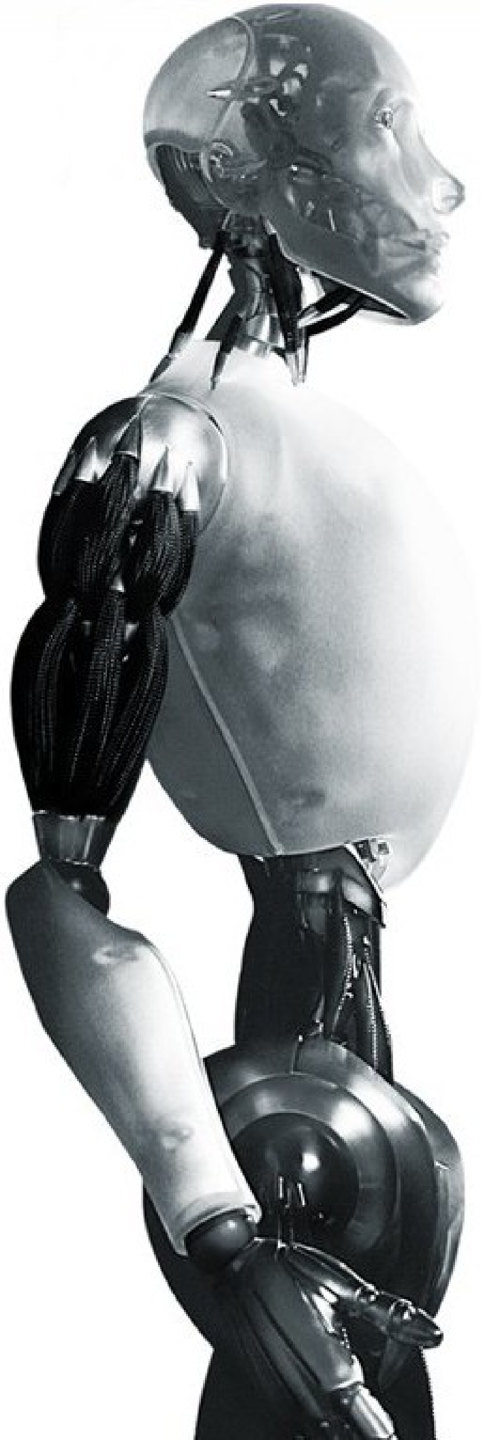
Disclaimer

These slides are intended as presentation aids for the lecture. They contain information that would otherwise be to difficult or time-consuming to reproduce on the board. But they are incomplete, not self-explanatory, and are not always used in the order they appear in this presentation. As a result, these slides should not be used as a script for this course. I recommend you take notes during class, maybe on the slides themselves. It has been shown that taking notes improves learning success.



Robotics

Computer Vision & Visual Servoing

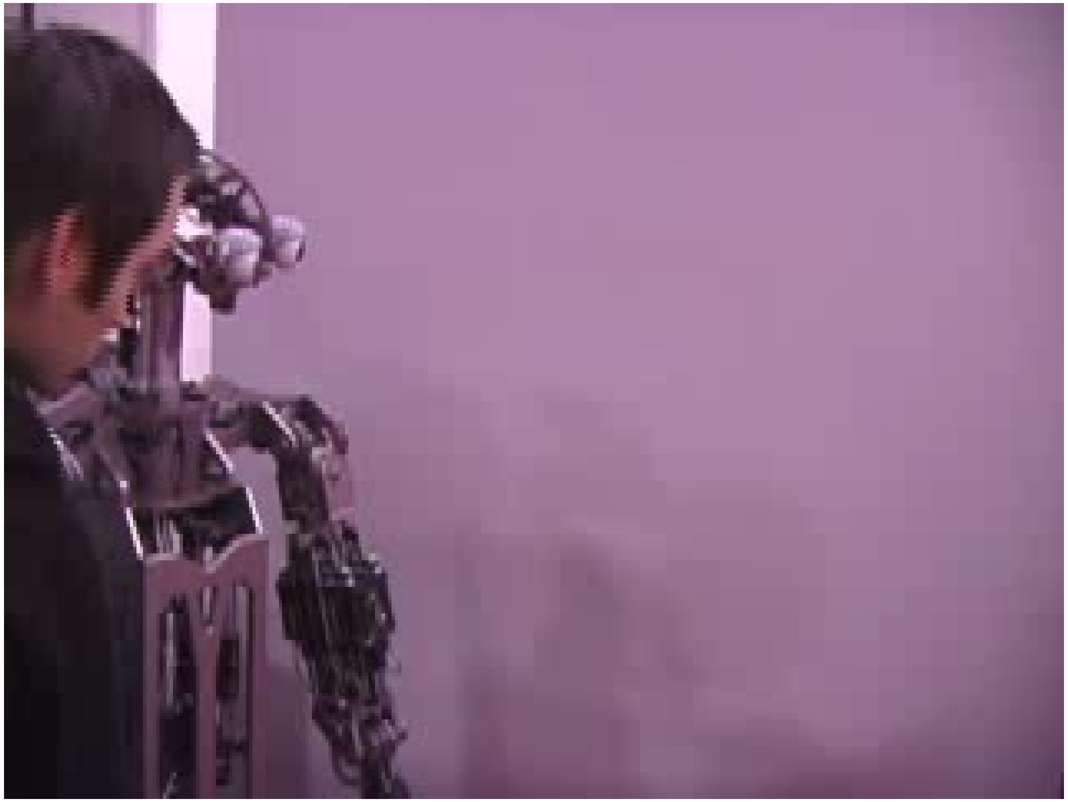
TU Berlin
Oliver Brock

Reading for this set of slides

- There is no reading for this set of slides. Please refer to the reading about Jacobians. We only cover very basic concepts of computer vision and the Web is a fine resource for this.

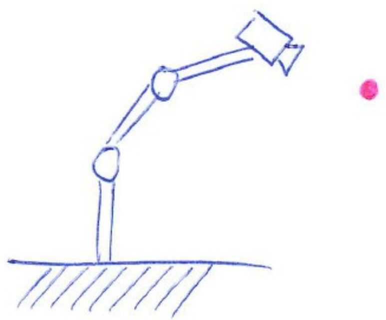
Please note that this set of slides is intended as support for the lecture, not as a stand-alone script. If you want to study for this course, please use these slides in conjunction with the indicated chapters in the text books. The textbooks are available online or in the TUB library (many copies that can be checked out for the entire semester. There are also some aspects of the lectures that will not be covered in the text books but can still be part of the homework or exam. For those It is important that you attend class or ask somebody about what was covered in class.

Visual Servoing

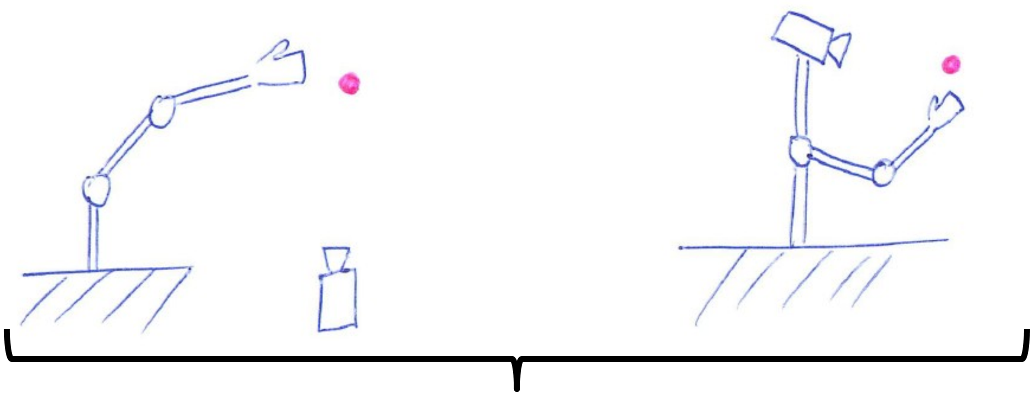


Visual Servo Control (Servoing)

- Using visual perception to control the motion of the robot



eye-in-hand system

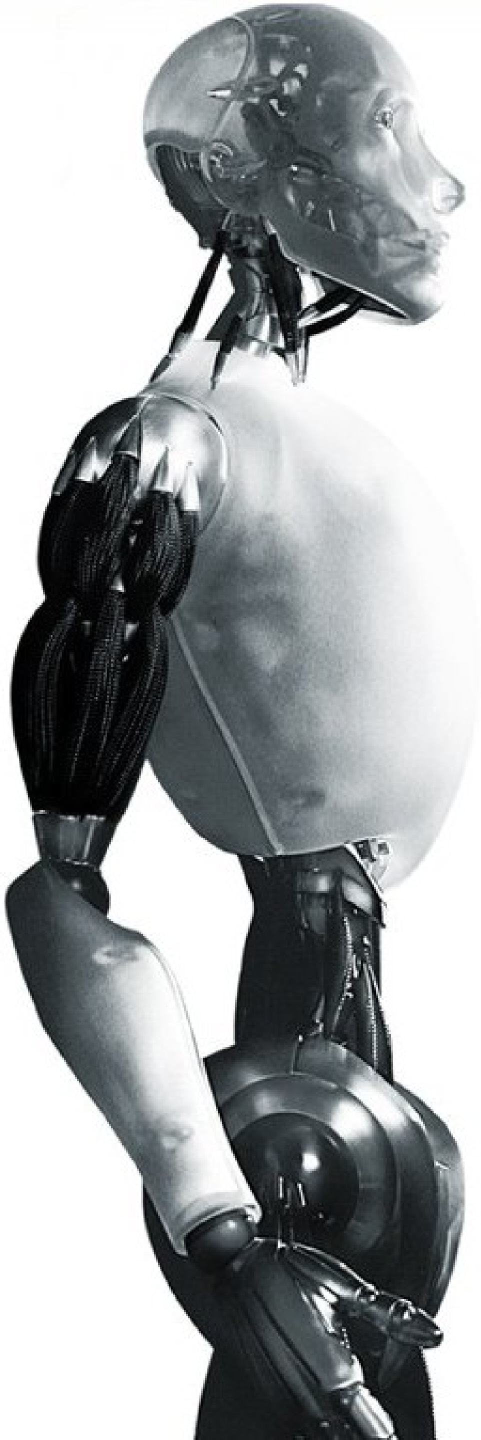


eye-to-hand system

endpoint closed-loop (ECL)
or
endpoint open-loop (EOL)

What do we need for Visual Servoing?

- Vision
- Servoing
- Of course…

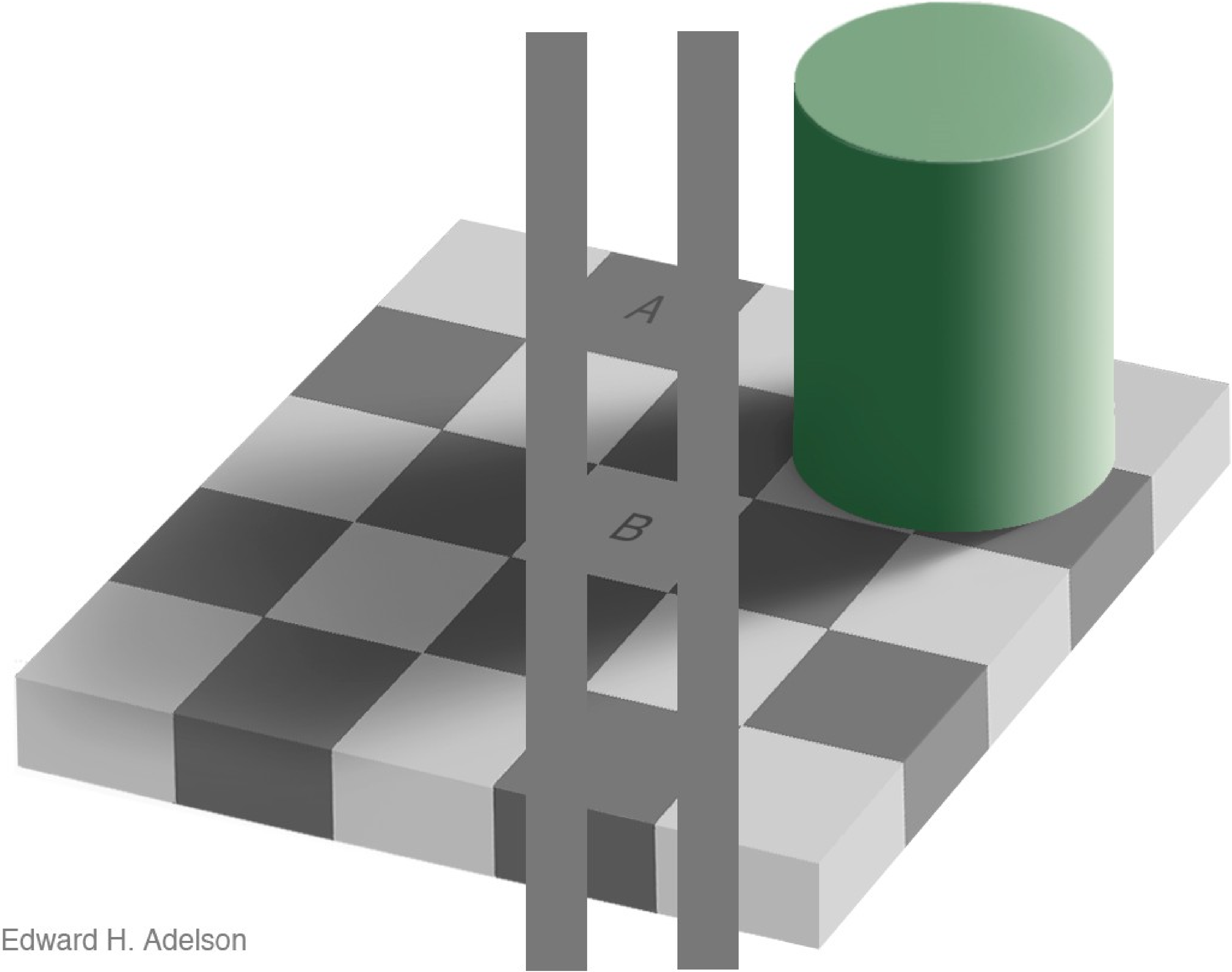


Robotics

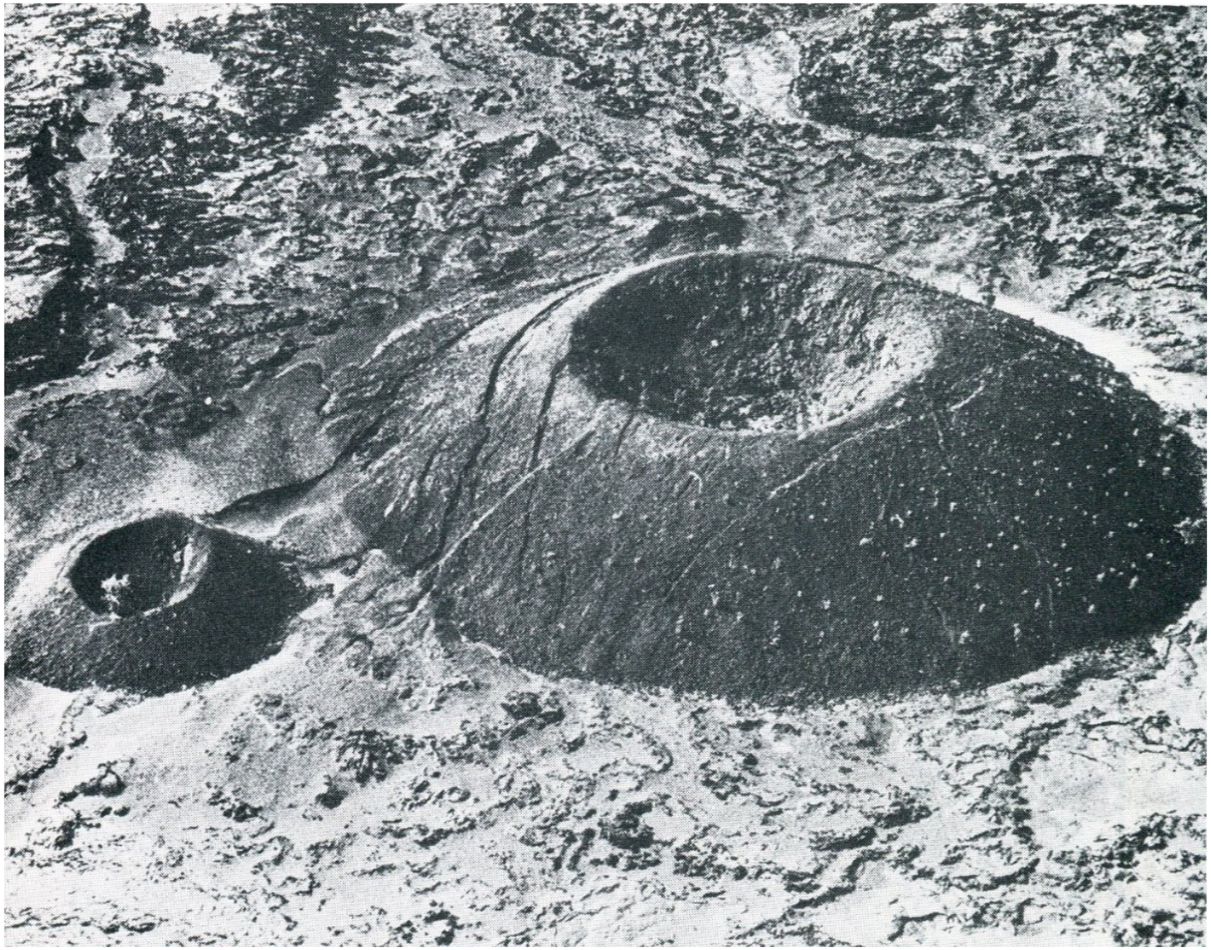
Human Vision

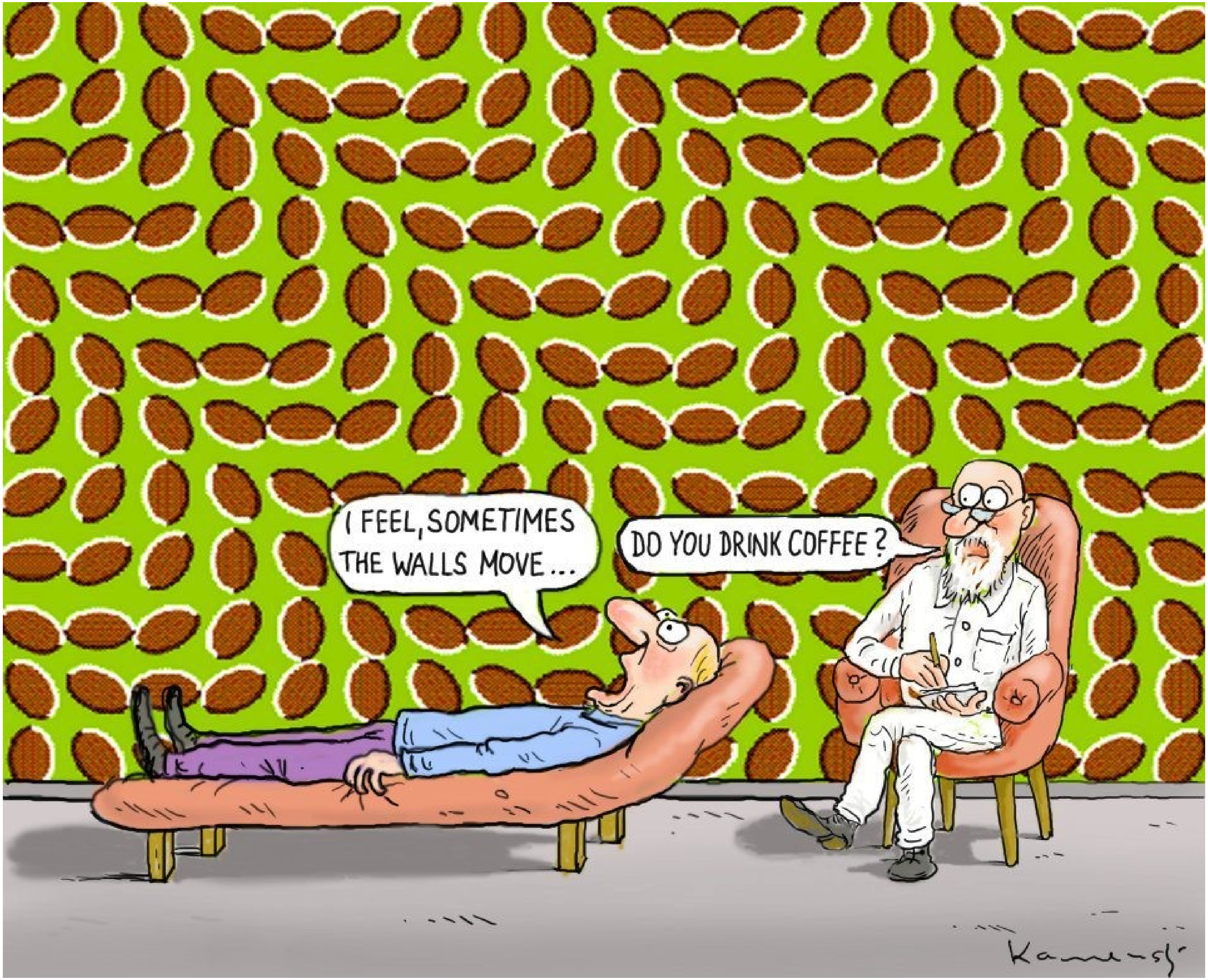
TU Berlin

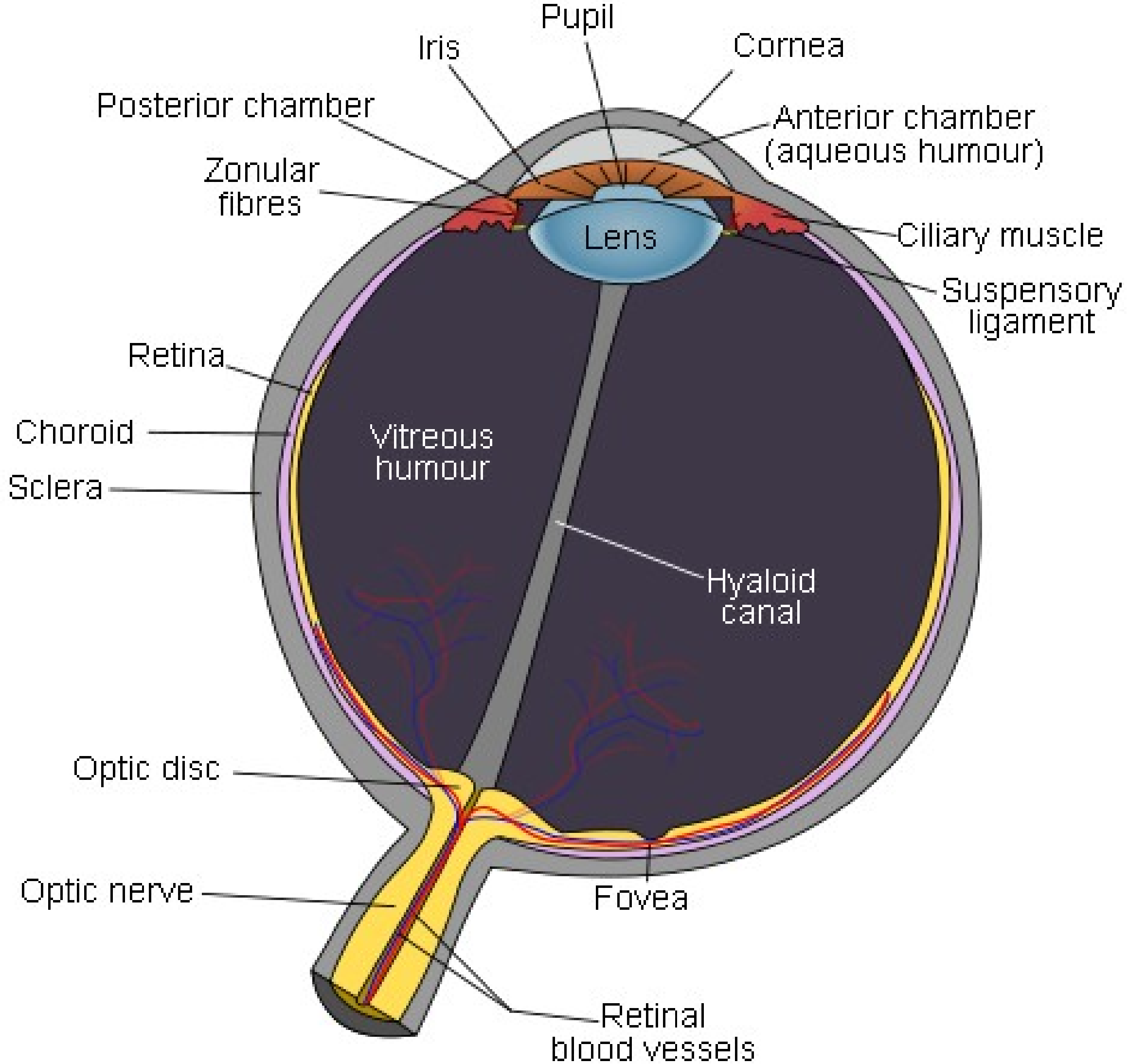
Oliver Brock

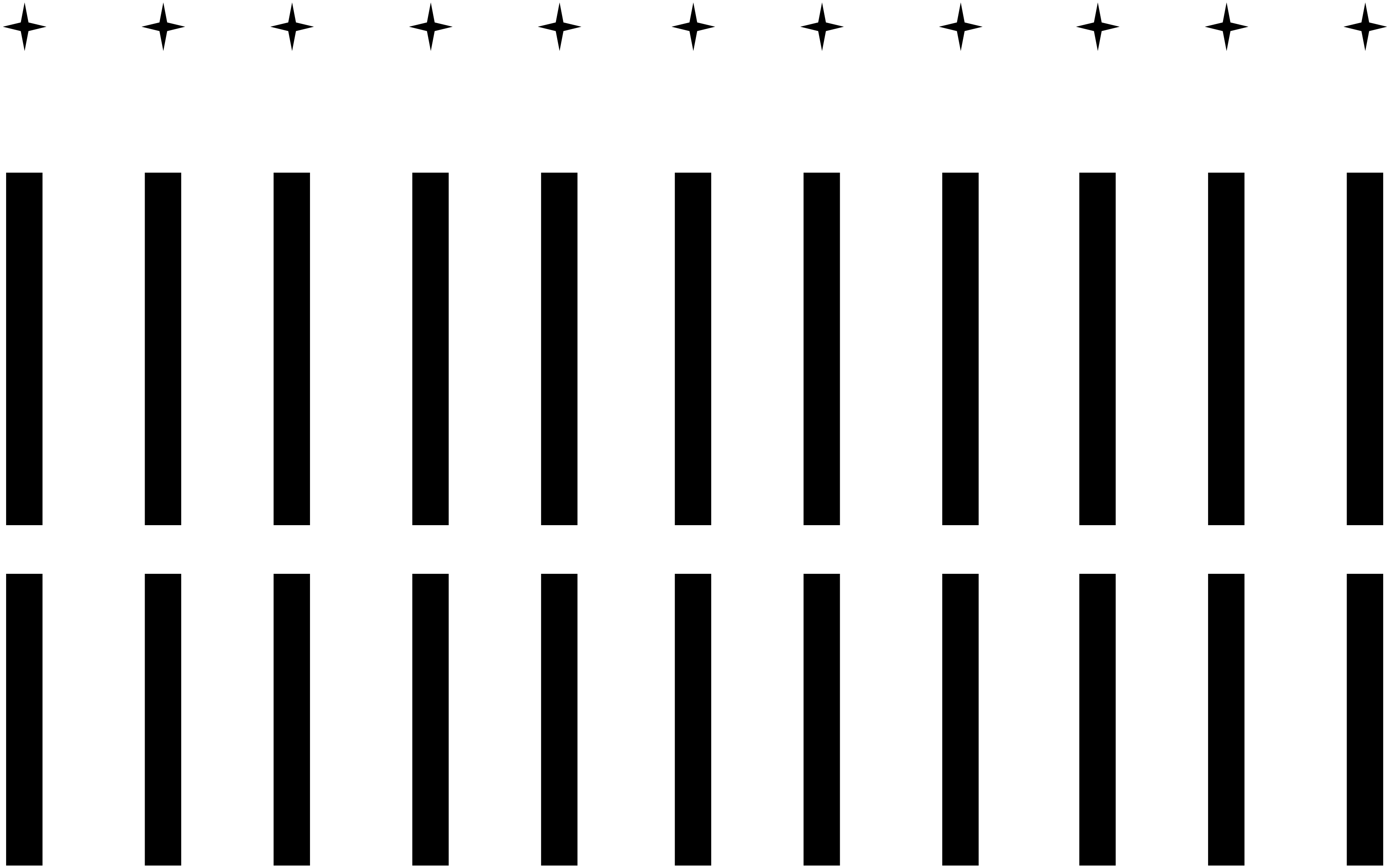


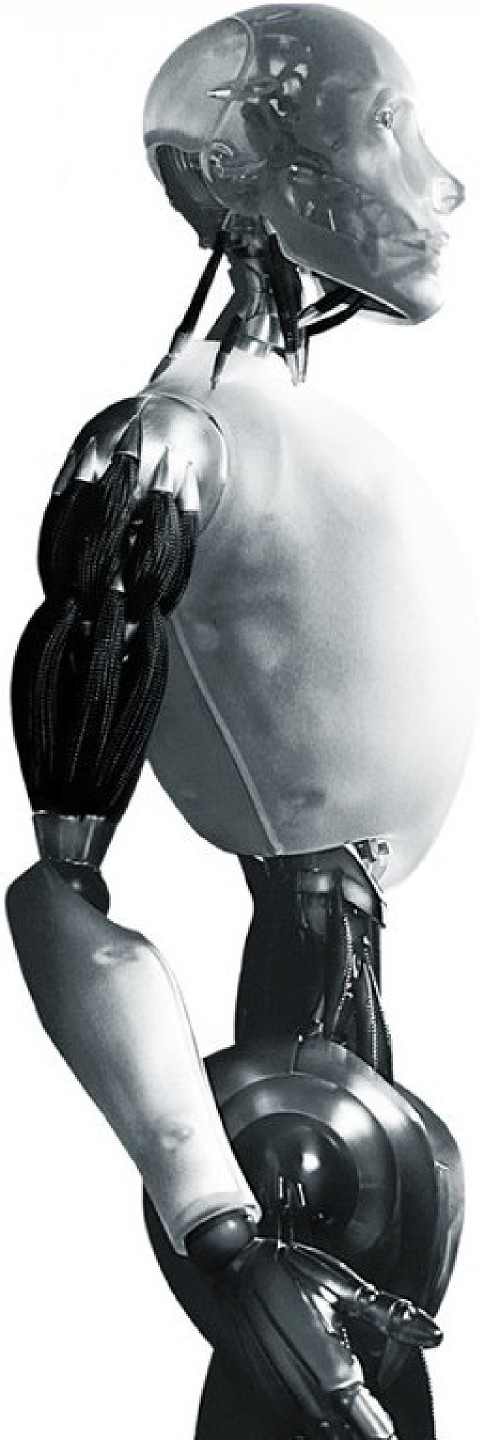
Edward H. Adelson




I FEEL, SOMETIMES THE WALLS MOVE...
DO YOU DRINK COFFEE?
Kamensky


Pupil
Iris
Cornea
Posterior chamber
Anterior chamber
(aqueous humour)
Zonular
fibres
Lens
Ciliary muscle
Suspensory
ligament
Retina
Choroid
Vitreous
humour
Sclera
Hyaloid
canal
Optic disc
Optic nerve
Fovea
Retinal
blood vessels



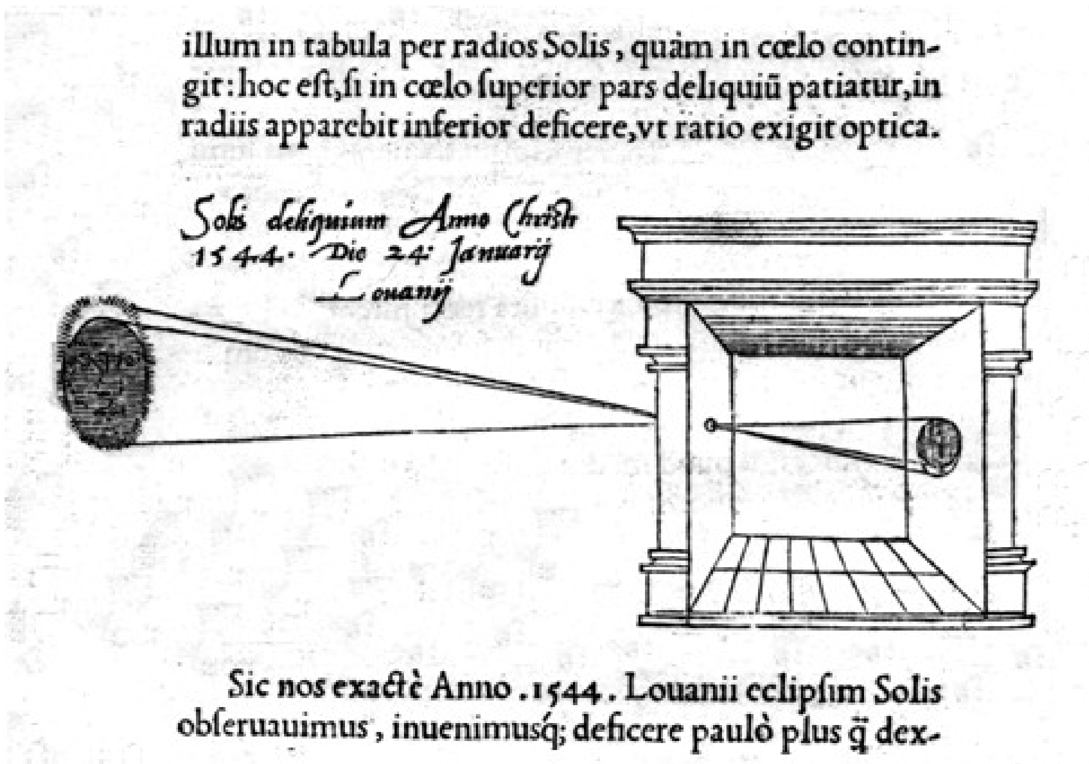


Robotics

Computer Vision

TU Berlin
Oliver Brock

Camera Obscura



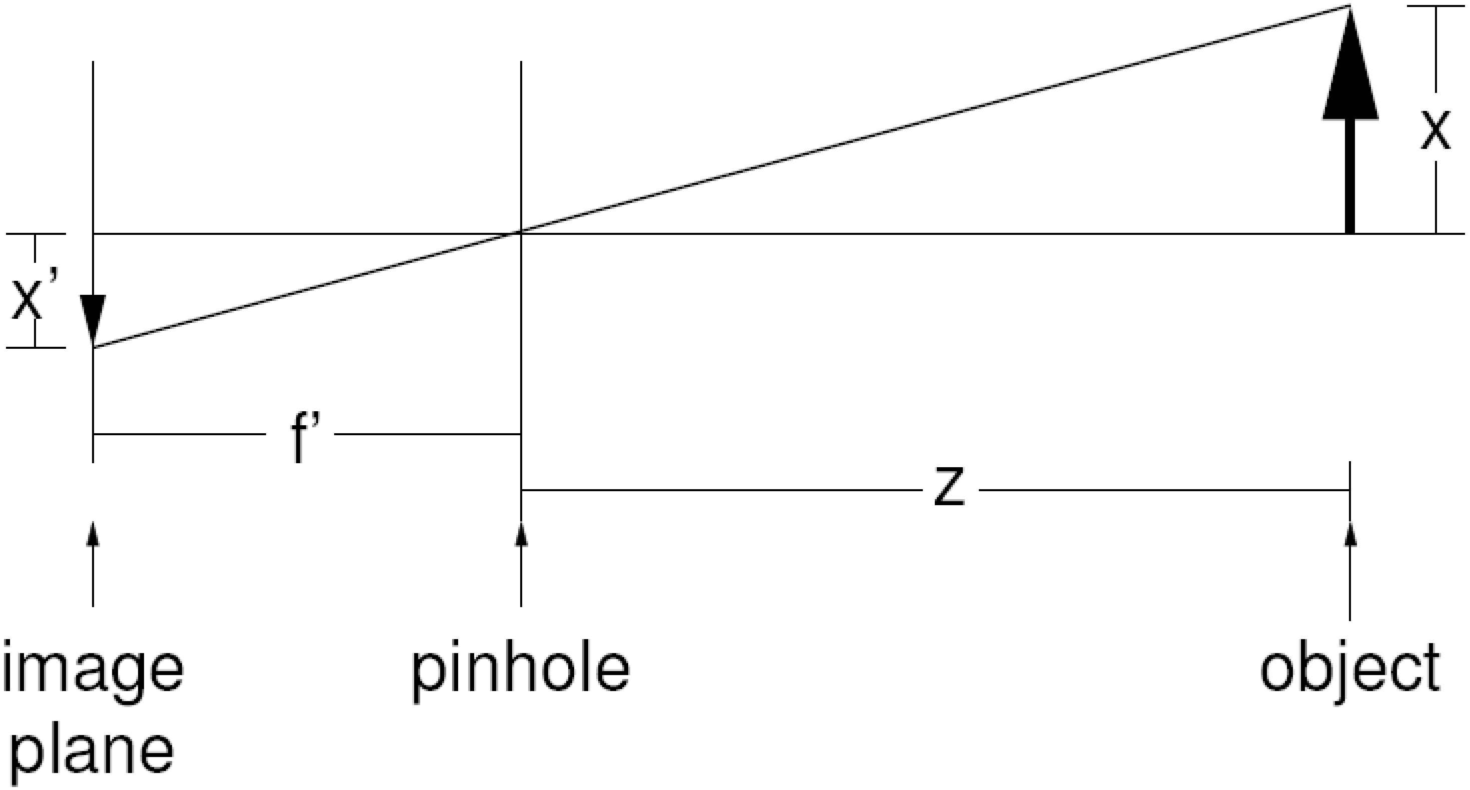
illum in tabula per radios Solis, quàm in cœlo contin-
git: hoc eſt, ſi in cœlo ſuperior pars deliquiũ patiatur, in
radiis apparebit inferior deficere, vt ratio exigit optica.

Sic nos exactè Anno .1544. Louanii eclipſim Solis
obſeruauimus, inuenimusq́; deficere paulò plus q̃ dex-

"When images of illuminated objects ... penetrate through a small hole into a very dark room ... you will see [on the opposite wall] these objects in their proper form and color, reduced in size ... in a reversed position, owing to the intersection of the rays".

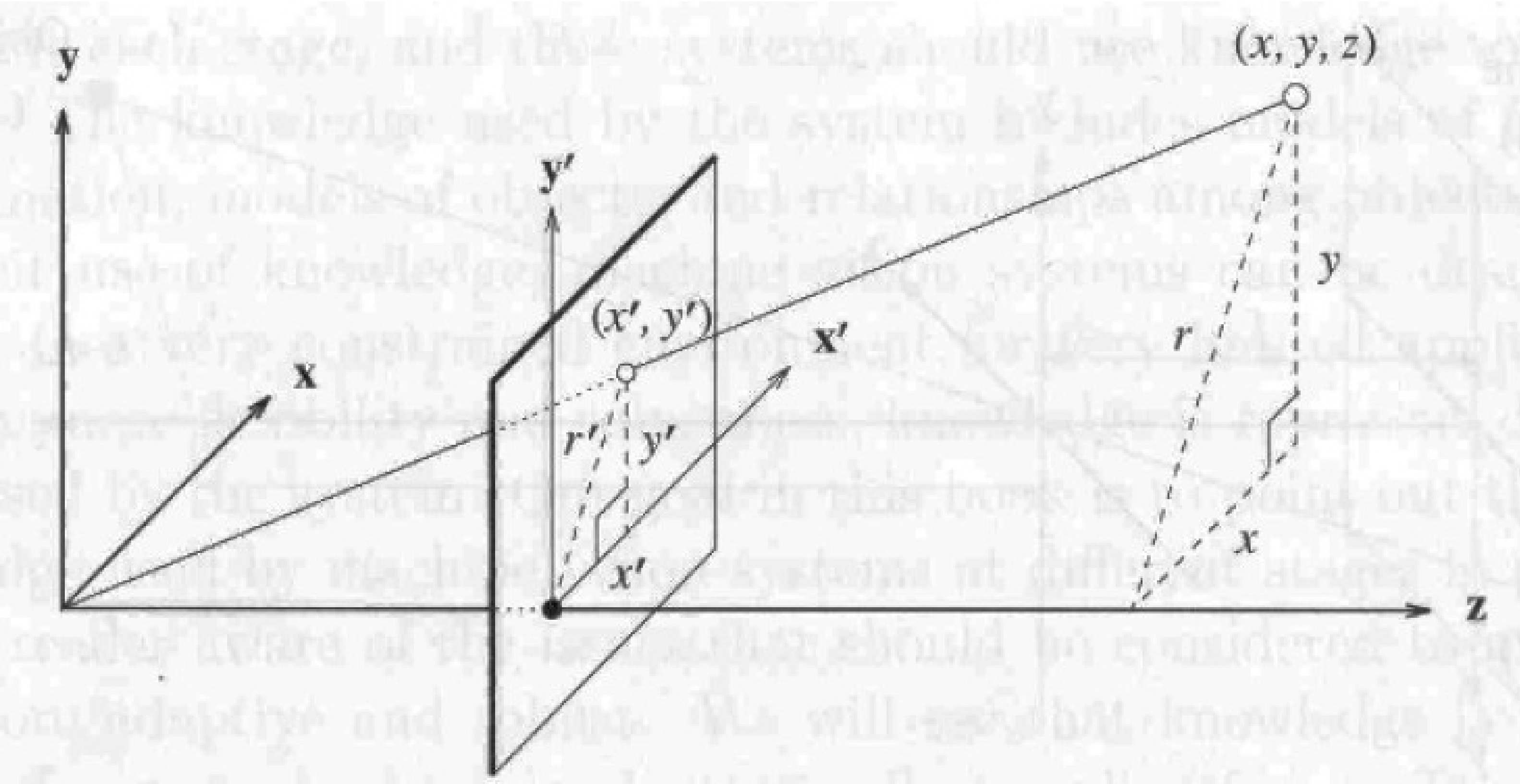
Leonardo Da Vinci

Pinhole Camera in 2D



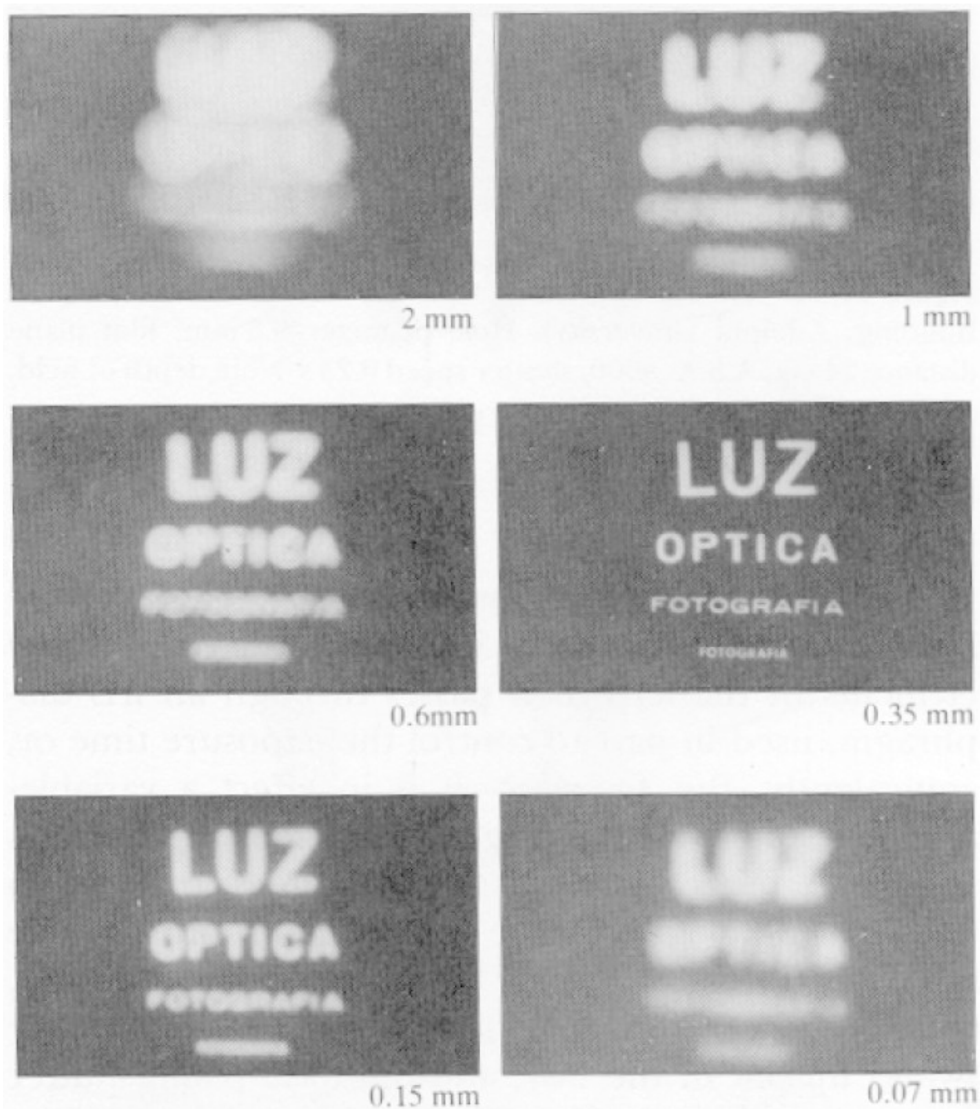
$$X' = f' \frac{X}{Z}$$

Flipped Pinhole Camera in 3D



$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \frac{f}{z} \begin{bmatrix} x \\ y \end{bmatrix}$$

Pinhole Size



Introducing Lenses

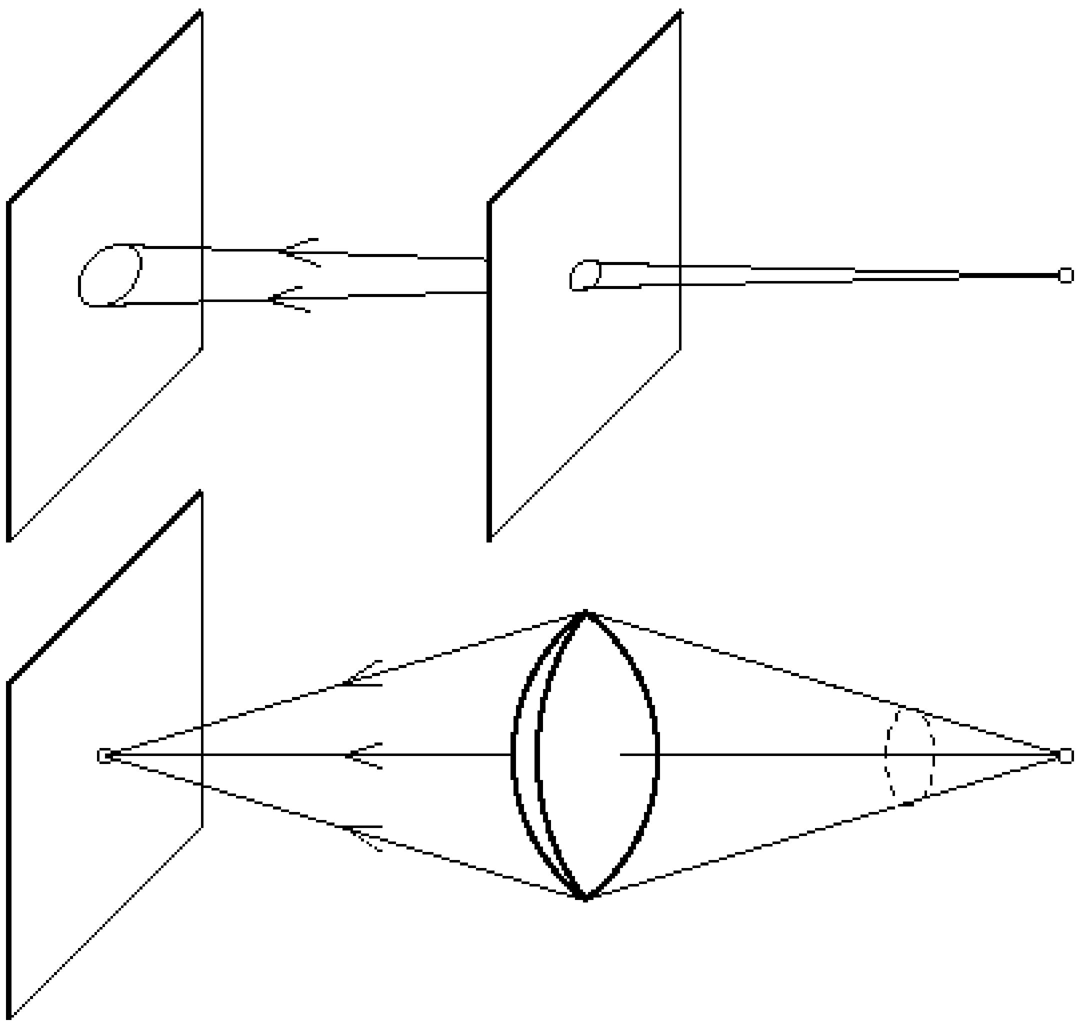
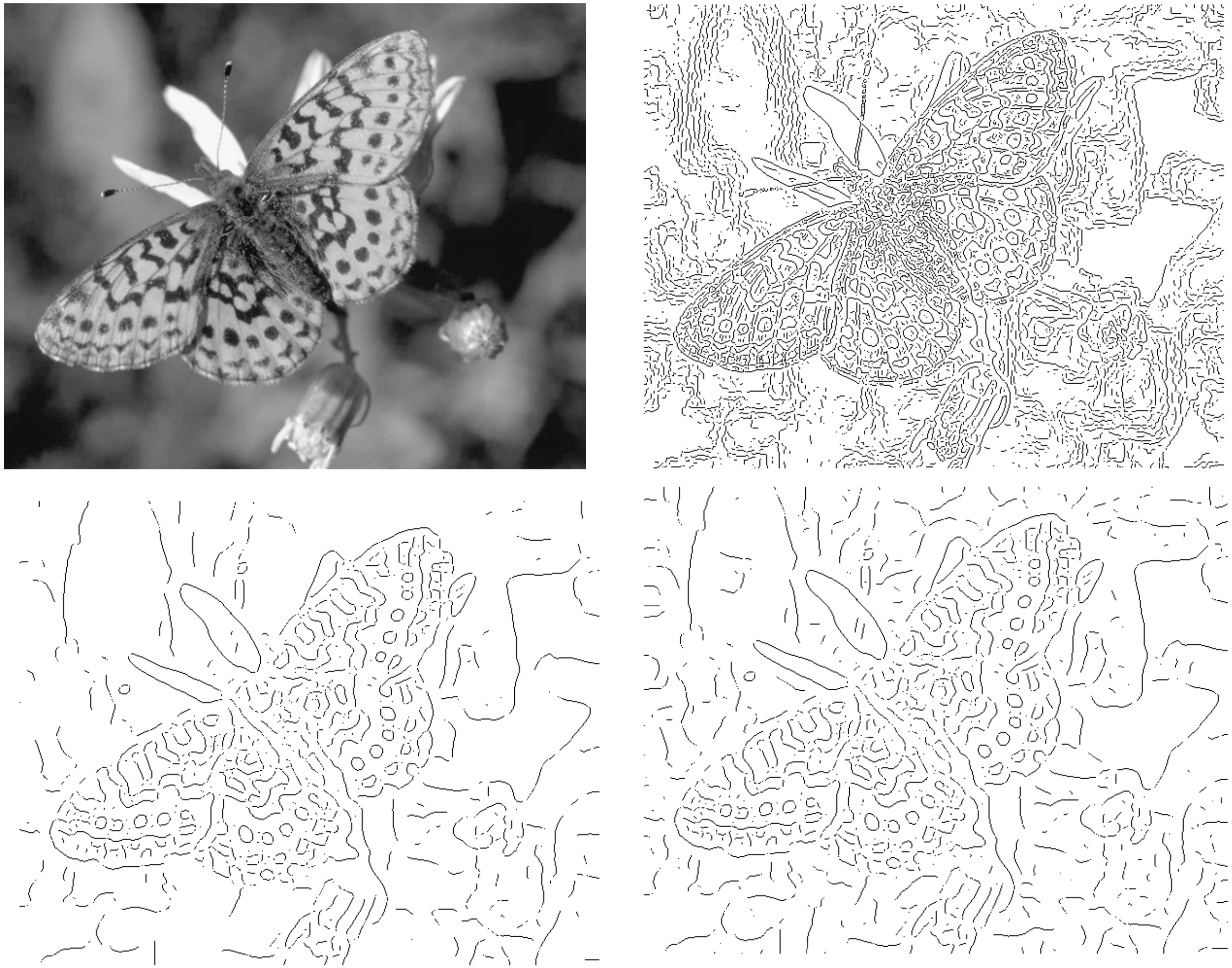


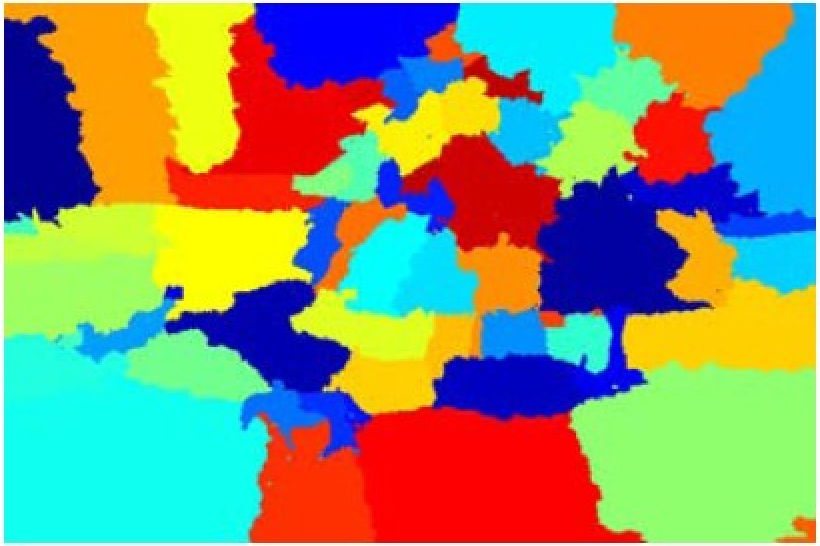
Image Features

Edges, Corners, SIFT, …



Edges

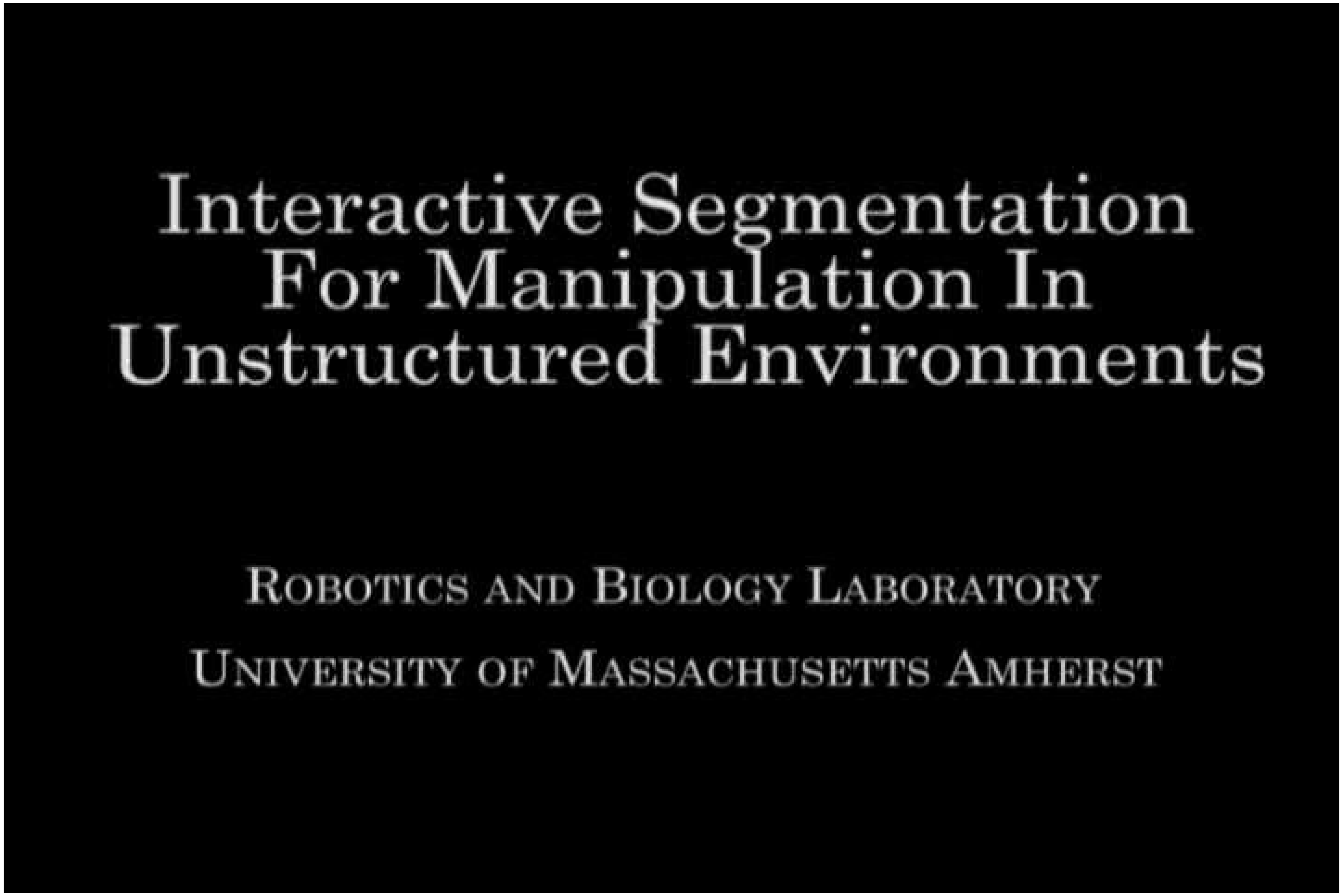
Segmentation



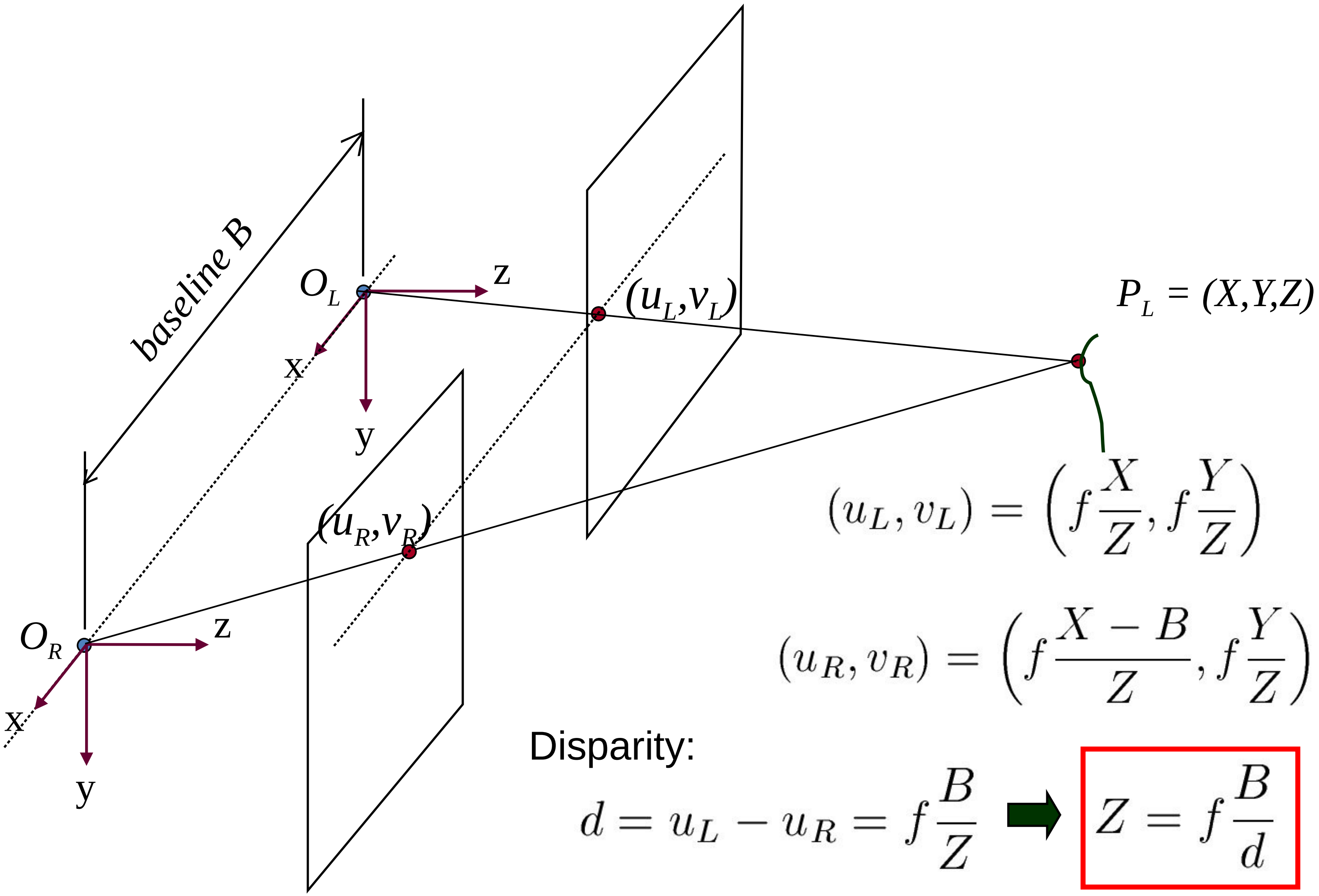
Active Segmentation

(a) (b)

Interactive Segmentation



Stereo Vision



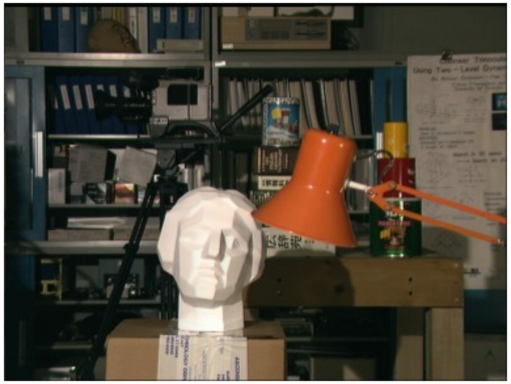
$$(u_L, v_L) = \left(f\frac{X}{Z}, f\frac{Y}{Z}\right)$$

$$(u_R, v_R) = \left(f\frac{X-B}{Z}, f\frac{Y}{Z}\right)$$

Disparity:

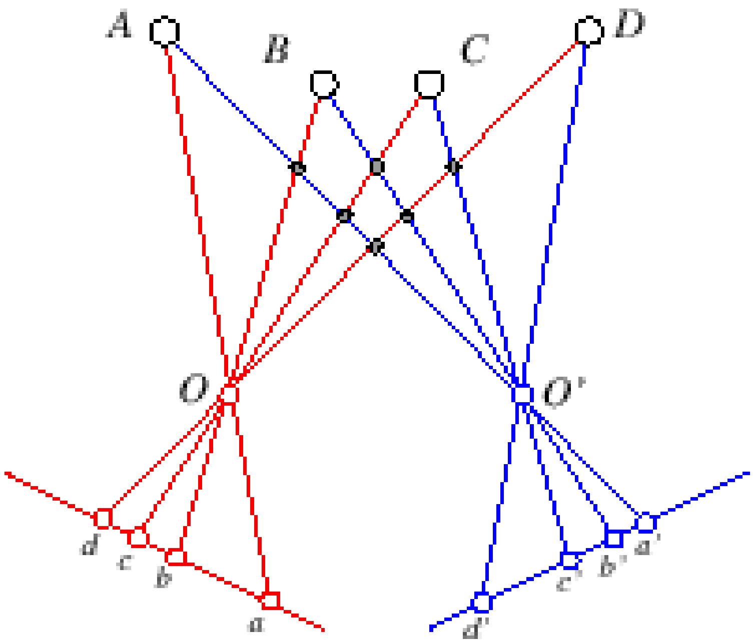
$$d = u_L - u_R = f\frac{B}{Z} \Rightarrow \boxed{Z = f\frac{B}{d}}$$

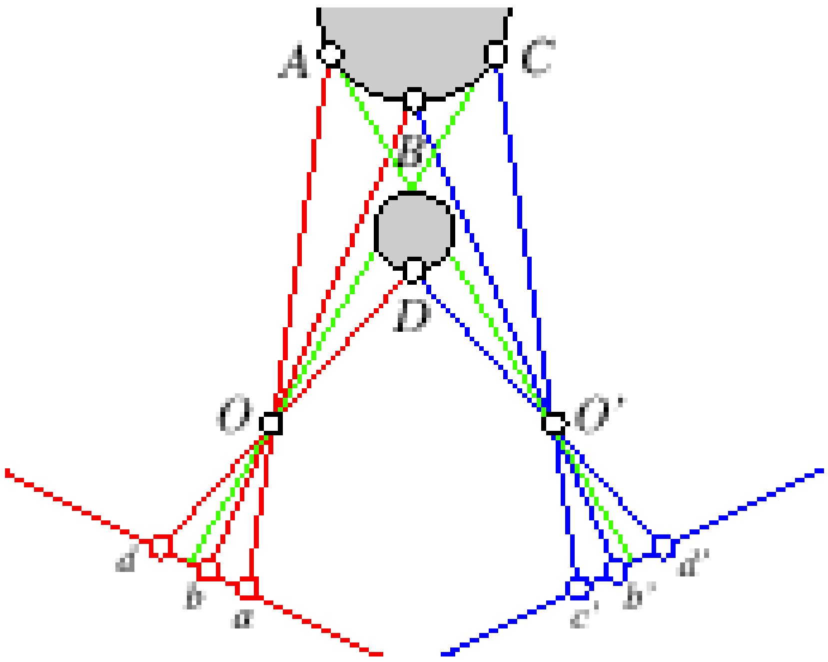
Stereo Results



ground truth

Matching Problem

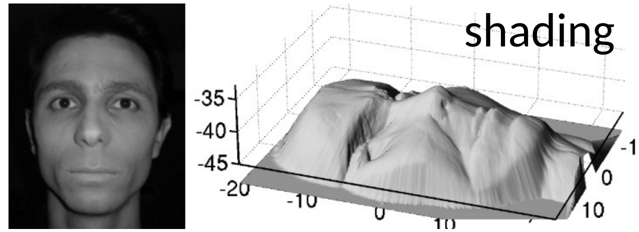




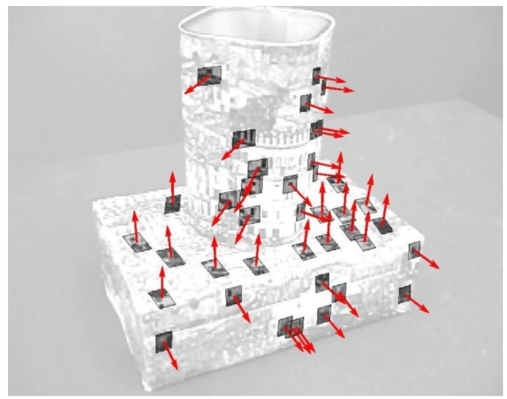
correspondence is inherently ambiguous (here for ordering constraint)

Shape from X

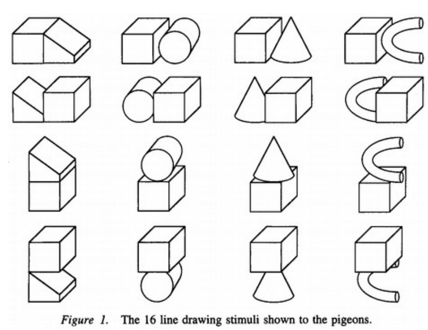
- Motion
  (structure from motion)
- Stereo
- Shading
- Texture
- Focus
- Line drawings



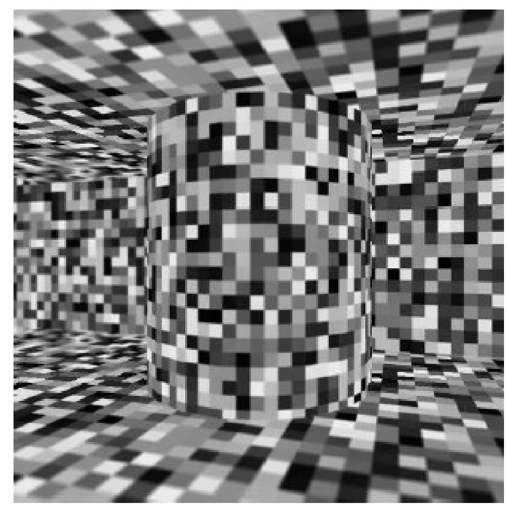
shading



motion

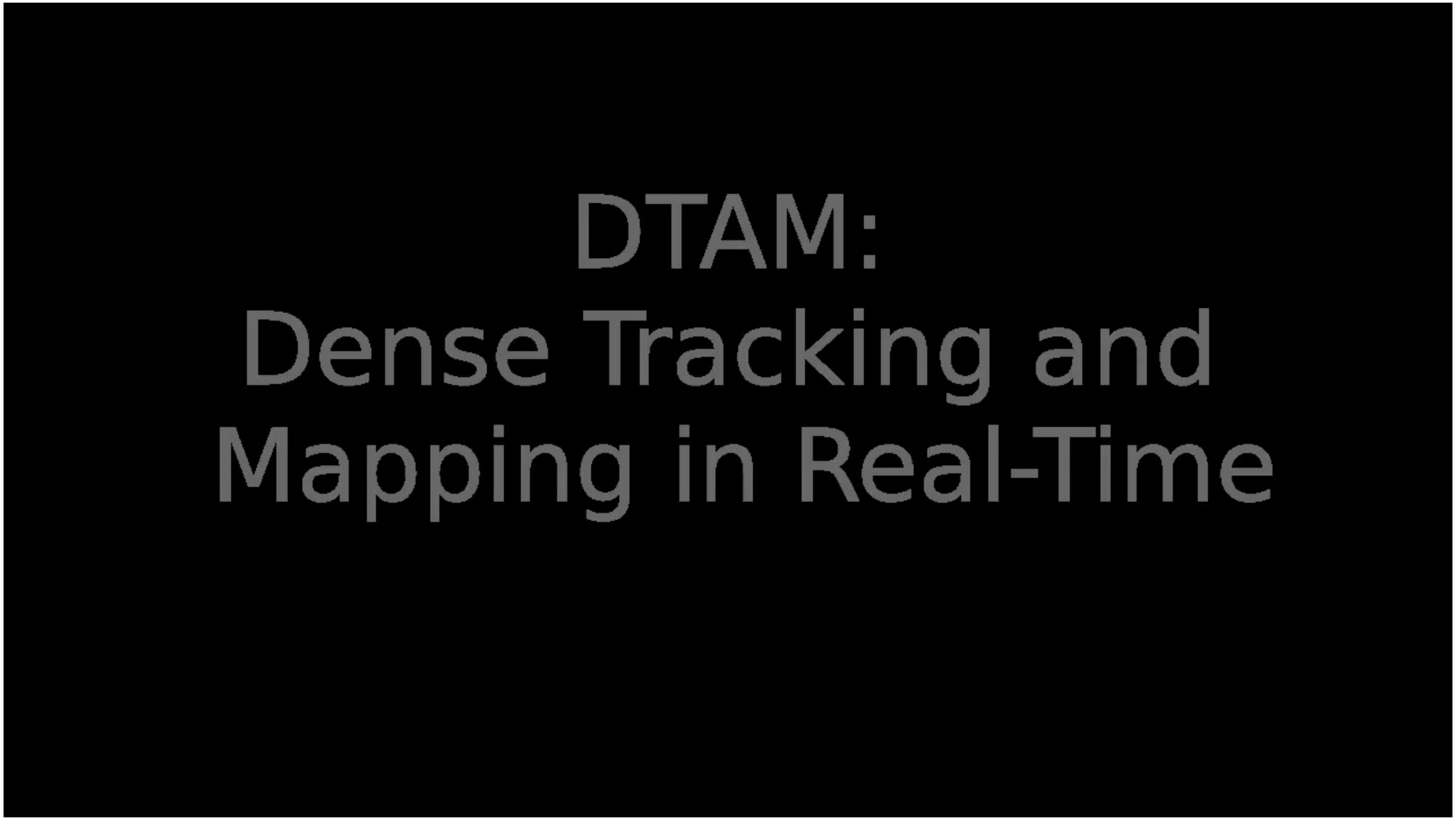


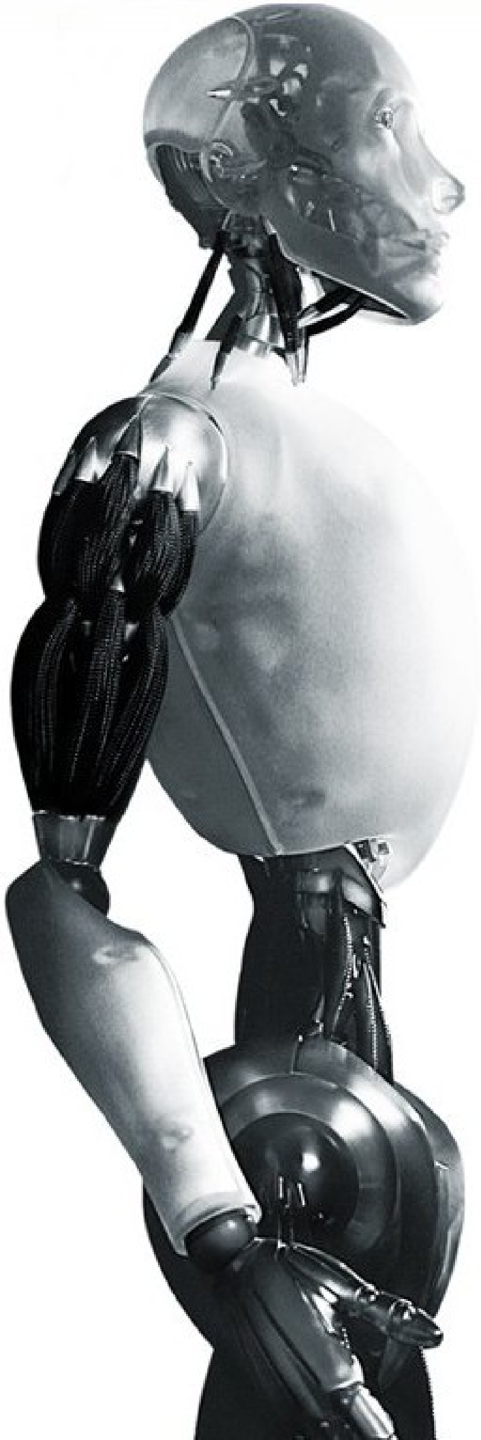
line drawings



texture

DTAM: Dense Tracking and Mapping





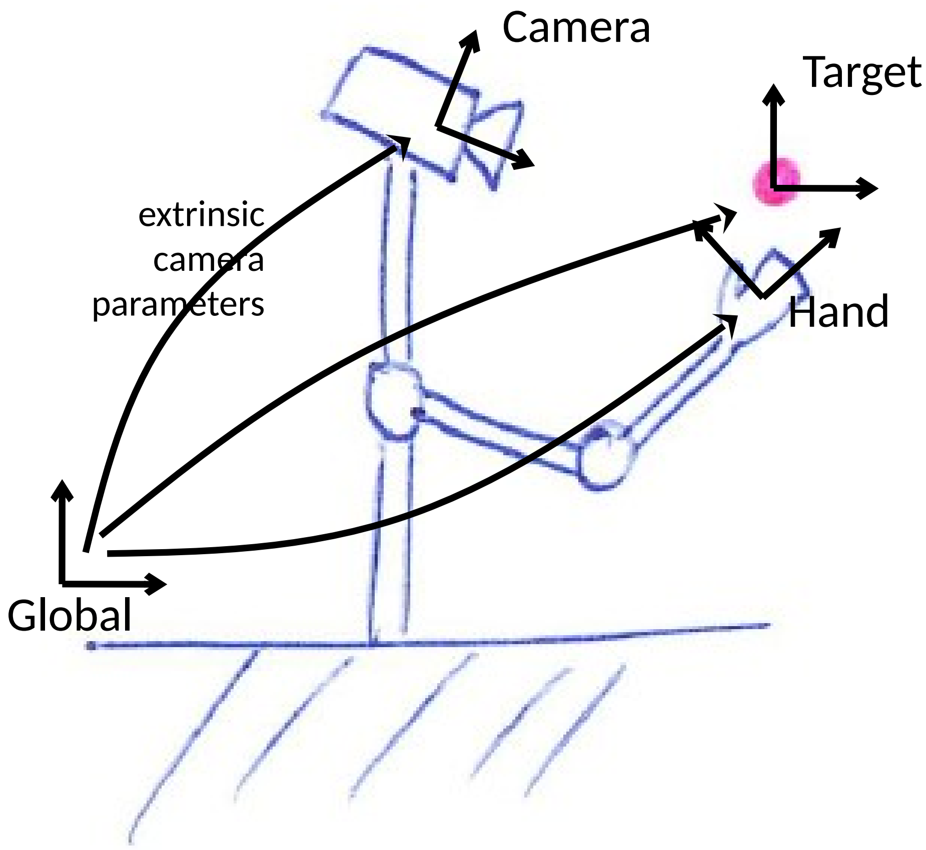
Robotics

Visual Servoing

TU Berlin

Oliver Brock

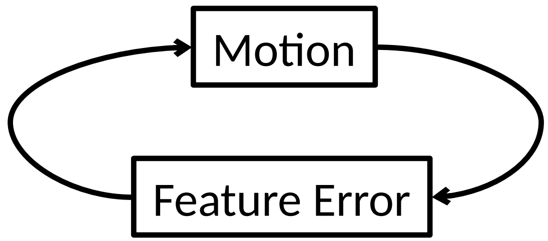
Frames and Parameters

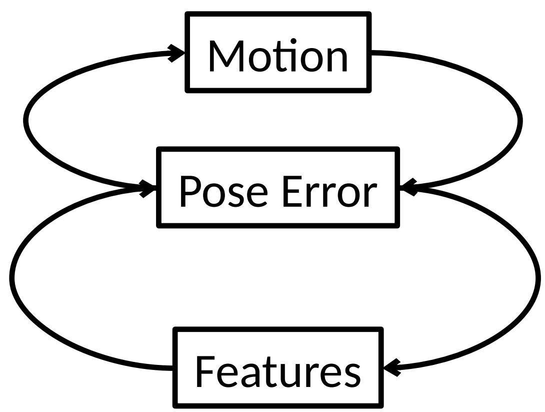


Intrinsic parameters

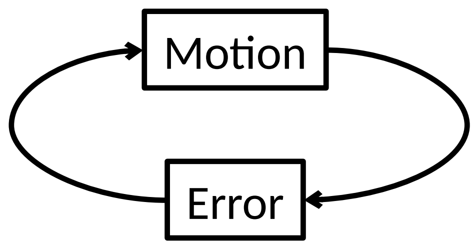
- focal length
- center of image plane
- axis scaling
- distortion
- pixel size
- …

Image-based versus Position-based

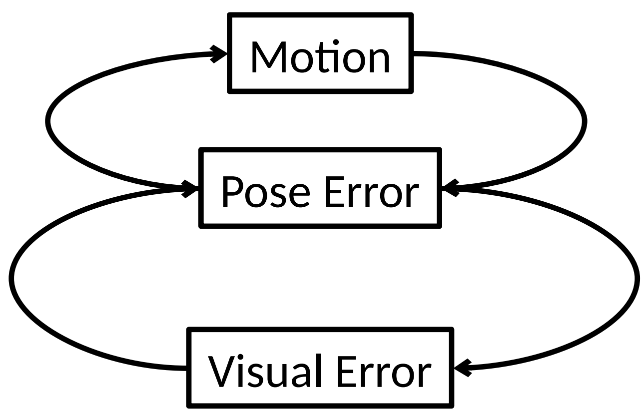




Dynamic Look-and-Move versus Direct Visual Servo



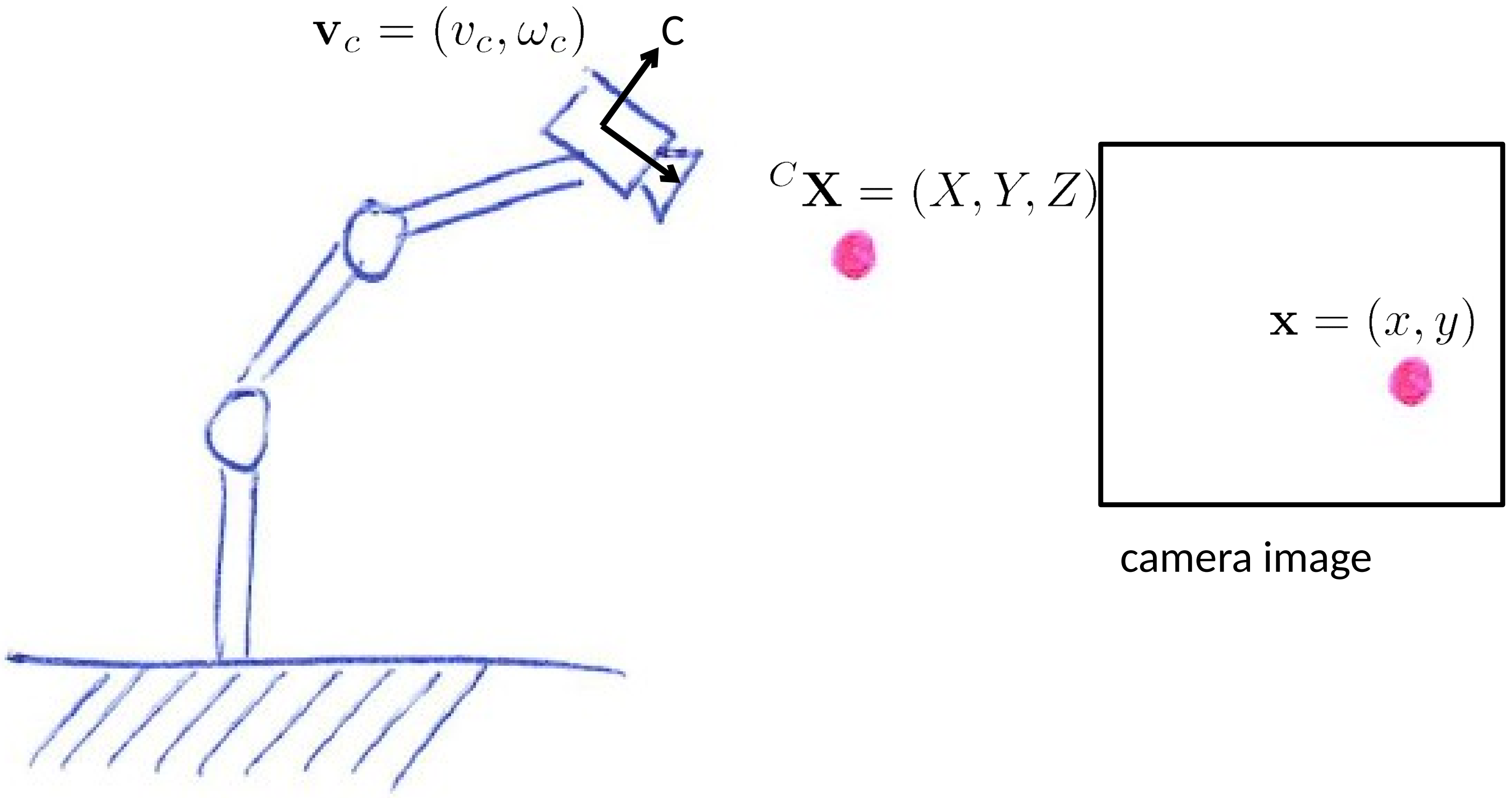
All in one servo loop



High-frequency motion loop

Low-frequency vision loop

How to servo? (image-based)



Servo to center the target in the image!

Key: Image Jacobian

vector of image features

$$\mathbf{s}$$

$$\dot{\mathbf{s}} = J_I \mathbf{v}_c$$

velocities of image features

camera velocity

$$\mathbf{e}(t) = \mathbf{s} - \mathbf{s}_{\text{des}}$$

$$\dot{\mathbf{e}} = J_I \mathbf{v}_c$$

$$\mathbf{v}_c = -\lambda J^{-1} \mathbf{e}$$

with $\dot{\mathrm{e}} = -\lambda \mathbf{e}$

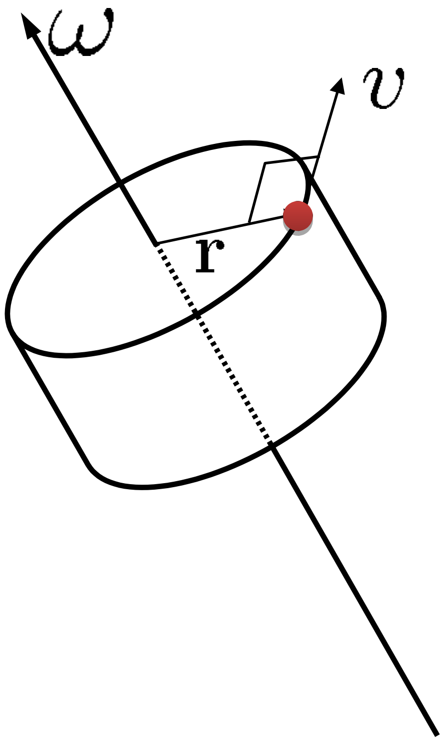
$$\mathbf{v}_c = -\lambda J^{+} \mathbf{e}$$

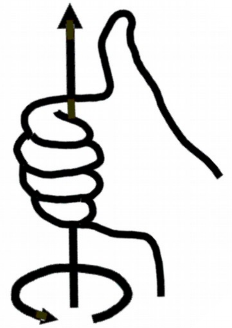
$J^+ = (J^T J)^{-1} J^T$

Velocity of a Rigid Body

$$v_r = \omega \times \mathbf{r}$$

$$v = v_t + v_r = v_t + \omega \times \mathbf{r}$$





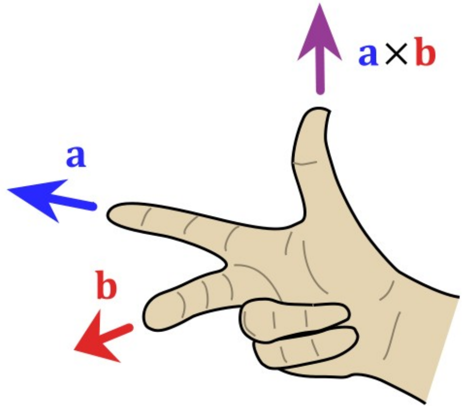
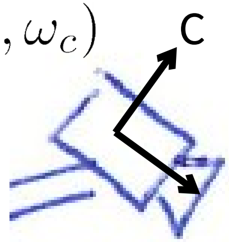


Image Jacobian/Interaction Matrix

$\mathbf{v}_c = (v_c, \omega_c)$



$^{C}\mathbf{X} = (X, Y, Z)$

$$\boxed{\begin{array}{c}\mathbf{x} = (x, y) \\ \\ \mathbf{m} = (u, v)\end{array}}$$

$$\begin{aligned} x &= X/Z = (u - c_u)/f\alpha \\ y &= Y/Z = (v - c_v)/f, \end{aligned}$$

$$\begin{aligned} \dot{x} &= \dot{X}/Z - X\dot{Z}/Z^2 &= (\dot{X} - x\dot{Z})/Z \\ \dot{y} &= \dot{Y}/Z - Y\dot{Z}/Z^2 &= (\dot{Y} - y\dot{Z})/Z. \end{aligned}$$

$$\dot{\mathbf{X}} = -\boldsymbol{v}_c - \omega_c \times \mathbf{X} \Leftrightarrow \begin{cases} \dot{X} = -v_x - \omega_y Z + \omega_z Y \\ \dot{Y} = -v_y - \omega_z X + \omega_x Z \\ \dot{Z} = -v_z - \omega_x Y + \omega_y X \end{cases}$$

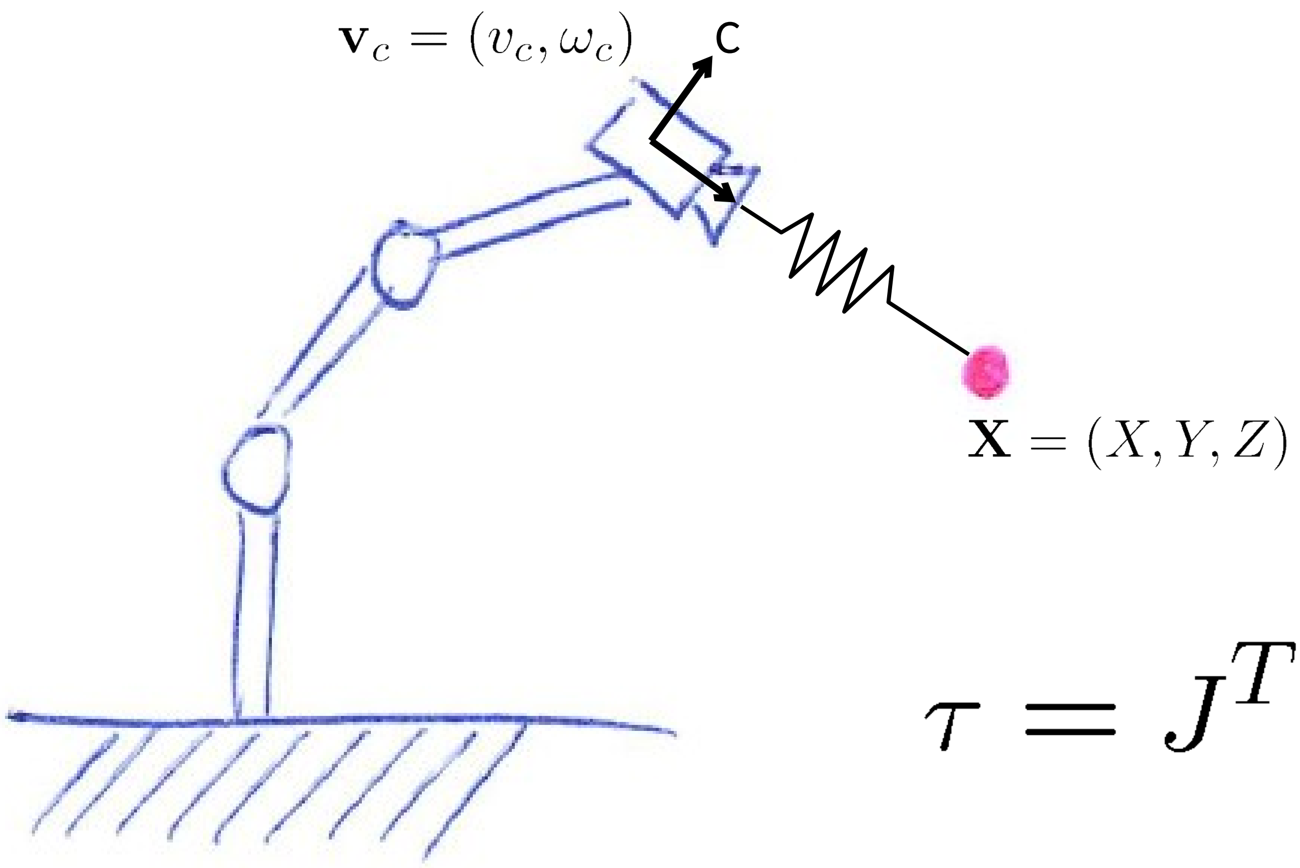
$$\begin{aligned} \dot{x} &= -v_x/Z + x v_z/Z + xy\omega_x - (1 + x^2)\omega_y + y\omega_z \\ \dot{y} &= -v_y/Z + y v_z/Z + (1 + y^2)\omega_x - xy\omega_y - x\omega_z \end{aligned}$$

$$\dot{\mathbf{x}} = \mathbf{L_x} \mathbf{v}_c \quad \mathbf{L_x} = \begin{bmatrix} \frac{-1}{Z} & 0 & \frac{x}{Z} & xy & -(1 + x^2) & y \\ 0 & \frac{-1}{Z} & \frac{y}{Z} & 1 + y^2 & -xy & -x \end{bmatrix}$$

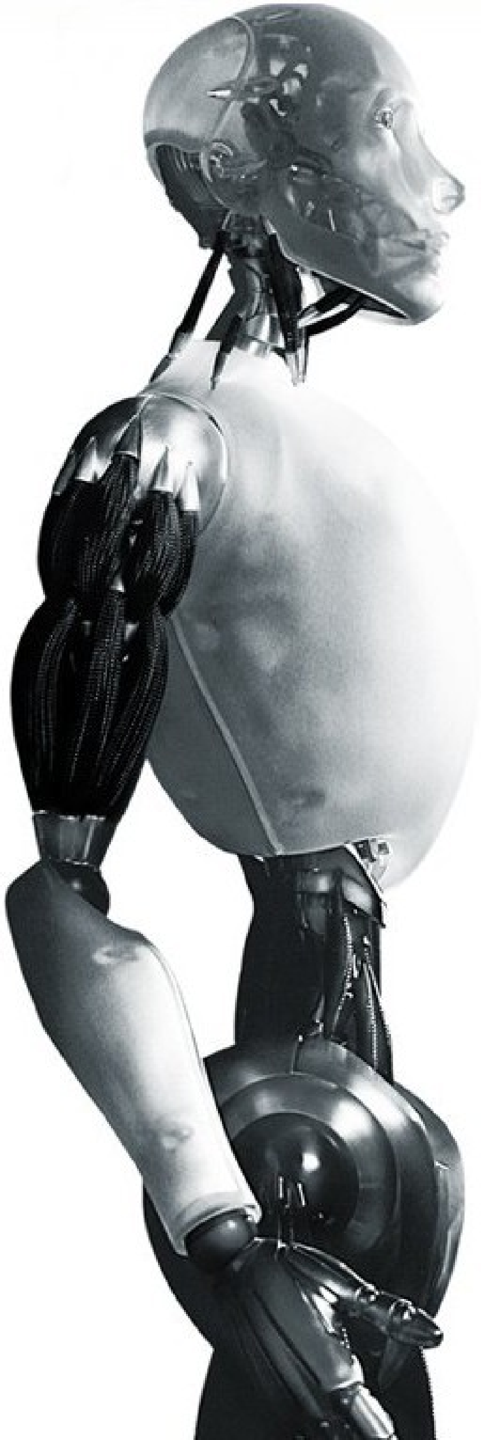
Things to Consider

- The equations require an estimate of Z (depth)
- We need a sufficient number of image features (not a problem usually, we can simply stack the Jacobians)
- Singularities may occur

How to servo? (image-based)



$$\tau = J^T\mathbf{F}?$$



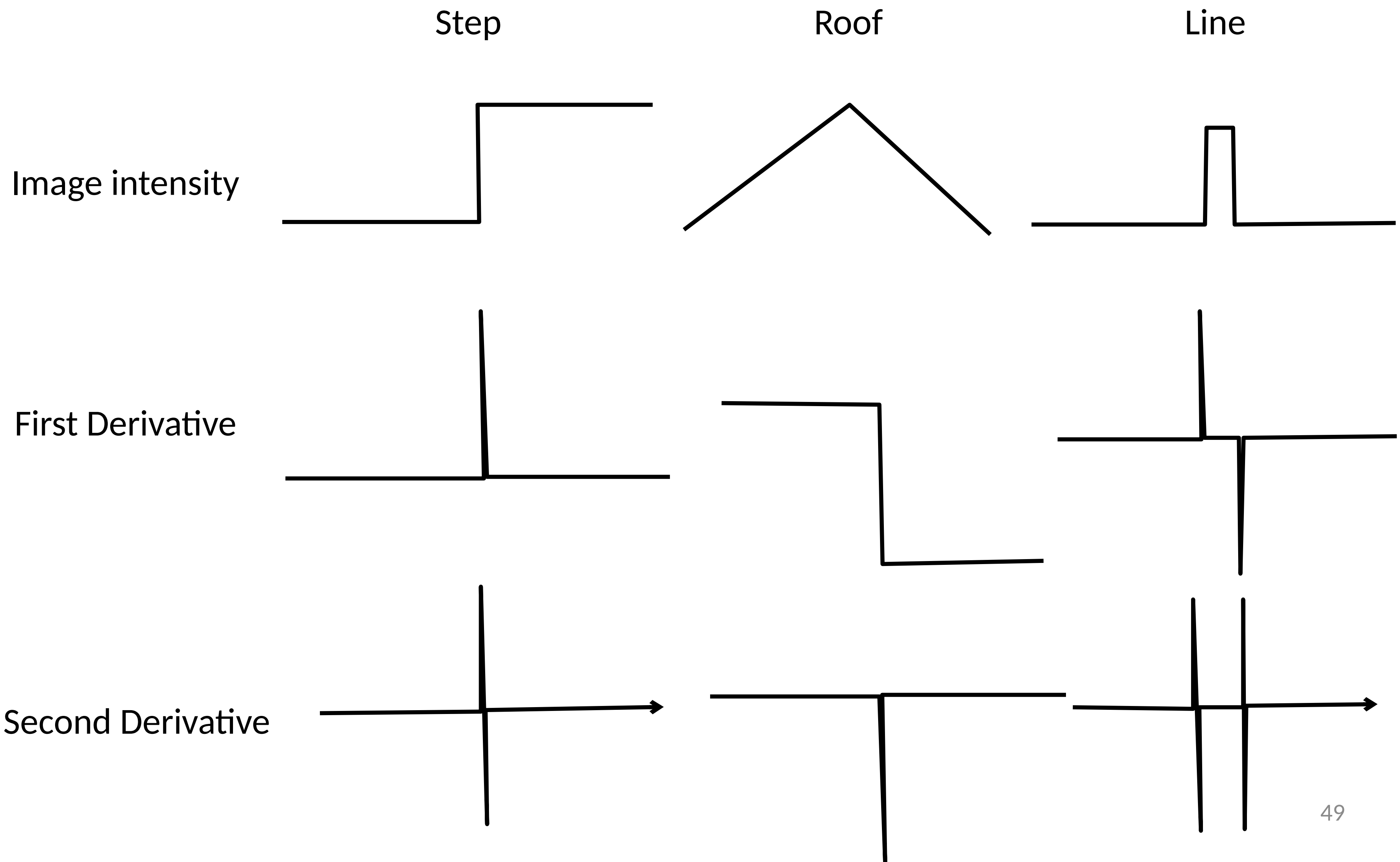
Robotics

Edge Detection, Hough Transform

TU Berlin
Oliver Brock

Edges

What is an edge?



First-Order Difference Operators

$$\begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} \qquad \begin{matrix} 1 & 0 \\ 0 & -1 \end{matrix}$$

Roberts’ cross operator

$$\begin{matrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{matrix} \qquad \begin{matrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{matrix}$$

3×3 Prewitt operator

$$\begin{matrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{matrix} \qquad \begin{matrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{matrix}$$

Sobel operator

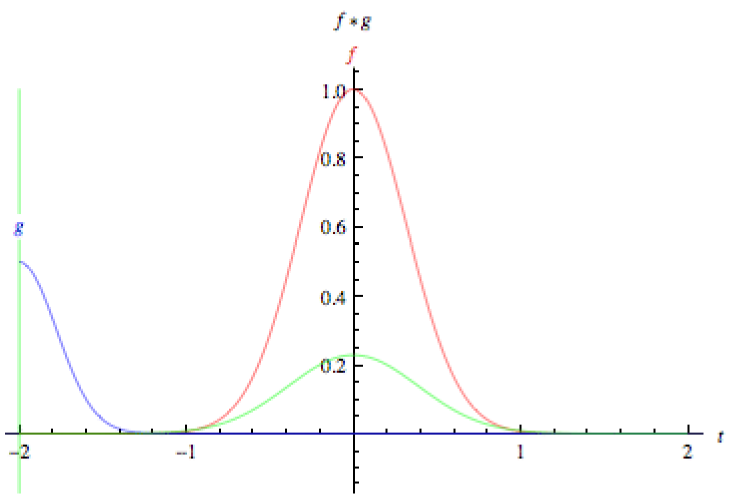
$$\begin{matrix} -3 & -1 & 1 & 3 \\ -3 & -1 & 1 & 3 \\ -3 & -1 & 1 & 3 \\ -3 & -1 & 1 & 3 \end{matrix} \qquad \begin{matrix} 3 & 3 & 3 & 3 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \\ -3 & -3 & -3 & -3 \end{matrix}$$

4×4 Prewitt operator

Convolution

$$f \circ g = \int_{-\infty}^{\infty} f(\tau) g(t-\tau) d\tau$$

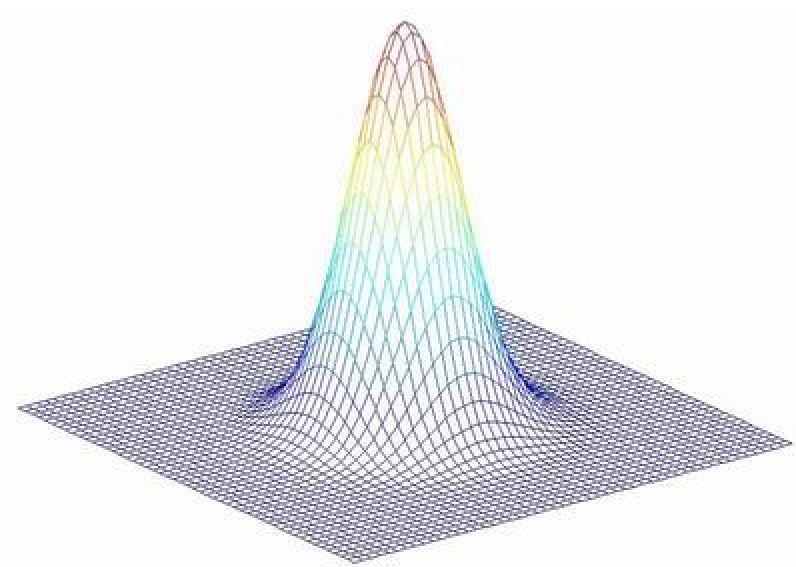
$$f \circ g = \int_{0}^{t} f(\tau) g(t-\tau) d\tau$$

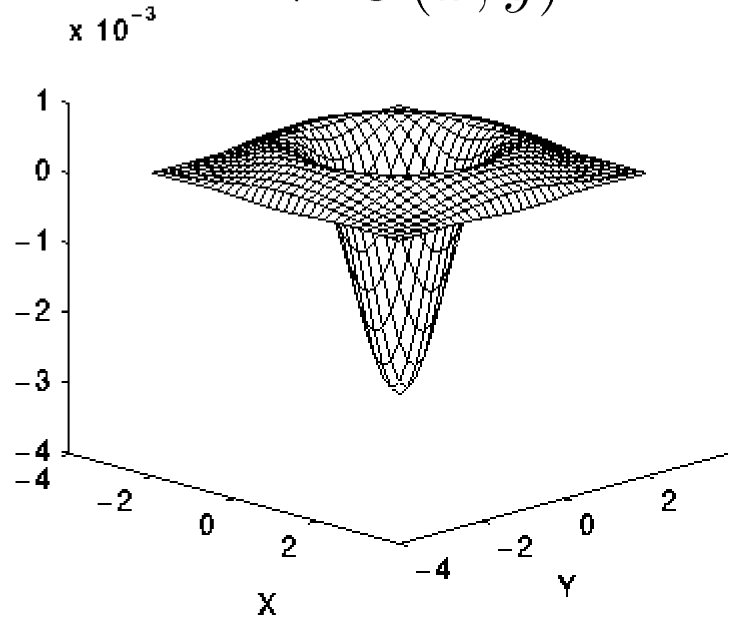


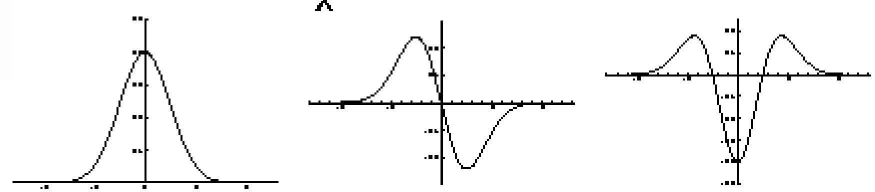
Edge Detection with Gaussians

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

$$\nabla^2 G(x,y)$$





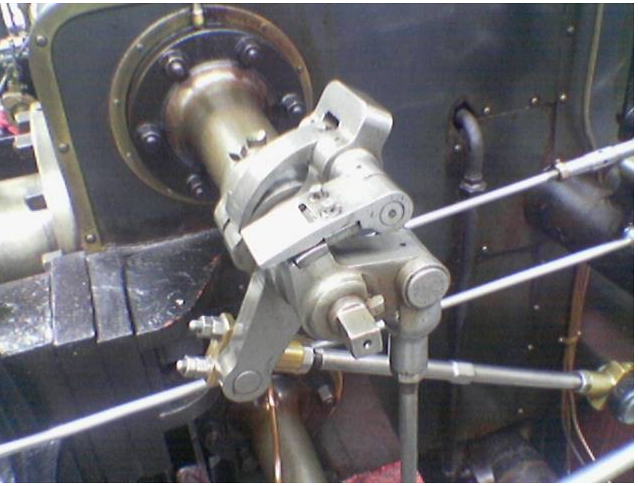


Second-order edge detection:
detection of zero-crossings in the second derivative captures local maxima in the gradient

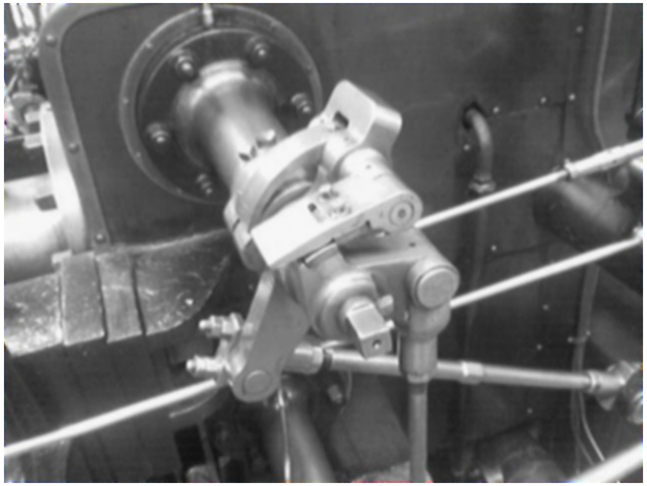
Canny Edge Detector

- Smoothing with a Gaussian
- Determine intensity gradients in four directions |– /\
- Find local maxima of gradient magnitude to determine the “direction” of a point
- Extract edges with hysteresis

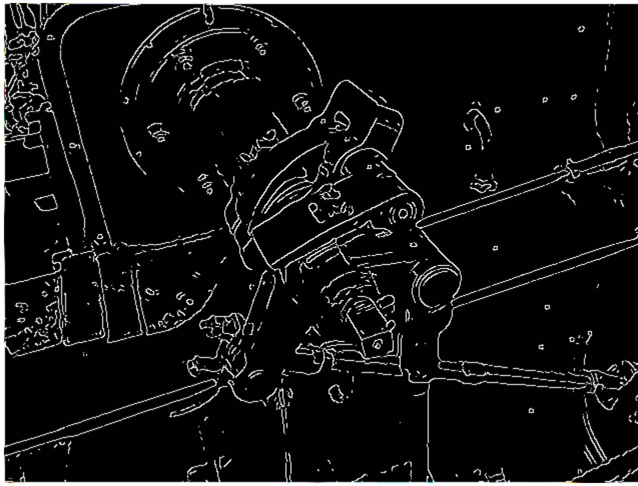
Canny Edge Detector: Example 1



original image



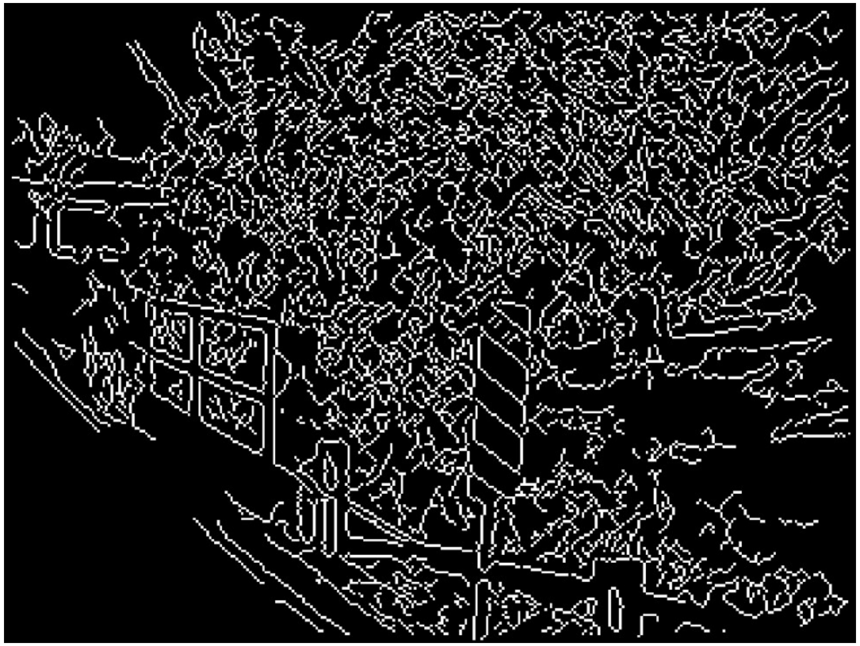
smoothed b/w image



result

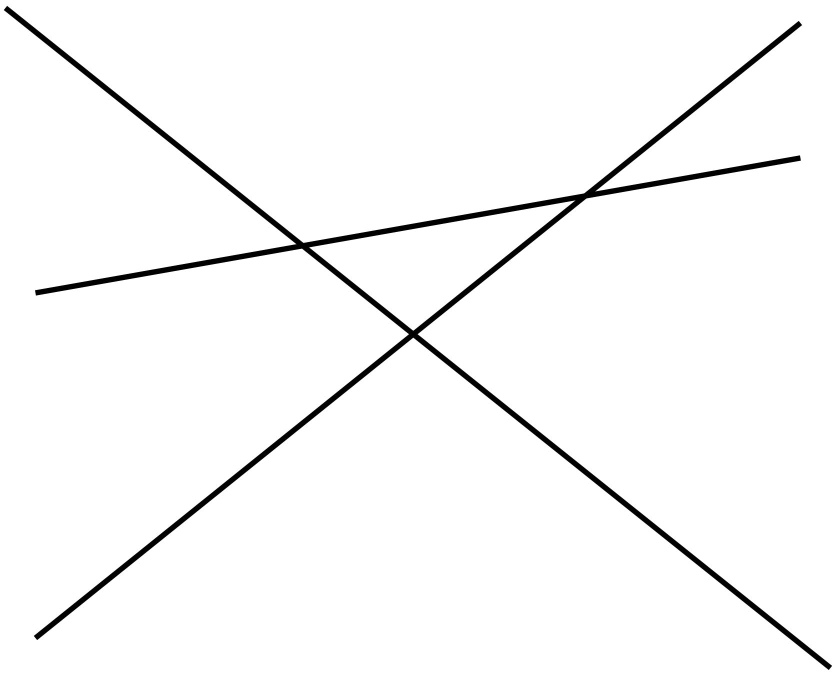
Canny Edge Detector: Example 2

original image

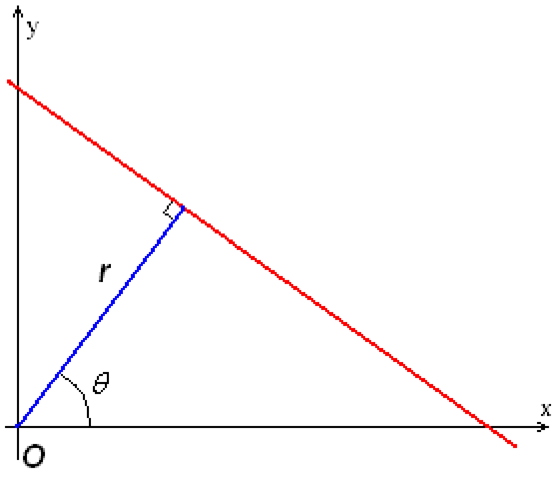


result

Hough Transform



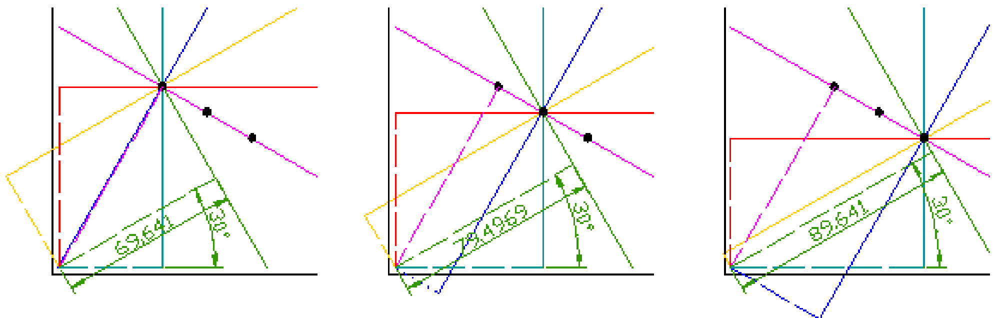
Parameterization of Line



$$y = \left(-\frac{\cos\theta}{\sin\theta}\right)x + \left(\frac{r}{\sin\theta}\right)$$

$$r(\theta) = x_0 \cdot \cos\theta + y_0 \cdot \sin\theta$$

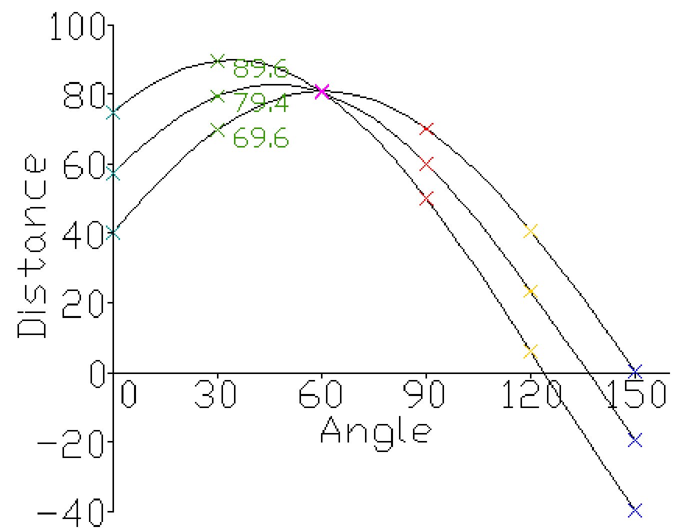
Counting Lines in Parameter Space



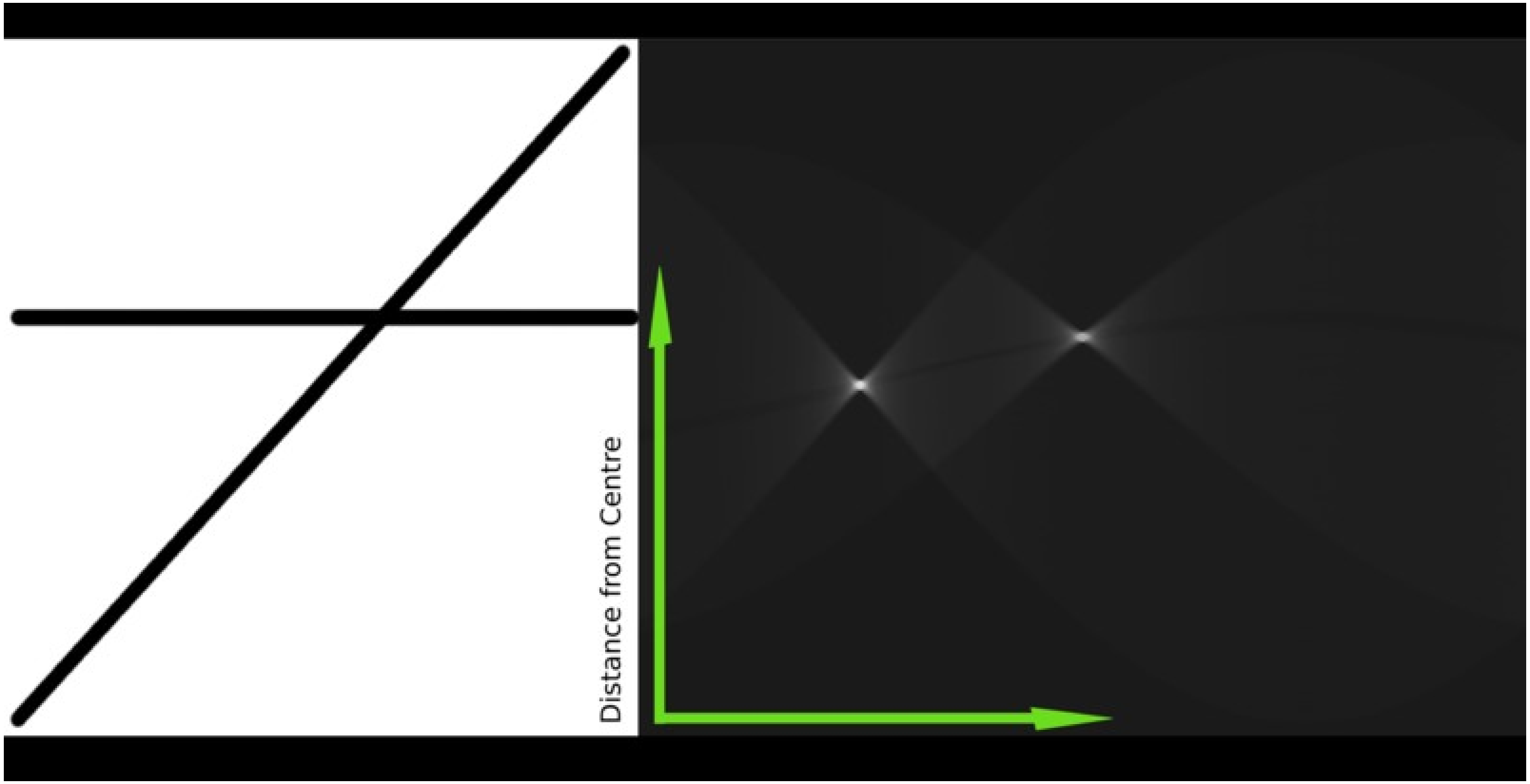
Angle	Dist.
0	40
30	69.6
60	81.2
90	70
120	40.6
150	0.4

Angle	Dist.
0	57.1
30	79.5
60	80.5
90	60
120	23.4
150	−19.5

Angle	Dist.
0	74.6
30	89.6
60	80.6
90	50
120	6.0
150	−39.6



Maxima in Parameter Space



Generalized Hough Transform