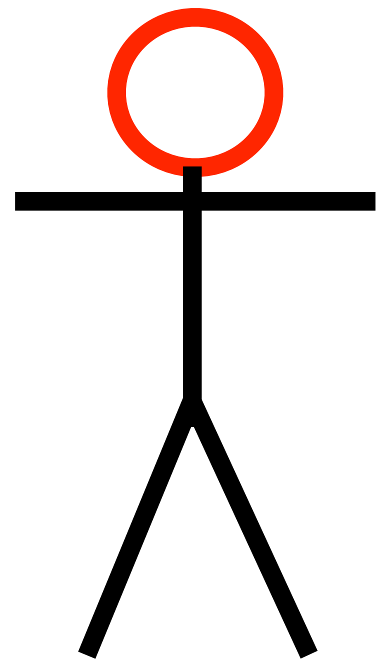
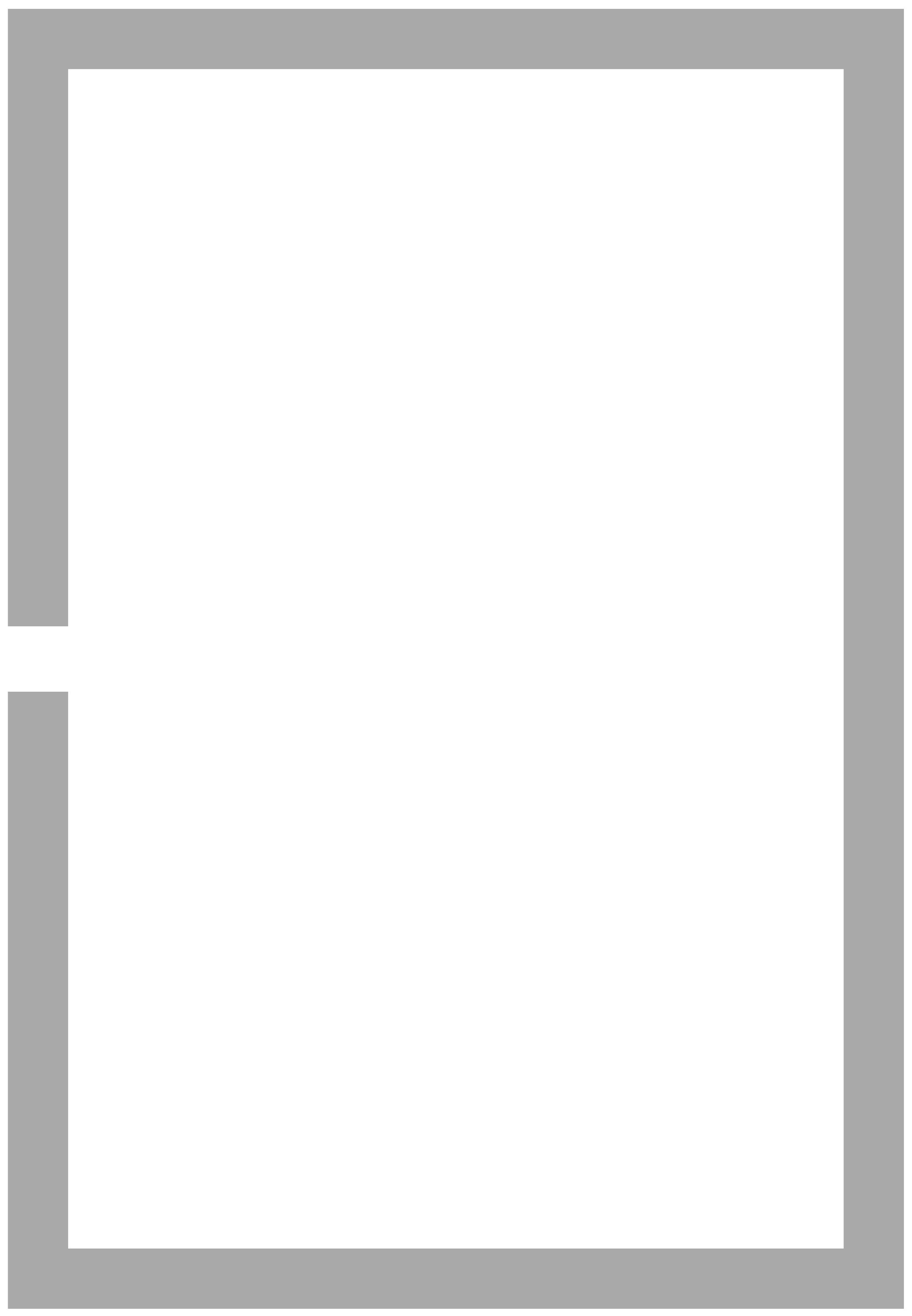
Camera and calibration

Karel Zimmermann

Camera model

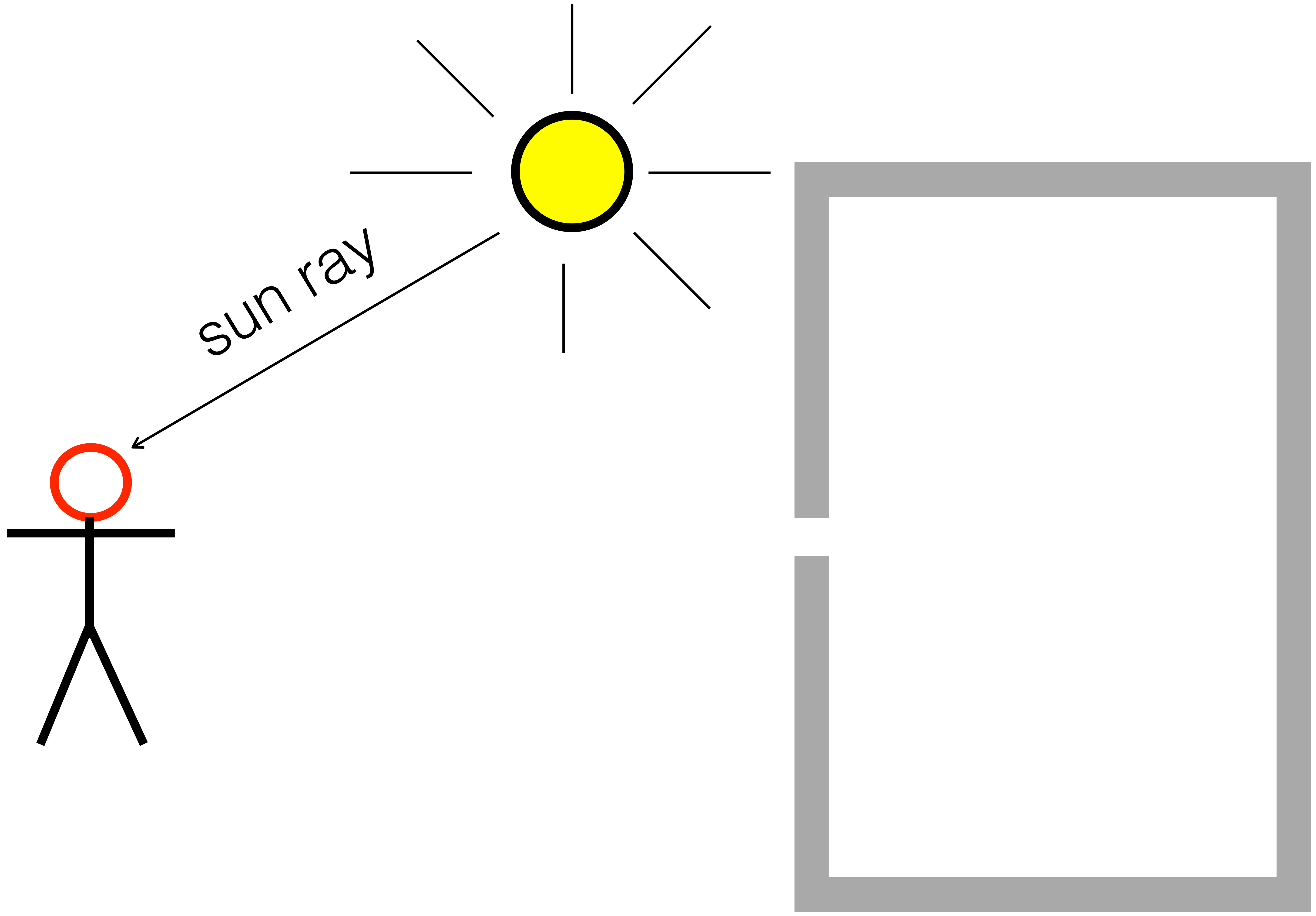


Object in front of camera

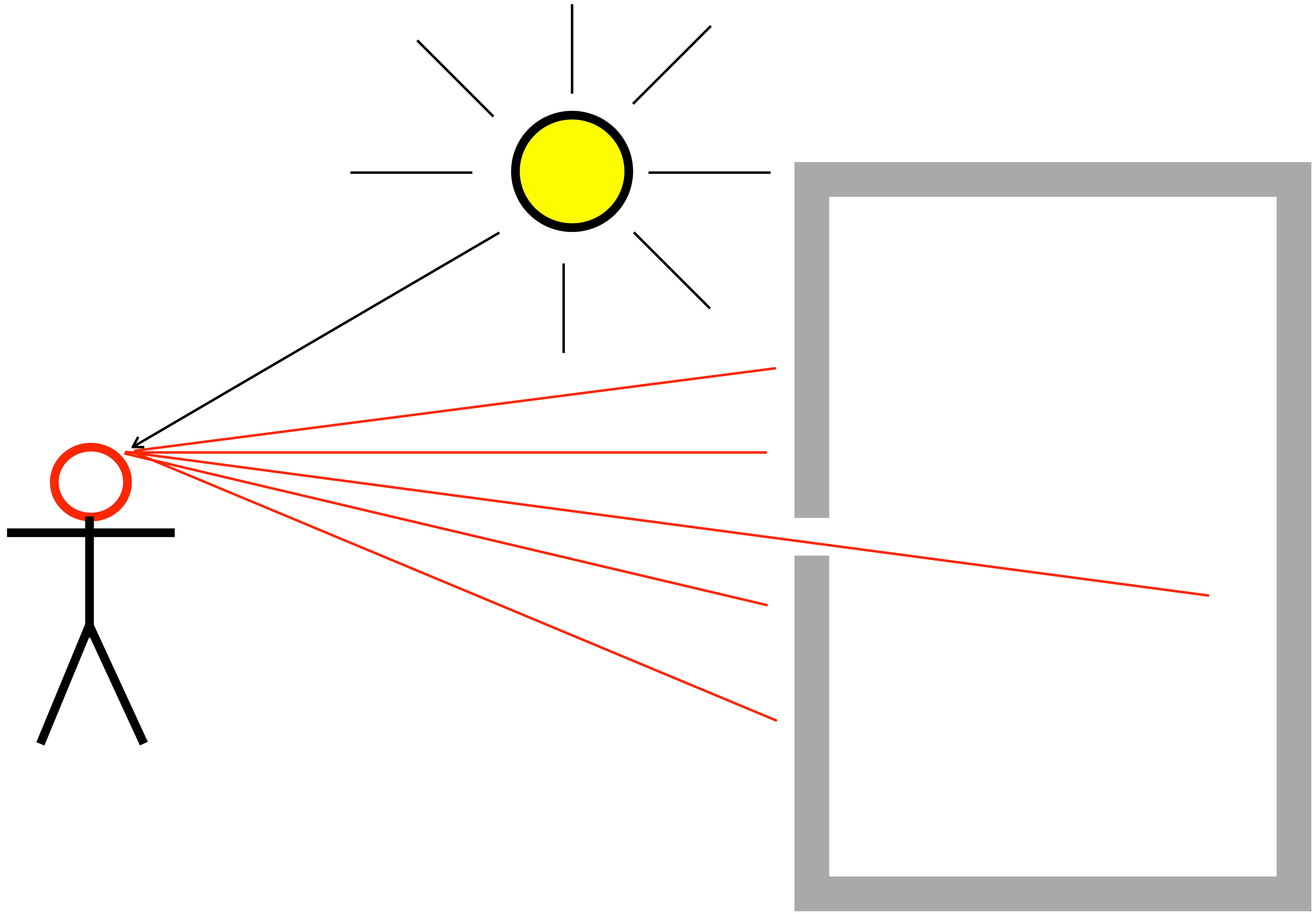


Pinhole camera model

Camera model

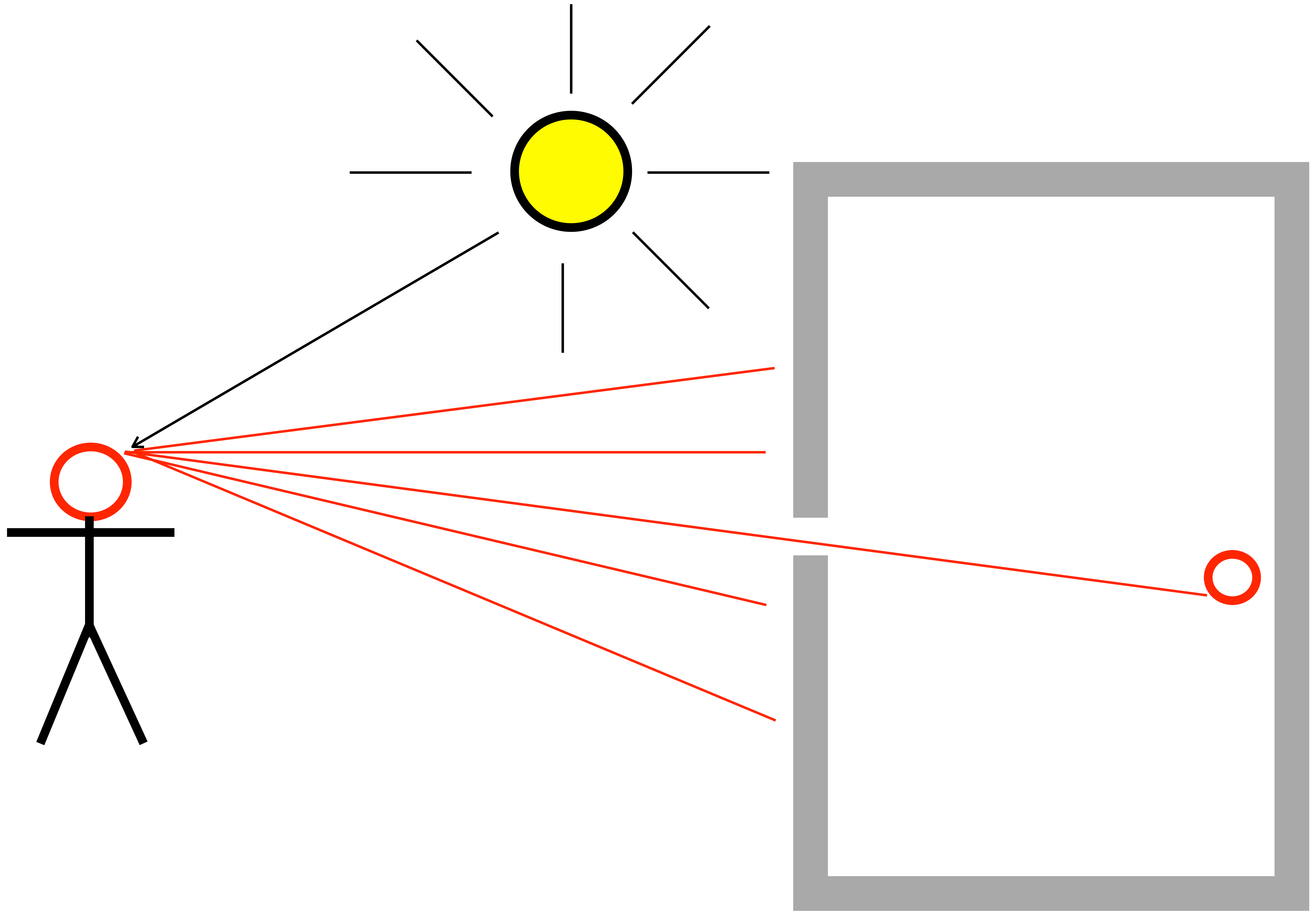


Camera model



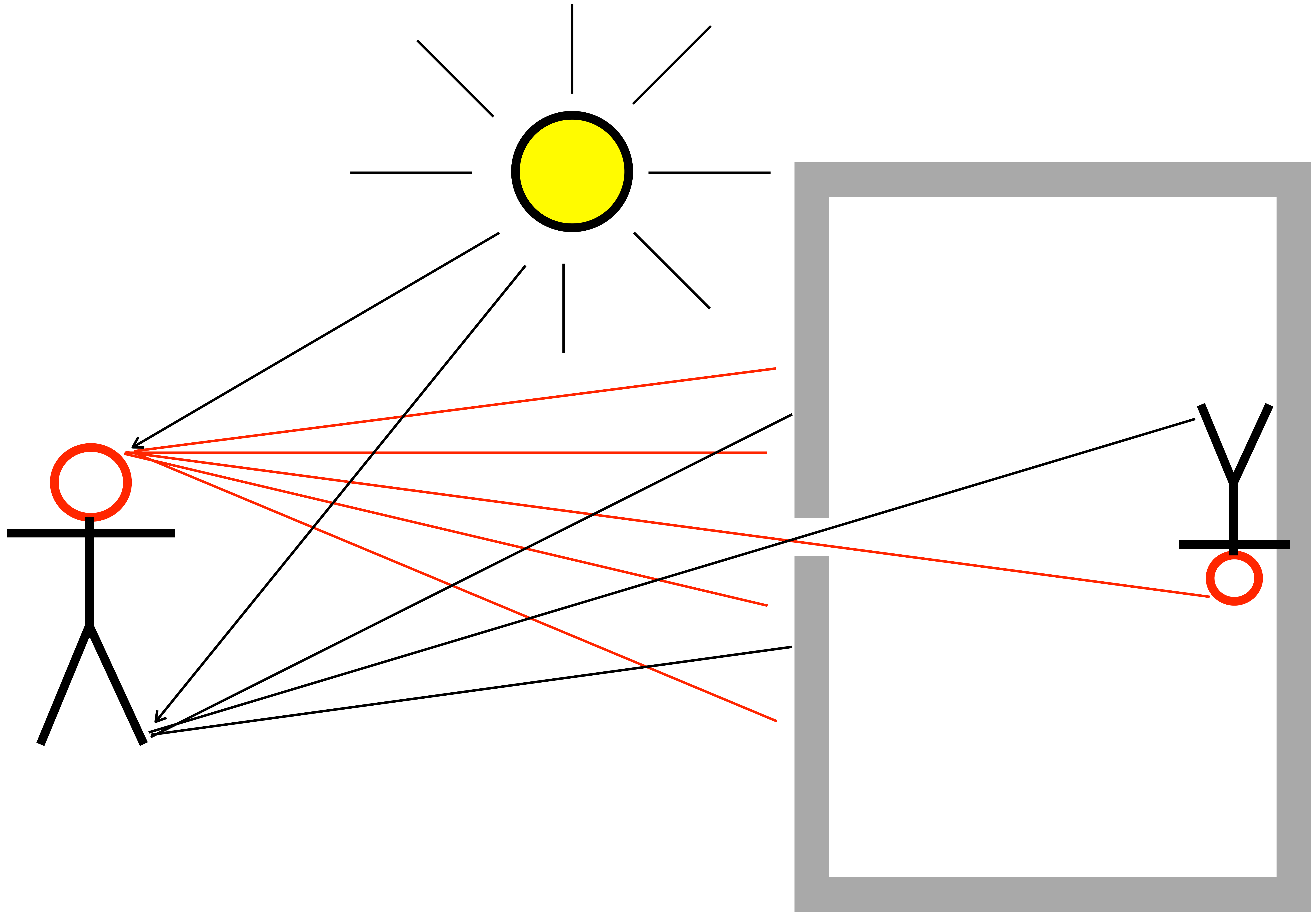
Sun ray is reflected from Lambertian surface in hemisphere

Camera model

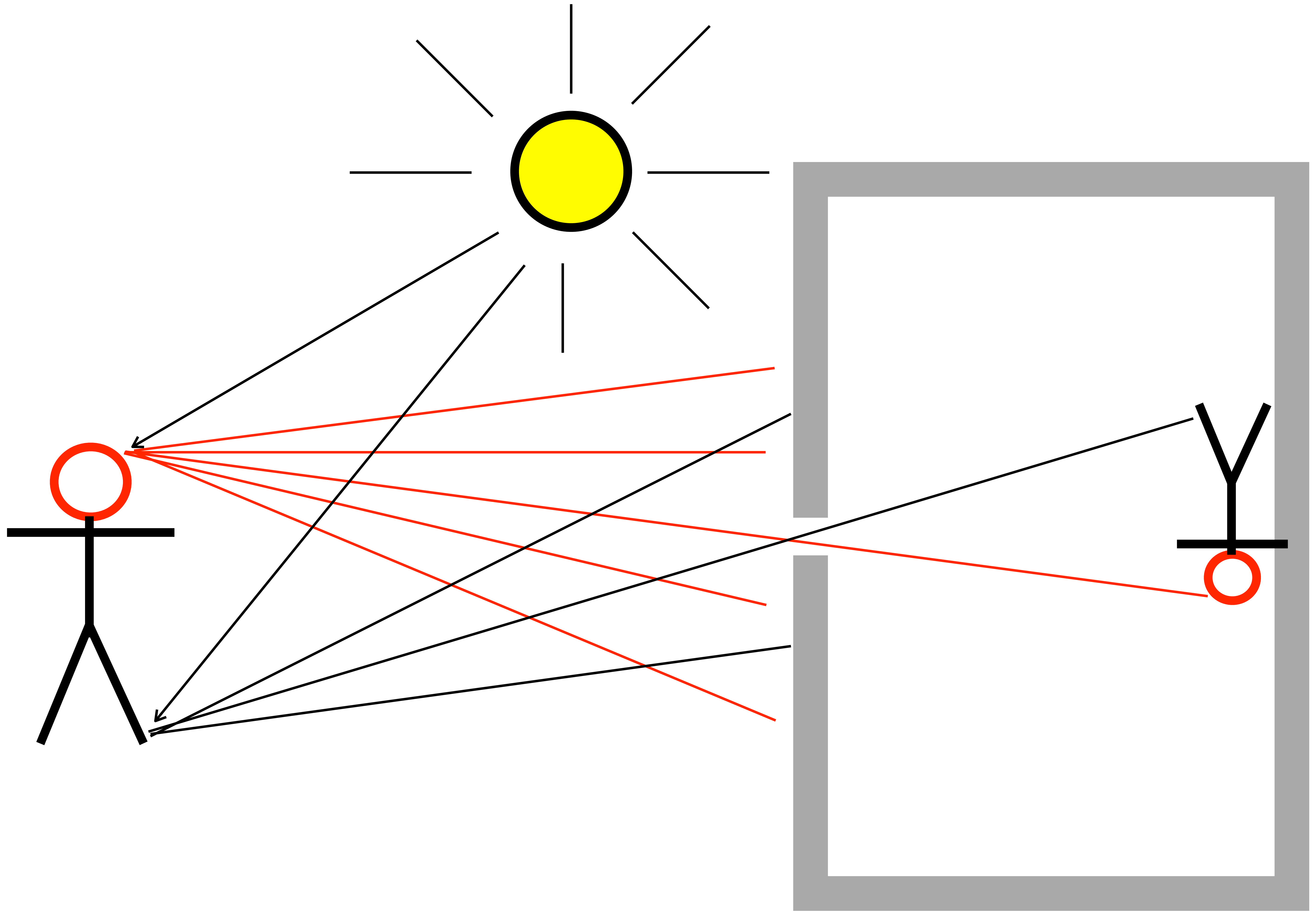


Reflected ray (red) forms inverted image of the object

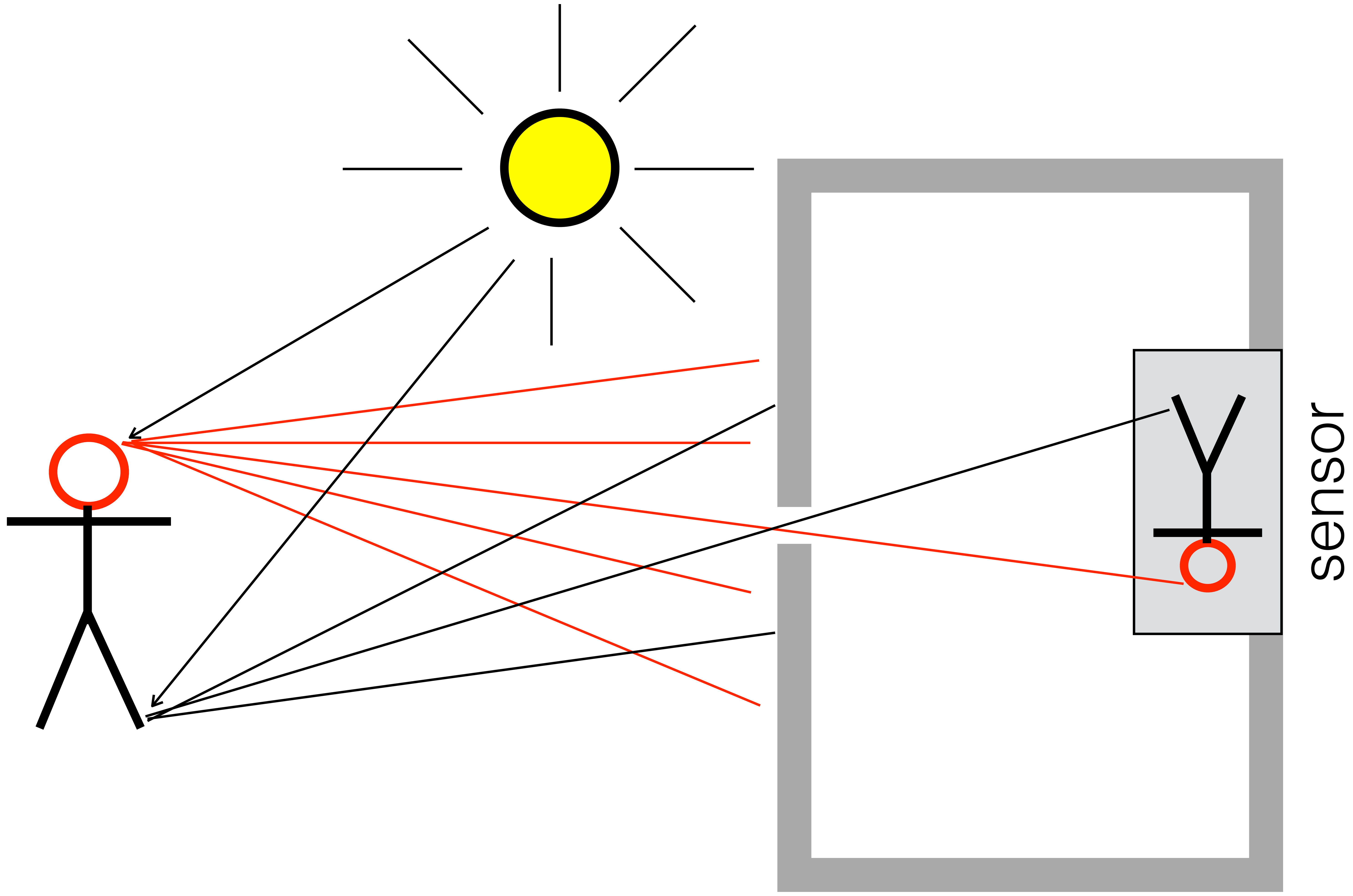
Camera model



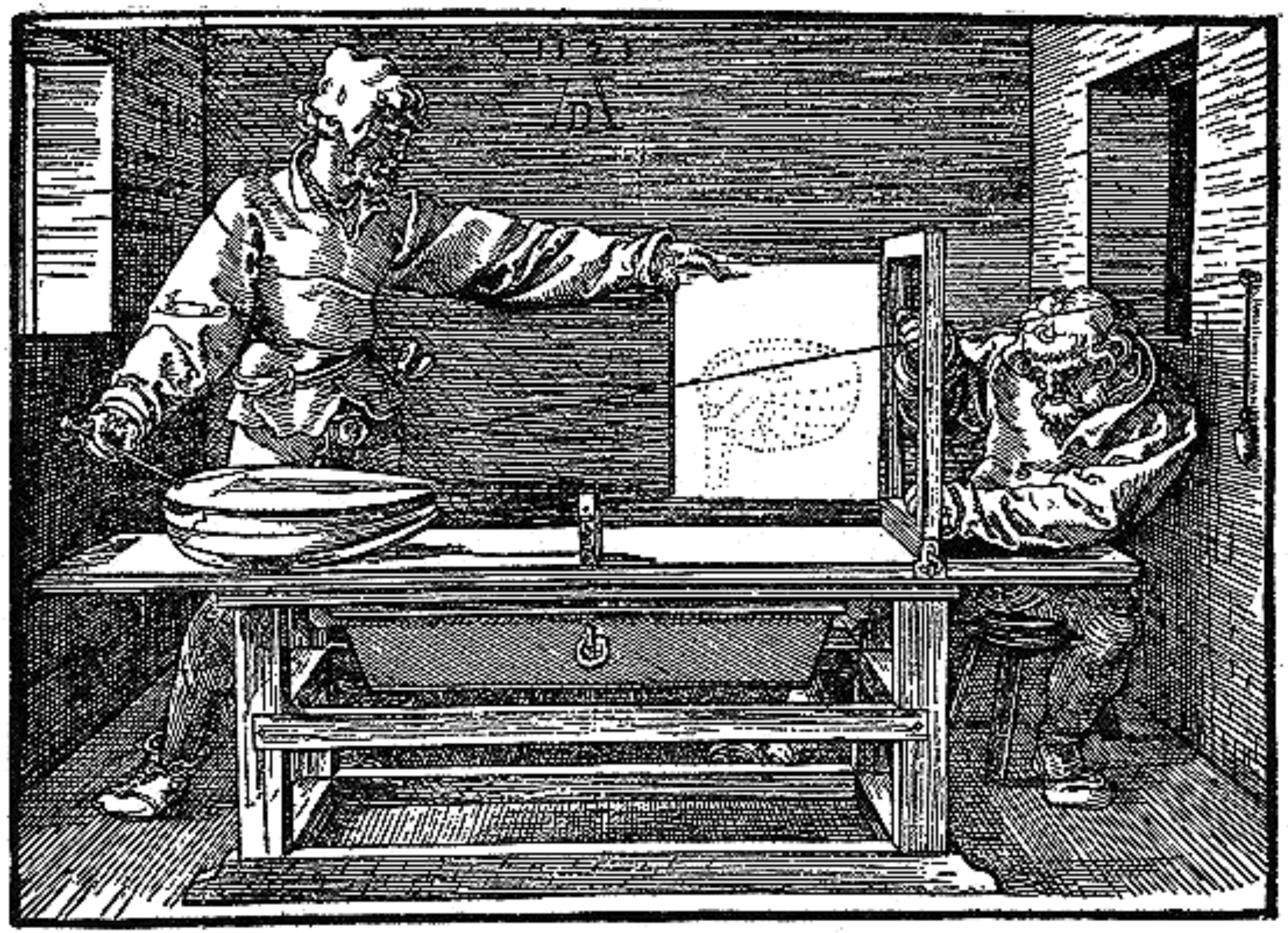
Camera model



Camera model

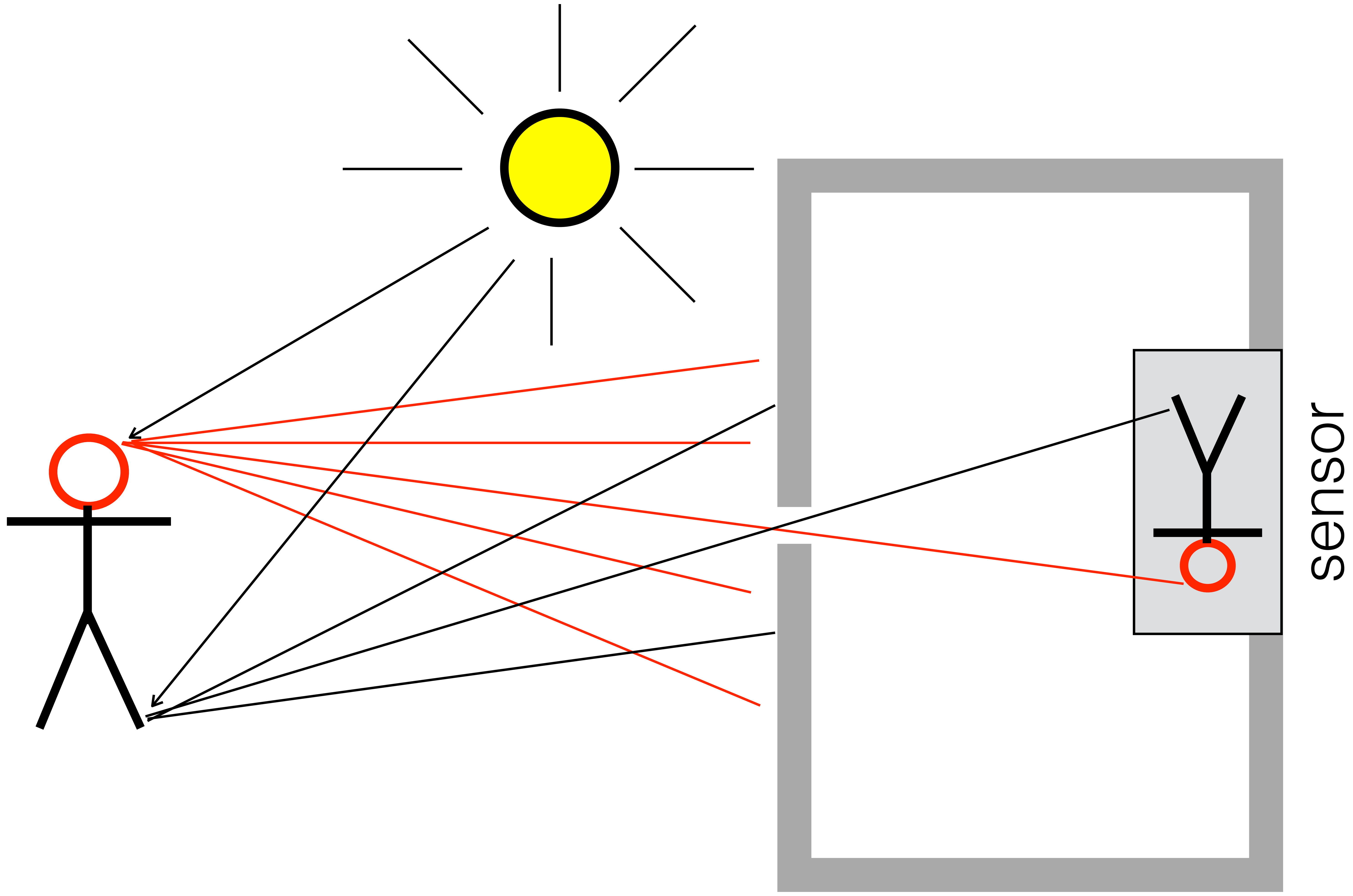


Projection of 3D point in \camera on image plane



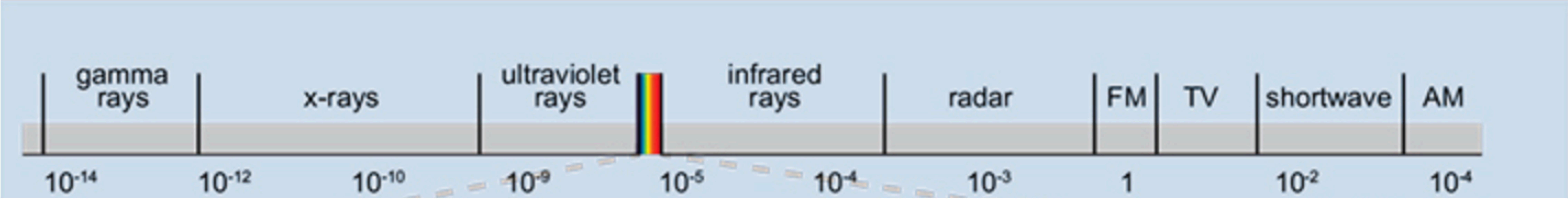
Albrecht Durer (1545), Hitachi Viewmuseum

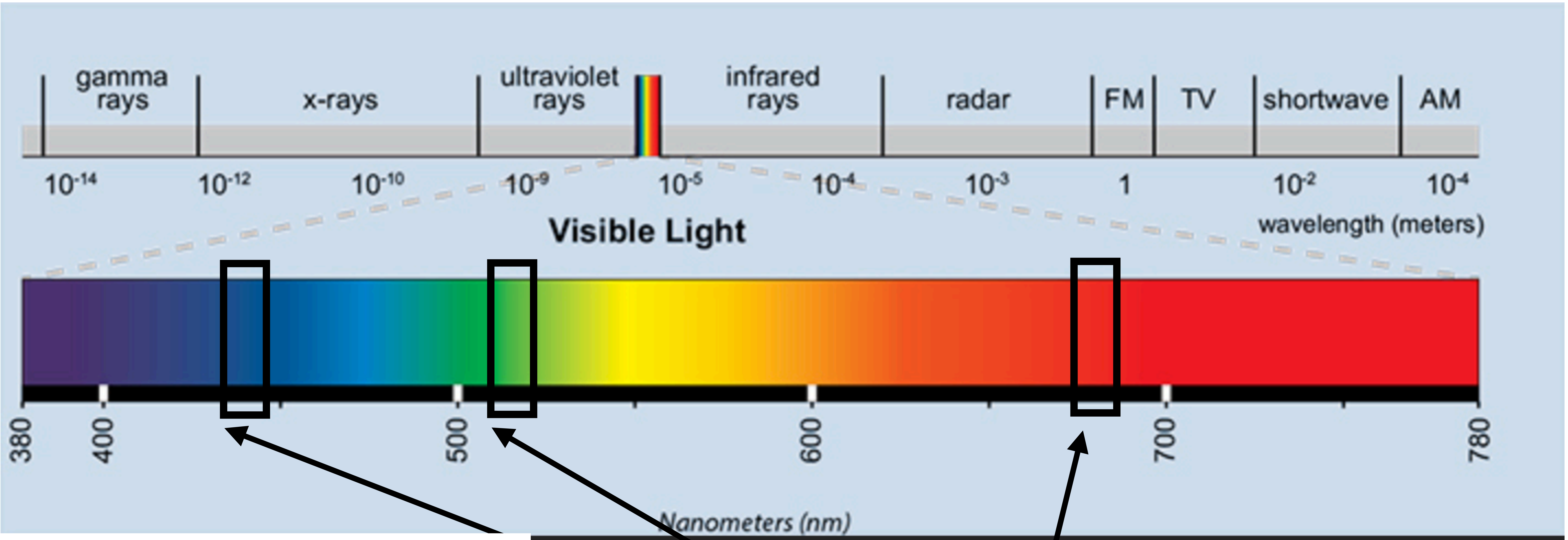
Camera model

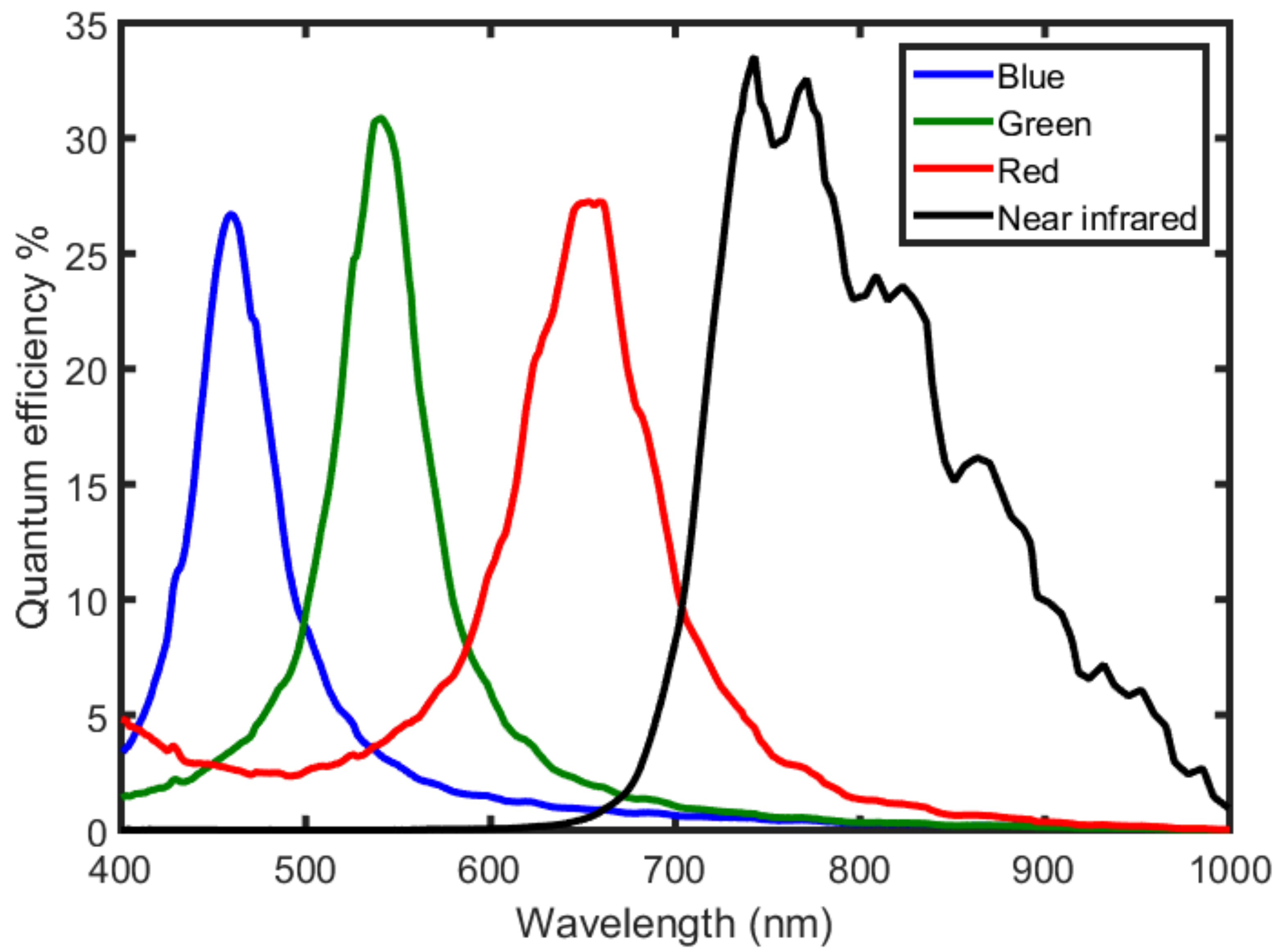


Resulting image is represented as array

Values corresponds to the signal captured by fotosensitive elements

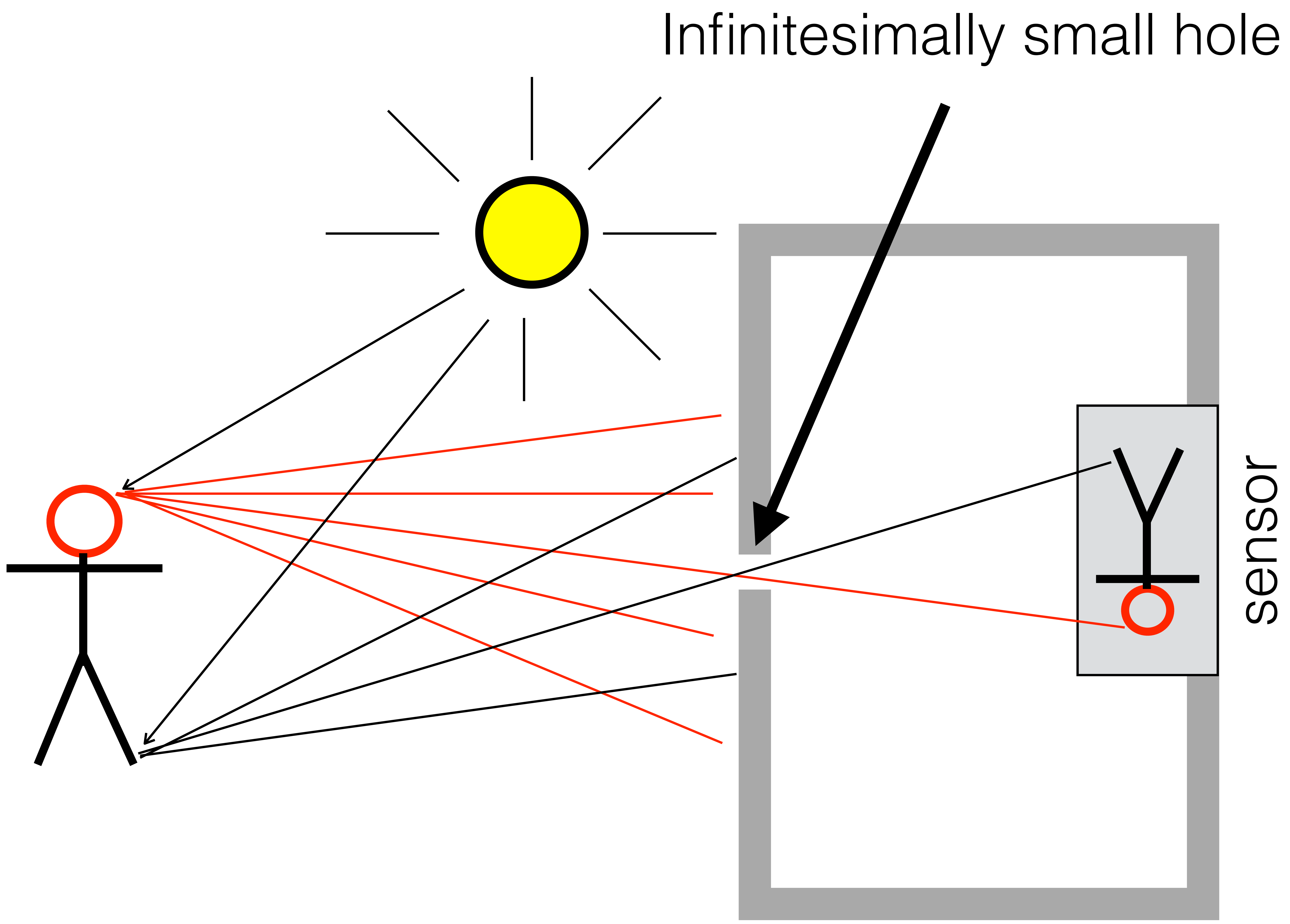

gamma rays
x-rays
ultraviolet rays
infrared rays
radar
FM
TV
shortwave
AM
10-14
10-12
10-10
10-9
10-5
10-4
10-3
1
10-2
10-4



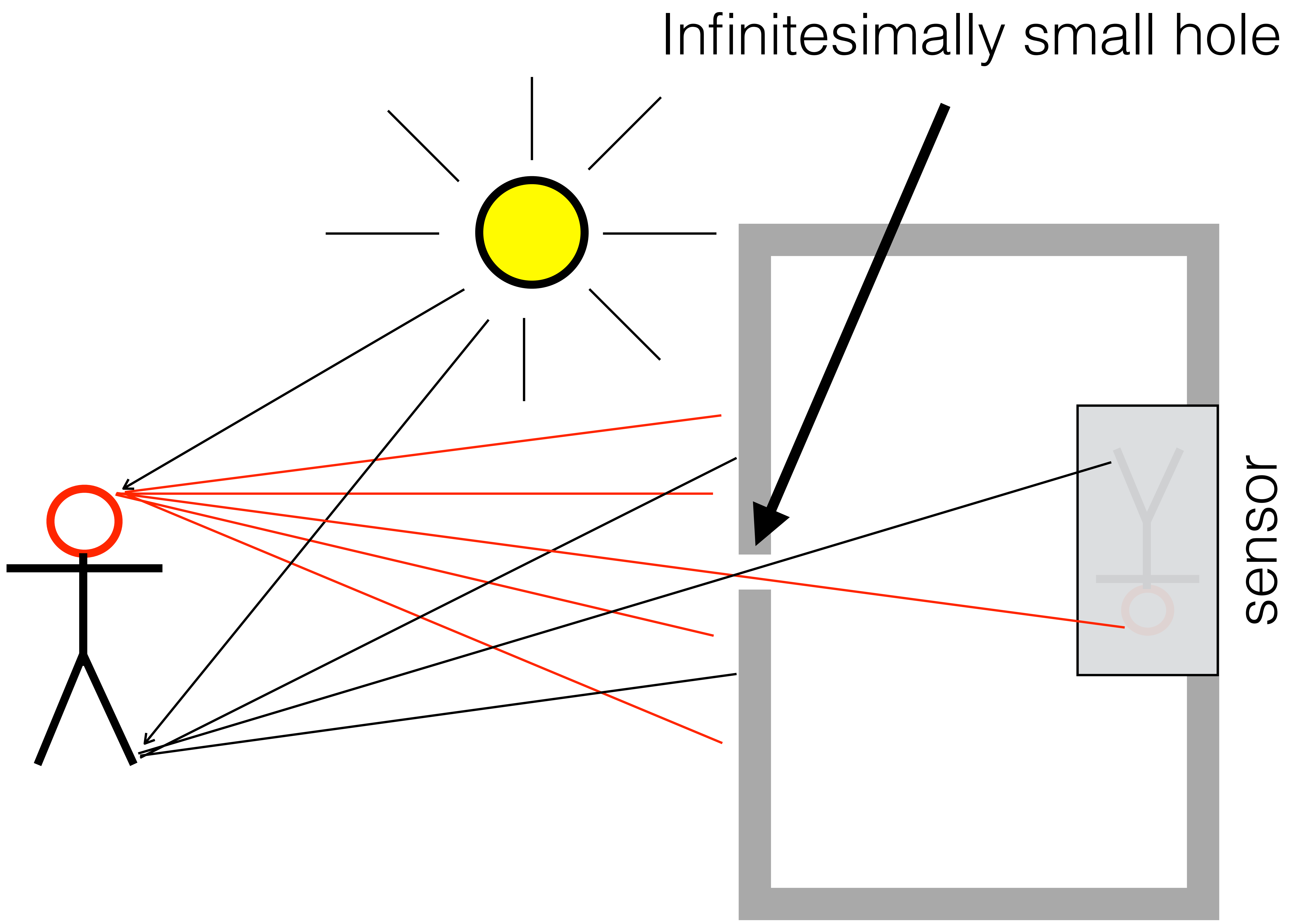


Each pixel contains 3 values:RGB

Camera model



Camera model



Light energy from rays traversed through pinhole is small

If the capturing time is short, the resulting image is underilluminated

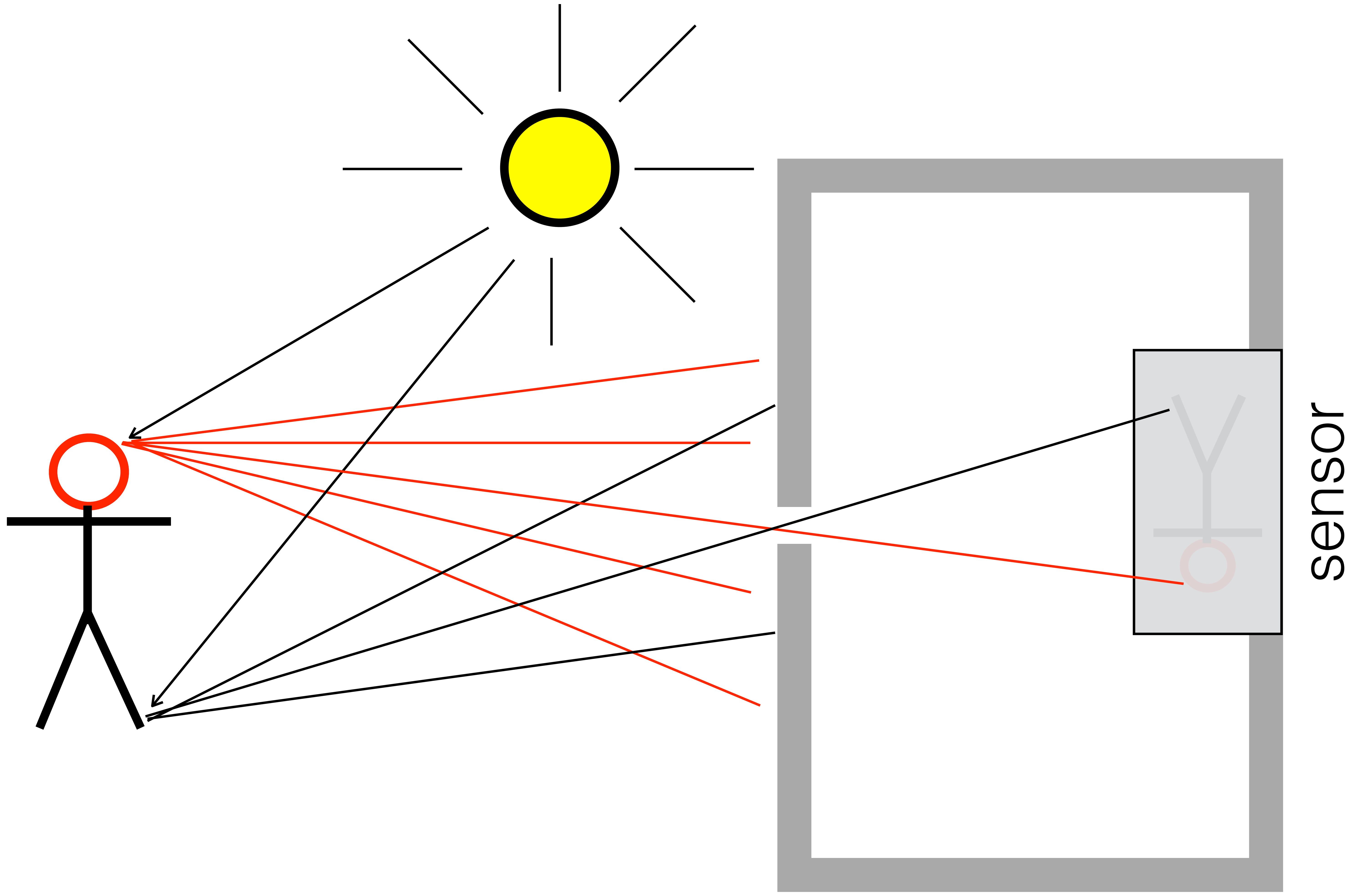
Postprocessing amplifies the noise

Postprocessing amplifies the noise

If signal is integrated sufficiently long, the result is better

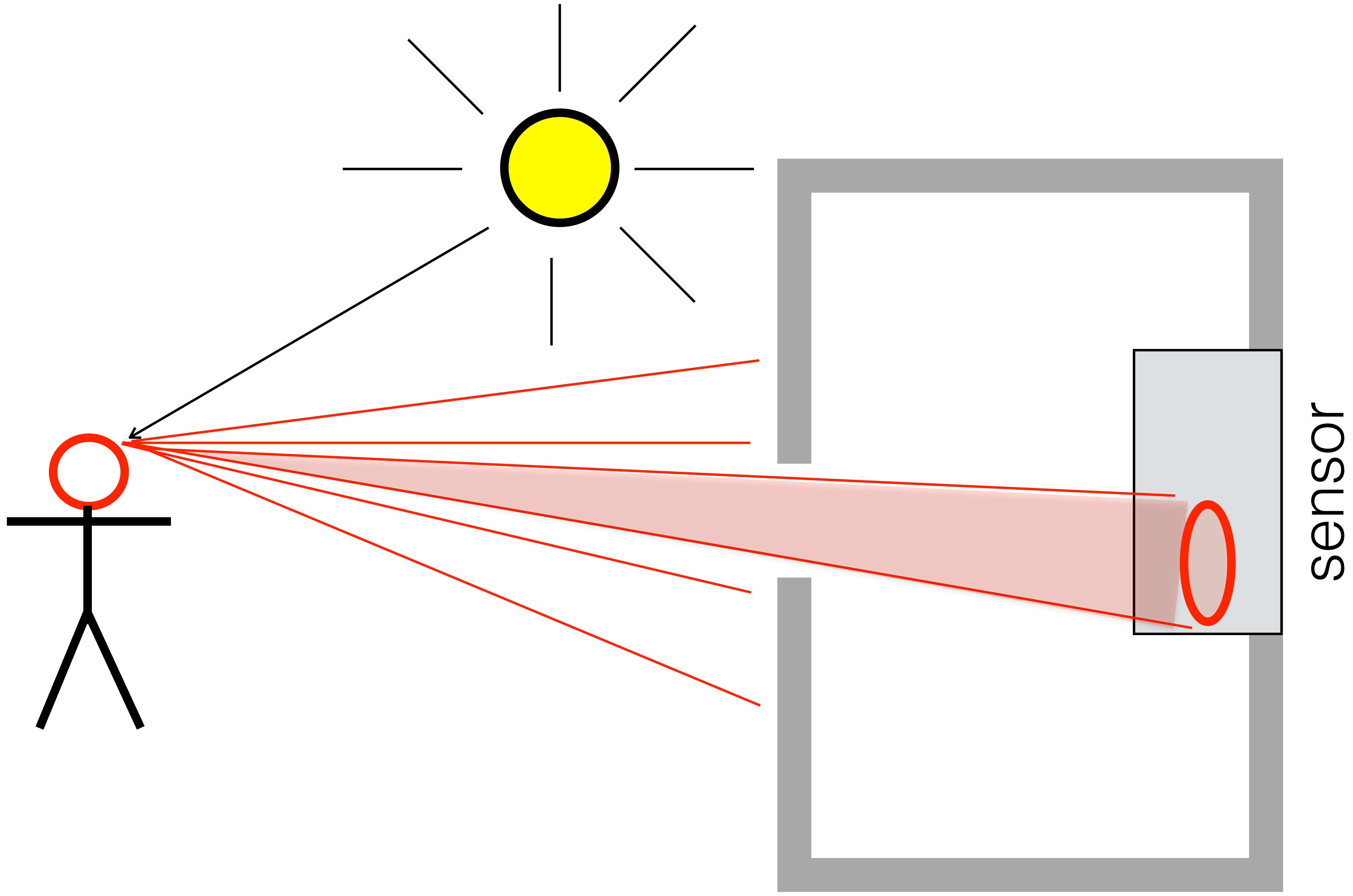

If the object or cameras moves during the capturing, the result is blurred

Camera model



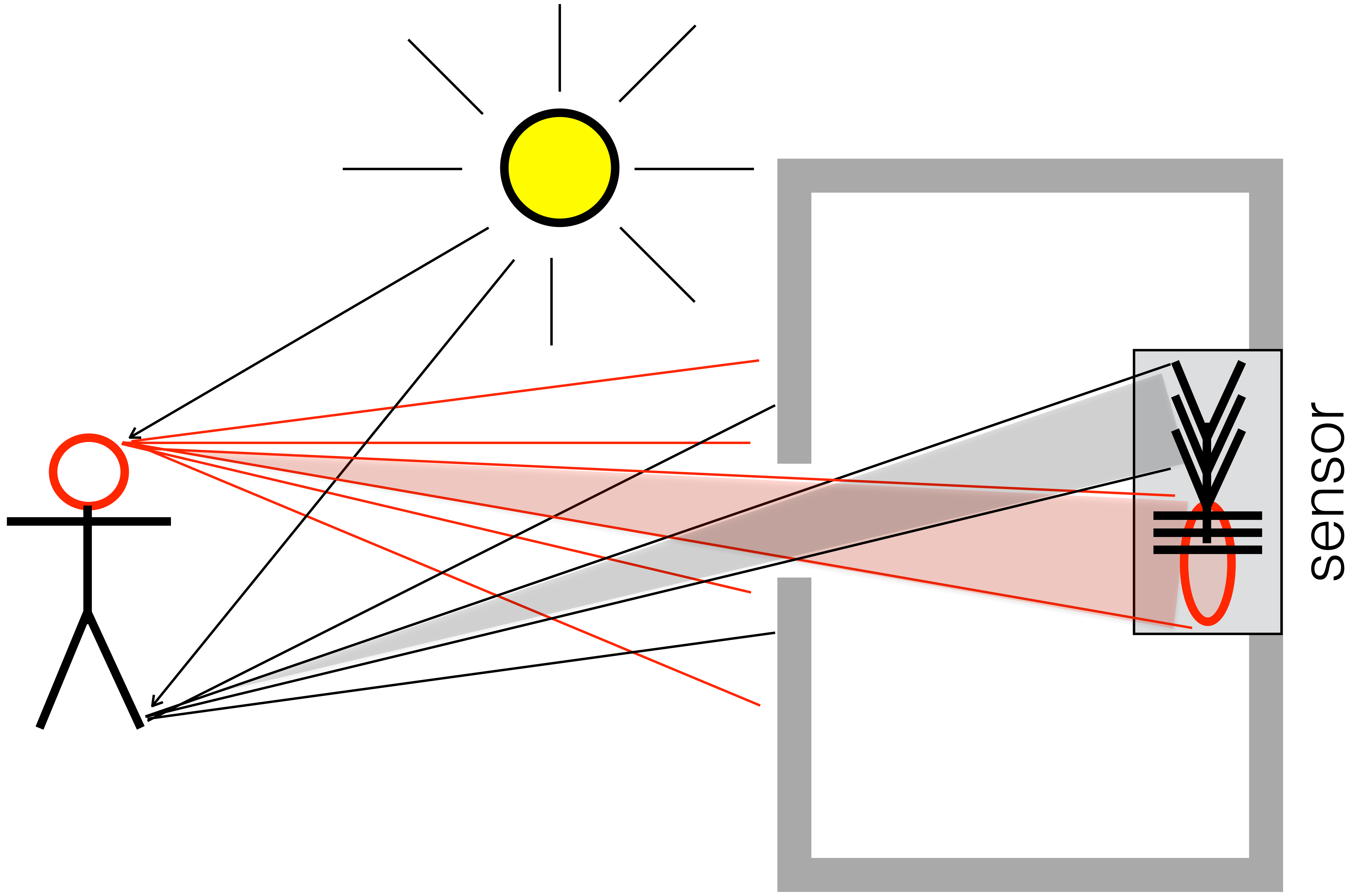
Light energy from rays traversed through pinhole is small

Camera model



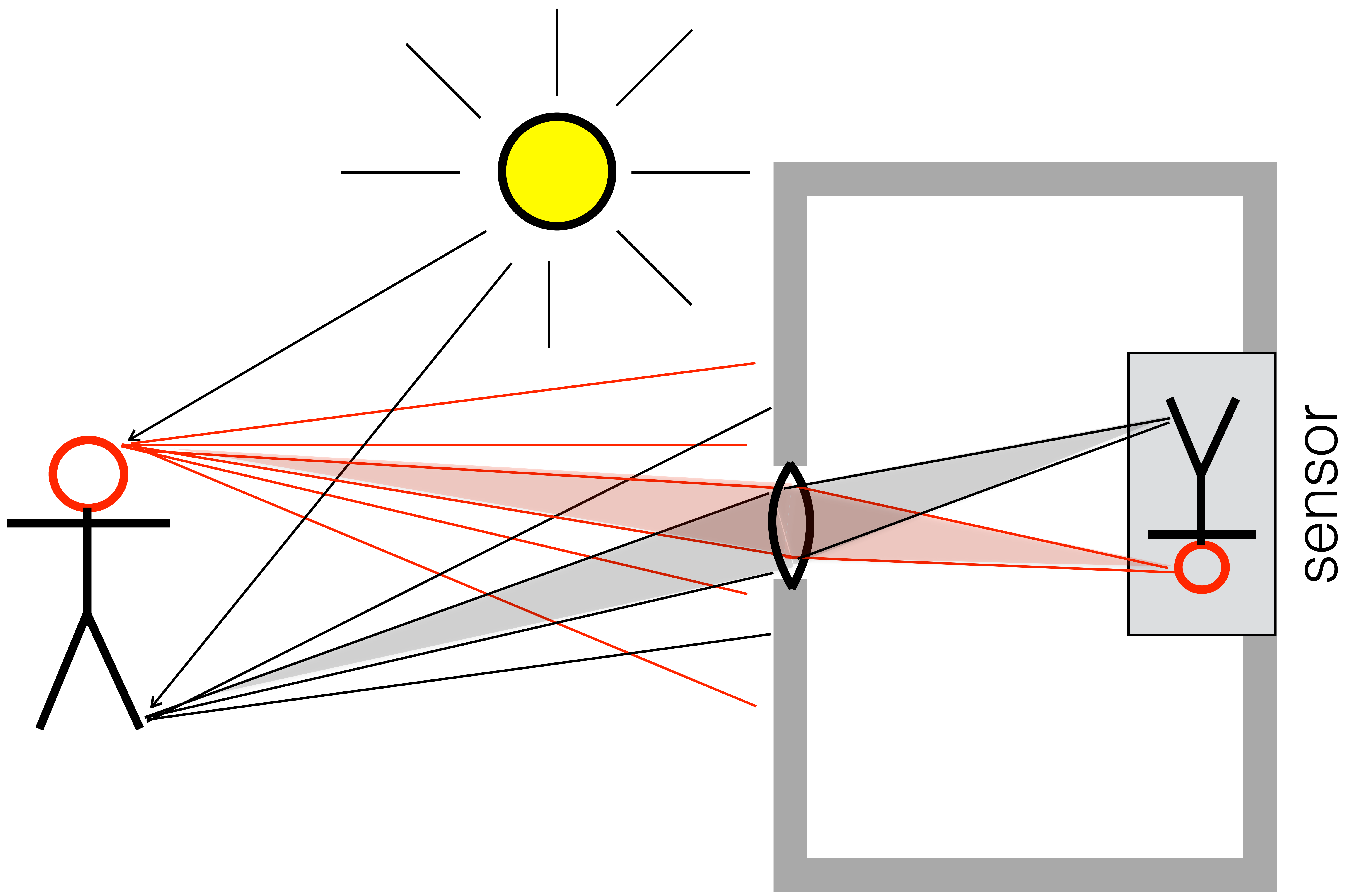
Increasing hole size yields more energy but blurs image

Camera model



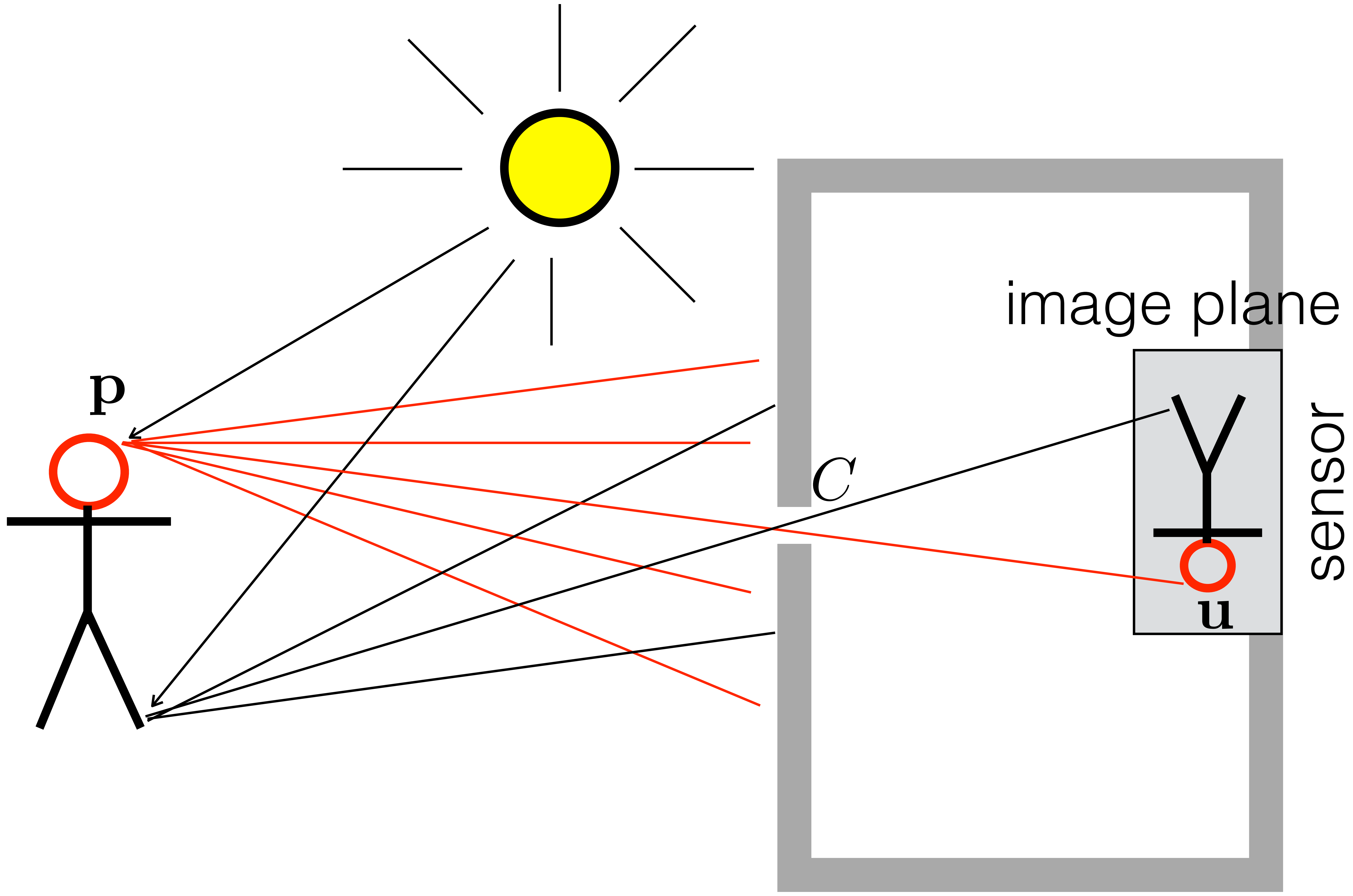
Increasing hole size yields more energy but blurs image

Camera model



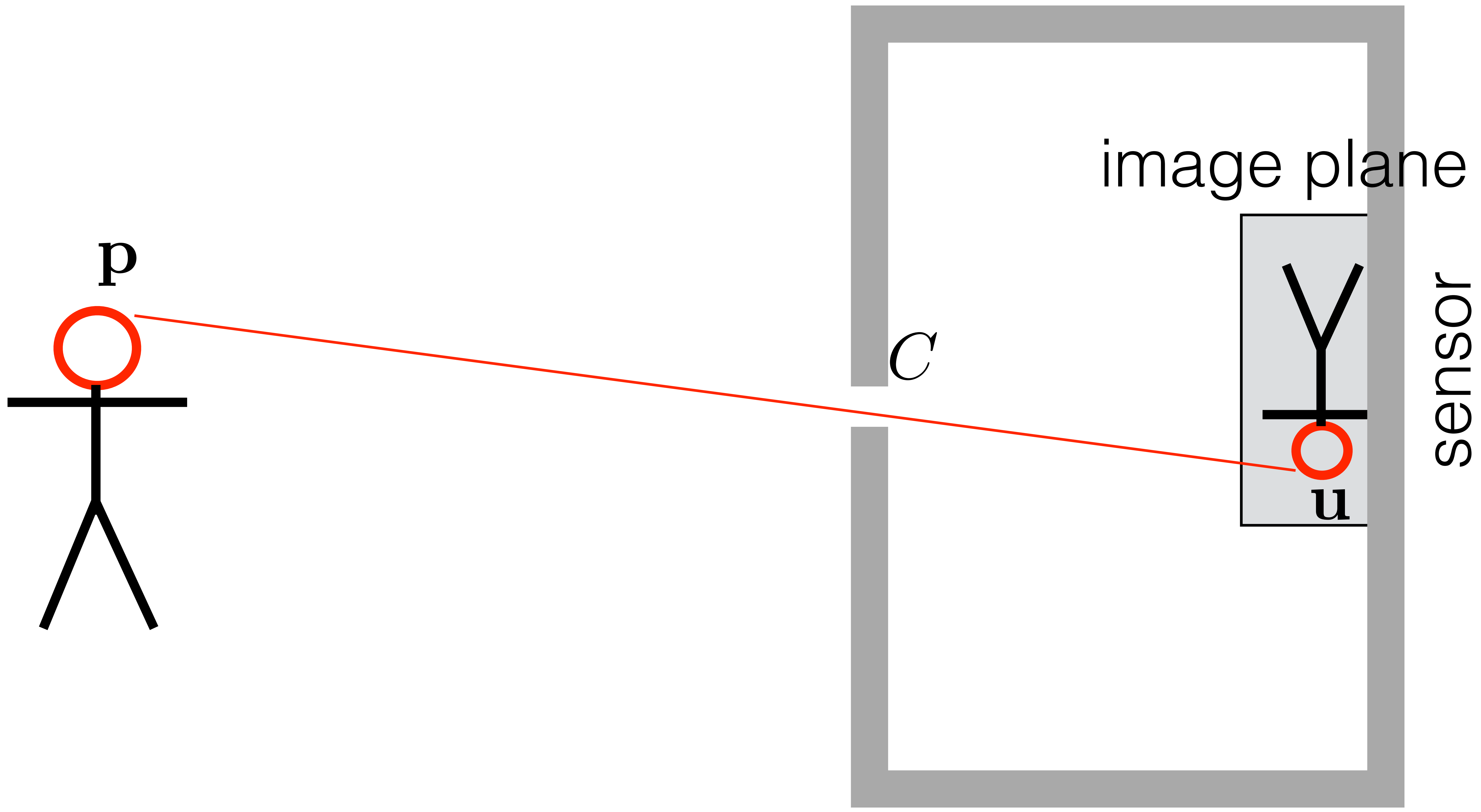
Lens focus cone of light rays in a single point

Camera model

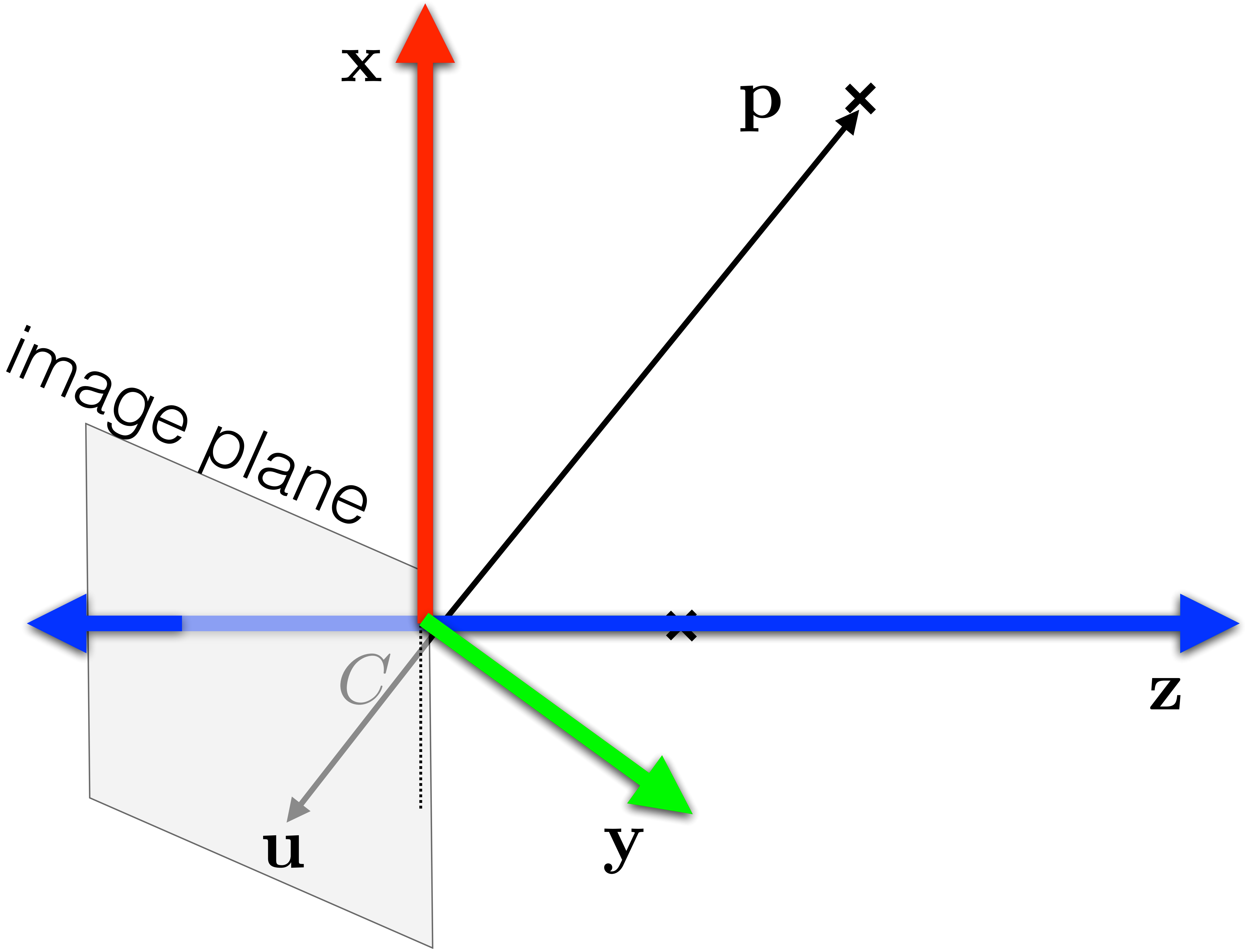


Since source and target of these cones are points, we omit the lens geometry.

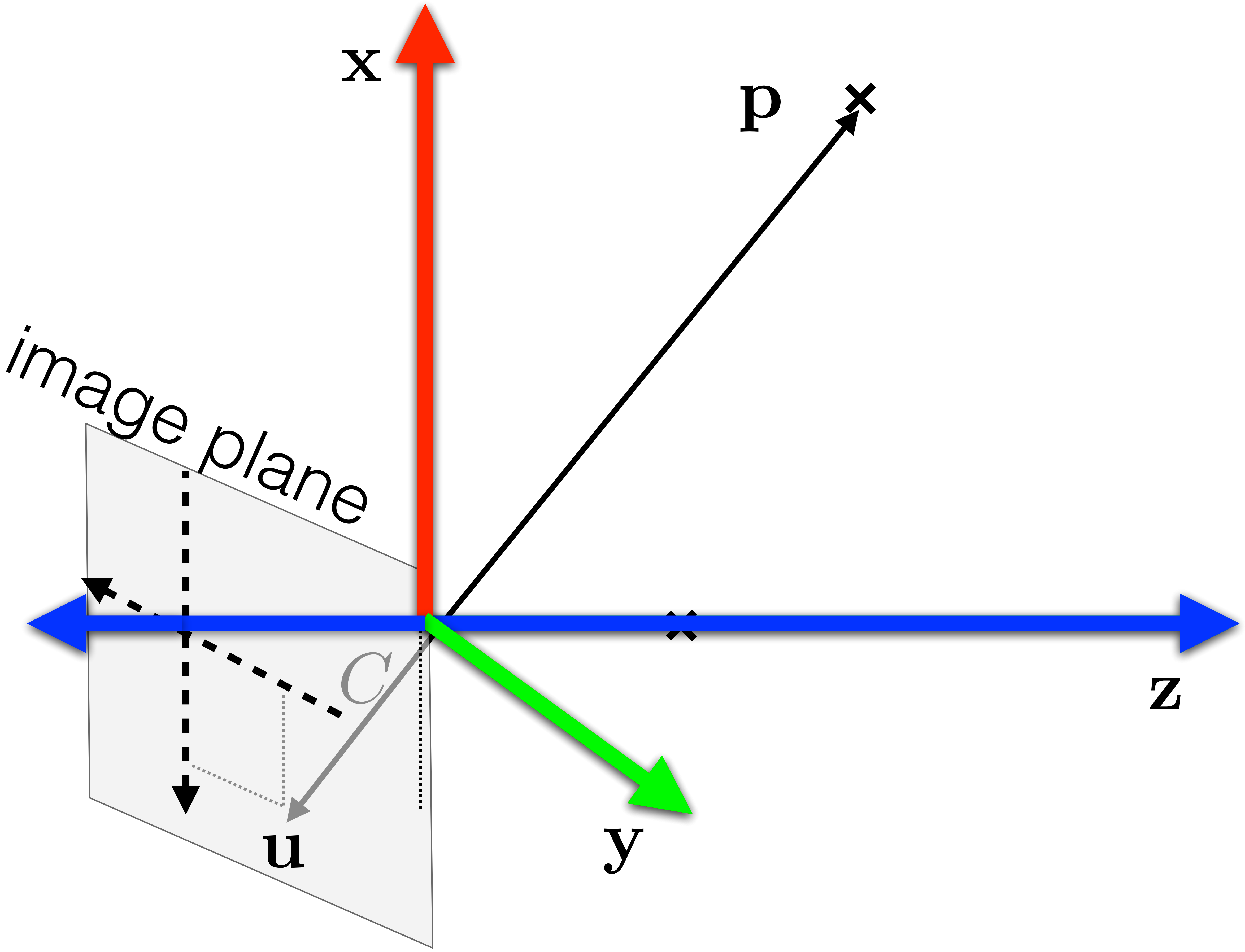
Camera model



Projection 3D points on the image plane

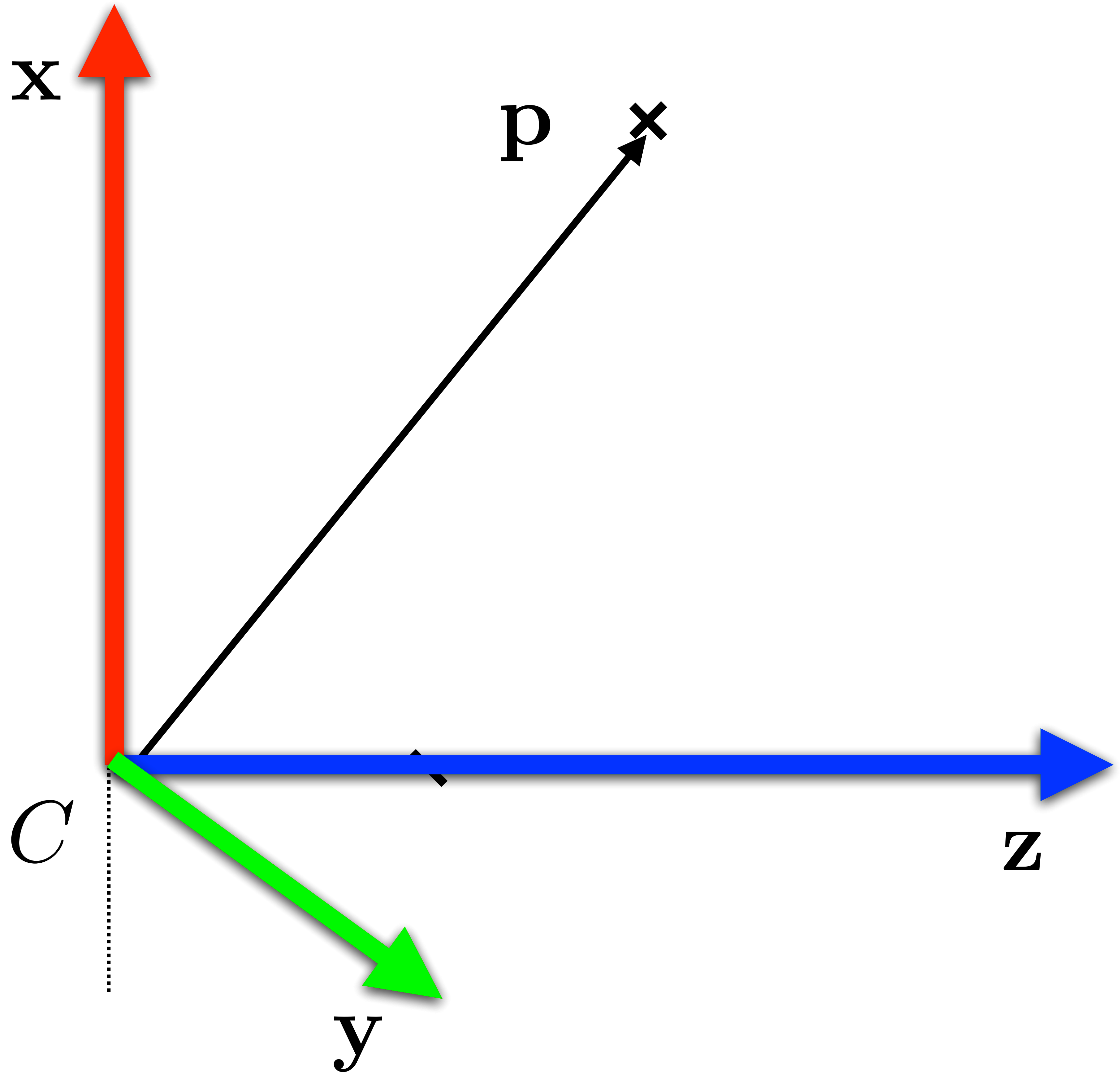


Projection 3D points on the image plane



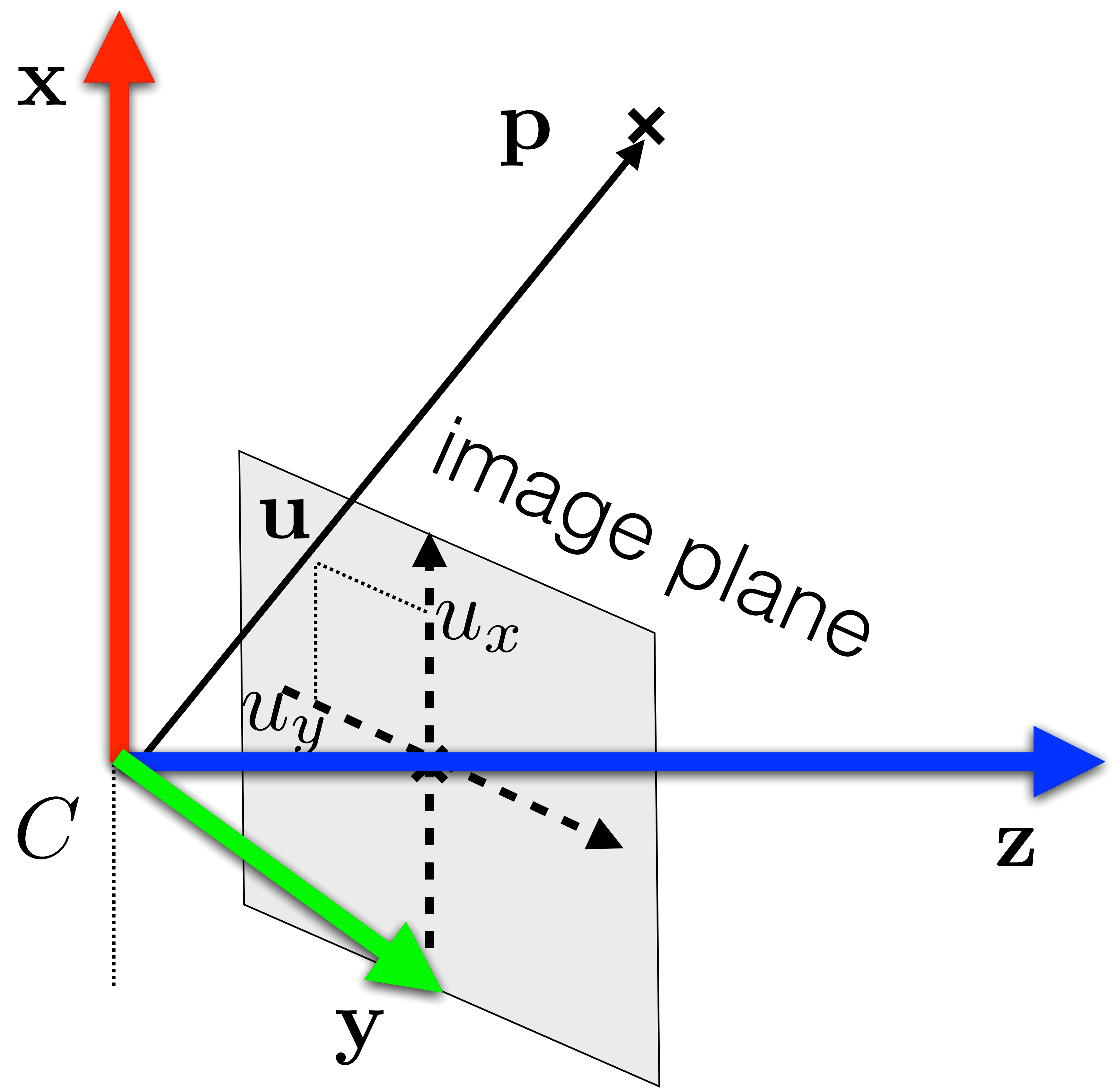
Pinhole camera model: Projection of 3D points on the image plane

Camera c.f.



Pinhole camera model: Projection of 3D points on the image plane

Camera c.f.

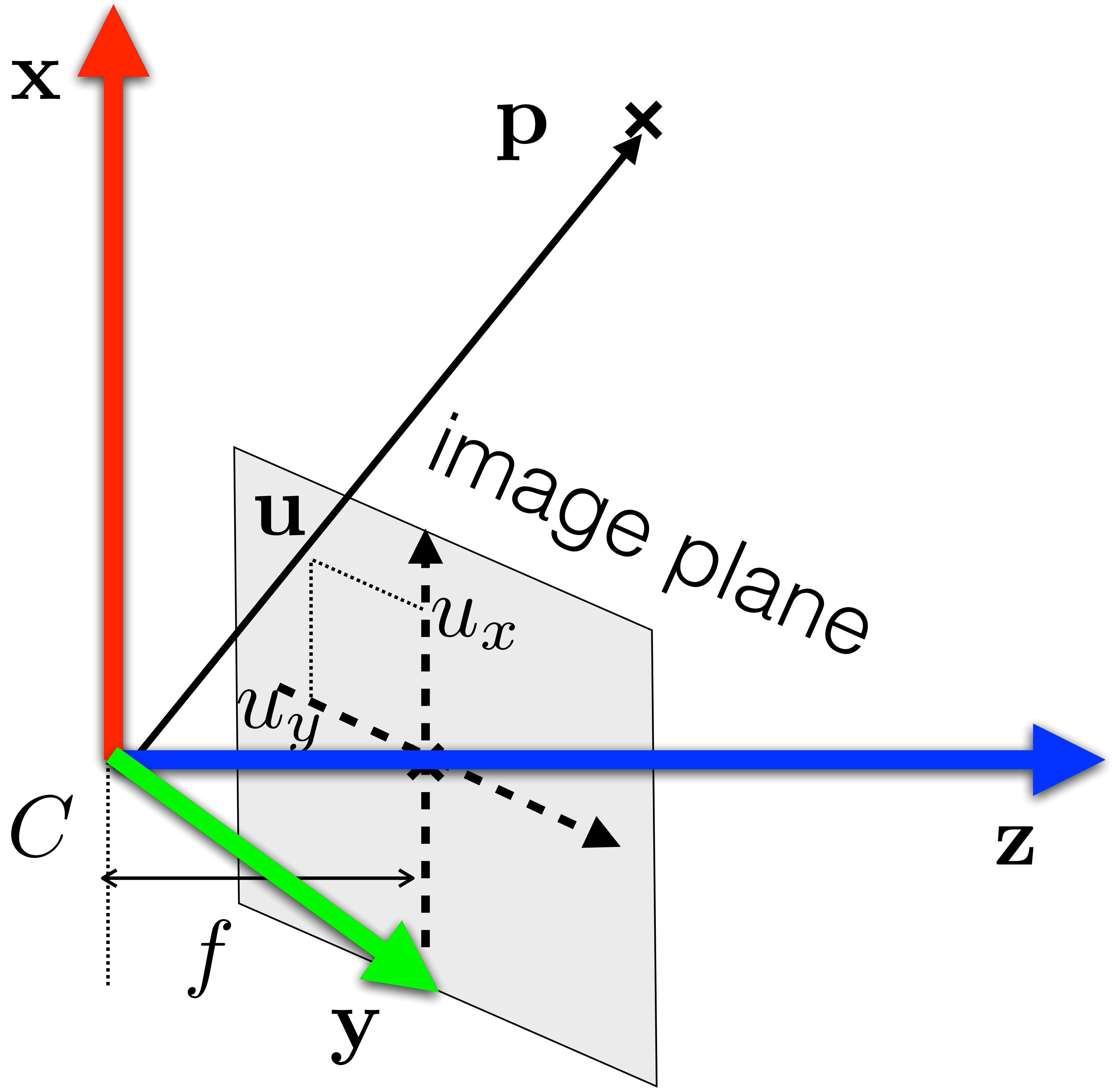


Pinhole camera model: Projection of 3D points on the image plane

$$u_x = f\frac{p_x}{p_z}$$

$$u_y = f\frac{p_y}{p_z}$$

Camera c.f.

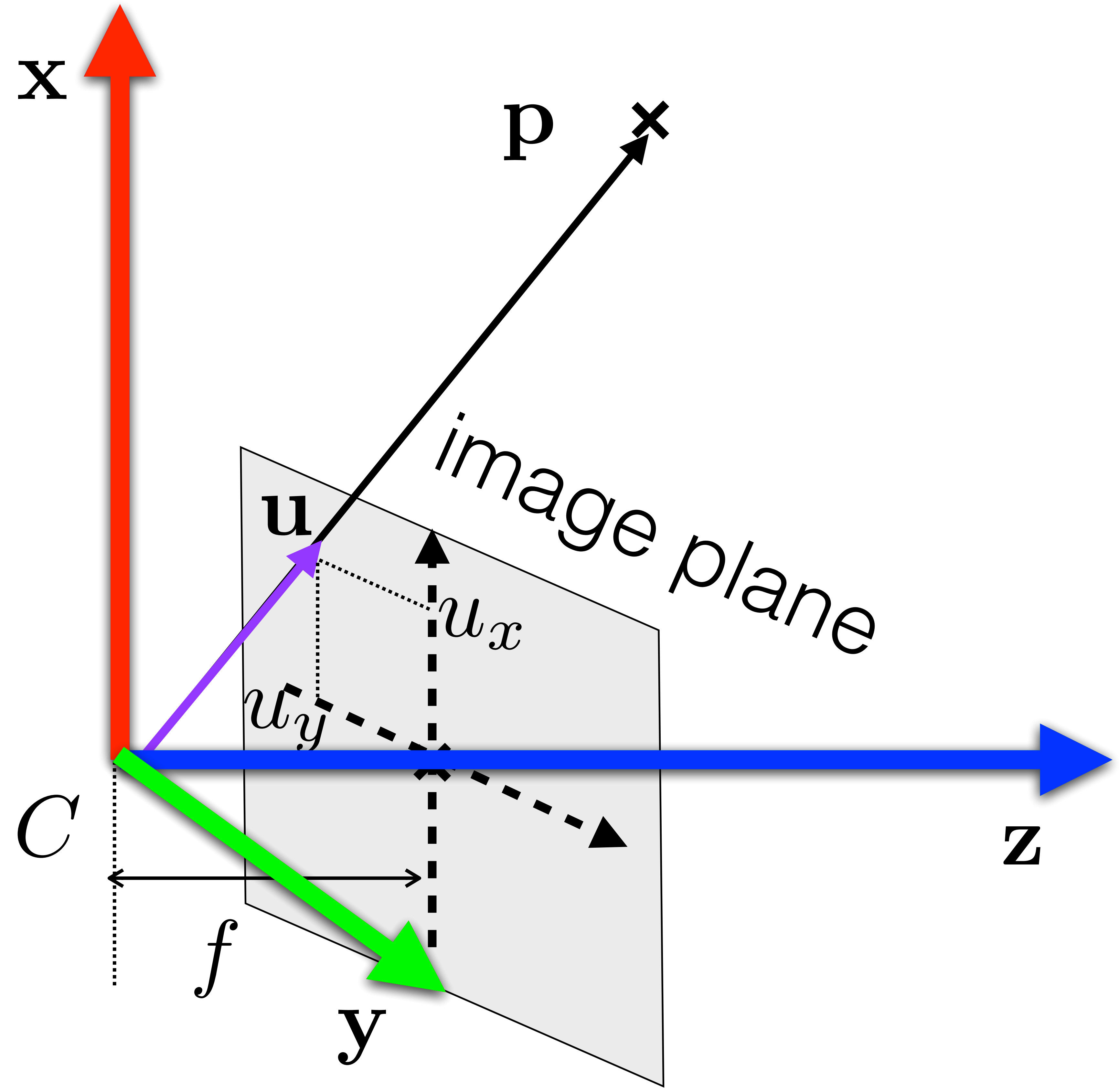


Pinhole camera model: Projection of 3D points on the image plane

$$u_x = f \frac{p_x}{p_z}$$

$$u_y = f \frac{p_y}{p_z}$$

Camera c.f.



Pinhole camera model: Projection of 3D points on the image plane

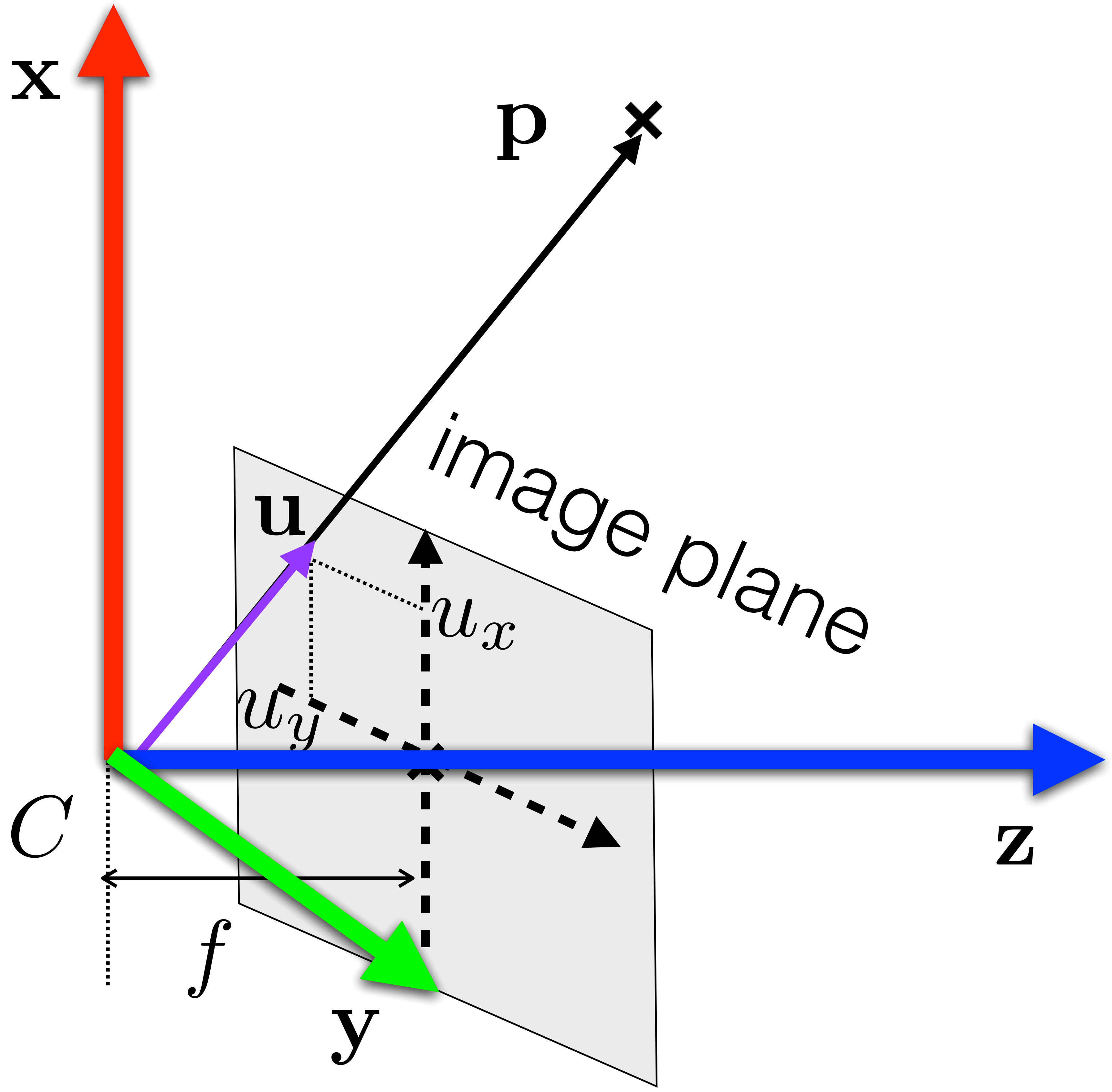
$$u_x = f\frac{p_x}{p_z}$$

$$u_y = f\frac{p_y}{p_z}$$

$$\Rightarrow$$

$$\lambda u_x = f p_x$$
$$\lambda u_y = f p_y$$
$$\lambda = p_z$$

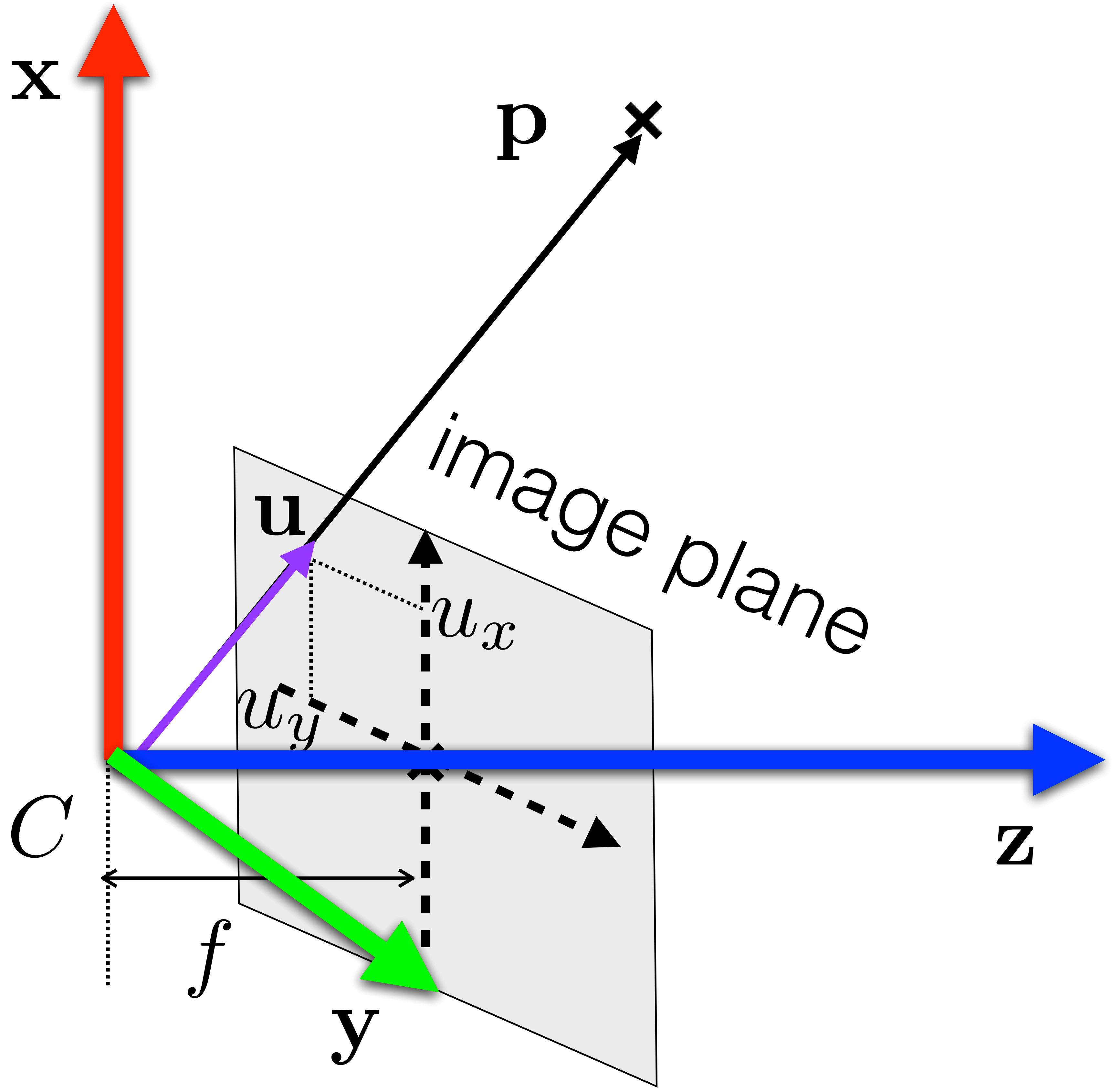
Camera c.f.



Pinhole camera model: Projection of 3D points on the image plane

$$u_x = f\frac{p_x}{p_z} \qquad u_y = f\frac{p_y}{p_z} \quad \Rightarrow \quad \begin{aligned} \lambda u_x &= f p_x \\ \lambda u_y &= f p_y \\ \lambda &= p_z \end{aligned} \quad \Rightarrow \quad \lambda \begin{bmatrix} u_x \\ u_y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix}$$

Camera c.f.



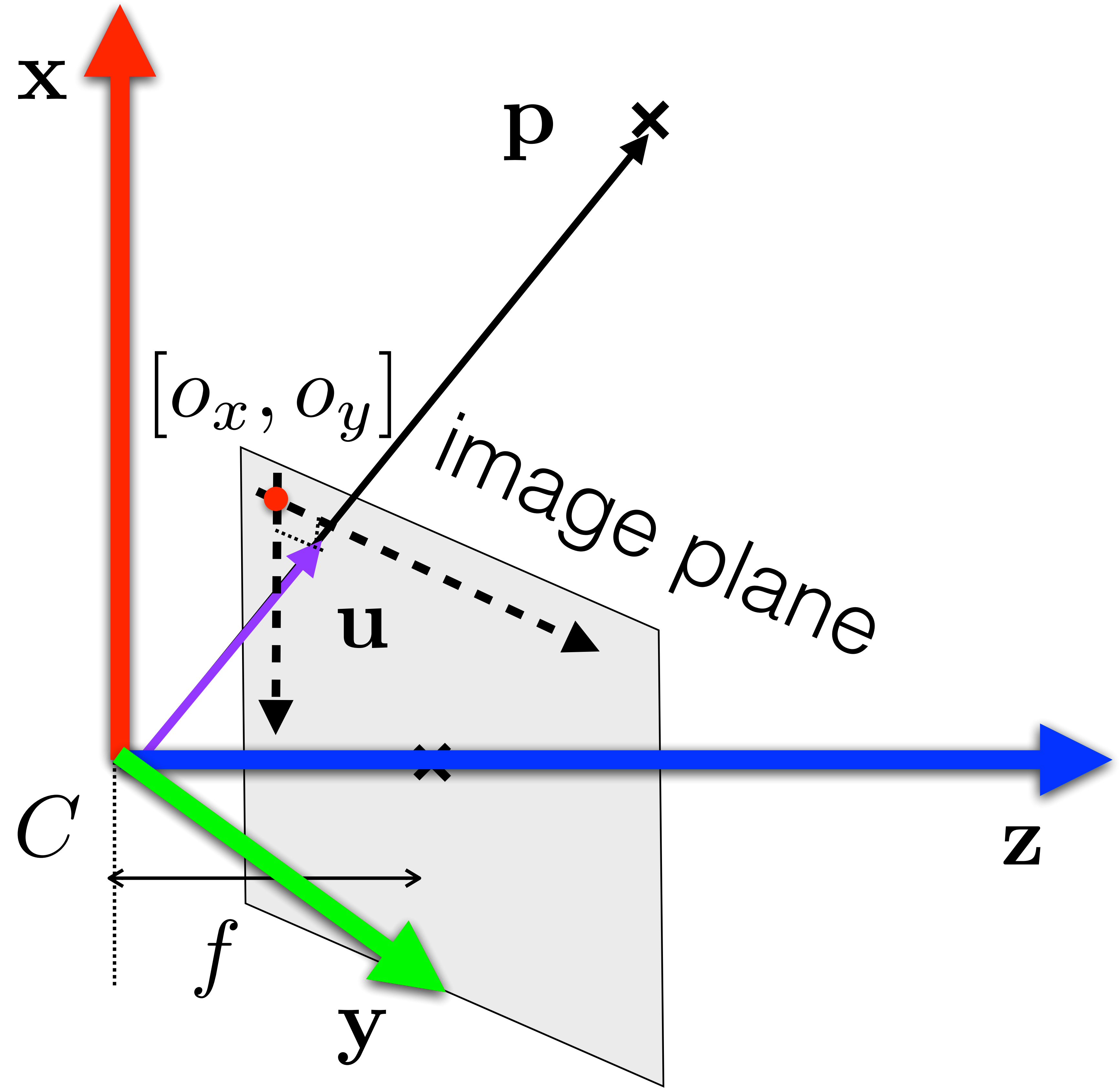
Pinhole camera model: Projection of 3D points on the image plane

$$u_x = o_x + s_x f \frac{p_x}{p_z}$$

$$u_y = o_y + s_y f \frac{p_y}{p_z}$$

$$\Rightarrow \quad \lambda \begin{bmatrix} u_x \\ u_y \\ 1 \end{bmatrix} = \begin{bmatrix} s_x f & 0 & o_x \\ 0 & s_y f & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix}$$

Camera c.f.



Pinhole camera model: Projection of 3D points on the image plane

$$u_x = o_x + s_x f \frac{p_x}{p_z}$$

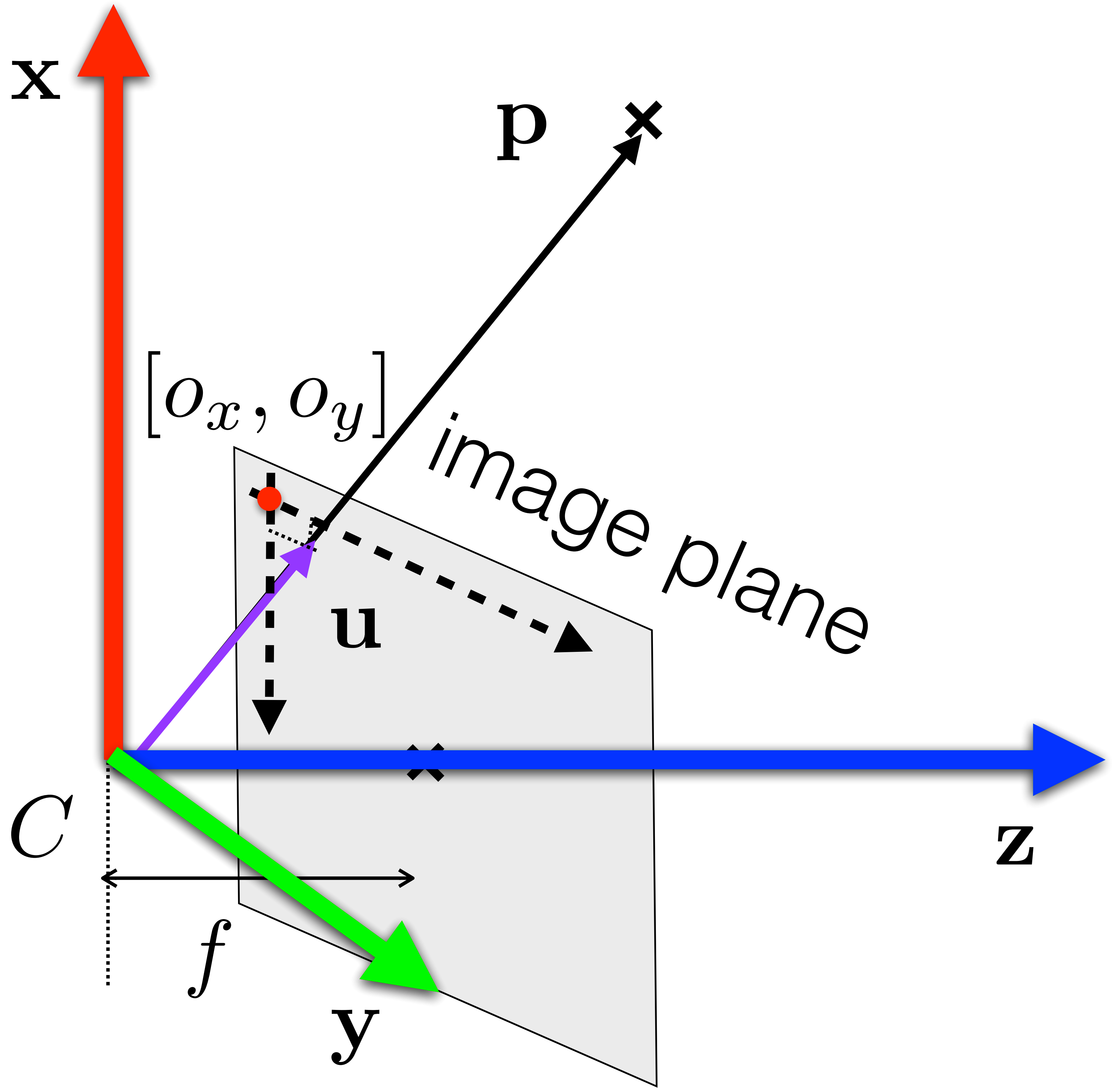
$$u_y = o_y + s_y f \frac{p_y}{p_z}$$

$$\Rightarrow \quad \lambda \begin{bmatrix} u_x \\ u_y \\ 1 \end{bmatrix} = \begin{bmatrix} s_x f & 0 & o_x \\ 0 & s_y f & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix}$$

$\mathbf{K} \in \mathcal{R}^{3\times 3}$

upper-triangular, regular matrix with intrinsic parameters of the camera

Camera c.f.



Pinhole camera model: Projection of 3D points on the image plane

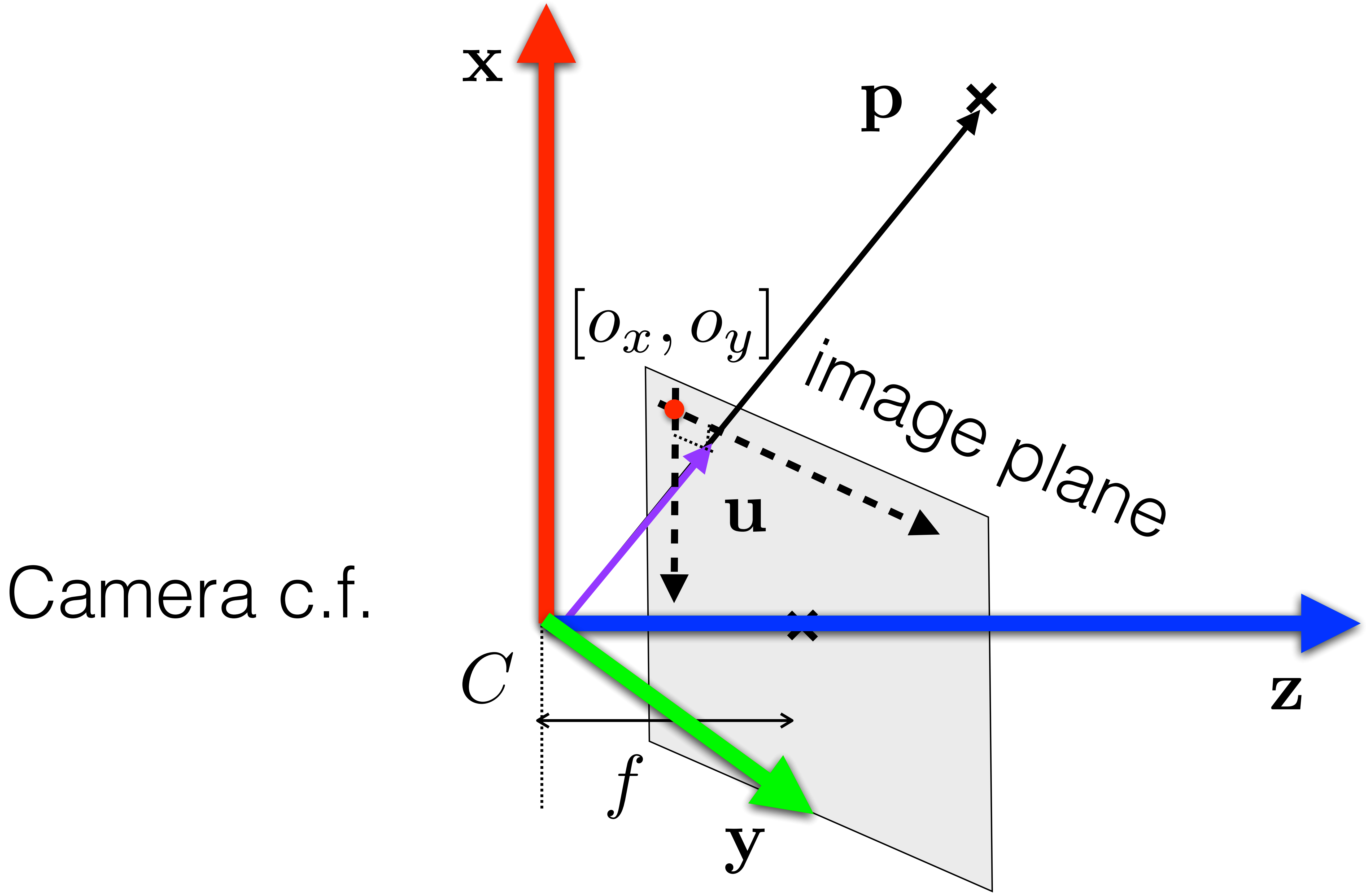
$$u_x = o_x + s_x f \frac{p_x}{p_z}$$

$$u_y = o_y + s_y f \frac{p_y}{p_z}$$

$$\Rightarrow \quad \lambda \begin{bmatrix} u_x \\ u_y \\ 1 \end{bmatrix} = \begin{bmatrix} s_x f & 0 & o_x \\ 0 & s_y f & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix}$$

$\mathbf{K} \in \mathcal{R}^{3\times3}$

upper-triangular, regular matrix with intrinsic parameters of the camera



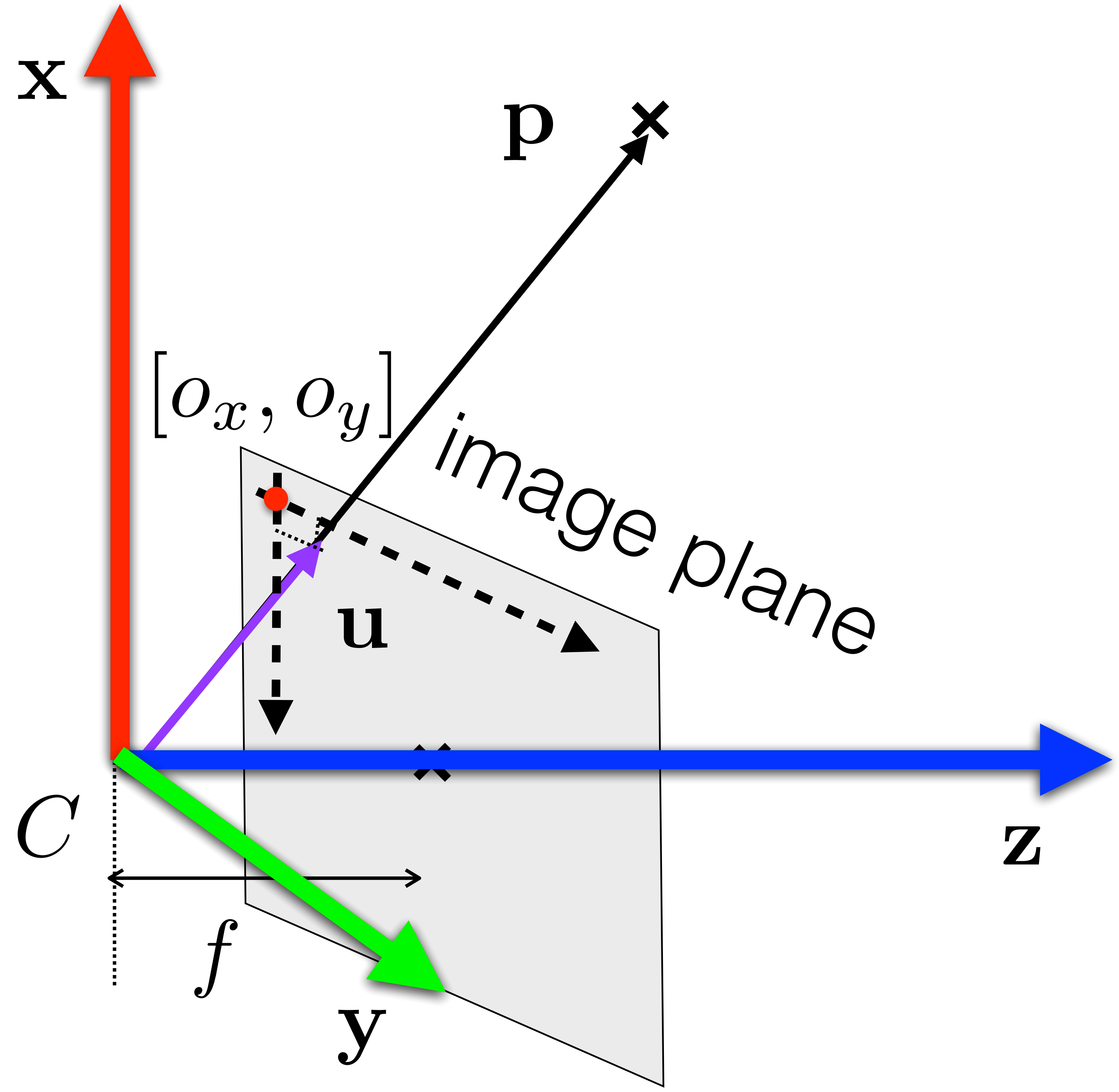
$$\lambda\bar{\mathbf{u}} = \mathbf{K}\mathbf{p}$$

Pinhole camera model: Projection of 3D points on the image plane

Applications:

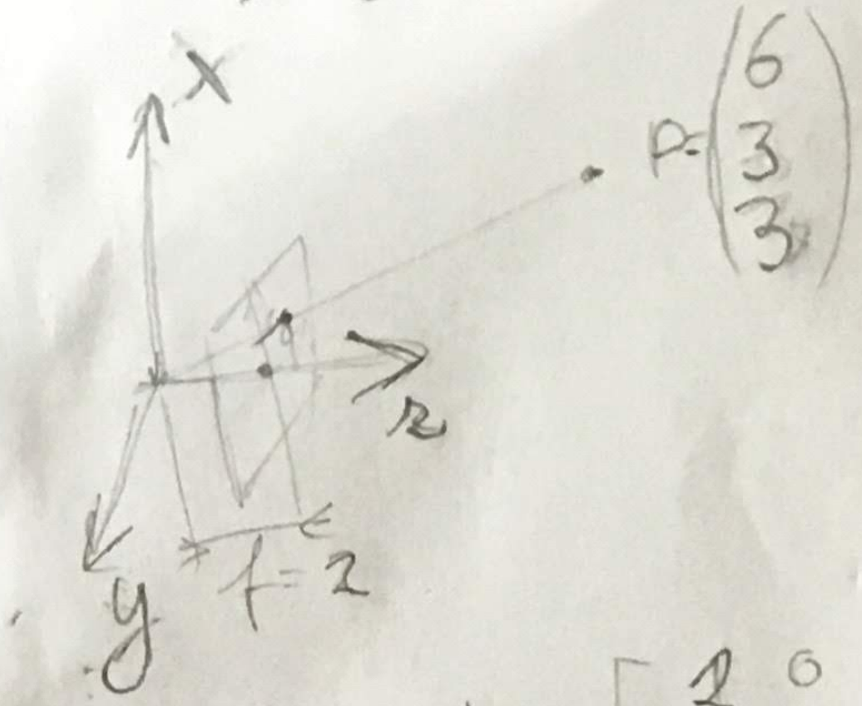
- 3D->2D projecting 3D PCL on image plane (colorizing)
- 2D->3D raycasting (projecting detections into 3D map)
- RGBD->3D PCL
- field-of-view, focal length and spatial resolution

Camera c.f.



$$\lambda\bar{\mathbf{u}} = \mathbf{K}\mathbf{p}$$

Projection & Raycasting



$$K = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

3D→2D:

$$\lambda \begin{pmatrix} u_x \\ u_y \\ 1 \end{pmatrix} = K p = \begin{bmatrix} 2 & & \\ & 2 & \\ & & 1 \end{bmatrix} \cdot \begin{pmatrix} 6 \\ 3 \\ 3 \end{pmatrix} = \begin{pmatrix} 12 \\ 6 \\ 3 \end{pmatrix}$$

$$\Rightarrow u = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$$

- Colorizing 3D pointcloud

2D→3D:

$$\lambda \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix} = \begin{bmatrix} 2 & & \\ & 2 & \\ & & 1 \end{bmatrix} \cdot p \Rightarrow p = \lambda \cdot \begin{bmatrix} 2 & & \\ & 2 & \\ & & 1 \end{bmatrix}^{-1} \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix} =$$

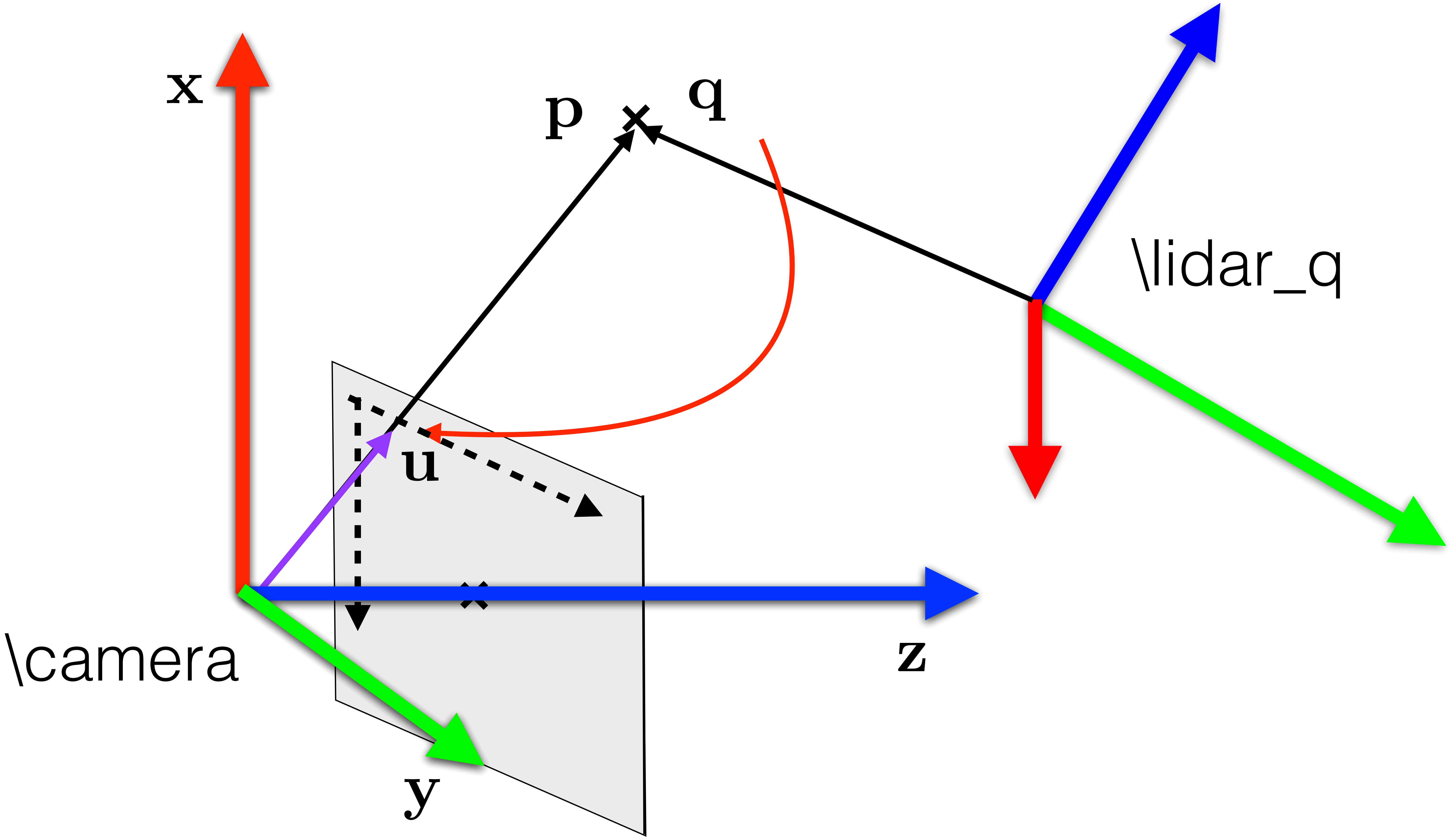
$$= \lambda \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

ray corresponding to pixel $u = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$ is

$$R = \left\{ p \in \mathbb{R}^3, \ p = \lambda \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} ; \lambda \in \mathbb{R}^+ \right\}$$

- raycasting (projecting detected objects from camera to map)

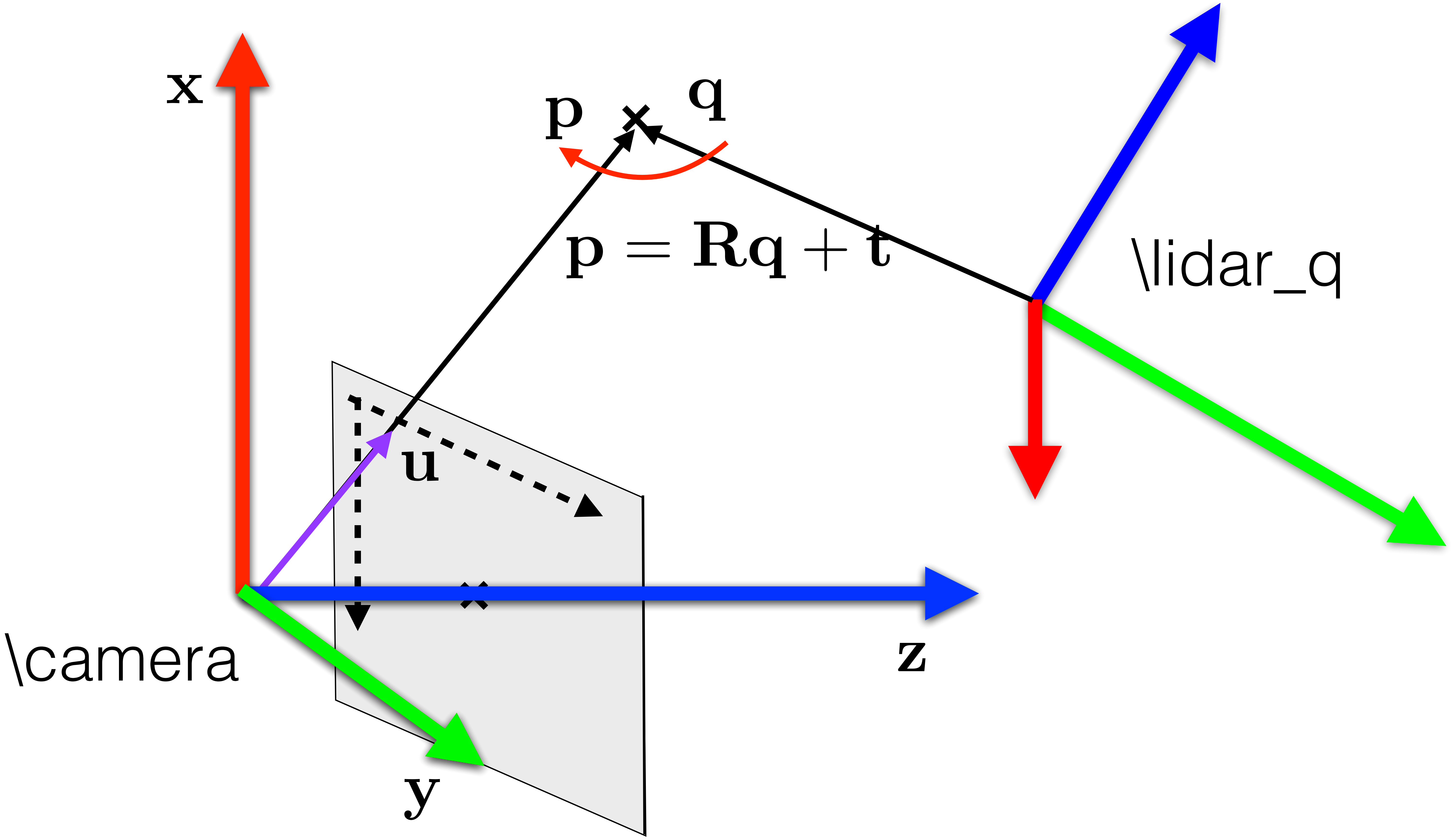
Let us have one lidar and one camera



Camera

Projection from lidar to image plane consists of two steps:

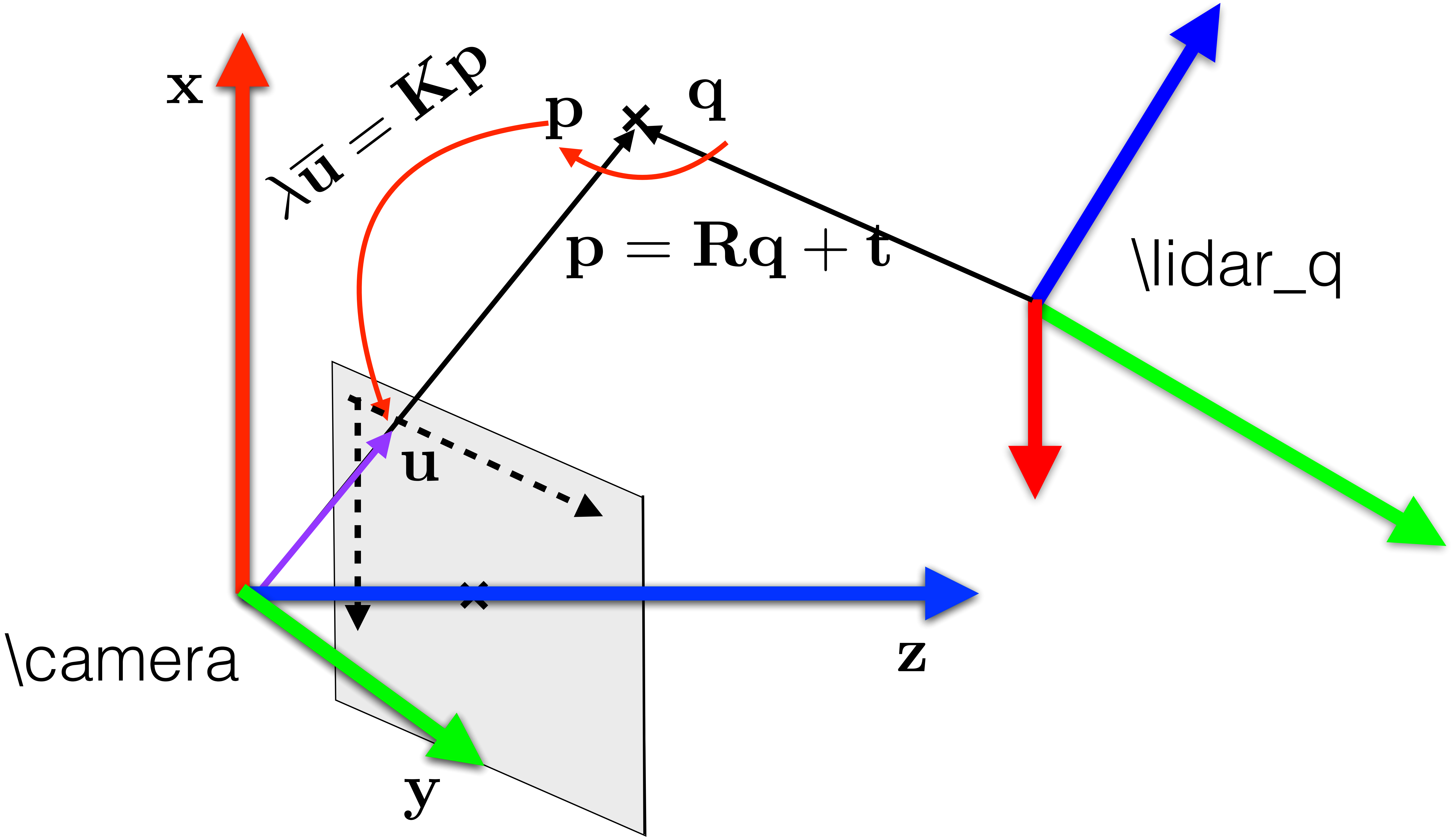
- transform of 3D point in \lidar_q $\mathbf{q} \in \mathcal{R}^3$ to \camera $\mathbf{p} \in \mathcal{R}^3$



Camera

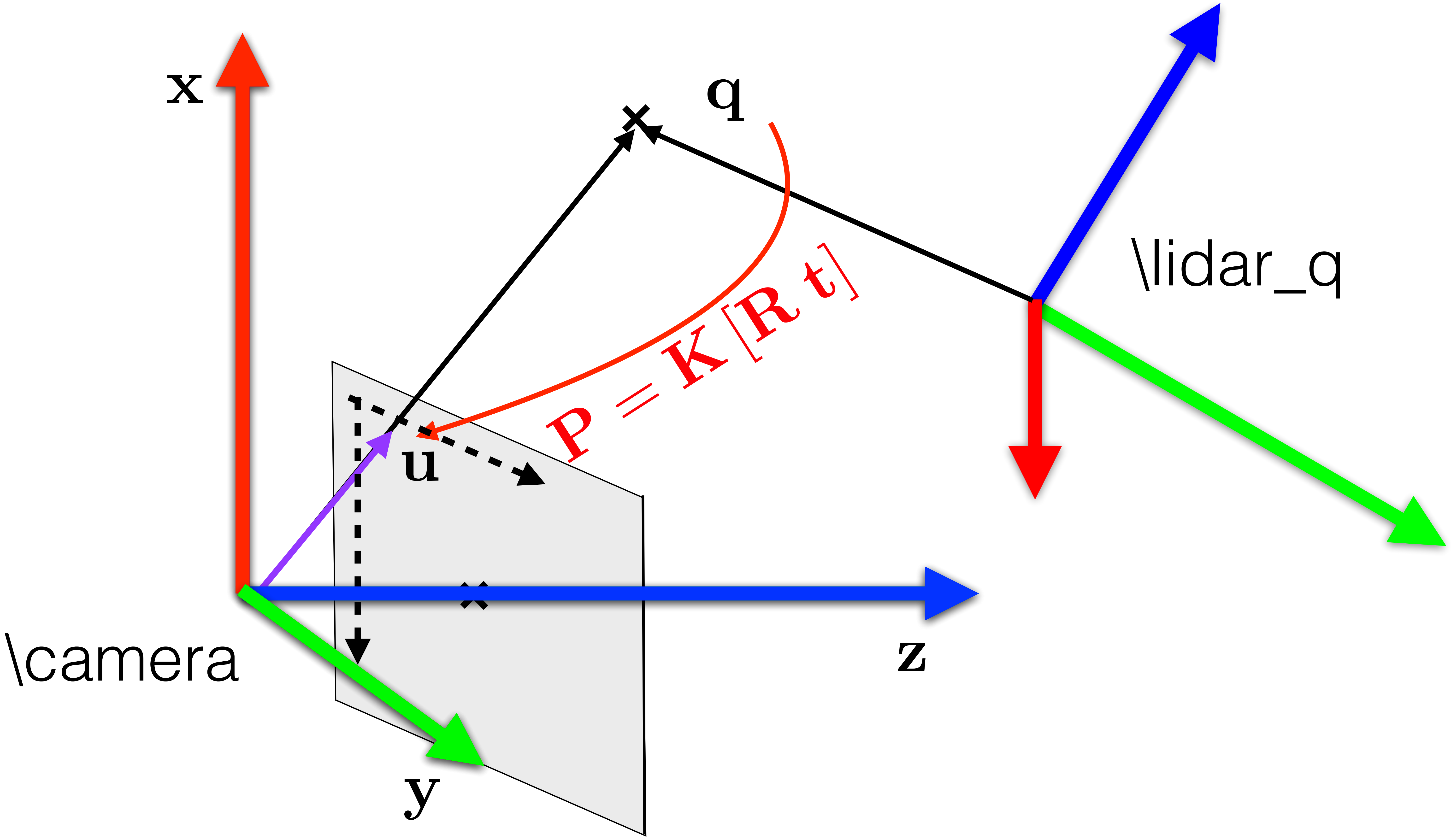
Projection from lidar to image plane consists of two steps:

- transform of 3D point in \lidar_q $\mathbf{q} \in \mathcal{R}^3$ to \camera $\mathbf{p} \in \mathcal{R}^3$
- projection of 3D point in \camera on image plane $\mathbf{u} \in \mathcal{R}^2$



Lidar to camera projection equation

$$\lambda \overline{\mathbf{u}} = \underbrace{\mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}}_{\mathbf{P}} \overline{\mathbf{q}}$$



Projection 3D points on the image plane

$$\lambda\overline{\mathbf{u}} = \begin{bmatrix} s_x & s_o & o_x \\ 0 & s_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix} \overline{\mathbf{q}}$$

$\mathbf{K} \in \mathcal{R}^{3\times3}$

$\mathbf{P} \in \mathcal{R}^{3\times4}$

$\mathbf{K} \in \mathcal{R}^{3\times3}$ .... intrinsic parameters

$\mathbf{R} \in \mathcal{SO}(3), \mathbf{t} \in \mathcal{R}^3$...... extrinsic parameters

$\mathbf{P} \in \mathcal{R}^{3\times4}$ .... camera projection matrix

Example 1: Project point to a given camera.

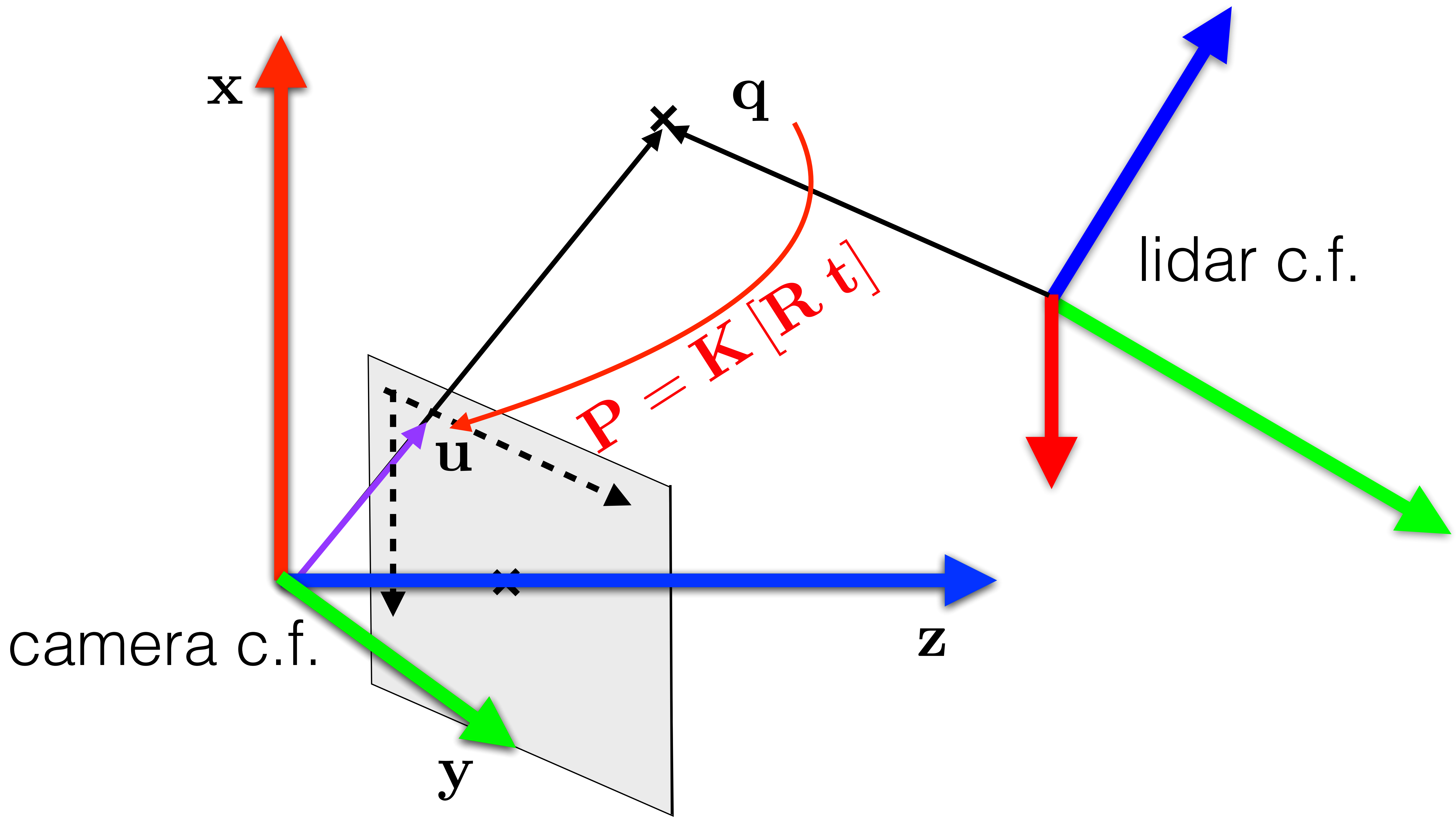
Example 2: What is a ray of a pixel?

Example 3: Depth to 3D point-cloud?

Camera to lidar calibration

$$\lambda \bar{\mathbf{u}} = \underbrace{\mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}}_{\mathbf{P}} \bar{\mathbf{q}}$$

Estimate this



Camera to lidar calibration

$$\lambda \overline{\mathbf{u}} = \mathbf{P}\overline{\mathbf{q}}$$

unknown

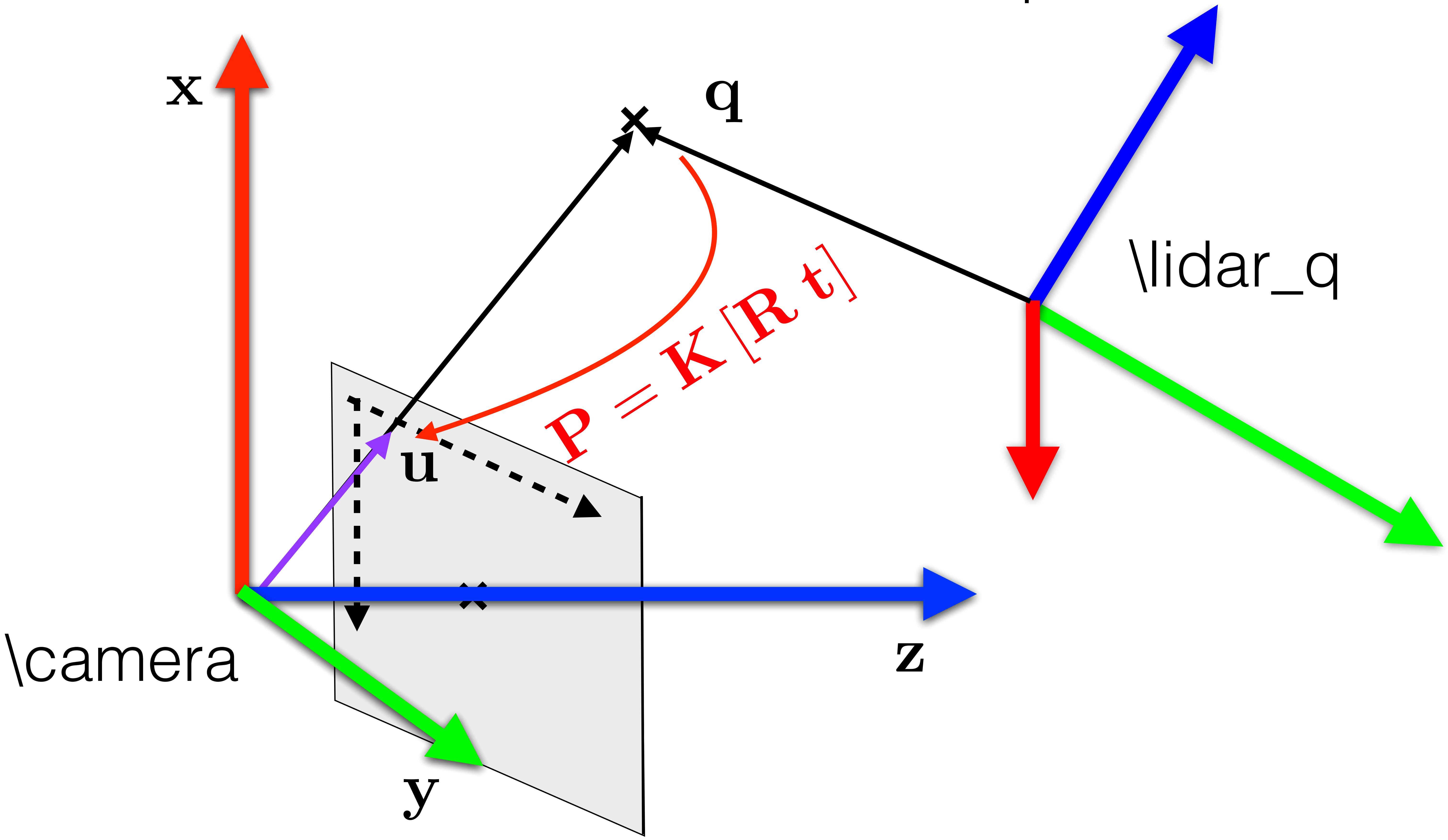
Solution:

(0) Get (u,q) corresponds.

(1) Get rid of lambda

(2) Express lin.eq. for P

calibration from 2D-3D correspondences



\camera to \lidar_q calibration

Each 2D-3D correspondence yields two equations:

$$\underbrace{\begin{bmatrix} -\overline{\mathbf{q}}_i^\top & \mathbf{0}^\top & u_{xi}\overline{\mathbf{q}}_i^\top \\ \mathbf{0}^\top & -\overline{\mathbf{q}}_i^\top & u_{yi}\overline{\mathbf{q}}_i^\top \end{bmatrix}}_{\mathbf{A}_{[2\times 12]}} \underbrace{\begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{bmatrix}}_{\mathbf{p}_{[12\times 1]}} = \mathbf{0}_{[2\times 1]}$$

For N 2D-3D correspondences, we obtain
(2Nx12) homogeneous linear system $\mathbf{Ap} = \mathbf{0}$

Assuming

- i.i.d. measurements and
- gaussian noise between left-hand-side and right-hand-side

$$\mathbf{p}^* = \operatorname{argmin}\|\mathbf{Ap}\|$$

\camera to \lidar_q calibration

Each 2D-3D correspondence yields two equations:

$$\underbrace{\begin{bmatrix} -\overline{\mathbf{q}}_i^\top & \mathbf{0}^\top & u_{xi}\overline{\mathbf{q}}_i^\top \\ \mathbf{0}^\top & -\overline{\mathbf{q}}_i^\top & u_{yi}\overline{\mathbf{q}}_i^\top \end{bmatrix}}_{\mathbf{A}_{[2\times 12]}} \underbrace{\begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{bmatrix}}_{\mathbf{p}_{[12\times 1]}} = \mathbf{0}_{[2\times 1]}$$

For N 2D-3D correspondences, we obtain (2Nx12) homogeneous linear system $\mathbf{A}\mathbf{p} = \mathbf{0}$

Assuming

- i.i.d. measurements and
- gaussian noise between left-hand-side and right-hand-side

$$\mathbf{p}^* = \operatorname{argmin}\|\mathbf{A}\mathbf{p}\| \ \text{ subject to } \ \|\mathbf{p}\| = 1$$

$$\mathbf{p}^* = \text{argmin}\|\mathtt{A}\mathbf{p}\| \;\; \text{subject to} \;\; \|\mathbf{p}\| = 1$$

Lagrange function:

$$L(\mathbf{p}, \lambda) = \|\mathtt{A}\mathbf{p}\| + \lambda(1 - \|\mathbf{p}\|) = \mathbf{p}^\top \mathtt{A}^\top \mathtt{A}\mathbf{p} + \lambda(1 - \mathbf{p}^\top \mathbf{p})$$

Critical points:

$$\frac{\partial L(\mathbf{p}, \lambda)}{\partial \mathbf{p}} = 2\mathtt{A}^\top \mathtt{A}\mathbf{p} - 2\lambda\mathbf{p} = \mathbf{0}$$

$$\frac{\partial L(\mathbf{p}, \lambda)}{\partial \lambda} = 1 - \mathbf{p}^\top \mathbf{p} = \mathbf{0}$$

First equation is characteristic equation $(\mathtt{A}^\top \mathtt{A} - \lambda \mathtt{I})\mathbf{p} = \mathbf{0}$

Every eigen-vector of $\mathtt{A}^\top \mathtt{A}$ is the critical point=> choose one

Cost function in these eigen vectors is equal to eigen-values

$$\|\mathtt{A}\mathbf{p}\| = \mathbf{p}^\top \mathtt{A}^\top \mathtt{A}\mathbf{p} = \mathbf{p}^\top \lambda \mathbf{p} = \lambda \mathbf{p}^\top \mathbf{p} = \lambda \|\mathbf{p}\| = \lambda$$

Solution is the eigen-vector of $\mathtt{A}^\top \mathtt{A}$ with the smallest eigen-value

Summary camera calibration

- Manually estimate 2D-3D correspondences
- Build matrix

$$\mathbf{A} = \begin{bmatrix} -\mathbf{q}^\top & \mathbf{0}^\top & u_x \mathbf{q}^\top \\ \mathbf{0}^\top & -\mathbf{q}^\top & u_y \mathbf{q}^\top \\ \vdots & \vdots & \vdots \end{bmatrix}$$

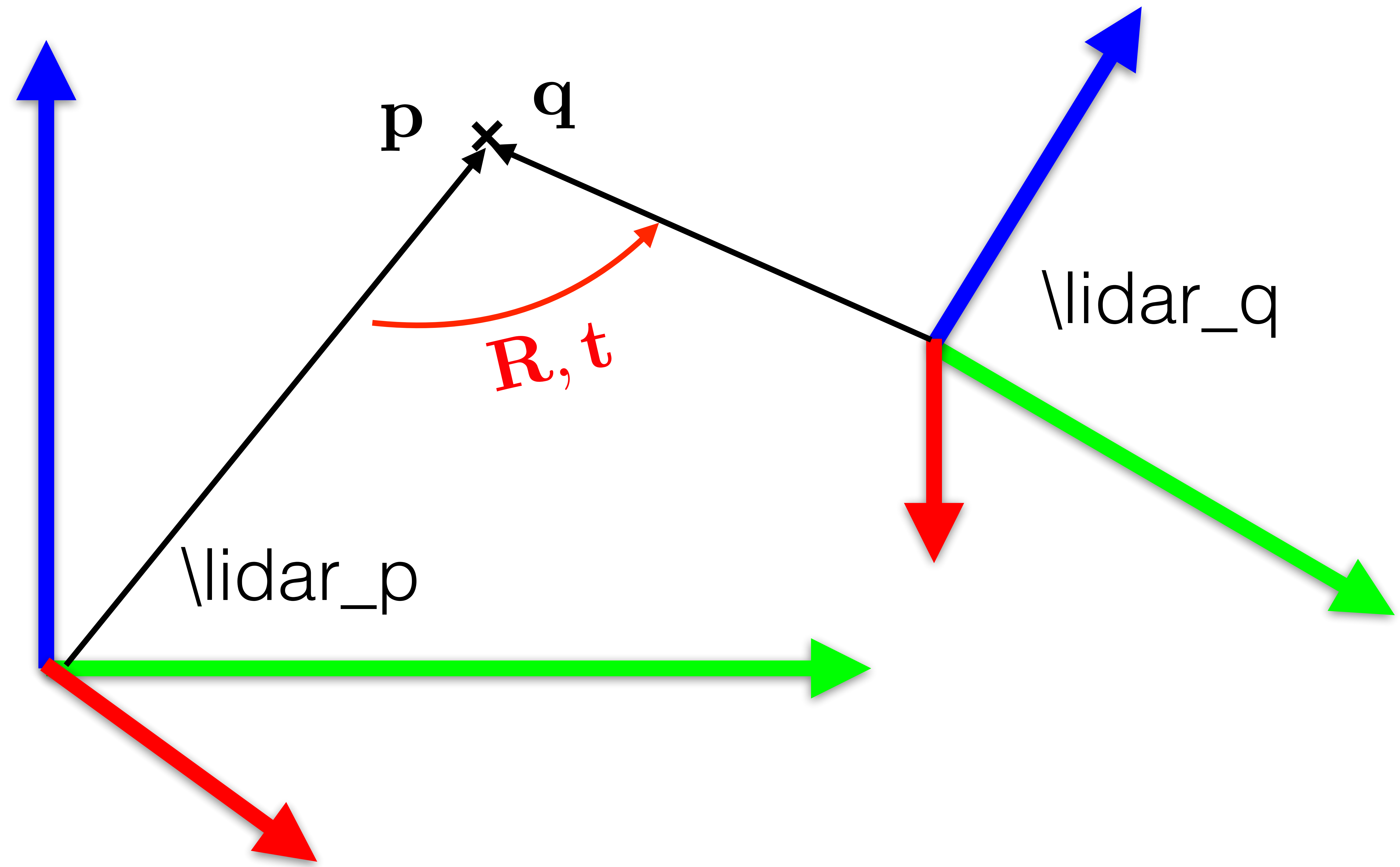
- Find eigen-values and eigen-vectors of $\mathbf{A}^\top \mathbf{A}$ (python: `numpy.linalg.eig`)
- Reshape the eigen-vector $\mathbf{p} \in \mathcal{R}^{12\times 1}$ with the smallest eigen-value to camera matrix $\mathbf{P} \in \mathcal{R}^{3\times 4}$
- Scale does not matter: $\mathbf{P} = \mathbf{P} / \| [\mathbf{p}_{31}, \mathbf{p}_{32}, \mathbf{p}_{33}]^\top \|$
- Optionally decompose:

$$\mathbf{P} = [\underbrace{\mathbf{KR}}_{\mathbf{B}} \ \ \underbrace{\mathbf{Kt}}_{\mathbf{c}}] = [\mathbf{B} \ \ \mathbf{c}] \qquad \mathbf{K}, \mathbf{R} = qr(\mathbf{B}) \qquad \mathbf{t} = \mathbf{K}^{-1}\mathbf{c}$$

(python: `numpy.linalg.qr`)

Summary

lidar-lidar calibration from 3D-3D correspondences

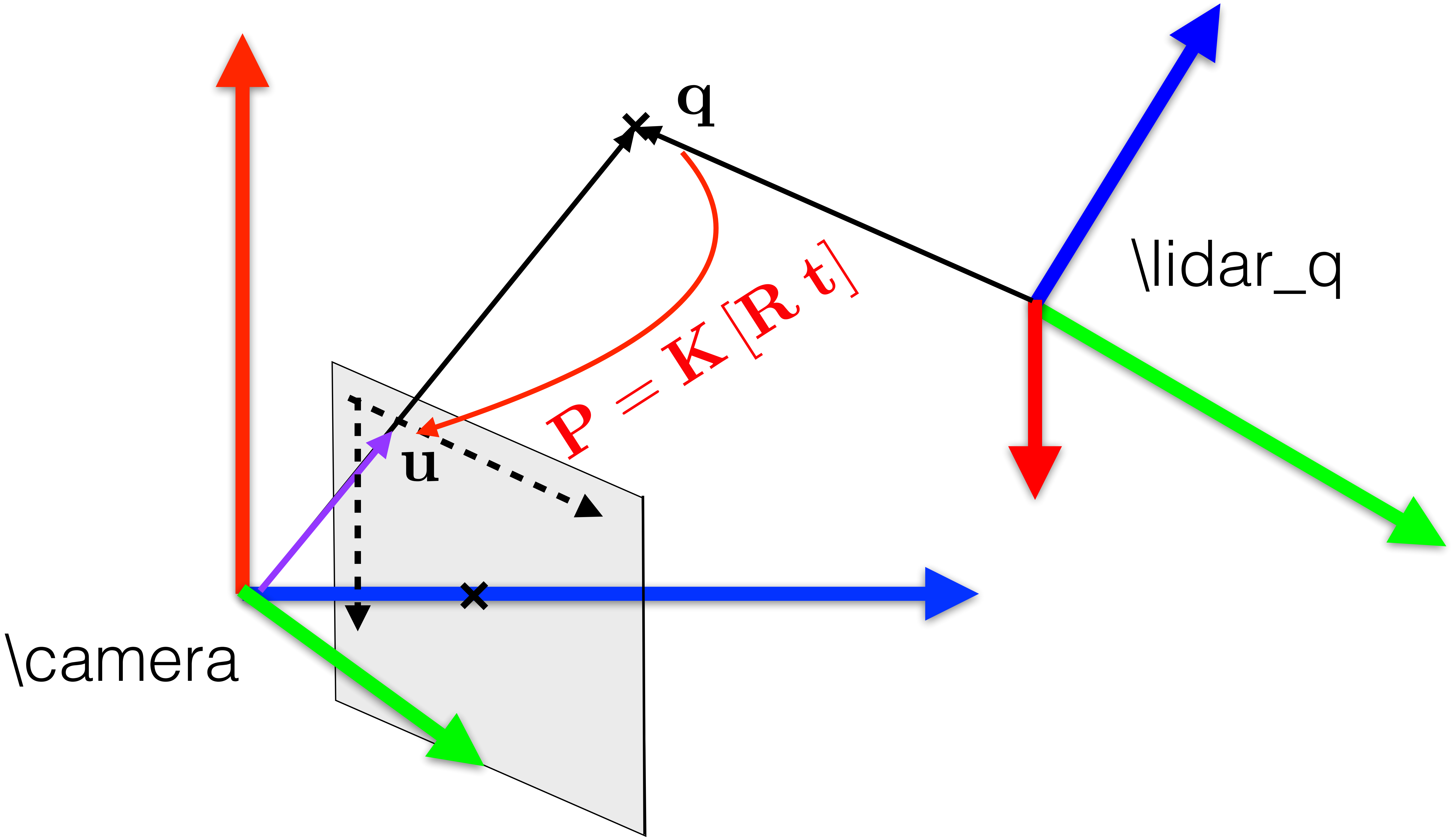


Solve: $$\mathbf{R}^*, \mathbf{t}^* = \underset{\mathbf{R} \in SO(3), \mathbf{t} \in \mathcal{R}^3}{\arg\min} \sum_i \|\mathbf{R}\mathbf{p}_i + \mathbf{t} - \mathbf{q}_i\|_2^2$$

Solution: $$\mathbf{R}^* = \mathbf{V}\mathbf{U}^\top$$
$$\mathbf{t}^* = \widetilde{\mathbf{q}} - \mathbf{R}^* \widetilde{\mathbf{p}}$$

Summary

camera-lidar calibration from 2D-3D correspondences



Solve: $\mathbf{p}^* = \text{argmin}||\mathbf{A}\mathbf{p}||$ subject to $||\mathbf{p}|| = 1$

Solution: smallest eigen-vector of $\mathbf{A}^\top\mathbf{A}$