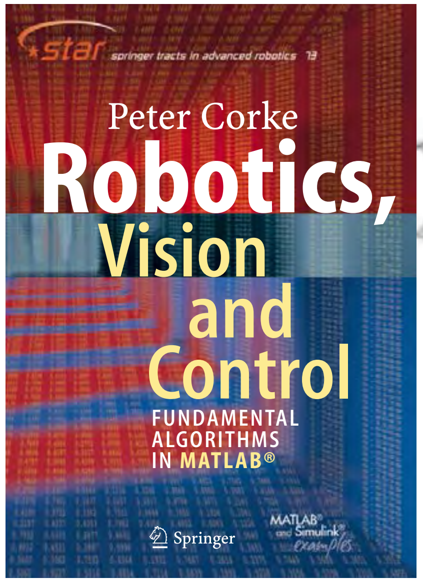
Position & motion in 2D and 3D

METR4202: Guest lecture
22 October 2014

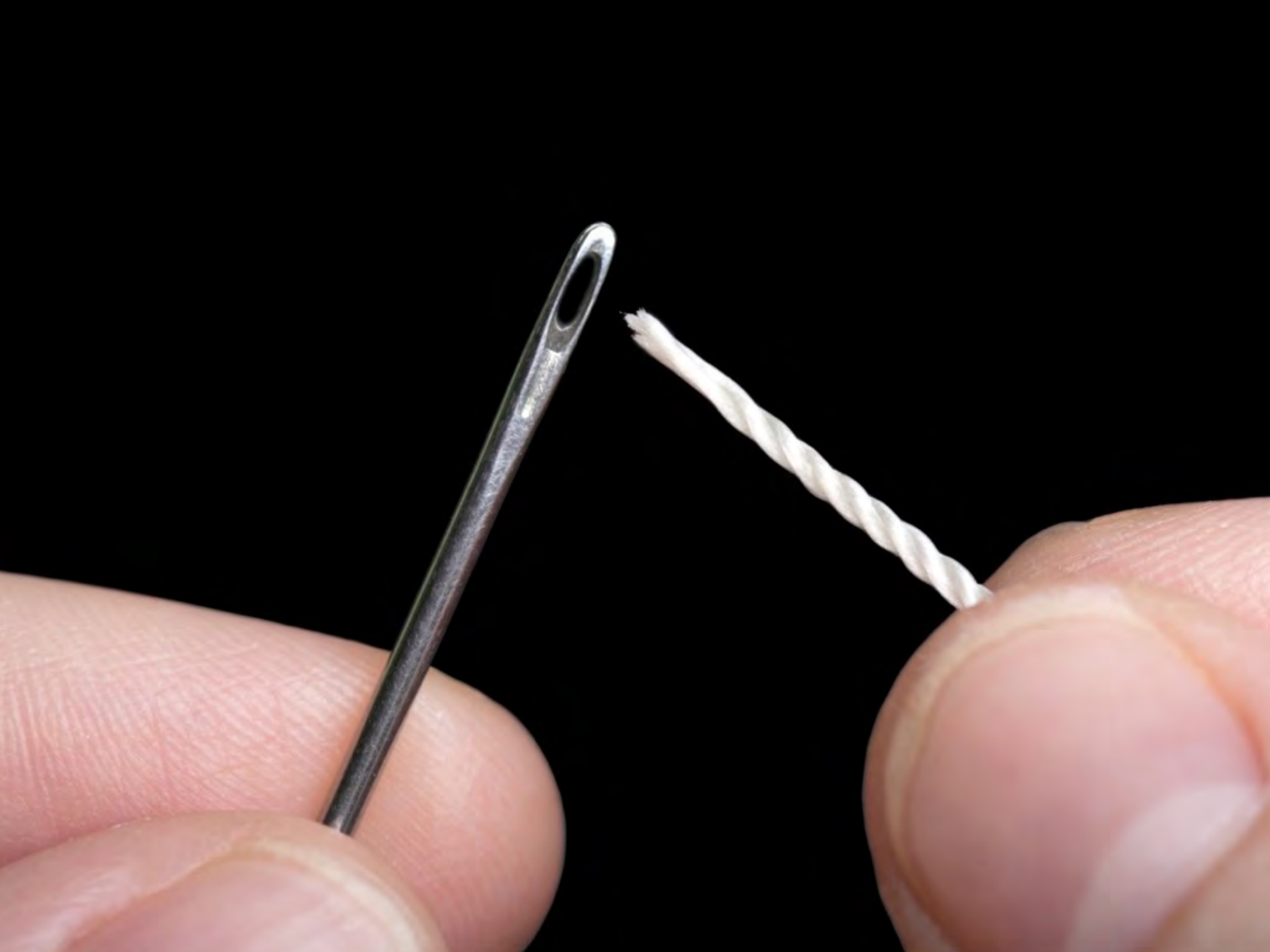
Peter Corke



Sections
11.1, 15.2

Queensland University of Technology


najoomi
.com



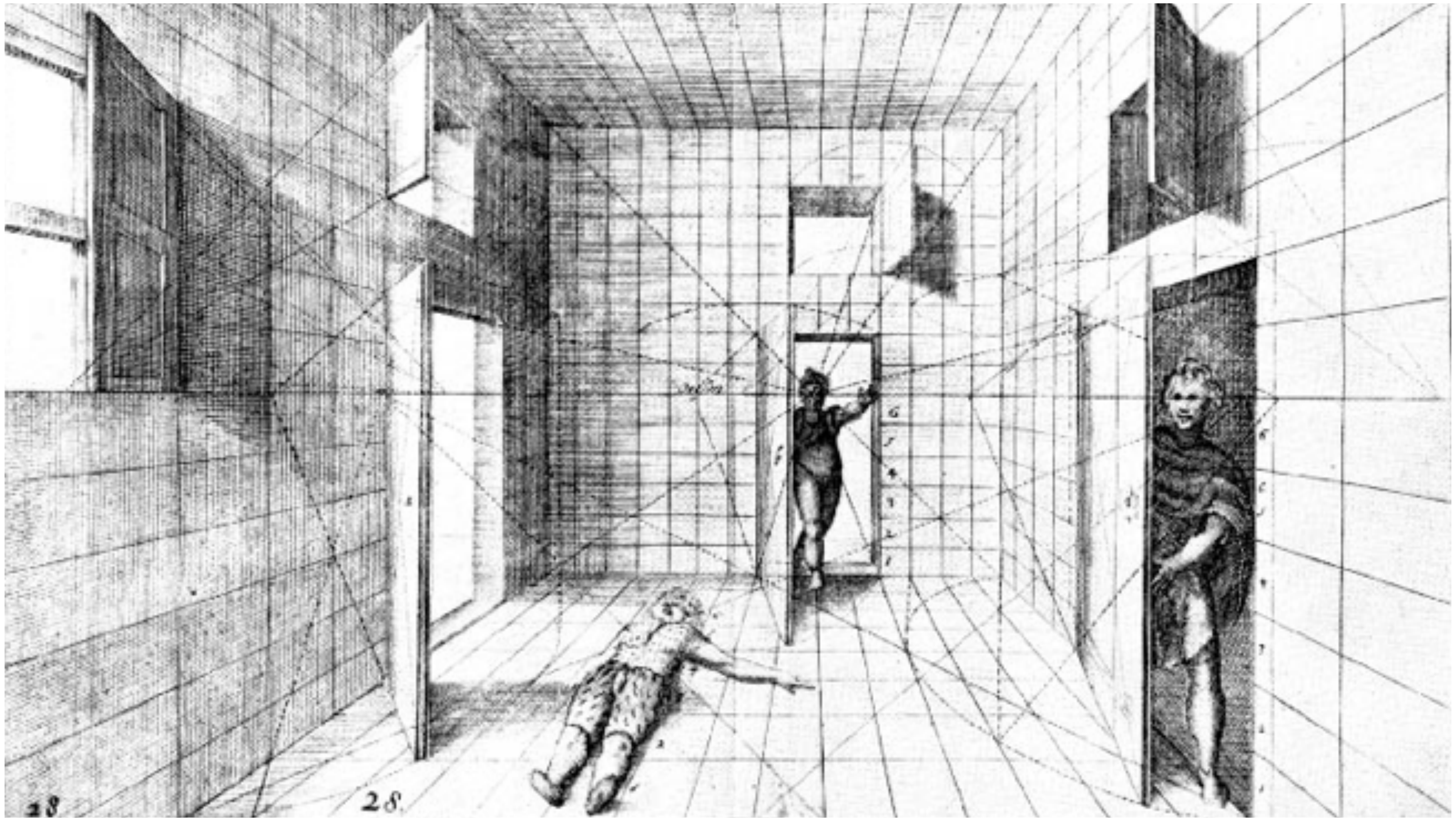
The (amazing) sense of vision

The trilobites were among the most successful of all early animals, appearing 521 million years ago and roaming the oceans for over 270 million years.

2D and 3D

Cave paintings ~40,000 years ago

Piero della Francesca (1415-1492)



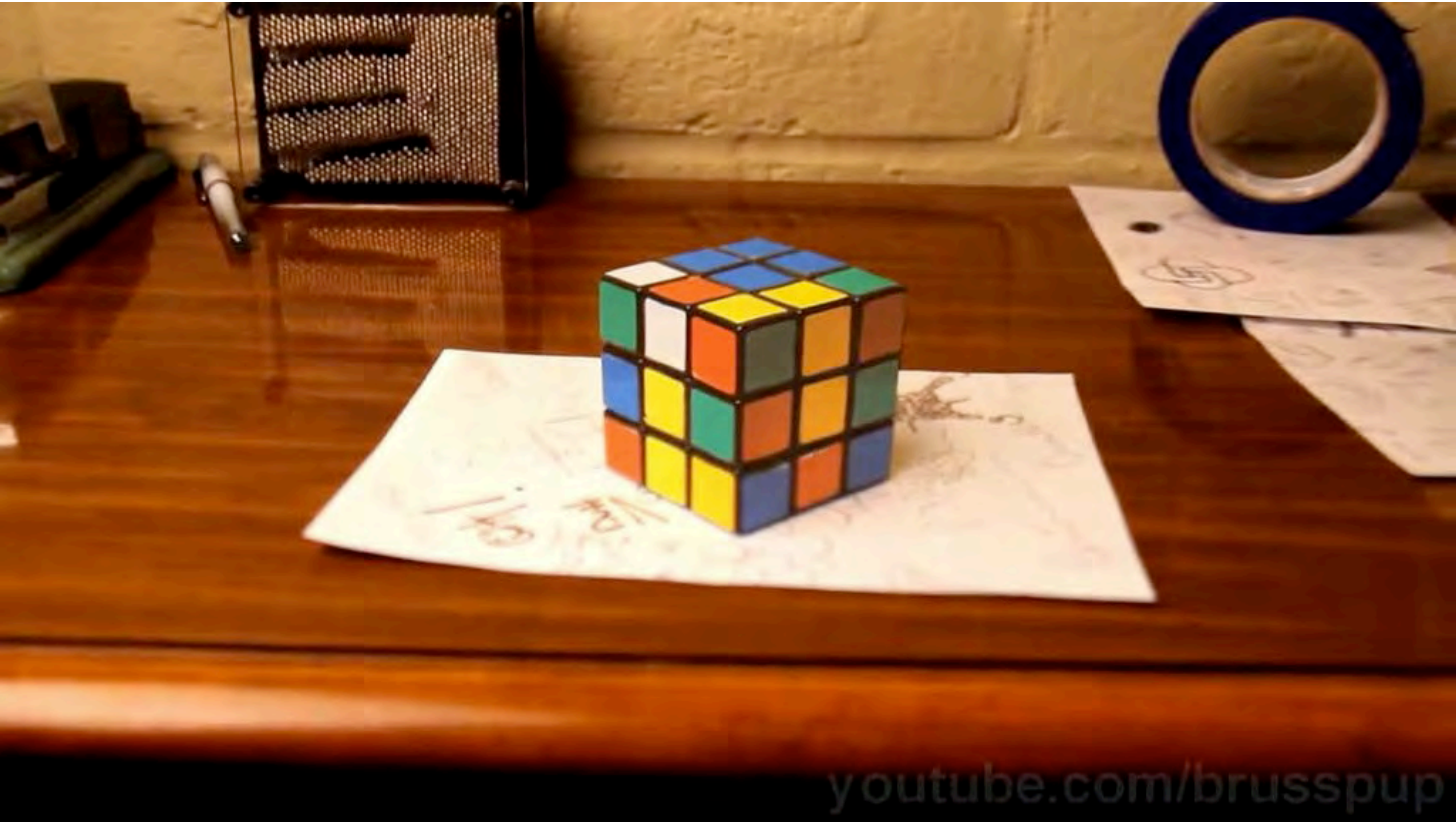
Jan Vredeman de Vries, 1604.

trompe l'oeil |ˌtrômp ˈloi|
noun (pl. **trompe l'oeils** pronunc. **same**)
visual illusion in art, esp. as used to trick the eye into perceiving a painted detail as a three-dimensional object.


TDC Netway
Jeg vil gerne ligne en under 28 år. TAK
tdc.dk/netway


DIES IRAE

Edgar Mueller
http://www.metanamorph.com


youtube.com/brusspup

Waterfall
M.C. Escher
1961

Image formation (in pictures)

points in the world

Image formation

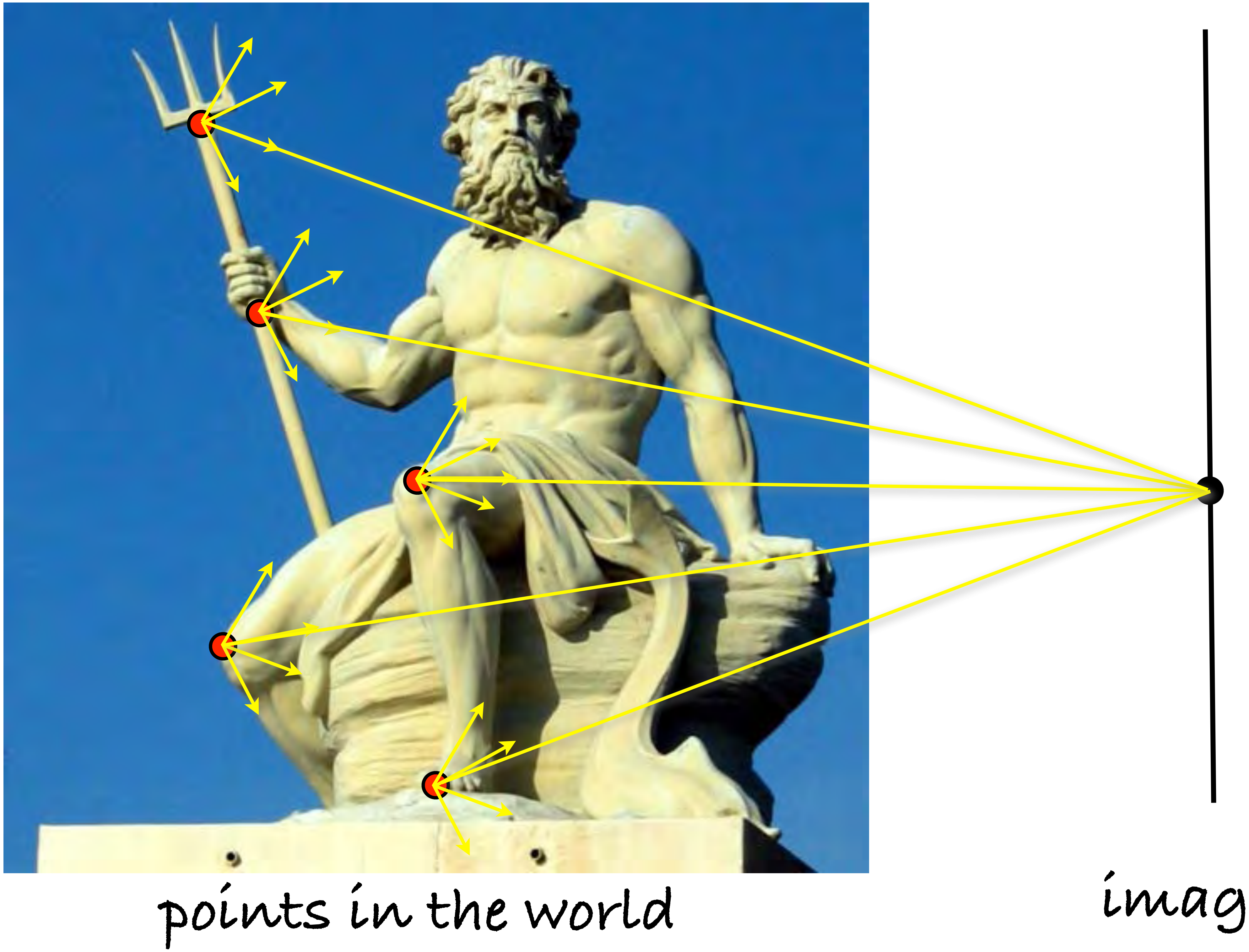
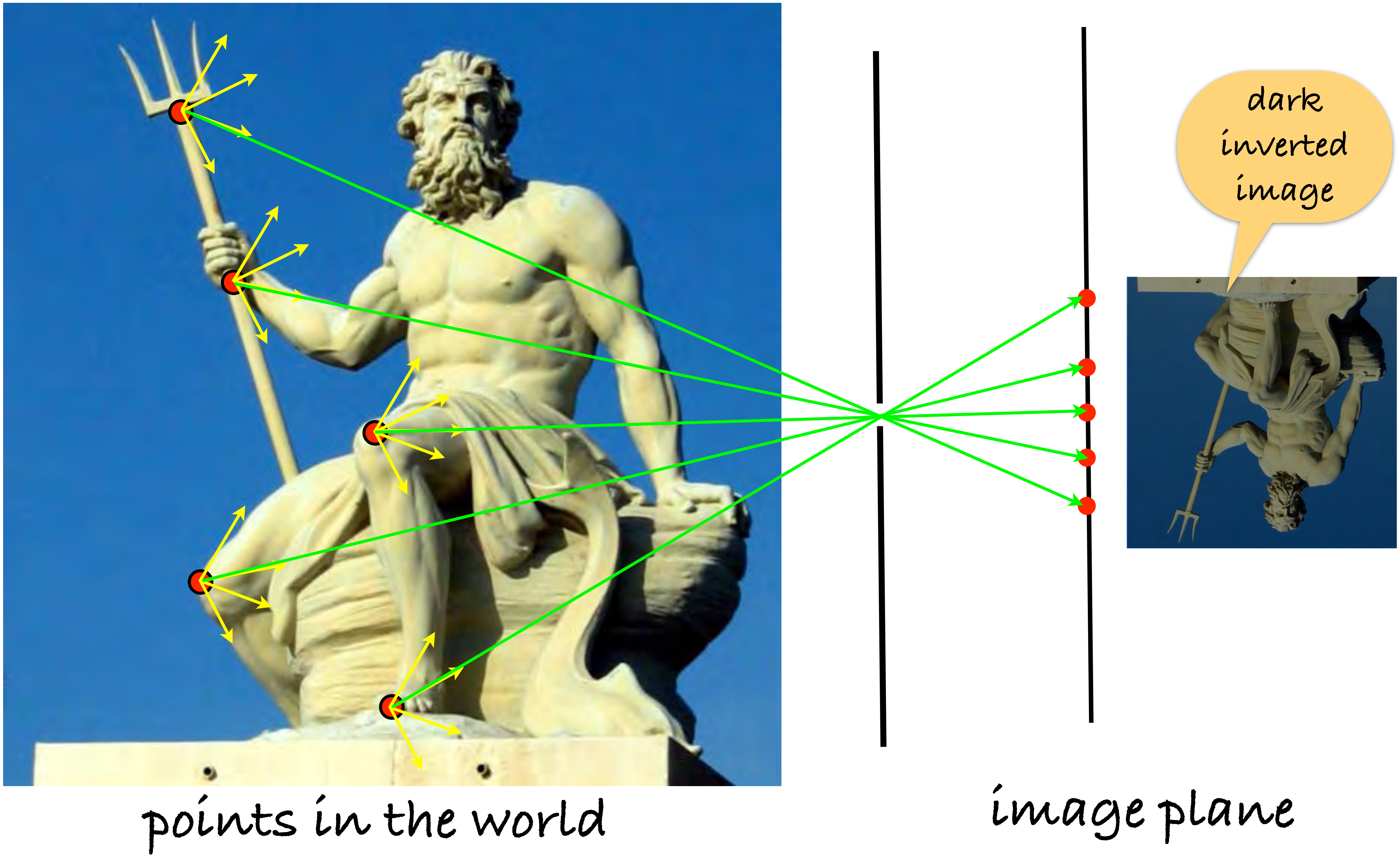
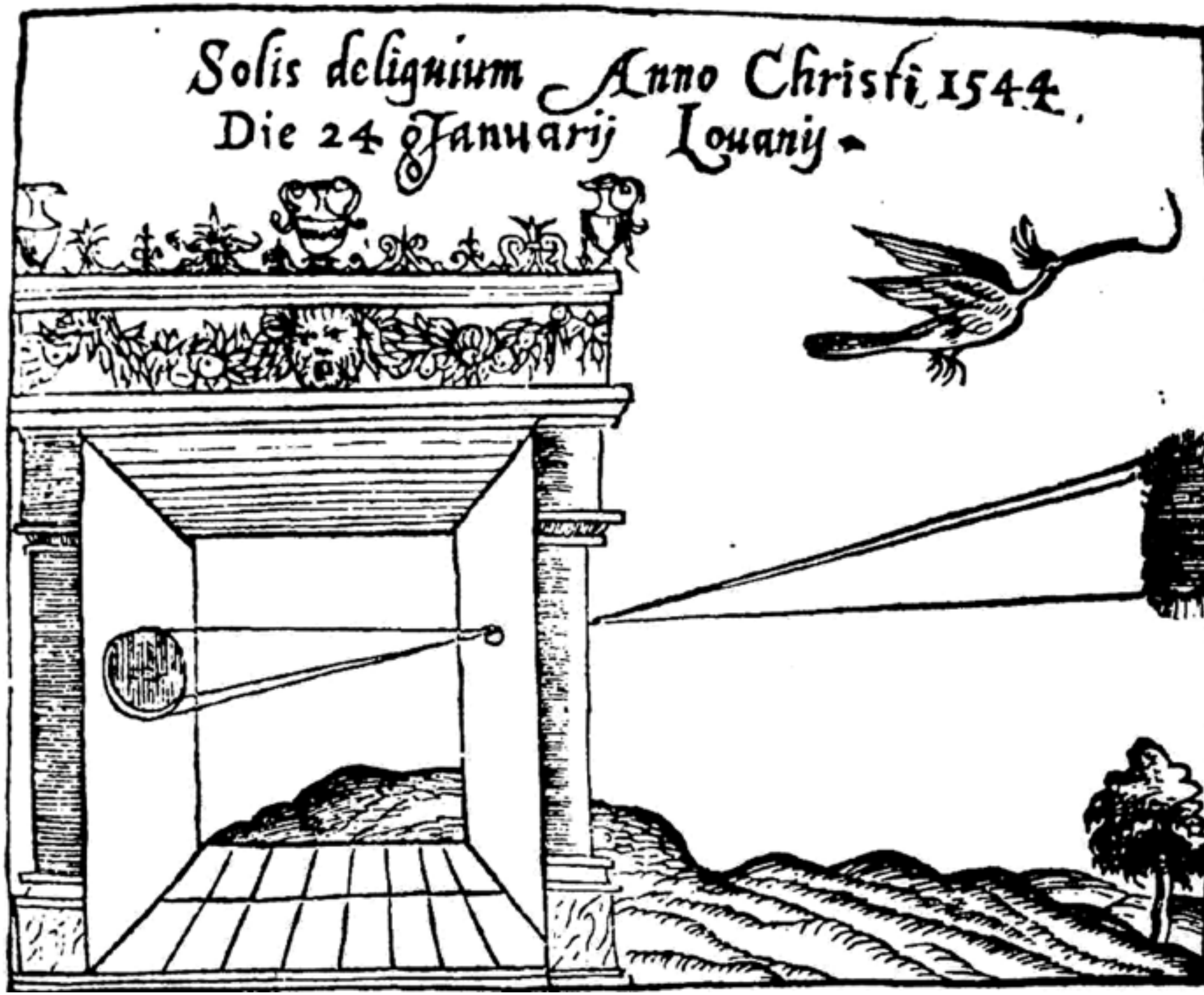


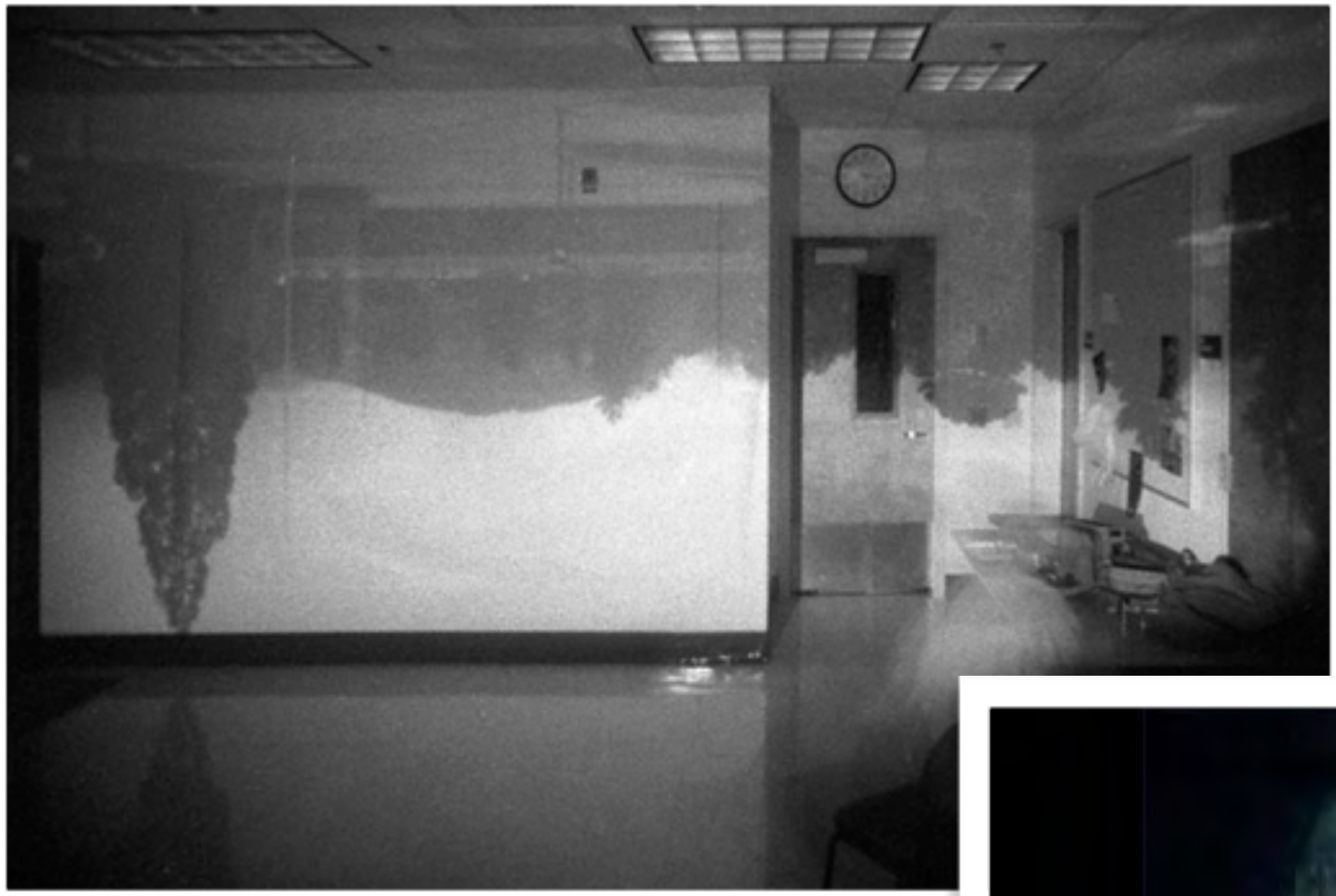
Image formation



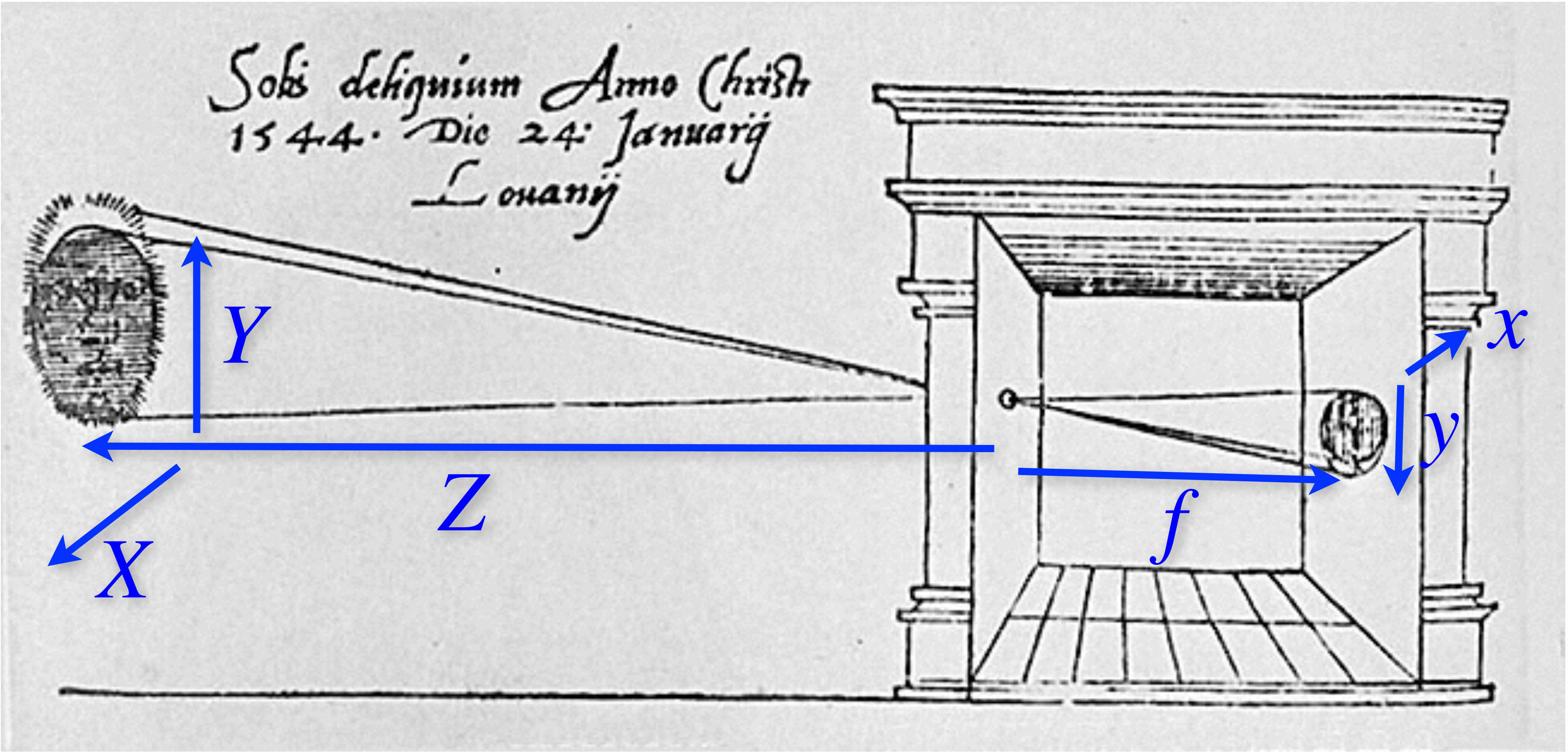
The pin hole camera



Pin hole images



Simple imaging



$$\frac{Y}{Z} = \frac{y}{f}$$

$$\frac{X}{Z} = \frac{x}{f}$$

- Image formation is the mapping of scene points to the image plane

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

$$(X, Y, Z) \mapsto (x, y)$$
$$\mathbb{R}^3 \mapsto \mathbb{R}^2$$

Image formation

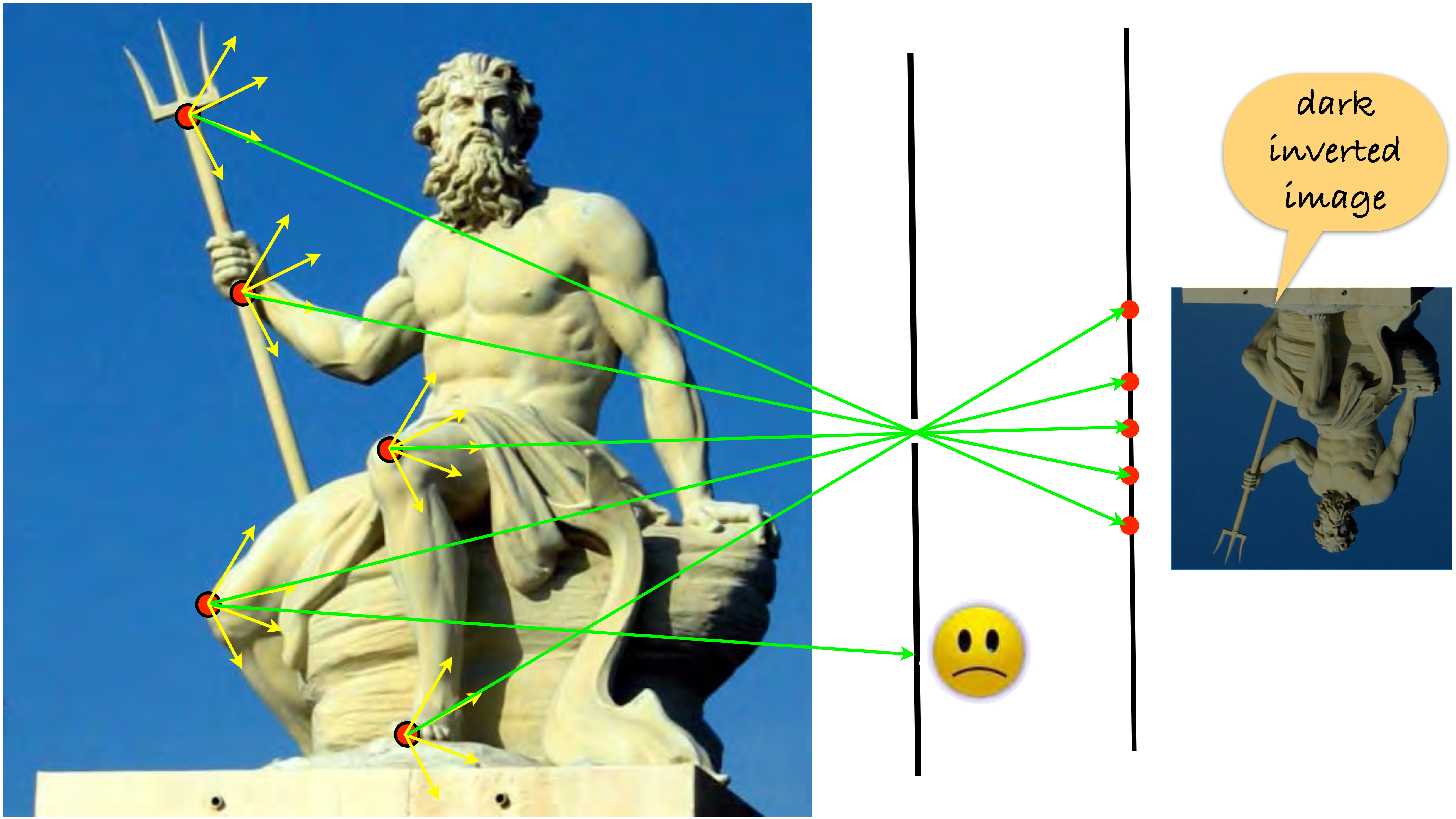
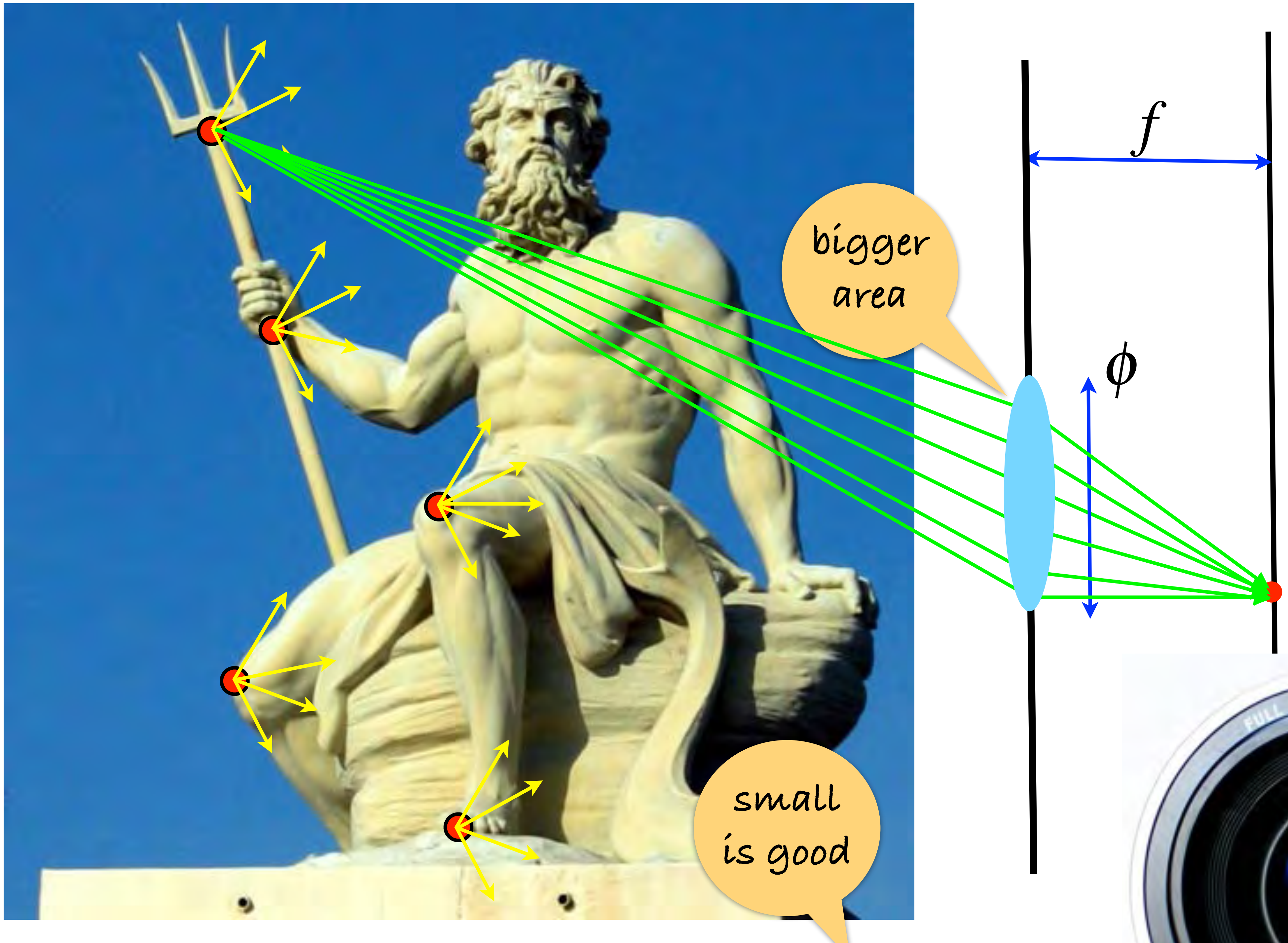
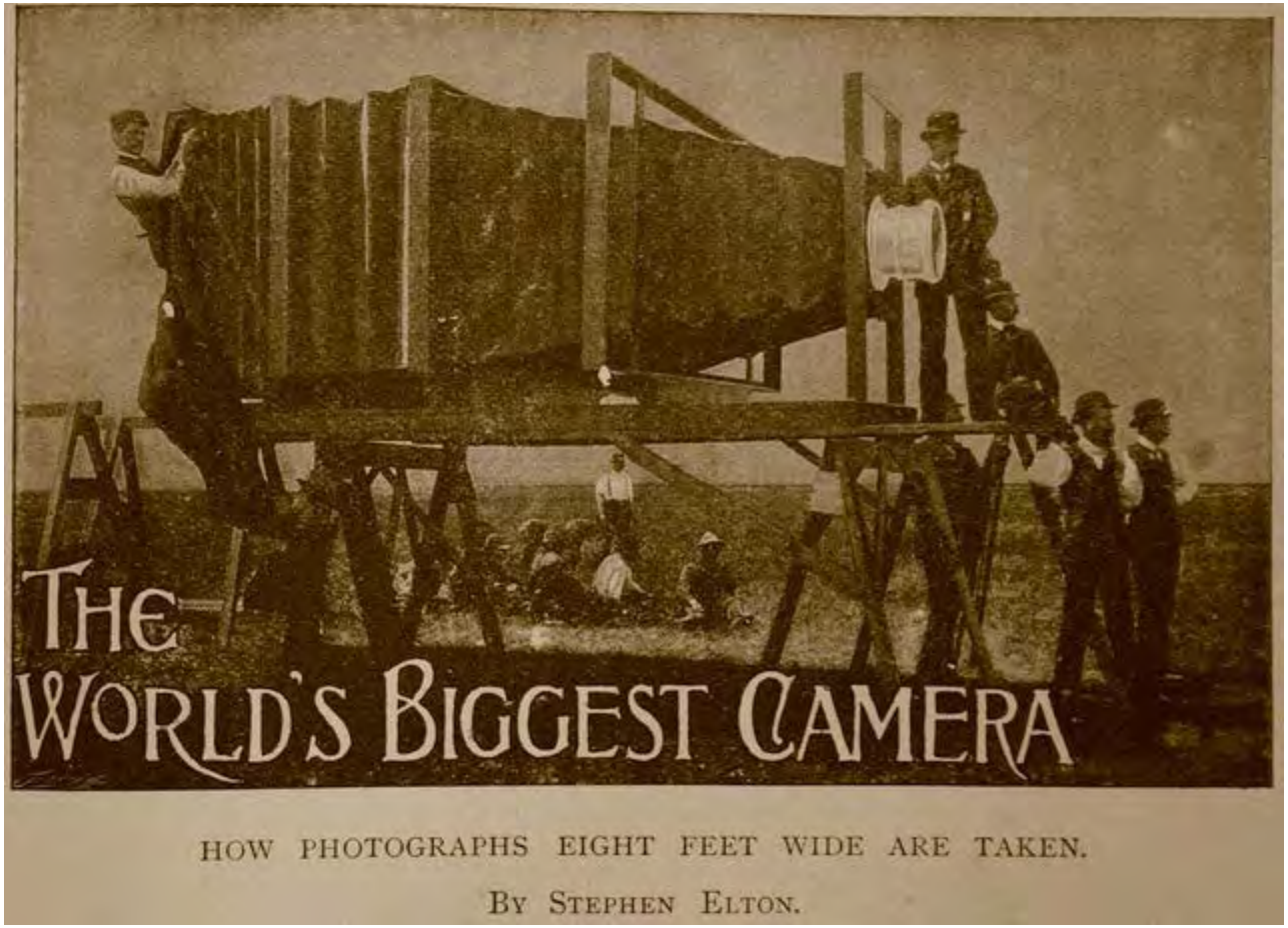


Image formation



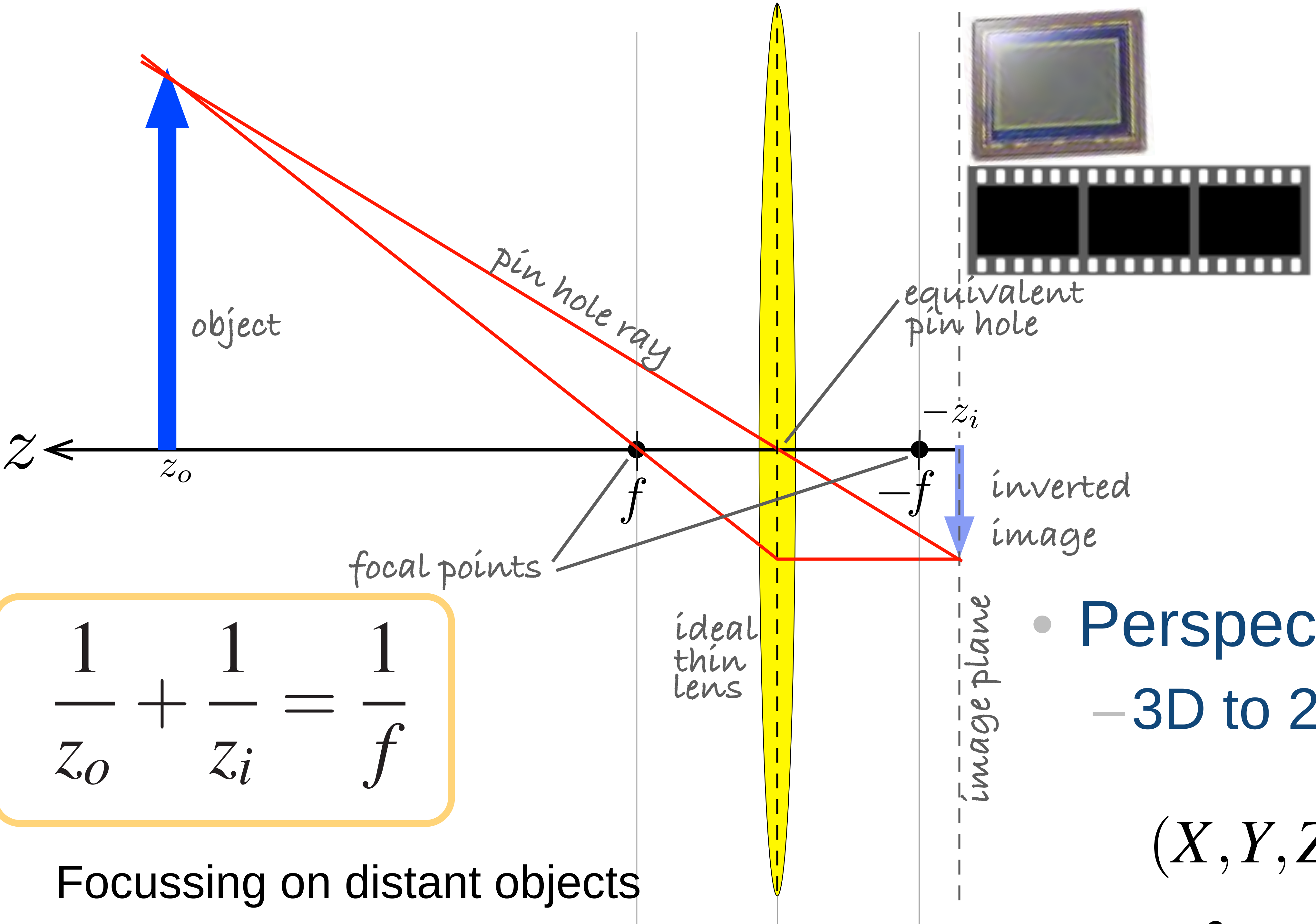
$$F = f / \phi$$

Use a lens to gather more light



George R. Lawrence 1900

Thin lens model



$$\boxed{\frac{1}{z_o}+\frac{1}{z_i}=\frac{1}{f}}$$

Focussing on distant objects

$$z_o \to \infty$$

$$z_i \to f$$

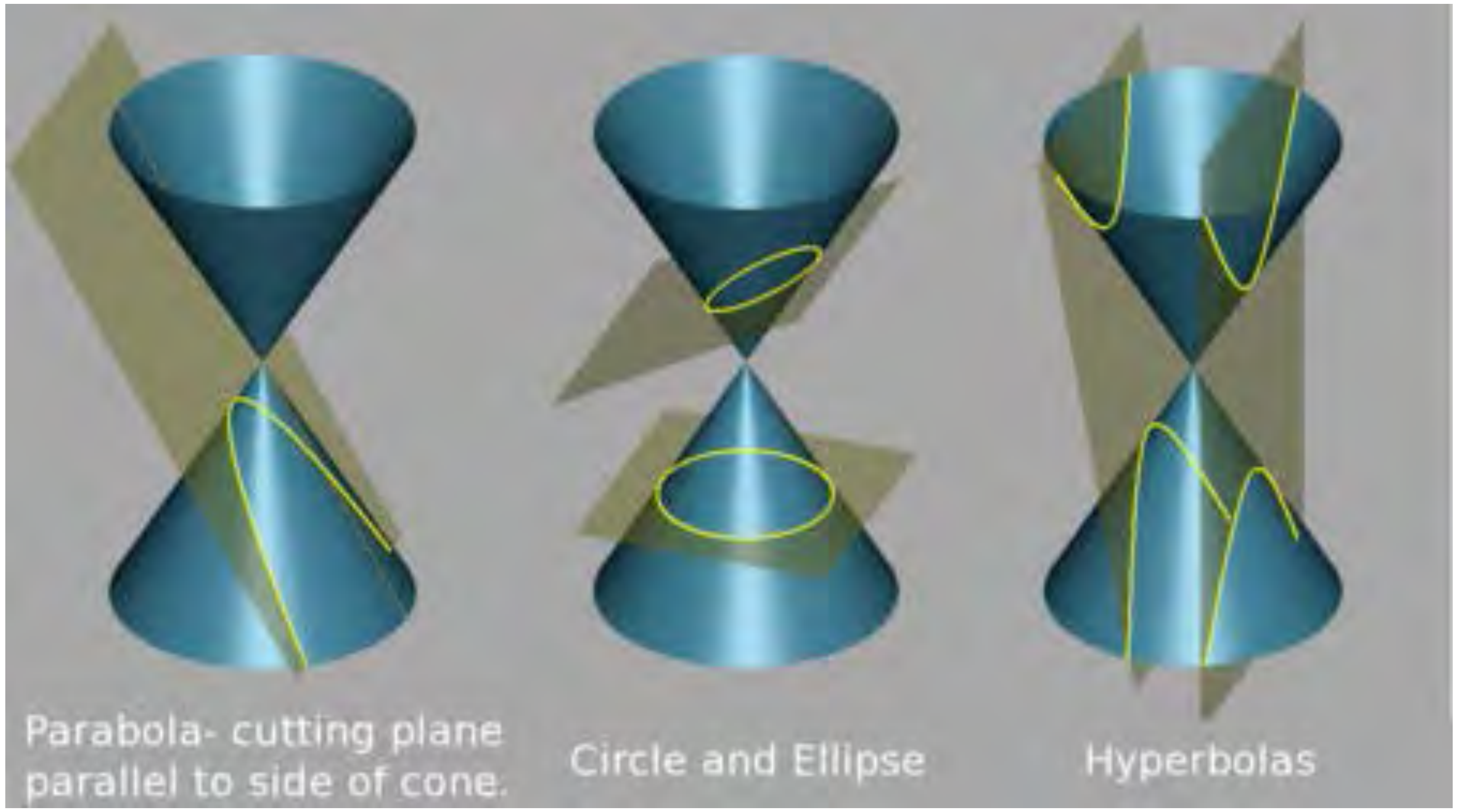
- Perspective projection
  - 3D to 2D

$$(X,Y,Z) \mapsto (x,y)$$

$$\mathbb{R}^3 \mapsto \mathbb{R}^2$$

Perspective projection

- Maps
  - Lines → lines
    - parallel lines not necessarily parallel
    - angles are not preserved
  - Conics → conics



No unique inverse

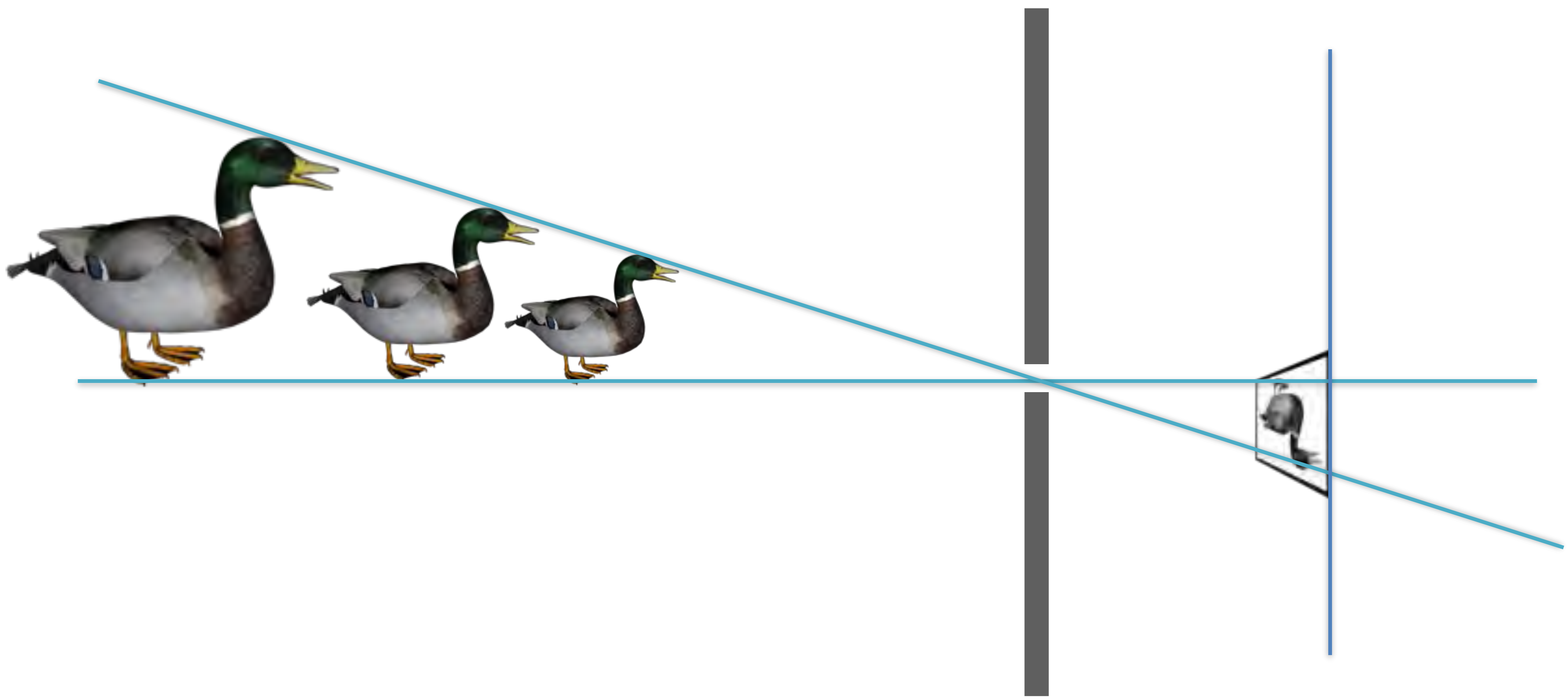


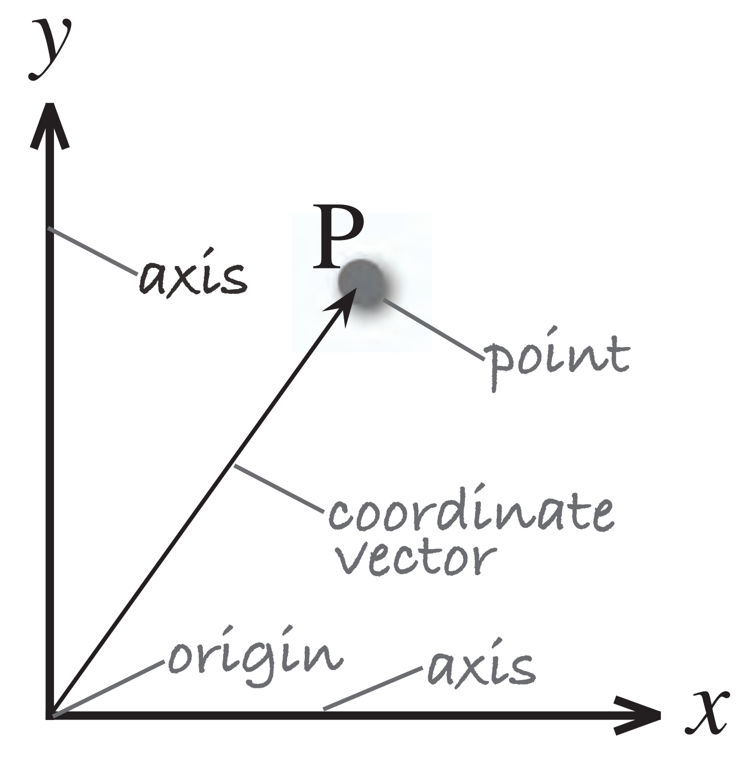

Image formation (in maths)

Homogeneous coordinates

- Cartesian → homogeneous

$$\boldsymbol{P} = (x, y) \qquad \tilde{\boldsymbol{P}} = (x, y, 1)$$

$$\boldsymbol{P} \in \mathbb{R}^2 \qquad \tilde{\boldsymbol{P}} \in \mathbb{P}^2$$



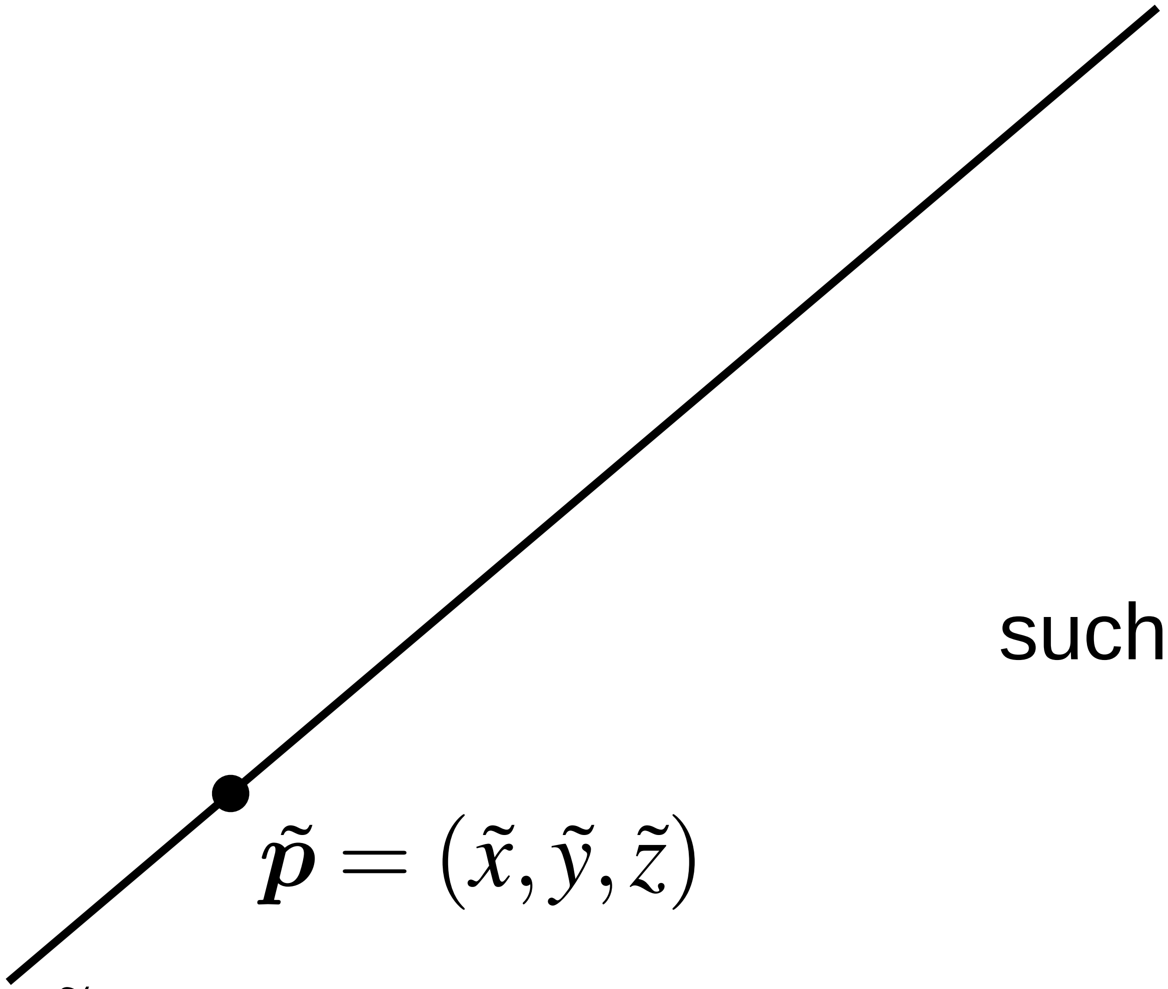
- homogeneous → Cartesian

$$\tilde{\boldsymbol{P}} = (\tilde{x}, \tilde{y}, \tilde{z}) \qquad \boldsymbol{P} = (x, y)$$

$$x = \frac{\tilde{x}}{\tilde{z}}, \; y = \frac{\tilde{y}}{\tilde{z}}$$

lines and points are duals

A line in homogeneous form



$$\tilde{\boldsymbol{p}} = (\tilde{x}, \tilde{y}, \tilde{z})$$

$$\tilde{\ell} = (l_1, l_2, l_3)$$

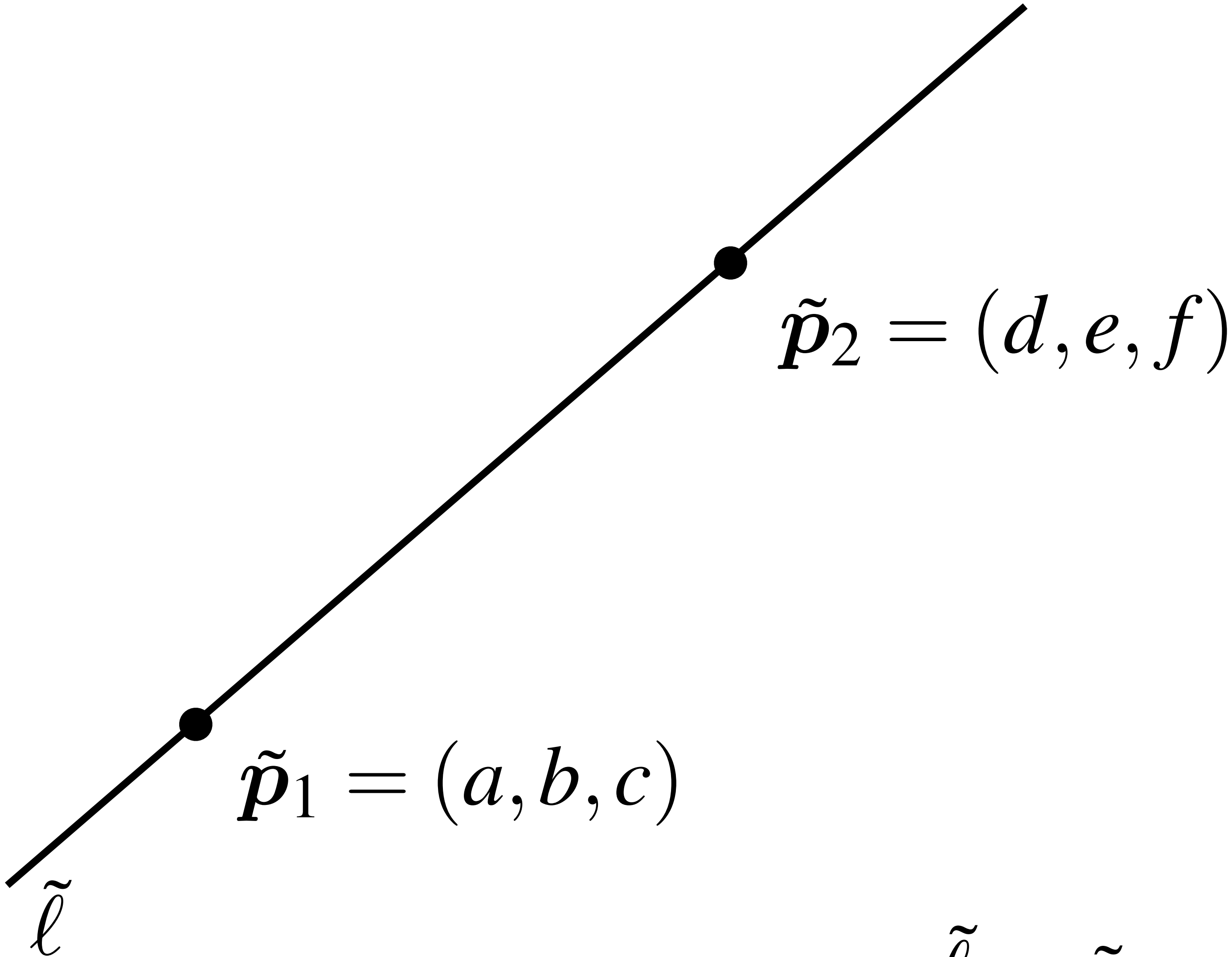
such that

$$\tilde{\ell}^T \tilde{\boldsymbol{p}} = 0$$

Point equation of a line

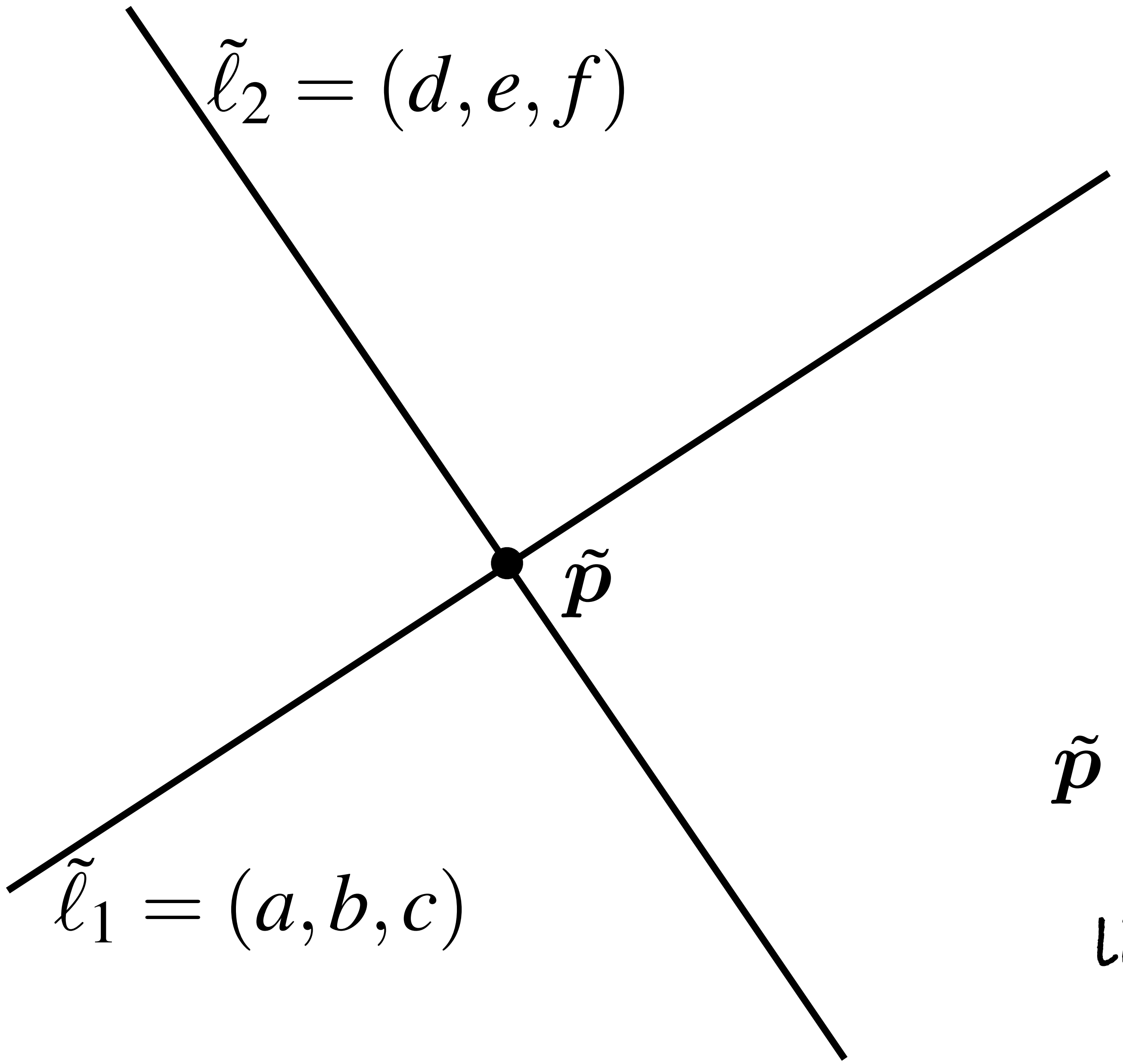
$$y = mx + c$$

Line joining points



$$\tilde{\ell} = \tilde{\boldsymbol{p}}_1 \times \tilde{\boldsymbol{p}}_2$$

Intersecting lines



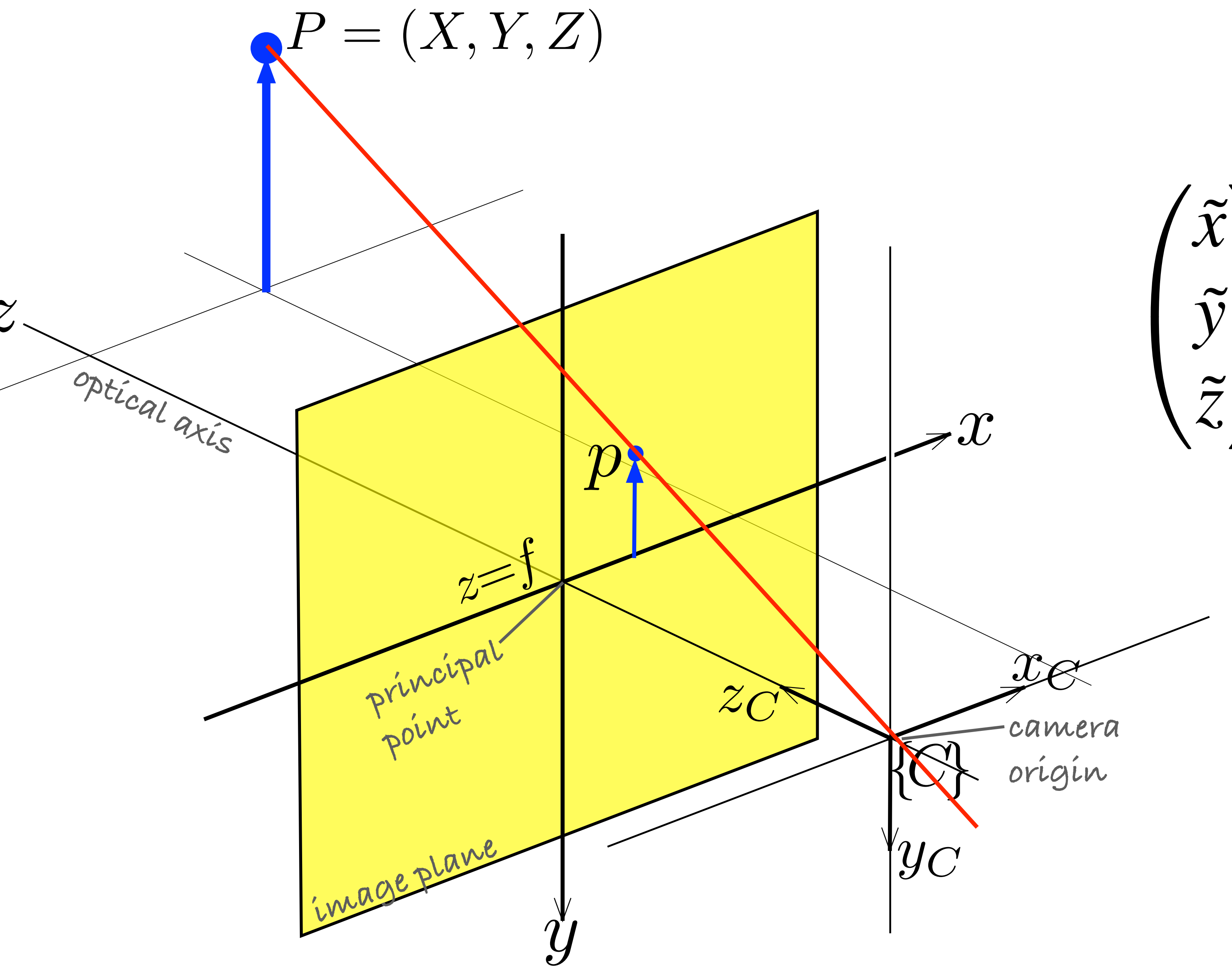
$$\tilde{\boldsymbol{p}} = \tilde{\ell}_1 \times \tilde{\ell}_2$$

line equation of a point

Central projection model

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Pin-hole model in homogeneous form



Central perspective model

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$\tilde{x} = fX, \tilde{y} = fY, \tilde{z} = Z$$

$$x = \frac{\tilde{x}}{\tilde{z}}, \; y = \frac{\tilde{y}}{\tilde{z}}$$

$$\Rightarrow x = \frac{fX}{Z}, y = \frac{fY}{Z}$$

- Perspective transformation, with the pesky divide by Z, is **linear** in homogeneous coordinate form.

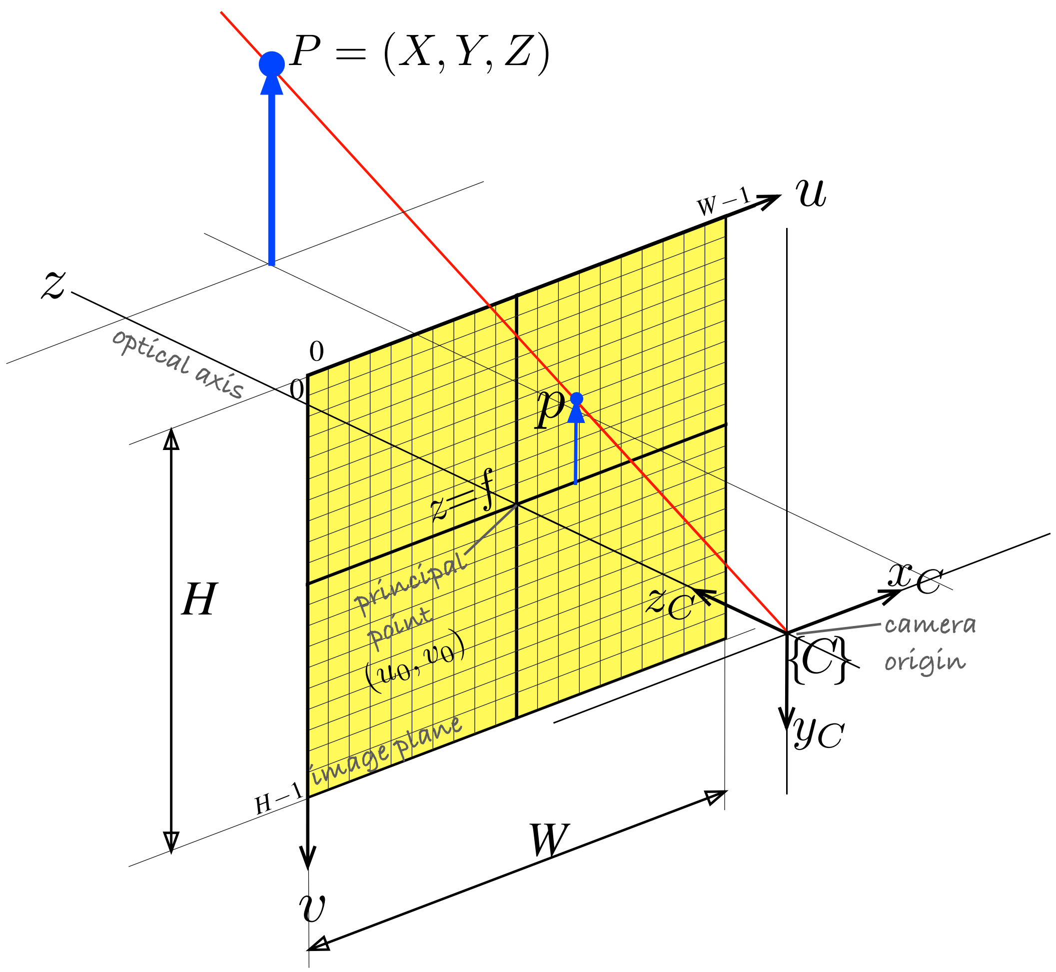
$$\begin{pmatrix}\tilde{x}\\ \tilde{y}\\ \tilde{z}\end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0\\ 0 & f & 0 & 0\\ 0 & 0 & 1 & 0\end{pmatrix}\begin{pmatrix}X\\ Y\\ Z\\ 1\end{pmatrix}$$

3D to 2D

scaling/ zooming

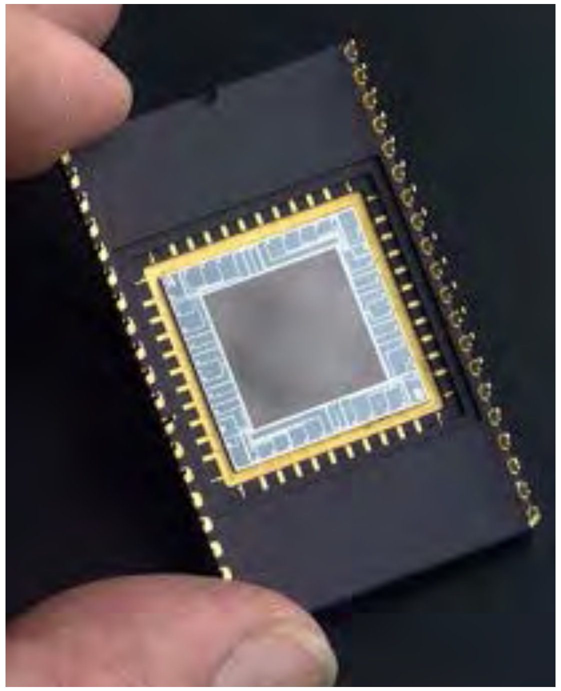
$$\begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\end{pmatrix}\begin{pmatrix} f & 0 & 0 & 0\\ 0 & f & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$$

Change of coordinates



$$u = \frac{x}{\rho_u} + u_0$$

$$v = \frac{y}{\rho_v} + v_0$$

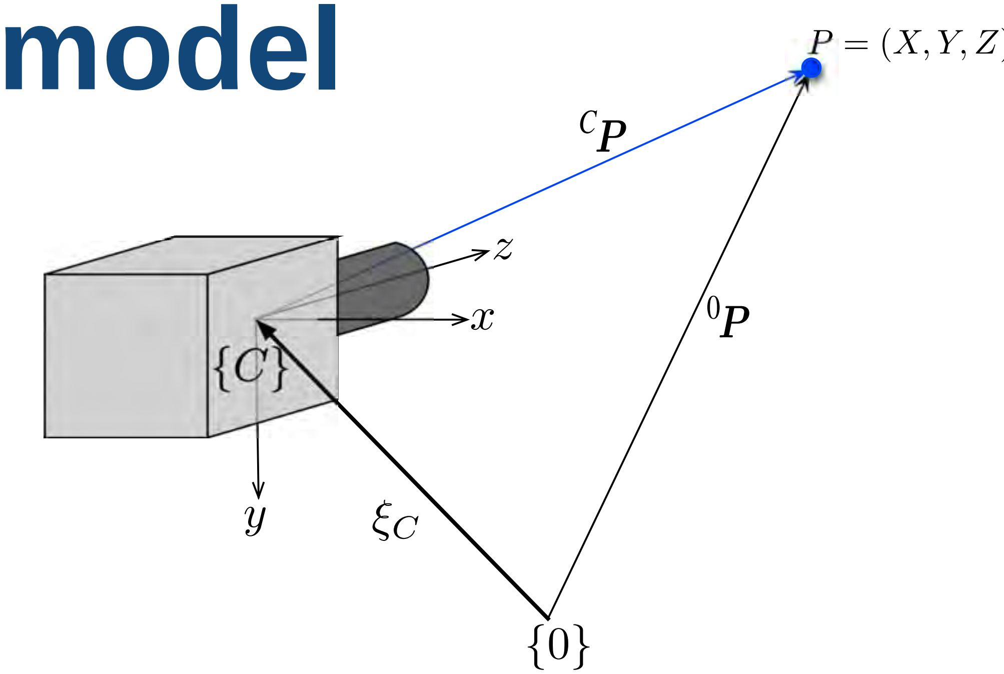


$$\begin{pmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{w} \end{pmatrix} = \begin{pmatrix} \frac{1}{\rho_u} & 0 & u_0 \\ 0 & \frac{1}{\rho_v} & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix}$$

- scale from metres to pixels
- shift the origin to top left corner

$$p = \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \tilde{u}/\tilde{w} \\ \tilde{v}/\tilde{w} \end{pmatrix}$$

Complete camera model



$$\begin{pmatrix}\tilde{x}\\ \tilde{y}\\ \tilde{z}\end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0\\ 0 & f & 0 & 0\\ 0 & 0 & 1 & 0\end{pmatrix}\begin{pmatrix}X\\ Y\\ Z\\ 1\end{pmatrix}$$

extrinsic parameters

$$\begin{pmatrix}\tilde{u}\\ \tilde{v}\\ \tilde{w}\end{pmatrix} = \underbrace{\underbrace{\begin{pmatrix} \frac{1}{\rho_u} & 0 & u_0\\ 0 & \frac{1}{\rho_v} & v_0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix} f & 0 & 0 & 0\\ 0 & f & 0 & 0\\ 0 & 0 & 1 & 0\end{pmatrix}}_{\mathbf{K}}\begin{pmatrix}\mathbf{R} & t\\ \mathbf{0}_{1\times 3} & 1\end{pmatrix}^{-1}}_{\mathbf{C}}\begin{pmatrix}X\\ Y\\ Z\\ 1\end{pmatrix}$$

intrinsic parameters

camera matrix

Camera matrix

- Mapping points from the world to an image (pixel) coordinate is simply a matrix multiplication using homogeneous coordinates

$$\begin{pmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{w} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$u = \frac{\tilde{u}}{\tilde{w}}, v = \frac{\tilde{v}}{\tilde{w}}$$

Scale invariance

- Consider an arbitrary scalar scale factor

$$\begin{pmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{w} \end{pmatrix} = \lambda \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

- $\tilde{u}, \tilde{v}, \tilde{w}$ will all be scaled by λ
- but

$$u = \frac{\tilde{u}}{\tilde{w}}, v = \frac{\tilde{v}}{\tilde{w}}$$

- so the result is unchanged

Normalized camera matrix

- Since scale factor is arbitrary we can fix the value of one element, typically C(3,4) to one.

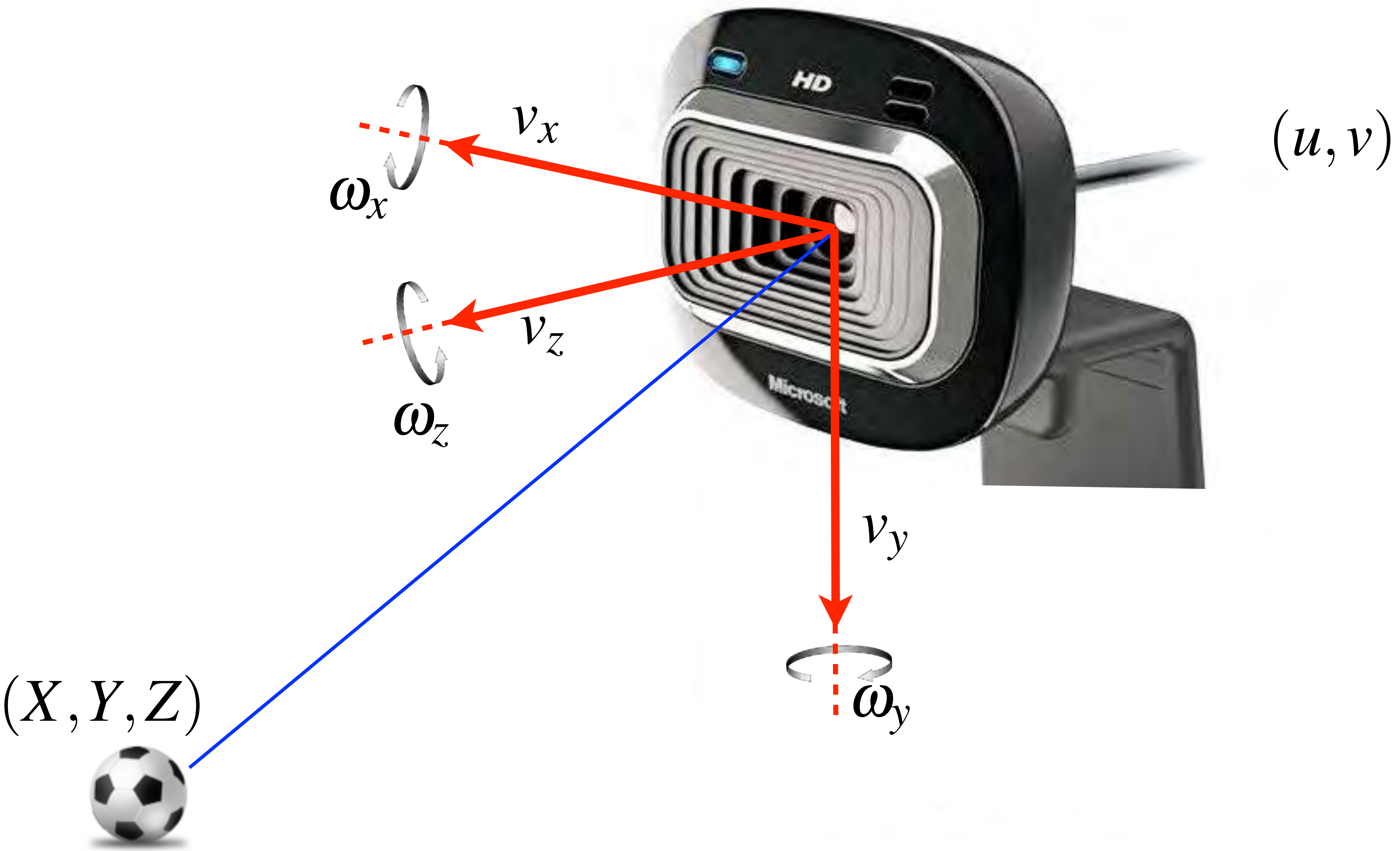
focal length
pixel size
camera position &
orientation

$$\begin{pmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{w} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$u = \frac{\tilde{u}}{\tilde{w}}, v = \frac{\tilde{v}}{\tilde{w}}$$

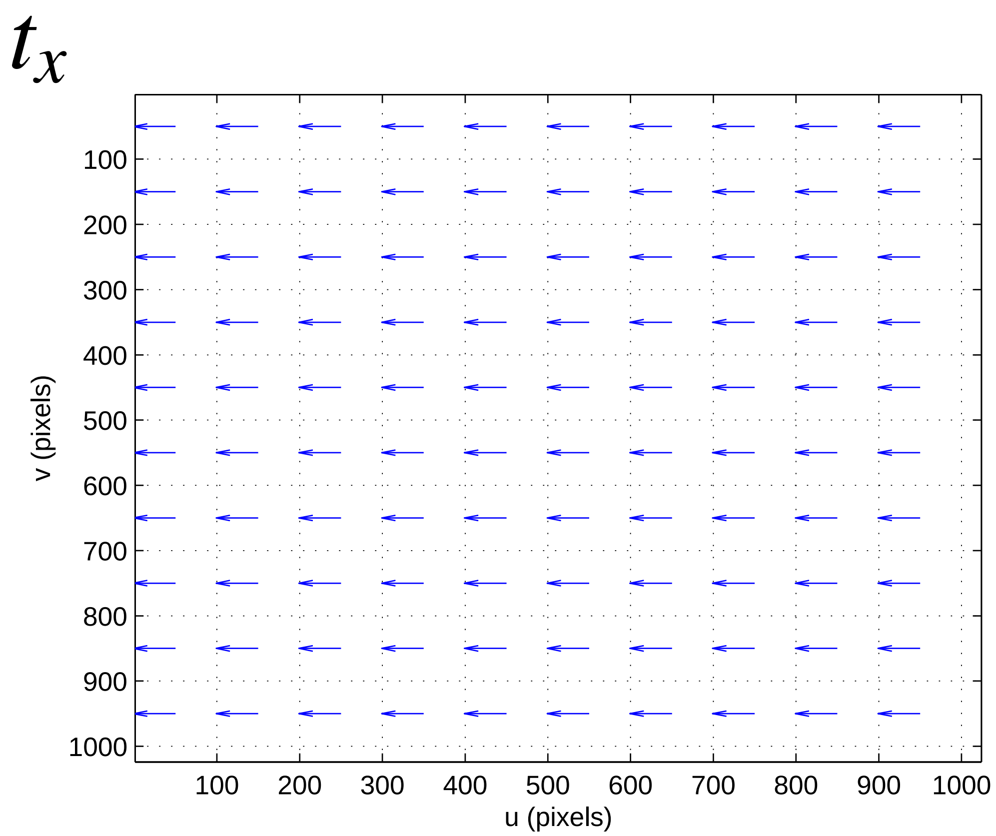
Consider a moving camera

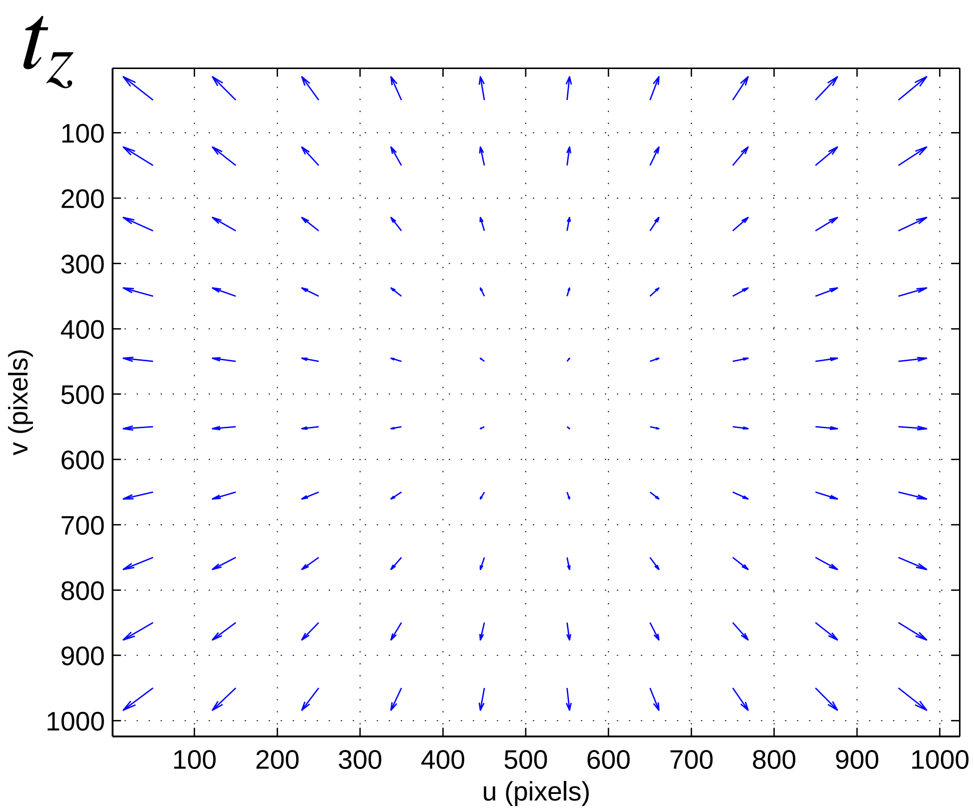
Motion of a camera

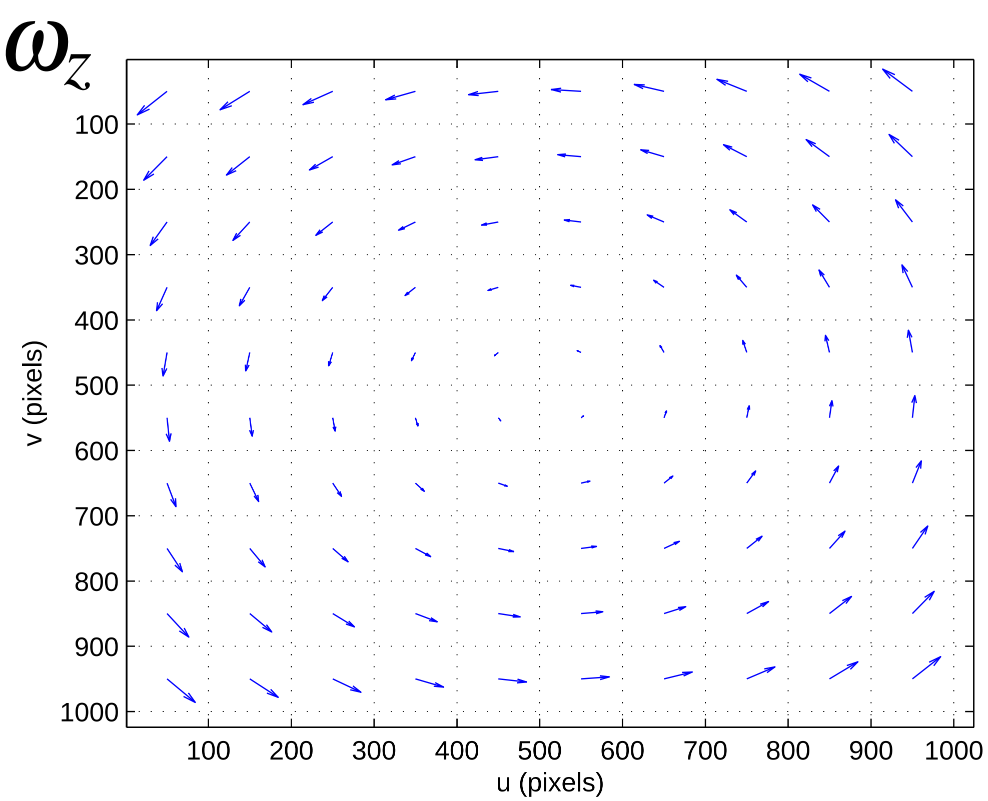


- 6 degrees of freedom
  - translate along x, y, z
  - rotate about x, y, z

Optical flow patterns



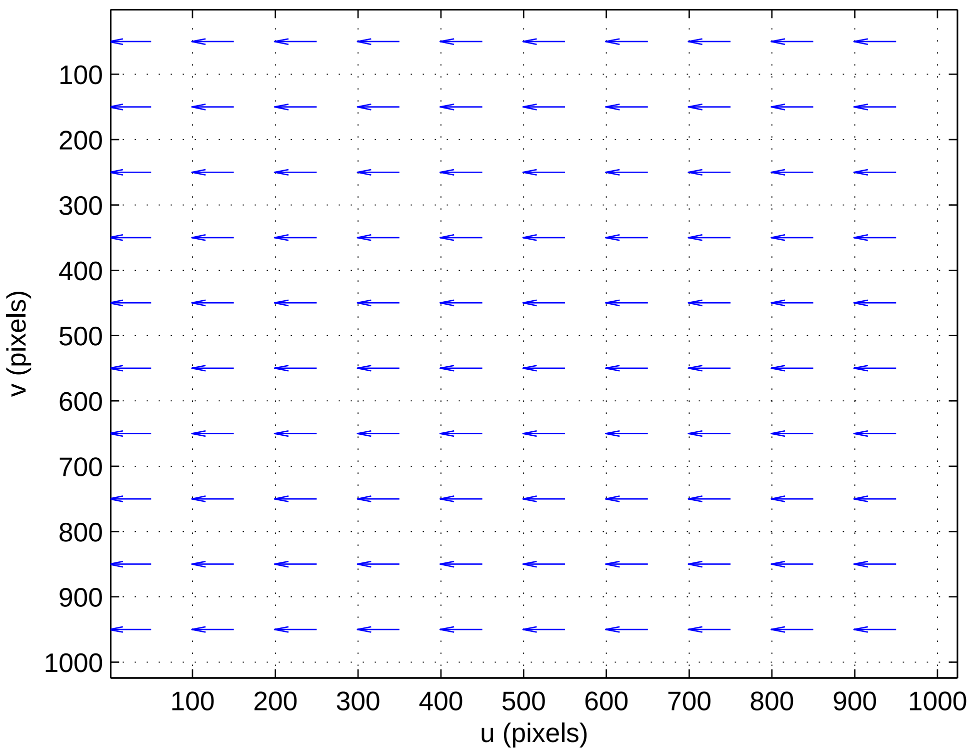




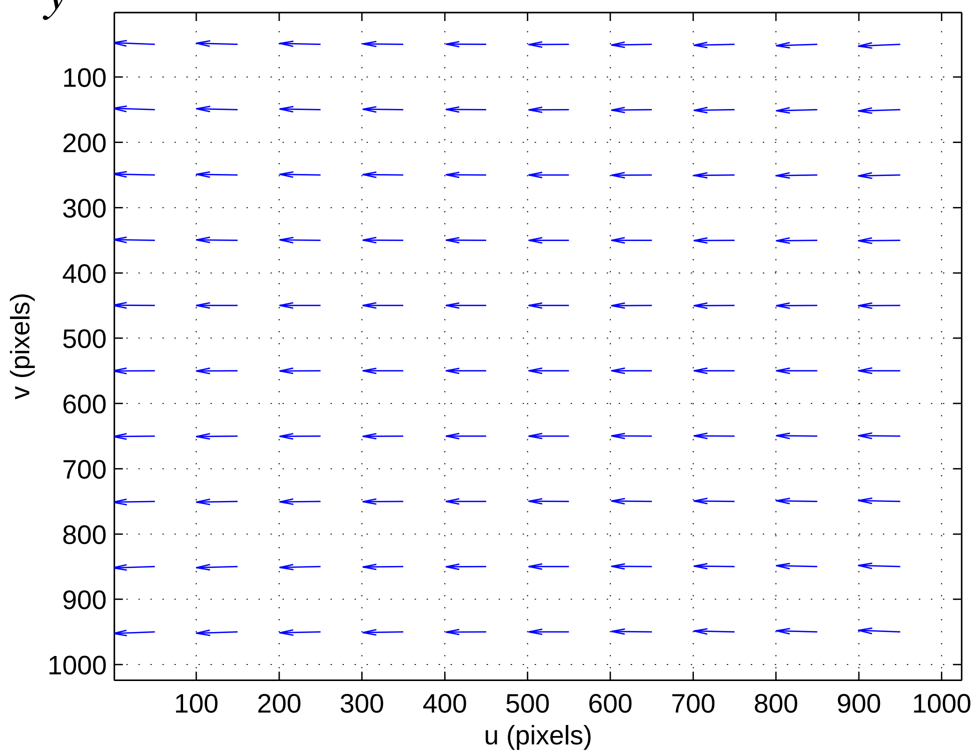
- Pixel motion depends on
  - pixel position
  - camera motion

Optical flow patterns

t_x



ω_y



- Rotation and translation in x and y cause very similar motion

Optical flow equation

speed of camera

speed of pixel

$$\begin{pmatrix} \dot{\bar{u}} \\ \dot{\bar{v}} \end{pmatrix} = \underbrace{\begin{pmatrix} -\frac{f}{\rho_u Z} & 0 & \frac{\bar{u}}{Z} & \frac{\rho_u \bar{u}\bar{v}}{f} & -\frac{f^2+\rho_u^2\bar{u}^2}{\rho_u f} & \bar{v} \\ 0 & -\frac{f}{\rho_v Z} & \frac{\bar{v}}{Z} & \frac{f^2+\rho_v^2\bar{v}^2}{\rho_v f} & -\frac{\rho_v \bar{u}\bar{v}}{f} & -\bar{u} \end{pmatrix}}_{J_p} \begin{pmatrix} v_x \\ v_y \\ v_z \\ \omega_x \\ \omega_y \\ \omega_z \end{pmatrix}$$

$(\bar{u}, \bar{v})$ are distances from principal point

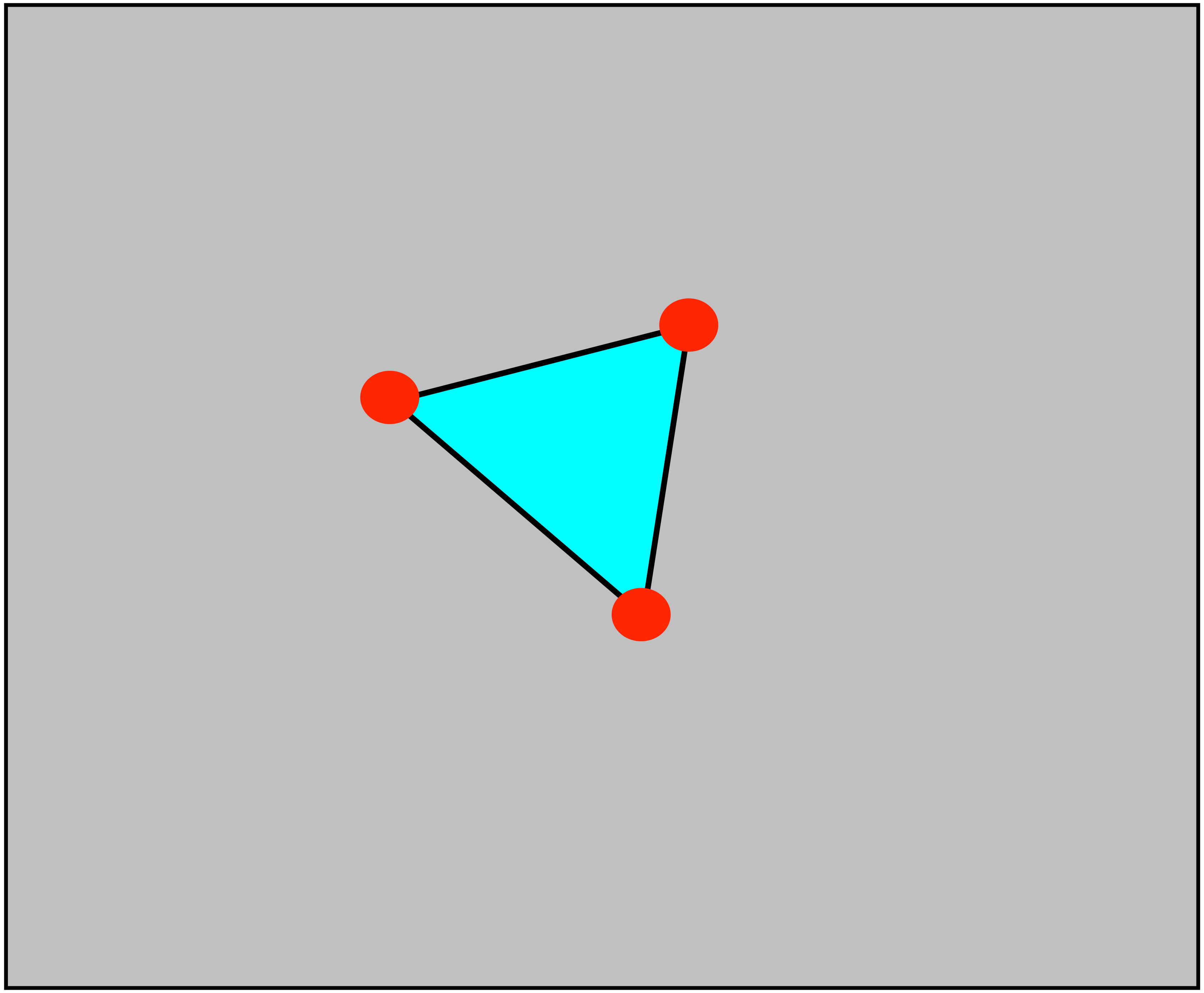
Motion of multiple points

- Consider the case of three points, in matrix form

$$\begin{pmatrix} \dot{u}_1 \\ \dot{v}_1 \\ \dot{u}_2 \\ \dot{v}_2 \\ \dot{u}_3 \\ \dot{v}_3 \end{pmatrix} = \begin{pmatrix} J_{p_1} \\ J_{p_2} \\ J_{p_3} \end{pmatrix} \boldsymbol{\nu}$$

$$\boldsymbol{\nu} = (v_x, v_y, v_z, \omega_x, \omega_y, \omega_z) \in \mathbb{R}^6$$

Desired view



Current view

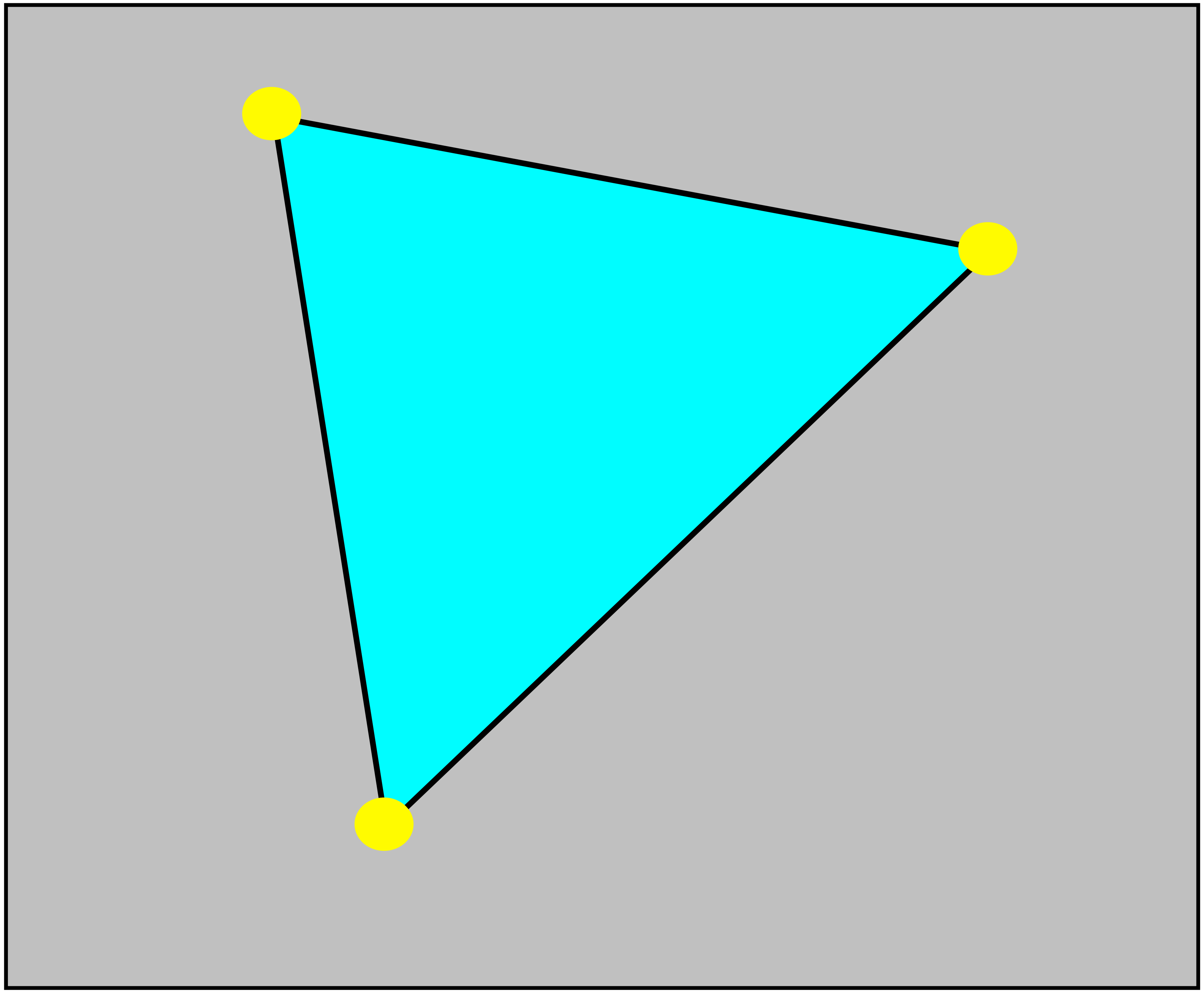


Image plane motion

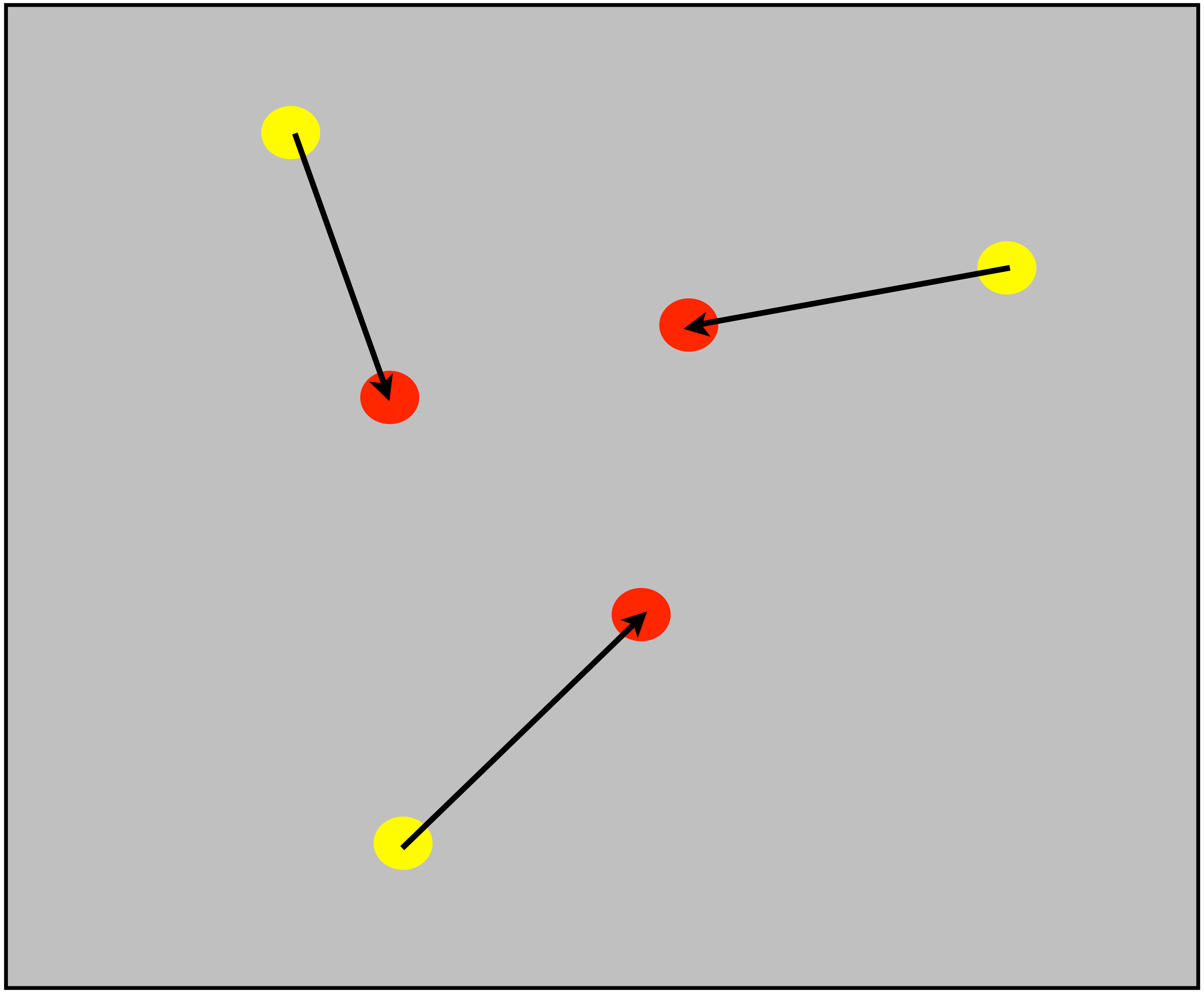
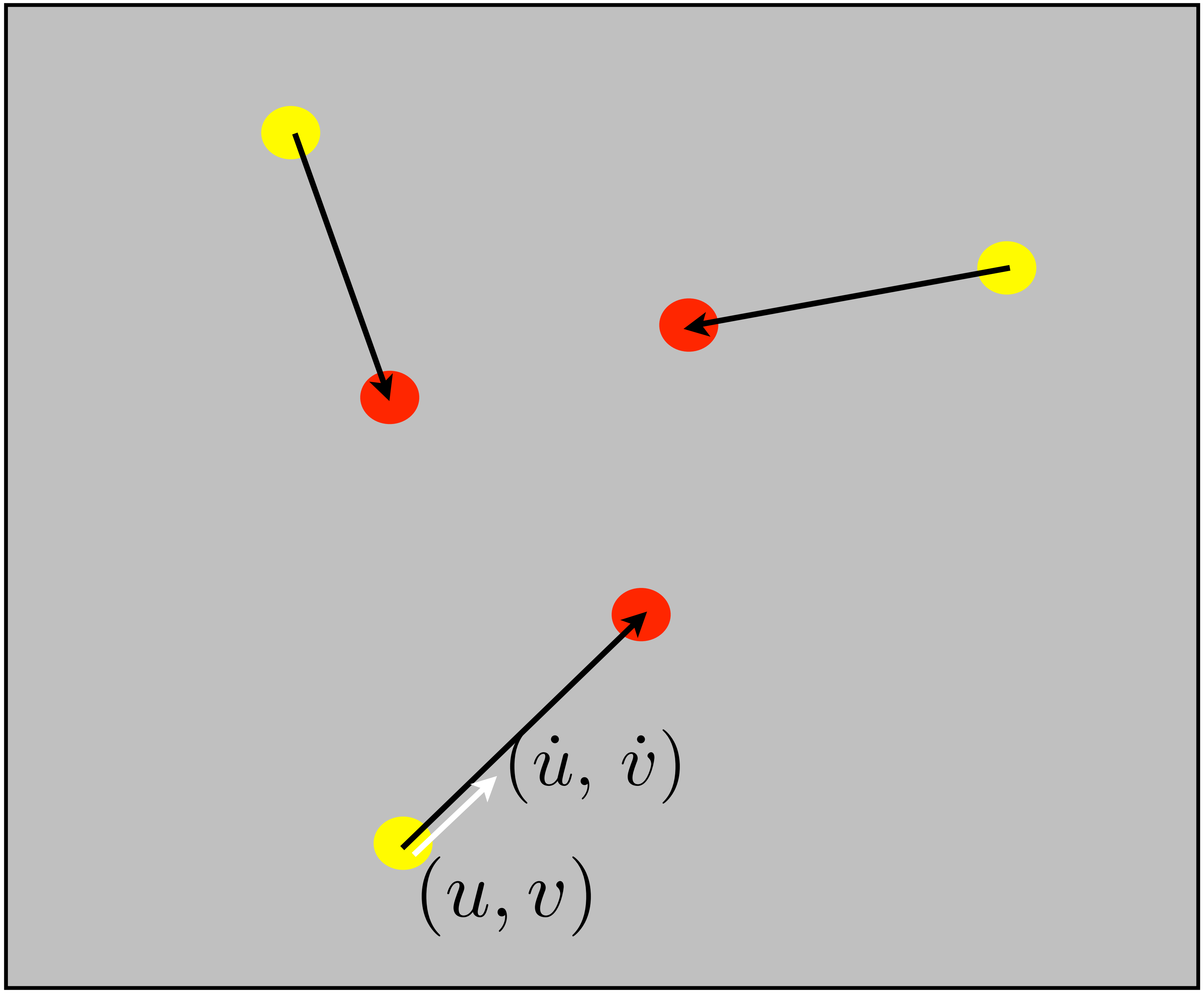


Image plane motion



Inverting the problem

$$\begin{pmatrix} \dot{u}_1 \\ \dot{v}_1 \\ \dot{u}_2 \\ \dot{v}_2 \\ \dot{u}_3 \\ \dot{v}_3 \end{pmatrix} = \begin{pmatrix} J_{p_1} \\ J_{p_2} \\ J_{p_3} \end{pmatrix} \nu$$

$$\nu = \begin{pmatrix} J_{p_1} \\ J_{p_2} \\ J_{p_3} \end{pmatrix}^{-1} \begin{pmatrix} \dot{u}_1 \\ \dot{v}_1 \\ \dot{u}_2 \\ \dot{v}_2 \\ \dot{u}_3 \\ \dot{v}_3 \end{pmatrix}$$

required point velocity to move from p to p*

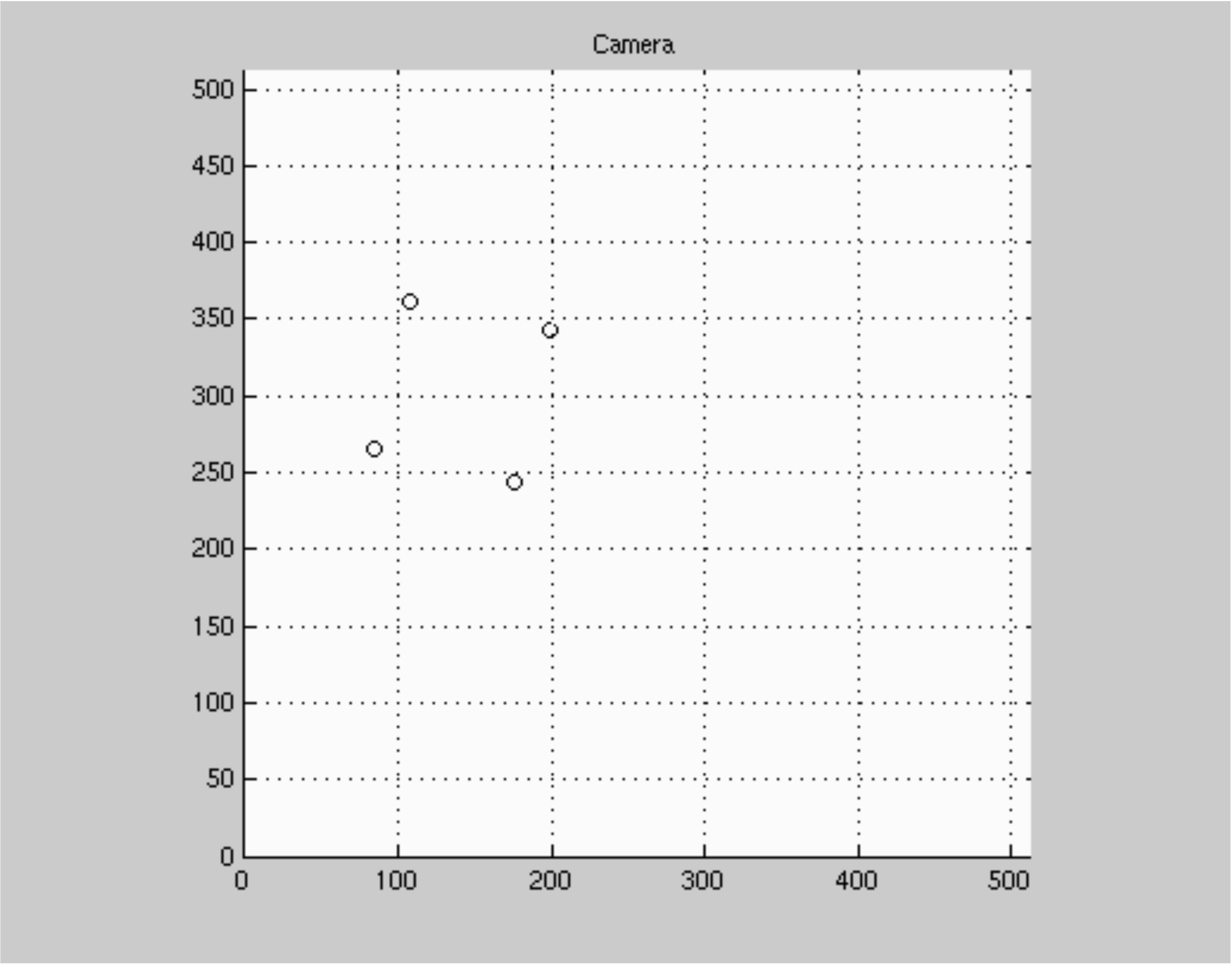
$$\dot{\boldsymbol{p}}^* = \lambda(\boldsymbol{p}^* - \boldsymbol{p})$$

desired point

current point

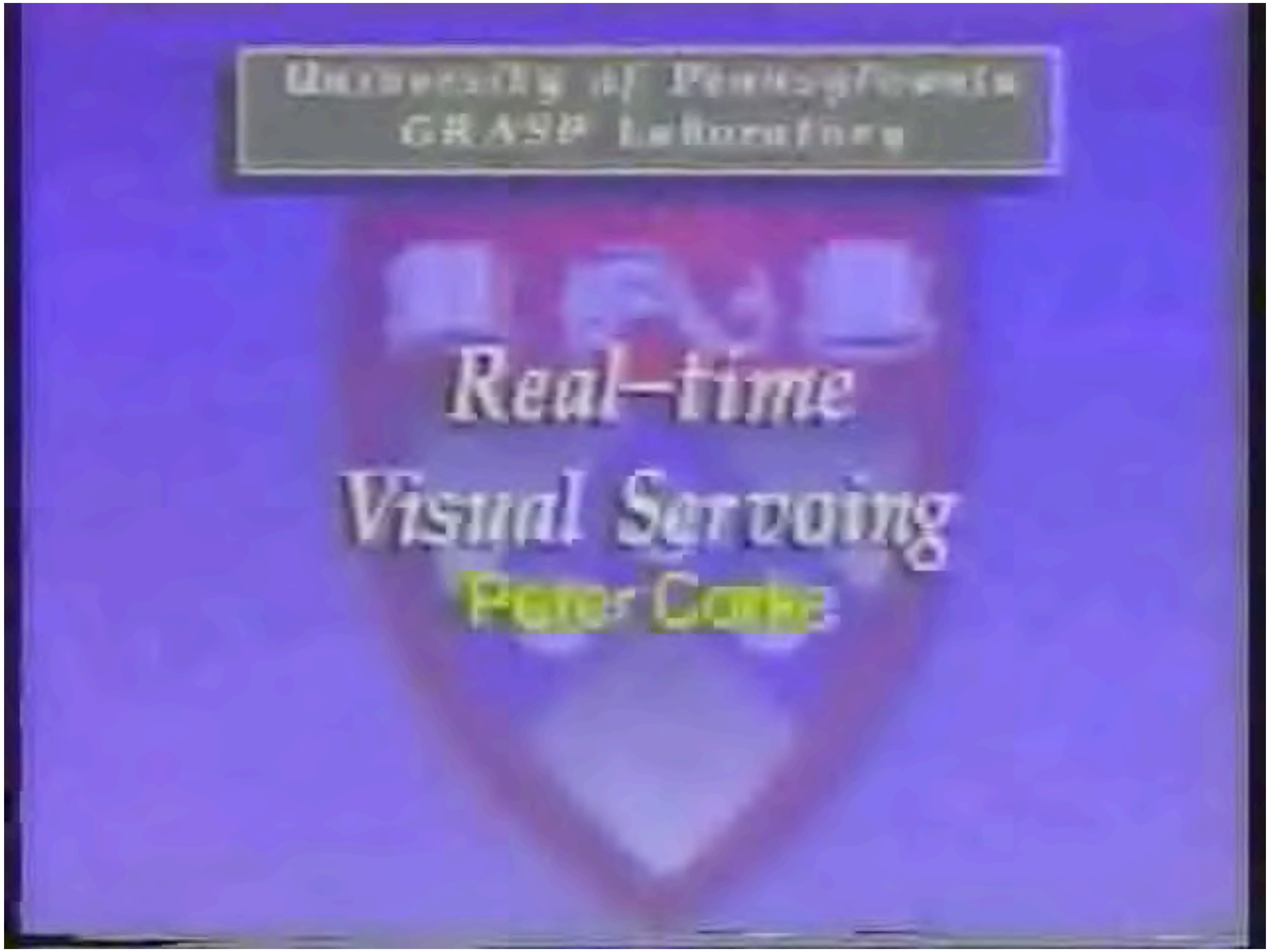
$$\boldsymbol{\nu} = \lambda \begin{pmatrix} J_{p_1} \\ J_{p_2} \\ J_{p_3} \end{pmatrix}^{-1} (\boldsymbol{p}^* - \boldsymbol{p})$$

IBVS simulation



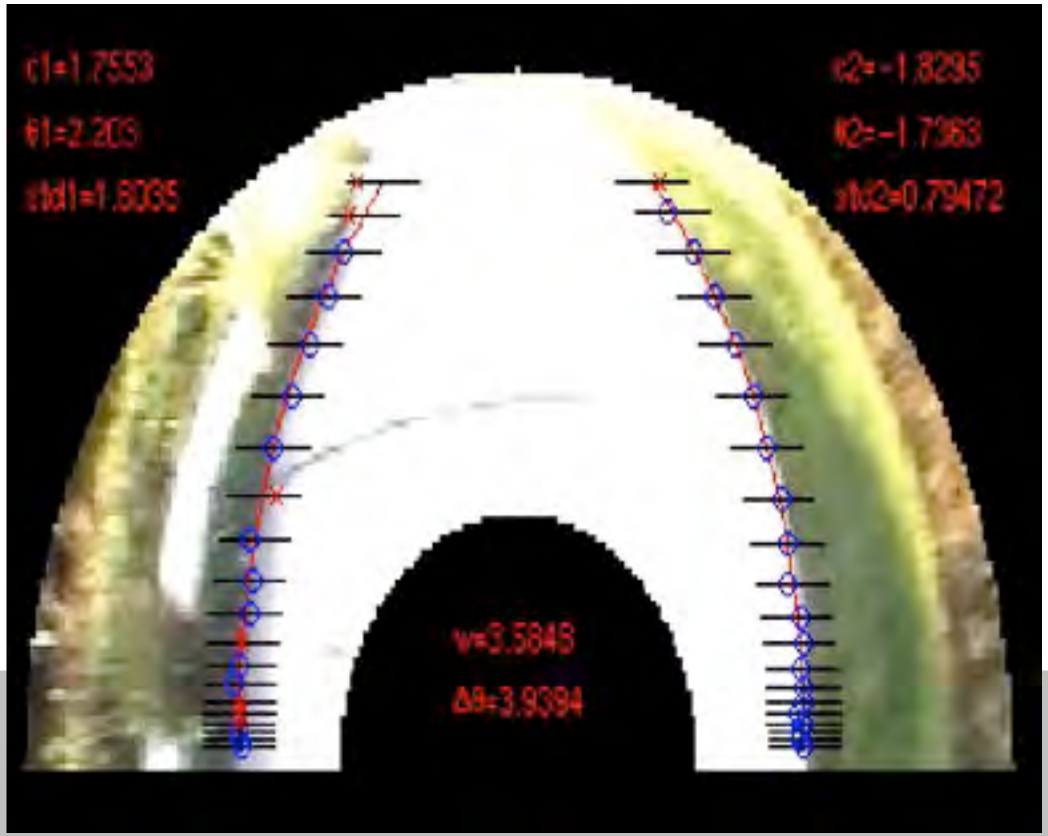
Applications

Visual servoing: simple feedback



Non-holonomic visual servoing

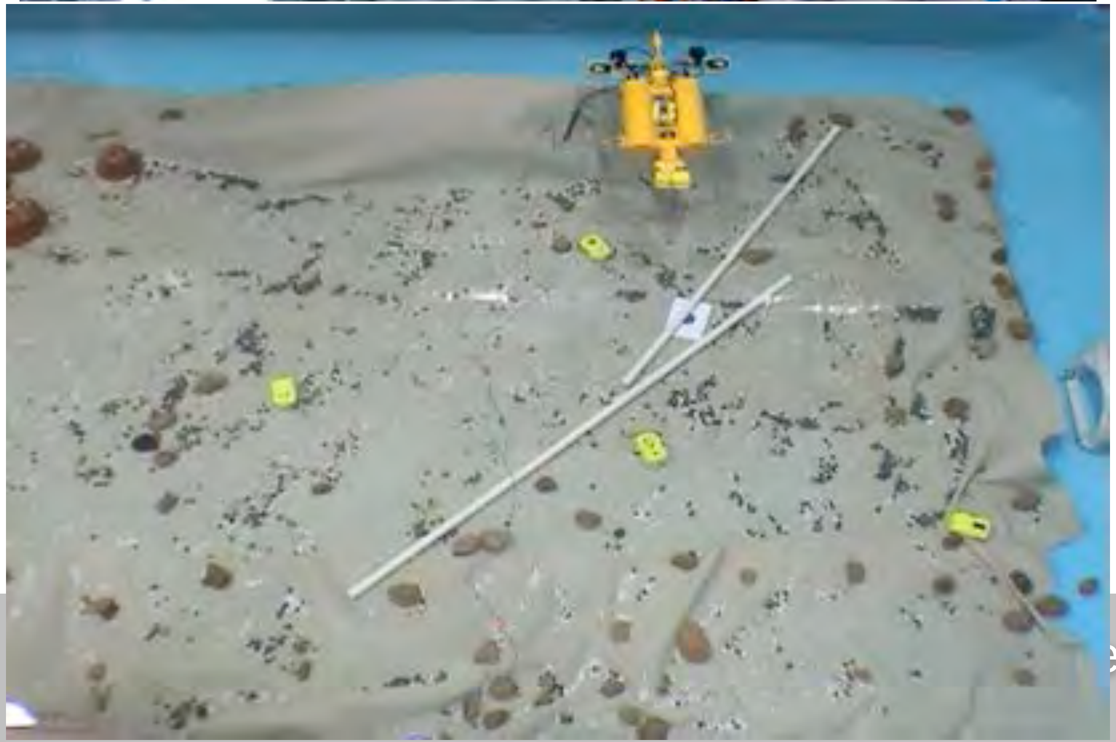
Direct processing of wide-angle imagery



K. Usher and J. Roberts and P. Corke and E. Duff, Vision-based navigational competencies for a car-like vehicles. In M. Ang and O. Khatib, editors, *Experimental Robotics IX*, volume 21 of *Springer Tracts in Advanced Robotics (STAR)*, 2006.

P. Corke and D. Symeonidis and K. Usher, Tracking road edges in the panospheric image plane. In *Proc. Int. Conf on Intelligent Robots and Systems (IROS)*, 2003.

AUV visual odometry



A hybrid AUV design for shallow water reef navigation. In *Proc. IEEE Int. Conf. Robotics and Automation*, 2005.

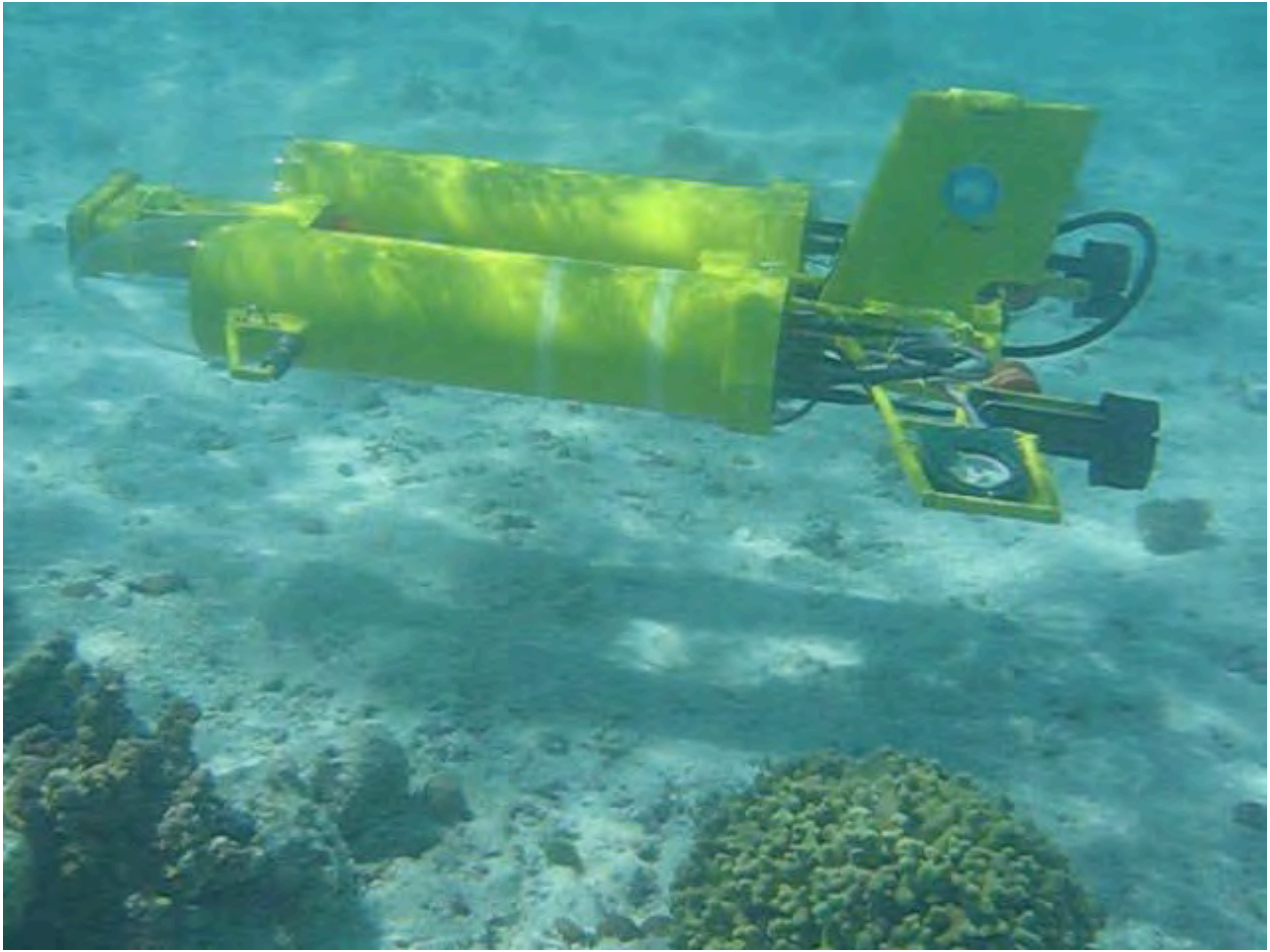
M. Dunbabin, P. Corke, Visual motion estimation for an autonomous underwater reef monitoring robot. In P. Corke and S. Sukkariah, editors, *Field and Service Robotics: Results of the 5th International Conference*, 2006.

I. Vasilescu, K. Kotay, D. Rus, M. Dunbabin, and P. Corke. Data collection, storage and retrieval with an underwater sensor network. In *Proc. IEEE SenSys*, 2005.

Crucible Finder

- The HMC must deal with uncertainty on crucible pose and vehicle approach position
- Use a pan/tilt/zoom camera
- Use visual fiducials on the crucible handle


Crucible pickup



Queensland University of Technology Knowledge Cloud

Participants needed for MOOC trial

Robotic Vision

Help us pioneer QUT's new *Massive Open Online Course* by joining us to test the first four weeks of this eight week course. *Robotic Vision,* by Professor Peter Corke, promises to be an in-depth look into the mechanics of sight and how robots receive and interpret visual stimulus.

Starting on 17 November

Pre-register at
qut.edcastcloud.com/learn/robotic-vision-trial

For more information contact **moocs@qut.edu.au**

QUT

a university for the real world®