



Universidad  
Rey Juan Carlos

Escuela de Ingeniería  
de Fuenlabrada



# Visión Artificial

## 7. Geometría y calibración

JOSÉ MIGUEL GUERRERO HERNÁNDEZ

E MAIL: [JOSEMIGUEL.GUERRERO@URJC.ES](mailto:JOSEMIGUEL.GUERRERO@URJC.ES)

# Índice de contenidos

---

1. Introducción
2. Geometría del modelo pinhole
3. Proyecciones
4. Parámetros intrínsecos
5. Parámetros extrínsecos
6. Matriz de proyección
7. Calibración

# Índice de contenidos

---

1. Introducción
2. Geometría del modelo pinhole
3. Proyecciones
4. Parámetros intrínsecos
5. Parámetros extrínsecos
6. Matriz de proyección
7. Calibración

# 1. Introducción

---

- Calibrar una cámara no es más que averiguar de qué forma podemos “corregir” lo que percibe el sensor para que coincida con la realidad
- Todos los sensores (aceleración, láser, GPS, micrófonos...) introducen “ruido”
- Lo que ellos captan es una versión distorsionada de la realidad
- La razón de por qué los sensores no reproducen el mundo real tiene origen en las restricciones de la tecnología, de los materiales con que se fabrica... Este tipo de ruido está relacionado con los **parámetros intrínsecos**

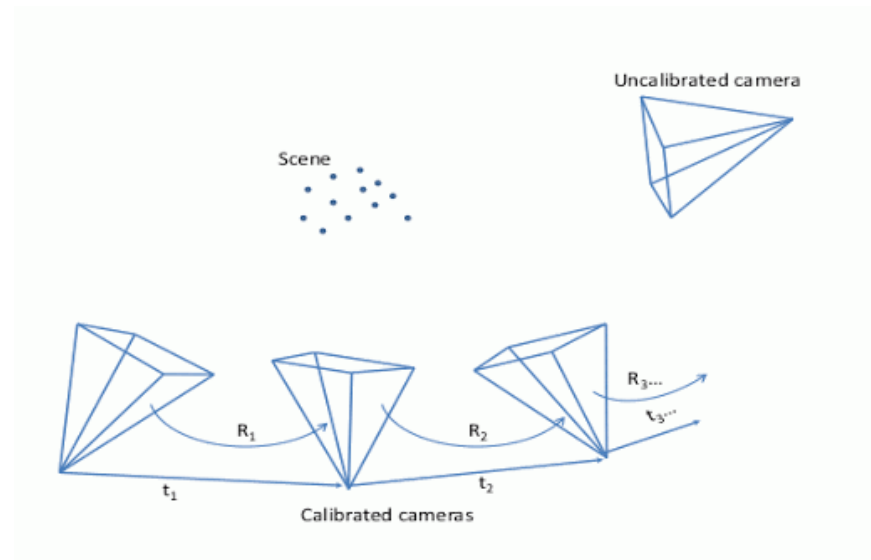
# 1. Introducción

---

- Antes hemos dicho “hay que averiguar de qué forma podemos corregir lo que percibe el sensor para que coincida con la realidad”
- Pero lo cierto es que hay múltiples realidades:
  - Lo que para mis ojos está a 3 metros, para los ojos del compañero está a 5 metros
- Cada uno expresa la realidad en función de su propio sistema de referencia

# 1. Introducción

- Calibrar un par de cámaras (o de cualquier otro sensor) involucra saber la transformación que hay entre sus dos sistemas de coordenadas, para que “un poco más arriba” signifique lo mismo para los dos observadores
- Esta transformación viene definida por los **parámetros extrínsecos**: aquellos que dependen de la posición del sistema de visión y de cada escenario



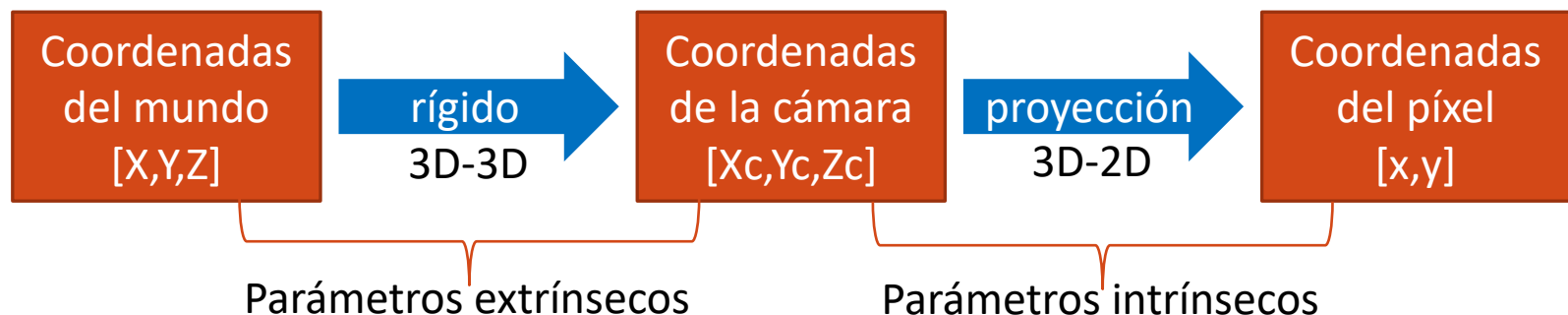
# 1. Introducción

---

- Como ya vimos al inicio del curso, hay dos tipos de parámetros que definen la calibración de la cámara:
  - **Parámetros intrínsecos:** definen la **geometría interna y la óptica de la cámara**. Éstos determinan cómo la cámara proyecta los puntos del mundo 3D al plano de la imagen en 2D, siendo **constantes** en tanto no varíen las características y posiciones relativas entre la óptica y el sensor imagen. **Los provee el fabricante de la cámara**, o bien se pueden estimar
  - **Parámetros extrínsecos:** Éstos **relacionan** los sistemas de referencia del **mundo real y la cámara**, describiendo la posición y orientación de la cámara en el sistema de coordenadas del mundo real

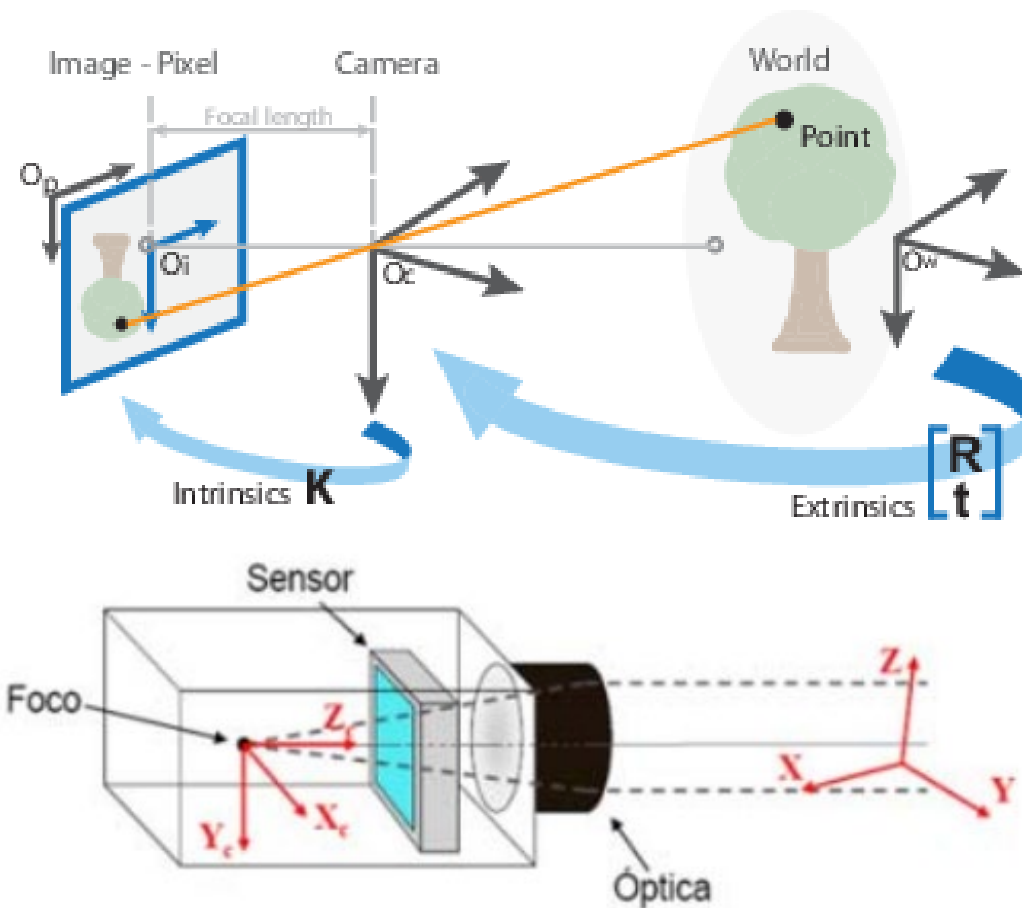
# 1. Introducción

- La relación entre estos parámetros es lo que nos sirve para pasar de las coordenadas 3D del mundo real, a las coordenadas 2D que serán los píxeles de nuestra imagen
  - Los parámetros extrínsecos representan una transformación rígida del sistema de coordenadas 3D del mundo real al sistema de coordenadas 3D de la cámara
  - Los parámetros intrínsecos representan la transformación del sistema de coordenadas 3D de la cámara al sistema de coordenadas 2D de la imagen



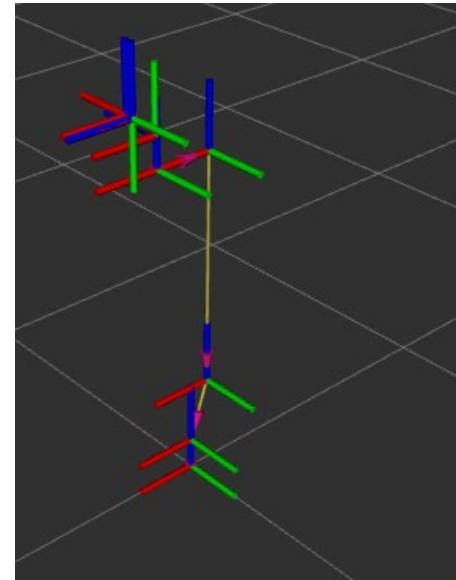
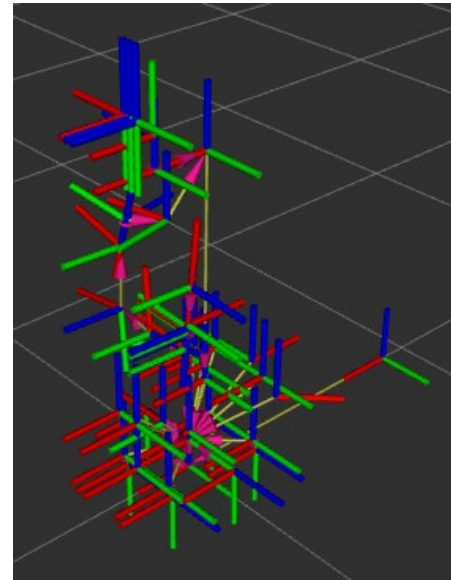
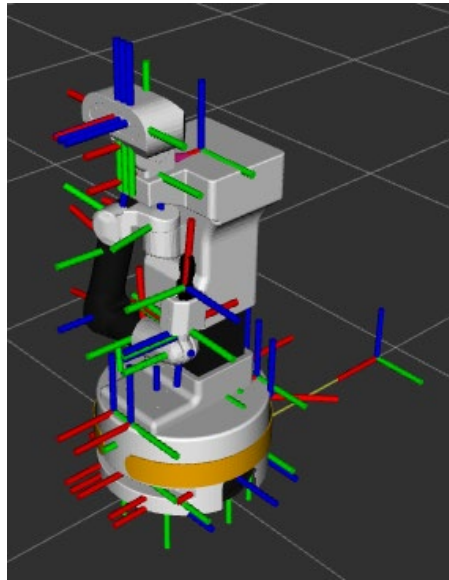
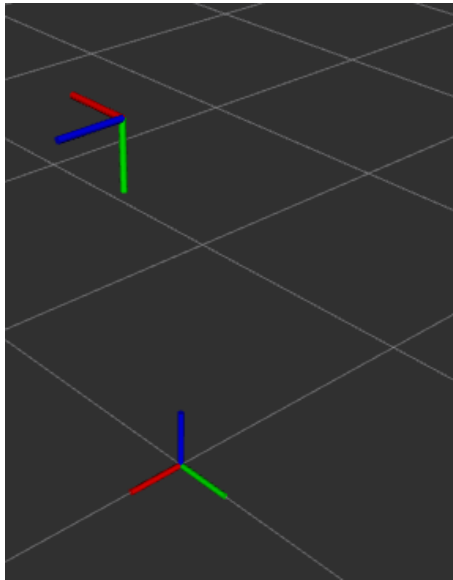


# 1. Introducción

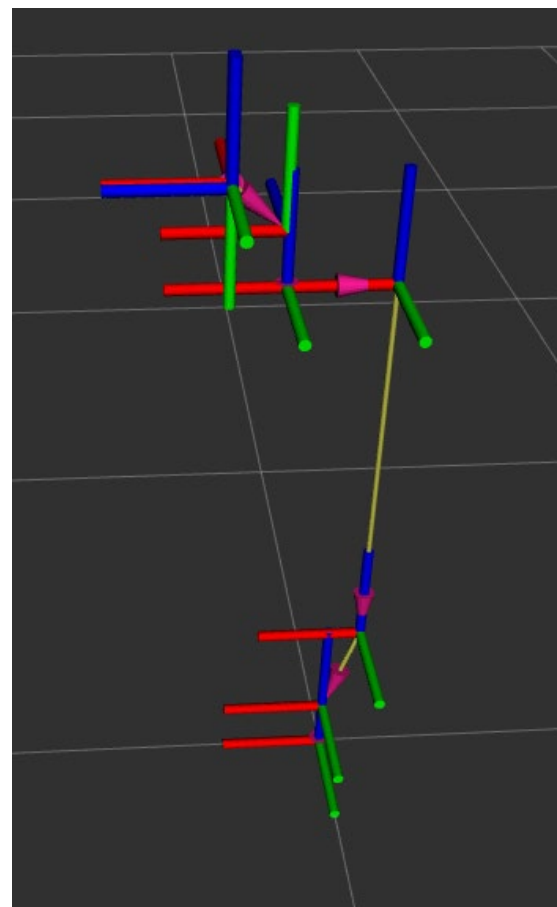
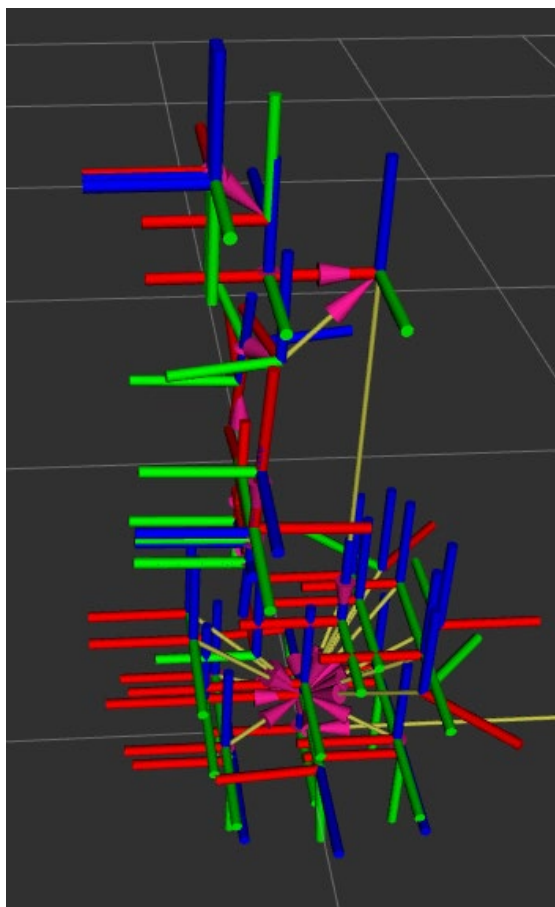
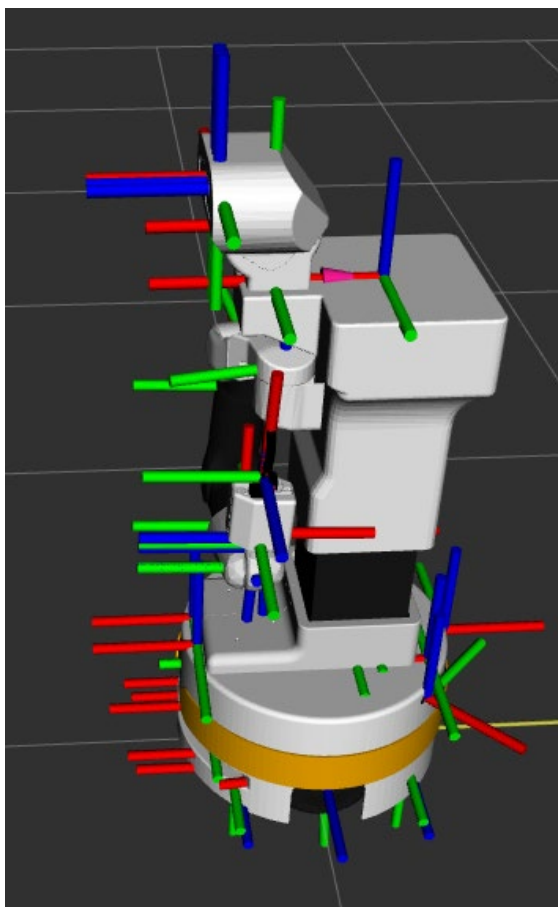


# 1. Introducción

---



# 1. Introducción



# 1. Introducción

- ¿Y para qué sirve? → Corresponder sensores



# 1. Introducción

- ¿Y para qué sirve? → Eliminar distorsión
  - Las cámaras deportivas tienen una lente de tipo “ojo de pez” lo cual les permite mayor ángulo de visión a costa de una mayor distorsión de la realidad



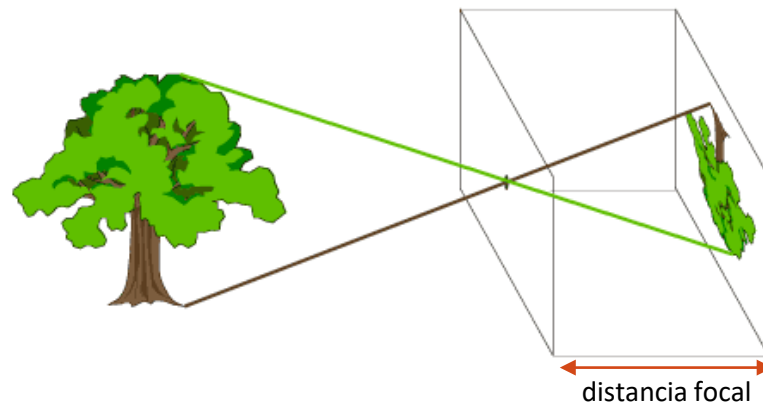
# Índice de contenidos

---

1. Introducción
2. Geometría del modelo pinhole
3. Proyecciones
4. Parámetros intrínsecos
5. Parámetros extrínsecos
6. Matriz de proyección
7. Calibración

## 2. Modelo pinhole

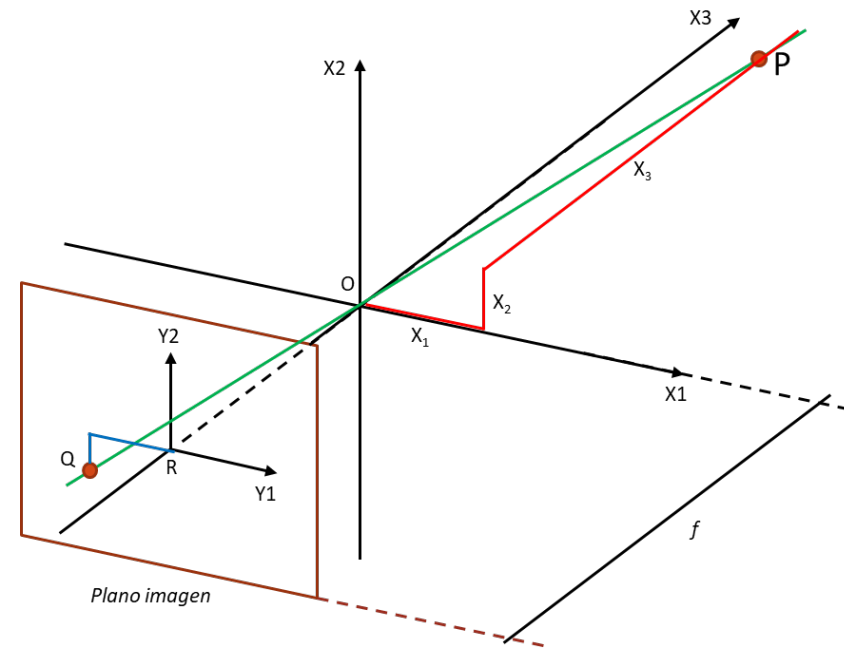
- El modelo pinhole describe una cámara (sin lente) y con una pequeña apertura. Los rayos de luz atraviesan la apertura y proyectan una versión invertida de lo que se capta en el fondo de la cámara (o el plano de la imagen)



- La distancia entre la apertura y el plano de la imagen se conoce como distancia focal

## 2. Modelo pinhole

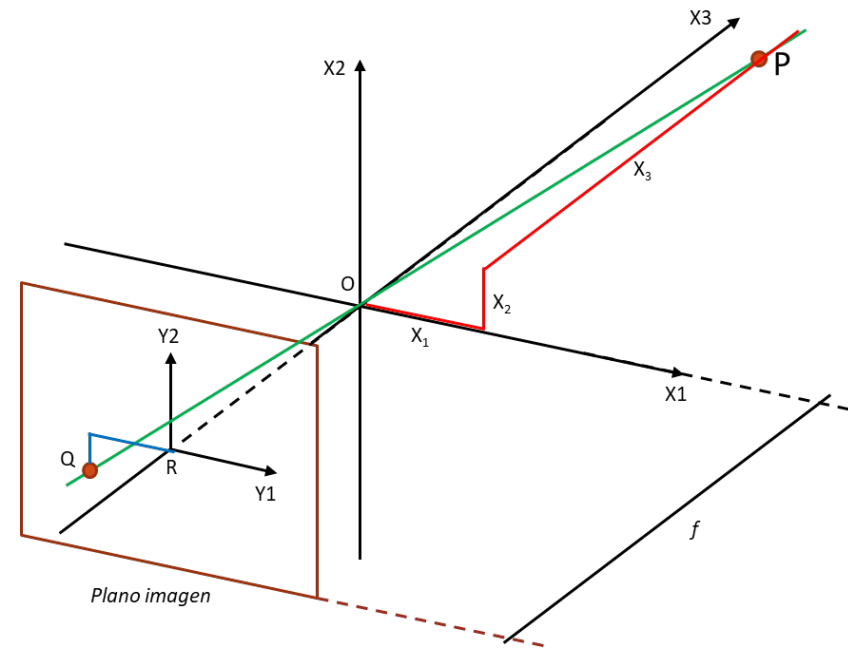
- **P** es un punto en el espacio 3D, el cual se va a proyectar en la imagen
- **X1X2X3** definen el sistema de coordenadas de la cámara, el cual tiene su origen en **O**, que es justo donde está la apertura de la cámara
- **X3** coincide con la dirección de la cámara
- El plano de la imagen es paralelo a los ejes **X1X2** y está desplazado a una distancia  $f$  en la dirección negativa del eje **X3**





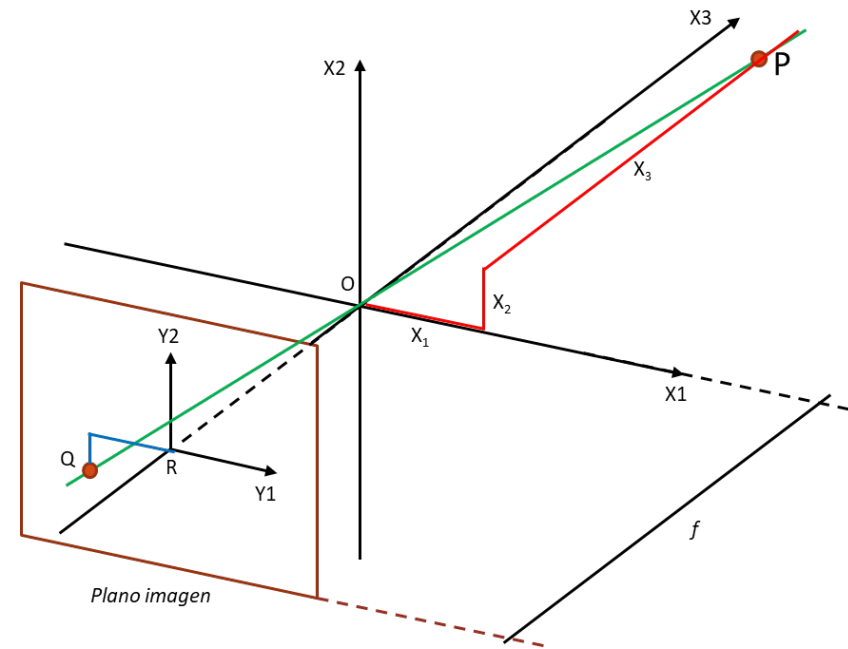
## 2. Modelo pinhole

- **R** es el punto en que intersecan el eje  $-X_3$  y el plano de la imagen, que está desplazado  $f$  unidades. **R** es también el centro óptico de la imagen
- **Q** es la intersección de la recta que parte de **P** y pasa por **O** en el plano de la imagen. Es decir, la proyección del punto 3D en el plano de la imagen



## 2. Modelo pinhole

- De esta forma, todos los puntos 3D del mundo real son proyectados en el plano de la imagen con coordenadas 2D  $Y_1Y_2$
- Nota que en la recta **PQ** hay un obstáculo, el cual se materializa en la imagen, pero no sabemos a qué profundidad se encuentra



# Índice de contenidos

---

1. Introducción
2. Geometría del modelo pinhole
- 3. Proyecciones**
4. Parámetros intrínsecos
5. Parámetros extrínsecos
6. Matriz de proyección
7. Calibración

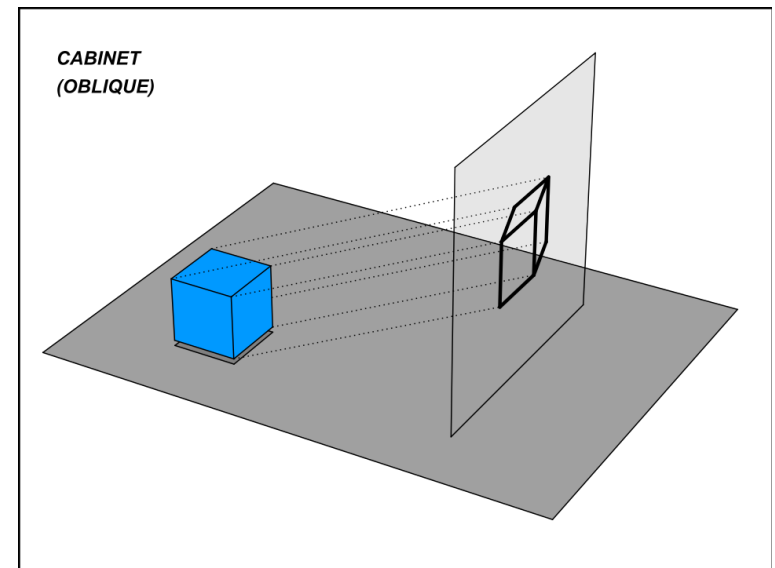
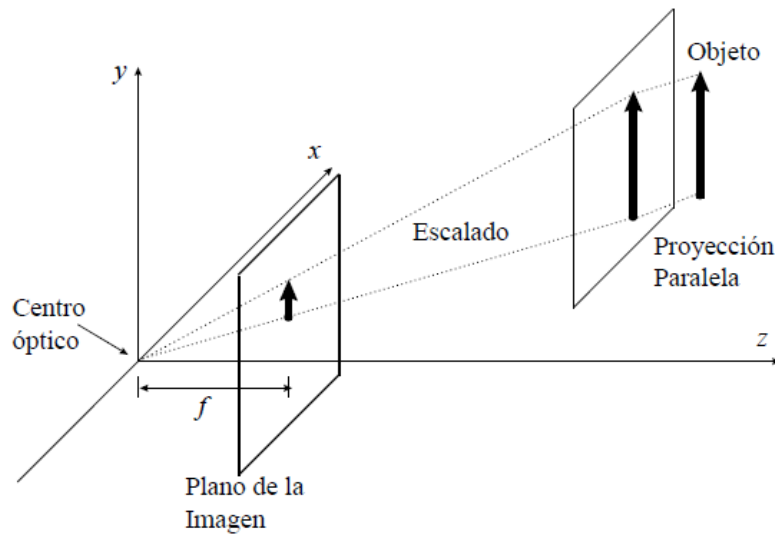
# 3. Proyecciones

---

- Hay dos tipos de proyecciones:
  - **Proyección paralela:**
    - El área de visión tiene forma paralelepípedica (un cubo o un prisma rectangular)
    - Todos los objetos dentro del volumen se proyectan sin distorsión de tamaño
    - Útil para la calibración de cámaras y medición de objetos
  - **Proyección en perspectiva:**
    - El área de visión tiene forma piramidal
    - Los objetos más lejanos aparecen más pequeños debido a la convergencia de las líneas de proyección
    - Reconstrucción 3D y detección de profundidad con múltiples vistas o cámaras estereoscópicas

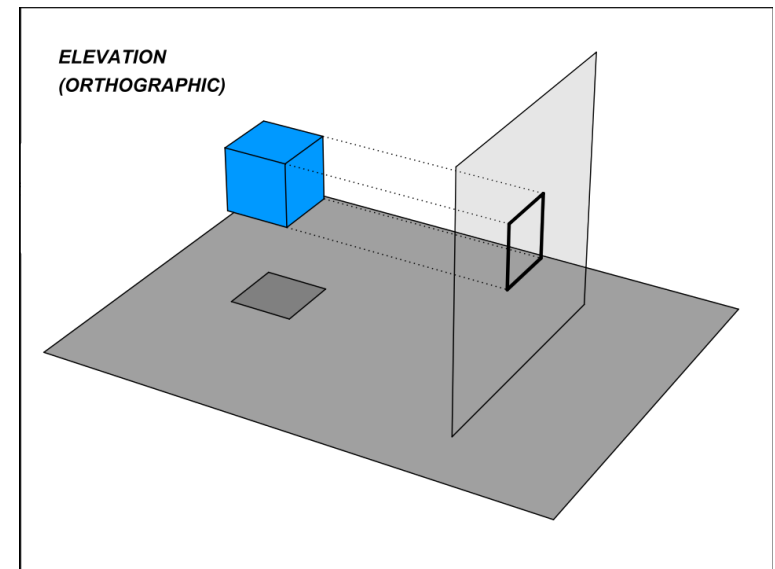
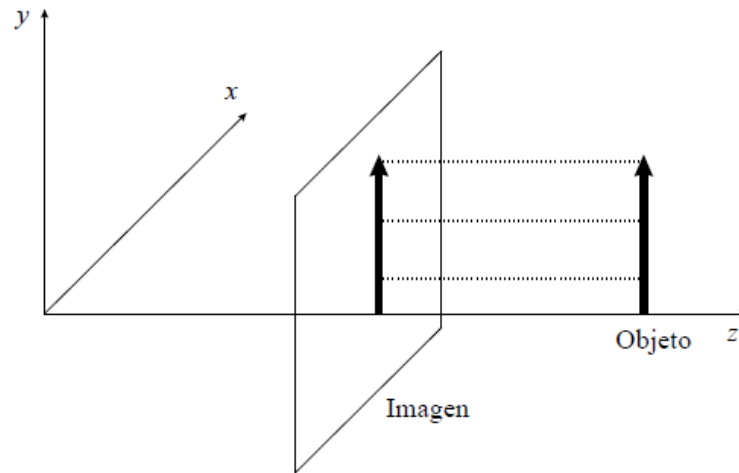
# 3. Proyecciones

- Proyección paralela (oblicua):
  - Los rayos son paralelos entre sí, pero no perpendiculares al plano
  - Da sensación de profundidad sin distorsión por perspectiva



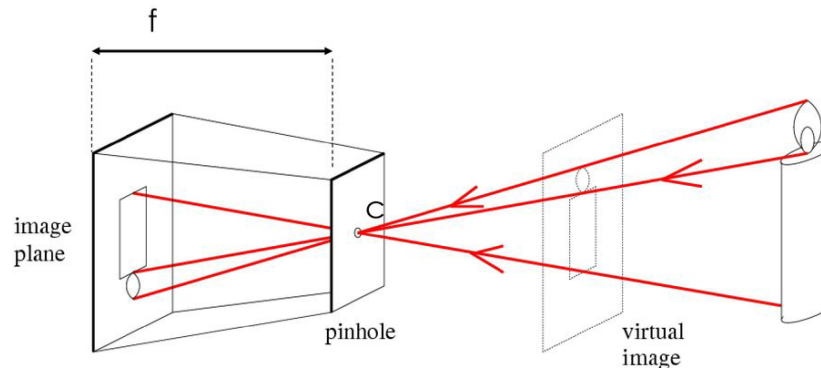
# 3. Proyecciones

- Proyección paralela (ortográfica):
  - Los rayos son perpendiculares al plano de proyección
  - Mantiene las proporciones exactas de los objetos



# 3. Proyecciones

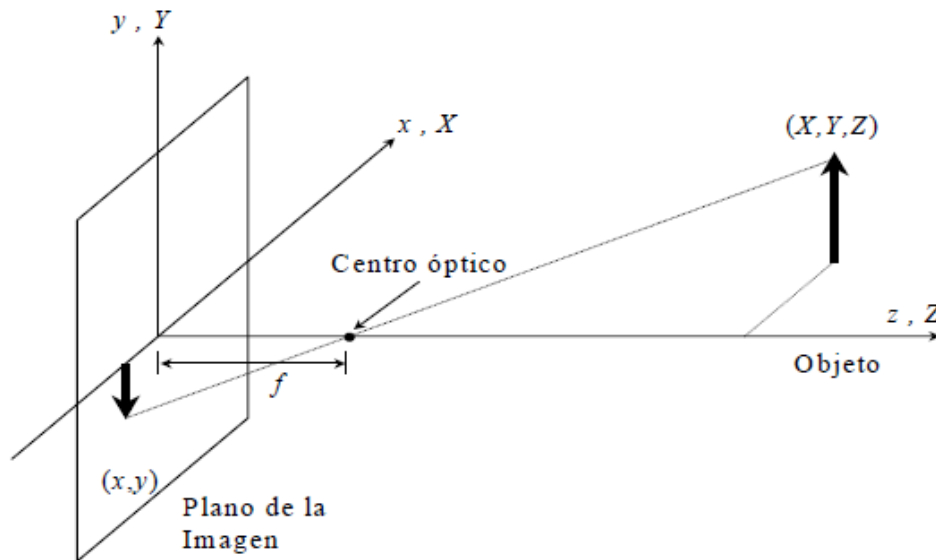
- Proyección en perspectiva:
  - Los rayos pasan por un punto llamado centro óptico (también conocido como centro de proyección o pinhole en el modelo de cámara estenopeica)
  - No pertenece al plano de imagen, sino que está en el espacio 3D, justo delante o detrás del plano dependiendo del modelo



$f$  = focal length  
 $c$  = center of the camera

# 3. Proyecciones

- Proyección en perspectiva (plano detrás):
  - La imagen capturada es una proyección invertida respecto a la escena real
  - Los objetos cercanos se muestran más pequeños y los lejanos más grandes



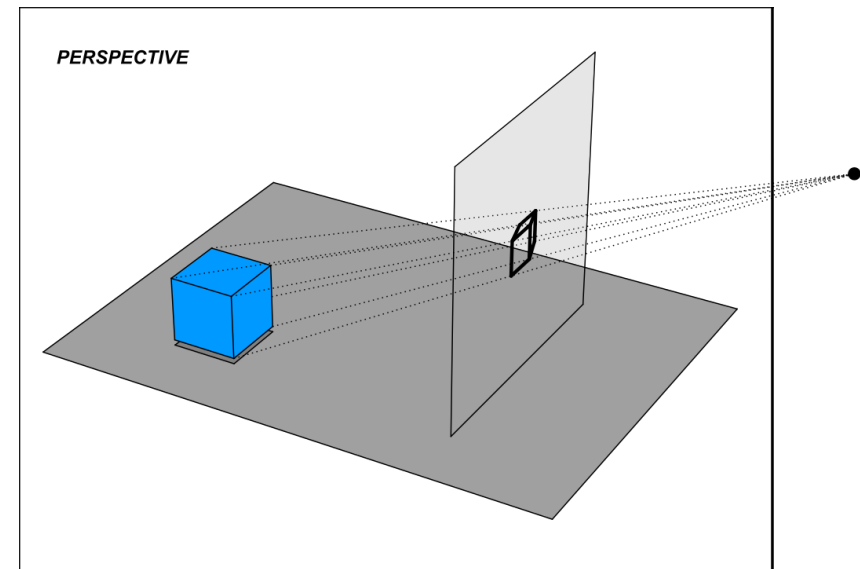
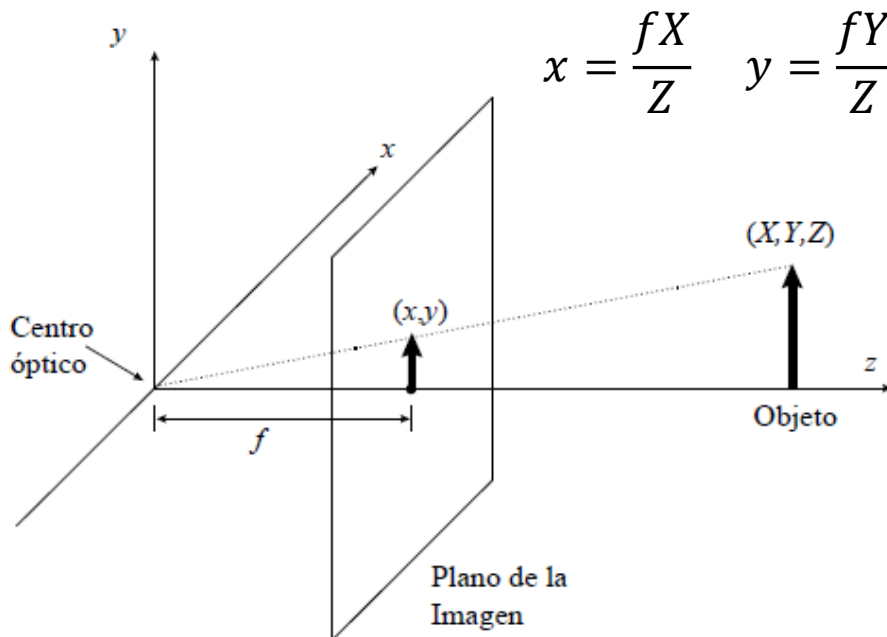
$$x = -\frac{fX}{Z}$$

$$y = -\frac{fY}{Z}$$



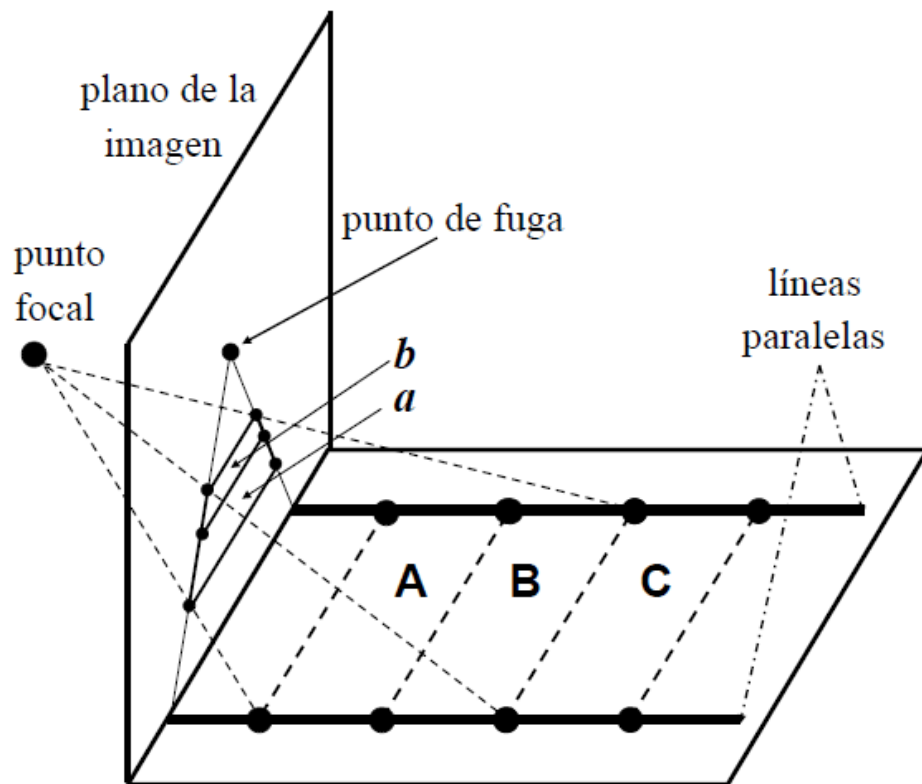
# 3. Proyecciones

- Proyección en perspectiva (plano delante):
  - La imagen capturada es una proyección directa (o normal) de la escena real
  - Los objetos cercanos se muestran más grandes y los lejanos más pequeños



# 3. Proyecciones

- Proyección en perspectiva, ejemplo:



# Índice de contenidos

---

1. Introducción
2. Geometría del modelo pinhole
3. Proyecciones
- 4. Parámetros intrínsecos**
5. Parámetros extrínsecos
6. Matriz de proyección
7. Calibración

## 4. Parámetros intrínsecos

- En el ejemplo del modelo pinhole, encontramos que los parámetros intrínsecos que nos permiten proyectar puntos 3D del mundo real a puntos 2D en el plano de la imagen son los siguientes:

- Distancia focal  $F$  ( $f_x = F \cdot k$ ,  $f_y = F \cdot l$ ;  $k$  y  $l$  son el ancho y alto de un pixel)
- Centro de la imagen  $c_x$  y  $c_y$

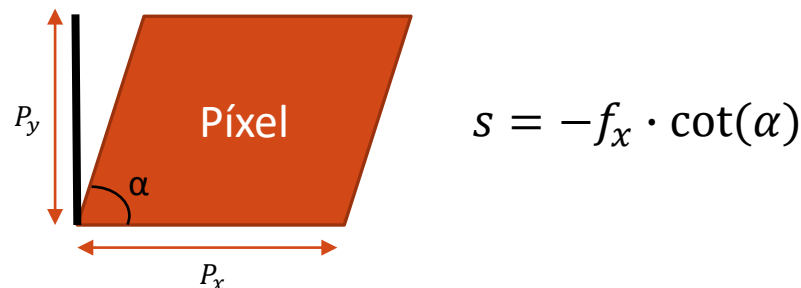
- Formalmente, estos valores vienen representados por una matriz  $K$ :

$$K = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

- En la literatura podemos encontrar que al centro de la imagen  $C$  también se le llama  $R$

## 4. Parámetros intrínsecos

- En la matriz  $K$  hay un valor  $s$  que hace referencia al factor de “skew”. Este valor está relacionado con el ángulo que toma el eje Y del sistema de coordenadas del plano de la imagen cuando los ejes no son perpendiculares
- En los modelos estándar los ejes siempre son perpendiculares por lo que este valor (casi) siempre nos lo encontraremos puesto a 0



## 4. Parámetros intrínsecos

- La relación que existe entre los **XYZ** puntos 3D del mundo y los **xy** puntos 2D de la imagen es la siguiente:

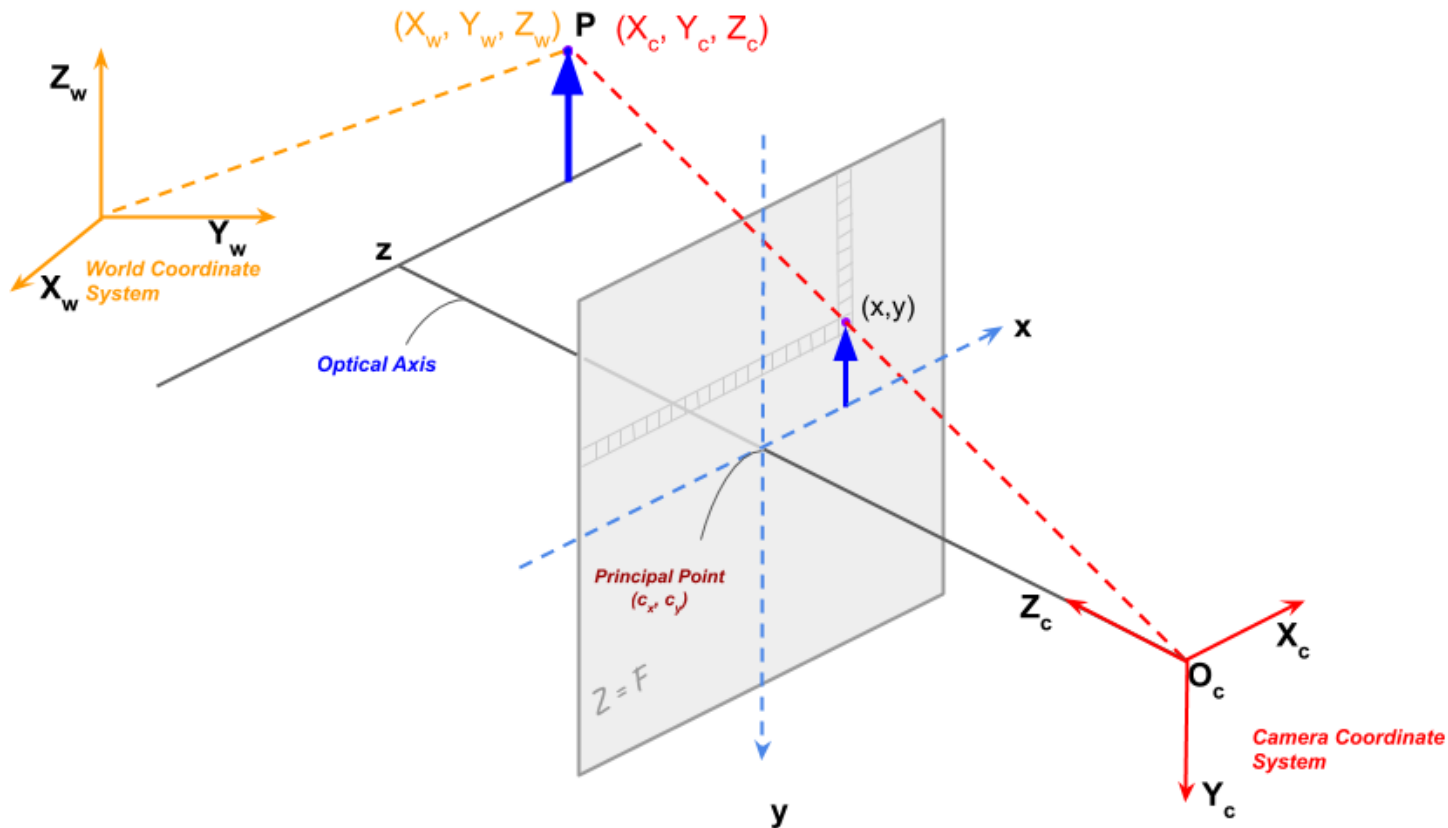
$$x = f_x \frac{X}{Z} - f_x \cot \alpha \frac{Y}{Z} + c_x \quad y = \frac{f_y}{\sin \alpha} \frac{Y}{Z} + c_y$$

- Por comodidad, los parámetros se resumen en la matriz **K** que se expresa en forma matricial, de forma que se puede operar muy fácilmente

$$[x \ y \ 1]^T = K[X \ Y \ Z]^T$$

- De esta forma podemos calcular la posición **xy** en la imagen para cualquier punto **XYZ** del mundo real

# 4. Parámetros intrínsecos



## 4. Parámetros intrínsecos

- Estos parámetros también nos permiten conocer la relación que existe entre los ***xy*** puntos 2D de la imagen y los ***XYZ*** puntos 3D del mundo

$$X = \frac{(x - c_x) * d}{f_x} \quad Y = \frac{(y - c_y) * d}{f_y} \quad Z = d$$

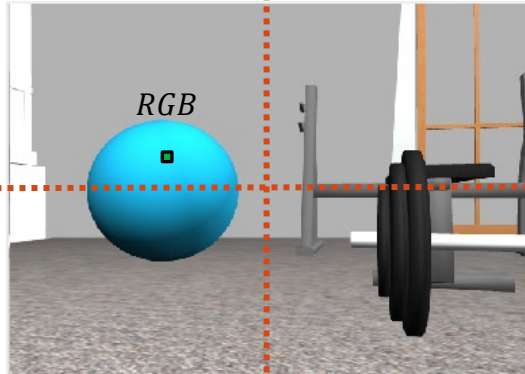
donde  $d$  es el valor de profundidad (*depth*) que debe conocerse a priori. Normalmente utilizando la imagen de profundidad, láser, o imagen de disparidad calculada por un sistema estéreo (lo veremos más adelante)

- De esta forma podemos calcular la posición ***XYZ*** del mundo real para cualquier punto ***xy*** de la imagen

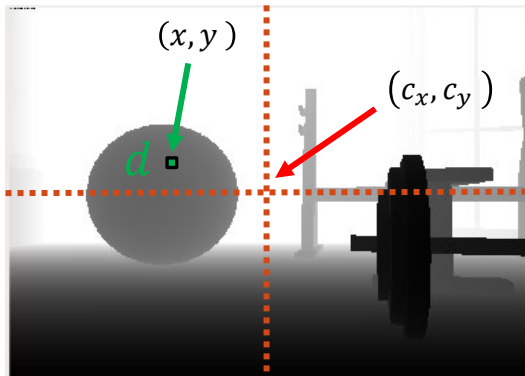


# 4. Parámetros intrínsecos

original

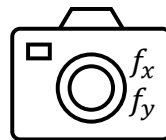
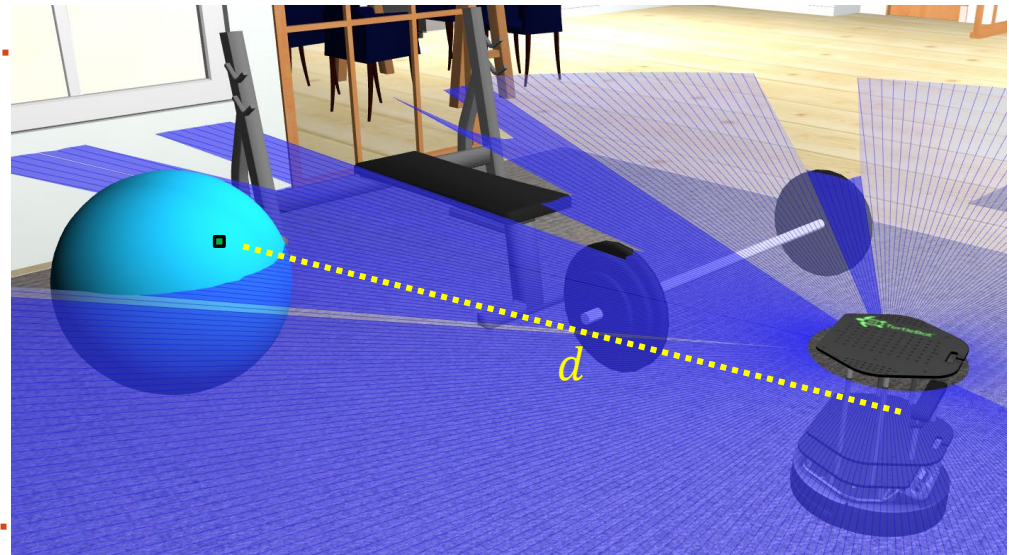


profundidad



Conversión 2D a 3D

$$(X, Y, Z) = \left( \frac{(x - c_x) * d}{f_x}, \frac{(y - c_y) * d}{f_y}, d \right)$$

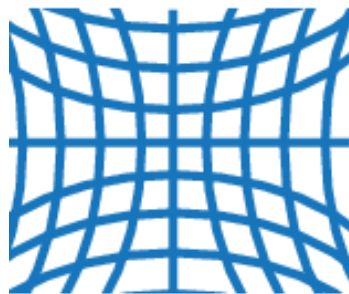


Conversión 3D a 2D

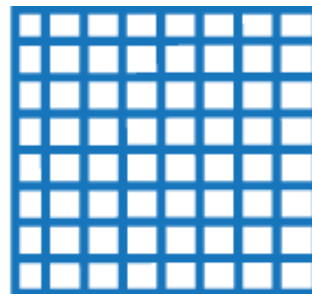
$$(x, y) = \left( f_x \frac{X}{Z}, f_y \frac{Y}{Z} \right)$$

## 4. Parámetros intrínsecos

- Como se ha indicado anteriormente, el modelo pinhole no contempla el uso de una lente. Sin embargo, las cámaras que se usan en la actualidad la integran
- Las lentes proporcionan una serie de ventajas como por ejemplo profundidad de campo definida (en contraste con el modelo pinhole puro que muestra una profundidad de campo infinito). También introduce desventajas como por ejemplo la distorsión radial
- Este tipo de distorsión no está contemplada en la matriz  $K$  de parámetros intrínsecos de la cámara, pero igualmente debemos tenerlo en cuenta



Negative radial distortion  
"pincushion"



No distortion



Positive radial distortion  
"barrel"

## 4. Parámetros intrínsecos

- Como último detalle sobre los parámetros intrínsecos, decir que éstos son fijos siempre y cuando no se modifiquen las características de la cámara
- ¿Qué características podrían hacer cambiar los parámetros intrínsecos?
  - Por ejemplo, cambiar el grupo óptico (el “objetivo”) de la cámara o modificar sus parámetros cambiando la apertura



**No tocar una vez calibrado**

# Índice de contenidos

---

1. Introducción
2. Geometría del modelo pinhole
3. Proyecciones
4. Parámetros intrínsecos
- 5. Parámetros extrínsecos**
6. Matriz de proyección
7. Calibración

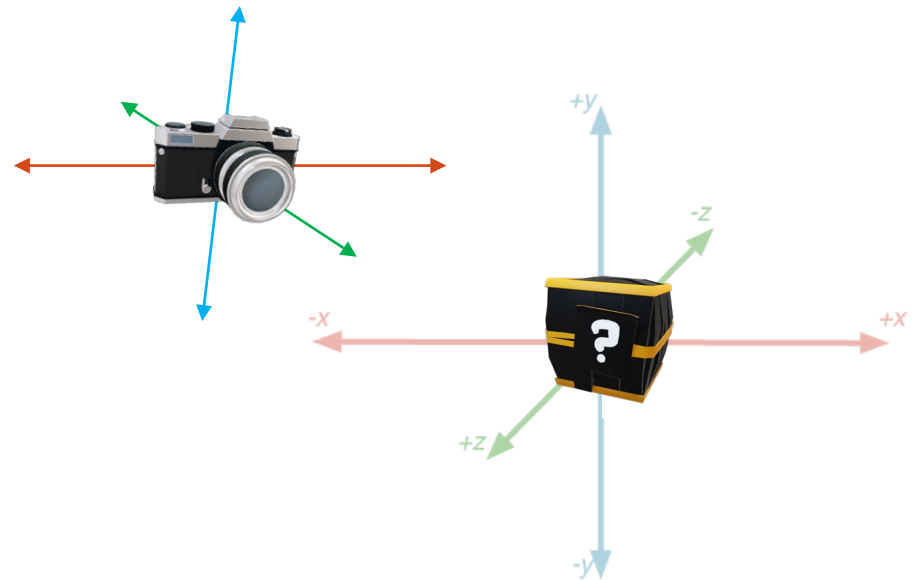
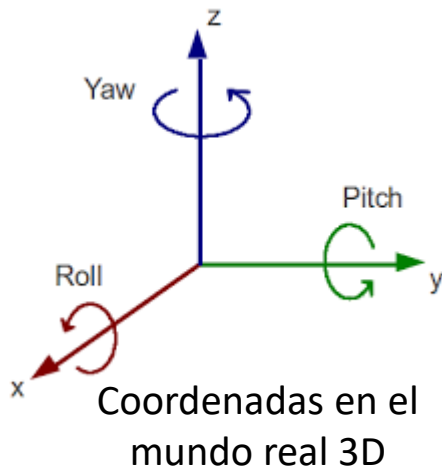
# 5. Parámetros extrínsecos

---

- Los parámetros extrínsecos definen la **posición** de la cámara en el sistema de coordenadas del **mundo real**
- Junto con los parámetros intrínsecos, permiten averiguar la posición 3D en el sistema de coordenadas del mundo de cierto punto 2D en el sistema de coordenadas de la imagen
- A diferencia de los intrínsecos, **estos parámetros no son fijos** y dependen de la pose de la cámara con respecto del mundo. Debido a esto, se tienen que estimar con algún método
- Estos parámetros **alinean** el sistema de coordenadas de los puntos de la escena tridimensional con respecto al sistema de coordenadas de la cámara

# 5. Parámetros extrínsecos

- Los parámetros extrínsecos vienen definidos por dos matrices  $R$  y  $t$ :
  - $R$ : rotación de los 3 ejes entre los sistemas de coordenadas
  - $t$ : desplazamiento en  $x, y, z$  entre los orígenes de los sistemas de coordenadas



# 5. Parámetros extrínsecos

- La matriz de rotación  $R$  está compuesta por la acumulación de la rotación independiente de cada eje para hacer cuadrar ambos sistemas de coordenadas

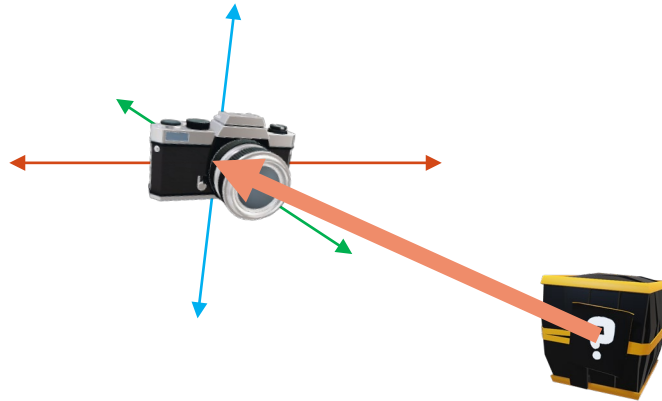
$$\begin{array}{ccc}
 \text{roll} & \text{pitch} & \text{yaw} \\
 R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix} & R_y(\theta) = \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} & R_z(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}
 \end{array}$$

- Como  $R = R_z(\Psi) \cdot R_y(\theta) \cdot R_x(\phi)$  es común encontrárselo de la siguiente manera directamente:

$$[R_0] = \begin{bmatrix} \cos(\psi)\cos(\theta) & \cos(\psi)\sin(\theta)\sin(\phi) - \sin(\psi)\cos(\phi) & \cos(\psi)\sin(\theta)\cos(\phi) + \sin(\psi)\sin(\phi) \\ \sin(\psi)\cos(\theta) & \sin(\psi)\sin(\theta)\sin(\phi) + \cos(\psi)\cos(\phi) & \sin(\psi)\sin(\theta)\cos(\phi) - \cos(\psi)\sin(\phi) \\ -\sin(\theta) & \cos(\theta)\sin(\phi) & \cos(\theta)\cos(\phi) \end{bmatrix}$$

# 5. Parámetros extrínsecos

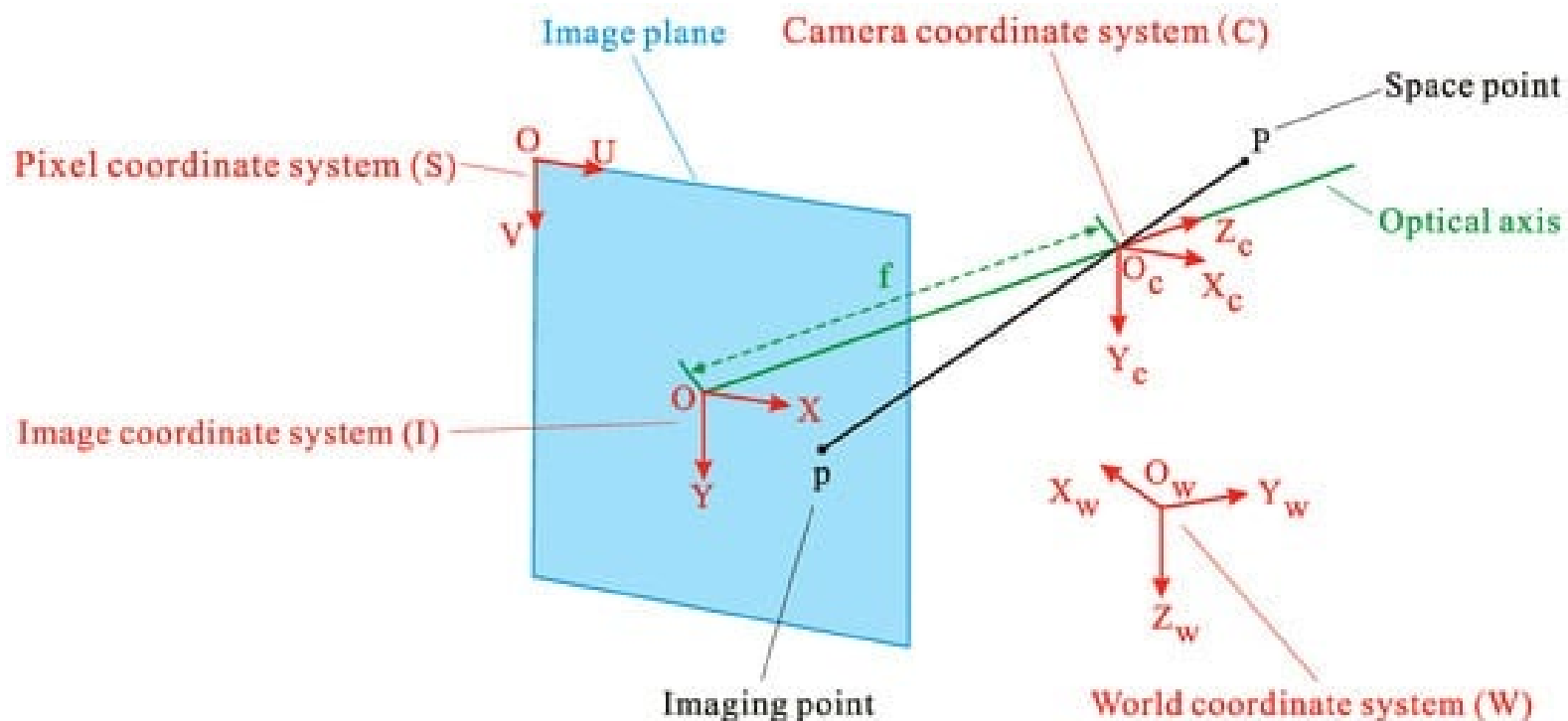
- La matriz de traslación  $t$  no es más que un vector que indica cuánto habría que mover el origen de un sistema de coordenadas para hacerlo coincidir con el otro



$$t = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$



# 5. Parámetros extrínsecos



# Índice de contenidos

---

1. Introducción
2. Geometría del modelo pinhole
3. Proyecciones
4. Parámetros intrínsecos
5. Parámetros extrínsecos
- 6. Matriz de proyección**
7. Calibración

## 6. Matriz de proyección

- Si unimos todas las matrices que hemos visto anteriormente, surge la operación que hay que hacer para finalmente proyectar un punto 3D del mundo real a un punto 2D en el plano de la imagen

$$x = K(RX + t)$$

- A veces, las matrices  $R$  y  $t$  se expresan conjuntamente en una matriz  $T$  de transformación

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

## 6. Matriz de proyección

- Combinando todas las matrices obtenemos:

$$w \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & [s] & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Parámetros intrínsecos
Parámetros extrínsecos

- $(X, Y, Z)$  son las coordenadas de un punto 3D en el mundo real
- $(u, v)$  son las coordenadas del punto proyectado en píxeles
- $f_x$  y  $f_y$  son las longitudes focales expresadas en píxeles
- $c_x$  y  $c_y$  son los puntos principales pertenecientes al centro de la imagen
- $r$  y  $t$  representan las matrices de rotación y traslación entre la cámara y la zona proyectada
- $s$  es el ángulo que tiene un píxel respecto al eje  $y$  (normalmente a 0)
- $w$  es el factor de escala que sirve para homogeneizar  $u$  y  $v$

# Índice de contenidos

---

1. Introducción
2. Geometría del modelo pinhole
3. Proyecciones
4. Parámetros intrínsecos
5. Parámetros extrínsecos
6. Matriz de proyección
- 7. Calibración**

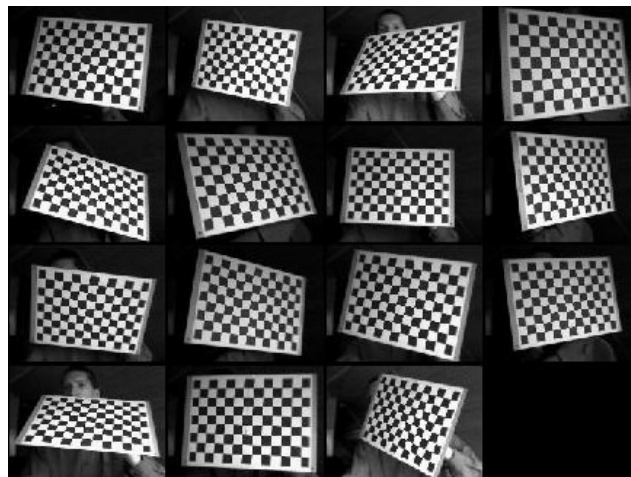
# 7. Calibración

---

- Los métodos de calibración nos permiten estimar los parámetros intrínsecos y extrínsecos de una cámara a partir de un conjunto de imágenes de un elemento de calibración
- El **elemento de calibración** normalmente es un **patrón** estilo **tablero de ajedrez** que está bien definido
- Las imágenes se pasan por un algoritmo que detecta las intersecciones de los cuadrados
- A partir de los puntos 2D extraídos en el paso anterior y sus correspondientes posiciones en el mundo real 3D, se minimiza el error de reproyección en todas las imágenes
- El tamaño de las casillas es conocido, por lo tanto, se puede calcular la correspondencia entre los puntos 2D y 3D fácilmente

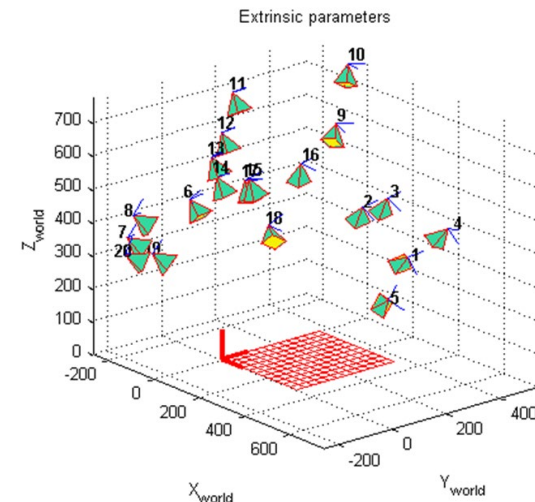
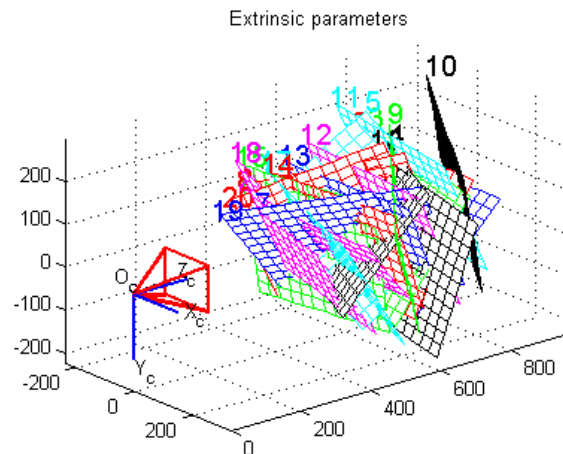
# 7. Calibración

1. Se capturan muchas imágenes del elemento de calibración
  - Para la calibración hacen falta puntos sobre un plano (muy importante que sea plano, ya veremos por qué dentro de poco)
  - Para hacer más fácil la selección de puntos, se usa un patrón estilo tablero de ajedrez porque se pueden obtener las intersecciones de las casillas de forma automática con algoritmo de detección de esquinas
  - Es muy importante que el plano sufra transformaciones de perspectiva, para dotar al sistema de diferentes valores de profundidad



# 7. Calibración

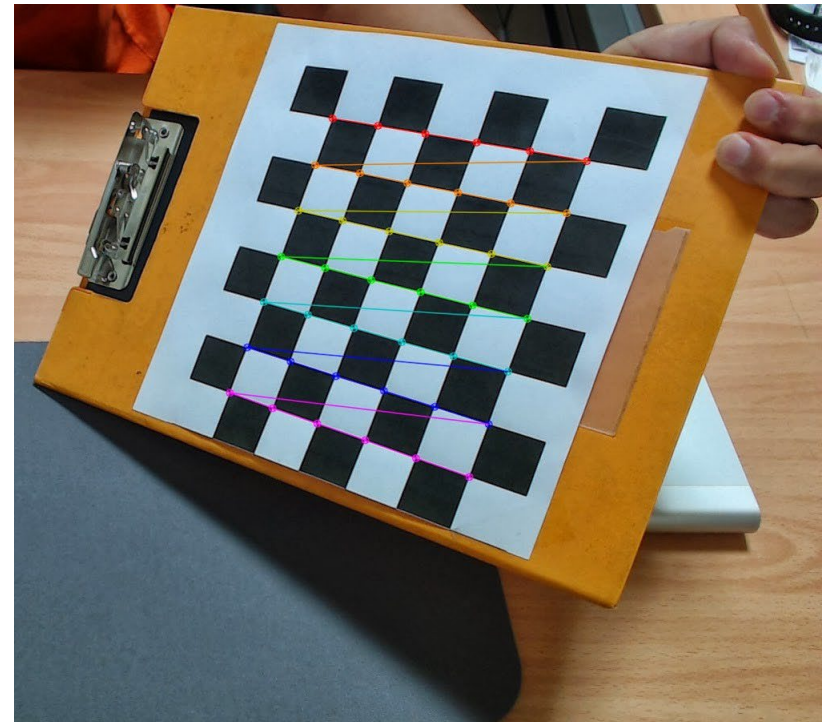
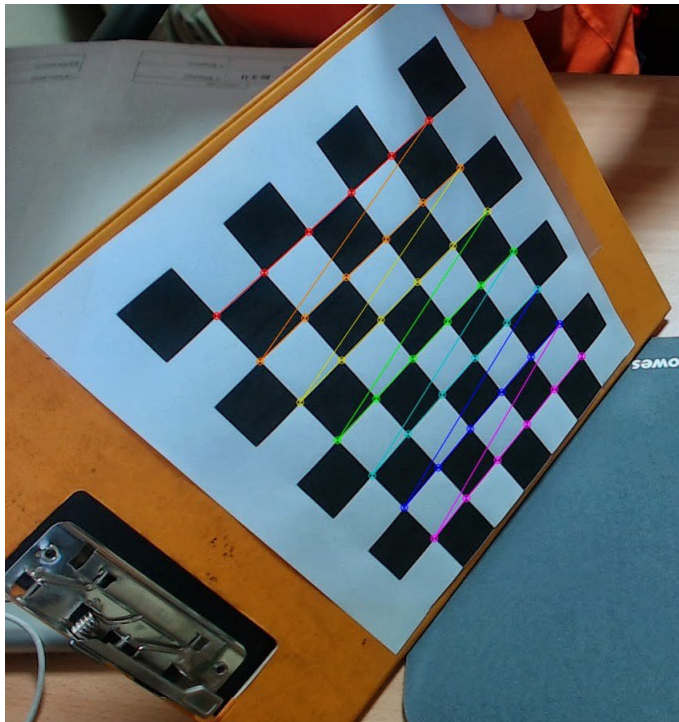
- Existen dos enfoques principales para obtener los parámetros de la cámara (intrínsecos y extrínsecos):
  - Mover el patrón
  - Mover la cámara
- Ambos enfoques difieren en cómo se logra, pero tienen el mismo objetivo:
  - Obtener las correspondencias entre puntos 3D del mundo real y puntos 2D en la imagen





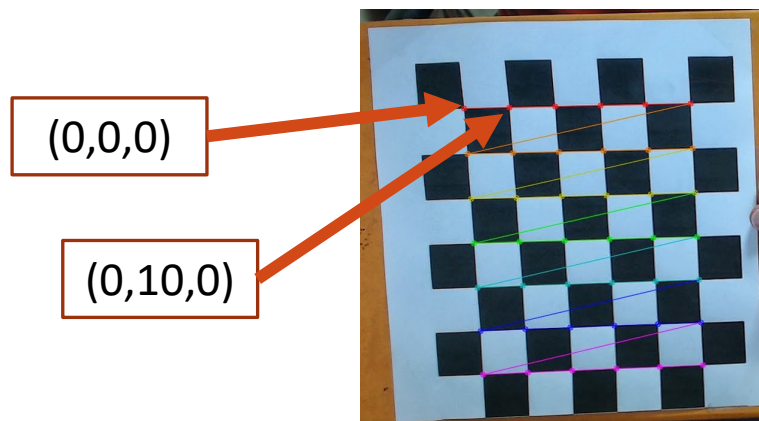
# 7. Calibración

2. Se computan los puntos 2D en el plano de la imagen que se van a usar para minimizar el error de reproyección con un algoritmo de detección de esquinas



# 7. Calibración

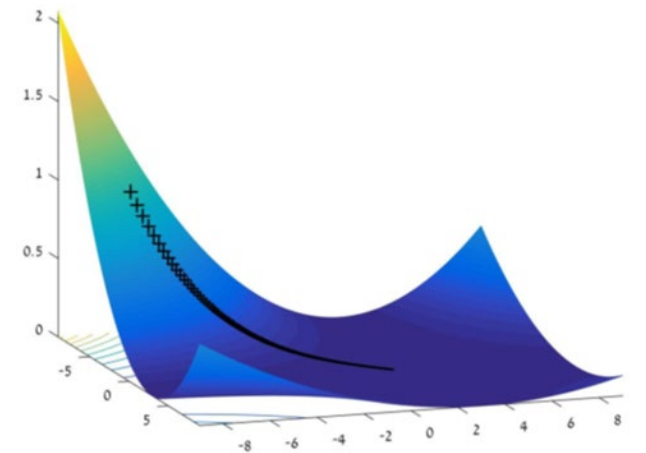
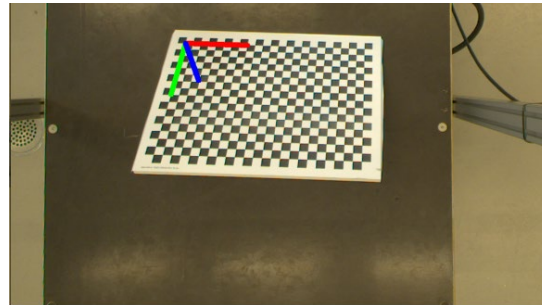
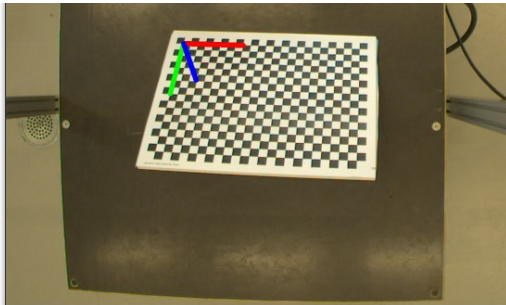
- Los **puntos 2D** en la imagen son las **intersecciones** de las casillas del tablero de ajedrez
- El tablero de ajedrez es conocido y tiene una estructura regular (sabemos cuánto mide el lado de cada cuadrado). Esto permite definir manualmente las **coordenadas 3D**
- Es muy complicado calcular la posición 3D de cada punto 2D de la imagen. Lo que se hace es establecer el **origen de coordenadas en el tablero**, y se **calcula la transformación** de la cámara con respecto del tablero para cada imagen
- El patrón (tablero) se coloca **en un plano 2D**, así todas las esquinas del tablero tienen la misma **profundidad  $Z = 0$**  (simplifica el cálculo y reducir la complejidad al poder trabajar solo con las coordenadas  $X$  e  $Y$ )



# 7. Calibración

## 3. Minimización del error de reproyección (Levenberg-Marquardt)

- A partir de las correspondencias 3D-2D:
  - Se proyectan los puntos 3D a 2D usando una  $K$ ,  $R$  y  $t$  iniciales
  - Se minimiza el error entre la proyección anterior y los puntos 2D observados (los que se obtienen del algoritmo detector de esquinas). Esto modifica  $K$ ,  $R$  y  $t$  para hacer coincidir la proyección con la posición observada
- El algoritmo devuelve  $K$ ,  $R$  y  $t$  optimizadas



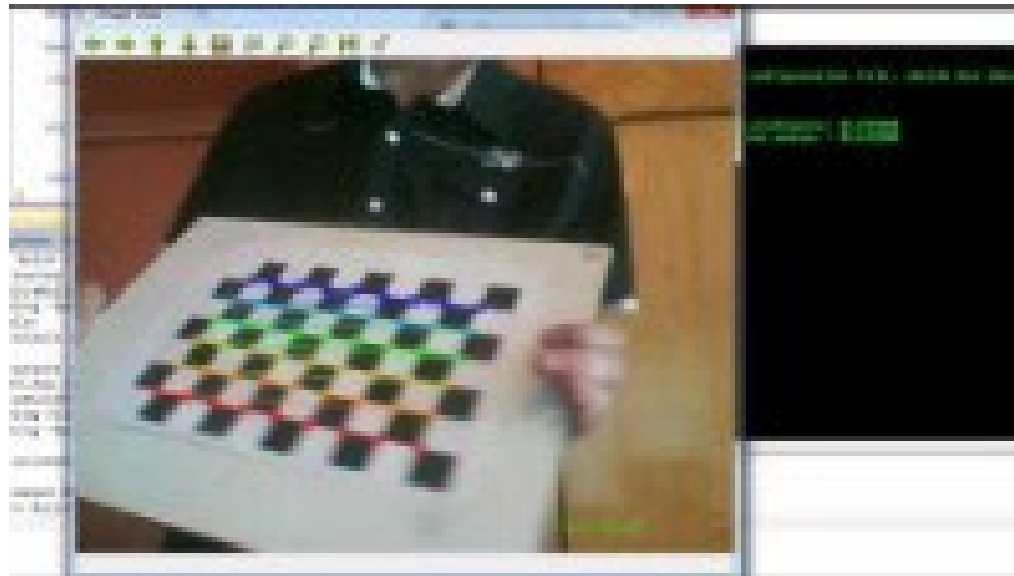
# 7. Calibración

---

- Este método es el que integra OpenCV
- Pero hay muchos pipelines de calibración:
  - Diferentes tableros
  - Diferentes formas de sacar puntos 2D del tablero
  - Diferentes formas de obtener correspondencias 3D
  - Diferentes algoritmos de optimización
- Cada uno tiene sus ventajas e inconvenientes

# 7. Calibración

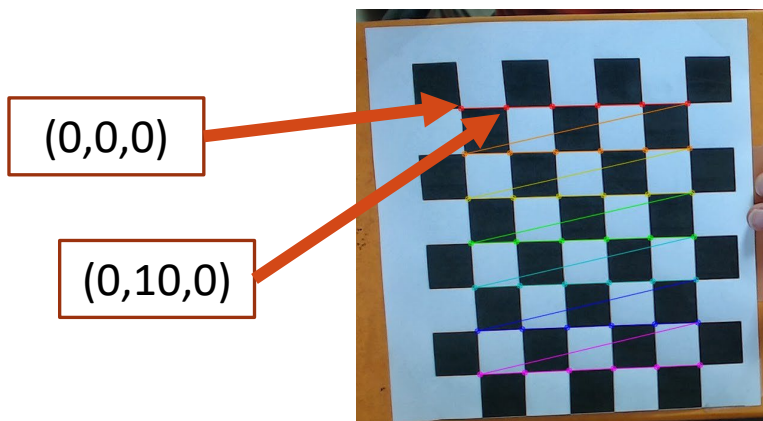
- Ejemplo OpenCV:



[https://docs.opencv.org/master/d4/d94/tutorial\\_camera\\_calibration.html](https://docs.opencv.org/master/d4/d94/tutorial_camera_calibration.html)

# 7. Calibración

- Esto se hace así porque para minimizar el error de reproyección es necesario un conjunto de correspondencias 2D de la imagen con puntos 3D del mundo
- Los puntos 2D los obtenemos de las intersecciones de las casillas del tablero de ajedrez
- Es imposible calcular la posición 3D con respecto de la cámara de cada punto 2D de la imagen. Lo que se hace es establecer el origen de coordenadas en el tablero, y se calcula la transformación de la cámara con respecto del tablero para cada imagen
- Como hemos dicho, el tablero es bien conocido, por lo que podemos crear el conjunto de puntos 3D “a mano” para las correspondientes posiciones de las esquinas en 2D



Se sabe que el lado de un cuadrado mide 10mm

Imposible ponerle profundidad de forma precisa con este método, por eso el patrón se sitúa en un plano forzando a que la  $Z$  sea siempre 0