Robot Perception: Camera Models

Advanced Robotics

Kostas Daniilidis

Can our aerial robots control their flight like gannets?

How can gannets' eyes and brain estimate distance to water!

Nature Vol. 293 24 September 1981

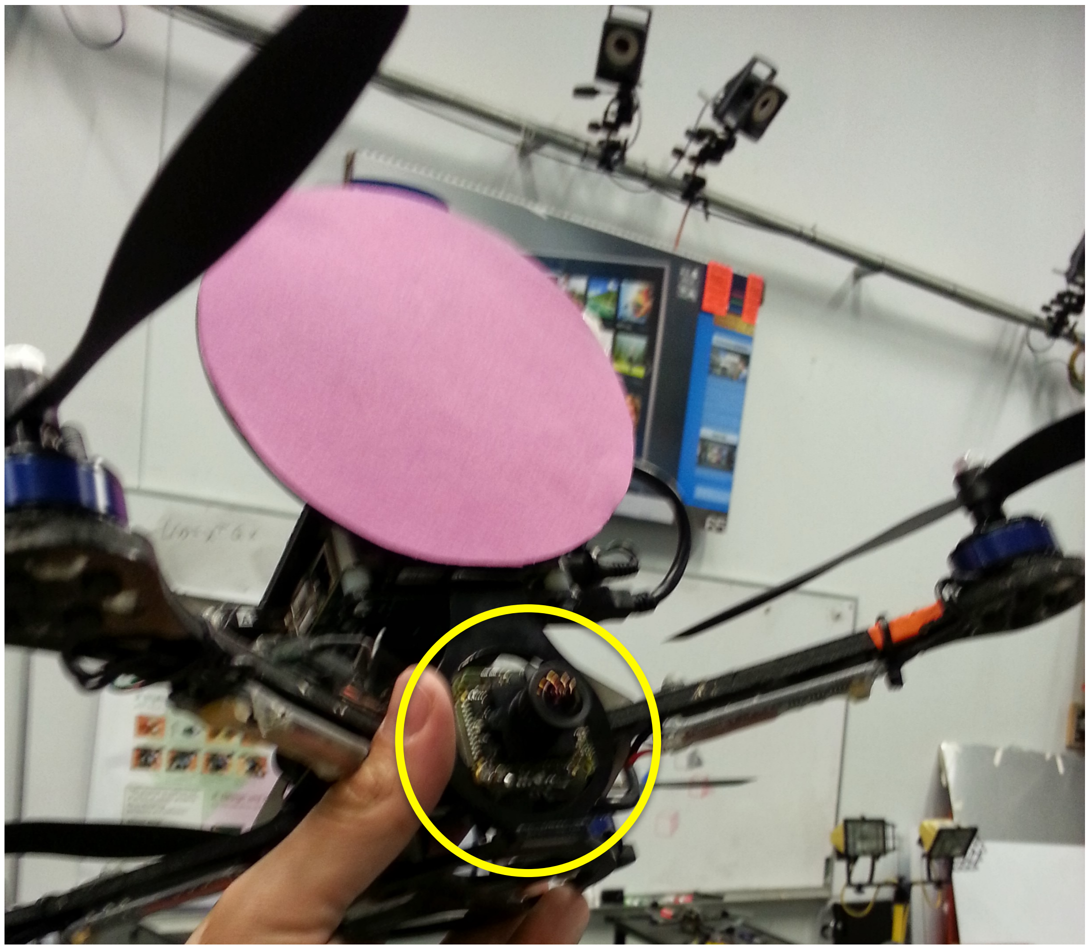
5. Spiess, F. N. *et al. Science* **207,** 1421–1433 (1980).
6. Corliss, J. B. *Science* **203,** 1073–1083 (1979).
7. Edmond, J. M. *et al. Earth planet. Sci. Lett.* **46,** 19–30 (1979).
8. Ruby, E. G., Wirsen, C. O. & Jannasch, H. W. *Appl. envir. Microbiol.* (in the press).
9. Cavanaugh, C. M., Gardiner, S. L., Jones, M. L., Jannasch, H. W. & Waterbury, J. B. *Science* **213,** 340–342 (1981).
10. Felbeck, H. *Science* **213,** 336–338 (1981).
11. Jones, M. L. *Science* **213,** 333–336 (1981).
12. Peck, H. D. Jr *Enzymes* **10,** 651–669 (1974).
13. Latzko, E. & Gibbs, M. *Pl. Physiol.* **44,** 295–300 (1969).
14. Reid, R. G. B. *Can. J. Zool.* **58,** 386–393 (1980).
15. Los Angeles County (California) Sanitation District files.
16. Rau, G. H. *Science* **213,** 338–340 (1981).
17. Rau, G. H. *Nature* **289,** 484–485 (1981).
18. Rau, G. H. & Hedges, J. I. *Science* **203,** 648–649 (1979).
19. Emery, K. O. & Hulsemann, J. *Deep-Sea Res.* **8,** 165–180 (1962).
20. Hartman, O. & Barnard, J. L. *Allan Hancock Pacific Exped.* **22,** (1958).
21. Nicholas, D. J. D., Ferrante, J. V. & Clarke, G. R. *Analyt. Biochem.* **95,** 24–31 (1979).
22. Lonsdale, P. *Nature* **281,** 531–534 (1979).

Plummeting gannets: a paradigm of ecological optics

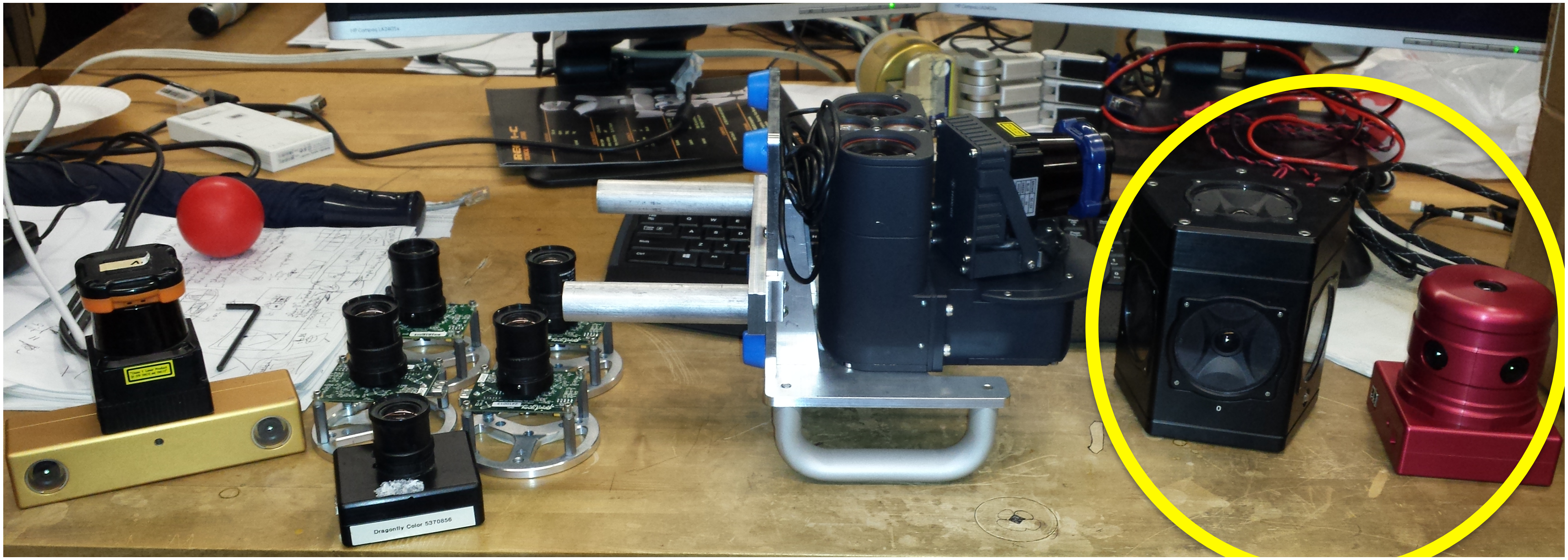
David N. Lee & Paul E. Reddish

Department of Psychology, University of Edinburgh, Edinburgh EH8 9JZ, UK

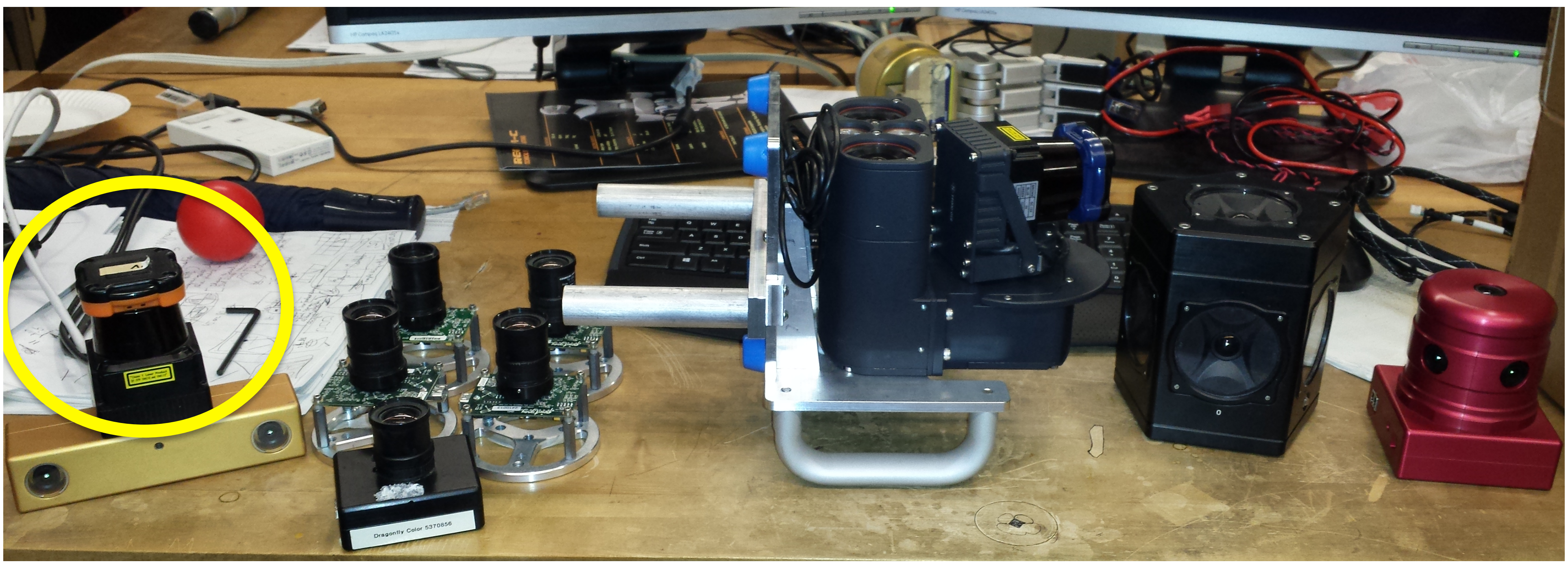
Quadrotor "sees" with a camera



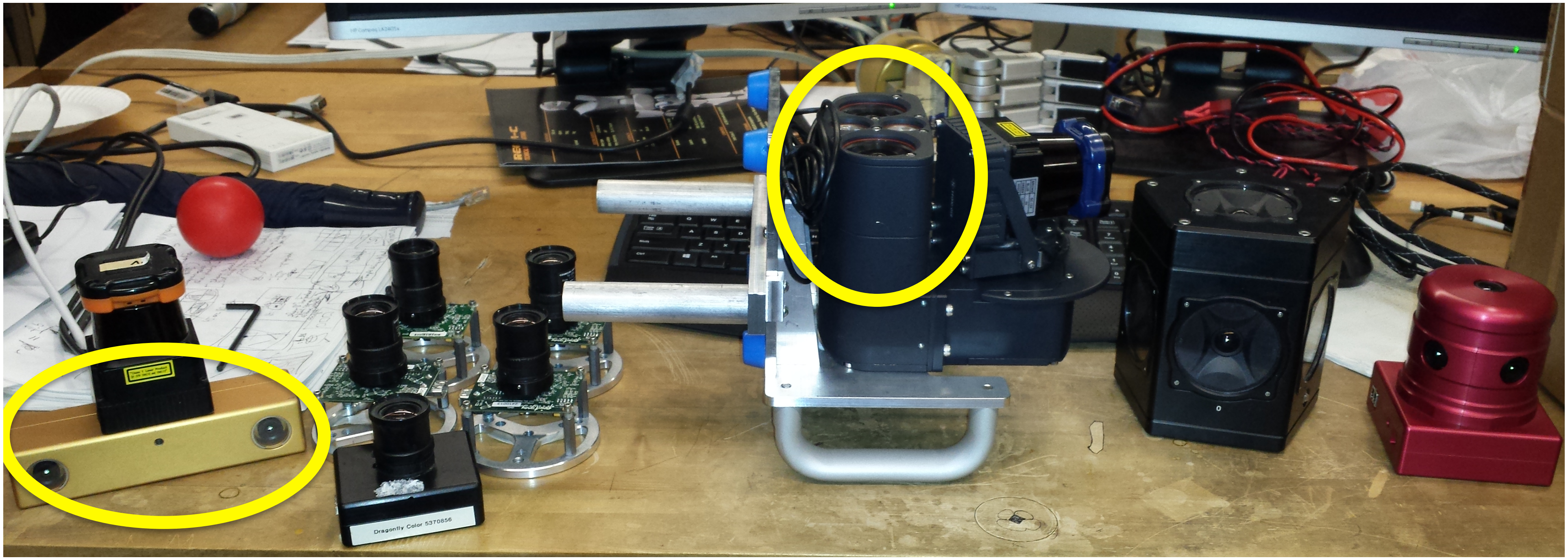
In robotics we use all kinds of cameras!



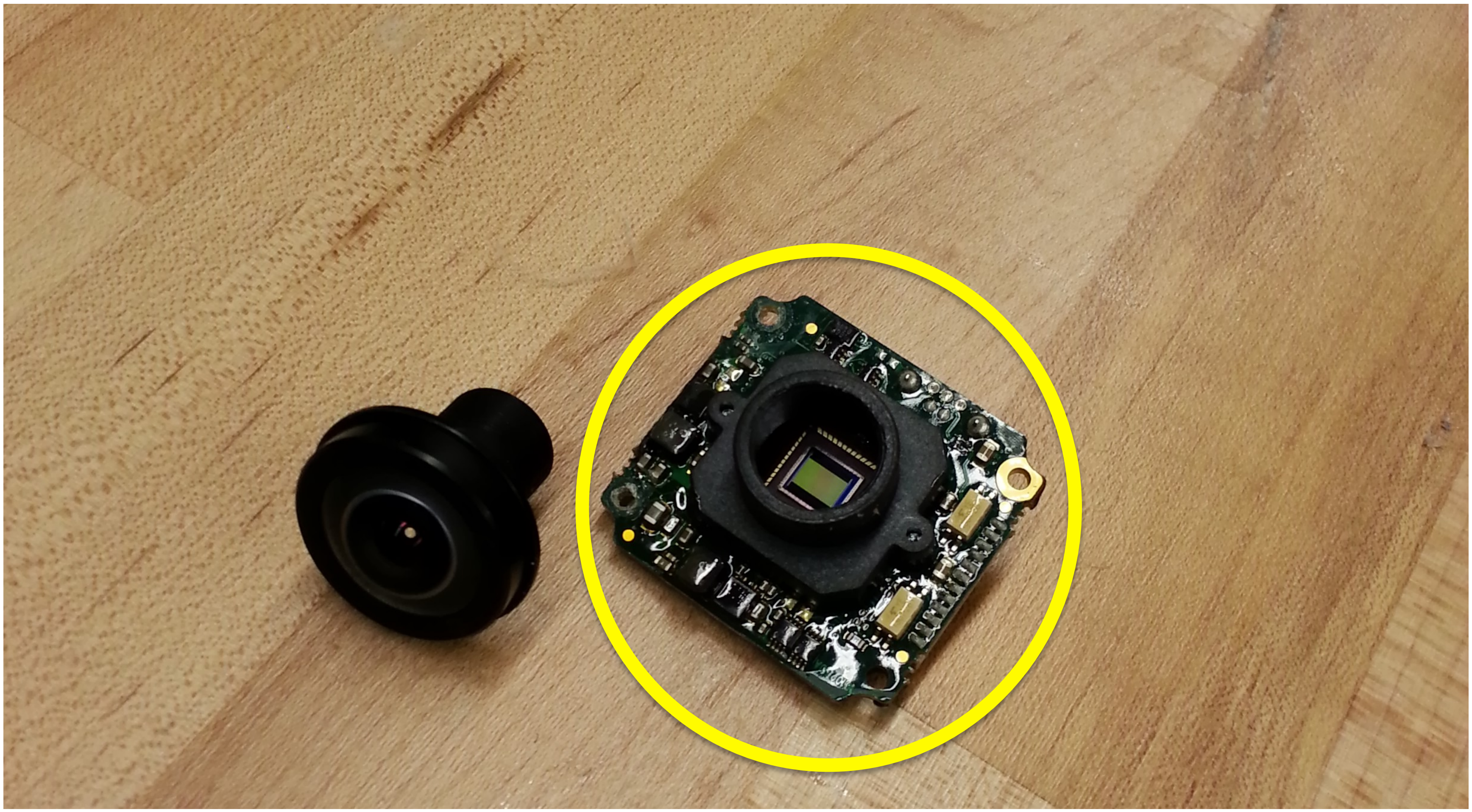
In robotics we use all kinds of cameras!



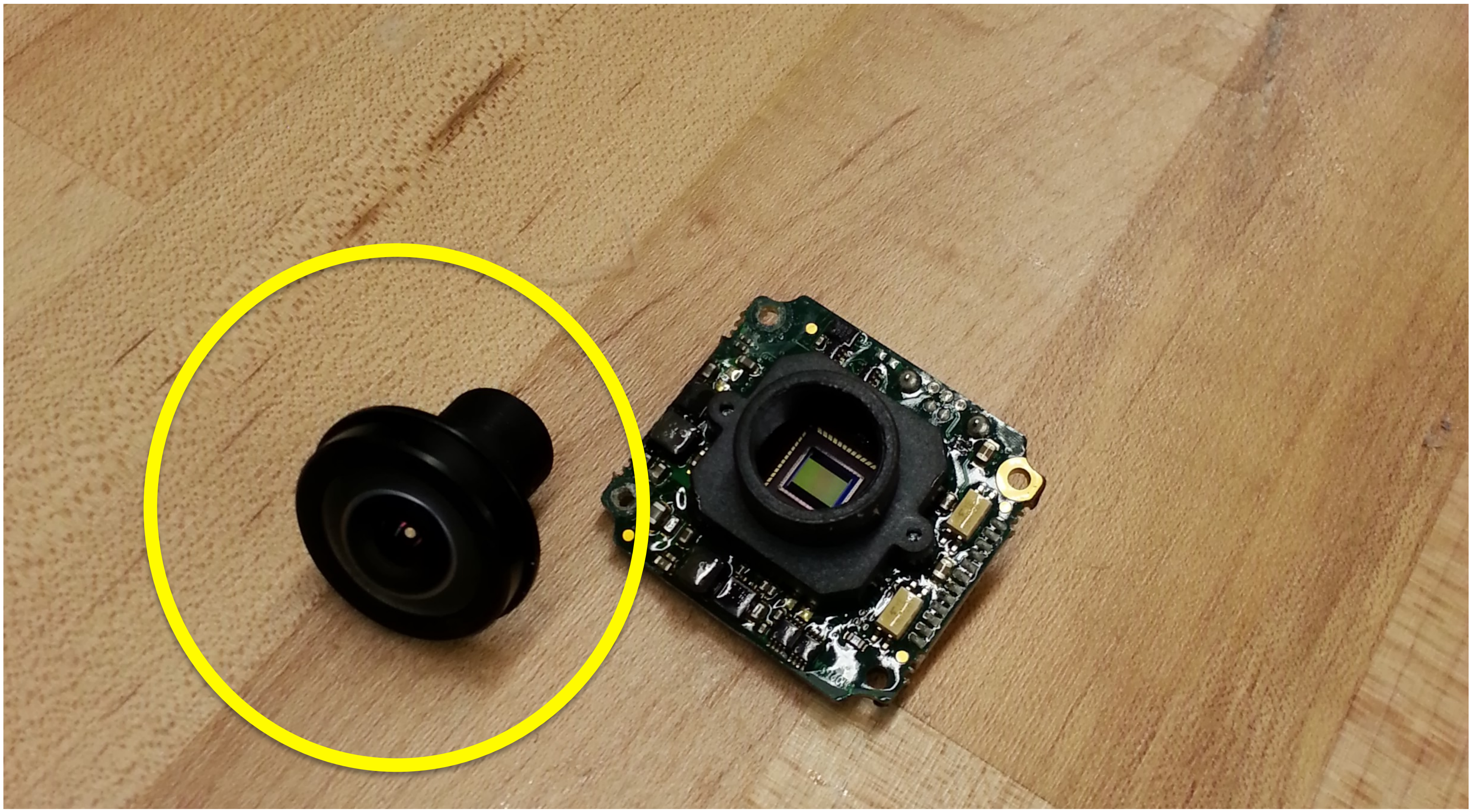
In robotics we use all kinds of cameras!



A camera is an imaging chip and a lens

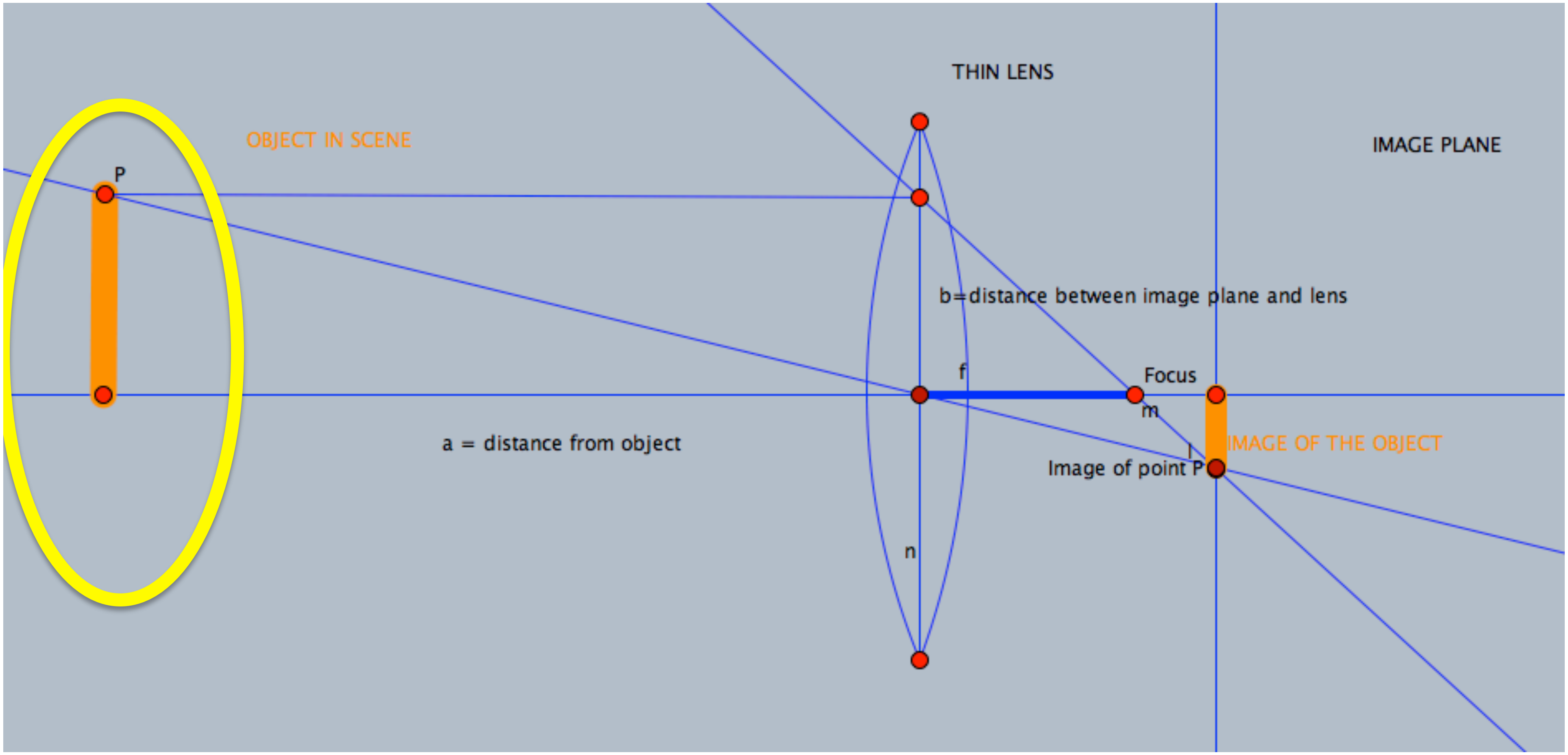


A camera is an imaging chip and a lens



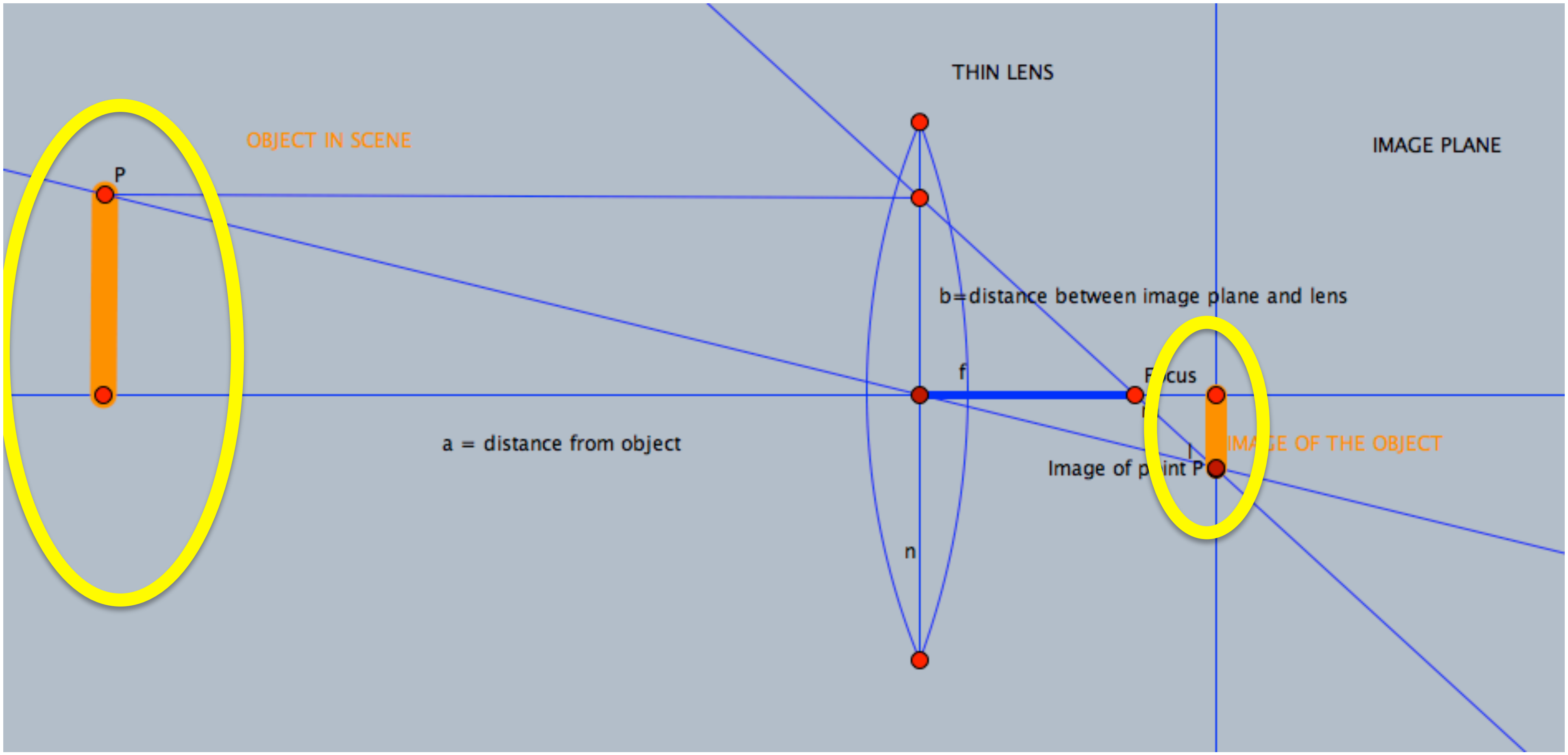
Magnifying glass

How does a thin lens work



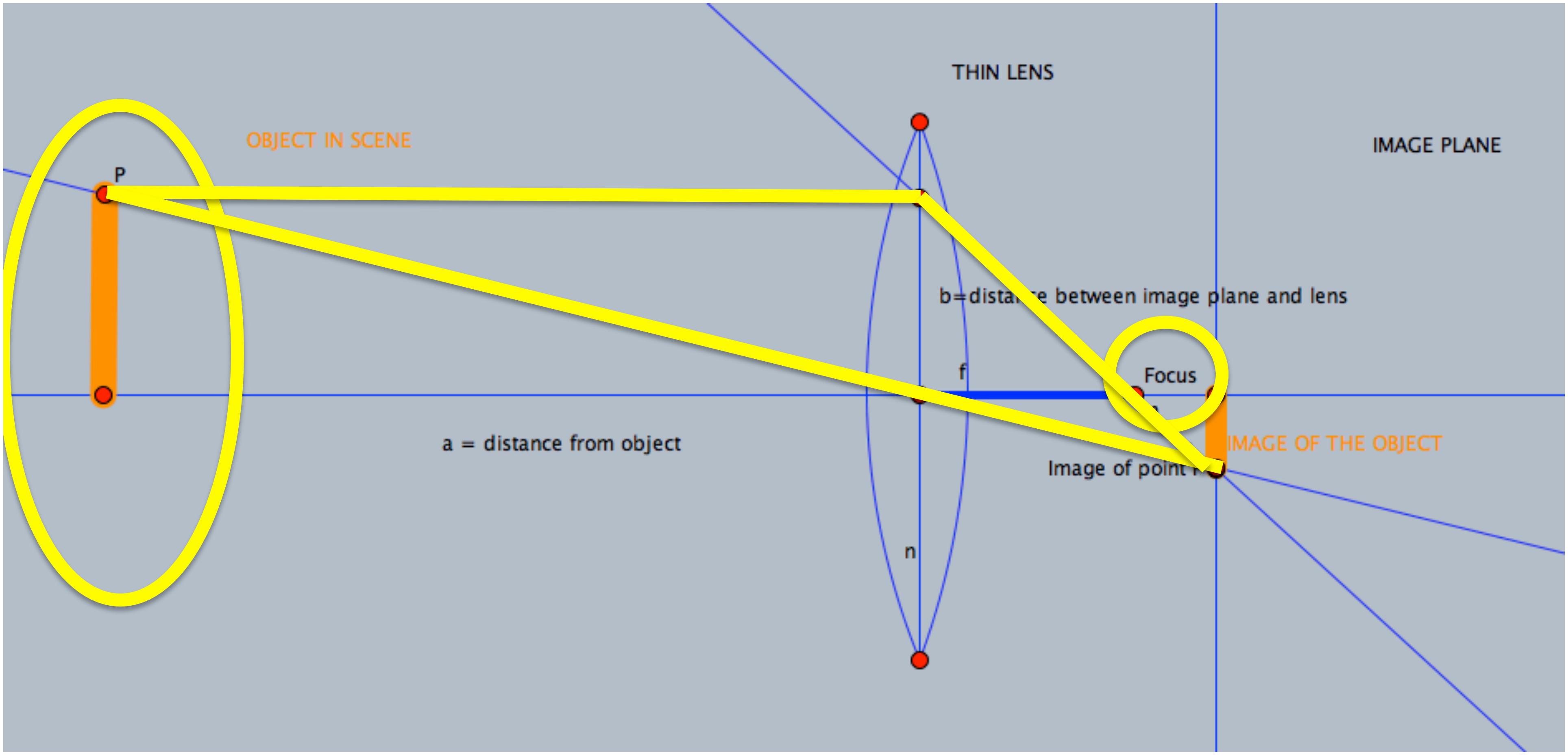
Rays from on object point P converge on a point p on the image plane

How does a thin lens work



Rays from a point P in the scene converge into a point in the image plane

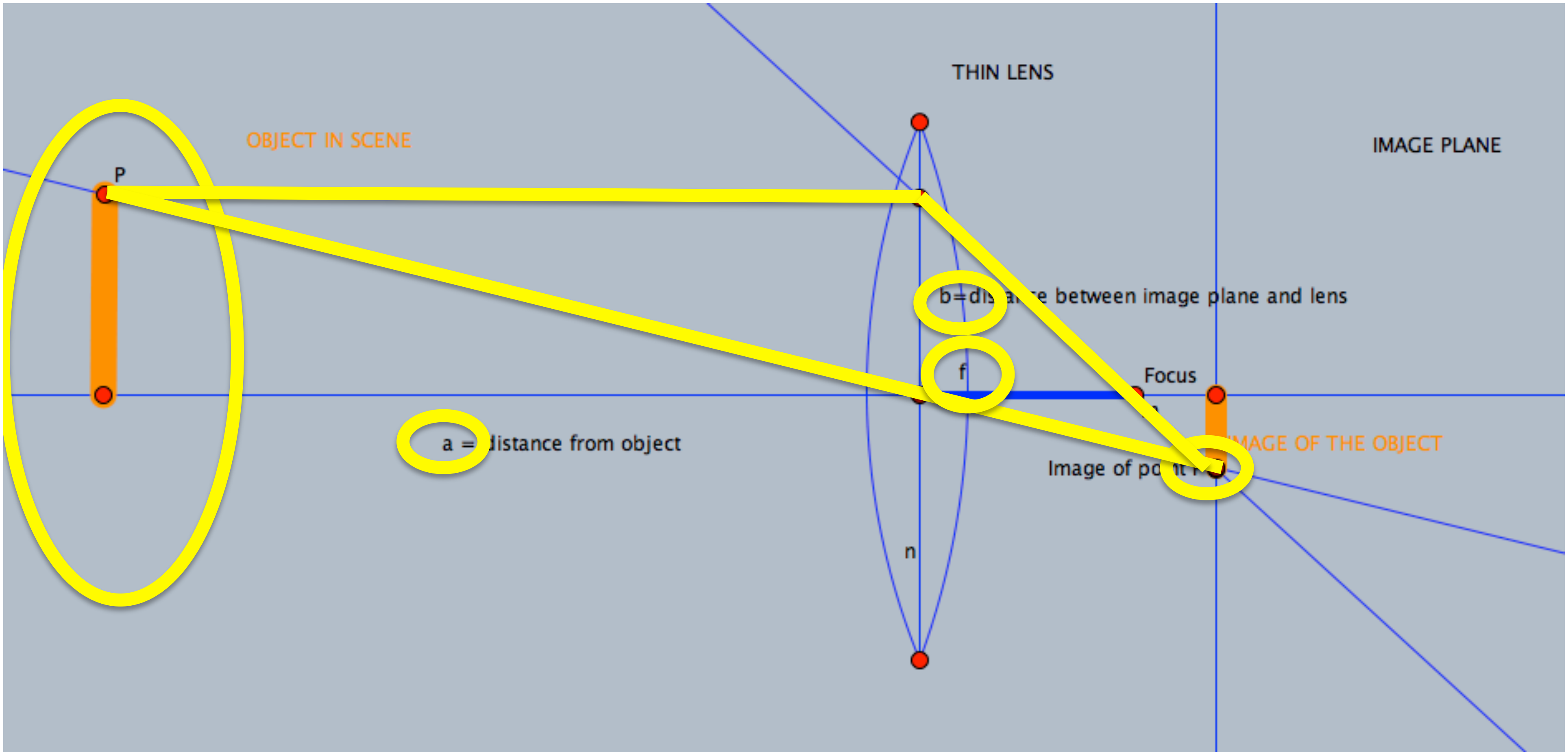
How does a thin lens work



Rays parallel to the optical axis meet the focus after leaving the lens.
Rays through center of the lens do not change direction.

$$\frac{1}{f} = \frac{1}{a} + \frac{1}{b}$$

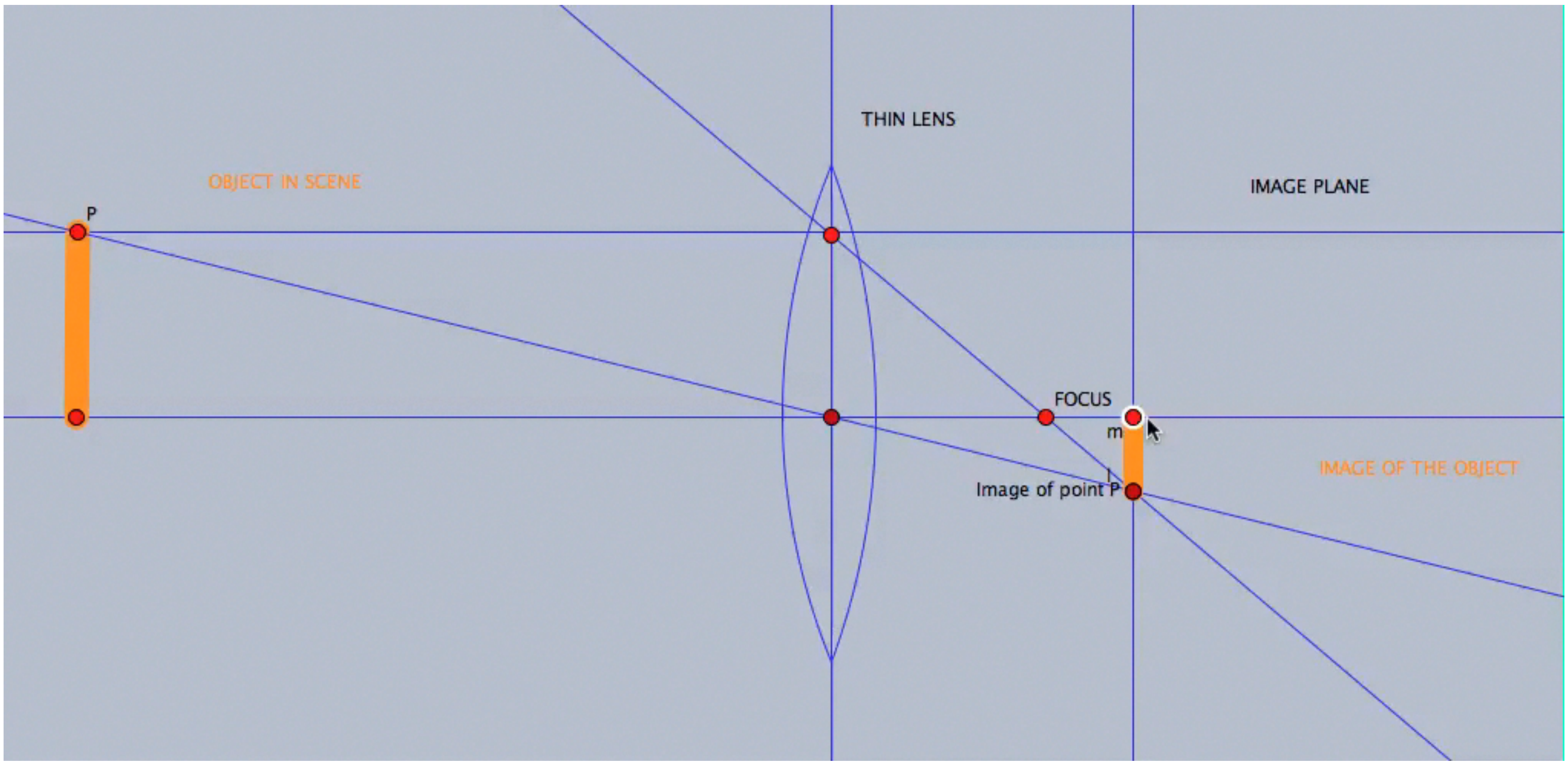
How does a thin lens work



These rays meet at one point if

$$\frac{1}{f} = \frac{1}{a} + \frac{1}{b}$$

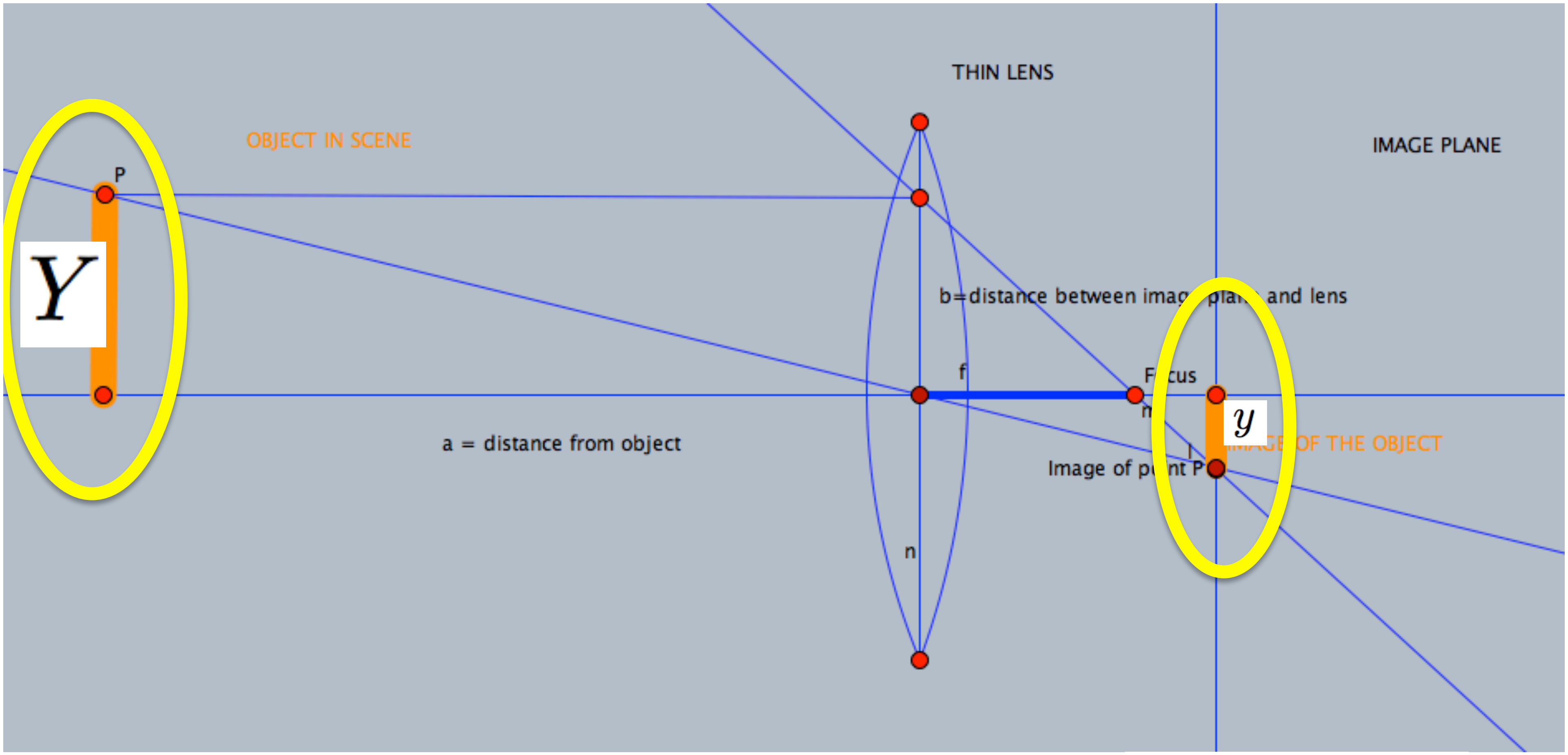
What happens when we move b, the image plane



Moving the image plane is what we call (de-) focusing !
Image starts blurring!

$$\frac{1}{f} \neq \frac{1}{a} + \frac{1}{b}$$

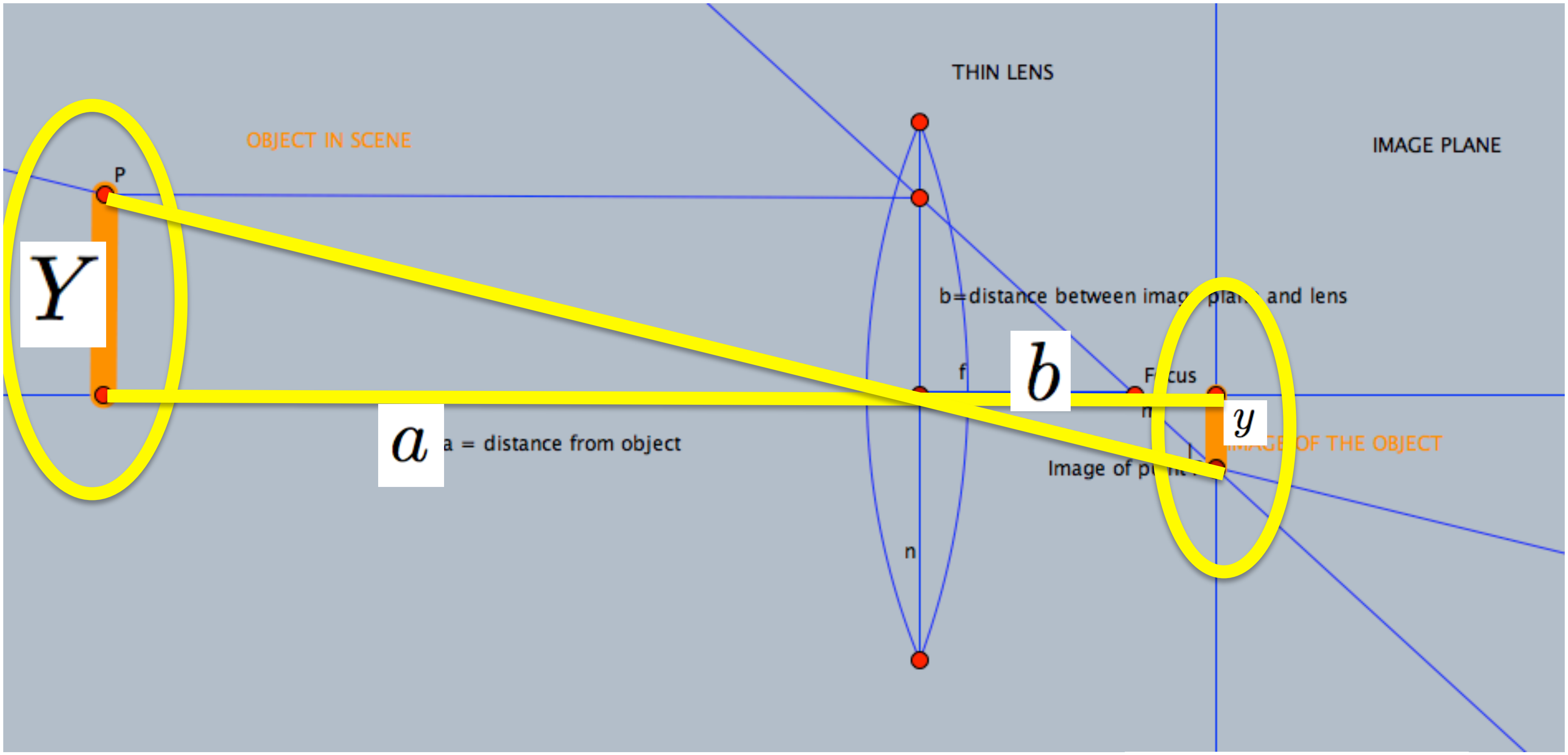
Perspective projection: size of object image



If you look only at the ray going through the center of the lens

$$\frac{Y}{a} = \frac{y}{b}$$

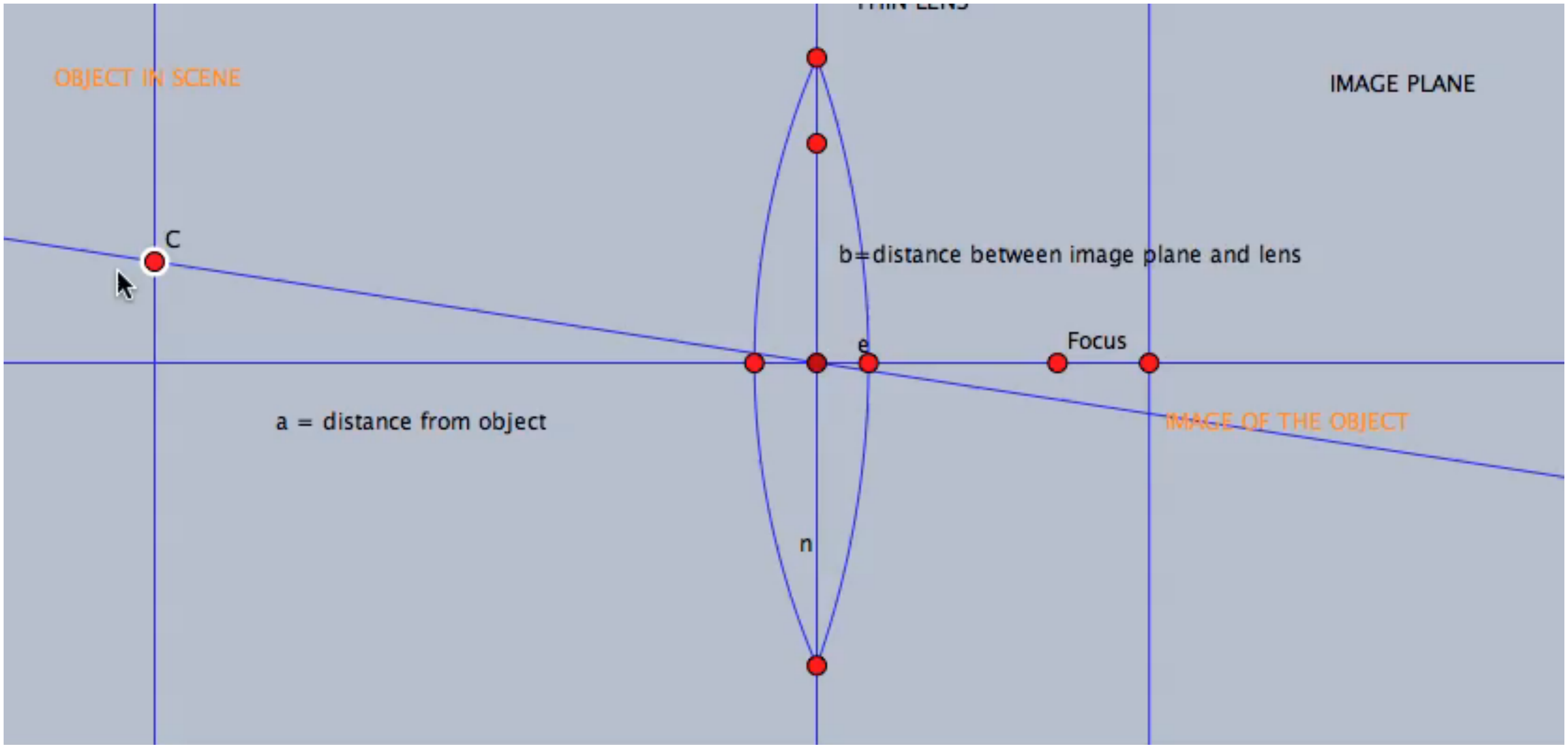
Perspective projection: size of object image



If you look only at the ray going through the center of the lens

$$\frac{Y}{a} = \frac{y}{b}$$

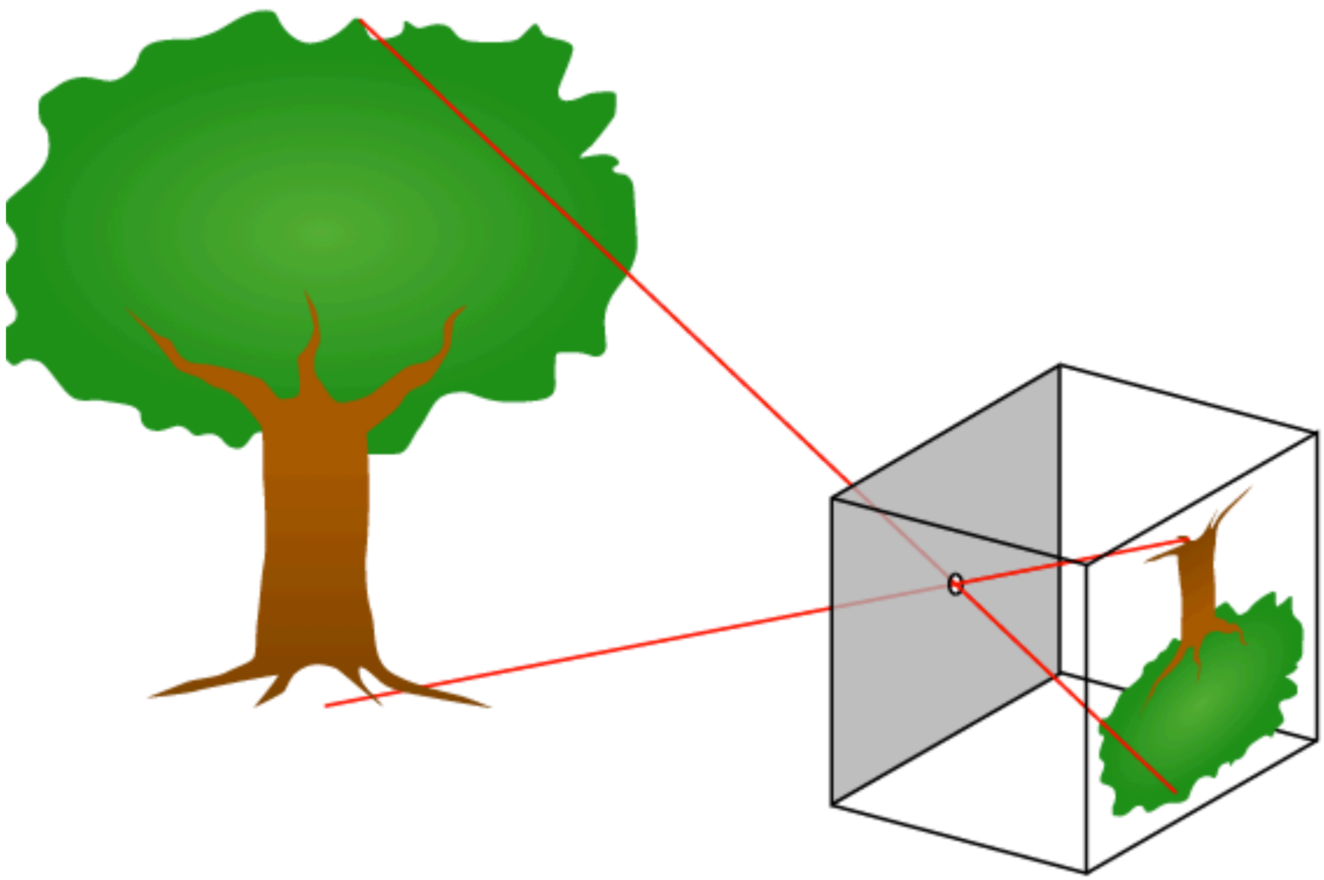
Perspective projection



- An object of the same size coming closer results on a larger image
- A point moving on the same ray does not change its image

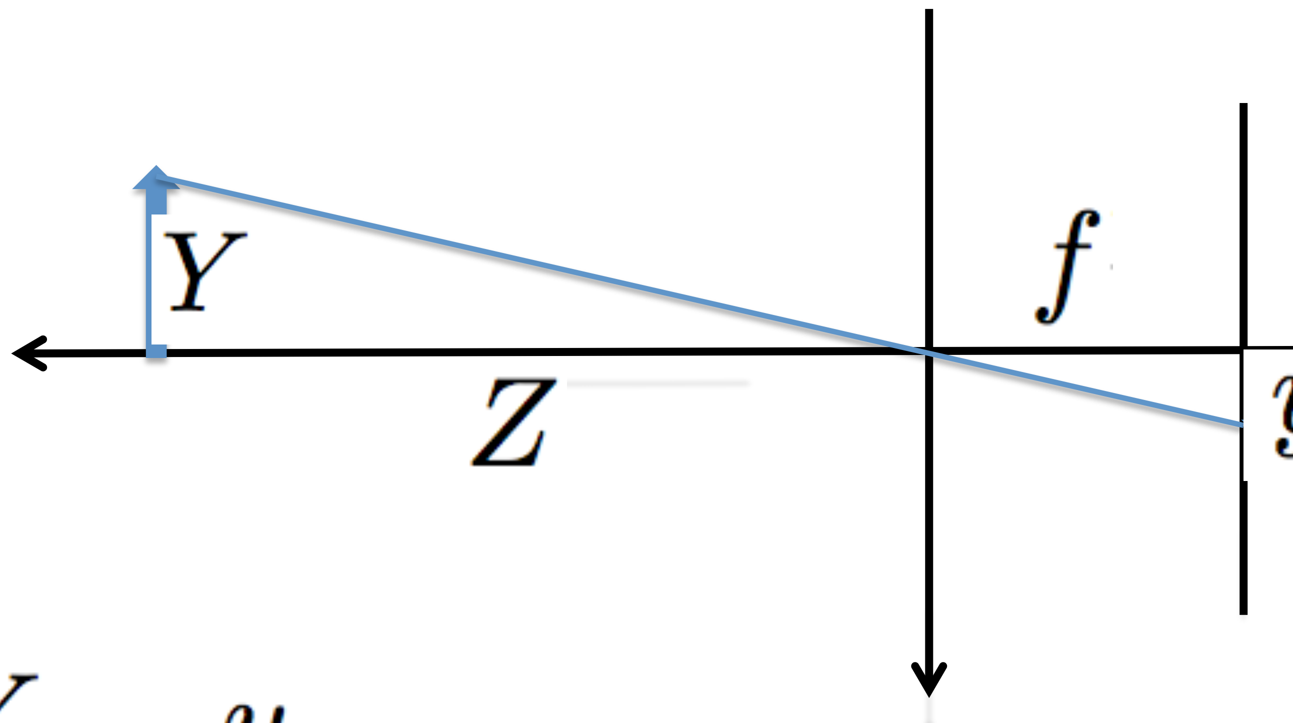
$$\frac{Y}{a} = \frac{y}{b}$$

Perspective projection = Pinhole Model



WIKIPEDIA
The Free Encyclopedia

Pinhole model



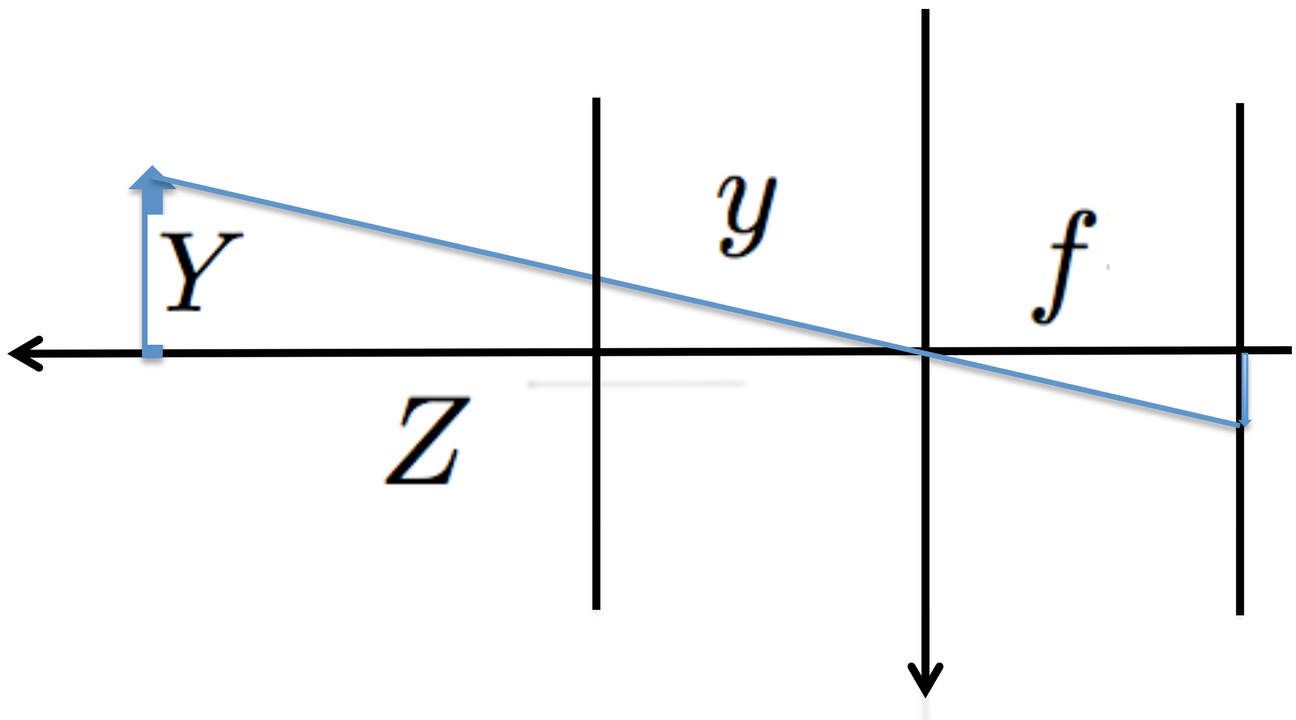
$$\frac{Y}{Z} = \frac{y}{b}$$

$$Z = a$$

If we replace b with f and include a minus because object image is upside down….

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

Pinhole model used in computer vision and robotics



… amd assum that image plane is in front of the lens

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

What is the effect of f=b ?

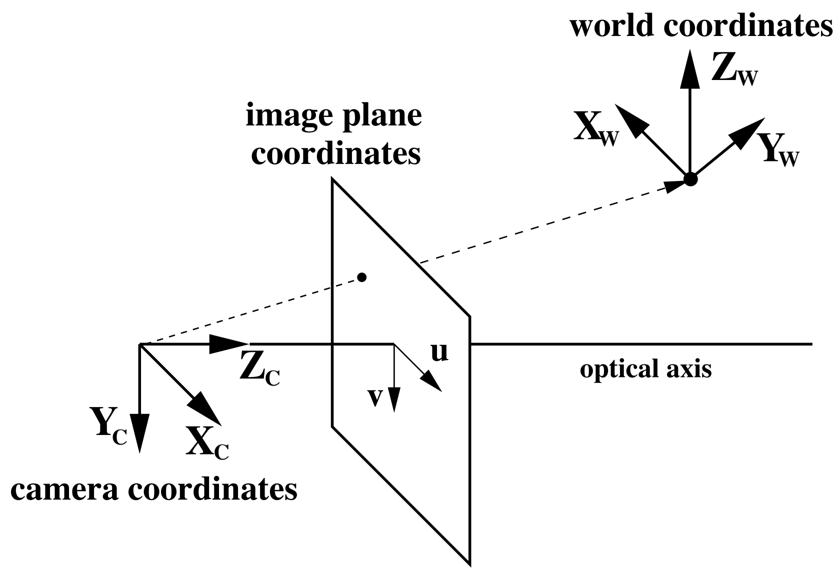
- Theoretically, we expect an offset in the x and y coordinates caused by the error (f-b).
- If the object is on focus:

$$\frac{b-f}{f} = \frac{b}{Z}$$

Relative error depends on the ratio of focal length to depth !
This would matter if we would actually use the f from specs of the camera.

In practice we use a process called calibration, yielding the f that best satisfies $y = f\frac{Y}{Z}$

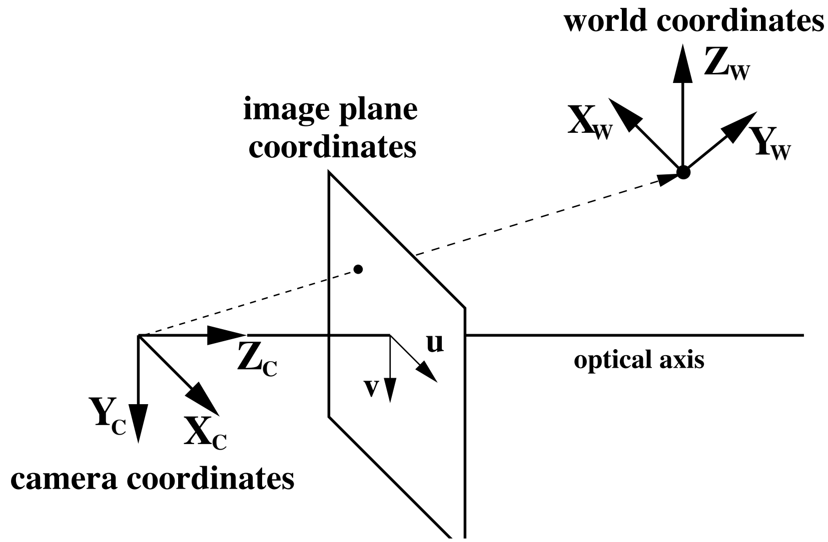
Camera Coordinate System



world coordinate system coordinates (X_w, Y_w, Z_w),

camera coordinate system coordinates (X_c, Y_c, Z_c).

Projection Equations

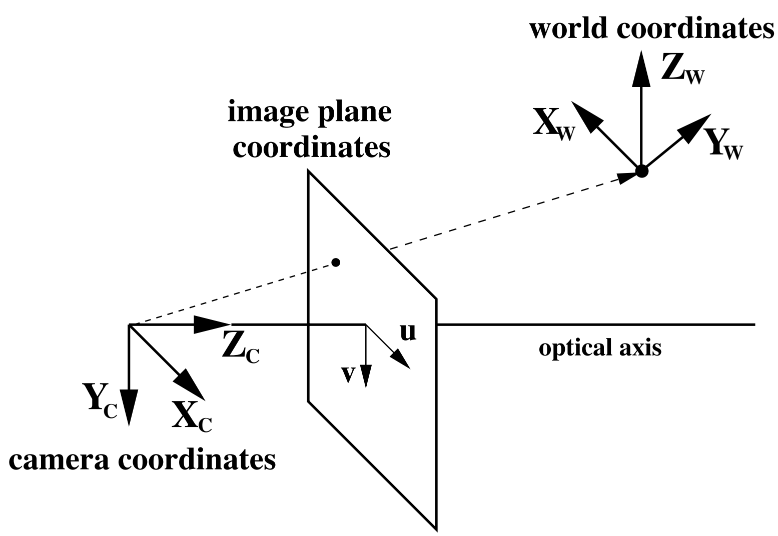


$$u = f\frac{X_c}{Z_c} + u_o \quad v = f\frac{Y_c}{Z_c} + v_o.$$

The optical axis is the z-axis.
The image plane (u, v) is perpendicular to the optical axis.
Intersection of the image plane with the optical axis is the *image center* (u_o, v_o)
f is the distance of the image plane from the origin (in pixels).

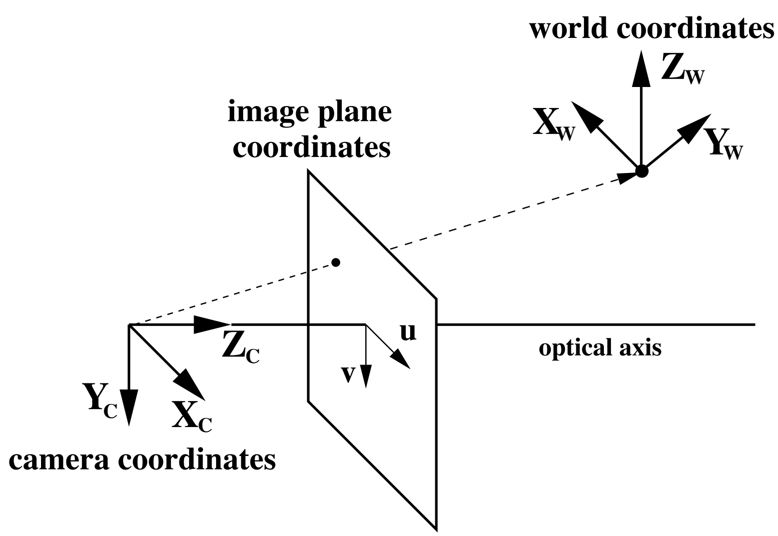
Perspective projection in matrix form

$$\lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & u_o \\ 0 & f & v_o \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix}$$



From camera to world

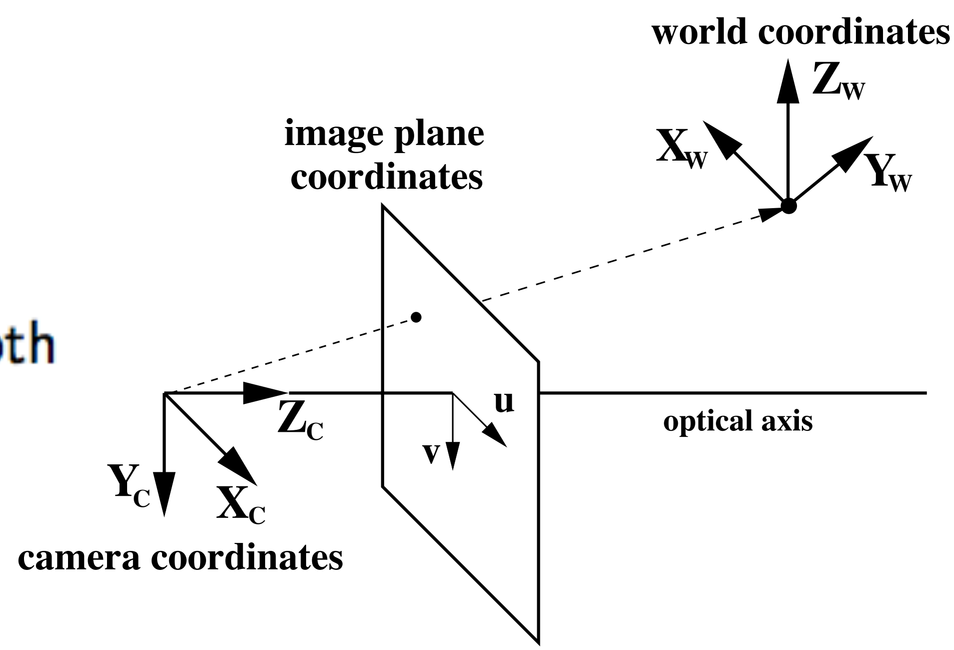
$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$



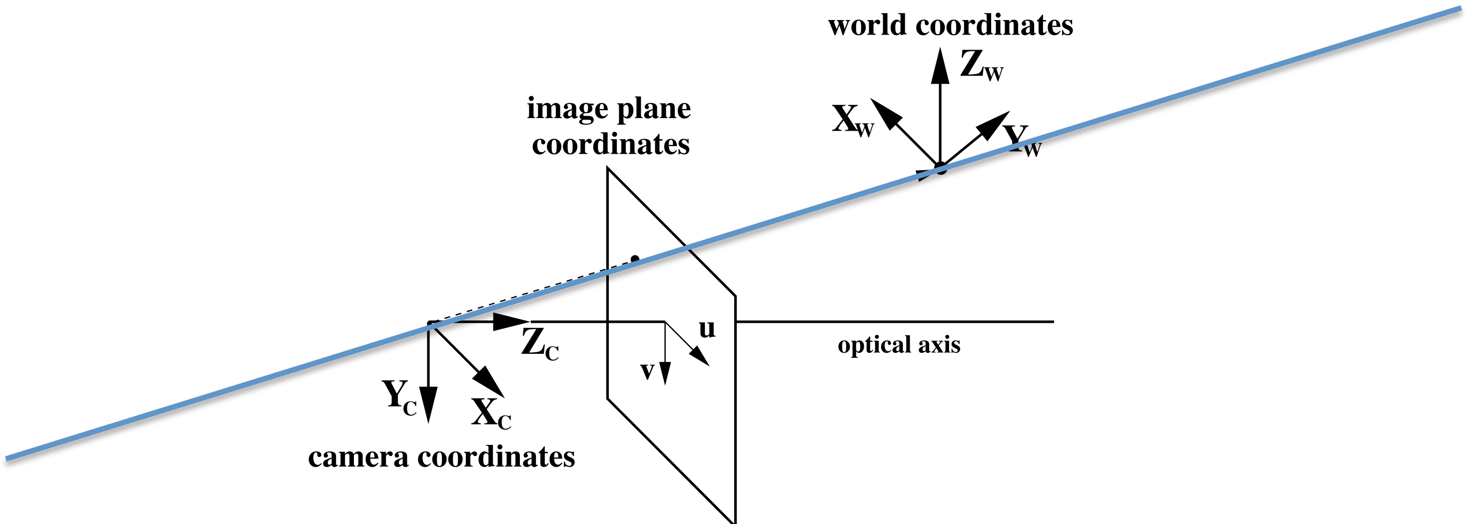
The 3x4 projection matrix P

$$\lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & u_o \\ 0 & f & v_o \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} R & t \end{pmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix} = P \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

λ is the unknown depth



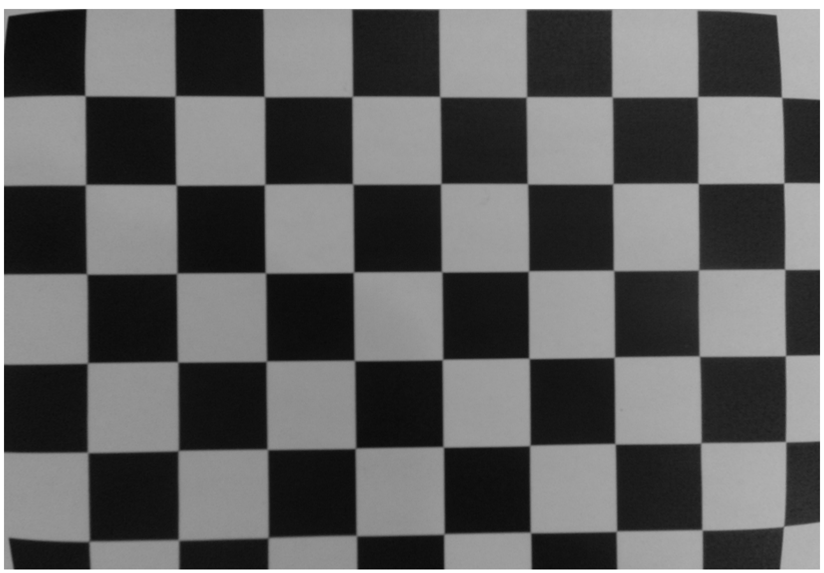
The meaning of the projection equation: It is the equation of a ray in world coordinates going through the camera center

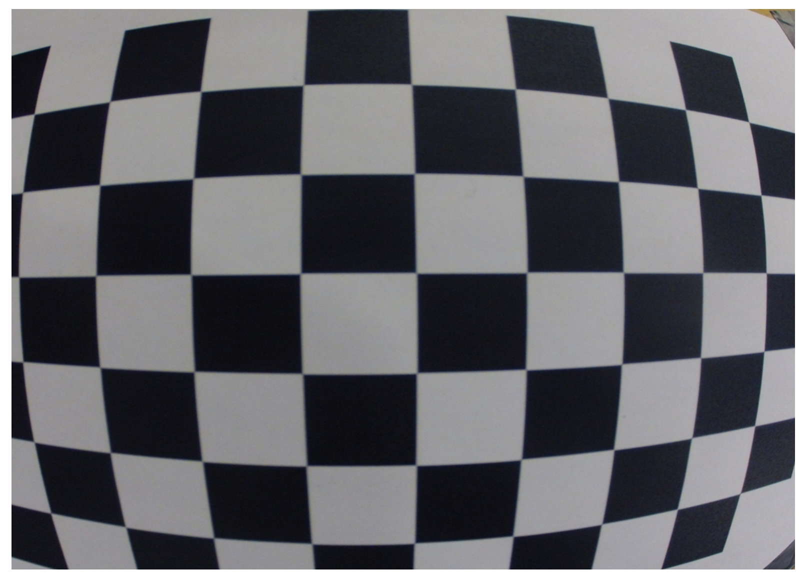


$$\lambda K^{-1}\begin{pmatrix}u\\ v\\ 1\end{pmatrix} = R\begin{pmatrix}X_w\\ Y_w\\ Z_w\end{pmatrix} + t \Longrightarrow \begin{pmatrix}X_w\\ Y_w\\ Z_w\end{pmatrix} = \overset{\text{Camera center}}{\boxed{-R^T T}} + \lambda \overset{\text{Ray direction}}{\boxed{R^T K^{-1}\begin{pmatrix}u\\ v\\ 1\end{pmatrix}}}$$

Cameras with large field of view have radial distortions

$$
\begin{aligned}
u^{dist} &= u(1 + k_1 r + k_2 r^2 + k_3 r^3 + \ldots) \\
v^{dist} &= v(1 + k_1 r + k_2 r^2 + k_3 r^3 + \ldots) \\
\text{where} \quad r^2 &= u^2 + v^2
\end{aligned}
$$

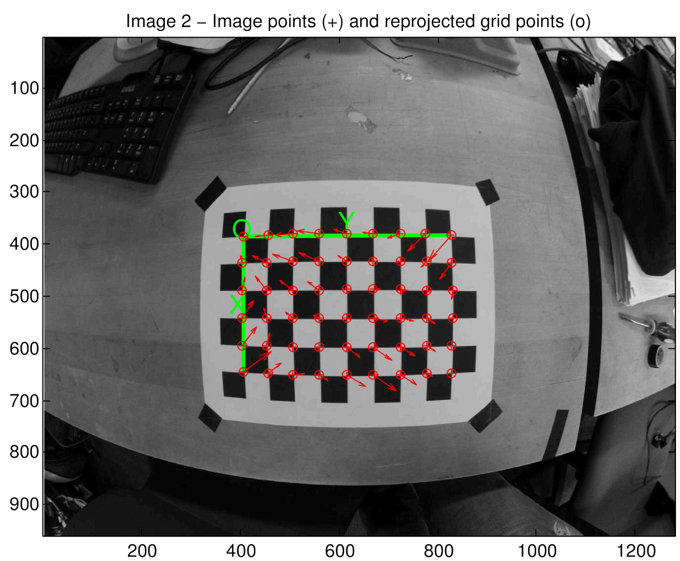


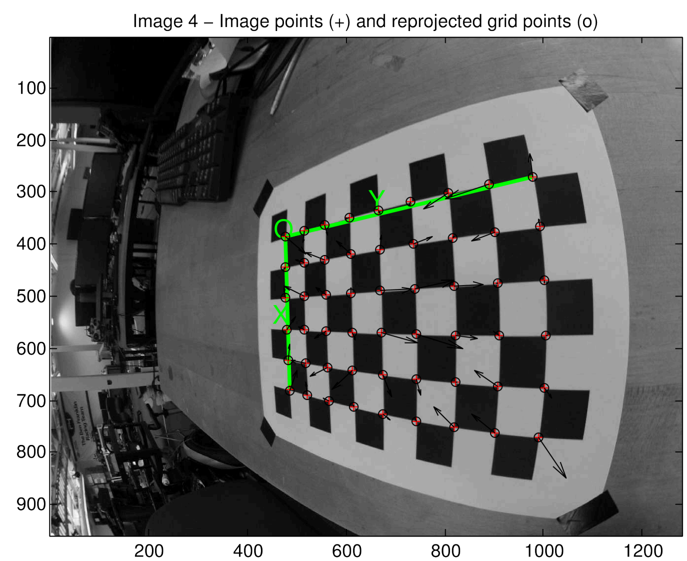


A procedure called calibration

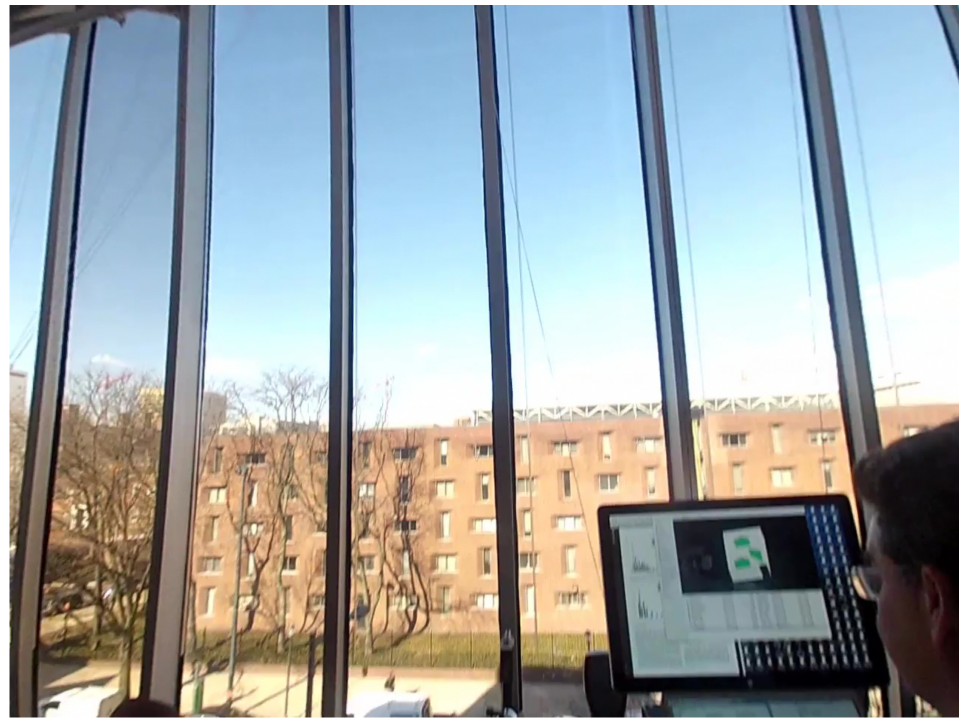
Estimates the *intrinsic parameters*

- f focal length
- (u_o, v_o) image center
- $k_1, k_2, \ldots$ radial distortion parameters

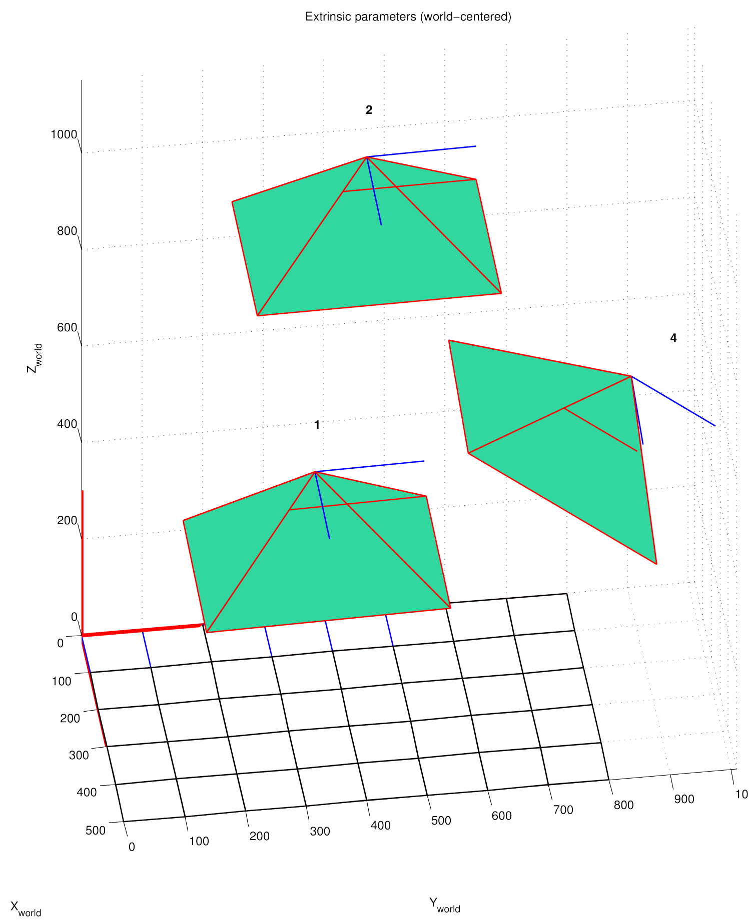




As a result of the calibration we have undistorted images and video



..as well as the poses of the camera and the projection rays in world coordinates



$$\begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} = \boxed{-R^T T} + \lambda \boxed{R^T K^{-1}} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

known

We will return later on the specifics of calibration….