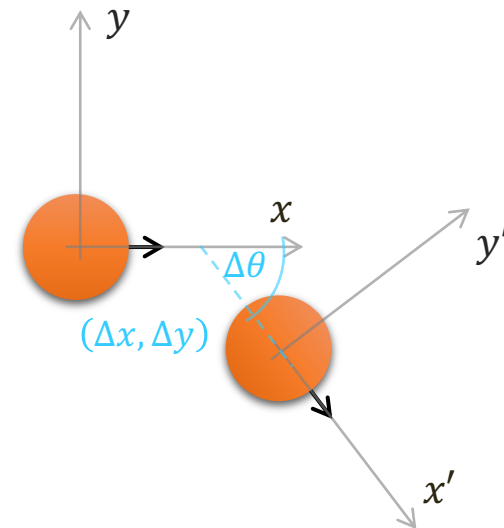
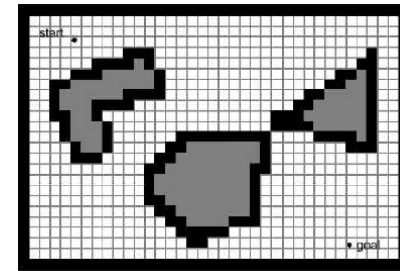


Intro Localización

Rodrigo Verschae

Localización

- Como el Robot puede saber donde esta en el mundo
 - Dado un “mapa” del mundo
 - o relativo a su posición actual



Descriptores Visuales

Ejemplo: ¿Cómo calzamos dos imágenes?



Ejemplo: ¿Cómo calzamos dos imágenes?

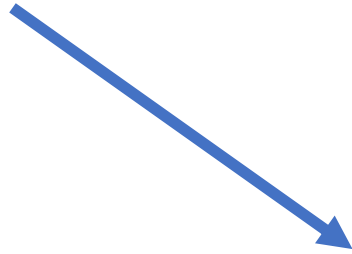
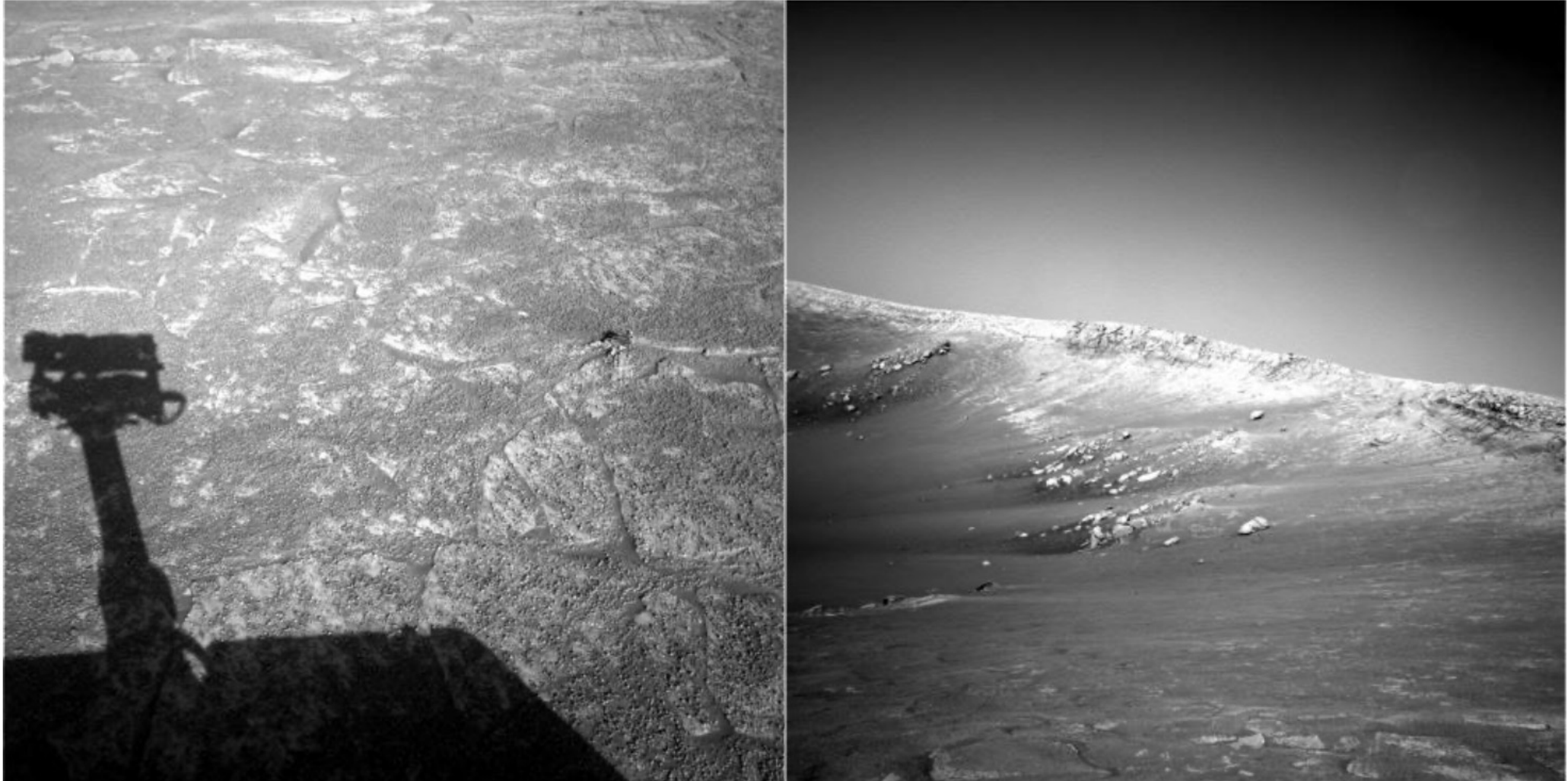


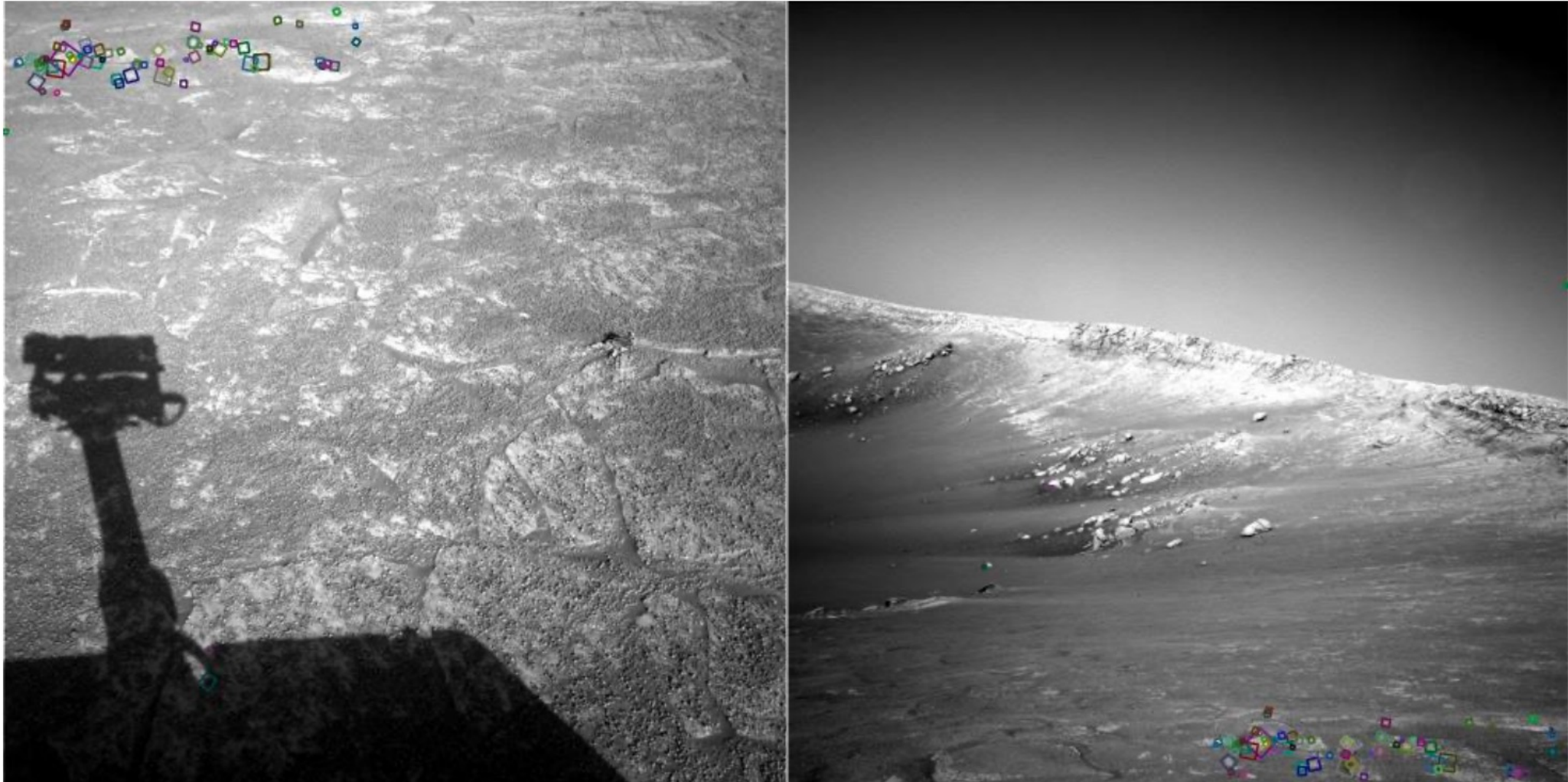
Image matching: why is it challenging?



NASA Mars Rover images

Image matching: why is it challenging?

Answer below



NASA Mars Rover images with SIFT feature matches

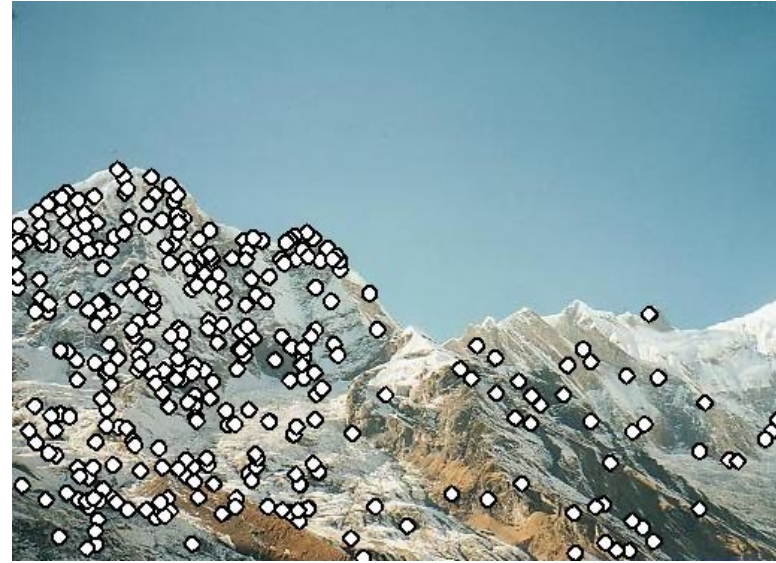
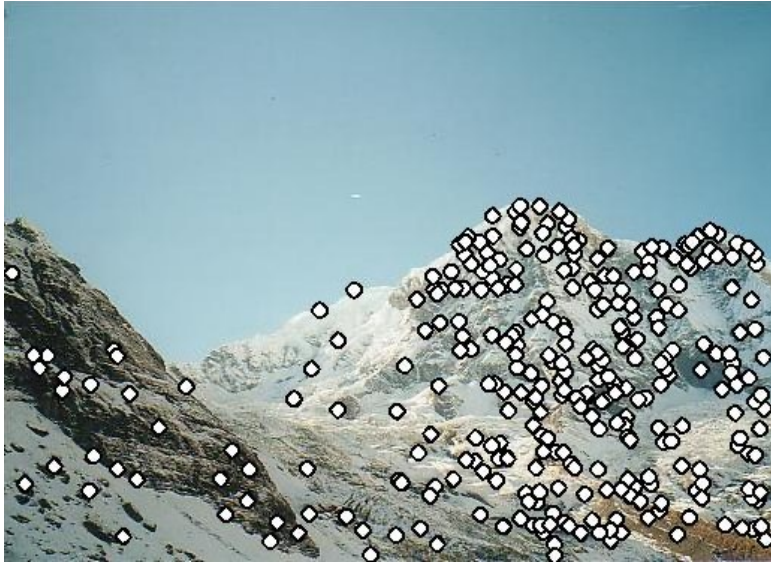
¿Cómo calzamos dos imágenes?

- Métodos globales sensibles a oclusión, iluminación, paralaje.
- Mejor buscar características locales que calcen.



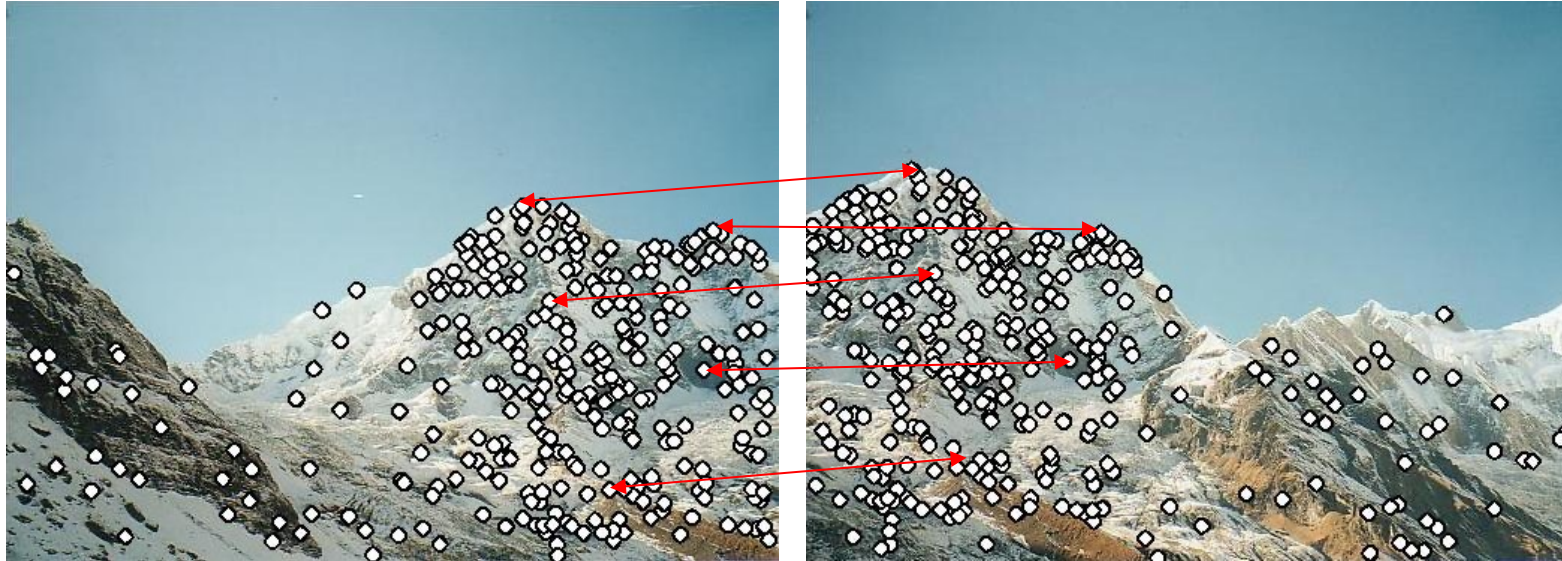
Calce con Características

- Detectar puntos característicos en ambas imágenes



Calce con Características

- Detectar puntos característicos en ambas imágenes
- Encontrar los pares que calzan



Calce con Características

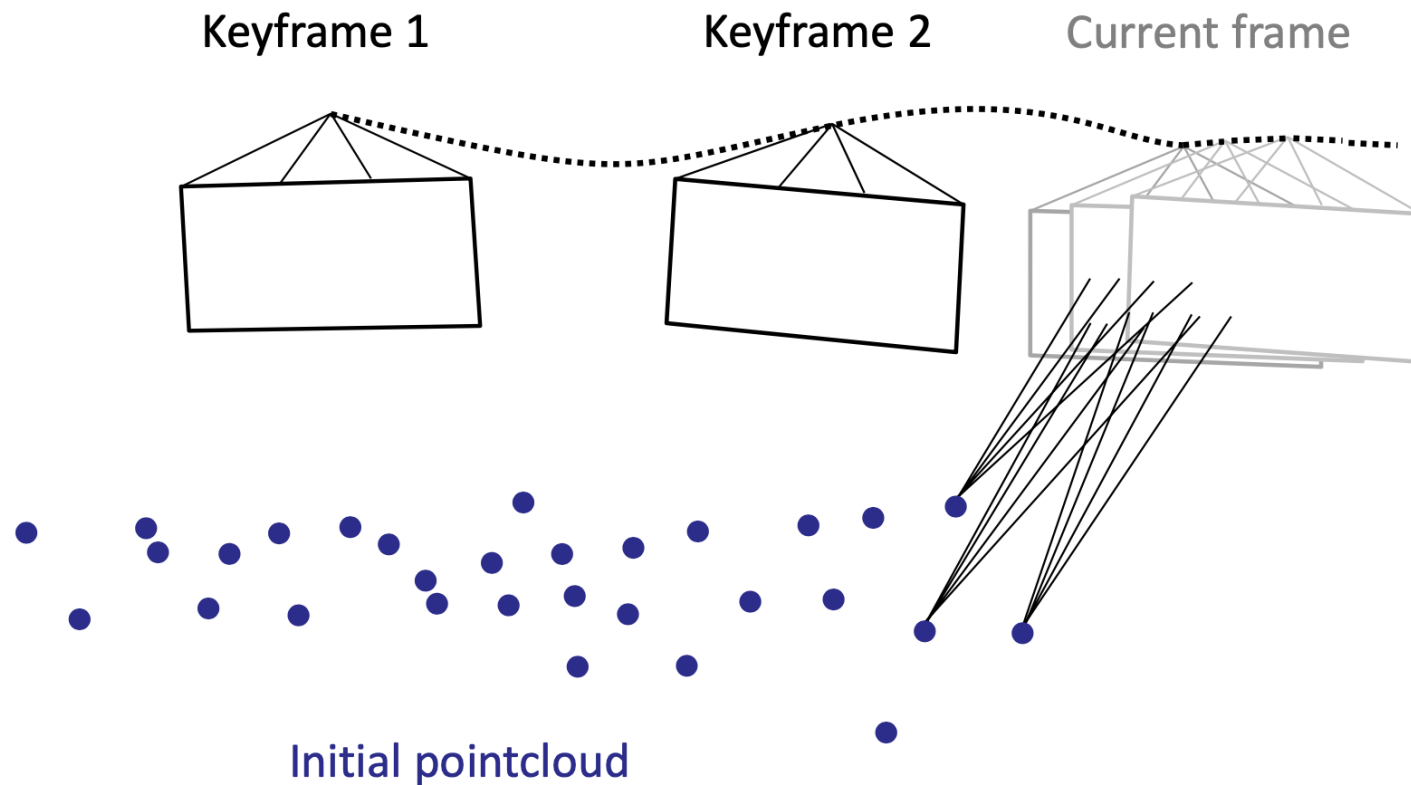
- Detectar puntos característicos en ambas imágenes
- Encontrar los pares que calzan
- Usar estos pares para calzar imágenes



Monocular VO (i.e., with a single camera)

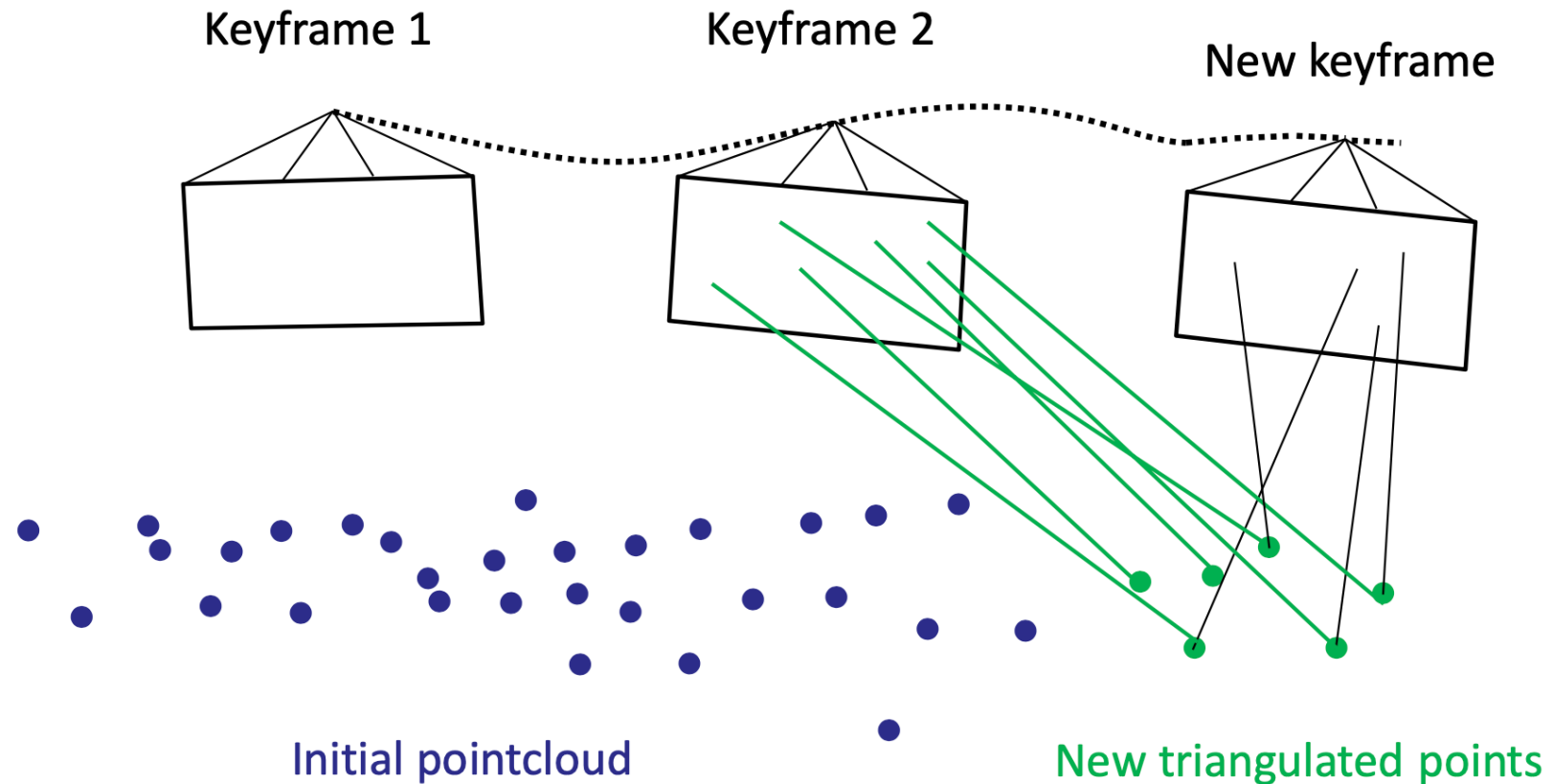
➤ Localization

- Given a 3D point cloud (map), determine the pose of each additional view
 - How?
 - How long can I do that?

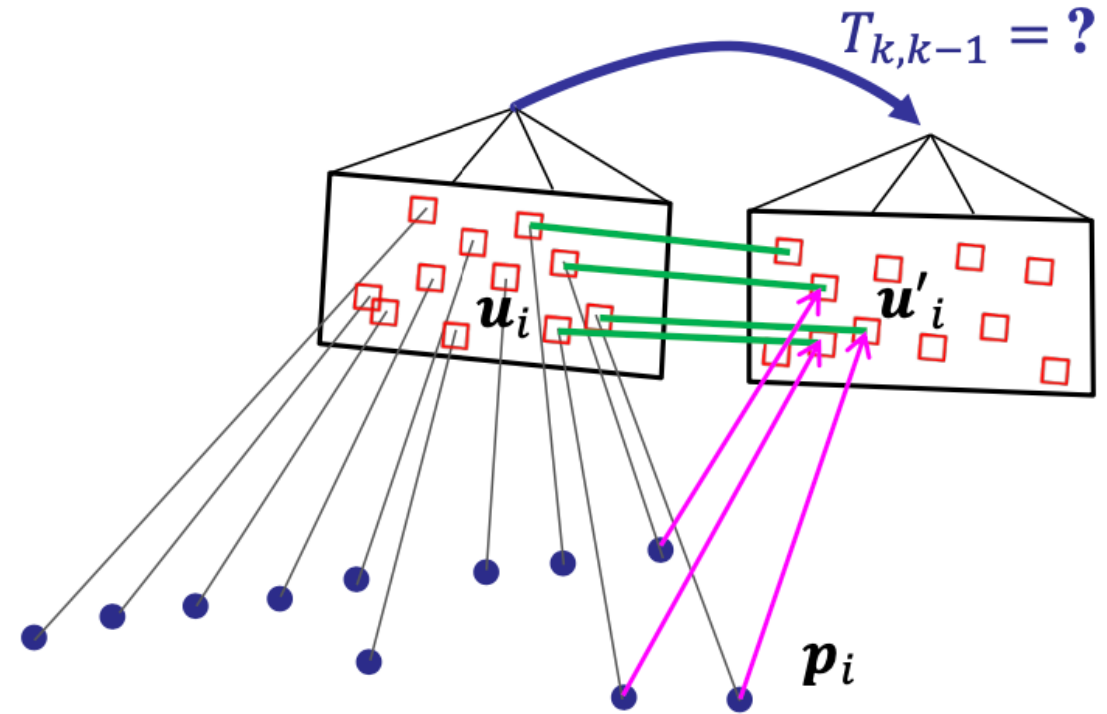
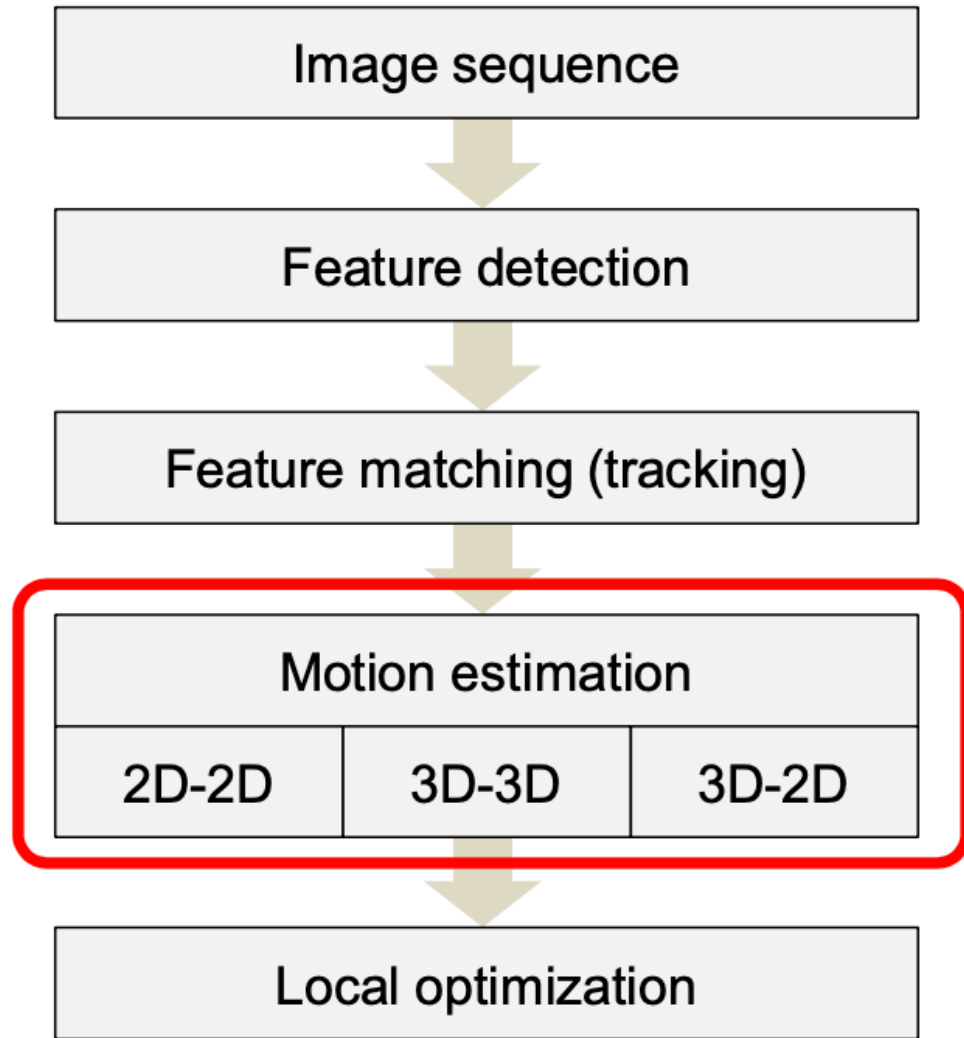


Extend Structure (i.e., mapping)

- Extract and triangulate new features

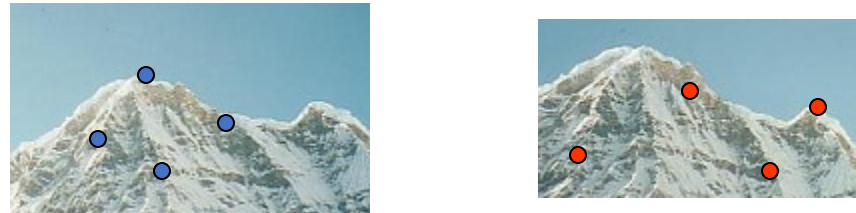


Keypoint extraction is the key ingredient of motion estimation!



Calce con Características

- Problema 1:
 - Detectar los *mismos* puntos *independientemente* en ambas imágenes

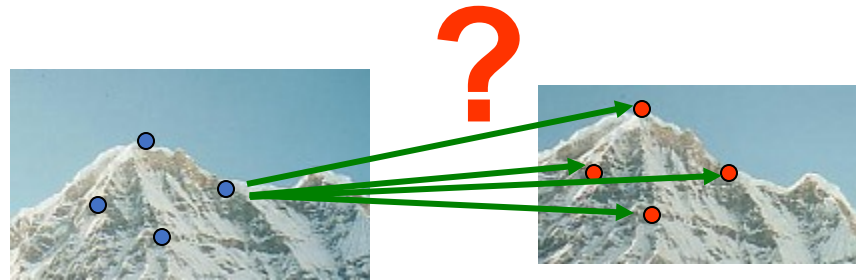


¡Imposible calzar!

Necesitamos un detector repetible

Calce con Características

- Problema 2:
 - Para cada punto, encontrar correctamente su par correspondiente



Necesitamos un descriptor confiable y repetible

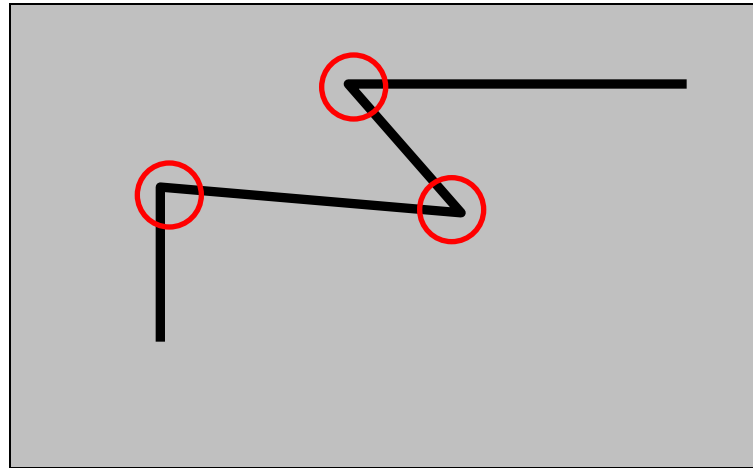
Más motivación

- Los puntos característicos (puntos de interés) son usados para:
 - Alineamiento de imágenes
 - Reconstrucción 3D
 - Tracking de Movimiento
 - Reconocimiento de Objetos
 - Bases de Datos 3D
 - Navegación de Robots
 - Panoramas
 - ... otros

Puntos de Interés

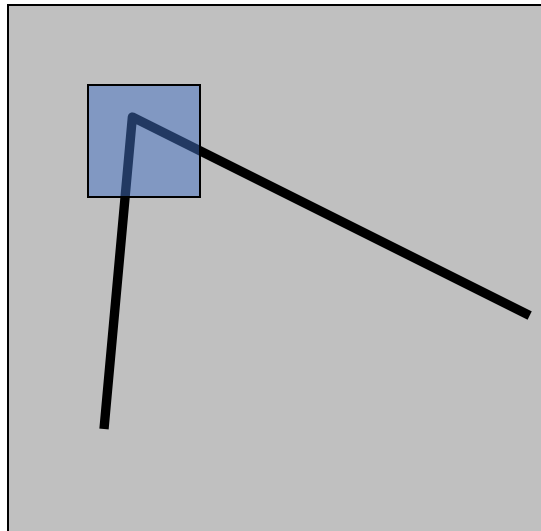
- Que son los puntos interés?
 - Satisfacen constancia de brillo – se ven iguales en ambas imágenes
 - Tienen suficiente variación de textura
 - No tienen demasiada variación de textura
 - Corresponden a una región “real”
 - No se deforman mucho en el tiempo

Ejemplo introductorio: *Detector de Esquinas de Harris*

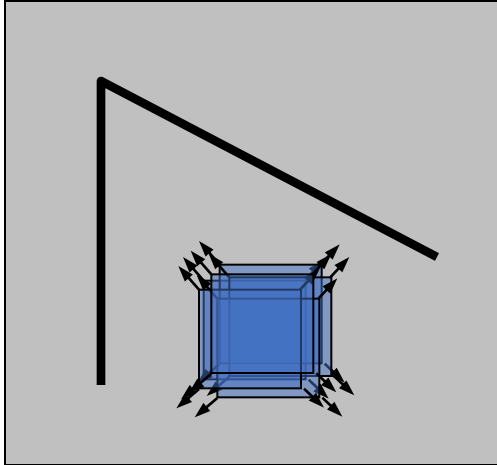


Idea Básica

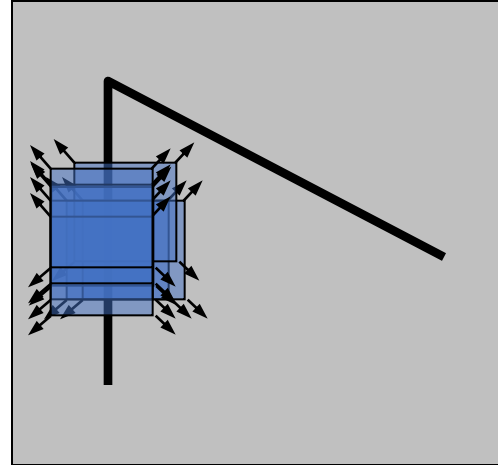
- Debe ser fácil localizar los puntos mirando una ventana pequeña
- Desplazar la ventana en cualquier dirección debe producir un gran cambio en intensidad



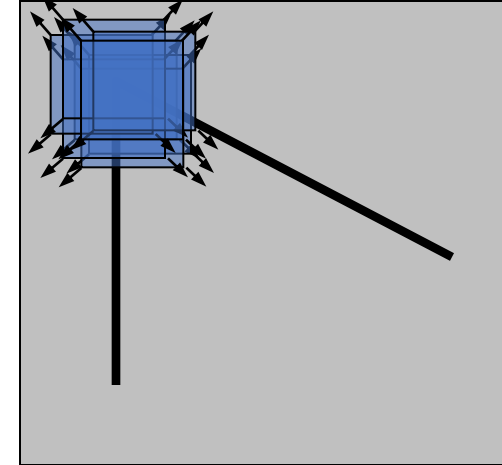
Detector Harris: Idea Básica



Región “plana”:
no hay cambios
al desplazar la
ventana en
ninguna
dirección



“borde”:
no hay cambios al
desplazar la
ventana en la
dirección del borde



“esquina”:
Un desplazamiento
en cualquier
dirección produce
un cambio
importante

Detector Harris: Matemática

Cambio de intensidad promediado en la ventana inducido por desplazar la imagen en $[u, v]$:

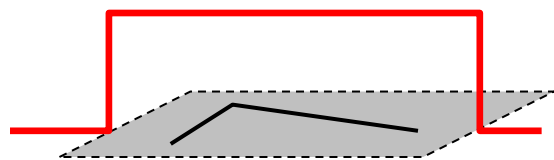
$$E(u, v) = \sum_{x, y} w(x, y) [I(x+u, y+v) - I(x, y)]^2$$

Función de ventana

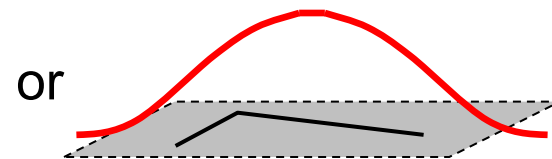
Intensidad desplazada

Intensidad

Función de ventana
 $w(x, y) =$



1 en ventana, 0 fuera



Gaussiana

Taylor de primer grado

$$\begin{aligned} E(u, v) &\approx \sum_{x, y} w(x, y) [I(x, y) + uI_x + vI_y - I(x, y)]^2 \\ &= \sum_{x, y} w(x, y) [uI_x + vI_y]^2 \\ &= \sum_{x, y} w(x, y) (u \quad v) \begin{bmatrix} I_x I_x & I_x I_y \\ I_x I_y & I_y I_y \end{bmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \end{aligned}$$

Detector Harris: Matemática

Usando la expansión de Taylor de $I(x+u,y+v)$ tenemos, para desplazamientos pequeños, $[u, v]$, una aproximación bilineal:

$$E(u, v) \cong [u, v] M \begin{bmatrix} u \\ v \end{bmatrix}$$

Donde M es una matriz 2×2 obtenida de las derivadas de la imagen:

$$M = \sum_{x,y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$

M también es llamada “tensor estructural”

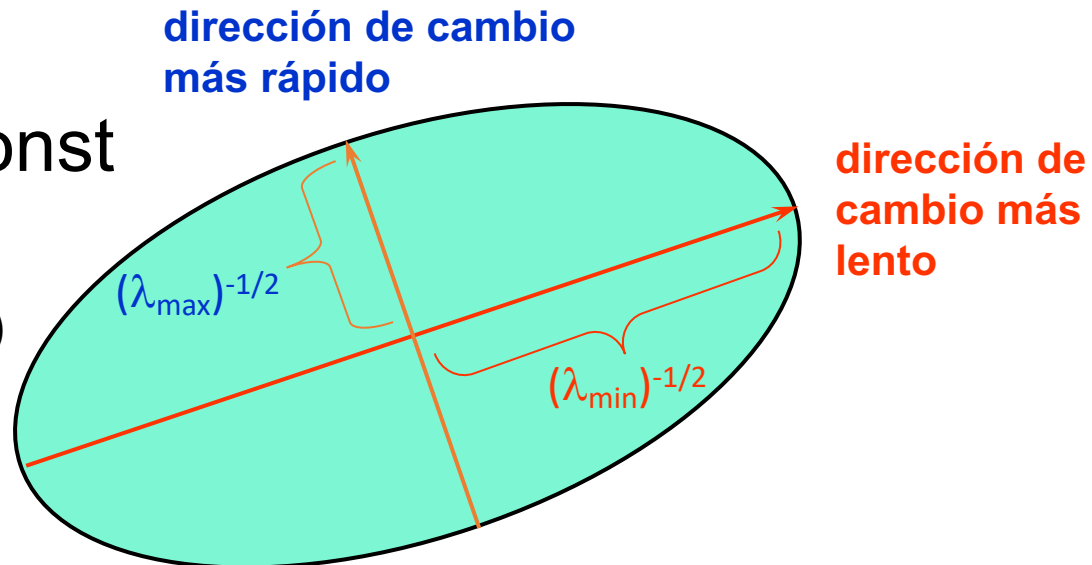
Detector Harris: Matemática

Cambio de intensidad al desplazar ventana: análisis de valores propios

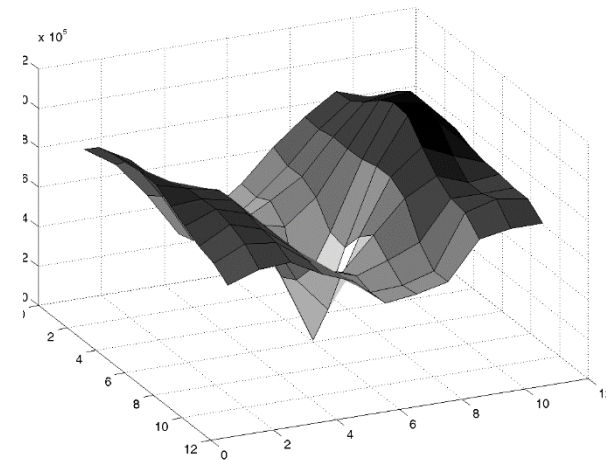
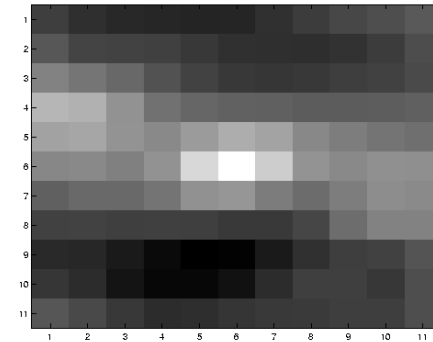
$$E(u, v) \cong [u, v] M \begin{bmatrix} u \\ v \end{bmatrix} \quad \lambda_1, \lambda_2 - \text{valores propios de } M$$

Elipse $E(u, v) = \text{const}$

Curva de nivel de $E(u, v)$

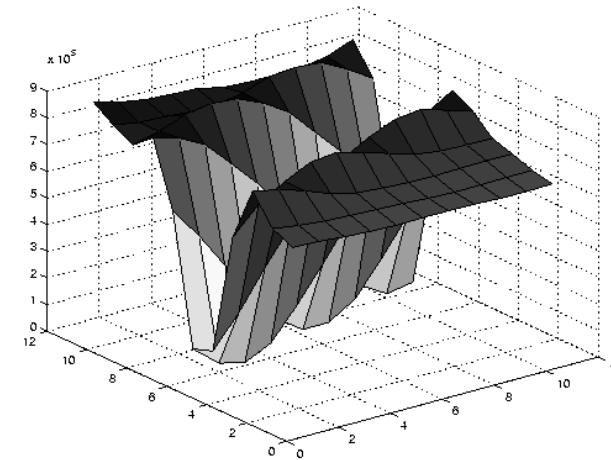
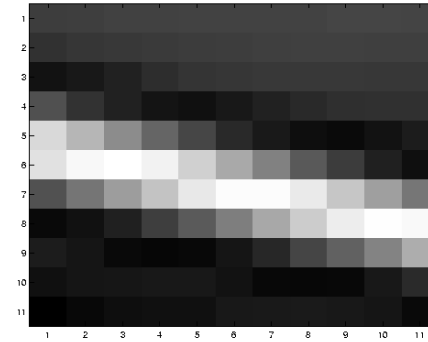


Selección de Buenos Puntos



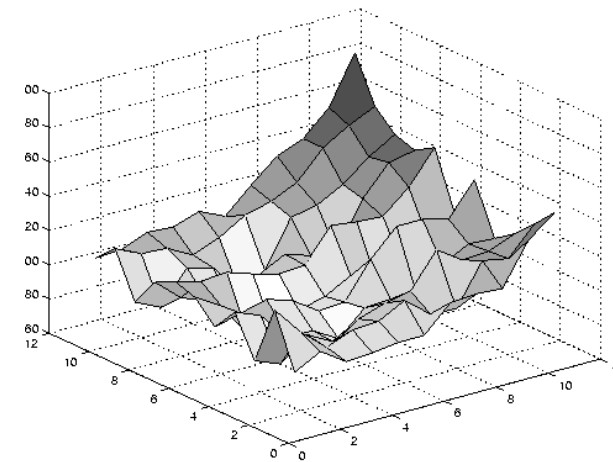
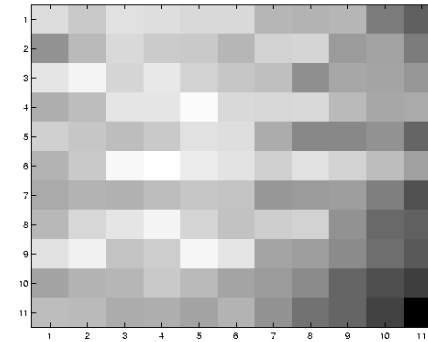
λ_1 y λ_2 son grandes

Selección de Buenos Puntos



λ_1 grande, λ_2 pequeño

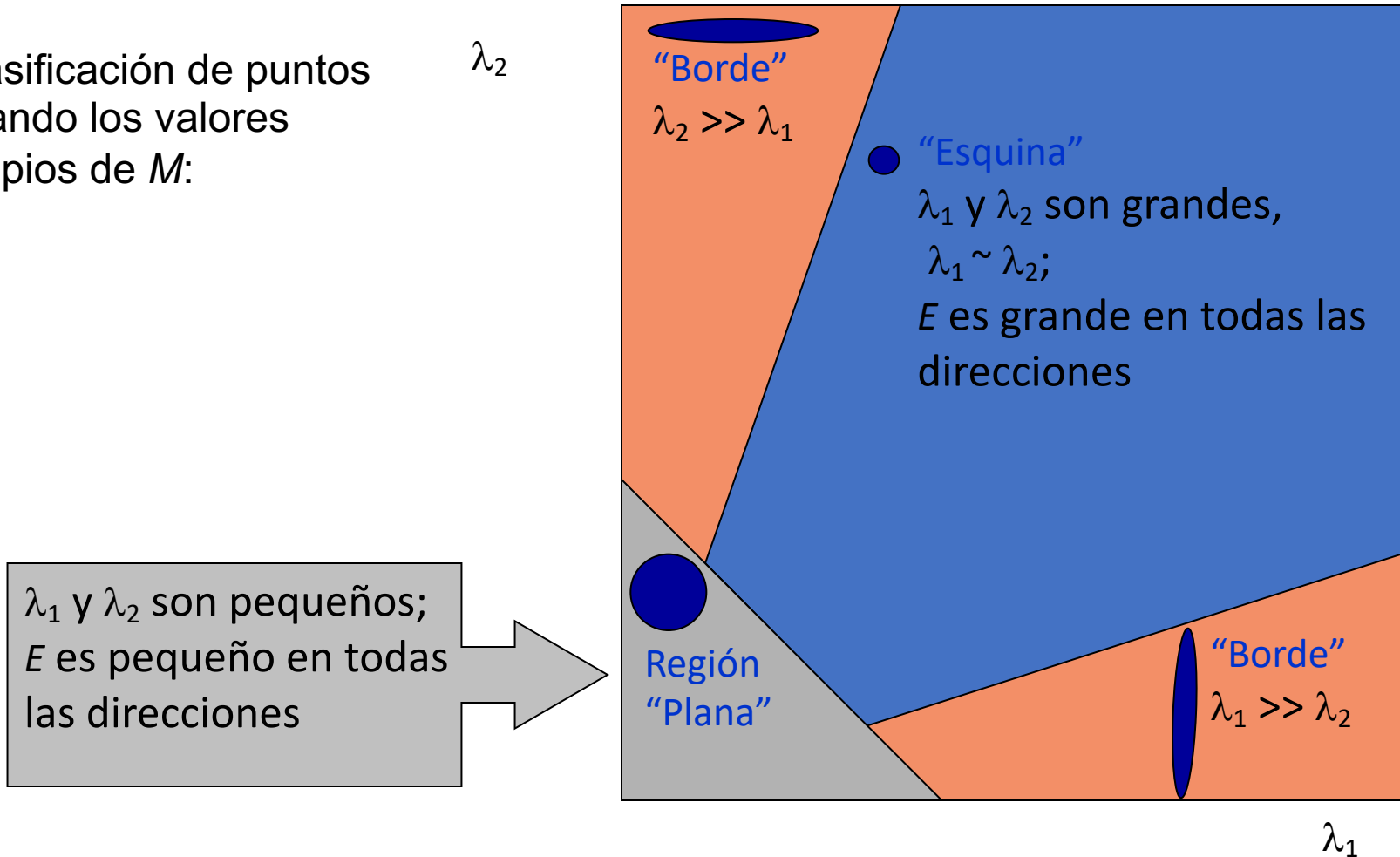
Selección de Buenos Puntos



λ_1 y λ_2 pequeños

Detector Harris: Matemática

Clasificación de puntos usando los valores propios de M :



Detector Harris: Eficiencia

Medida de la respuesta de esquina:

$$R = \det M - k (\text{trace } M)^2$$

Esta expresión no requiere el cálculo de los valores propios

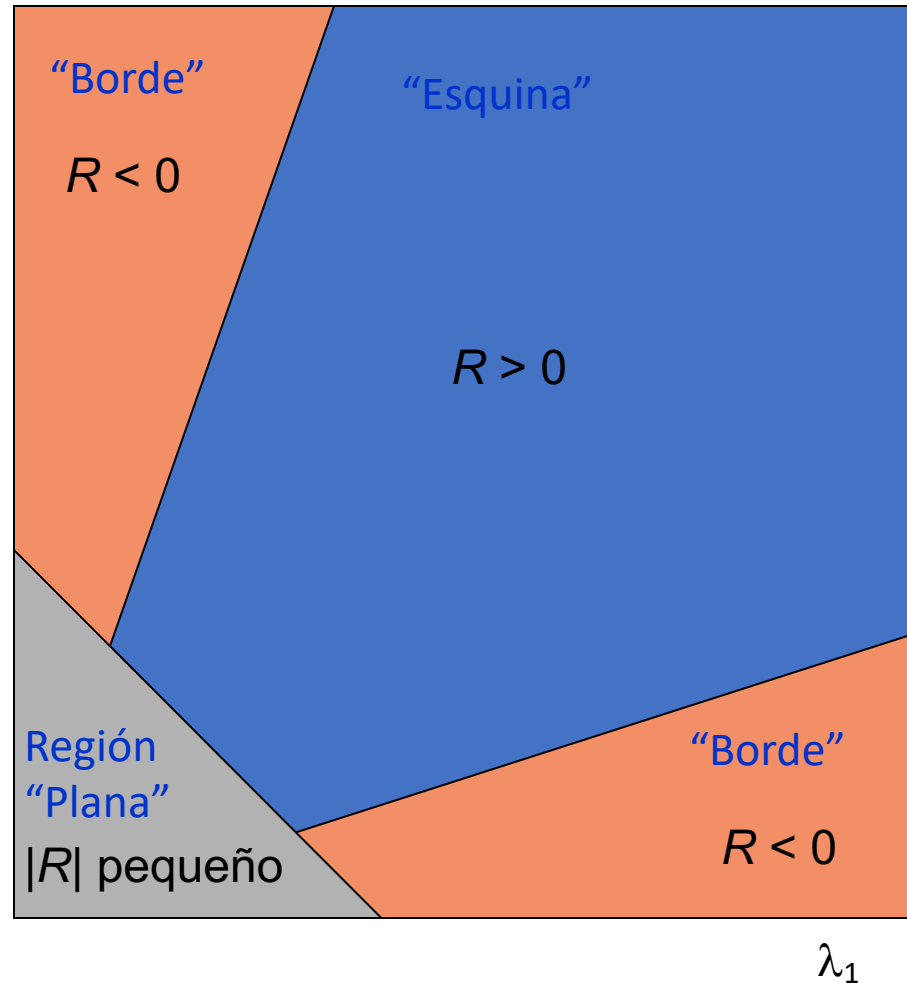
$$\det M = \lambda_1 \lambda_2$$

$$\text{trace } M = \lambda_1 + \lambda_2$$

(k – constante empírica, $k = 0.04-0.06$)

Detector Harris: Eficiencia

- R depende solo de los valores propios de M
- R es grande para una **esquina**
- R es muy negativo para un **borde**
- $|R|$ es pequeño para una región **plana**



Detector Harris

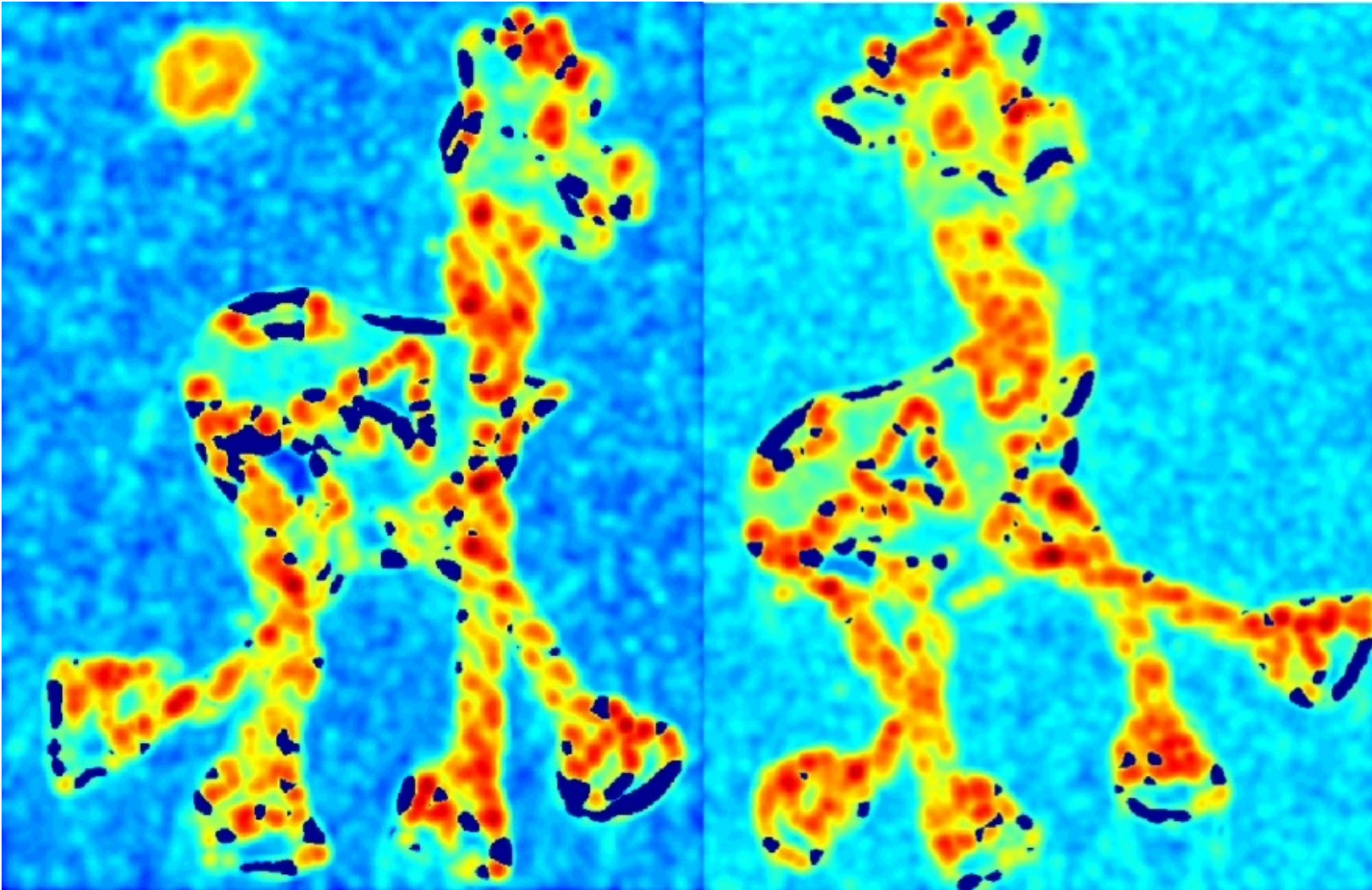
- El Algoritmo:
 - Encuentra puntos con Respuesta de Esquina grande ($R > \text{umbral}$)
 - Toma los máximos locales de R

Detector Harris: Paso a Paso



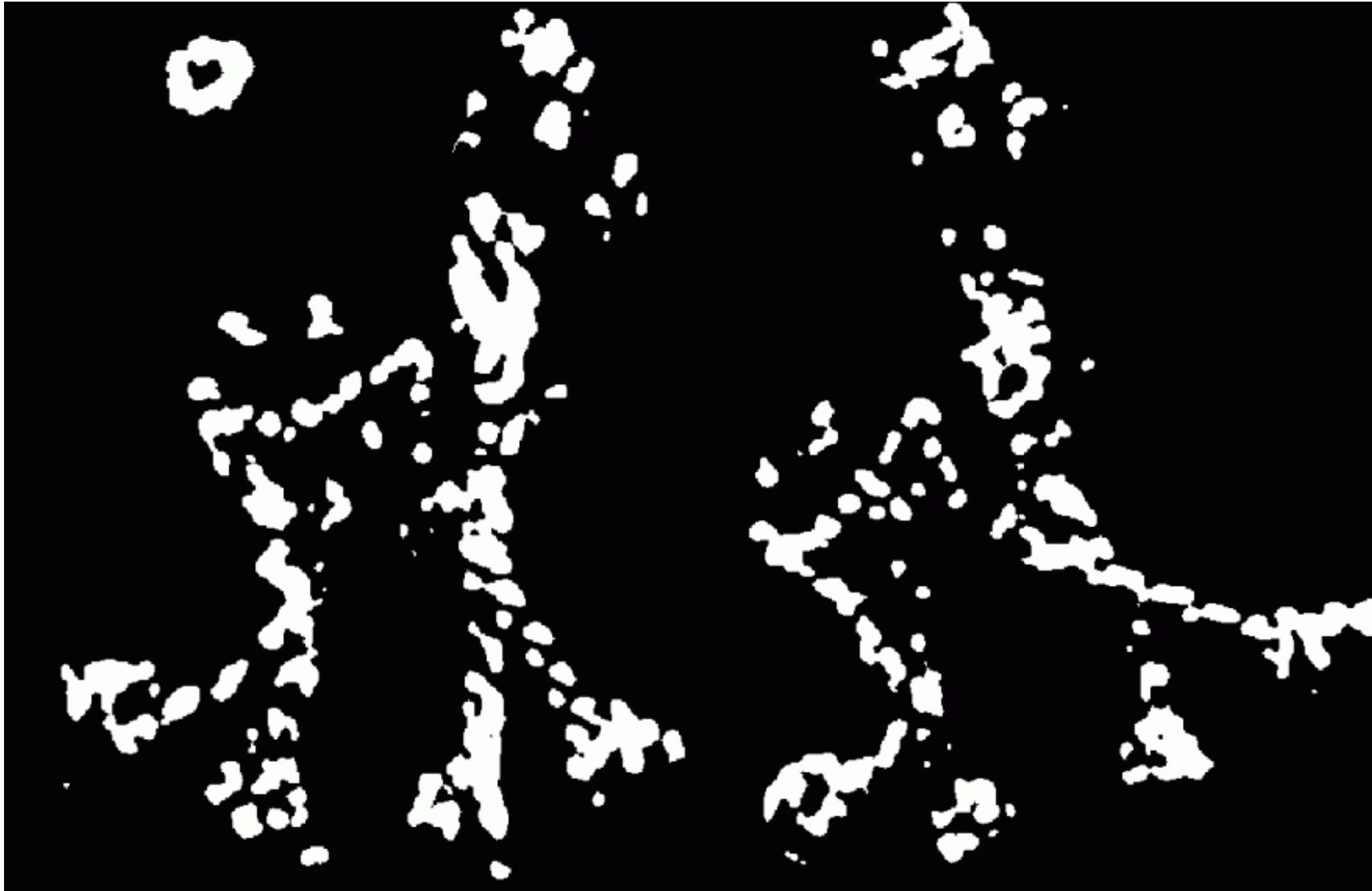
Detector Harris: Paso a Paso

Calcular la Respuesta de Esquina R



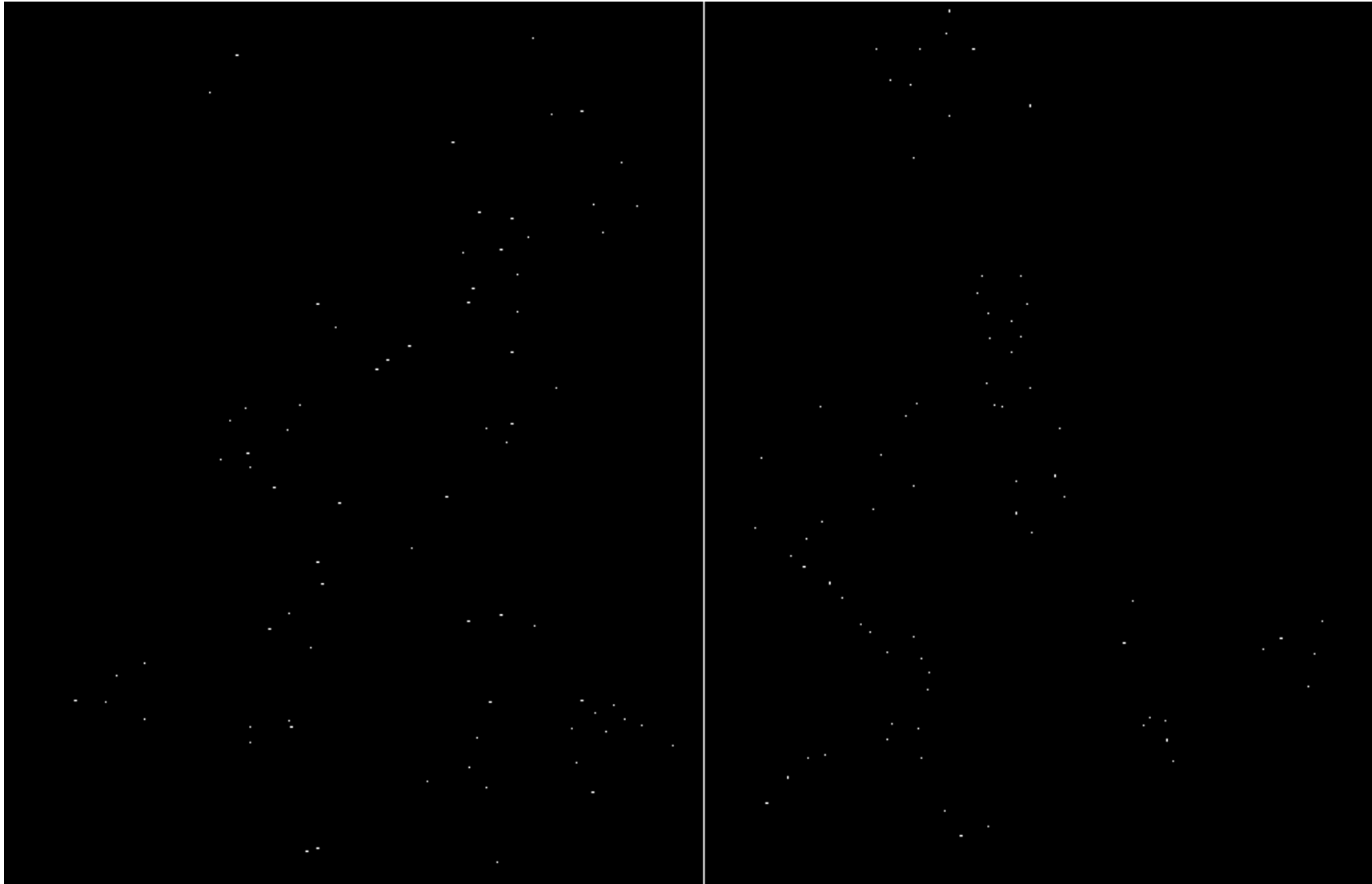
Detector Harris: Paso a Paso

Encontrar puntos con Respuesta de
Esquina grande: $R > \text{umbral}$



Detector Harris: Paso a Paso

Tomar máximos locales de R



Detector Harris: Resultado



Detector Harris: Resumen

- Cambio promedio en dirección $[u, v]$ puede ser expresado como una forma bilineal:

$$E(u, v) \cong [u, v] M \begin{bmatrix} u \\ v \end{bmatrix}$$

- Describe un punto en términos de valores propios de M : *medida de respuesta de esquina*

$$R = \lambda_1 \lambda_2 - k (\lambda_1 + \lambda_2)^2$$

- Un buen punto (esquina) debe tener un *gran cambio de intensidad en todas las direcciones*, i.e. R debe ser muy positivo

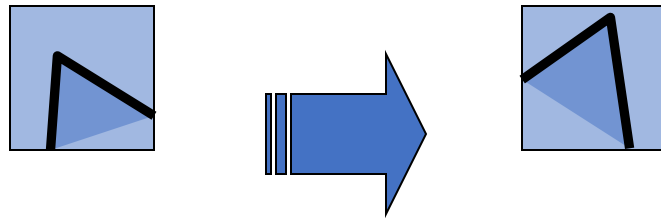
Detector Ideal de Puntos Característicos

Siempre encuentra los mismos puntos en un objeto, sin importar los cambios en la imagen, i.e.:

- invariante a:
 - cambios en:
 - Escala
 - Iluminación
 - Perspectiva
- Robusto a
 - oclusión parcial

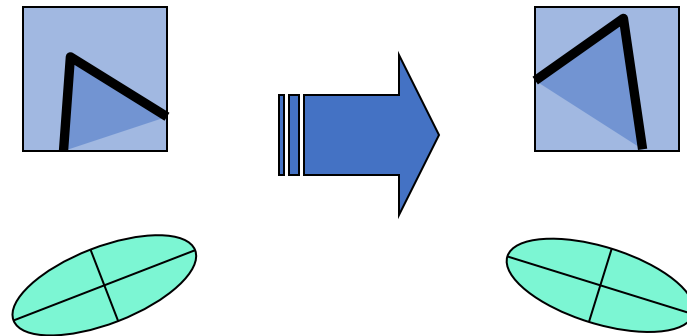
Detector Harris: Algunas Propiedades

- ¿Invariancia a la Rotación?



Detector Harris: Algunas Propiedades

- Invariancia a la Rotación



La elipse rota pero su forma (i.e. valores propios) se mantiene

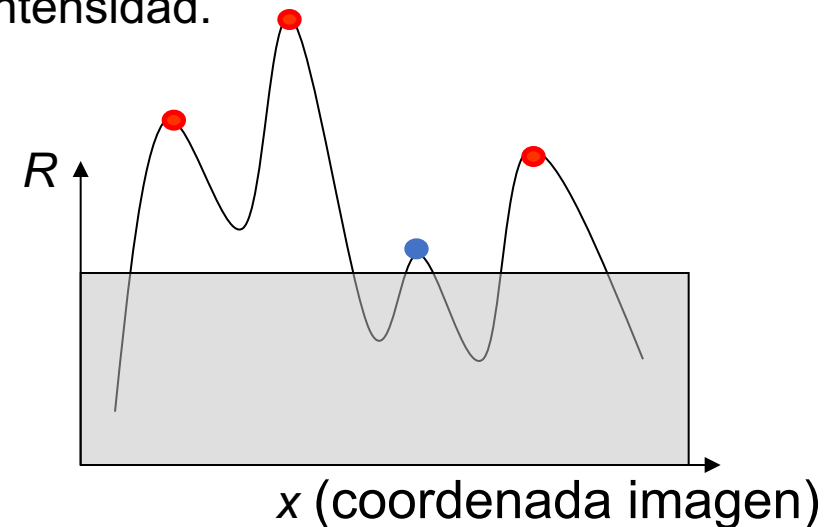
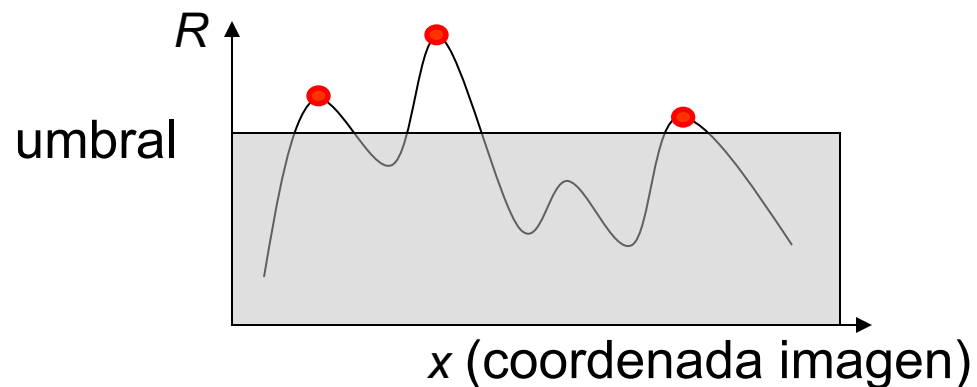
La respuesta de esquina R es invariante a la rotación de la imagen

Detector Harris: Algunas Propiedades

- ¿Invariancia a cambios en la intensidad de la imagen ?

Detector Harris: Algunas Propiedades

- Invariancia parcial a cambios de intensidad aditivos y multiplicativos
 - ✓ Sólo se usan las derivadas => invariancia a desplazamientos en intensidad $I \rightarrow I + b$
 - ✓ Escalamiento intensidad: $I \rightarrow a I$ Dado que el umbral de intensidad para máximos locales es fijo, sólo invariancia parcial a cambios multiplicativos de intensidad.

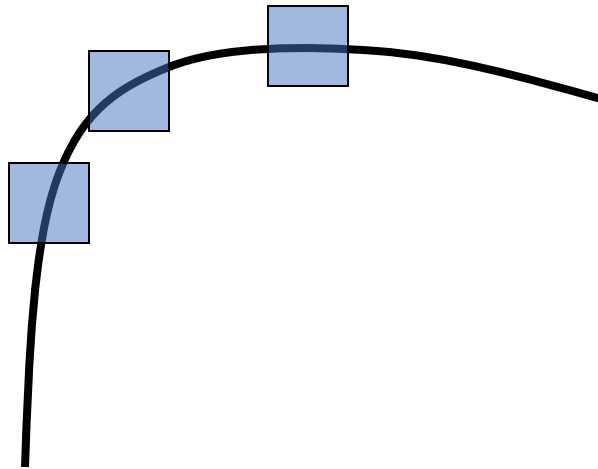


Detector Harris: Algunas Propiedades

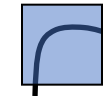
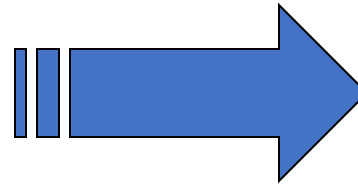
- ¿Invariante a escalamiento imagen?

Detector Harris: Algunas Propiedades

- ¡No es invariante al *escalamiento de la imagen!*



Todos los puntos
serán clasificados
como **Bordes**



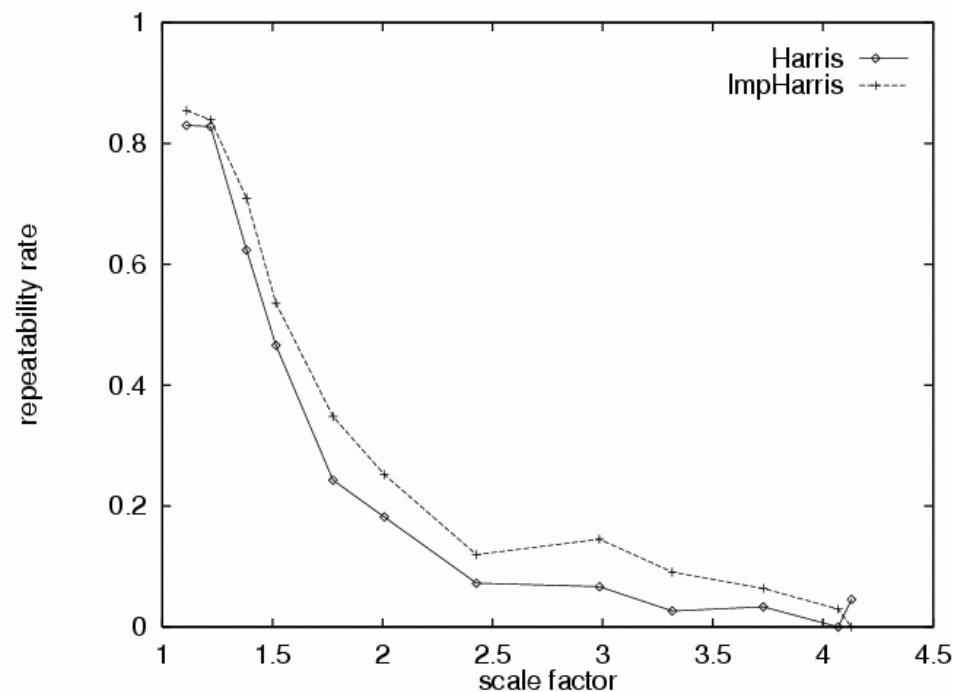
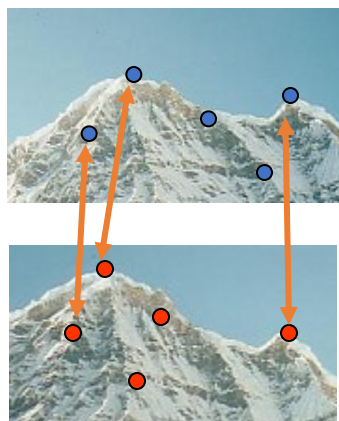
¡ Esquina !

Detector Harris: Algunas Propiedades

- Calidad del Detector Harris para diferentes cambios de escala

Tasa de repetibilidad:

$$\frac{\# \text{ correspondencias}}{\# \text{ correspondencias posibles}}$$



Gráficos son de este paper



International Journal of Computer Vision 37(2), 151–172, 2000
© 2000 Kluwer Academic Publishers. Manufactured in The Netherlands.

Evaluation of Interest Point Detectors

CORDELIA SCHMID, ROGER MOHR AND CHRISTIAN BAUCKHAGE

INRIA Rhône-Alpes, 655 av. de l'Europe, 38330 Montbonnot, France

Cordelia.Schmid@inrialpes.fr

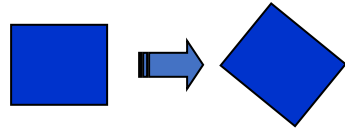
Abstract. Many different low-level feature detectors exist and it is widely agreed that the evaluation of detectors is important. In this paper we introduce two evaluation criteria for interest points: repeatability rate and information content. Repeatability rate evaluates the geometric stability under different transformations. Information content measures the distinctiveness of features. Different interest point detectors are compared using these two criteria. We determine which detector gives the best results and show that it satisfies the criteria well.

Queremos: **detectar *los mismos* puntos de interés sin importar los *cambios en la imagen***

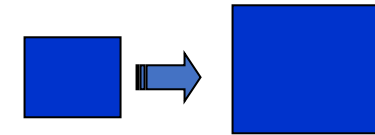
Modelos de cambios en la imagen

- Geometría

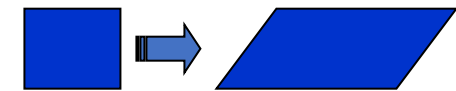
- Rotación



- Similitud (Rotación + Escala uniforme)

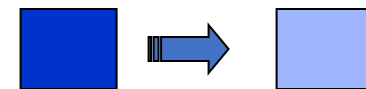


- Afín (Escalamiento dependiente de dirección) válido para: cámara ortográfica, objeto localmente plano



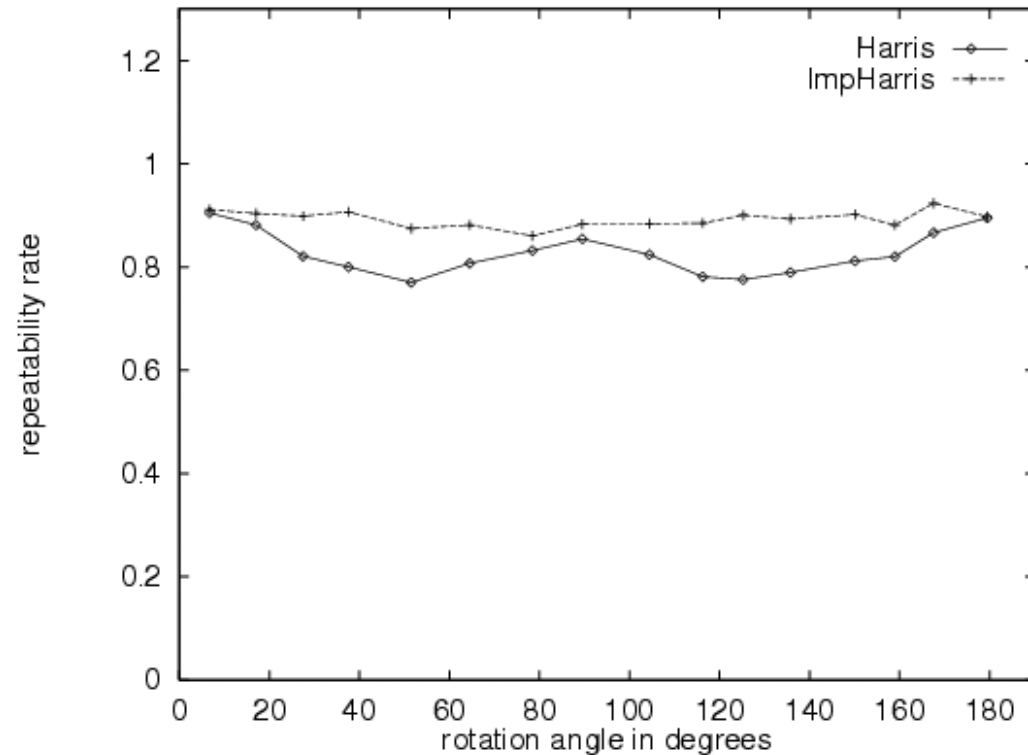
- Fotometría

- Cambio de intensidad afín ($I \rightarrow a I + b$)



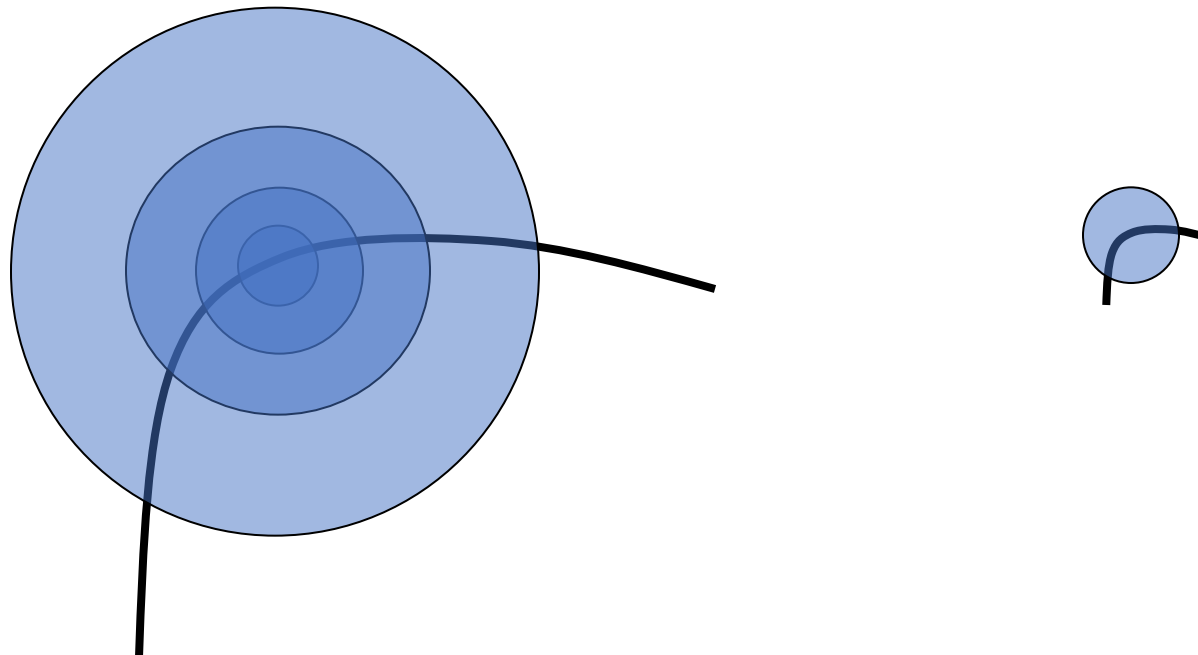
Detección Invariante a Rotación

- Detector de Esquinas Harris



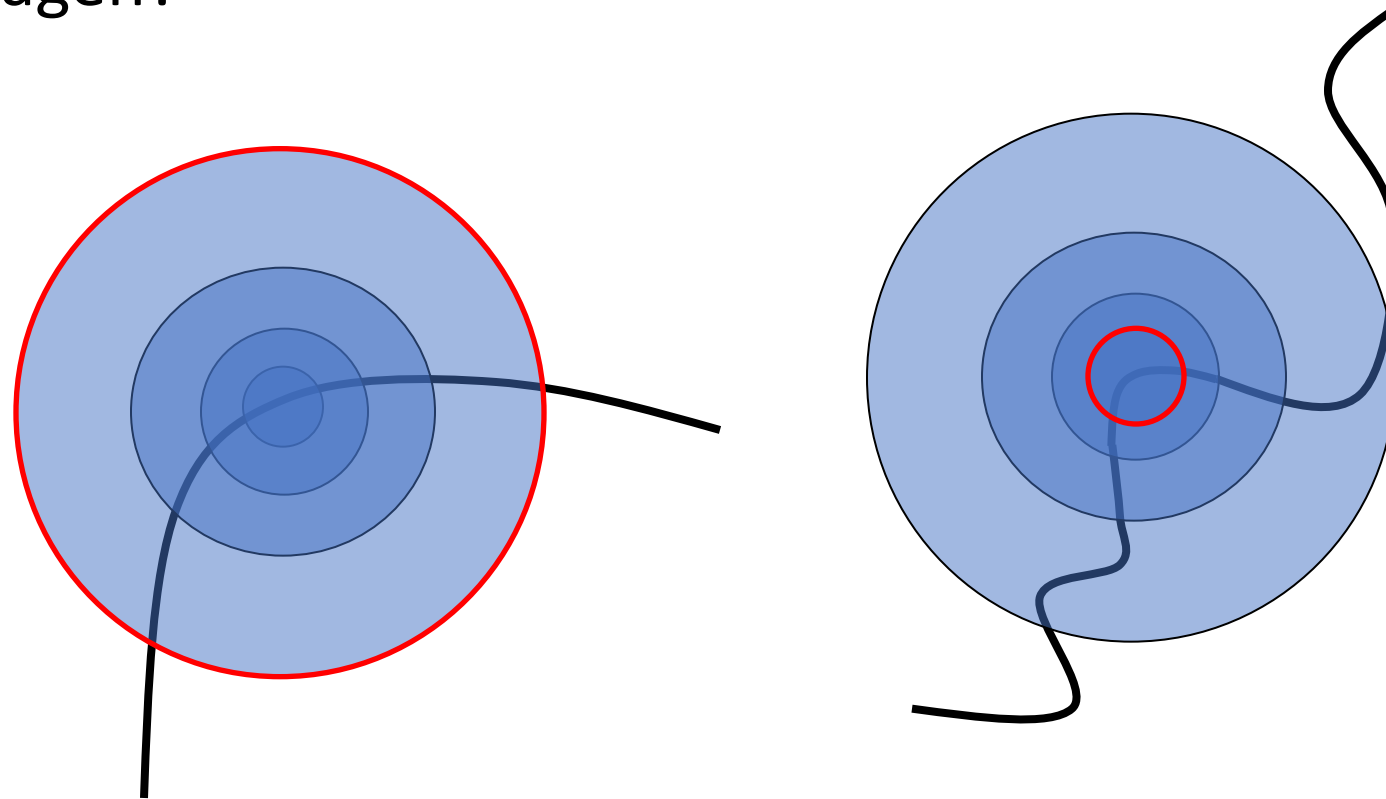
Detección Invariante a Escalamiento

- Consideremos regiones (Ej: círculos) de diferentes tamaños alrededor de un punto
- Regiones de tamaños correspondientes se verán iguales en ambas imágenes



Detección Invariante a Escalamiento

- El problema: cómo escogemos círculos correspondientes *independientemente* en cada imagen?

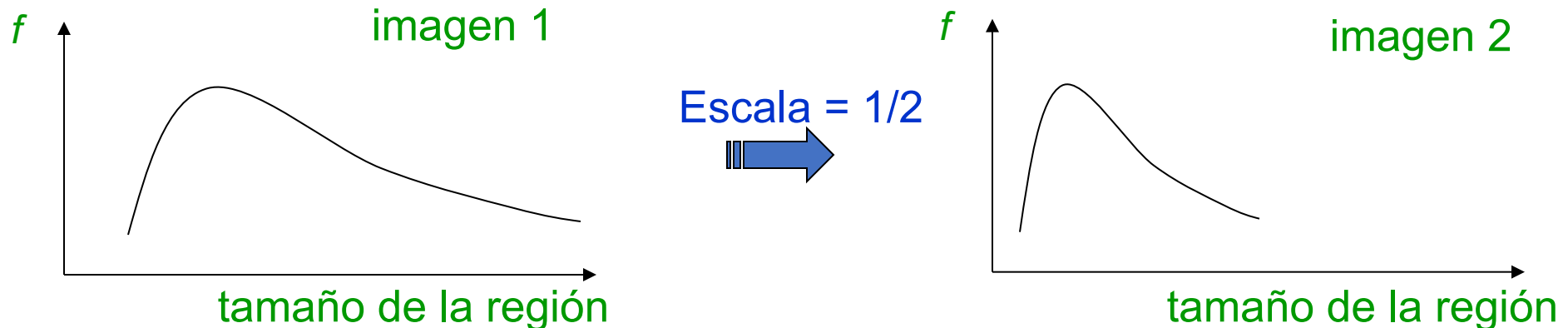


Detección Invariante a Escalamiento

- Solución:
 - Diseñar una función en la región (círculo), que sea “Invariante a Escalamiento” (la misma para regiones correspondientes, incluso si están a diferentes escalas)

Ejemplo: intensidad promedio. Para regiones correspondientes (incluso de diferentes tamaños) será la misma.

- Para un punto en una imagen, podemos considerarla como una función del tamaño de la región (radio del círculo)



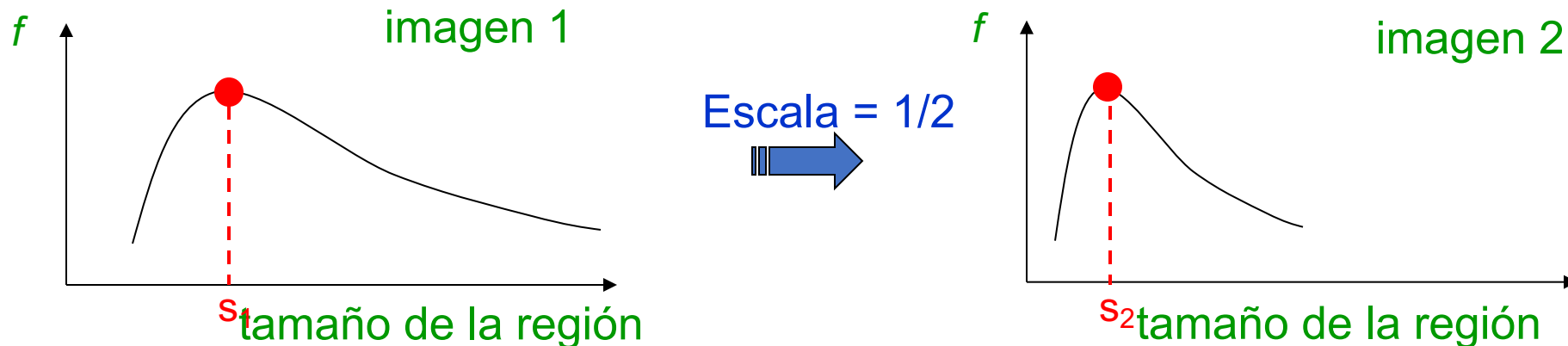
Detección Invariante a Escalamiento

- Solución común:

Tomar un máximo local de esta función

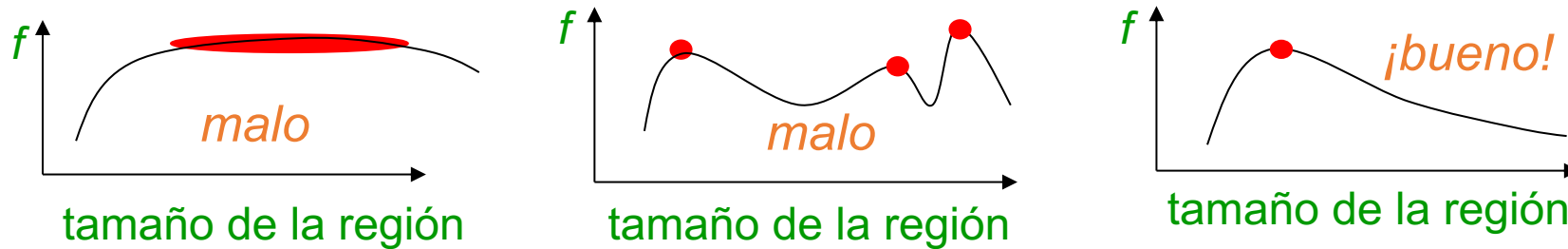
Observación: el tamaño de la región, para el cual el máximo se alcanza, debería ser *invariante* a la escala de la imagen.

Importante: ¡Este tamaño de la región invariante a escalamiento es encontrado en cada imagen **independientemente!**



Detección Invariante a Escalamiento

- Una “buena” función para detección de escala: tiene un máximo afilado y estable



- Para imágenes usuales: una buena función debería responder al contraste (cambio local rápido de intensidad)

Detección Invariante a Escalamiento

- Funciones para determinar escala

Kernels:

$$L = \sigma^2 \left(G_{xx}(x, y, \sigma) + G_{yy}(x, y, \sigma) \right)$$

(Laplaciano)

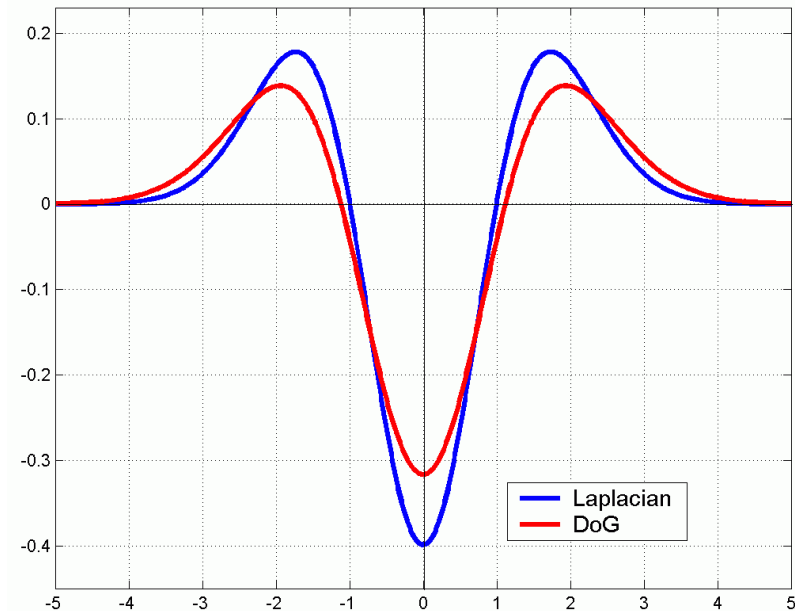
$$DoG = G(x, y, k\sigma) - G(x, y, \sigma)$$

(Diferencia de Gaussianas)

Donde:

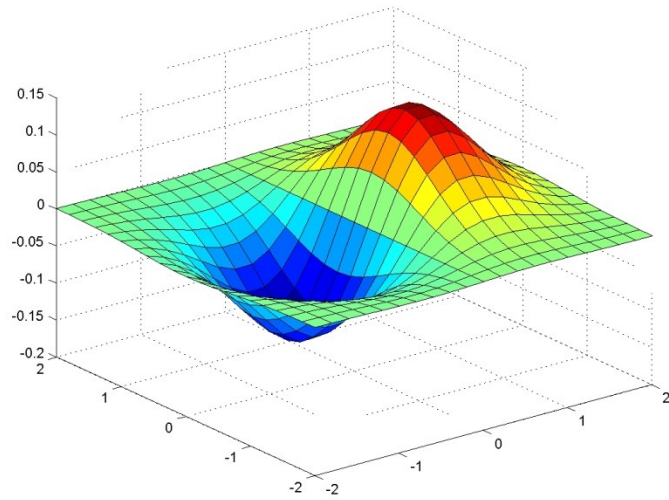
$$G(x, y, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2 + y^2}{2\sigma^2}}$$

$$f = \text{Kernel} * \text{Image}$$

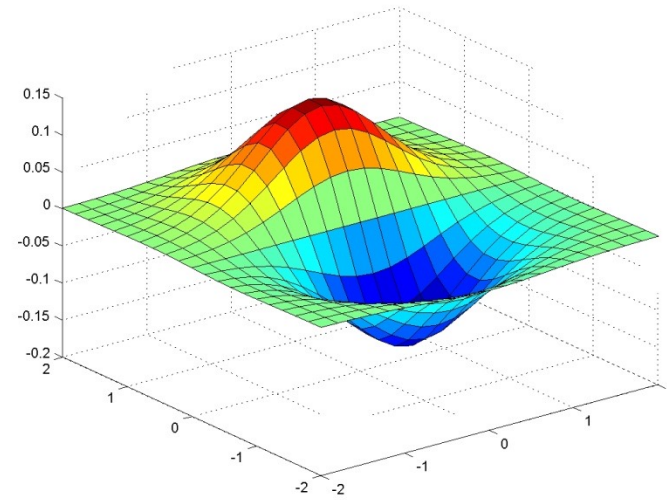


Nota: ambos kernels son invariantes a *escalamiento* y *rotación*

Derivada de Gaussiana

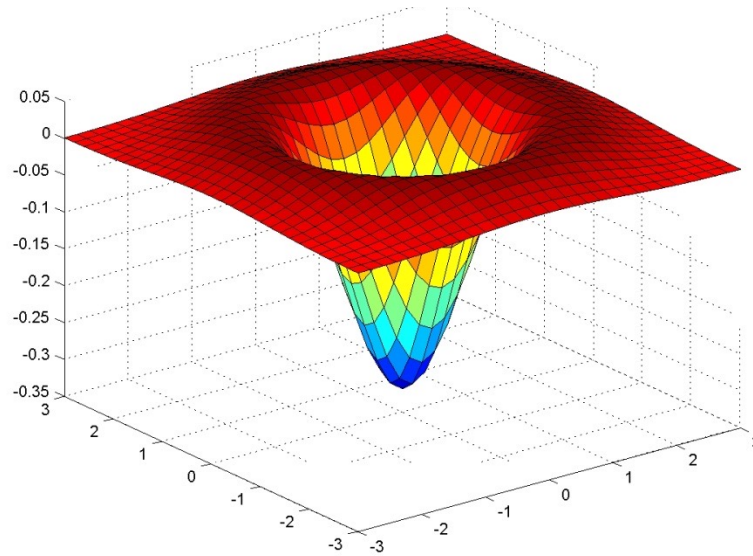


$$\frac{\partial}{\partial x} G_{\sigma}$$



$$\frac{\partial}{\partial y} G_{\sigma}$$

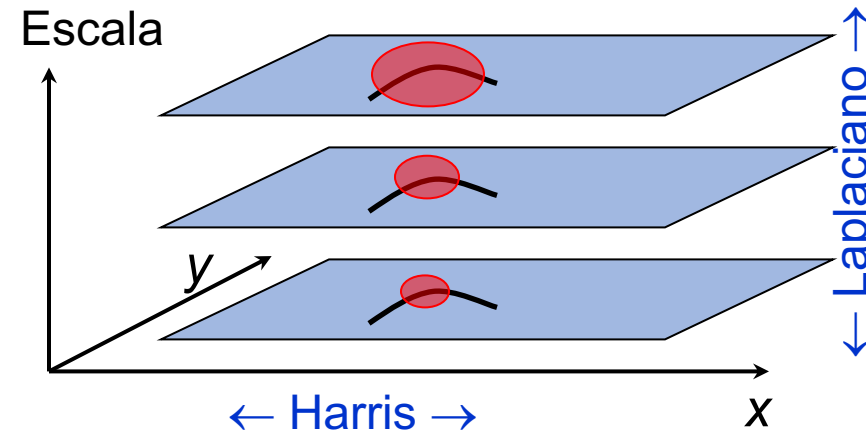
Laplaciano de Gaussiana



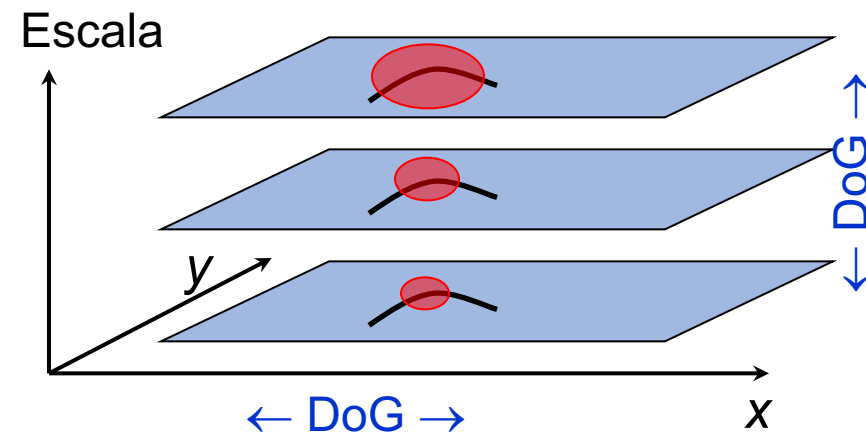
$$\text{Lo}G_{\sigma} = -\frac{1}{\pi\sigma^4} \left[1 - \frac{x^2 + y^2}{2\sigma^2} \right] e^{-\frac{(x^2+y^2)}{2\sigma^2}}$$

Detectores Invariantes a Escalamiento

- **Harris-Laplaciano**¹
Encuentra máximo local de:
 - Detector de Esquinas Harris en espacio (coordenadas imagen)
 - Laplaciano en Escala



- **SIFT (Lowe)**²
Encuentra máximo local de:
 - Diferencia de Gaussianas en espacio y Escala



¹ K.Mikolajczyk, C.Schmid. "Indexing Based on Scale Invariant Interest Points". ICCV 2001

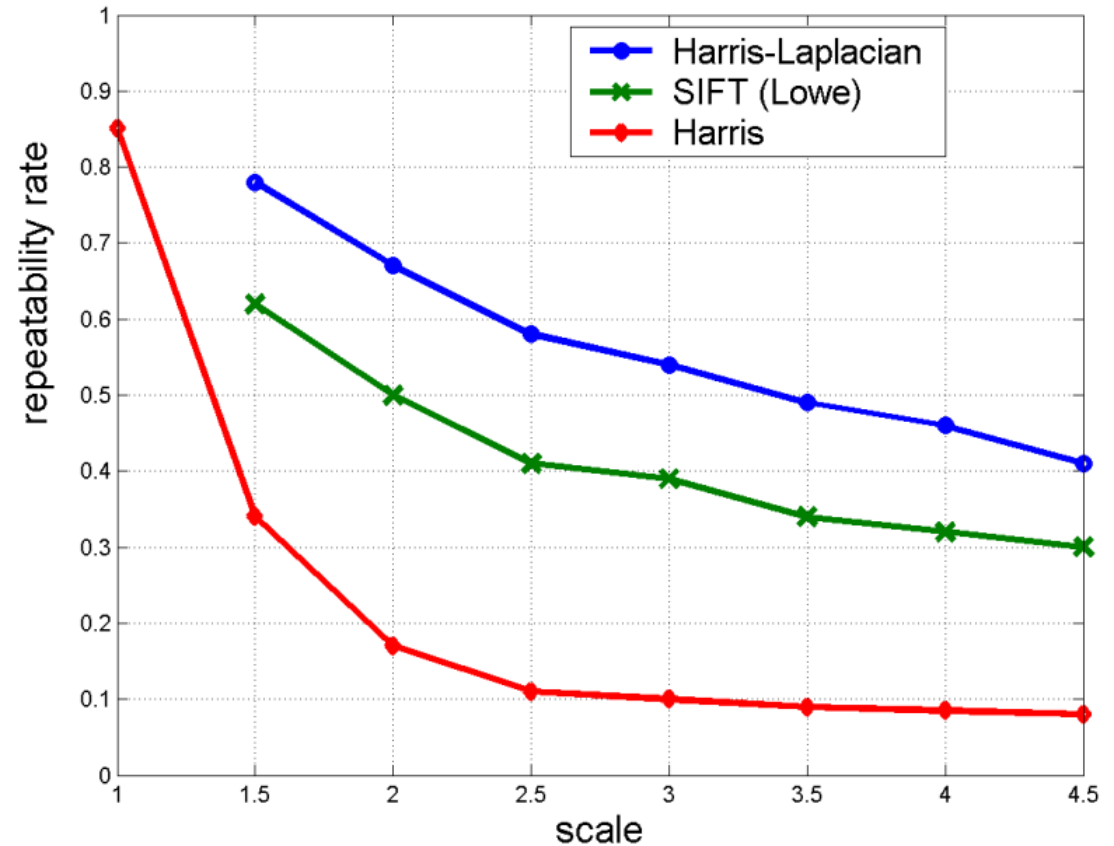
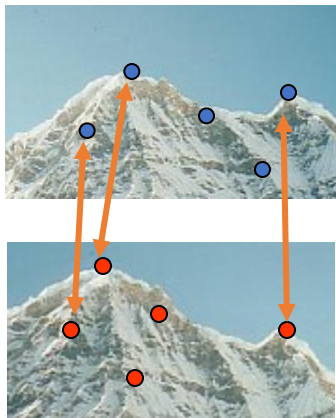
² D.Lowe. "Distinctive Image Features from Scale-Invariant Keypoints". IJCV 2004

Detectores Invariante a Escalamiento

- Evaluación experimental de detectores c.r. a cambio de escala

Tasa de repetibilidad:

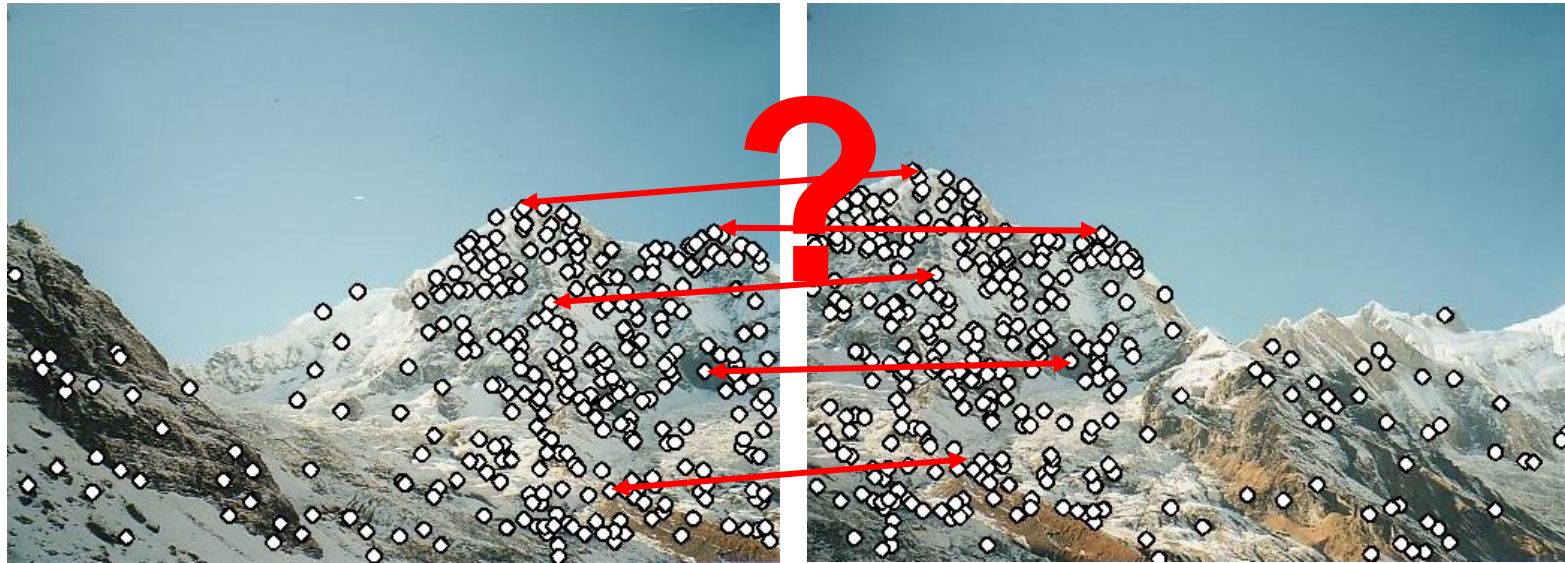
$$\frac{\# \text{ correspondencias}}{\# \text{ posibles correspondencias}}$$



Descriptores

- Sabemos cómo detectar puntos
- Luego:

¿Cómo los calzamos?



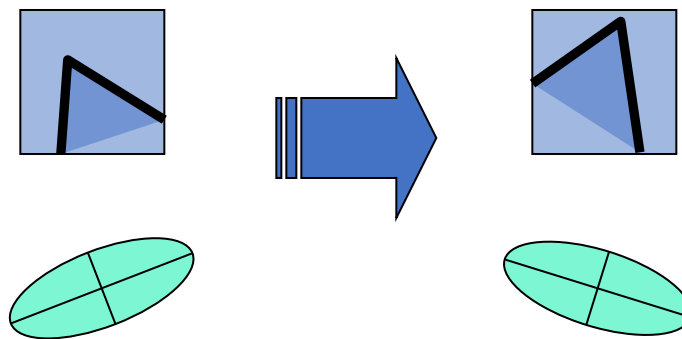
El descriptor debe ser:

1. Invariante
2. Distintivo
3. Abundante
4. Eficiente

Descriptores Invariantes a Rotación

- (1) Respuesta de Esquina Harris (R):
depende sólo de los valores propios de la matriz M

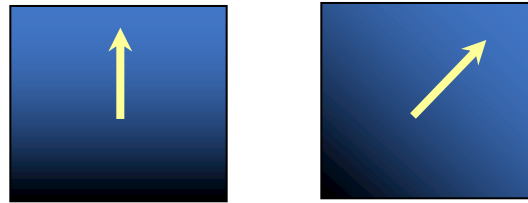
$$M = \sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$



Descriptores Invariantes a Rotación

- (3) Encontrar la orientación local

Dirección dominante del gradiente



- Calcular las derivadas de la imagen relativas a esta orientación

¹ K.Mikolajczyk, C.Schmid. "Indexing Based on Scale Invariant Interest Points". ICCV 2001

² D.Lowe. "Distinctive Image Features from Scale-Invariant Keypoints". IJCV 2004

Descriptores Invariantes a Escala

- Usar la escala determinada por el detector para calcular el descriptor en una ventana normalizada

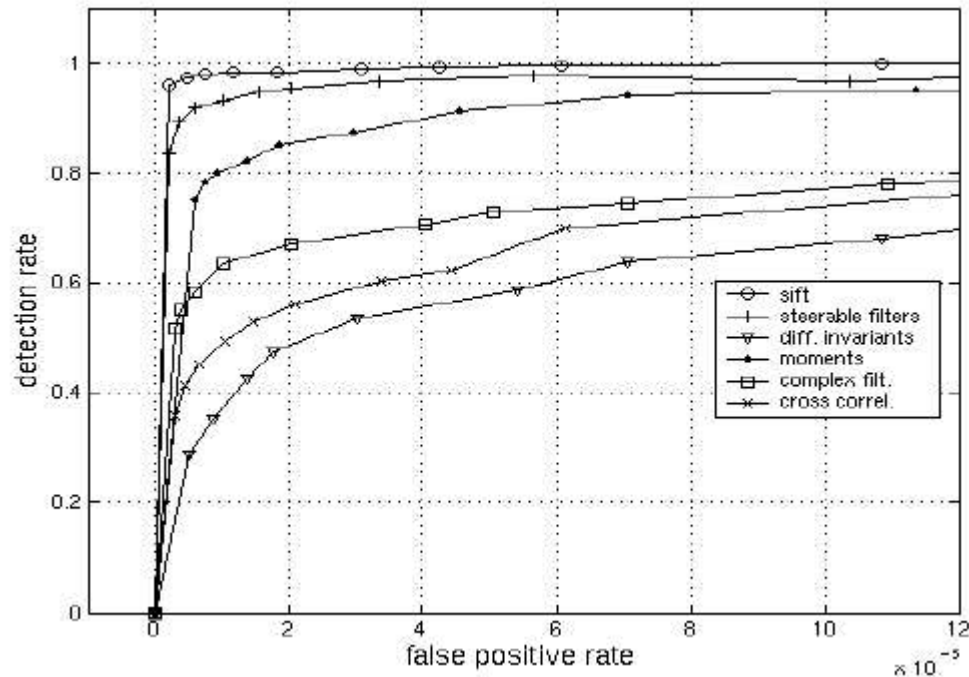
por Ejemplo:

- derivadas adaptadas a escala: S/x

SIFT – Scale Invariant Feature Transform¹

- Muestra empíricamente² muy buenos resultados. Invariante a *rotación, escalamiento, cambio de intensidad, y a transformaciones afines moderadas*

Escala = 2.5
Rotación = 45°

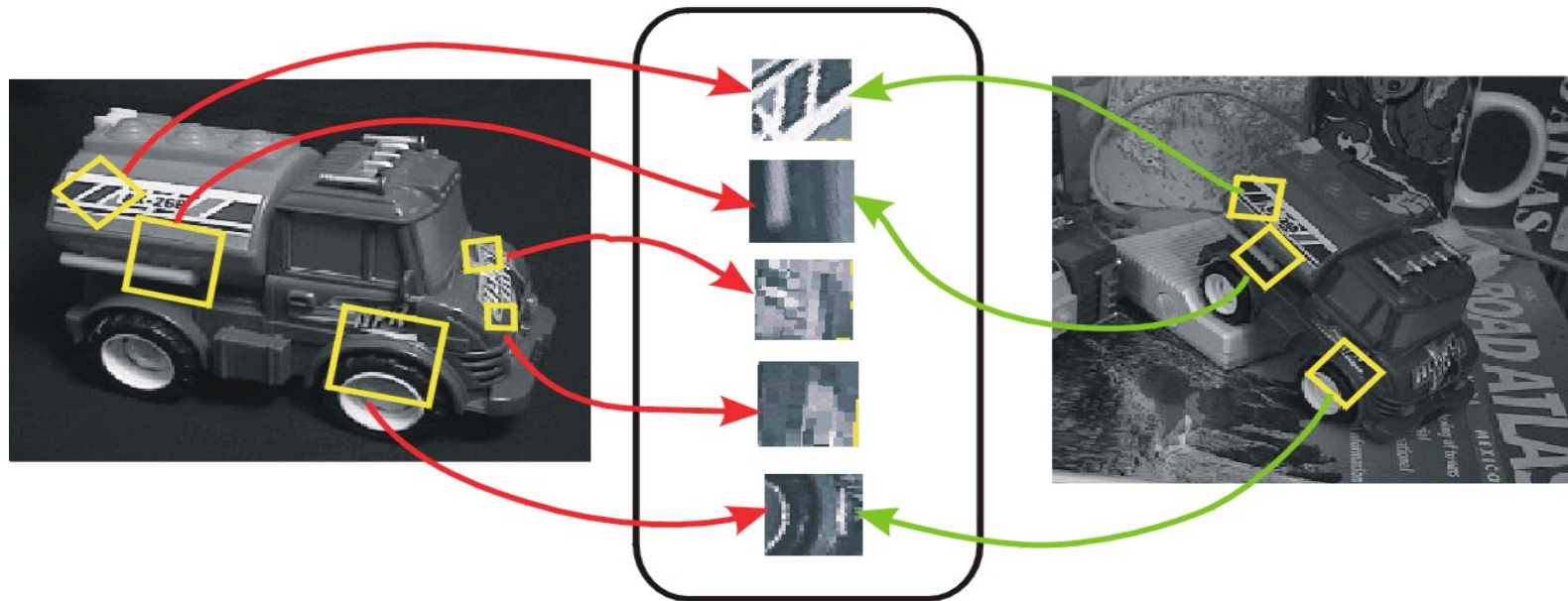


¹ D.Lowe. “Distinctive Image Features from Scale-Invariant Keypoints”. IJCV 2004

² K.Mikolajczyk, C.Schmid. “A Performance Evaluation of Local Descriptors”. CVPR 2003

Descriptores Locales Invariantes

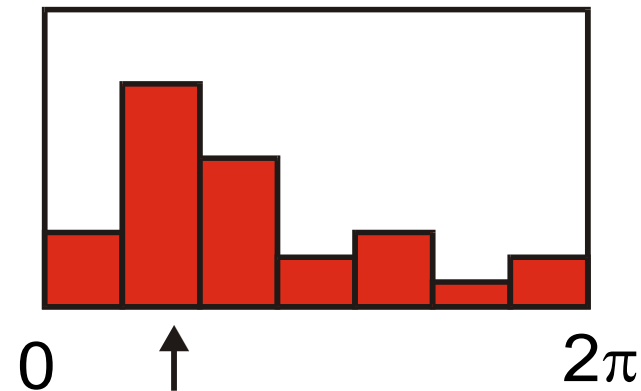
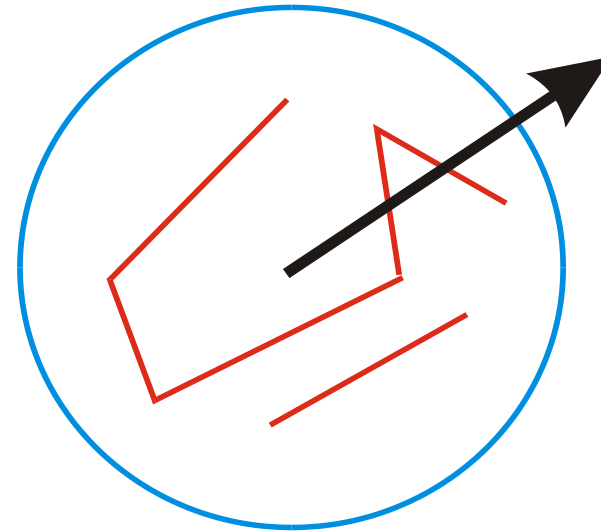
- El contenido de la imagen es transformado en puntos de interés con sus descriptores que son invariantes a traslación, rotación, escala, y otros parámetros de la formación de la imagen.



Regiones SIFT

Seleccionar orientación canónica

- Crear un histograma de direcciones de gradientes locales calculados en la escala seleccionada
- Asignar la orientación canónica al máximo del histograma suavizado
- Cada identificador especifica coordenadas estables (x , y , escala, orientación)



Ejemplo de Detección de Puntos de Interés

Umbral en el valor de los máximos de DOG y en la razón de curvaturas principales (similar a Harris)



- (a) 233x189 imagen
- (b) 832 máximos DOG
- (c) 729 quedan después de aplicar umbral a DOG
- (d) 536 quedan después de aplicar umbral a razón de curvaturas principales

Formación de Vector SIFT

- Se recopilan los gradientes a lo largo de una ventana de 16x16 en la escala seleccionada
- Se crea un arreglo de histogramas de orientaciones
- 8 orientaciones x 4x4 histogramas = 128 dimensiones

