Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

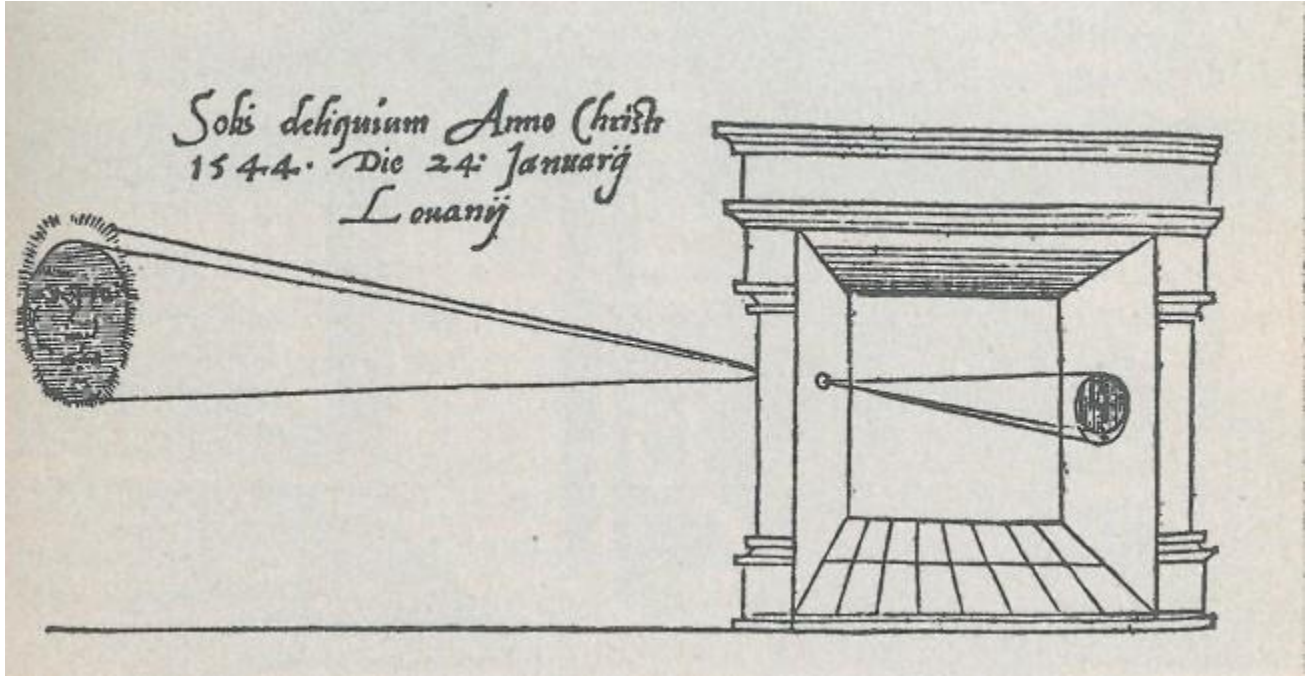
CS7GV1 Computer vision

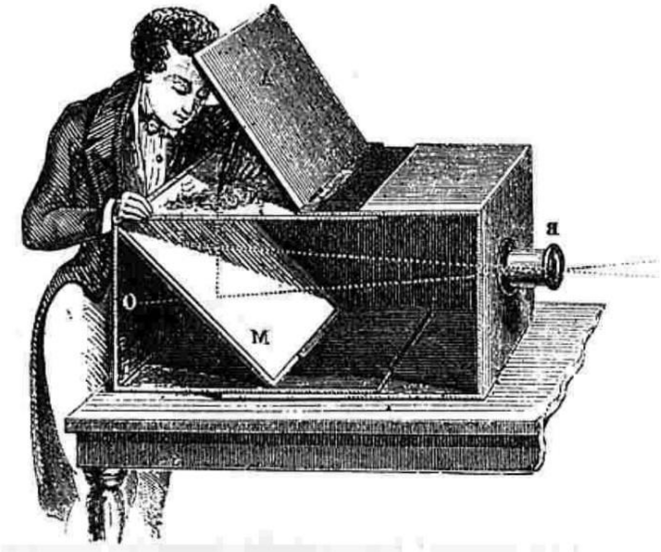
Camera models

Dr. Martin Alain

Introduction – Image formation

- Camera obscura

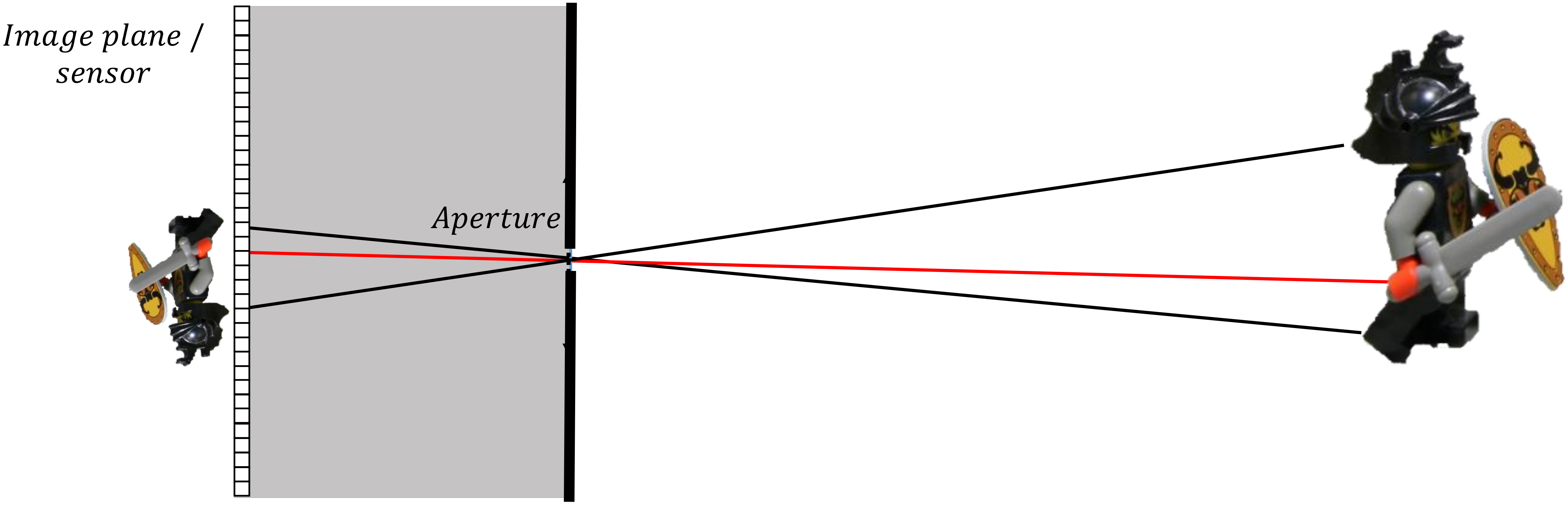




https://en.wikipedia.org/wiki/Camera_obscura

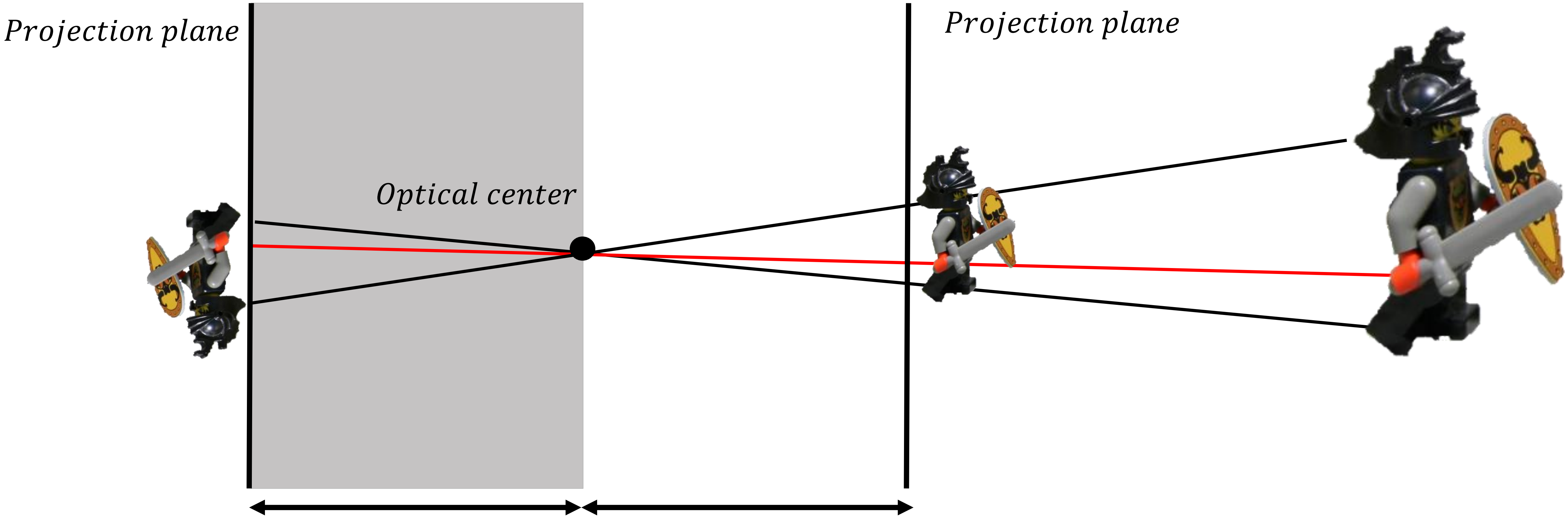
The pinhole camera model

- How to link points in 3D world coordinates to 2D image coordinates



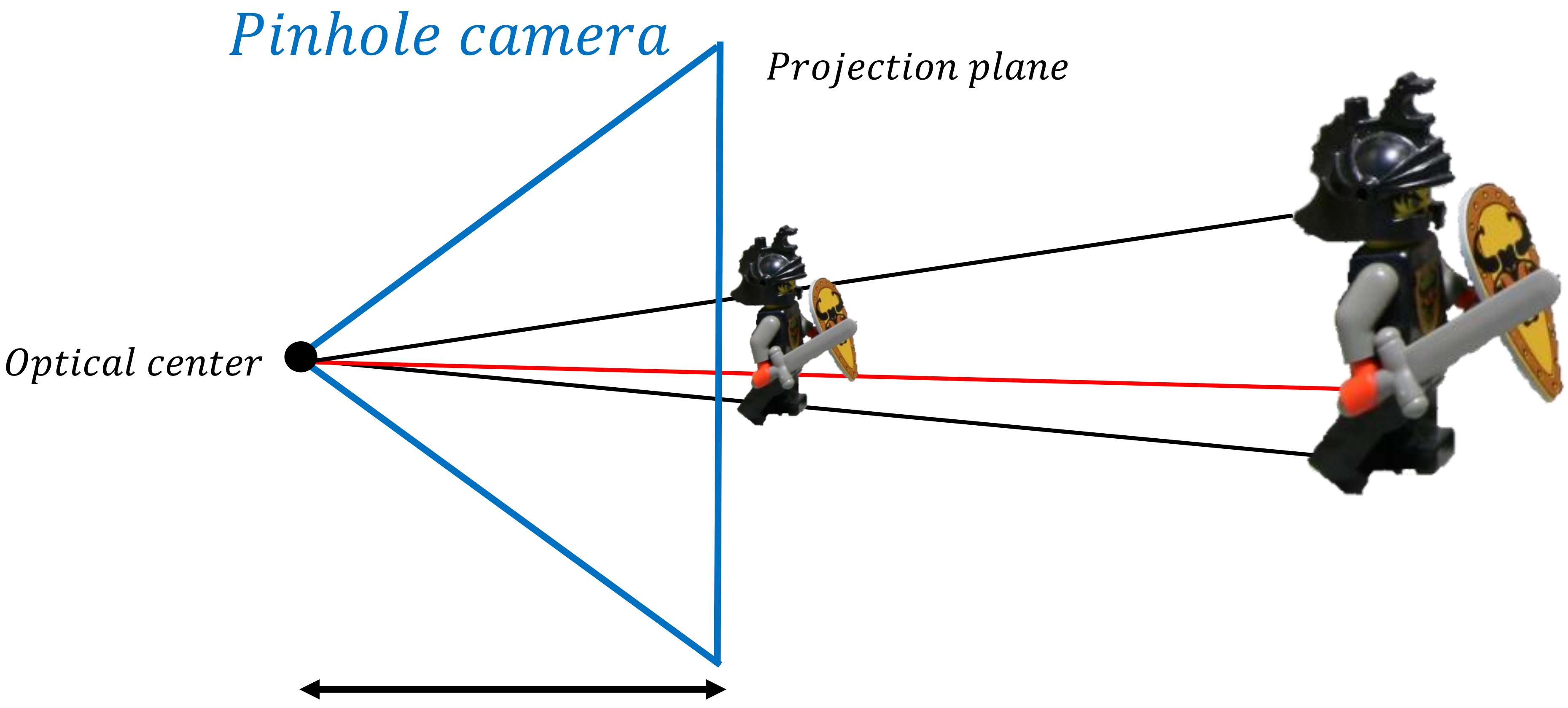
The pinhole camera model

- How to link points in 3D world coordinates to 2D image coordinates



The pinhole camera model

- How to link points in 3D world coordinates to 2D image coordinates

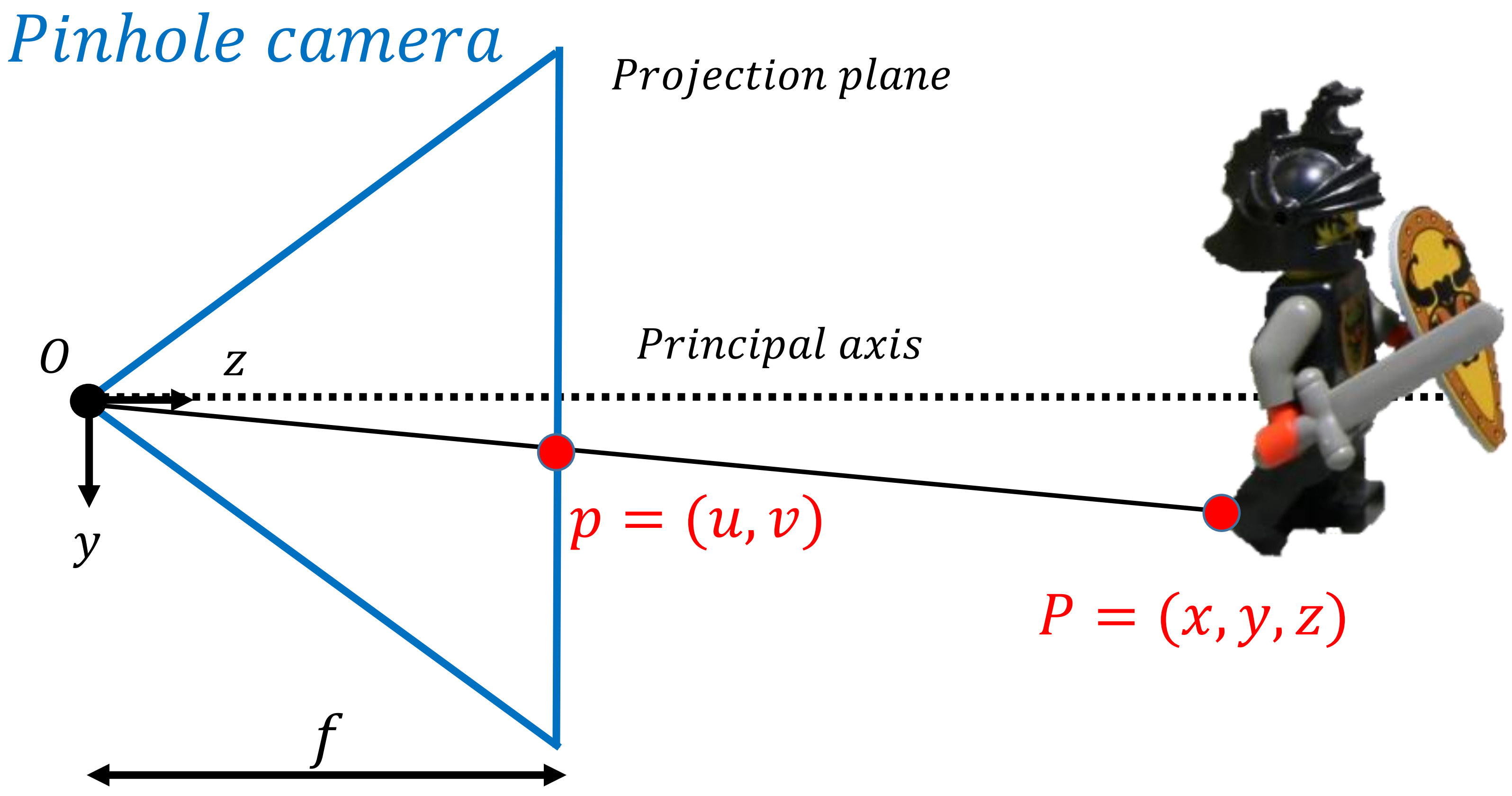


The pinhole camera model

- How to link points in 3D world coordinates to 2D image coordinates

$$u = \frac{f}{z}x$$

$$v = \frac{f}{z}y$$



Homogeneous coordinates

$$u = \frac{f}{z}x$$

$$v = \frac{f}{z}y$$

Not linear

Homogeneous coordinates to the rescue!

$$(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

homogeneous image coordinates

$$(x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

homogeneous scene coordinates

Converting *from* homogeneous coordinates

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)$$

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)$$

Source: [Raquel Urtasun (lecture notes)](), N. Snavely

Projection matrix

Projection is a matrix multiply using homogeneous coordinates

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z/f \end{bmatrix} \Rightarrow \left(f\frac{x}{z}, \; f\frac{y}{z} \right) = (u, v)$$

divide by third coordinate

This is known as **perspective projection**

The matrix is called the **projection matrix**

Can also represent as a 4x4 matrix

Source: Raquel Urtasun (lecture notes), N. Snavely

Projection matrix

How does scaling the projection matrix change the transformation?

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z/f \end{bmatrix} \Rightarrow \left(f\frac{x}{z}, \quad f\frac{y}{z} \right) = (u, v)$$

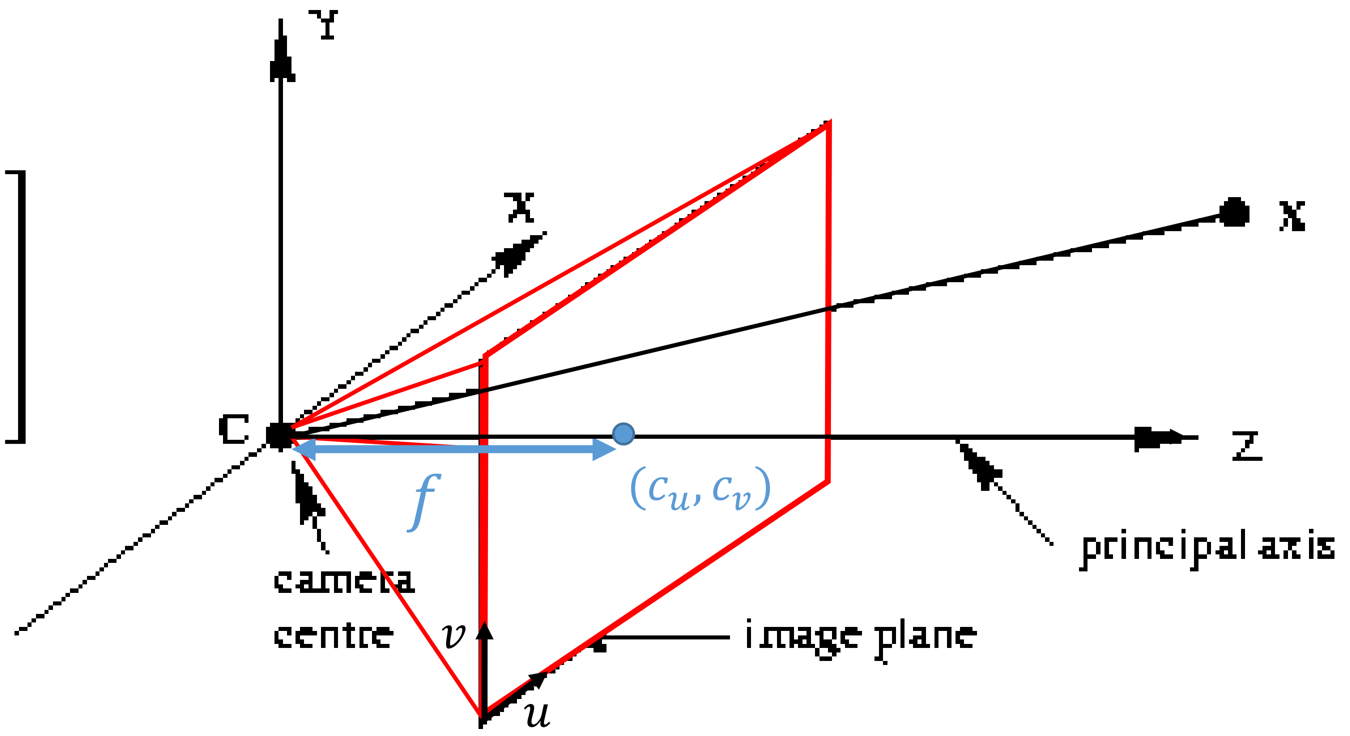
$$\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} f\,x \\ f\,y \\ z \end{bmatrix} \Rightarrow \left(f\frac{x}{z}, \quad f\frac{y}{z} \right) = (u, v)$$

It is not possible to recover the distance of the 3D point from the image.

Source: [Raquel Urtasun (lecture notes)](), N. Snavely

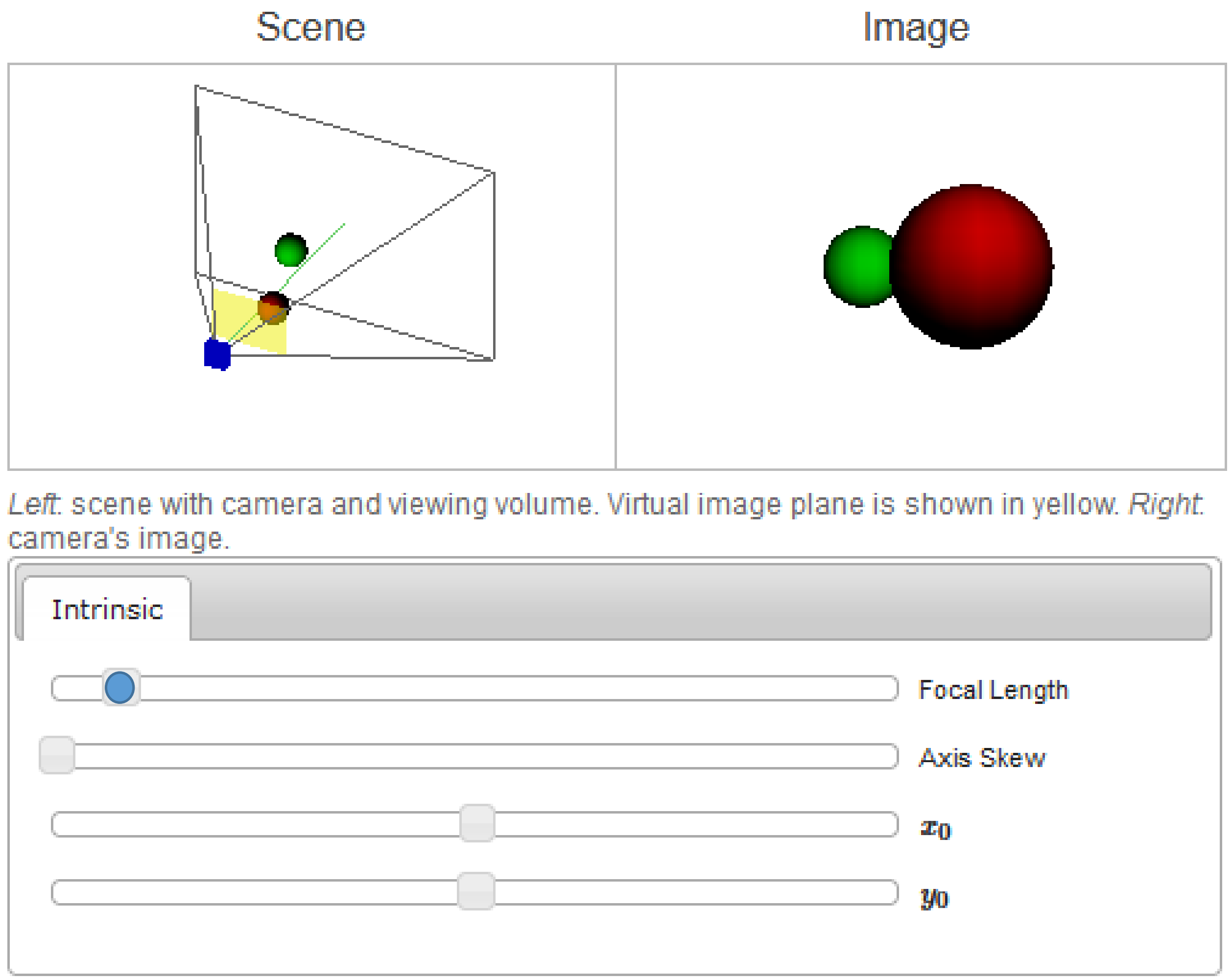
Intrinsic camera parameters

$$K = \begin{bmatrix} f_x & s_x & 0 & c_u \\ s_y & f_y & 0 & c_v \\ 0 & 0 & 1 & 0 \end{bmatrix}$$



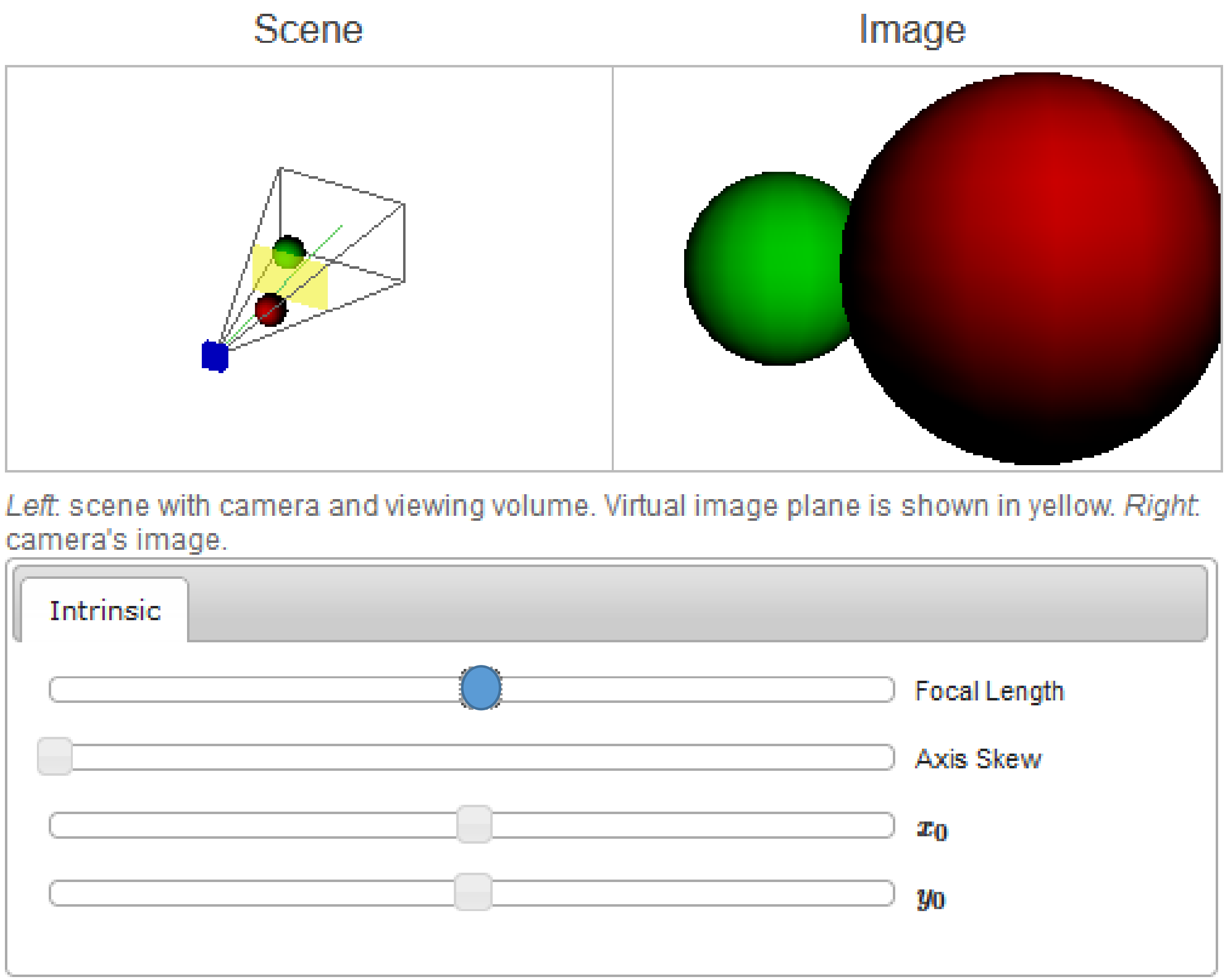
Hartley and Zisserman ["Multiple View Geometry in Computer Vision"]()

Intrinsic camera parameters



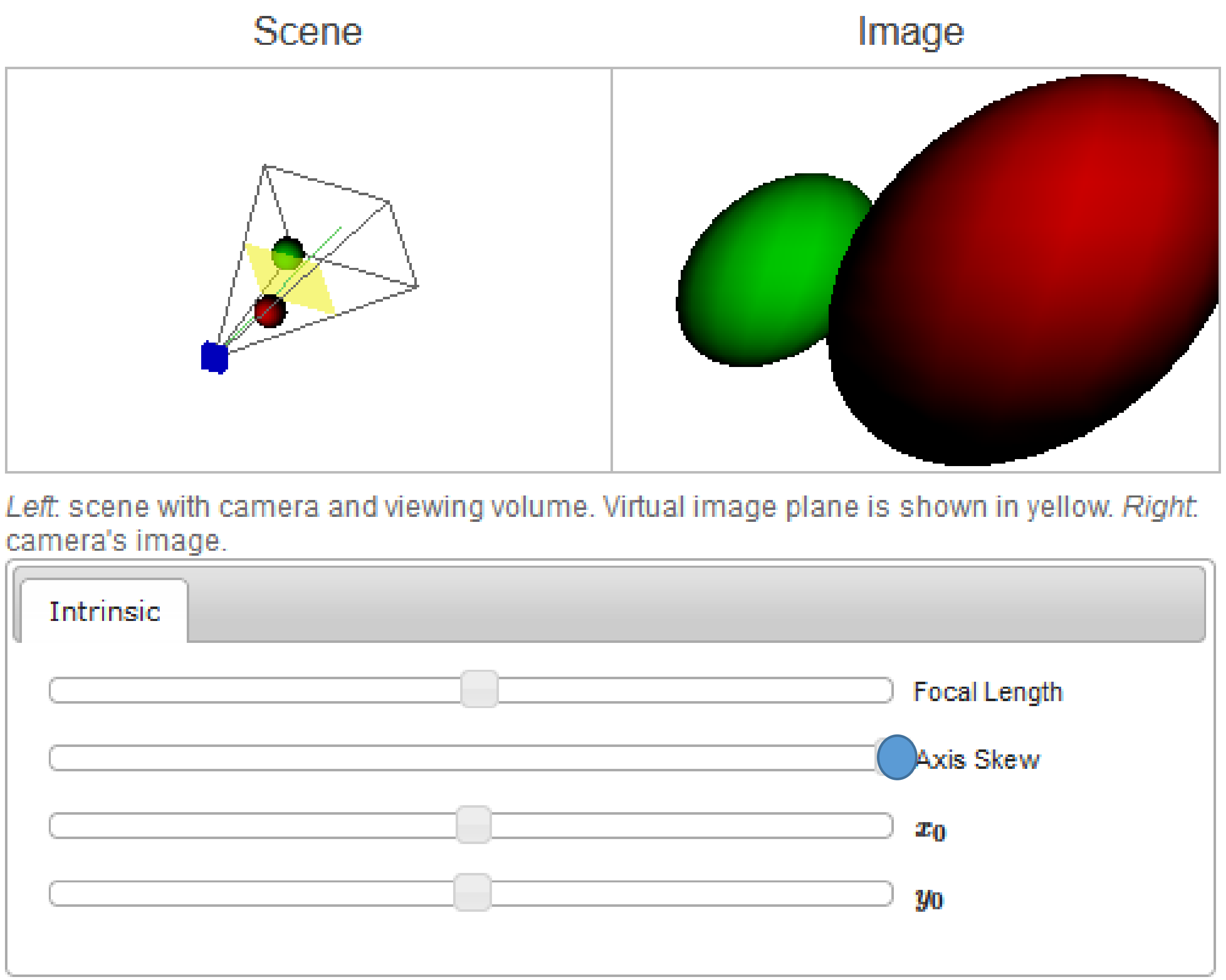
https://ksimek.github.io/2013/08/13/intrinsic/

Intrinsic camera parameters



https://ksimek.github.io/2013/08/13/intrinsic/

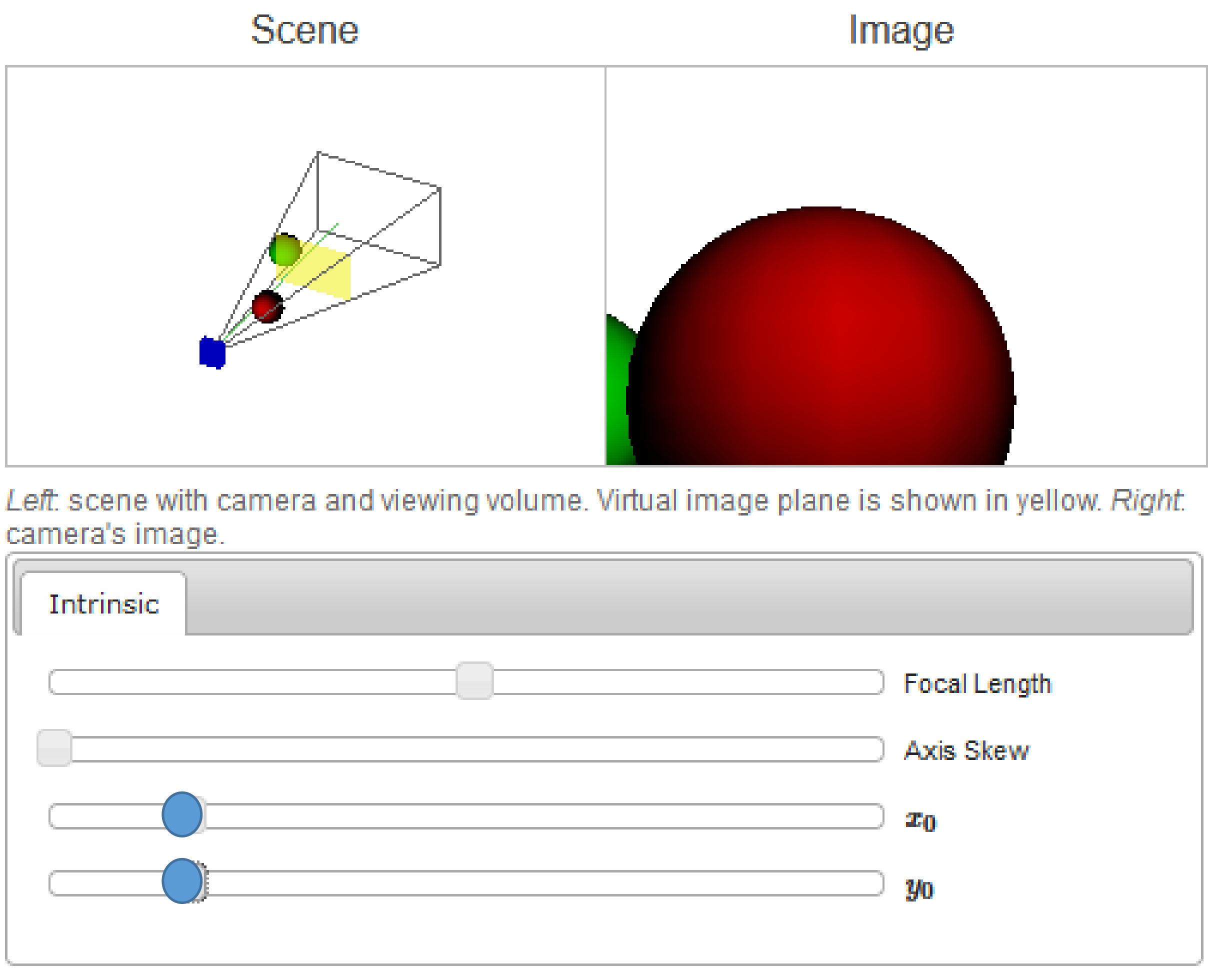
Intrinsic camera parameters



Left: scene with camera and viewing volume. Virtual image plane is shown in yellow. *Right*: camera's image.

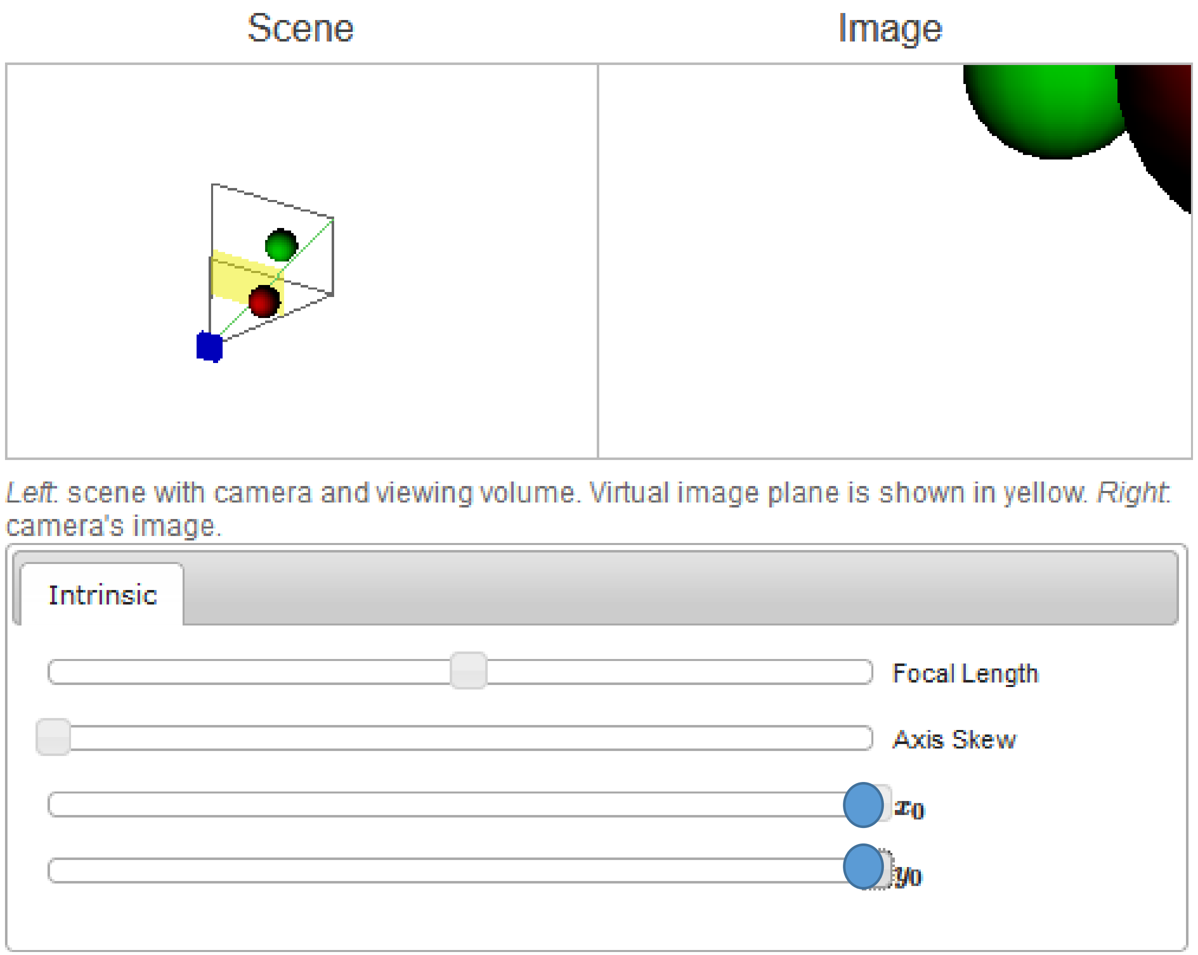
https://ksimek.github.io/2013/08/13/intrinsic/

Intrinsic camera parameters



https://ksimek.github.io/2013/08/13/intrinsic/

Intrinsic camera parameters

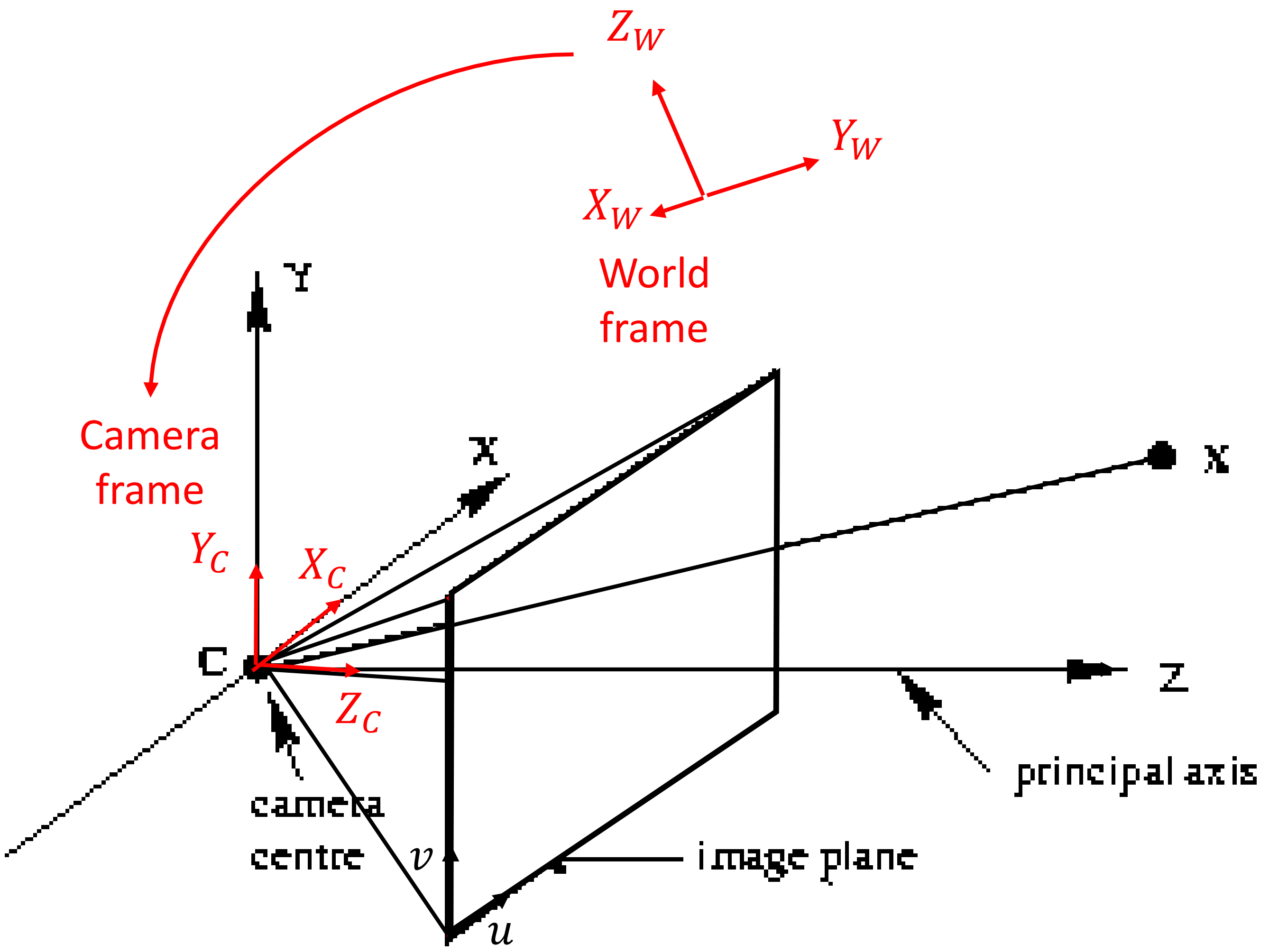


https://ksimek.github.io/2013/08/13/intrinsic/

Try it yourself!

Extrinsic camera parameters

$$[R \mid t] = \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} & t_x \\ r_{2,1} & r_{2,2} & r_{2,3} & t_y \\ r_{3,1} & r_{3,2} & r_{3,3} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



Hartley and Zisserman ["Multiple View Geometry in Computer Vision"]

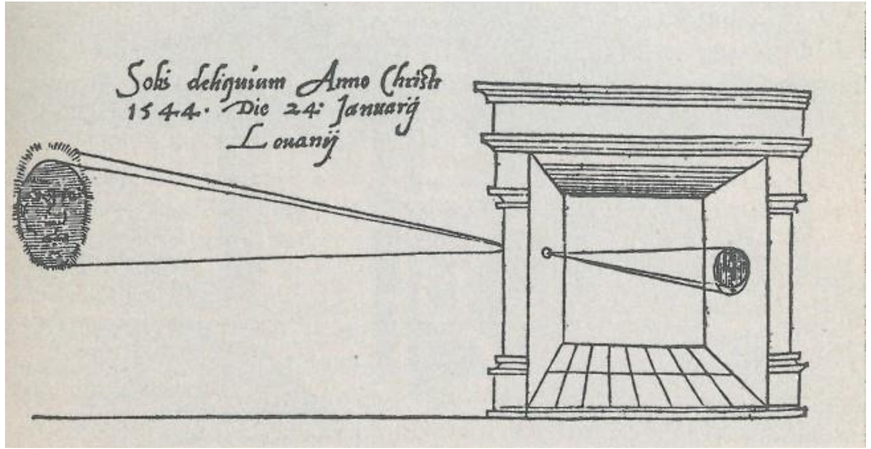
The pinhole camera model

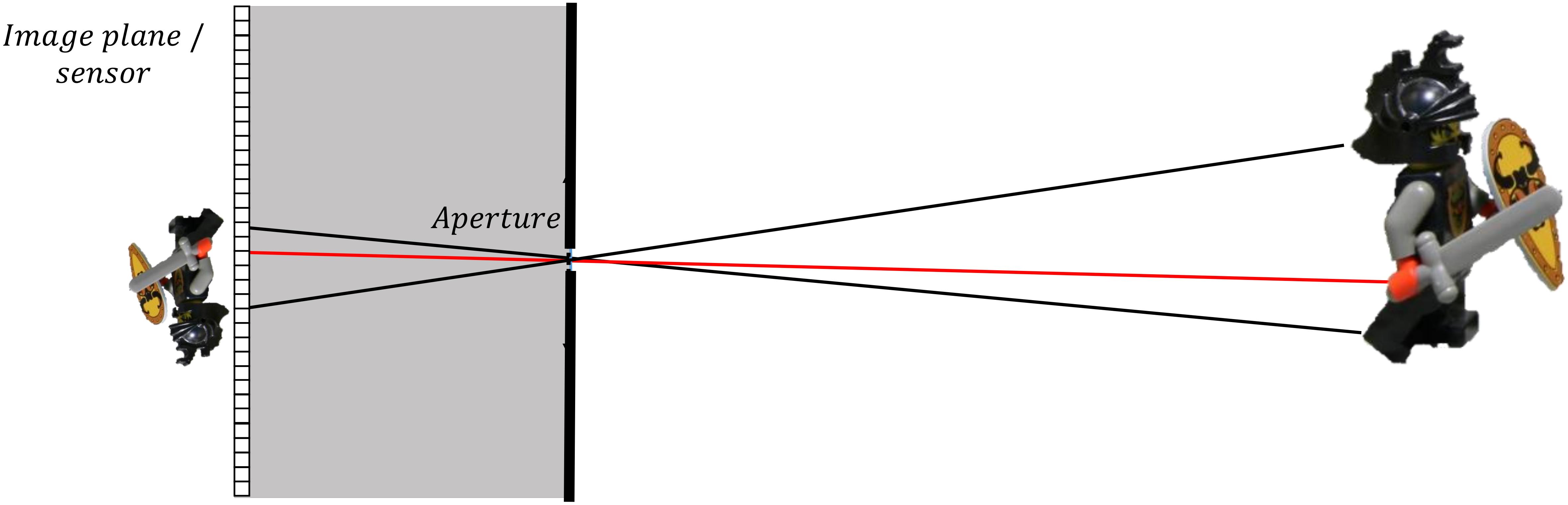
- How to link points in 3D world coordinates to 2D image coordinates

$$K[R \mid \ t]\begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} = \begin{bmatrix} u \\ v \\ w \end{bmatrix} \Rightarrow \begin{bmatrix} u/w \\ v/w \\ 1 \end{bmatrix}$$

The pinhole camera model

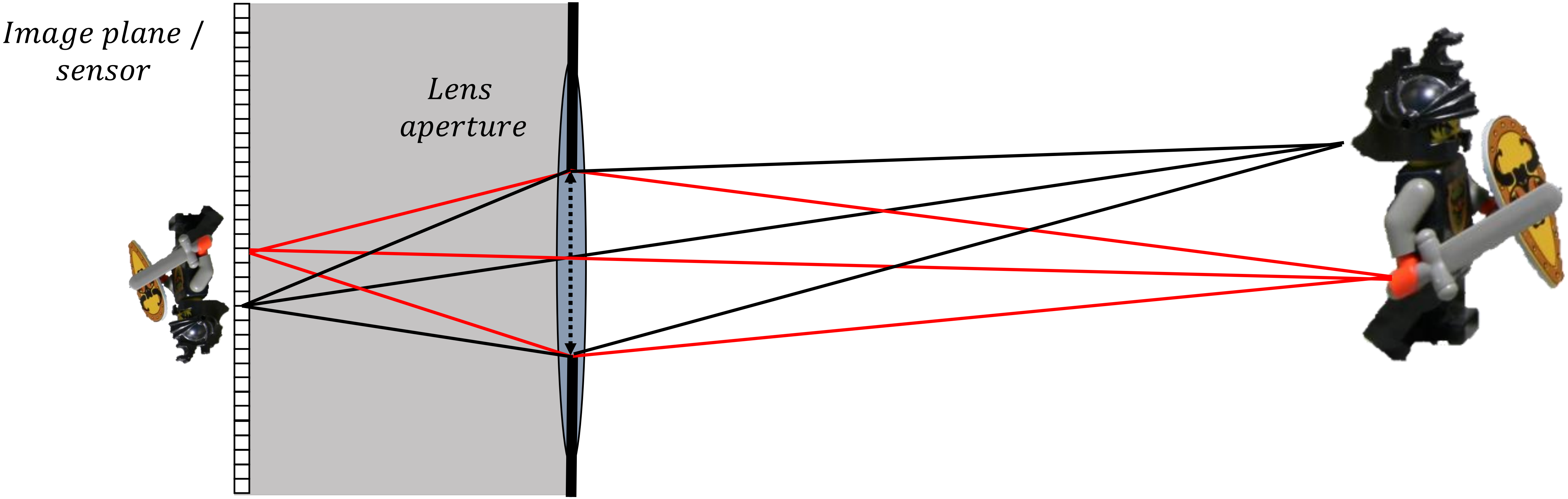
- Why does it hold?





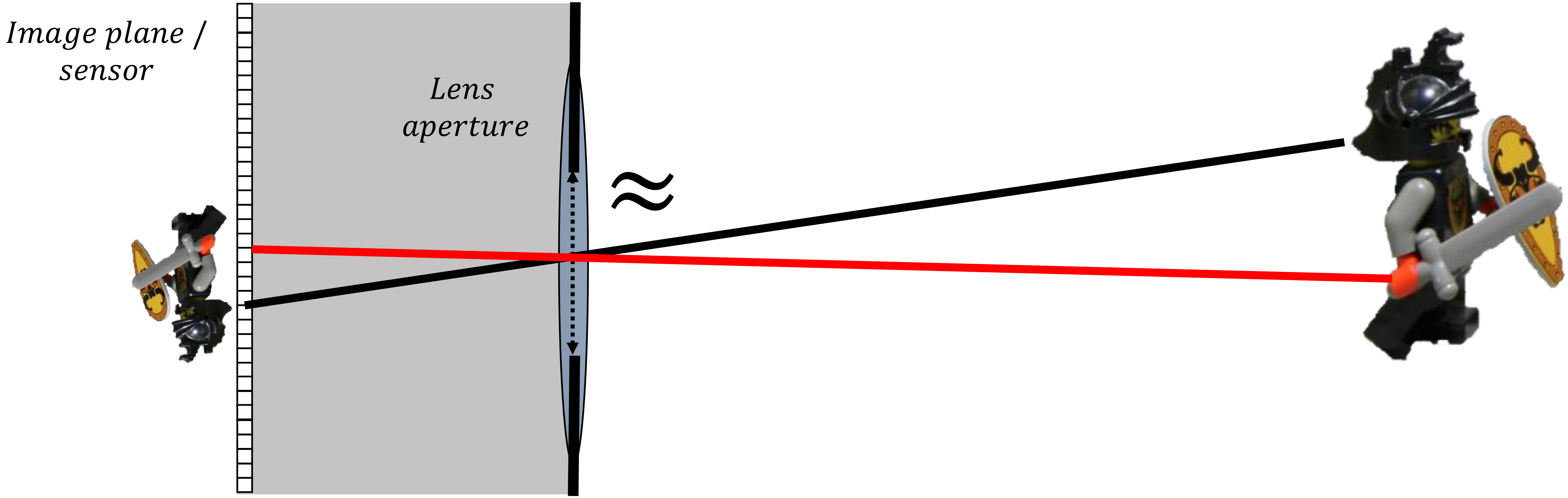
The pinhole camera model

- Why does it hold?



The pinhole camera model

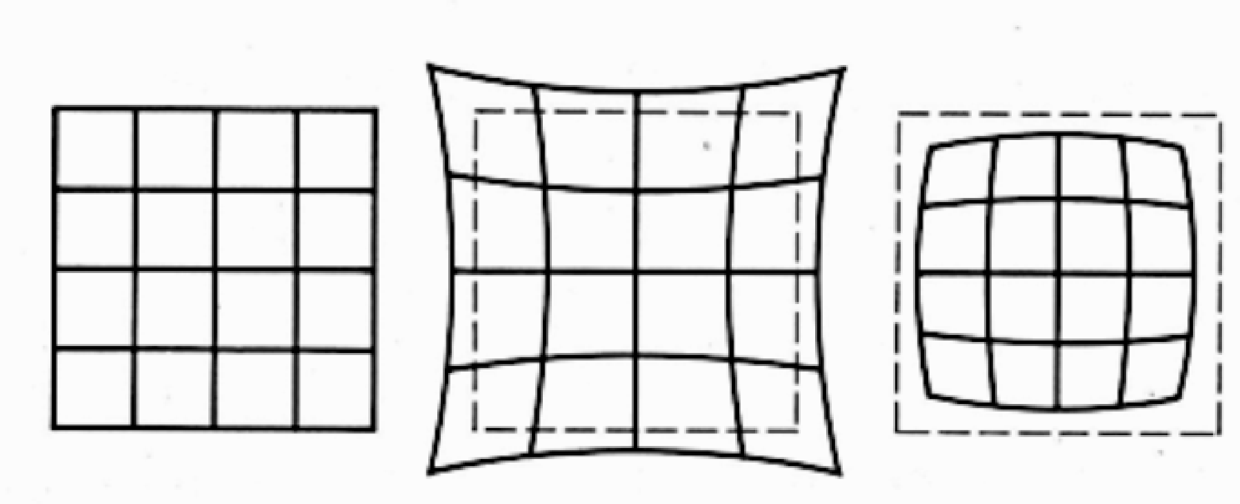
- Why does it hold?



Radial distortions

Caused by imperfect lenses

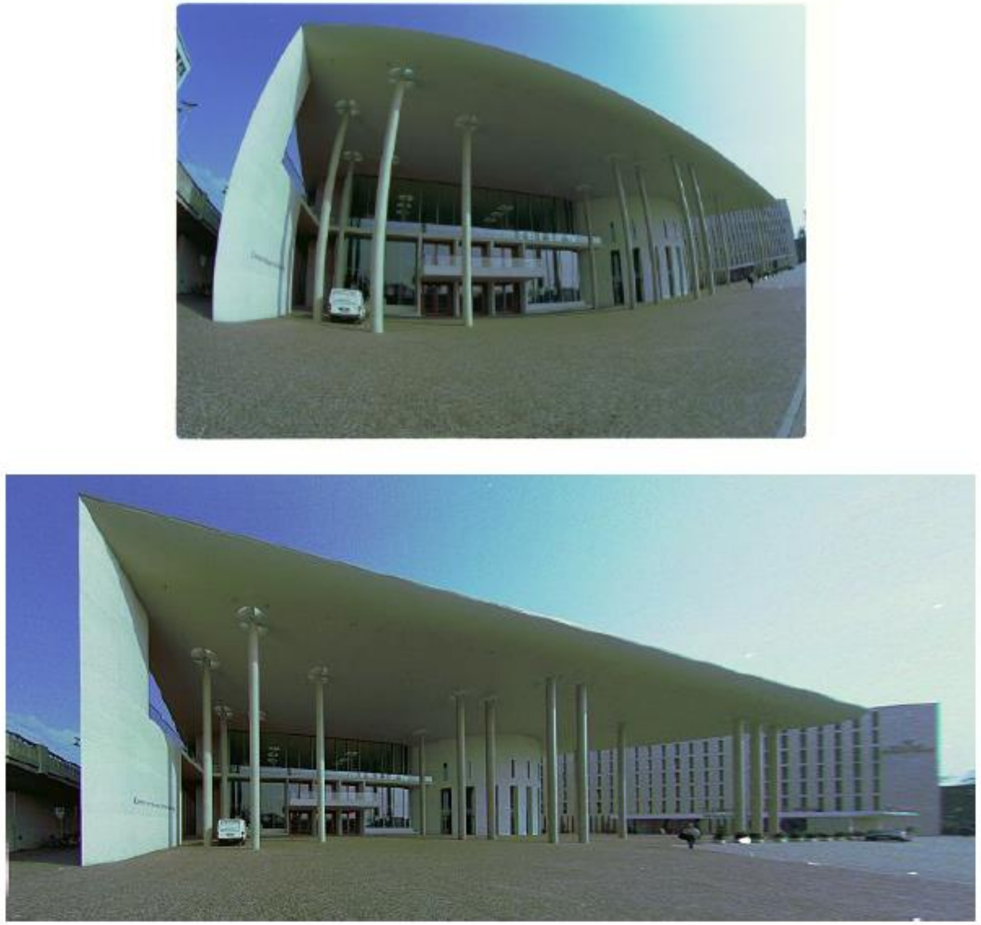
Deviations are most noticeable for rays that pass through the edge of the lens



No distortion   Pin cushion   Barrel

Source: [Raquel Urtasun (lecture notes)](), N. Snavely

Radial distortions



from Helmut Dersch

Source: Raquel Urtasun (lecture notes), Helmut Dersh

Modeling distortions

Project point to normalized image coordinates

$$x_n = \frac{x}{z}$$
$$y_n = \frac{y}{z}$$

Apply radial distortion

$$r^2 = x_n^2 + y_n^2$$
$$x_d = x_n(1 + \kappa_1 r^2 + \kappa_2 r^4)$$
$$y_d = y_n(1 + \kappa_1 r^2 + \kappa_2 r^4)$$

Apply focal length and translate image center

$$x' = f x_d + x_c$$
$$y' = f y_d + y_c$$

To model lens distortion use above projection operation instead of standard projection matrix multiplication

Source: Raquel Urtasun (lecture notes), N. Snavely

Modeling distortions

Project point to normalized image coordinates

$$\begin{aligned} x_n &= \frac{x}{z} \\ y_n &= \frac{y}{z} \end{aligned}$$

Apply radial distortion

$$\begin{aligned} r^2 &= x_n^2 + y_n^2 \\ x_d &= x_n(1 + \kappa_1 r^2 + \kappa_2 r^4) \\ y_d &= y_n(1 + \kappa_1 r^2 + \kappa_2 r^4) \end{aligned}$$

Apply focal length and translate image center

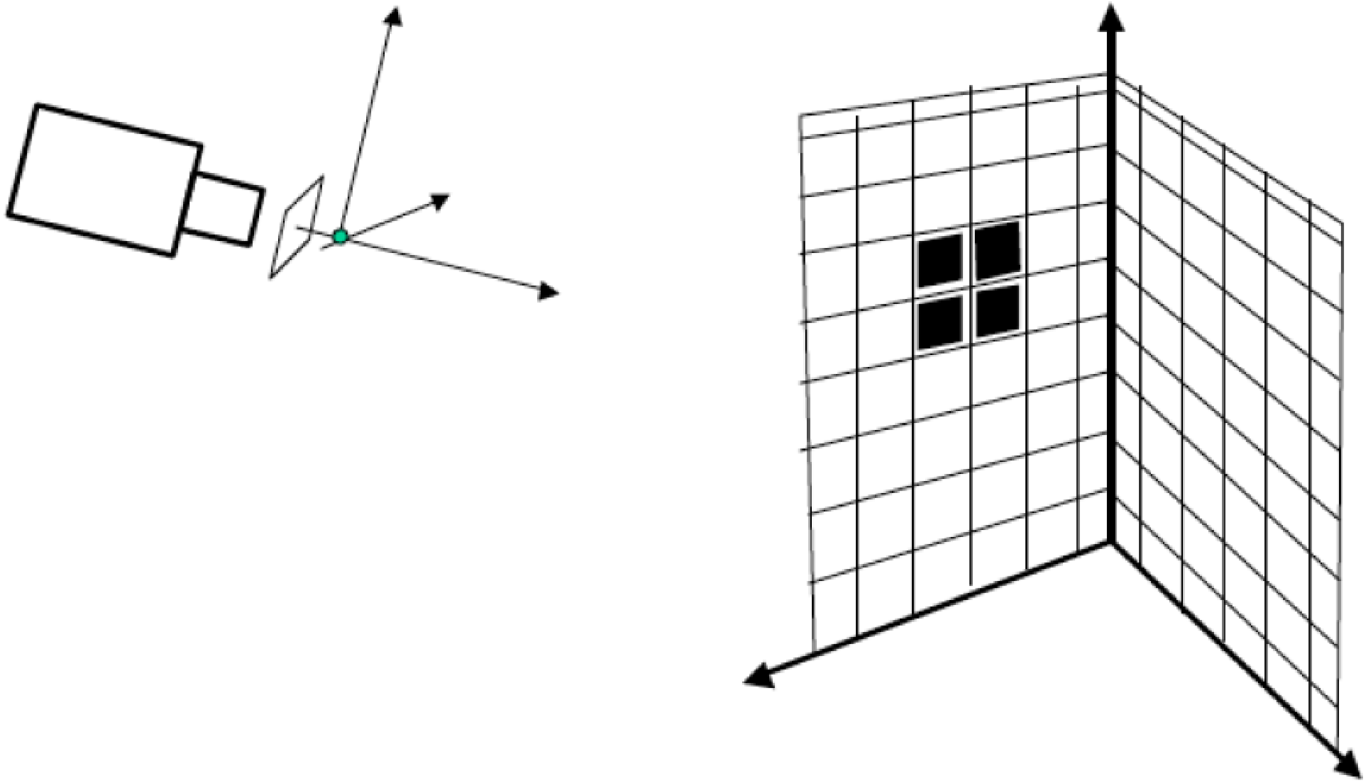
$$\begin{aligned} x' &= f x_d + x_c \\ y' &= f y_d + y_c \end{aligned}$$

To model lens distortion use above projection operation instead of standard projection matrix multiplication

Source: Raquel Urtasun (lecture notes), N. Snavely

Camera calibration

Most methods assume that we have a known 3D target in the scene



Source: [Raquel Urtasun (lecture notes),](#) Ramani

Camera calibration

Calibration target: 2 planes at right angle with checkerboard patterns (Tsai grid)

We know positions of pattern corners only with respect to a coordinate system of the target

We position camera in front of target and find images of corners

We obtain equations that describe imaging and contain internal parameters of camera

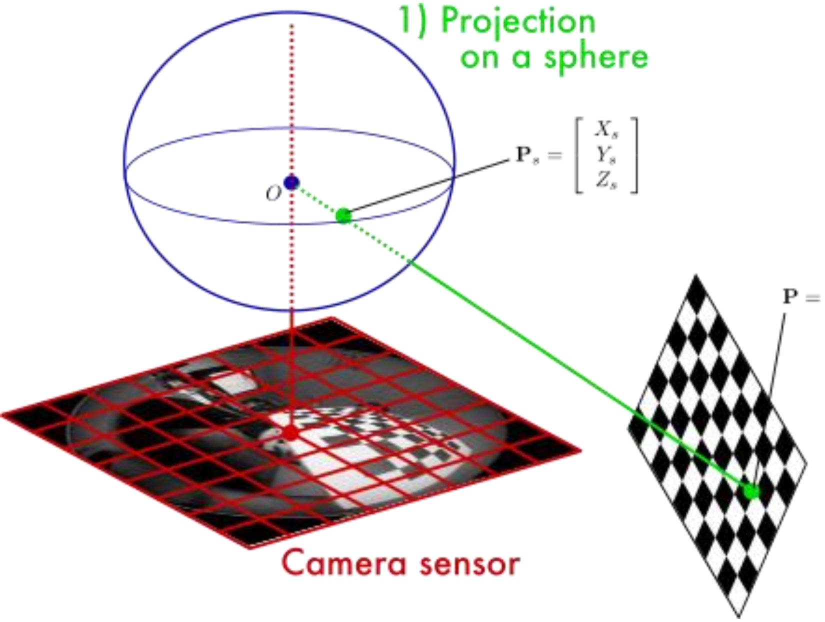
We also find position and orientation of camera with respect to target (camera pose)

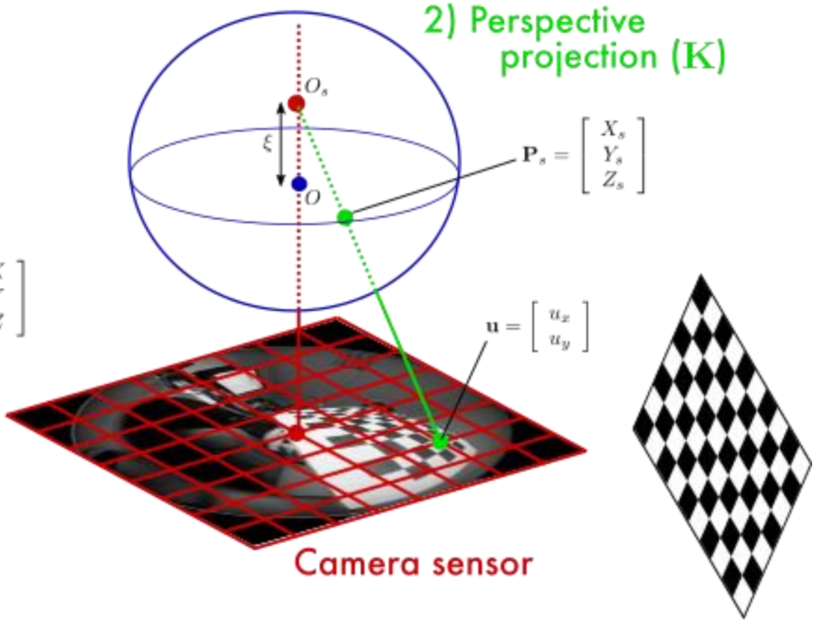
Source: Raquel Urtasun (lecture notes), Ramani

Beyond the pinhole camera

[Thomas Maugey, Laurent Guillo, Cédric Le Cam, FTV360: a Multiview 360-degree Video Dataset with Calibration Parameters, ACM Multimedia Systems Conference, June 2019]

Unified Spherical Model





$$\mathbf{P}_s = \frac{1}{||\mathbf{P}||}\mathbf{P} = \begin{bmatrix} \frac{X}{\sqrt{X+Y+Z}} \\ \frac{Y}{\sqrt{X+Y+Z}} \\ \frac{Z}{\sqrt{X+Y+Z}} \end{bmatrix}$$

$$\begin{bmatrix} u_x \\ u_y \\ 1 \end{bmatrix} \equiv \mathbf{K}\left(\mathbf{P}_s + \begin{bmatrix} 0 \\ 0 \\ \xi \end{bmatrix}\right) = \mathbf{K}\begin{bmatrix} \frac{X}{\sqrt{X+Y+Z}} \\ \frac{Y}{\sqrt{X+Y+Z}} \\ \frac{Z}{\sqrt{X+Y+Z}} + \xi \end{bmatrix}$$

[*Thomas Maugey, Laurent Guillo, Cédric Le Cam, FTV360: a Multiview 360-degree Video Dataset with Calibration Parameters, ACM Multimedia Systems Conference, June 2019*]

Summary

- A camera model links 3D world coordinates to 2D image coordinates
- The pinhole camera model is the default camera model
- Camera parameters are divided in intrinsic and extrinsic parameters
- It correspond to a perspective projection
- Other camera models exist