

## **Lección 8: Visión 3D**

---

- 1. Introducción**
- 2. Modelo de la cámara**
- 3. Calibración**
- 4. Visión estéreo**
- 5. Búsqueda de correspondencias**
- 6. Otras técnicas de visión 3D**

## **1. Introducción**

---

- **Una cámara obtiene una proyección 2D de una escena 3D**
- **Para reconstruir la información 3D son necesarias 2 o más cámaras**
  - **2 cámaras: estéreo binocular**
  - **3 cámaras: estéreo trinocular**
- **Se necesita un modelo de la geometría y la óptica de las cámaras**
- **Problema fundamental: búsqueda de correspondencias**
  - **dado un elemento de una imagen (punto, recta, región, ...) encontrar el elemento correspondiente en las otras imágenes**

## 2. Modelo de la cámara (Tsai)

- **Proyección de un punto en la imagen**

$$\mathbf{u} = \text{Proy}(\mathbf{x}_w, \theta_c)$$

$$\mathbf{x}_w = \begin{pmatrix} x_w \\ y_w \\ z_w \end{pmatrix} \quad \text{Coordenadas de un punto en la referencia absoluta W}$$

$$\mathbf{u} = \begin{pmatrix} u \\ v \end{pmatrix} \quad \text{Coordenadas de su proyección en la imagen (en pixels)}$$

$$\theta_c = \begin{pmatrix} \theta_{int} \\ \theta_{ext} \end{pmatrix} \quad \text{Parámetros de la cámara intrínsecos y extrínsecos}$$

- **Parámetros intrínsecos**

- Geometría y óptica de la cámara
- Muestreo de la tarjeta de adquisición

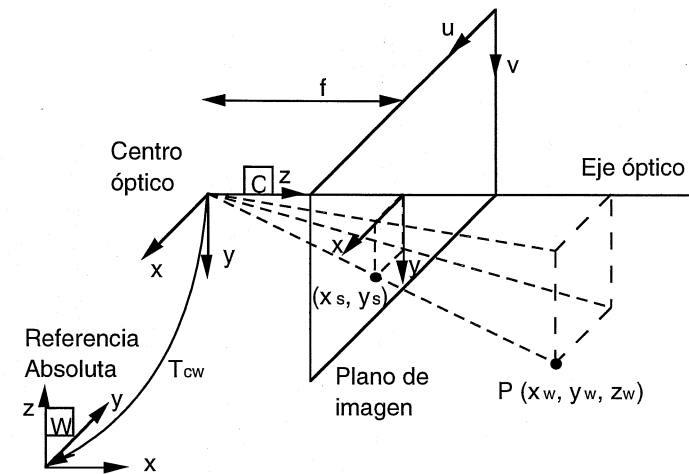
- **Parámetros extrínsecos**

- Posición y orientación de la cámara en el espacio

- **Calibración: determinación experimental de los parámetros de la cámara**

## Modelo de Tsai de la cámara

### 1) Paso a la referencia de la cámara



$$\mathbf{x}_c = R_{cw} \mathbf{x}_w + \mathbf{t}_{cw}$$

$$\mathbf{x}_c = \begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} \quad \text{Coordenadas de un punto en la referencia de la cámara}$$

$$\mathbf{t}_{cw} = \begin{pmatrix} x_{cw} \\ y_{cw} \\ z_{cw} \end{pmatrix} \quad \text{Vector de traslación}$$

$$\mathbf{R}_{cw} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \quad \text{Matriz de rotación (ortonormal)}$$

## Modelo de Tsai de la cámara

- En coordenadas homogéneas:

$$\begin{pmatrix} x_c \\ y_c \\ z_c \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{R}_{cw} & \mathbf{t}_{cw} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} = \mathbf{T}_{cw} \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix}$$

$\mathbf{T}_{cw}$  Posición y orientación de la referencia absoluta respecto de la cámara

$\mathbf{T}_{wc} = \mathbf{T}_{cw}^{-1}$  Posición y orientación de la cámara en la referencia absoluta

$$\mathbf{T}_{wc} = \begin{pmatrix} \mathbf{R}_{cw}^T & -\mathbf{R}_{cw}^T \mathbf{t}_{cw} \\ 0 & 1 \end{pmatrix}$$

- La orientación viene dada por 3 ángulos:

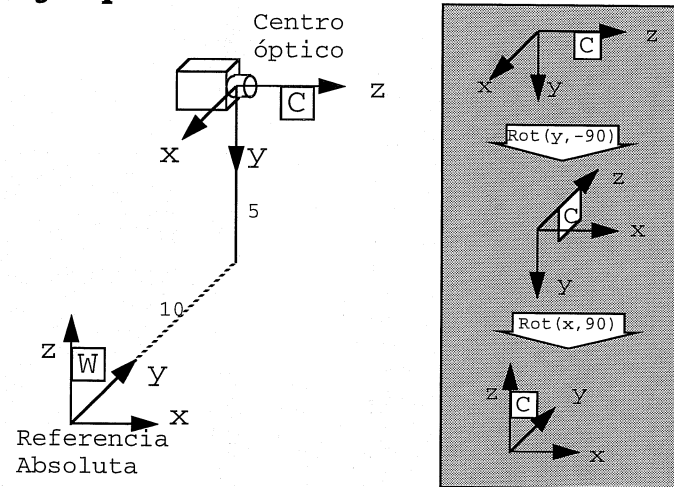
Los giros se hacen en este orden

$$RPY(\phi, \theta, \psi) = Rot(z, \phi)Rot(y, \theta)Rot(x, \psi) =$$

$$= \begin{pmatrix} C\phi C\theta & C\phi S\theta S\psi - S\phi C\psi & C\phi S\theta C\psi + S\phi S\psi \\ S\phi C\theta & S\phi S\theta S\psi + C\phi C\psi & S\phi S\theta C\psi - C\phi S\psi \\ -S\theta & C\theta S\psi & C\theta C\psi \end{pmatrix}$$

## Modelo de Tsai de la cámara

- Ejemplo:



$$\mathbf{R}_{cw} = Rot(z, 0)Rot(y, -90)Rot(x, 90) = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$\mathbf{T}_{cw} = \begin{pmatrix} 0 & -1 & 0 & 10 \\ 0 & 0 & -1 & 5 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

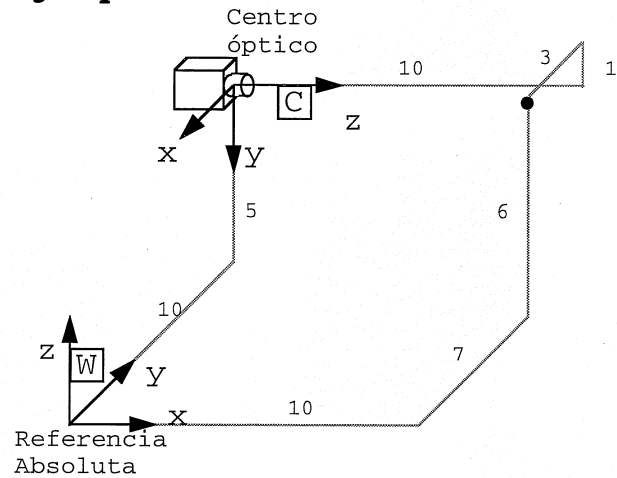
Proyectar los ejes x y z de W sobre C

$$\mathbf{T}_{wc} = \mathbf{T}_{cw}^{-1} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 10 \\ 0 & -1 & 0 & 5 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Proyectar los ejes x y z de C sobre W

## Modelo de Tsai de la cámara

• Ejemplo:

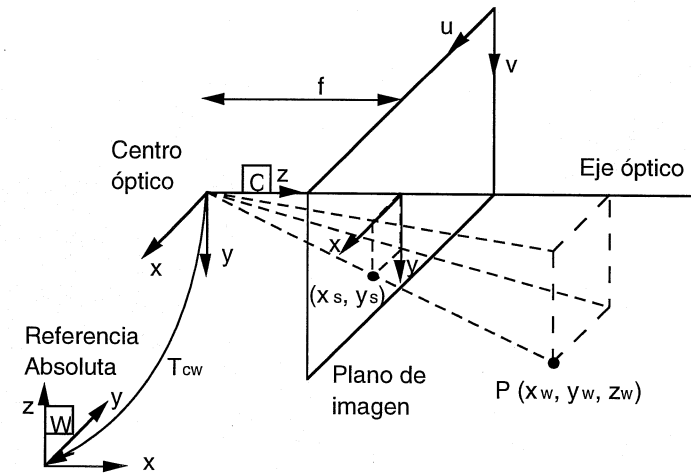


$$\begin{pmatrix} x_c \\ y_c \\ z_c \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{R}_{cw} & \mathbf{t}_{cw} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} = \mathbf{T}_{cw} \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 3 \\ -1 \\ 10 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 & 10 \\ 0 & 0 & -1 & 5 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 10 \\ 7 \\ 6 \\ 1 \end{pmatrix}$$

## Modelo de Tsai de la cámara

2) Proyección teórica sobre el sensor



$$x_{st} = f \frac{x_c}{z_c} \quad y_{st} = f \frac{y_c}{z_c}$$

$\mathbf{x}_{st} = \begin{pmatrix} x_{st} \\ y_{st} \end{pmatrix}$  Coordenadas de la proyección teórica del punto sobre el sensor (en mm)

$f$  Longitud focal del objetivo (mm)

## Modelo de Tsai de la cámara

### 3) Distorsión radial

- Puede ser importante con objetivos de tipo gran angular  $f < 12 \text{ mm}$

$$x_{st} = x_{sd}(1 + k_1 r^2)$$

$$y_{st} = y_{sd}(1 + k_1 r^2)$$

$$r^2 = x_{sd}^2 + y_{sd}^2$$

$\mathbf{x}_{sd} = \begin{pmatrix} x_{sd} \\ y_{sd} \end{pmatrix}$  Proyección distorsionada del punto sobre el sensor (en mm)

$k_1$  Factor de distorsión radial

## Modelo de Tsai de la cámara

### 4) Paso a coordenadas de la imagen (pixels)

$$u = C_x + \frac{x_{sd}}{d_x} s_x \quad v = C_y + \frac{y_{sd}}{d_y}$$

$\mathbf{u} = \begin{pmatrix} u \\ v \end{pmatrix}$  Coordenadas en pixels del punto en la imagen

$\mathbf{C} = \begin{pmatrix} C_x \\ C_y \end{pmatrix}$  Coordenadas en pixels del centro óptico en la imagen

$\mathbf{d} = \begin{pmatrix} d_x \\ d_y \end{pmatrix}$  Dimensiones teóricas del pixel de la imagen adquirida, en mm

$s_x$  Factor de corrección del muestreo horizontal (debe ser  $\approx 1$ )

#### • Cálculo de $d_x$ y $d_y$ para una cámara CCD:

tomando todas las filas

$$d_y = d_{cy}$$

ó

tomando la mitad de las filas

$$d_y = 2d_{cy}$$

$$d_x = d_{cx} \frac{N_{sx}}{N_{tx}}$$

$\mathbf{d}_c = \begin{pmatrix} d_{cx} \\ d_{cy} \end{pmatrix}$  Dimensiones de la celda del CCD

$N_{sx}$  Pixels por fila del sensor

$N_{tx}$  Pixels por fila adquiridos por la tarjeta

## Modelo de Tsai de la cámara

- **Ejemplo: cámara Pulnix TM-6**
  - **Pixels:** 752(H) x 582 (V)
  - **Cell size:** 8.4µm x 8.2 µm

### a) Muestreo 512(H) x 512(V)

$$d_x = 8.4 \times 10^{-3} \frac{752}{512} = 12.34 \times 10^{-3} \text{ mm}$$

$$d_y = 8.2 \times 10^{-3} \text{ mm}$$

$$PAR = \frac{d_x}{d_y} = 1.5$$

### b) Muestreo 768(H) x 512(V)

"square pixel"  
PAR = 1

$$d_x = 8.4 \times 10^{-3} \frac{752}{768} = 8.2 \times 10^{-3} \text{ mm}$$

$$d_y = 8.2 \times 10^{-3} \text{ mm}$$

## Modelo de Tsai de la cámara

- **Resumen del modelo de Tsai:**

$$\mathbf{u} = \text{Proy}(\mathbf{x}_w, \theta_c)$$

### Paso a referencia de la cámara

$$\mathbf{x}_c = \mathbf{R}_{cw} \mathbf{x}_w + \mathbf{t}_{cw}$$

### Proyección teórica

$$x_{st} = f \frac{x_c}{z_c}$$

$$y_{st} = f \frac{y_c}{z_c}$$

### Distorsión radial

$$x_{st} = x_{sd} (1 + k_1 r^2)$$

$$y_{st} = y_{sd} (1 + k_1 r^2)$$

$$r^2 = x_{sd}^2 + y_{sd}^2$$

### Paso a pixels en la imagen

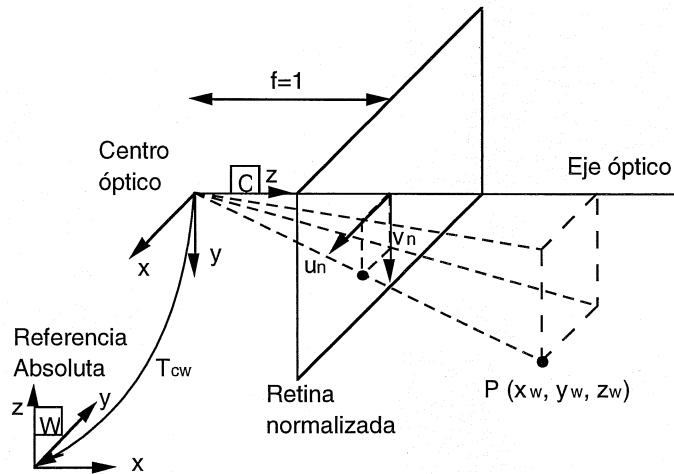
$$u = C_x + \frac{x_{sd}}{d_x} s_x$$

$$v = C_y + \frac{y_{sd}}{d_y} s_y$$

- **Constantes conocidas:  $d_x, d_y$**
- **Parámetros a calibrar:**
  - » **Extrínsecos:**  $\mathbf{t}_{cw}, \mathbf{R}_{cw}$
  - » **Intrínsecos:**  $f, k_1, C_x, C_y, s_x$

## Retina normalizada

- Es una retina ideal (sin distorsión) situada a una distancia focal  $f = 1$



- Depende de los parámetros extrínsecos:

$$\mathbf{u}_n = \text{Proy\_norm}(\mathbf{x}_w, \theta_{ext})$$

$$\mathbf{x}_c = \mathbf{R}_{cw} \mathbf{x}_w + \mathbf{t}_{cw}$$

$$u_n = \frac{x_c}{z_c} \quad v_n = \frac{y_c}{z_c}$$

$$\mathbf{u}_n = \begin{pmatrix} u_n \\ v_n \end{pmatrix} \quad \text{Coordenadas normalizadas, en mm}$$

## Retina normalizada

- Paso de la retina real a la normalizada:

$$\mathbf{u}_n = \text{Normal}(\mathbf{u}, \theta_{int})$$

$$x_{sd} = (u - C_x) \frac{d_x}{s_x} \quad y_{sd} = (v - C_y) d_y$$

$$r^2 = x_{sd}^2 + y_{sd}^2$$

$$x_{st} = x_{sd} (1 + k_1 r^2) \quad y_{st} = y_{sd} (1 + k_1 r^2)$$

$$u_n = \frac{x_{st}}{f} \quad v_n = \frac{y_{st}}{f}$$

### 3. Calibración

- **Modelo de la cámara:**

$$u = \text{Proy}(x_w, \theta_c)$$

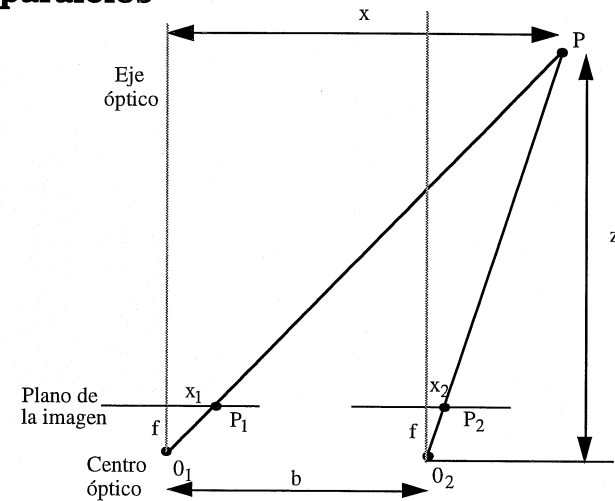
- **Colocar ante la cámara un patrón con N puntos, de coordenadas conocidas:**  $(x_{w1}, \dots, x_{wN})$
- **Tomar una imagen y obtener en ella las coordenadas de cada punto:**  $(u_1, \dots, u_N)$
- **Obtener el vector de parámetros  $\theta_c$  que minimiza el error cuadrático:**

$$\sum_{i=1}^N [u_i - \text{Proy}(x_{wi}, \theta_c)]$$

Para calcular todos los parámetros, el patrón no puede ser plano

### 4. Visión estéreo

- **Cámaras sin distorsión, con ejes ópticos paralelos**



$$x_1 = f \frac{x_{c1}}{z_{c1}} = f \frac{x}{z} \quad x_2 = f \frac{x_{c2}}{z_{c2}} = f \frac{x-b}{z}$$

$$\text{operando: } z = \frac{b f}{x_1 - x_2}$$

- **La profundidad es función de la disparidad:  $x_1 - x_2$**



## Visión estéreo

- **Cámaras reales, en cualquier situación**
  - Trabajamos con las retinas normalizadas

$$\mathbf{x}_{c_k} = R_{c_k w} \mathbf{x}_w + \mathbf{t}_{c_k w} = \begin{pmatrix} r_{11k} & r_{12k} & r_{13k} \\ r_{21k} & r_{22k} & r_{23k} \\ r_{31k} & r_{32k} & r_{33k} \end{pmatrix} \mathbf{x}_w + \begin{pmatrix} x_{c_k w} \\ y_{c_k w} \\ z_{c_k w} \end{pmatrix}$$

$$\begin{pmatrix} x_{c_k} \\ y_{c_k} \\ z_{c_k} \end{pmatrix} = \begin{pmatrix} \mathbf{r}_{1k} \mathbf{x}_w + x_{c_k w} \\ \mathbf{r}_{2k} \mathbf{x}_w + y_{c_k w} \\ \mathbf{r}_{3k} \mathbf{x}_w + z_{c_k w} \end{pmatrix}; \quad \mathbf{r}_{ik} = (r_{i1k} \quad r_{i2k} \quad r_{i3k})$$

$$u_{n_k} = \frac{x_{c_k}}{z_{c_k}} \quad v_{n_k} = \frac{y_{c_k}}{z_{c_k}}$$

de donde, operando:

$$\mathbf{r}_{1k} \mathbf{x}_w + x_{c_k w} = u_{n_k} (\mathbf{r}_{3k} \mathbf{x}_w + z_{c_k w})$$

$$\mathbf{r}_{2k} \mathbf{x}_w + y_{c_k w} = v_{n_k} (\mathbf{r}_{3k} \mathbf{x}_w + z_{c_k w})$$

es decir:

$$\mathbf{A}_k \mathbf{x}_w + \mathbf{b}_k = 0$$

con:

$$\mathbf{A}_k = \begin{pmatrix} \mathbf{r}_{1k} - u_{n_k} \mathbf{r}_{3k} \\ \mathbf{r}_{2k} - v_{n_k} \mathbf{r}_{3k} \end{pmatrix} \quad \mathbf{b}_k = \begin{pmatrix} x_{c_k w} - u_{n_k} z_{c_k w} \\ y_{c_k w} - v_{n_k} z_{c_k w} \end{pmatrix}$$

## Visión estéreo

- **Reconstrucción 3D de un punto, a partir de su proyección en n imágenes**
  - Tres incógnitas:  $x_w, y_w, z_w$
  - Cada cámara aporta dos ecuaciones:

$$\mathbf{A}_k \mathbf{x}_w + \mathbf{b}_k = 0$$

- **Sistema de ecuaciones:**

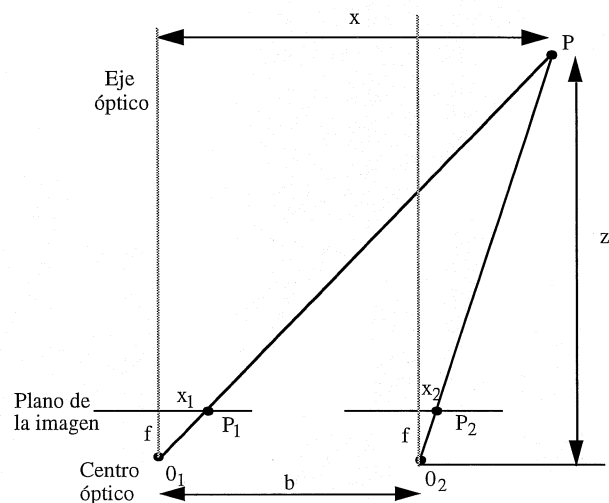
$$\mathbf{A} \mathbf{x}_w + \mathbf{b} = 0$$

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 \\ \dots \\ \mathbf{A}_n \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} \mathbf{b}_1 \\ \dots \\ \mathbf{b}_n \end{pmatrix}$$

- **Solución:**

$$\mathbf{x}_w = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$$

## Ejemplo: Cámaras paralelas



- Tomamos referencia absoluta  $W = C_1$

$$R_{C_1W} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad t_{C_1W} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$R_{C_2W} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad t_{C_2W} = \begin{pmatrix} -b \\ 0 \\ 0 \end{pmatrix}$$

## Ejemplo: Cámaras paralelas

- Sistema de ecuaciones:

$$\begin{pmatrix} 1 & 0 & -u_{n1} \\ 0 & 1 & -v_{n1} \\ 1 & 0 & -u_{n2} \\ 0 & 1 & -v_{n2} \end{pmatrix} \begin{pmatrix} x_w \\ y_w \\ z_w \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ -b \\ 0 \end{pmatrix} = 0$$

- Solución analítica:

$$z_w = \frac{b}{u_{n1} - u_{n2}}$$

$$x_w = u_{n1} \frac{b}{u_{n1} - u_{n2}}$$

$$y_w = v_{n1} \frac{b}{u_{n1} - u_{n2}}$$

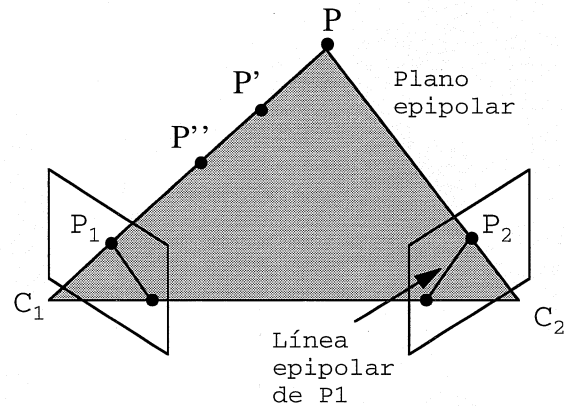
$$v_{n1} = v_{n2}$$



- Restricción epipolar: los puntos correspondientes deben estar en la misma fila en ambas imágenes

## 5. Búsqueda correspondencias

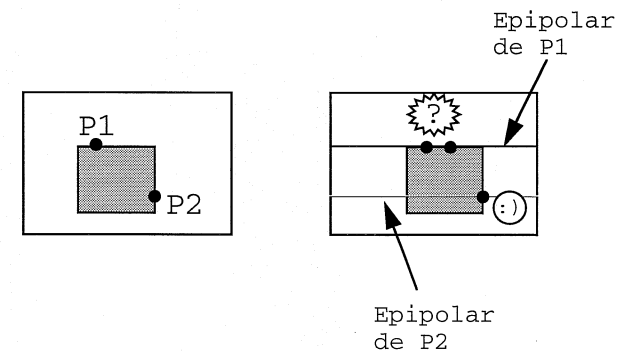
- Dado un punto  $P$  en el espacio, se llama plano epipolar de  $P$  al plano  $PC_1C_2$
- Las proyecciones  $P_1$  y  $P_2$  pertenecen al plano epipolar de  $P$



- **Restricción epipolar para  $P_1$ :**
  - Conocidos  $C_1$ ,  $C_2$  y  $P_1$ , se calcula el plano epipolar
  - La línea epipolar de  $P_1$  en la imagen 2 es la intersección del plano epipolar con el plano de imagen 2
  - El correspondiente de  $P_1$  en la imagen 2 debe estar sobre la línea epipolar
- Puede haber varios candidatos  $P_2, P_2', \dots$

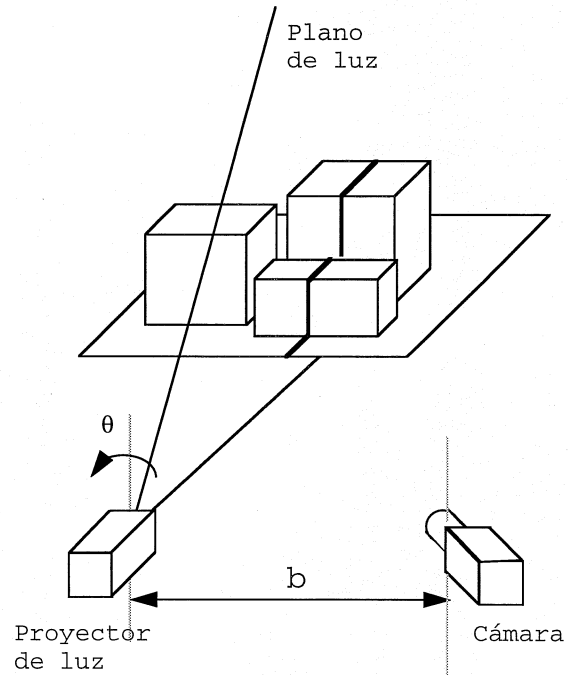
## Correspondencias de contornos

- Detectar puntos de contorno en ambas imágenes
- Para cada punto de contorno de la imagen 1, buscar a lo largo de su línea epipolar un punto de contorno con
  - magnitud de gradiente similar
  - orientación de gradiente similar
- Los puntos de contorno orientados como su línea epipolar no se pueden emparejar



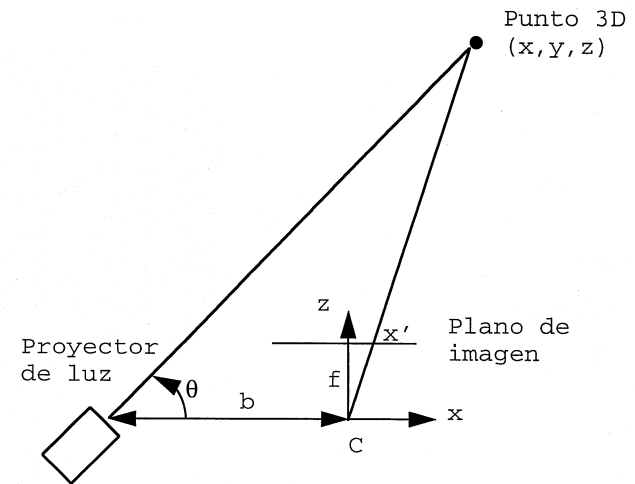
## 6. Otras técnicas de visión 3D: Iluminación estructurada

- Barrido de la escena con un punto o un plano de luz láser



## Iluminación estructurada

- Información 3D por triangulación



$$x' = f \frac{x}{z} \quad y' = f \frac{y}{z} \quad \tan \theta = \frac{z}{x+b}$$

operando:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{b}{f \cot \theta - x'} \begin{pmatrix} x' \\ y' \\ f \end{pmatrix}$$