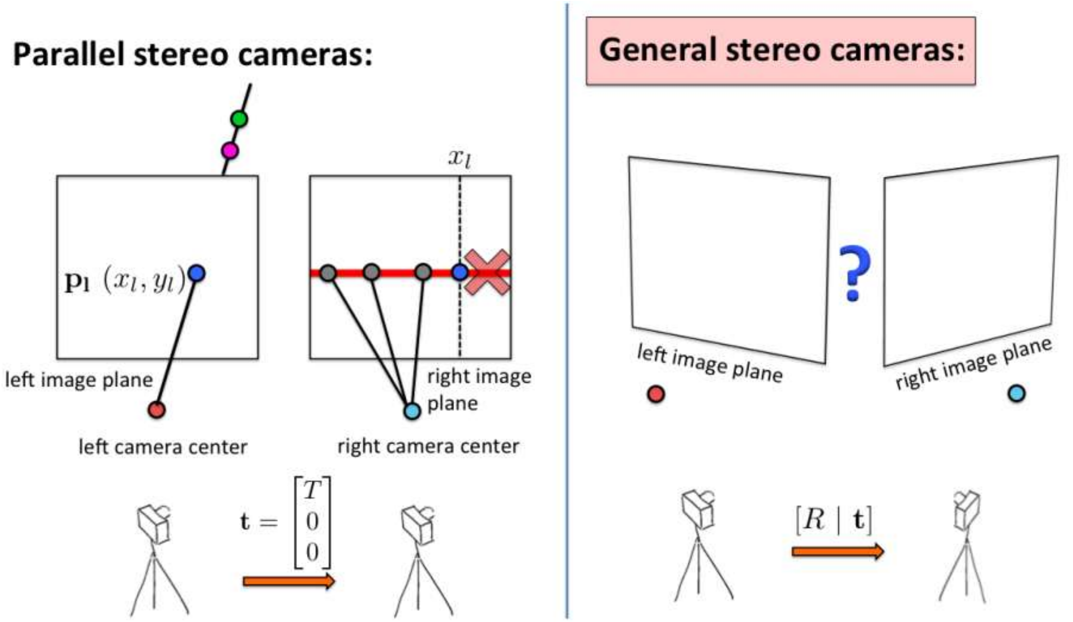
# Visión Computacional

Ivan Sipiran

# Geometría Epipolar

Geometría epipolar

- Caso con dos cámaras con ejes ópticos paralelos
- Caso general



# Geometría Epipolar

Para qué se necesita el caso general, si se puede usar cámaras paralelas?

# Geometría Epipolar

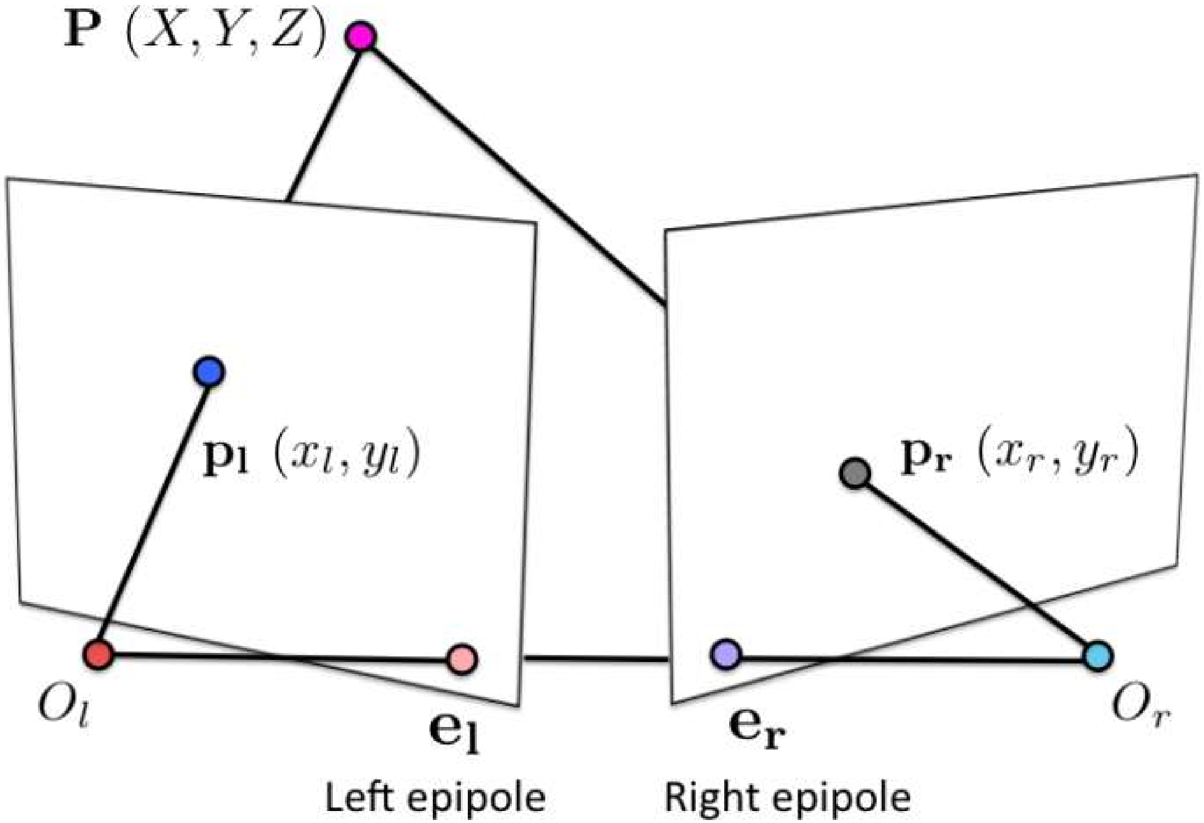
Se quiere reconstruir una torre en 3D

- No se pueden obtener fotos que cubran la torre entera, solo con cámaras paralelas
- Pero puedes descargar imágenes de la web, pero no son con cámaras paralelas...

# Cámaras

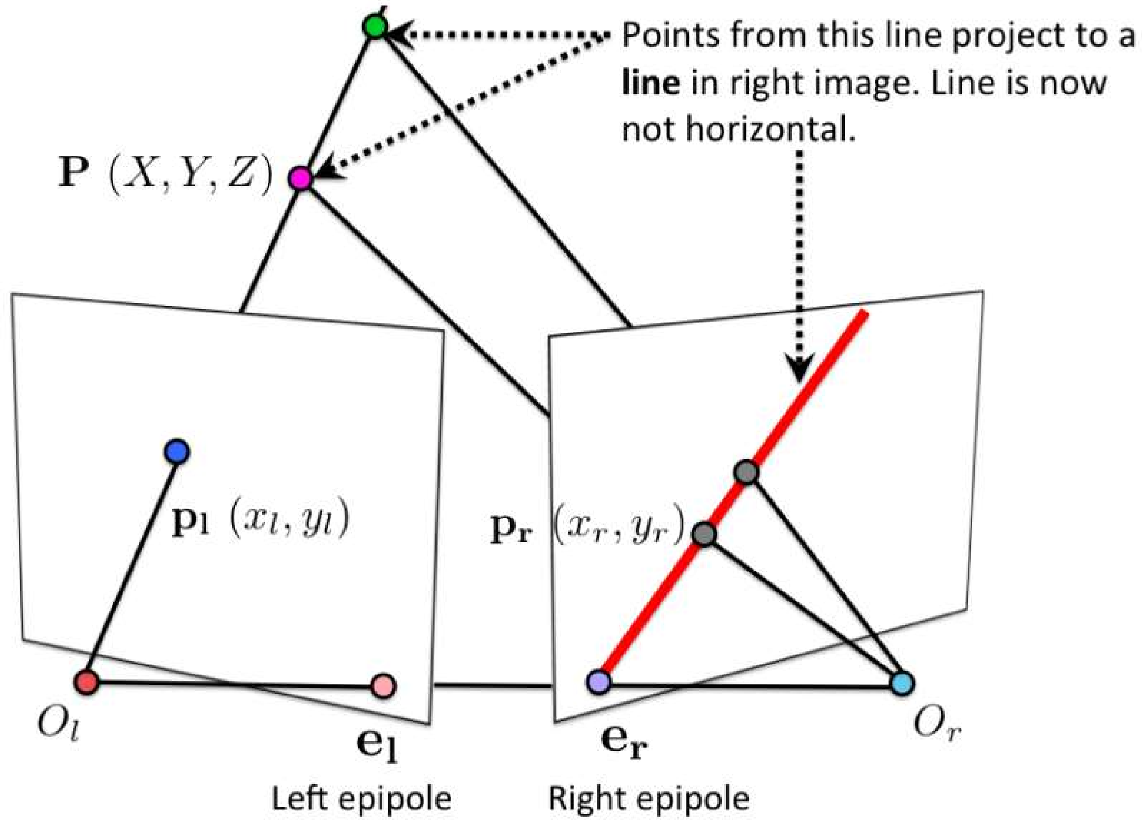
## Alguna notación

Línea $O_l O_r$ intersecta los planos de imágenes



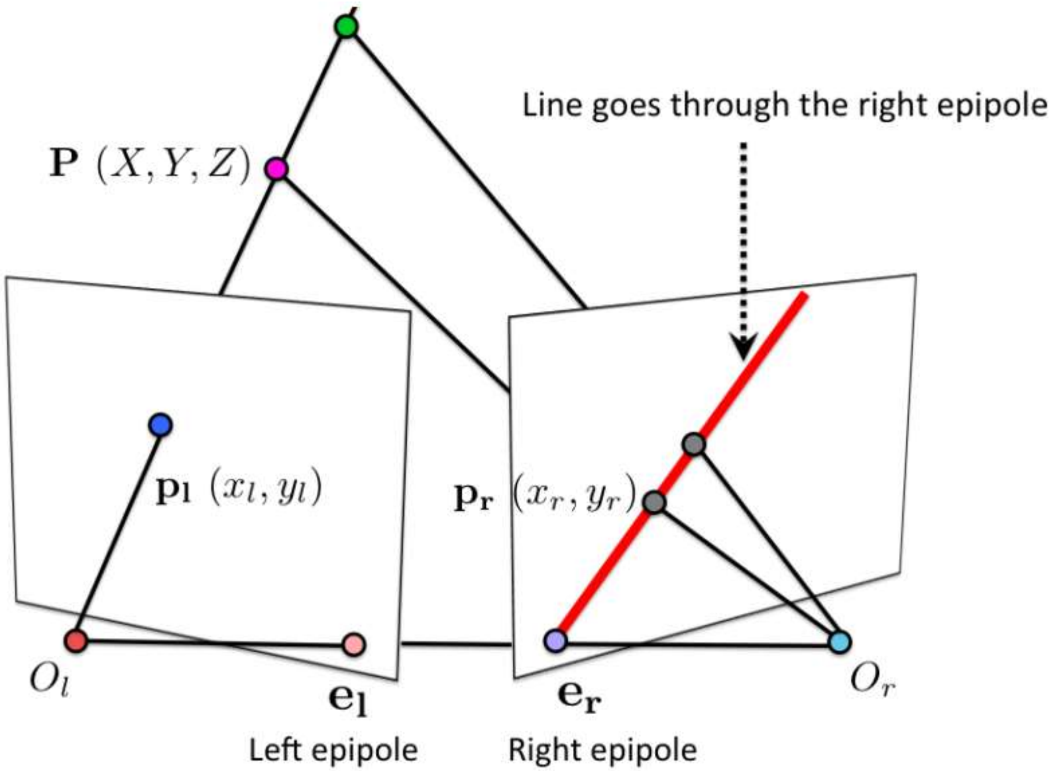
# Cámaras

Todos los puntos desde la línea proyectiva $O_l p_l$ proyectan a una línea en la imagen de la derecha. Esta vez la línea no es horizontal.



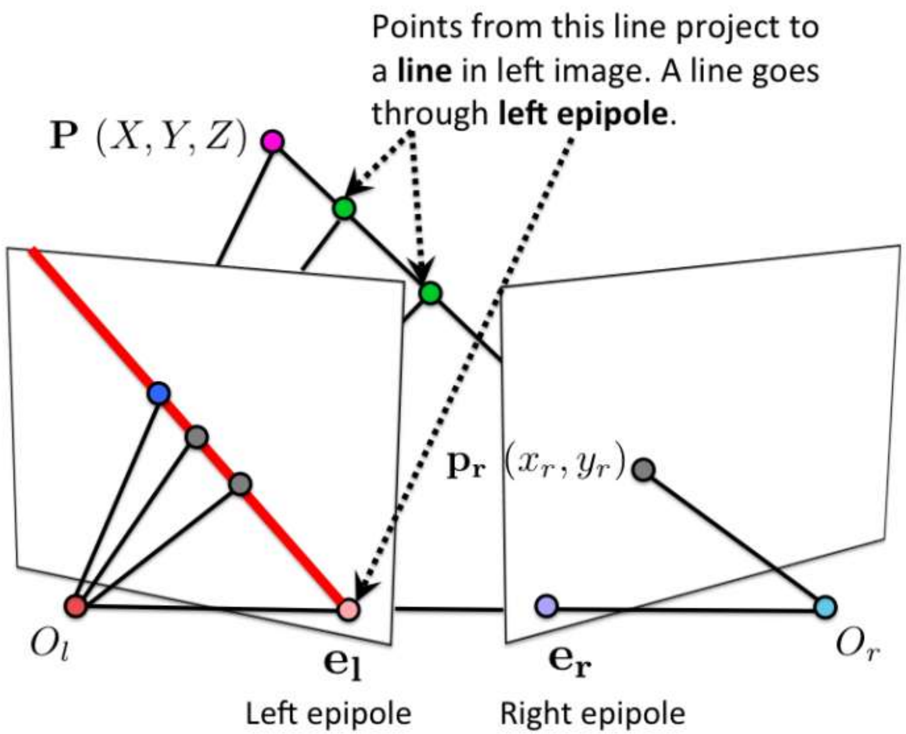
# Cámaras

La línea pasa por el epipolo derecho



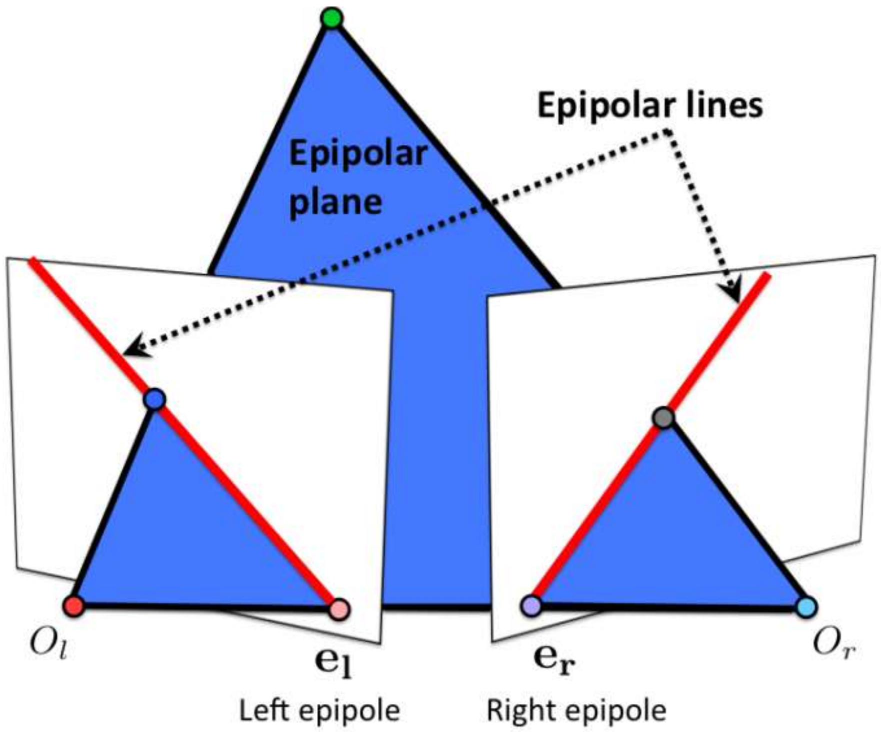
# Cámaras

Similarmente, todos los puntos en la línea proyectiva $O_r p_r$ proyectan a una línea en la imagen de la izquierda. Esta línea pasa a través del epipolo izquierdo



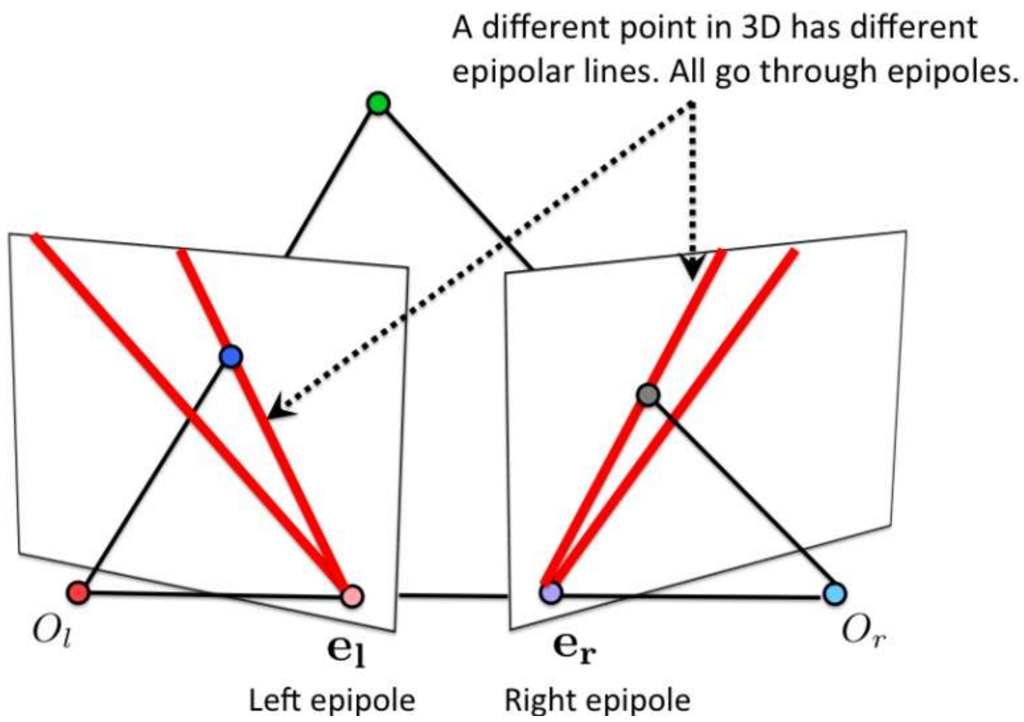
# Cámaras

Puntos $O_l, O_r$ y un punto $P$ en 3D caen en un plano. Lo llamamos el **plano epipolar.** Este plano intersecta cada imagen en una línea. Le llamamos **líneas epipolares.**



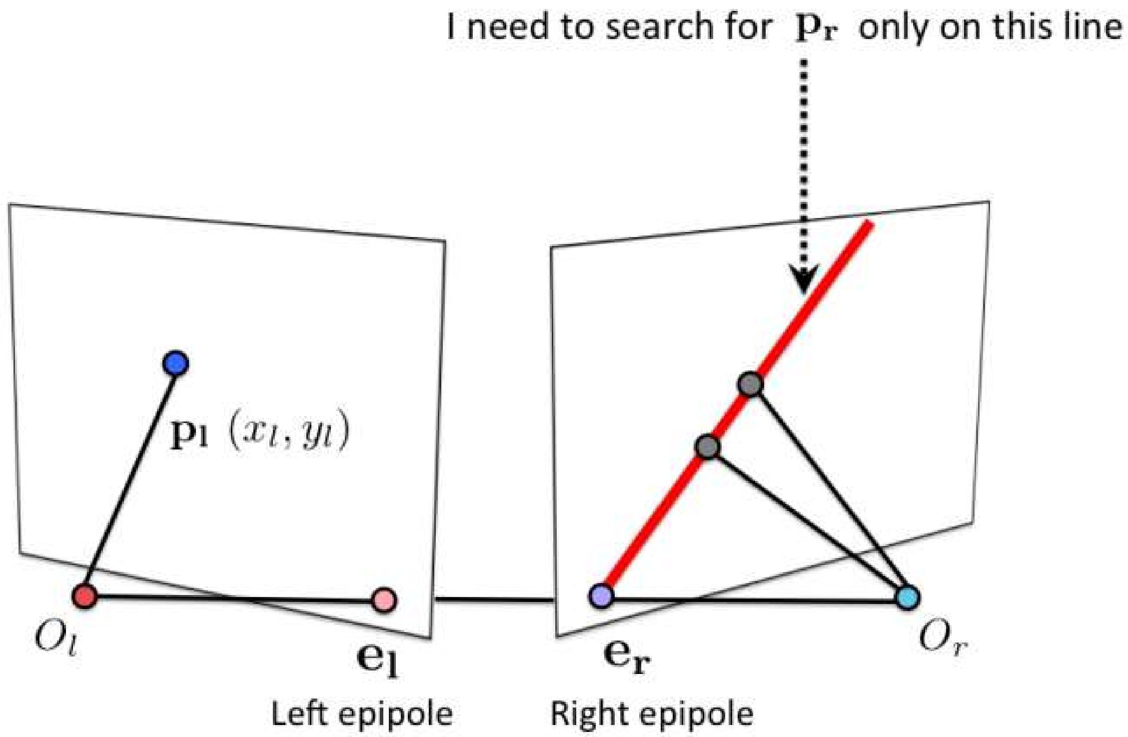
# Cámaras

Obviamente un punto diferente en 3D formará un plano epipolar diferente y diferentes líneas epipolares. Pero también pasan por los epipolos.



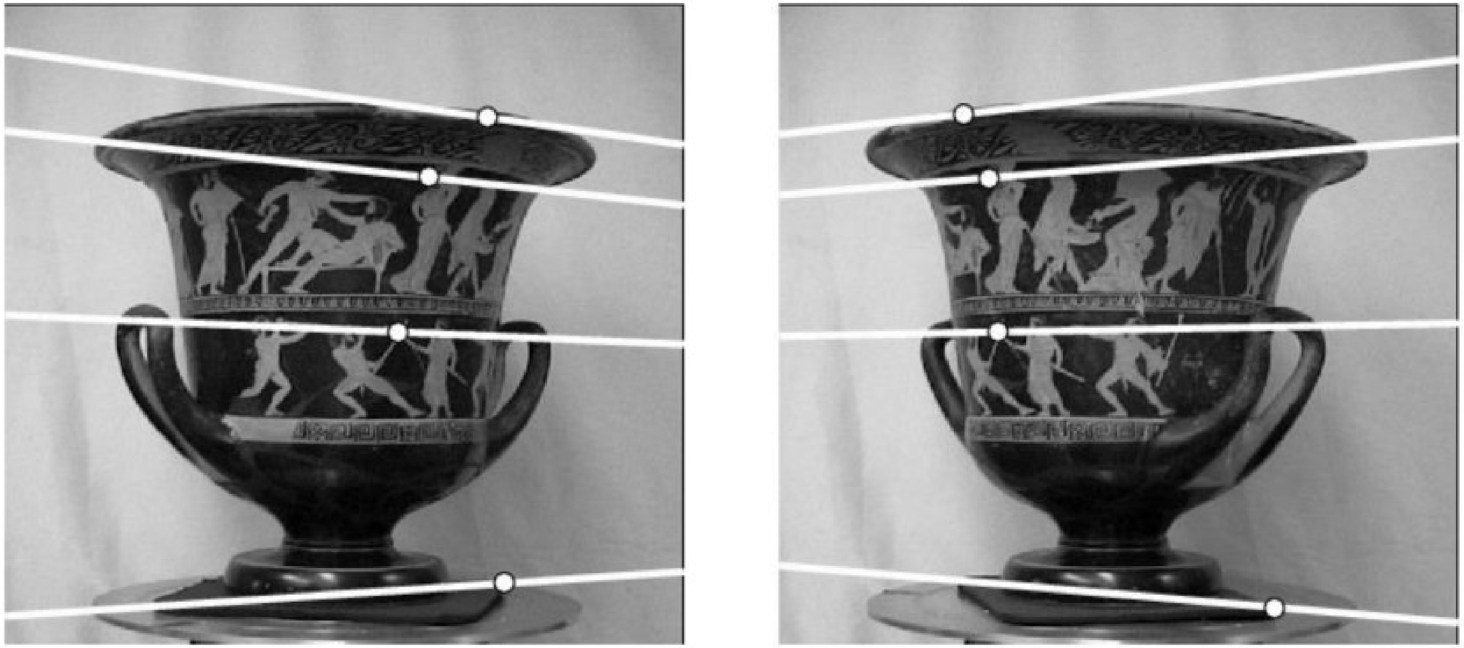
# Cámaras

- Geometría epipolar es útil porque restringe la búsqueda de matches.
- Para cada punto $p_l$, necesitamos buscar $p_r$ solo en la línea epipolar
- Todo match debe caer en una línea que intersecta en epipolos



# Cámaras

- Ejemplo de líneas epipolares en cámaras convergentes



# Estéreo para cámaras generales

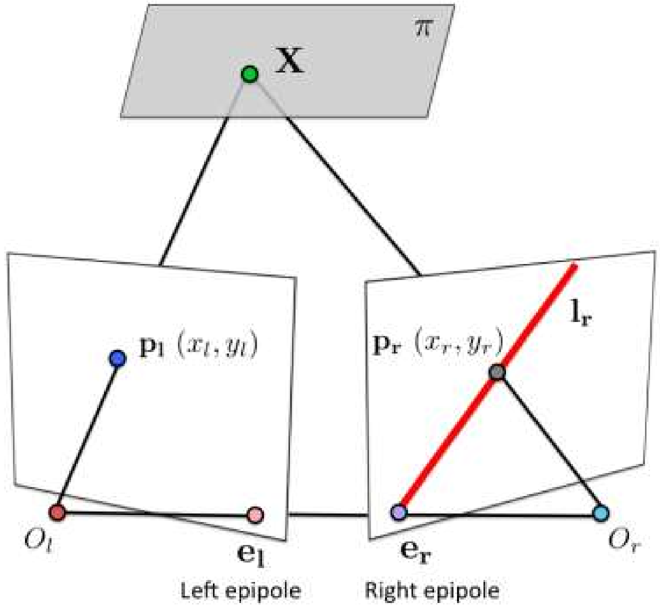
- Primero tenemos que saber en qué línea buscar matches
- Cada punto en la imagen de la izquierda mapea a una línea en la derecha. Veremos que este mapping puede ser descrito por una matriz $\boldsymbol{F}$ de 3x3 , llamada la **matriz fundamental**.
- Dada $\boldsymbol{F}$, puedes rectificar las imágenes tal que las líneas epipolares sean horizontales
- En ese punto se parece al análisis de cámaras paralelas

# Estéreo para cámaras generales

- La matriz fundamental $\boldsymbol{F}$ es definida como $\boldsymbol{I_r} = \boldsymbol{F}\boldsymbol{p_l}$ donde $\boldsymbol{I_r}$ es la línea epipolar de la derecha que corresponde a $\boldsymbol{p_l}$.
- $\boldsymbol{F}$ es una matriz de $3 \times 3$
- Para cada punto $\boldsymbol{p_l}$ la línea epipolar está definida por la misma matriz

# La matriz fundamental

- Extiende la línea $O_l p_l$ hasta llegar a un plano arbitrario $\pi$
- Encuentra $p_r$ en la cámara de la derecha
- Obtener la línea epipolar $I_r$ desde $e_r$ a $p_r$: $I_r = e_r \times p_r$
- Puntos $p_l$ y $p_r$ están relacionados con una homografía: $p_r = H_\pi p_l$
- Entonces: $I_r = e_r \times p_r = e_r \times H_\pi p_l = \mathrm{F} p_l$



# La matriz fundamental

- La matriz fundamental $F$ es definida como $I_r = Fp_l$
- Hacemos un truco

$$p_r^T \cdot I_r = p_r^T \, Fp_l$$

# La matriz fundamental

- La matriz fundamental $F$ es definida como $I_r = Fp_l$
- Hacemos un truco

$$\underbrace{p_r^T \cdot I_r}_{=\,0} = p_r^T\, F p_l$$

porque $p_r$ cae en la línea $I_r$

# La matriz fundamental

- La matriz fundamental $F$ es definida como $I_r = Fp_l$
- Hacemos un truco

$$\underbrace{p_r^T \cdot I_r}_{=0} = p_r^T \; F p_l$$

porque $p_r$ cae en la línea $I_r$

- Entonces

$$\boxed{p_r^T \; F p_l = 0}$$

Para cualquier match $(p_l, p_r)$

# La matriz fundamental

- Supongamos que encontramos algunos matches

$$(x_{l,1}, y_{l,1}) \leftrightarrow (x_{r,1}, y_{r,1}), \ \ldots, \ (x_{l,n}, y_{l,n}) \leftrightarrow (x_{r,n}, y_{r,n})$$

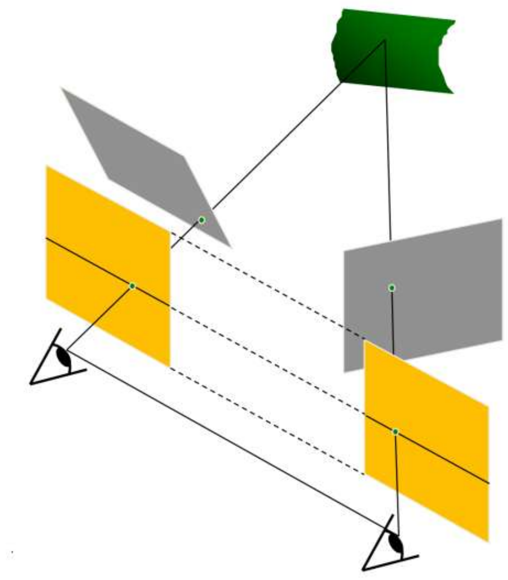
- Podemos calcular los valores de $F$ así

$$\begin{bmatrix} x_{r,1}\, x_{l,1} & x_{r,1}\, y_{l,1} & x_{r,1} & y_{r,1}\, x_{l,1} & y_{r,1}\, y_{l,1} & y_{r,1} & x_{l,1} & y_{l,1} & 1 \\ & & & & \vdots & & & & \\ x_{r,n}\, x_{l,n} & x_{r,n}\, y_{l,n} & x_{r,n} & y_{r,n}\, x_{l,n} & y_{r,n}\, y_{l,n} & y_{r,n} & x_{l,n} & y_{l,n} & 1 \end{bmatrix} \mathbf{f} = \mathbf{0}$$

- Cuántas correspondencias se necesitan?
- $F$ tiene 9 elementos, pero descartando escala, son 8
- En realidad se puede con 7 correspondencias. Pero con más, mejor

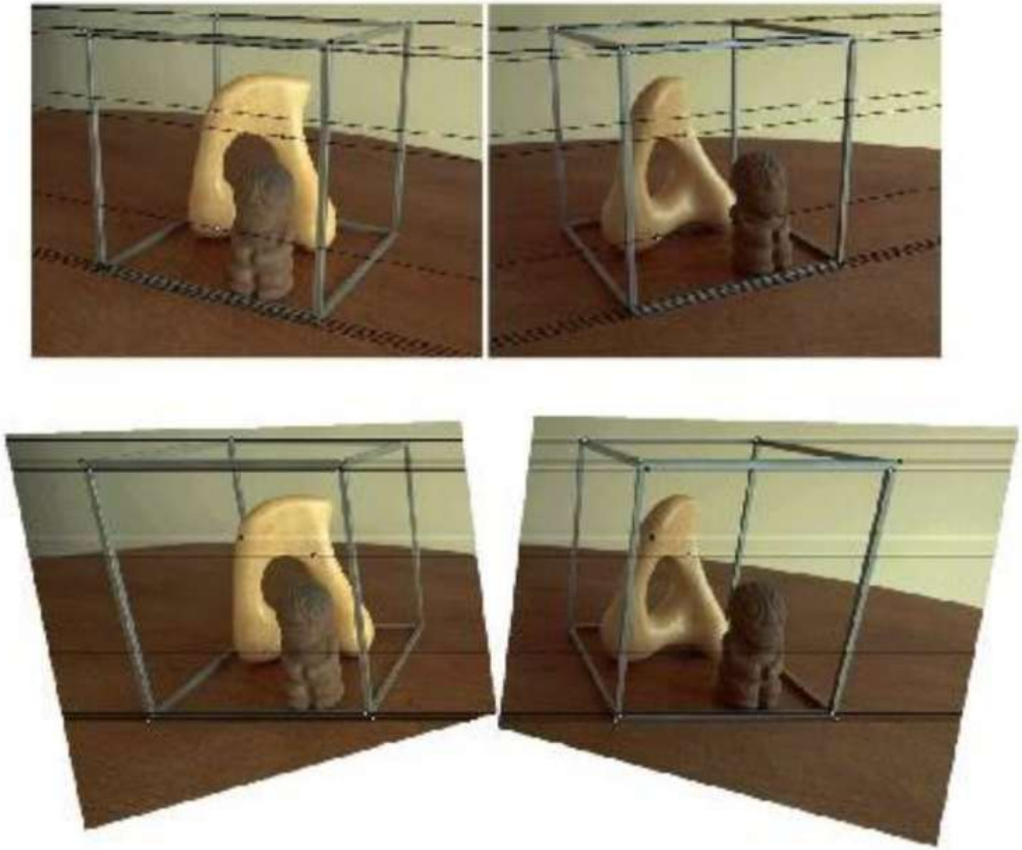
# Rectificación

- Una vez que tenemos $F$ podemos computar homografías que transforman los planos de la imagen para que sean paralelos
- Una vez que son paralelos, sabemos cómo hacer



# Rectificación

- Ejemplo de rectificación



# Matriz fundamental

- Una vez que tienes $F$, puedes computar matrices de proyección de cámara $P_l$ y $P_r$. Puede hacerse así:

$$P_{left} = [\, I_{3\times 3} \mid \mathbf{0} \,] \qquad P_{right} = \left[\, [\mathbf{e_r}]_\times F \mid \mathbf{e_r} \,\right]$$

$$[\mathbf{a}]_\times = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

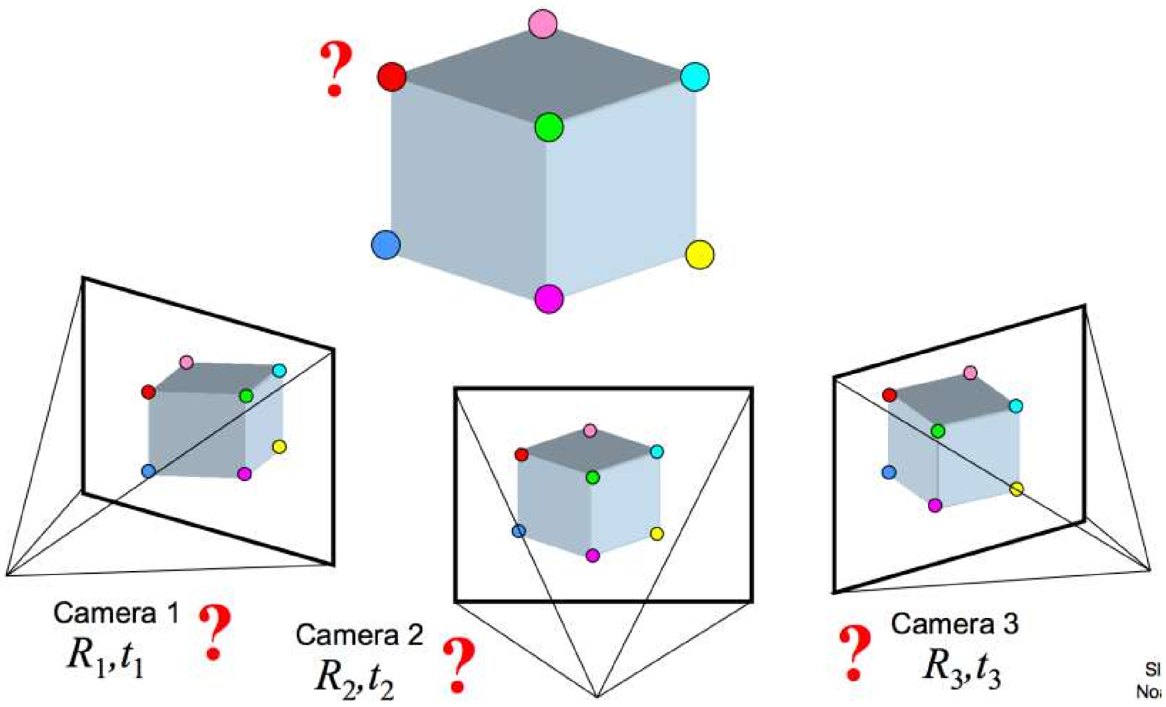
- Poses relativas de cámaras se pueden calcular!
- Útil cuando las imágenes tienen origen desconocido

# Matriz fundamental

- Para computar $P_{right}$ necesito $e_r$
- Sabemos que $e_r^T I_r = 0$
- Sabemos que $I_r = Fx_l$, así que $e_r F x_l = 0$ para todo $x_l$
- Esto solo es posible si $e_r^T F = 0$. Así que $e_r$ es una vector que mapea $F$ a cero.

# Structure from Motion (SfM)

- Que tal si tienes más de dos vistas?
- El problema es llamado structure-from-motion



# Structure from Motion (SfM)

- Resolver un problema de optimización no lineal que minimice el error de re-proyección

$$E(\mathbf{P}, \mathbf{X}) = \sum_{i=1}^{\#cameras} \sum_{j=1}^{\#points} \mathrm{dist}(\mathbf{x_{ij}}, P_i X_j)$$

- Se puede resolver con una técnica llamada bundle adjustment