

## TEMA 3

### *- FUNDAMENTOS DE LA VISIÓN ARTIFICIAL -*

- 3.1.- Etapas en un proceso de visión por computador.
- 3.2.- Técnicas básicas de procesamiento de imágenes.
- 3.3.- Segmentación de objetos basada en histograma
- 3.4.- Introducción al problema de la visión dinámica
- 3.5.- Métodos de detección de movimiento

## TEMA 3

### – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL –

3.1.- Etapas en un proceso de visión por computador.

3.2.- Técnicas básicas de procesamiento de imágenes.

3.3.- Segmentación de objetos basada en histograma

3.4.- Introducción al problema de la visión dinámica

3.5.- Métodos de detección de movimiento

## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.1.- Etapas en un Proceso de Visión por Computador

#### ➤ Visión por computador (visión artificial): *definición*

- ⇒ Ciencia que estudia la interpretación de imágenes mediante computadores digitales.
- ⇒ Proceso de extracción de información del mundo físico a partir de imágenes utilizando para ello un computador.
- ⇒ Sistema autónomo que realiza algunas tareas que realiza el sistema de visión humano:
  - Información o tareas que este sistema de visión puede llegar a extraer o realizar: detección de objetos, interpretación de escenas.

### CAMPOS DE APLICACIÓN

- |   |   |
|---|---|
| <ul style="list-style-type: none"><li>• Medio-ambiente<ul style="list-style-type: none"><li>– Detección y medición de incendios forestales</li><li>– Detección y medición de vertidos</li></ul></li><li>• Medicina<ul style="list-style-type: none"><li>– Evaluación y diagnóstico automático</li><li>– Análisis médicos</li></ul></li><li>• Biología<ul style="list-style-type: none"><li>– Detección y seguimiento de especies</li><li>– Análisis de la biomasa</li></ul></li></ul> | <ul style="list-style-type: none"><li>• Industria<ul style="list-style-type: none"><li>– Control de calidad</li><li>– Inspección</li><li>– Medición</li><li>– Automatización de procesos</li></ul></li><li>• Robótica<ul style="list-style-type: none"><li>– Guiado</li><li>– Detección y evitación de obstáculos</li><li>– Generación de mapas</li></ul></li><li>• Seguridad<ul style="list-style-type: none"><li>– Detección</li><li>– Monitorización</li><li>– Seguimiento</li><li>– Inspección y medida en ambientes peligrosos</li></ul></li></ul> |
|---|---|

## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.1.- Etapas en un Proceso de Visión por Computador

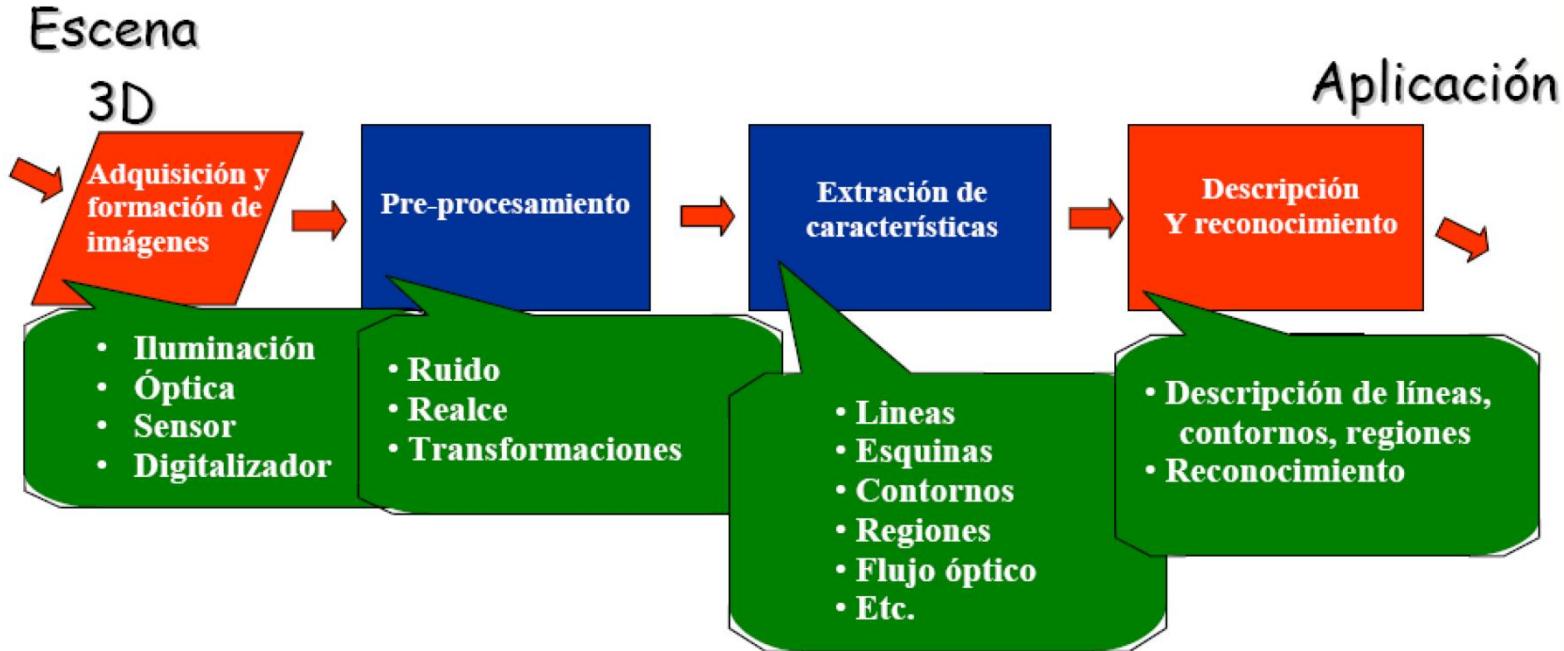
➤ Distintas etapas que suelen considerarse en un proceso de Visión:



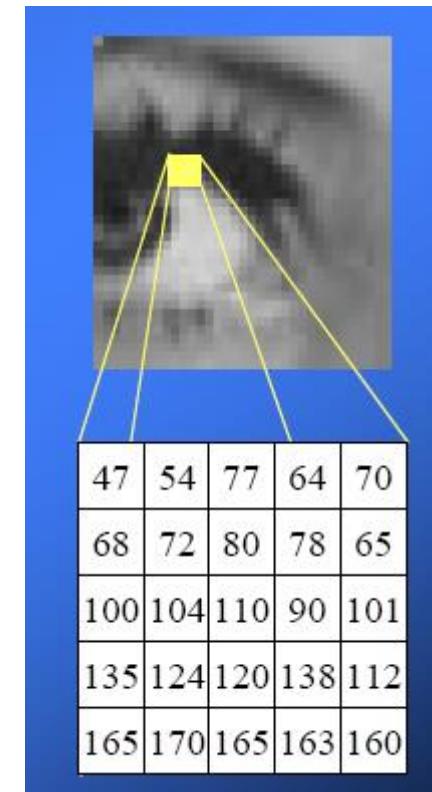
## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.1.- Etapas en un Proceso de Visión por Computador

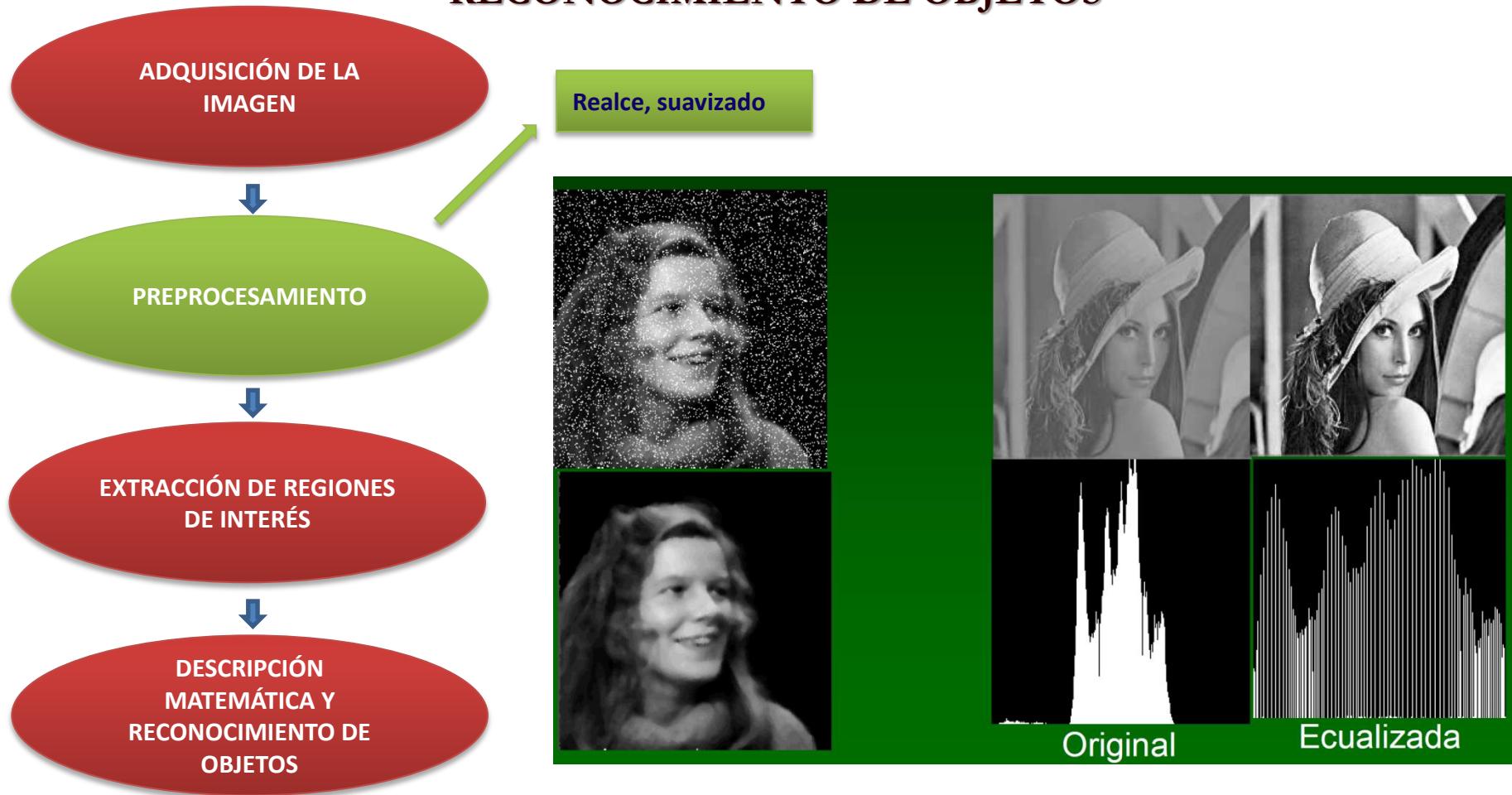
➤ Distintas etapas que suelen considerarse en un proceso de Visión:



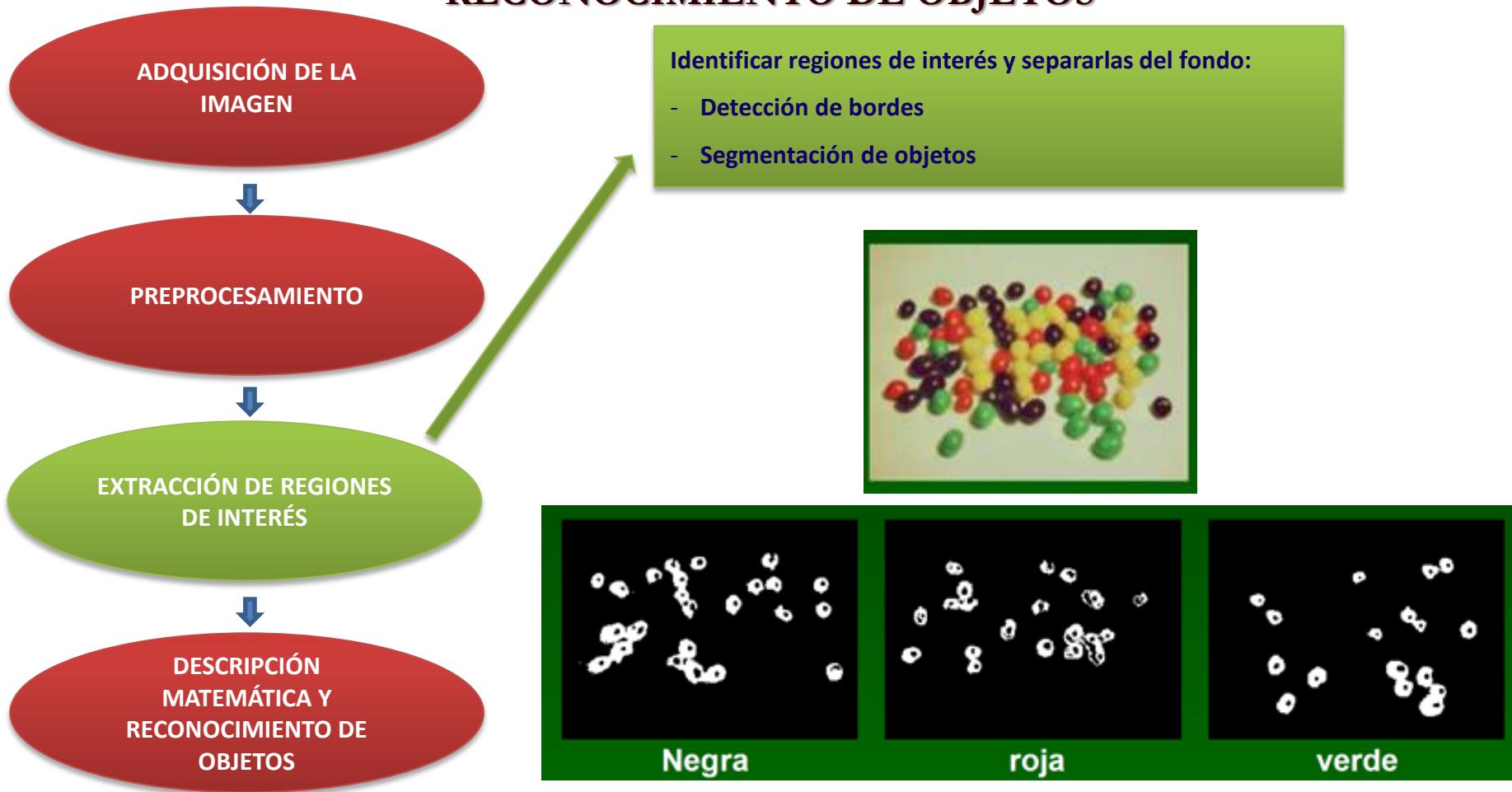
# ETAPAS METODOLÓGICAS GENERALES DE UN PROCESO DE VISIÓN POR COMPUTADOR QUE TIENE COMO OBJETIVO EL RECONOCIMIENTO DE OBJETOS



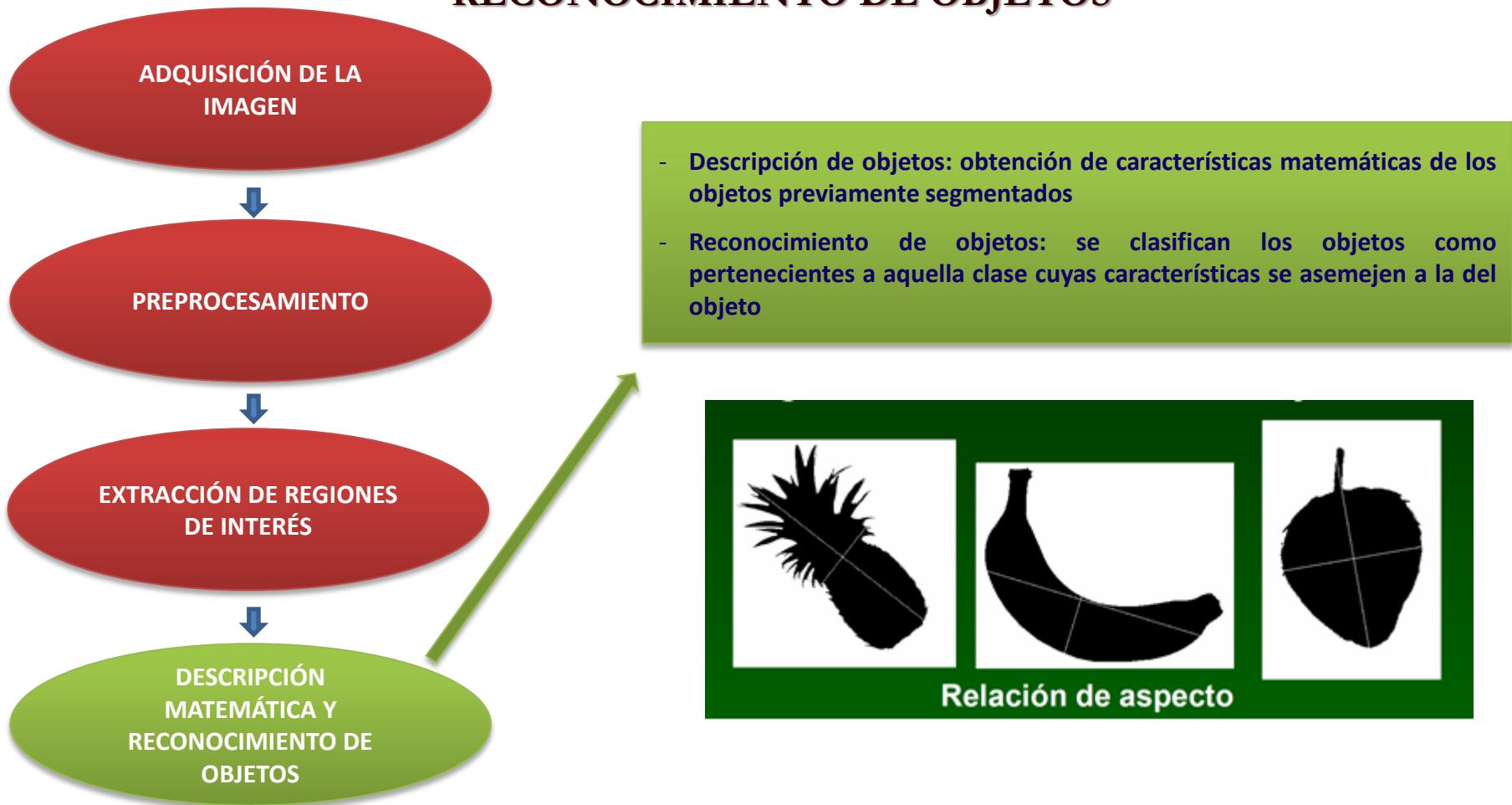
## **ETAPAS METODOLÓGICAS GENERALES DE UN PROCESO DE VISIÓN POR COMPUTADOR QUE TIENE COMO OBJETIVO EL RECONOCIMIENTO DE OBJETOS**



## ETAPAS METODOLÓGICAS GENERALES DE UN PROCESO DE VISIÓN POR COMPUTADOR QUE TIENE COMO OBJETIVO EL RECONOCIMIENTO DE OBJETOS



## ETAPAS METODOLÓGICAS GENERALES DE UN PROCESO DE VISIÓN POR COMPUTADOR QUE TIENE COMO OBJETIVO EL RECONOCIMIENTO DE OBJETOS



# TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

## 3.1.- Etapas en un Proceso de Visión por Computador

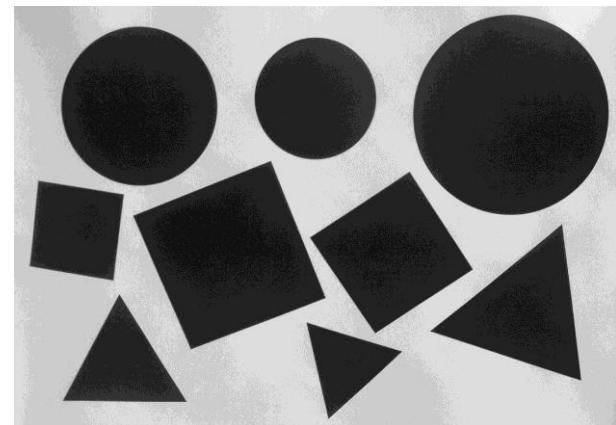
### EJEMPLO

Reconocimiento de objetos en una imagen como perteneciente a alguna de las siguientes clases u objetos: triángulo equilátero, círculo y cuadrado.



Planteamiento:

- ❖ Descriptor matemático:
  - Compactidad (relación perímetro<sup>2</sup> y área).



Análisis teórico de la idoneidad del descriptor

	Área	Perímetro	Perímetro <sup>2</sup> / Área
O	$\pi r^2$	$2\pi r$	<b>12.56</b>
□	$l^2$	$4l$	<b>16</b>
Δ	$\sqrt{3}l^2/4$	$3l$	<b>20.8</b>



❖ Reconocimiento:

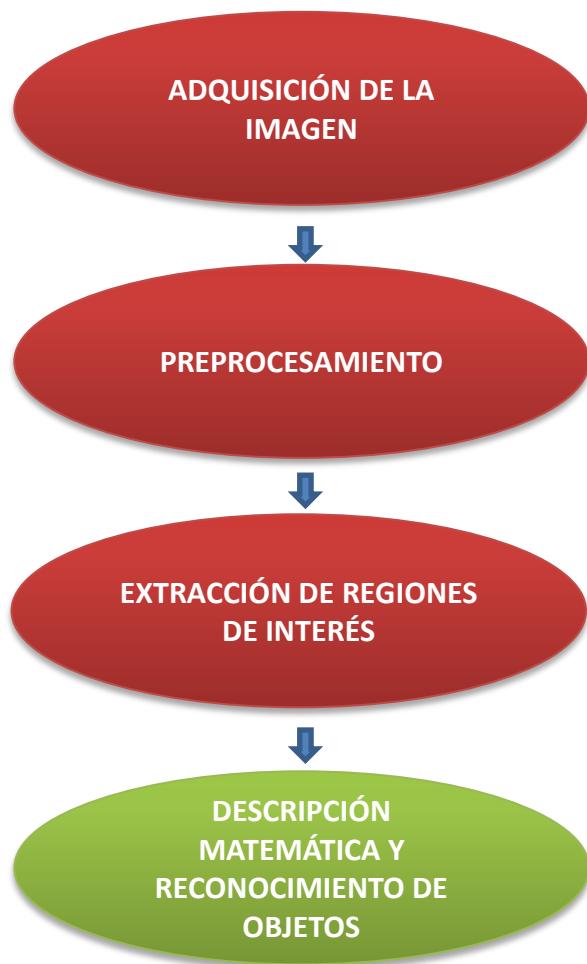
- Regla de decisión
  - Si per.<sup>2</sup> / área < 14.3 entonces objeto = círculo.
  - Si 14.3 < per.<sup>2</sup> / área < 18.4 entonces objeto = cuadrado
  - Si 18.4 < per.<sup>2</sup> / área entonces objeto = triángulo

# TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

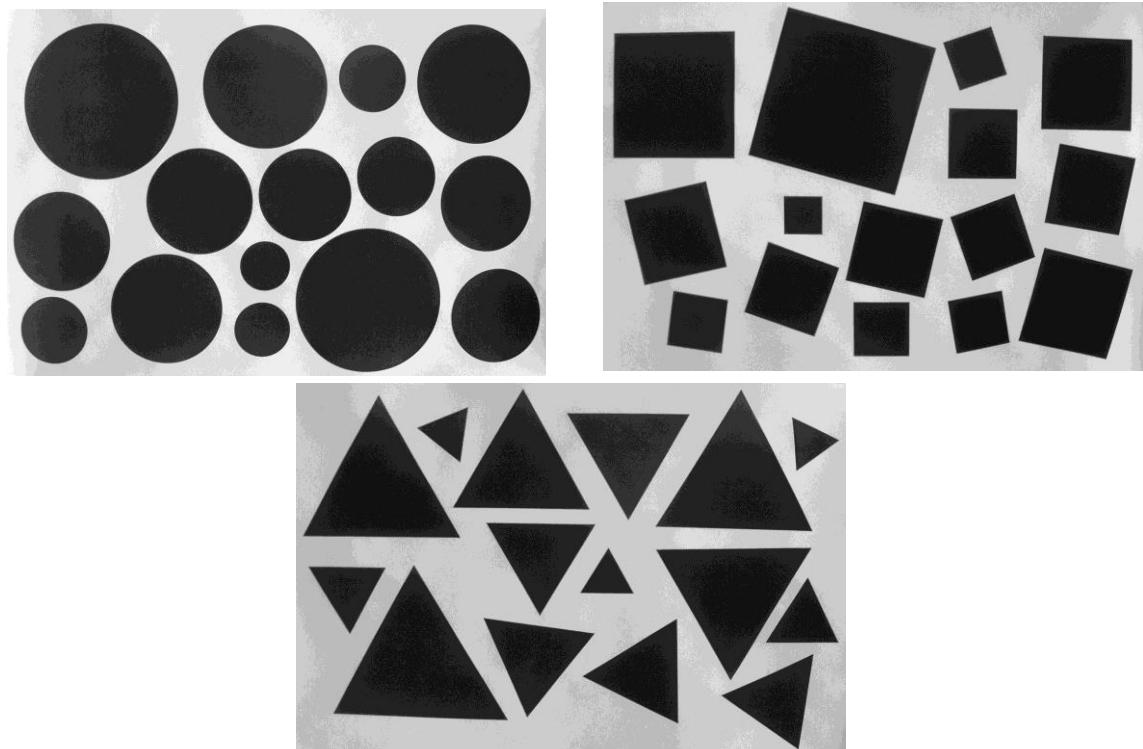
## 3.1.- Etapas en un Proceso de Visión por Computador

### EJEMPLO

Reconocimiento de objetos en una imagen como perteneciente a alguna de las siguientes clases u objetos:  
triángulo equilátero, círculo y cuadrado.



- SELECCIÓN DE CARÁCTERISTICAS Y DISEÑO DEL CLASIFICADOR
  - ❖ IMÁGENES DE ENTRENAMIENTO
  - ❖ EXTRACCIÓN DE PATRONES DE ENTRENAMIENTO  
(representación matemática de objetos de clase conocida)

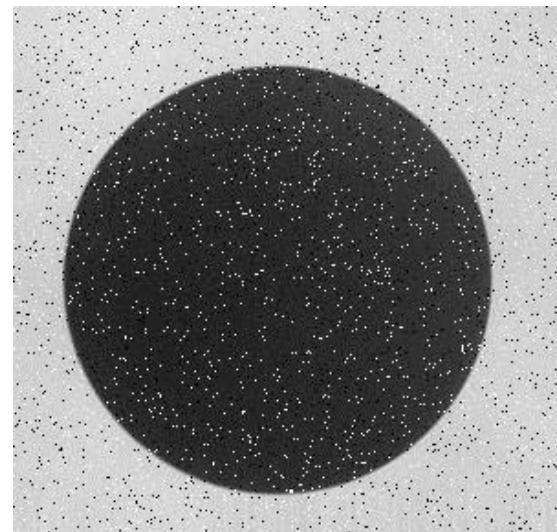
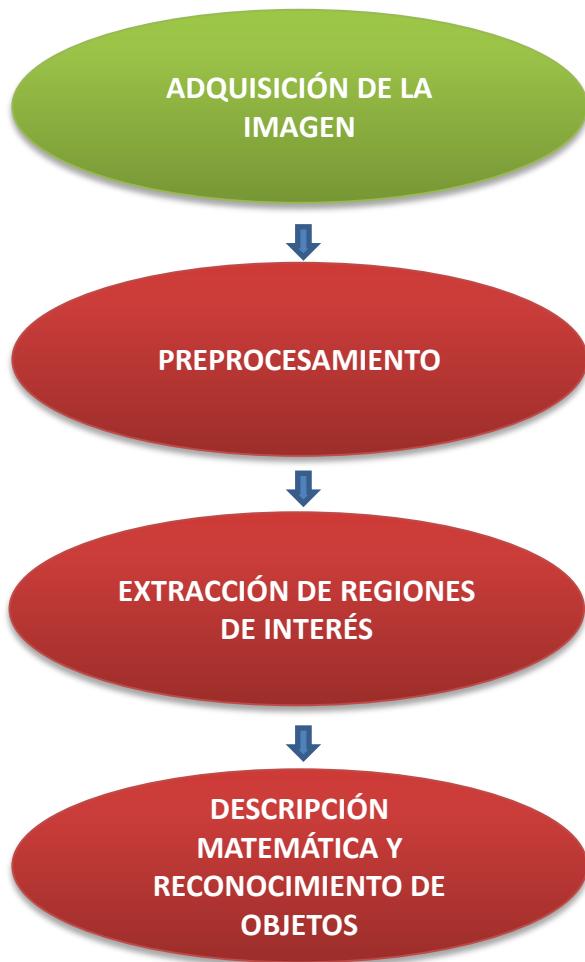


## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.1.- Etapas en un Proceso de Visión por Computador

#### EJEMPLO

Reconocimiento de objetos en una imagen como perteneciente a alguna de las siguientes clases u objetos:  
triángulo equilátero, círculo y cuadrado.

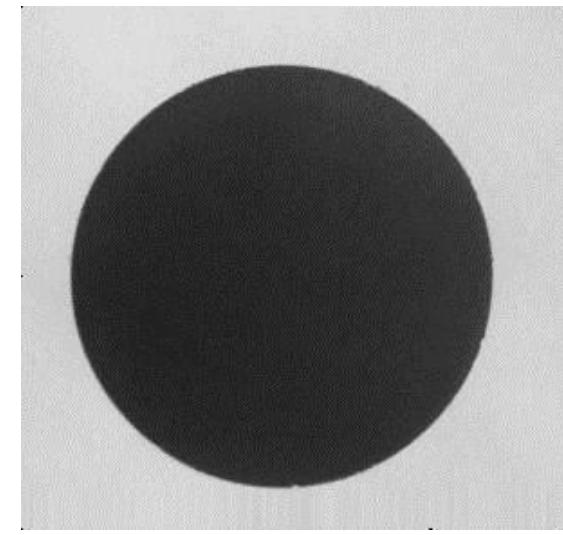
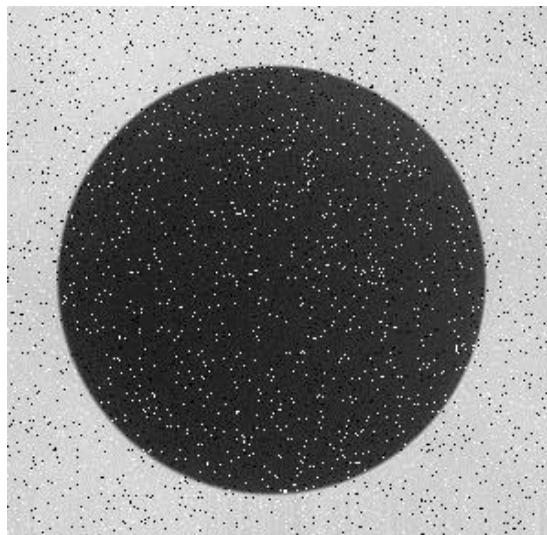


## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.1.- Etapas en un Proceso de Visión por Computador

#### EJEMPLO

Reconocimiento de objetos en una imagen como perteneciente a alguna de las siguientes clases u objetos:  
triángulo equilátero, círculo y cuadrado.

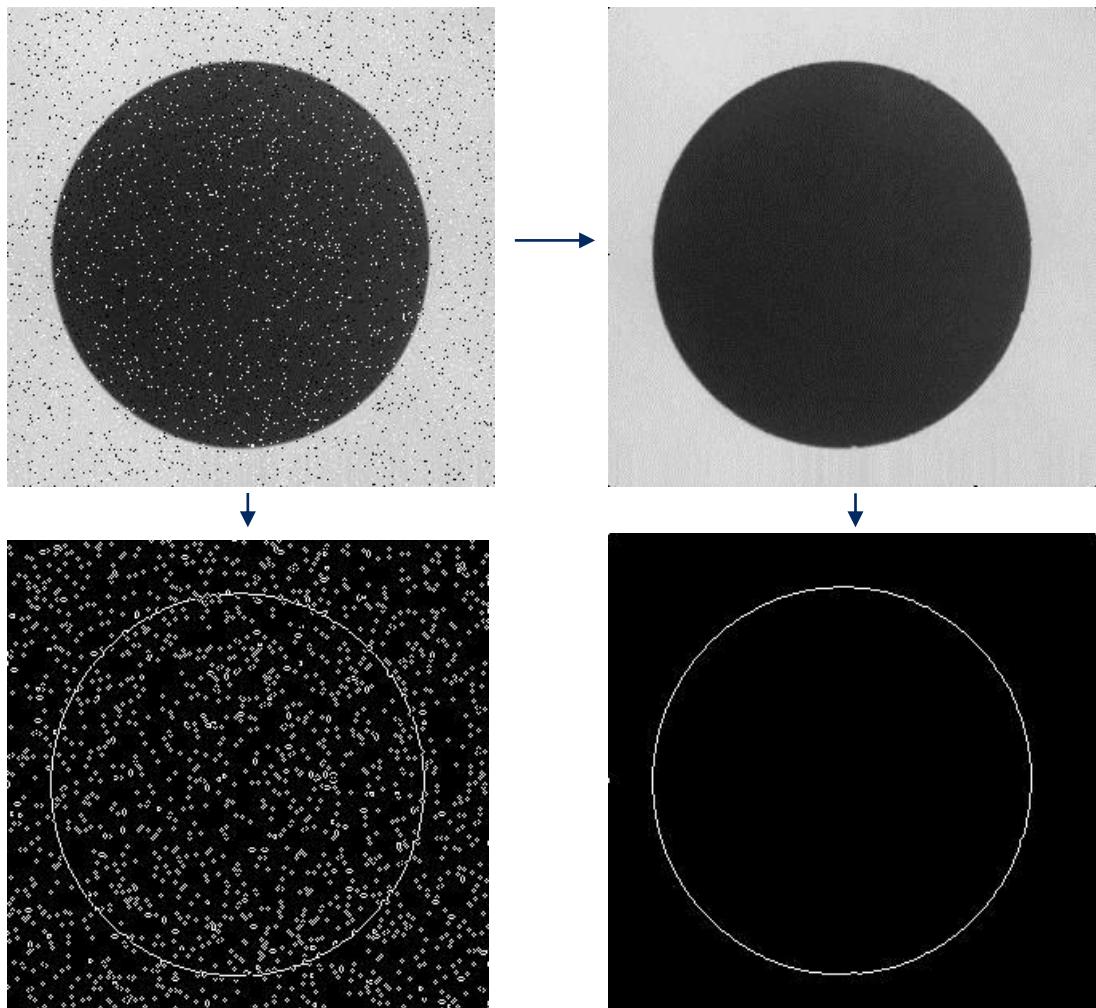


# TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

## 3.1.- Etapas en un Proceso de Visión por Computador

### EJEMPLO

Reconocimiento de objetos en una imagen como perteneciente a alguna de las siguientes clases u objetos:  
triángulo equilátero, círculo y cuadrado.

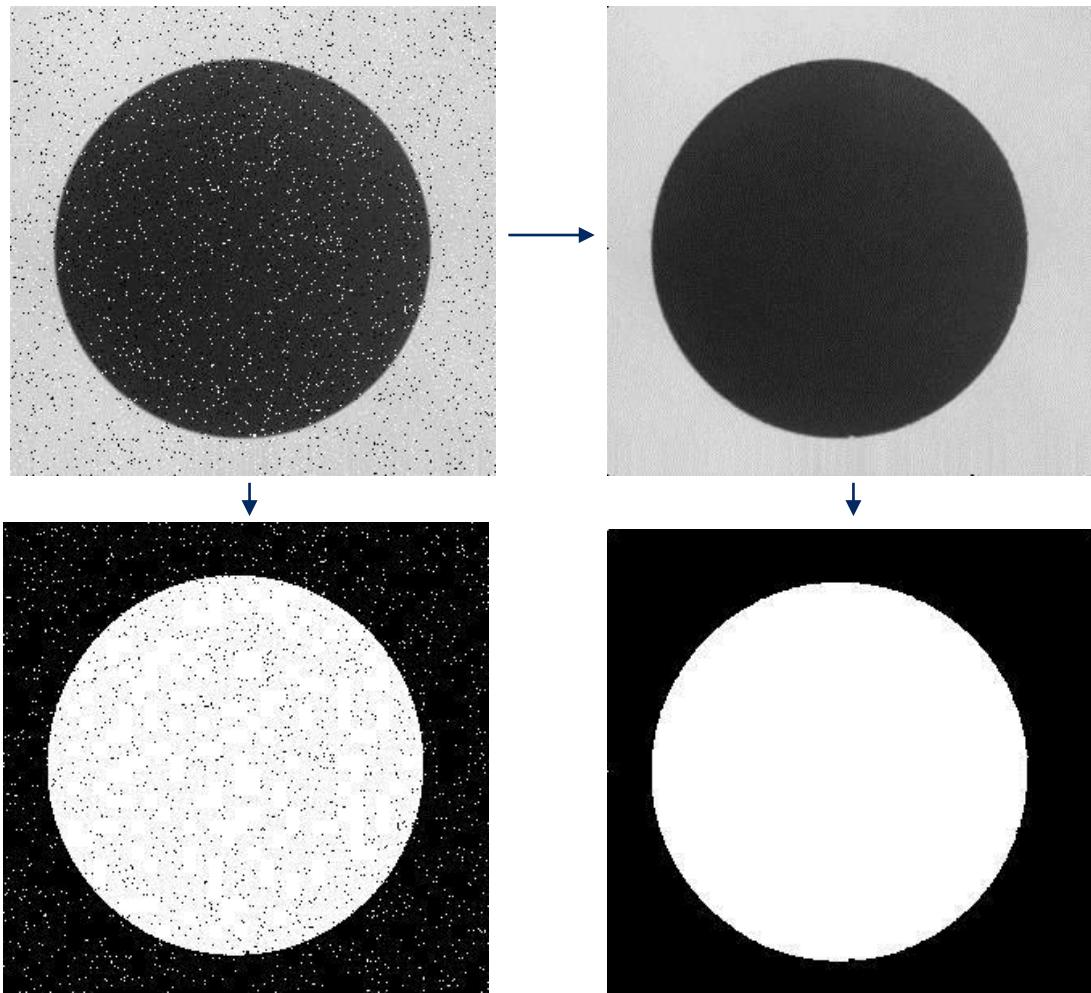


# TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

## 3.1.- Etapas en un Proceso de Visión por Computador

### EJEMPLO

Reconocimiento de objetos en una imagen como perteneciente a alguna de las siguientes clases u objetos:  
triángulo equilátero, círculo y cuadrado.



# TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

## 3.1.- Etapas en un Proceso de Visión por Computador

### EJEMPLO

Reconocimiento de objetos en una imagen como perteneciente a alguna de las siguientes clases u objetos: triángulo equilátero, círculo y cuadrado.



1. Calcular perímetro a partir del borde previamente detectado del objeto.
2. Calcular área a partir de la segmentación previa del objeto.
3. Calcular descriptor:  $\text{perímetro}^2 / \text{área}$ ,
4. Aplicar el clasificador diseñado en base al conocimiento a priori del problema.

❖ Ejemplo anterior:

➤ Regla de decisión

- Si  $\text{per.}^2 / \text{área} < 14.3$  entonces objeto = círculo.
- Si  $14.3 < \text{per.}^2 / \text{área} < 18.4$  entonces objeto = cuadrado
- Si  $18.4 < \text{per.}^2 / \text{área}$  entonces objeto = triángulo

## TEMA 3

### *– FUNDAMENTOS DE LA VISIÓN ARTIFICIAL –*

3.1.- Etapas en un proceso de visión por computador.

3.2.- Técnicas básicas de procesamiento de imágenes.

3.3.- Segmentación de objetos basada en histograma

3.4.- Introducción al problema de la visión dinámica

3.5.- Métodos de detección de movimiento

## TEMA 3

### – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL –

3.1.- Etapas en un proceso de visión por computador.

3.2.- Técnicas básicas de procesamiento de imágenes:

- Procesamiento de imágenes en el dominio espacial
- Transformaciones lógicas
- Transformaciones geométricas

3.3.- Segmentación de objetos basada en histograma

3.4.- Introducción al problema de la visión dinámica

3.5.- Métodos de detección de movimiento

## **TÉCNICAS BÁSICAS DE PROCESAMIENTO BÁSICO DE IMÁGENES**

- **Procesamiento de imágenes:** transformación de una imagen en otra, mediante la aplicación de una función u operación sobre los valores de niveles de gris de los píxeles.
- **2 enfoques:**

### **1. TRANSFORMACIONES EN EL DOMINIO FRECUENCIAL:**

**Operaciones en el dominio de la frecuencia** (se aborda como una extensión del procesamiento digital de señales e implica una transformación de la imagen al dominio de la frecuencia).

### **2. TRANSFORMACIONES EN EL DOMINIO ESPACIAL:**

**Operaciones aritméticas** sobre los elementos de la matriz de muestras discretas que representa a la imagen digital.

## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

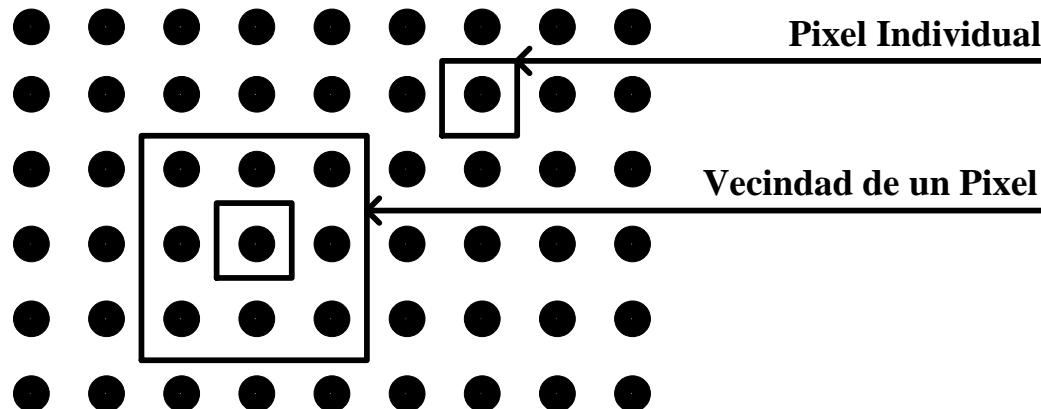
### 3.2.- Técnicas Básicas de Procesamiento de Imágenes

#### PROCESAMIENTO BÁSICO DE IMÁGENES: Transformaciones en el dominio espacial

El procesamiento de datos en un sistema de visión puede enfocarse desde dos perspectivas:

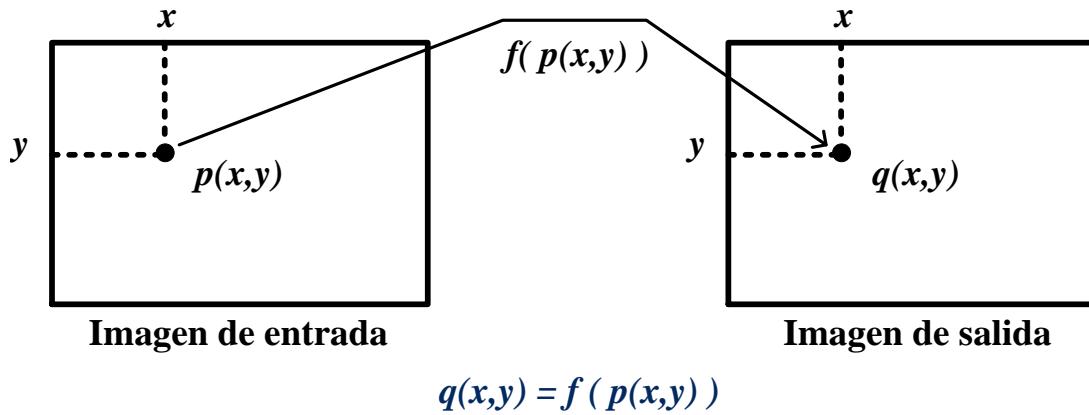
- Operaciones Individuales: *Alteración píxel a píxel de los datos en una escala global.*
- Operaciones de Vecindad: *Operaciones basadas en múltiples puntos .*

La generación de un nuevo píxel en una nueva imagen será una función bien del valor de cada píxel en su localización individual o bien de los valores de los píxeles en la vecindad de un píxel dado:



## **OPERACIONES INDIVIDUALES**

- Implican la generación de una nueva imagen modificando el valor del píxel en una simple localización basándose en una regla global aplicada a cada localización en la imagen original.
- Consisten en obtener el valor del píxel de una localización dada en la imagen, modificándolo por una operación lineal o no lineal y colocando el valor del nuevo píxel en la correspondiente localización de la nueva imagen. El proceso se repite para todas y cada una de las localizaciones de los píxeles en la imagen original.



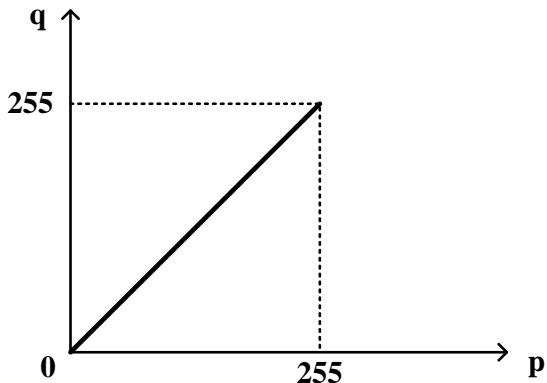
- La función transforma el valor del nivel de gris de cada píxel en la imagen obteniéndose una imagen de la misma dimensión que la original.
- Los valores de los píxeles de la imagen resultante dependen únicamente de la magnitud del correspondiente píxel de entrada (es independiente de los píxeles adyacentes).

**OPERACIONES INDIVIDUALES: Operador Identidad**

- Crea una imagen de salida idéntica a la imagen de entrada.

- Función de transformación:

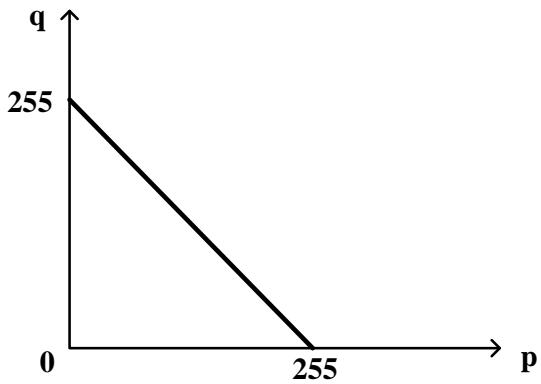
$$q = p$$

**OPERACIONES INDIVIDUALES: Operador Inverso o Negativo**

- Crea una imagen de salida inversa a la imagen de entrada.

- Función de transformación:

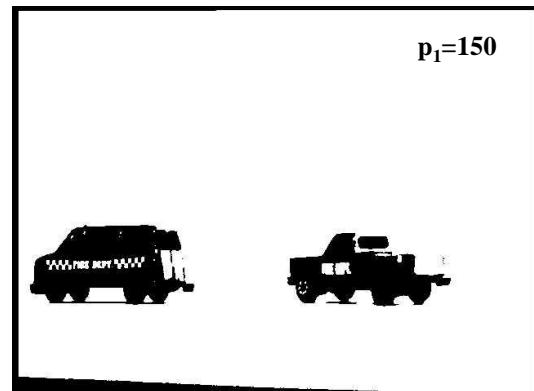
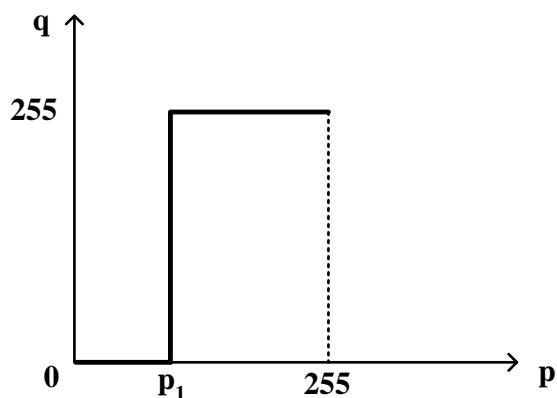
$$q = 255 - p$$



**OPERACIONES INDIVIDUALES: Operador Umbral**

➤ Función de transformación:

$$q = \begin{cases} 0 & \text{para } p \leq p_1 \\ 255 & \text{para } p > p_1 \end{cases}$$

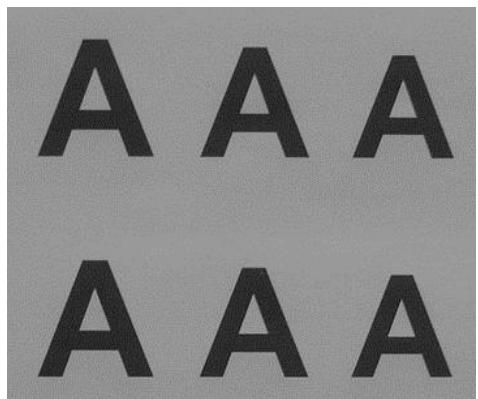


Se puede invertir el operador umbral:  $q = \begin{cases} 255 & \text{para } p \leq p_1 \\ 0 & \text{para } p > p_1 \end{cases}$

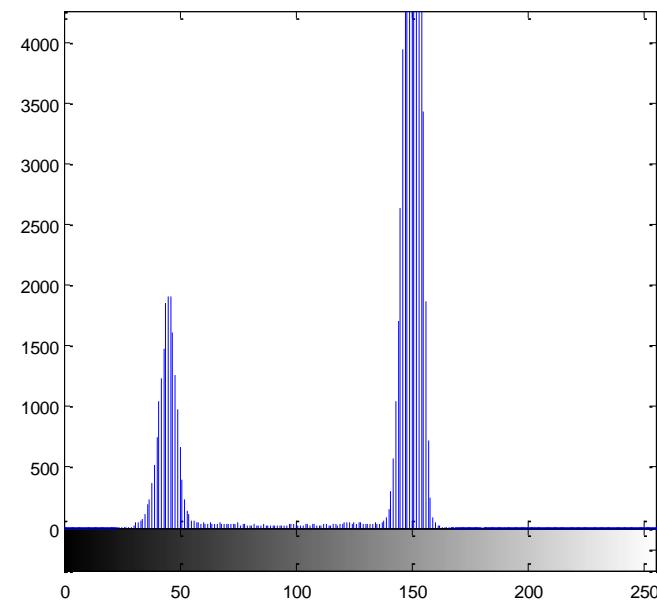
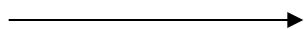
## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.2.- Técnicas Básicas de Procesamiento de Imágenes

#### EJEMPLO: Segmentación de objetos



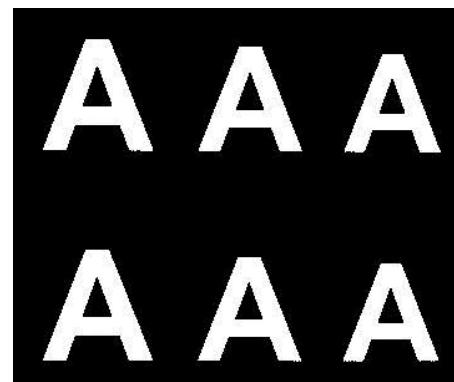
Histograma



#### Función de transformación:

$$q = \begin{cases} '1' \text{ (255 para visualización como imagen) para } p \leq 100 \\ 0 \text{ para } p > 100 \end{cases}$$

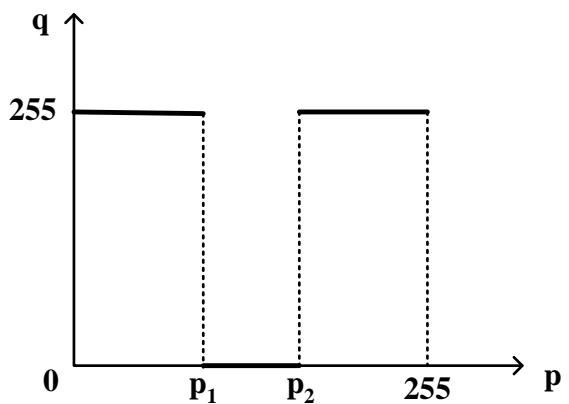
Imagen Segmentada



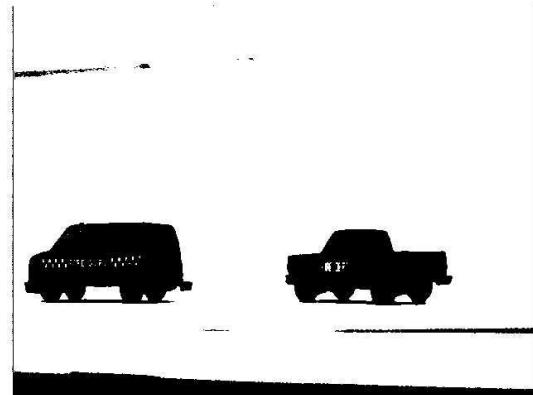
OPERACIONES INDIVIDUALES: Operador Intervalo de Umbral Binario

➤ Función de transformación:

$$q = \begin{cases} 255 & \text{para } p \leq p_1 \text{ o } p \geq p_2 \\ 0 & \text{para } p_1 < p < p_2 \end{cases}$$

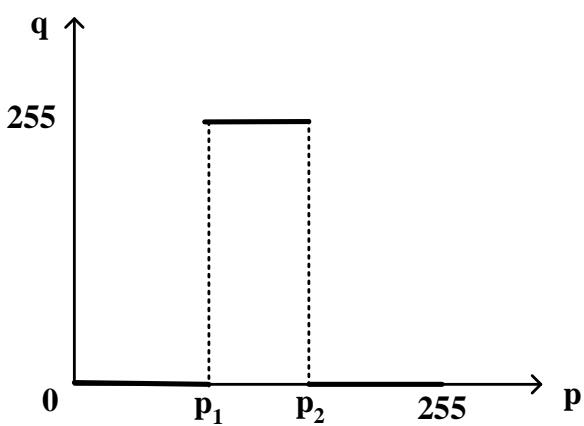


$$p_1=40 ; p_2=190$$

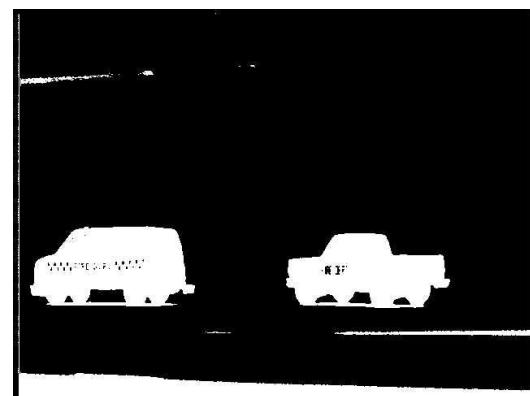
OPERACIONES INDIVIDUALES: Operador Intervalo de Umbral Binario Invertido

➤ Función de transformación:

$$q = \begin{cases} 0 & \text{para } p \leq p_1 \text{ o } p \geq p_2 \\ 255 & \text{para } p_1 < p < p_2 \end{cases}$$



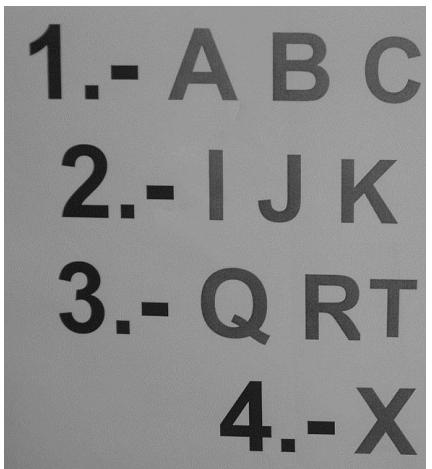
$$p_1=40 ; p_2=190$$



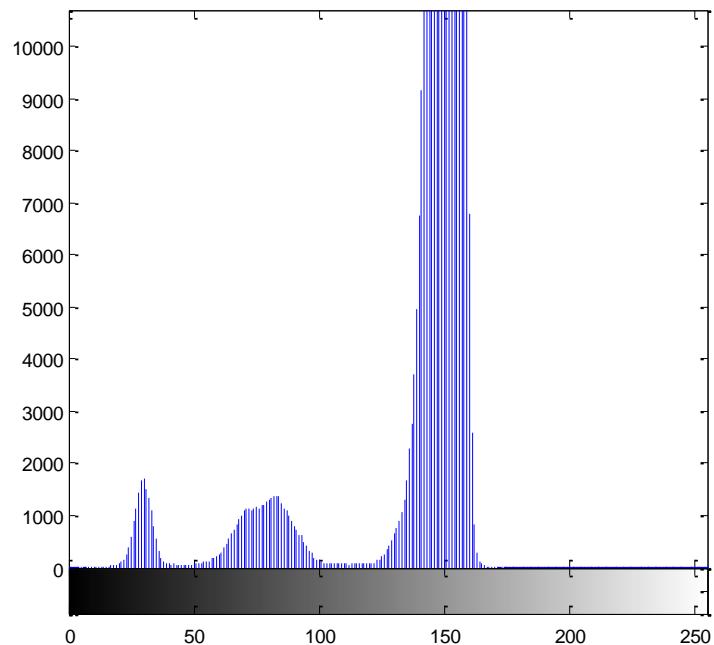
# TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

## 3.2.- Técnicas Básicas de Procesamiento de Imágenes

### EJEMPLO: Segmentación de objetos



Histograma  
→



### Función de transformación:

$$q = \begin{cases} '1' & (255 \text{ para visualización como imagen}) \text{ para } 45 < p < 110 \\ 0 & \text{otro caso} \end{cases}$$

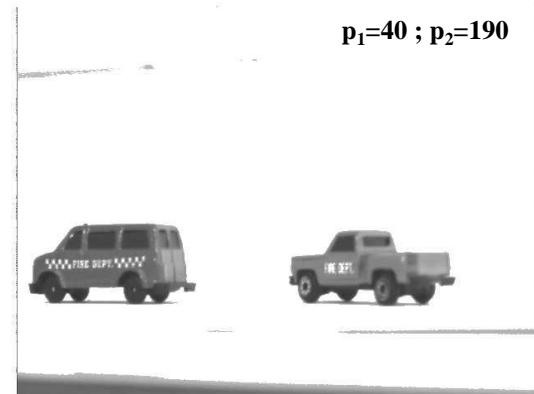
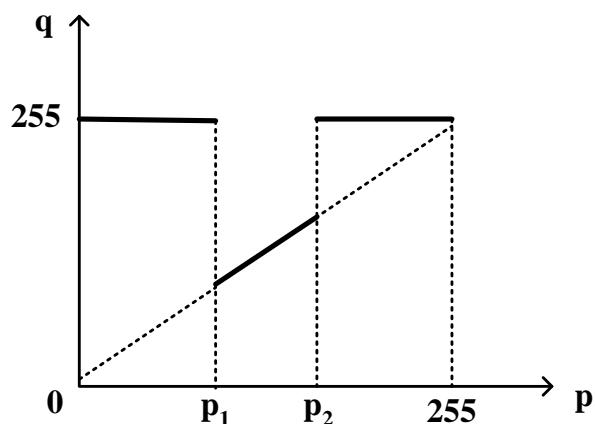
Imagen Segmentada  
→



OPERACIONES INDIVIDUALES: Operador de Umbral de la Escala de Grises

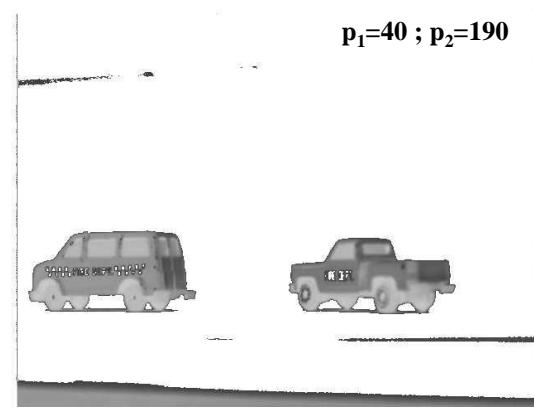
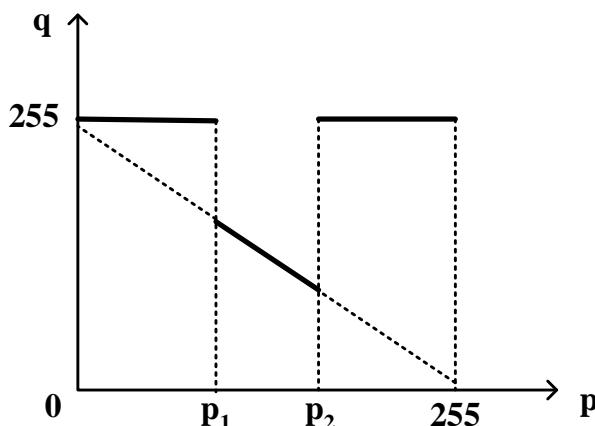
➤ Función de transformación:

$$q = \begin{cases} 255 & \text{para } p \leq p_1 \text{ o } p \geq p_2 \\ p & \text{para } p_1 < p < p_2 \end{cases}$$

OPERACIONES INDIVIDUALES: Operador de Umbral de la Escala de Grises Invertido

➤ Función de transformación:

$$q = \begin{cases} 255 & \text{para } p \leq p_1 \text{ o } p \geq p_2 \\ 255-p & \text{para } p_1 < p < p_2 \end{cases}$$



# TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

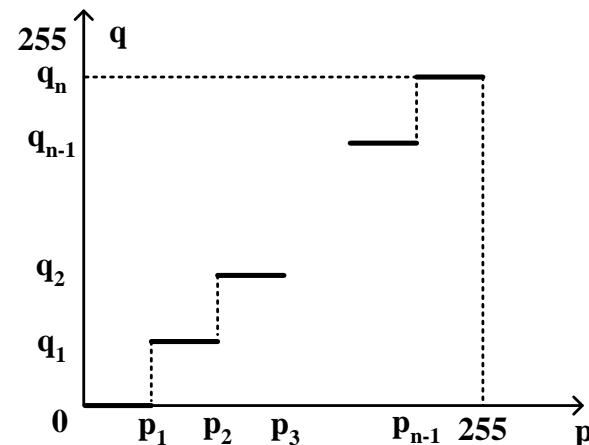
## 3.2.- Técnicas Básicas de Procesamiento de Imágenes

### OPERACIONES INDIVIDUALES: Operador Reducción del Nivel de Gris

➤ Proporciona una imagen de salida con un menor número de niveles de gris respecto de la imagen original de entrada (ésta es reducida a  $n+1$  niveles de gris).

➤ Función de transformación:

$$q = \begin{cases} 0 & \text{para } p \leq p_1 \\ q_1 & \text{para } p_1 < p < p_2 \\ \vdots \\ q_n & \text{para } p_{n-1} < p < 255 \end{cases}$$



$n=10$



# TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

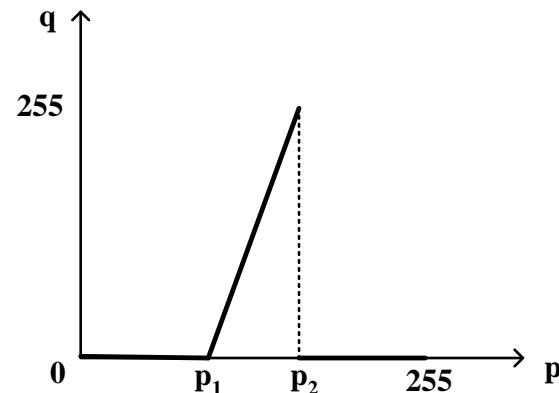
## 3.2.- Técnicas Básicas de Procesamiento de Imágenes

### OPERACIONES INDIVIDUALES: Operador de Extensión

➤ Proporciona una imagen de salida con la escala de grises completa correspondiente al intervalo de entrada definido por  $p_1$  y  $p_2$ , y suprime todos los valores fuera de este rango.

➤ Función de transformación:

$$q = \begin{cases} 0 & \text{para } p \leq p_1 \text{ o } p \geq p_2 \\ (p - p_1) \frac{255}{p_2 - p_1} & \text{para } p_1 < p < p_2 \end{cases}$$



$p_1=40 ; p_2=190$



# TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

## 3.2.- Técnicas Básicas de Procesamiento de Imágenes

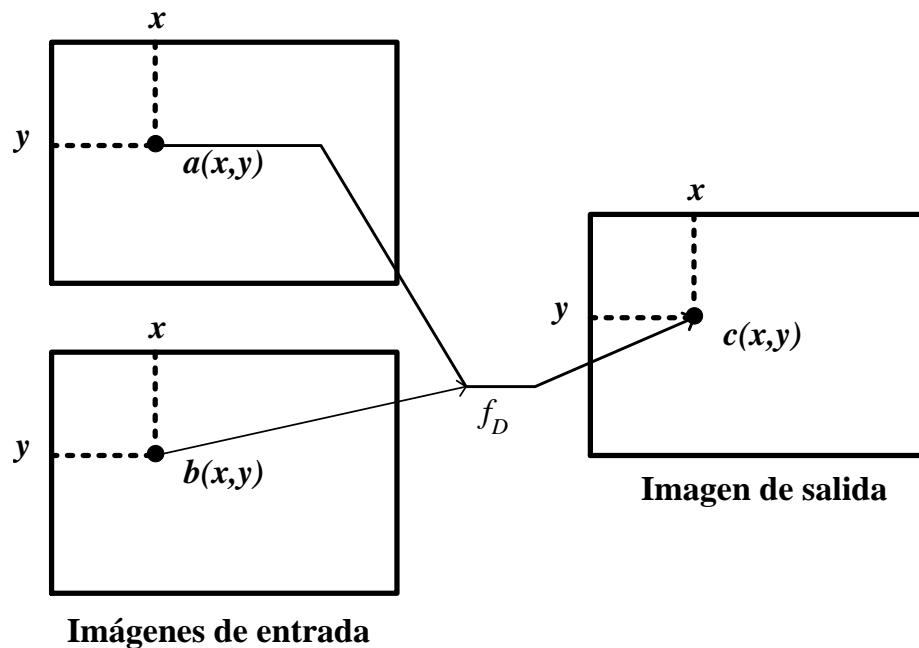
### OPERACIONES INDIVIDUALES: *Transformaciones de dos imágenes punto a punto*

➤ Utiliza la información contenida en la misma localización (posición de los píxeles) de dos imágenes de entrada A y B para crear un nueva imagen C

- La información en una localización de píxel de una imagen se combina con la información de la correspondiente localización de la segunda para obtener el valor, también en la misma localización de píxel de la tercera imagen.
- La dimensión de las imágenes es la misma.

➤ La función de la transformación  $c_{x,y} = f_D(a_{x,y}, b_{x,y})$  se aplica a todos los píxeles en las imágenes de entrada.

- $f_D$  puede ser adición, sustracción, multiplicación, división, exponenciación, máximo, ...



**OPERACIONES INDIVIDUALES: Transformaciones de dos imágenes punto a punto**

- **En el procesamiento de imágenes, los niveles de gris de los píxeles son números positivos:** después de una operación y antes de escribir los valores en la imagen, hay que normalizar los valores entre 0 y 255, es decir, es necesario reescalar de tal forma que el valor más negativo se corresponda con el 0 y el más positivo con el 255.

➤ **ADICIÓN**

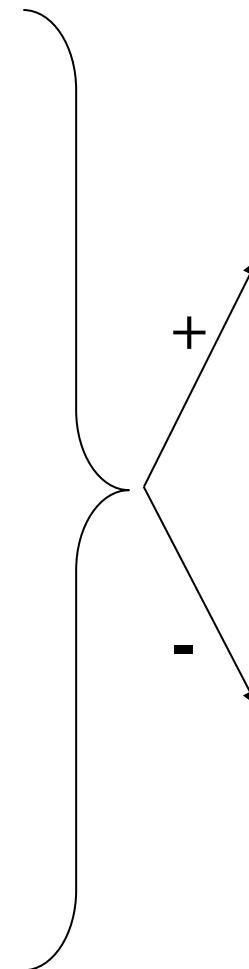
- **Efecto pixel overflow o desbordamiento de pixel:** las imágenes tienen una resolución que suele ser de 8 bits, por lo que el valor más alto que puede tener un píxel como nivel de gris es de 255. Si por efecto de sumar un valor se sobrepasa este máximo, el valor del nivel de gris del píxel resultante tendrá el valor de la suma menos 255.
- **Media:**  $c_{x,y} = (a_{x,y} + b_{x,y}) / 2$ , donde los valores de salida finales deben redondearse por defecto o por defecto. (Puede utilizarse para reducir los efectos del ruido en una imagen).

➤ **SUSTRACCIÓN**

- **Técnica útil para detectar el cambio producido en dos imágenes que han sido captadas en dos instantes de tiempo diferentes.**
- $c_{x,y} = f(a_{x,y} - b_{x,y})$ , donde  $f$  es una función de forma que el valor mínimo que toma  $c_{x,y}$  es 0 y máximo 255.

### OPERACIONES INDIVIDUALES: Transformaciones de dos imágenes punto a punto

➤ Ejemplo ADICIÓN y SUSTRACCIÓN:

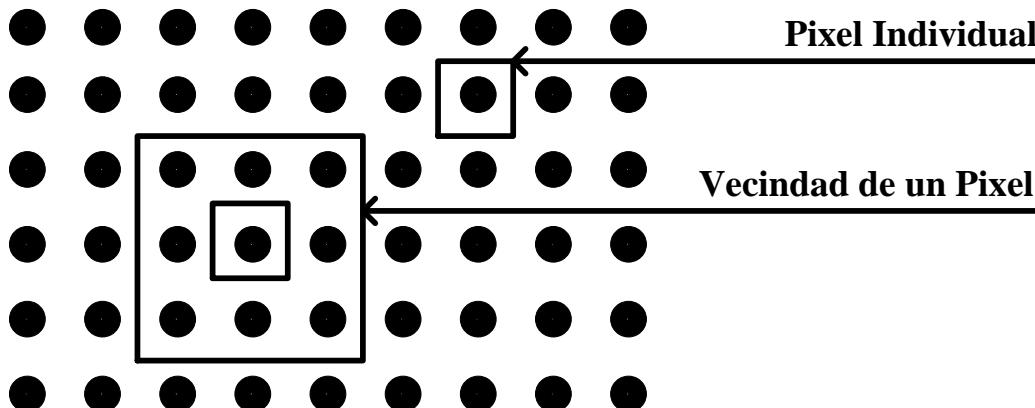


**PROCESAMIENTO BÁSICO DE IMÁGENES:** Transformaciones en el dominio espacial

El procesamiento de datos en un sistema de visión puede enfocarse desde dos perspectivas:

- Operaciones Individuales: *Alteración píxel a píxel de los datos en una escala global.*
- Operaciones de Vecindad: *Operaciones basadas en múltiples puntos.*

- El nuevo valor del píxel en la imagen de salida depende de una combinación de los valores de los píxeles en la vecindad del píxel de la imagen original que está siendo transformada.



# TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

## 3.2.- Técnicas Básicas de Procesamiento de Imágenes

### OPERACIONES DE VECINDAD

#### Operaciones de Convolución

- Considerando un entorno de vecindad  $E_8(p)$ , se realiza una suma ponderada con los valores de los 8 vecinos, siendo el resultado de esa suma el valor del nuevo píxel q de la imagen de salida en la misma posición (x,y).
- Los valores de ponderación se definen por medio de una máscara con valores constantes (filtros).
- La imagen de salida es de dimensión menor a la original: esta operación no se puede realizar sobre los píxeles extremos de la imagen original al no tener los 8 vecinos.

w1	w2	w3
w4	w5	w6
w7	w8	w9

Máscara

0,0	1,0	2,0	3,0	4,0	5,0
0,1	1,1	2,1	3,1	4,1	5,1
0,2	1,2	2,2	3,2	4,2	5,2
0,3	1,3	2,3	3,3	4,3	5,3
0,4	1,4	2,4	3,4	4,4	5,4
0,5	1,5	2,5	3,5	4,5	5,5

Imagen

$$G(2,2) = \begin{array}{l} f(1,1)*w1 + \\ f(2,1)*w2 + \\ f(3,1)*w3 + \\ f(1,2)*w4 + \\ f(2,2)*w5 + \\ f(3,2)*w6 + \\ f(1,3)*w7 + \\ f(2,3)*w8 + \\ f(3,3)*w9 \end{array}$$

# TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

## 3.2.- Técnicas Básicas de Procesamiento de Imágenes

**Convolución de dos funciones bidimensionales discretas:**  $f(x,y) * h(x,y) = [f * h](x,y) = \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} f(m,n)h(x-m, y-n)$

⇒ Ejemplo: Convolución en un punto  $(x,y)$  de una imagen digitalizada  $f(x,y)$  de dimensiones  $M \times N$  elementos y una función  $h(x,y)$  de dimensiones  $3 \times 3$ :

$$\begin{bmatrix} f(0,0) & f(1,0) & & \dots & f(M-2,0) & f(M-1,0) \\ f(0,1) & f(1,1) & & \dots & f(M-2,1) & f(M-1,1) \\ \vdots & \vdots & f(x-1,y-1) & f(x,y-1) & f(x+1,y-1) & \vdots & \vdots \\ & & f(x-1,y) & f(x,y) & f(x+1,y) & & \\ & & f(x-1,y+1) & f(x,y+1) & f(x+1,y+1) & & \\ f(0,N-2) & f(1,N-2) & & \dots & f(M-2,N-2) & f(M-1,N-2) \\ f(0,N-1) & f(1,N-1) & & \dots & f(M-2,N-1) & f(M-1,N-1) \end{bmatrix} * \begin{bmatrix} h(-1,-1) & h(0,-1) & h(1,-1) \\ h(-1,0) & h(0,0) & h(1,0) \\ h(-1,1) & h(0,1) & h(1,1) \end{bmatrix}$$

### 1. Obtención de $h(-m,-n)$ :

$$h(m,n) = \begin{bmatrix} h(-1,-1) & h(0,-1) & h(1,-1) \\ h(-1,0) & h(0,0) & h(1,0) \\ h(-1,1) & h(0,1) & h(1,1) \end{bmatrix}; \quad h(m,-n) = \begin{bmatrix} h(-1,1) & h(0,1) & h(1,1) \\ h(-1,0) & h(0,0) & h(1,0) \\ h(-1,-1) & h(0,-1) & h(1,-1) \end{bmatrix}; \quad h(-m,-n) = \begin{bmatrix} h(1,1) & h(0,1) & h(-1,1) \\ h(1,0) & h(0,0) & h(-1,0) \\ h(1,-1) & h(0,-1) & h(-1,-1) \end{bmatrix}$$

### 2. Desplazar $h(-m,-n)$ al punto $(x,y)$ y sumar los productos de aquellas celdas donde $f(x,y)$ y $h(x-m,y-n)$ se solapan:

$$f_0(x,y) = f(x,y) * h(x,y) = \begin{cases} = h(1,1)f(x-1,y-1) + h(0,1)f(x,y-1) + h(-1,1)f(x+1,y-1) + \\ + h(1,0)f(x-1,y) + h(0,0)f(x,y) + h(-1,0)f(x+1,y) + \\ + h(1,-1)f(x-1,y+1) + h(0,-1)f(x,y+1) + h(-1,-1)f(x+1,y+1) \end{cases}$$

## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.2.- Técnicas Básicas de Procesamiento de Imágenes

Operación de convolución de la función imagen por  $h(x,y)$ :  $fo(x,y) = h(x,y) * f(x,y)$

- En Visión por Computador, se habla, en lugar de convolucionar la imagen  $f(x,y)$  con la transformación  $h(x,y)$ , de convolucionar la imagen con una máscara  $h$ .
  - ✓  $fo(x,y) = \text{imagen resultante}$  ;  $f(x,y) = \text{imagen a procesar}$
  - ✓  $h(x,y) = \text{máscara o núcleo de convolución}$ : su tamaño puede ser muy variado: desde  $3x3$  (ejemplo anterior) hasta dimensiones mucho mayores (por ejemplo,  $63x63$ ).
- Por definición, antes de desplazar la máscara hasta el punto donde se va a calcular la convolución, hay que realizar dos giros para obtener  $h(-m,-n)$ . Muy a menudo, las máscaras son simétricas y estas rotaciones no son necesarias.
- En el caso de imágenes digitales, la operación de convolución se realiza para todos los píxeles de la imagen menos para los píxeles del borde, en los que no es posible aplicar la convolución al no solaparse totalmente la máscara y la imagen. Para imágenes de dimensiones  $MxN$  y máscaras  $mxn$ , la convolución resulta en una imagen más pequeña de dimensiones  $M-(m-1)xN-(n-1)$
- Para cada píxel de la imagen  $f(x,y)$ , se suman de forma ponderada los niveles de grises de su entorno de vecindad. La ponderación para cada píxel considerado viene dada por el valor almacenado en la máscara de convolución. El resultado obtenido es el nivel de gris resultante del píxel de la imagen resultante  $fo(x,y)$ .

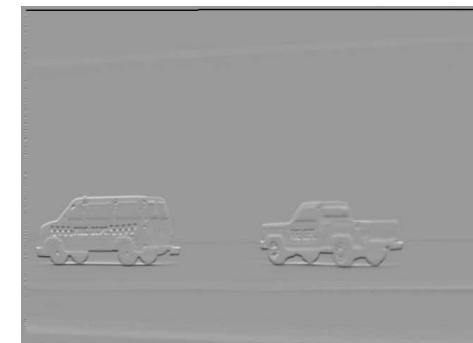
# TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

## 3.2.- Técnicas Básicas de Procesamiento de Imágenes

### CONVOLUCIÓN: Ejemplos

1.-

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$



$$q(x,y) = 1 \times p(x-1,y-1) + 2 \times p(x,y-1) + 1 \times p(x+1,y-1) + \\ + 0 \times p(x-1,y) + 0 \times p(x,y) + 0 \times p(x+1,y) - \\ - 1 \times p(x-1,y+1) - 2 \times p(x,y+1) - 1 \times p(x+1,y+1)$$

Se marcan los bordes respecto al resto de la imagen  
(la máscara utilizada es uno de los operadores de Sobel)

2.-

$$\begin{bmatrix} -0.1667 & -0.6667 & -0.1667 \\ -0.6667 & 4.3333 & -0.6667 \\ -0.1667 & -0.6667 & -0.1667 \end{bmatrix}$$



Obtención de una imagen de mayor contraste mediante una transformación de vecindad

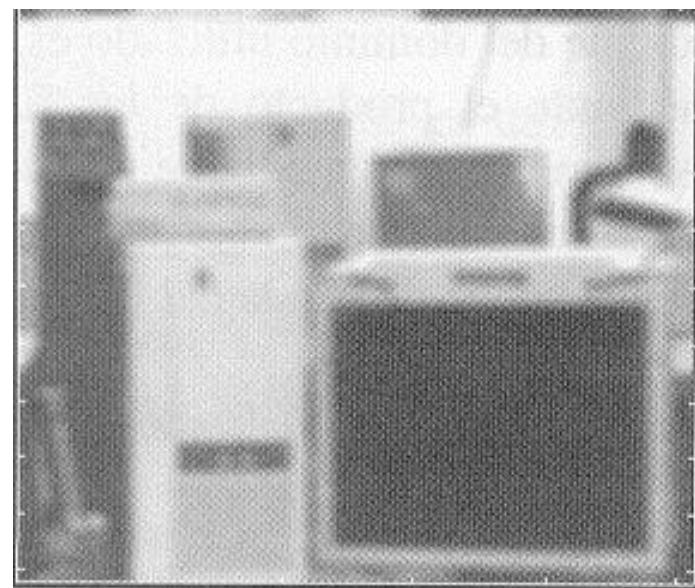
#### CONVOLUCIÓN: Ejemplos

3.-

$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

⇒ Es un filtro de suavizado, que promedia el entorno de vecindad.

➤ Resultado del Filtrado: imagen suavizada, se muestra borrosa con los bordes difuminados.



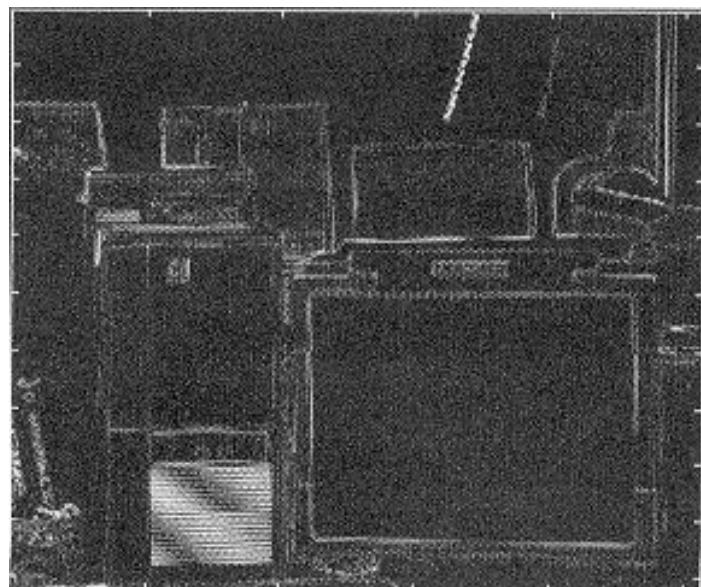
#### CONVOLUCIÓN: Ejemplos

4.-

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

⇒ Es un filtro Laplaciana utilizado para la extracción de bordes.

➤ Resultado del Filtrado: imagen de bordes obtenida mediante filtrado paso alto.



## TEMA 3

### – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL –

3.1.- Etapas en un proceso de visión por computador.

3.2.- Técnicas básicas de procesamiento de imágenes:

- *Procesamiento de imágenes en el dominio espacial*
- *Transformaciones lógicas*
- *Transformaciones geométricas*

3.3.- Segmentación de objetos basada en histograma

3.4.- Introducción al problema de la visión dinámica

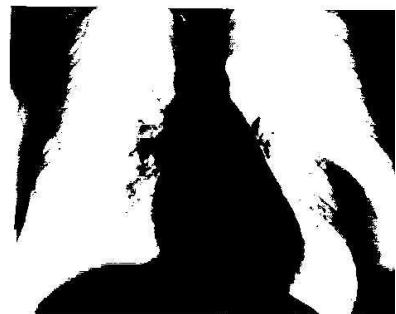
3.5.- Métodos de detección de movimiento

## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

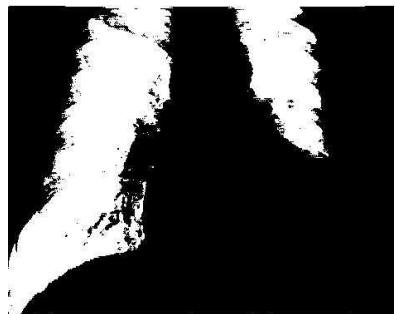
### 3.2.- Técnicas Básicas de Procesamiento de Imágenes

#### TRANSFORMACIONES LÓGICAS

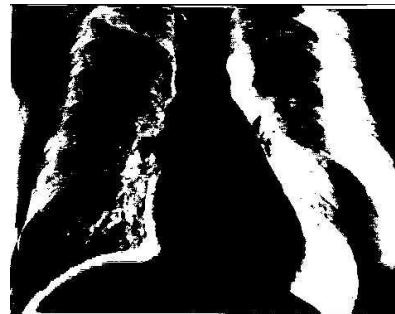
- Sólo se pueden aplicar sobre imágenes binarizadas.
- “NOT”, “OR”, “AND”, “XOR”
- Ejemplo:



OR Lógica



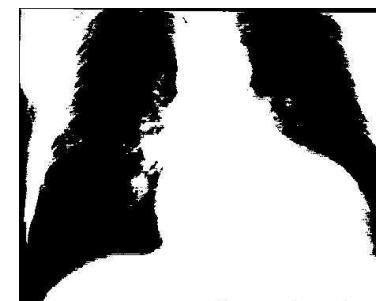
AND Lógica



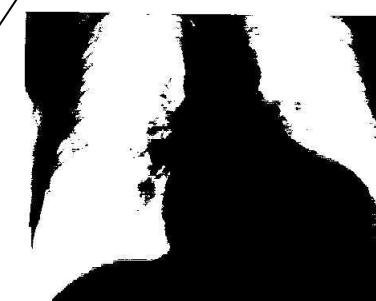
XOR Lógica



Imágenes originales con 256 niveles de gris



Binarización con umbral 128



## TEMA 3

### – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL –

3.1.- Etapas en un proceso de visión por computador.

3.2.- Técnicas básicas de procesamiento de imágenes:

- *Procesamiento de imágenes en el dominio espacial*
- *Transformaciones lógicas*
- *Transformaciones geométricas*

3.3.- Segmentación de objetos basada en histograma

3.4.- Introducción al problema de la visión dinámica

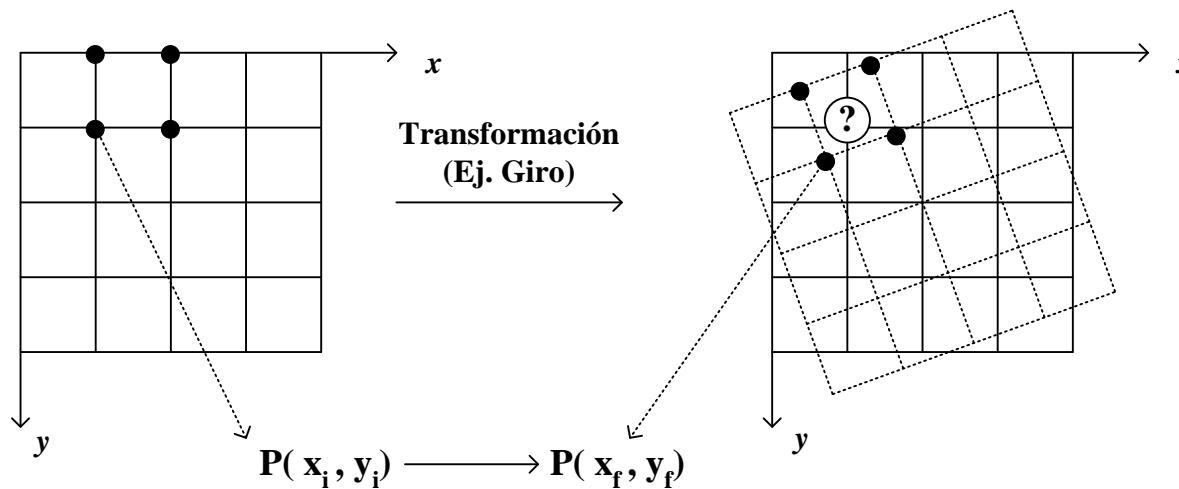
3.5.- Métodos de detección de movimiento

# TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

## 3.2.- Técnicas Básicas de Procesamiento de Imágenes

### TRANSFORMACIONES GEOMÉTRICAS

- Operaciones que modifican las coordenadas espaciales de los píxeles de una imagen.



#### Se necesitan dos tipos de algoritmos:

- La imagen digital es discreta, no existen valores de intensidad entre los valores discretos de las coordenadas  $x$  e  $y$ , que coinciden con las intersecciones de los valores discretos horizontales y verticales. Mediante una transformación, se modifican las coordenadas de los píxeles, por lo que los valores de éstos ya no tienen por que quedar sobre tales intersecciones ⇒ **Transformación geométrica**.
- La imagen final debe asumir una estructura similar a la de la imagen inicial, por lo que es necesario calcular los valores de los píxeles en las coordenadas  $x$  e  $y$  donde estaban definidos antes de la transformación ⇒ **Interpolación**.

## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.2.- Técnicas Básicas de Procesamiento de Imágenes

#### Transformaciones Geométricas:

➤ Traslación : si se quiere trasladar el origen de la imagen, se trasladan los valores de los píxeles en coordenadas  $x_i$  e  $y_i$  de la imagen original, a las coordenadas  $x_i+x_0$  e  $y_i+y_0$  de la imagen resultante.

$$x_f = x_i + x_0 \quad ; \quad \begin{bmatrix} x_f \\ y_f \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

#### ➤ Magnificación:

$$x_f = \frac{x_i}{a} \quad ; \quad \begin{bmatrix} x_f \\ y_f \\ 1 \end{bmatrix} = \begin{bmatrix} 1/a & 0 & 0 \\ 0 & 1/b & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

#### ➤ Rotación respecto al origen:

$$\begin{bmatrix} x_f \\ y_f \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

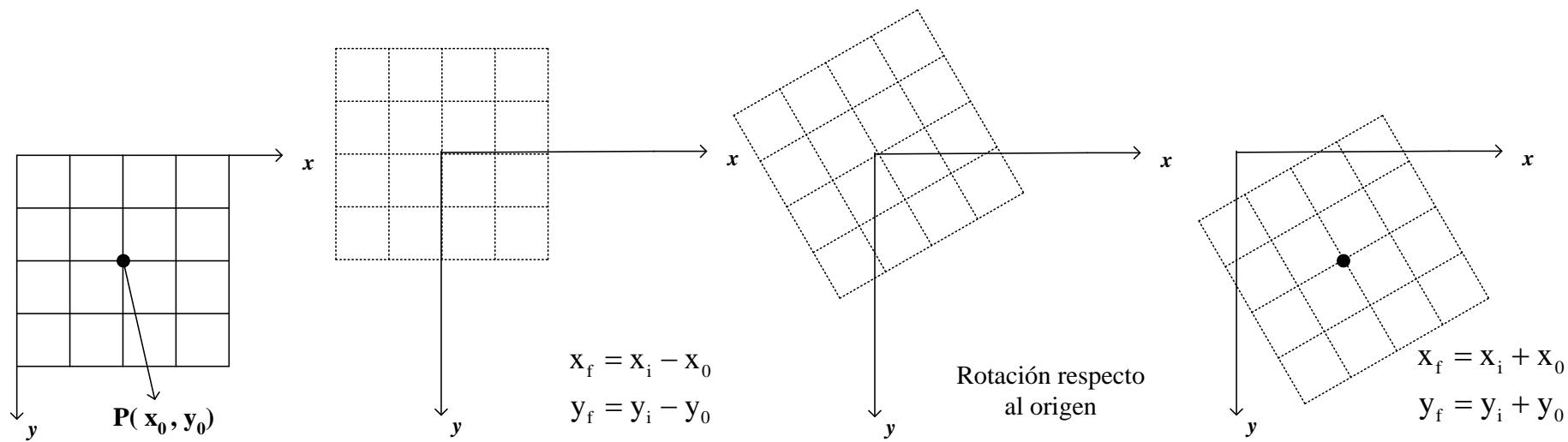
# TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

## 3.2.- Técnicas Básicas de Procesamiento de Imágenes

### Transformaciones Geométricas:

➤ Rotación respecto a un punto cualquiera: equivale a trasladar la imagen de forma que el punto desde el que se quiere rotar se sitúe en el origen, realizar una rotación respecto el origen y deshacer la traslación.

$$\begin{bmatrix} x_f \\ y_f \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -x_0 \\ 0 & 1 & -y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$



# TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

## 3.2.- Técnicas Básicas de Procesamiento de Imágenes

### Interpolación:

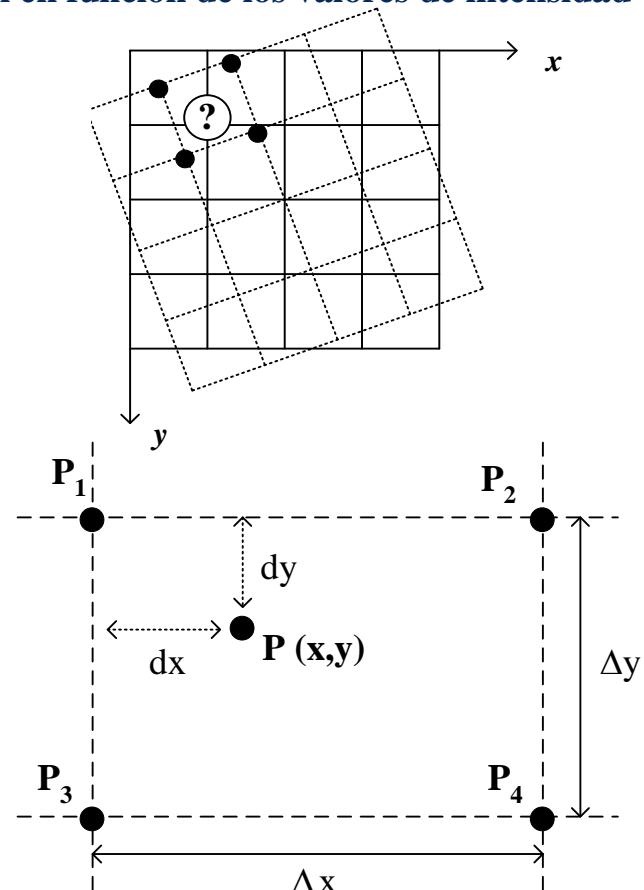
➤ Puede considerarse como el cálculo del valor de intensidad de un píxel en función de los valores de intensidad de los píxeles que le rodean:

### Distintas formas de hacerlo:

- Suponer que el píxel toma el mismo valor que el más cercano de entre los cuatro que le rodean (para decidir cuál es el más cercano se puede utilizar la distancia Euclídea).
- Asociar la intensidad media de los dos píxeles más cercanos.
- **Interpolación bilineal:** Asignar al píxel en cuestión un valor medio ponderado de las intensidades de los cuatro píxeles que le rodean. Los factores de ponderación vienen dados por la distancia entre el píxel y los del entorno:

$$P(x, y) = a_1 P_1 + a_2 P_2 + a_3 P_3 + a_4 P_4$$

$$a_1 = \left(1 - \frac{dx}{\Delta x}\right) \left(1 - \frac{dy}{\Delta y}\right) ; \quad a_2 = \frac{dx}{\Delta x} \left(1 - \frac{dy}{\Delta y}\right) ; \quad a_3 = \left(1 - \frac{dx}{\Delta x}\right) \frac{dy}{\Delta y} ; \quad a_4 = \frac{dx}{\Delta x} \frac{dy}{\Delta y}$$



## TEMA 3

### – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL –

3.1.- Etapas en un proceso de visión por computador.

3.2.- Técnicas básicas de procesamiento de imágenes.

3.3.- Segmentación de objetos basada en histograma

3.4.- Introducción al problema de la visión dinámica

3.5.- Métodos de detección de movimiento

## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.3.- Segmentación de objetos basada en histograma

**Segmentación de objetos por umbralización de nivel de gris:** técnica de segmentación empleada cuando hay una clara diferencia entre los niveles de gris representativos de los objetos a extraer y aquellos del fondo.

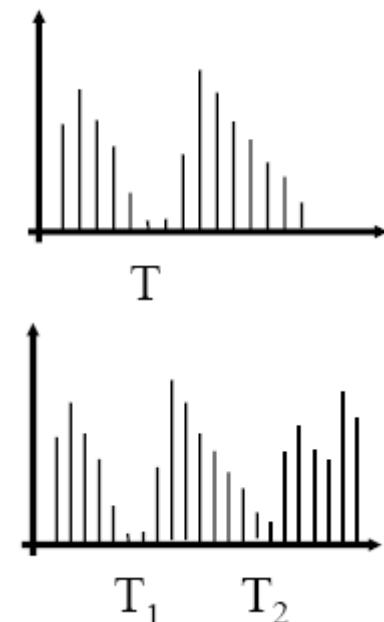
- **Objetivo:** convertir una imagen con varios niveles de gris a una nueva con sólo dos, de manera que los objetos queden separados del fondo.
- La umbralización se basa en que los píxeles de un determinado objeto tienen el mismo nivel de gris, o parecidos ⇒ por medio de umbrales, se seleccionan aquellos intervalos de gris que pertenecen a un objeto.

$$\Rightarrow g(x,y) = \begin{cases} 1 & \text{si } f(x,y) \geq T \\ 0 & \text{cualquier otro caso} \end{cases} \quad (\text{objeto claro sobre fondo oscuro})$$

→ Objeto claro sobre fondo oscuro: se supone que todos los píxeles correspondientes al objeto se encuentran por encima de un determinado umbral.

$$\Rightarrow g(x,y) = \begin{cases} 1 & \text{si } f(x,y) \leq T \\ 0 & \text{cualquier otro caso} \end{cases} \quad (\text{objeto oscuro sobre fondo claro})$$

$$\Rightarrow g(x,y) = \begin{cases} 1 & \text{si } T_1 \leq f(x,y) \leq T_2 \\ 0 & \text{cualquier otro caso} \end{cases} \quad (\text{niveles de gris de los píxeles del objeto comprendidos entre } T_1 \text{ y } T_2)$$



## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.3.- Segmentación de objetos basada en histograma

#### Segmentación de objetos por umbralización de nivel de gris:

**Objetivo:** encontrar los valores de gris adecuados como umbrales (debido al ruido y que el objeto y el fondo no tienen un único valor de gris, sino un intervalo, se pueden solapar valores).

- **Técnicas de selección de umbral:** es posible elegir el umbral directamente, basándose en conocimiento del contenido de la imagen. Si no es así hay que buscar métodos de selección automática. La mayoría de ellos se basan en el histograma:

→ Técnicas de detección de umbral basadas en histograma

#### Si se conoce la proporción de píxeles de cada tipo:

- ✓ **Método del percentil:** se calcula el histograma, y se va aumentando un nivel cada vez, comenzando en cero. Cada umbral se sitúa en el nivel que haya llegado a la proporción requerida.

## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.3.- Segmentación de objetos basada en histograma

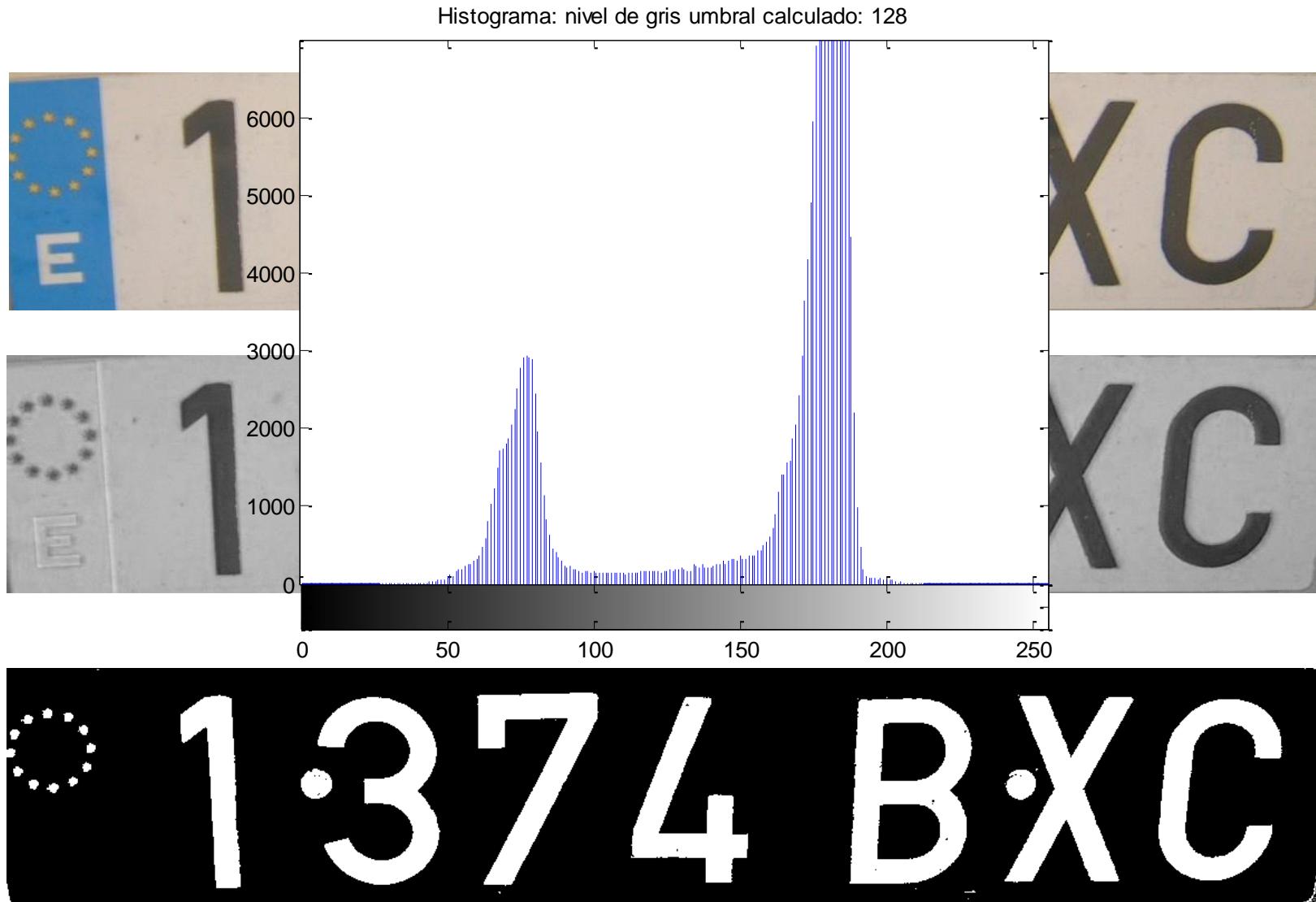
#### Si el histograma es claramente bimodal:

- ✓ Estadísticos simples: nivel de gris medio de la imagen, mediana, valor medio entre el nivel mínimo y máximo...
- ✓ Métodos de mínimo: escoger como umbral el nivel de gris correspondiente al mínimo entre los dos picos máximos del histograma,
  - Primero se encuentran los dos picos más altos y después el valor mínimo entre ellos.
  - Nivel de gris del primer pico: corresponde al nivel de gris en el que el histograma alcanza el mayor valor
  - Nivel de gris del segundo pico: multiplicar los valores del histograma por el cuadrado de la distancia al primer pico y tomar el máximo:  $\max [ (k - f)^2 * h(k) ], 0 \leq k \leq 255$ 
    - ❖ K: nivel de gris considerado; h(k): el valor del histograma en k; f: nivel de gris del primer pico más alto.
    - ❖ Hay que asegurarse que la separación entre los dos máximos sea suficiente para evitar que ambos se encuentren en la misma clase.
  - Problema: método muy sensible al ruido - necesario antes suavizar el histograma, y después se exige además que la separación en niveles entre dos máximos consecutivos sea suficiente.

## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.3.- Segmentación de objetos basada en histograma

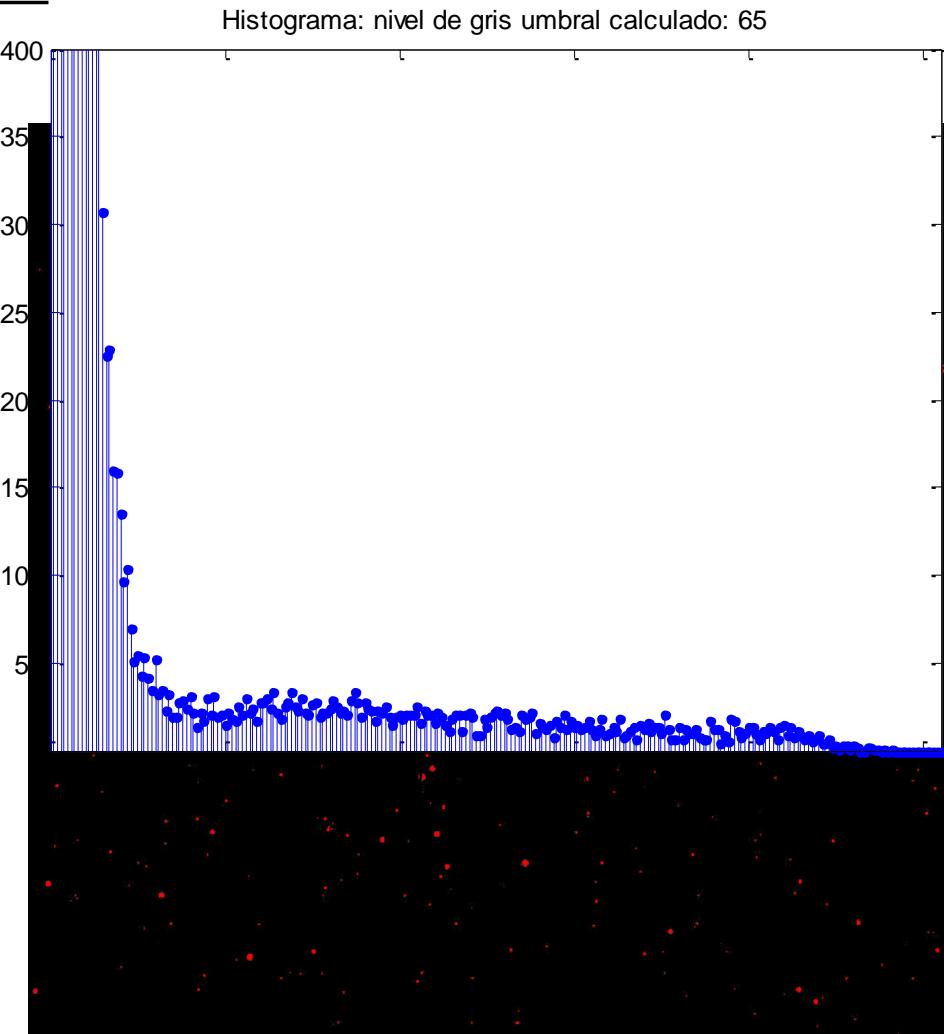
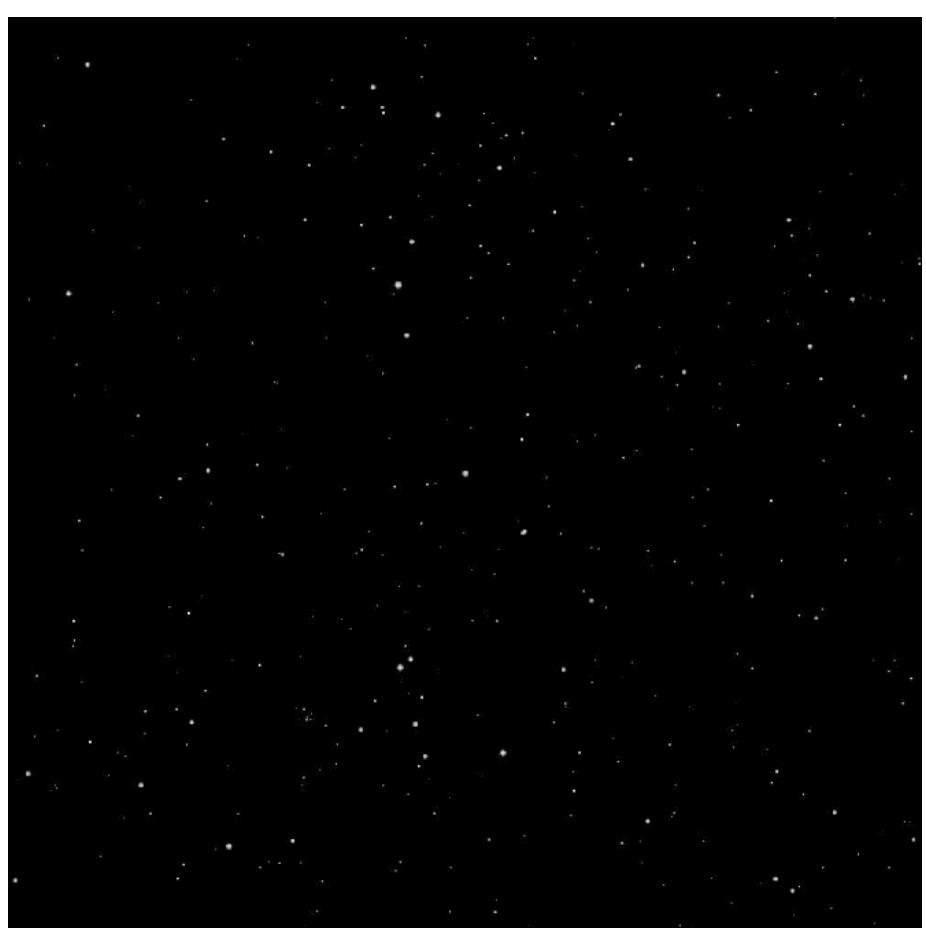
#### Ejemplos de binarización: selección automática de umbral. Histograma Bimodal



## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.3.- Segmentación de objetos basada en histograma

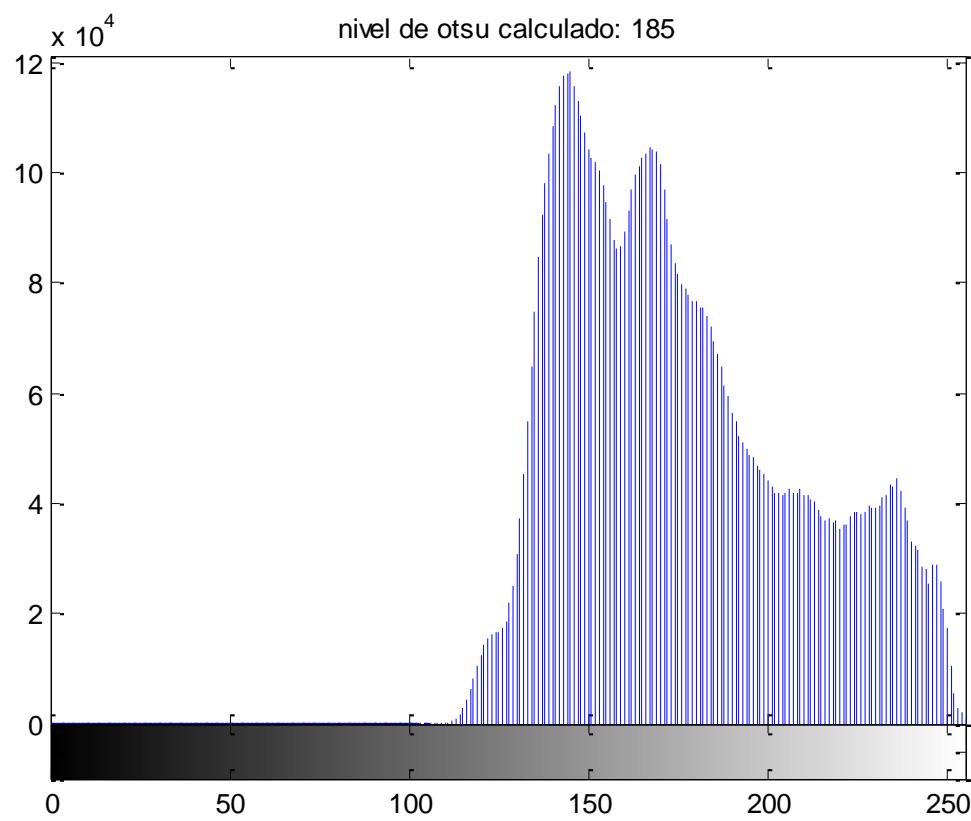
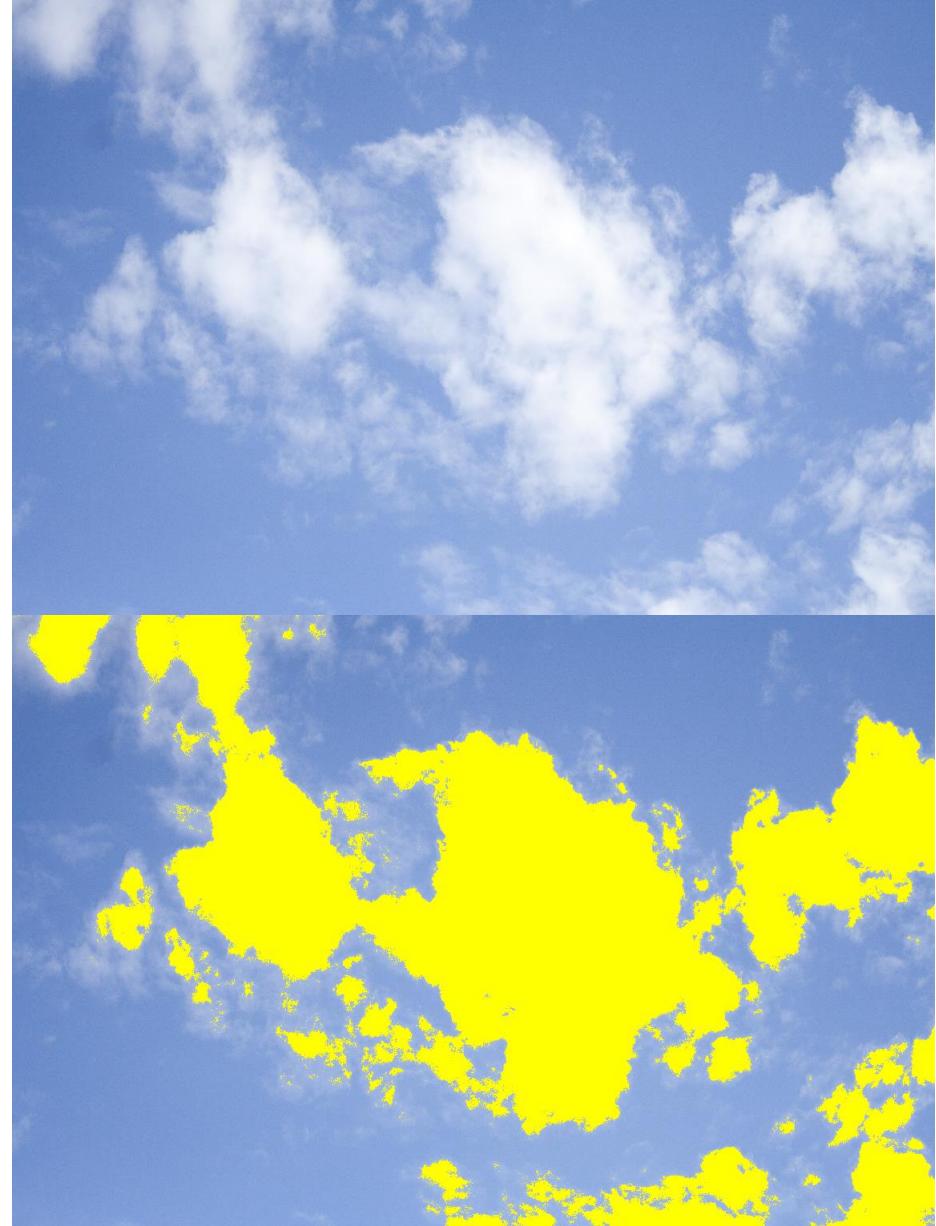
#### Ejemplos de binarización: selección automática de umbral



## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.3.- Segmentación de objetos basada en histograma

#### Ejemplos de binarización: selección automática de umbral



## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.3.- Segmentación de objetos basada en histograma

#### Binarización mediante detección de umbral basadas en histograma

**Si el histograma presenta solapamiento objetos/fondo:** técnicas de umbralización óptima - los umbrales se obtienen de forma óptima de acuerdo a algún criterio.

✓ **Método de Ridler y Calvard (1978) - ISODATA:**

1. Estima un valor de umbral inicial  $T$  ( por ejemplo, media de intensidad).
2. Establecer dos agrupaciones de píxeles:  $G_1$ : píxeles cuyo nivel de gris es menor o igual a dicho valor umbral  $T$ ;  $G_2$ : píxeles cuyo nivel de gris es mayor a  $T$ .
3. Calcula la media de intensidad  $m_1$  y  $m_2$  de los grupos  $G_1$  y  $G_2$ , respectivamente.
4. Calcula un nuevo valor umbral  $T$ :  $T = (1/2) * (m_1 + m_2)$
5. Repite los pasos 2, 3 y 4 hasta que no haya diferencia entre los valores  $T$  en sucesivas iteraciones ( $\Delta T = 0$ ) o ésta sea más pequeña que valor predefinido ( $\Delta T < \text{valor}$ )

## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.3.- Segmentación de objetos basada en histograma

#### Binarización mediante detección de umbral basadas en histograma

**Si el histograma presenta solapamiento objetos/fondo:** técnicas de umbralización óptima - los umbrales se obtienen de forma óptima de acuerdo a algún criterio.

- ✓ Método de Otsu (1979) – (IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, VOL. SMC-9, NO. 1, JANUARY 1979)

Consideraciones generales:

Sea una imagen  $I$  con  $N$  píxeles y  $L$  niveles de gris ( $g = 1, 2, \dots, L$ ), cuyo histograma viene dado por  $h$

- $h(g)$ : número de píxeles de  $I$  con nivel de gris  $g$
- Probabilidad de ocurrencia del nivel de gris  $g$  en la imagen:  $p(g) = \frac{h(g)}{N}$
- Valor medio:  $\mu = \frac{\sum_{g=1}^L g h(g)}{N} = \sum_{g=1}^L g p(g)$
- Varianza:

$$\sigma^2 = \frac{\sum_{g=1}^L (g - \mu)^2 h(g)}{N} = \sum_{g=1}^L (g - \mu)^2 p(g)$$

## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.3.- Segmentación de objetos basada en histograma

- ✓ Método de Otsu (1979) – (IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, VOL. SMC-9, NO. 1, JANUARY 1979)

Consideraciones generales:

- Si fijásemos un umbral de binarización en un nivel de gris  $g = k$ , la imagen y su histograma quedarían divididos en dos clases o agrupaciones de píxeles:

C1: píxeles de I cuyo nivel de gris  $g$  es menor o igual a  $k$ :  $[1, 2, \dots, k]$

C2: píxeles de I cuyo nivel de gris es mayor a  $k$ :  $[k+1, k+2, \dots, 256]$

- Número de píxeles de cada clase:  $N1(k) = \sum_{g=1}^k h(g)$  ;  $N2(k) = \sum_{g=k+1}^L h(g)$
- Probabilidad de ocurrencia de un nivel de gris de cada clase:  $\omega1(k) = \frac{N1}{N}$  ;  $\omega2(k) = \frac{N2}{N}$
- Nivel de gris medio de cada clase:

$$\mu1(k) = \frac{\sum_{g=1}^k gh(g)}{N1(k)} ; \mu2(k) = \frac{\sum_{g=k+1}^L gh(g)}{N2(k)}$$

- Varianza entre clases ponderada por la probabilidad de ocurrencia (asumiendo que cada clase está representada por su nivel de gris medio):

$$\sigma_B^2(k) = \omega1(k) * (\mu1(k) - \mu)^2 + \omega2(k) * (\mu2(k) - \mu)^2$$

## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.3.- Segmentación de objetos basada en histograma

- ✓ **Método de Otsu (1979) – (IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, VOL. SMC-9, NO. 1, JANUARY 1979)**

**Caso binarización en dos niveles:**

- **El Umbral es el nivel de gris que maximiza la varianza entre clases:**

$$T = \max_k \left\{ \sigma_B^2(k) \right\} ; \quad k = 1, 2, \dots, L$$

- ✓ Con este umbral, se dispersan lo máximo posible los niveles de intensidad de las dos clases.
- ✓ Si el máximo no es único, se define T como la media de los niveles de gris que proporcionan todos los máximos posibles.

**Caso binarización en M niveles:**

- **En este caso, al existir M clases, existirán M-1 umbrales distintos.**
- **Generalizando el caso de dos niveles, estos umbrales serán el conjunto de niveles de gris que maximiza la varianza entre clases de la forma:**

$$\{T_1, T_2, \dots, T_{M-1}\} = \max_{\{k_1, k_2, \dots, k_{M-1}\}} \left\{ \sigma_B^2(k_1, k_2, \dots, k_{M-1}) \right\} ; \quad 1 < k_1 < k_2 < k_{M-1} < L$$

$$\sigma_B^2(k_1, k_2, \dots, k_{M-1}) = \sum_{i=1}^M (\omega_i(k_1, k_2, \dots, k_{M-1}) * (\mu_i(k_1, k_2, \dots, k_{M-1}) - \mu)^2)$$

## TEMA 3

### – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL –

3.1.- Etapas en un proceso de visión por computador.

3.2.- Técnicas básicas de procesamiento de imágenes.

3.3.- Segmentación de objetos basada en histograma

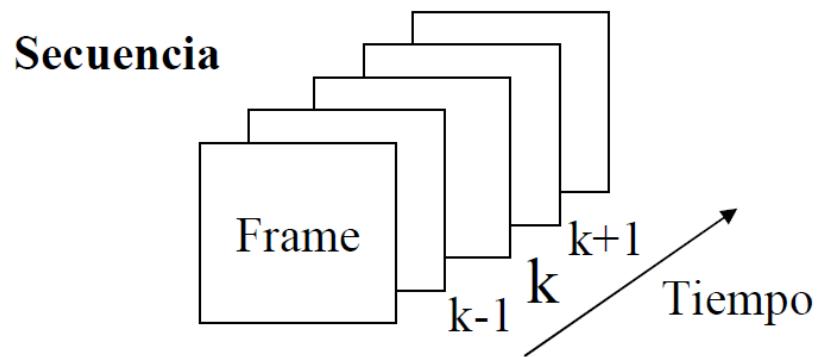
3.4.- Introducción al problema de la visión dinámica

3.5.- Métodos de detección de movimiento

## **INTRODUCCIÓN AL PROBLEMA DE LA VISIÓN DINÁMICA**

### **CONSIDERACIONES:**

- ✓ Una escena dinámica se compone de una secuencia de frames interrelacionados.
  - ✓ Cada frame representa la escena cambiante en un determinado instante de tiempo t.
  - ✓ En general, la entrada a un sistema de visión dinámica será una función  $f(x,y,t)$  donde x,y representarán las coordenadas espaciales en el frame que representa la escena en el instante de tiempo t.
- 
- ✓ Los cambios en la escena pueden ser debidos al movimiento de la propia escena, al movimiento de la cámara o al de ambos.
  - ✓ Para procesar información dinámica podemos aplicar técnicas estáticas a cada frame por separado o analizar una secuencia de forma continua.
  - ✓ Cualquier cambio en la iluminación de la escena puede provocar cambios en los niveles de intensidad de la imagen que no se traducirá en movimiento real de los objetos.
  - ✓ La mayoría de las técnicas de análisis de escenas dinámicas están basadas en la detección de cambios en la secuencia de frames, ya sea frame a frame, región a región o píxel a píxel.



## **INTRODUCCIÓN AL PROBLEMA DE LA VISIÓN DINÁMICA**

- ✓ El análisis del movimiento está relacionado con la obtención de información dinámica, que varía en el tiempo, de una escena. Sus principales objetivos son:
- **Detección de movimiento.**
    - ❖ El objetivo es detectar si hay movimiento en la escena.
    - ❖ Tiene aplicaciones a seguridad, video vigilancia...
  - **Detección y localización de objetos en movimiento.**
    - ❖ Es más complicado que el anterior y puede incluir también la detección de trayectorias y predicción de futuras trayectorias.
    - ❖ Tiene aplicaciones de seguimiento.

### **MÉTODOS DE DETECCIÓN DE MOVIMIENTO**

Existen diferentes técnicas de detección de movimiento – las que se estudian en el siguiente apartado son:

- **Sustracción del fondo (Background subtraction)**
- **Imagen diferencia**
- **Imagen de diferencias acumuladas.**

## **INTRODUCCIÓN AL PROBLEMA DE LA VISIÓN DINÁMICA**

### **MÉTODOS DE DETECCIÓN DE MOVIMIENTO**

- **Ajuste de bloques (block matching):**
  - Se divide la imagen en bloques y se estima el desplazamiento bajo una medida de error.
  - El desplazamiento que minimice el error se considera el movimiento de la imagen.
  - Las técnicas más usuales de búsqueda inteligente son:
    - ✓ Método Three-Step
    - ✓ Método conjugado modificado.
- **Detección de cambios en el flujo óptico**
  - Flujo Óptico (Optical Flow): Campo de velocidades asociado con los cambios en los niveles de intensidad en una imagen. Puede dar información sobre la localización espacial de los objetos vistos y la velocidad de cambio de esta localización.
  - Las técnicas más utilizadas para el cálculo del flujo óptico:
    - ✓ Métodos diferenciales (basados en computar derivadas espacio-temporales de la intensidad de la imagen - Algoritmo de Horn & Schunck).
    - ✓ Métodos de ajuste de regiones (definen la velocidad como el desplazamiento necesario que mejor ajuste nos da entre regiones de imágenes a diferentes tiempos. Para realizar el ajuste utilizamos una función o medida de similitud).

## **INTRODUCCIÓN AL PROBLEMA DE LA VISIÓN DINÁMICA**

### **DETECCIÓN Y LOCALIZACIÓN DE OBJETOS EN MOVIMIENTO: SEGUIMIENTO**

#### **Objetivos:**

- Seguimiento espacio-temporal de un móvil a través de la estimación óptima de su estado.
- Predicción-actualización del estado futuro del modelo a partir de una secuencia de observaciones.

#### **Consideraciones:**

- Una proyección de un movimiento 3D suave sigue siendo suave
- El movimiento de un objeto en cualquier punto de la imagen en una secuencia no cambiará abruptamente. 3 asunciones que ayudan a resolver problema de la correspondencia:
  - La posición de un punto dado será relativamente invariante de un frame al siguiente.
  - La magnitud de velocidad de un punto dado será relativamente invariante de un frame al siguiente
  - La dirección de movimiento de un punto dado es relativamente invariante de un frame al siguiente
- El problema puede ser sencillo si tenemos un único objeto en movimiento en la secuencia. Se complica si tenemos varios móviles: problema del ajuste o correspondencia.
- Entre otras técnicas, destacan:
  - Filtro predictivo de Kalman
  - Filtros de Partículas

## TEMA 3

### *– FUNDAMENTOS DE LA VISIÓN ARTIFICIAL –*

3.1.- Etapas en un proceso de visión por computador.

3.2.- Técnicas básicas de procesamiento de imágenes.

3.3.- Segmentación de objetos basada en histograma

3.4.- Introducción al problema de la visión dinámica

*3.5.- Métodos de detección de movimiento*

APROXIMACIÓN BÁSICA: Imágenes diferencia

- Simple comparación píxel a píxel de dos imágenes tomadas en dos instantes de tiempo, implementada como una resta en valor absoluto.
- Comparar una imagen de una escena dinámica con la siguiente mediante la diferencia de ambas, implica la cancelación de los elementos estáticos (que no varían en el tiempo); las contribuciones distintas de cero se corresponden con las componentes de la imagen en movimiento.
- IMAGEN DIFERENCIA BINARIA entre dos imágenes tomadas en instantes de tiempo  $t_i$  y  $t_j$ :

$$DI_{ij}(x, y) = \begin{cases} 1 & \text{si } |f(x, y, t_i) - f(x, y, t_j)| > T \\ 0 & \text{en otro caso} \end{cases}$$

$f(x, y, t_i)$ ,  $f(x, y, t_j)$ : imágenes tomadas en los instantes de tiempo  $t_i$  y  $t_j$   
 $T$ : umbral para considerar que la diferencia es apreciablemente diferente

- En un procesamiento de imágenes dinámicas, todos los píxeles en  $DI_{ij}(x, y)$  con valor 1 se consideran el resultado del movimiento, lo cual sólo puede asumirse si la iluminación es relativamente constante dentro de los límites establecidos por  $T$ .
- En la práctica, los valores en  $DI_{ij}(x, y)$  con valor 1 pueden aparecer como puntos ruidosos asilados, por lo que esta aproximación suele ir acompañada de una eliminación de agrupaciones conexas que no tengan un número mínimo de píxeles (hay que tener especial cuidado en casos en los que se busque la detección de movimientos lentos o de pequeños objetos).

**APROXIMACIÓN BÁSICA: Imágenes diferencias acumulativas**

- **Sea la siguiente secuencia de frames:**

$f(x, y, t_1), f(x, y, t_2), \dots, f(x, y, t_n)$ ,      *siendo     $R(x, y) = f(x, y, t_1)$  una imagen de referencia*

- Una imagen diferencias acumulativa (ADI) se forma comparando una imagen de referencia con cada uno de los frames siguientes de la secuencia:

- Se obtiene incrementando un contador para cada pixel cada vez que ocurre una diferencia significativa entre la imagen de la secuencia y la de referencia.
- Cuando el frame  $k$ -ésimo se compara con el de referencia, la entrada de un pixel dado en la imagen acumulativa representa el número de veces que la intensidad en esa posición fue significativa (según el umbral  $T$  considerado) respecto al valor en la imagen referencia en los frames anteriores.

- Se definen tres tipos de ADIs: absolutas, positivas y negativas (asumiendo objeto en movimiento más oscuro que el fondo – si fuera más claro, la ADI positiva y negativa se intercambiarían):

$$\text{ADI Absoluta de la secuencia } k - \text{ésima: } A_k(x, y) = \begin{cases} A_{k-1}(x, y) + 1 & \text{si } |R(x, y) - f(x, y, t_k)| > T \\ A_{k-1}(x, y) & \text{en otro caso} \end{cases}$$

$$\text{ADI Positiva de la secuencia } k - \text{ésima: } P_k(x, y) = \begin{cases} P_{k-1}(x, y) + 1 & \text{si } (R(x, y) - f(x, y, t_k)) > T \\ P_{k-1}(x, y) & \text{en otro caso} \end{cases}$$

$$\text{ADI Negativa de la secuencia } k - \text{ésima: } N_k(x, y) = \begin{cases} N_{k-1}(x, y) + 1 & \text{si } (R(x, y) - f(x, y, t_k)) < -T \\ N_{k-1}(x, y) & \text{en otro caso} \end{cases}$$

(ADIs inicializadas a cero)

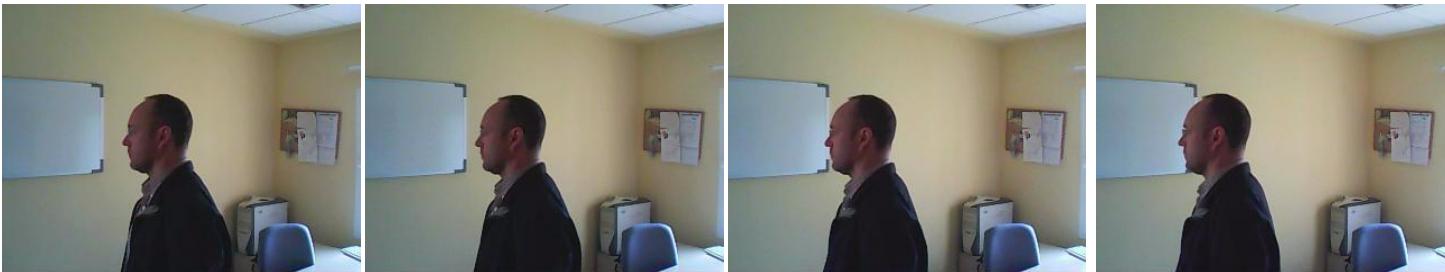
## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.5.- Métodos de detección de movimiento

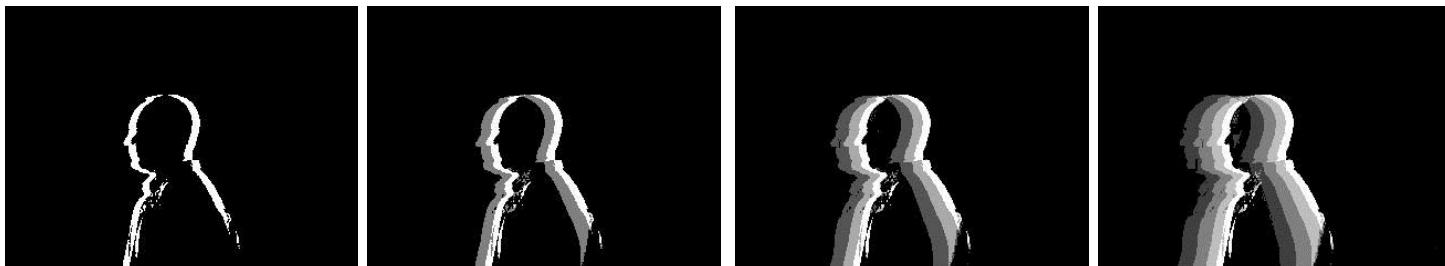
#### APROXIMACIÓN BÁSICA: Imágenes diferencias acumulativas - EJEMPLOS



FRAME  
REFERENCIA



ADI ABSOLUTA



ADI POSITIVA



ADI NEGATIVA



## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.5.- Métodos de detección de movimiento

#### APROXIMACIÓN BÁSICA: Imágenes diferencias acumulativas. CONSIDERACIONES

- El número de píxeles con valor distinto de cero en la ADI positiva representa el número de píxeles del objeto que han sufrido movimiento.
- La localización de los valores en la ADI positiva se corresponde con la localización del objeto en movimiento en el frame actual.
- El número de cuentas en la ADI positiva deja de incrementarse cuando el objeto en movimiento se desplaza completamente con respecto a esa zona en la imagen de referencia.
- La ADI absoluta contiene las regiones de la ADI positiva y negativa.
- La dirección y velocidad del objeto en movimiento pueden determinarse desde las valores de las ADI absoluta y negativa.
- De esta forma, cuando el frame k-ésimo se compara con el de referencia, la entrada de un pixel dado en la imagen acumulativa representa el número de veces que la intensidad en esa posición fue significativa (según el umbral T considerado) respecto al valor en la imagen referencia.
- *Por tanto: por medio de las imágenes diferencias acumulativas se puede extraer información de cómo ha sido el movimiento: dirección y desplazamiento del movimiento (la ADI positiva contiene el estado del objeto en su estado final o actual mientras que la negativa contiene el estado del mismo en su posición inicial), velocidad.*

**APROXIMACIÓN BÁSICA: Imagen de Referencia – Sustracción de fondo (elementos estáticos)**

- Si es posible generar una imagen de referencia que contenga únicamente elementos estáticos, la diferencia de intensidad entre los frames analizados de una escena dinámica (con elementos en movimiento) y esta referencia, muestra directamente los píxeles afectados por el movimiento presente (píxeles que presentan un valor diferencia de intensidad superior al umbral previamente establecido para considerar variaciones significativas).
- Cuando no es posible obtener esta imagen de fondo de referencia de forma directa (escenas con movimiento continuo, necesidad de actualización por condiciones cambiantes,...) hay que recurrir a técnicas cuyo objetivo sea su generación a partir de frames afectados por movimiento. En este contexto, puede ser de gran utilidad la información que aporta las imágenes diferencia definidas en su forma positiva y negativa.
- Asumiendo objeto más oscuro que el fondo, estas imágenes diferencia se definen para dos frames tomados en instantes de tiempo  $t_i$  y  $t_j$ :

$$\text{Imagen Diferencia Absoluta; } DI_{ij}(x, y) = \begin{cases} 1 & \text{si } |f(x, y, t_i) - f(x, y, t_j)| > T \\ 0 & \text{en otro caso} \end{cases}$$

$$\text{Imagen Diferencia Positiva ; } DI_{ij}(x, y) = \begin{cases} 1 & \text{si } [f(x, y, t_i) - f(x, y, t_j)] > T \\ 0 & \text{en otro caso} \end{cases}$$

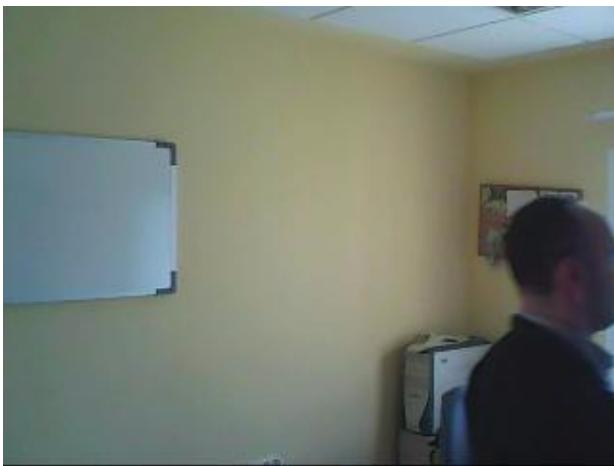
$$\text{Imagen Diferencia Negativa; } DI_{ij}(x, y) = \begin{cases} 1 & \text{si } [f(x, y, t_i) - f(x, y, t_j)] < -T \\ 0 & \text{en otro caso} \end{cases}$$

## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.5.- Métodos de detección de movimiento

#### APROXIMACIÓN BÁSICA: Imagen de Referencia – Sustracción de fondo (elementos estáticos)

EJEMPLO: Obtención de imagen de fondo estático a partir de dos frames con elementos en movimiento más oscuros que el fondo:



**FRAMES TOMADOS EN DOS INSTANTES DE TIEMPO**

**DIs NEGATIVA Y POSITIVA**

#### **IMAGEN DE FONDO:**

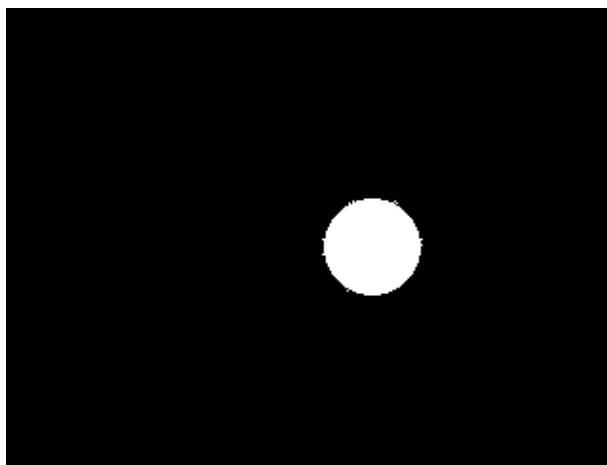
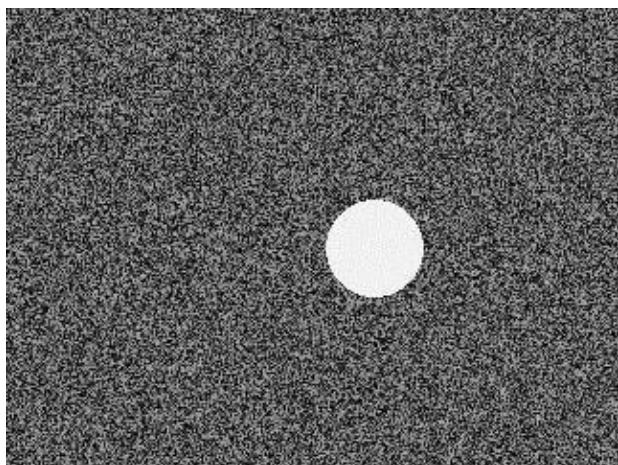
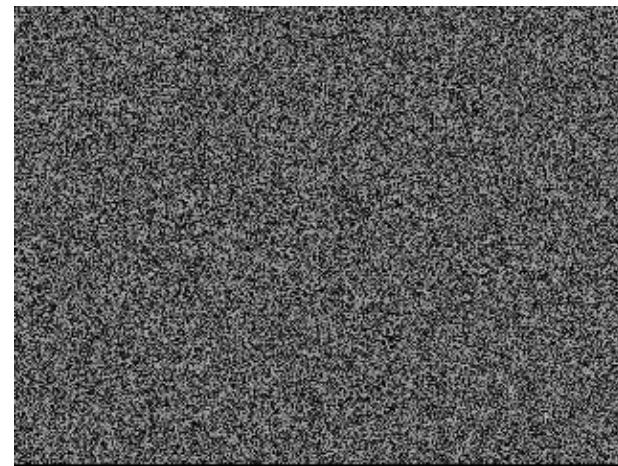
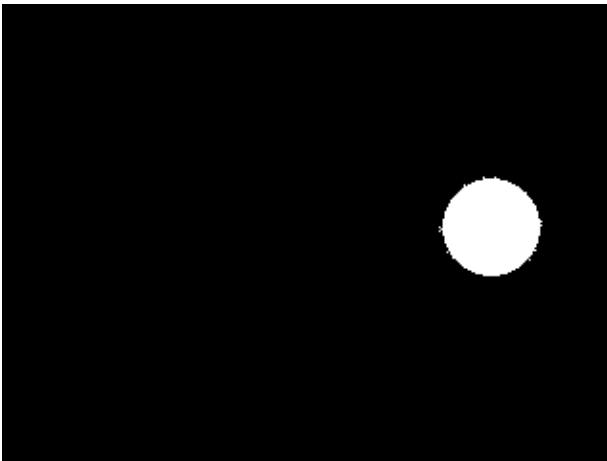
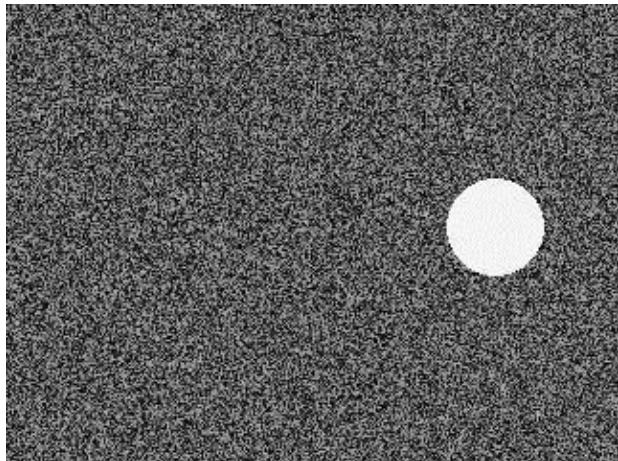
Obtenida a partir del segundo frame, reemplazando sus valores en los píxeles que tienen un nivel alto de la DI Positiva, por los que tienen esos píxeles en el primer frame.

## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.5.- Métodos de detección de movimiento

#### APROXIMACIÓN BÁSICA: Imagen de Referencia – Sustracción de fondo (elementos estáticos)

EJEMPLO: Obtención de imagen de fondo estático a partir de dos frames con elementos en movimiento más claros que el fondo:



**FRAMES TOMADOS EN DOS INSTANTES DE TIEMPO**

**DIs NEGATIVA Y POSITIVA**

#### **IMAGEN DE FONDO:**

Obtenida a partir del segundo frame, reemplazando sus valores en los píxeles que tienen un nivel alto de la DI Positiva, por los que tienen esos píxeles en el primer frame.

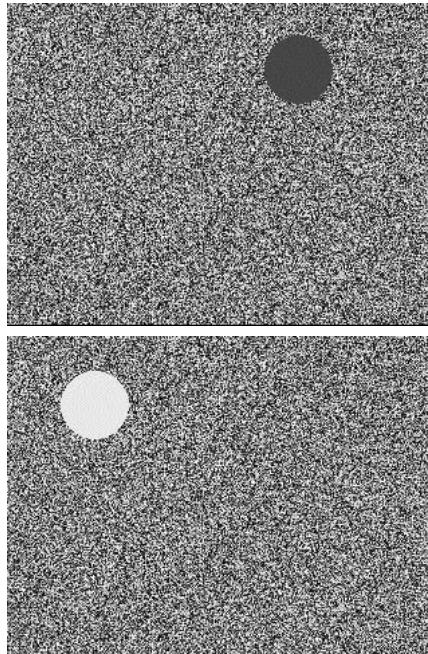
## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.5.- Métodos de detección de movimiento

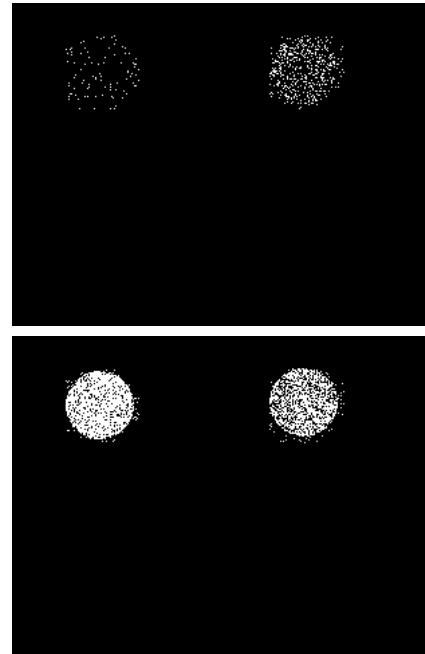
#### APROXIMACIÓN BÁSICA: Imagen de Referencia – Sustracción de fondo (elementos estáticos)

**LIMITACIÓN:** se asume que los objetos en movimiento son más claros u oscuros que el fondo.

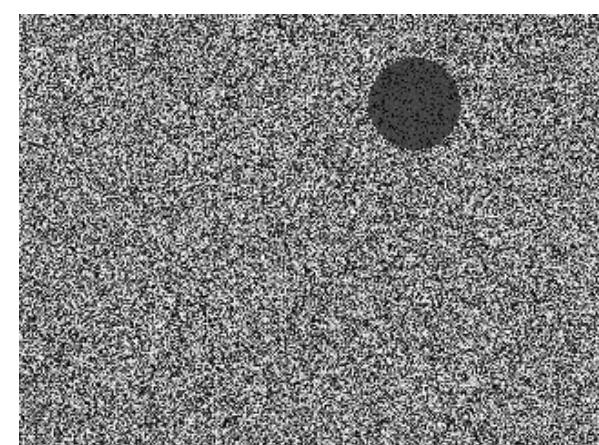
- En los dos últimos ejemplos, la obtención de la imagen de fondo podía llevarse a cabo porque se conocía que los valores de intensidad del objeto en movimiento eran más claros o más oscuros que el fondo.
- De esa forma, el contenido de la imagen diferencia positiva y negativa podía ser asociado e identificado con el objeto en movimiento en cada frame: la imagen positiva señala el objeto en el segundo frame mientras que la negativa lo hace en el primer frame.
- Esta suposición no puede asumirse en muchas aplicaciones reales, ya que el fondo estático puede presentar gran variabilidad en sus valores de intensidad.



**FRAMES DOS INSTANTES**



**DIs NEGATIVA Y POSITIVA**



**IMAGEN DE FONDO:**

Obtenida a partir del segundo frame, reemplazando sus valores en los píxeles que tienen un nivel alto de la DI Positiva, por los que tienen esos píxeles en el primer frame.

## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.5.- Métodos de detección de movimiento

#### APROXIMACIÓN BÁSICA: Imagen de Referencia – Sustracción de fondo (elementos estáticos)

##### TÉCNICAS ALTERNATIVAS PARA OBTENER LA IMAGEN DE REFERENCIA:

- Por ejemplo, la mediana de un número suficiente de frames puede ser una buena estimación del fondo.

10 FRAMES DE LA ESCENA

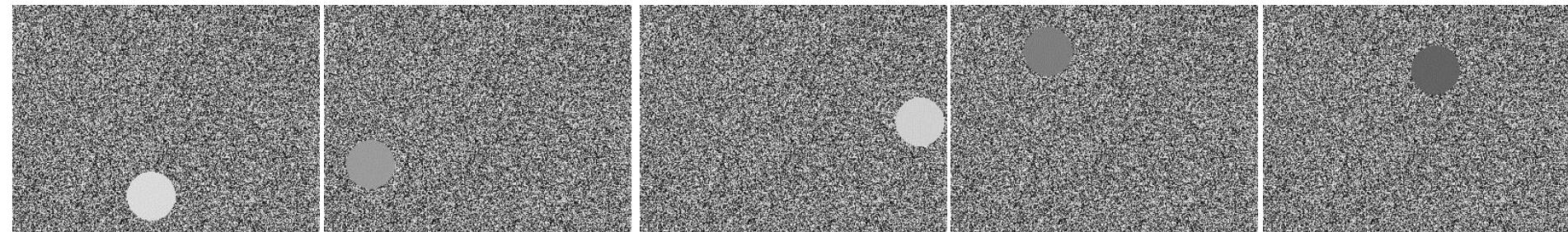
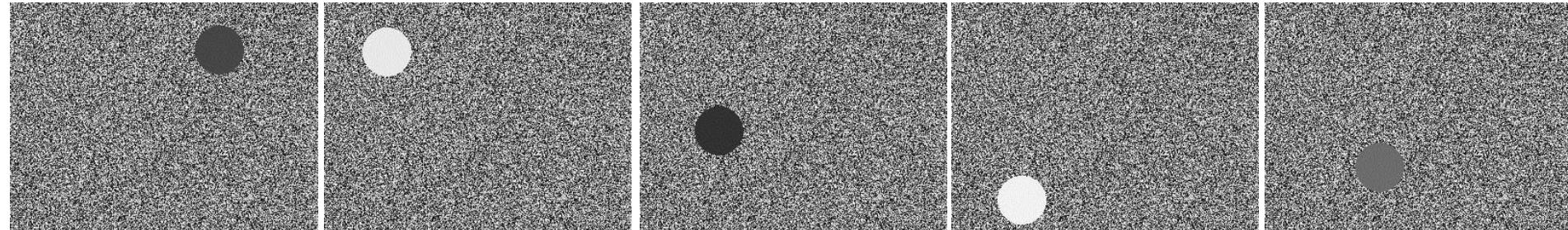
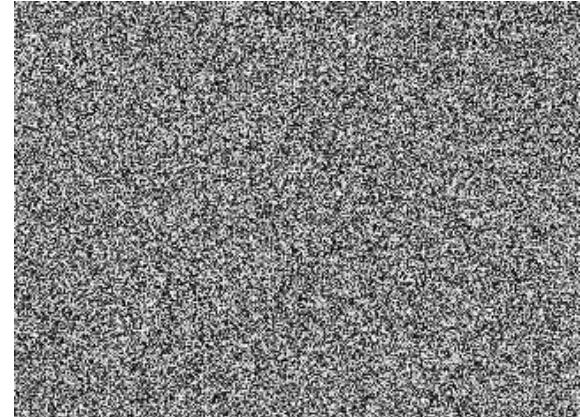


IMAGEN DE FONDO:

- A cada píxel se le asigna el valor de la mediana de todos los valores de intensidad que tiene el píxel en cuestión en los 10 frames considerados.



## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.5.- Métodos de detección de movimiento

#### APROXIMACIÓN BÁSICA: Imagen de Referencia – Sustracción de fondo (elementos estáticos)

#### TÉCNICAS ALTERNATIVAS PARA OBTENER LA IMAGEN DE REFERENCIA:

- EJEMPLO DE OBTENCIÓN DE IMAGEN DE FONDO POR MEDIANA EN CÁMARAS DE TRÁFICO DE SEVILLA – Imágenes cada 2-3 minutos (se muestran 4 imágenes – la mediana se ha obtenido con 10).

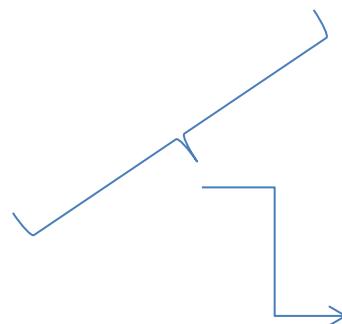


IMAGEN DE FONDO POR  
MEDIANA



## TEMA 3 – FUNDAMENTOS DE LA VISIÓN ARTIFICIAL

### 3.5.- Métodos de detección de movimiento

#### APROXIMACIÓN BÁSICA: Imagen de Referencia – Sustracción de fondo (elementos estáticos)

#### TÉCNICAS ALTERNATIVAS PARA OBTENER LA IMAGEN DE REFERENCIA:

- EJEMPLO DE OBTENCIÓN DE IMAGEN DIFERENCIAS ABSOLUTAS RESPECTO IMAGEN DE FONDO

