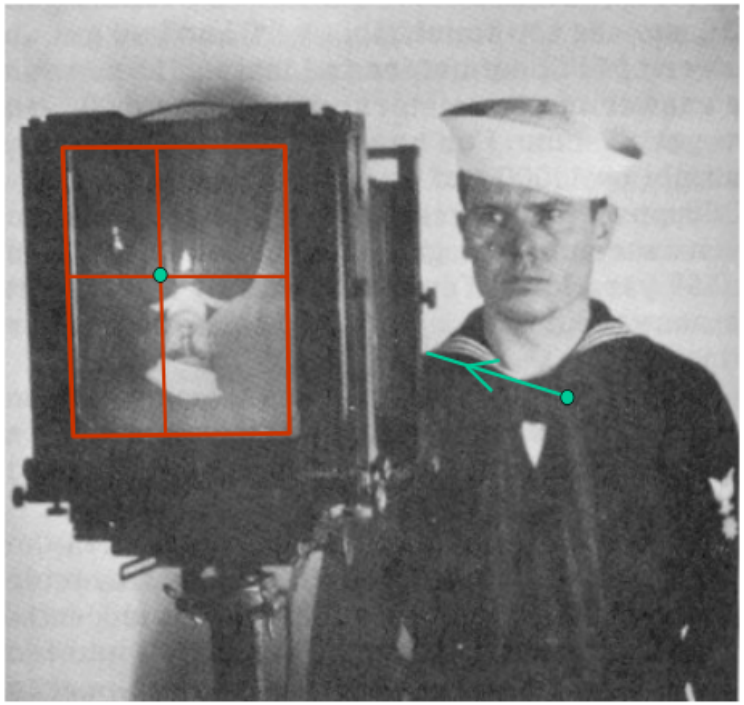
Lección 2
Fundamentos

Hugo Estrada, Reynaldo Martell, Julio Cruz, Jesús Savage, Marco Negrete, Luis Nava

5 de marzo de 2021

Modelo de Cámara Pinhole

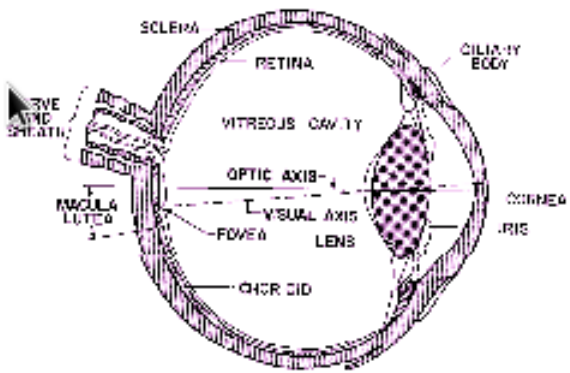
Las imágenes son arreglos bidimensionales de valores de intensidad.



Están formadas por proyecciones de objetos 3D.

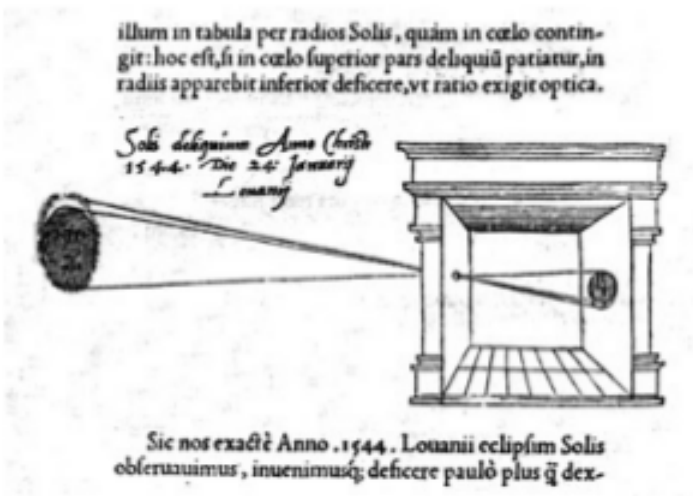
Proyecciones Cámara Pinhole

Dispositivos de imagen:



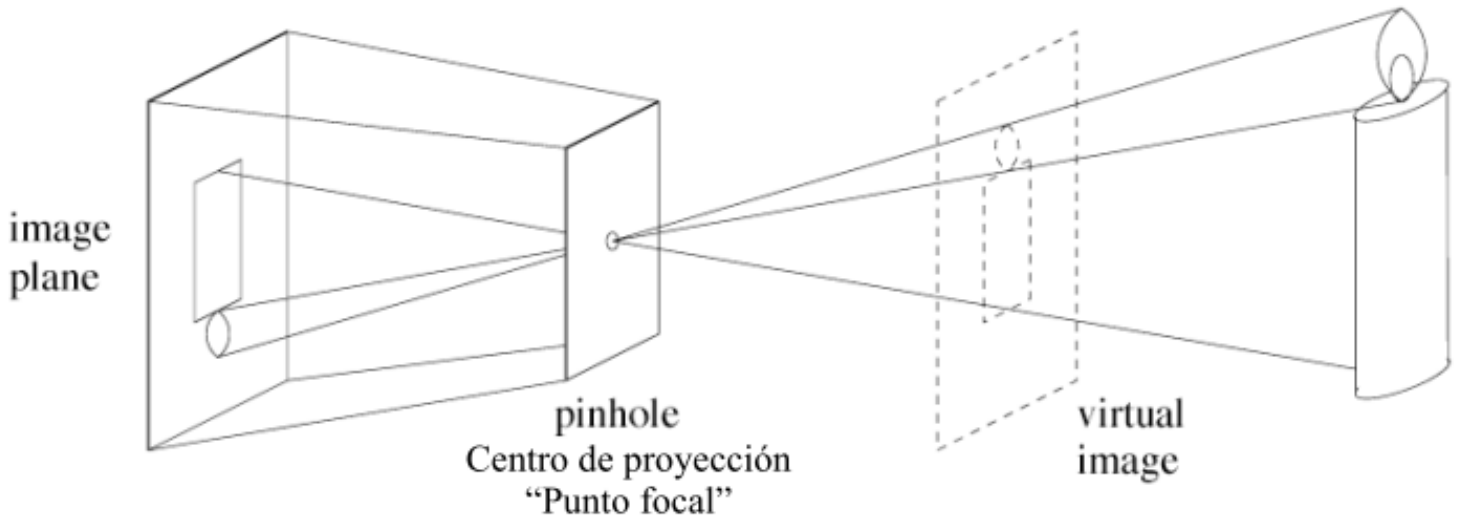
Ojos

Cámara fotográfica: Niepce, 1816



Proyecciones Cámara Pinhole

La proyección de perspectiva crea imágenes invertidas, algunas veces conviene considerar en su lugar a la imagen virtual asociada al plano que está enfrente del "pinhole" a la misma distancia que el plano de la imagen. Dependiendo del contexto.



Modelo de Cámara Pinhole

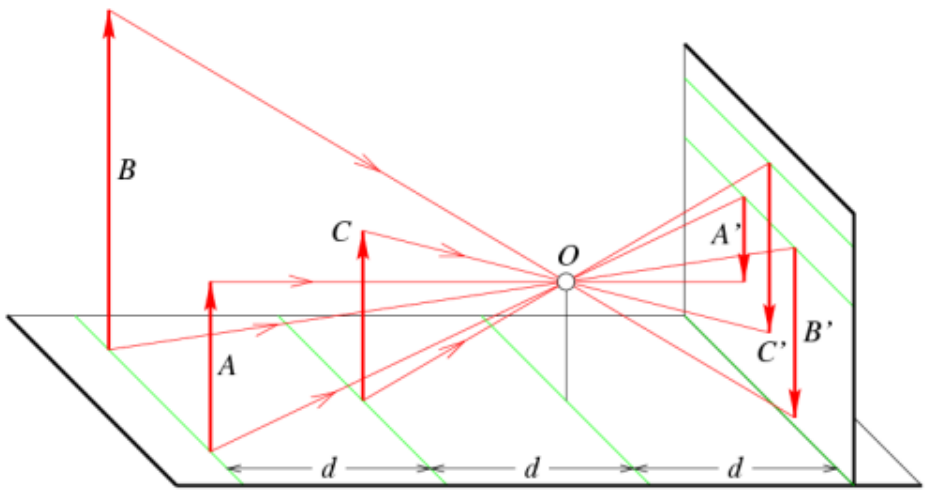
Efecto: Los objetos distantes parecen más pequeños.

Líneas

C' y B' parecen del mismo tamaño

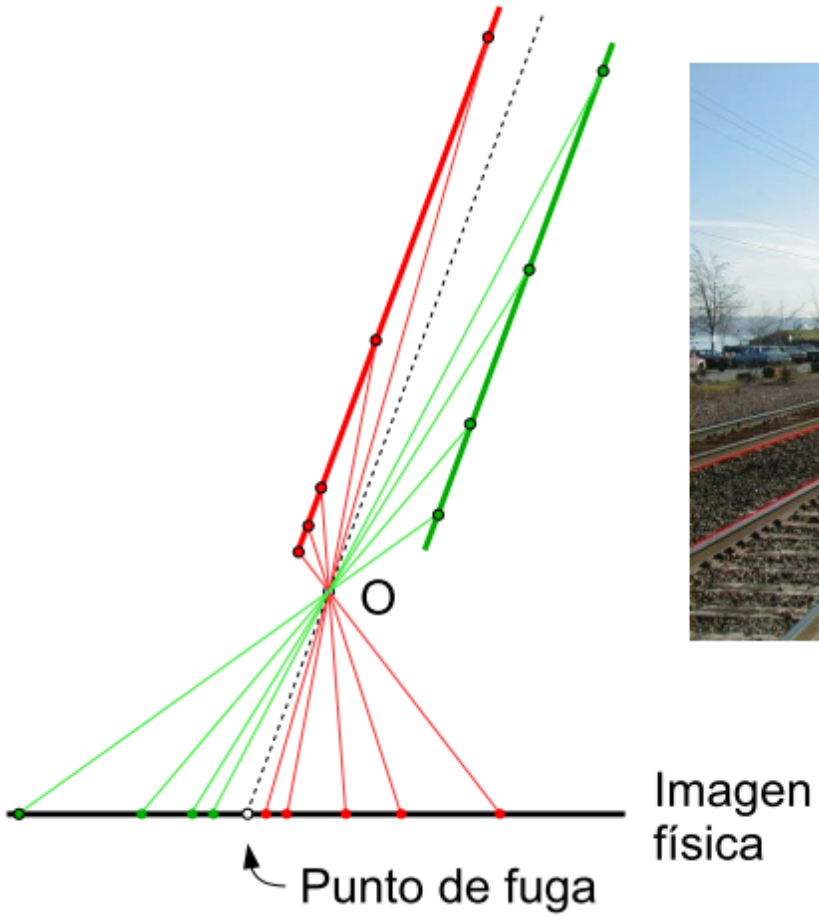
Sin embargo, $C = \frac{1}{2} B$

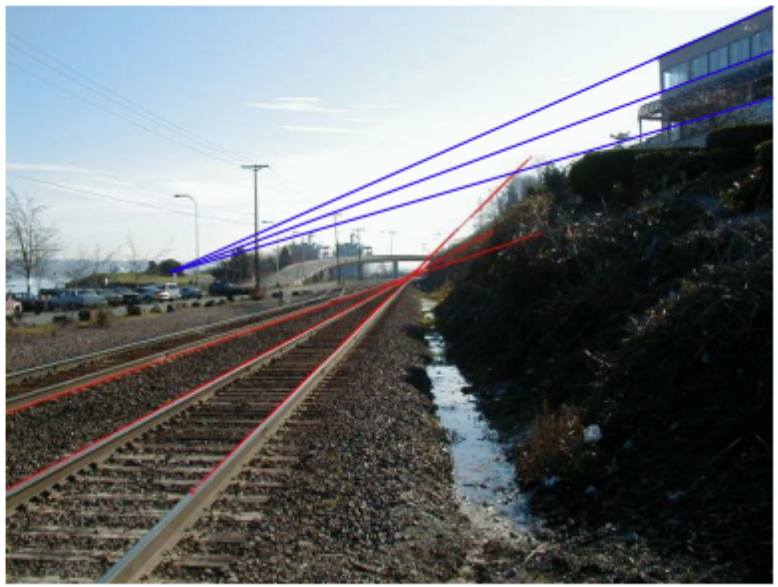
Debido a que: La distancia de B a O es 2d y la distancia C a O es d



Modelo de Cámara Pinhole

Efecto: Líneas paralelas se encuentran.

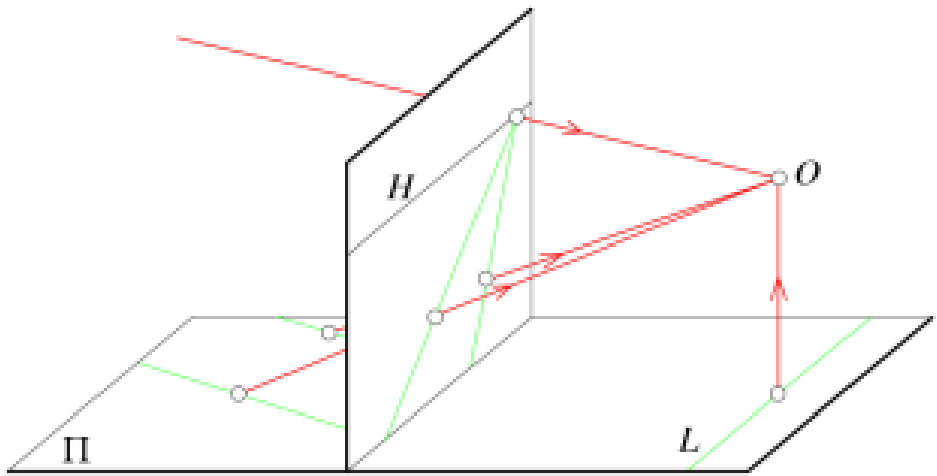




Modelo de Cámara Pinhole

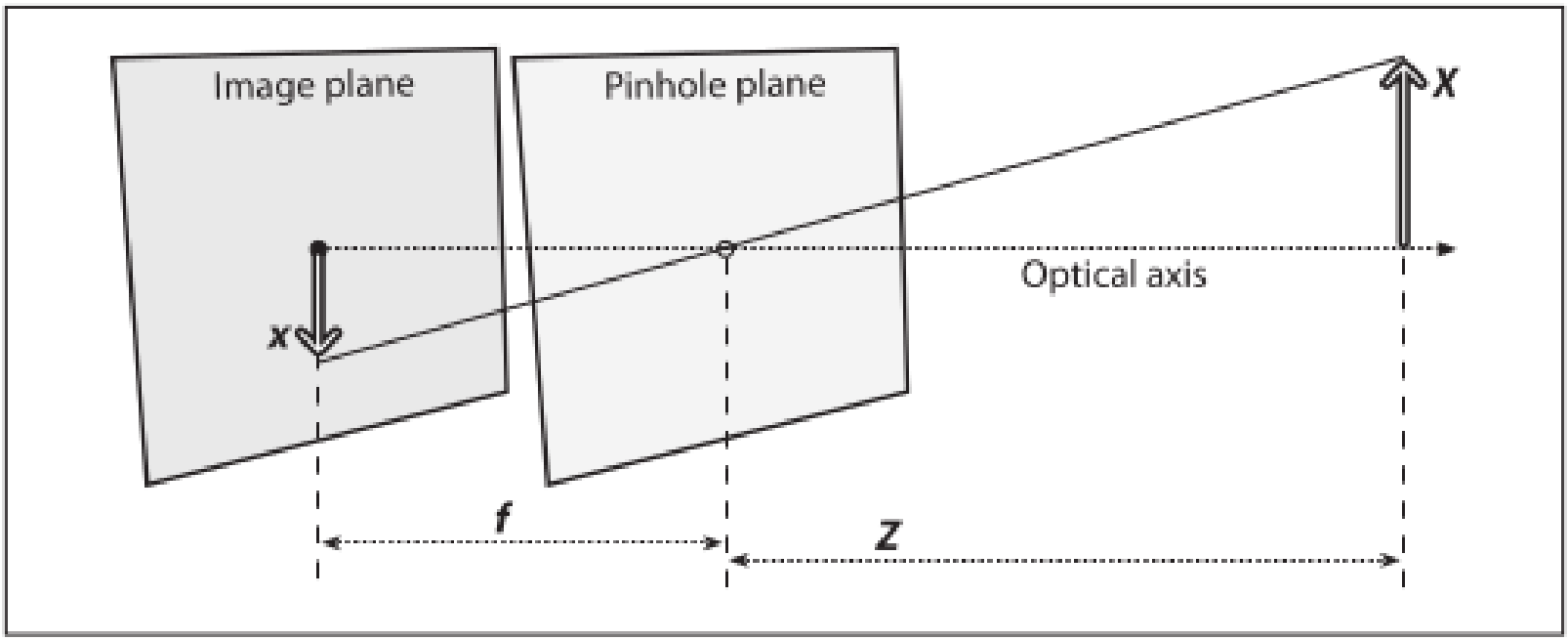
Propiedades geométricas de la proyección:

1. Puntos que van a puntos
2. Líneas que van a líneas
3. Planos que van a toda la imagen o mitad del plano
4. Polígonos que van a polígonos
5. Casos degenerados:
6. línea a través del punto focal da un punto
7. plano a través del punto focal de una línea



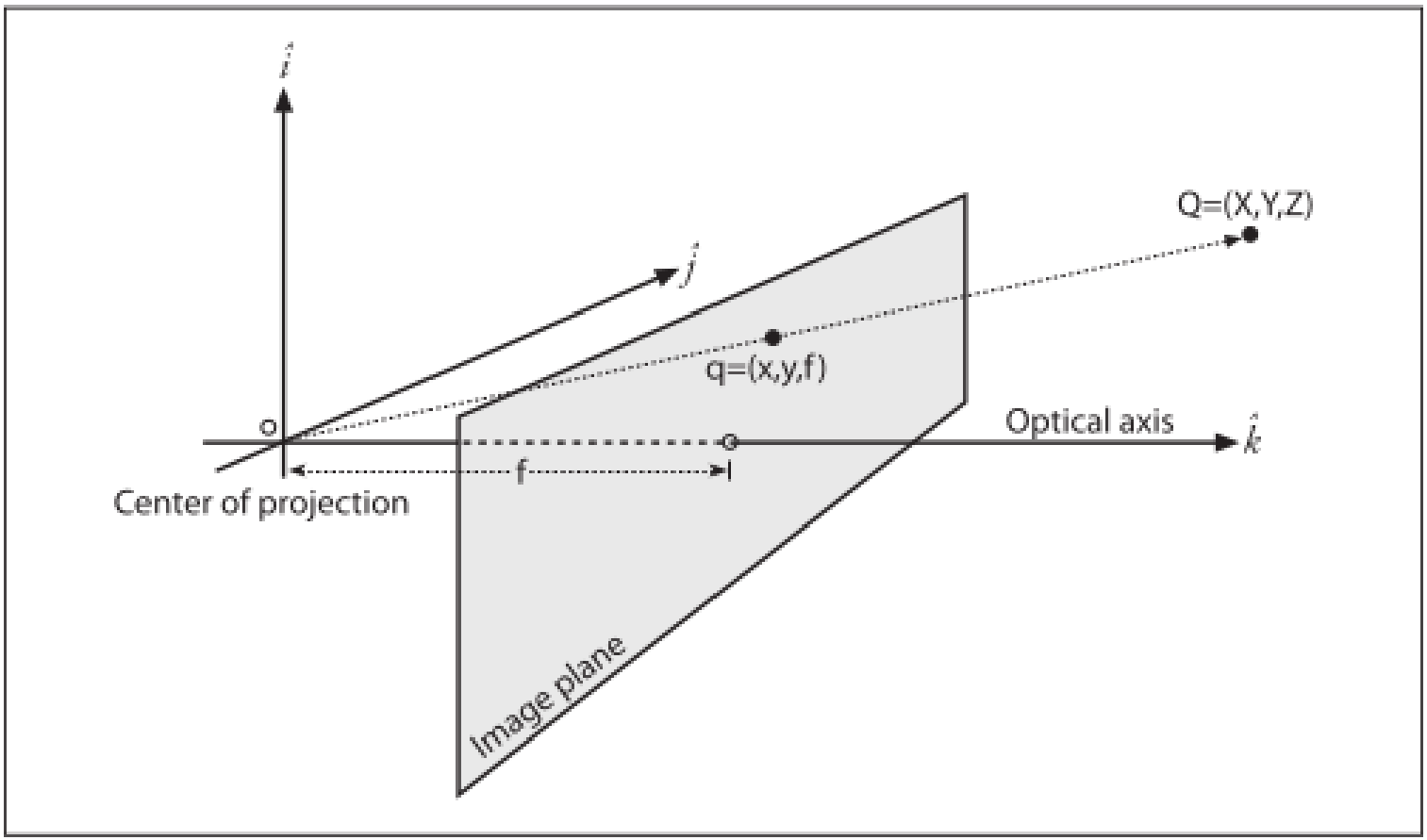
Modelo de Cámara Pinhole

Modelo del pinhole, un objeto externo se proyecta en el plano de la imágen a una distancia f, llamado punto focal, del pinhole.
El objeto real con tamaño X se encuentra a una distancia Z. Por la relación de triangulos semejantes $-x/f = X/Z$



Modelo de Cámara Pinhole

Para evitar trabajar con valores negativos y facilitar los calculos se proyecta el plano de la imágen hacia adelante, entonces un punto en el mundo real $Q = (X, Y, Z)$ se representa en la imágen como $q(x, y, f)$.



Modelo de Cámara Pinhole

Se tiene que un punto Q en el mundo físico, el cual tiene como coordenadas (X, Y, Z), es proyectado en plano de la pantalla en un punto que tiene una posición en pixeles dado por x_{screen}, y_{screen}:

$$x_{screen} = f_x \cdot (\frac{X}{Z}) + c_x$$

$$y_{screen} = f_y \cdot (\frac{Y}{Z}) + c_y$$

Se introducen los puntos focales $f_x = F \cdot s_x$ y $f_y = F \cdot s_y$ ya que la mayoría de las camaras son rectangulares y no cuadrangulares, s_x y s_y corresponden a los tamaños físicos de los elementos que capturan la imágen. También se introducen los terminos c_x y c_y los cuales corresponden al desplazamiento del centro de las coordenadas al lugar en donde se está proyectando el punto Q.

Modelo de Cámara Pinhole

Usando notación matricial:

$$q = \begin{bmatrix} x \\ y \\ w \end{bmatrix}; \quad Q = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}; \quad M = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

Entonces:

$$q = M \cdot Q$$

Defectos en la captura de las imágenes

Dos tipos de distorsión aparecen en una gran cantidad de lentes, especialmente en lentes del tipo gran angular, que son la distorsión de Barril (Barrel) y la distorsión de Cojín (Pincushion).

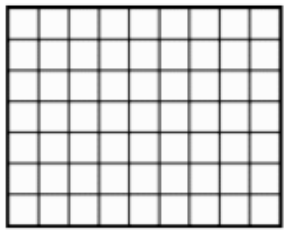


Figura 1.6: Imagen sin distorsión

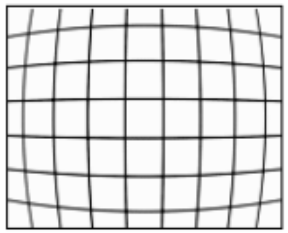


Figura 1.7: Imagen con distorsión de barril

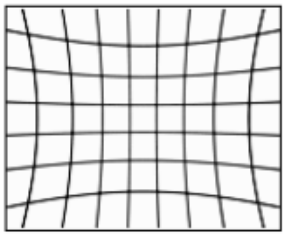


Figura 1.8: Imagen con distorsión de cojín

Además de los errores de los lentes se tienen los errores tangenciales generados por la forma en se ensamblaron los sensores de captura de las imágenes dentro de la cámara.

Defectos en la captura de las imágenes

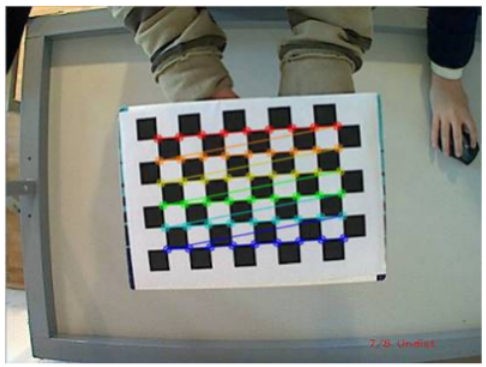
Los errores de los lentes y los errores tangenciales pueden ser corregidos cambiando la posición de los pixeles (x, y) a otra usando las siguientes expresiones:

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

$$\text{Distortion}_{\text{coefficients}} = (k_1 \quad k_2 \quad p_1 \quad p_2 \quad k_3)$$

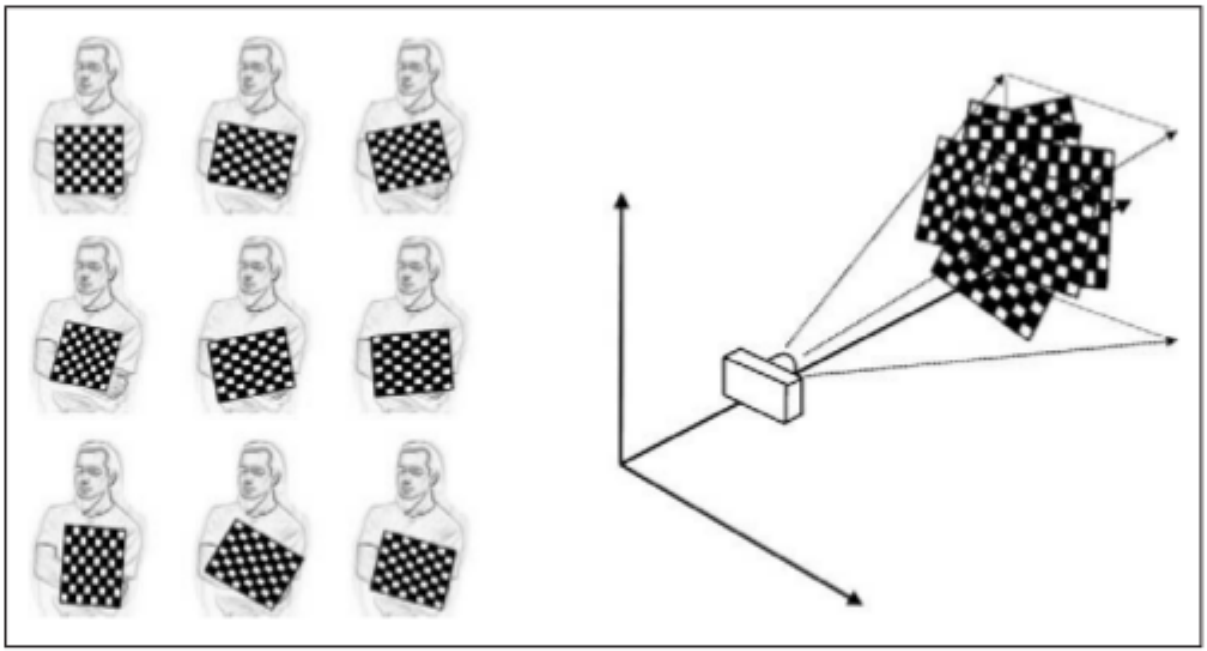
$$x_{\text{corrected}} = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)$$
$$y_{\text{corrected}} = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)$$

$$x_{\text{corrected}} = x + [2p_1 xy + p_2(r^2 + 2x^2)]$$
$$y_{\text{corrected}} = y + [p_1(r^2 + 2y^2) + 2p_2 xy]$$



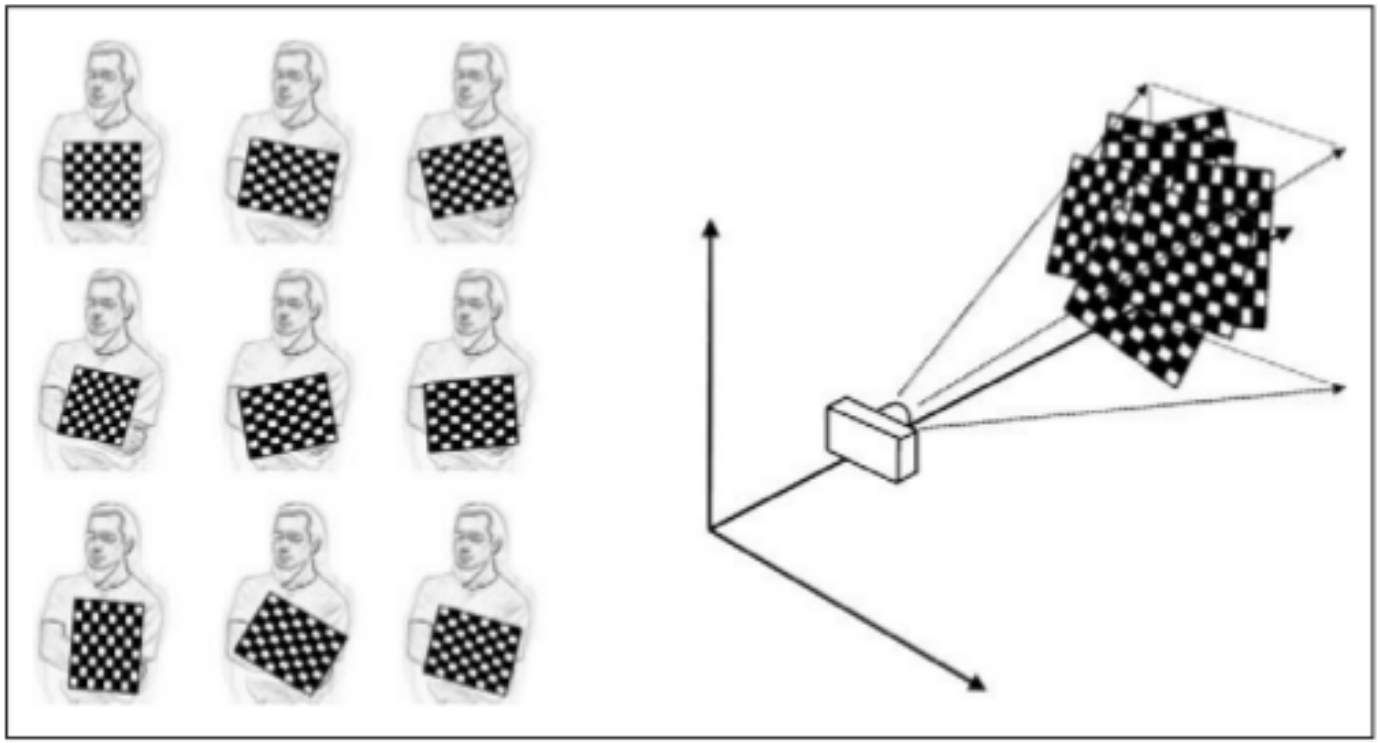
Calibración

Donde x y y es la posición de los pixeles en la imágen distorsionada y r es el radio ponderado que se utiliza para mover los pixeles a su posición corregida en la imágen nueva sin distorsión.
El objetivo de la calibración es encontrar los parámetros intrínsecos de las cámaras, así como los parámetros que eliminaran la distorsión en las imágenes.



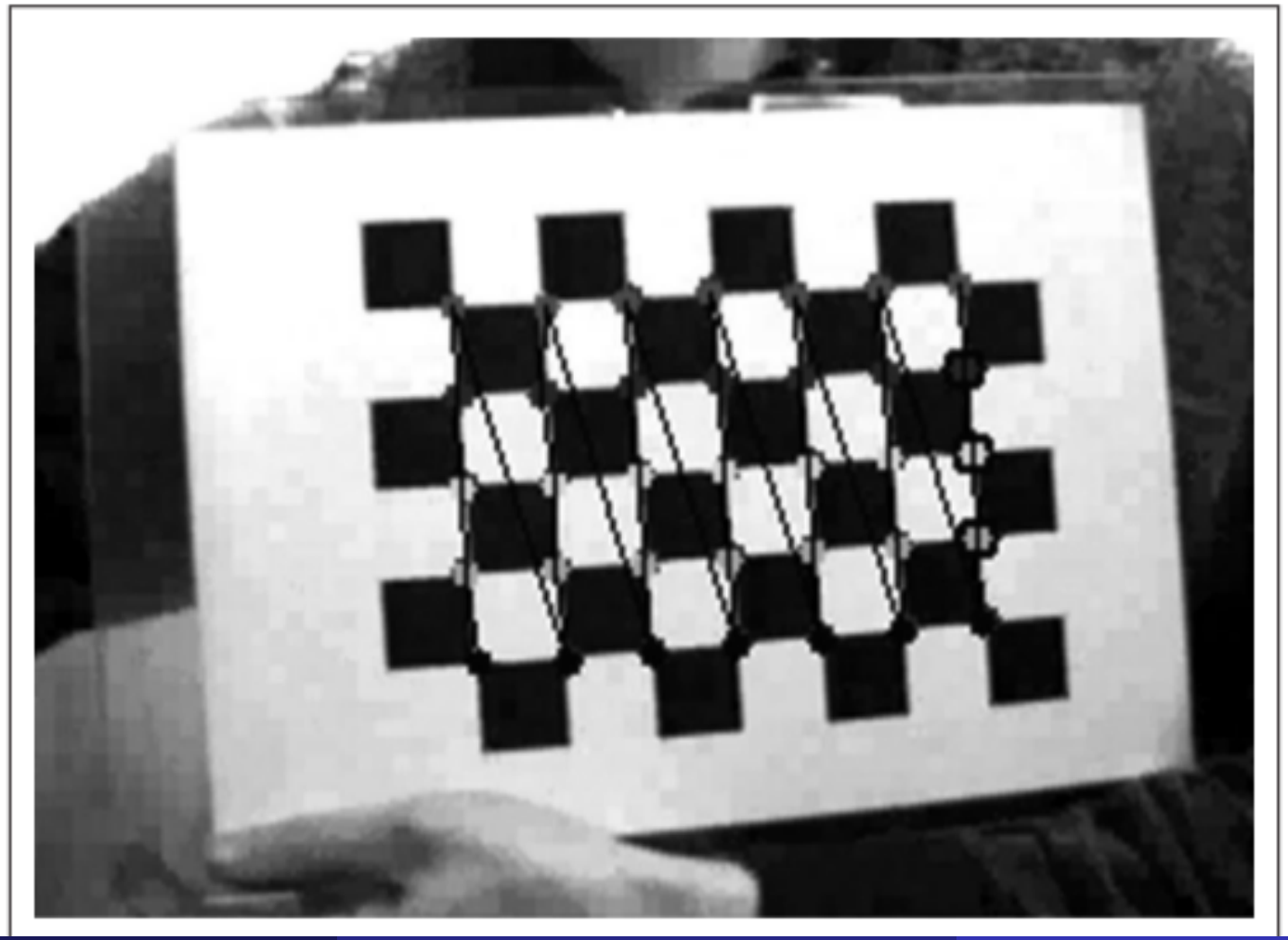
Calibración

Es decir, utilizando imágenes conocidas, un tablero de ajedrez mostrado en diferentes posiciones, se desea encontrar los parámetros intrínsecos de la cámara (f_x, f_y, c_x, c_y) y los parámetros de distorsión de los lentes (k_1, k_2, k_3) y los de la distorsión tangencial (p_1, p_2).



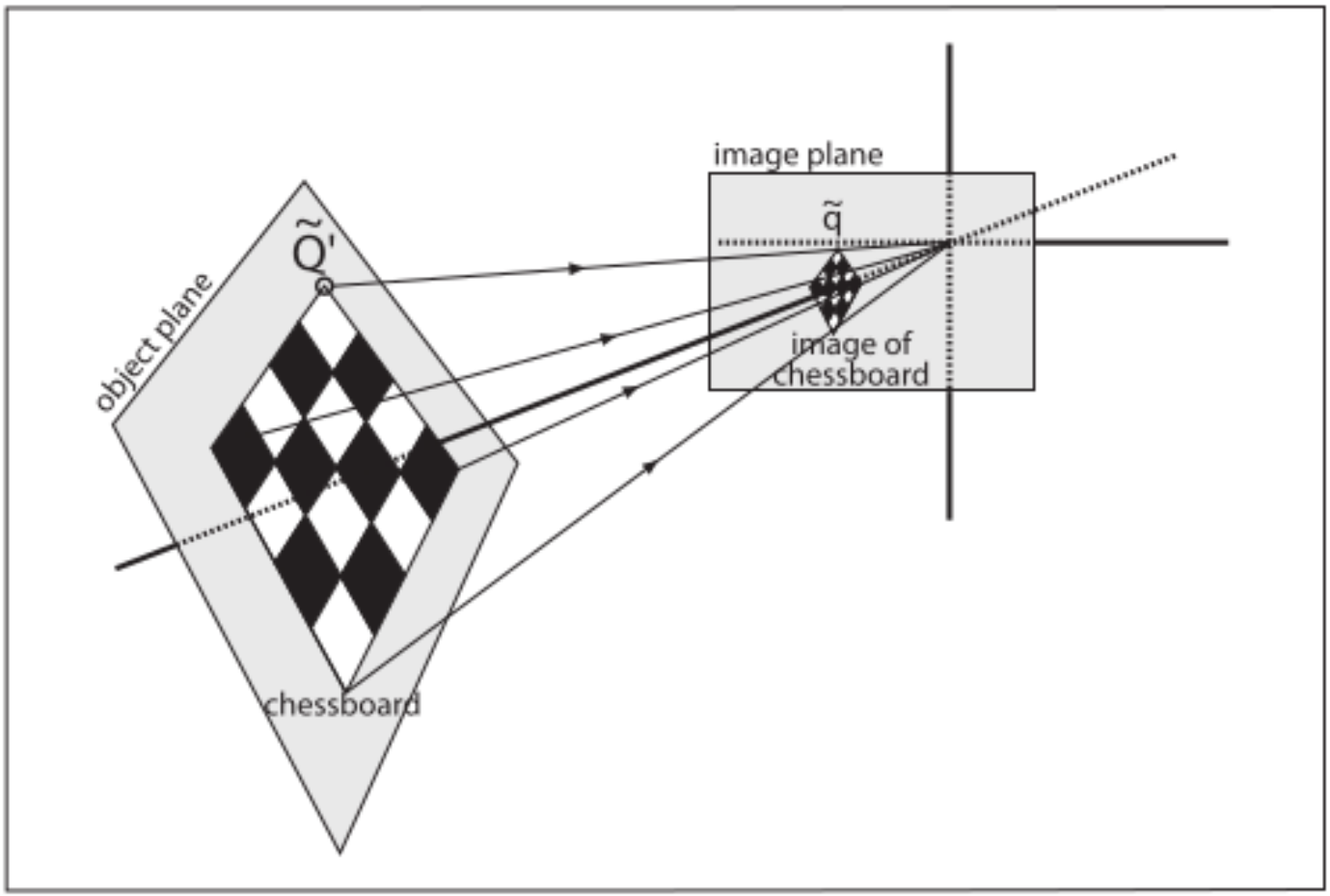
Calibración

Por cada imágen del tablero mostrada se encuentran las esquinas de cada uno de los cuadros del tablero.



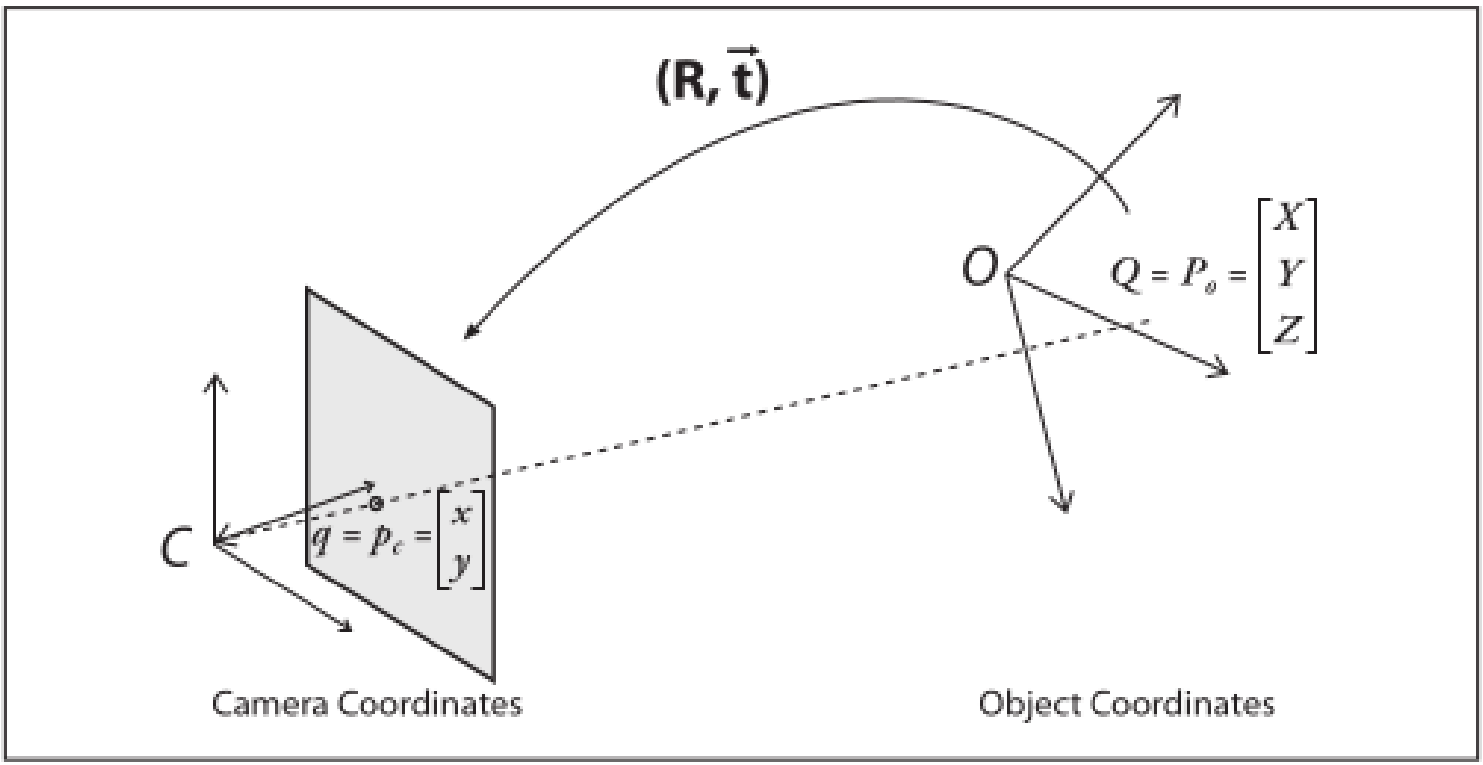
Calibración

Se encuentra también los limites de las esquinas superiores e inferiores.



Calibración

Se encuentran los angulos de giros ψ, α and θ de las imágenes capturadas con la imágen del tablero con su posición original.



Calibración

Se rotan las imágenes capturadas para que tengan la misma orientación que la original y en forma iterativa se van modificando los parámetros de distorsión de los lentes (k_1, k_2, k_3) y los de la distorsión tangencial (p_1, p_2) hasta que las imágenes capturadas tengan la misma forma que la original. Despues se encuentran f_x y f_y de las siguientes ecuaciones:

$$x_{screen} = f_x \cdot (\frac{X}{Z}) + c_x$$

$$y_{screen} = f_y \cdot (\frac{Y}{Z}) + c_y$$

¿Qué es una imagen?

Una imagen es un registro de valores organizados en forma bidimensional, generalmente representando intensidad de radiación electromagnética. Tiene las siguientes características:

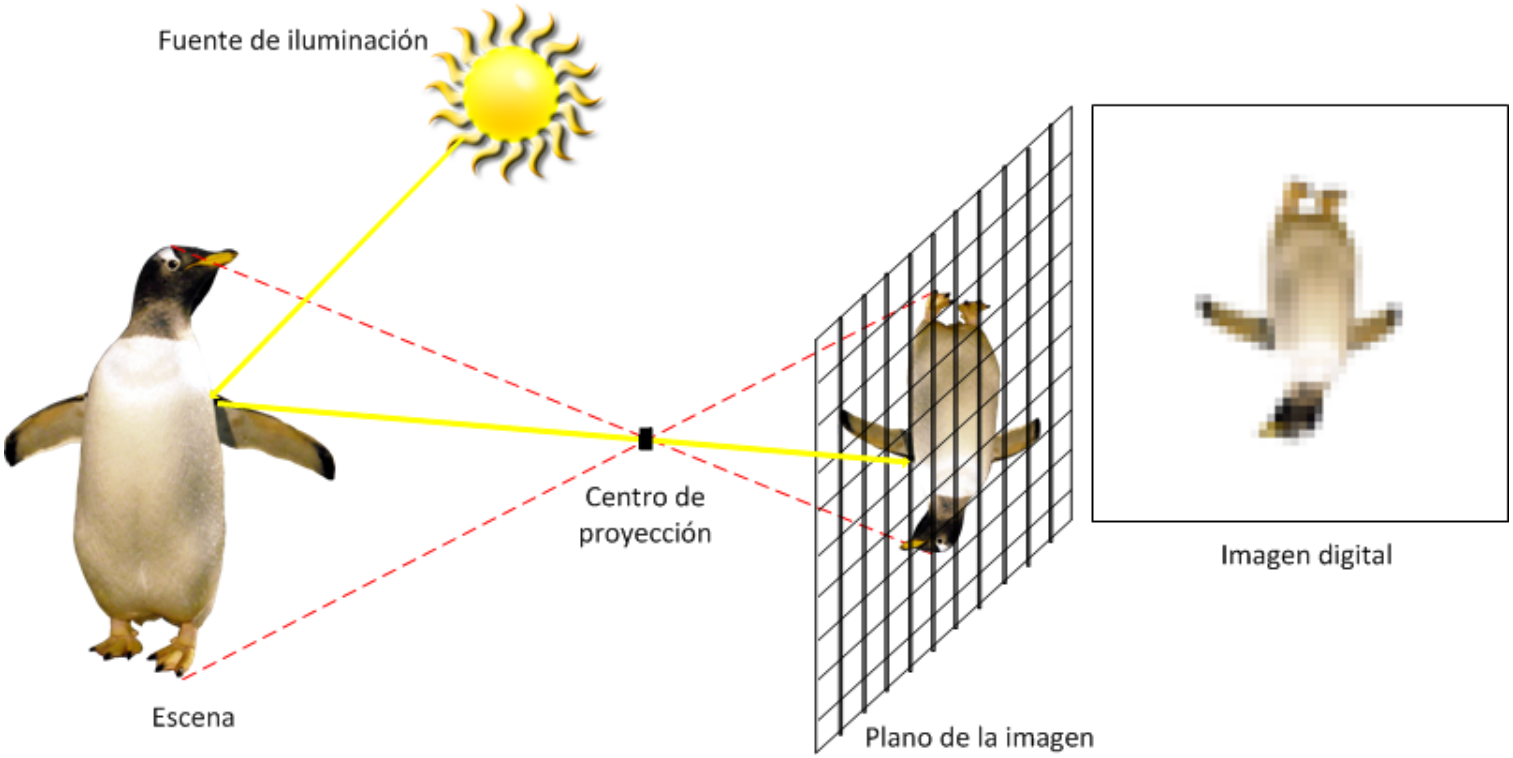
1. Muestreo espacial:
2. Cuantización numérica: e.g. Negro $= 0$, Blanco $= 255$
3. Cuantización espectral:

Sensores

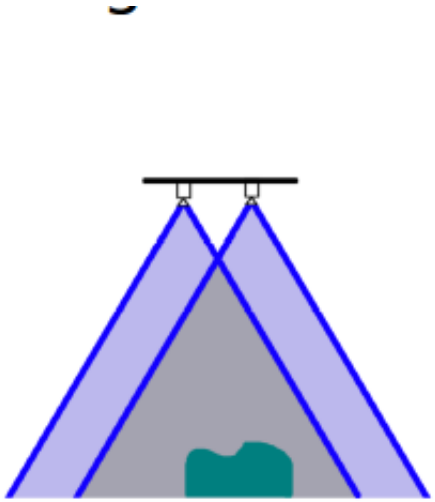
1. Ojo humano
2. Cámaras Digitales: proyección 2-D de la escena
3. Cámaras RGB-D

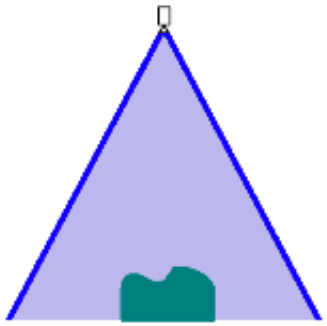
Cámara Digital

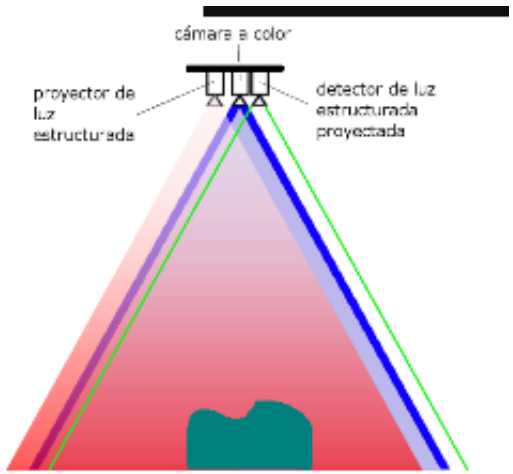
CCD (charge-coupled device),
CMOS (complementary metal-oxide-semiconductor)



Cámaras RGB-D







Estéreo

ToF

Luz estructurada