DTU

Daniel Olesen, DTU Space

30540 – Photogrammetry (1)

Lecture Plan - Photogrammetry

Lecture	Topics
Module 1 (3) 13/09	- Introduction - Camera and lens basics - Projective Geometry (2D)
Module 2 (9) 01/11	- Coordinate Systems (World and Camera) - Pinhole-Camera model and calibration - Spatial Resection (PnP)
Module 3 (10) 08/11	- Epipolar Geometry - Essential and Fundamental Matrix - Stereo-rectification - Triangulation
Module 4 (11) 15/11	- Feature Detection - RANSAC
Module 5 (12) 22/11	- Optimization and Bundle-Adjustment
Module 6 (13) 29/11	- Summary, applications and outlook

Today's lecture

- Photogrammetry
  - Definition and History
  - Types of photogrammetry
  - Products
    - Orthophotos
    - Digital Surface Model (DSM)
- Camera-basics
  - The Pinhole camera
  - Elements of a Modern Camera
  - Aberrations and distortions

- Homogeneous Coordinates and 2D projective transformations

Photogrammetry

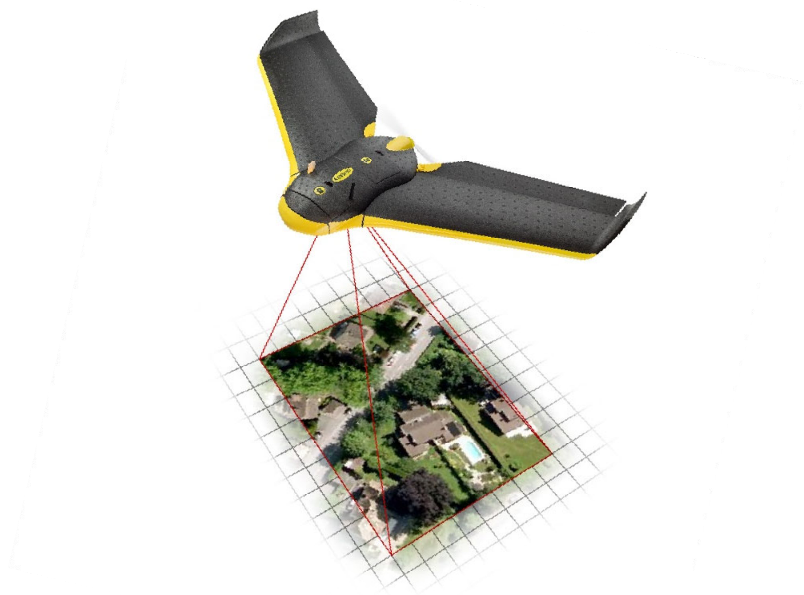
- The art and science of determining position and shape of objects from photographs
- The process of reconstructing objects without touching them

- Objectives
  – Invert the process of photographing (2D -> 3D)
  – Reconstruct the object space from imagery

- Typical Applications
  – Photomosaic
  – Orthophoto / Maps
  – Digital Surface Models (DSMs)
  – 3D models

Types of Photogrammetry

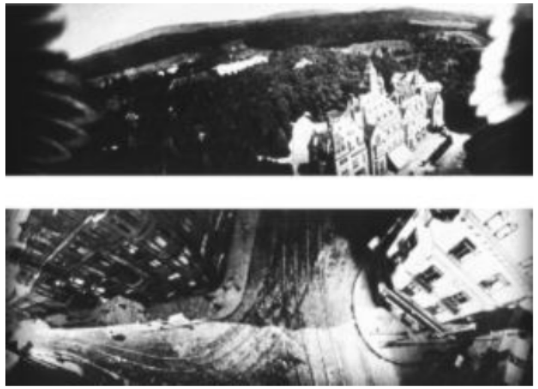
- Aerial Photogrammetry
  - Camera is mounted on an airplane or UAV
  - Camera is usually pointed nadir (vertically towards the ground)
  - Often used for generating maps, orthophotograps, Digital Surface Models…
- Terrestial and close-range photogrammetry
  - Used for generating 3D models and measurements
  - Camera is hand-held or mounted on tripod



https://bitfab.io/blog/photogrammetry/

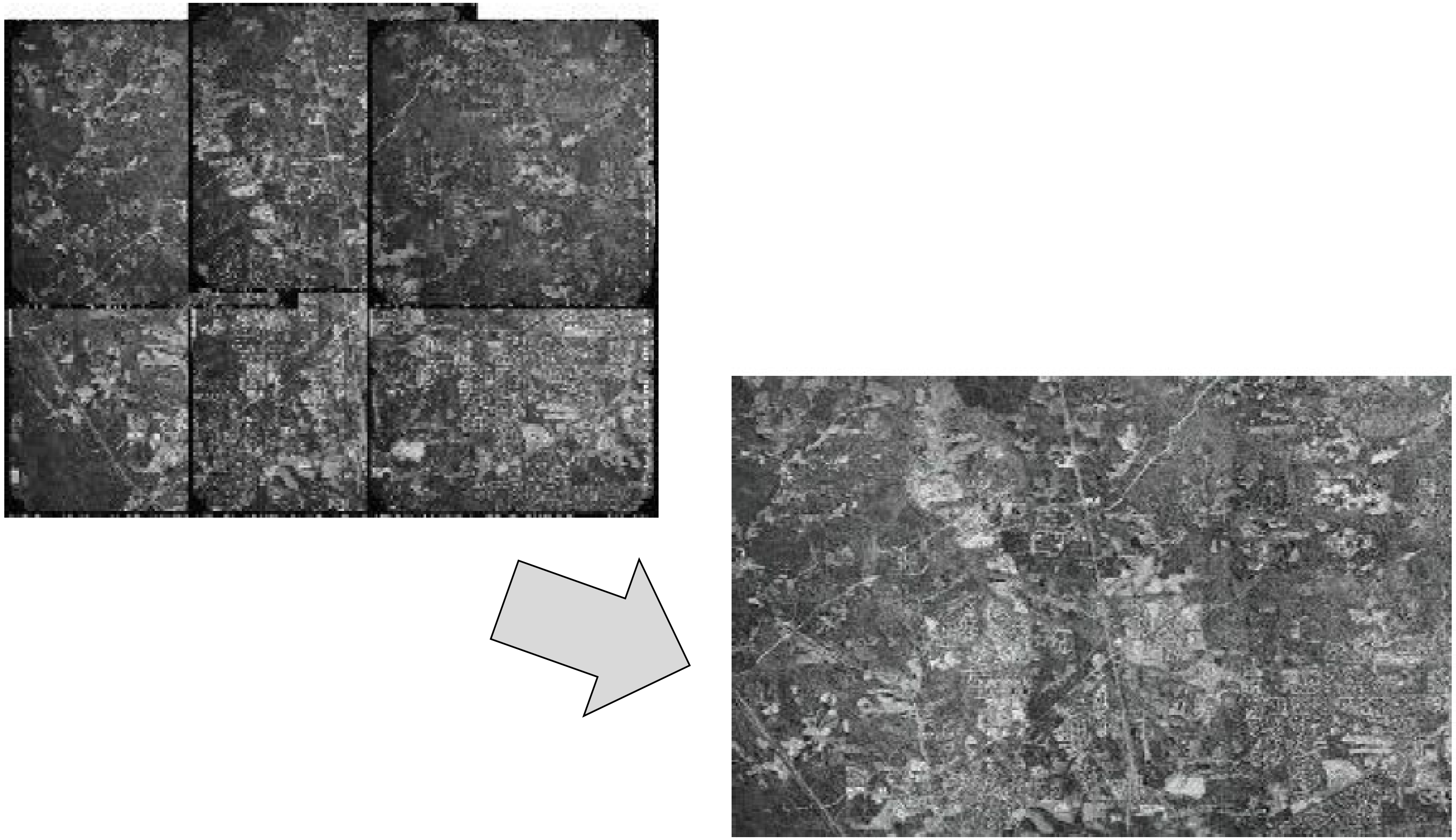
Historical perspective

Laussedat (1862)

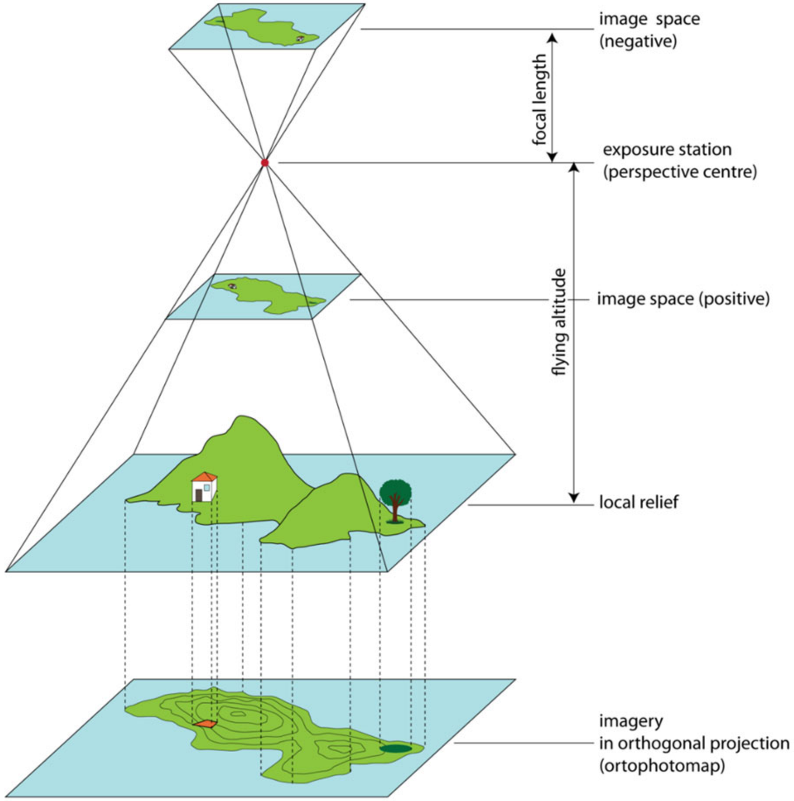


Bavarian Pigeon Corp (1903)

Photogrammetric Products: Photomosaics



Photogrammetric Products: Orthophotos



https://ncsu-geoforall-lab.github.io/uav-lidar-analytics-course/lectures/HM_Photogrammetry_and_SfM.html

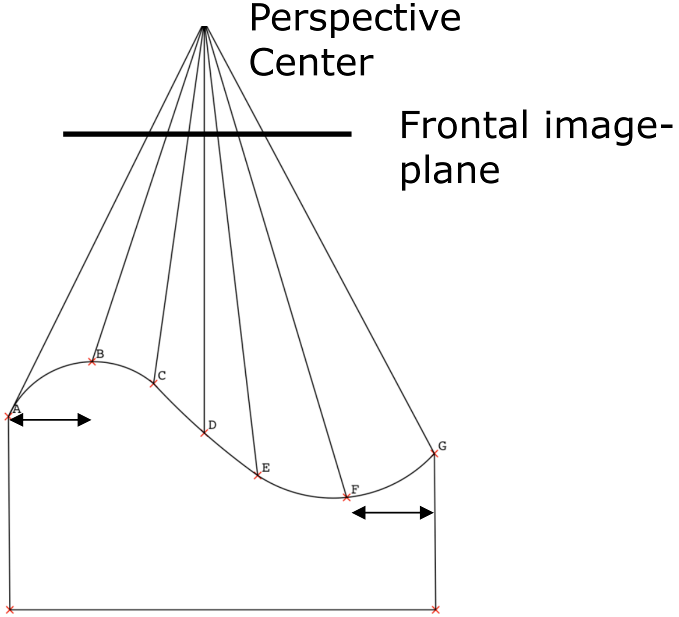
An Orthophoto is a rectified aerial image, which has a uniform horizontal scale and hence can be be used as a ”map”

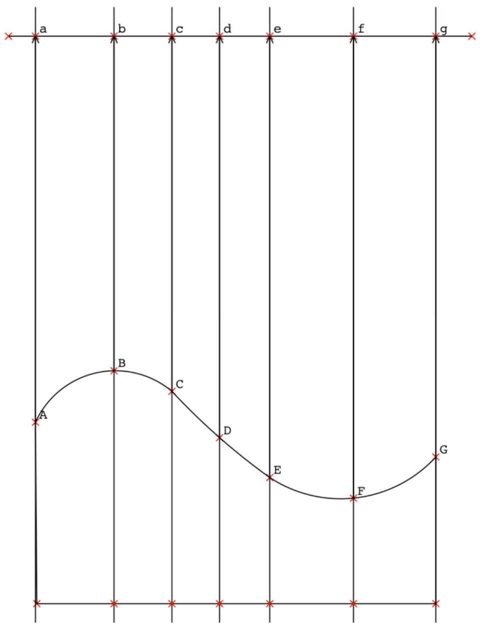
An orthographic projection requires that 3D information of objects, buildings and terrain is known in order to ensure a uniform **horizontal scale**.

Perspective vs Orthographic Projection (Image versus Map)

Images
- Perspective Projection
- Non-uniform scale

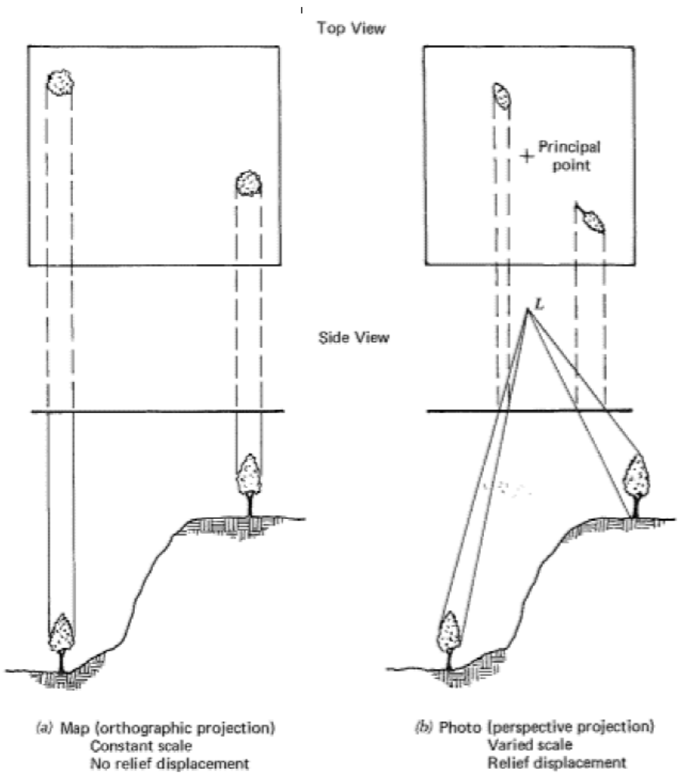
Maps / Orthographic Projections
- Orthographic (parallel projections)
- Uniform scale





Horizontal distance between A&B same as F&G

Aerial Photos: Relief Displacement



For aerial images we encounter Relief displacement, which makes objects of certain heights appear tilted in the photographs.

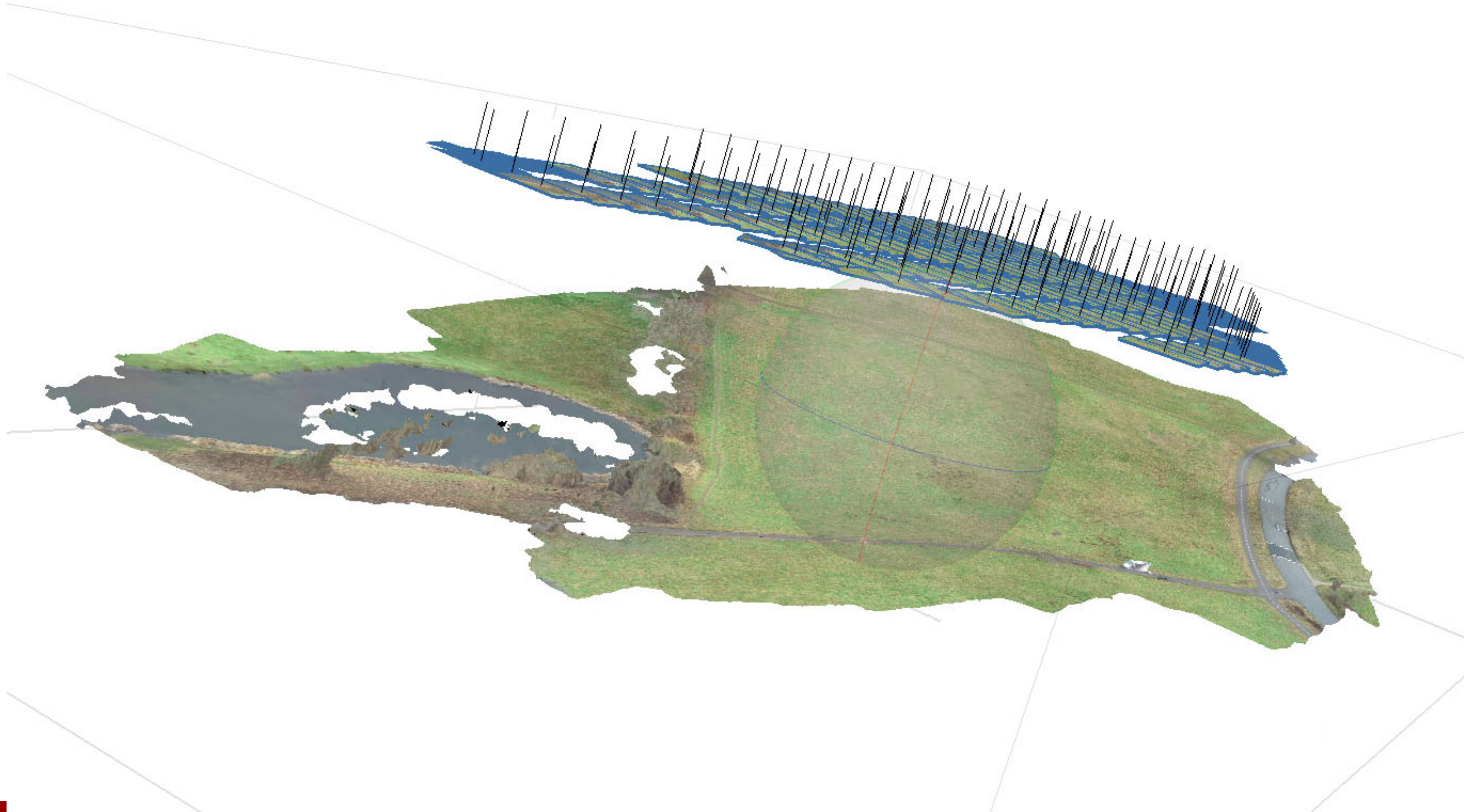
It depends on the height of the object and is more pronounced near edges of the photograph

For a true orthophoto this effect is also mitigated

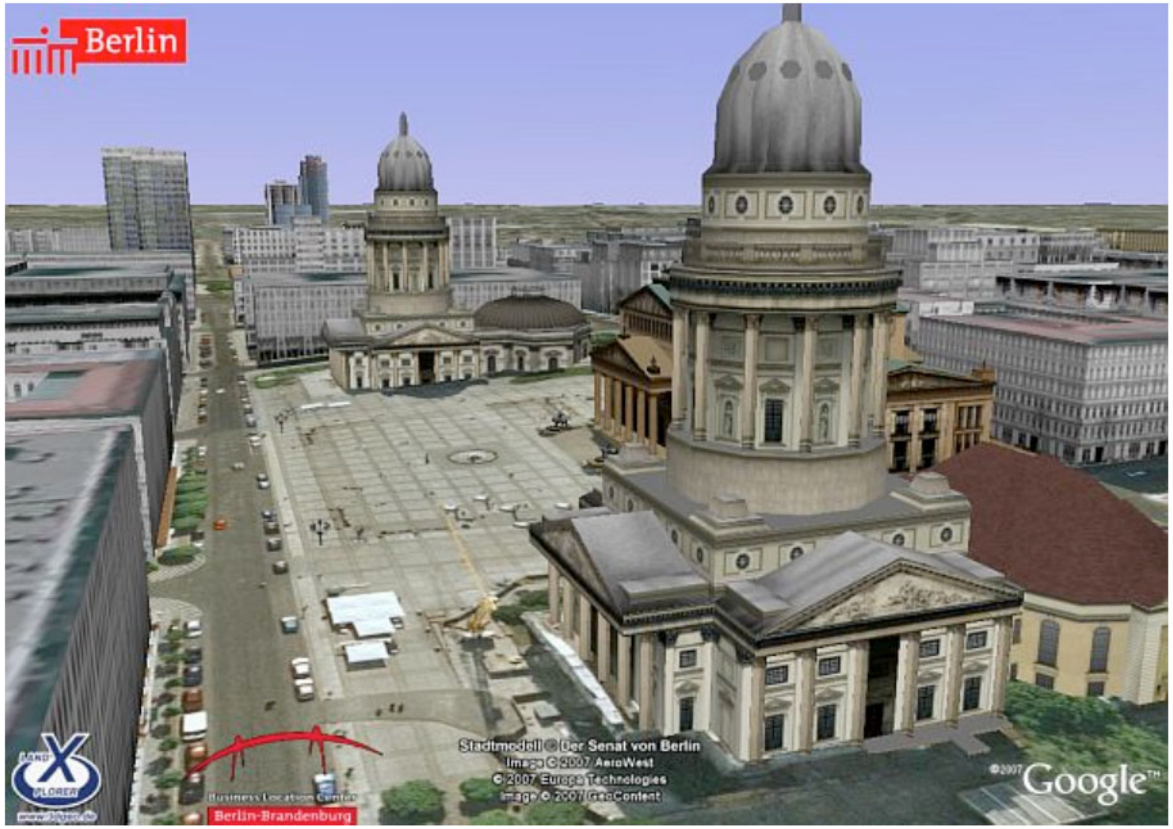
Orthophoto

Source: Fugro EarthData

Photogrammetric Products: Digital Surface Model (Agisoft Metashape)



Photogrammetric Products: 3D Citymodels



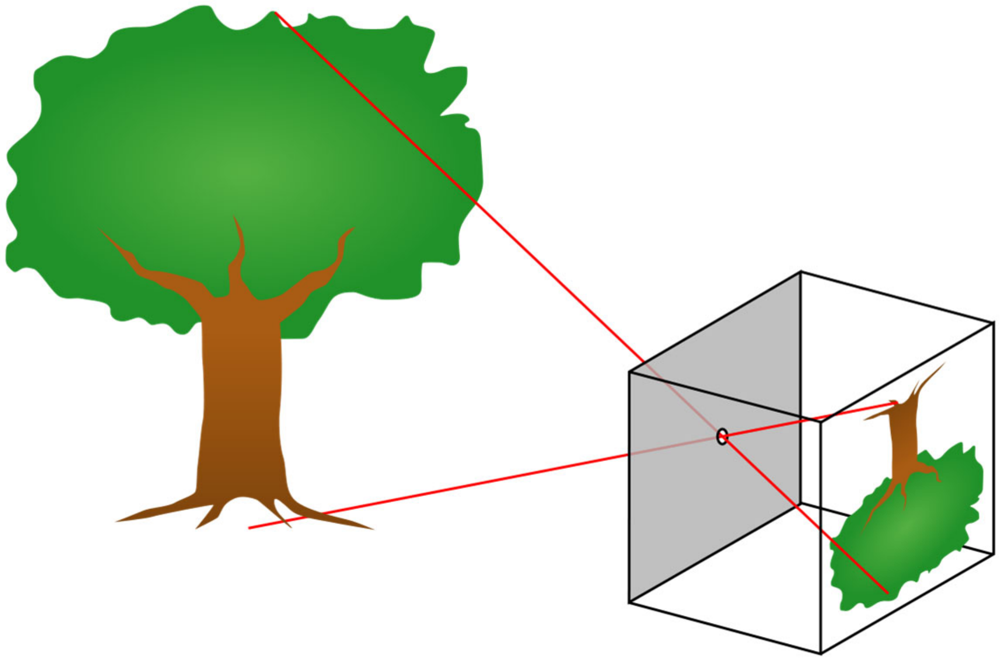
Today's lecture

- ✓ Photogrammetry
  - ✓ Definition and History
  - ✓ Types of photogrammetry
  - ✓ Products
    - ✓ Orthophotos
    - ✓ Digital Surface Models
- Camera-basics
  - The Pinhole camera
  - Elements of a Modern Camera
  - Aberrations and distortions

- Homogeneous Coordinates and 2D projective transformations

The Pinhole Camera

- Principle first discovered in 1544 by Reinerus Gemma-Frisius, who observed an projected “image” of a solar eclipse (Camera-Obscura)
- All rays from different object points has to pass through the pinhole.
- An image is formed upside down on the back wall of the camera-box

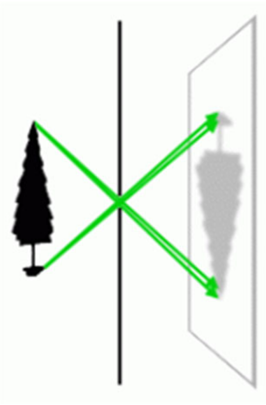


The Pinhole Camera

- The main **disadvantage** of a true pinhole camera is that only a tiny fraction of lightrays propogate to the film/sensor. In order to capture ”enough” light the shutter has to remain open for a long time.

- A larger opening would allow more light (in a given timeframe) to hit the sensor, but diverging rays from the same point source would result in image blur.

- Modern cameras relies on lenses, which can focus the light rays and hence provide a bigger opening



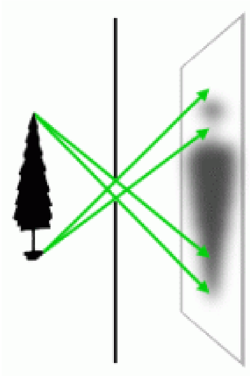
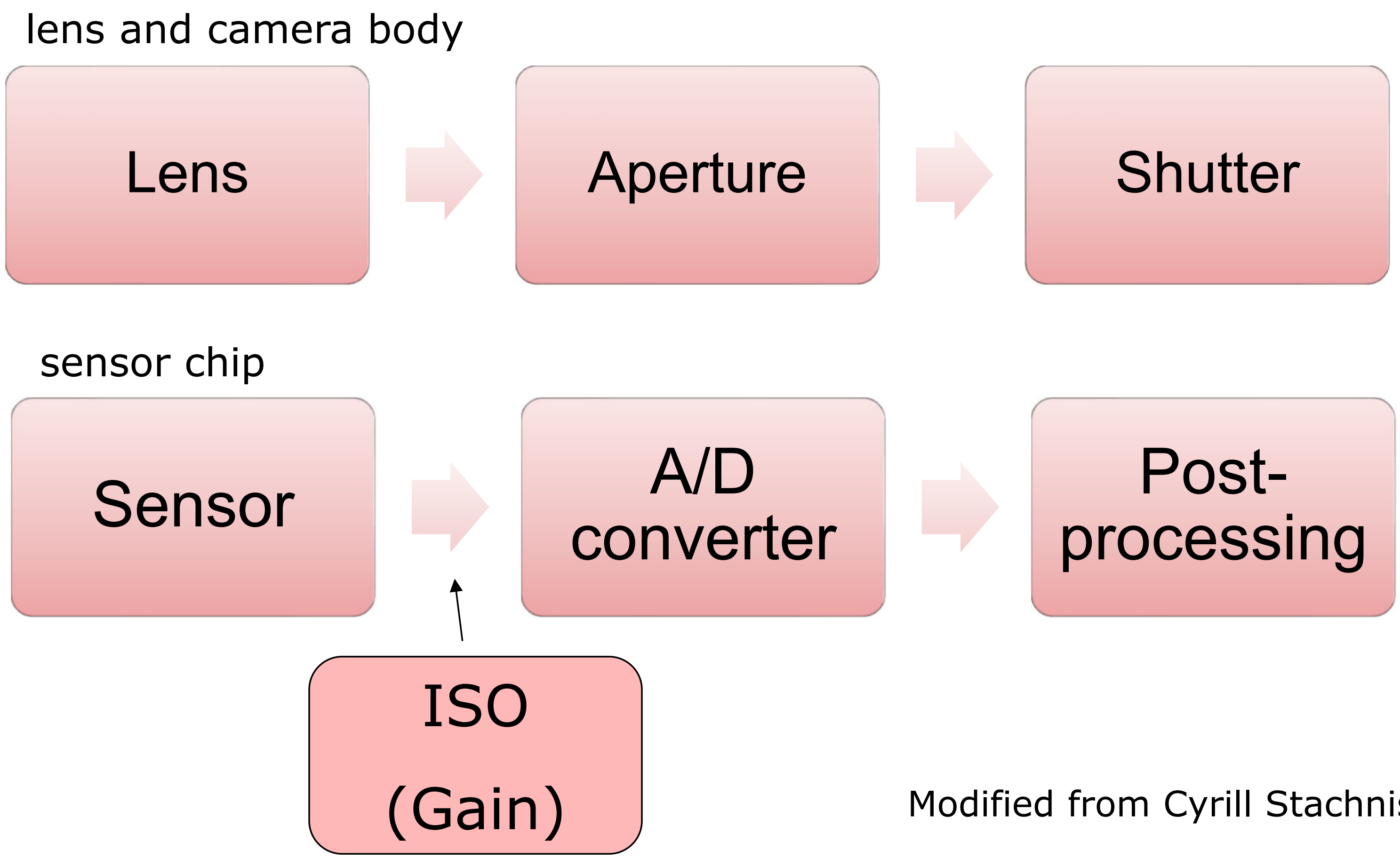


Image Courtesy: https://digital-photography-school.com/photography-101-light-and-the-pinhole-camera/

Elements of a modern (Digital) Camera



Modified from Cyrill Stachniss

Basics of a digital camera – How are images formed?

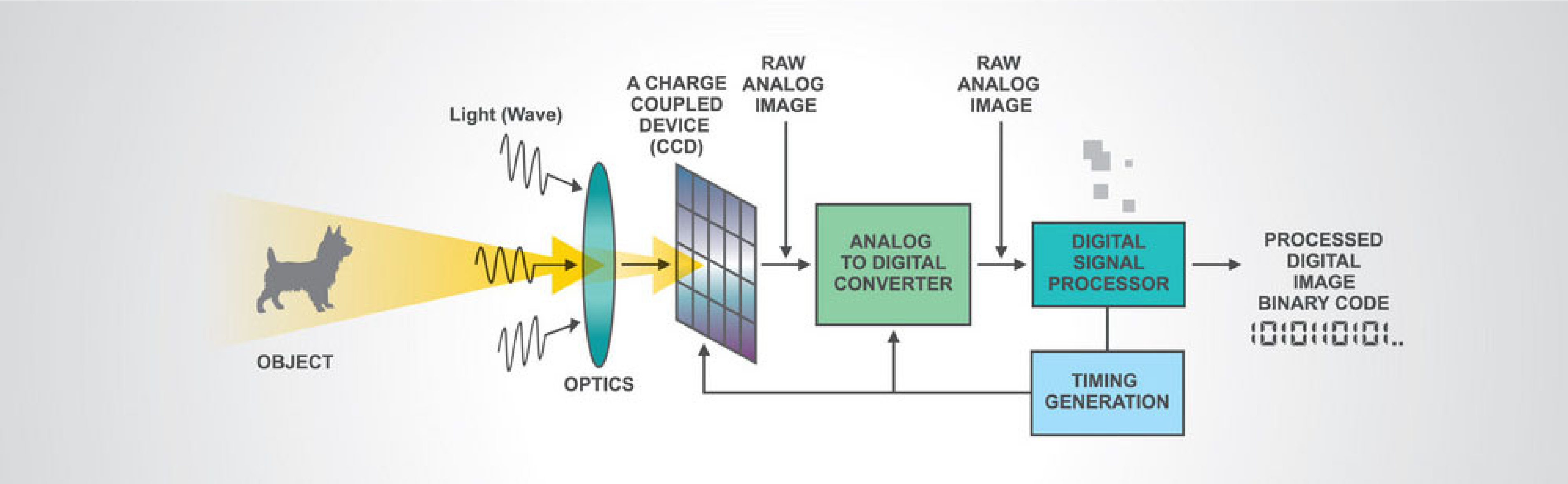
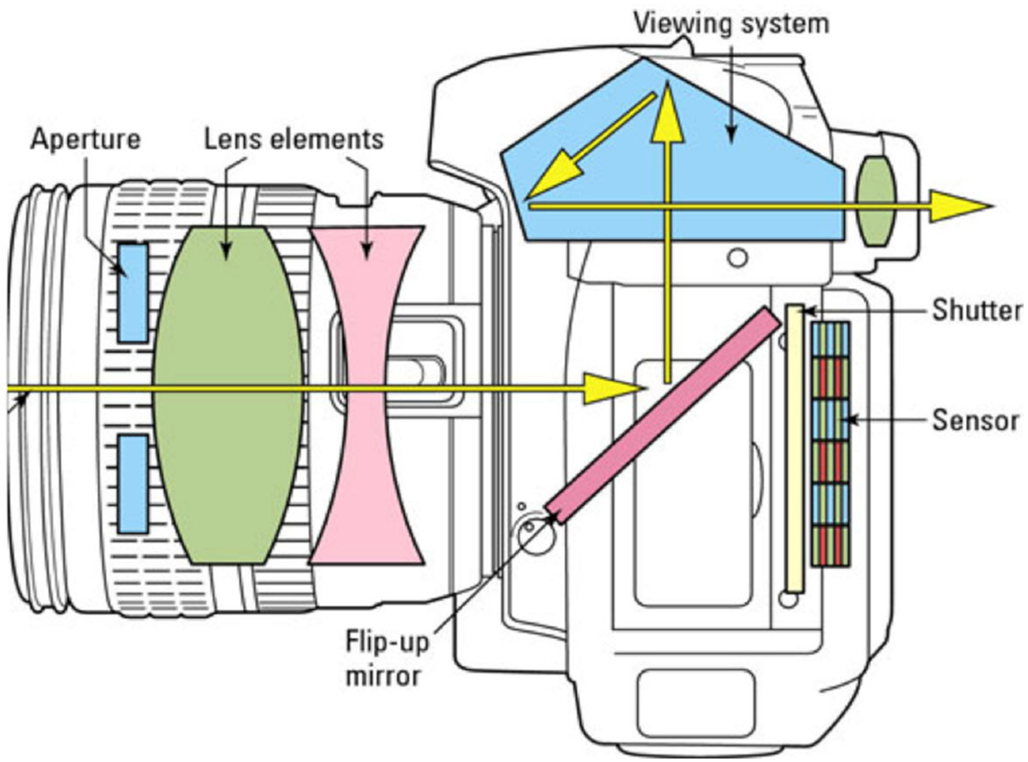


Illustration from: https://www.vectorstock.com/royalty-free-vector/digital-camera-working-with-diagrams-vector-15452894

A CCD sensor converts photons to electrons (charge)

Basics of a digital camera – How are images formed?



http://howthingswork.org/electronics-how-digital-camera-works/

Lens

- The Lens of a modern camera tries to ”gather” as much light as possible without affecting (losing) focus

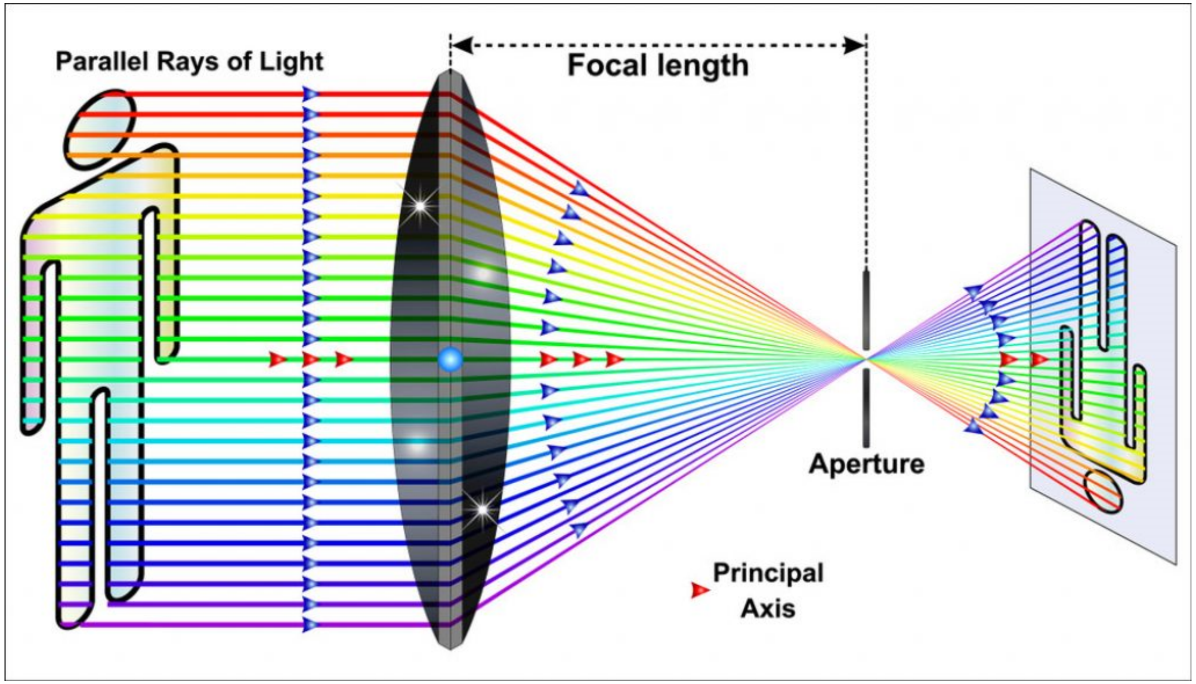


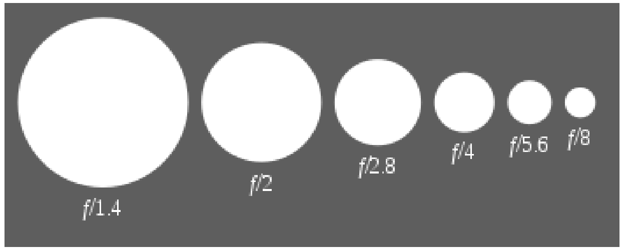
Image Sensor

https://thegadgetnerds.com/how-does-a-camera-work/

Aperture

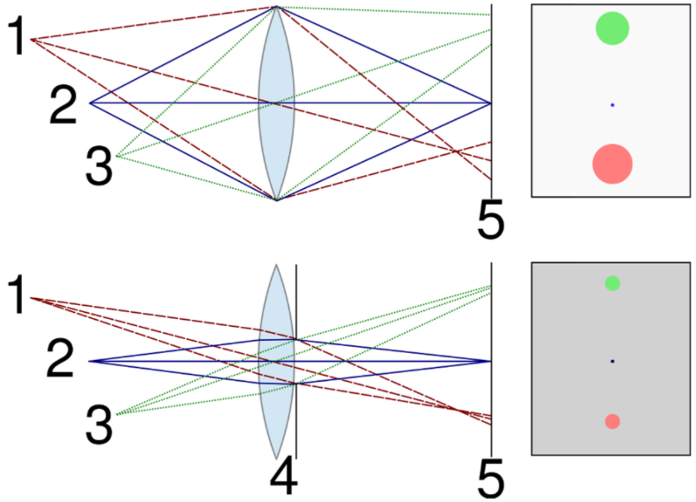
The aperture determined how big an opening you have between the lens and the image-sensor

f-number = focal-length / aperture diameter



https://en.wikipedia.org/wiki/Aperture

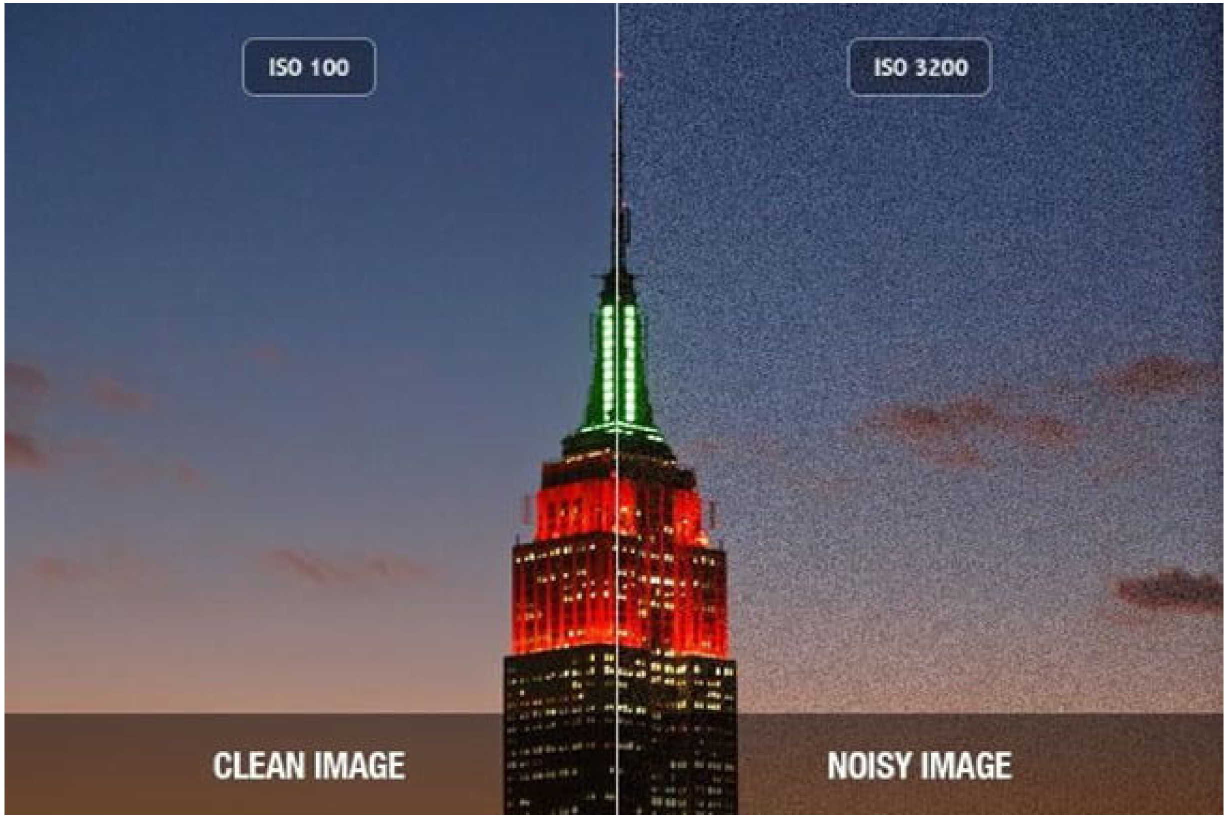
Depth Of Field



https://en.wikipedia.org/wiki/Depth_of_field

Aperture = f/1.4. DOF=0.8 cm Aperture = f/4.0. DOF=2.2 cm Aperture = f/22. DOF=12.4 cm

ISO sensitivity



https://www.digitaltrends.com/photography/what-is-iso/

Exposure Triangle

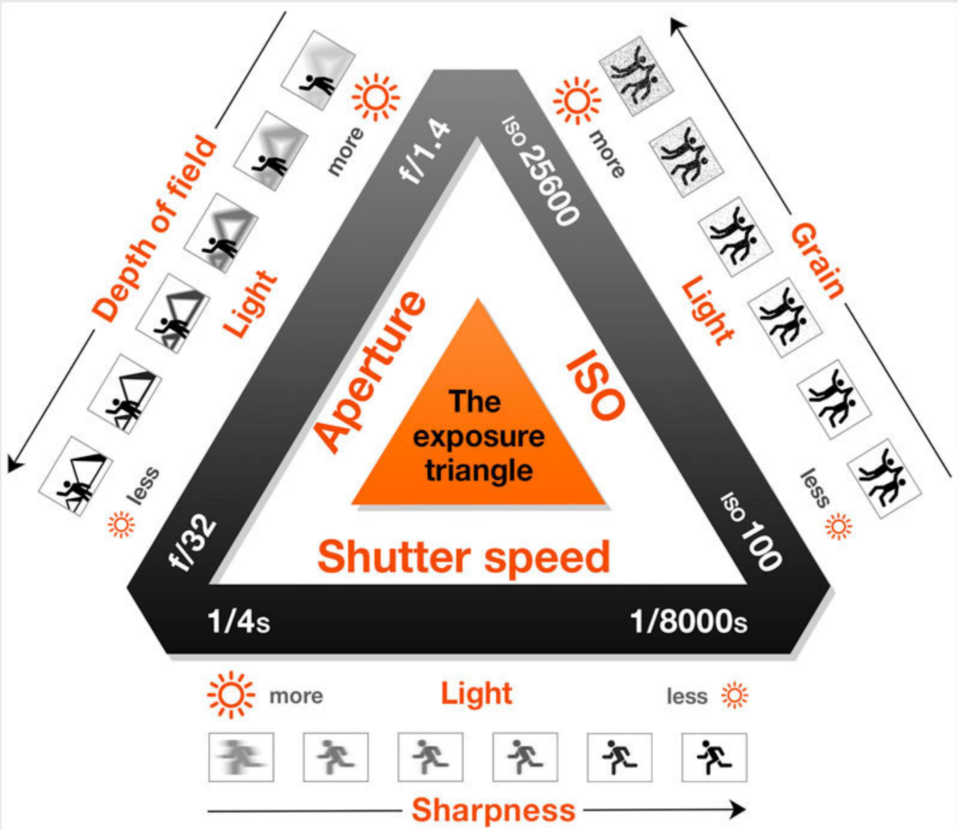
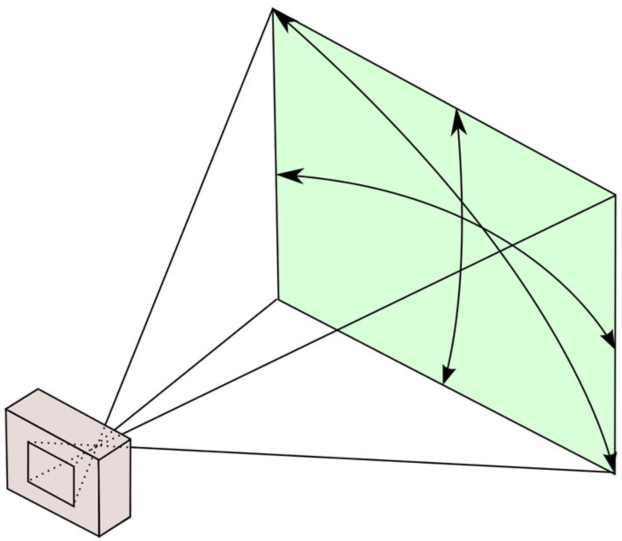


Image Courtesy: https://petapixel.com/2017/03/25/exposure-triangle-making-sense-aperture-shutter-speed-iso/

Field-Of-View (FOV)



The field (or angle) of view of a camera is determined by the physical sensor area and the focal-length of the lens.

A long focal-length would imply a narrow field-of-view

$$\theta = 2\arctan\left(\frac{l/2}{f}\right)$$

Where, l, is the sensor-diagonal and f is the lens focal-length

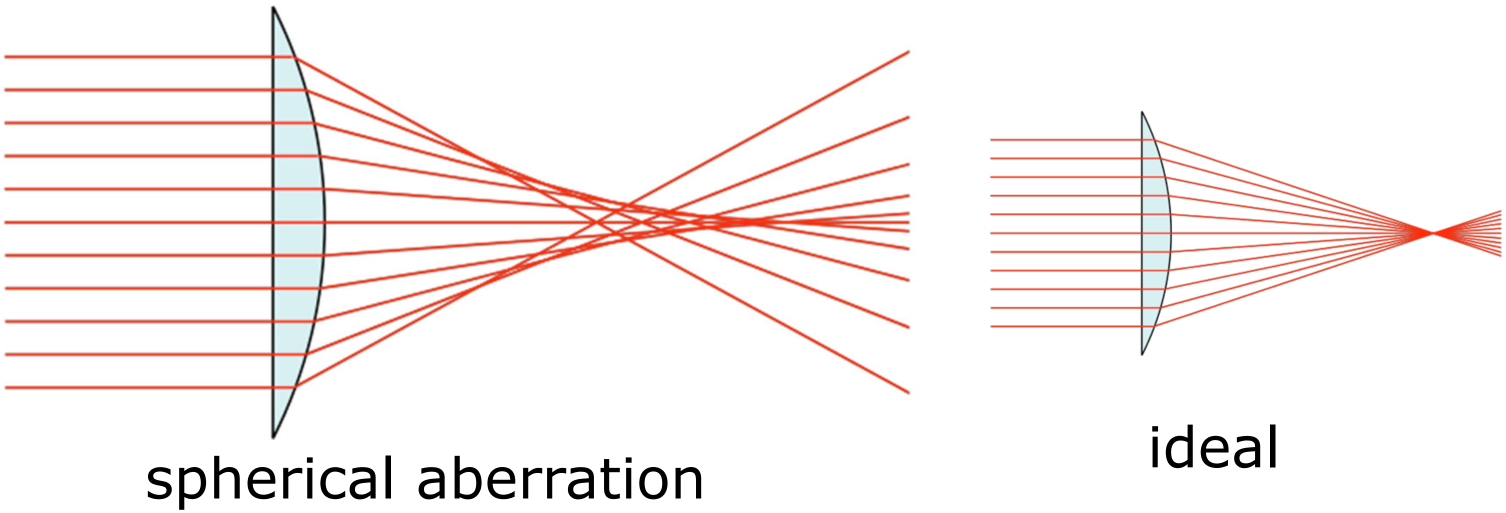
https://en.wikipedia.org/wiki/Angle_of_view#Measuring_a_camera's_field_of_view

Lens Distortion and Abberations

- Spherical
- Chromatic
- Vignetting
- Radial distortion

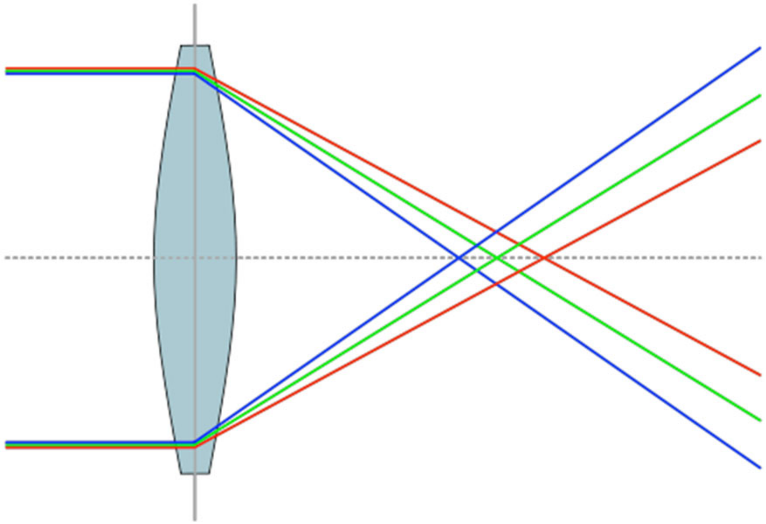
Spherical Abberation

Spherical abberation is caused by non-idealites in the lens-shape, so collimated light (parallel rays) is focussed in different points



https://en.wikipedia.org/wiki/Spherical_aberration

Chromatic Aberrations



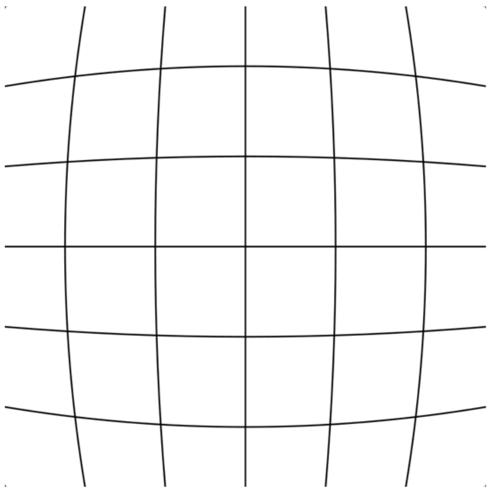
Chromatic aberrations is caused by wave-length dependent refractions in the lens element(s)

https://en.wikipedia.org/wiki/Chromatic_aberration

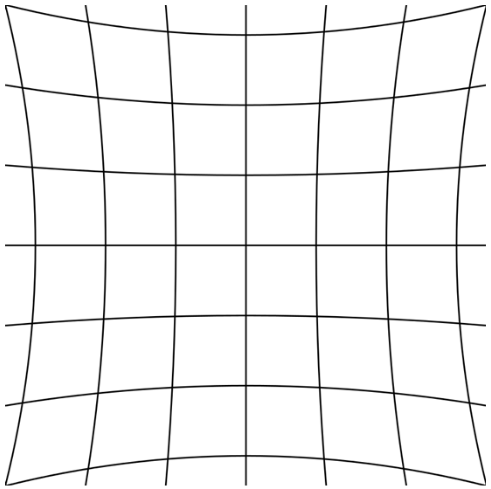
Vignetting

Stronger intensity in the center of the image.

Radial Distortion



Barrel Distortion



Pincushion Distortion

Typical Distortion Model:

$$r^2 = x^2 + y^2$$

$$x_d = x(1 + a_1 r^2 + a_2 r^4 + a_3 r^6)$$

$$y_d = y(1 + a_1 r^2 + a_2 r^4 + a_3 r^6)$$

Radial Distortion

Barrel-distorted image (left) and corrected (right)

From: https://github.com/bbenligiray/lens-distortion-rectification

Today’s lecture

- ✓ Photogrammetry
  - ✓ Definition and History
  - ✓ Types of photogrammetry
  - ✓ Products
    - ✓ Orthophotos
    - ✓ Digital Surface Models
- ✓ Camera-basics
  - ✓ Camera-Obscura and the Pinhole camera
  - ✓ Elements of a Camera
  - ✓ Aberrations and distortions

- Homogeneous Coordinates and 2D projective transformations

Hands-on Exercise (10 min + 5 mins break)

- Use your mobile-phone cameras and practise the fine art of taking selfies
  - Specifically try to take a very-close up selfie and also take a selfie where your head is as far from the camera as possible
  - Do your selfies in an area with a non-uniform background – for instance outside or in front of a wall with multiple posters or photographs

- After you have taken your selfies reflect on the outcome
  - Is your face proportions similar for the two selfies?
  - If you crop the far-away selfie so that it has roughly the same size as the other one, do you see any difference in the background of the images?

Pinhole Camera Model

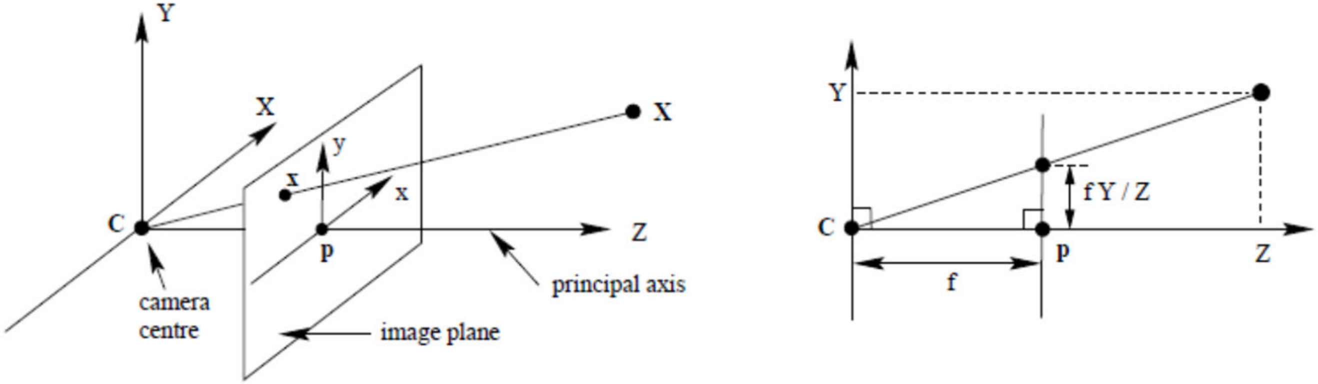


Fig. 6.1. **Pinhole camera geometry.** C *is the camera centre and* p *the principal point. The camera centre is here placed at the coordinate origin. Note the image plane is placed in front of the camera centre.*

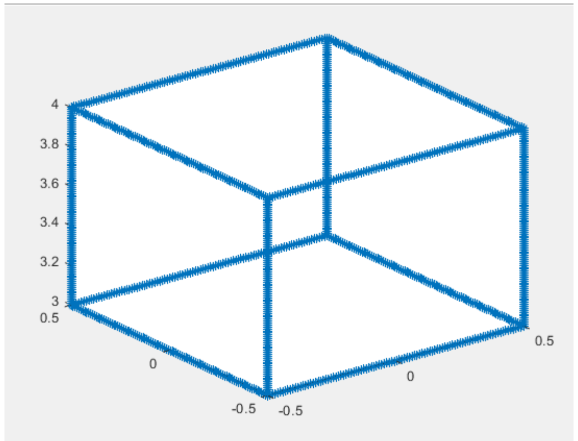
$$\boxed{(X, Y, Z)^{\mathsf{T}} \mapsto (fX/Z, fY/Z)^{\mathsf{T}}}$$

Source: Hartley & Zisserman: Multiple View Geometry

MATLAB Exercise (10 min)

- Go to DTU Learn and download Box3D.m, which is a matlab function that generates outer points of a (1 x 1 x 1) meter 3D box

- Call this function as X=Box3D(Xoff,Yoff Zoff); where you can use the initial values of Xoff=-0.5, Yoff=-0.5 and Zoff=-3. The offsets denotes the initial X, Y and Z values of the 3Dbox.
- Visualize the box, X, with plot3 -> plot3(X(1,:), X(2,:), X(3,:),’*’)



- Generate a 2D projection (an image) of the 3D box using the pinhole camera model, i.e. obtain pixel coordinates as

$$[x, y]^T = \left[f \cdot \frac{X}{Z}, f \cdot \frac{Y}{Z}\right]^T$$

- Experiment with the object distance to the camera (0,0,0) by varying Zoff up and down. Also try to change the X and Y offsets. Reflect on your findings.

Perspective

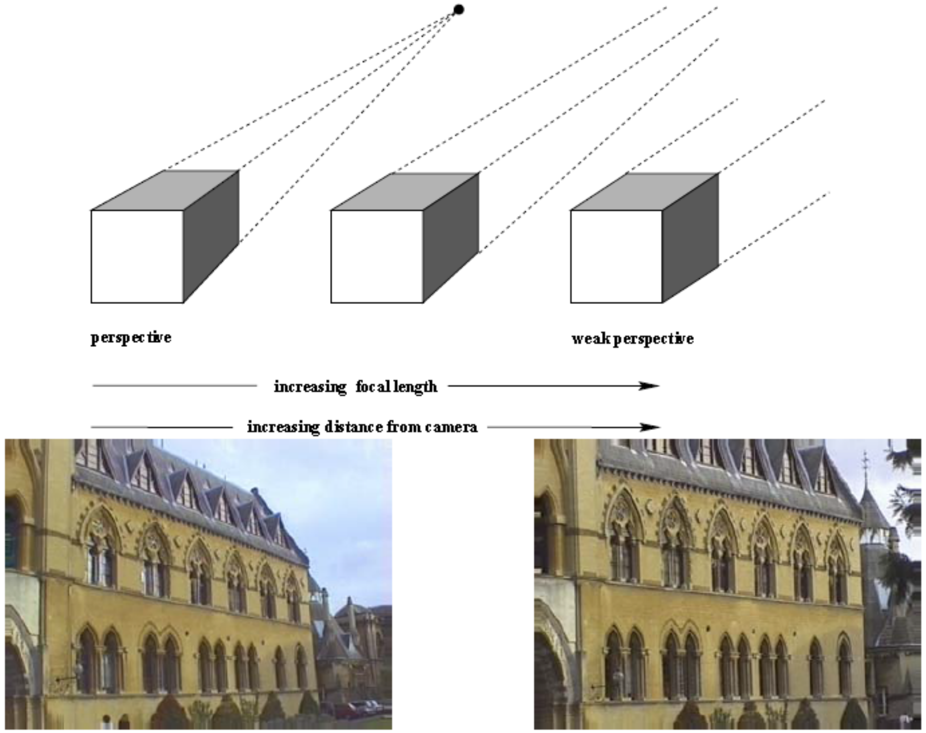
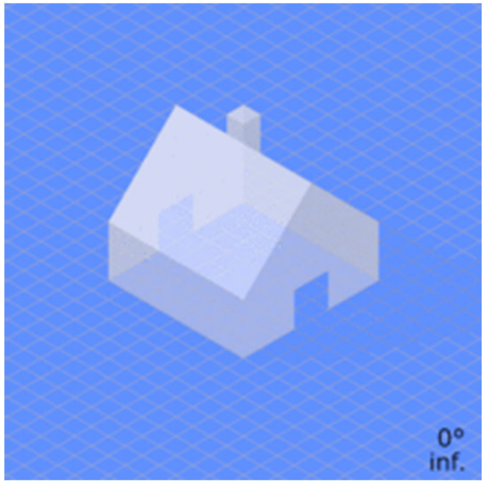


Fig. 6.7. *As the focal length increases and the distance between the camera and object also increases, the image remains the same size but perspective effects diminish.*

Source: Hartley & Zisserman: Multiple View Geometry

Perspective ”Distortion”

The camera’s distance to the object and focal-length determines the ”amount” of perspective distortion – worst when using small focal-lengths and close to the depicted object



https://en.wikipedia.org/wiki/Perspective_distortion_(photography)

Properties of the Pinhole-Camera Model

- Straight lines are mapped to straight lines
- (Ratio of) Length is not preserved
- Angles are not preserved
  - Parallel lines may intersect

Vanishing point

Photo credit: Förstner

Vanishing Points

- Parallel lines are not parallel anymore
- Parallel lines intersect in a vanishing point at infinity

Projective Geometry and Homogeneous Coordinates

- Euclidean geometry is suboptimal to describe points and shapes in imagery (due to ”distortions” arising from central projection)
- Homogeneous coordinates (H.C.) can make transformations such as translation, rotation and scaling easier than for euclidean geometry
- H.C. is also convenient for working with lines and points (intersection of lines, distance between lines and points…)
- Points at infinity can be represented using finite coordinates

Homogeneous Coordinates (2D)

- H.C. uses an 2+1 dimensional vector to represent an 2-dimensional (image)point

$$\overline{\boldsymbol{p}} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

As the name implies this representation is homogeneous, e.g.

$$\overline{\boldsymbol{p}} = \lambda \overline{\boldsymbol{p}} = \begin{bmatrix} \lambda x \\ \lambda y \\ \lambda \end{bmatrix}, (\lambda \neq 0)$$

Transformation back to cartesian coordinates (inhomogeneous coordinates) is done as…

$$\boldsymbol{p} = \begin{bmatrix} \frac{\bar{p}(1)}{\bar{p}(3)} \\ \frac{\bar{p}(2)}{\bar{p}(3)} \end{bmatrix}$$

Line representation in H. C. (2D)

- Line representation in (regular) inhomogenous coordinates

A point $\boldsymbol{p} = [x\ y]^T$, is located on a line, $\boldsymbol{l}$ if:

$$ax + by + c = 0$$

We normally represent lines in slope-intercept form, i.e. $y = 5x + 7$ – this can easily be converted to the above representation by collecting all terms one side of the equation, i.e. $5x - 1y + 7 = 0$

In homogeneous representation, we define the point as $\bar{\mathbf{p}} = [x\ y\ 1]^{\mathrm{T}}$ and denotes the line coefficients as: $\boldsymbol{l} = [\mathrm{a\ b\ c}]^{\mathrm{T}}$, hence

$$\bar{\boldsymbol{p}}^T \boldsymbol{l} = \boldsymbol{l}^T \bar{\boldsymbol{p}} = 0$$

Intersection of lines and points

Intersection of lines. The homegeneous representation is convenient for finding the intersection of lines, e.g

$$\boldsymbol{l_1} = [\mathrm{a}_1\ \mathrm{b}_1\ \mathrm{c}_1]^{\mathrm{T}}$$
$$\boldsymbol{l_2} = [\mathrm{a}_2\ \mathrm{b}_2\ \mathrm{c}_2]^{\mathrm{T}}$$

Intuitively, a intersection point, $\overline{\boldsymbol{p}}$, must satisfy

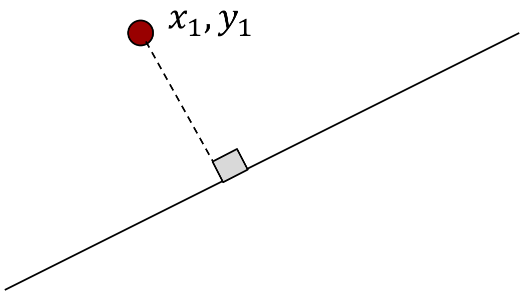
$$\begin{bmatrix} \boldsymbol{l}_1^T \\ \boldsymbol{l}_2^T \end{bmatrix} \overline{\boldsymbol{p}} = 0$$

Hence the point can be identified as the right nullspace of $\begin{bmatrix} \boldsymbol{l}_1^T \\ \boldsymbol{l}_2^T \end{bmatrix}$ and can be computed as $\overline{\boldsymbol{p}} = [\boldsymbol{l_1} \times \boldsymbol{l_2}]$

A line that connects/intersects two points, $\overline{\boldsymbol{p}}_1, \overline{\boldsymbol{p}}_2$, can in analogous way be found as $\boldsymbol{l} = [\overline{\boldsymbol{p_1}} \times \overline{\boldsymbol{p_2}}]$

Distance between a point and a line

- In cartesian/inhomogenous coordinates, we can find the closest distance from a point (x_1,y_1) to a line (a,b,c)



$$d = \frac{|ax_1 + by_1 + c|}{\sqrt{a^2 + b^2}}$$

- In Homogenous coordinates, we can find the distance by

$$d = \bar{l}^T p$$

Where $p = \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$ and $\bar{l}^T = \begin{bmatrix} \bar{a} \\ \bar{b} \\ \bar{c} \end{bmatrix}^T = \frac{1}{\sqrt{a^2+b^2}} l^T$

Points at infinity (Ideal points)

Points at infinity can be expressed finitely in H.C., e.g.

$$\bar{p} = \begin{bmatrix} 23 \\ 10 \\ 0 \end{bmatrix}$$

In cartesian coordinates, information is lost since…

$$p = \begin{bmatrix} \frac{23}{0} \\ \frac{10}{0} \end{bmatrix} = \begin{bmatrix} \infty \\ \infty \end{bmatrix}$$

The line at infinity

The line at infinity, has the following coefficients:

$$\mathbf{l}_\infty = [0\ 0\ 1]^T$$

All ideal points, $[x\ y\ 0]^T$, can be shown to lie on this line since

$$[x\ y\ 0]\begin{bmatrix}0\\0\\1\end{bmatrix} = 0$$

2D Transformations (translation)

In inhomogeneuos coordinates a collections of points can be translated as

$$\boldsymbol{p_2} = \boldsymbol{p_1} + \mathbf{t} = \begin{bmatrix} x_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}$$

In Homogeneous coordinates we have,

$$\overline{\boldsymbol{p_2}} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \overline{\boldsymbol{p_1}} = \overline{\boldsymbol{T}} \cdot \overline{\boldsymbol{p_1}}$$

2D Transformations (translation + rotation)

In inhomogeneuos coordinates a collections of points can be rotated and translated as

$$\boldsymbol{p_2} = \mathbf{R} \cdot \boldsymbol{p_1} + \mathbf{t} = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} = \begin{bmatrix} \cos(\phi)\, x_1 - \sin(\phi)\, y_1 + t_x \\ \sin(\phi)\, x_1 + \cos(\phi)\, y_1 + t_y \end{bmatrix}$$

In Homogeneous coordinates we have,

$$\overline{\boldsymbol{p_2}} = \mathbf{P} \cdot \overline{\boldsymbol{p_1}} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{bmatrix} \cdot \overline{\boldsymbol{p_1}} = \begin{bmatrix} \cos(\phi) & -\sin(\phi) & t_x \\ \sin(\phi) & \cos(\phi) & t_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = \begin{bmatrix} \cos(\phi)\, x_1 - \sin(\phi)\, y_1 + t_x \\ \sin(\phi)\, x_1 + \cos(\phi)\, y_1 + t_y \\ 1 \end{bmatrix}$$

$\mathbf{P}$ can further be decomposed as: $\mathbf{P} = \overline{\mathbf{T}}\overline{\mathbf{R}} = \begin{bmatrix} \mathbf{I} & \mathbf{t} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{0} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \boldsymbol{t} \\ 0 & 1 \end{bmatrix}$

2D Transformations (Similarity)

A transformation which involves a translation, rotation and scaling of points is known as a similarity transform, the transformation $\mathbf{P}$ has the general structure:

$$\mathbf{P} = \bar{\mathbf{T}} \cdot \bar{\mathbf{R}} \cdot \bar{\mathbf{S}} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\phi) & -\sin(\phi) & 0 \\ \sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathrm{s} & 0 & t_x \\ 0 & \mathrm{s} & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

2D Transformations (Affine)

An affine transformation $\mathbf{P}$ has the general structure:

$$\mathbf{P} = \begin{bmatrix} \mathbf{A} & \mathbf{t} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

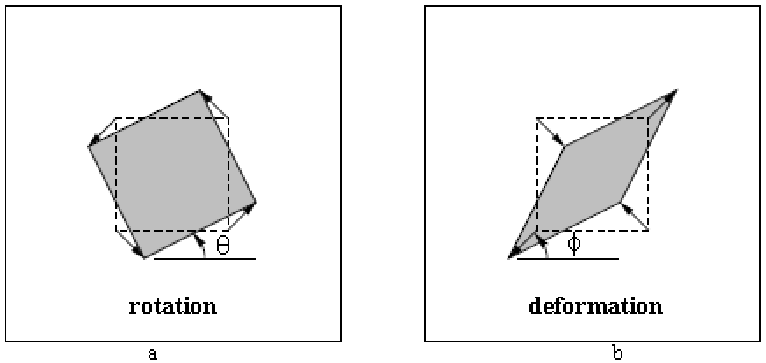


Fig. 2.7. **Distortions arising from a planar affine transformation.** *(a) Rotation by* $R(\theta)$*. (b) A deformation* $R(-\phi)\,D\,R(\phi)$*. Note, the scaling directions in the deformation are orthogonal.*

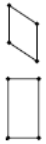
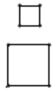
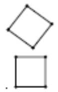
$$\mathbf{D} = \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix}$$

2 DOF extra than similarity

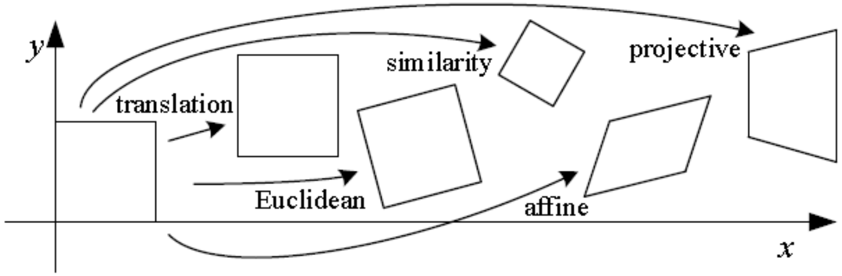
- Imposed from ratio of λ_1, λ_2 and scaling direction ϕ

(Hartley & Zisserman)

Projective Transformations (2D)

Group	Matrix	Distortion	Invariant properties
Projective 8 dof	$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$		Concurrency, collinearity, **order of contact**: intersection (1 pt contact); tangency (2 pt contact); inflections (3 pt contact with line); tangent discontinuities and cusps. cross ratio (ratio of ratio of lengths).
Affine 6 dof	$\begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		Parallelism, ratio of areas, ratio of lengths on collinear or parallel lines (e.g. midpoints), linear combinations of vectors (e.g. centroids). The line at infinity, $\mathbf{l}_\infty$.
Similarity 4 dof	$\begin{bmatrix} sr_{11} & sr_{12} & t_x \\ sr_{21} & sr_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		Ratio of lengths, angle. The circular points, **I**, **J** (see section 2.7.3).
Euclidean 3 dof	$\begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		Length, area

Hartley & Zisserman



Projections of points between planes

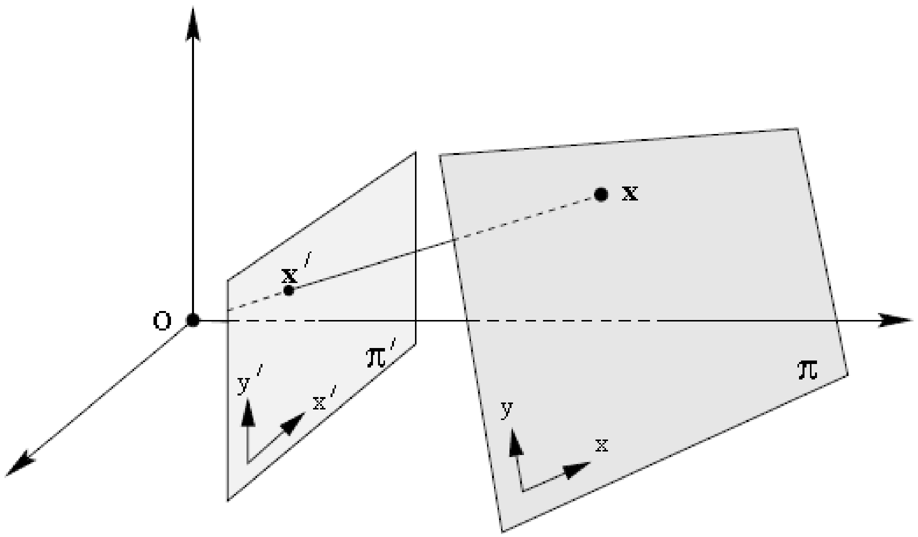


Fig. 2.3. **Central projection maps points on one plane to points on another plane.** *The projection also maps lines to lines as may be seen by considering a plane through the projection centre which intersects with the two planes π and π'. Since lines are mapped to lines, central projection is a projectivity and may be represented by a linear mapping of homogeneous coordinates* $\mathbf{x}' = \mathrm{H}\mathbf{x}$.

(Hartley & Zisserman)

Removing perspective ”distortion”

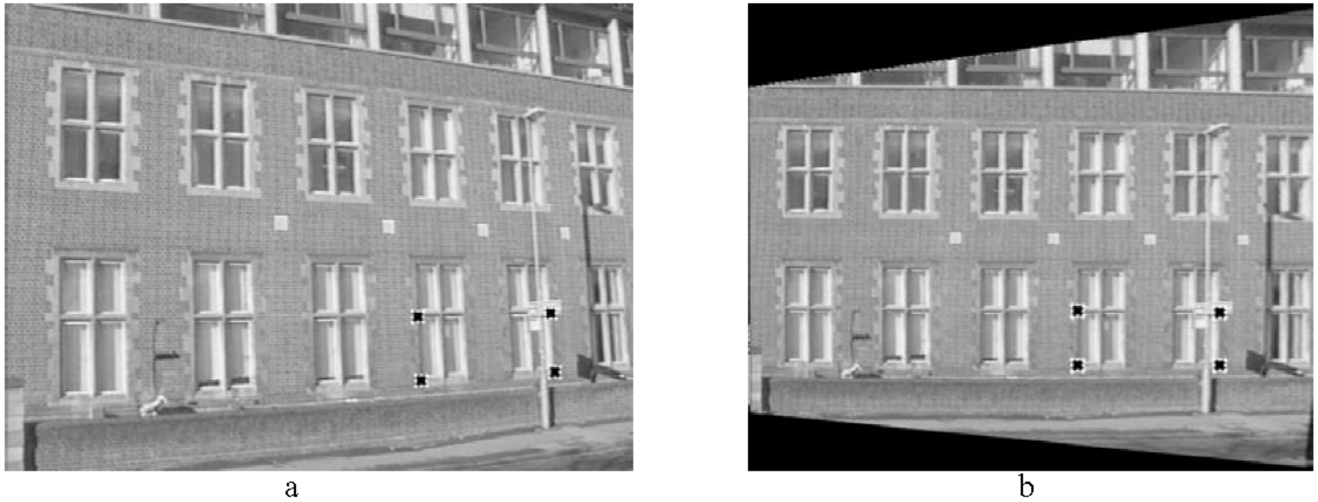
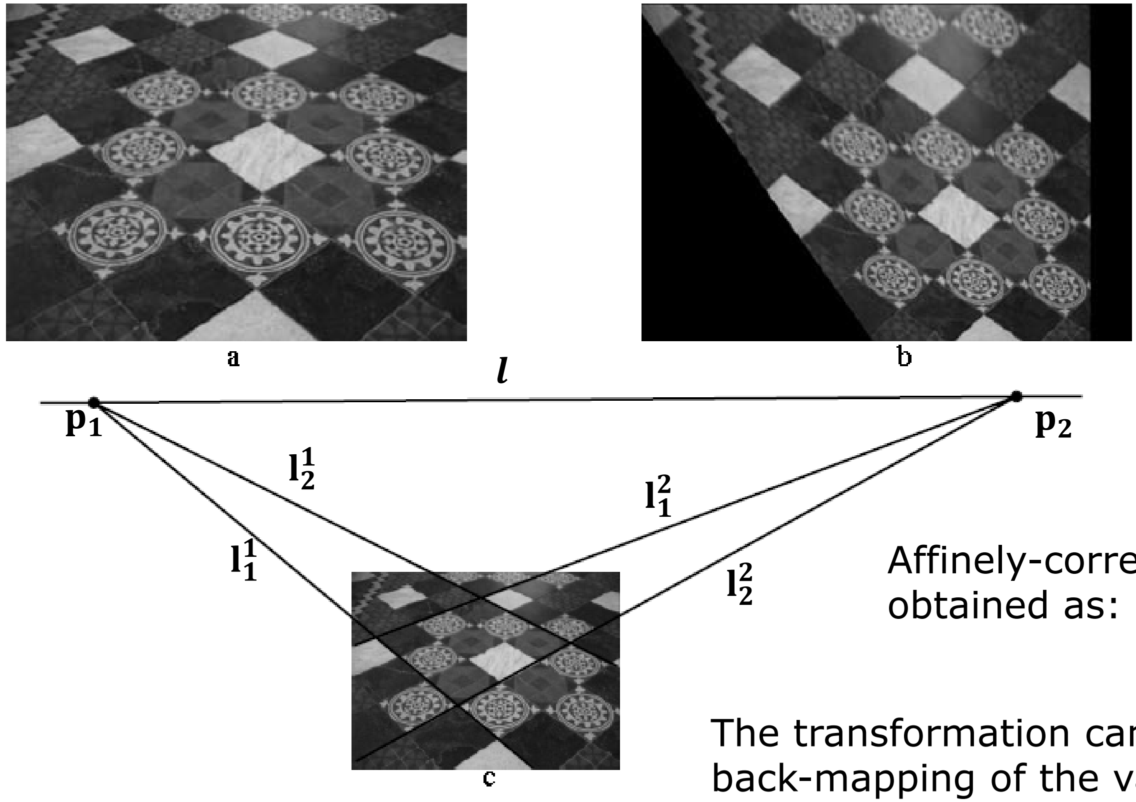


Fig. 2.4. **Removing perspective distortion.** *(a) The original image with perspective distortion – the lines of the windows clearly converge at a finite point. (b) Synthesized frontal orthogonal view of the front wall. The image (a) of the wall is related via a projective transformation to the true geometry of the wall. The inverse transformation is computed by mapping the four imaged window corners to corners of an appropriately sized rectangle. The four point correspondences determine the transformation. The transformation is then applied to the whole image. Note that sections of the image of the ground are subject to a further projective distortion. This can also be removed by a projective transformation.*

(Hartley & Zisserman)

Recover affine properties from vanishing-line



Vanishing-points

$$\mathbf{p}_1 = \mathbf{l}_1^1 \times \mathbf{l}_2^1$$

$$\mathbf{p}_2 = \mathbf{l}_1^2 \times \mathbf{l}_2^2$$

Vanishing-line

$$\mathbf{l} = \mathbf{p}_1 \times \mathbf{p}_2$$

$$\mathbf{H} = \begin{bmatrix} 1 & & \\ & 1 & \\ l(1) & l(2) & l\,(3) \end{bmatrix}$$

Affinely-corrected image can be obtained as:

$$\mathbf{x}' = \mathbf{H}\mathbf{x}$$

The transformation can be thought of as a back-mapping of the vanishing line to the line at infinity $l_\infty = \mathbf{H}^{-\mathbf{T}}\mathbf{l}$ Eq. 2.6 (H-Z)

Recover affine properties from vanishing-line

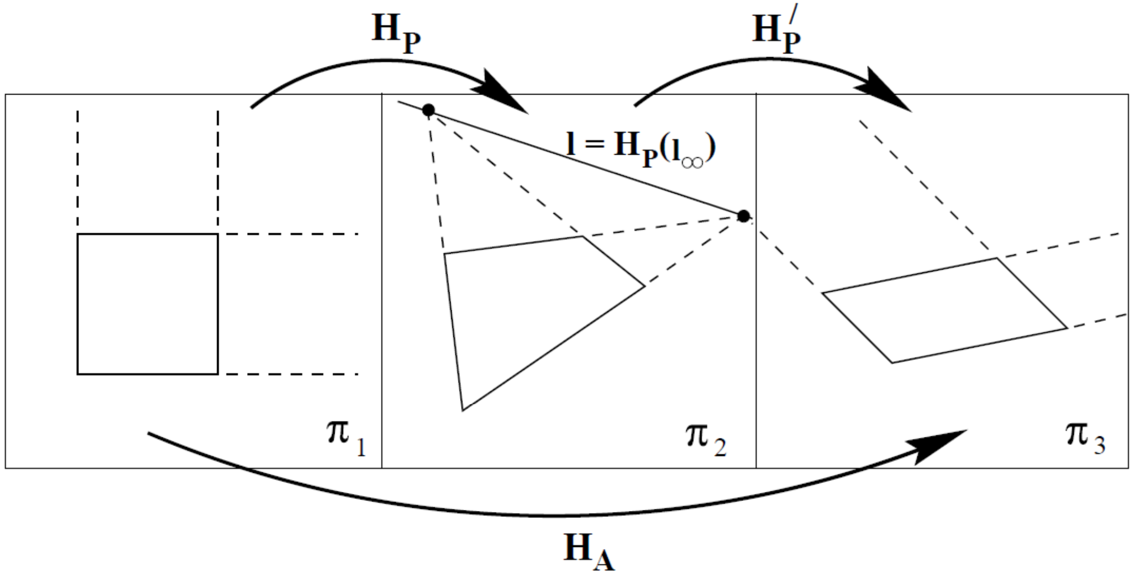


Fig. 2.12. **Affine rectification.** *A projective transformation maps* $\mathbf{l}_\infty$ *from* $(0,0,1)^\mathsf{T}$ *on a Euclidean plane* π_1 *to a finite line* $\mathbf{l}$ *on the plane* π_2. *If a projective transformation is constructed such that* $\mathbf{l}$ *is mapped back to* $(0,0,1)^\mathsf{T}$ *then from result 2.17 the transformation between the first and third planes* must be an affine transformation *since the canonical position of* $\mathbf{l}_\infty$ *is preserved. This means that affine properties of the first plane can be measured from the third, i.e. the third plane is within an affinity of the first.*

(Hartley & Zisserman)

Recover metric properties

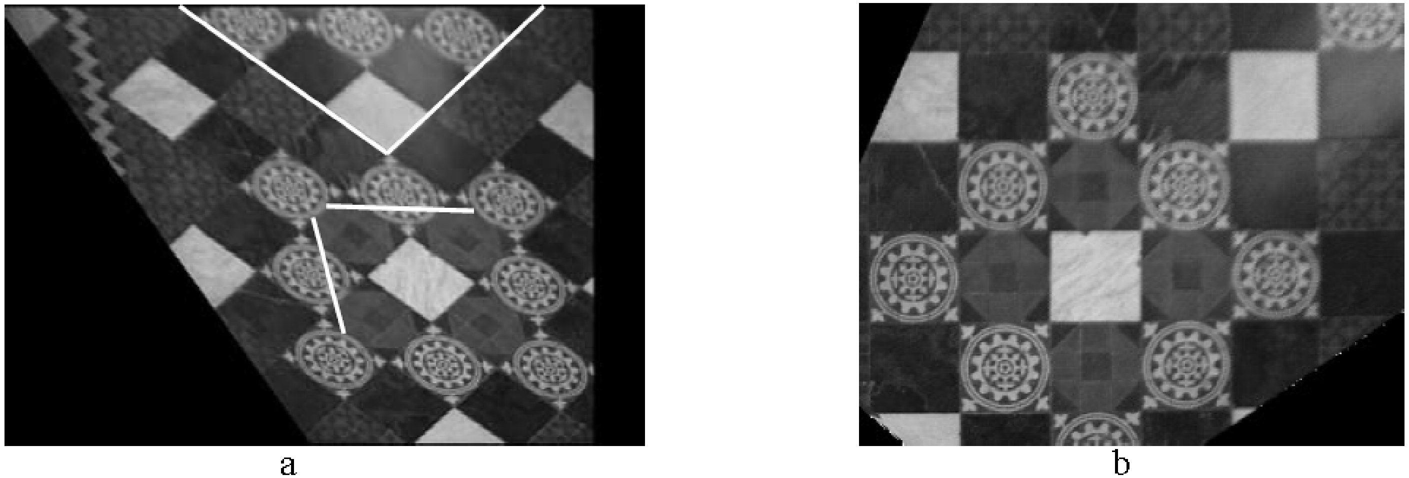


Fig. 2.17. **Metric rectification via orthogonal lines I.** *The affine transformation required to metrically rectify an affine image may be computed from imaged orthogonal lines. (a) Two (non-parallel) line pairs identified on the affinely rectified image (figure 2.13) correspond to orthogonal lines on the world plane. (b) The metrically rectified image. Note that in the metrically rectified image all lines orthogonal in the world are orthogonal, world squares have unit aspect ratio, and world circles are circular.*

See p. 56-57 H-Z (chp 2.)