

# Procesamiento de Imágenes

## Representación y Descripción

Gonzalo Sad  
gonzalosad@gmail.com

# Representación y Descriptores

Definir la forma de un objeto puede resultar difícil...

**Forma** = Figura exterior (o geometría) de un cuerpo u objeto.

Algunos descriptores:

- Área
- Perímetro
- Diámetro
- Distancias: máxima y mínima al centro de masas
- Ejes mayor y menor, ángulo del eje mayor
- Envolvente (bounding box)





# Esquemas de representación

Propiedades deseables:

1. **Unicidad:** cada objeto debe tener una única representación.
2. **Invariancia** frente a transformaciones geométricas, como traslaciones, rotaciones, cambios de escala y reflexiones.
3. **Sensibilidad** o capacidad para diferenciar objetos casi iguales.
4. **Abstracción del detalle** o capacidad para representar los rasgos característicos básicos de los objetos y abstraer los detalles.



# Esquemas de representación

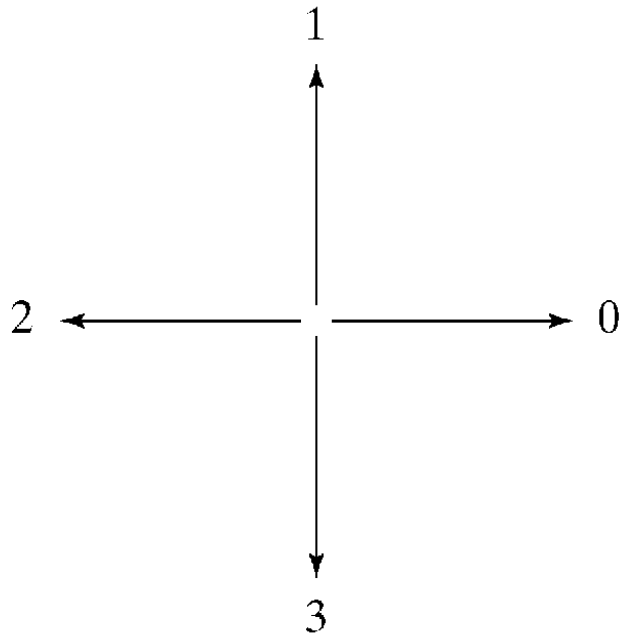
- **Esquemas de representación externa**, que utilizan el contorno de los objetos y sus rasgos característicos, como son los códigos de cadena, los descriptores de Fourier y las aproximaciones poligonales.
- **Esquemas de representación interna**, que describen la región ocupada por el objeto en la imagen binaria, como son el área, los momentos, el esqueleto.



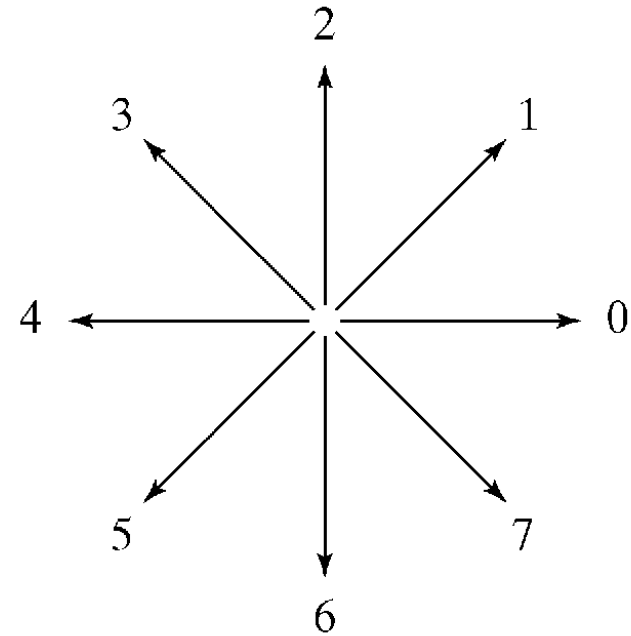
# Chain Codes

- Es una de las técnicas más antiguas, usadas actualmente.
- Es un descriptor de contornos de un objeto o región.
- Trabaja sobre imágenes binarias.
- Se almacena la posición relativa de un pixel, con respecto al adyacente, con lo que se obtiene una cadena o vector de números.
- Existen dos maneras de realizar la conexión de los pixeles para formar la cadena.

# Chain Codes

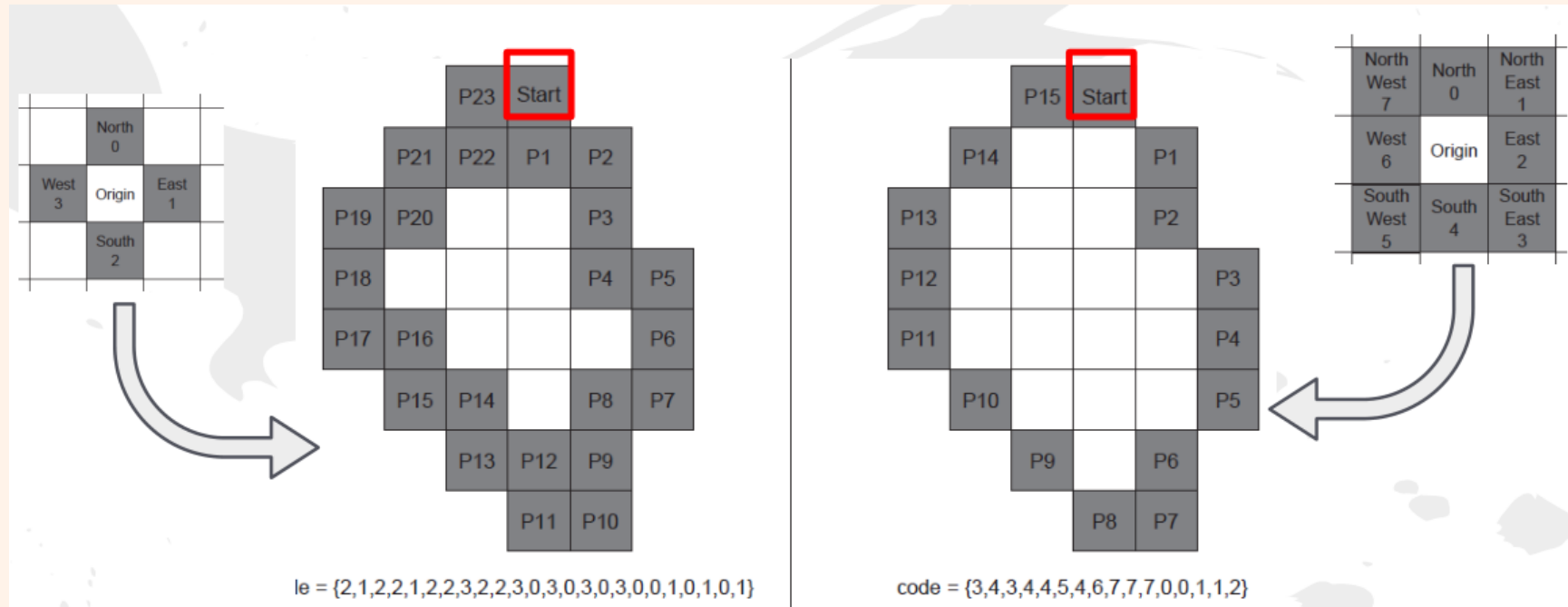


4-vecinos



8-vecinos

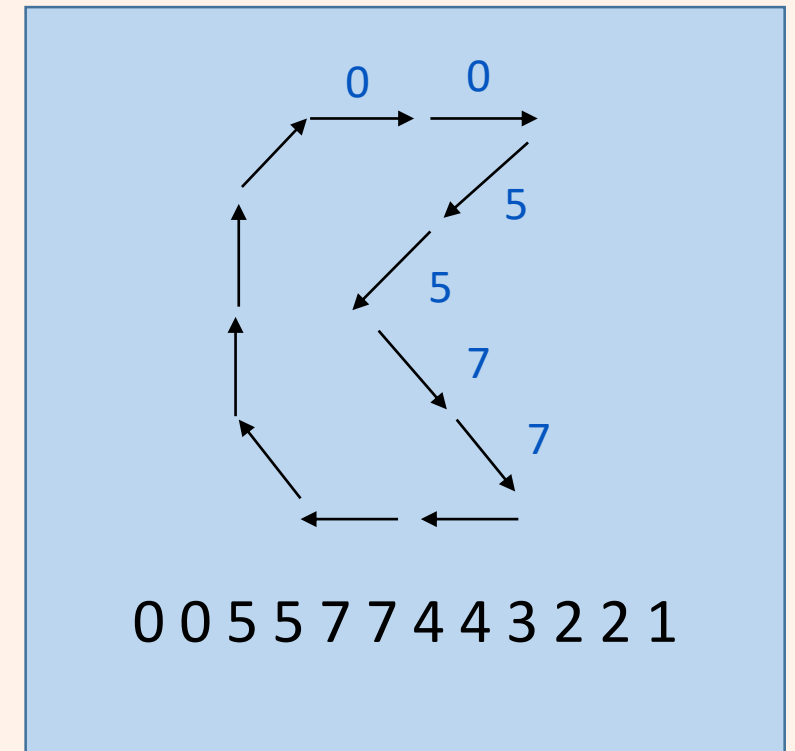
# Chain Codes



# Chain Codes

Se puede hacer invariante a la posición inicial:

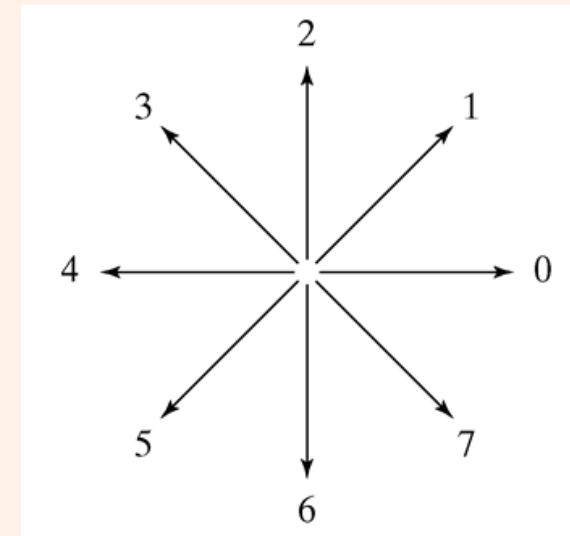
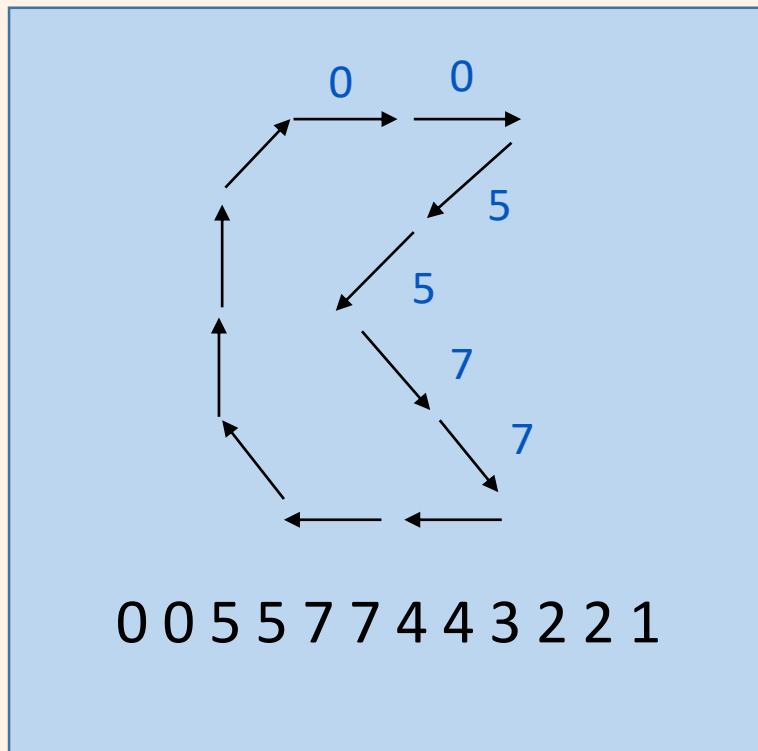
- Secuencia circular.
- Elegir la posición inicial cómo el código que forme el entero de menor magnitud.





# Chain Codes

Se puede hacer invariante a la rotación tomando la diferencia del código:



Diferencia:  
7 0 5 0 2 0 5 0 7 7 0 7



# Chain Codes

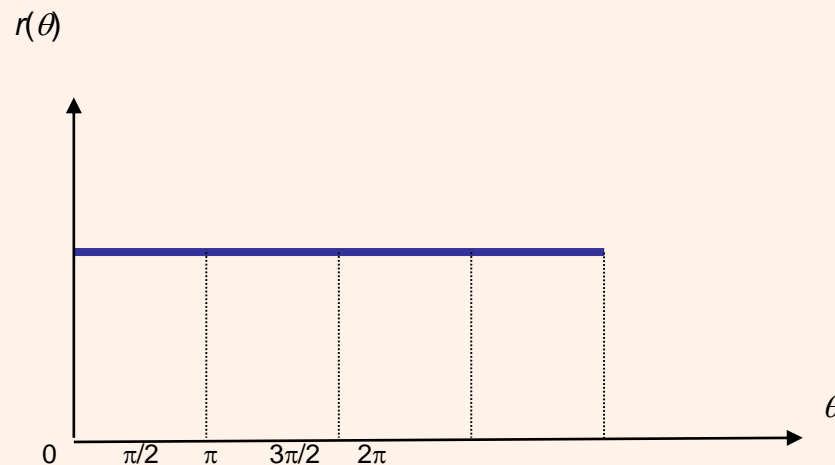
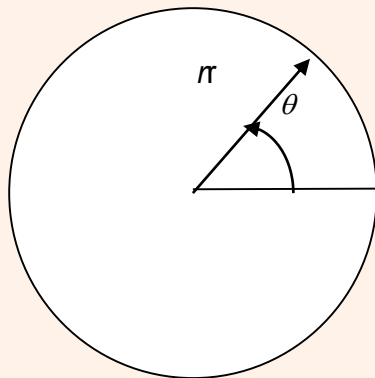
Esta representación se mantiene invariante frente a traslaciones, lo que facilita la comparación de objetos.

- A partir del resultado que éste entrega, se pueden obtener datos del contorno como:
  - Perímetro
  - Área
  - Descriptores de Fourier
- Cualquier ruido o perturbación en la imagen, puede inducir a errores.
- La cadena obtenida puede llegar a ser demasiado larga en objetos grandes.

# Signatures

Una **signature** es una representación de un contorno mediante una función real unidimensional que sea más sencilla que la función bidimensional que define el contorno.

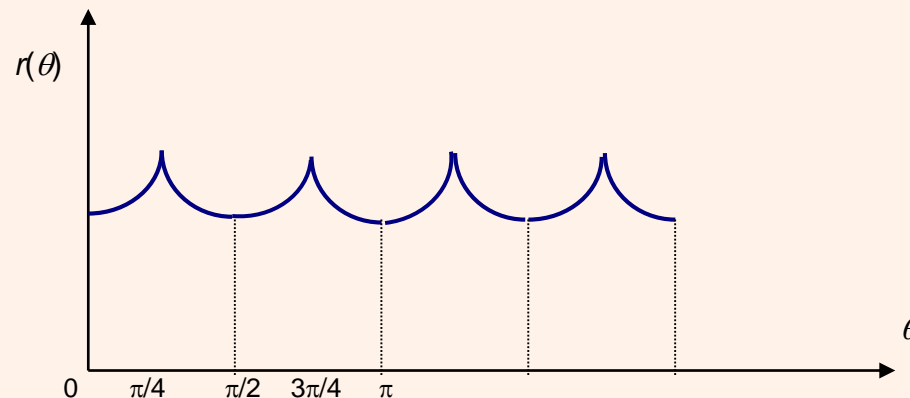
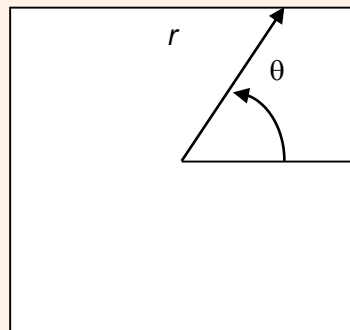
Hay varias maneras de definir una firma. Una de las más simples es a través de la distancia desde un punto interior, como puede ser el centroide del contorno, a cada uno de los puntos del contorno como una función del ángulo



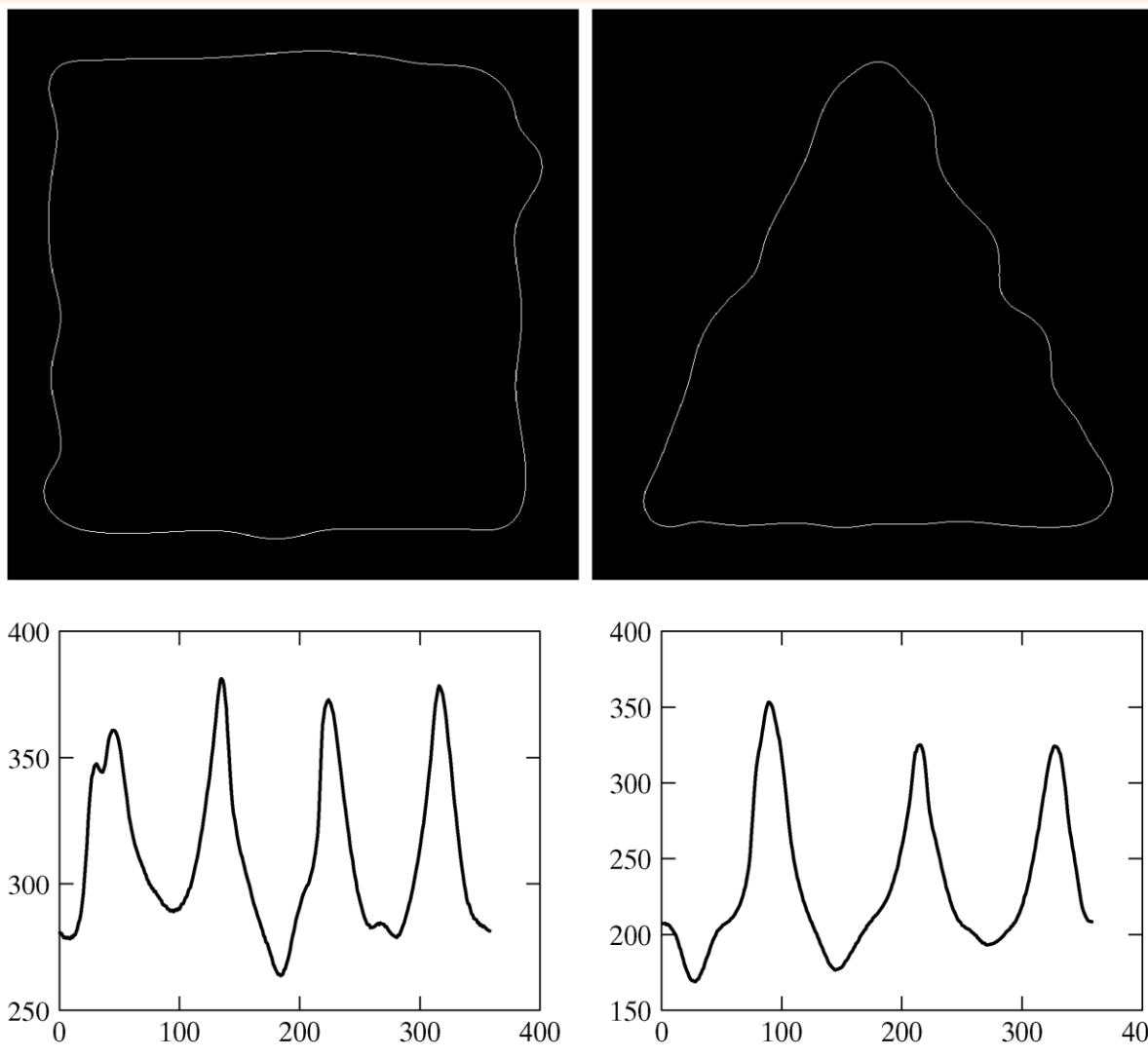
# Signatures

La firma es **invariante frente a traslaciones** pero no lo es frente a rotaciones o cambios de escala.

Sin embargo, se puede conseguir la **invariancia frente a rotaciones** cuando se encuentra un punto característico del contorno a partir del cual se comienza a generar la firma. Dicho punto puede ser, por ejemplo, el más cercano al centroide, siempre que sea único, o un punto del contorno determinado por la intersección de este con su eje mayor.

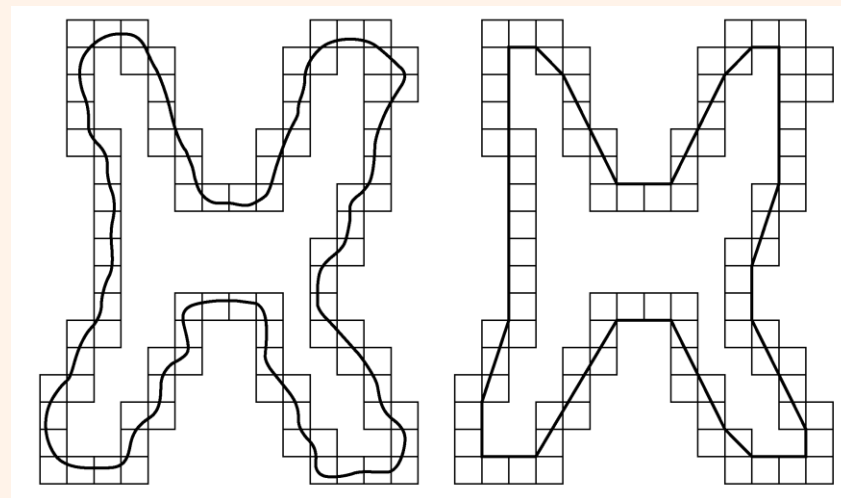
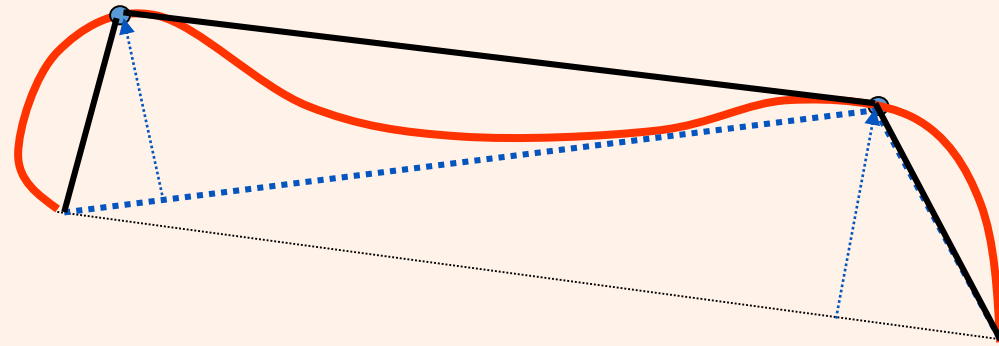


# Signatures



# Ajuste paramétrico

- Segmentos lineales
- Arcos
- Curvas
- Splines



# Descriptores de Fourier

- La representación de contornos dados por una curva cerrada puede pensarse como una secuencia periódica.
- Chain codes → secuencia real
- Puntos del contorno → secuencia compleja

$$s(k) = x(k) + jy(k) \quad k = 0, 1, \dots, N-1$$

## Descriptores de Fourier

DFT →  $a(u) = \sum_{k=0}^{N-1} s(k) e^{-j2\pi uk/N}$

IDFT →  $s(k) = \frac{1}{N} \sum_{u=0}^{N-1} a(u) e^{j2\pi uk/N}$

# Descriptores de Fourier

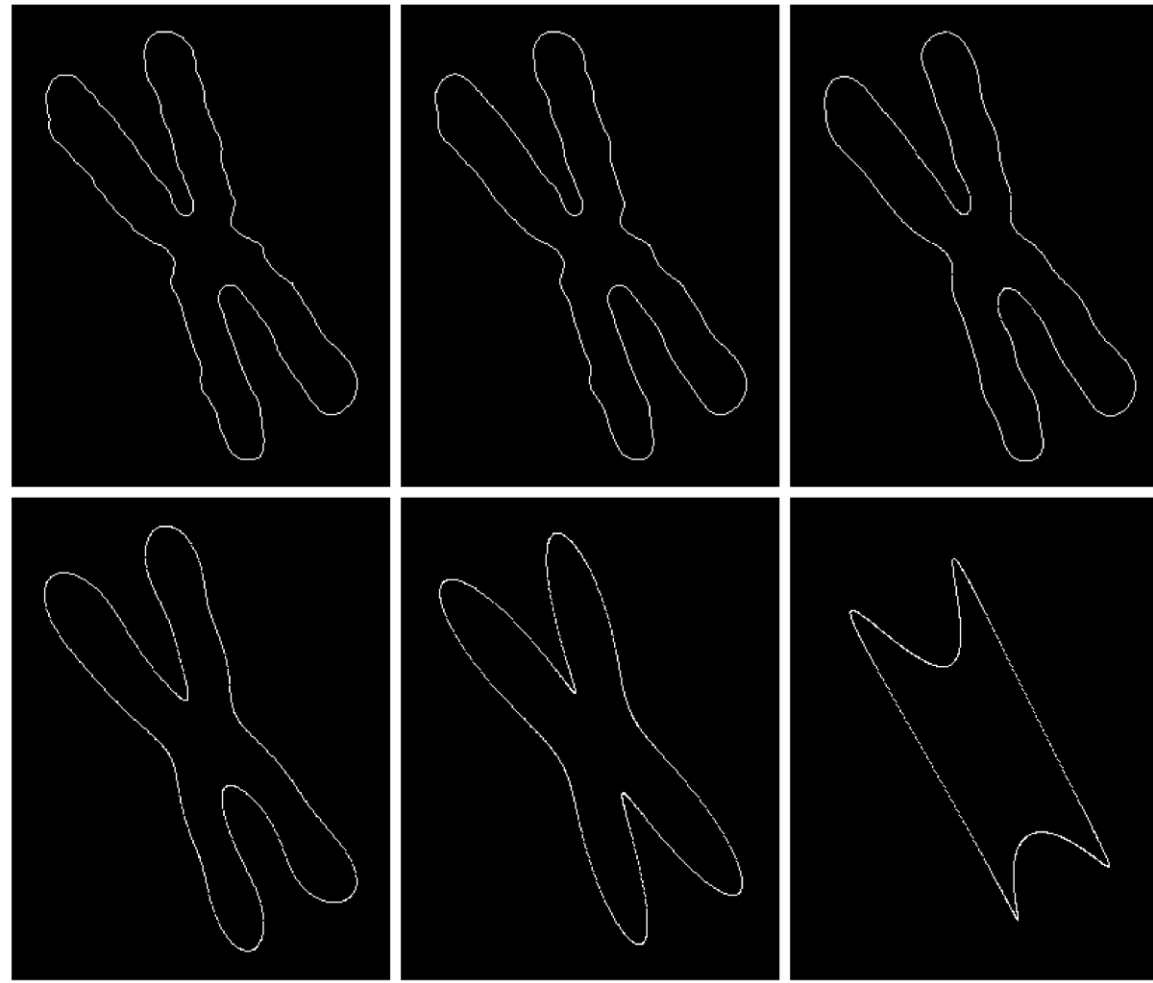
Usualmente se calcula un número de descriptores de Fourier  $P < N$ , por lo que resulta la siguiente aproximación de  $s(k)$ :

$$\hat{s}(k) = \frac{1}{P} \sum_{u=0}^{P-1} a(u) e^{j2\pi uk/N}, \quad k = 0, 1, \dots, N-1$$

Las componentes de **alta frecuencia** tienen en cuenta los **detalles más finos** del contorno, en tanto que las de **baja frecuencia** representan la **forma global** del mismo.



# Descriptores de Fourier

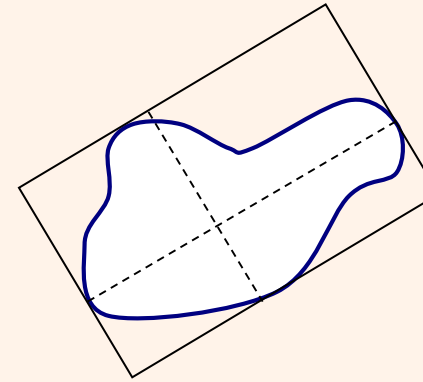


a	b	c
d	e	f

**FIGURE 11.17** (a)–(f) Boundary reconstructed using 546, 110, 56, 28, 14, and 8 Fourier descriptors out of a possible 1090 descriptors.

# Descriptores geométricos

Estudiar la forma geométrica de los contornos de las regiones (objetos).



**Descriptores:** valoraciones numéricas que nos van a permitir identificar y reconocer los objetos de dicha imagen

- El **perímetro** viene dado por el número total de píxeles que configuran su contorno pero los píxeles de bordes diagonales se ponderan con raíz de dos.
- El **diámetro** de un contorno viene dado por la distancia Euclídea entre los dos píxeles del contorno más alejados. La recta que pasan por dichos puntos se llama eje mayor de la región.

# Descriptores geométricos

- El **área** se trata del número de píxeles que componen la región.
- El **rectángulo base**, con dos lados paralelos al eje mayor, que tiene la propiedad de que es el menor rectángulo que contiene al contorno
- El cociente entre la longitud del lado mayor y la longitud del lado menor se llama **excentricidad** del contorno.
- El **centro de gravedad** o **centroide** de un contorno determinado por el conjunto de píxeles  $\{(x_i, y_i), i=1,2,\dots,N\}$

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N}$$

$$\bar{y} = \frac{\sum_{i=1}^N y_i}{N}$$

- El **eje menor** del contorno viene definido por la recta perpendicular al eje mayor que pasa por el centro de gravedad del contorno.



# Descriptores geométricos

Todos los parámetros anteriores son **invariantes frente a traslaciones** pero no lo son frente a transformaciones de escala.

La **curvatura** se define como la tasa de cambio de la pendiente (tangente) del contorno, pero es difícil de obtener medidas fiables en una imagen digital porque los bordes suelen presentar “muescas”. Sin embargo, se pueden obtener descriptores de la curvatura bastante útiles mediante diferencia de las pendientes de segmentos adyacentes del contorno.

# Descriptores geométricos

La topología es el estudio de configuraciones geométricas con propiedades específicas como la invariancia bajo ciertas transformaciones (cambios de escala).

La **compacidad** (circularidad) es un parámetro que no depende del tamaño de la región y viene dado por el cociente entre el área y el perímetro al cuadrado

$$c = \frac{A}{p^2} \longrightarrow c = 4\pi \frac{A}{p^2}$$

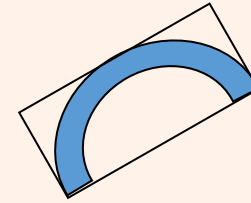
Círculo:  $c = 1/(4\pi)$  ( $\approx 0.07957$ )

Triángulo equilátero vale  $1/(12)$  ( $\approx 0.048$ )

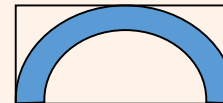
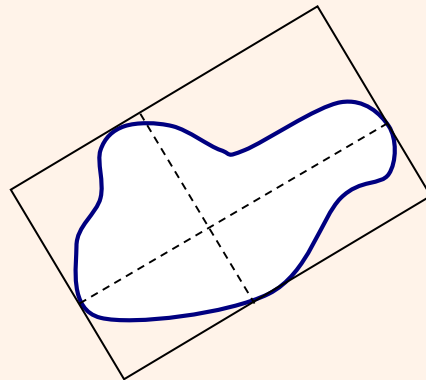
# Descriptores geométricos

La **rectangularidad** de una región se define como el cociente entre el área de la región y el área de su rectángulo base.

$$c = \frac{\text{Área de la región}}{\text{Área de su rectángulo base}}$$



El **alargamiento** de una región se puede definir por el cociente entre la longitud del lado mayor y el lado menor de su rectángulo base.



Sin embargo, no es una medida adecuada para regiones curvadas.

# Descriptores topológicos

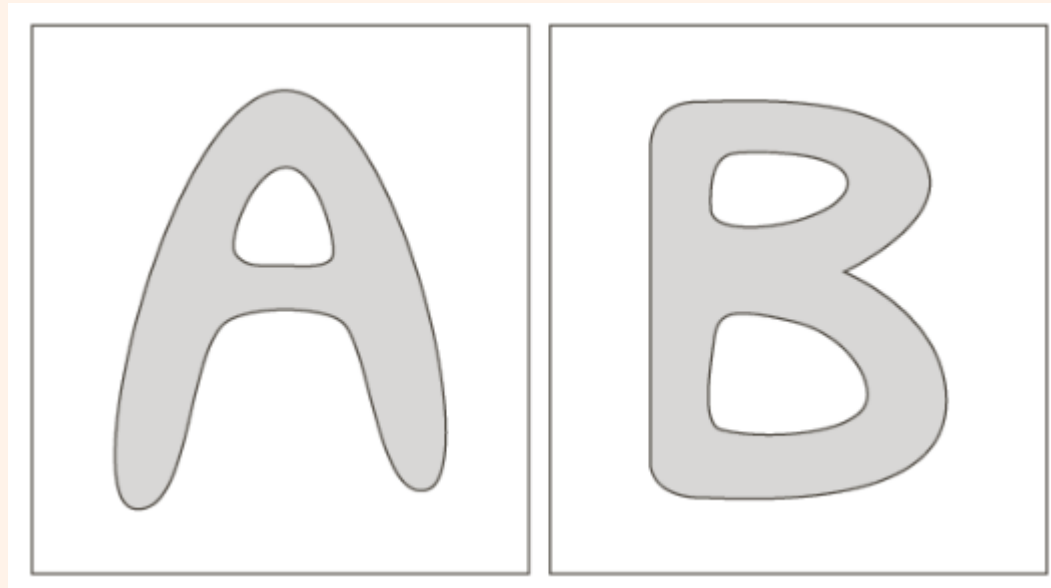
Otras características topológicas importantes son la **conectividad** y los **huecos** en los objetos.

Una imagen segmentada puede estar compuesta por regiones que tienen **componentes conexas** que configuran los objetos, es decir, regiones tales que dos puntos cualesquiera de ellas se pueden unir por una curva contenida en ellas.

Un **hueco** es una región de la imagen que está completamente encerrada por una componente conexa de la imagen.

El **número de Euler** de una imagen se define como:  $E = C - H$ , donde  $C$  es el número de componentes conexas y  $H$  el número de huecos de la imagen. Este número es invariante frente a traslaciones, rotaciones y cambios de escala, y nos permite de forma sencilla discriminar entre ciertas clases de objetos.

# Descriptores topológicos



$E = 0$

$E = -1$



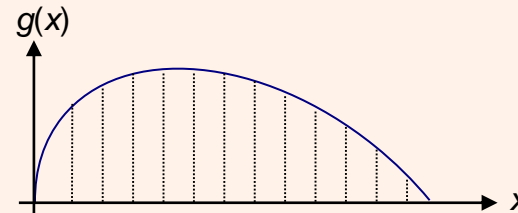
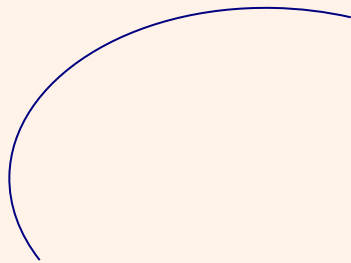
# Descriptores topológicos

'Area'	The number of pixels in a region.
'BoundingBox'	$1 \times 4$ vector defining the smallest rectangle containing a region. BoundingBox is defined by [ul_corner width], where ul_corner is in the form [x y] and specifies the upper-left corner of the bounding box, and width is in the form [x_width y_width] and specifies the width of the bounding box along each dimension. Note that the BoundingBox is aligned with the coordinate axes and, in that sense, is a special case of the basic rectangle discussed in Section 11.3.1.
'Centroid'	$1 \times 2$ vector; the center of mass of the region. The first element of Centroid is the horizontal coordinate (or x-coordinate) of the center of mass, and the second element is the vertical coordinate (or y-coordinate).
'ConvexArea'	Scalar; the number of pixels in 'ConvexImage'.
'ConvexHull'	$p \times 2$ matrix; the smallest convex polygon that can contain the region. Each row of the matrix contains the x- and y-coordinates of one of the $p$ vertices of the polygon.
'ConvexImage'	Binary image; the convex hull, with all pixels within the hull filled in (i.e., set to on). (For pixels that the boundary of the hull passes through, regionprops uses the same logic as roipoly to determine whether the pixel is inside or outside the hull.) The image is the size of the bounding box of the region.
'Eccentricity'	Scalar; the eccentricity of the ellipse that has the same second moments as the region. The eccentricity is the ratio of the distance between the foci of the ellipse and its major axis length. The value is between 0 and 1, with 0 and 1 being degenerate cases (an ellipse whose eccentricity is 0 is a circle, while an ellipse with an eccentricity of 1 is a line segment).
'EquivDiameter'	Scalar; the diameter of a circle with the same area as the region. Computed as $\sqrt{4 \cdot \text{Area} / \pi}$ .
'EulerNumber'	Scalar; equal to the number of objects in the region minus the number of holes in those objects.
'Extent'	Scalar; the proportion of the pixels in the bounding box that are also in the region. Computed as Area divided by the area of the bounding box.
'Extrema'	$8 \times 2$ matrix; the extremal points in the region. Each row of the matrix contains the x- and y-coordinates of one of the points. The format of the vector is [top-left, top-right, right-top, right-bottom, bottom-right, bottom-left, left-bottom, left-top].
'FilledArea'	The number of on pixels in FilledImage.
'FilledImage'	Binary image of the same size as the bounding box of the region. The on pixels correspond to the region, with all holes filled.
'Image'	Binary image of the same size as the bounding box of the region; the on pixels correspond to the region, and all other pixels are off.
'MajorAxisLength'	The length (in pixels) of the major axis <sup>1</sup> of the ellipse that has the same second moments as the region.
'MinorAxisLength'	The length (in pixels) of the minor axis <sup>1</sup> of the ellipse that has the same second moments as the region.
'Orientation'	The angle (in degrees) between the x-axis and the major axis <sup>1</sup> of the ellipse that has the same second moments as the region.
'PixelList'	A matrix whose rows are the [x, y] coordinates of the actual pixels in the region.
'Solidity'	Scalar; the proportion of the pixels in the convex hull that are also in the region. Computed as $\text{Area} / \text{ConvexArea}$ .

# Descriptores estadísticos

**Momentos estadísticos.** La forma de una representación unidimensional de un contorno a través de una función real,  $g(x)$ , se puede describir utilizando momentos estadísticos, como la **media**, la **varianza** o **momentos de orden superior**.

Dicha función puede ser la firma del contorno. En el caso de contornos abiertos, se puede utilizar la función que se obtiene de las distancias de los puntos del contorno al segmento que une los dos puntos extremos de dicho contorno



$$m = \sum_{i=0}^{N-1} x_i g(x_i)$$

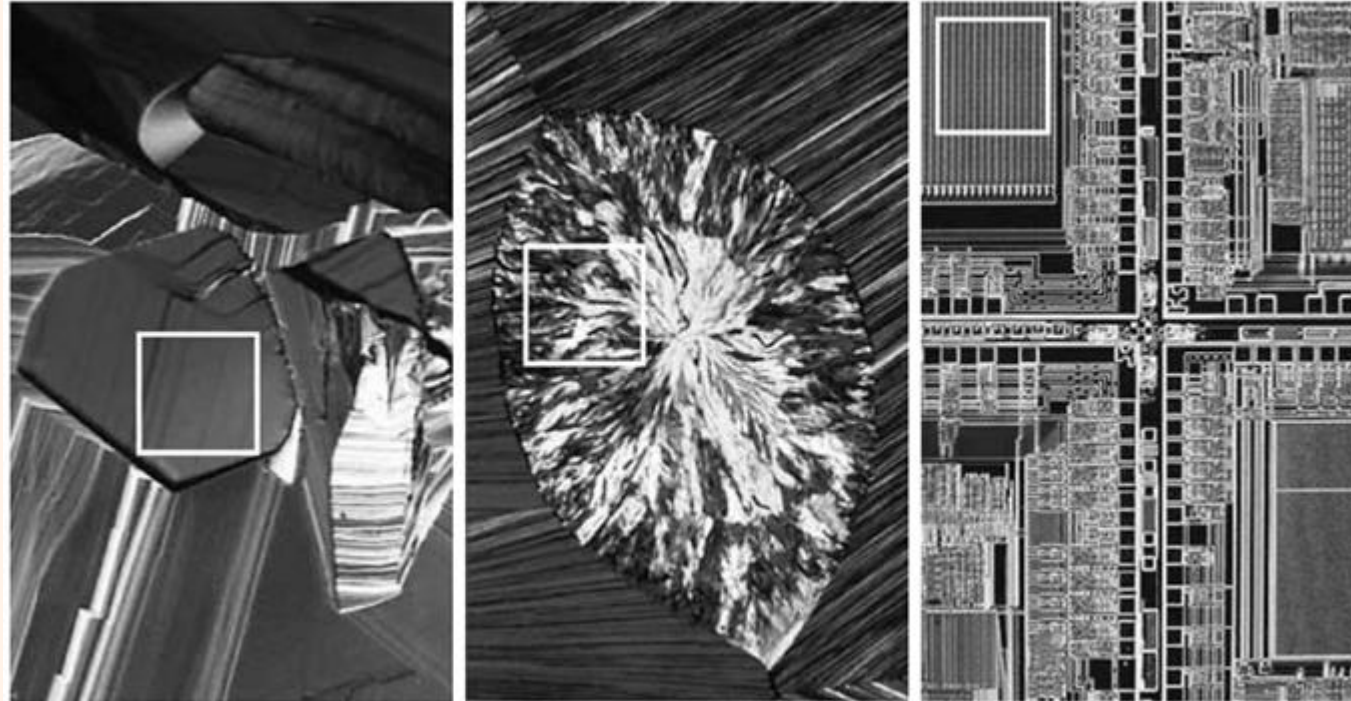
$$\sigma^2 = \sum_{i=0}^{N-1} (x_i - m)^2 g(x_i)$$

$$\mu_3 = \sum_{i=0}^{N-1} (x_i - m)^3 g(x_i)$$

# Descriptores estadísticos

Moment	Expression	Measure of Texture
Mean	$m = \sum_{i=0}^{L-1} z_i p(z_i)$	A measure of average intensity.
Standard deviation	$\sigma = \sqrt{\mu_2(z)} = \sqrt{\sigma^2}$	A measure of average contrast.
Smoothness	$R = 1 - 1/(1 + \sigma^2)$	Measures the relative smoothness of the intensity in a region. $R$ is 0 for a region of constant intensity and approaches 1 for regions with large excursions in the values of its intensity levels. In practice, the variance used in this measure is normalized to the range $[0, 1]$ by dividing it by $(L - 1)^2$ .
Third moment	$\mu_3 = \sum_{i=0}^{L-1} (z_i - m)^3 p(z_i)$	Measures the skewness of a histogram. This measure is 0 for symmetric histograms, positive by histograms skewed to the right (about the mean) and negative for histograms skewed to the left. Values of this measure are brought into a range of values comparable to the other five measures by dividing $\mu_3$ by $(L - 1)^2$ also, which is the same divisor we used to normalize the variance.
Uniformity	$U = \sum_{i=0}^{L-1} p^2(z_i)$	Measures uniformity. This measure is maximum when all gray levels are equal (maximally uniform) and decreases from there.
Entropy	$e = - \sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i)$	A measure of randomness.

# Descriptores estadísticos



Texture	Average Intensity	Average Contrast	R	Third Moment	Uniformity	Entropy
Smooth	87.02	11.17	0.002	-0.011	0.028	5.367
Coarse	119.93	73.89	0.078	2.074	0.005	7.842
Periodic	98.48	33.50	0.017	0.557	0.014	6.517

# Descriptores estadísticos

## Medidas invariantes de momentos

Se define el momento de orden p y q de la imagen digital f(i,j) por la expresión:

$$m_{pq} = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} i^p j^q f(i, j)$$

Si lo calculamos para el objeto determinado por la región S de una imagen binaria vale:

$$m_{pq} = \sum_{(i,j) \in S} i^p j^q$$

Se puede observar que m00 nos da el área del objeto y que (m10 /m00 , m01 /m00 ) es el **centroide** (centro de gravedad) del objeto.

# Descriptores estadísticos

## Medidas invariantes de momentos

Los momentos de orden superior no son invariantes a traslaciones, por ello vamos a realizar una traslación del origen al centroide y obtenemos así los **momentos centrales** de orden p y q mediante la expresión:

$$\mu_{pq} = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (i - \bar{i})^p (j - \bar{j})^q f(i, j)$$

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{1 + \frac{p+q}{2}}}$$



Invariante frente a  
cambios de escala  
para  $p+q = 2, 3, \dots$

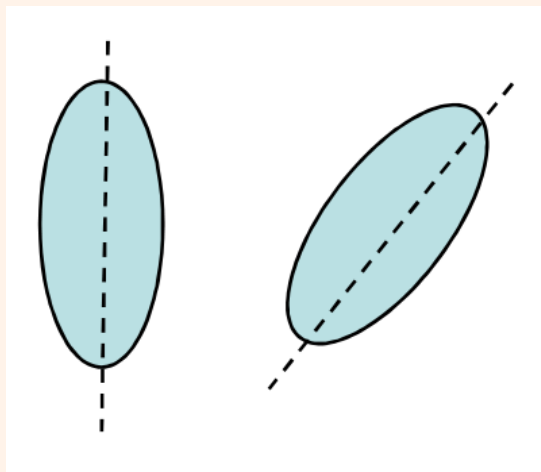


# Descriptores estadísticos

## Medidas invariantes de momentos

$$\mu_{pq} = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (i - \bar{i})^p (j - \bar{j})^q f(i, j)$$

- Son invariantes a la traslación.
- $\mu_{10}$  y  $\mu_{01}$  son cero.
- Los valores de  $\mu_{20}$  y  $\mu_{02}$  aumentan cuanto mayor sea la componente horizontal y vertical de una figura, respectivamente.



La orientación del eje de mínima inercia es:

$$\theta = \frac{1}{2} \arctg \left( \frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right)$$

# Descriptores estadísticos

## Medidas invariantes de momentos

- $\mu_{20}$  es una medida de la **dispersión horizontal** del objeto con respecto al centroide.
- $\mu_{02}$  es una medida de la **dispersión vertical** del objeto con respecto al centroide.
- $\mu_{12}$  es una medida de la **divergencia horizontal**; indica la extensión de la región izquierda del objeto frente a la derecha
- $\mu_{21}$  es una medida de la **divergencia vertical**; indica la extensión de la región inferior del objeto frente a la superior
- $\mu_{30}$  es una medida del **desequilibrio** (o asimetría) horizontal e indica si el objeto tiene mayor extensión a la izquierda o a la derecha del centroide.
- $\mu_{03}$  es una medida del **desequilibrio** (o asimetría) vertical.



# Descriptores estadísticos

## Medidas invariantes de momentos

Un conjunto de seis **invariantes de momentos** que son insensibles a traslaciones, cambios de escala, rotaciones y transformaciones especulares viene dado por las siguientes expresiones:

$$\phi_1 = \eta_{20} + \eta_{02}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

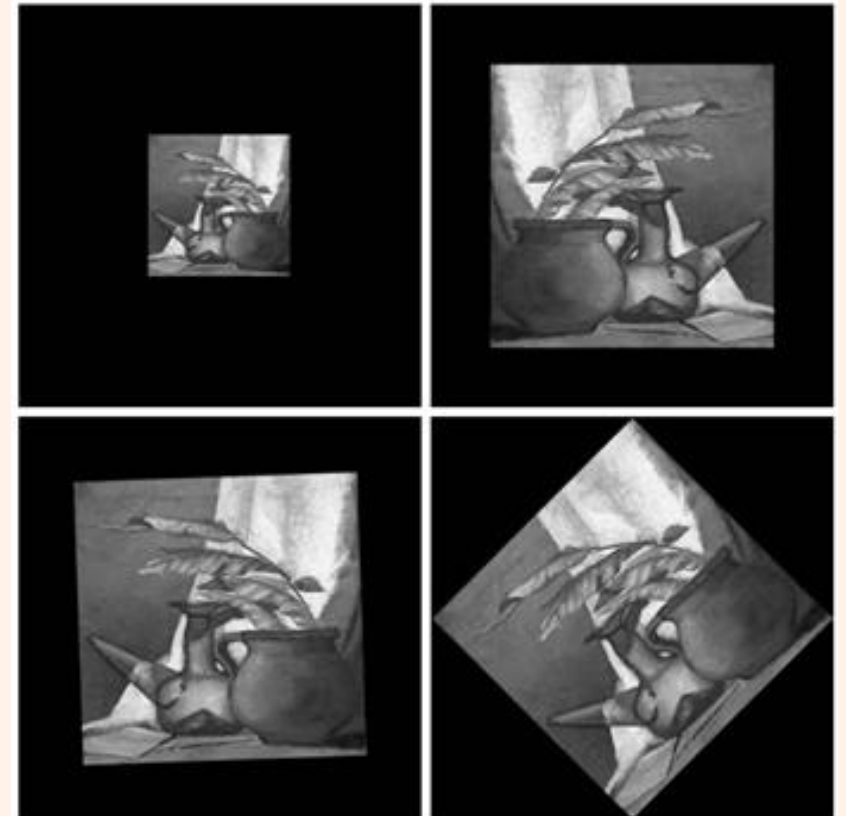
$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) + \left[ (\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})$$

$$\phi_6 = (\eta_{20} - 3\eta_{02}) \left[ (\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] + 4\eta_{11} (\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

# Descriptores estadísticos

## Medidas invariantes de momentos



Invariant ( log )	Original	Half Size	Mirrored	Rotated 2°	Rotated 45°
$\phi_1$	6.600	6.600	6.600	6.600	6.600
$\phi_2$	16.410	16.408	16.410	16.410	16.410
$\phi_3$	23.972	23.958	23.972	23.978	23.973
$\phi_4$	23.888	23.882	23.888	23.888	23.888
$\phi_5$	49.200	49.258	49.200	49.200	49.198
$\phi_6$	32.102	32.094	32.102	32.102	32.102
$\phi_7$	47.953	47.933	47.850	47.953	47.954

# Características locales

Patrón que difiere a su cercanía inmediata en la imagen

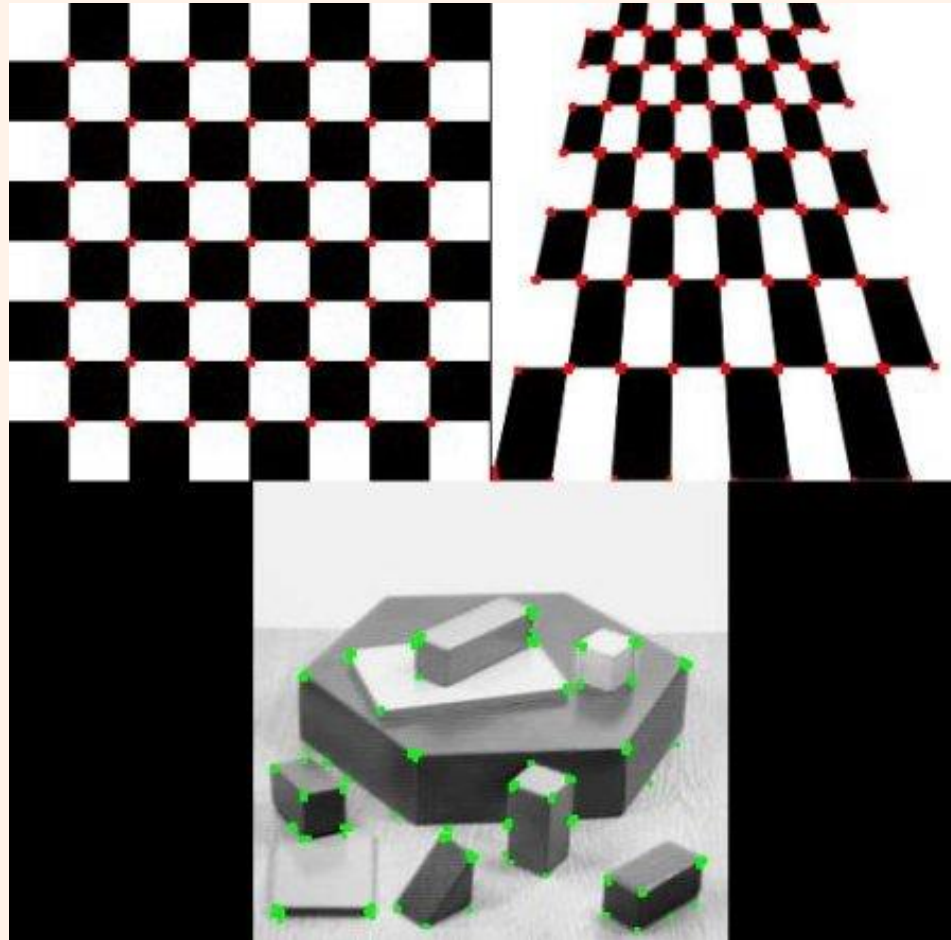
**Detección:** consiste en obtener la ubicación de puntos o regiones relevantes y distinguibles en una imagen, tales como esquinas, bordes o regiones con ciertas características.

**Descripción:** consiste en computar y asignar un vector a cada uno de los puntos o regiones detectadas tal que la vecindad de los mismos en la imagen quede caracterizada, en lo posible, en forma única.

**Matching:** consiste en la comparación de descriptores pertenecientes a dos imágenes de un mismo entorno. De esta forma, si el proceso de descripción es correcto y dos descriptores (pertenecientes a imágenes distintas) son lo suficientemente similares, se podrá relacionar unívocamente cada punto de una imagen, con otros puntos perteneciente a la otra.

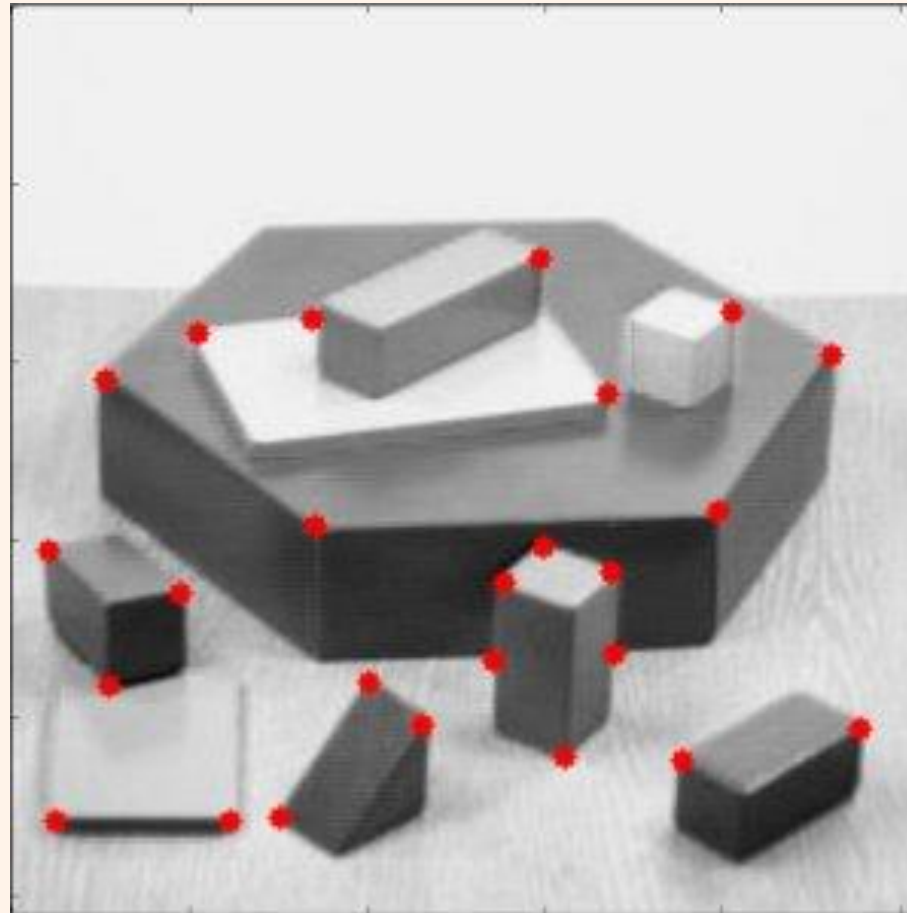
# Características locales

## Harris Corner Detection



# Características locales

## Shi-Tomasi Corner Detector



# Características locales

SIFT (Scale-Invariant Feature Transform)





# Características locales

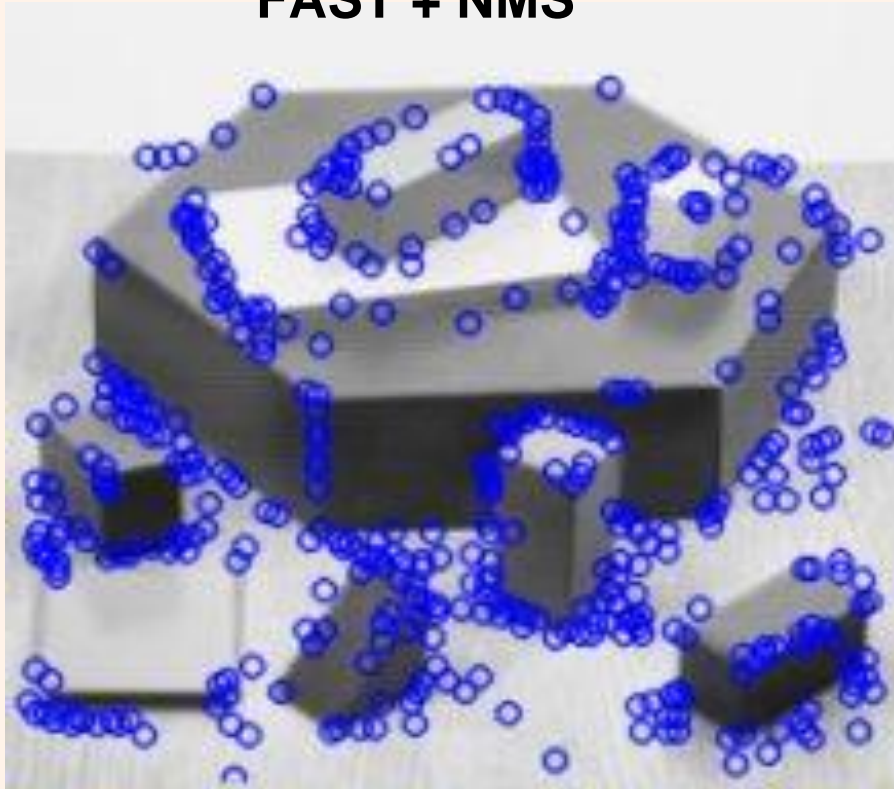
SURF (Speeded-Up Robust Features)



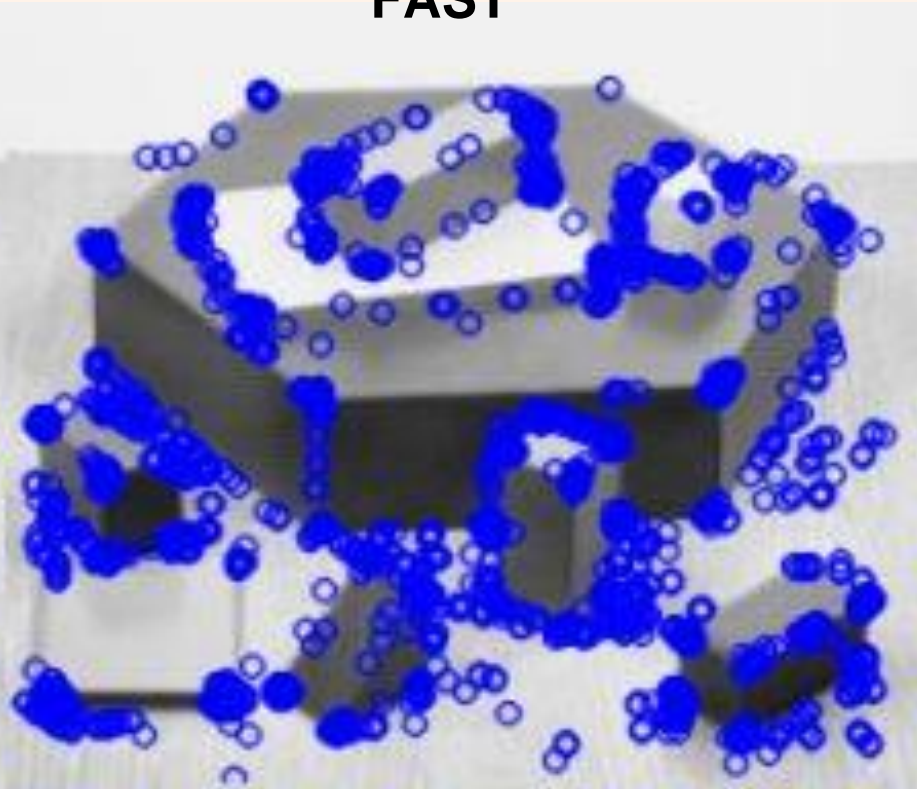
# Características locales

## FAST Algorithm for Corner Detection

**FAST + NMS**



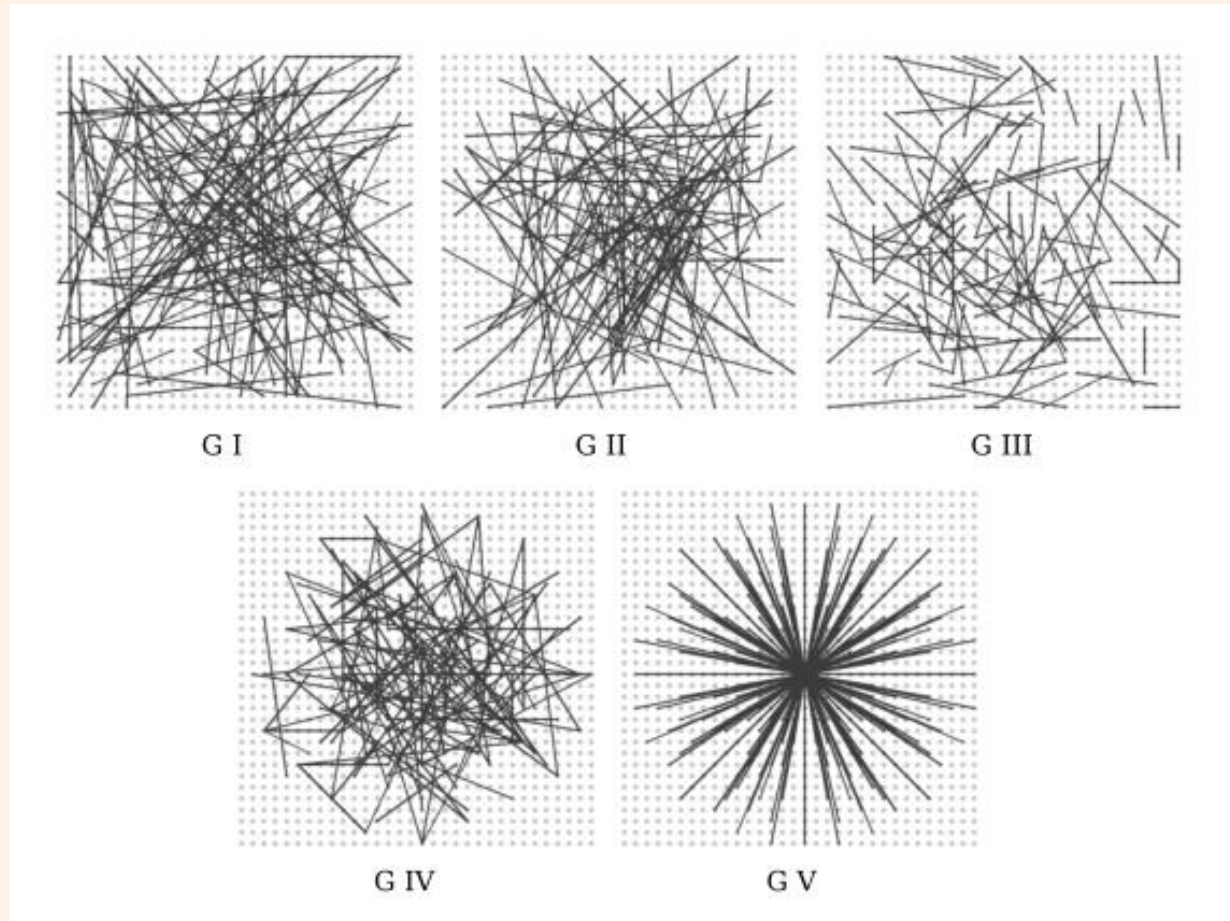
**FAST**





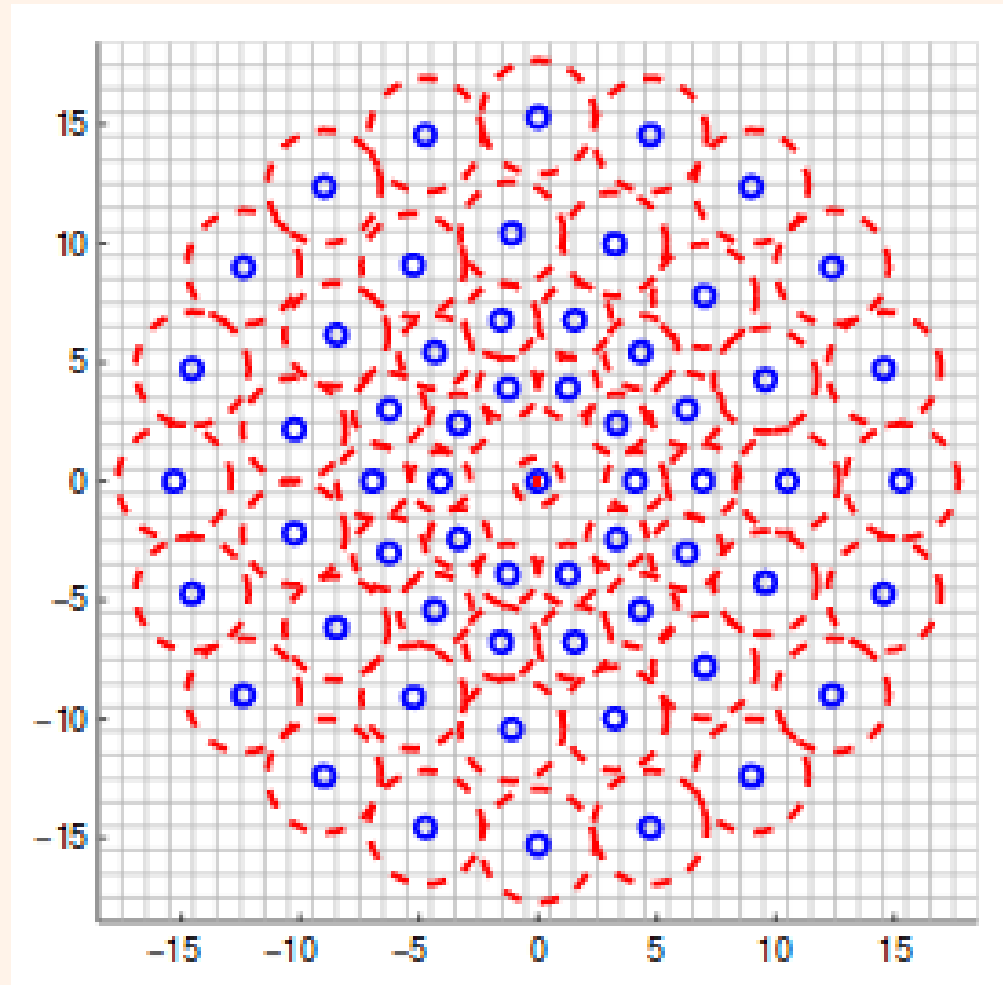
# Características locales

BRIEF (Binary Robust Independent Elementary Features)



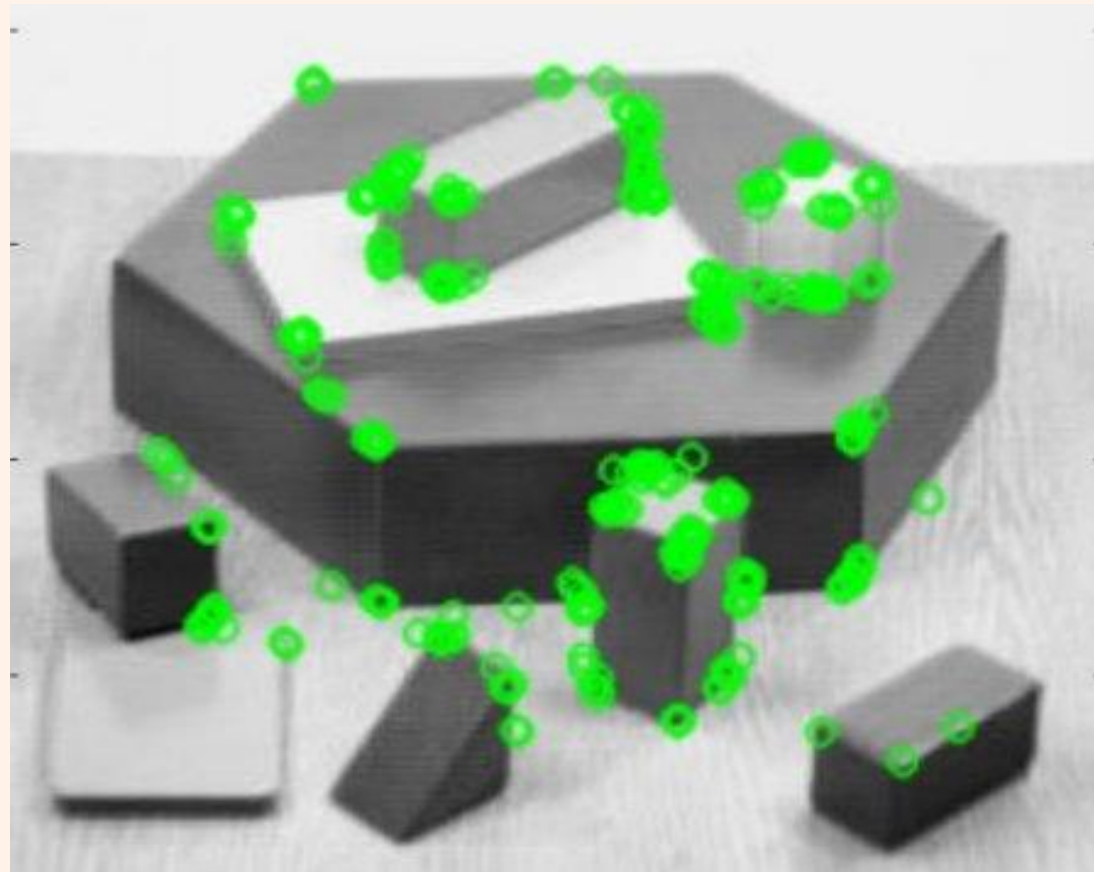
# Características locales

BRISK (Binary Robust Invariant Scalable Keypoints)



# Características locales

ORB (Oriented FAST and Rotated BRIEF)



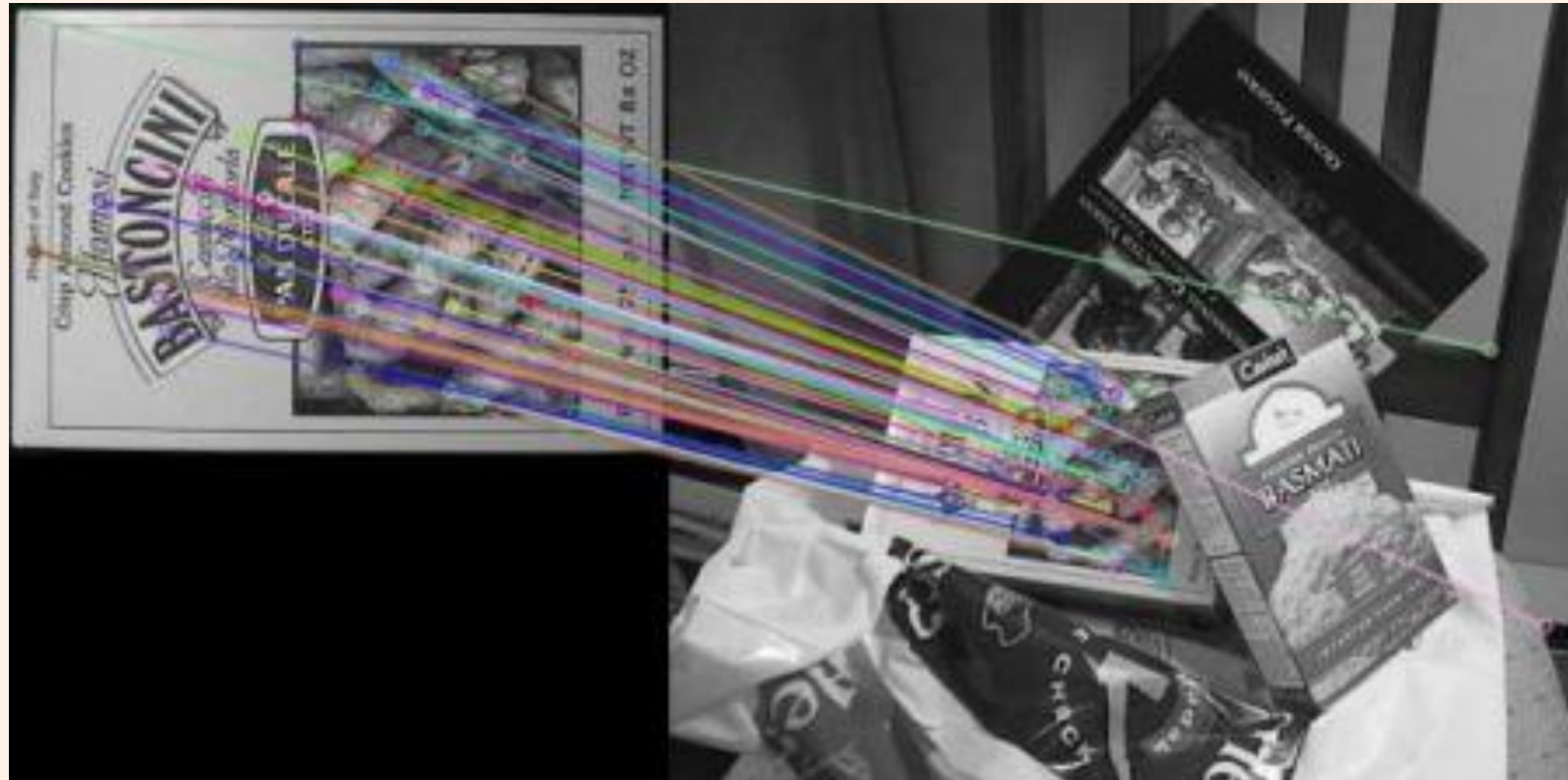
# Características locales

Brute-Force Matcher



# Características locales

Brute-Force Matcher

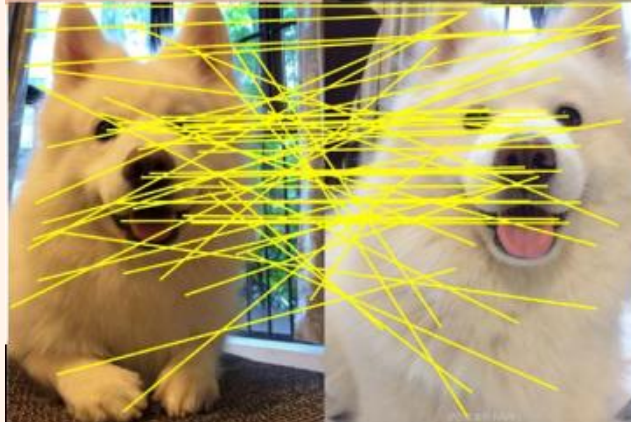
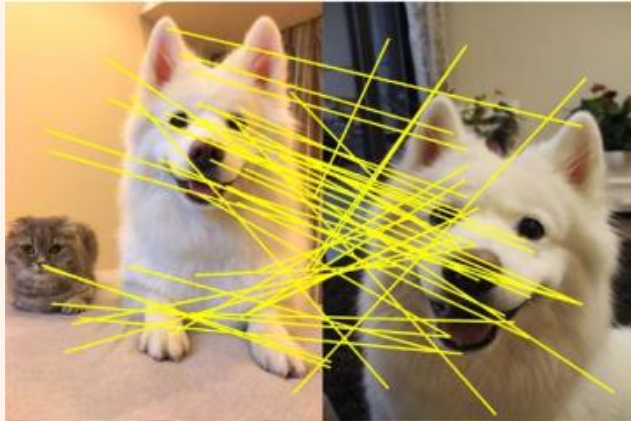




# Características locales

GMS (Grid-based Motion Statistics)

**Brute-Force Matcher**



**GMS**

