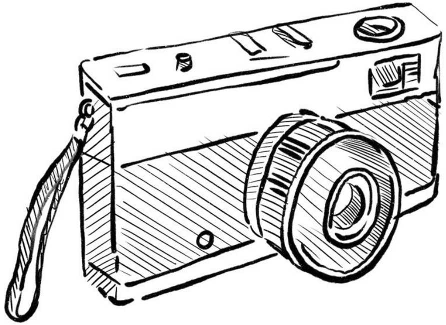
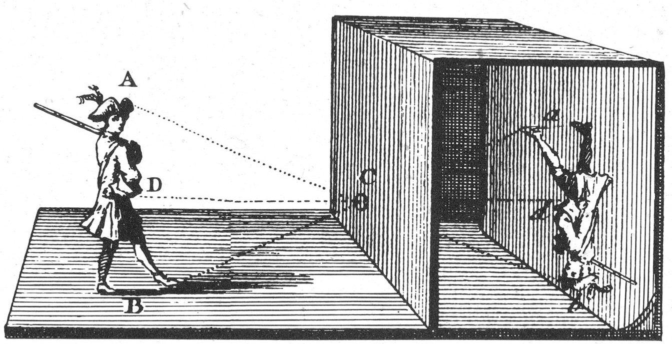


Computer Vision

Class 02



Raquel Frizera Vassallo

Class 02

01 Pinhole Model

02 Calibration Matrix

03 Radial Distortion

04 Experiment

01

Image Formation

Image Formation – Geometric Model

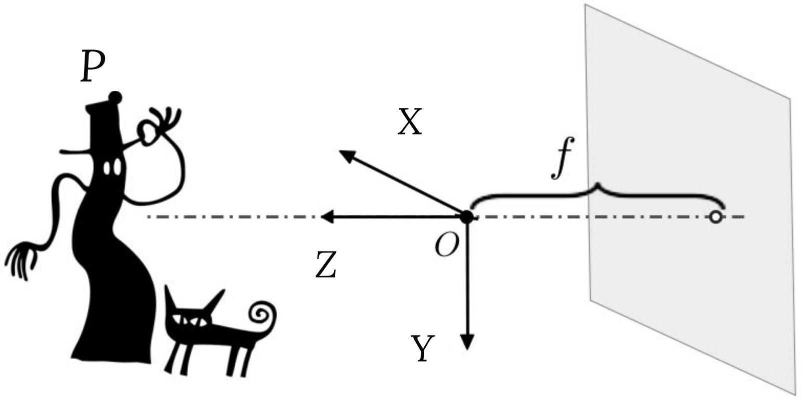
Assumptions:

- Pinhole Camera
- Lambertian Surface (uniform reflection)

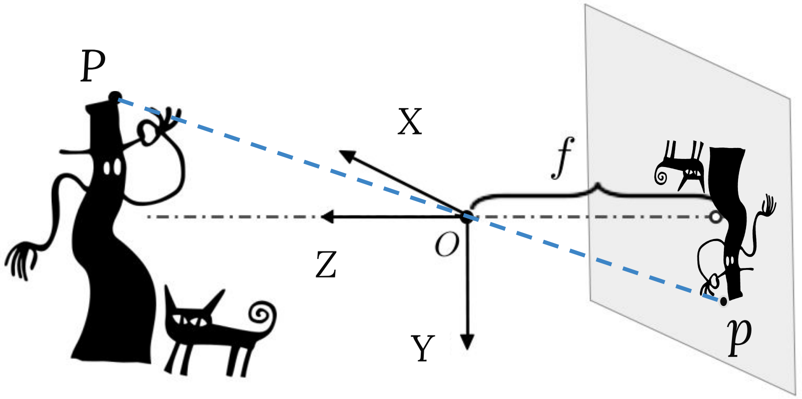
Mathematical Model accounts for 3 types of transformations:

- Coordinate transformations between the camera frame and the world frame;
- Projection of 3D coordinates onto a 2D image;
- Coordinate transformation between possible choices of image frame.

Pinhole Model

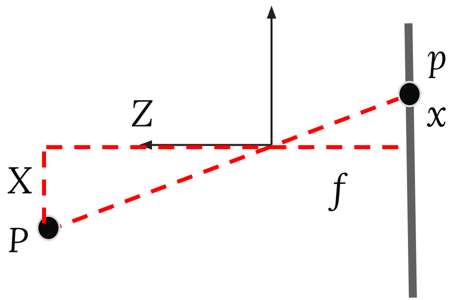


Pinhole Model

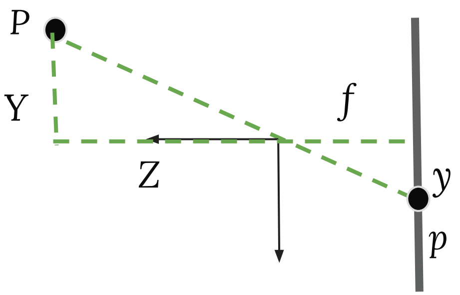


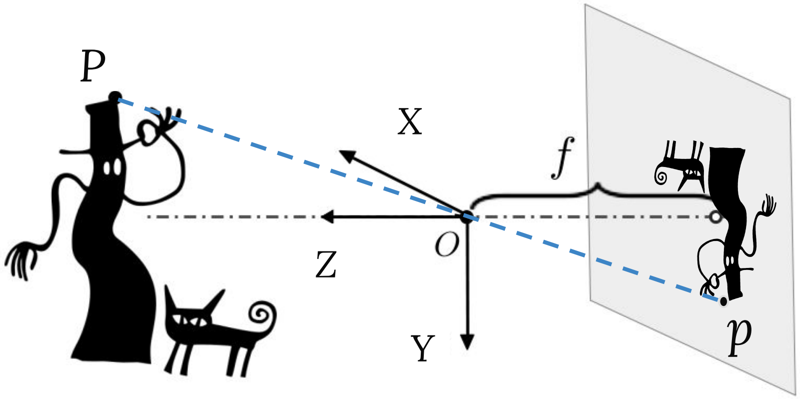
Pinhole Model

Top view



Side view





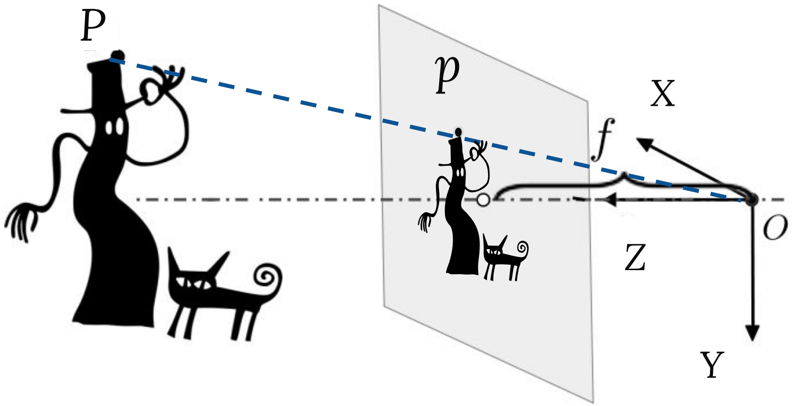
$$\frac{y}{Y} = -\frac{f}{Z} \qquad y = -\frac{fY}{Z}$$

$$\frac{x}{X} = -\frac{f}{Z} \qquad x = -\frac{fX}{Z}$$

Frontal Pinhole Model

$$\frac{y}{Y} = + \frac{f}{Z} \qquad y = + \frac{fY}{Z}$$

$$\frac{x}{X} = + \frac{f}{Z} \qquad x = + \frac{fX}{Z}$$

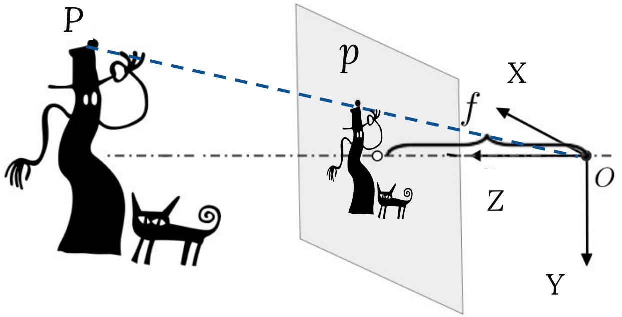


$$P = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \longrightarrow p = \begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$$

~~Frontal~~ Pinhole Model

Cartesian coordinates: $p = \begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$

Homogeneous coordinates: $\widetilde{P} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$ $\quad \widetilde{p} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} fX \\ fY \\ Z \end{bmatrix}$



$$Z \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{K_f} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\Pi_o} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Image point

3D point

Standard or canonical projection matrix

Pinhole Model

Since Z is usually unknown, it may be represented by an arbitrary positive scalar.

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Attention: P (X,Y,Z) are coordinates represented in the Camera Frame.

We usually know the coordinates of 3D points in the World Frame.
The Camera Frame and the World Frame are related by a rotation and translation.

Pinhole Model

Since Z is usually unknown, it may be represented by an arbitrary positive scalar.

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

Attention: $P\,(X,Y,Z)$ are coordinates represented in the Camera Frame.

We usually know the coordinates of 3D points in the World Frame.
The Camera Frame and the World Frame are related by a rotation and translation.

Extrinsic Parameters

Pinhole Model

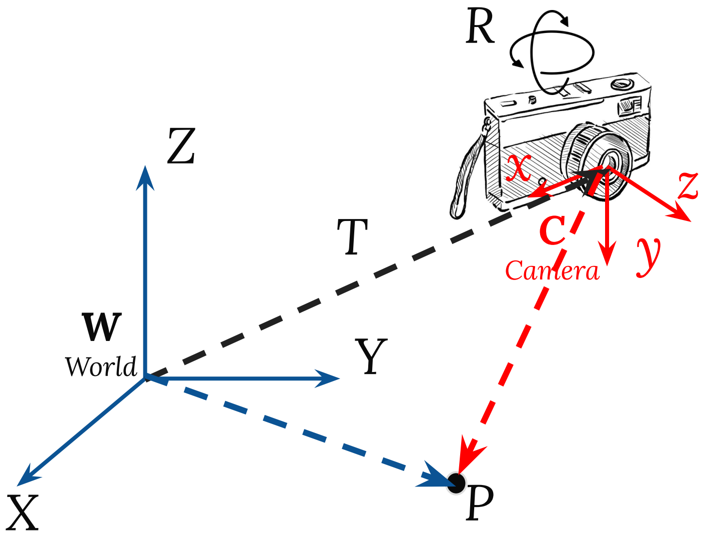
We already have performed two of the three transformations involved in image formation when using a pinhole camera:

- Coordinate transformations between the camera frame and the world frame;
- Projection of 3D coordinates onto a 2D image.

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\text{Projects 3D coordinates onto 2D coordinates losing depth information}} \underbrace{\begin{bmatrix} R & T \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}}_{\text{Converts from world to camera frame}}$$

Important to Notice

From the previous class...



If you are using R and T to represent the initial rotation and translation applied to the camera in the world frame, note that they **are not directly the Extrinsic Parameters,** since they convert from the Camera Frame to the World Frame.

$$\widetilde{P}_w = \begin{bmatrix} R & T \\ \mathbf{0} & 1 \end{bmatrix} .\widetilde{P}_c$$

In this case, the Extrinsic Parameters will actually be:

$$\widetilde{P}_c = \begin{bmatrix} R & T \\ \mathbf{0} & 1 \end{bmatrix}^{-1} .\widetilde{P}_w$$

02

Calibration Matrix

Pinhole Model

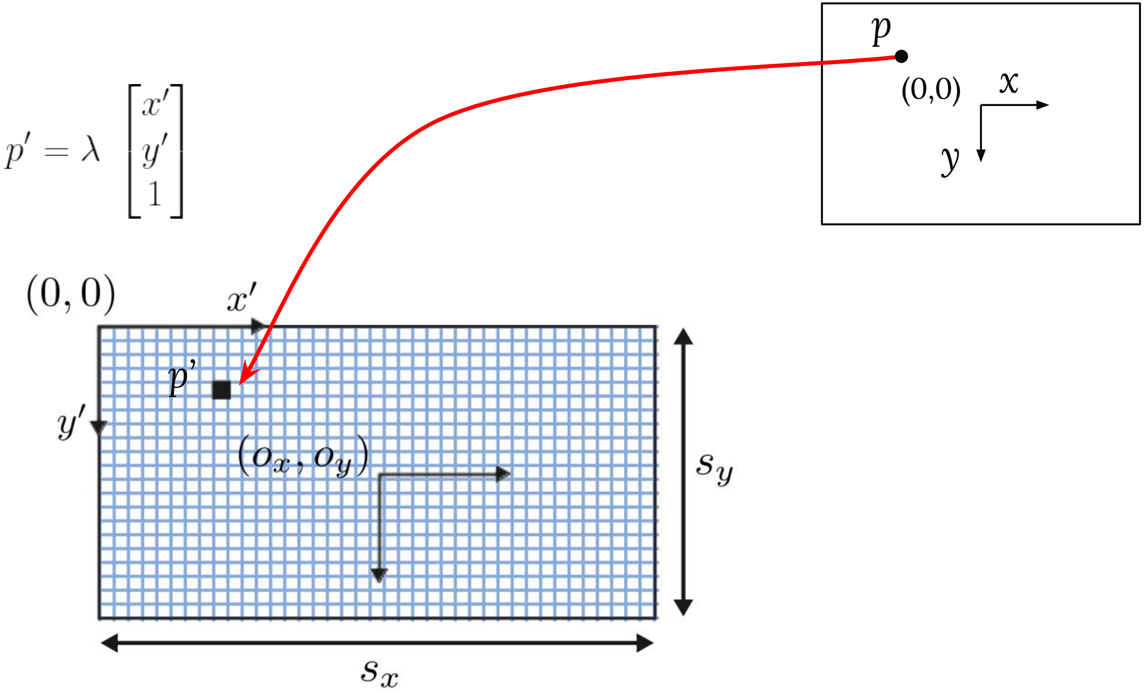
Points in images are given in pixels, not in centimeters or millimeters.

$$p = \lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Metric Coordinates

Pixel Coordinates

$$p' = \lambda \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}$$

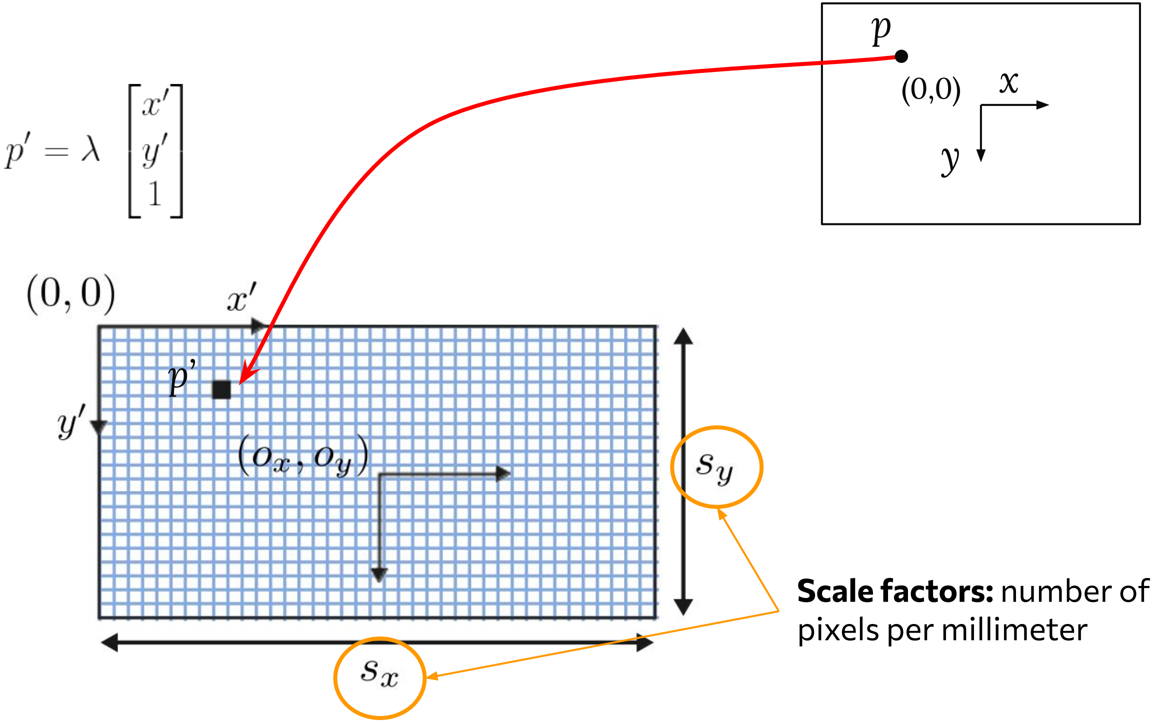


Pinhole Model

Points in images are given in pixels, not in centimeters or millimeters.

Pixel Coordinates $p' = \lambda \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}$

$p = \lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$ Metric Coordinates



Scale factors: number of pixels per millimeter

Pinhole Model

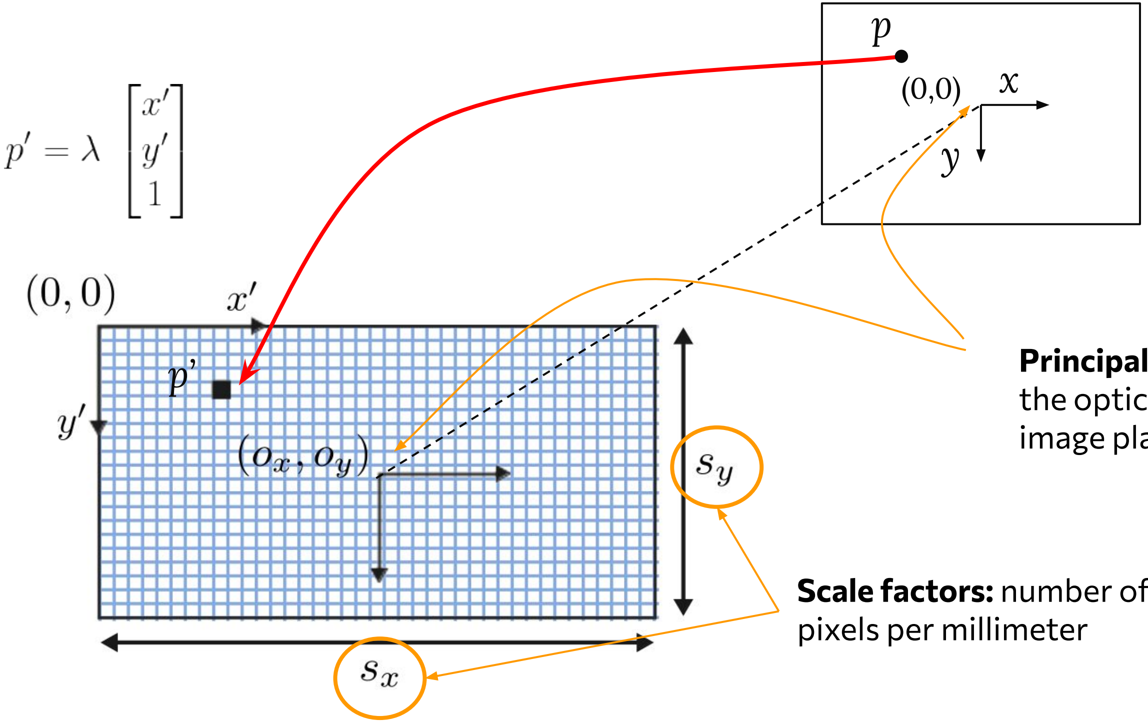
Points in images are given in pixels, not in centimeters or millimeters.

Pixel Coordinates

$$p' = \lambda \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}$$

$$p = \lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Metric Coordinates



Principal point: where the optical axis meets the image plane

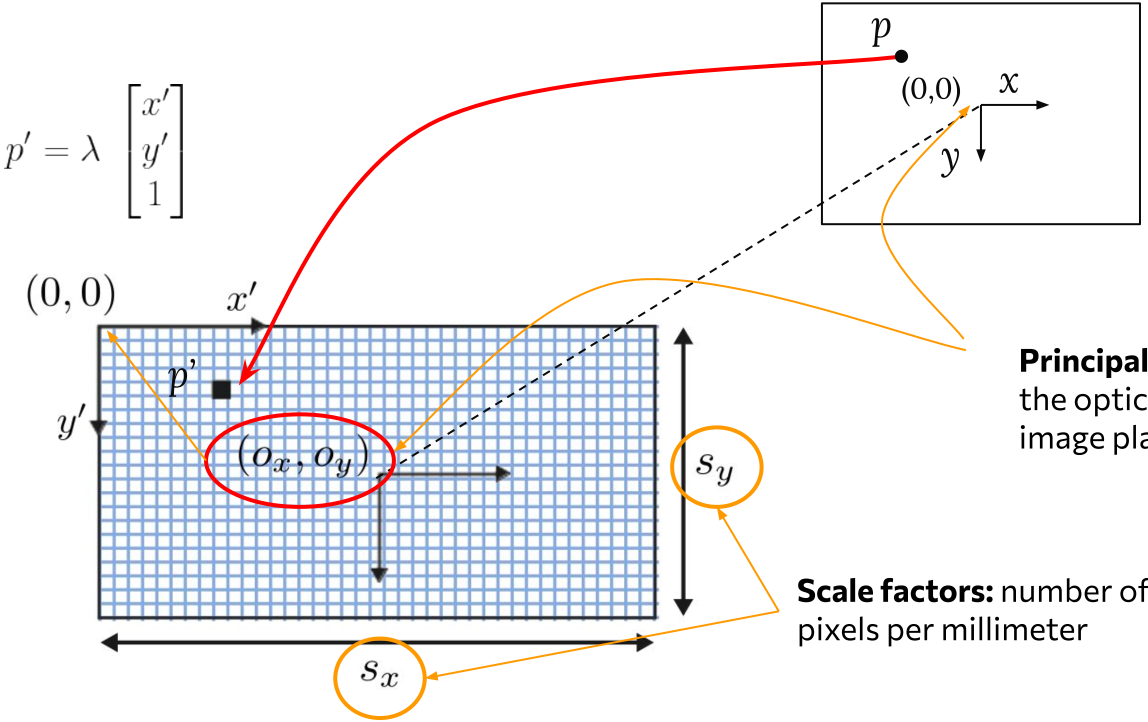
Scale factors: number of pixels per millimeter

Pinhole Model

Points in images are given in pixels, not in centimeters or millimeters.

Pixel Coordinates $p' = \lambda \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}$

Metric Coordinates $p = \lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$



Principal point: where the optical axis meets the image plane

Scale factors: number of pixels per millimeter

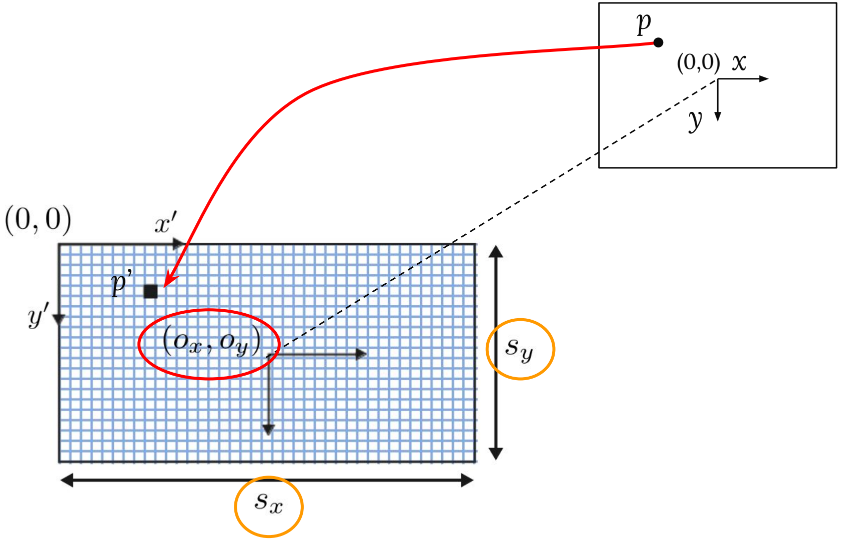
Pinhole Model

Points in images are given in pixels, not in centimeters or millimeters.

Scale factors　　Skew factor　　Principal point

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} s_x & s_\theta & o_x \\ 0 & s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}}_{K_s} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Metric Coordinates



Pixel Coordinates

Skew pixel

$s_\theta \longrightarrow$ 

Usually s_θ~0
Square pixels

Pinhole Model

$$\lambda \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & s_\theta & o_x \\ 0 & s_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

$$\lambda \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} fs_x & fs_\theta & o_x \\ 0 & fs_y & o_y \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Intrinsic Parameters Matrix or Calibration Matrix}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

Pinhole Model

$$\lambda \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} fs_x & fs_\theta & o_x \\ 0 & fs_y & o_y \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Intrinsic Parameters}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\text{Canonical Projection Matrix}} \underbrace{\begin{bmatrix} R & T \\ \mathbf{0} & 1 \end{bmatrix}}_{\text{Extrinsic Parameters}} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

Image point

3D point

$$\lambda p' = K \, \Pi_0 \, g(R, T) \, P_w$$

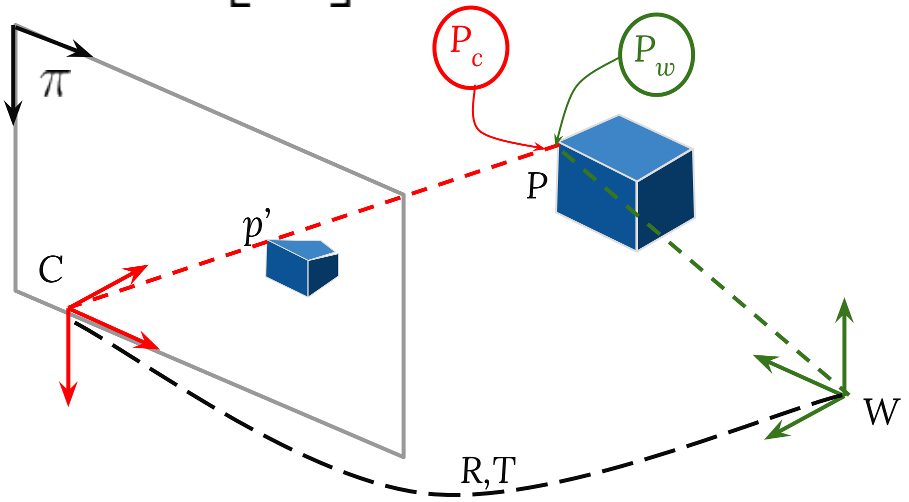
PS: fs_x and fs_y are called the focal distance in pixels at the horizontal and vertical directions

Pinhole Model

$$\lambda \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} fs_x & fs_\theta & o_x \\ 0 & fs_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

$$\lambda p' = \underbrace{K\, \Pi_0\, g(R,T)}\; P_w$$

$$\lambda p' = \Pi\, P_w$$



Pinhole Model

Considering just the extrinsic parameters, if we are not working in pixels:

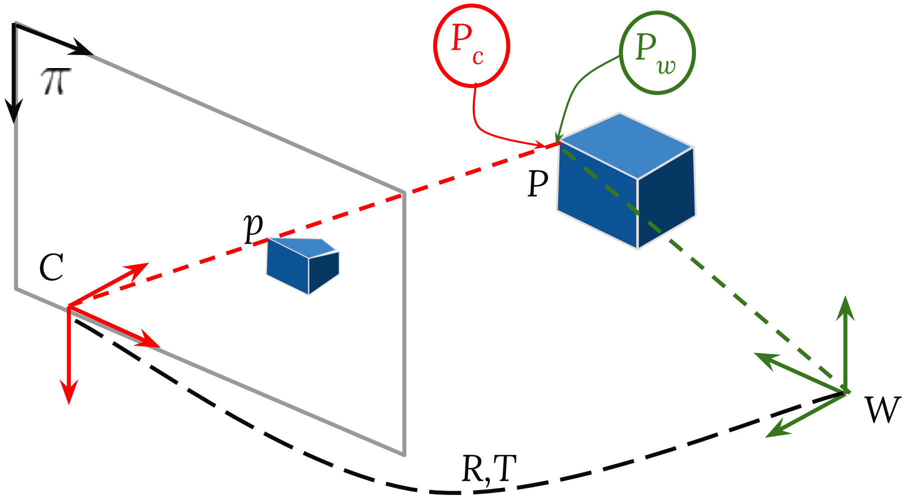
$$\lambda p = \Pi_0 \; g(R, T) \; P_w$$

$$\lambda p = \Pi \; P_w \;\; \rightarrow \boxed{\lambda p = [R, T]_{3\times4} \; P_w}$$

Considering the complete model, if we are working in pixels:

$$\lambda p' = K \; \Pi_0 \; g(R, T) \; P_w$$

$$\lambda p' = \Pi \; P_w \;\; \rightarrow \boxed{\lambda p' = [KR, KT]_{3\times4} \; P_w}$$

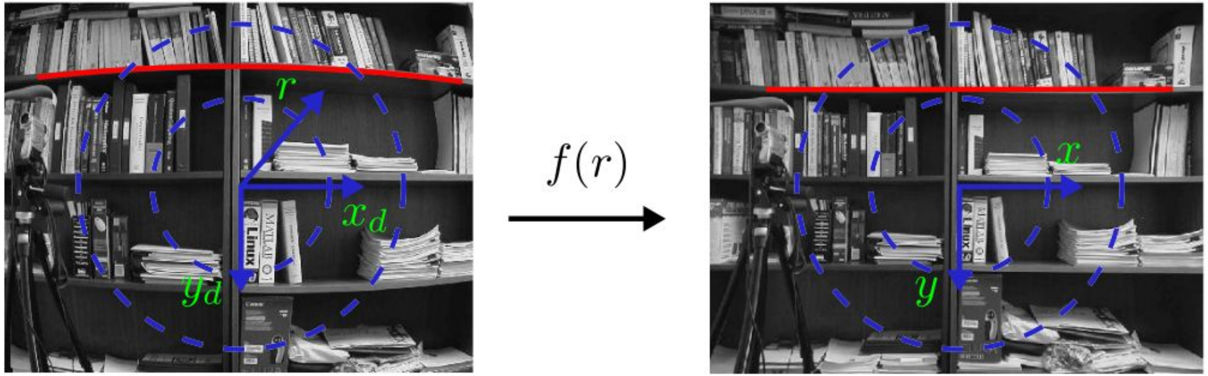


03

Radial Distortion

Radial Distortion

Non-linear Transformation along radial direction



Correction: straight lines

Corrected point $\rightarrow$ $p = c + f(r)(p_d - c), \quad \text{with} \quad r = ||p_d - c||$

Distorted point

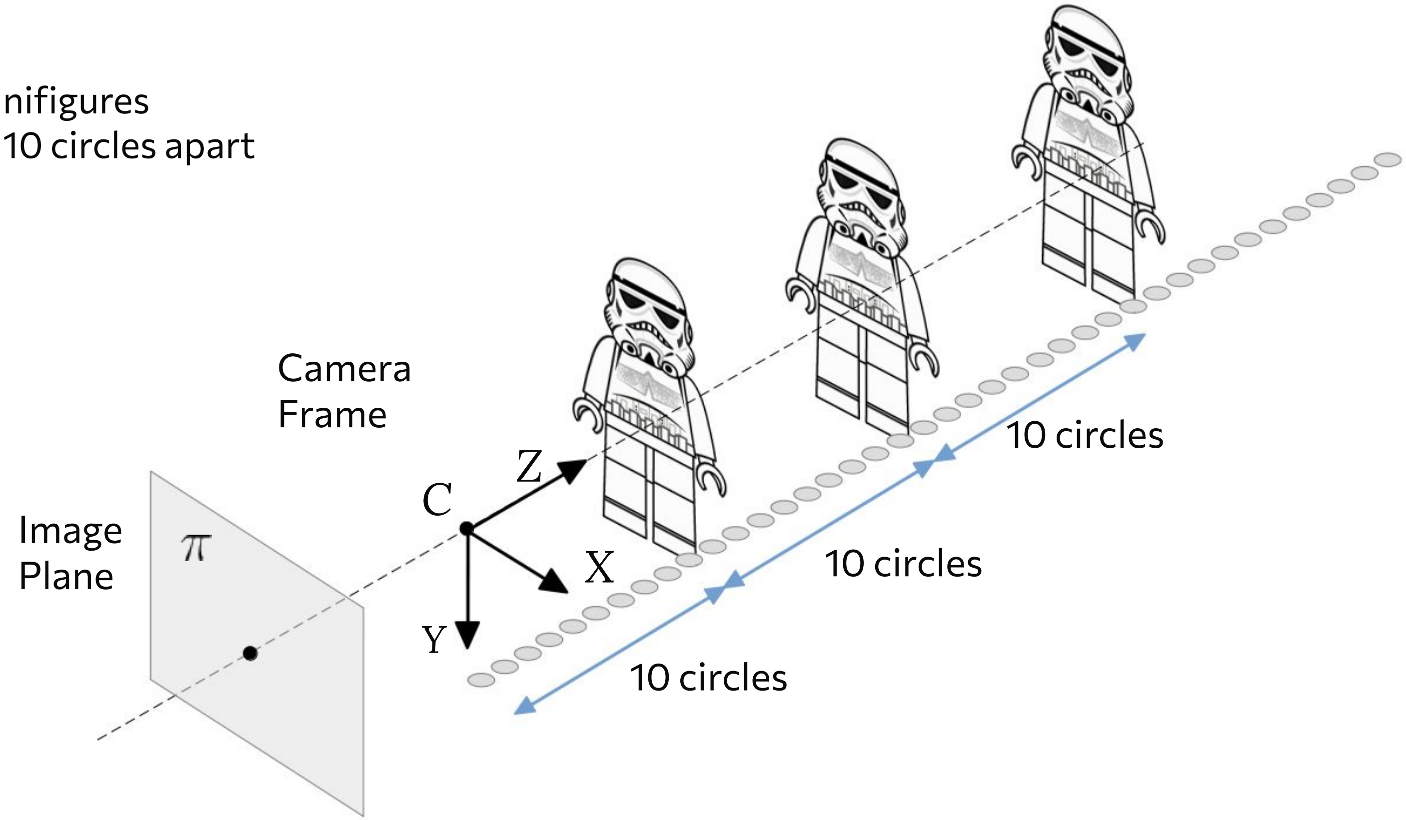
$$f(r) = 1 + a_1 r + a_2 r^2 + a_3 r^3 + a_4 r^4 + \cdots$$

04

Experiment

Experiment

- 3 Lego Minifigures
- Distance: 10 circles apart



Experiment

- 3 Lego Minifigures
- Same dimensions

Experiment

- 3 Lego Minifigures
- Same dimensions
- 10 circles apart to each other

Experiment

Basic Perspective Projection

$$x = \frac{fX}{Z} \qquad y = \frac{fY}{Z}$$

Experiment

Basic Perspective Projection

All 3 minifigures have the same dimensions, but different depth (distance) from the camera.

$$x = \frac{fX}{Z} \qquad y = \frac{fY}{Z}$$

Experiment

Basic Perspective Projection

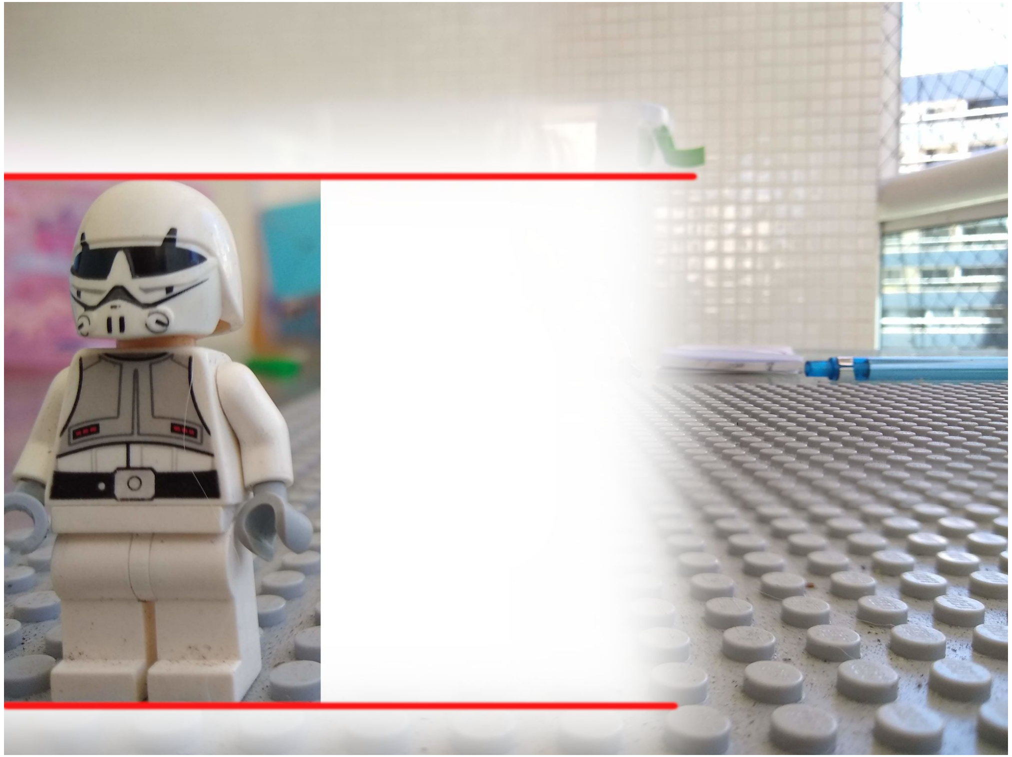
All 3 minifigures have the same dimensions, but different depth (distance) from the camera.

$$x = \frac{fX}{Z} \qquad y = \frac{fY}{Z}$$

Inversely proportional to the depth.

Experiment

$$x = \frac{fX}{Z} \qquad y = \frac{fY}{Z}$$

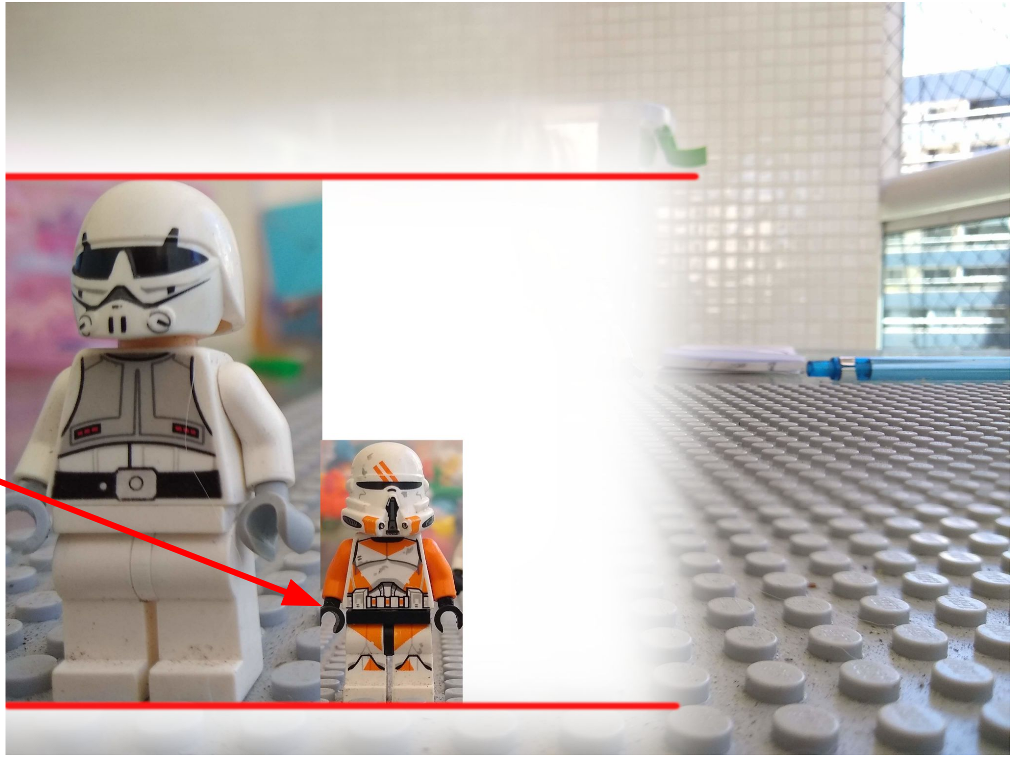


Experiment

$$x = \frac{fX}{Z} \qquad y = \frac{fY}{Z}$$

Twice the distance

$$x' = \frac{fX}{2Z} \qquad y = \frac{fY}{2Z}$$



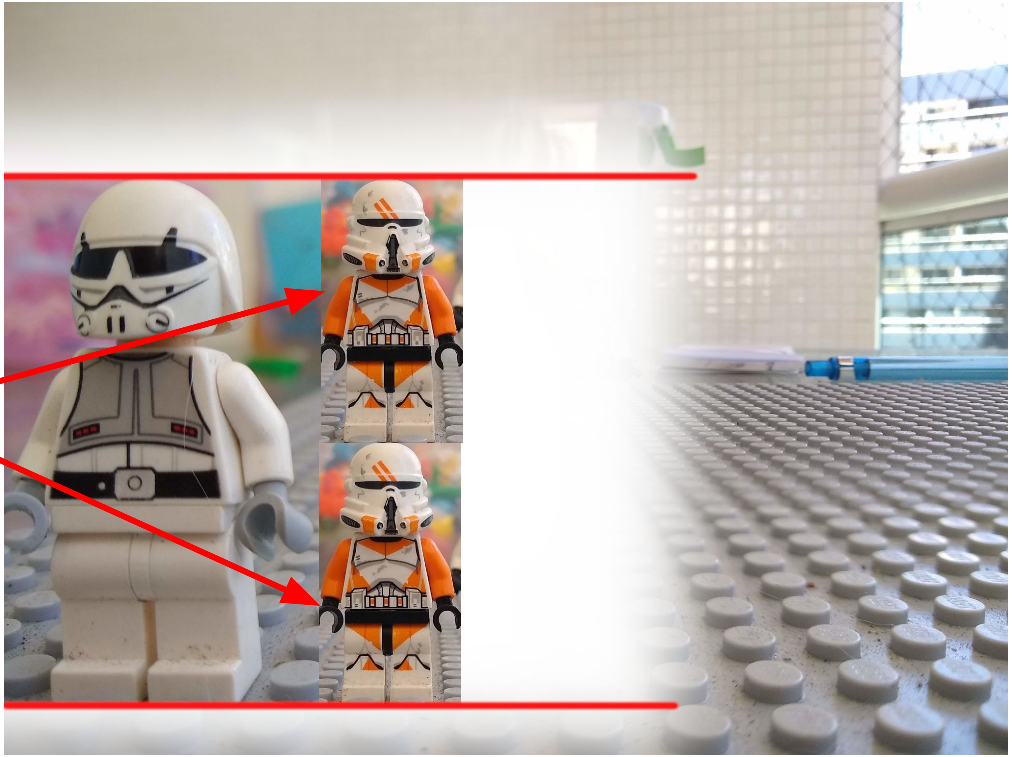
Experiment

$$x = \frac{fX}{Z} \qquad y = \frac{fY}{Z}$$

Twice the distance

Half the size

$$x' = \frac{fX}{2Z} \qquad y = \frac{fY}{2Z}$$



Experiment

$$x = \frac{fX}{Z} \qquad y = \frac{fY}{Z}$$

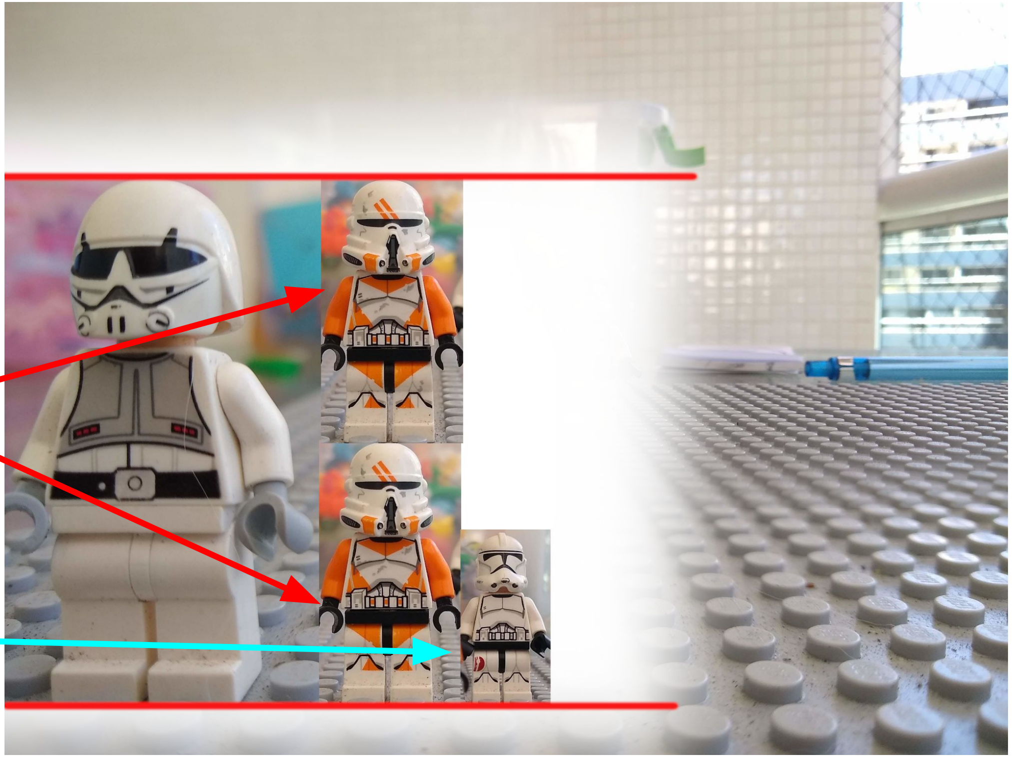
Twice the distance

$\boxed{\text{Half the size}}$

$$x' = \frac{fX}{2Z} \qquad y = \frac{fY}{2Z}$$

Triple the distance

$$x'' = \frac{fX}{3Z} \qquad y = \frac{fY}{3Z}$$



Experiment

$$x = \frac{fX}{Z} \qquad y = \frac{fY}{Z}$$

Twice the distance

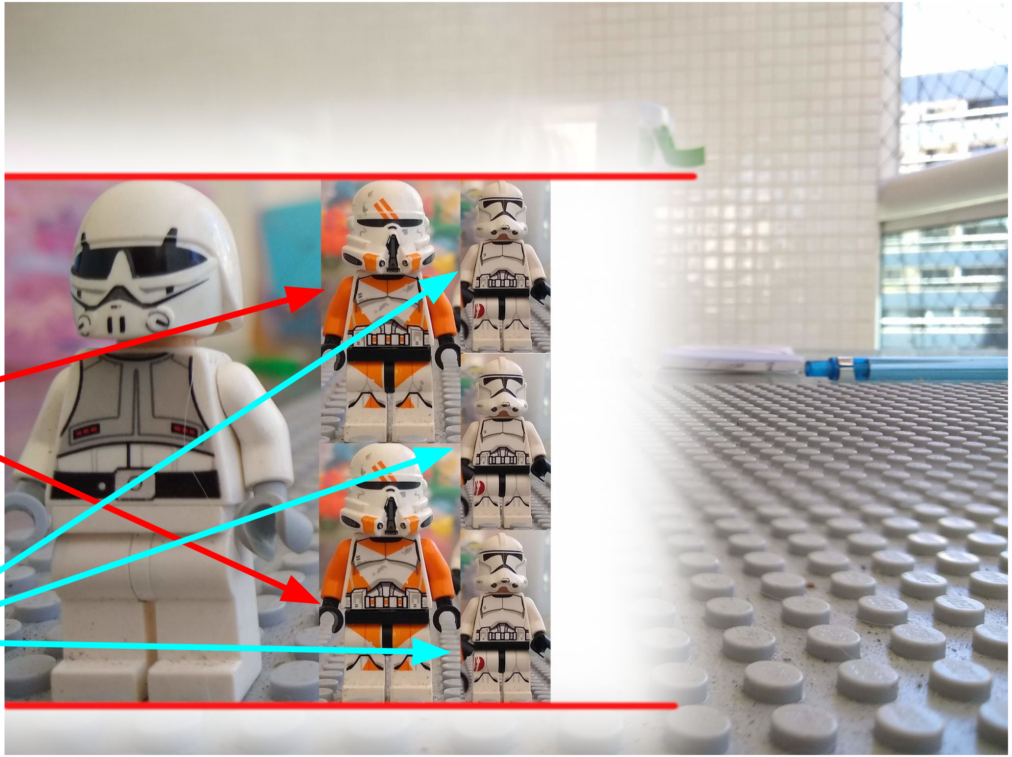
Half the size

$$x' = \frac{fX}{2Z} \qquad y = \frac{fY}{2Z}$$

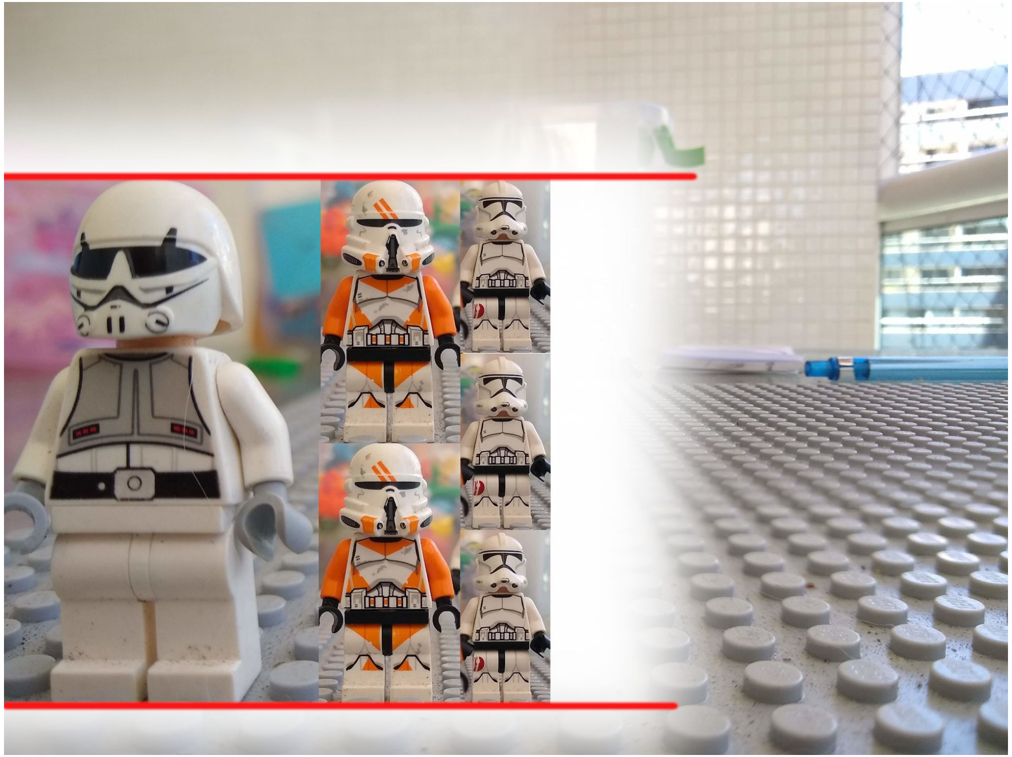
Triple the distance

Third the size

$$x'' = \frac{fX}{3Z} \qquad y = \frac{fY}{3Z}$$



Experiment



Credits

Yi Ma, Stefano Soatto, Jana Kosecka e S. Shankar Sastry. An Invitation to 3D Vision: From Images to Geometric Models.
Springer, ISBN 0387008934