises:

1 píxel = 1 byte

- Permite 256 niveles de gris
- 0 = negro; 255 = blanco

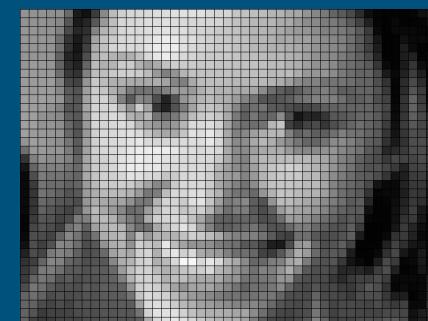
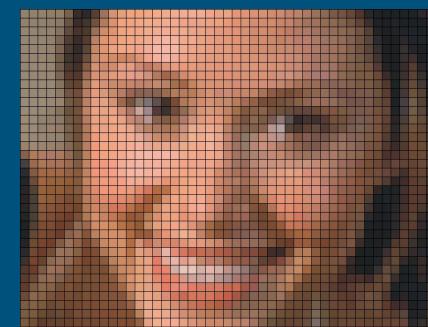


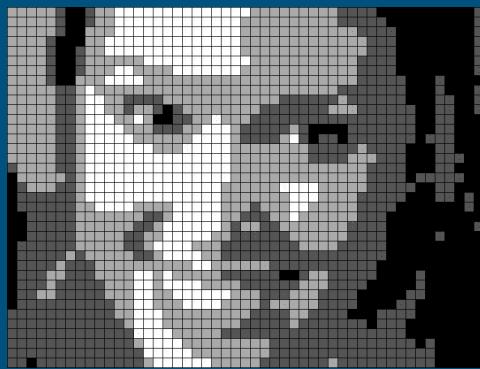
Imagen en color:

1 píxel = 3 bytes

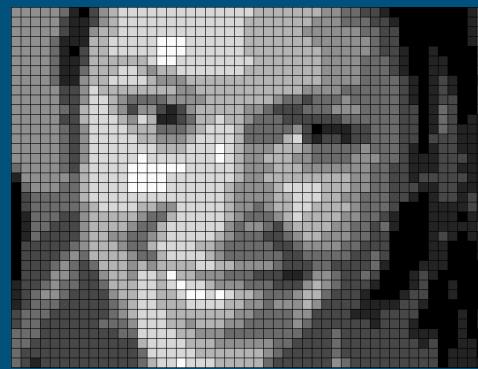
- Cada píxel consta de 3 valores: (Rojo, Verde, Azul)
- Un byte por color
- 16,7 millones de colores posibles



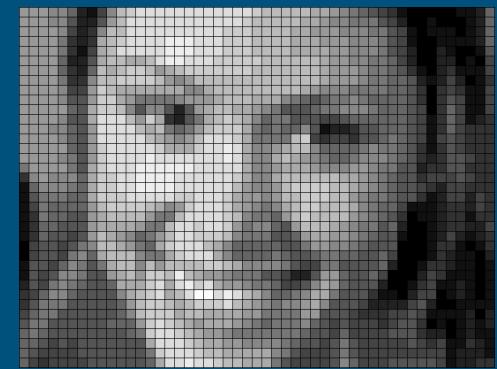
Un nivel de gris, o un color, se puede representar con más o menos bits: profundidad de color (*depth*).



2 bits por píxel



3 bits por píxel



4 bits por píxel

Hi-color: método reducido para representar colores

1 píxel = 2 bytes

- 5 bits por cada color (Rojo, Verde, Azul)

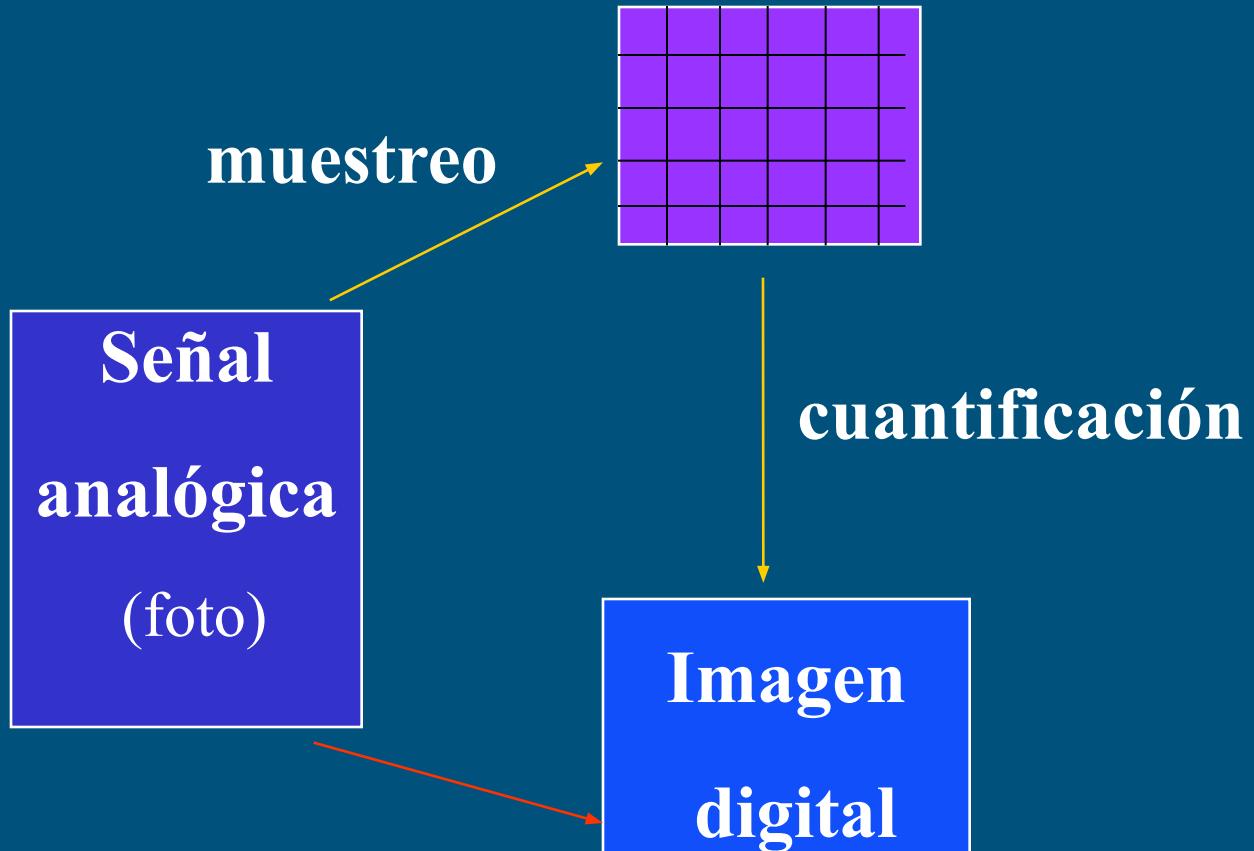
Imágenes en punto flotante: útiles en procesos intermedios

1 píxel = 1 float o un double

Muestreo y Cuantización

- Una función $f(x,y)$ continua debe ser digitalizada tanto espacialmente como en su amplitud.
- **Muestreo:** digitalización de las coordenadas espaciales.
- **Cuantificación** del nivel de gris: digitalización de la amplitud.

analógico a digital



Qué muestreo y cuantificación son necesarios para una buena aproximación?

- La **Resolución** es el grado de detalle discernible y depende de **muestreo y cuantificación**.
- Una vez que una función es **muestreada**, se puede recuperar, y cómo?
- Cuán **finamente** hay que muestrear una función para preservarla?
- Efectos del muestreo: ***Aliasing***

Muestreo



Resolución

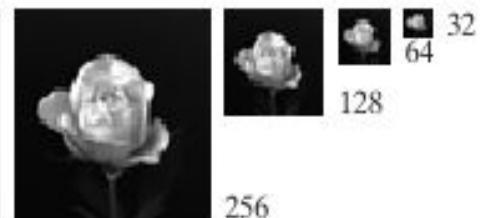
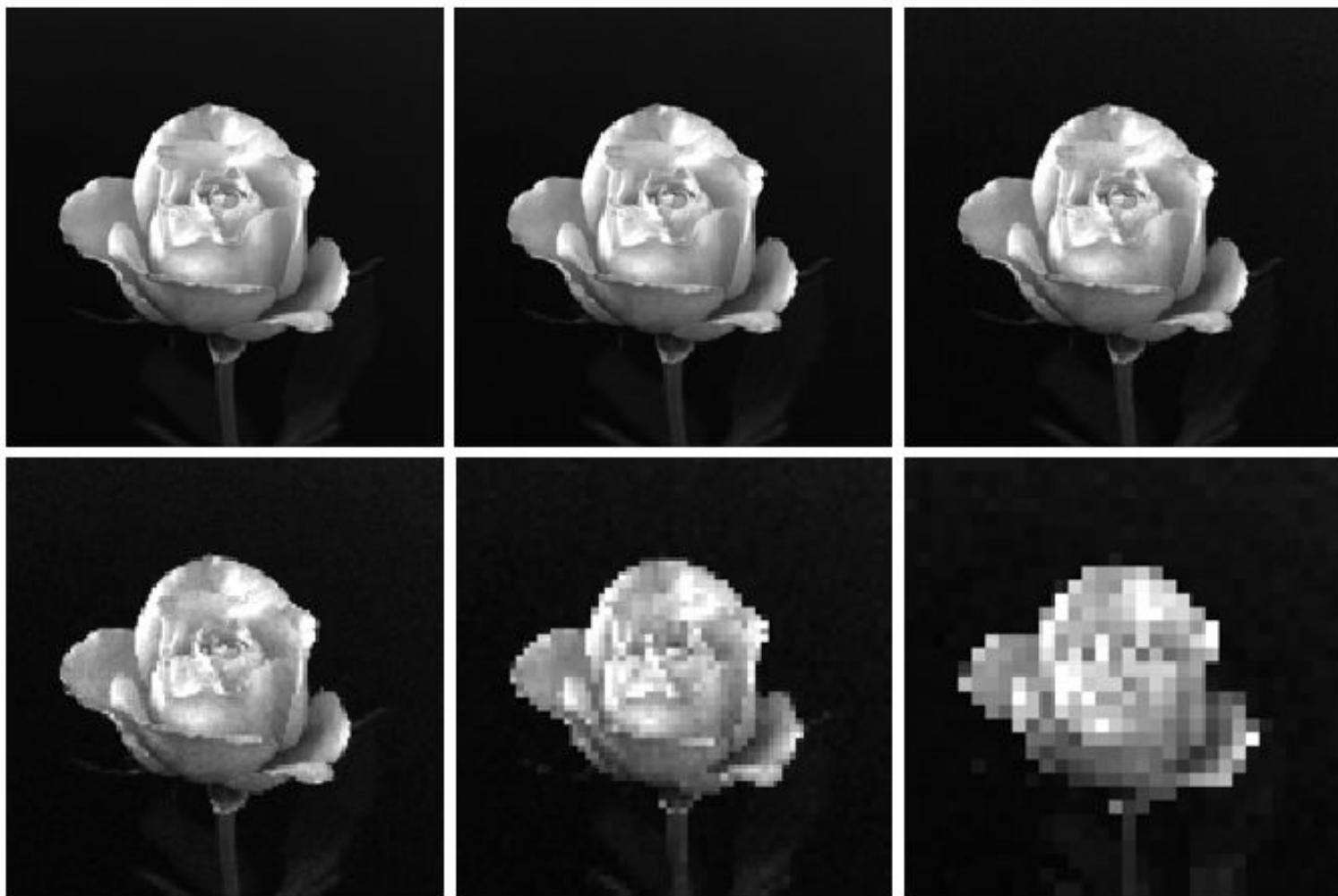


FIGURE 2.19 A 1024×1024 , 8-bit image subsampled down to size 32×32 pixels. The number of allowable gray levels was kept at 256.

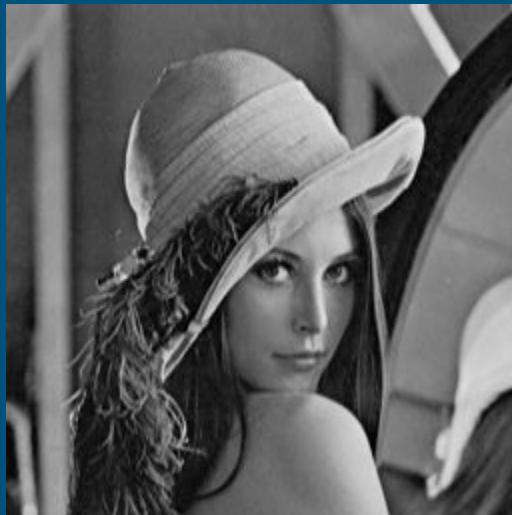
Cantidad de pixels que tiene la imagen



a b c
d e f

FIGURE 2.20 (a) 1024×1024 , 8-bit image. (b) 512×512 image resampled into 1024×1024 pixels by row and column duplication. (c) through (f) 256×256 , 128×128 , 64×64 , and 32×32 images resampled into 1024×1024 pixels.

Cuantización



8 bits - 256 colores



4 bits - 16 colores



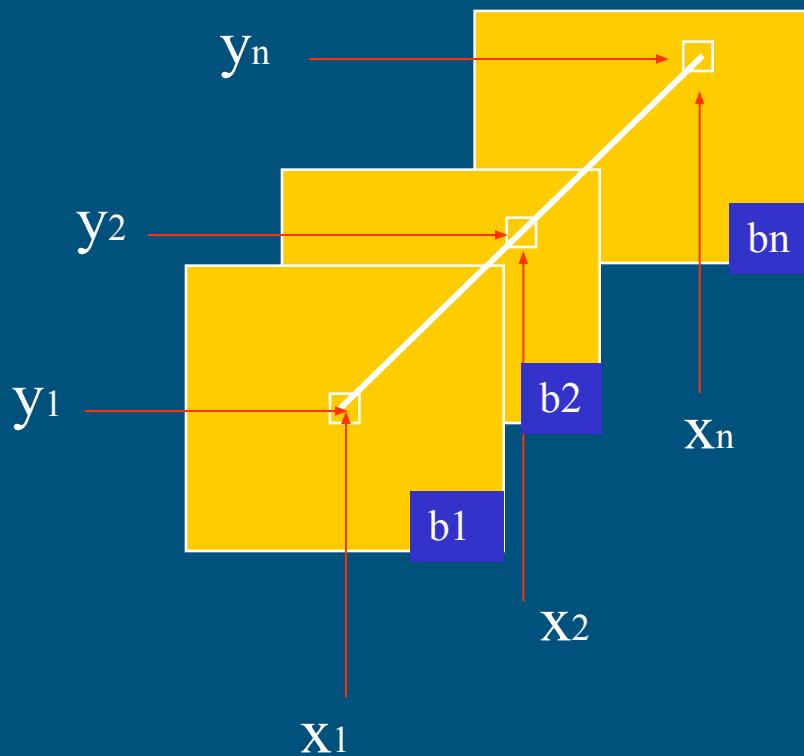
1 bit - 2 colores

Imagen digital multidimensional

(x , y) coordenadas espaciales

n bandas, s = nivel de gris

pixel=(x,s)

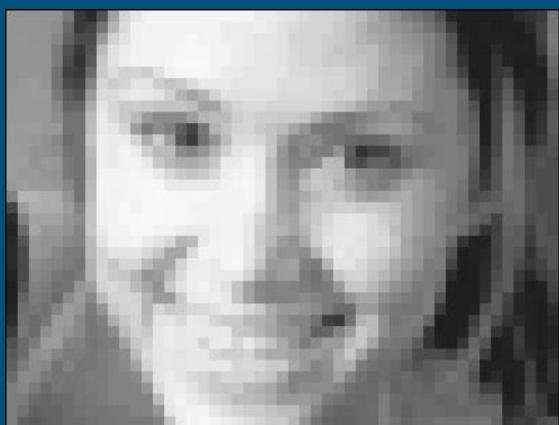


- Es un array multidimensional
- Pixel vector n-dimensional
- s: vector n-dimensional de números reales

Imágenes multibanda o multicanal:

Cuando los pixels representan magnitudes en distintos dominios físicos, decimos que la imagen es multicanal.

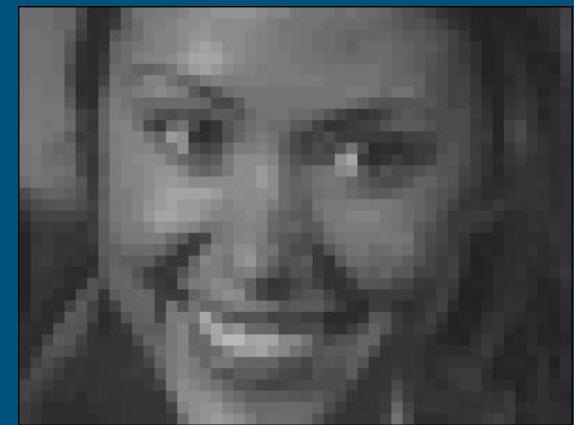
Imagen en color → Imagen con 3 canales: canal R (rojo), canal G (verde), canal B (azul).



Canal R

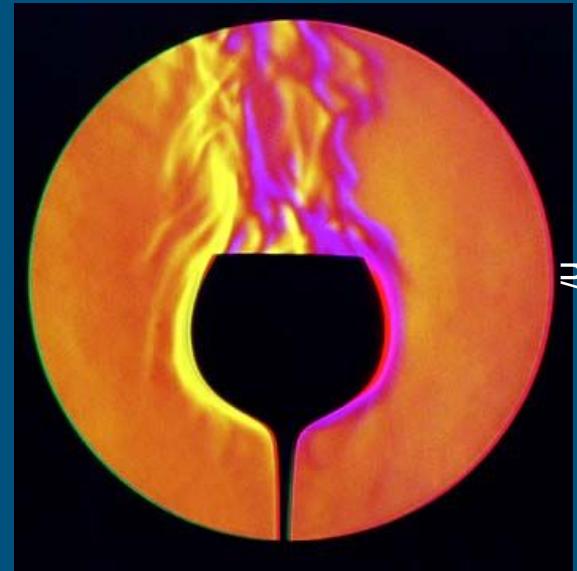


Canal G



Canal B

En algunas aplicaciones imágenes de satélite, visión nocturna, etc., suelen usarse canales para frecuencias no visibles, infrarrojo, ultravioleta, entre otras



Qué podría representar cada píxel de una imagen?

- Brillo de los objetos de una escena (foto tomada con una cámara común)
- Las características de absorción de un tejido humano (imágenes de rayos X)
- La señal de retorno de radar (imágenes de radar)
- Perfil de temperatura de una región (imágenes infrarrojas)
- Campos gravitacionales en un área (imágenes geofísicas)

Tópicos

- Mejora de la imagen
- Transformaciones de Intensidad
- Filtros Espaciales y Realce
- Procesamiento en el Dominio de las Frecuencias
- Ruido y Restauración

Tópicos

- Color
- Bordes
- Segmentación
- Extracción de Características
- Reconocimiento de Objetos

Un ejemplo

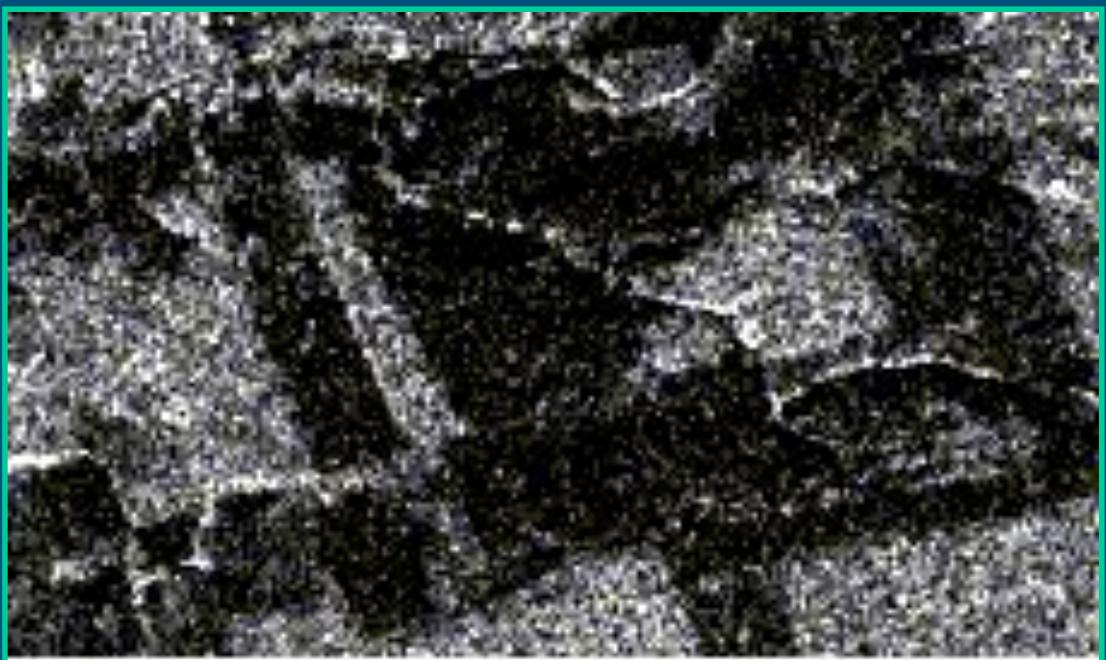
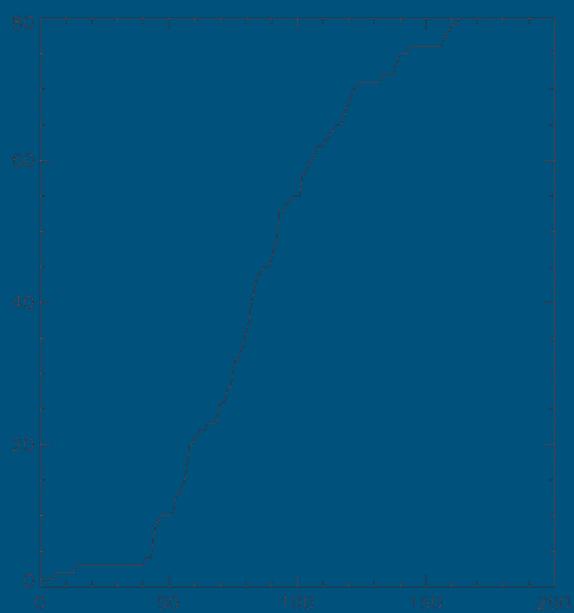
- Mejorar el contraste de la imagen



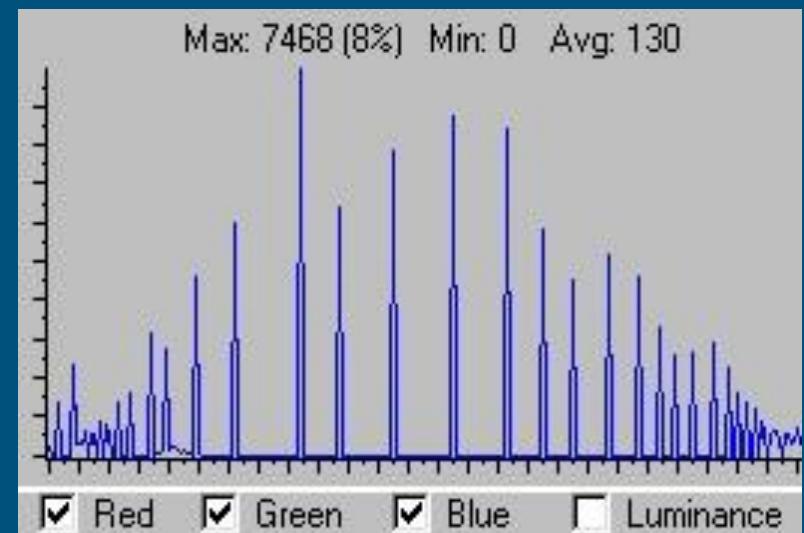
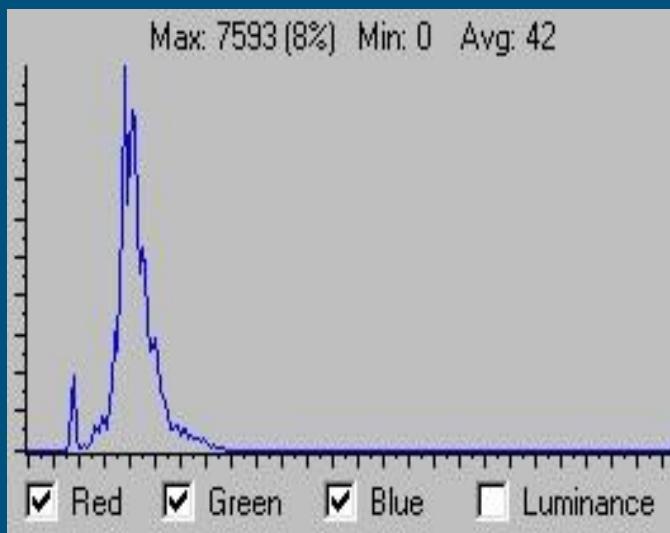
Solución: modelar

- Contraste (perceptual, cualitativo)
 - Varianza (estadística, cuantitativa)
 - Se desea aumentar el contraste
- Máxima varianza \leftrightarrow distribución uniforme

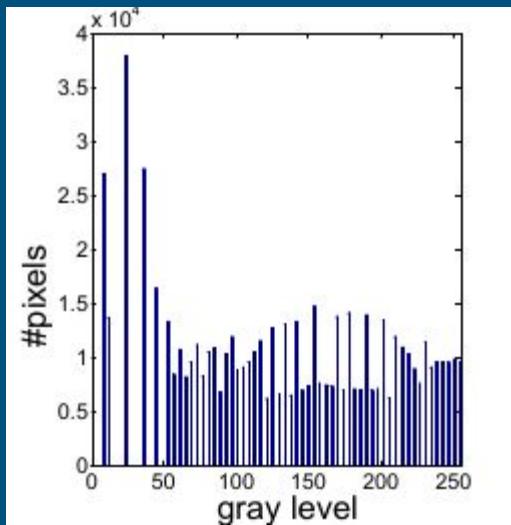
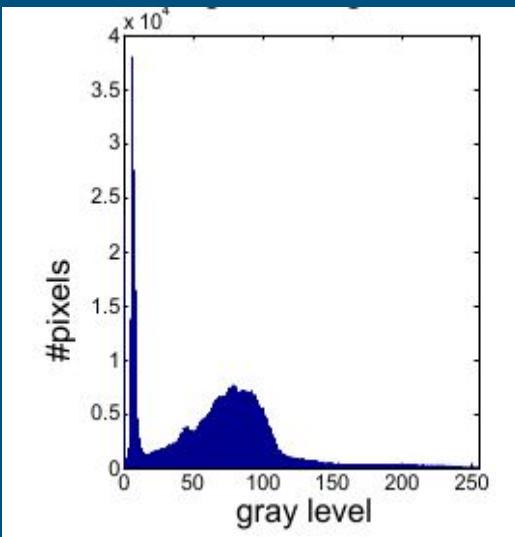
Resultado



Ecualización de Histogramas



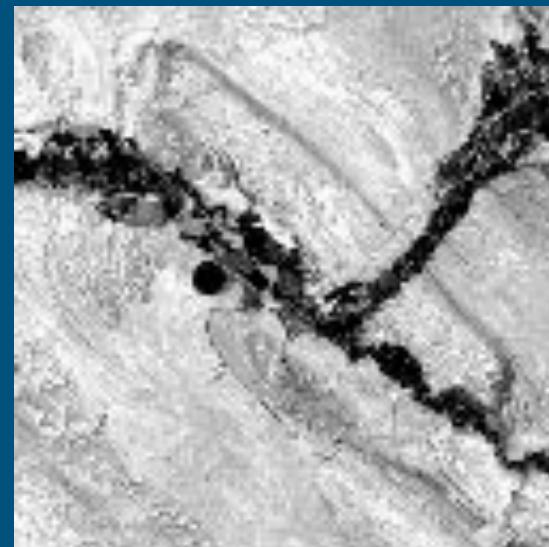
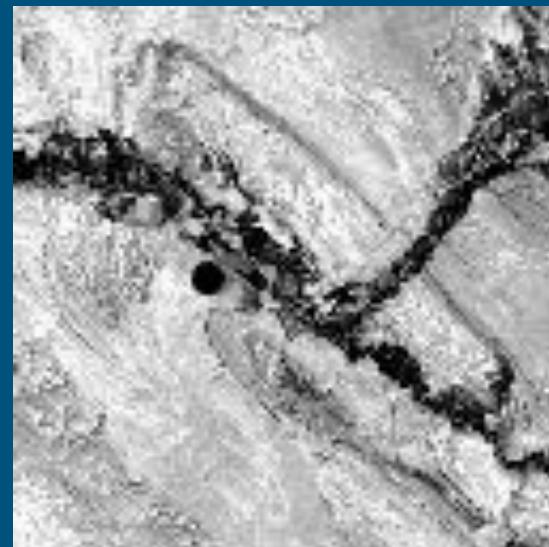
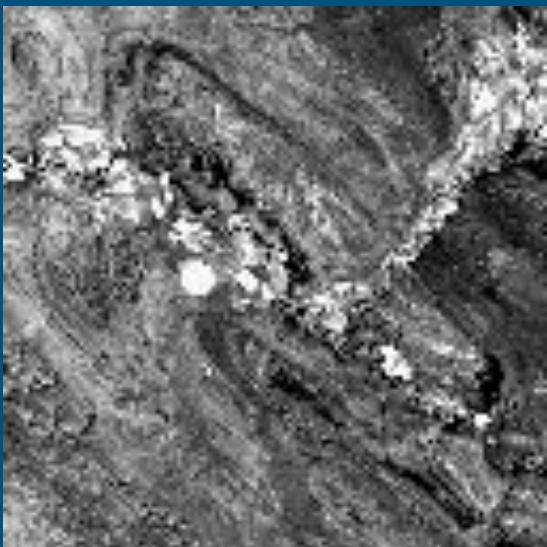
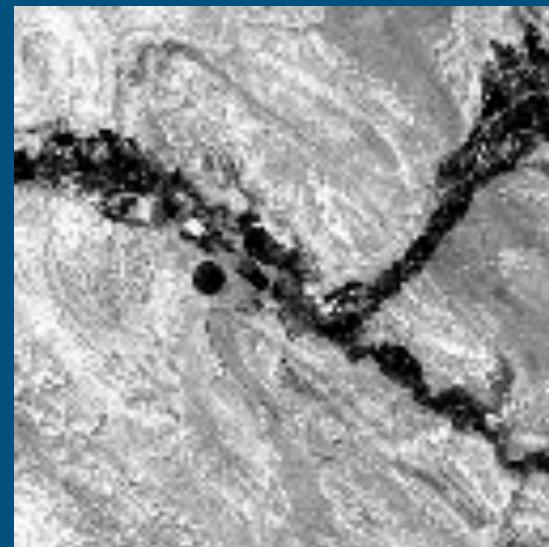
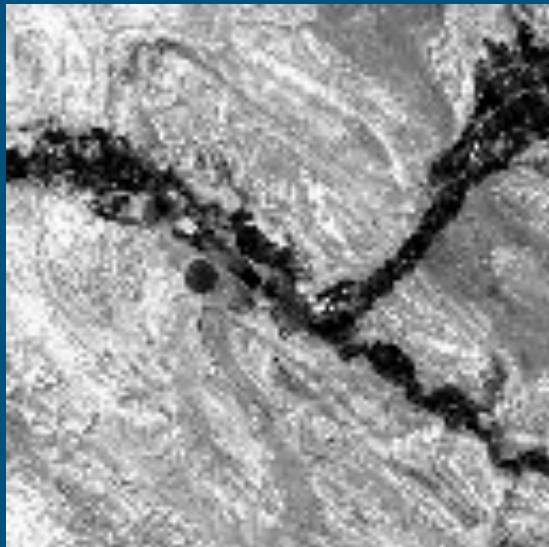
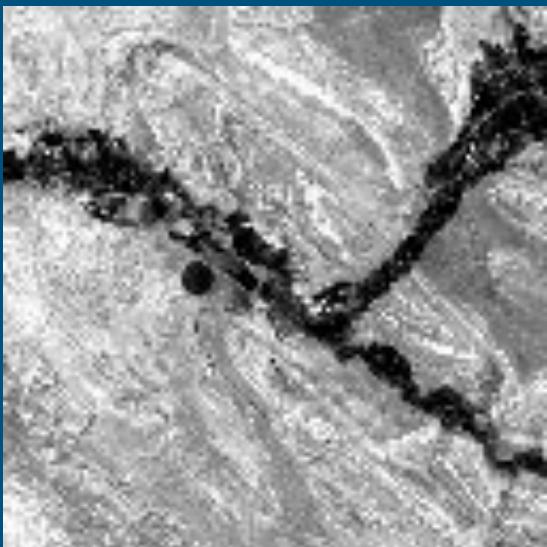
Histogramas en la mejora de la imagen



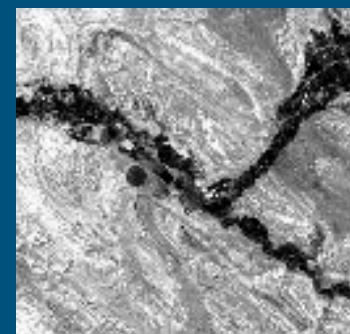
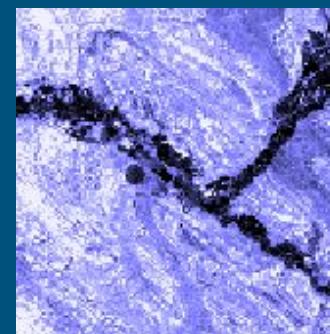
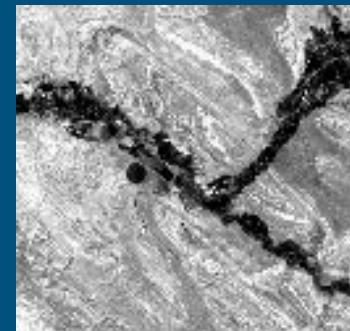
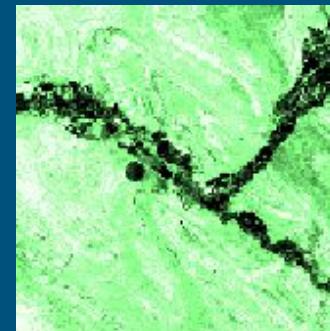
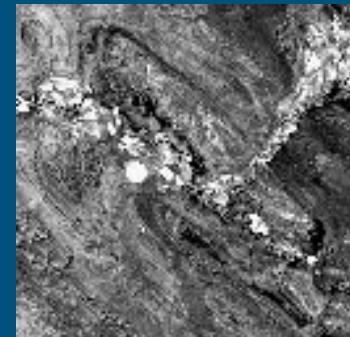
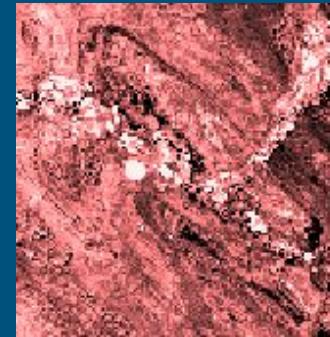
Cómo encontrar la mejor composición de color?

Ejemplo

Imágenes mulitespectrales (6 bandas)

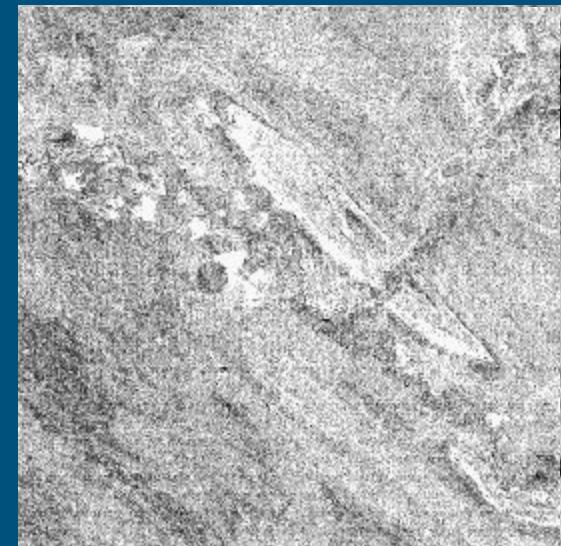
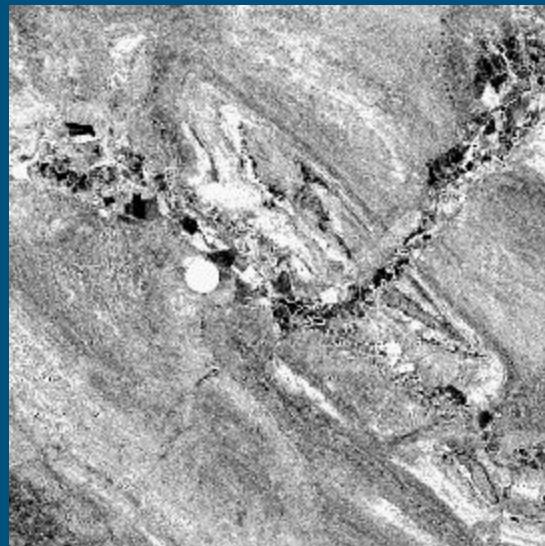
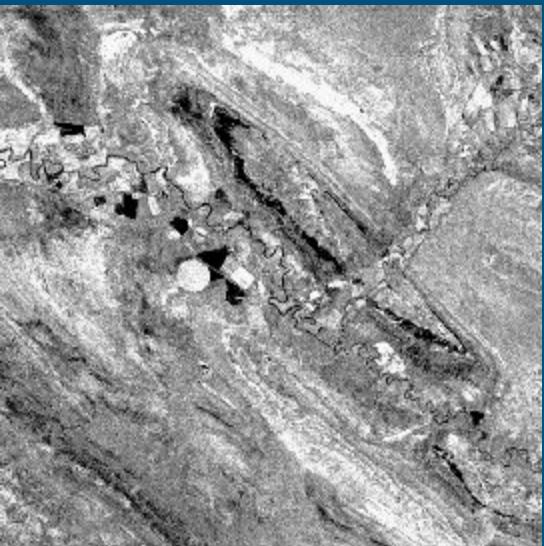
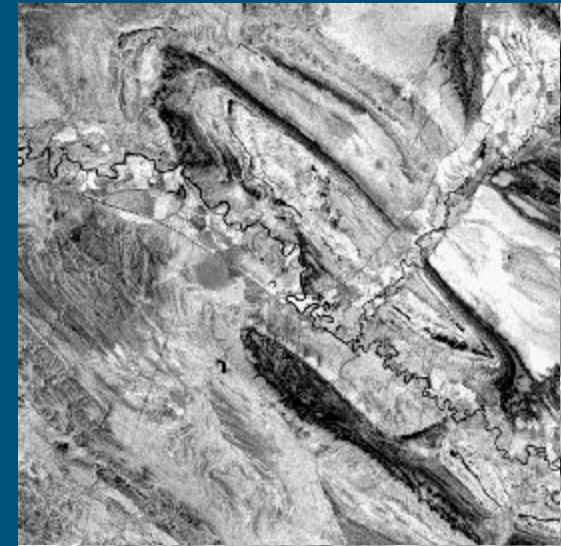
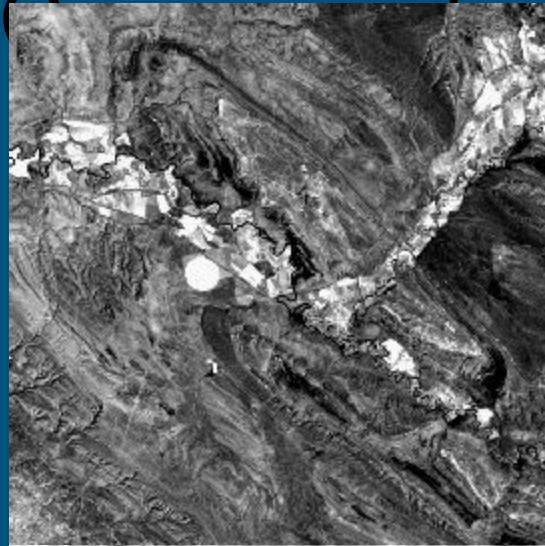
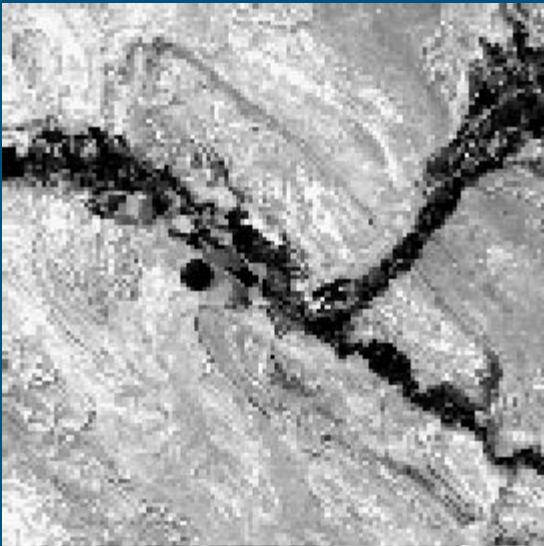


Composición *ad hoc*



Componentes principales

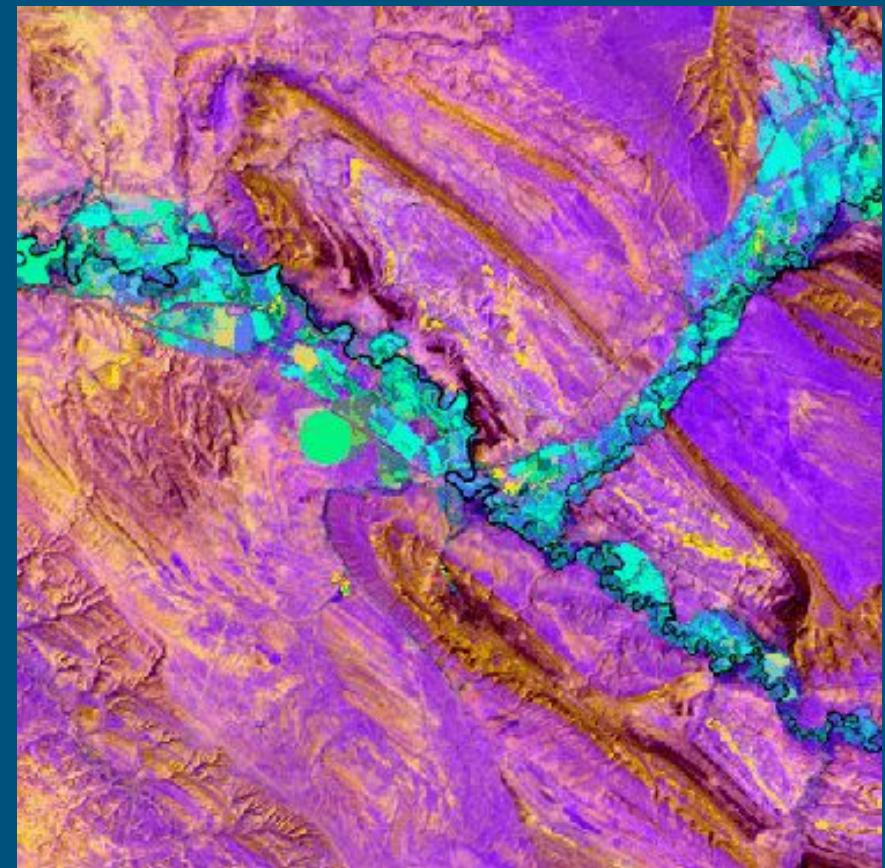
- Conjunto de transformaciones lineales
- Descomposición de la matriz de covariancia
- Ordenar por la variancia



A mano



Las dos composiciones



Estadística

EIGENFACES

Componentes Principales



- Base de imágenes
- Hallar las componentes principales de las imágenes
- Reducir la dimensionalidad del problema
- Los eigenfaces representan las variaciones de las imágenes
- Se calcula la distancia entre la imagen y su proyección al espacio de caras

Filtros pasa alto y
pasa bajo



Imagen Original



Realce de bordes



Suavizado

Realce utilizando combinación de filtros



a. Original

b. Laplaciano

$c = a + b$ (filtro pasa alto)

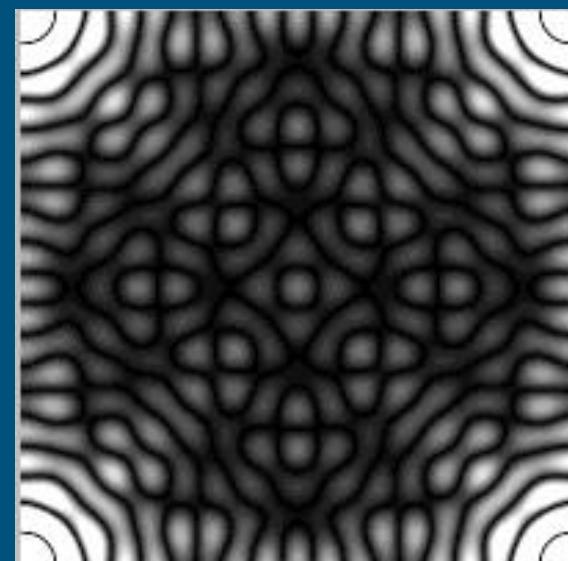
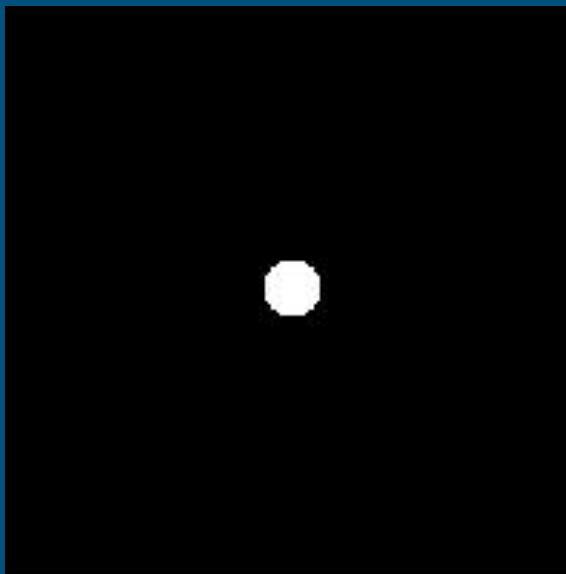
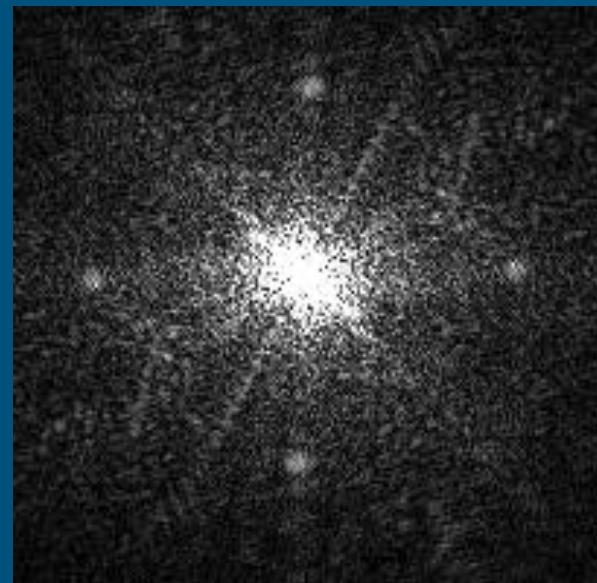


d. máscara

$e = b \times d$

$f = a + e$

Transformada de Fourier



Caracterización de la clase de texto basada en la transformada de Fourier



Ruido filtros no lineales



Imagen
original



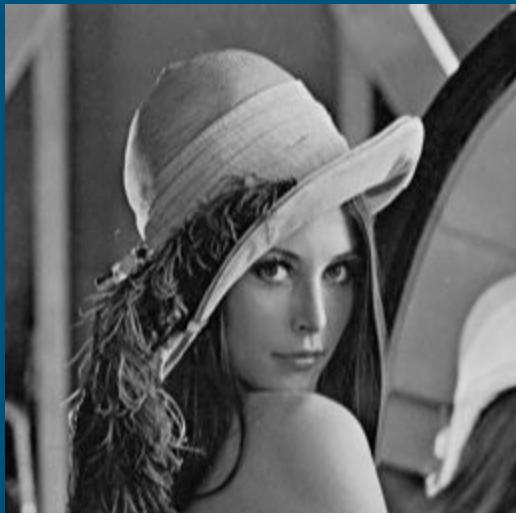
Imagen degradada
con ruido uniforme

Aplicación de filtro de la mediana



Imagen filtrada

Filtro stack entrenado con Lena

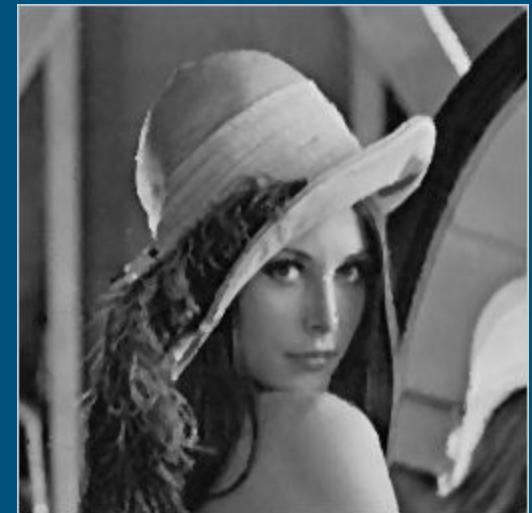


Lena sin ruido

MAE: 0.238922



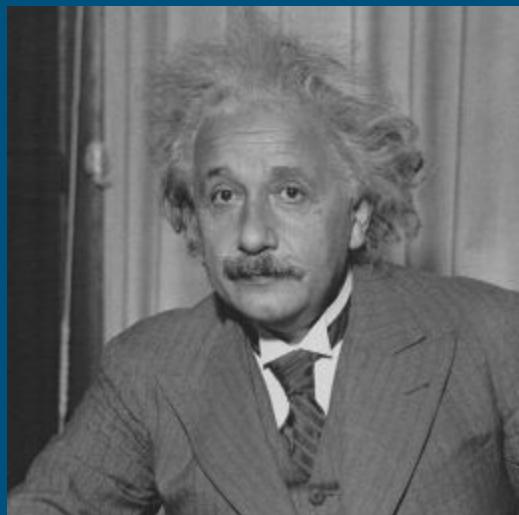
Lena con ruido
impulsivo



Lena filtrada

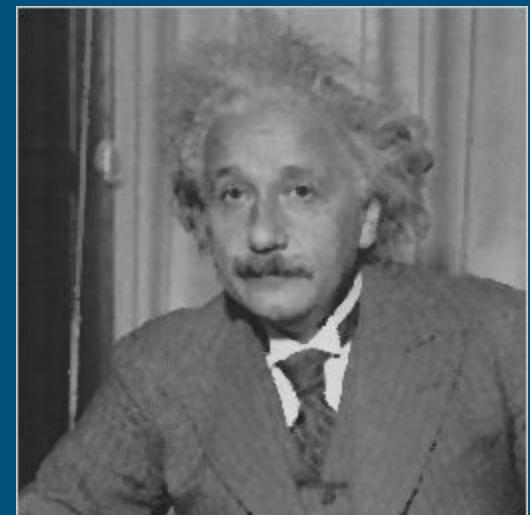
Filtro no lineal adaptativo

Filtro stack luego de entrenar con Lena



Albert sin ruido

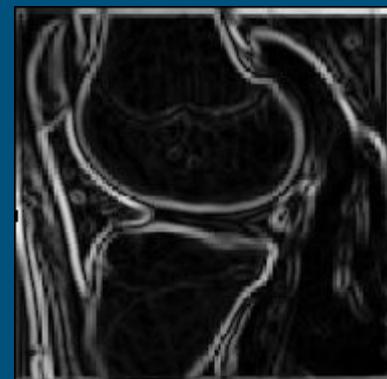
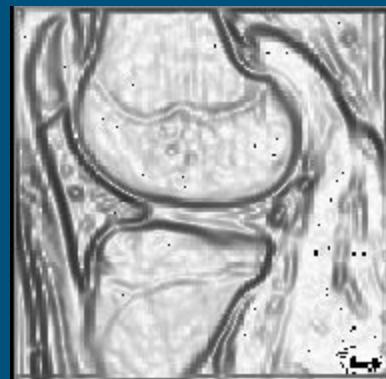
Albert con ruido
impulsivo



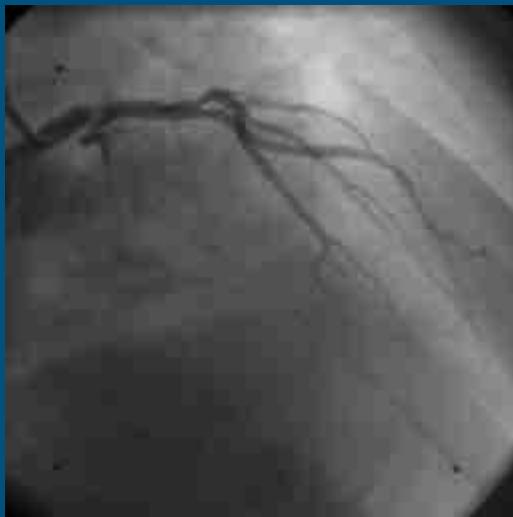
Albert filtrada

MAE 0.462448

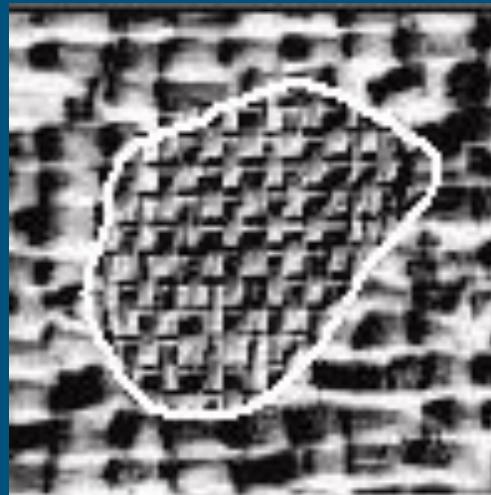
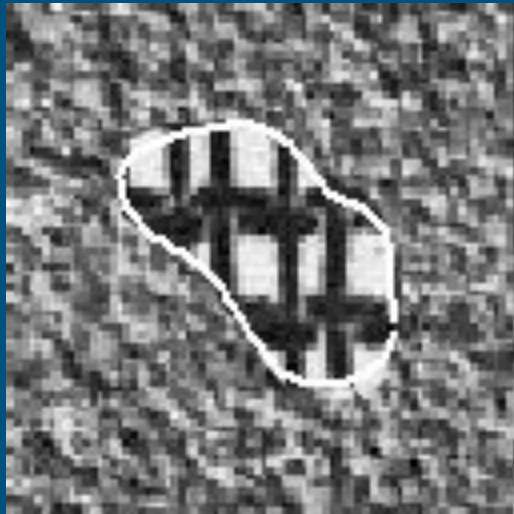
Ejemplos: Segmentación de Imágenes médicas.



Ejemplos: Detección de fallas en soldaduras



Ejemplos: Segmentación de Texturas

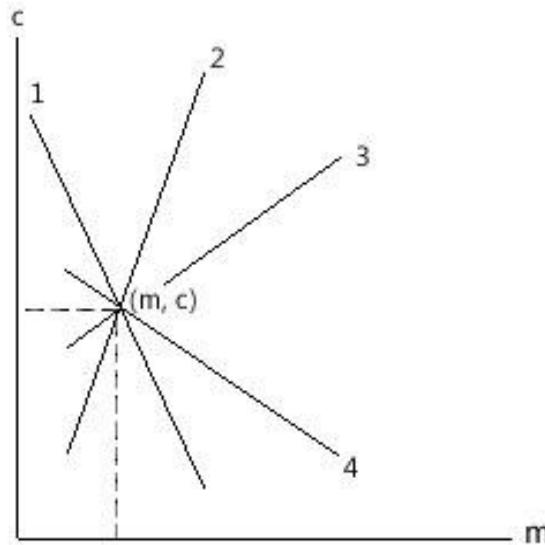
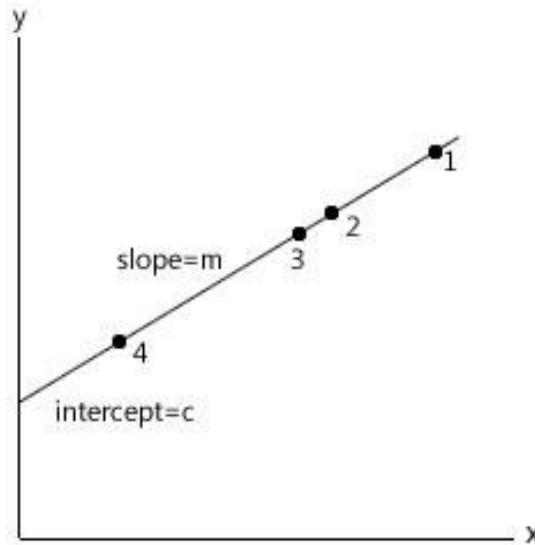


Transformada de Hough



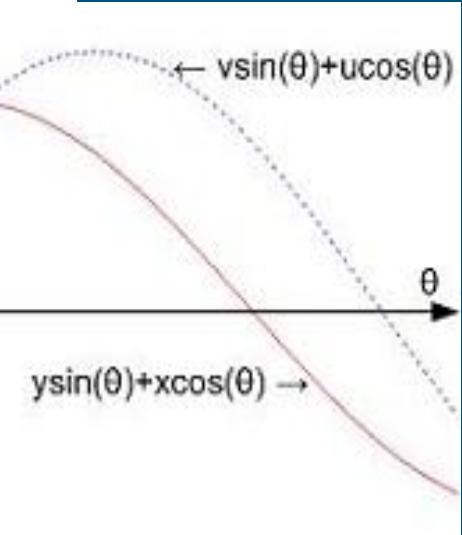
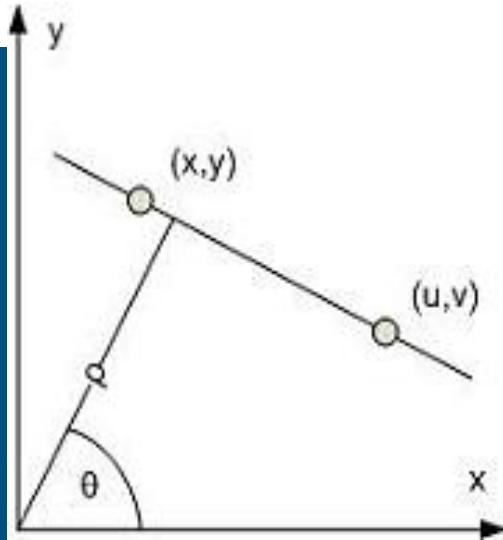
Encuentra líneas rectas

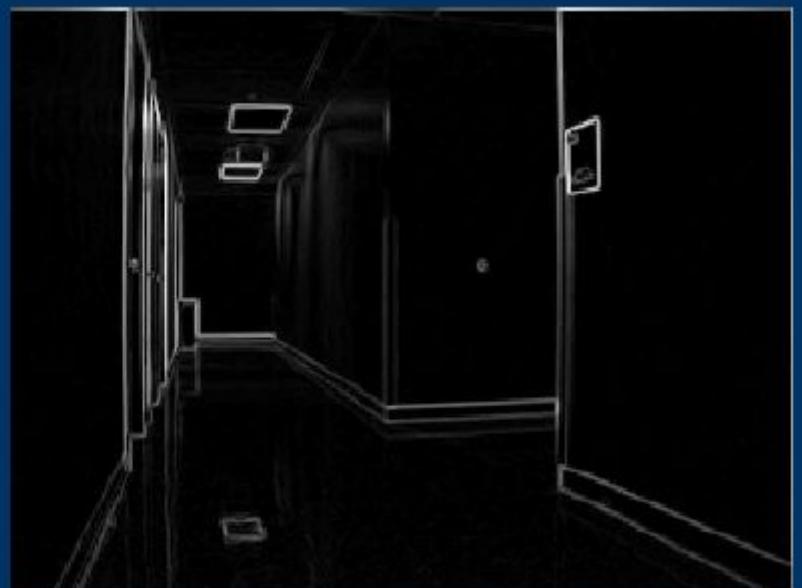
Transformada de Hough



xy space

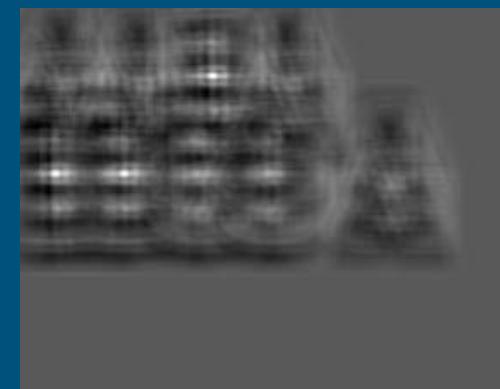
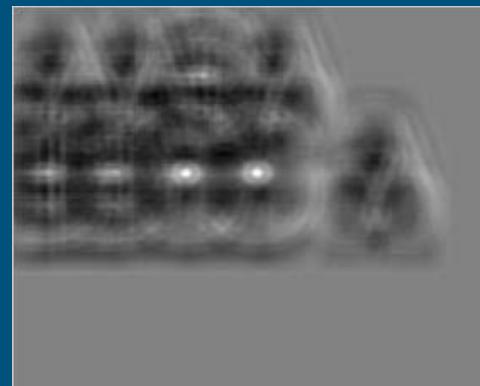
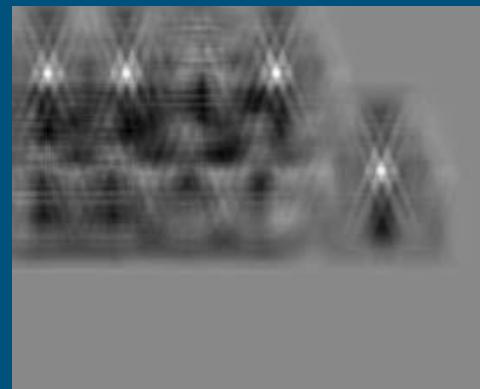
mc space



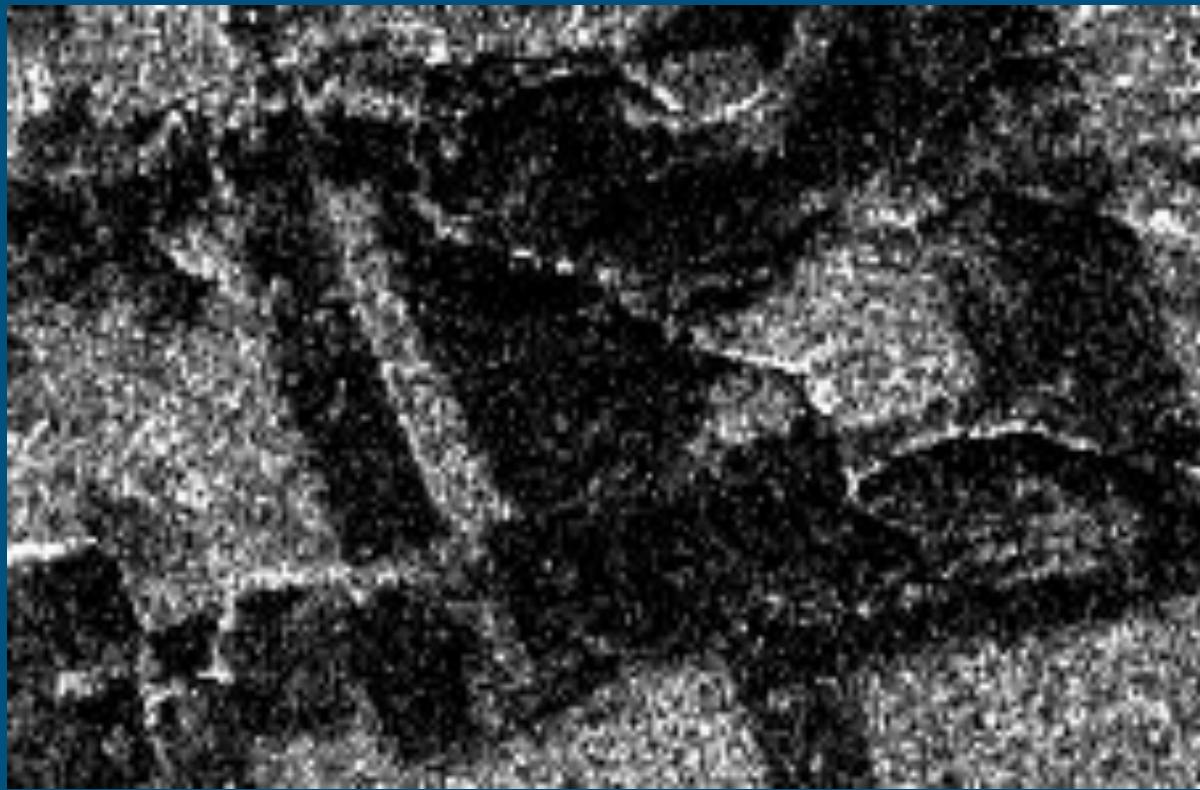


Transformada de Hough generalizada

reconoce un símbolo particular del alfabeto en aquellos lugares donde esté presente en la imagen



Clasificación



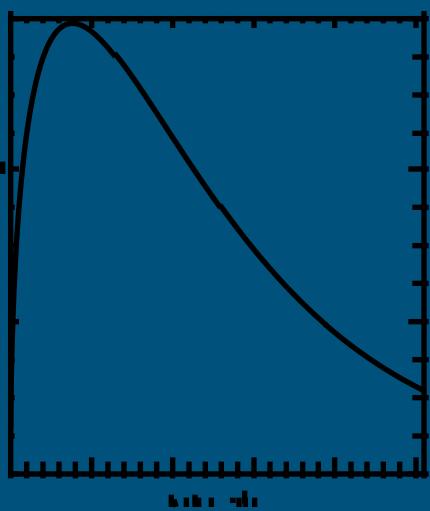
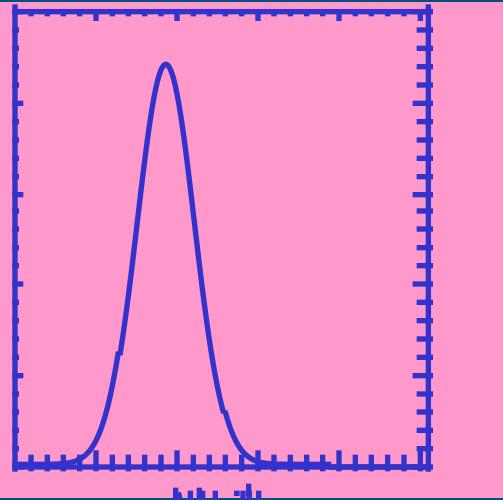
$$N(\mu_1, \sigma_1^2)$$

$$N(\mu_2, \sigma_2^2)$$

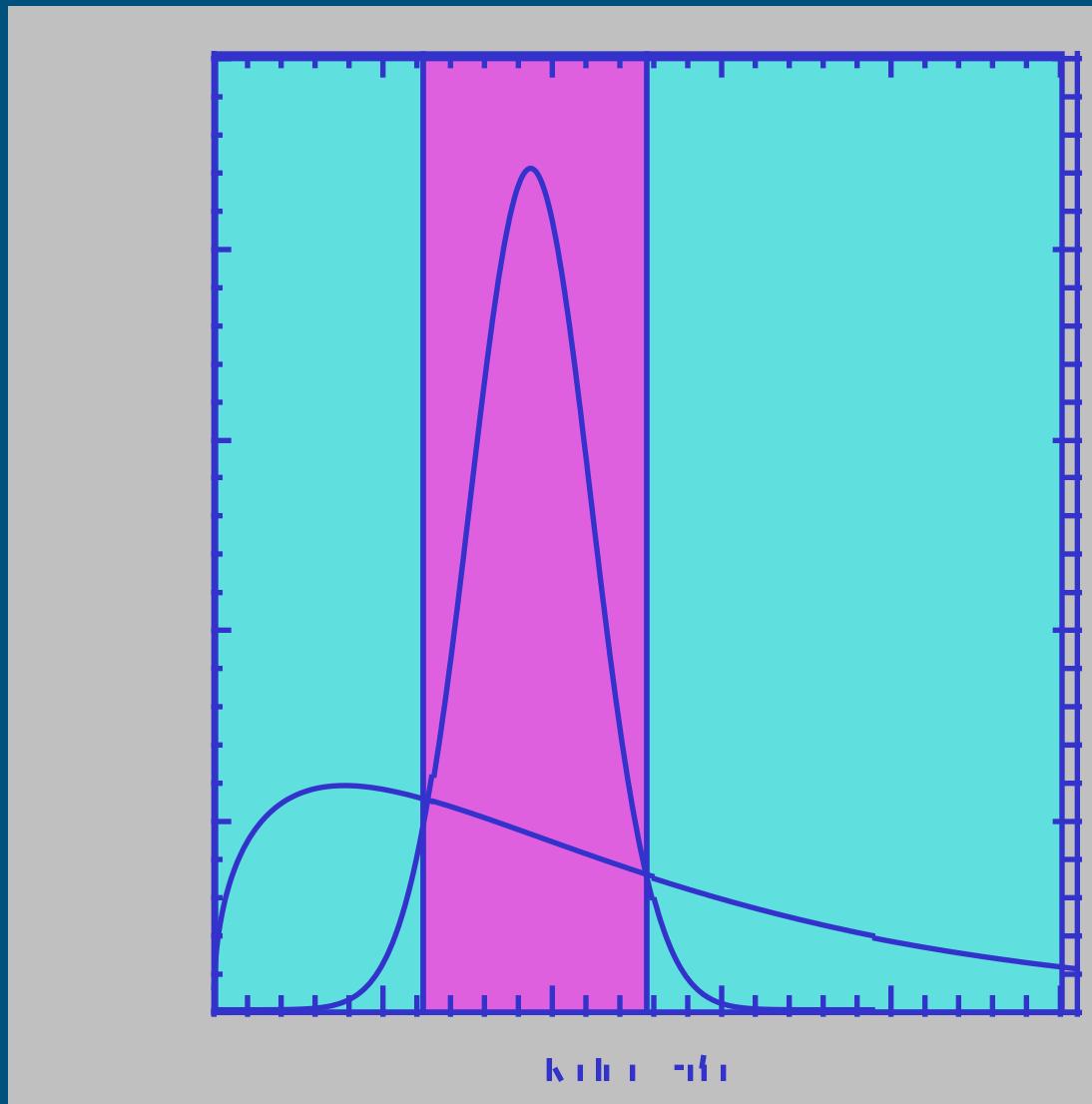
Imagen de valores de grises se transforman en mapa de clases

Entrenamiento

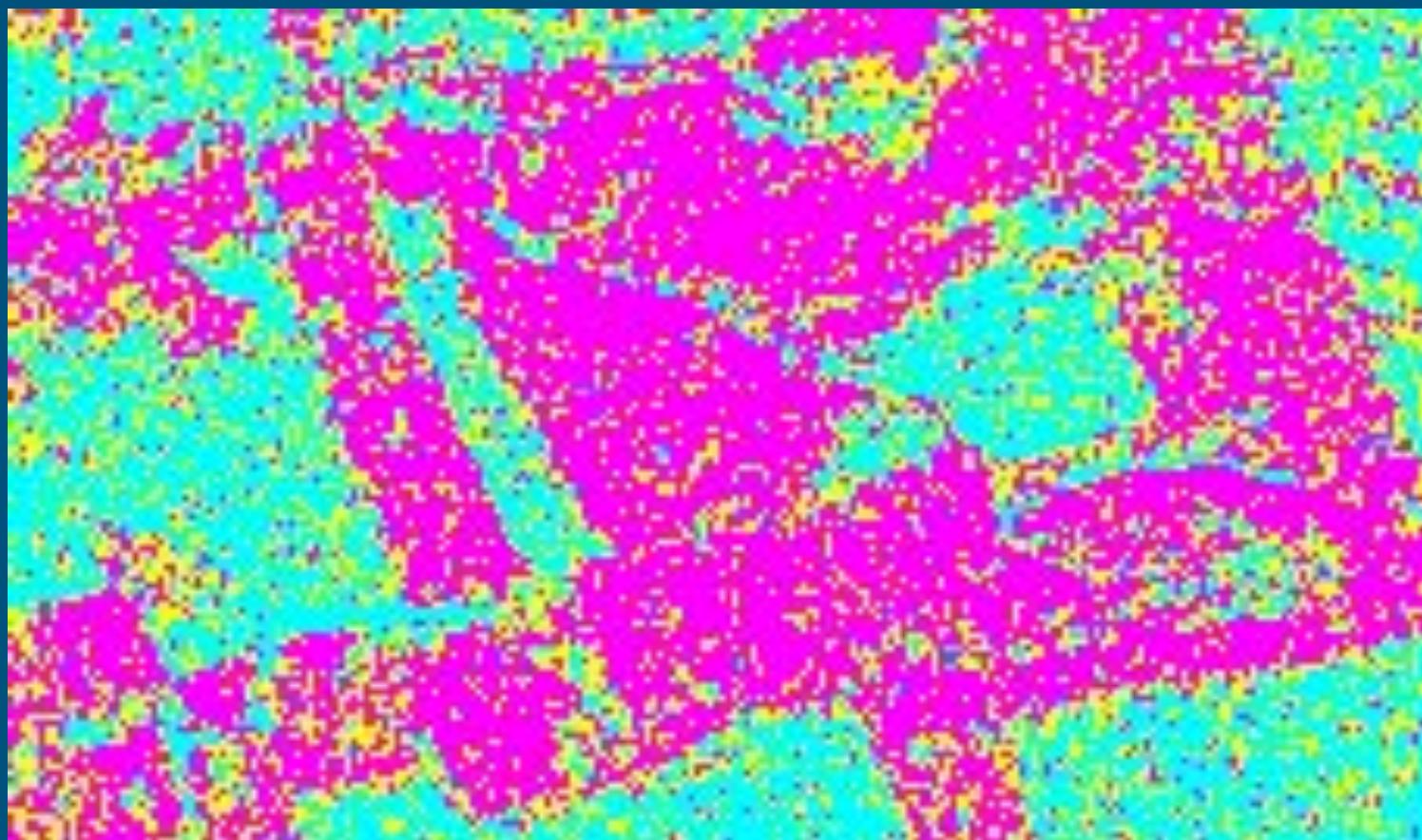
$$N(\mu_1, \sigma_1^2), \\ \hat{\mu}_1 = 94, 32, \hat{\sigma}_1^2 = 325, 87$$



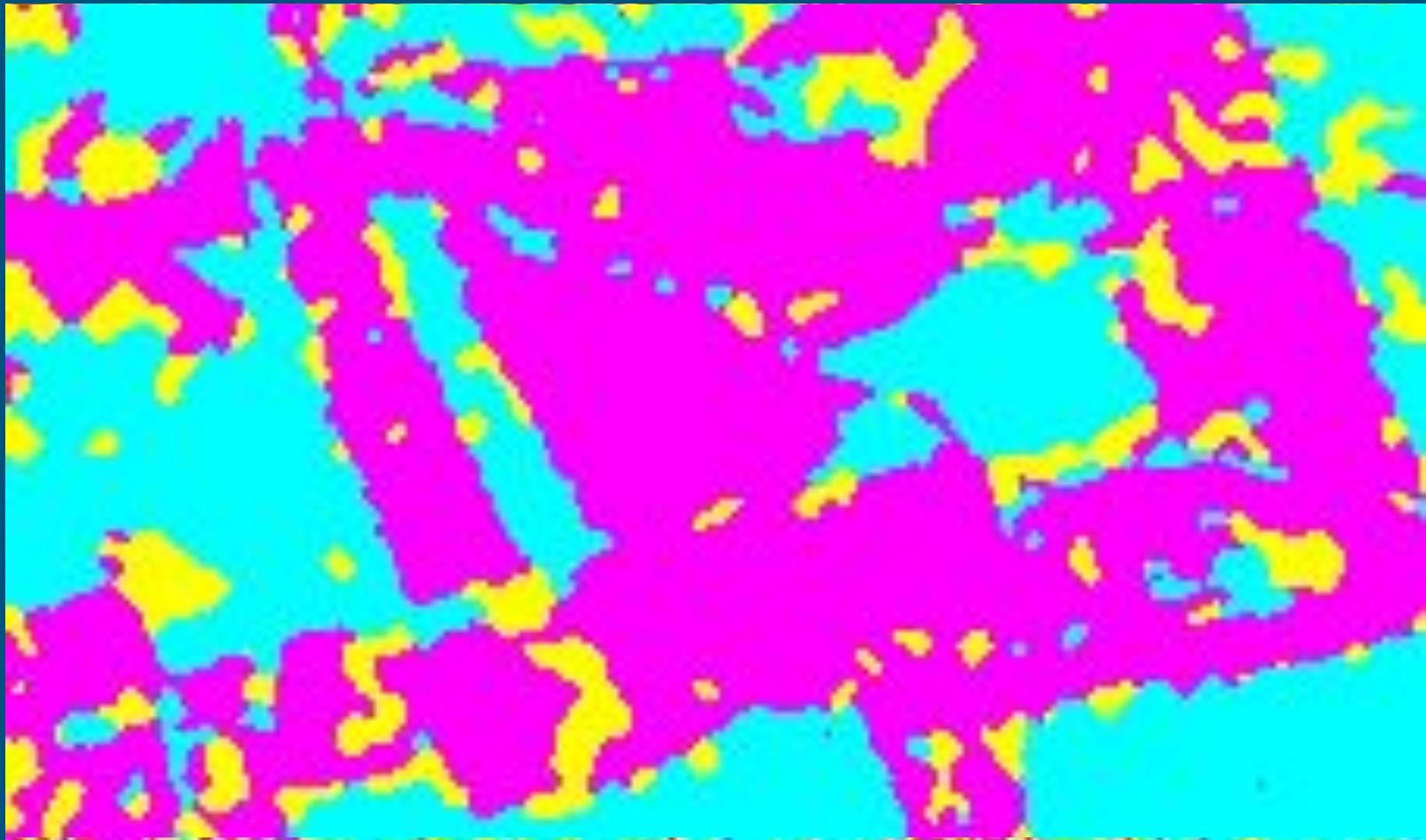
Clasificación



Resultado

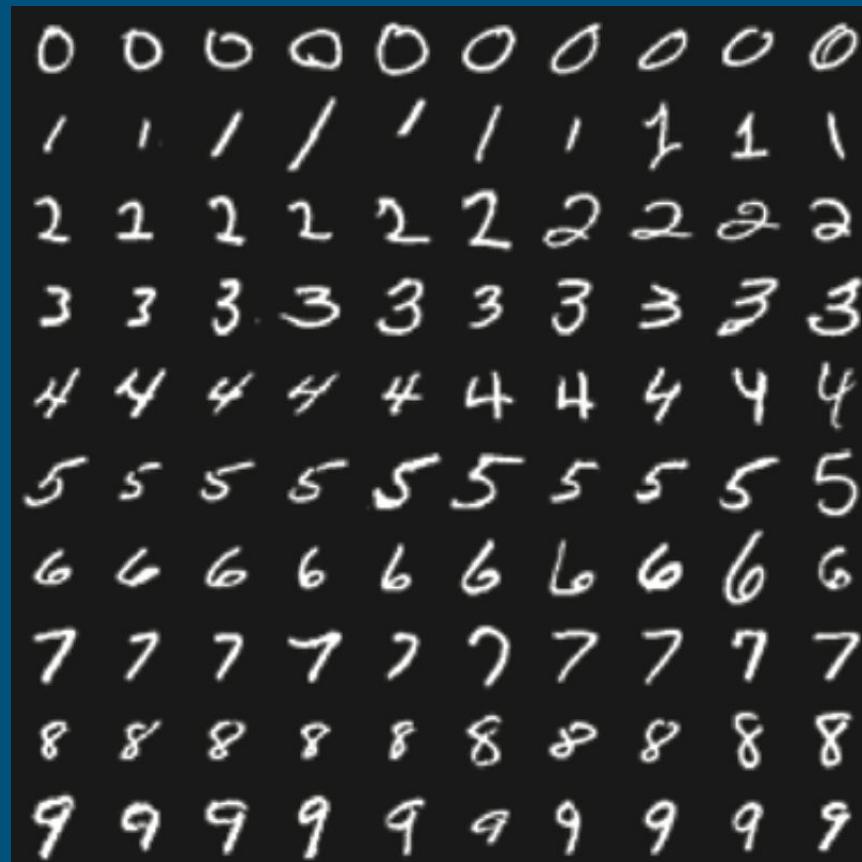


Mejorar la clasificación



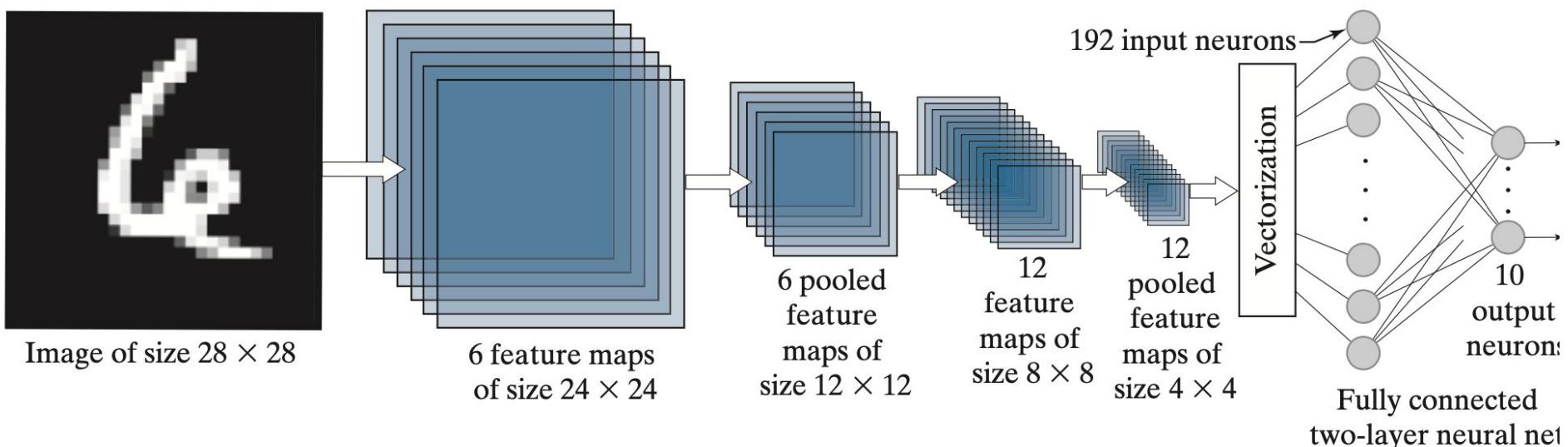
Redes convolucionales

Reconocimiento de dígitos manuscritos



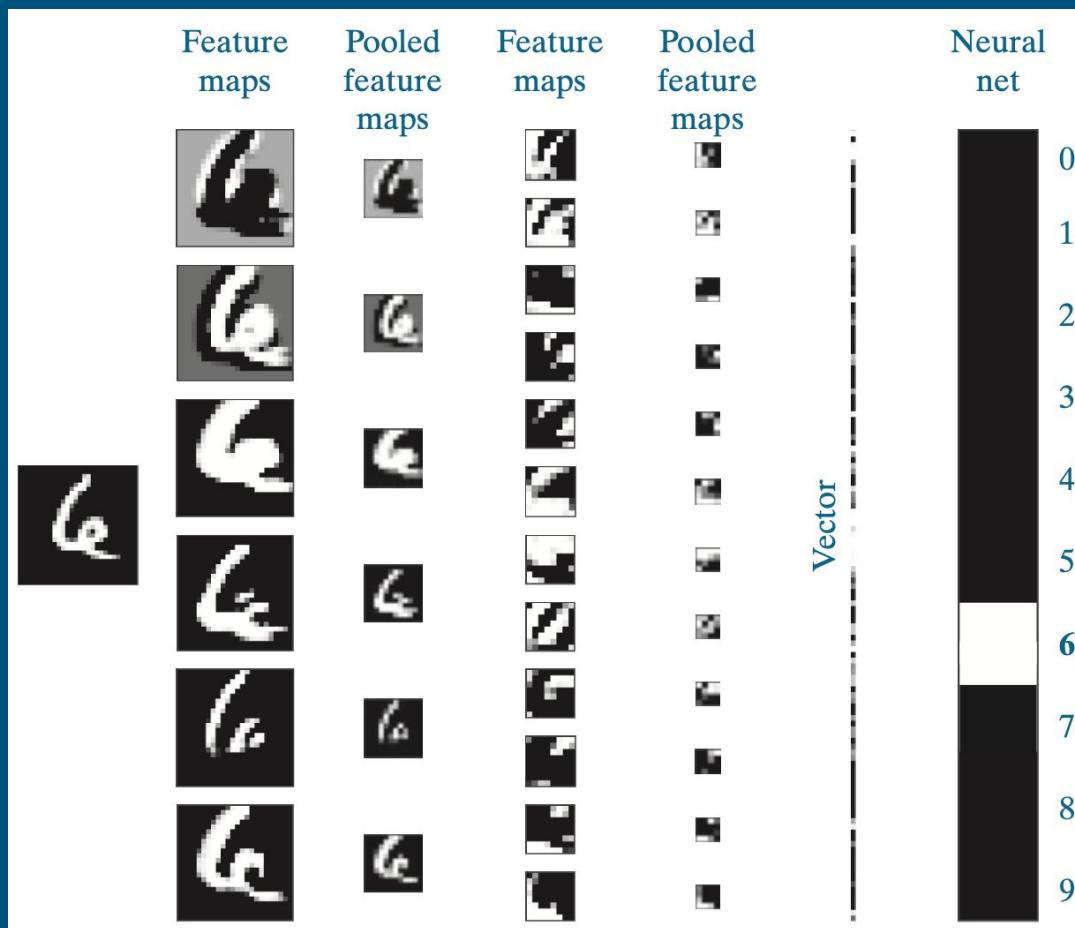
Redes convolucionales

Reconocimiento de dígitos manuscritos



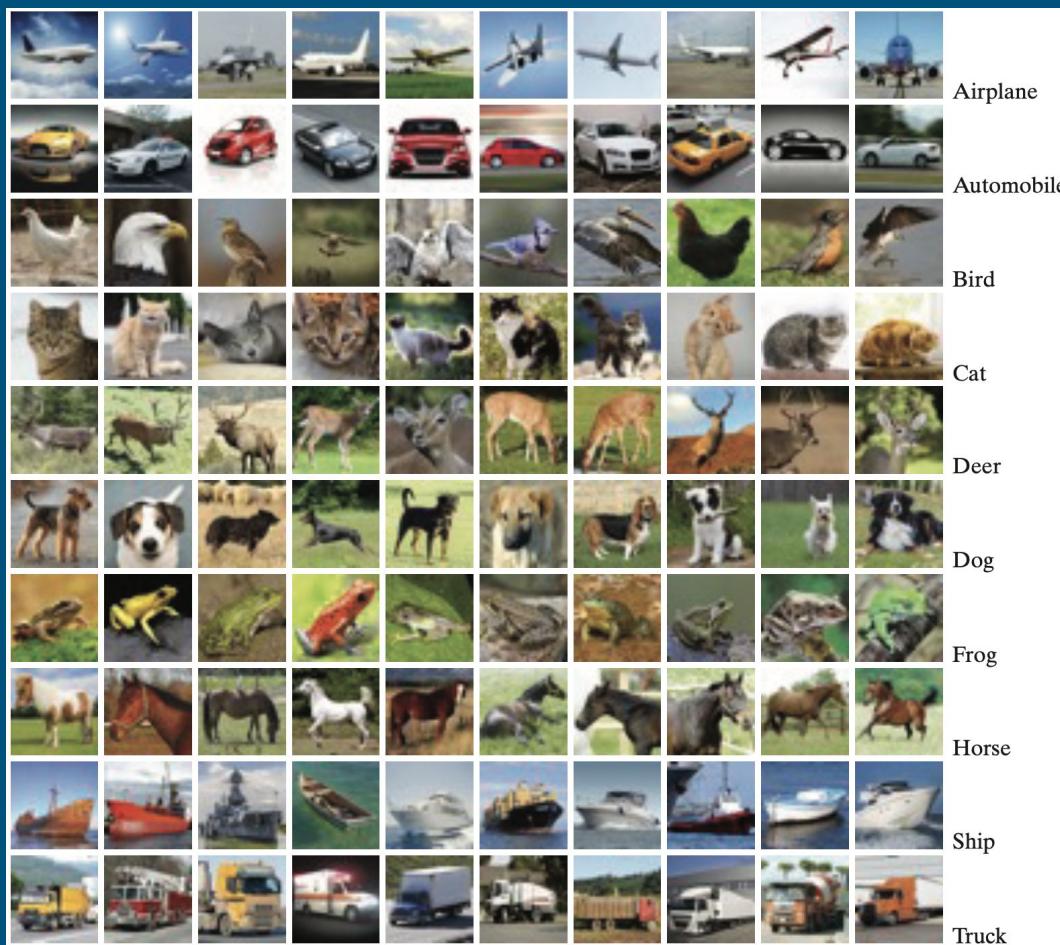
Redes convolucionales

Reconocimiento de dígitos manuscritos



Redes convolucionales

Reconocimiento de objetos



Qué necesito saber para procesar imágenes?

- Hacer Algoritmos
- Programar en cualquier lenguaje bien.