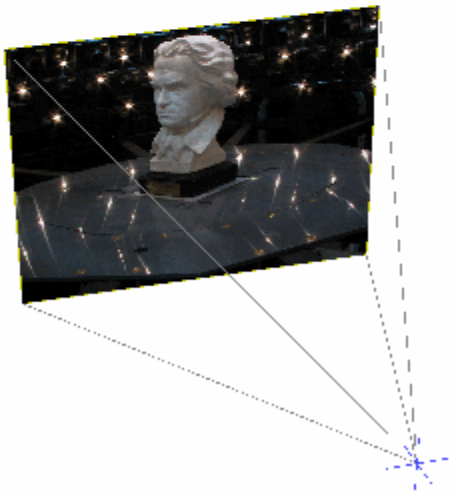
Visión por Computador

Tema 8. Geometría, estructura y forma



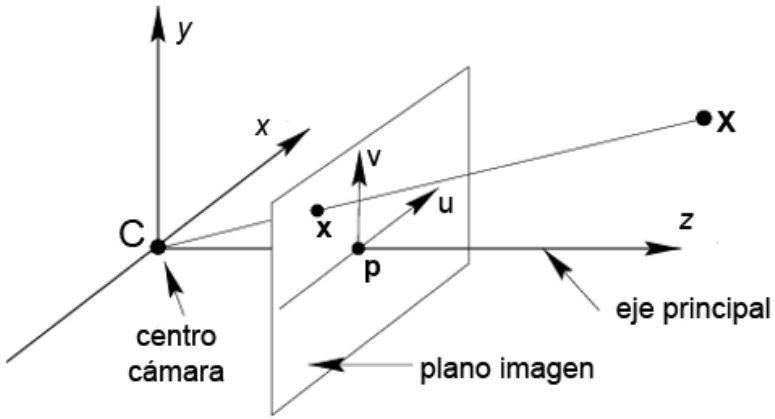
EII, GII. *Modesto Castrillón Santana*

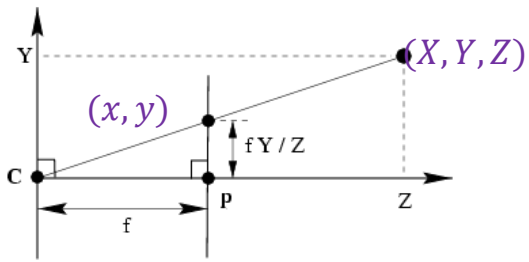
Contenidos

- Calibración geométrica
- Metrología
- Geometría epipolar
- Correspondencia
- Estructura desde movimiento
- Reconstrucción 3D

Perspectiva

Proyección de coordenadas de cámara al plano imagen





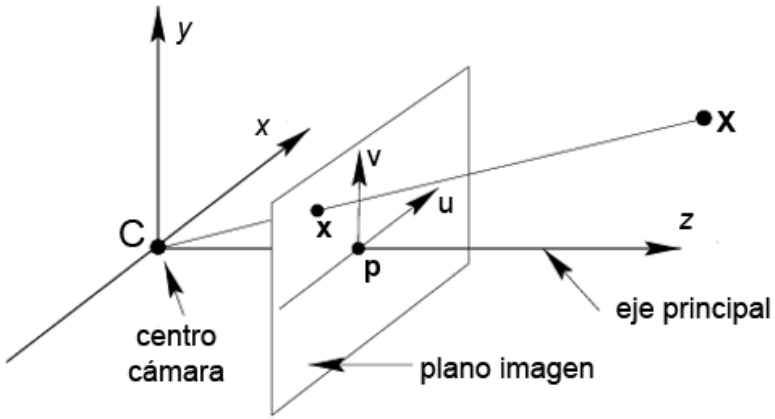
$$x = f\frac{X}{Z}, y = f\frac{Y}{Z}$$

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \cong \begin{pmatrix} fX \\ fY \\ Z \end{pmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

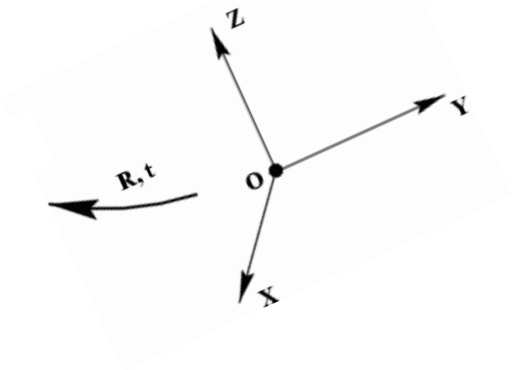
$\boldsymbol{x}$	$\boldsymbol{P}$	$\boldsymbol{X}$
Punto de imagen en coordenadas homogéneas	Matriz de proyección	Punto 3D en coordenadas homogéneas

Calibración

Estimar transformación de coordenadas del mundo a cámara

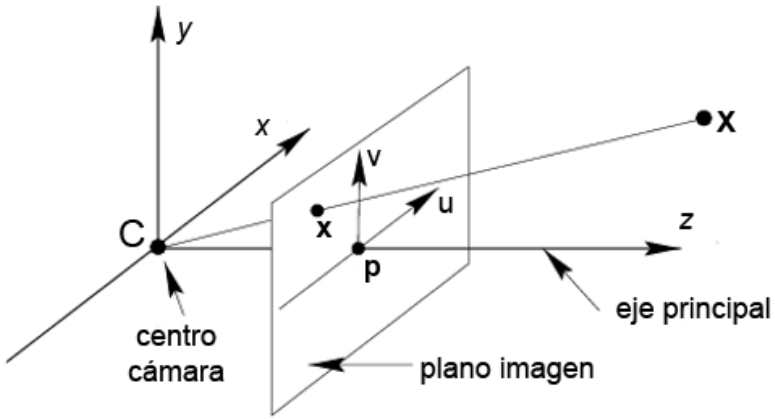


sistema de coordenadas del mundo

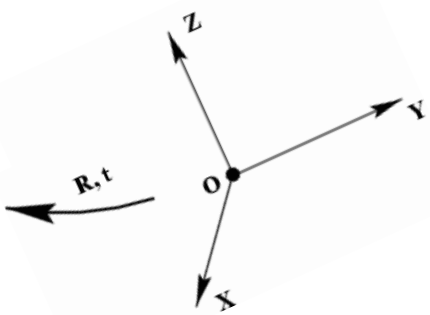


Calibración

Estimar transformación de coordenadas del mundo a cámara



sistema de coordenadas del mundo



Coordenadas homogéneas

$$\begin{pmatrix} \text{Punto} \\ \text{2D} \\ \text{(3x1)} \end{pmatrix} \cong \begin{pmatrix} \text{Matriz de} \\ \text{transformación} \\ \text{cámara-píxeles} \\ \boldsymbol{K}\ \text{(3x3)} \end{pmatrix} \begin{pmatrix} \textit{Matriz de} \\ \textit{proyección} \\ [\boldsymbol{I} \mid \boldsymbol{0}]\ \text{(3x4)} \end{pmatrix} \begin{pmatrix} \text{Matriz de} \\ \text{transformación} \\ \text{mundo-cámara} \\ \begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ \boldsymbol{0}^T & 1 \end{bmatrix}\ \text{(4x4)} \end{pmatrix} \begin{pmatrix} \text{Punto} \\ \text{3D} \\ \text{(4x1)} \end{pmatrix}$$

$\boldsymbol{x}$

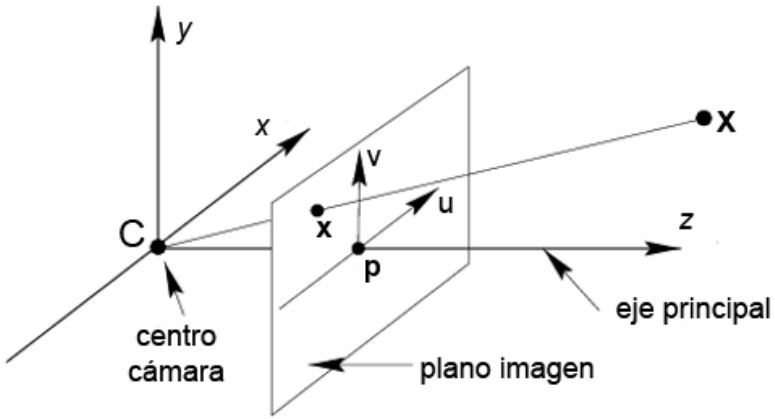
Parámetros ***intrínsecos***: punto principal, escala, ...

Parámetros ***extrínsecos***: rotación, traslación

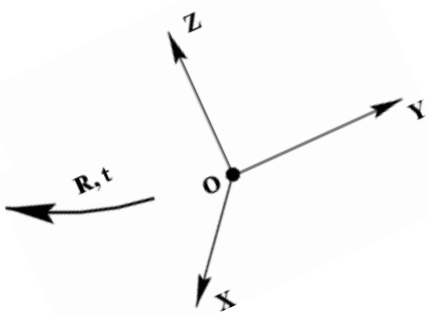
$\boldsymbol{X}$

Calibración

Estimar transformación de coordenadas del mundo a cámara



sistema de coordenadas del mundo



Coordenadas homogéneas

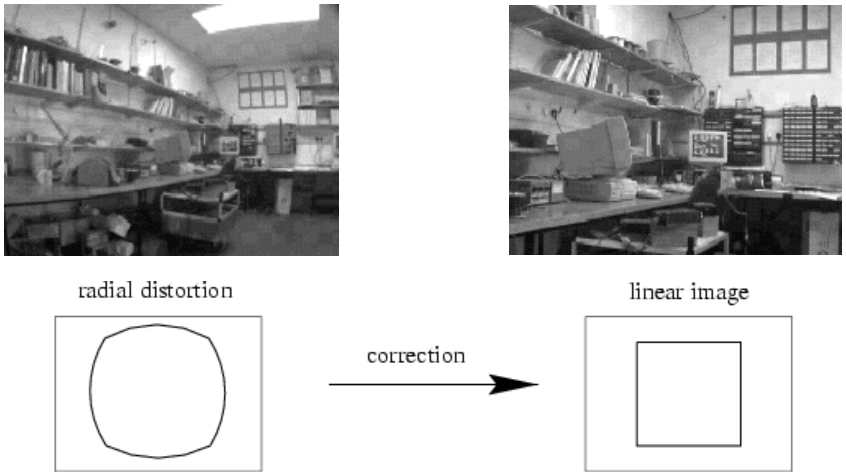
$$\underbrace{\begin{pmatrix}\text{Punto}\\ \text{2D}\\ \text{(3x1)}\end{pmatrix}}_{\boldsymbol{x}} \cong \underbrace{\begin{pmatrix}\text{Matriz de}\\ \text{transformación}\\ \text{cámara-píxeles}\\ \boldsymbol{K}\ \text{(3x3)}\end{pmatrix}\begin{pmatrix}\textit{Matriz de}\\ \textit{proyección}\\ [\boldsymbol{I}\mid\boldsymbol{0}]\ \text{(3x4)}\end{pmatrix}\begin{pmatrix}\text{Matriz de}\\ \text{transformación}\\ \text{mundo-cámara}\\ \begin{bmatrix}\boldsymbol{R} & \boldsymbol{t}\\ \boldsymbol{0}^T & 1\end{bmatrix}\ \text{(4x4)}\end{pmatrix}}_{\boldsymbol{P}=\boldsymbol{K}[\boldsymbol{R}|\boldsymbol{t}]}\underbrace{\begin{pmatrix}\text{Punto}\\ \text{3D}\\ \text{(4x1)}\end{pmatrix}}_{\boldsymbol{X}}$$

Matriz de proyección de cámara

Intrínsecos

Punto principal
Focal
Escalado
Píxeles no rectangulares
Distorsión radial

$$\boldsymbol{K} = \begin{bmatrix} m_x & 0 & 0 \\ 0 & m_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & \beta_x \\ 0 & \alpha_y & \beta_y \\ 0 & 0 & 1 \end{bmatrix}$$



Intrínsecos

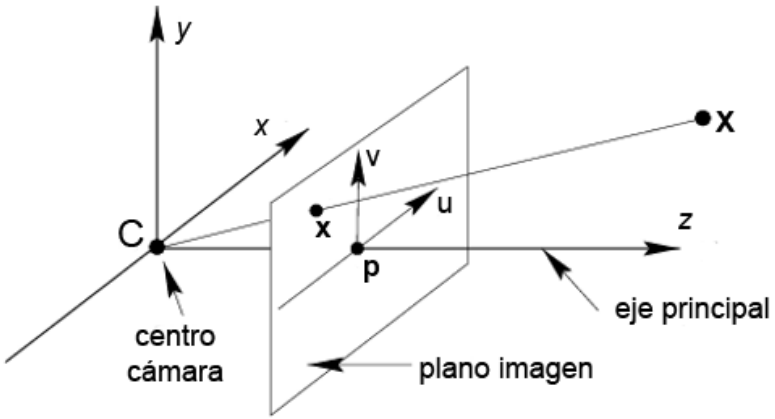
Punto principal

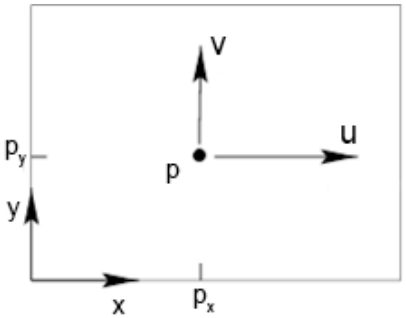
Intersección eje principal con plano imagen

Origen en coordenadas de la cámara, $p=(p_x,p_y)$

Coordenadas de la imagen, origen esq. sup. izquierda

Suele asumirse el centro de la imagen





$$\begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Matriz de calibración $\boldsymbol{K}$

Matriz de Proyección $[\boldsymbol{I} \mid \boldsymbol{0}]$

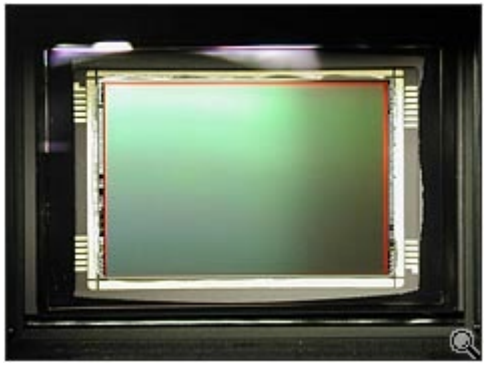
$$\boldsymbol{P} = \boldsymbol{K}[\boldsymbol{I}|\boldsymbol{0}]$$

Intrínsecos

Factor de escala

Mapeo del mundo a unidades de píxeles

m_x y m_y píxeles por metro



Tamaño del píxel (m): $\frac{1}{m_x} \times \frac{1}{m_y}$

Factores de escala	Matriz de calibración en unidades métricas		Matriz de calibración en píxeles
$\begin{bmatrix} m_x & 0 & 0 \\ 0 & m_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix}$	$=$	$\begin{bmatrix} \alpha_x & 0 & \beta_x \\ 0 & \alpha_y & \beta_y \\ 0 & 0 & 1 \end{bmatrix}$
píxeles/m	m		píxeles

Extrínsecos

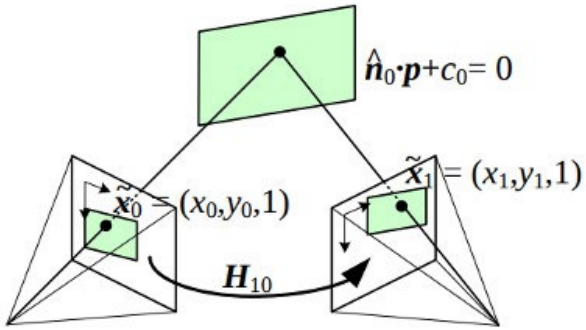
Homografía

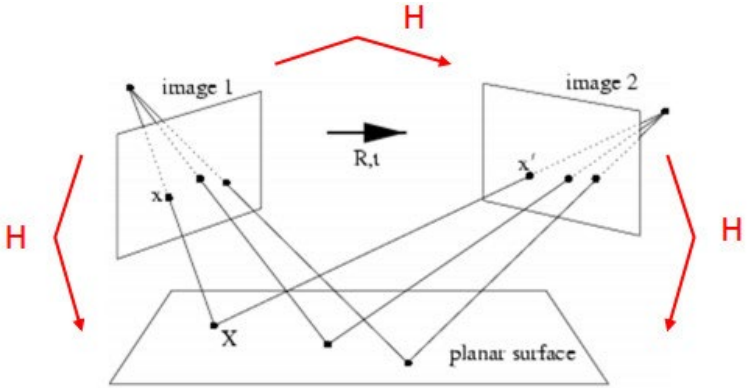
Relaciona puntos sobre un plano en dos imágenes

Matriz 3x3, nueve valores, pero 8 incógnitas

Necesarios al menos 4 puntos en ambas imágenes

Cada par, dos ecuaciones

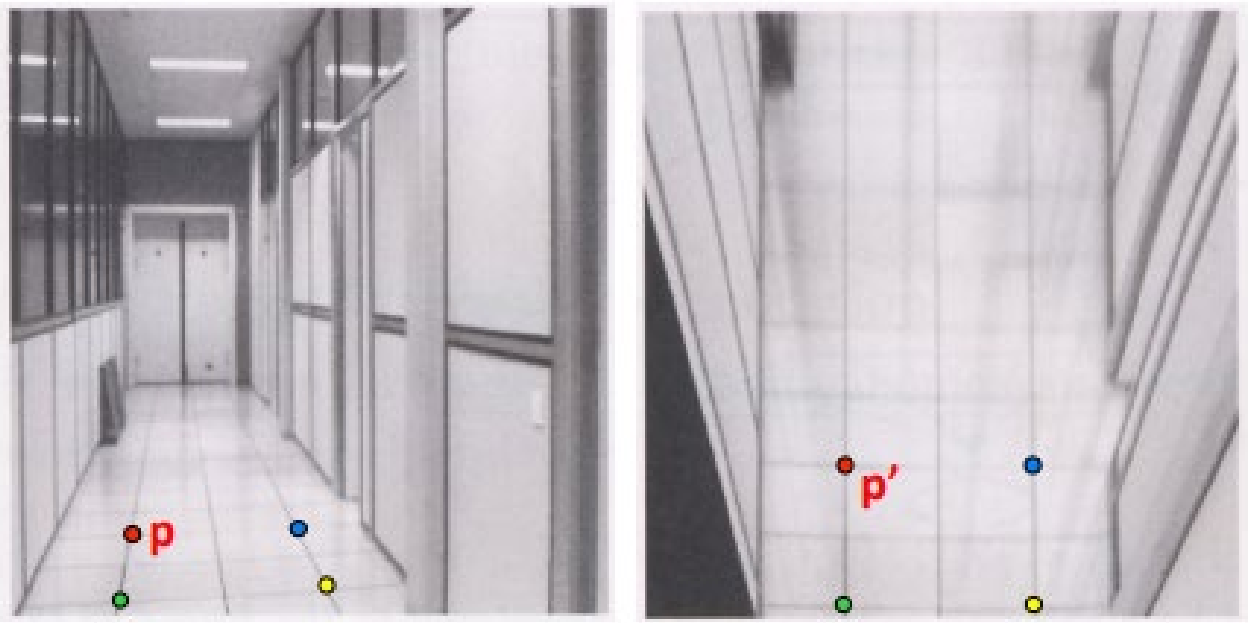




$$s \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Extrínsecos

Homografía



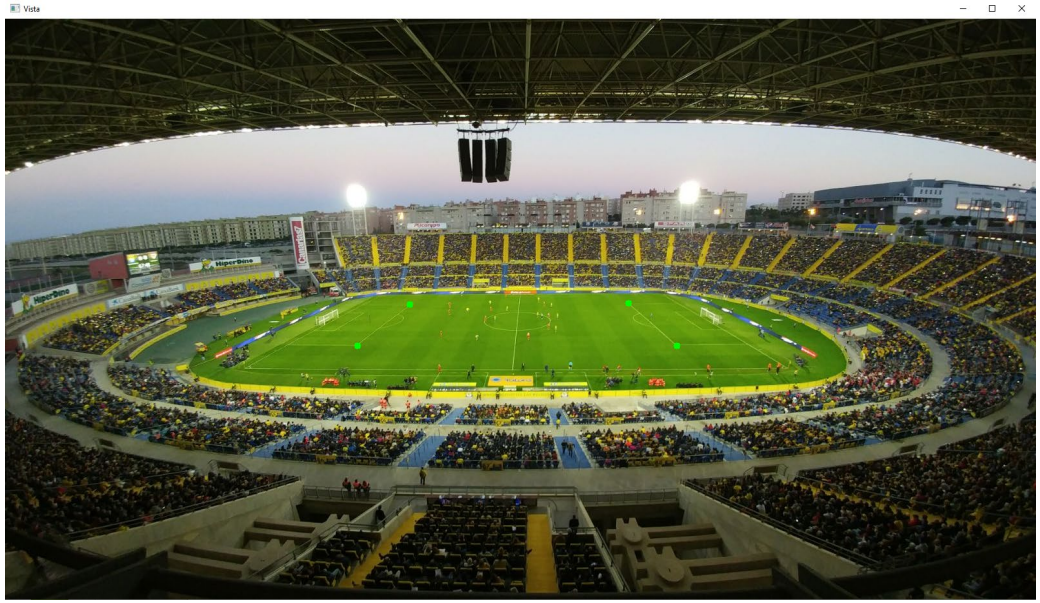
Extrínsecos

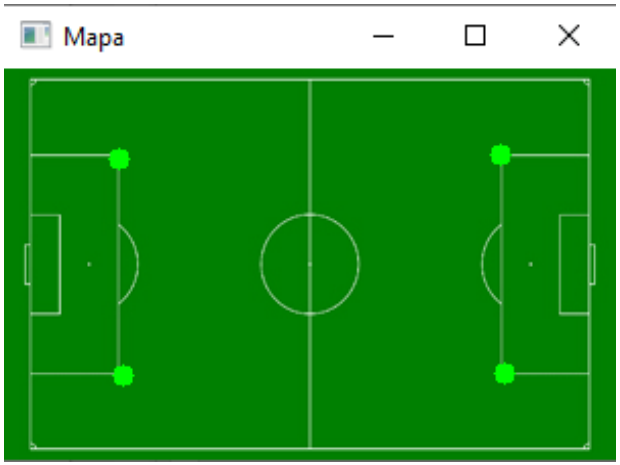
Homografía

Extrínsecos

Homografía

```
from skimage import transform
tform = transform.estimate_transform('projective', np.array(puntosA), np.array(puntosB))
tf_img = transform.warp(vista, tform.inverse, output_shape=(mapa.shape))
```



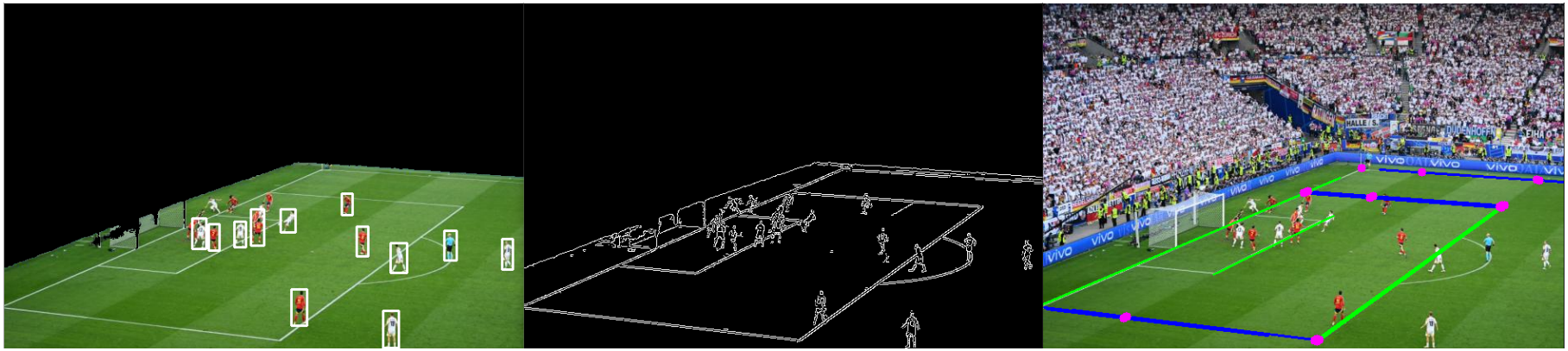


Extrínsecos

Homografía

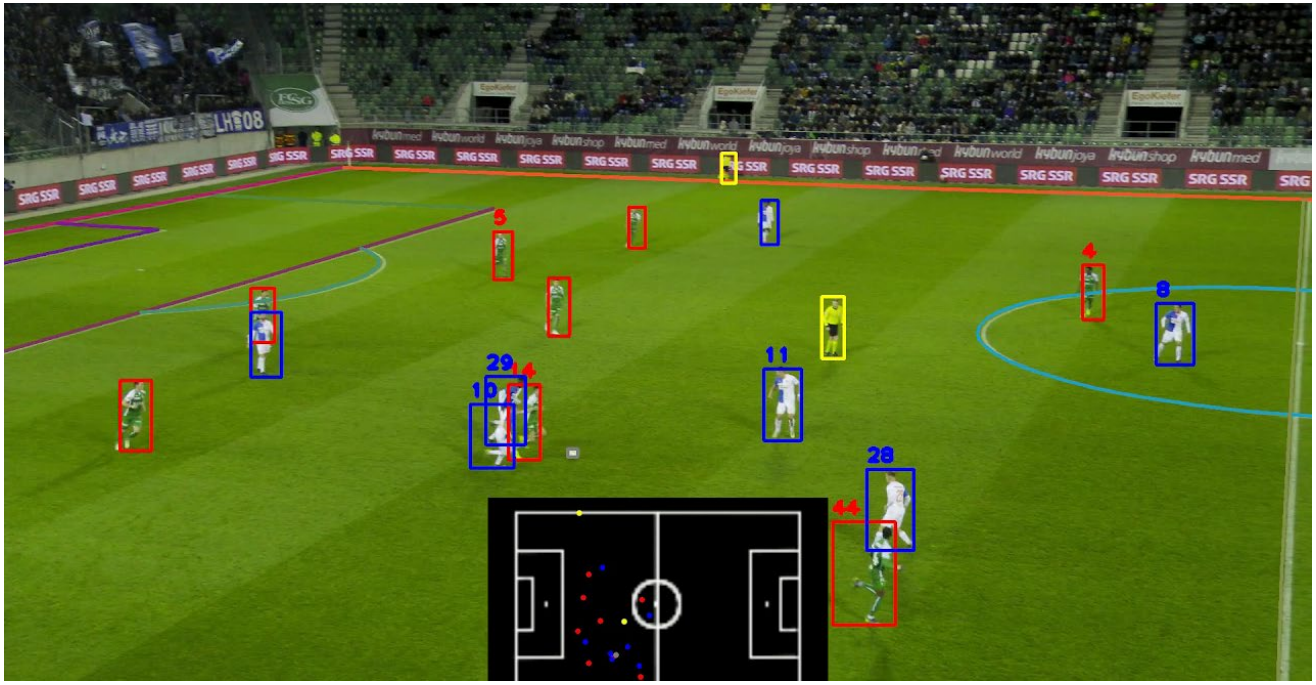
Cámara no estática

Automatizar



Extrínsecos

SoccerNet 2025 Challenges, ¿TFG?



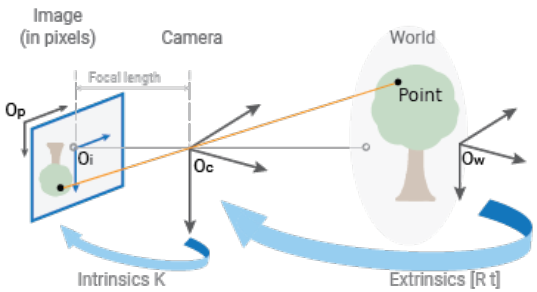
Extrínsecos

Calibrar la cámara, tb conocido como estimación de pose

Relacionar sistemas de coordenadas de cámara y mundo

Determinar traslación y rotación

Seis incógnitas



$$\widetilde{\boldsymbol{X}}_{\text{cam}} = \boldsymbol{R}\left(\widetilde{\boldsymbol{X}} - \widetilde{\boldsymbol{C}}\right) = \boldsymbol{R}\widetilde{\boldsymbol{X}} + \boldsymbol{t}$$

Coordenadas de un punto en la cámara

Matriz de rotación 3x3

Coordenadas de un punto en el mundo

Coordenadas del centro de la cámara en el mundo

Extrínsecos

Método lineal, ***Direct Linear Transform*** (DLT)

Homografía entre dos sistemas de coordenadas

Al menos seis pares de puntos (no coplanares)

Ecuaciones por cada correspondencia 3D-2D

Resolución con **Descomposición en Valores Singulares**

Solución lineal proporciona la matriz de transformación

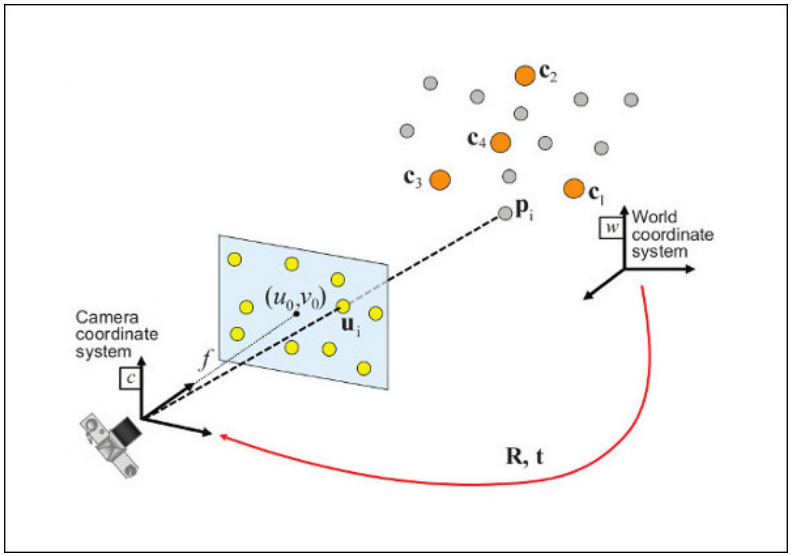
Posteriormente parámetros intrínsecos

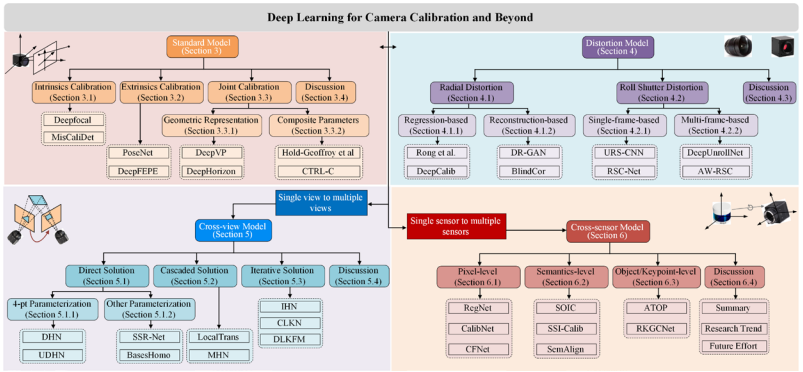
Extrínsecos

Varias técnicas en OpenCV con *solvePnP*

Integración de RANSAC en el proceso

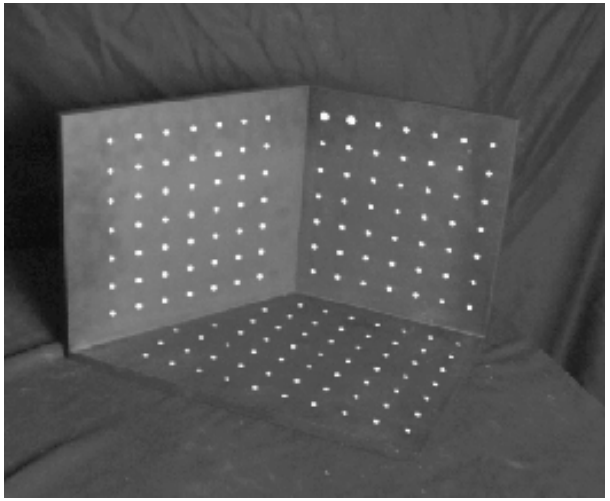
Actualmente tb DL

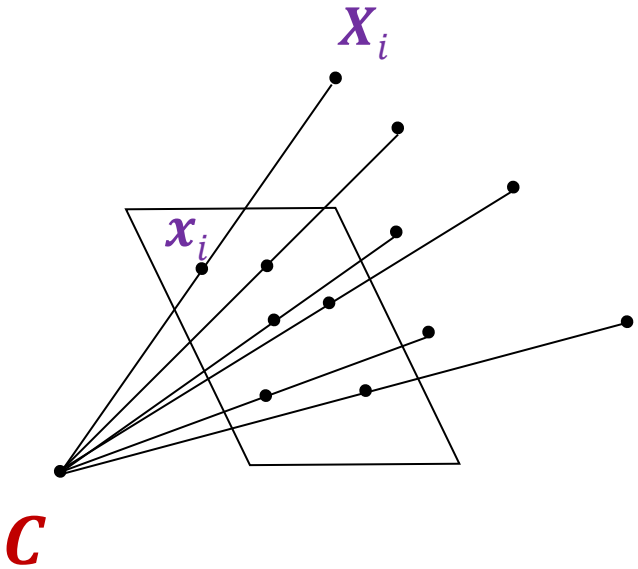




Extrínsecos

Dados *n* puntos (>=6), estimar parámetros de cámara





$$\boldsymbol{x} \cong \boldsymbol{K}[\boldsymbol{R} \;\; \boldsymbol{t}]\boldsymbol{X}$$

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \cong \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Extrínsecos

Dados *n* puntos (>=6), estimar parámetros de cámara

J. Hays

Coordenadas 2D en imagen conocidas	Posiciones 3D conocidas
880 214	312.747 309.140 30.086
43 203	305.796 311.649 30.356
270 197	307.694 312.358 30.418
886 347	310.149 307.186 29.298
745 302	311.937 310.105 29.216
943 128	311.202 307.572 30.682
476 590	307.106 306.876 28.660
419 214	309.317 312.490 30.230
317 335	307.435 310.151 29.318
783 521	308.253 306.300 28.881
235 427	306.650 309.301 28.905
665 429	308.069 306.831 29.189
655 362	309.671 308.834 29.029
427 333	308.255 309.955 29.267
412 415	307.546 308.613 28.963
746 351	311.036 309.206 28.913
434 415	307.518 308.175 29.069
525 234	309.950 311.262 29.990
716 308	312.160 310.772 29.080
602 187	311.988 312.709 30.514

Extrínsecos

Soluciones no lineales

Plantean función objetivo sobre SSD (*sum of squared distances*)

$$\sum_i \|\mathrm{proj}(\boldsymbol{K}[\boldsymbol{R} \mid \boldsymbol{t}]\boldsymbol{X_i}) - \boldsymbol{x}_i\|_2^2$$

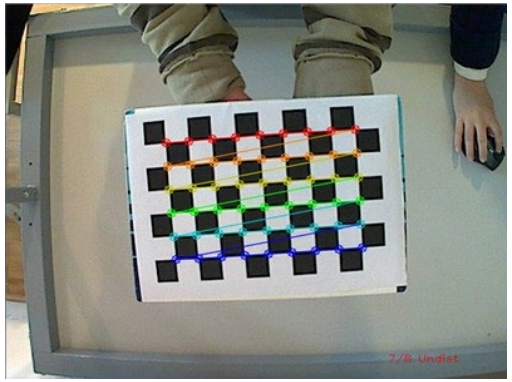
Integran distorsiones e incluyen restricciones, p.e. focal

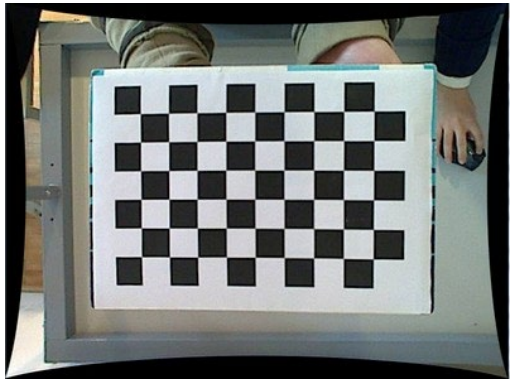
Minimización del error

Frecuente inicialización con solución lineal

Calibración

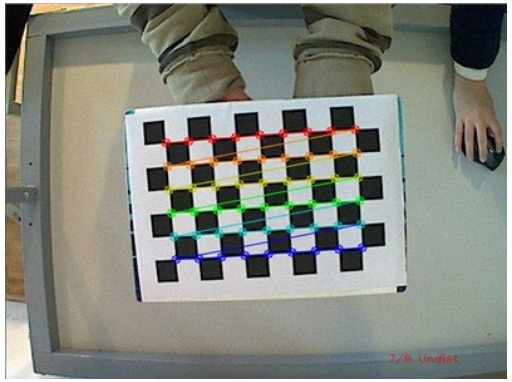
OpenCV con patrón de geometría conocida
Matriz de transformación (3x3) y distorsión

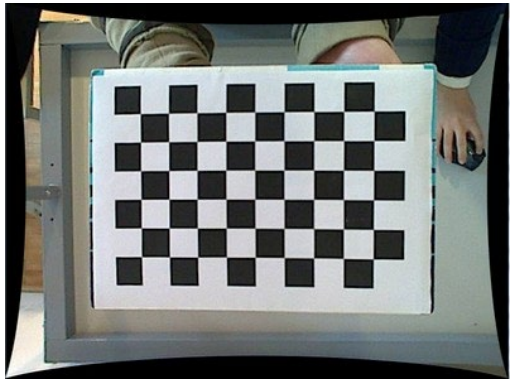


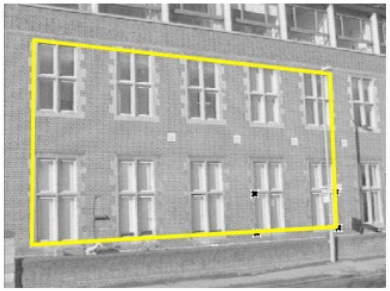


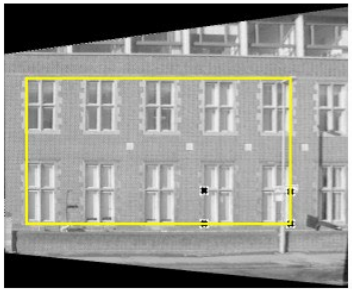
Calibración

OpenCV con patrón de geometría conocida
Matriz de transformación (3x3) y distorsión









from Hartley & Zisserman

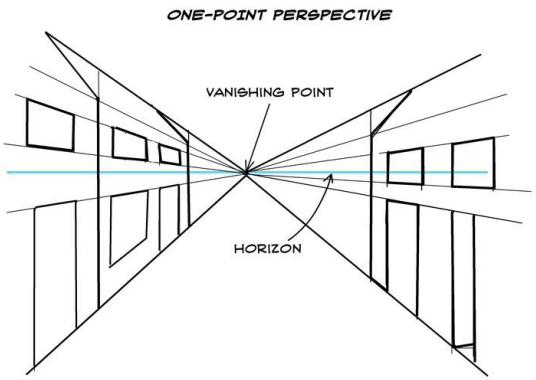
Si no hay patrón, buscar estructura en imagen

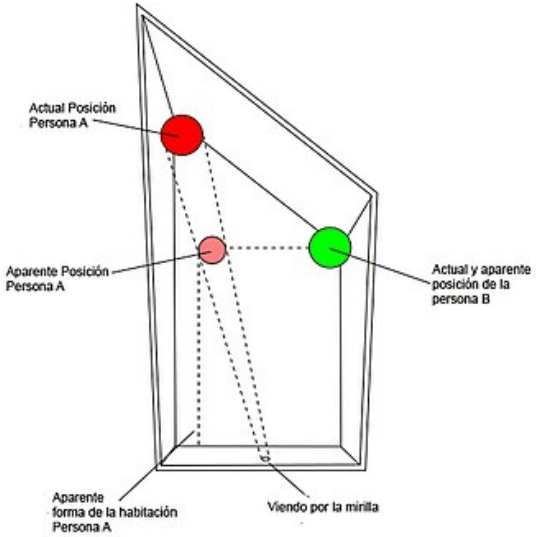
Extrínsecos

Deporte

Metrología

Estimar información 3D de una imagen
Perspectiva
Puntos de fuga
Objetos de referencia





Habitación de Ames (1946)

Metrología

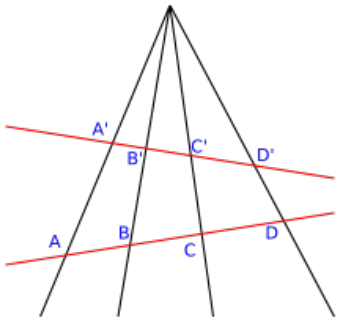
Invarianzas de la perspectiva

- Las líneas rectas permanecen
- Razón anarmónica o doble (*cross-ratio*)

Sean A, B, C y D colineales

$$CR(A,B,C,D) = \frac{|C-A||D-B|}{|C-B||D-A|}$$

Dos grupos de puntos relacionados tienen el mismo CR



Metrología

Invarianzas de la perspectiva

- Las líneas rectas permanecen
- Razón anarmónica o doble (*cross-ratio*)

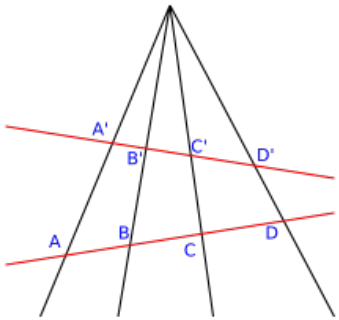
Sean A, B, C y D colineales

$$CR(A,B,C,D) = \frac{|C-A||D-B|}{|C-B||D-A|}$$

Dos grupos de puntos relacionados tienen el mismo CR

Si D (o A) están en ∞, se simplifica

$$CR(A,B,C,D) = \frac{|C-A|}{|C-B|}$$



Metrología

Puntos de fuga

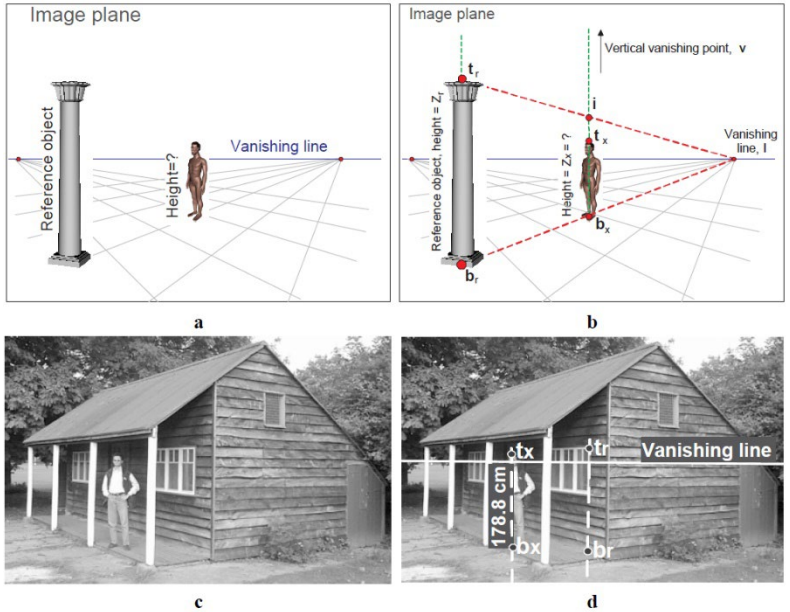
Objetos artificiales como edificios, etc.

Intersectar líneas 2D que se corresponden a paralelas

Obtención de puntos de fuga

Aproxima focal

Permite estimar medidas



Antonio Criminisi. Single-View Metrology: Algorithms and Applications, 2002

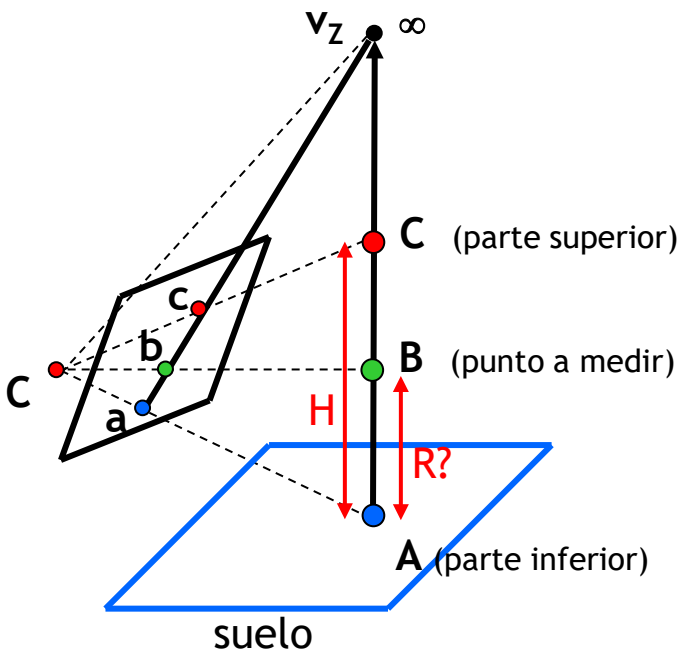
Metrología

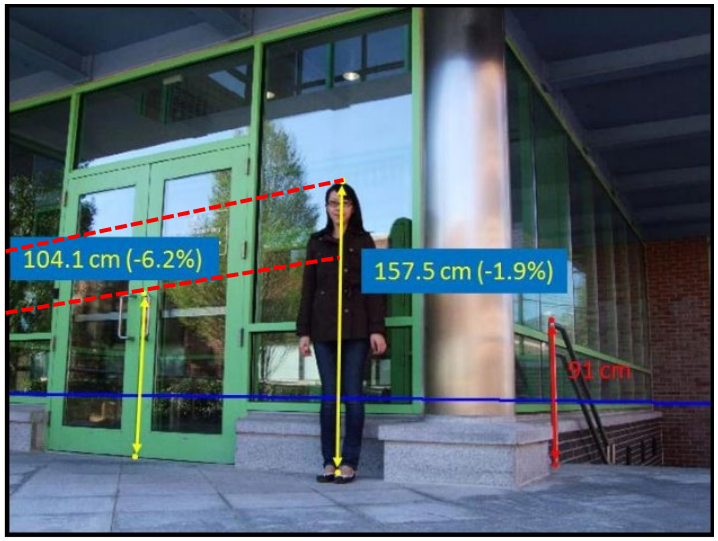
Midiendo alturas

Relación razón anarmónica en escena e imagen

Puntos de fuga en ∞ (v_x, v_y, v_z)

$$CR(A,B,C,D) = \frac{|C-A|}{|C-B|} = \frac{H}{R}$$





$$\frac{|c-a||v_z-b|}{|c-b||v_z-a|} = \frac{|c-a|}{|c-b|} = \frac{H}{R}$$

Metrología

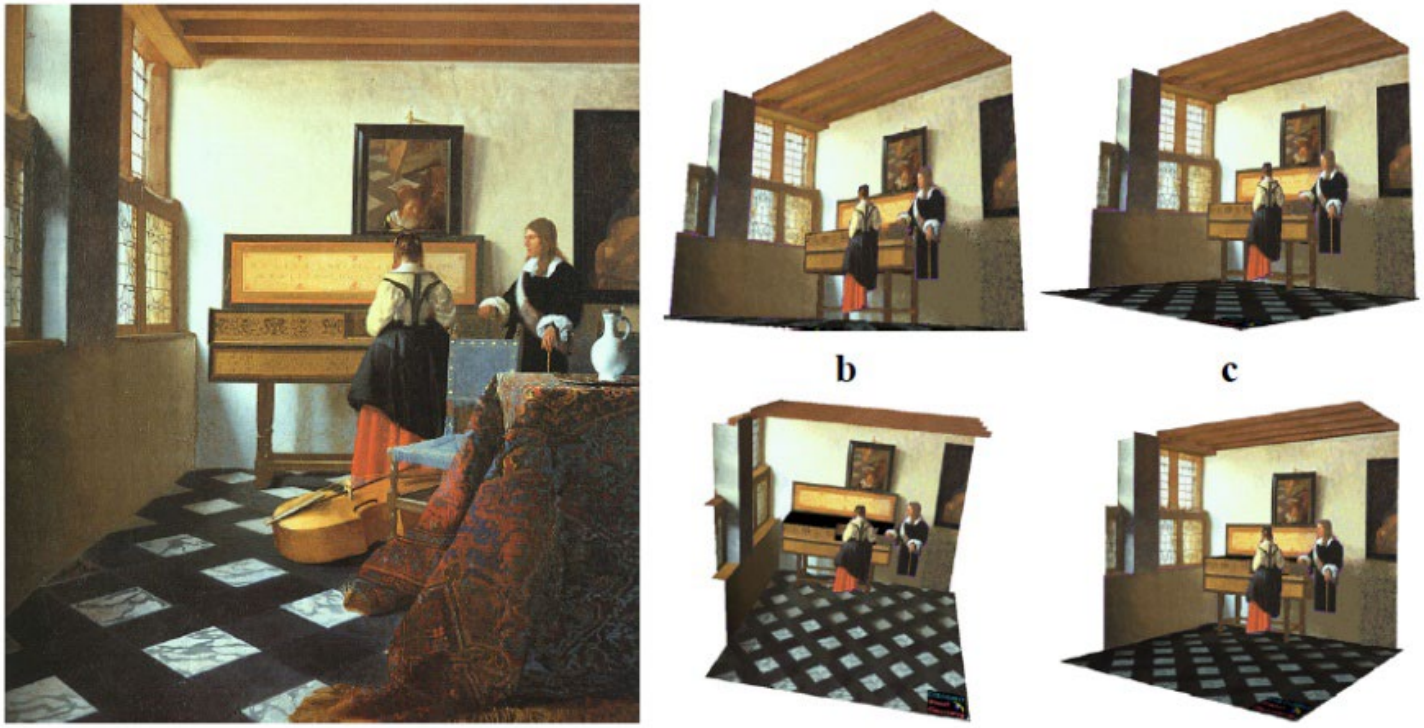
Puntos de fuga

1. Detección de bordes (Canny)
2. Repetir
   - Tomar dos segmentos y calcular intersección p
   - S_p conjunto de segmentos que pasan por p
3. Escoger p con mayor S_p como punto de fuga dominante
4. Eliminar segmentos que tocan a p, volver a 2

Antonio Criminisi. Single-View Metrology: Algorithms and Applications, 2002

Metrología

Cambiar la vista del espacio



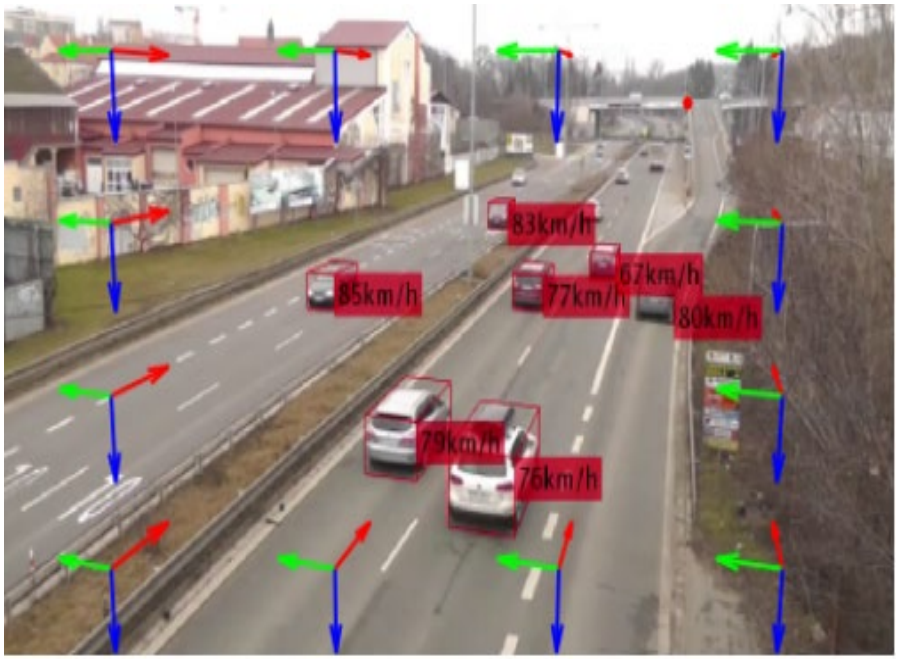
Antonio Criminisi. Single-View Metrology: Algorithms and Applications, 2002

Casos de uso

Realidad aumentada

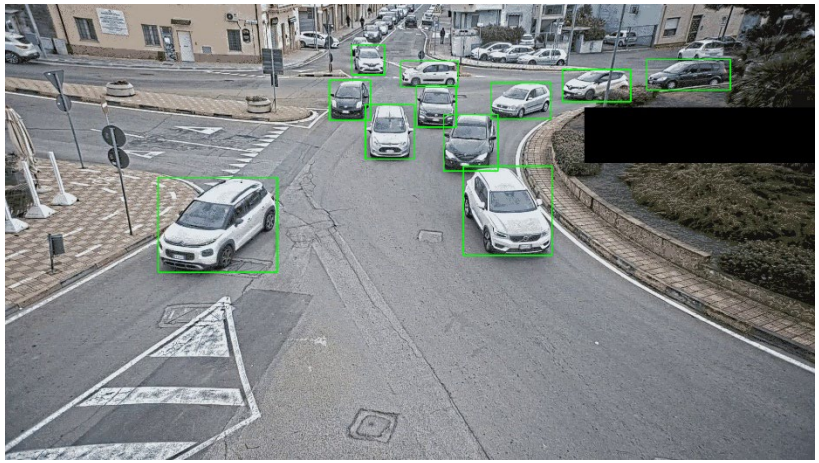
Casos de uso

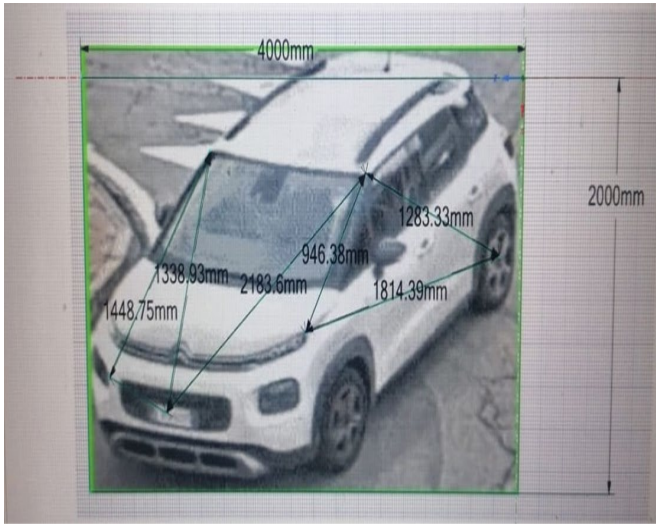
Determinación posición 3D vehículos
Puntos de fuga en rectas



Casos de uso

Determinación posición 3D vehículos
No siempre rectas
Modelos concretos para calibración



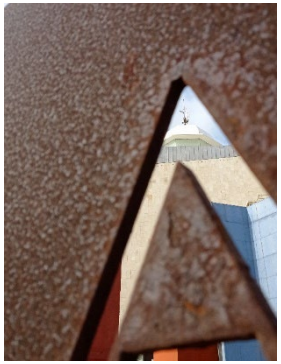


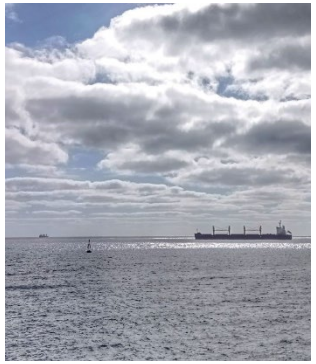
Profundidad monocular

A partir de una imagen

Diversas pistas:

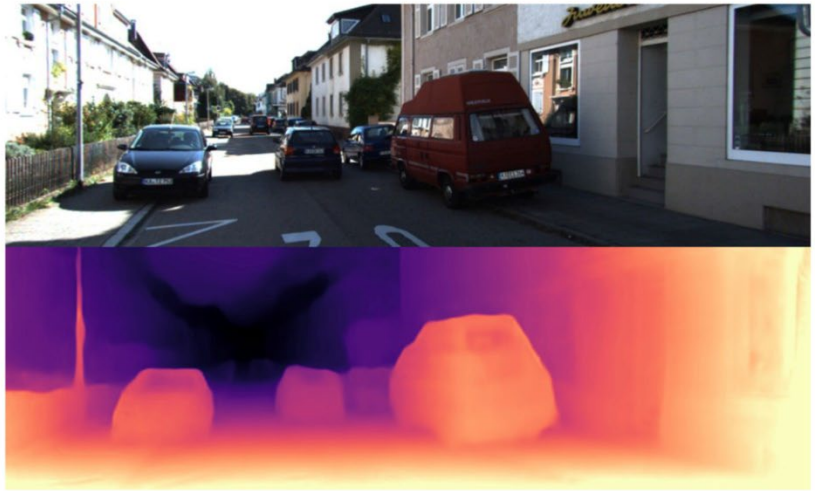
- Intensidad (*shading*)
- Perspectiva
- Tamaño relativo
- Tamaño conocido
- Atenuación atmosférica
- Oclusión
- Gradientes de textura





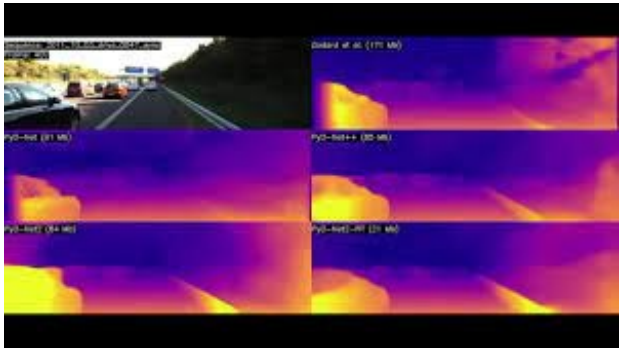
Profundidad monocular

Métodos supervisados
Sensores que captan profundidad y RGB alineado
RGB-D, LIDAR



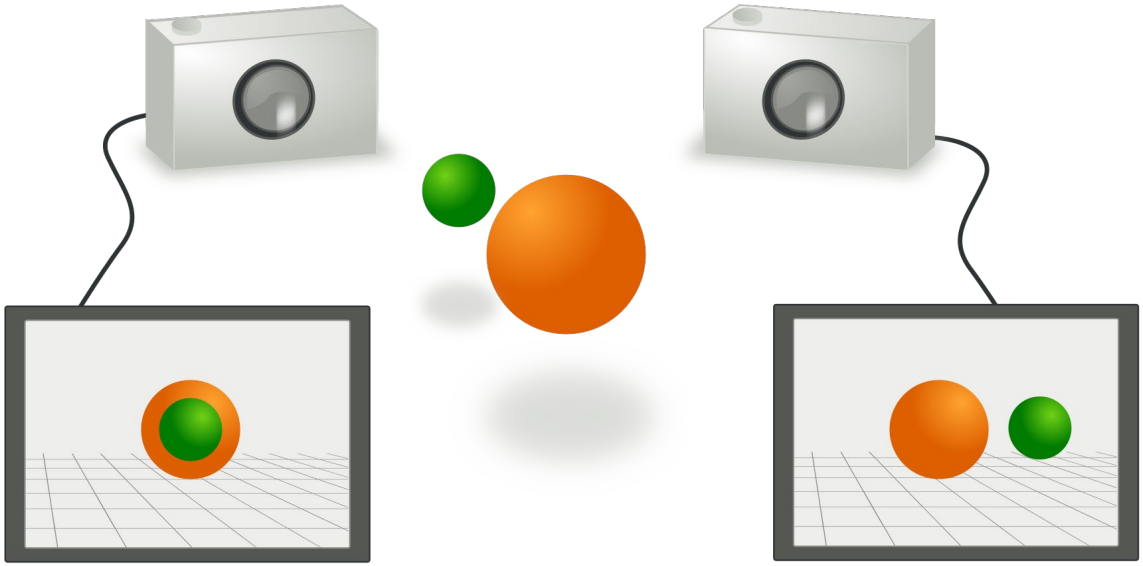
Profundidad monocular

Métodos supervisados



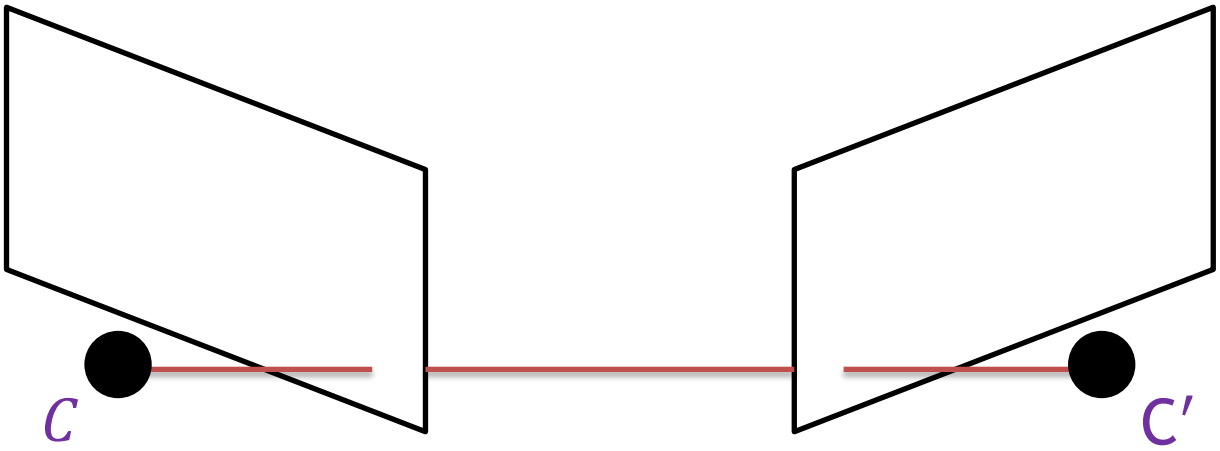
Epipolar

Dos cámaras, una escena, dos imágenes
Geometría epipolar



Epipolar

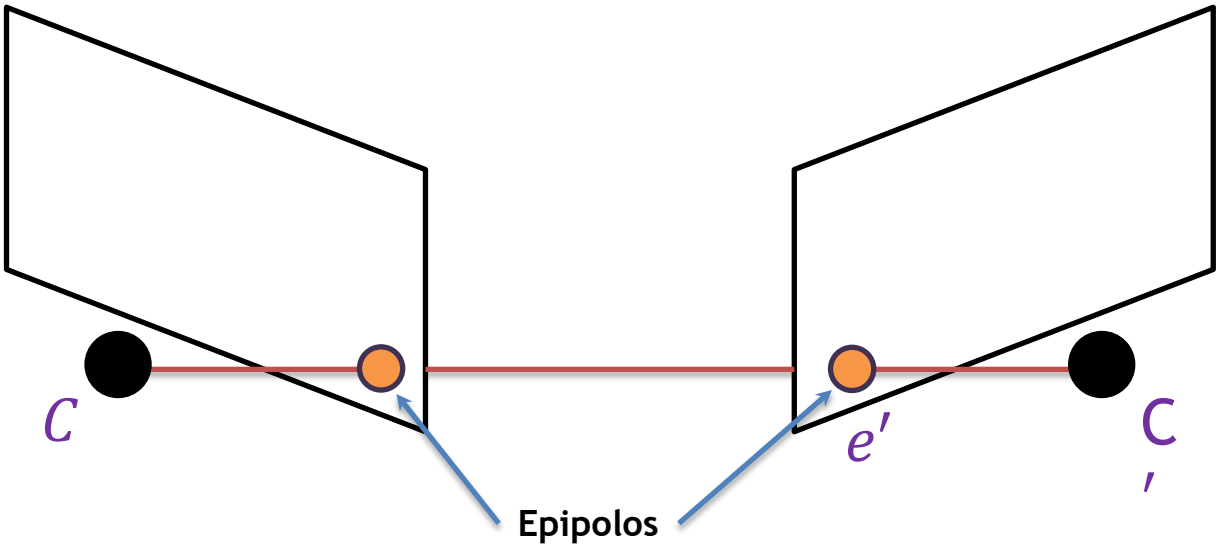
Dos cámaras y la recta que une sus centros C y C’



Imágenes adaptadas de S. Lazebnik

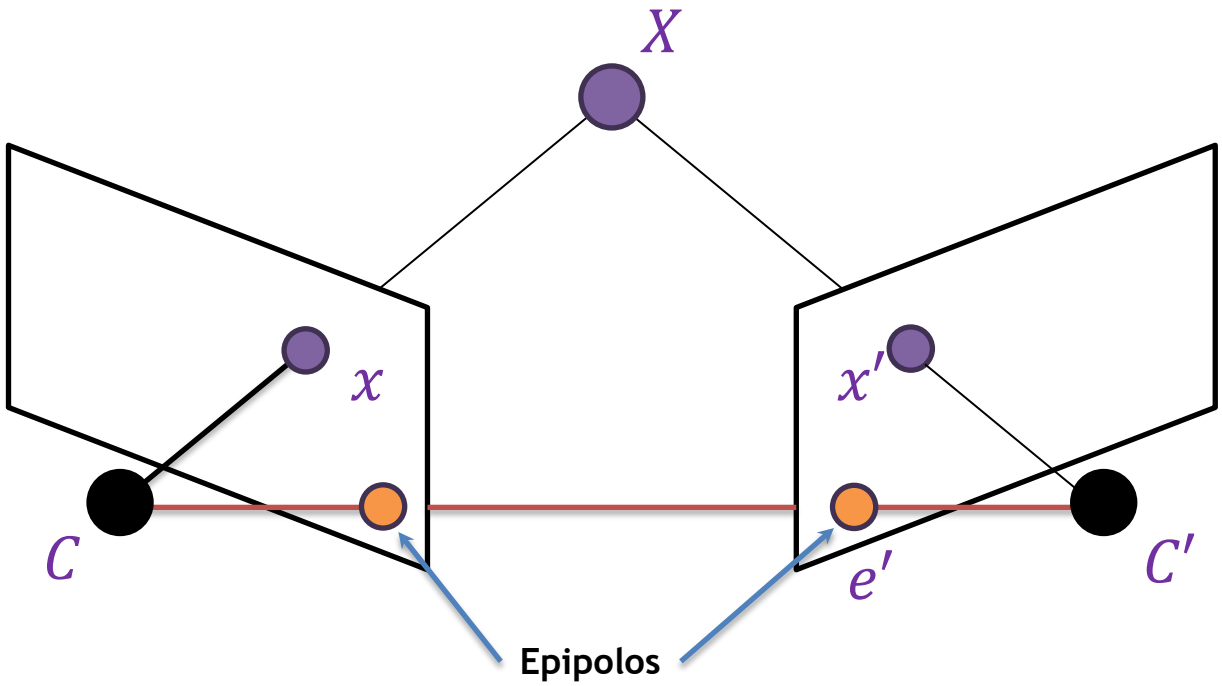
Epipolar

Dos cámaras y la recta que une sus centros C y C’
Epipolos: intersección recta con planos imagen
Cada epipolo es la proyección de la otra cámara



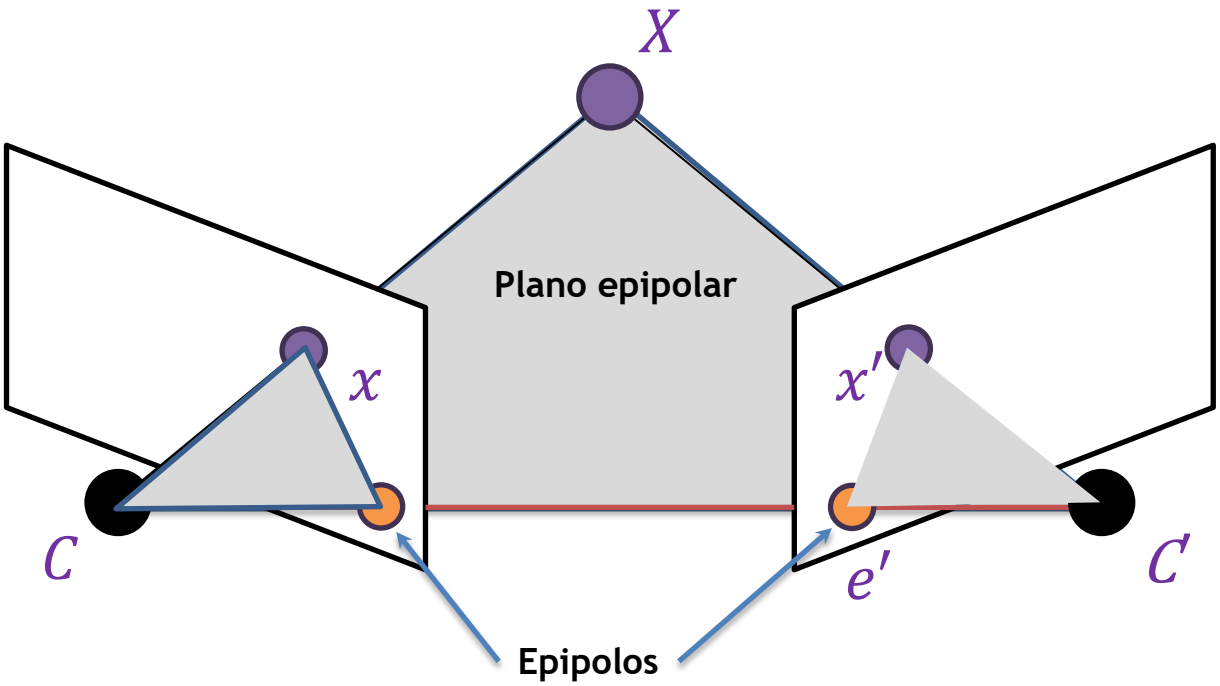
Epipolar

Dos cámaras y la recta que une sus centros C y C’
Dado un punto X, y sus proyecciones en cada cámara



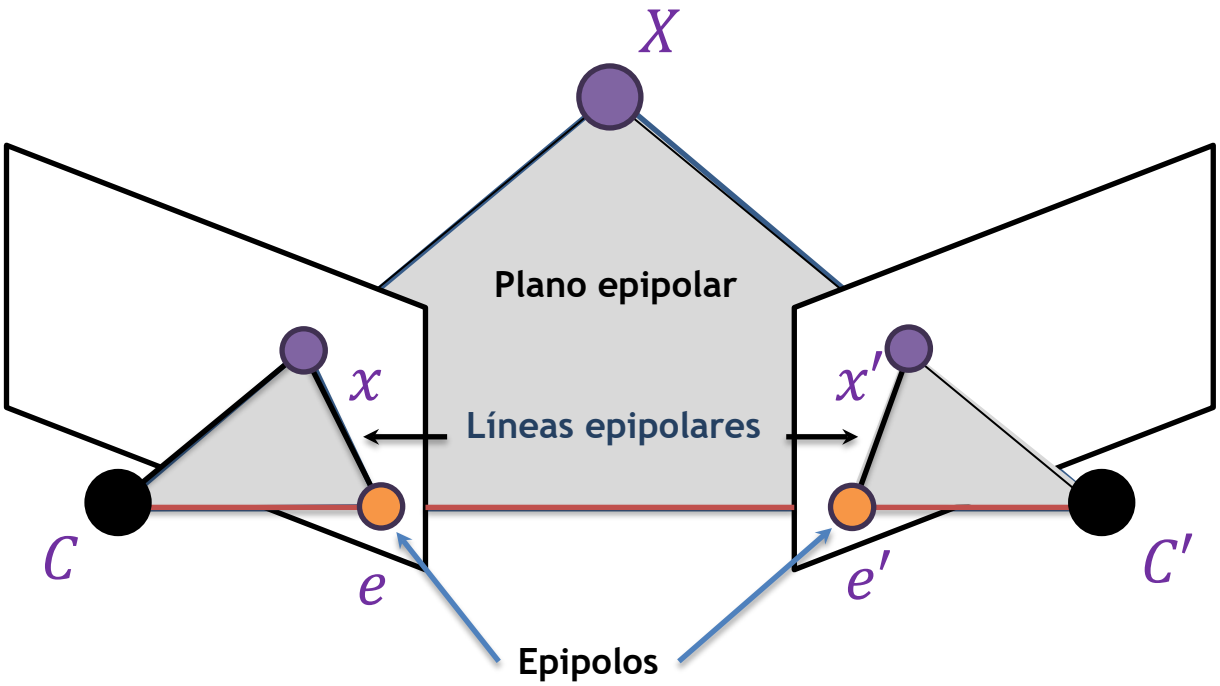
Epipolar

Dos cámaras y la recta que une sus centros C y C’
Dado un punto X, y sus proyecciones en cada cámara
Conforman el **plano epipolar**, intersecta con planos imagen
Uno de los posibles planos que pasan por C y C’



Epipolar

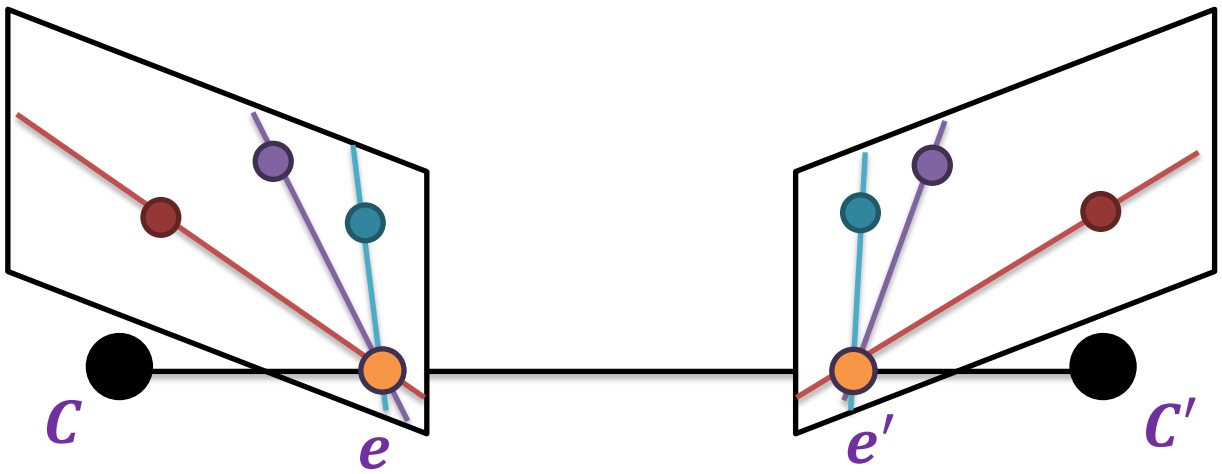
Líneas epipolares conectan epipolos y proyección de X



Epipolar

El aspecto de las líneas epipolares, difiere según los epipolos

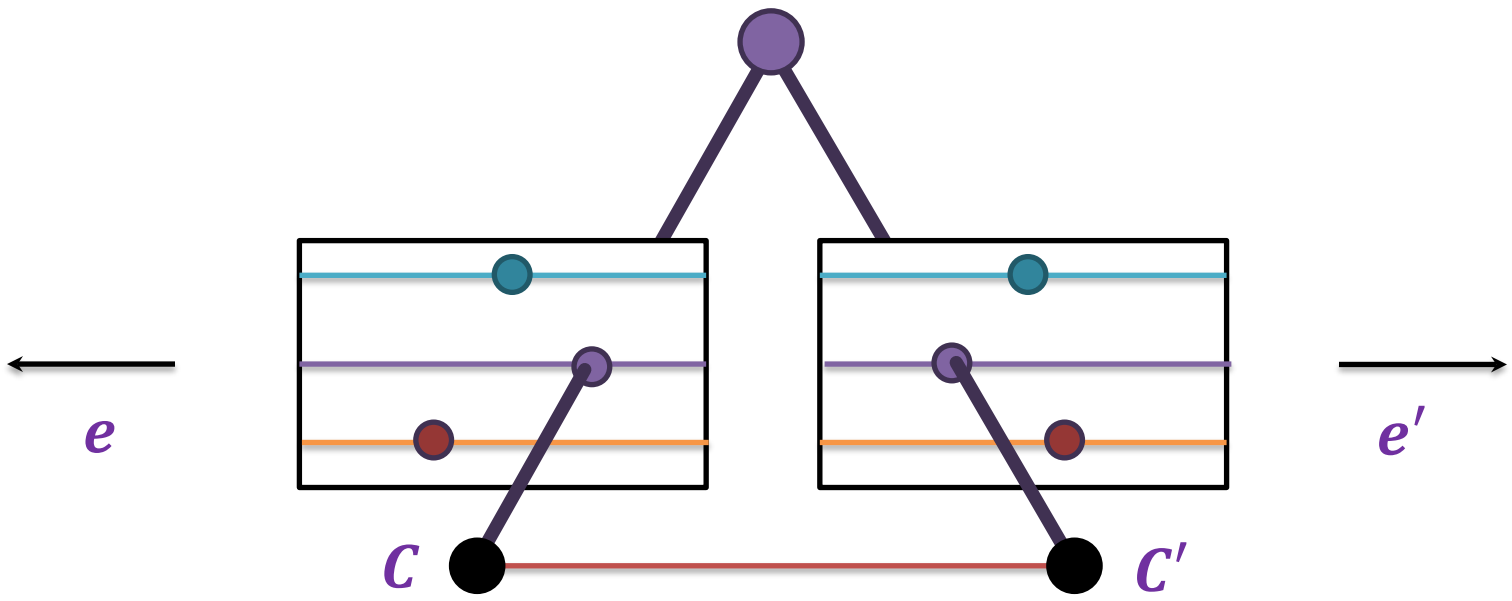
Epipolos finitos, cuando aparecen en la imagen



Epipolar

Movimiento paralelo al plano imagen

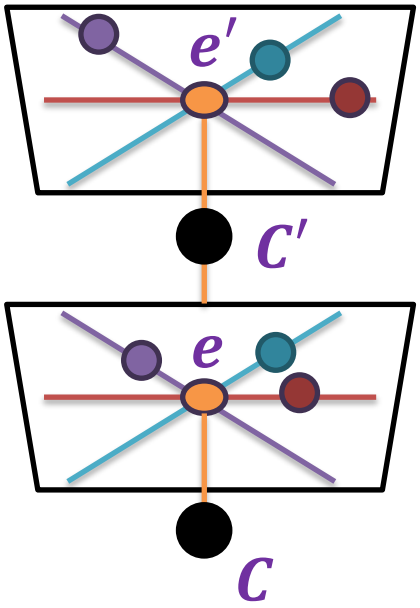
Epipolos muy alejados, líneas paralelas



Epipolar

Movimiento perpendicular al plano imagen (robot en pasillo)

Epipolos coincidentes con punto principal



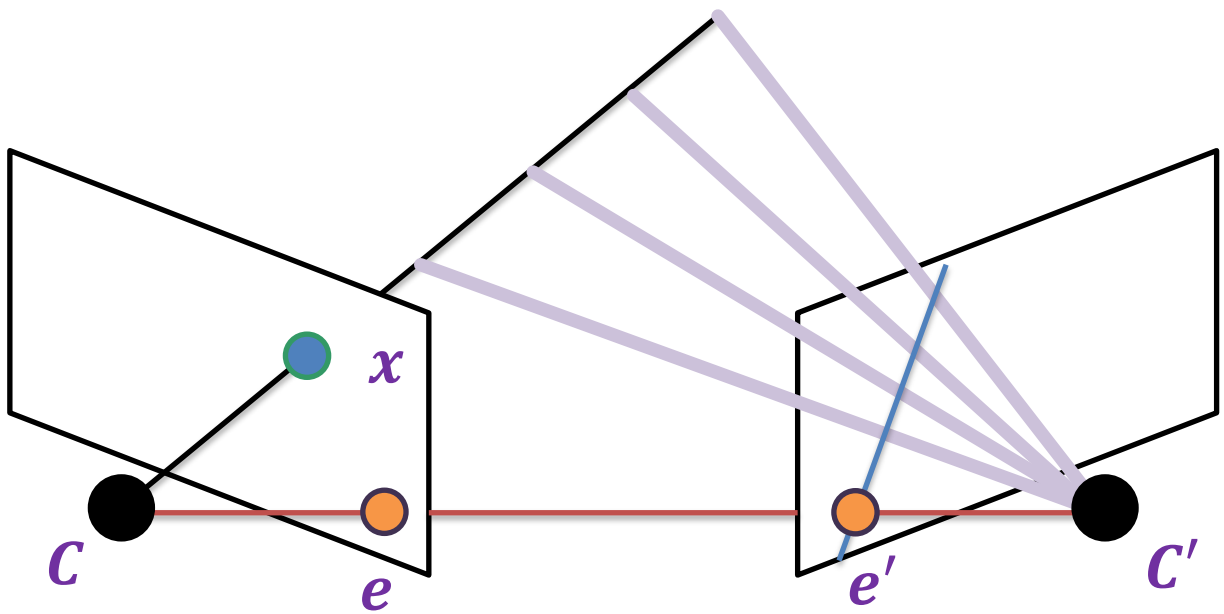
Epipolar

Restricción epipolar

Dado un punto proyectado, *x*, en una cámara

Se encuentra en otra cámara en la línea epipolar (y a la inversa)

Posibles correspondencias deben caer en la línea epipolar



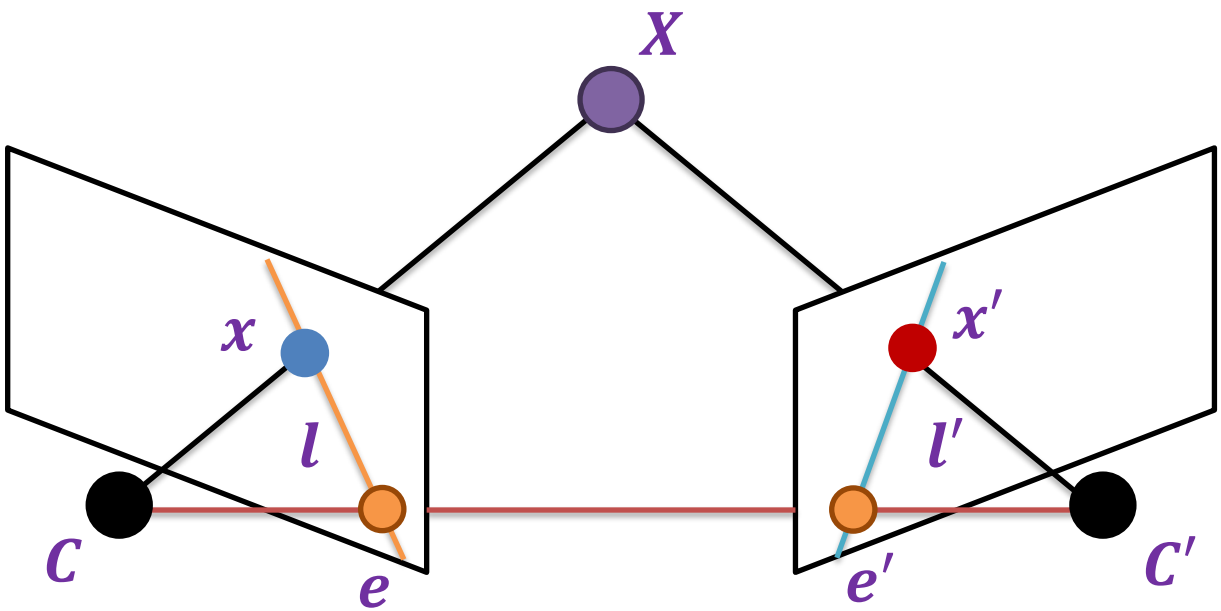
Epipolar

Restricción epipolar

Dos puntos *x* y *x'* están en líneas epipolares *l* y *l'* que encajan

Sus rayos visuales se cruzan en el espacio (en general no sucede)

Pueden ser proyecciones del mismo punto 3D *X*



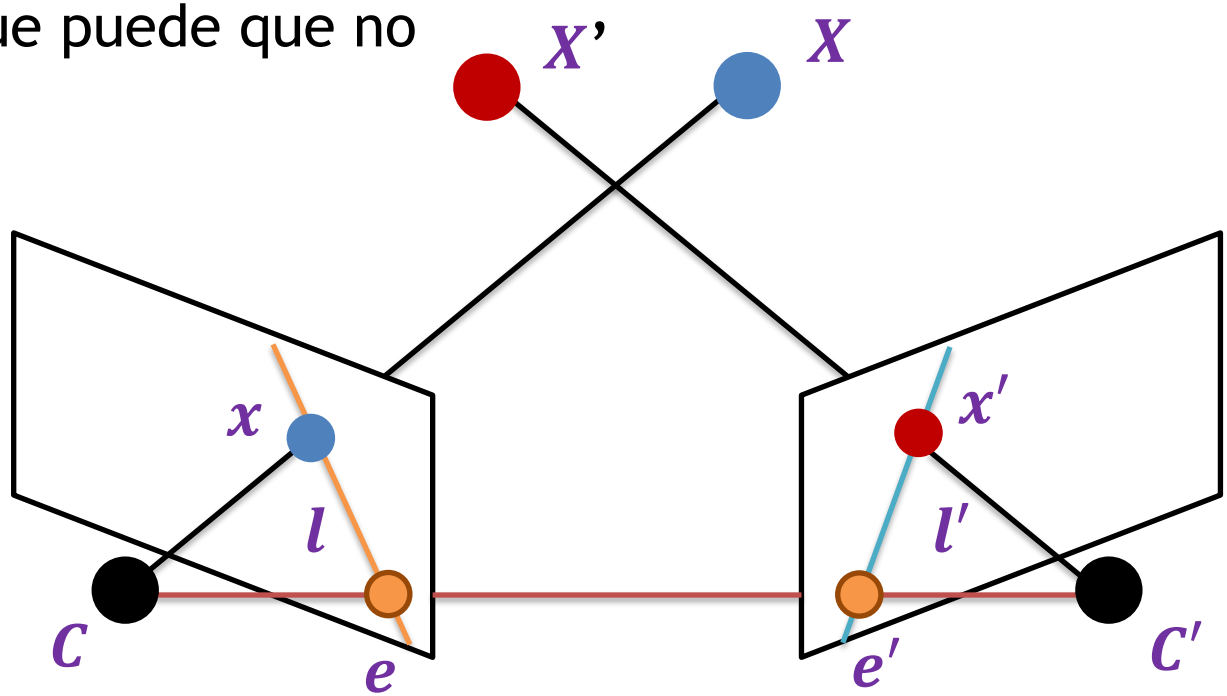
Epipolar

Restricción epipolar

Dos puntos *x* y *x'* están en líneas epipolares *l y l'* que encajan

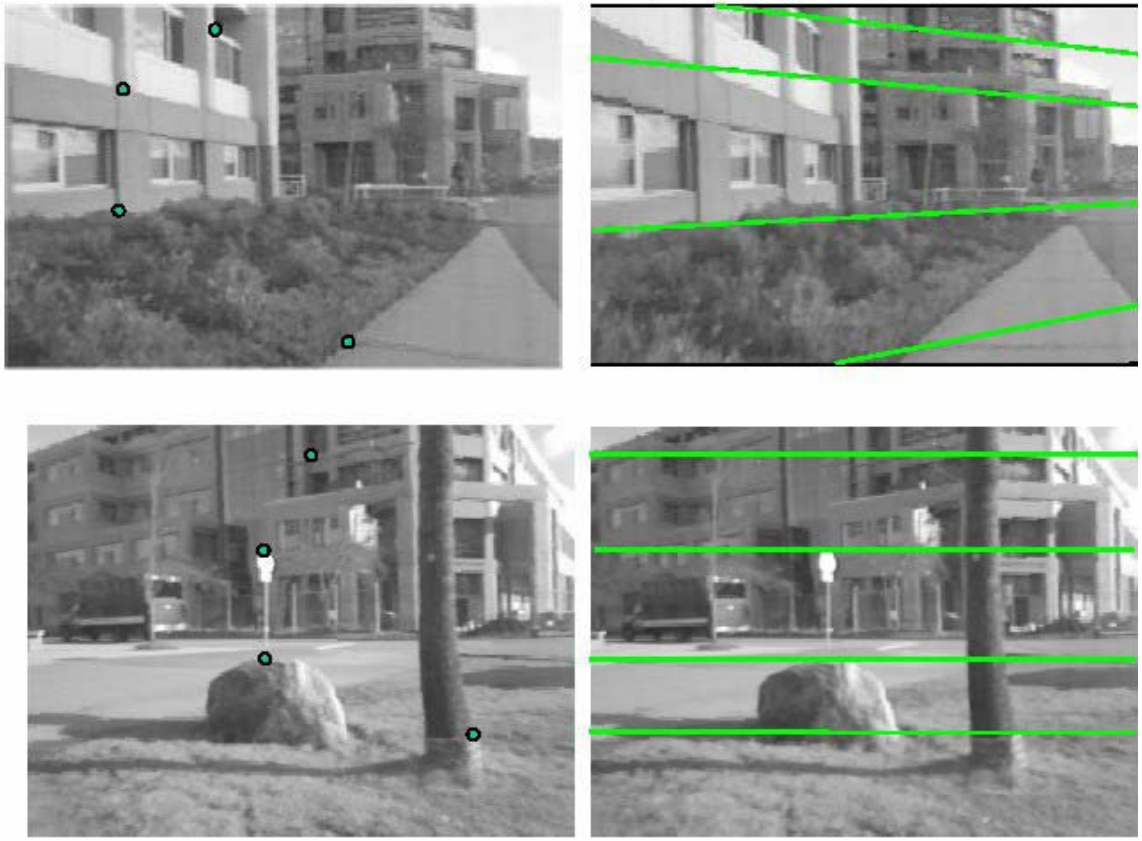
Sus rayos visuales se cruzan en el espacio (en general no sucede)

Aunque puede que no



Epipolar

Restricción epipolar



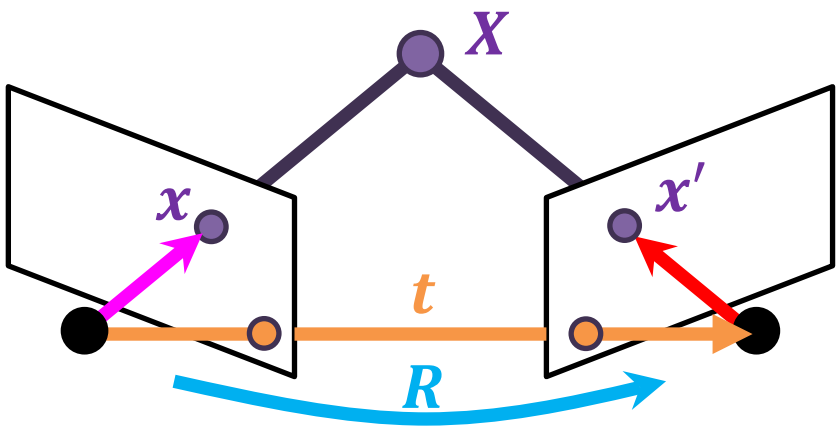
Epipolar

Restricción epipolar en cámaras calibradas

K y *K'* conocidas (intrínsecos), $\boldsymbol{K}[\boldsymbol{I} \mid \boldsymbol{0}]$ y $\boldsymbol{K}'[\boldsymbol{R} \mid \boldsymbol{t}]$

Coordenadas de cámara

$$\boldsymbol{x}_{\text{norm}} = \boldsymbol{K}^{-1}\boldsymbol{x} \quad \cong [\boldsymbol{I} \mid \boldsymbol{0}]\boldsymbol{X}, \qquad \boldsymbol{x}'_{\text{norm}} = \boldsymbol{K}'^{-1}\boldsymbol{x}' \quad \cong [\boldsymbol{R} \mid \boldsymbol{t}]\boldsymbol{X}$$



H. C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. Nature 293 (5828): 133–135, 1981

Epipolar

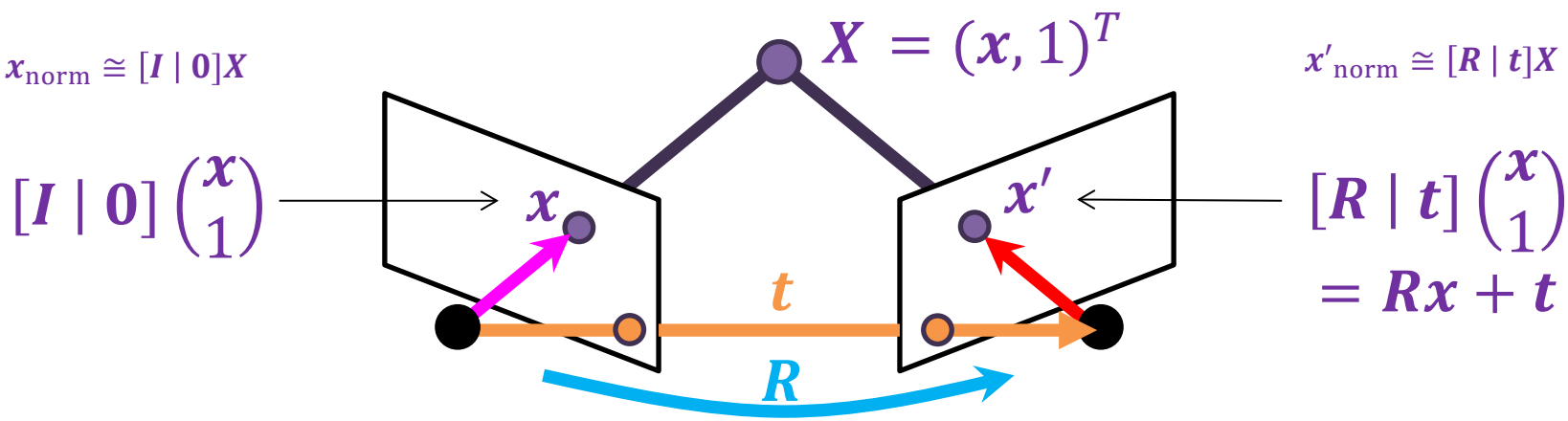
Restricción epipolar en cámaras calibradas

$\boldsymbol{x}'$, $\boldsymbol{R}\boldsymbol{x}$, y $\boldsymbol{t}$ son coplanares, o linealmente dependientes

$$\boldsymbol{x}' \cong \boldsymbol{R}\boldsymbol{x} + \boldsymbol{t}$$

Puede escribirse como

$$\boldsymbol{x}' \cdot [\boldsymbol{t} \times (\boldsymbol{R}\boldsymbol{x})] = 0$$



H. C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. Nature 293 (5828): 133–135, 1981

Epipolar

Restricción epipolar en cámaras calibradas

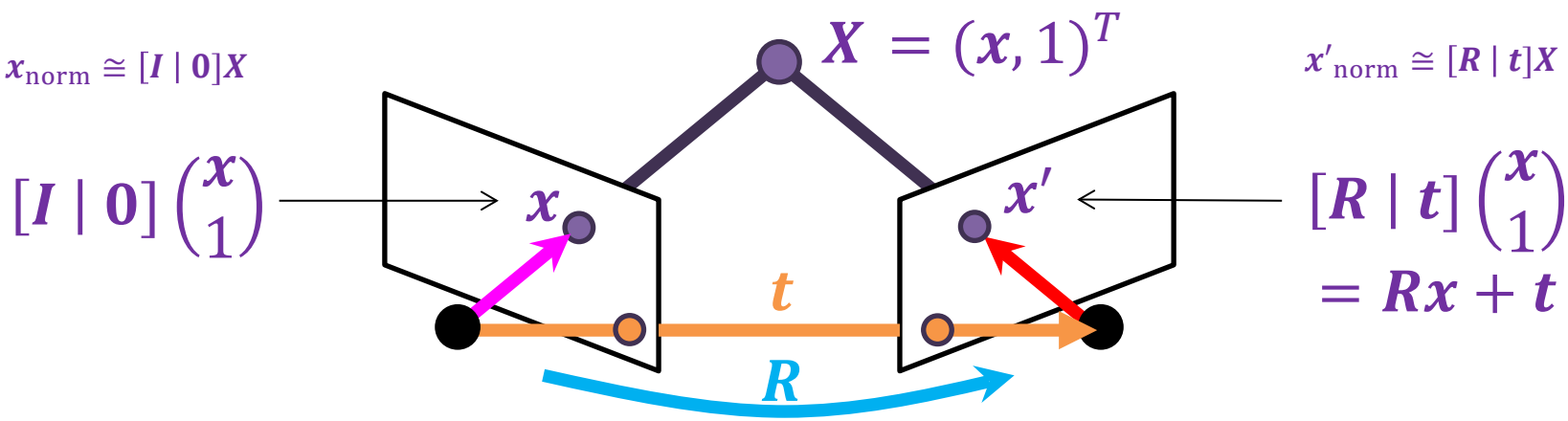
Producto vectorial en forma matricial

$$\boldsymbol{a} \times \boldsymbol{b} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = [\boldsymbol{a}_\times]\boldsymbol{b}$$

En forma matricial

$$\boldsymbol{x}' \cdot [\boldsymbol{t} \times (\boldsymbol{R}\boldsymbol{x})] = 0 \longrightarrow \boldsymbol{x}'^T [\boldsymbol{t}_\times] \boldsymbol{R}\boldsymbol{x} = 0 \longrightarrow \boldsymbol{x}'^T \boldsymbol{E}\boldsymbol{x} = 0$$

Matriz esencial



H. C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. Nature 293 (5828): 133–135, 1981

Epipolar

Restricción epipolar en cámaras calibradas

Propiedades de la matriz esencial

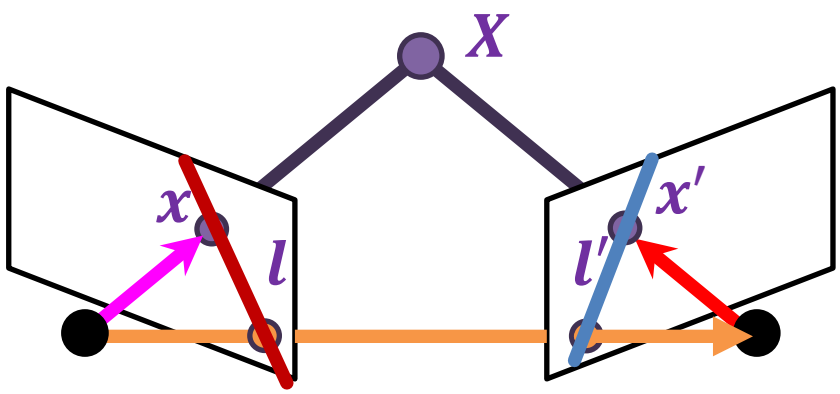
$$\boldsymbol{x}'^T \boldsymbol{E}\boldsymbol{x} = 0 \qquad (x', y', 1)\begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$$

$\boldsymbol{E}\boldsymbol{x}$ es la línea epipolar asociada con $\boldsymbol{x}$ $(\boldsymbol{l}' = \boldsymbol{E}\boldsymbol{x})$

$\boldsymbol{E}^T\boldsymbol{x}'$ es la línea epipolar asociada con $\boldsymbol{x}'$ $(\boldsymbol{l} = \boldsymbol{E}^T\boldsymbol{x}')$

$\boldsymbol{E}\boldsymbol{e} = \boldsymbol{0}$ y $\boldsymbol{E}^T\boldsymbol{e}' = \boldsymbol{0}$

$\boldsymbol{E}$ es singular (rango 2) y tiene cinco grados de libertad



H. C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. Nature 293 (5828): 133–135, 1981

Epipolar

Restricción epipolar en cámaras no calibradas

K y *K'* no conocidas

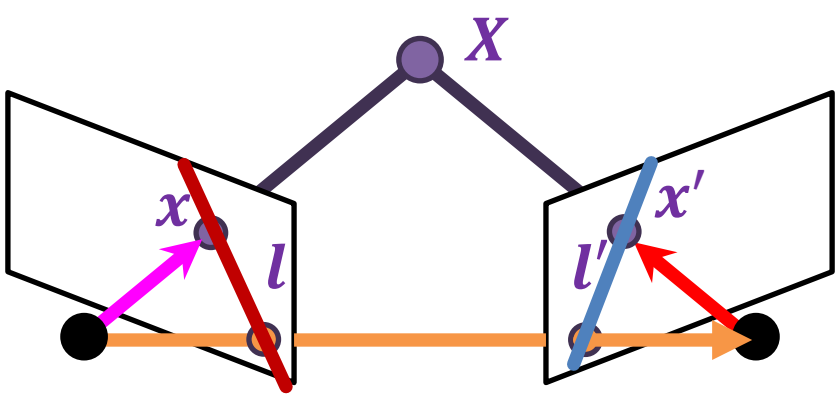
Restricción epipolar en términos de coordenadas no conocidas

$$\boldsymbol{x}'^{T}_{\text{norm}}\boldsymbol{E}\boldsymbol{x}_{\text{norm}} = 0$$

$$\boldsymbol{x}_{\text{norm}} = \boldsymbol{K}^{-1}\boldsymbol{x},\ \boldsymbol{x}'_{\text{norm}} = \boldsymbol{K}'^{-1}\boldsymbol{x}'$$

$$\boldsymbol{x}'^{T}\boldsymbol{F}\boldsymbol{x} = 0 \text{ donde } \boldsymbol{F} = \boldsymbol{K}'^{-T}\boldsymbol{E}\boldsymbol{K}^{-1}$$

Matriz fundamental



Epipolar

Restricción epipolar en cámaras no calibradas

Propiedades de la matriz fundamental

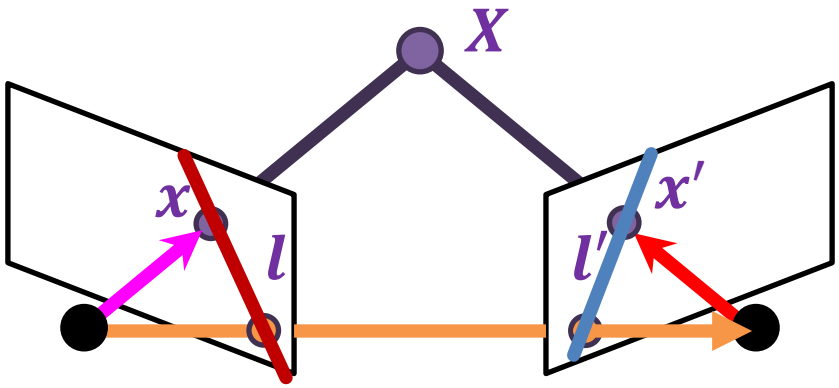
$$\boldsymbol{x}'^T \boldsymbol{F} \boldsymbol{x} = 0 \qquad (x', y', 1) \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$$

$\boldsymbol{F}\boldsymbol{x}$ es la línea epipolar asociada con $\boldsymbol{x}$ $(\boldsymbol{l}' = \boldsymbol{F}\boldsymbol{x})$

$\boldsymbol{F}^{\boldsymbol{T}}\boldsymbol{x}'$ es la línea epipolar asociada con $\boldsymbol{x}'$ $(\boldsymbol{l} = F^T \boldsymbol{x}')$

$\boldsymbol{F}\boldsymbol{e} = \boldsymbol{0}$ y $\boldsymbol{F}^{\boldsymbol{T}}\boldsymbol{e}' = \boldsymbol{0}$

$\boldsymbol{F}$ es singular (rango 2) y tiene siete grados de libertad

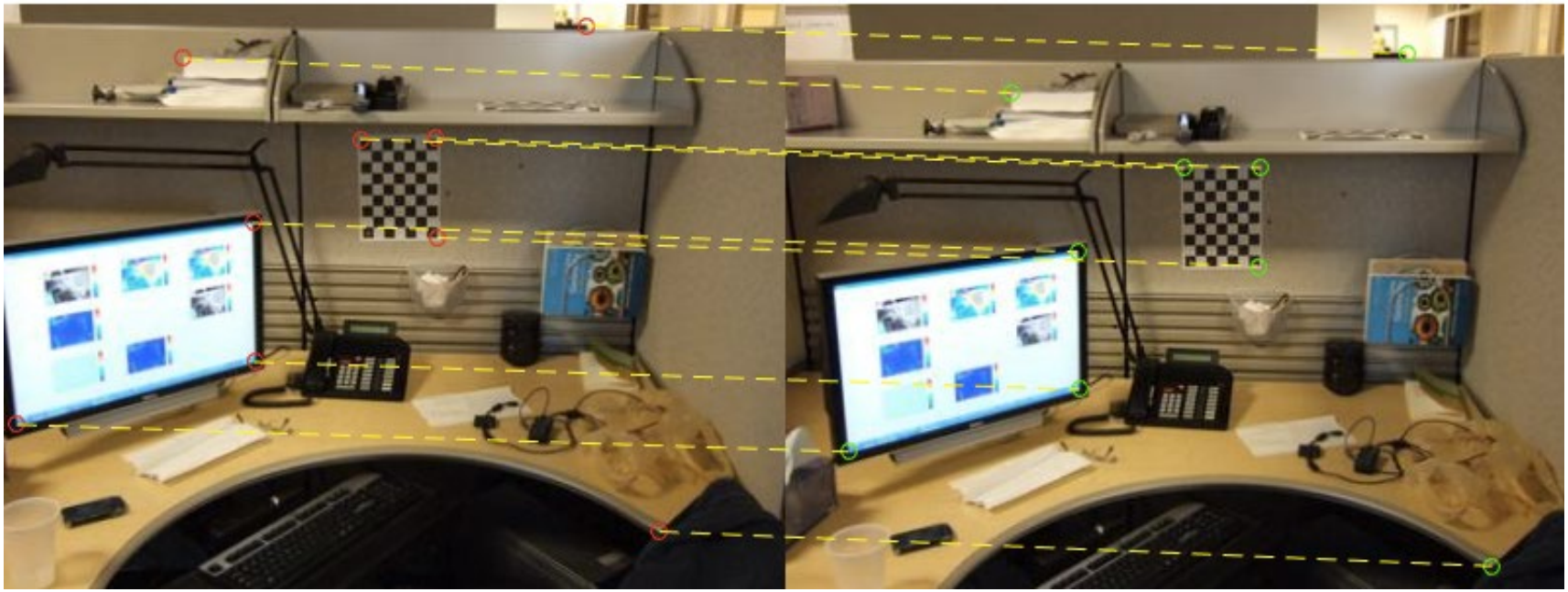


Epipolar

Estimación de la matriz fundamental

A partir de correspondencias de puntos

$$\boldsymbol{x} = (x, y, 1)^T \text{ y } \boldsymbol{x}' = (x', y', 1)^T$$



Epipolar

Estimación de la matriz fundamental

A partir de correspondencias de puntos

$$\boldsymbol{x} = (x, y, 1)^T \text{ y } \boldsymbol{x}' = (x', y', 1)^T$$

$$\boldsymbol{x}'^{\boldsymbol{T}} \boldsymbol{F} \boldsymbol{x} = 0$$

$$(x', y', 1) \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0 \quad (x'x, x'y, x', y'x, y'y, y', x, y, 1) \begin{pmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{pmatrix} = 0$$

Epipolar

Estimación de la matriz fundamental

Algoritmo de ocho puntos

$$\begin{bmatrix} & & & \vdots & & & & & \\ x'x & x'y & x' & y'x & y'y & y' & x & y & 1 \\ & & & \vdots & & & & & \end{bmatrix} \begin{pmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{pmatrix} = \mathbf{0}$$

Solución lineal

Mínimos cuadrados para encontrar f

$$\arg\min_{\|\boldsymbol{f}\|=1} \|\boldsymbol{U}\boldsymbol{f}\|_2^2$$

Autovector de U^TU con menor autovalor

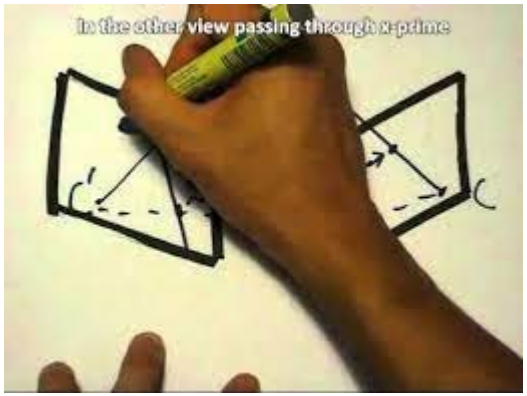
Epipolar

Estimación de la matriz fundamental

Conocida como calibración débil

Si se conocen *K* y *K'*:

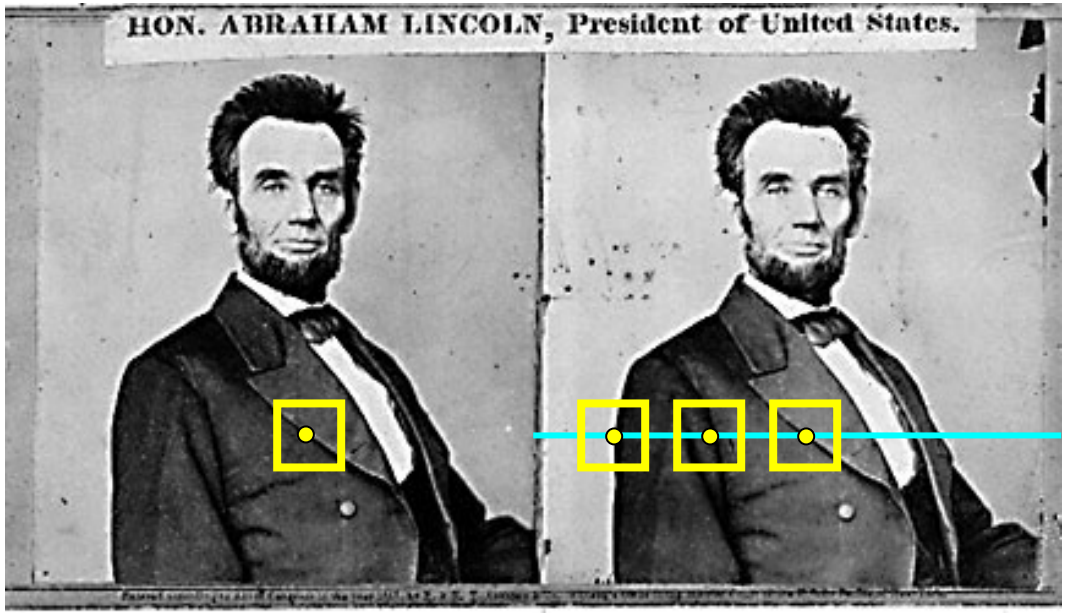
- Estimar la matriz esencial a partir de *F*
- O usar algoritmo de cinco puntos



Estéreo

Estéreo

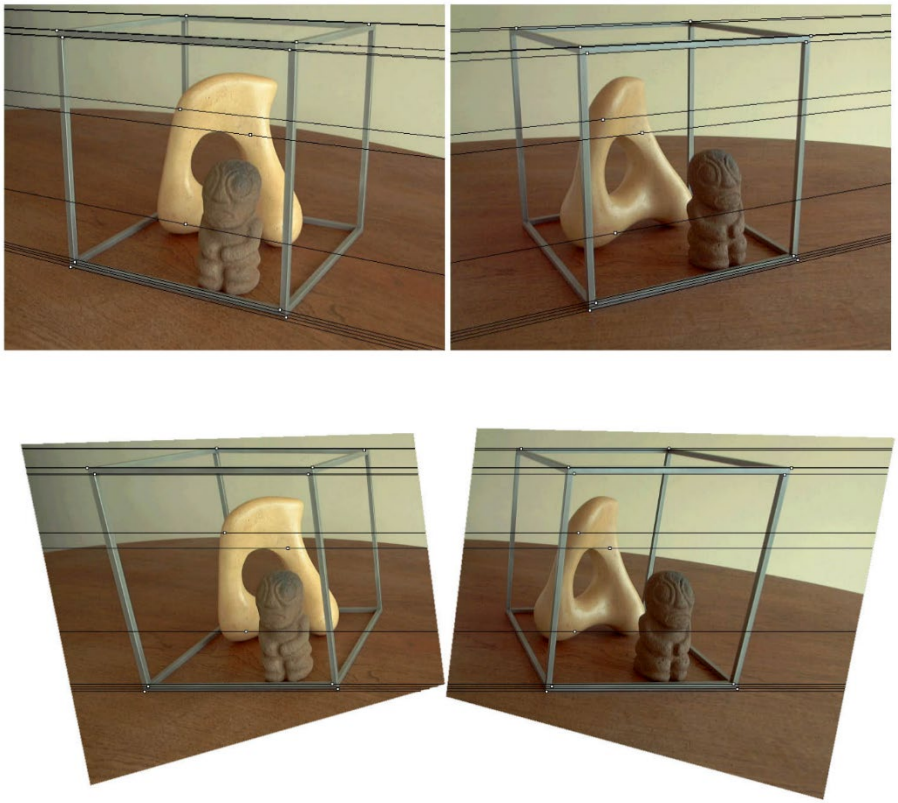
Para cada píxel en la imagen, buscar encaje en línea epipolar

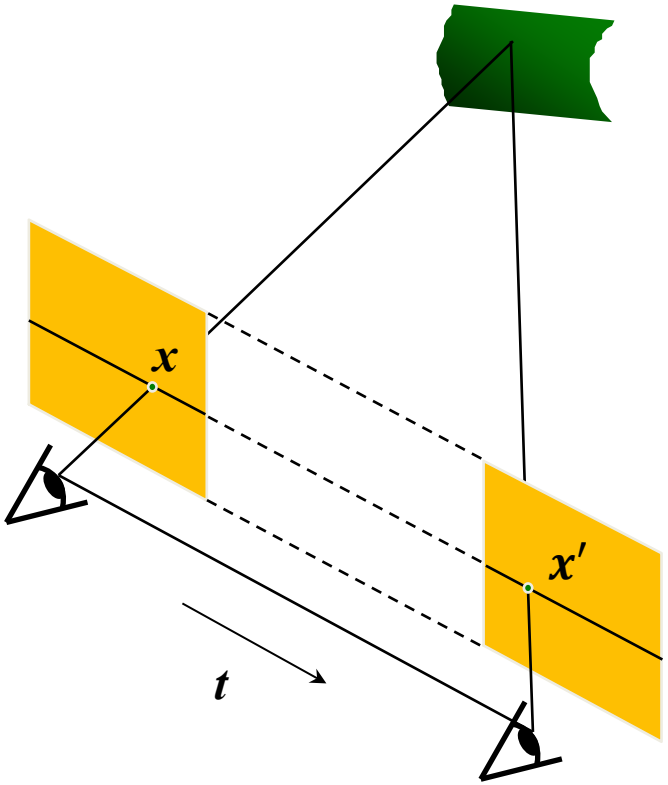


¿Líneas epipolares coincidentes con las de barrido?

Estéreo

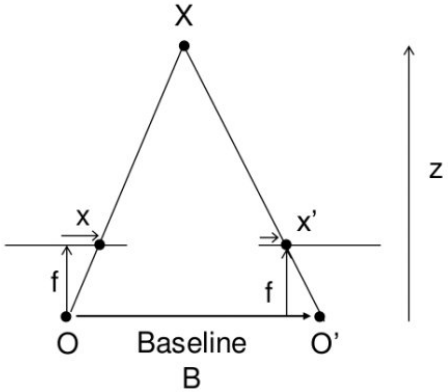
Par estéreo
Rectificar con homografía si fuera necesario
Encaje de píxeles, triangulación





Triangulación

Profundidad desde la disparidad entre par estéreo calibrado
Disparidad proporcional a profundidad



$$disparidad = x - x' = \frac{B \cdot f}{Z}$$

Triangulación

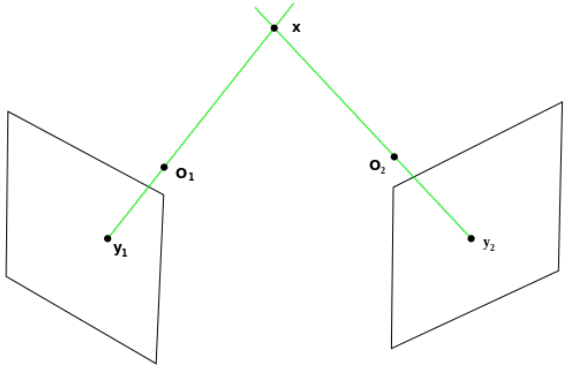
Posición 3D a partir de:

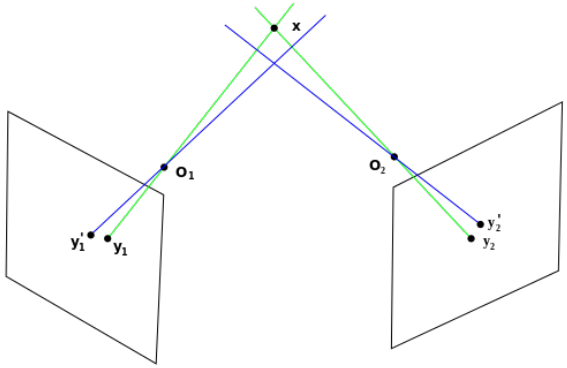
- posiciones en imagen
- localizaciones conocidas de las cámaras

En general, no habrá plano epipolar ideal

Errores numéricos y ruido

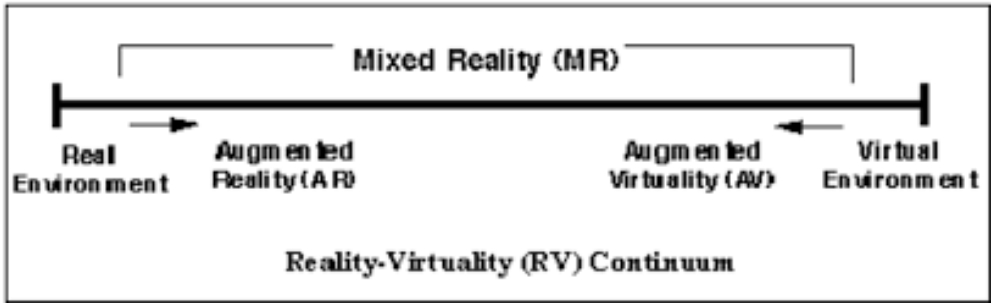
Punto que queda más próximo a todos los rayos ópticos





AR

Reality-virtuality continuum (RV)



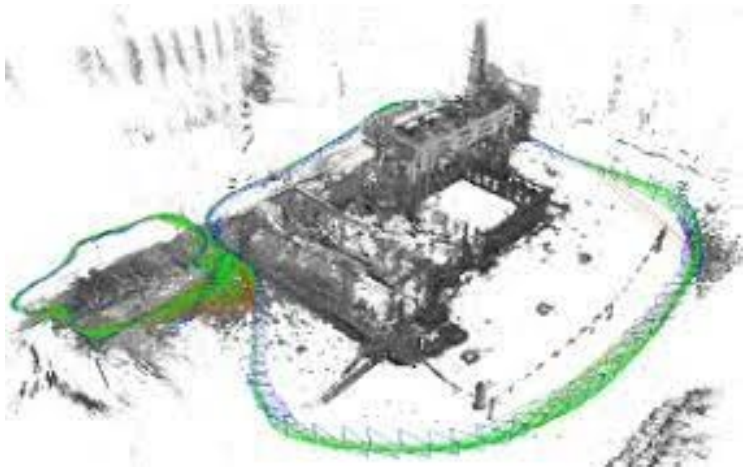
Paul Milgram y Fumio Kishino

	VR Reemplaza	AR Aumenta
Generación escena	Realismo	No necesario realismo
Dispositivo de visualización	Inmersivo, fov amplio	No inmersivo
Precisión *tracking*	Baja-media basta	Alta

SLAM

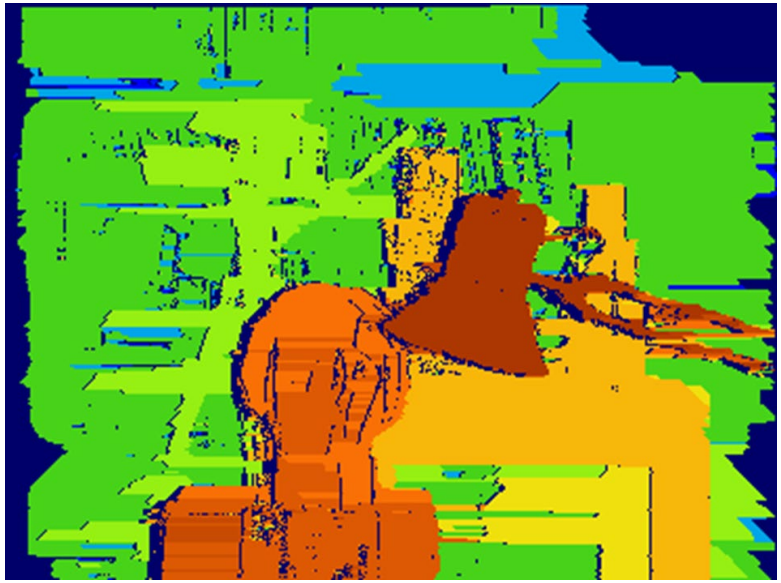
Simoultaneous localization and mapping (SLAM)

Mapas 3D desde robots autónomos
Multisensorial: láser, sónar, imagen
Multirobot
Zonas amplias



Densa

Profundidad de todos los píxeles
Encaje en línea epipolar
Triangulación para obtener profundidad
Problemas con oclusiones, zonas sin texturas, etc.



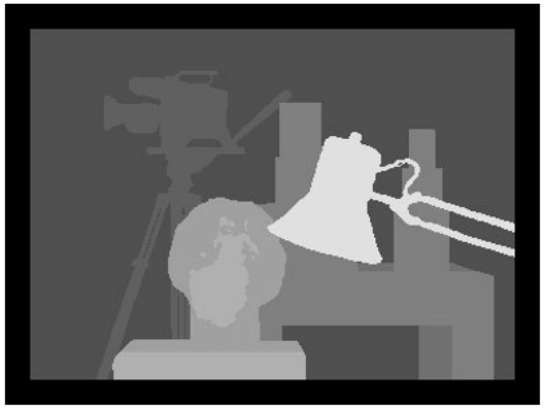
Densa

Disparidad
Tsukuba

Left image

Right image

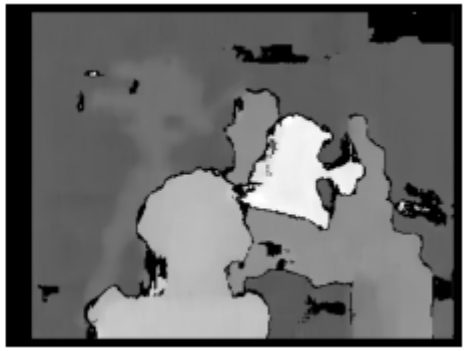
Disparity map

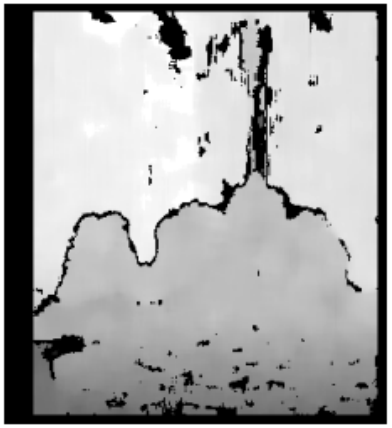


Densa

Disparidad

```
imgL = cv.imread('tsukuba_l.png', cv.IMREAD_GRAYSCALE)
imgR = cv.imread('tsukuba_r.png', cv.IMREAD_GRAYSCALE)
stereo = cv.StereoBM.create(numDisparities=16, blockSize=15)
disparity = stereo.compute(imgL,imgR)
```



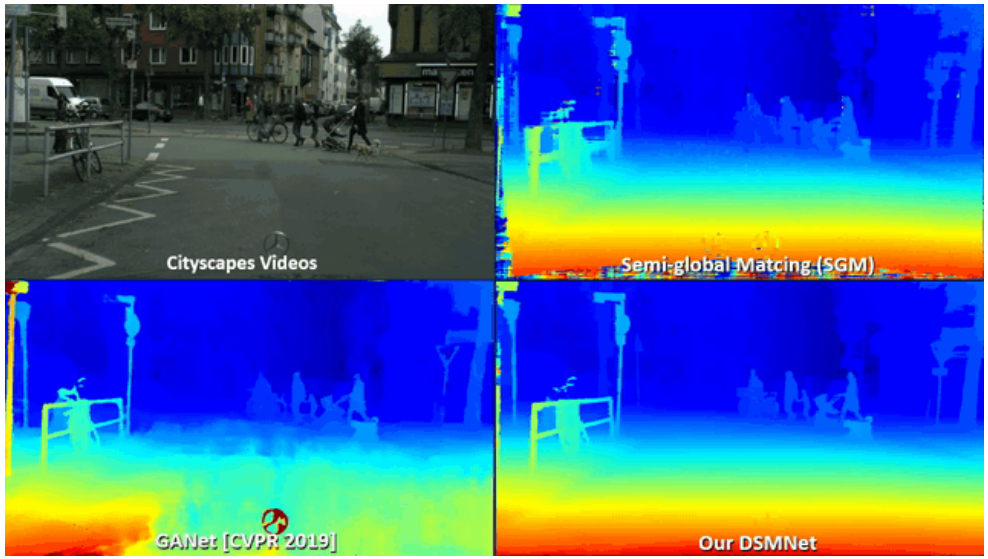


Densa

Aprendizaje

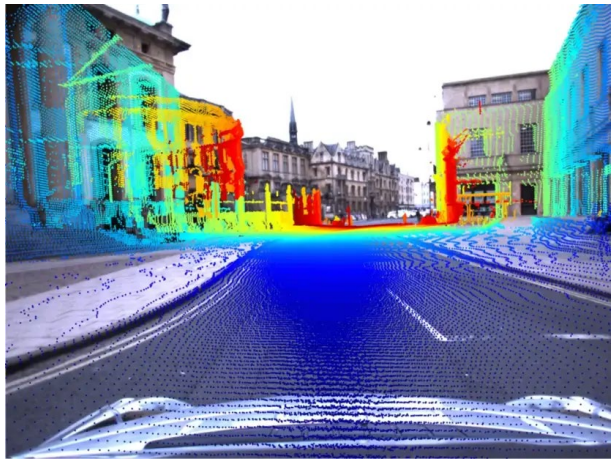
MC CNN

DSMNet

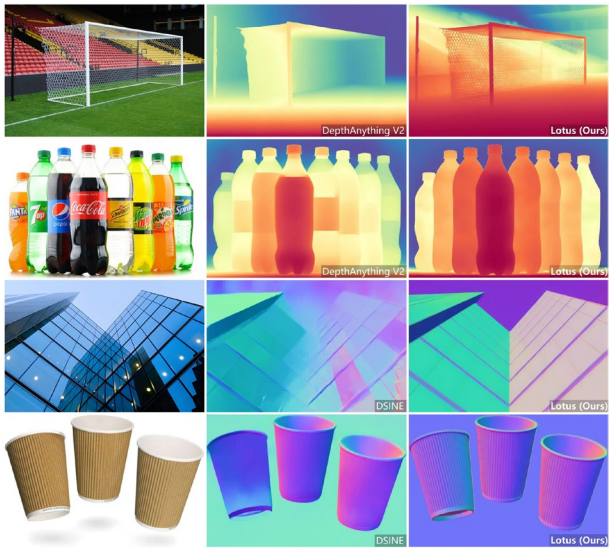


Densa

Multicue

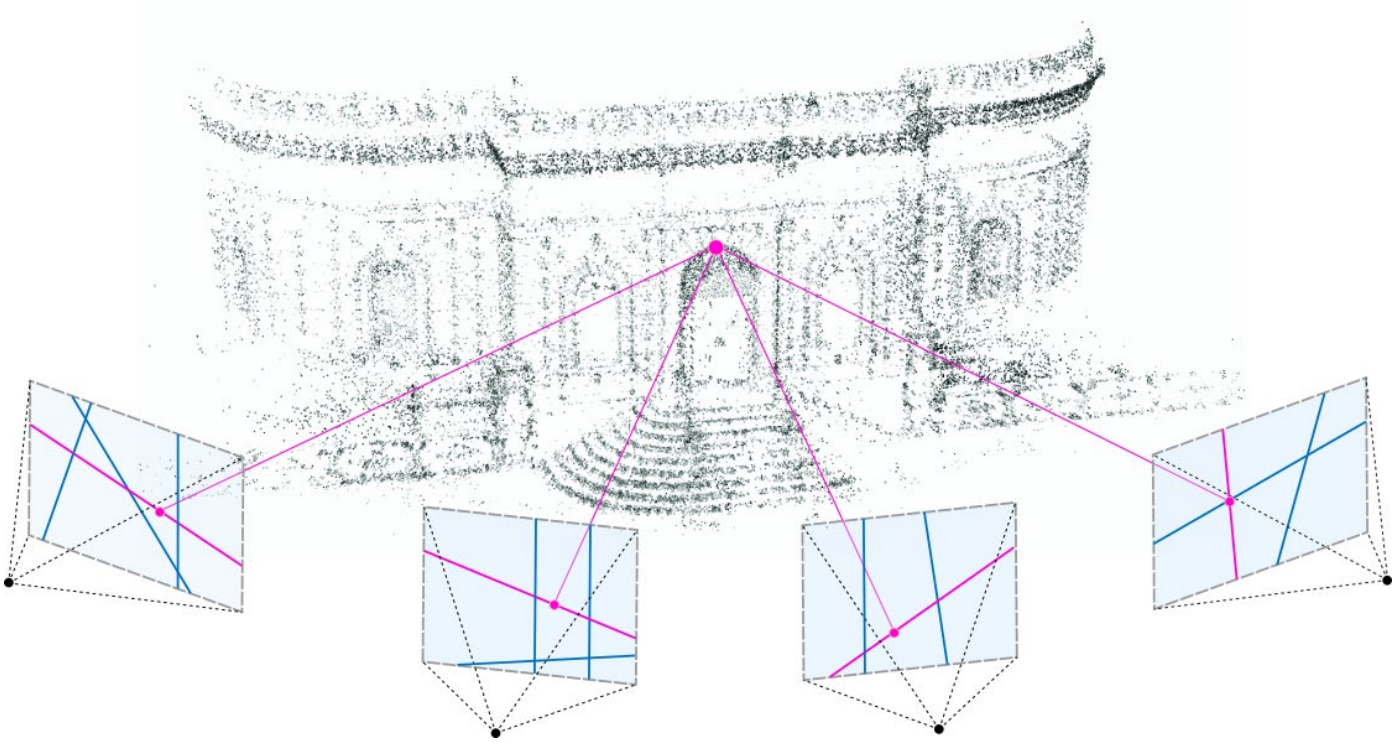


Lotus (modelo difusión)



Estructura desde movimiento (SfM)

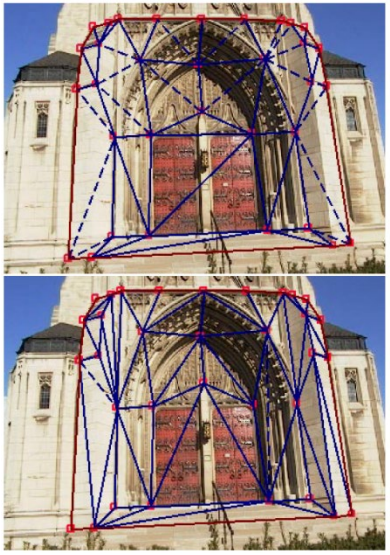
Geometría y movimiento de cámara con dos o más imágenes



Estructura desde movimiento

Proceso:

- Detección de buenos puntos
- Correspondencia entre imágenes (LK, correlación, …)
- Estimación movimiento y estructura 3D
- Refinamiento
- Ajuste a superficies



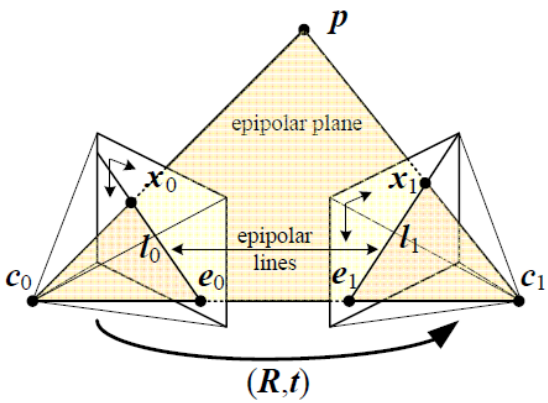
D. Morris y T. Kanade. Image-consistent surface triangulation. CVPR 2000

Estructura desde movimiento

Recuperar posición 3D de cámara y puntos de interés

p_0 punto 3D observado en la primera imagen con $t = c_1 - c_0$

$$p_0 = d_0 \hat{x}_0$$
$$p_1 = d_1 \hat{x}_1 = R p_0 + t$$



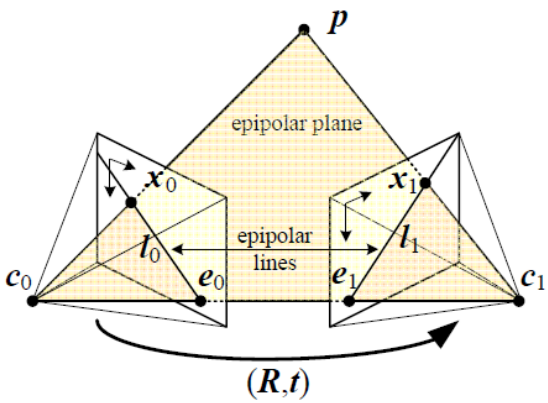
Szeliski

Estructura desde movimiento

Recuperar posición 3D de cámara y puntos de interés

p_0 punto 3D observado en la primera imagen con $t = c_1 - c_0$

$$p_0 = d_0 \hat{x}_0$$
$$p_1 = d_1 \hat{x}_1 = R p_0 + t$$



Szeliski

La restricción epipolar

$$\hat{x}_1{}^T E \hat{x}_0 = 0$$

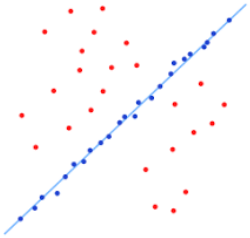
Con E, matriz esencial (3x3) que relaciona dos imágenes

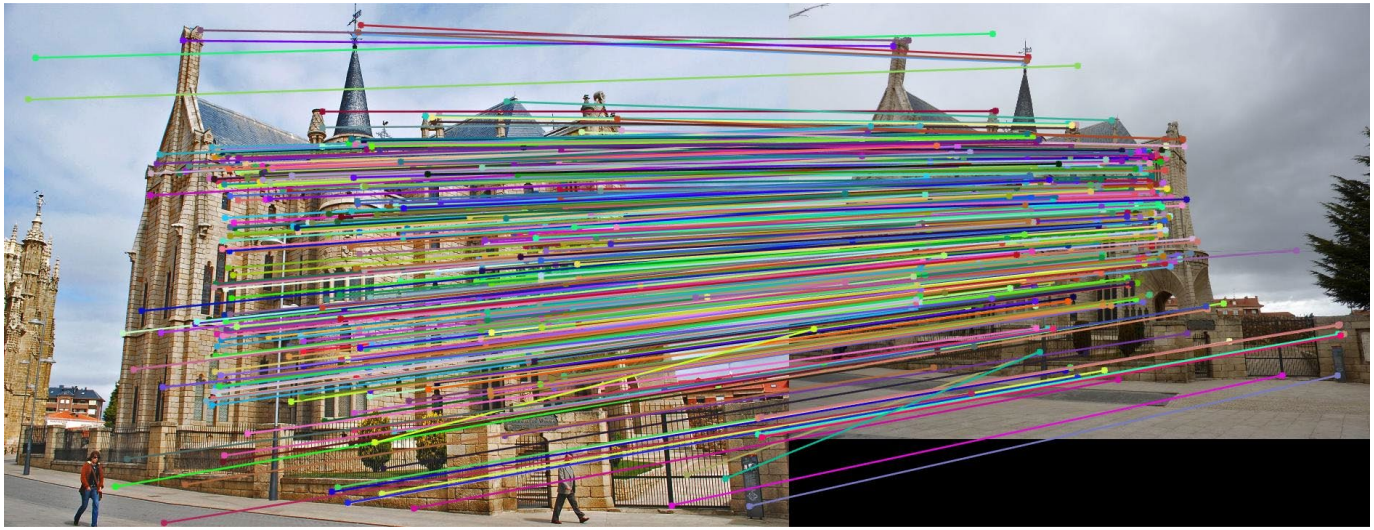
n puntos definen ecuaciones para estimar los nueve elementos de *E*

Estructura desde movimiento

RANSAC (*RANdom SAmple Consensus*)

Seleccionar un subconjunto número de correspondencias
Estimar estructura y movimiento
Contar número de *inliers*
Tomar subconjunto con mayor número
Lanzar proceso completo de mínimos cuadrados





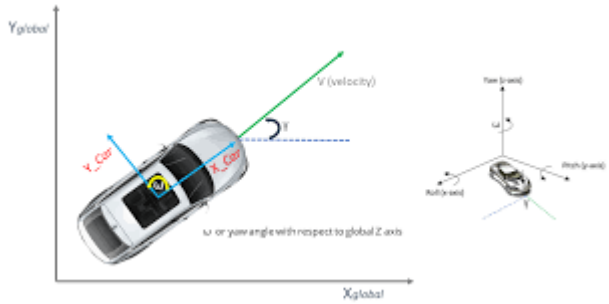
Estructura desde movimiento

Buscar puntos con información previa de movimiento

Estimación de localización

Integración de incertidumbre:

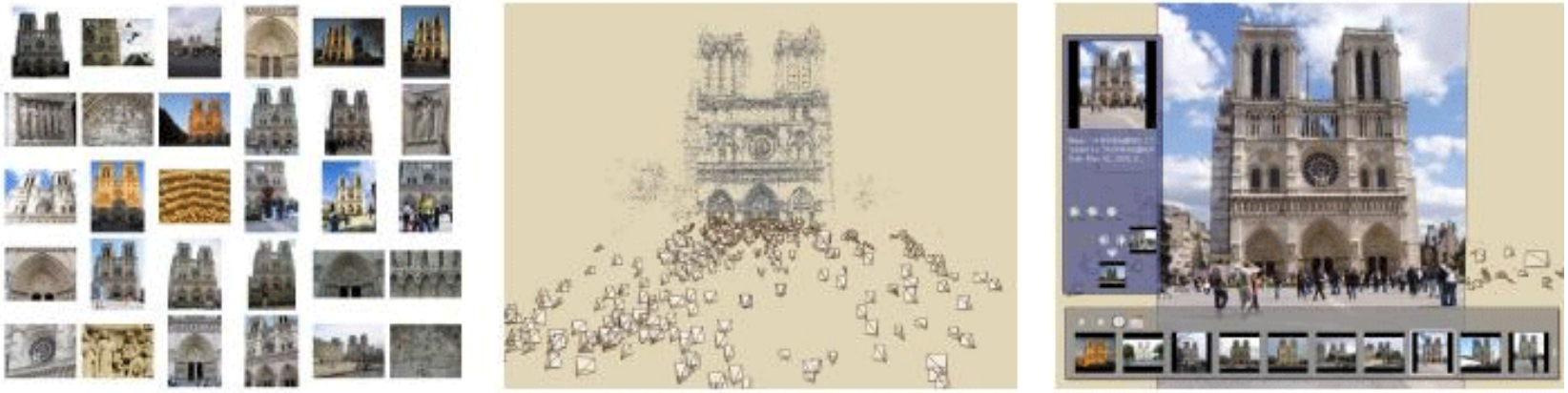
- Filtro de Kalman
  - Actualiza estado del sistema
  - Modelo gaussiano
- Filtro de partículas (también *Condensation*)
  - Actualiza distribución de probabilidades
  - Permite múltiples hipótesis



Reconstrucción 3D

Photo Tourism

Exploring photo collections in 3D



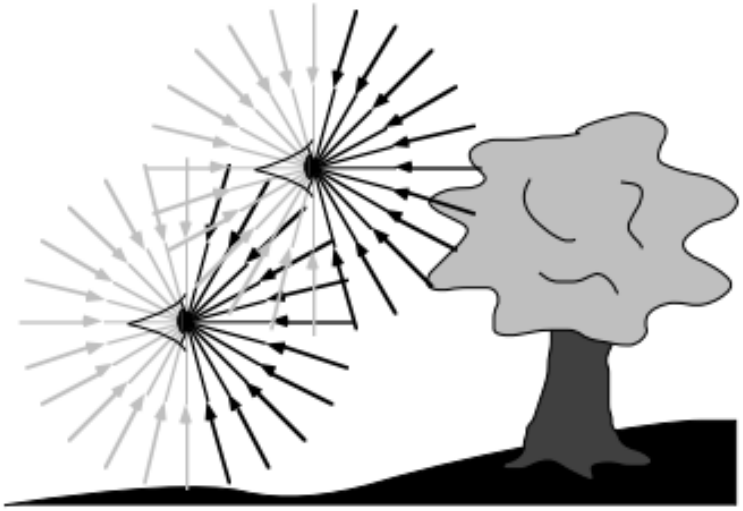
(a) (b) (c)

Reconstrucción 3D

Función plenóptica

Intensidad de la luz pasando por un punto 3D

$$L(X, Y, Z, \psi, \phi, \lambda, t\ )$$



Adelson y Bergen

Habitualmente no calculada

Si conocida, permite reproducir la escena desde cualquier punto de vista, p.e. *randiance field*

Reconstrucción 3D

Radiance field (campo de radiancia/resplandor)
Escena como nube de partículas con distinta transparencia
Reproducción volumétrica, con densidad como transparencia

$$L(X, Y, Z, \psi, \phi, \lambda, t\,) \longrightarrow r, g, b, \sigma$$

Neural Radiance fields (NeRFs*, 2020*)
Modela *L* con una red reuronal
Irrupción aprendizaje automático

Reconstrucción 3D

Without the camera, there would be no …

Explosión generadores de escenas 3D a partir de imágenes
Nuevas vistas desde imágenes no controladas

Reconstrucción 3D

NeRF-W (2020)

Reducir número de imágenes, p.e. zip-NeRF, ReconFusion
Tiempo real

Reconstrucción 3D

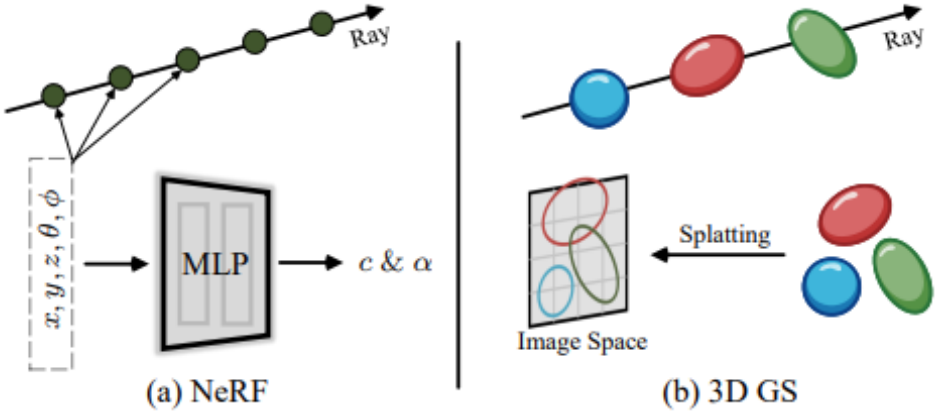
Escenarios específicos

Reconstrucción 3D

Gaussian Splatting **(dispersión gaussiana)**

Desde nubes de puntos (p.e. SfM)

Discretiza en campo de radiación

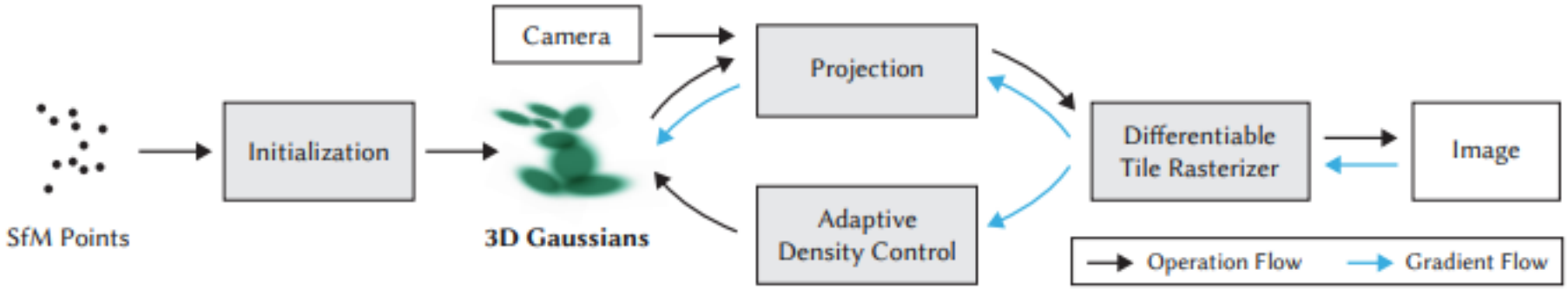


Reconstrucción 3D

Gaussian Splatting

Discretiza en campo de radiación

Optimización

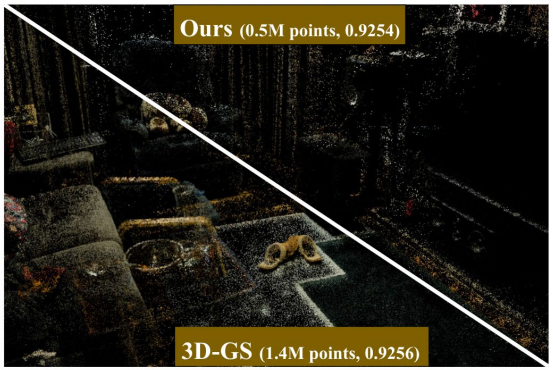


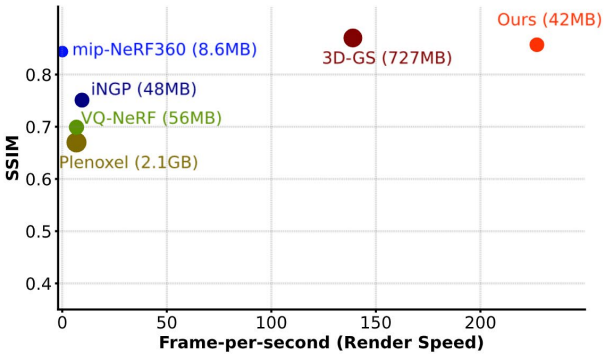
Reconstrucción 3D

Gaussian Splatting

Mayor rapidez

Menos imágenes necesarias

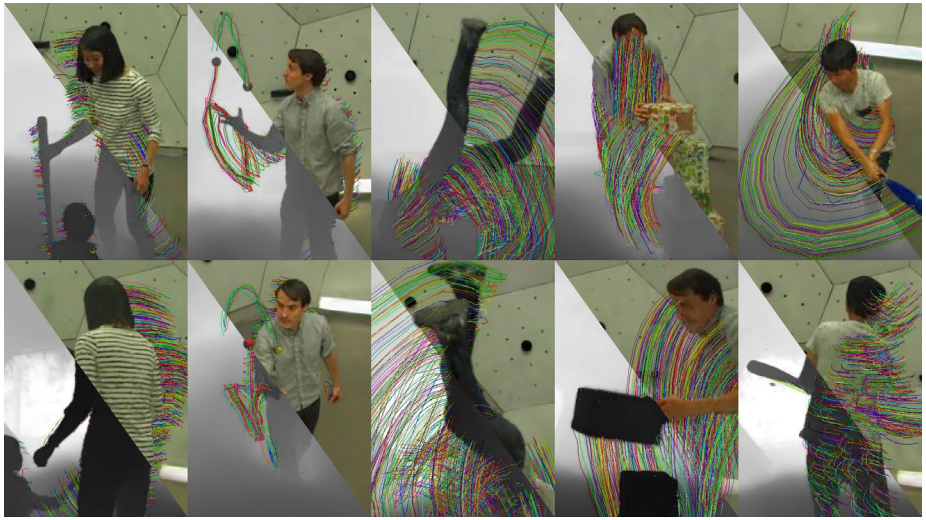




Reconstrucción 3D

Gaussian Splatting

No solo estáticas



J. Luiten et al. Dynamic 3D Gaussians: Tracking by Persistent Dynamic View Synthesis, 3DV, 2024

Bibliografía

- R. Szeliski. Computer Vision: Algorithms and Applications, 2nd ed. 2022
- A. Torralba et al. Foundations of Computer Vision, 2024