Computer Vision I

Image Formation - 24.04.2013

TECHNISCHE UNIVERSITÄT DARMSTADT

Credits: Slides adapted from Konrad Schindler
including many slides from:

Svetlana Lazebnik
Steve Seitz
Fredo Durand
Alyosha Efros

visual inference

Announcements

TECHNISCHE UNIVERSITÄT DARMSTADT

- Remember: No class next week!
  - May 1st, public holiday
- First homework assignment
  - out next week
  - please make sure to be subscribed to the mailing list
- If you still need a partner to do the homework assignments
  - meet at the front of the classroom after the lecture

Computer Vision

TECHNISCHE UNIVERSITÄT DARMSTADT

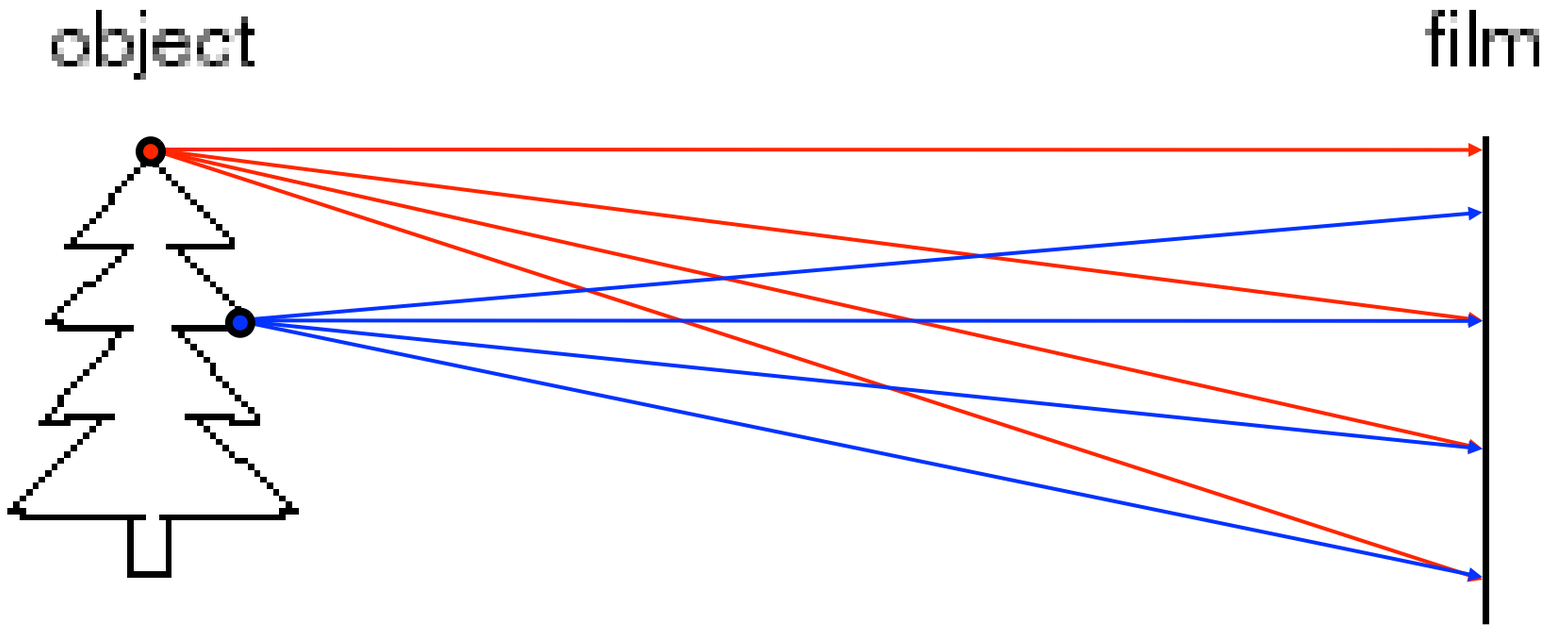
- We want to devise computer algorithms to understand the visual world much like we do.
- This means:
  - We have to use a lot of different cues...
  - We have to deal with ambiguity...
  - We have to exploit what is a plausible and meaningful interpretation...
  - ... in order to extract information about the 3D visual world from a small amount of data.
- CV is an **inverse problem**.

Overview

TECHNISCHE UNIVERSITÄT DARMSTADT

- The **pinhole projection** model
  - Properties of pinhole cameras
  - Algebraic model: Perspective projection matrix
- **Single view geometry**
  - Projection from world to image coordinates
  - Camera calibration
- Cameras with **lenses**
  - Depth of focus, field of view, lens aberrations
- **Digital cameras**
  - Types of sensors, color

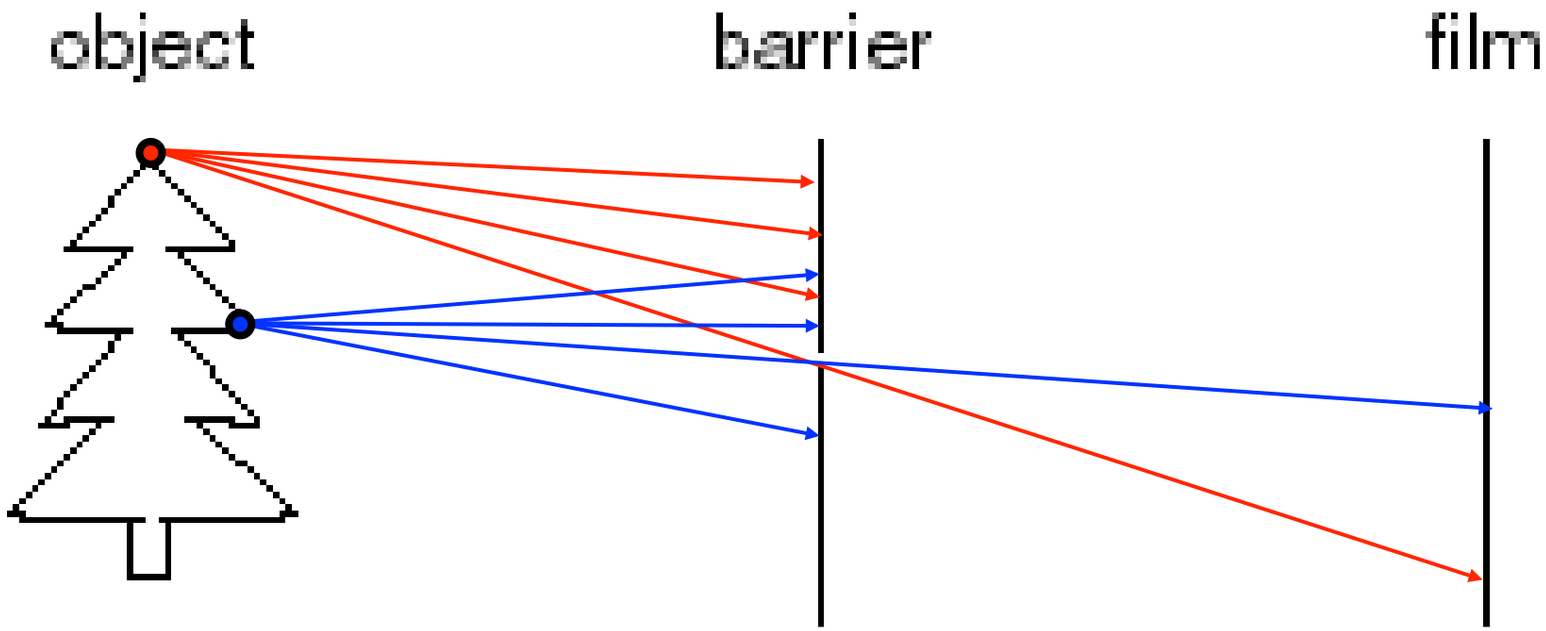
How do we see the world?



- Let’s design a camera:
  - Idea 1: Put a piece of film (or a CCD) in front of an object
  - Do we get a reasonable image?

Slide by Steve Seitz

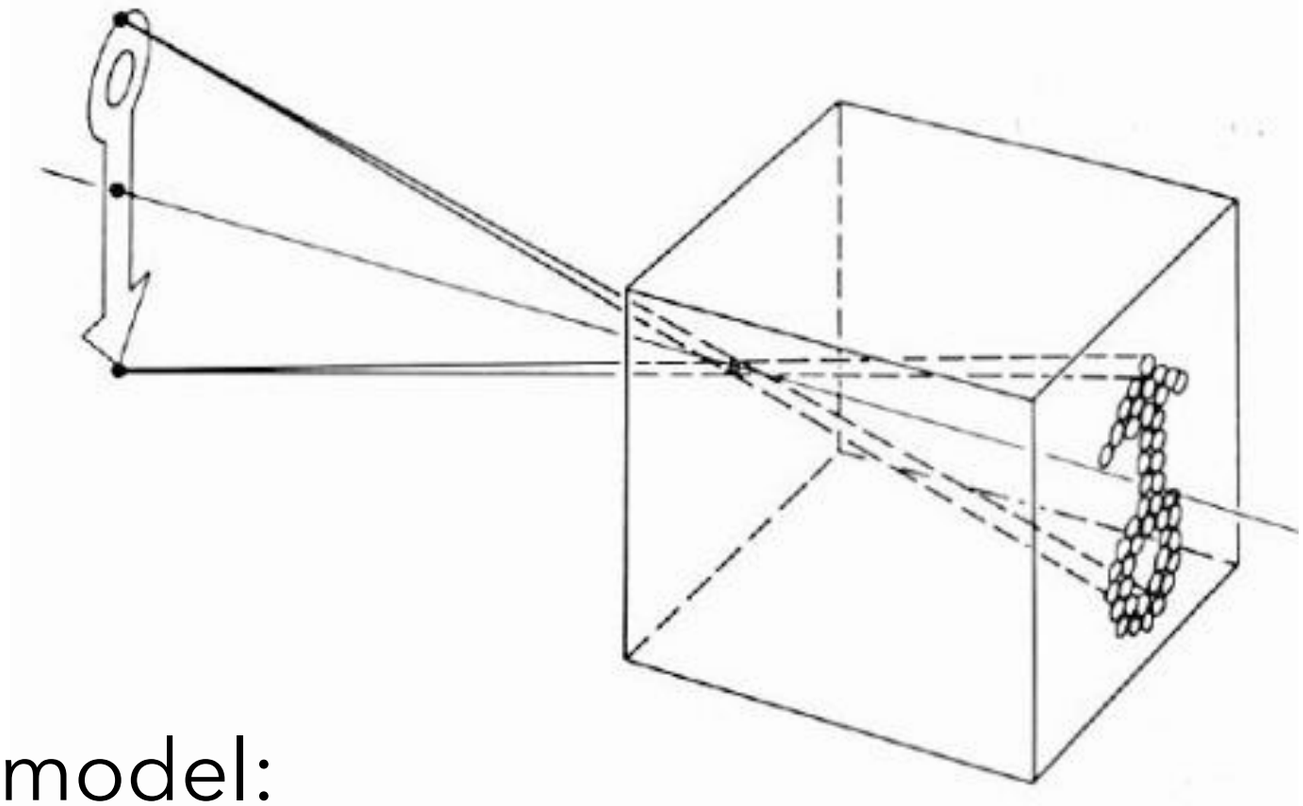
Pinhole camera



- Add a barrier to block off most of the rays:
  - This reduces blurring
  - The opening is known as the **aperture**

Slide by Steve Seitz

Pinhole camera model

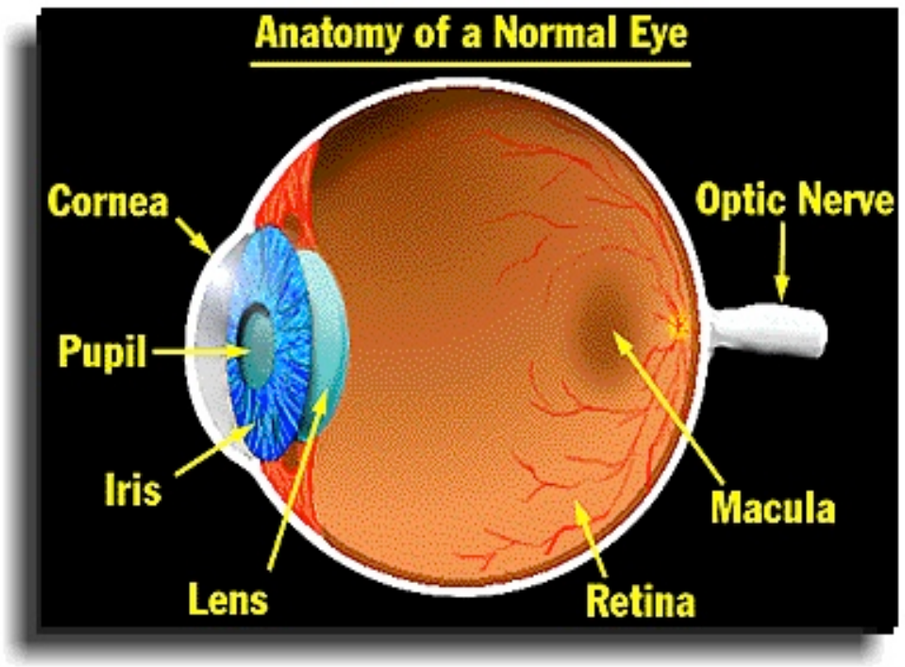


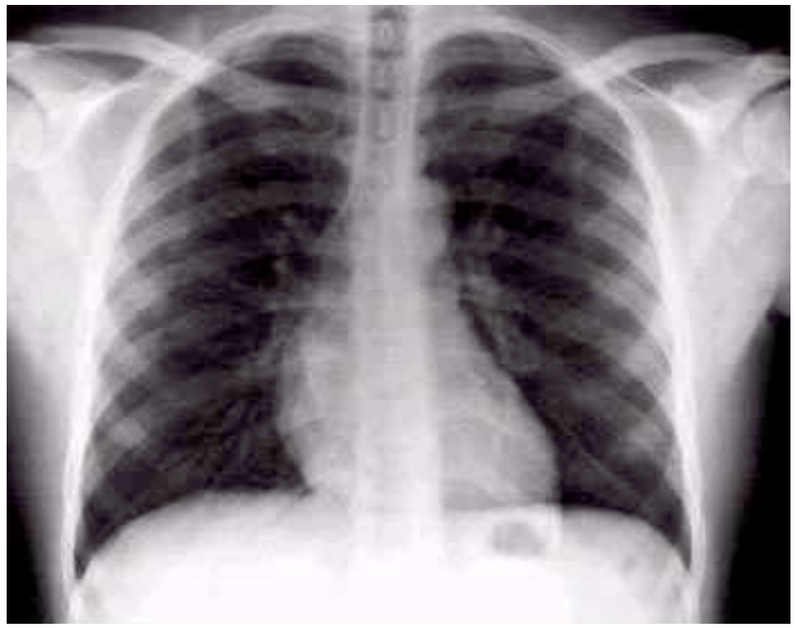
- Pinhole model:
  - Captures **pencil of rays** – all rays through a single point
  - Projection rays are **straight** lines
  - The point is called **center of projection (focal point)**
  - The image is formed on the **image plane**
  - Two equivalent projections: **positive, negative**

Pinhole camera

TECHNISCHE UNIVERSITÄT DARMSTADT

- A model for many common sensors:
  - Photographic cameras
  - human eye
  - X-ray machines

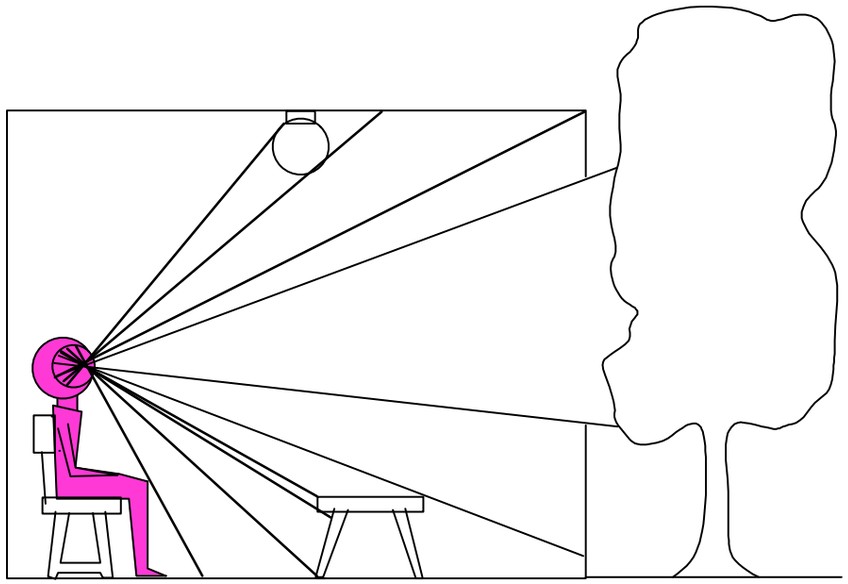




Dimensionality Reduction Machine (3D to 2D)

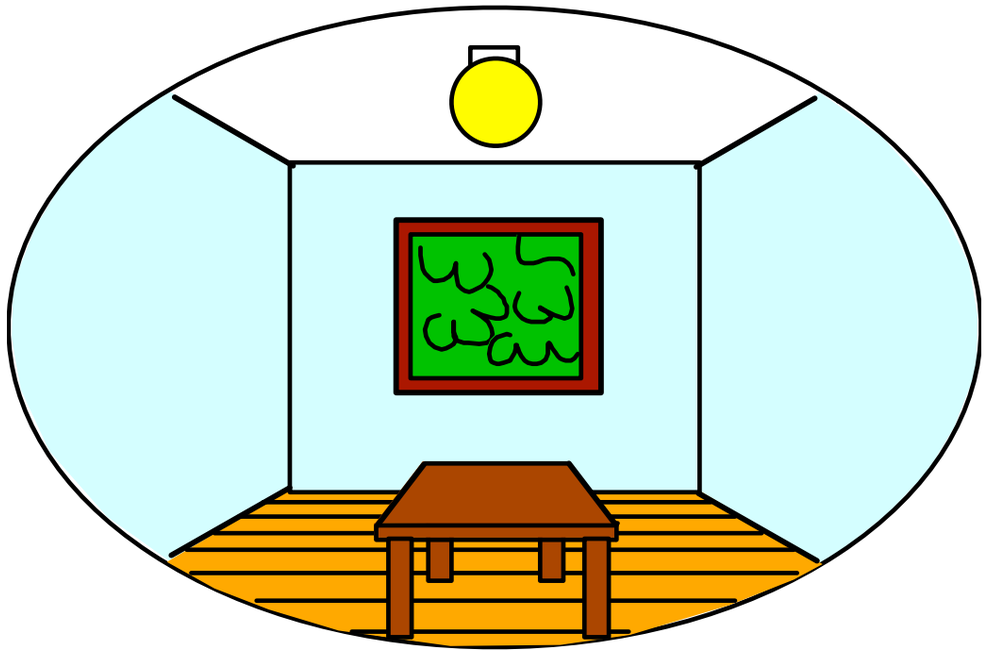
TECHNISCHE UNIVERSITÄT DARMSTADT

3D world



Point of observation

2D image



- What have we lost?
  - Angles
  - Distances (lengths)

Slide by A. Efros

Figures © Stephen E. Palmer, 2002

Projection properties

TECHNISCHE UNIVERSITÄT DARMSTADT

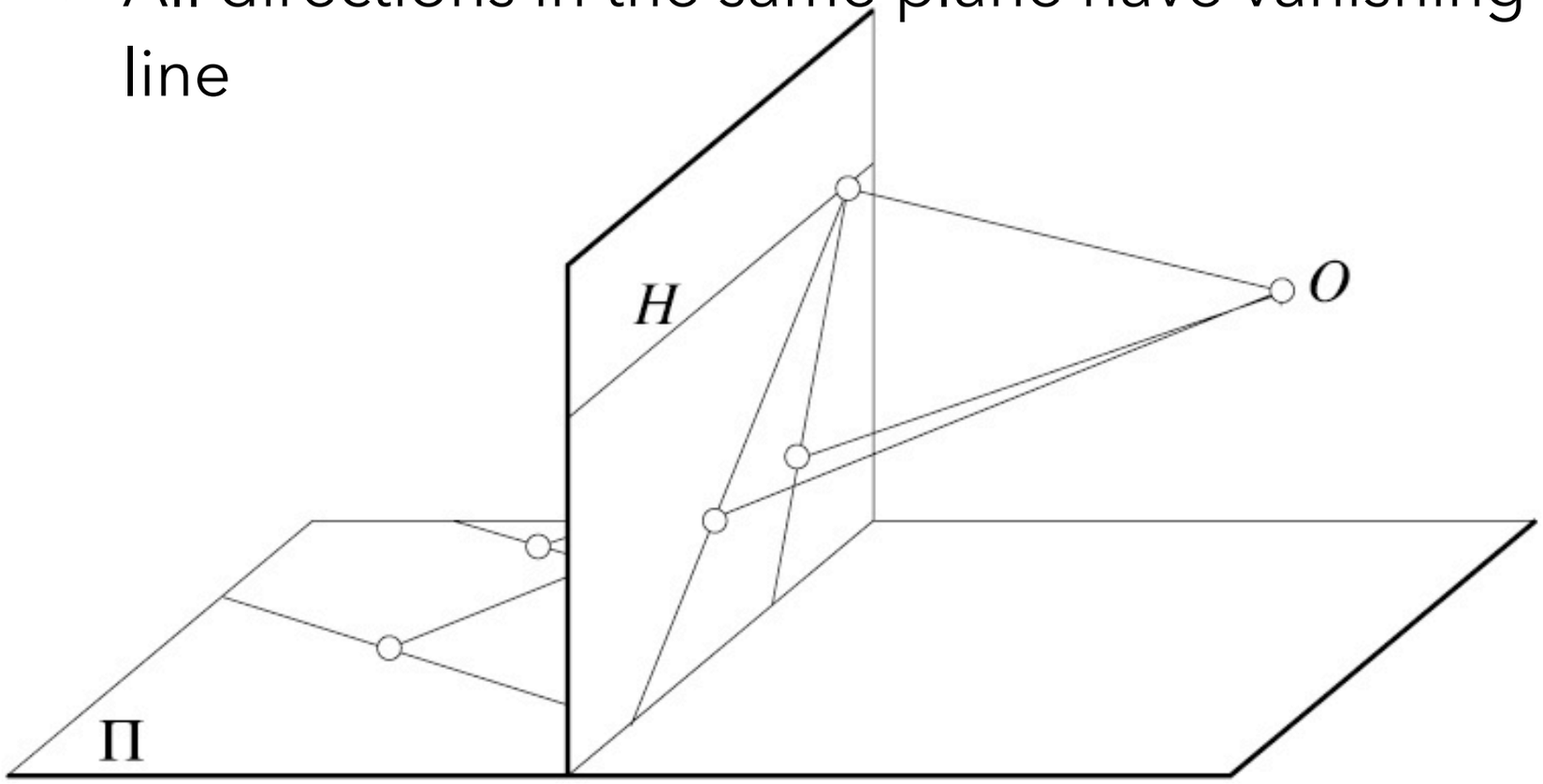
- **Many-to-one:**
  - any points along the same ray map to the same point in image
- Points → points
  - But projection of points on focal plane is undefined
- Lines → lines (collinearity is preserved)
  - But line through focal point projects to a point
- Planes → planes (or half-planes)
  - But plane through focal point projects to line

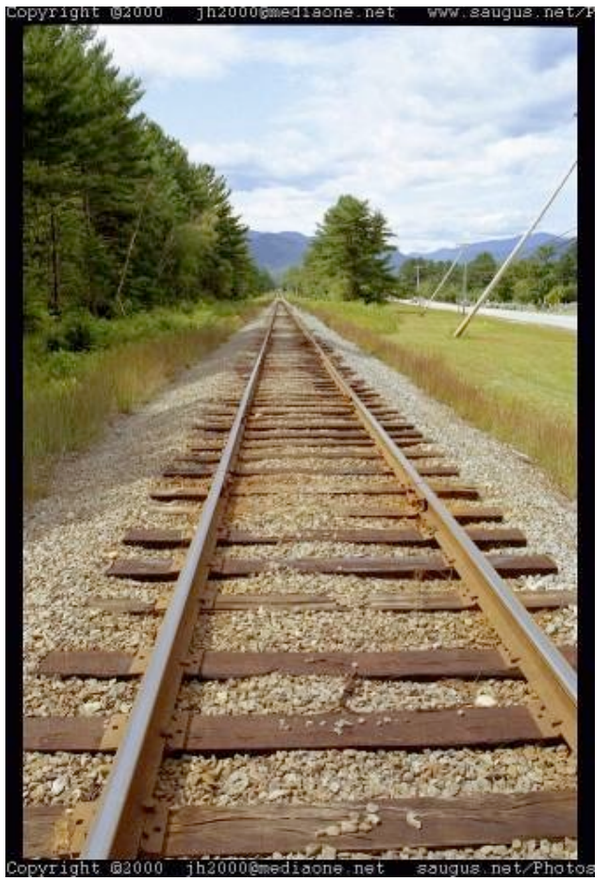
Slide by Lana Lazebnik

Projection properties

TECHNISCHE UNIVERSITÄT DARMSTADT

- Parallel lines converge at a **vanishing point**
  - Each direction in space has its own vanishing point
  - But parallels also parallel to the image plane remain parallel
  - All directions in the same plane have vanishing points on the same line



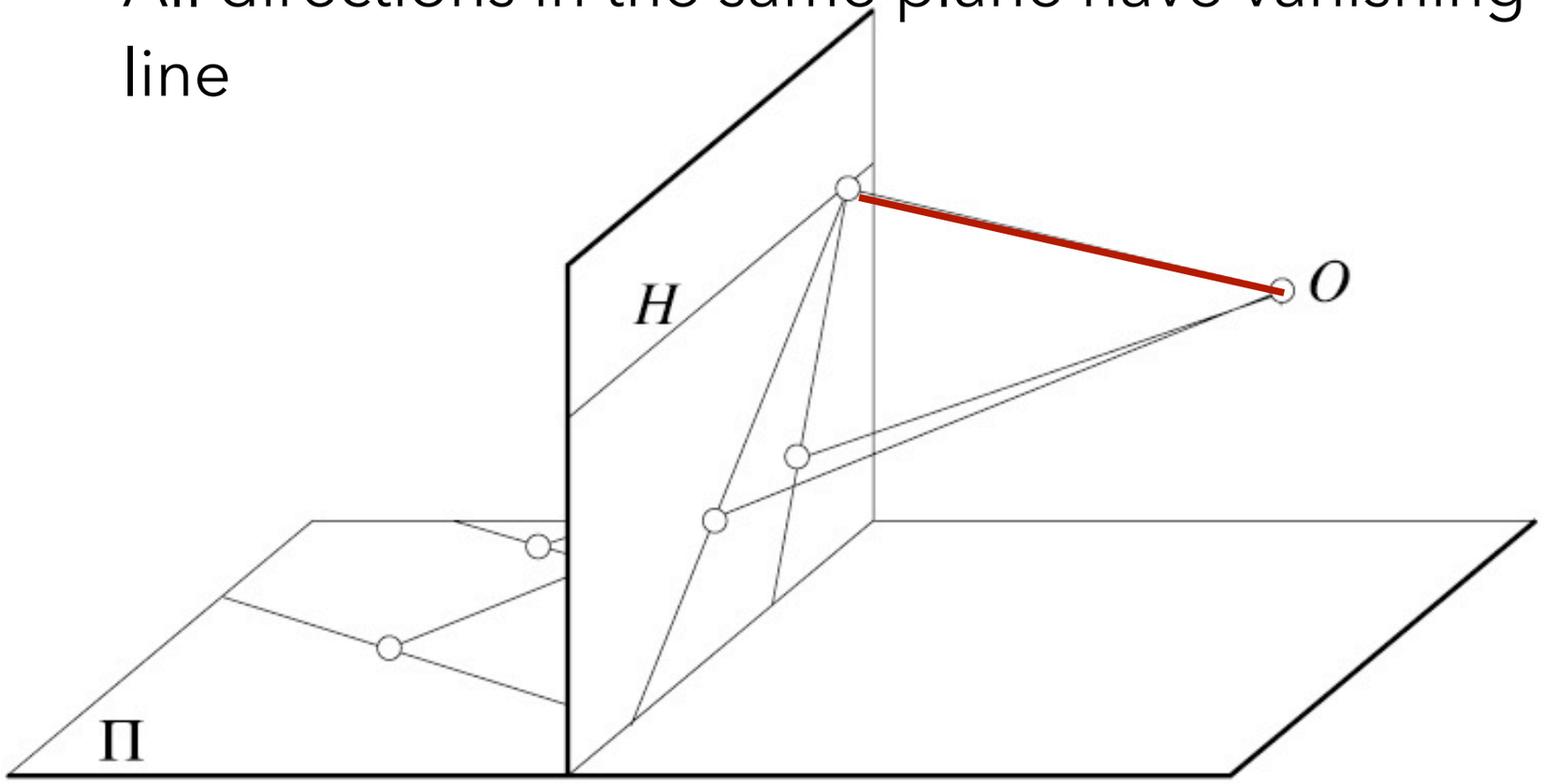


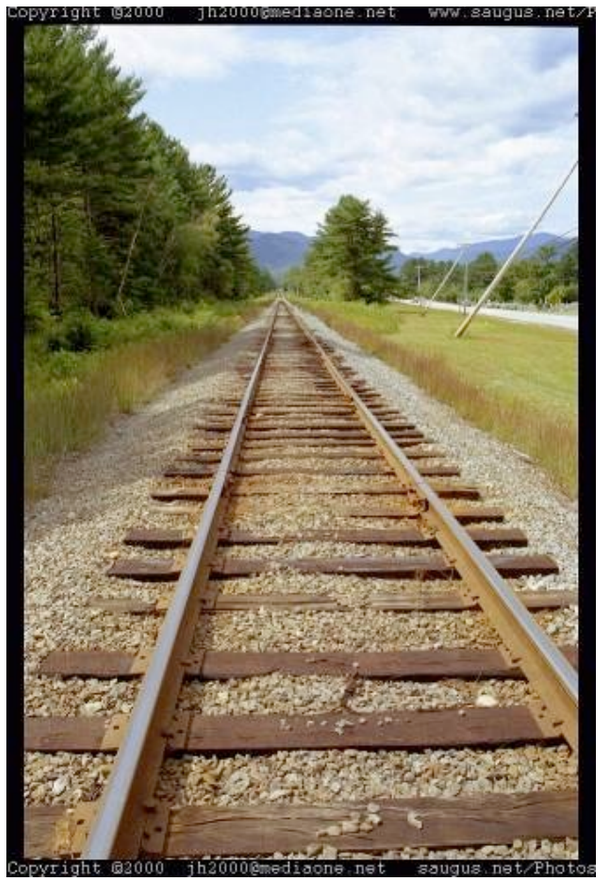
How do we construct the vanishing point/line?

Projection properties

TECHNISCHE UNIVERSITÄT DARMSTADT

- Parallel lines converge at a **vanishing point**
  - Each direction in space has its own vanishing point
  - But parallels also parallel to the image plane remain parallel
  - All directions in the same plane have vanishing points on the same line

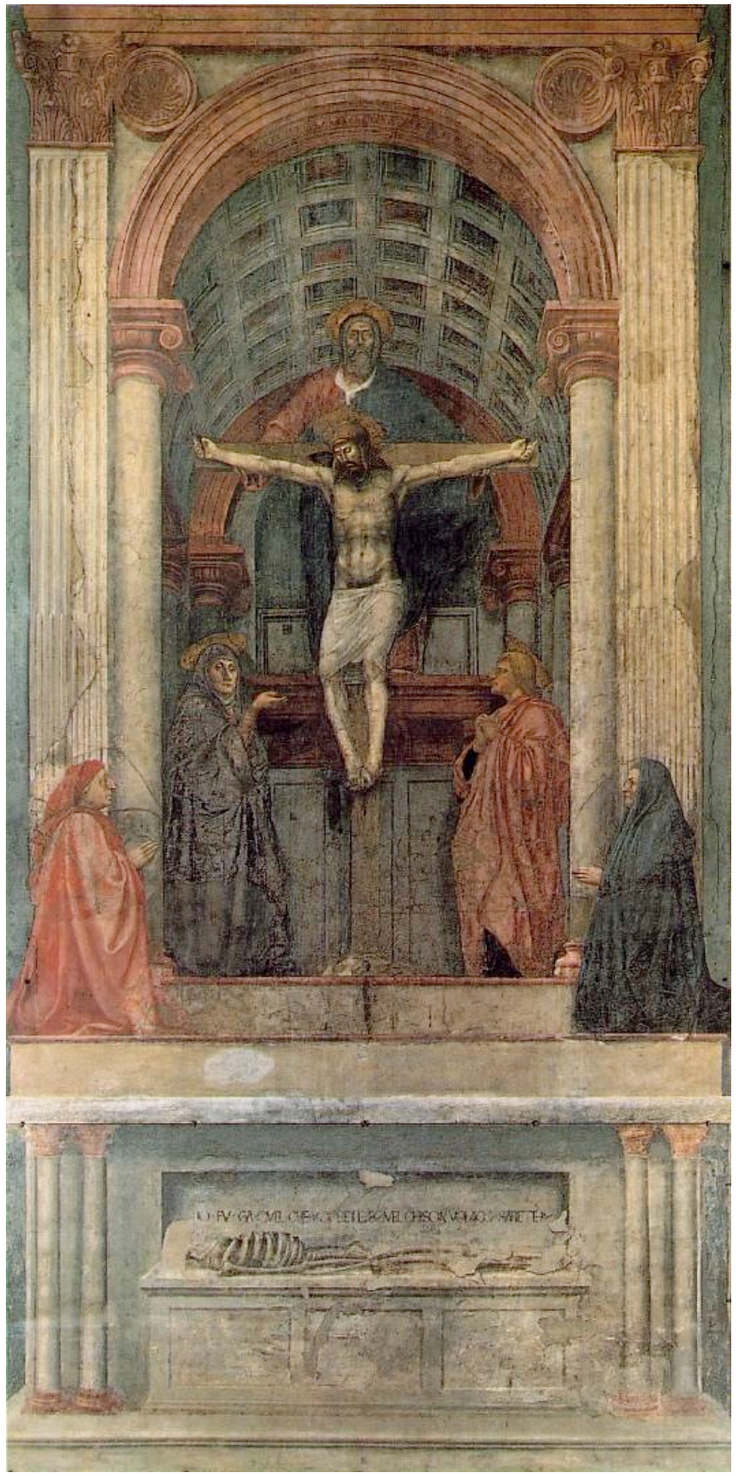


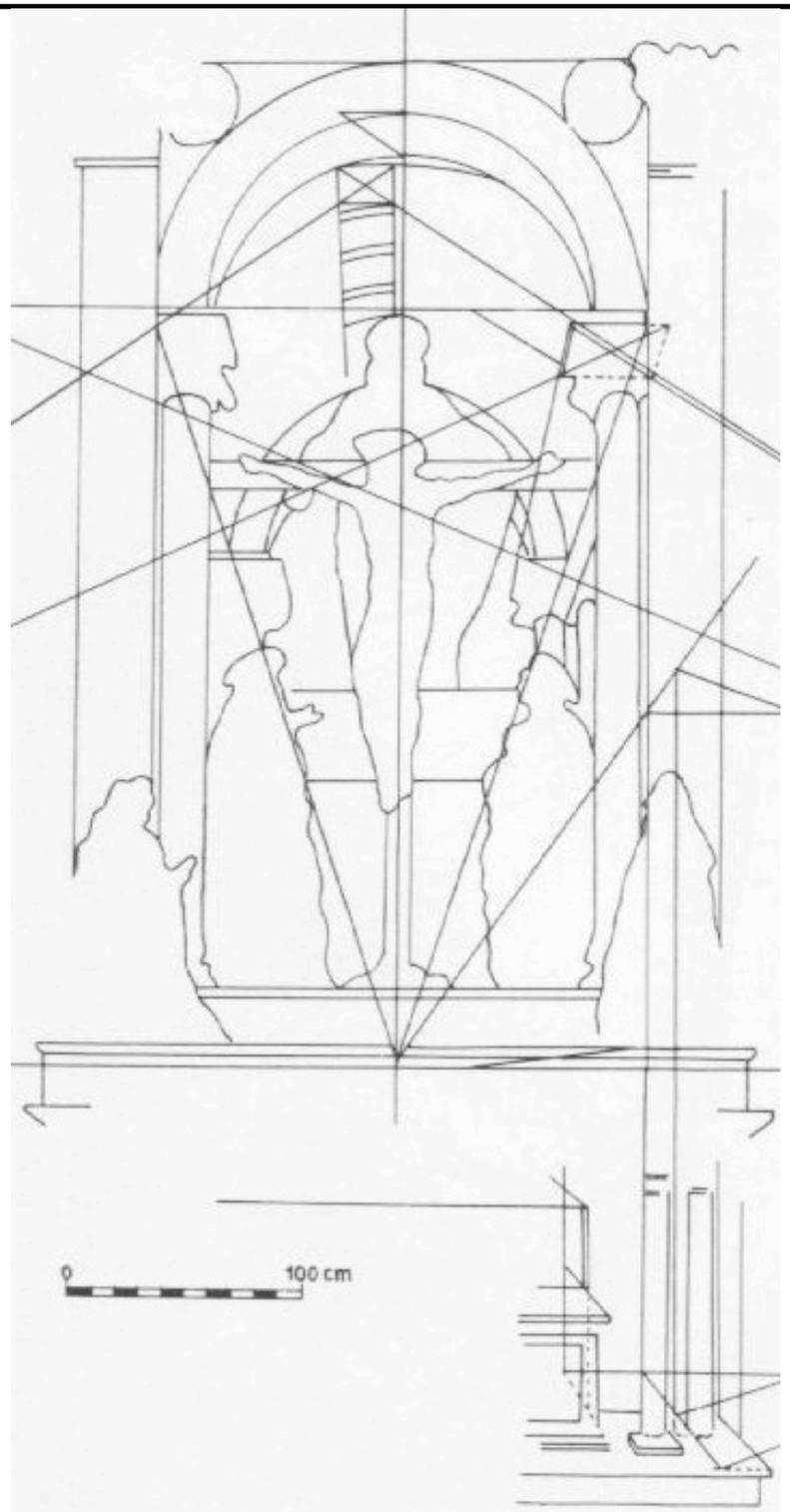


How do we construct the vanishing point/line?

One-point perspective

TECHNISCHE UNIVERSITÄT DARMSTADT





- Masaccio, Trinity, Santa Maria Novella, Florence, 1425-28

- First consistent use of perspective in Western art?

Perspective distortion

TECHNISCHE UNIVERSITÄT DARMSTADT

- What does a **sphere** project to?

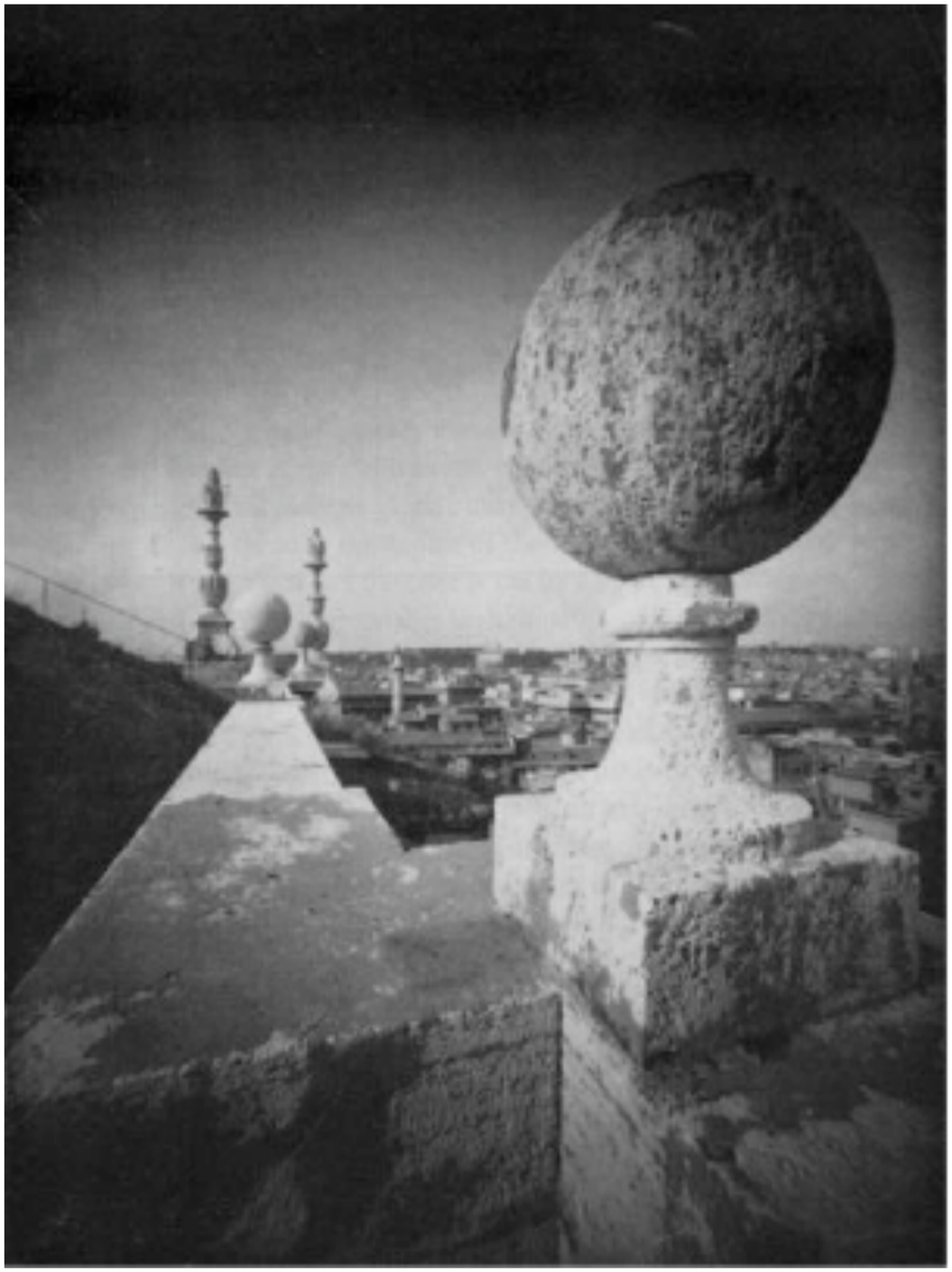
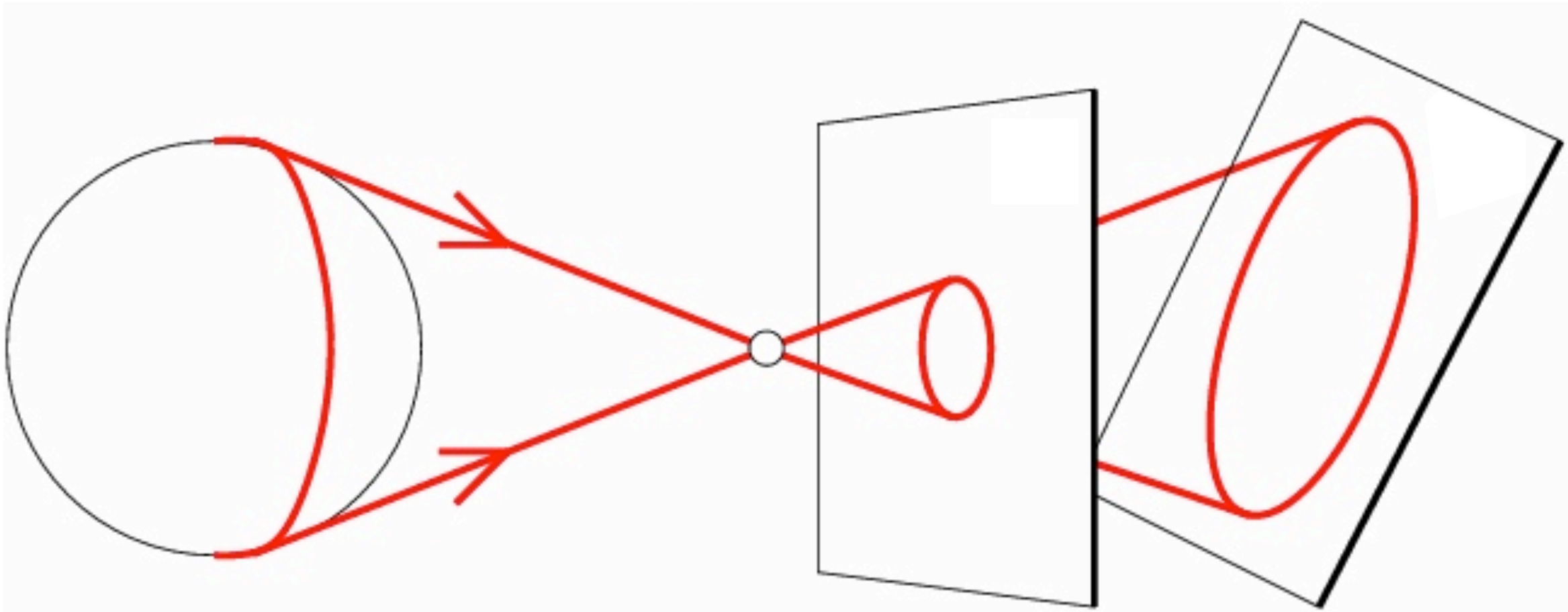


Image source: F. Durand

Perspective distortion

TECHNISCHE UNIVERSITÄT DARMSTADT

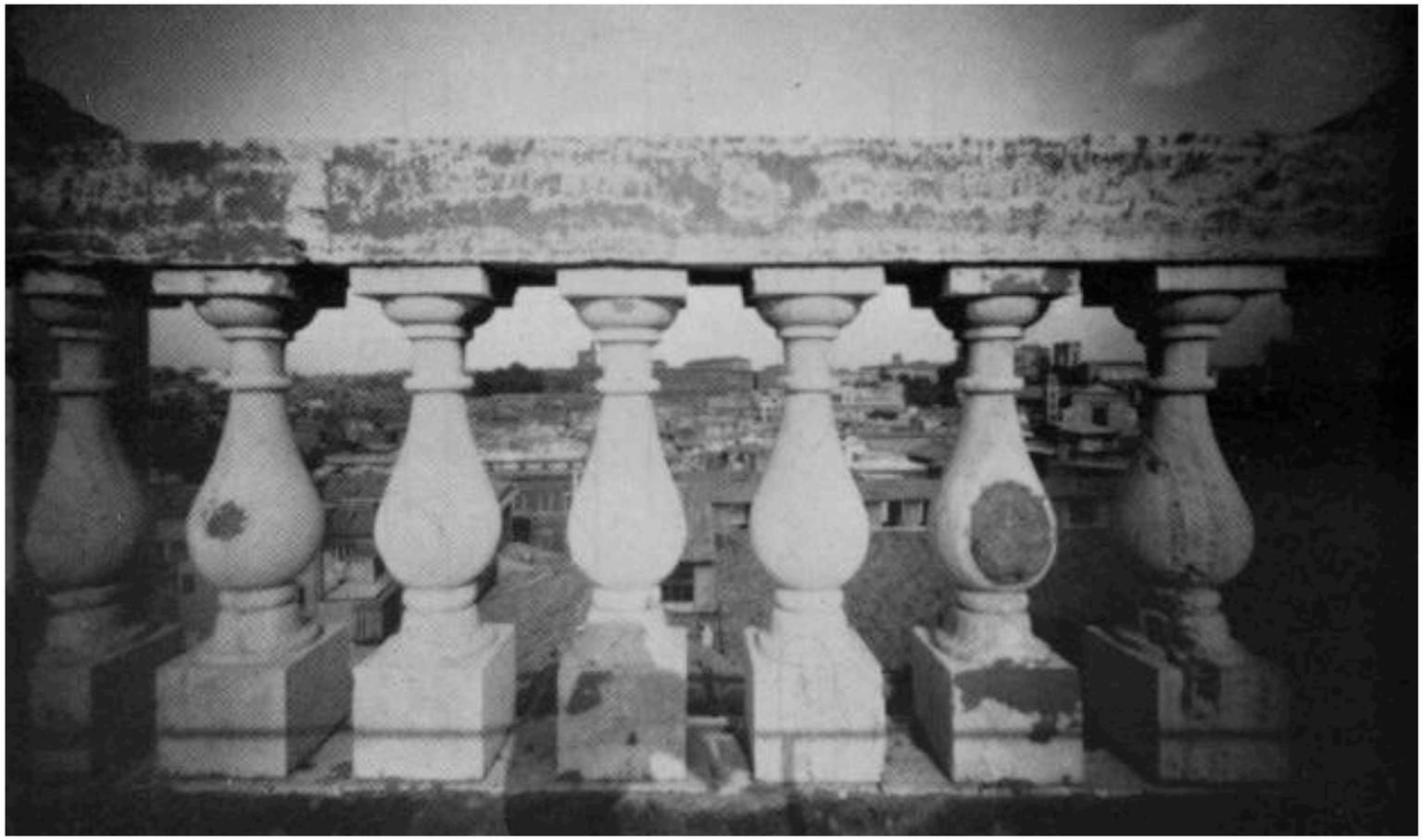
- What does a **sphere** project to?

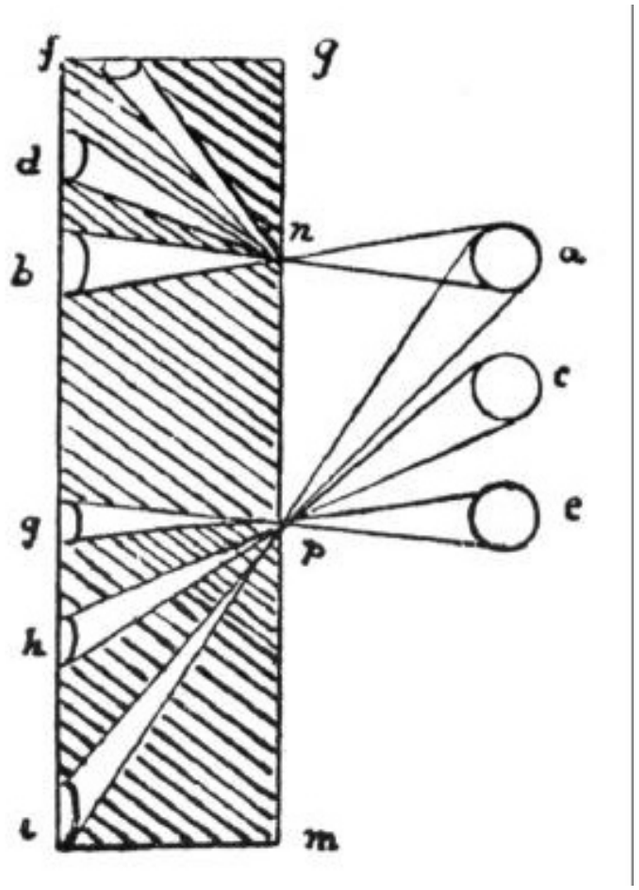


Perspective distortion

TECHNISCHE UNIVERSITÄT DARMSTADT

- The exterior columns appear bigger
- The distortion is **not** due to lens flaws
- Problem pointed out by Da Vinci



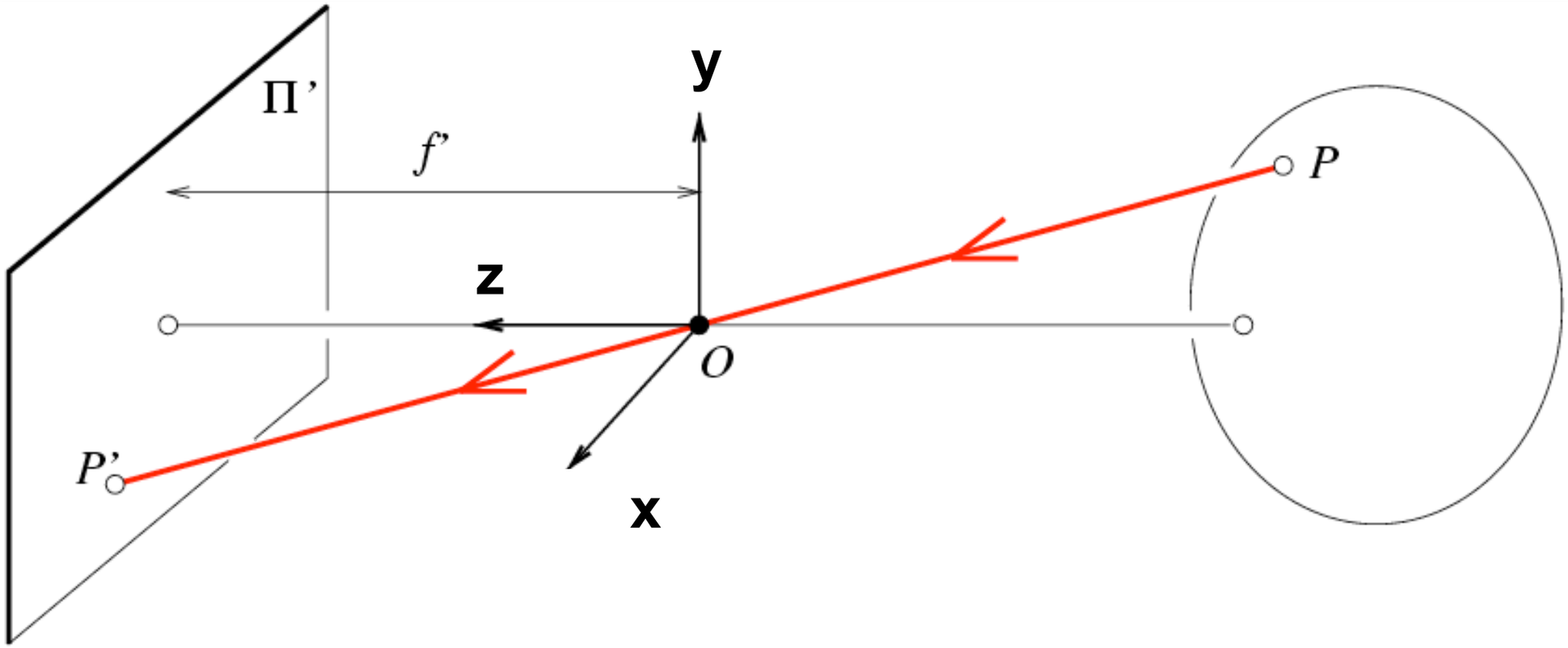


Slide by F. Durand

Perspective distortion: People

TECHNISCHE UNIVERSITÄT DARMSTADT

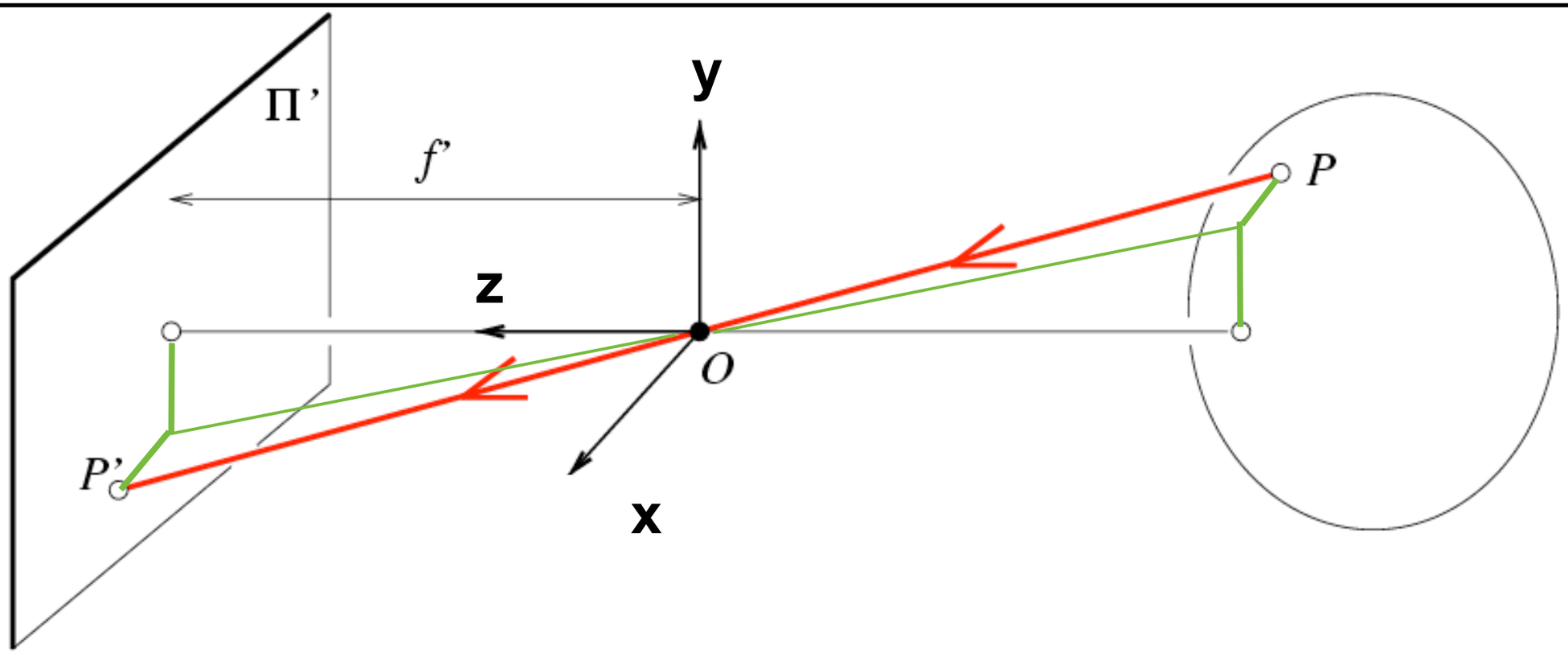
Modeling projection



- The coordinate system
  - We will use the pinhole model as an approximation
  - Put the optical center (O) at the origin
  - Put the image plane (Π') in front of O

Source: J. Ponce, S. Seitz

Modeling projection



- Projection equations
  - Compute intersection with Π' of ray from $P = (x,y,z)$ to O
  - Derived using similar triangles $(x, y, z) \rightarrow (f' \frac{x}{z}, f' \frac{y}{z}, f')$
  - To get the projection, throw away last coordinate $(x, y, z) \rightarrow (f' \frac{x}{z}, f' \frac{y}{z})$

Source: J. Ponce, S. Seitz

Homogeneous coordinates

TECHNISCHE UNIVERSITÄT DARMSTADT

$$(x, y, z) \rightarrow (f' \frac{x}{z}, f' \frac{y}{z})$$

- Is this a linear transformation?

Slide by Steve Seitz

Homogeneous coordinates

$$(x, y, z) \rightarrow (f' \frac{x}{z}, f' \frac{y}{z})$$

- Is this a linear transformation?
  - no - division by z is nonlinear
- Trick: add one more coordinate

$$(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

homogeneous image coordinates

$$(x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

homogeneous scene coordinates

- Converting from homogeneous coordinates

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w) \qquad \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)$$

Slide by Steve Seitz

Perspective Projection Matrix

- Projection is a matrix multiplication in homogeneous coordinates:

$$\begin{pmatrix} f' & 0 & 0 & 0 \\ 0 & f' & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} f'x \\ f'y \\ z \end{pmatrix} \quad \Rightarrow \quad (f'\frac{x}{z}, f'\frac{y}{z})$$

divide by the third coordinate

Slide by Lana Lazebnik

Perspective Projection Matrix

TECHNISCHE UNIVERSITÄT DARMSTADT

- Projection is a matrix multiplication in homogeneous coordinates:

$$\begin{pmatrix} f' & 0 & 0 & 0 \\ 0 & f' & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} f'x \\ f'y \\ z \end{pmatrix} \quad \Rightarrow \quad (f'\frac{x}{z}, f'\frac{y}{z})$$

divide by the third coordinate

- By changing the algebraic representation all relevant transformations become linear

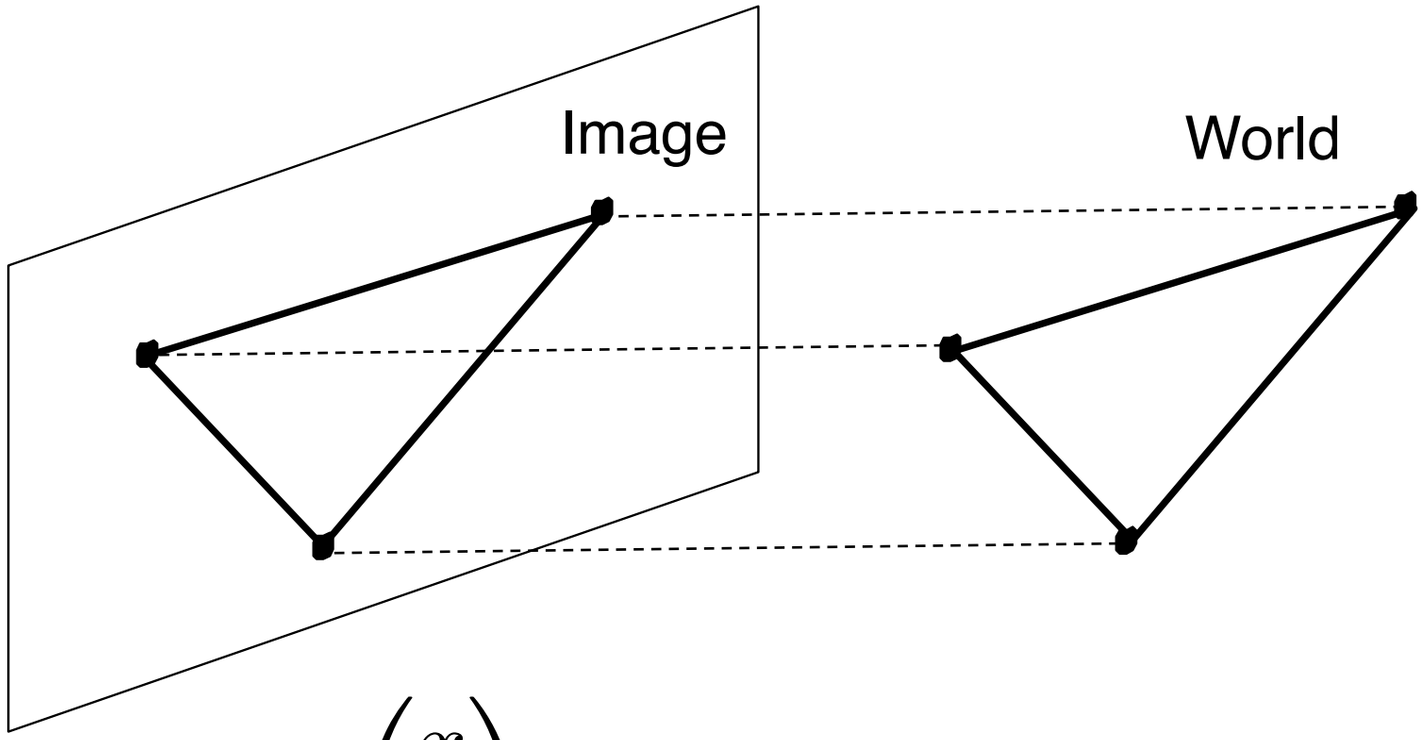
$$\begin{pmatrix} \text{2D point (3x1)} \end{pmatrix} = \begin{pmatrix} \text{Camera to pixel coord. trans. matrix (3x3)} \end{pmatrix} \begin{pmatrix} \text{Perspective projection matrix (3x4)} \end{pmatrix} \begin{pmatrix} \text{World to camera coord. trans. matrix (4x4)} \end{pmatrix} \begin{pmatrix} \text{3D point (4x1)} \end{pmatrix}$$

Slide by Lana Lazebnik

Orthographic Projection

TECHNISCHE UNIVERSITÄT DARMSTADT

- Special case of perspective projection
  - Distance from center of projection to image plane is infinite
  - Also called "parallel projection"



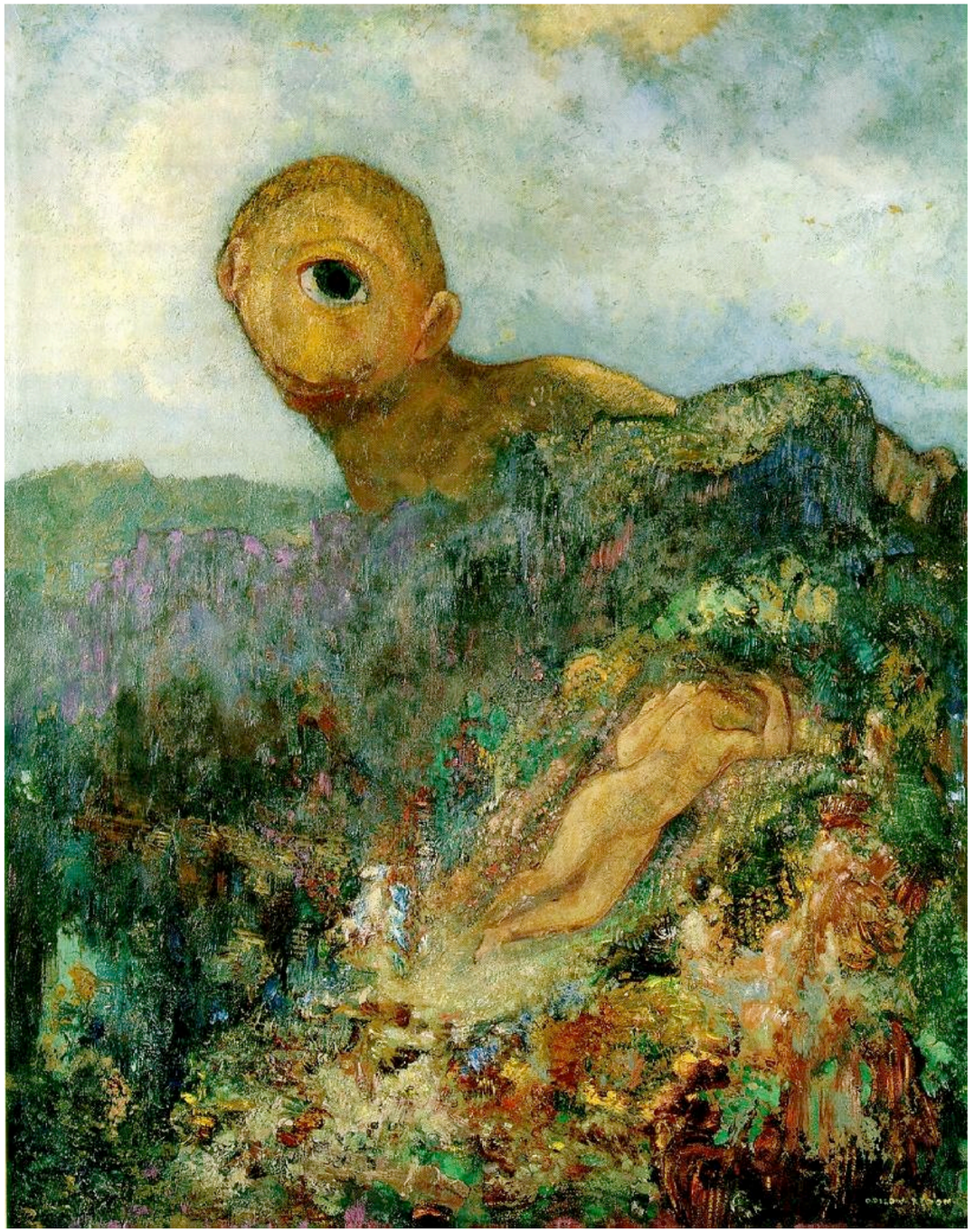
- What's the projection matrix?

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad \Rightarrow \quad (x, y)$$

Slide by Steve Seitz

Single-view geometry

TECHNISCHE UNIVERSITÄT DARMSTADT



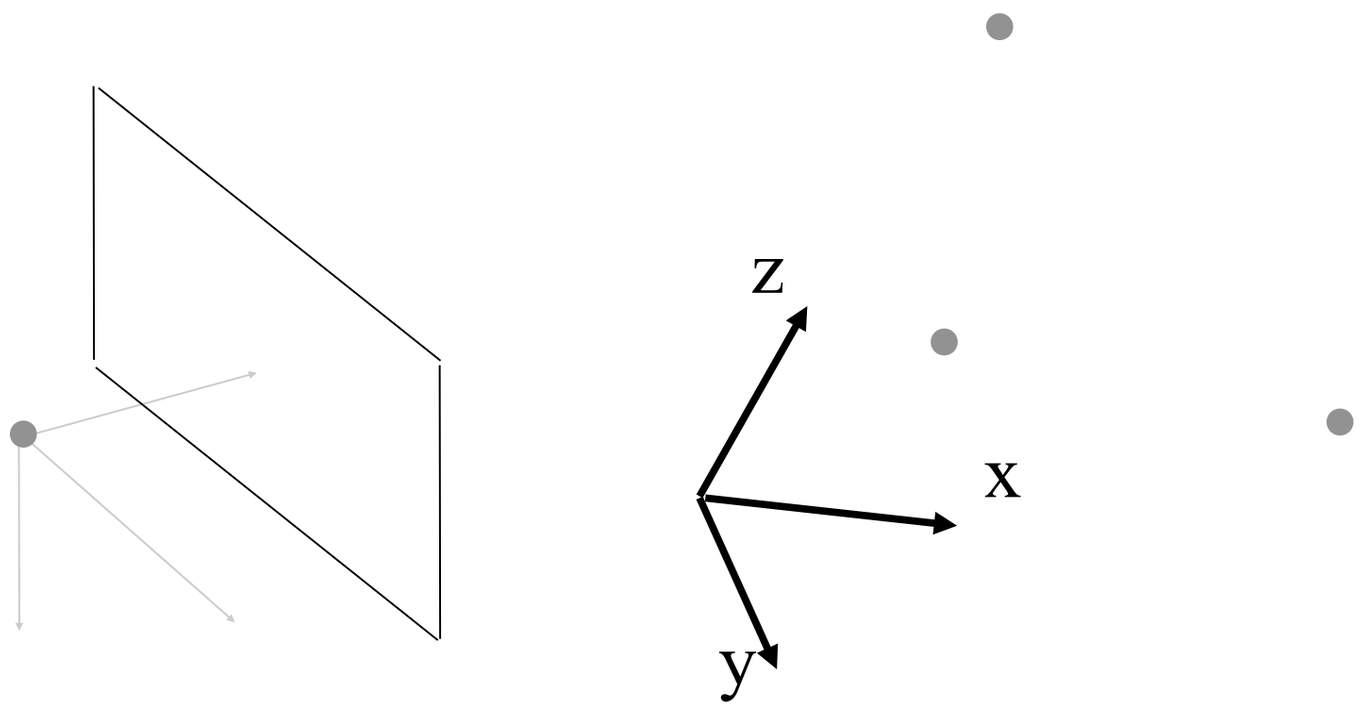
Odilon Redon, Cyclops, 1914

Slide by Lana Lazebnik

Camera calibration

TECHNISCHE UNIVERSITÄT DARMSTADT

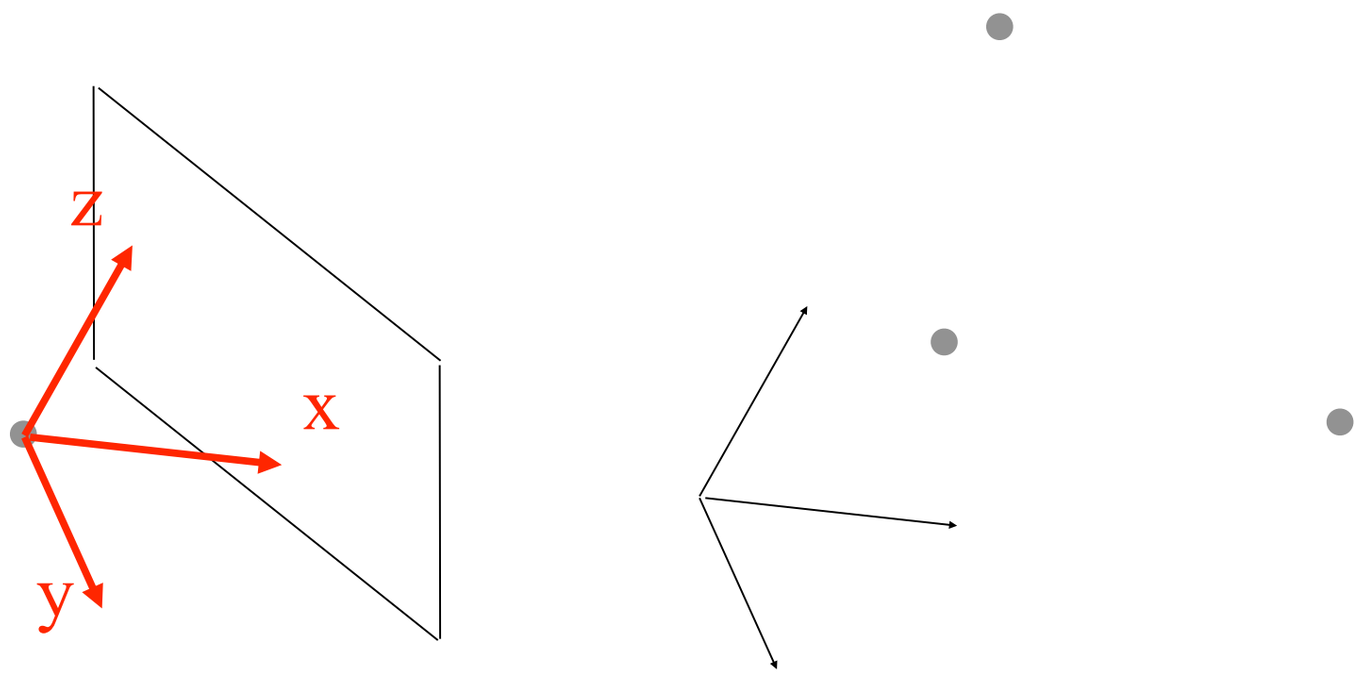
- Let's place the camera in the world
  - **Translation**
  - **Rotation**
  - **Pinhole projection**



Camera calibration

TECHNISCHE UNIVERSITÄT DARMSTADT

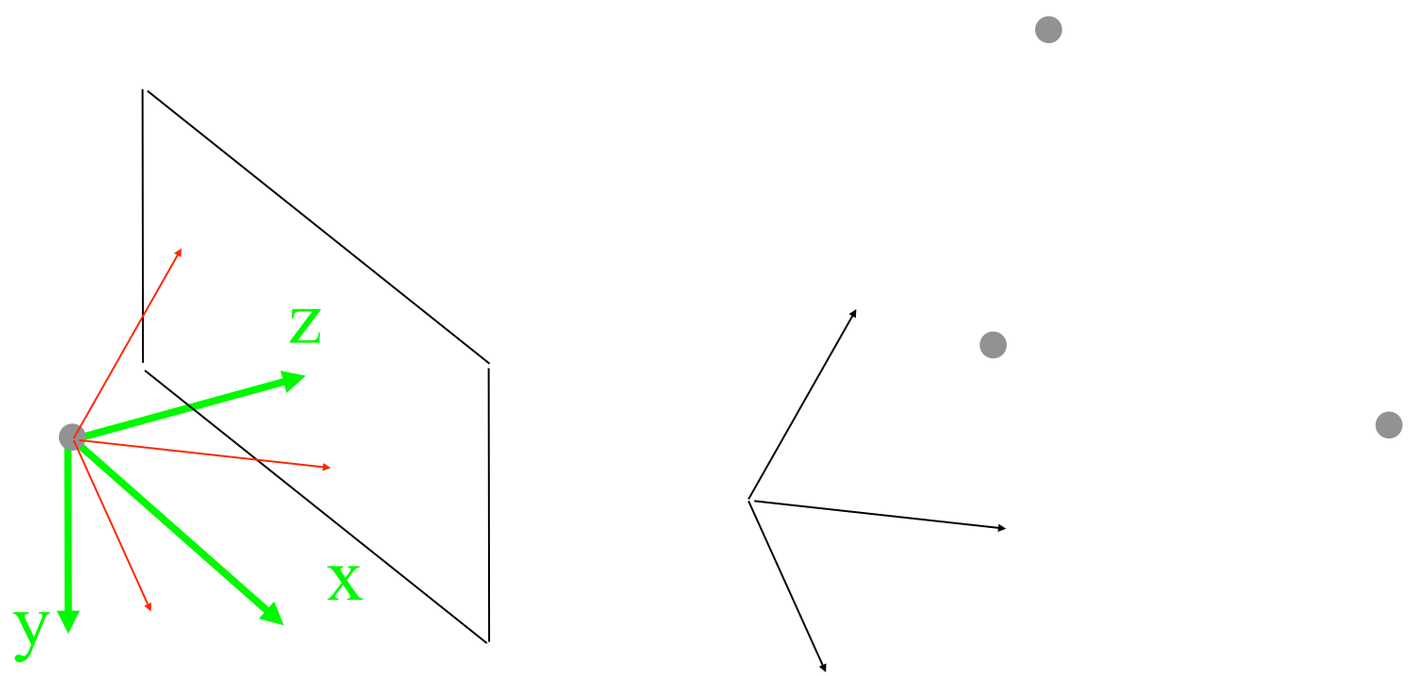
- Let's place the camera in the world
  - **Translation**
  - **Rotation**
  - **Pinhole projection**



Camera calibration

TECHNISCHE UNIVERSITÄT DARMSTADT

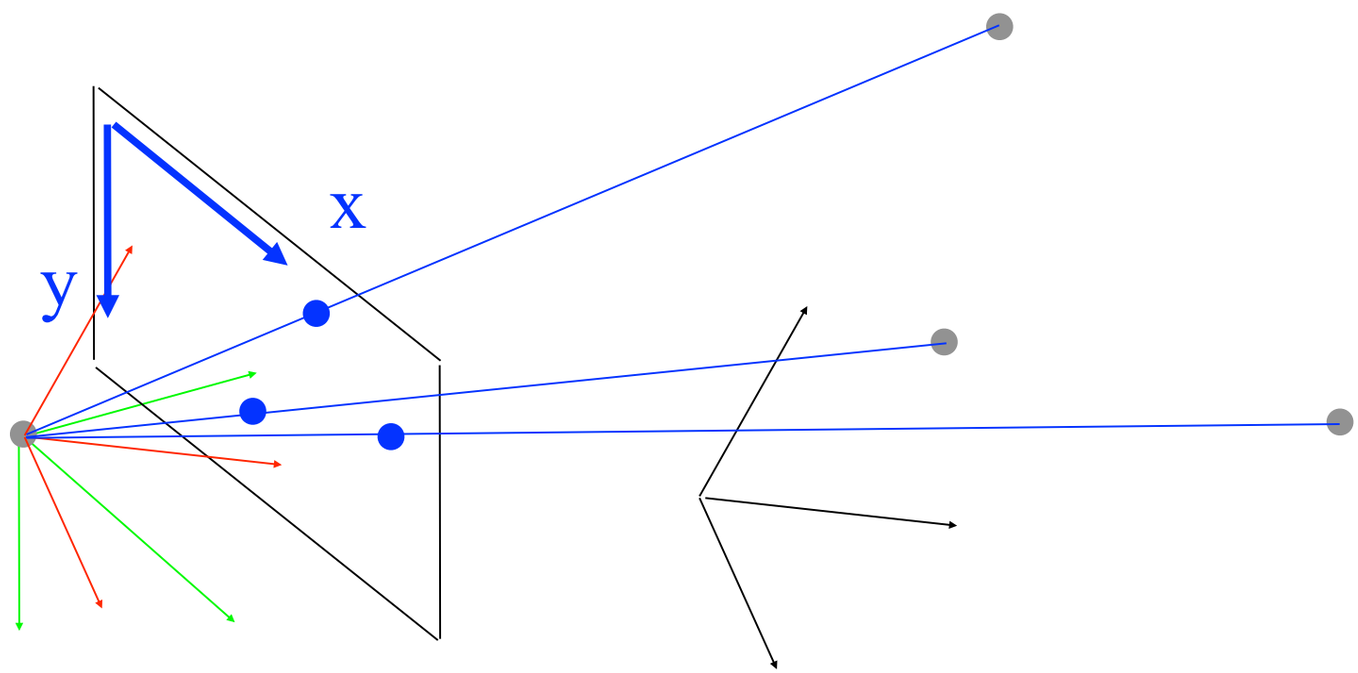
- Let's place the camera in the world
  - Translation
  - Rotation
  - Pinhole projection



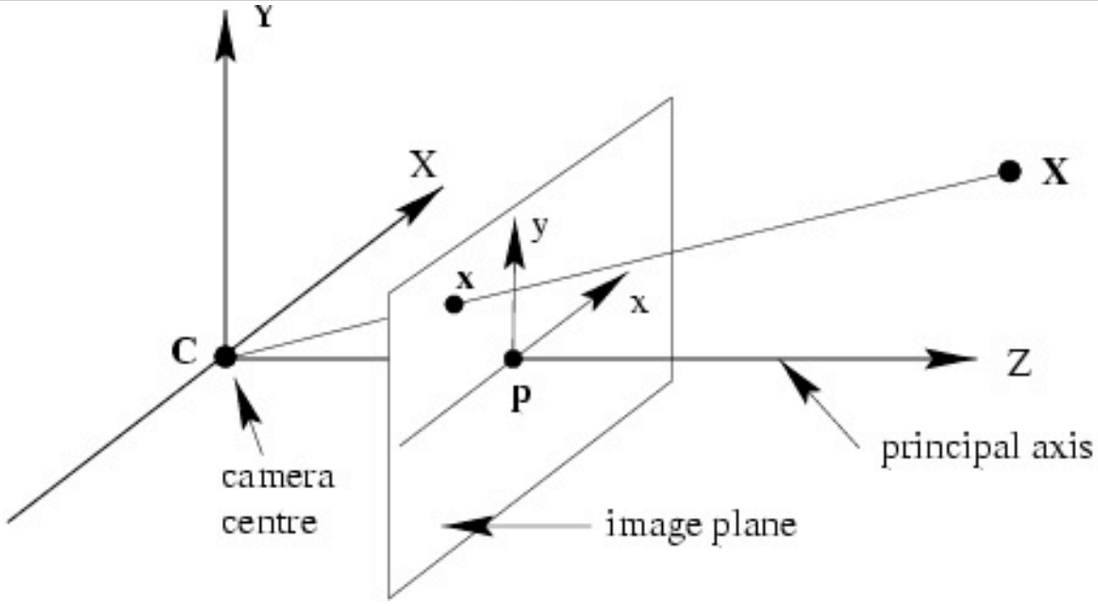
Camera calibration

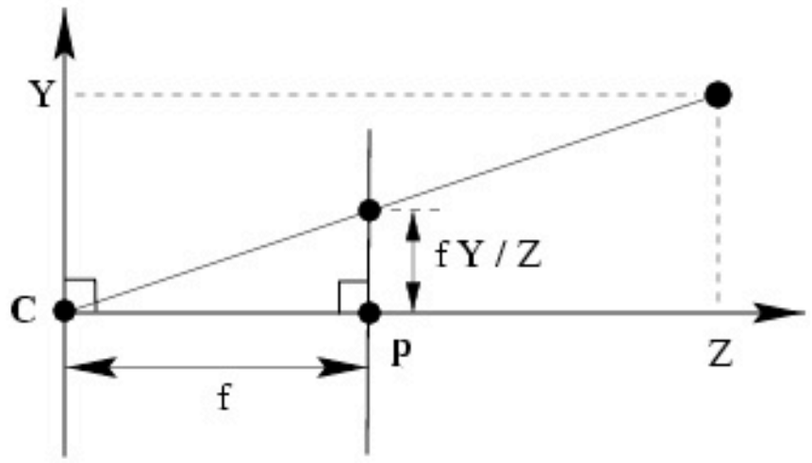
TECHNISCHE UNIVERSITÄT DARMSTADT

- Let's place the camera in the world
  - **Translation**
  - **Rotation**
  - **Pinhole projection**



Recall: Pinhole camera model



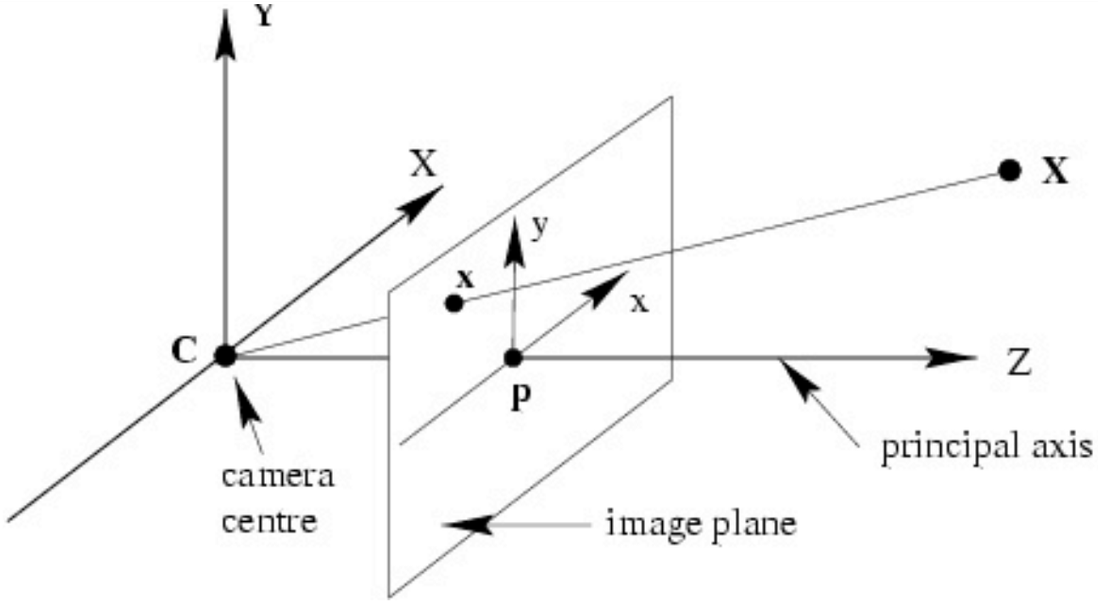


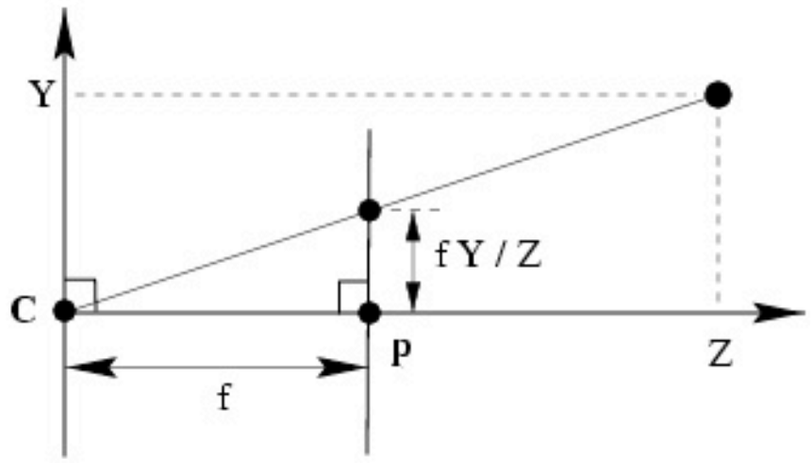
$$(X, Y, Z) \rightarrow (fX/Z, fY/Z)$$

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} fX \\ fY \\ Z \end{pmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$\mathbf{x} = P\mathbf{X}$$

Recall: Pinhole camera model



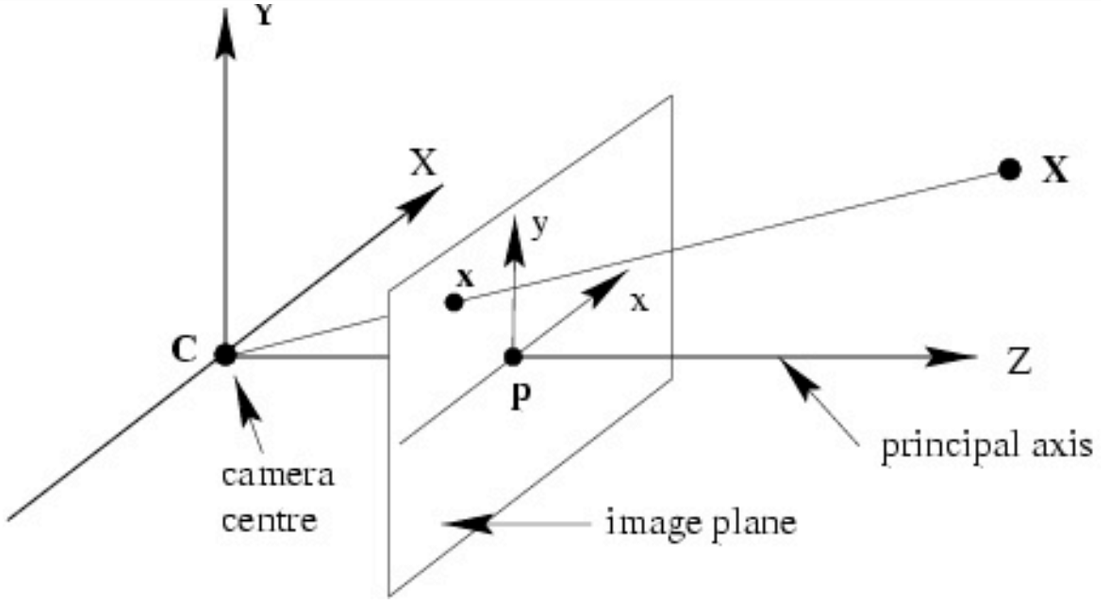


$$(X, Y, Z) \rightarrow (fX/Z, fY/Z)$$

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} fX \\ fY \\ Z \end{pmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$\mathbf{x} = P\mathbf{X} \qquad P = \mathrm{diag}(f, f, 1)[I|0]$$

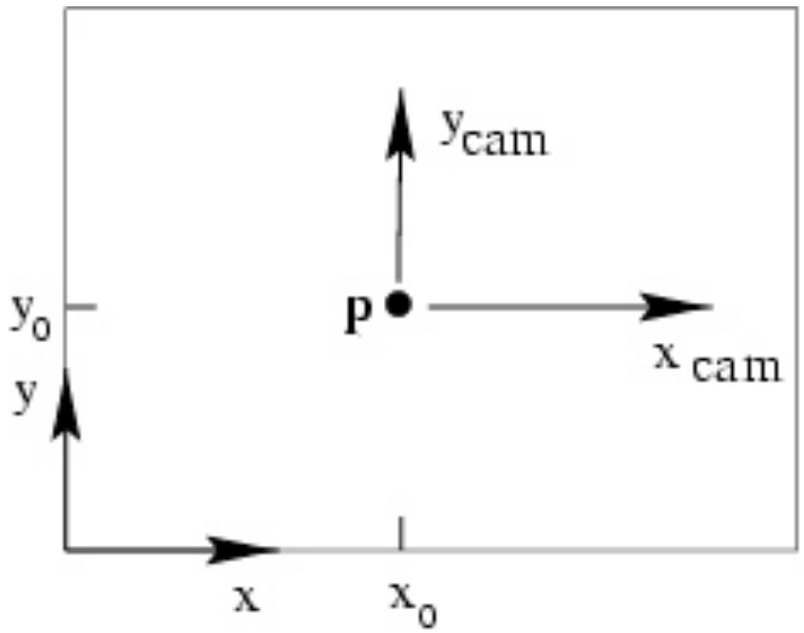
Camera coordinate system



- **Principal axis:** line from the camera center perpendicular to the image plane
- **Normalized (camera) coordinate system:** camera center is at the origin and the principal axis is the z-axis
- **Principal point** ($\mathbf{p}$): point where principal axis intersects the image plane (origin of normalized coordinate system)

Slide by Lana Lazebnik

Principal point offset

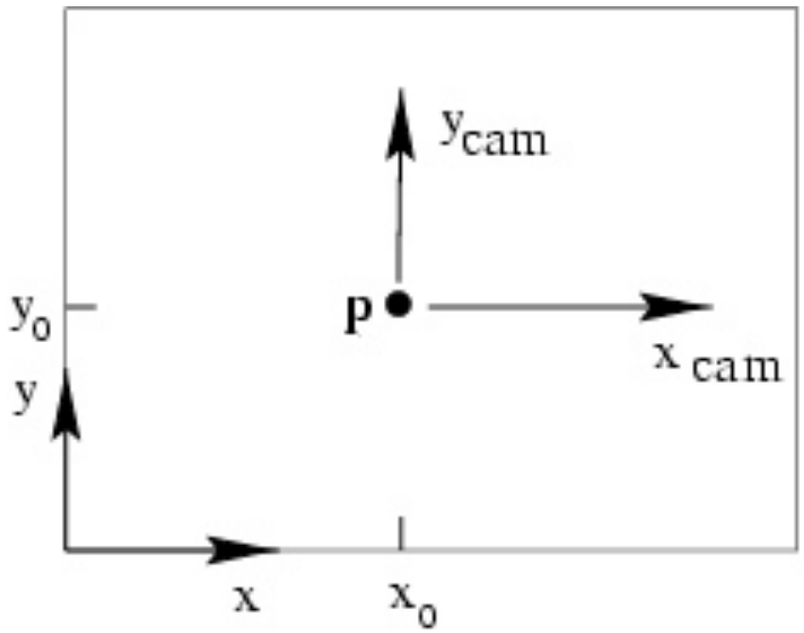


principal point: (p_x, p_y)

- **Camera coordinate system:** origin is at the principal point
- **Image coordinate system:** origin is in the corner

Slide by Lana Lazebnik

Principal point offset

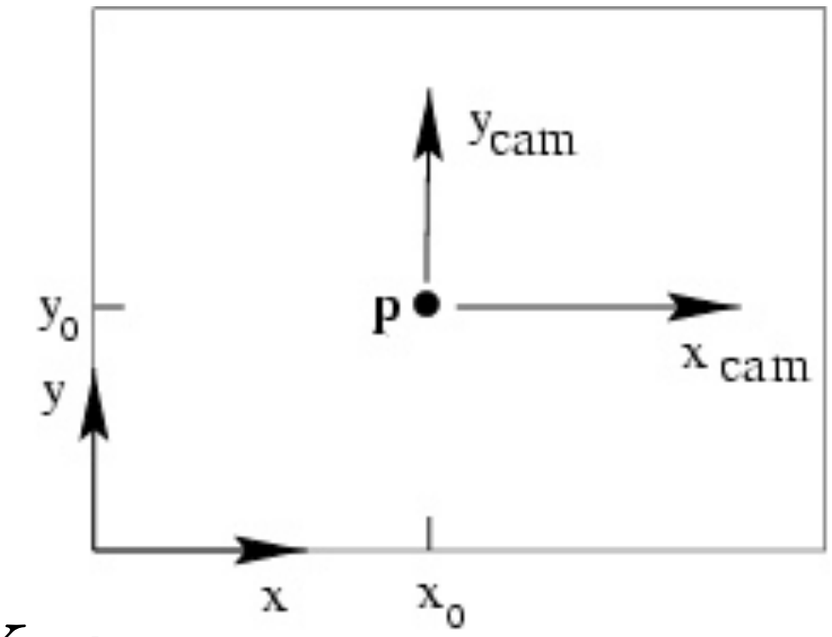


principal point: (p_x, p_y)

$$(X, Y, Z) \rightarrow (fX/Z + p_x, fY/Z + p_y)$$

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} fX + Zp_x \\ fY + Zp_y \\ Z \end{pmatrix} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Principal point offset



principal point: (p_x, p_y)

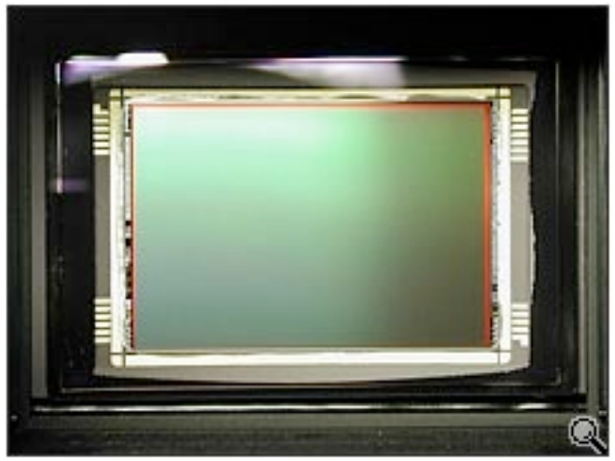
$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} fX + Zp_x \\ fY + Zp_y \\ Z \end{pmatrix} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$K = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix}$$

calibration matrix

$$P = K[I|0]$$

Pixel coordinates



Pixel size: $\frac{1}{m_x} \times \frac{1}{m_y}$

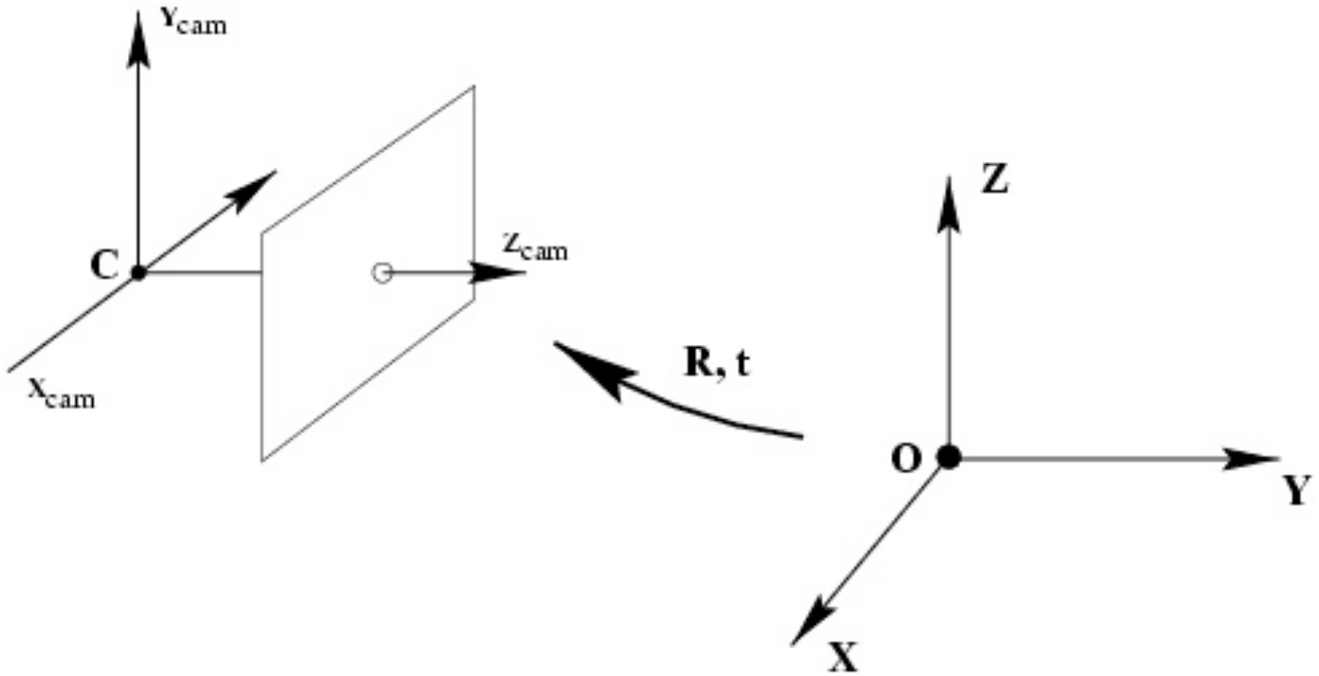
m_x pixels per unit (m,mm,inch,...) in horizontal direction,

m_y pixels per unit in vertical direction

$$K = \underbrace{\begin{bmatrix} m_x & & \\ & m_y & \\ & & 1 \end{bmatrix}}_{\text{pixels/unit}} \underbrace{\begin{bmatrix} f & & p_x \\ & f & p_y \\ & & 1 \end{bmatrix}}_{\text{units}} = \underbrace{\begin{bmatrix} \alpha_x & & \beta_x \\ & \alpha_y & \beta_y \\ & & 1 \end{bmatrix}}_{\text{pixels}}$$

Slide by Lana Lazebnik

Camera rotation and translation



- The camera coordinate frame is related to the world coordinate frame by a rotation and a translation

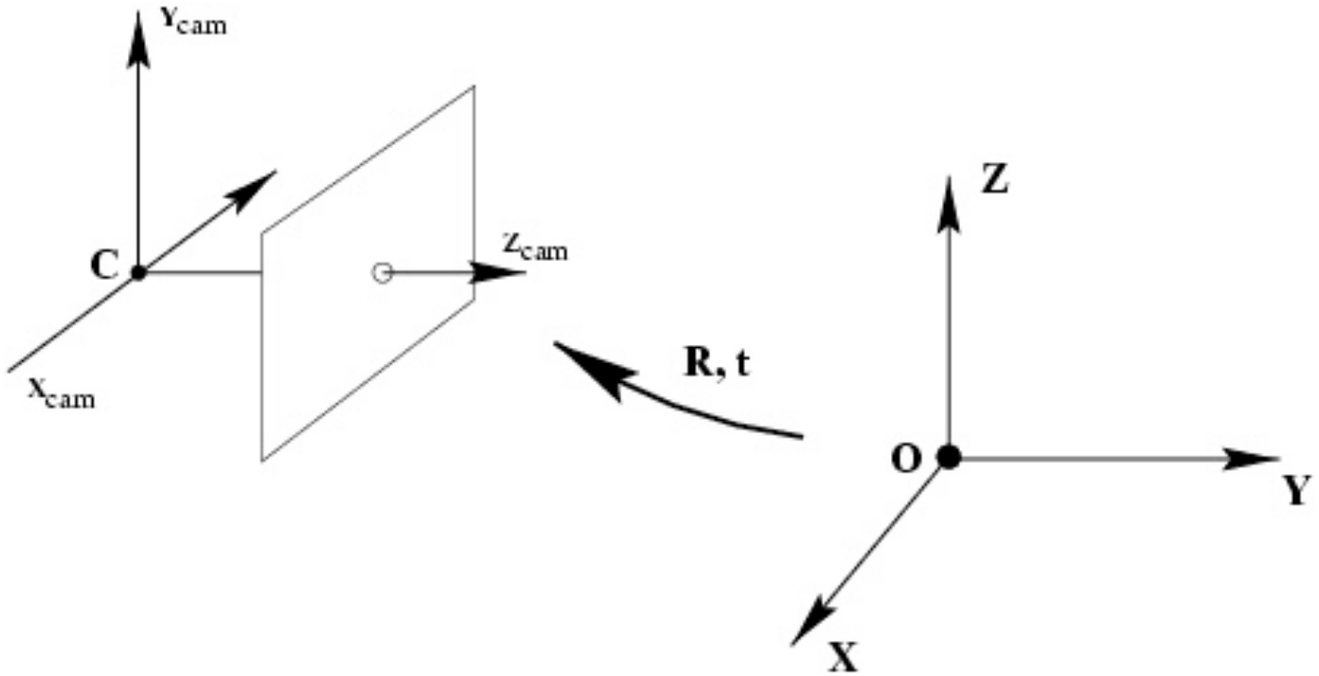
$$\tilde{X}_{cam} = R(\tilde{X} - \tilde{C})$$

coords. of point in camera frame

coords. of a point in world frame (non-homogeneous)

coords. of camera center in world frame

Camera rotation and translation



In non-homogeneous coordinates:

$$\tilde{X}_{cam} = R(\tilde{X} - \tilde{C})$$

$$X_{cam} = \begin{bmatrix} R & -R\tilde{C} \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \tilde{X} \\ 1 \end{pmatrix} = \begin{bmatrix} R & -R\tilde{C} \\ 0 & 1 \end{bmatrix} X$$

$$x = K[I|0]X_{cam} = K[R|-R\tilde{C}]X \qquad P = K[R|t] \qquad t = -R\tilde{C}$$

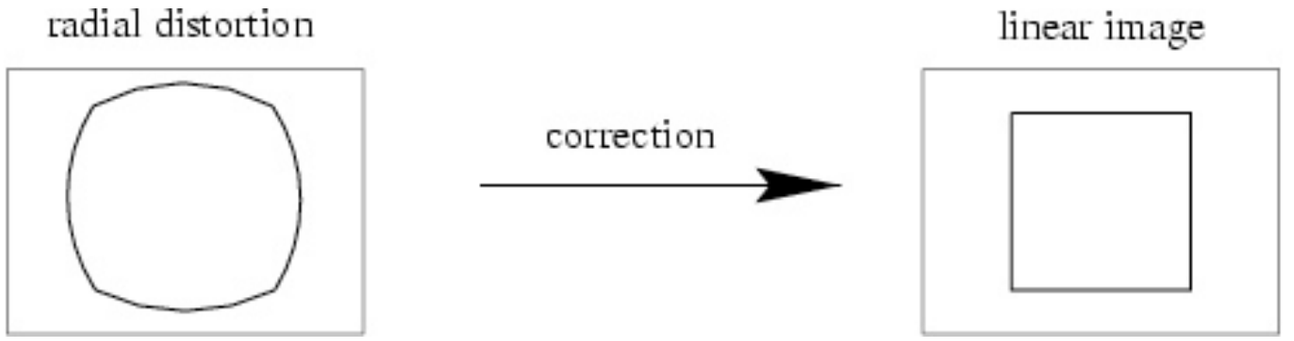
Note: C is the null space of the camera projection matrix ($PC = 0$)

Camera parameters

TECHNISCHE UNIVERSITÄT DARMSTADT

- Intrinsic parameters
  - Principal point coordinates
  - Focal length
  - Pixel magnification factors
  - Skew (non-rectangular pixels)
  - Radial distortion

$$K = \begin{bmatrix} m_x & & \\ & m_y & \\ & & 1 \end{bmatrix} \begin{bmatrix} f & & p_x \\ & f & p_y \\ & & 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & & \beta_x \\ & \alpha_y & \beta_y \\ & & 1 \end{bmatrix}$$



Camera parameters

TECHNISCHE UNIVERSITÄT DARMSTADT

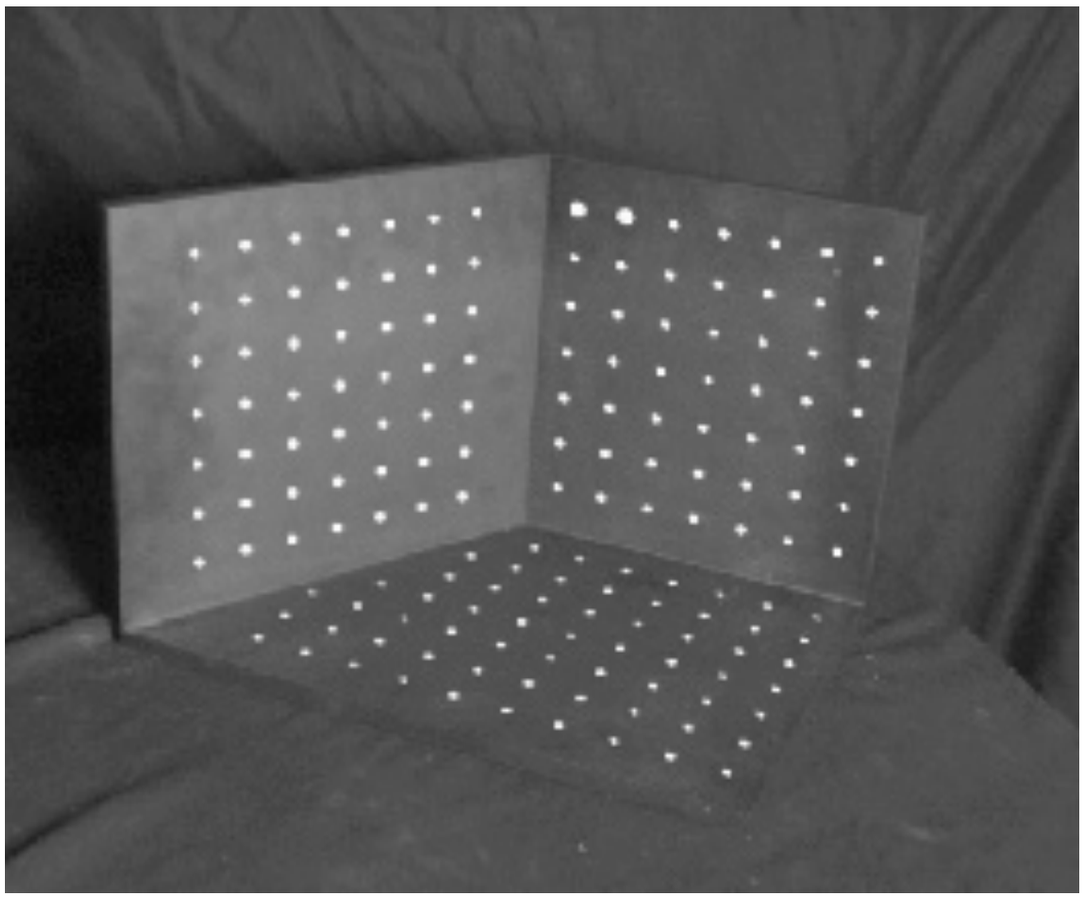
- Intrinsic parameters
  - Principal point coordinates
  - Focal length
  - Pixel magnification factors
  - Skew (non-rectangular pixels)
  - Radial distortion
- Extrinsic parameters
  - Rotation and translation relative to world coordinate system

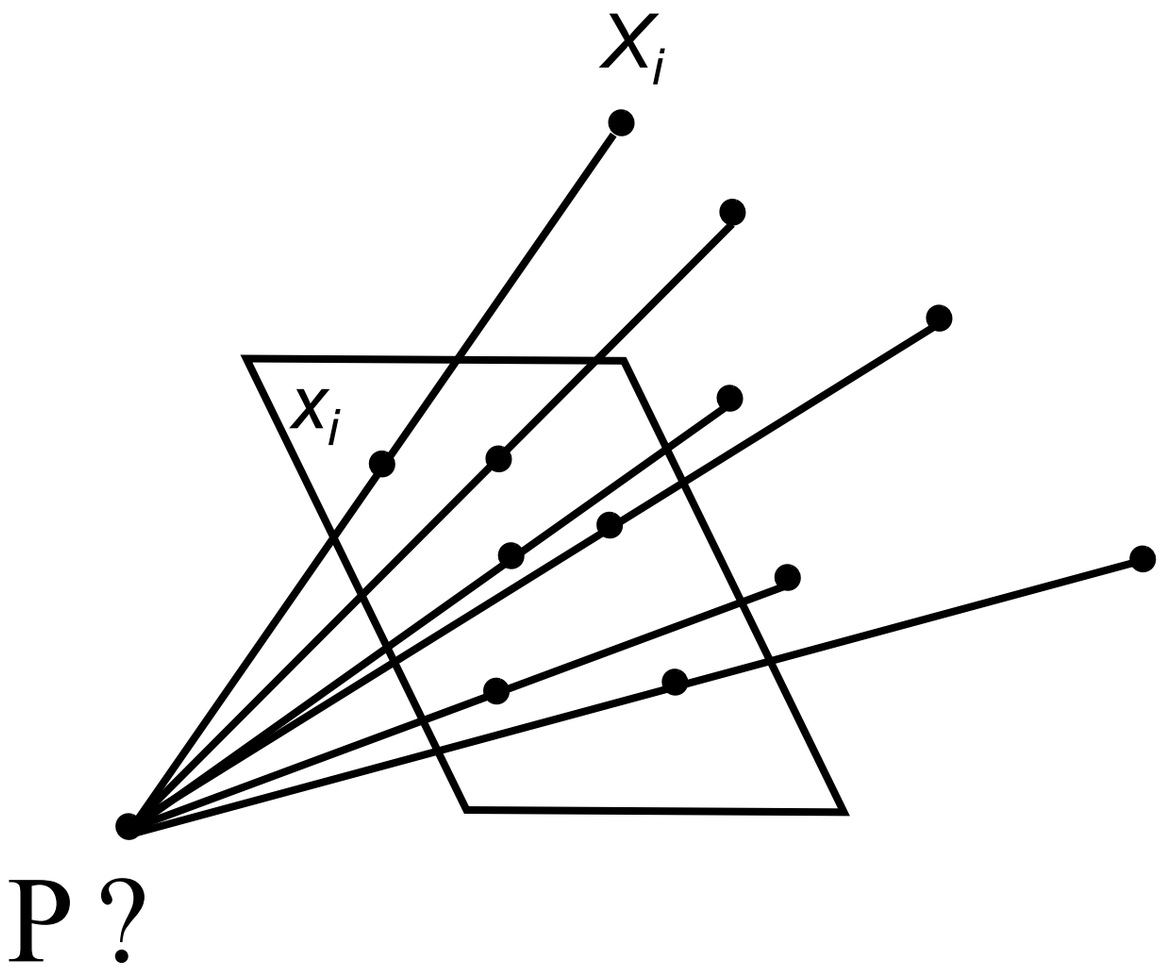
Slide by Lana Lazebnik

Camera calibration

TECHNISCHE UNIVERSITÄT DARMSTADT

- Given *n* points with known 3D coordinates X_i and known image projections x_i, estimate the camera parameters





Slide by Lana Lazebnik

Camera calibration

TECHNISCHE UNIVERSITÄT DARMSTADT

- What have we got now?
  - We have a projection matrix P, which maps any point in the 3D world coordinate frame onto the (infinite) image plane of the camera
  - Given a 3D point in the world, we can find out where its projection is in the image
  - One goal of vision: given a point in the images, find out where there pre-image is in the world

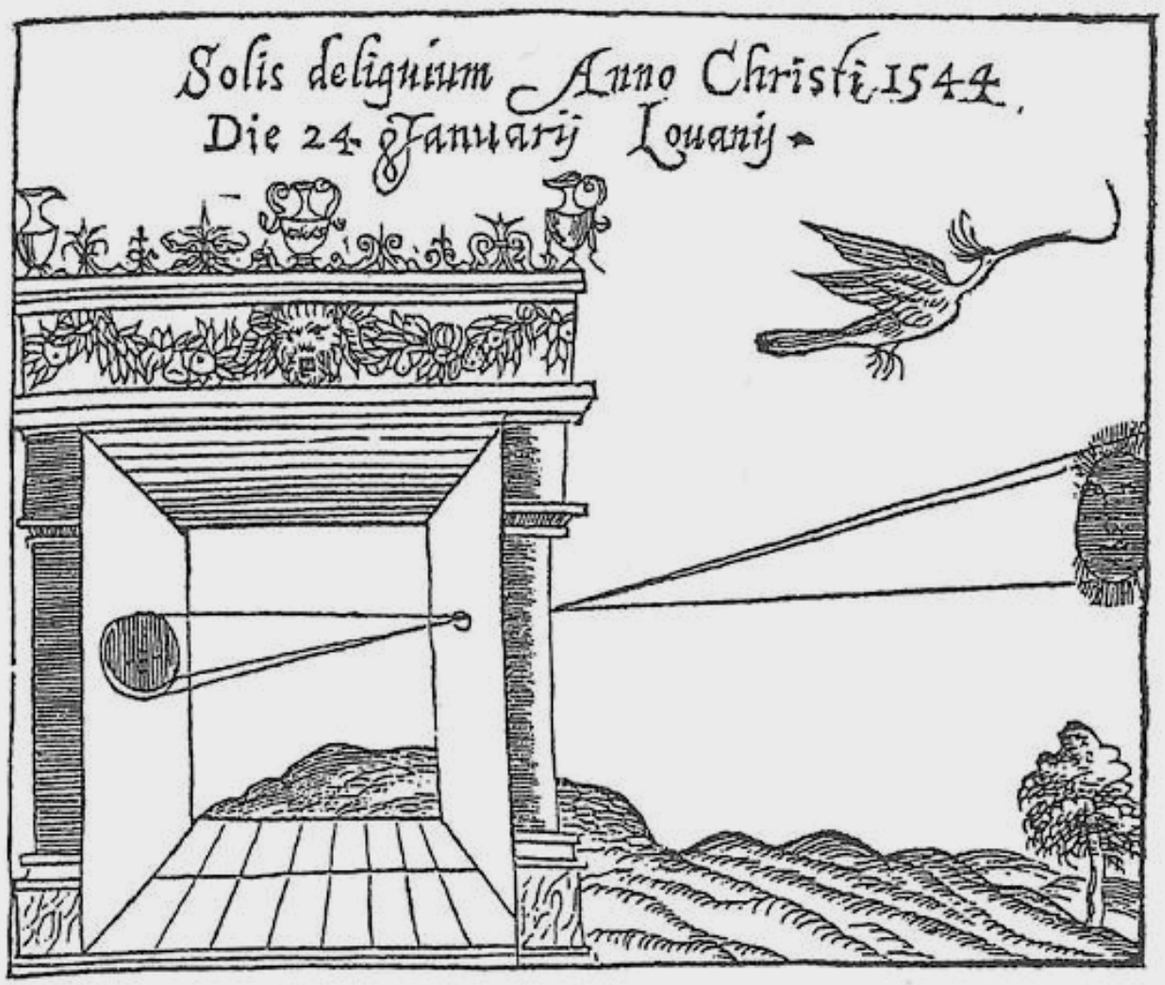
- We cannot do this from one image
  - The depth is lost, the projection matrix can only tell us the ray through an image point x

Building a real camera

TECHNISCHE UNIVERSITÄT DARMSTADT

Slide by Lana Lazebnik

Camera Obscura



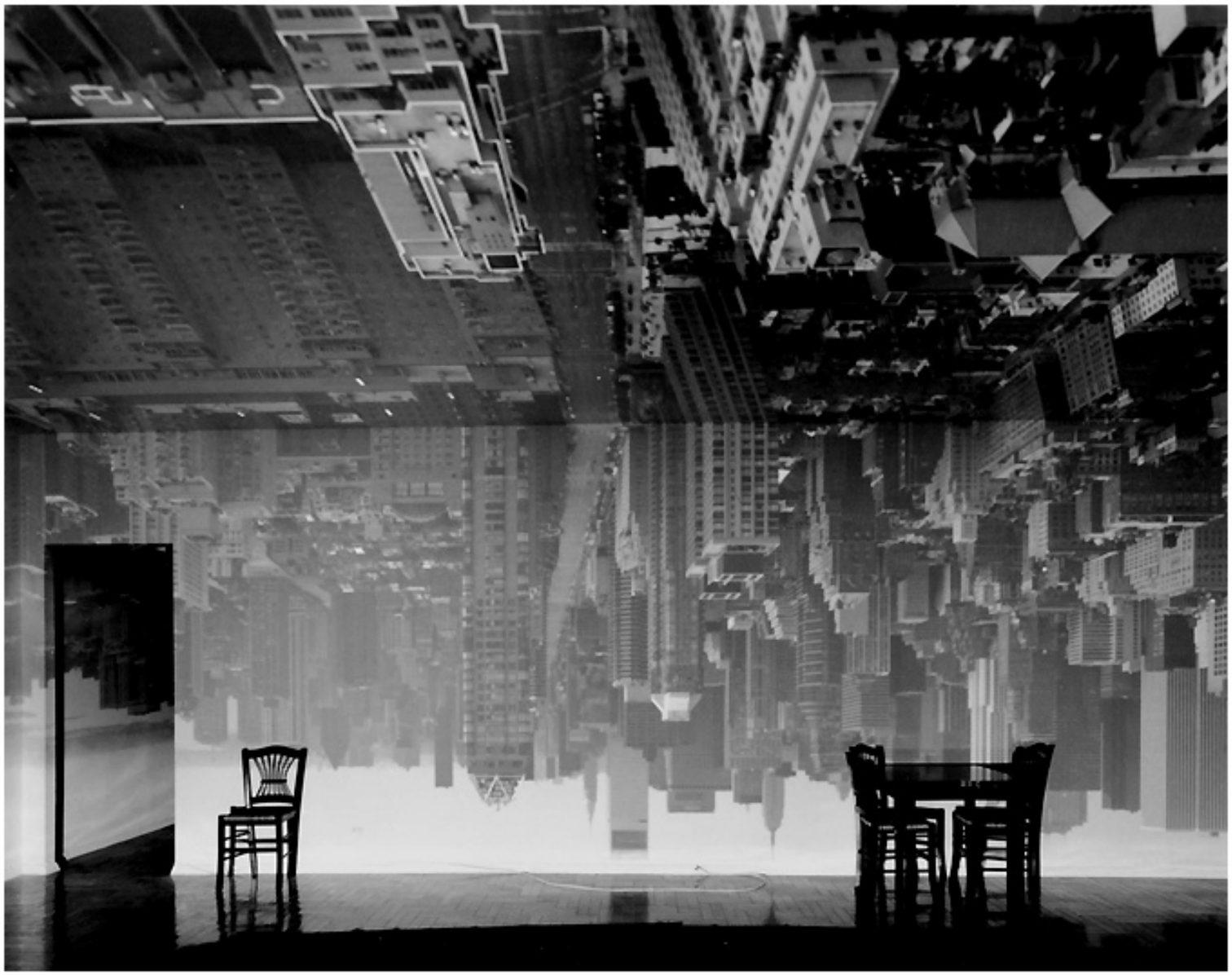
Gemma Frisius, 1558

- Basic principle known to Mozi (470-391 BC) and Aristoteles (384-322 BC)
- Described by Ibn al-Haytham (Alhazen) in his "Book of optics" (around 1000AD).
- Drawing aid for artists: described by Leonardo da Vinci (1452-1519)

Source: A. Efros

Abelardo Morell

TECHNISCHE UNIVERSITÄT DARMSTADT



After scouting rooms and reserving one for at least a day, Morell masks the windows except for the aperture. He controls three elements: the size of the hole, with a smaller one yielding a sharper but dimmer image; the length of the exposure, usually eight hours; and the distance from the hole to the surface on which the outside image falls and which he will photograph. He used 4 x 5 and 8 x 10 view cameras and lenses ranging from 75 to 150 mm.

After he's done inside, it gets harder. "I leave the room and I am constantly checking the weather, I'm hoping the maid reads my note not to come in, I'm worrying that the sun will hit the plastic masking and it will fall down, or that I didn't trigger the lens."

From *Grand Images Through a Tiny Opening*, **Photo District News,** February 2005

- Camera Obscura Image of Manhattan View Looking South in Large Room, 1996

http://www.abelardomorell.net/camera_obscura1.html

Slide by Lana Lazebnik

Home-made pinhole camera

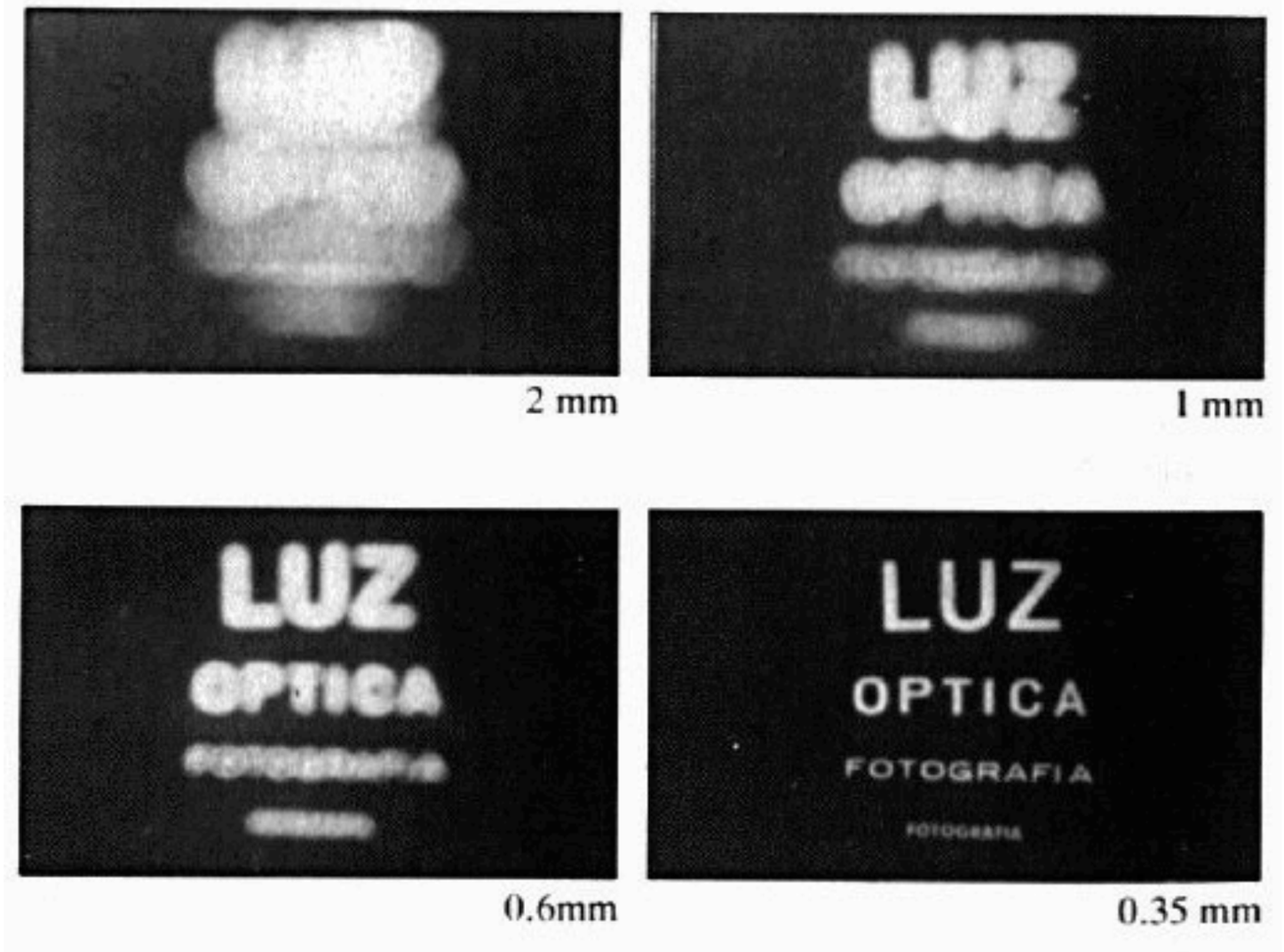
TECHNISCHE UNIVERSITÄT DARMSTADT

Why so blurry?

http://www.debevec.org/Pinhole/

Slide by A. Efros

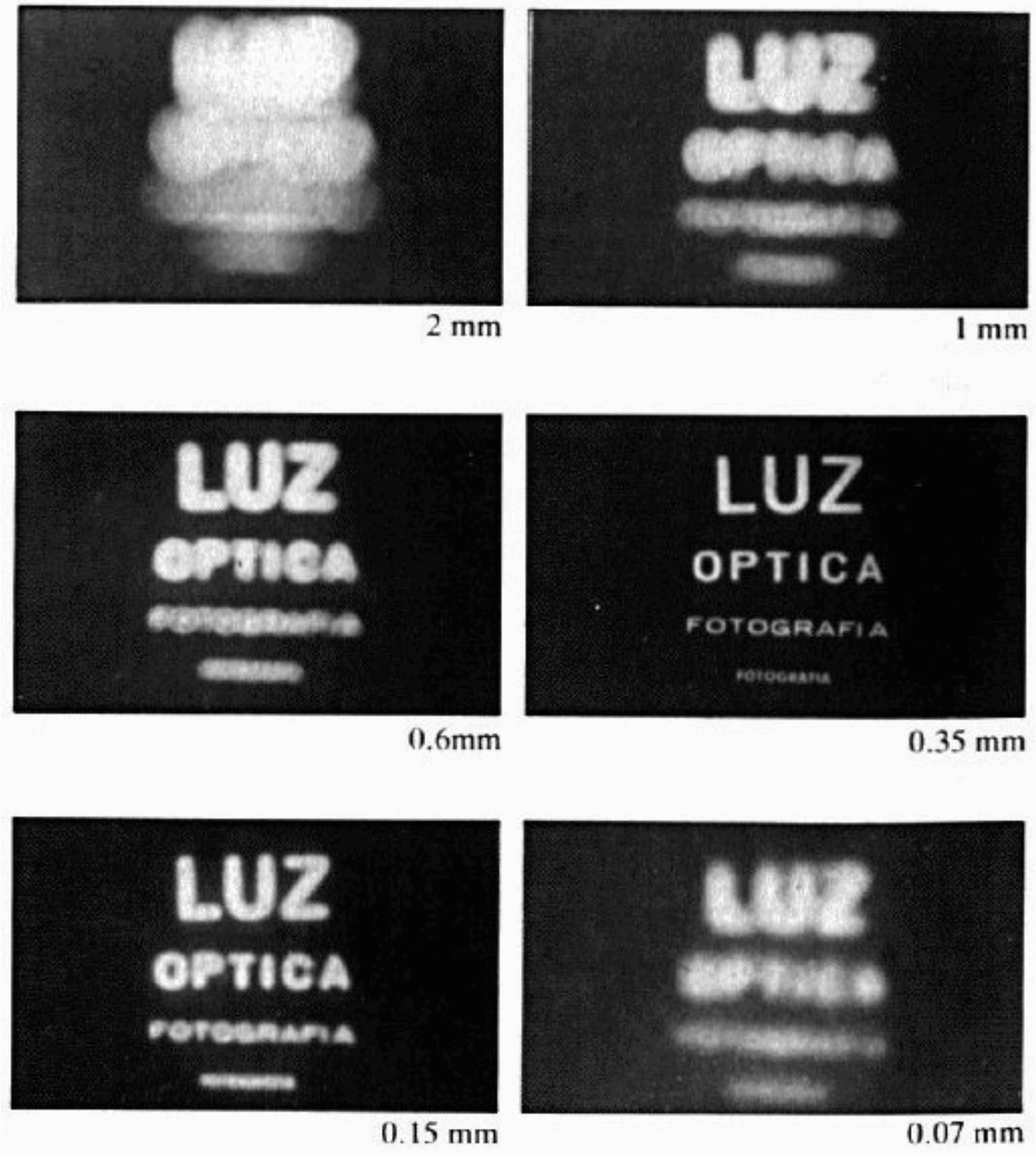
Shrinking the aperture



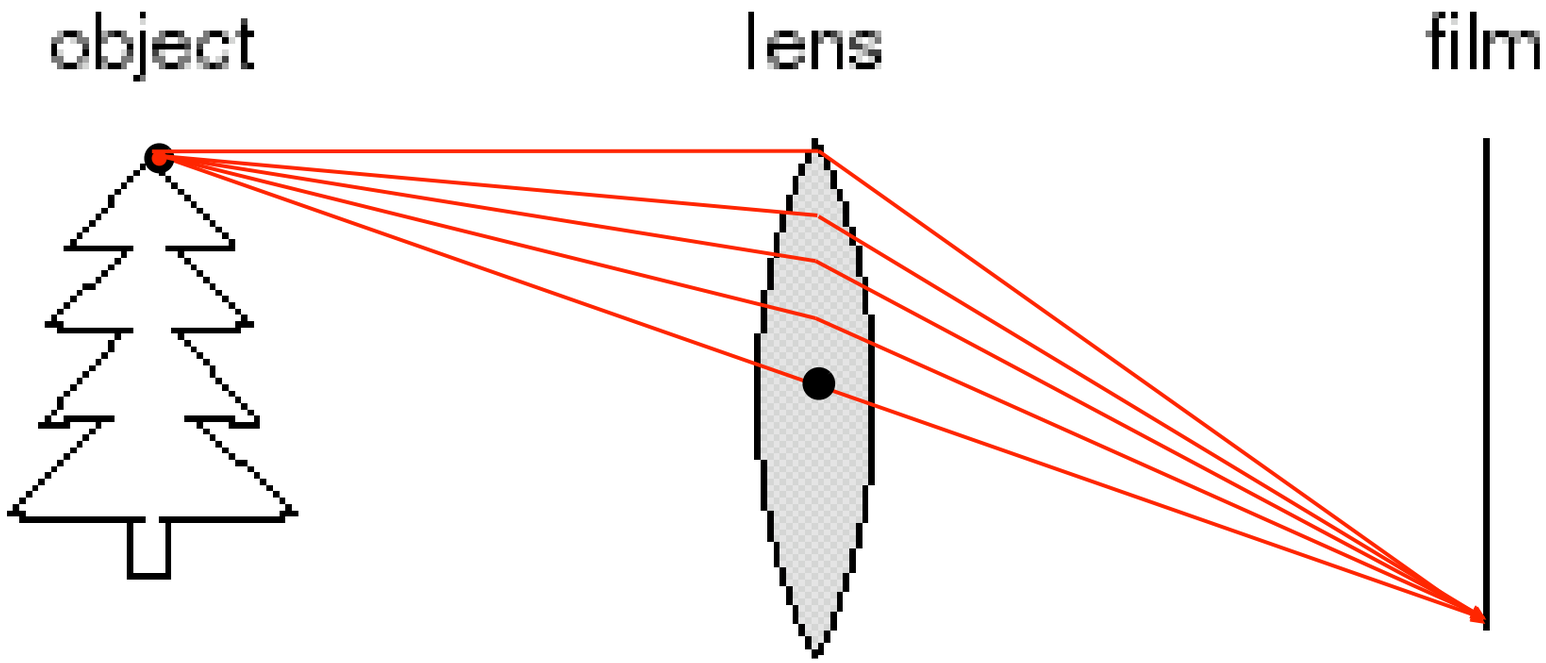
- Why not make the aperture as small as possible?
  - Less light gets through
  - Diffraction effects…

Slide by Steve Seitz

Shrinking the aperture



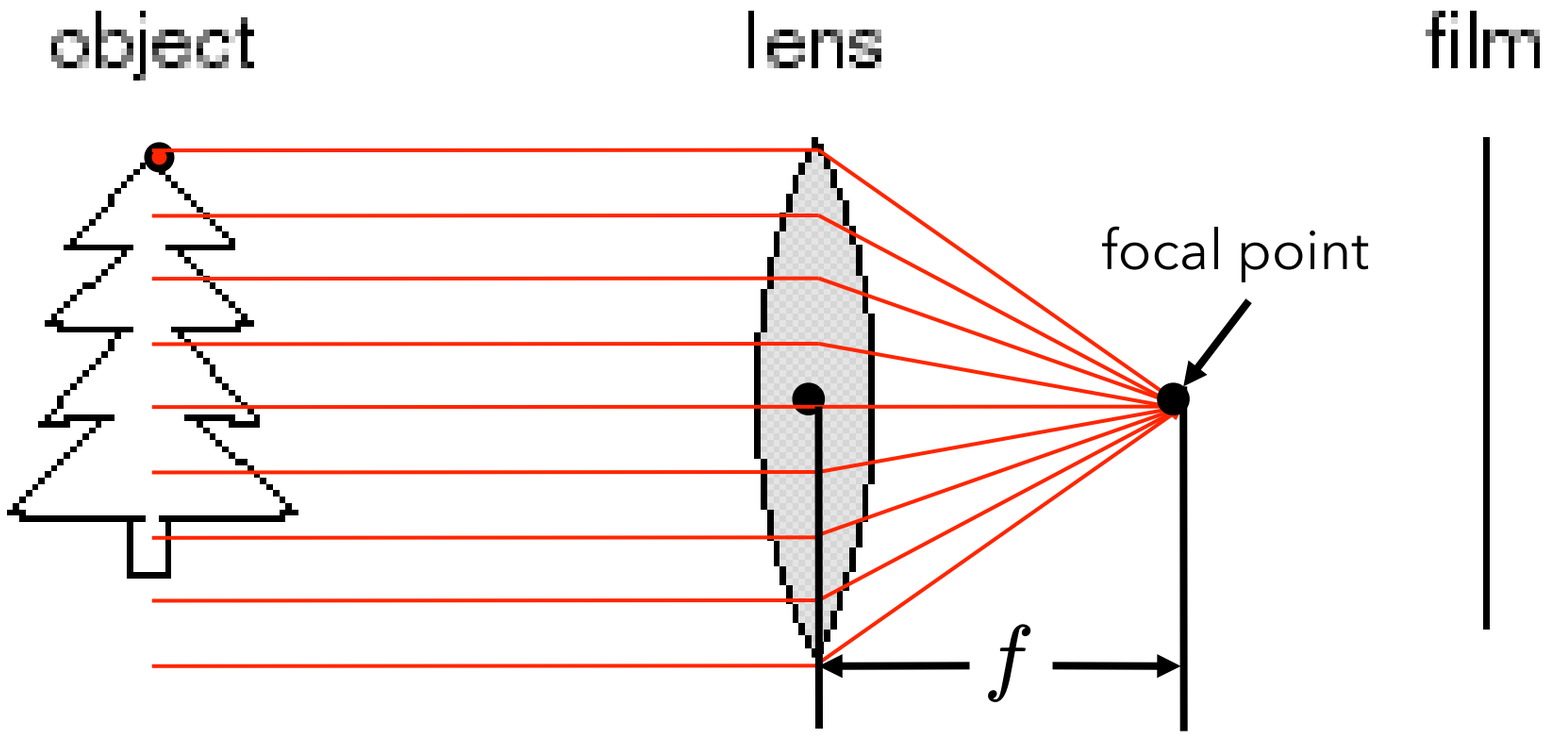
Adding a lens



- A lens focuses light onto the film
  - Rays passing through the center are not deviated

Slide by Steve Seitz

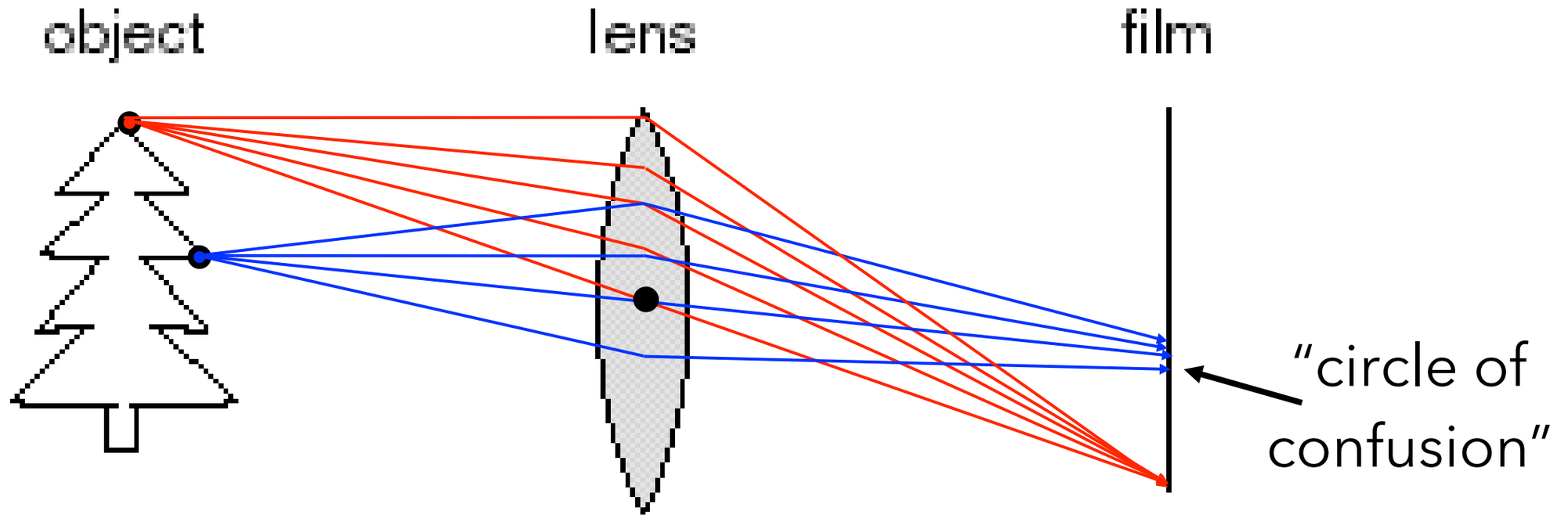
Adding a lens



- A lens focuses light onto the film
  - Rays passing through the center are not deviated
  - All parallel rays converge to one point on a plane located at the **focal length** f

Slide by Steve Seitz

Adding a lens



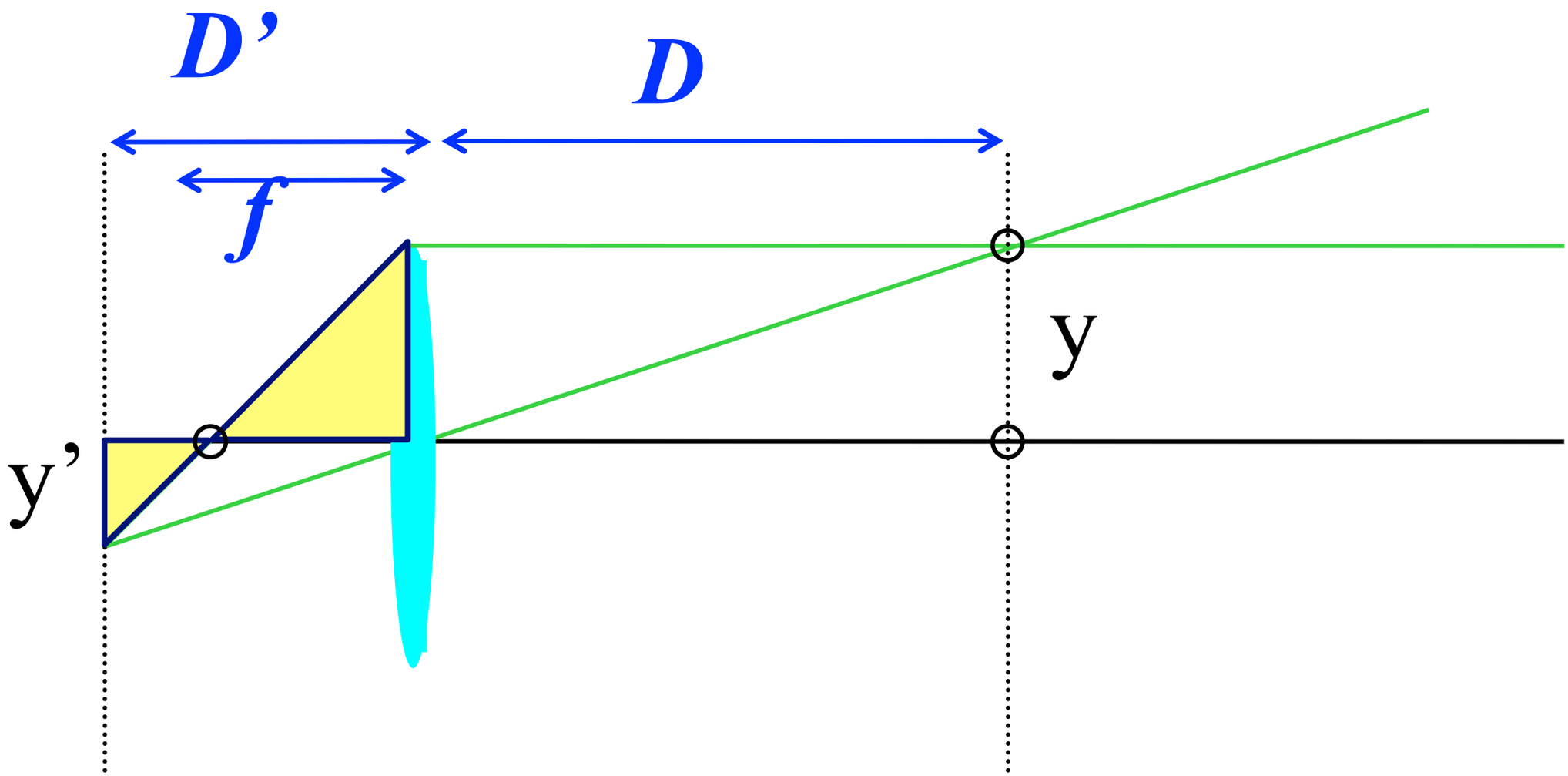
- A lens focuses light onto the film
  - There is a specific distance at which objects are "in focus"
  - other points project to a "circle of confusion" in the image

Slide by Steve Seitz

Thin lens formula

Similar triangles everywhere!

$$\frac{y'}{y} = \frac{D'}{D} \qquad \frac{y'}{y} = \frac{D' - f}{f}$$

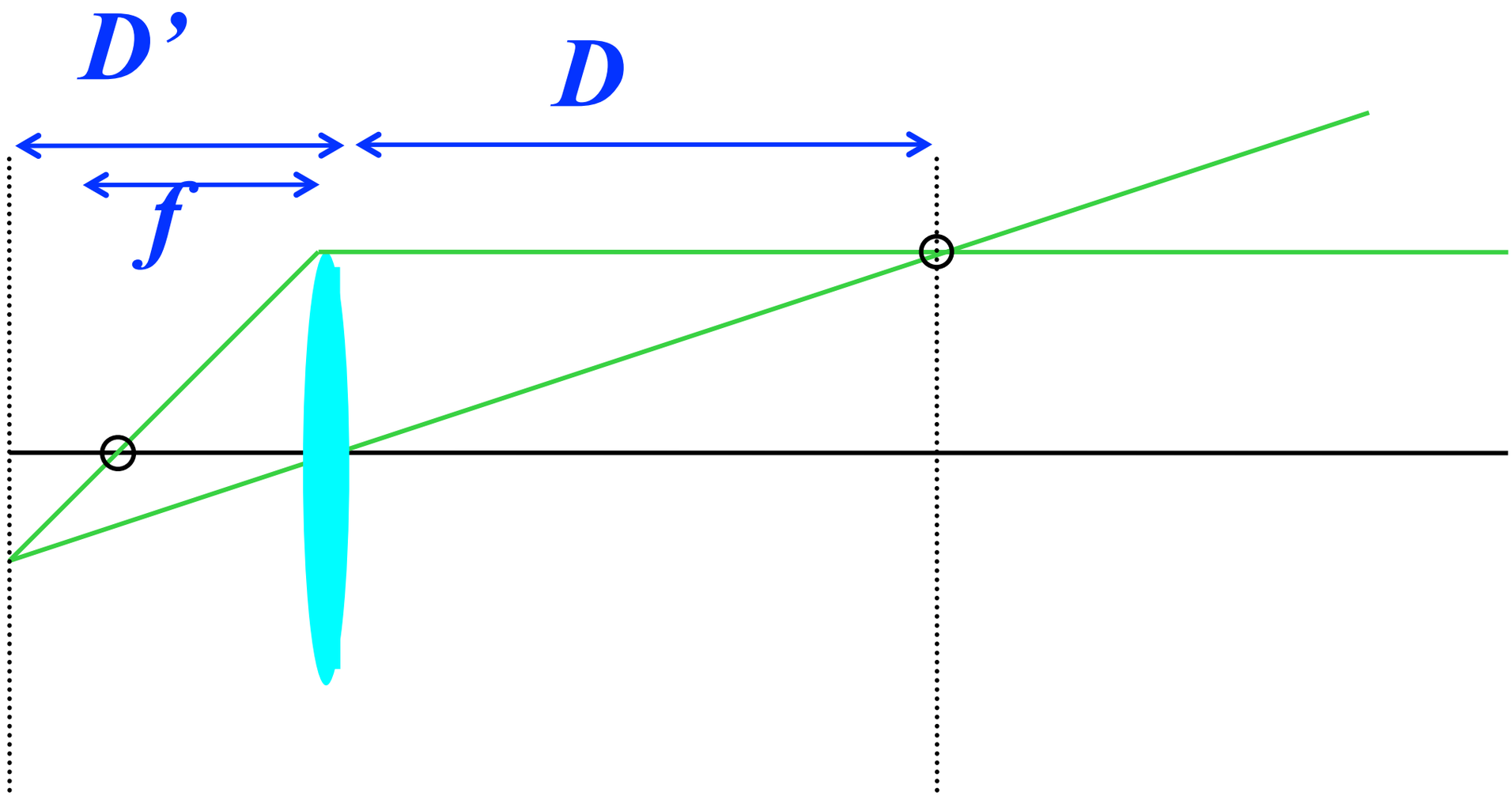


Slide from Frédo Durand

Thin lens formula

$$\frac{1}{D'} + \frac{1}{D} = \frac{1}{f}$$

Any point satisfying the **thin lens equation** is in focus.



Slide from Frédo Durand

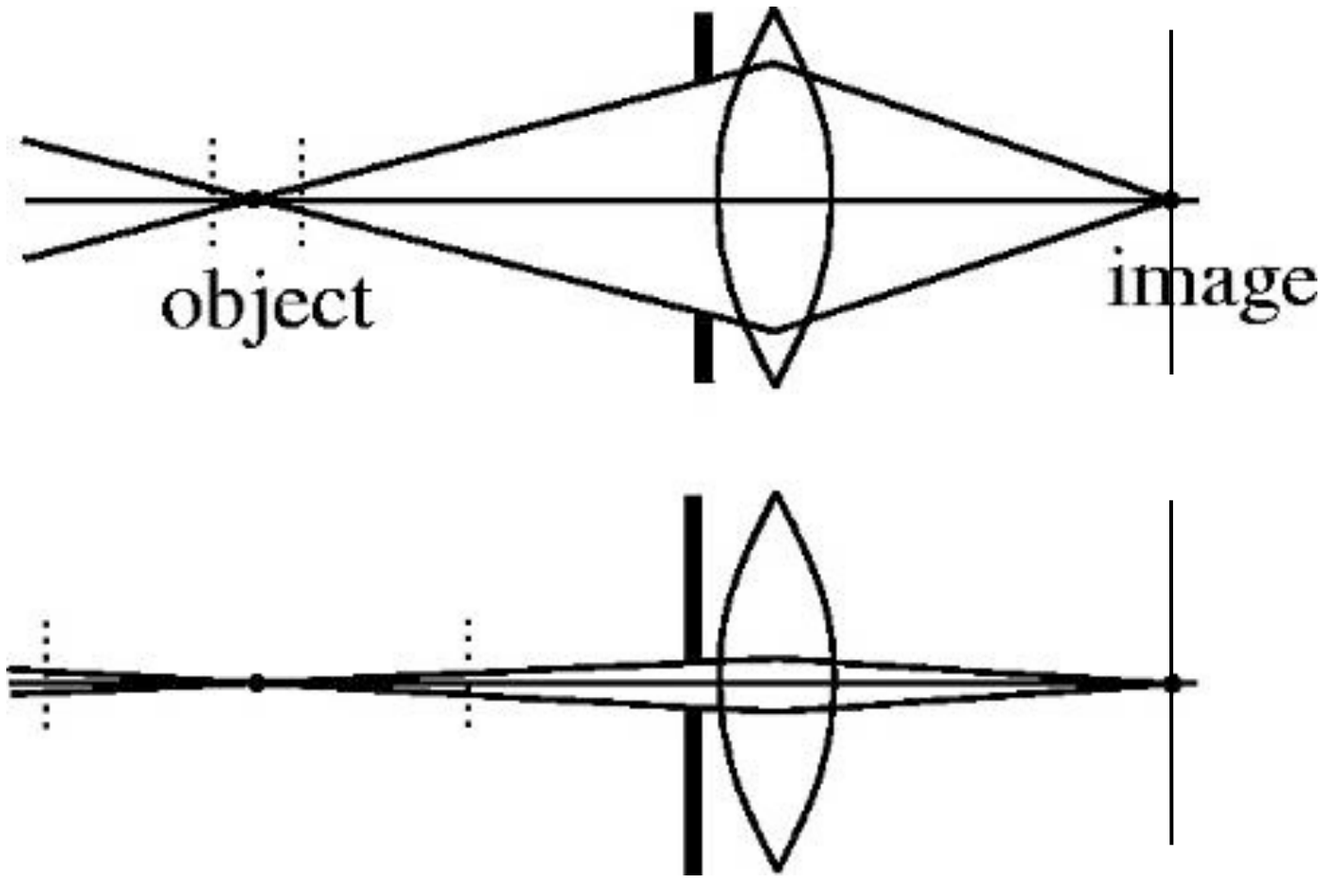
Depth of Field

TECHNISCHE UNIVERSITÄT DARMSTADT

http://www.cambridgeincolour.com/tutorials/depth-of-field.htm

Slide by A. Efros

How can we control the depth of field?

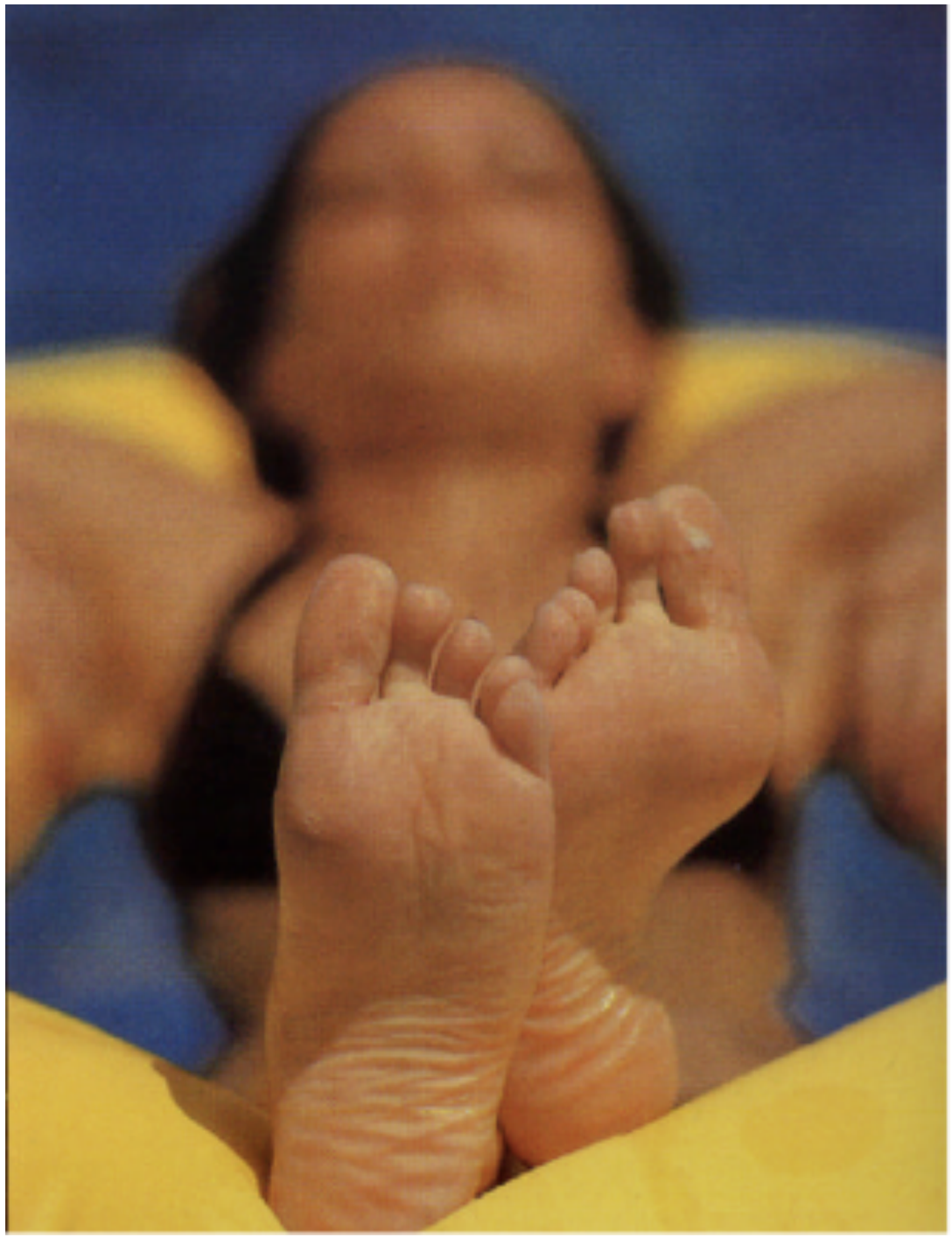


- Changing the aperture size affects depth of field
  - A smaller aperture increases the range in which the object is approximately in focus
  - But small aperture reduces amount of light – need to increase exposure

Slide by A. Efros

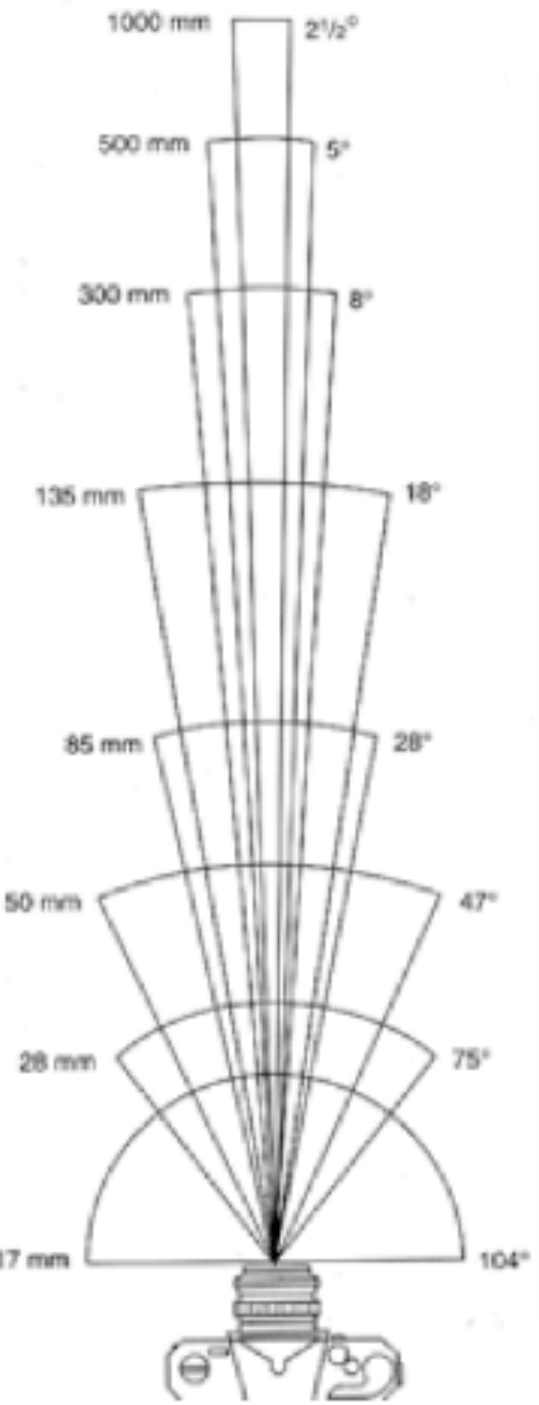
Depth of Field Effect

TECHNISCHE UNIVERSITÄT DARMSTADT



Source: F. Durand

Field of View (Zoom)



17mm

28mm

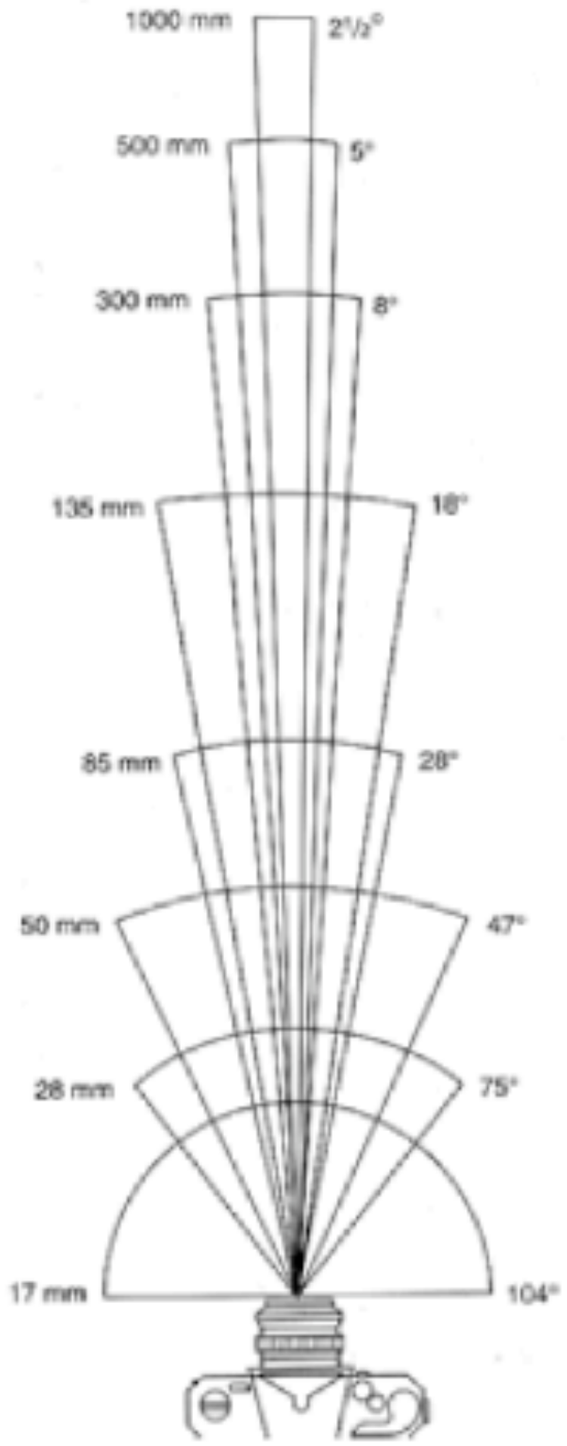
50mm

85mm

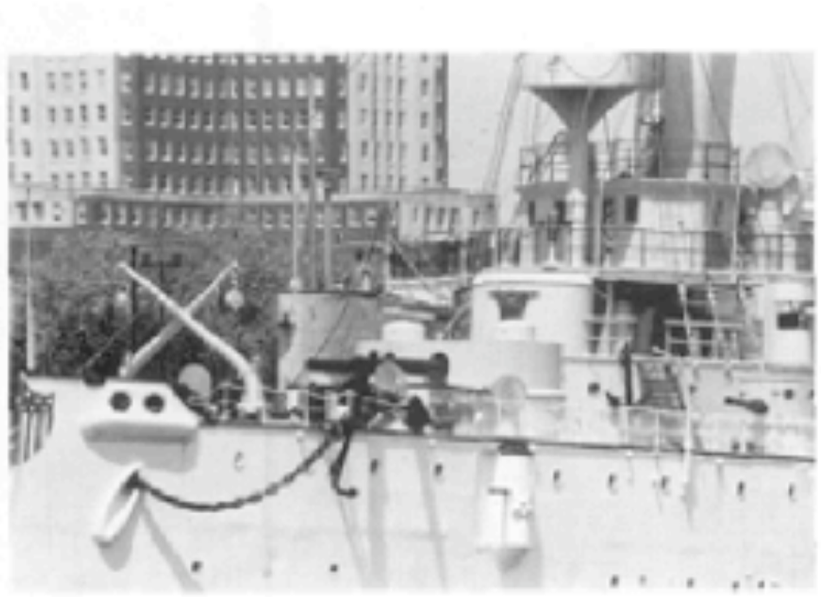
From London and Upton

Slide by A. Efros

Field of View (Zoom)



135mm

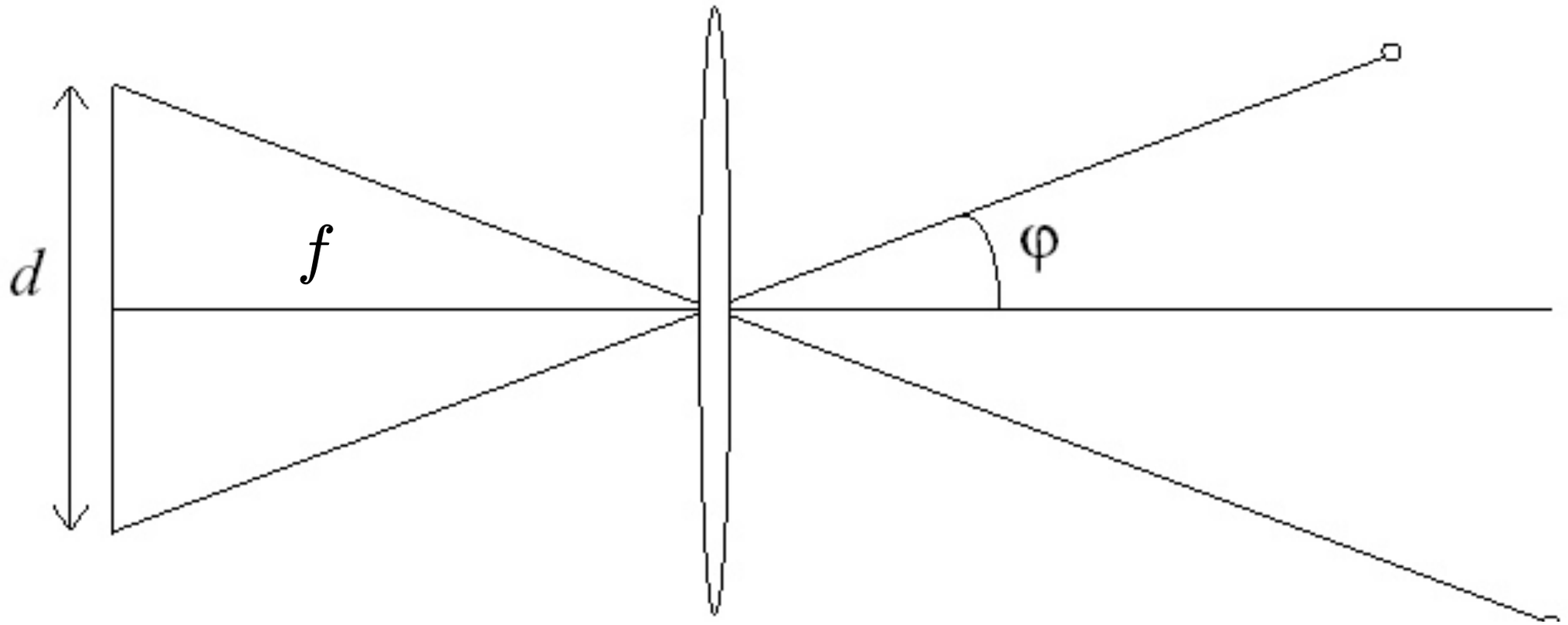


300mm

From London and Upton

Slide by A. Efros

Field of View



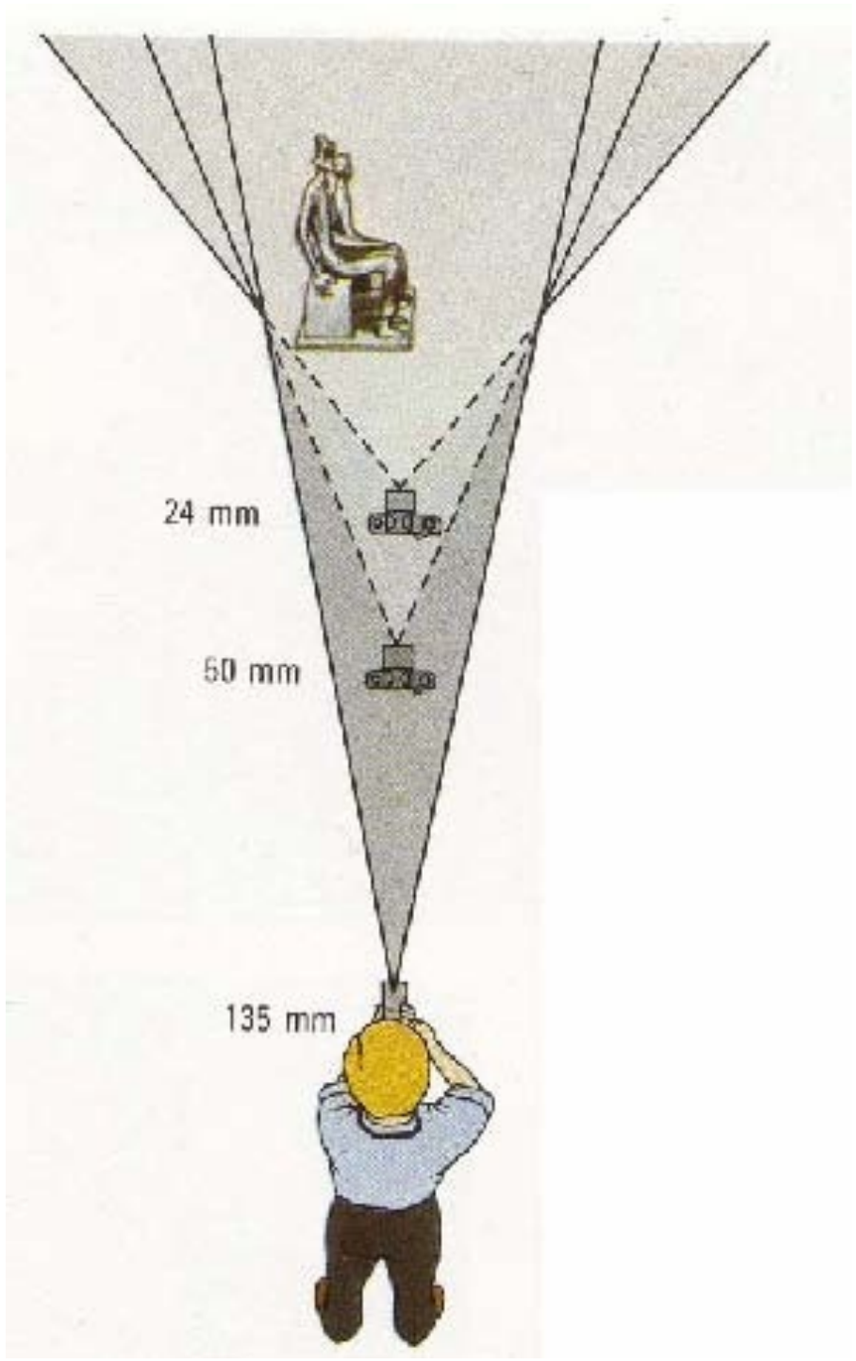
FOV depends on focal length *and*
the size of the camera retina (sensor)

$$\phi = \tan^{-1}\left(\frac{d}{2f}\right)$$

Smaller FOV = larger Focal Length

Slide by A. Efros

Field of View / Focal Length



Large FOV, small f

Camera close to car

Small FOV, large f

Camera far from the car

Sources: A. Efros, F. Durand

Same effect for faces

TECHNISCHE UNIVERSITÄT DARMSTADT

wide-angle

standard

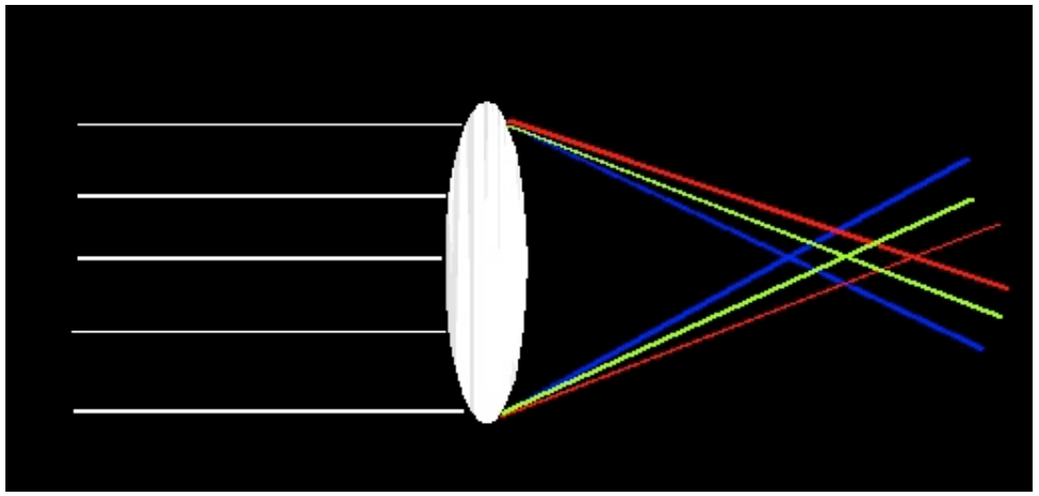
telephoto

Source: F. Durand

Lens Flaws: Chromatic Aberration

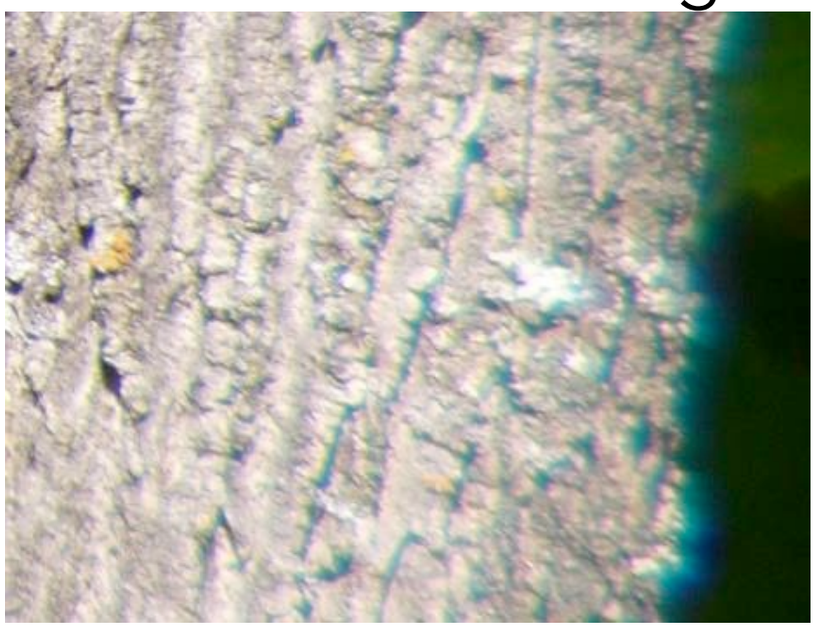
TECHNISCHE UNIVERSITÄT DARMSTADT

- Lens has different refractive indices for different wavelengths: causes color fringing



near lens center

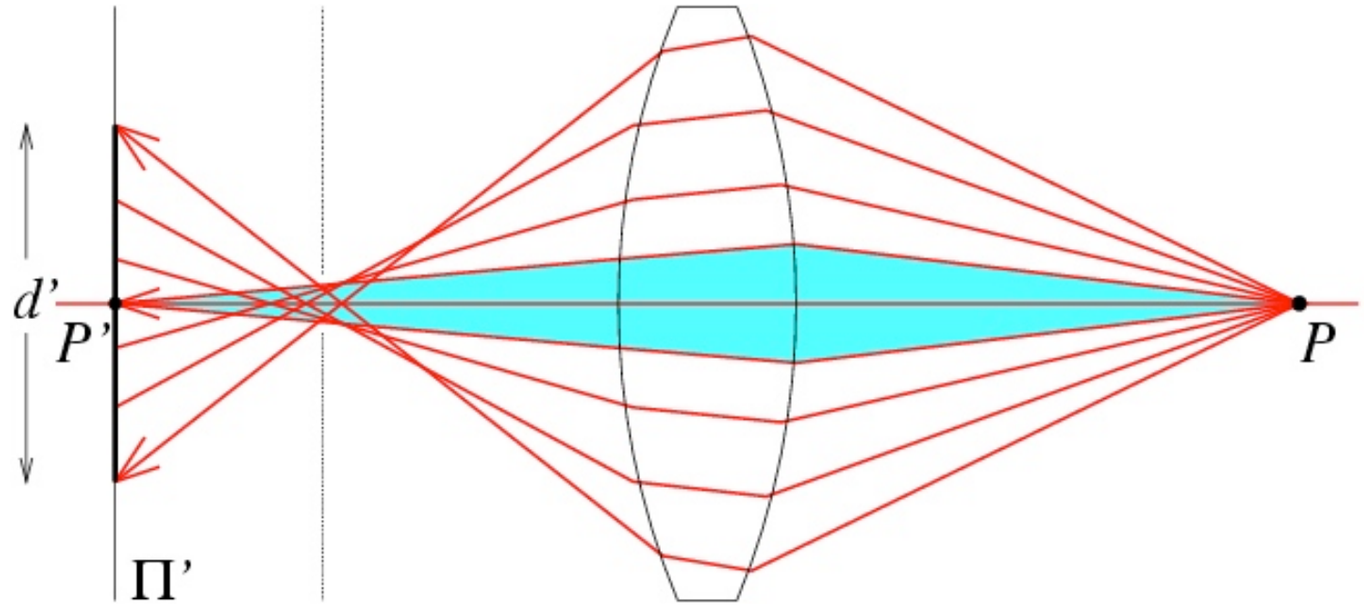
near lens outer edge



Lens flaws: Spherical aberration

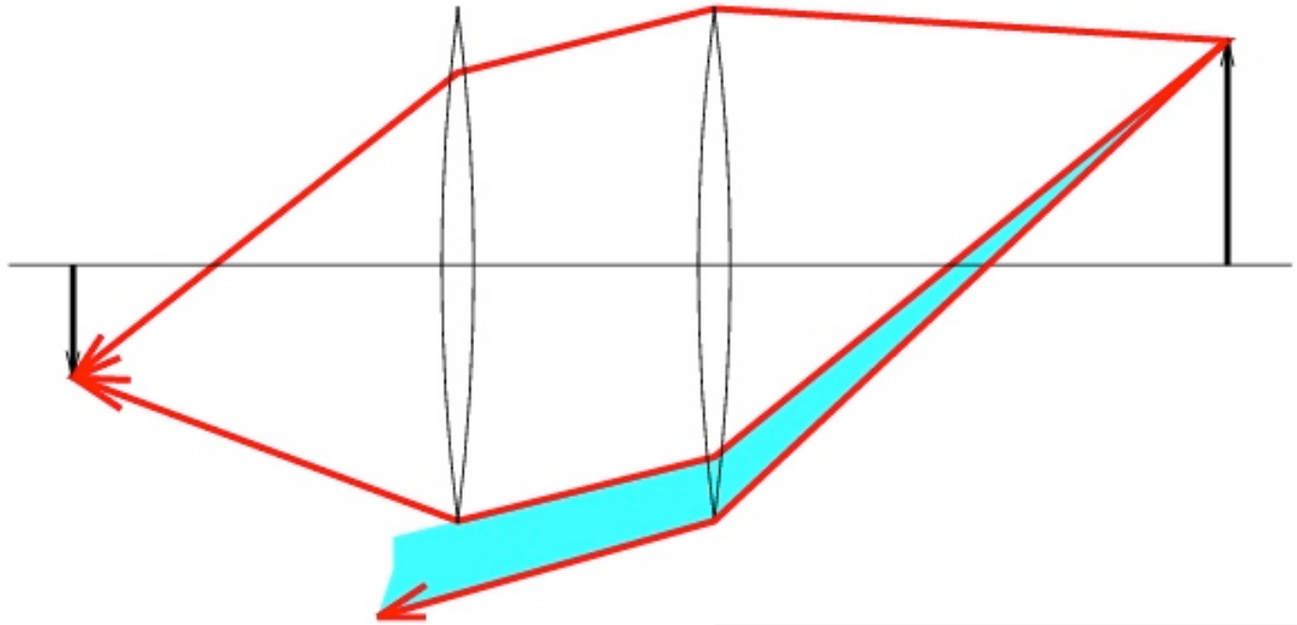
TECHNISCHE UNIVERSITÄT DARMSTADT

- Spherical lenses don't focus light perfectly
  - Rays farther from the optical axis focus closer



Slide by Lana Lazebnik

Lens flaws: Vignetting

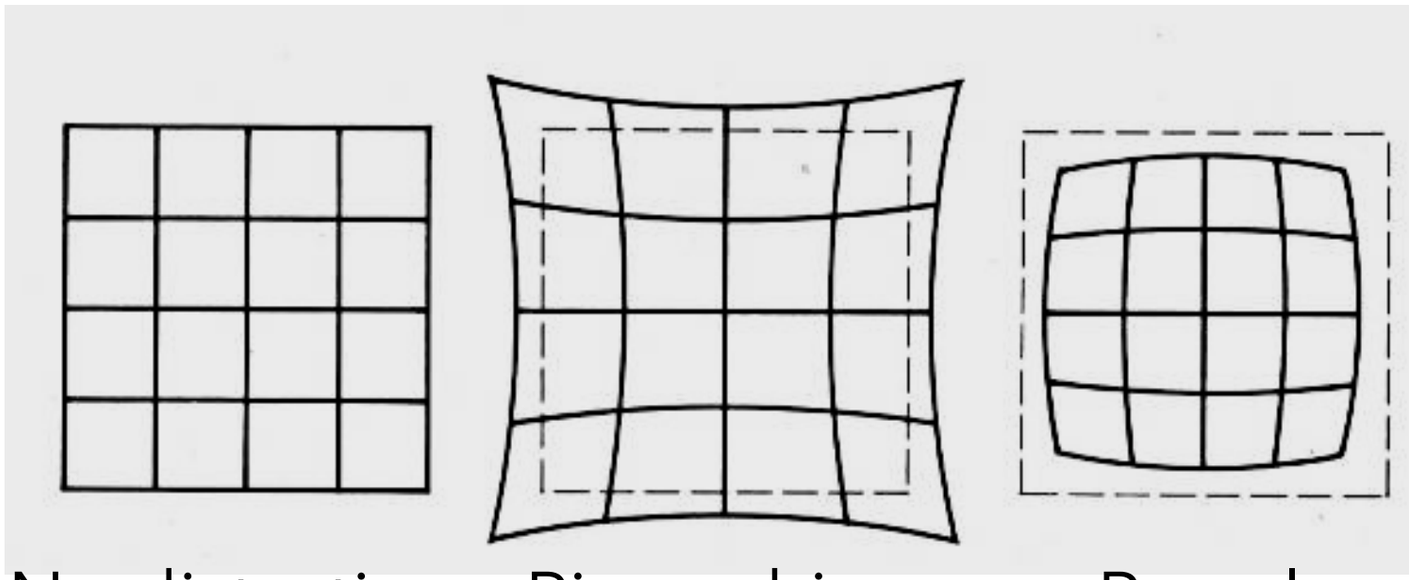


Slide by Lana Lazebnik

Radial Distortion

TECHNISCHE UNIVERSITÄT DARMSTADT

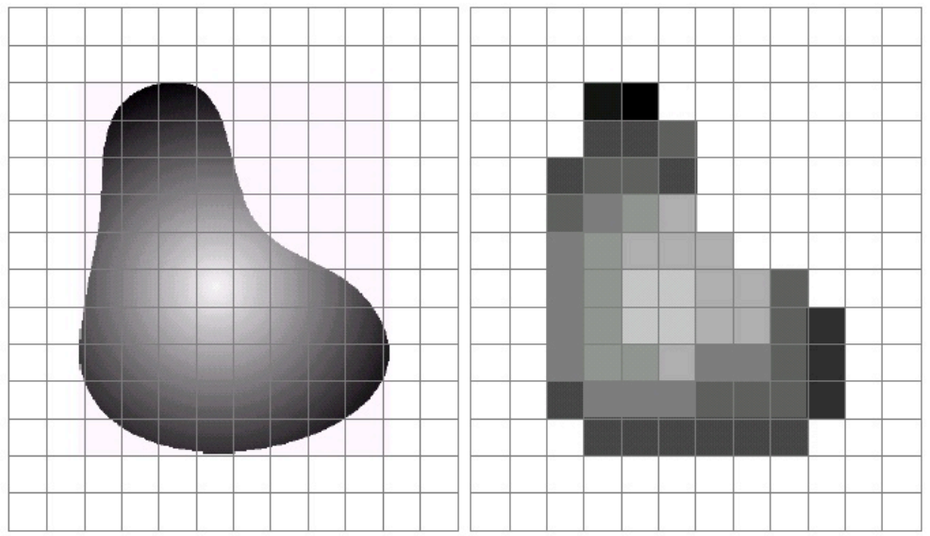
- Caused by imperfect lenses
- Deviations are most noticeable for rays that pass through the edge of the lens



No distortion Pin cushion Barrel

Slide by Lana Lazebnik

Digital camera



a b

FIGURE 2.17 (a) Continuos image projected onto a sensor array. (b) Result of image sampling and quantization.

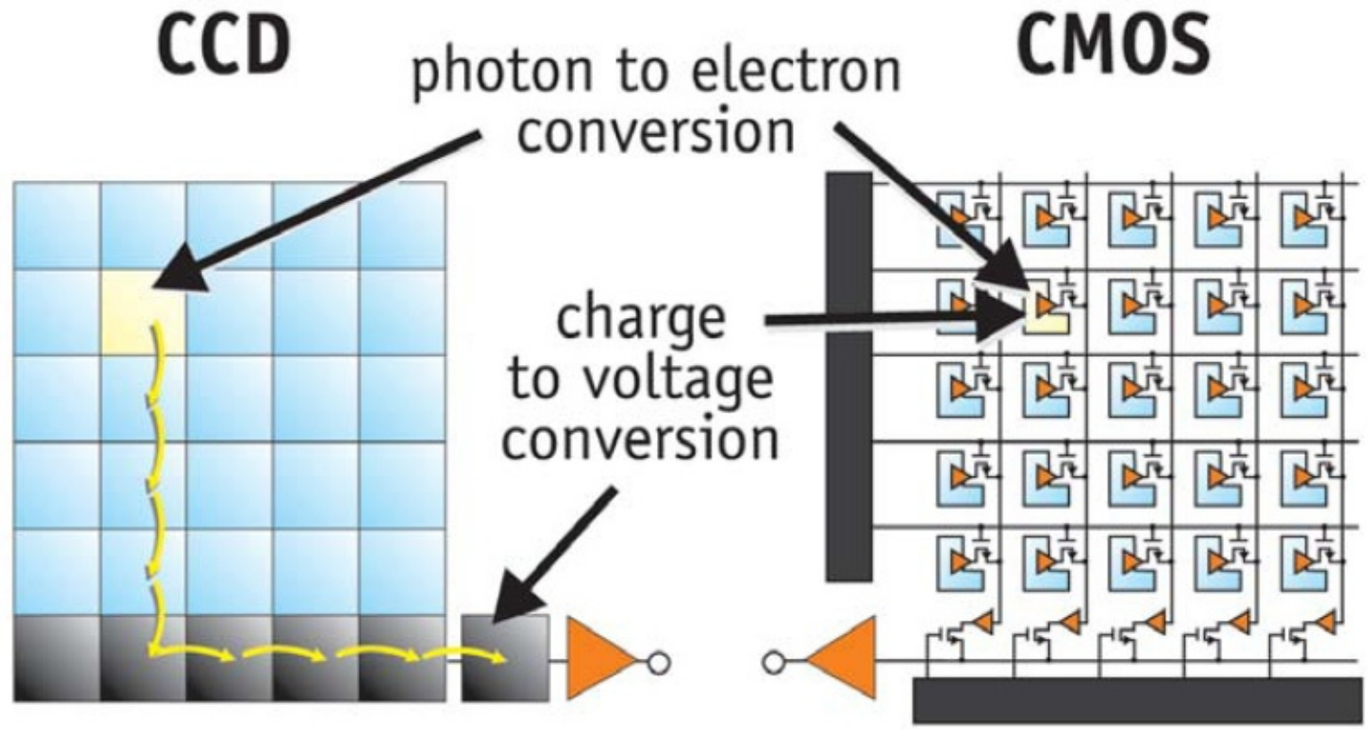
- A digital camera replaces film with a sensor array
  - Each cell in the array is light-sensitive diode that converts photons to electrons
  - Two common types
    - Charge Coupled Device (CCD)
    - Complementary metal oxide semiconductor (CMOS)

Slide by Steve Seitz

CCD vs. CMOS

TECHNISCHE UNIVERSITÄT DARMSTADT

- CCD: transports the charge across the chip and reads it at one corner of the array. An analog-to-digital converter (ADC) then turns each pixel's value into a digital value by measuring the amount of charge at each photosite and converting that measurement to binary form



CCDs move photogenerated charge from pixel to pixel and convert it to voltage at an output node. CMOS imagers convert charge to voltage inside each pixel.

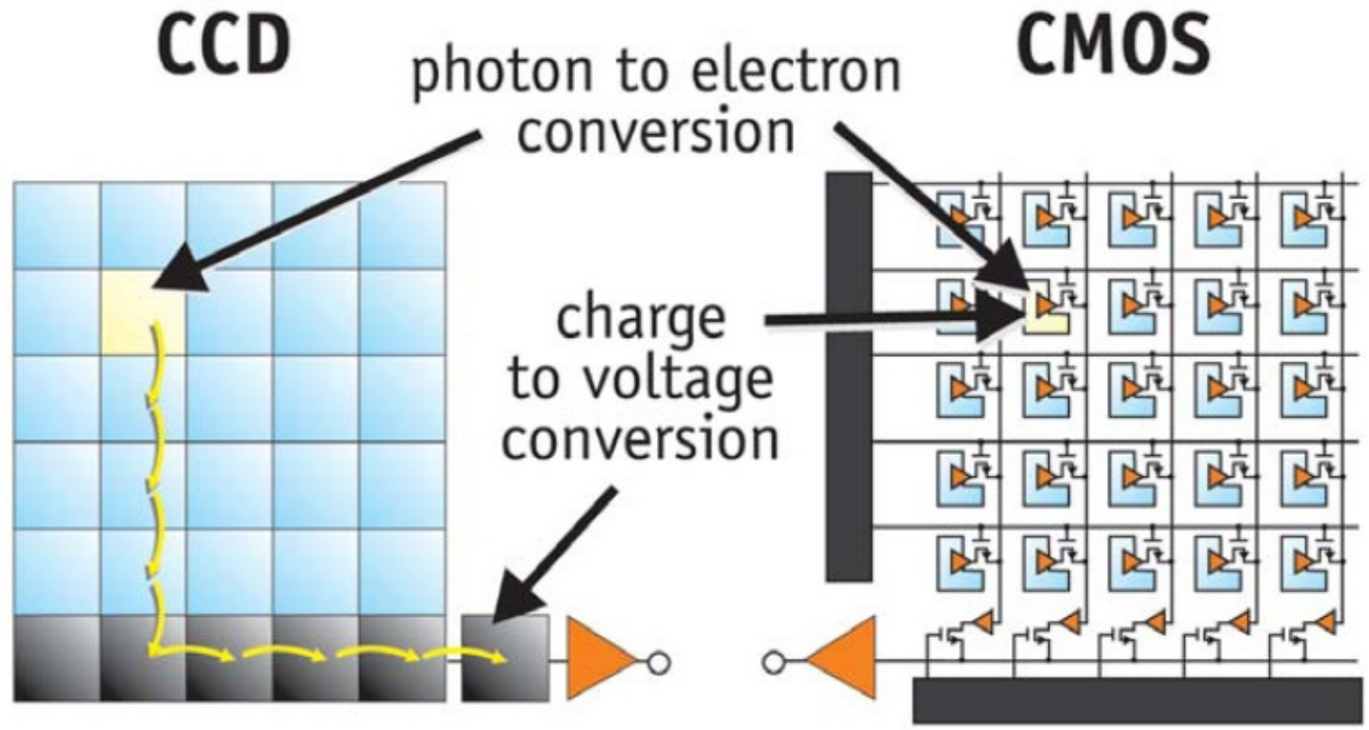
http://www.dalsa.com/shared/content/pdfs/CCD_vs_CMOS_Litwiller_2005.pdf

Slide by Lana Lazebnik

CCD vs. CMOS

TECHNISCHE UNIVERSITÄT DARMSTADT

- CMOS: uses several transistors at each pixel to amplify and move the charge using more traditional wires. The CMOS signal is digital, so it needs no ADC.



CCDs move photogenerated charge from pixel to pixel and convert it to voltage at an output node. CMOS imagers convert charge to voltage inside each pixel.

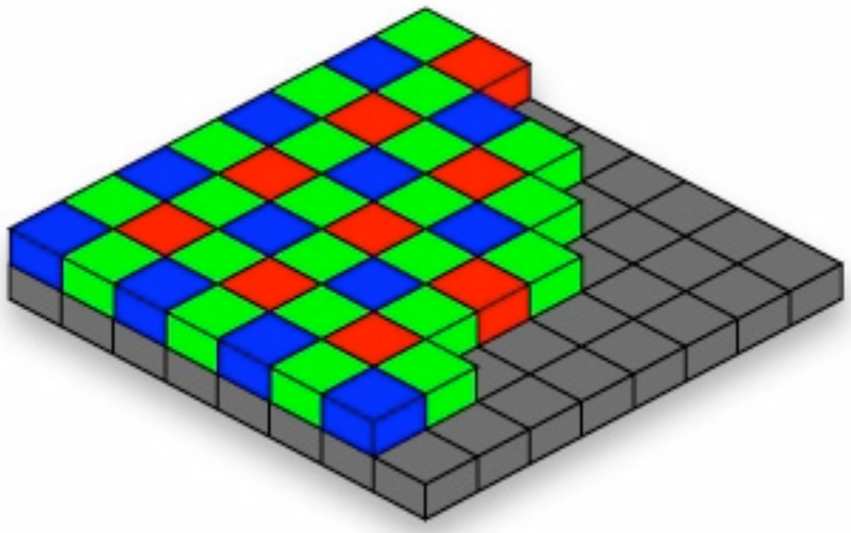
http://www.dalsa.com/shared/content/pdfs/CCD_vs_CMOS_Litwiller_2005.pdf

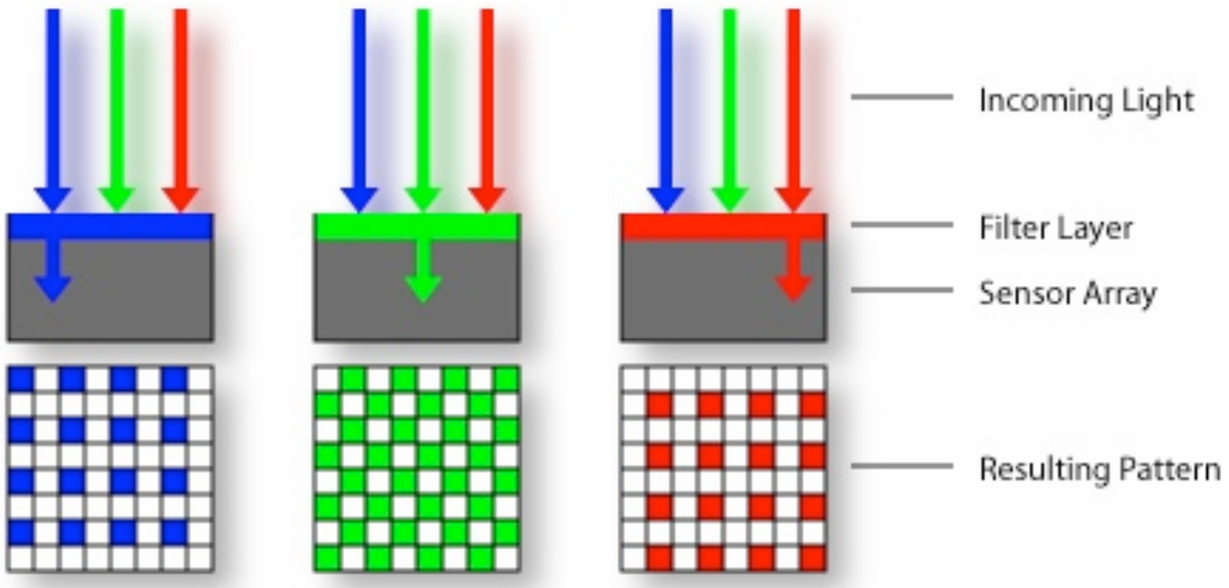
Slide by Lana Lazebnik

Color sensing in camera: Color filter array

TECHNISCHE UNIVERSITÄT DARMSTADT

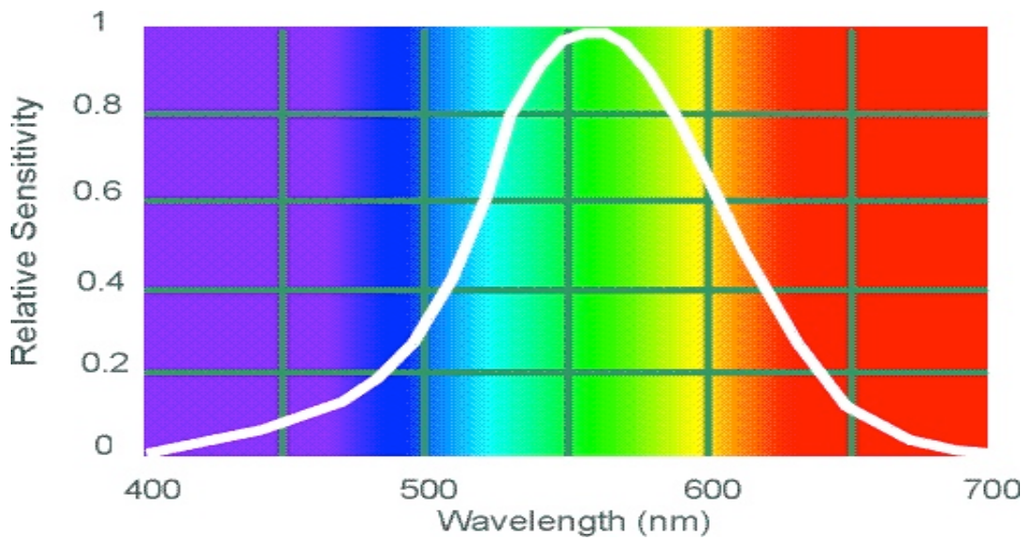
Bayer grid





Estimate missing components from neighboring values (demosaicing)

Why more green?



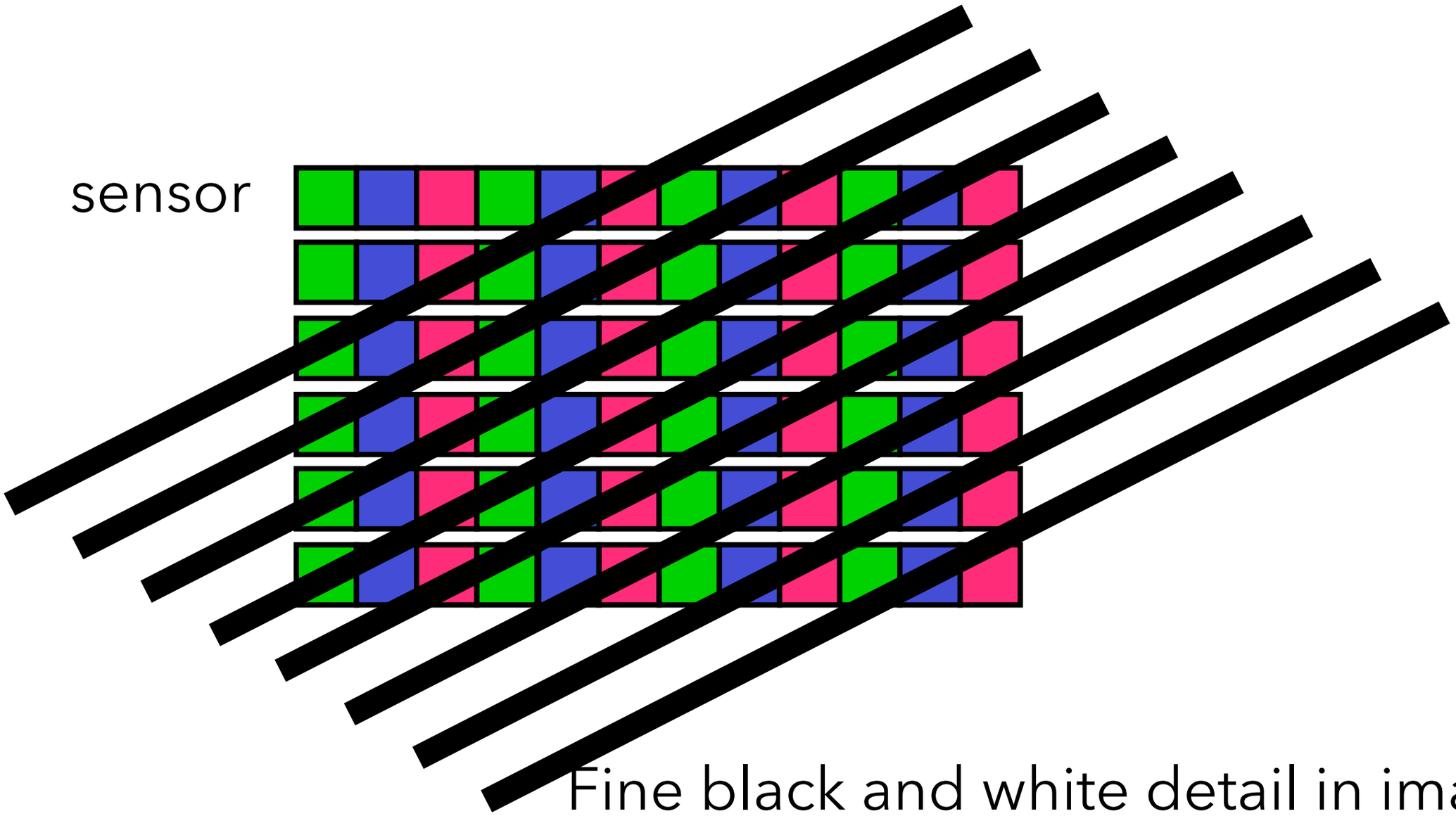
Human Luminance Sensitivity Function

Source: Steve Seitz

Problem with demosaicing: color moiré

Slide by F. Durand

The cause of color moiré



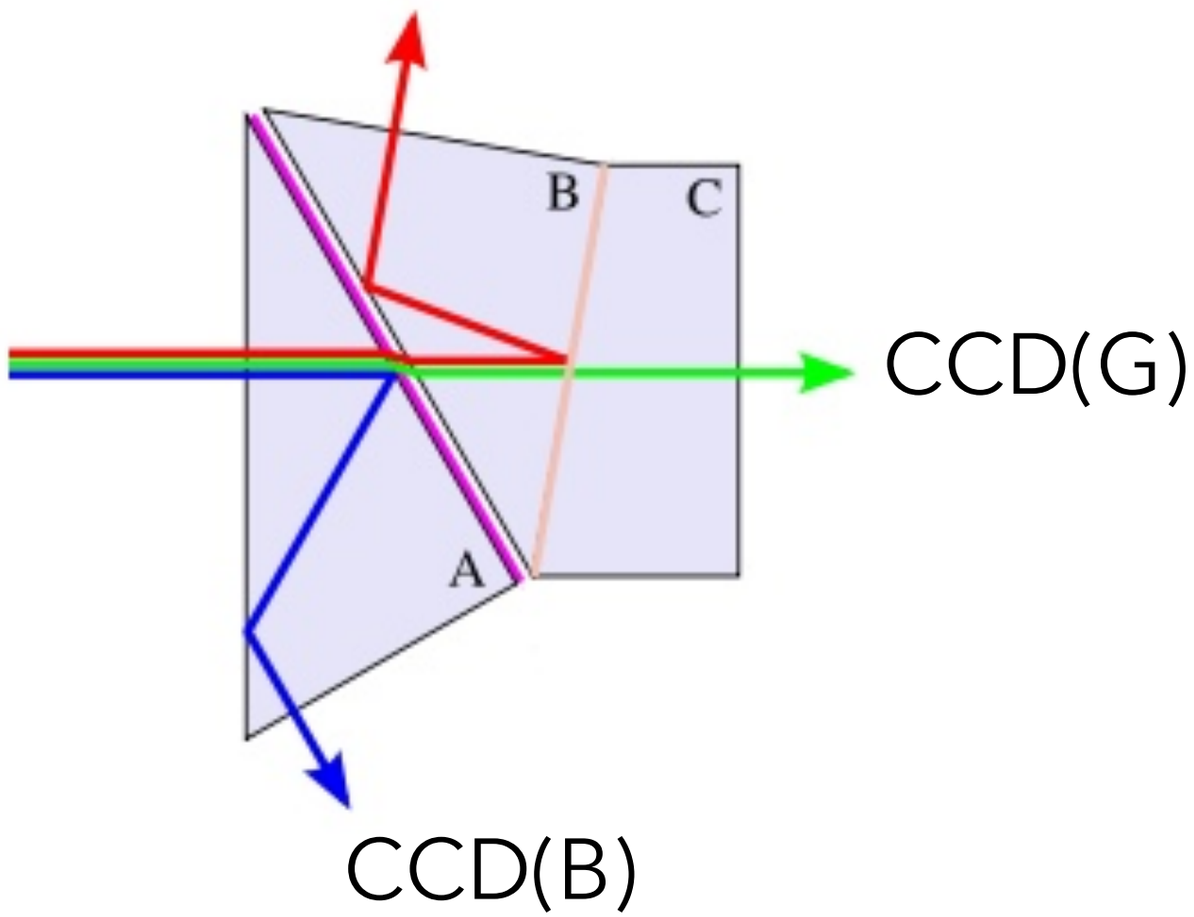
Fine black and white detail in image misinterpreted as color information

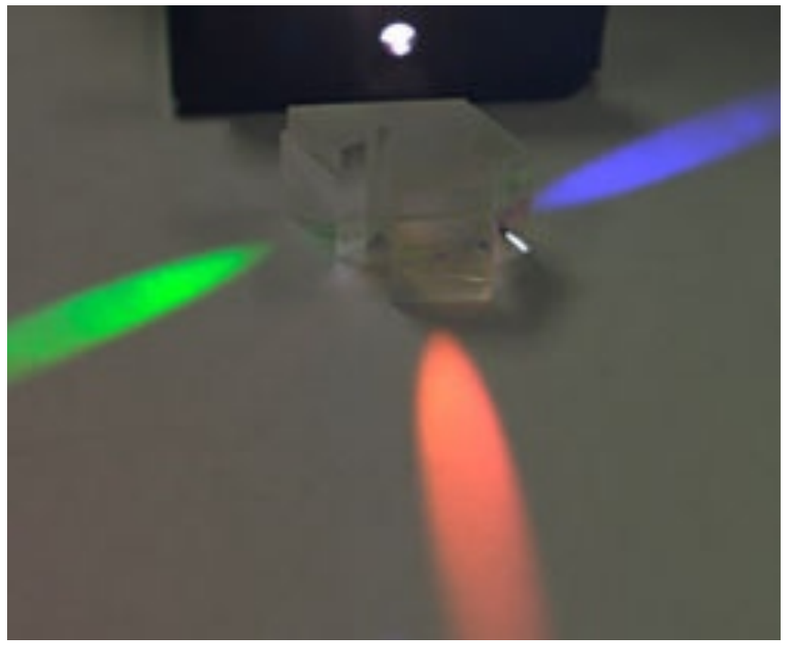
Slide by F. Durand

Color sensing in camera: Prism

TECHNISCHE UNIVERSITÄT DARMSTADT

- Requires three chips and precise alignment
- More expensive

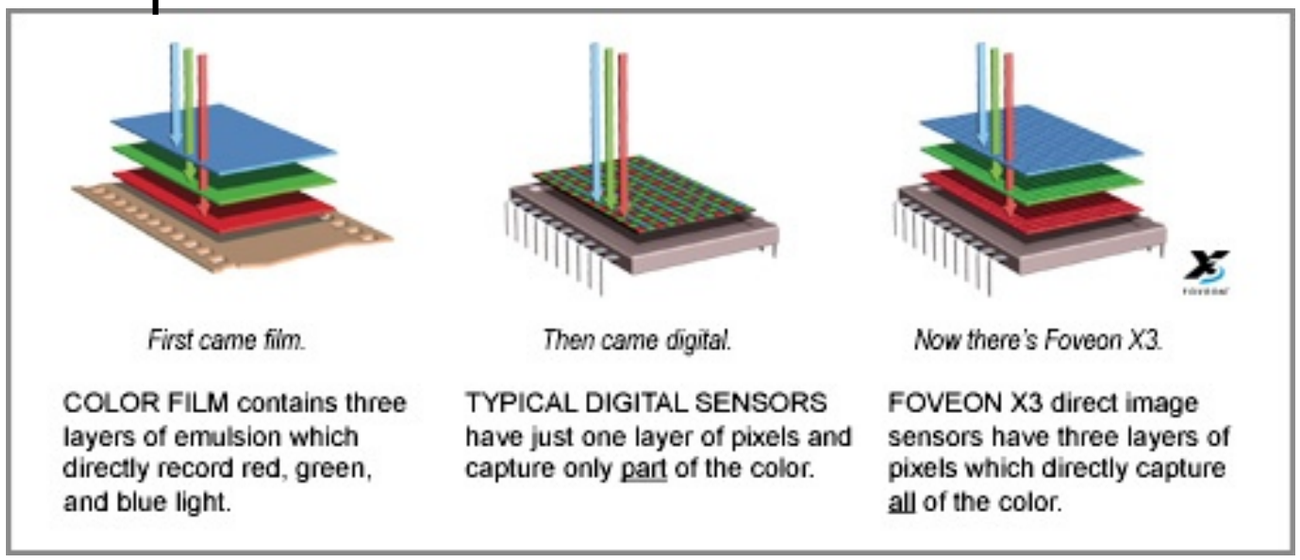


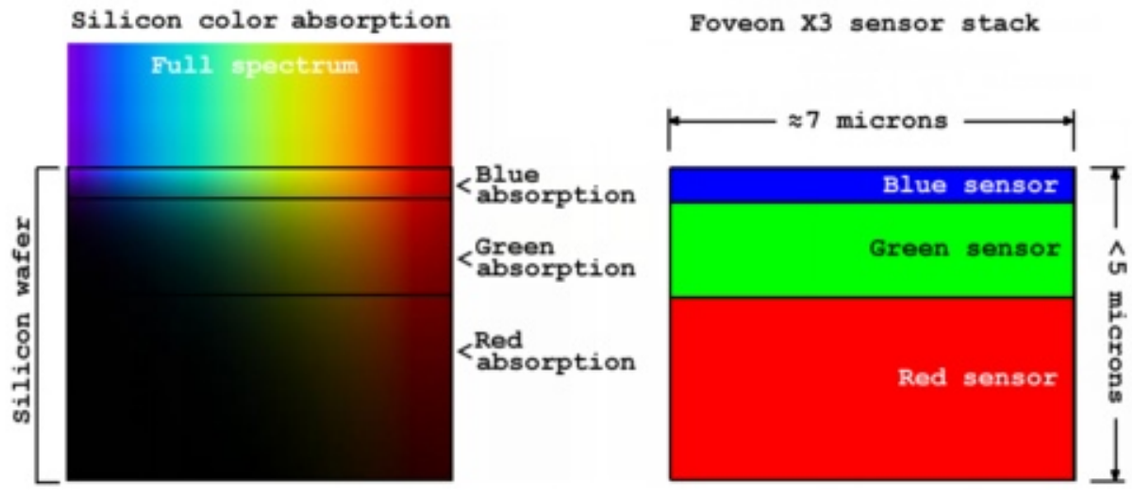


Slide by Lana Lazebnik

Color sensing in camera: Foveon X3

- CMOS sensor
- Takes advantage of the fact that red, blue and green light penetrate silicon to different depths

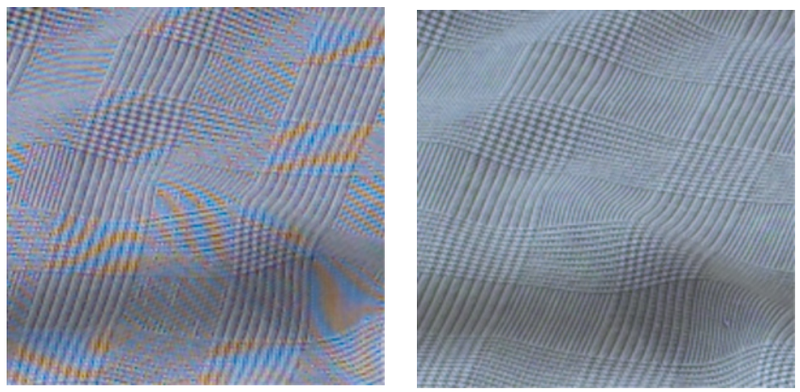




http://www.foveon.com/article.php?a=67

http://en.wikipedia.org/wiki/Foveon_X3_sensor

better image quality



Source: M. Pollefeys

Issues with digital cameras

TECHNISCHE UNIVERSITÄT DARMSTADT

- Noise
  - low light is where you most notice noise
  - light sensitivity (ISO) / noise tradeoff
  - stuck pixels

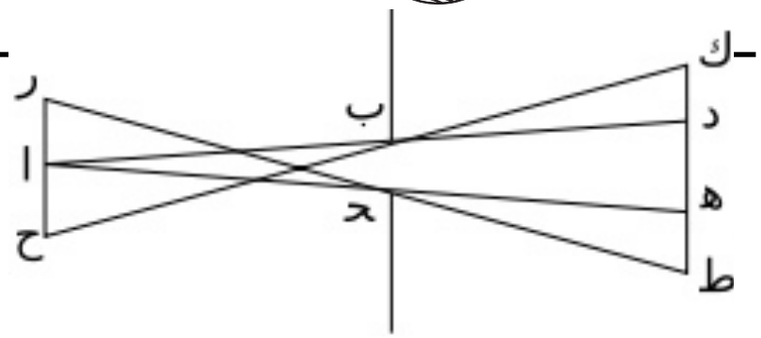
- Resolution: Are more megapixels better?
  - requires higher quality lens
  - noise issues

- In-camera processing
  - oversharpening can produce halos
- RAW vs. compressed
  - file size vs. quality tradeoff
- Blooming
  - charge overflowing into neighboring pixels
  - white balance

- More info online:
  - http://electronics.howstuffworks.com/digital-camera.htm
  - http://www.dpreview.com/
  - http://www.dxomark.com/

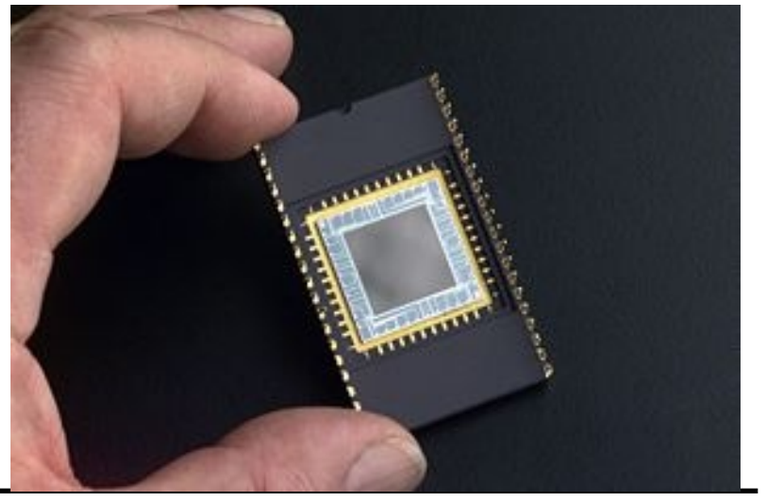
Historical context

- Pinhole model: Mozi (470-390=1 BCE), Aristoteles (384-322 BCE)
- Principles of optics (including lenses): Alhazen (965-1039 CE)
- Camera obscura: Leonardo da Vinci (1452-1519), Johann Zahn (1631-1707)
- First photo: Joseph Nicephore Niepce (1822)
- Daguerréotypes (1839)
- Photographic film (Eastman, 1889)
- Cinema (Lumière Brothers, 1895)
- Color Photography (Lumière Brothers, 1908)
- Television (Baird, Farnsworth, Zworykin, 1920s)
- First consumer camera with CCD: Sony Mavica (1981)
- First fully digital camera: Kodak DCS100 (1990)



Alhazen's notes

Niepce, "La Table Servie," 1822



CCD chip

Readings for next week

TECHNISCHE UNIVERSITÄT DARMSTADT

- Image processing (Ch. 3)
  - Point operations - 3.1
  - Linear filters - 3.2
  - Neighborhood operators - 3.3
  - Fourier transform - 3.4

Computer Vision I

Basic Image Processing - 24.04.2013

TECHNISCHE UNIVERSITÄT DARMSTADT

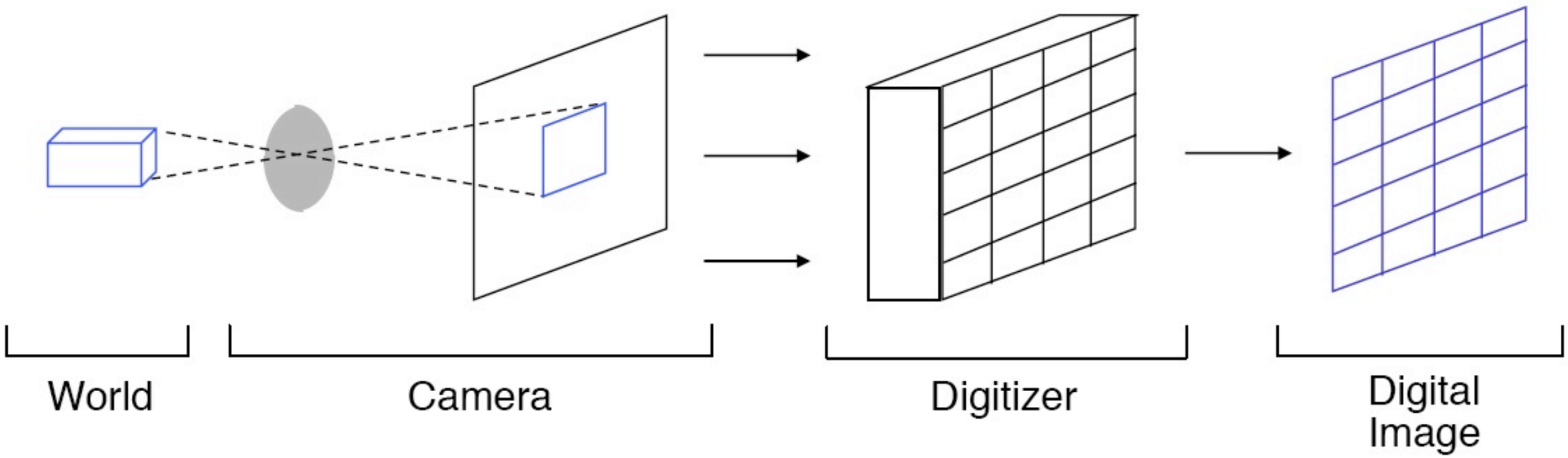
visual inference

Stefan Roth, FB Informatik, TU Darmstadt – 24.04.2013

Overview

TECHNISCHE UNIVERSITÄT DARMSTADT

- Computer Vision: “invert” the imaging process
  - Today: 2D (2-dimensional) digital image processing
  - Later: ‘pattern recognition’ / 3D image analysis
  - Later: image understanding



Digital Image Processing

TECHNISCHE UNIVERSITÄT DARMSTADT

- Some Basics
  - (digital signal processing, FFT, …)
  - Image Filtering

Credits: slides adapted from
Bernt Schiele
Michael Black

- Image Filtering
  - to reduce noise
  - to fill-in missing values/information
  - to extract image features (e.g. edges/corners)
  - …

10	5	3
4	5	1
1	1	7

Local image data

Some function →

	7	

Modified image data

Today - Basics of Digital Image Processing

TECHNISCHE UNIVERSITÄT DARMSTADT

- Images
- Filtering
  - Linear filtering
  - Non-linear filtering & morphology
- Multi-scale image representation
  - Gaussian pyramid
  - Laplacian pyramid
- Edge detection
  - ‘Recognition using line drawings’
  - Image derivatives (1st and 2nd order)

Linear Operations / Systems

TECHNISCHE UNIVERSITÄT DARMSTADT

- Basic properties:
  - homogeneity $T[a \cdot X] = a \cdot T[X]$
  - additivity $T[X + Y] = T[X] + T[Y]$
  - superposition $T[a \cdot X + b \cdot Y] = a \cdot T[X] + b \cdot T[Y]$
- Examples:
  - matrix-vector operations
  - convolutions

Convolution

TECHNISCHE UNIVERSITÄT DARMSTADT

- Replace each pixel by a linear combination of its neighbors (and itself).
- 2D convolution (discrete):

$$g = f * h \qquad g(i,j) = \sum_{k,l} f(i-k, j-l)h(k,l) = \sum_{k,l} f(k,l)h(i-k, j-l),$$

45	60	98	127	132	133	137	133
46	65	98	123	126	128	131	133
47	65	96	115	119	123	135	137
47	63	91	107	113	122	138	134
50	59	80	97	110	123	133	134
49	53	68	83	97	113	128	133
50	50	58	70	84	102	116	126
50	50	52	58	69	86	101	120

*

0.1	0.1	0.1
0.1	0.2	0.1
0.1	0.1	0.1

=

69	95	116	125	129	132
68	92	110	120	126	132
66	86	104	114	124	132
62	78	94	108	120	129
57	69	83	98	112	124
53	60	71	85	100	114

smaller output?

image $f(i,j)$ | filter (kernel) $h(i,j)$ | filtered image $g(i,j)$

Convolution / Correlation

TECHNISCHE UNIVERSITÄT DARMSTADT

- Convolution is
  - linear $h * (f_0 + f_1) = h * f_0 + h * f_1$
  - associative $(f * g) * h = f * (g * h)$
  - commutative $f * h = h * f$
  - shift-invariant $g(i,j) = f(i+k, j+l) \Leftrightarrow (h * g)(i,j) = (h * f)(i+k, j+l)$
  - can be represented as matrix-vector product $\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f}$
  - Continuous versions exist, skip them here...
- Correlation:
  - Closely related, but do not mirror filter

$$g = f \otimes h \qquad g(i,j) = \sum_{k,l} f(i+k, j+l) h(k,l)$$

Filtering to Reduce Noise

TECHNISCHE UNIVERSITÄT DARMSTADT

- "Noise" is what we are not interested in
  - low-level noise: light fluctuations, sensor noise, quantization effects, finite precision, …
  - complex noise (not today): shadows, extraneous objects.

- Assumption:
  - The pixel's neighborhood contains information about its intensity

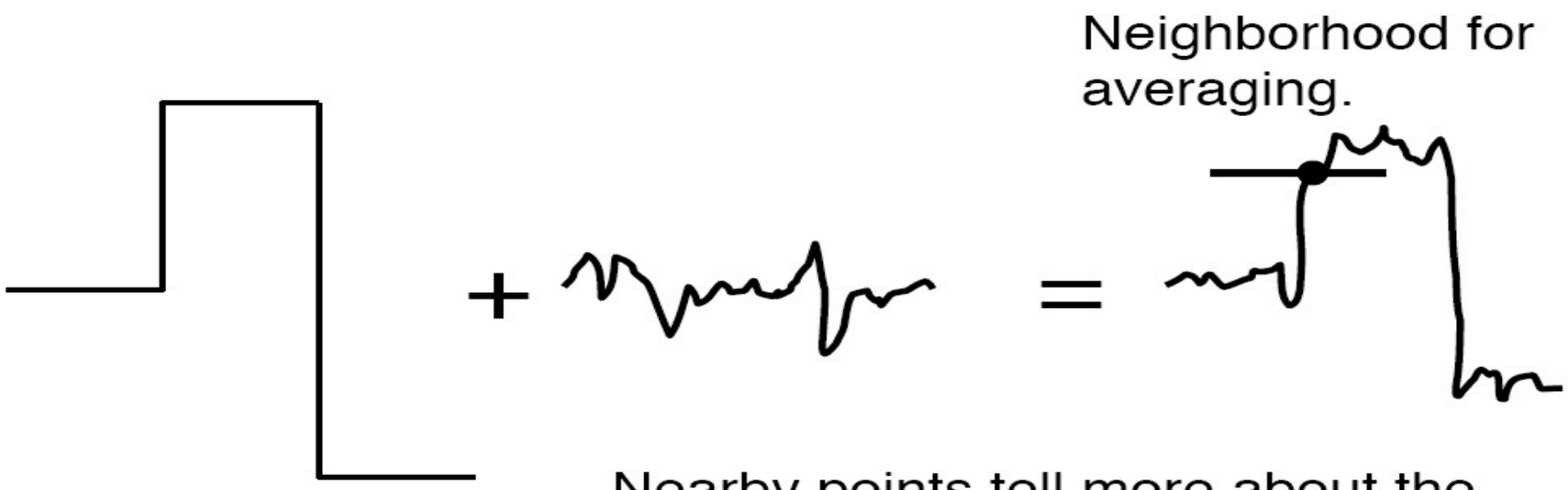
$$\begin{array}{|c|c|c|} \hline 2 & 3 & 3 \\ \hline 3 & 20 & 2 \\ \hline 3 & 2 & 3 \\ \hline \end{array} \rightarrow \begin{array}{|c|c|c|} \hline 2 & 3 & 3 \\ \hline 3 & 3 & 2 \\ \hline 3 & 2 & 3 \\ \hline \end{array}$$

Model: Additive Noise

TECHNISCHE UNIVERSITÄT DARMSTADT

Signal + Noise = Image

$$S + N = I$$



Average Filter

TECHNISCHE UNIVERSITÄT DARMSTADT

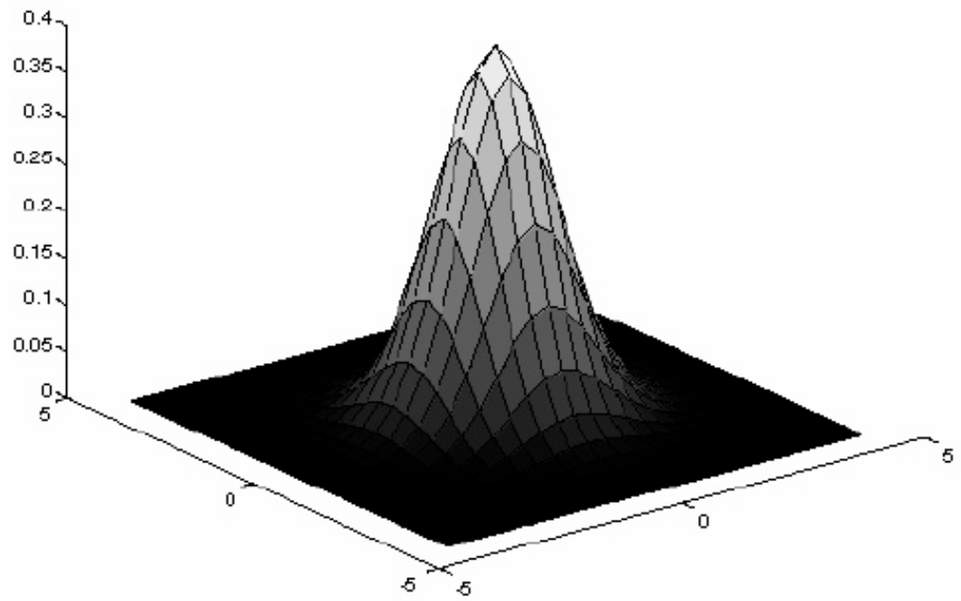
- Average Filter
  - replaces each pixel with an average of its neighborhood
  - Mask with positive entries that sum to 1
- If all weights are equal, it is called a box filter

$$\frac{1}{9} \cdot \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$$

Gaussian Averaging

TECHNISCHE UNIVERSITÄT DARMSTADT

- Isotropic Gaussian (rotationally symmetric)
- Weighs nearby pixels more than distant ones



- Smoothing kernel proportional to

$$G_\sigma(x,y) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$$