

INTRODUCCIÓN A LA DETECCIÓN DE OBJETOS

Template matching

Maria Vanrell

DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

... hasta ahora

Descriptores de píxeles (COLOR)



Conclusión: Los descriptores basados en la información del píxel son muy **sensibles a cambios** y sobre todo **no consideran la información espacial** que es relevante en la descripción de un objeto.

Template Matching, la imagen como descriptor

Si continuamos con el **detector de manos**

Imagen de entrada:

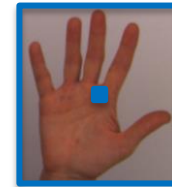


Objetivo:



Descriptor total: Una mano es una imagen, un patrón (*template*), como este

Template



En resumen

Descriptor simple:

EL COLOR DEL PÍXEL

Descriptor total:

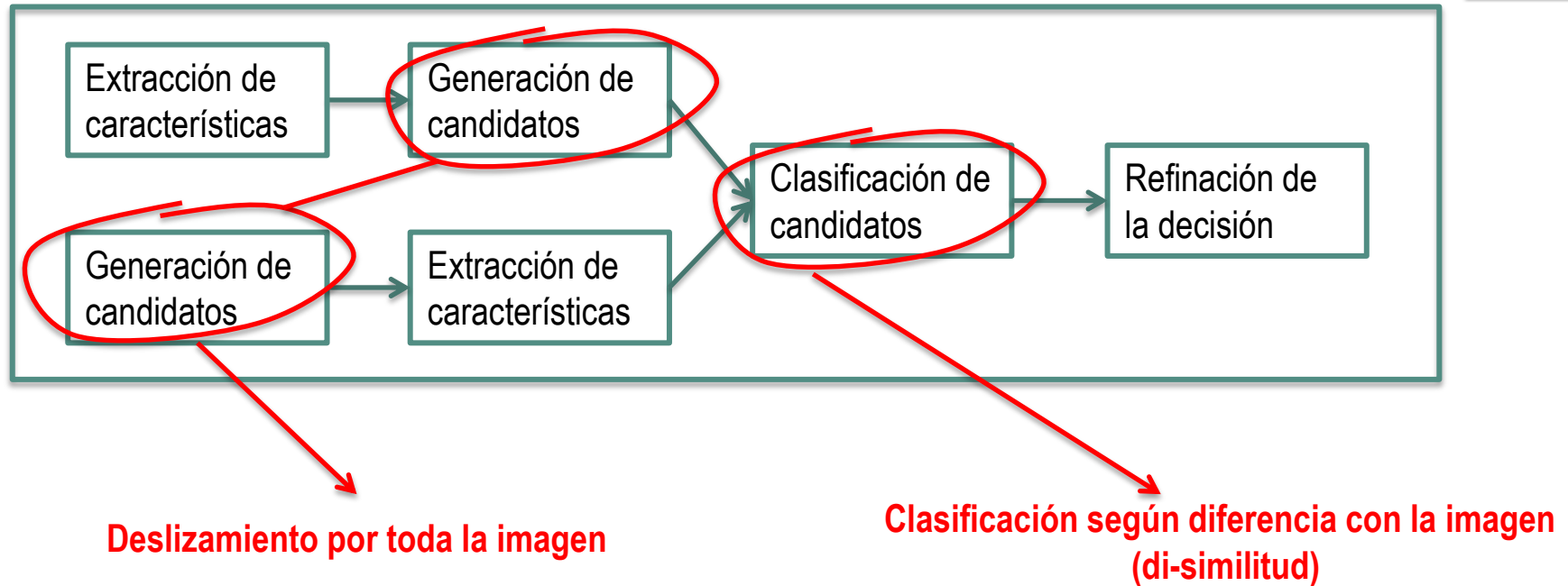
LA IMAGEN DEL OBJETO

Template matching (correspondencia de patrón). Es una técnica para detectar una parte de una imagen, aquella ventana que más se parece a un template o patrón dado. Este patrón representa el objeto que queremos detectar.

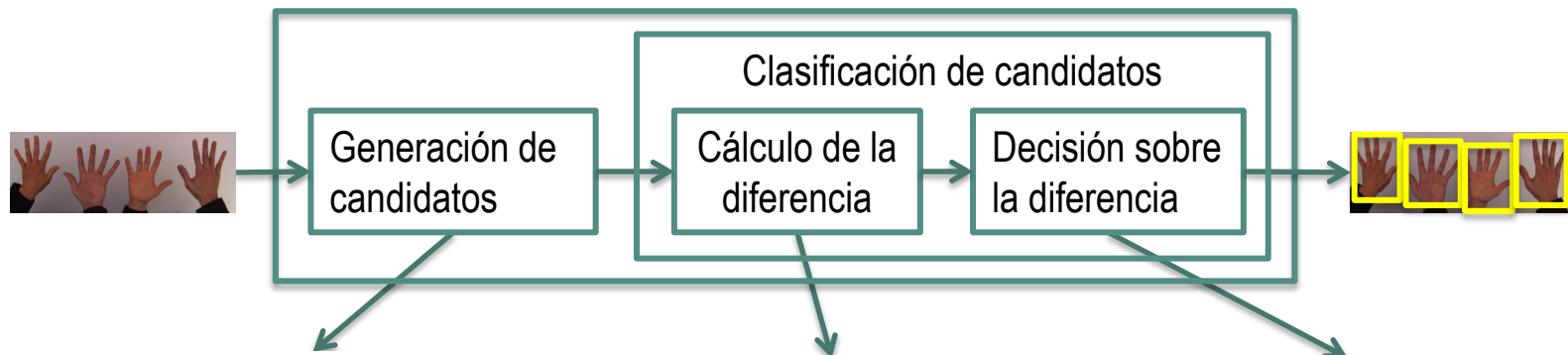
Algoritmo básico:

- **Cálculo de la diferencia.** Para cada punto de la imagen se asigna un valor que representa la semejanza entre el patrón y la ventana correspondiente a cada punto y que tiene tamaño igual al patrón
- **Extracción de mínimos.** El punto de la imagen que presente el valor mínimo de las diferencias nos dará la localización del objeto encontrado en la imagen. Si queremos detectar varias apariciones, otros mínimos locales que aparezcan en el mapa de distancia.

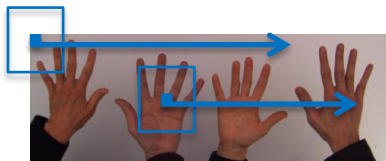
Esquema general de un sistema detector de objetos



Esquema simplificado:



Para cada píxel de la imagen se desplaza la ventana del *template*



Cálculo de la diferencia entre el patrón que buscamos y la ventana de la imagen



—

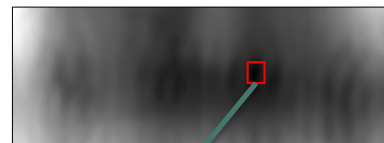


= Diferencia

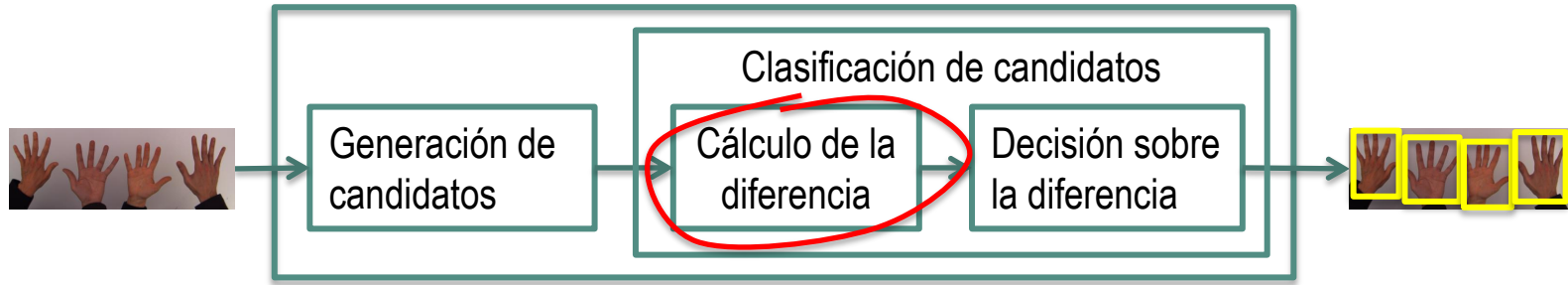
Template

Ventana Imagen

Extracción de mínimos sobre las diferencias calculadas



Diferencia mínima → Mayor similitud



Dos maneras de estimar la semejanza entre el template y cada ventana de la Imagen:



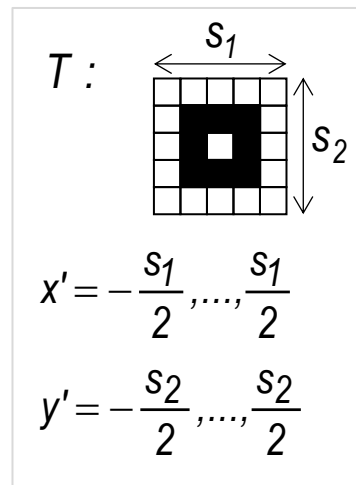
- Cálculo de las diferencias (**DISIMILITUD**)
- Cálculo de similitudes (**SIMILITUD**)

Template matching basado en diferencia

$T(x, y)$: Template de tamaño $s_1 \times s_2$

$I(x, y)$: Imagen

$$TM_{I,T}(x, y) = \sum_{x', y'} (T(x', y') - I(x + x', y + y'))^2$$



Ejemplo:

Template:

$T(x, y)$



Imagen:

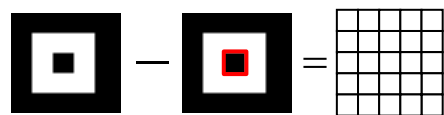
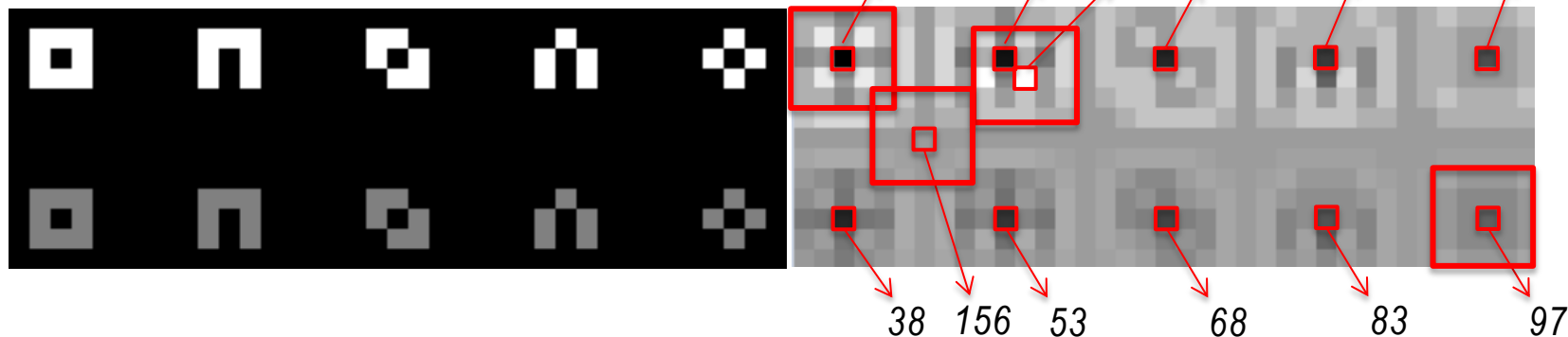
$I(x, y)$



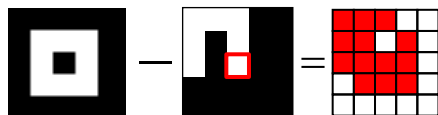
Negro: 0
Blanco: 255

Ejemplo: Cálculo de diferencias por cada ventana

$$TM_{I,T}(x,y) = \sum_{x',y'} (T(x',y') - I(x+x',y+y'))^2$$



$$TM(x,y) = 0$$



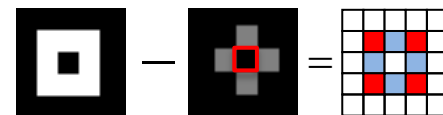
$$TM(x,y) = 255$$

$$13 * (255)^2$$



$$TM(x,y) = 156$$

$$8 * (255)^2$$



$$TM(x,y) = 97$$

$$4 * (255)^2 + 4 * (128)^2$$

Ejemplo: con una escena real

Template:



Imagen



Otros mínimos locales

Diferencias



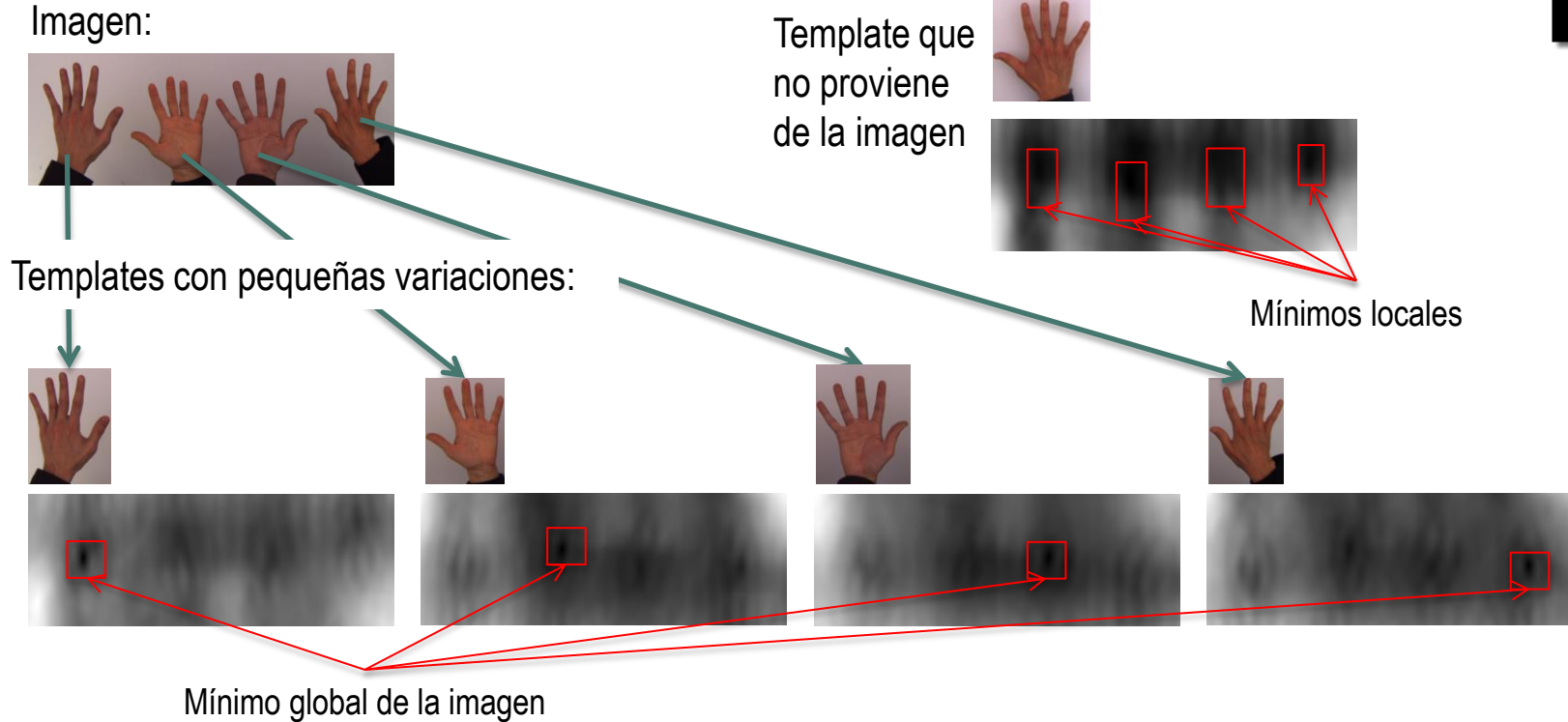
Mínimo global de la imagen

Otras detecciones

Detección



Ejemplo: variaciones entre la imagen y el template



Ejemplo: cuando la imagen presenta variaciones de intensidad

Imagen:



Template de la imagen:



Template que no proviene de la imagen:



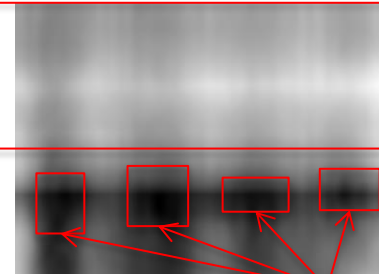
Mínimo global de
la imagen



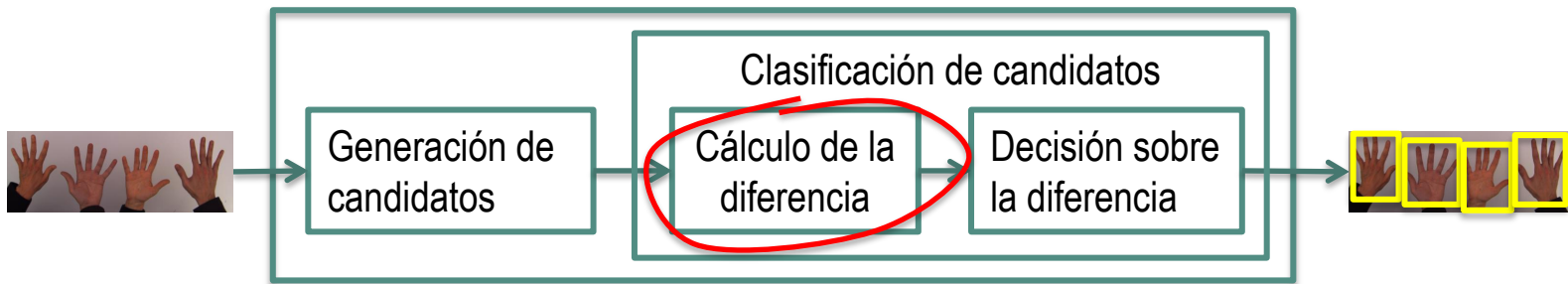
Mínimo global de
la imagen



Pocos mínimos
importantes en
la zona oscura
de la imagen

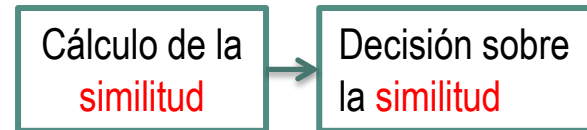


Zonas de mínimos locales



Dos maneras de estimar la semejanza entre el template y cada ventanas de la Imagen:

- ✓ • Cálculo de las diferencias (**DISIMILITUD**)
- ✓ • Cálculo de similitudes (**SIMILITUD**)

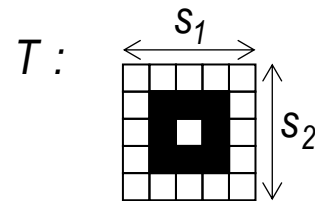


Template matching basado en similitud (correlación normalizada)

$T(x, y)$: Template de tamaño $s_1 \times s_2$

$I(x, y)$: Imagen

$$TM_{I,T}(x, y) = \frac{\sum_{x', y'} (T(x', y') \cdot I(x + x', y + y'))}{\sqrt{\sum_{x', y'} T(x', y')^2 \cdot \sum_{x', y'} I(x + x', y + y')^2}}$$



$$x' = -\frac{s_1}{2}, \dots, \frac{s_1}{2}$$

$$y' = -\frac{s_2}{2}, \dots, \frac{s_2}{2}$$

Ejemplo:

Template:

$T(x, y)$



Imagen:

$I(x, y)$

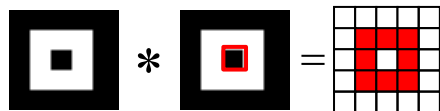
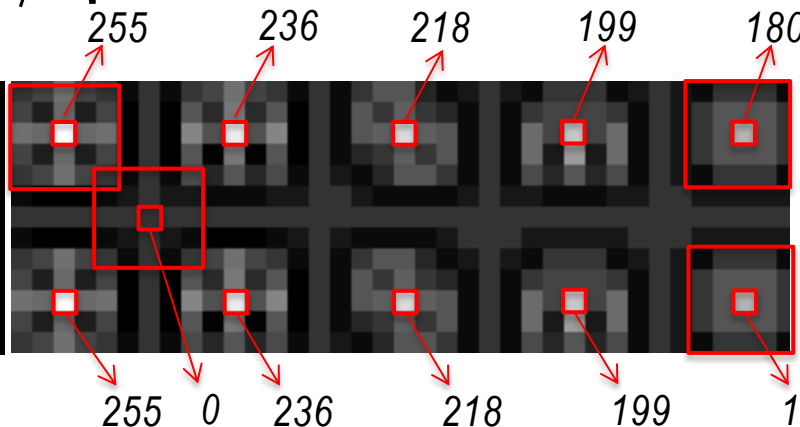


Ejemplo: Cálculo de SIMILITUDES por cada ventana

$$TM_{I,T}(x,y) = \sum_{x',y'} (T(x',y') \cdot I(x+x',y+y')) / \sqrt{\sum_{x',y'} T(x',y')^2 \cdot \sum_{x',y'} I(x+x',y+y')^2}$$

UAB

Universitat Autònoma
de Barcelona

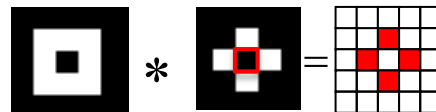


$$TM(x,y) = 255$$

$$8(255 \cdot 255) / \sqrt{8 \cdot 255^2 \cdot 8 \cdot 255^2} = 1$$

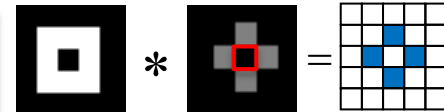


$$TM(x,y) = 0$$



$$TM(x,y) = 180$$

$$4(255 \cdot 255) / \sqrt{8 \cdot 255^2 \cdot 4 \cdot 255^2} = \frac{4}{\sqrt{32}}$$



$$TM(x,y) = 180$$

$$4(255 \cdot 128) / \sqrt{8 \cdot 255^2 \cdot 4 \cdot 128^2} = \frac{4}{\sqrt{32}}$$

Ejemplo: con una escena real

Template:



Imagen



Otros mínimos locales

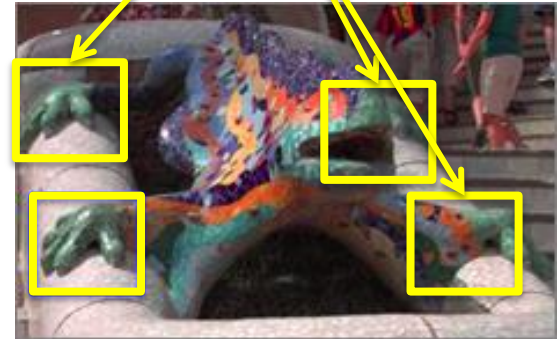
Diferencias



Máximo global de la imagen

Otras detecciones

Detección

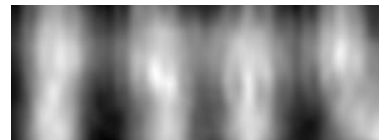


Ejemplo: variaciones entre la imagen y el template

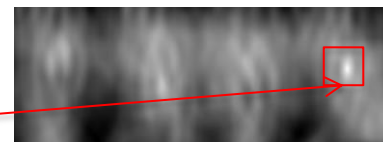
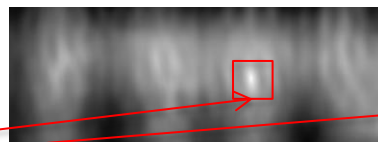
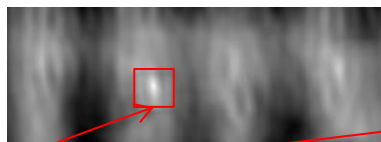
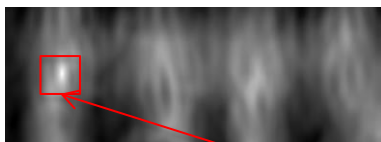
Imagen:



Template que no proviene de la imagen



Templates con pequeñas variaciones:



Máximo global de la imagen

Ejemplo: cuando la imagen presenta variaciones de intensidad

Imagen:



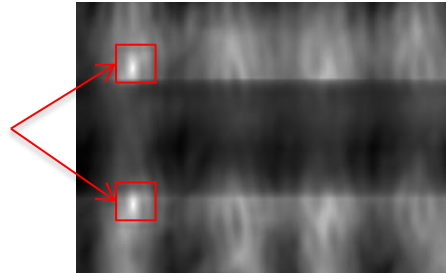
Template de la imagen:



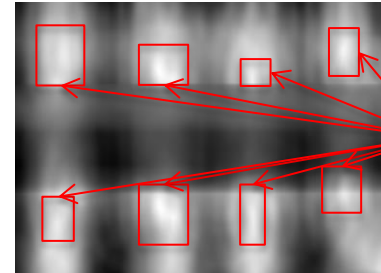
Template que no proviene de la imagen:



Máximo global
de la imagen



Zonas de máximos
locales



En resumen:

- Se ha introducido un método muy básico de detección de objetos: el Template Matching.
- La imagen del objeto como descriptor que nos ha permitido trabajar con la forma del objeto.
- Hemos visto que la diferencia es una medida de disimilitud que no es invariante a la intensidad.
- En cambio la correlación normalizada, como medida de similitud, nos permite detectar objetos independientemente de variaciones de la intensidad.