University at Buffalo
The State University of New York

Image Formation, Light and Color

CSE473/573 - Computer Vision and Image Processing

Junsong Yuan

jsyuan@buffalo.edu

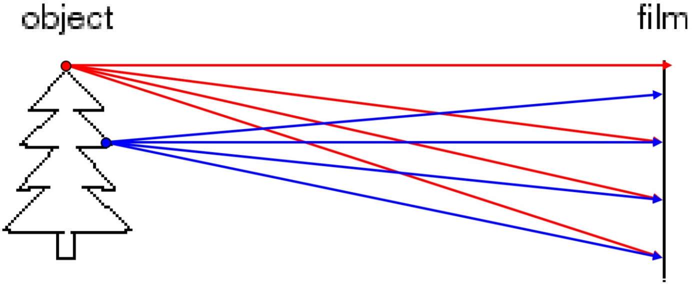
2018 Fall

Overview

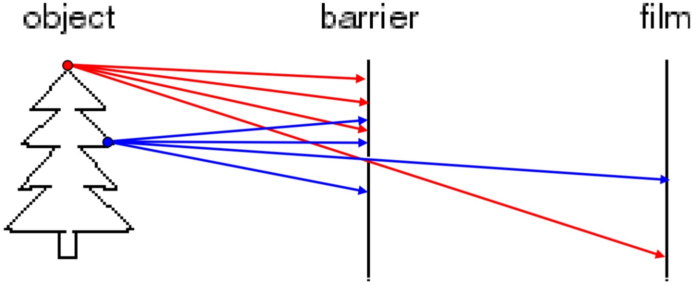
1. Image Formation
2. Color
3. Image Enhancement

How to capture an Image

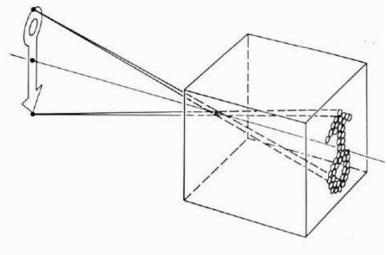
Why can't we capture an image this way?



Pinhole Camera



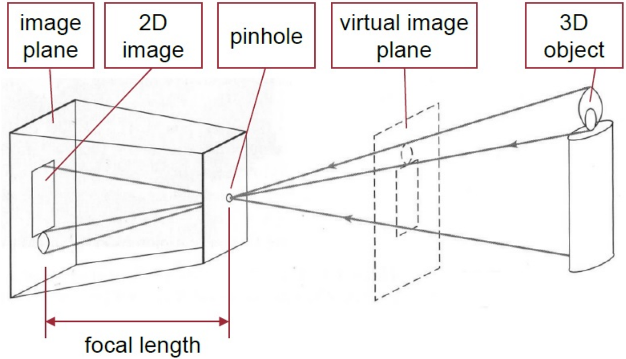
Pinhole Camera Model



Pinhole model:

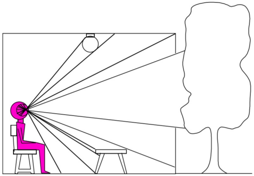
- Captures **pencil of rays** – all rays through a single point
- The point is called **Center of Projection (COP)**
- The image is formed on the **Image Plane**
- **Effective focal length *f*** is distance from COP to Image Plane

Pinhole Camera Model



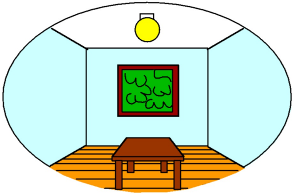
3D to 2D

3D world



Point of observation

2D image



What have we lost?

- Angles
- Distances (lengths)

Camera Model

In order to have a quantitative description of how a given camera views the real world, we need a **camera model**.

This is a mapping C: $R^3 \rightarrow R^2$ which specifies how the 3D scene will appear on the 2D image plane of the camera. A camera model supports two types of parameters:

Intrinsic and Extrinsic Parameters

Camera Parameters

Intrinsic parameters

properties of the camera itself, which do not change as the position and orientation of the camera in space are changed.

Extrinsic parameters

those which change with position and orientation of the camera.

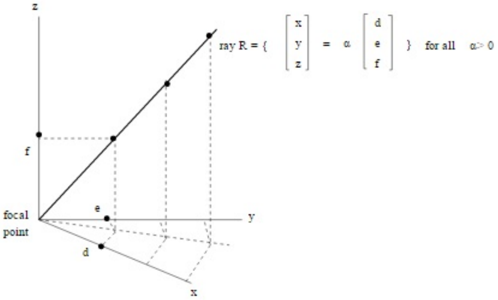
- E.g.: focal length and lens magnification factor are intrinsic parameters, focal point location and vector orientation of the optical axis are extrinsic parameters.

Both sets of parameters must be completely specified before the camera model C: $R^3 \rightarrow R^2$ is known and we can predict the 2D image on the image plane that derives from any given 3D scene.

Ray maps to a Point on Image Plane

The geometry of perspective projection is used to develop the camera model.

Define a **ray** as a half-line beginning at the origin.

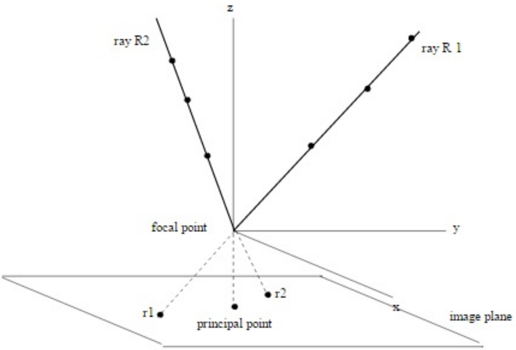


All points on a given ray R will map into the single point r in the image plane.

Perspective Projection

All points on a given ray R will map into the single point r in the image plane.

The set of all such mappings is the perspective projection.



Projective Space

Formally, let x and x' be nonzero vectors in R^{n+1} (we will use n=2) and define $x \equiv x'$ (x is equivalent to x') if and only if $x' \propto x$.

Then the quotient space of this equivalence relation (set of equivalence classes) is P^n, the **projective space** associated with the Euclidean space R^{n+1}.

Note that the projective space is lower-dimensional than the Euclidean space (dim = 2 vs 3).

Projective Space - Example

Informally, points in P^2 can be thought of, and represented, as follows. Take a point in R^3, say $[x, y, z]^T$, write as $z[x/z, y/z, 1]^T$ ($z \neq 0$). Then $[x, y, z]^T$ in R^3 maps into the point $r = [x/z, y/z]$ in P^2.

E.g.:

$[30, 15, 5]^T \to [6, 3]^T$

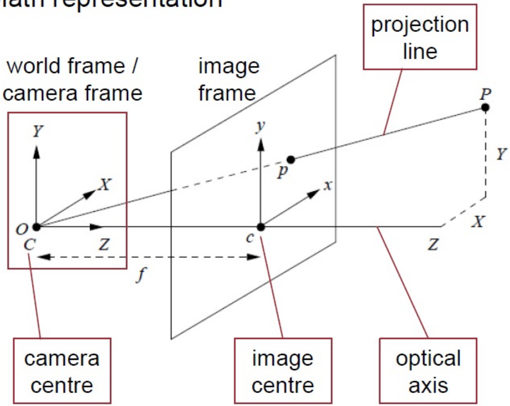
$[3, 3/2, 1/2]^T \to [6, 3]^T$

So these two points in R^3 project to the same point [6 3] in the projective space P^2.

A good example to understand why homogeneous co-ordinates are used in Camera Model

Camera geometry

- Math representation



Perspective Projection Equations

- 3D point $\mathbf{P} = (X, Y, Z)^\mathsf{T}$ projects to 2D image point $\mathbf{p} = (x, y)^\mathsf{T}$.
- By symmetry,

$$\frac{X}{Z} = \frac{x}{f}, \quad \frac{Y}{Z} = \frac{y}{f}$$

i.e.,

$$x = f\frac{X}{Z}, \quad y = f\frac{Y}{Z}$$

- Simplest form of perspective projection.

Perspective Projection - Example

Source: S. Lazebnik

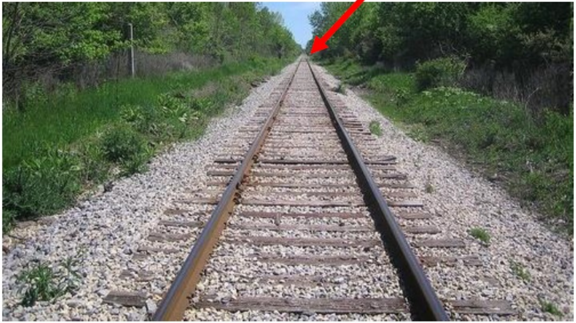
Perspective Projection - Properties

- Points project to points
- Lines project to lines
- Distant objects look smaller

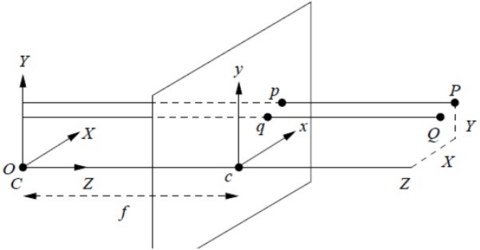
Perspective Projection - Properties

- Angles are not preserved
- Parallel lines meet!

Vanishing point



Orthographic Projection

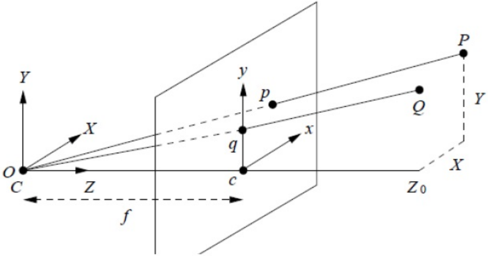


- 3D scene is at infinite distance from camera.
- All projection lines are parallel to optical axis.
- So,

$$x = X, \quad y = Y$$

Very unrealistic. You need a camera sensor as big as the object

Weak Perspective Projection



- Scene depth << distance to camera.
- Z is the same for all scene points, say Z_0

$$x = sX, \quad y = sY, \quad s = \frac{f}{Z_0} \text{ for all scene points.}$$

When you zoom your camera very far, the depth changes in the scene are negligible when compared to the distance from camera

Weak Perspective Projection - Example

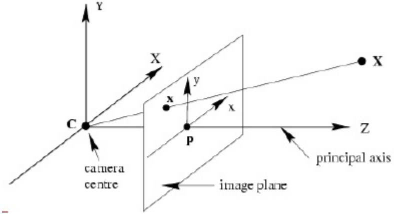
Qingming Festival by the Riverside Zhang Zeduan ~900 AD

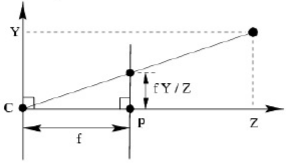
Projection Matrix

- Geometrically saw x=fX/Z, y=fY/Z

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

3x4 Projection Matrix

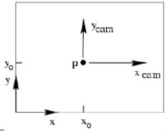




Projection Matrix - Image Origin

- Intersection of principal axis with image plane often not at image origin

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

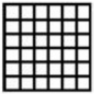


$$K = \begin{bmatrix} f & & p_x \\ & f & p_y \\ & & 1 \end{bmatrix}$$ (Intrinsic) Calibration matrix

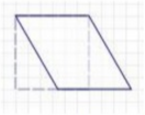
Camera Calibration Matrix

- Camera Calibration Matrix ‘K’ – 3x3 Upper triangular Matrix
- Constitutes – Focal length of the camera ‘f’ , Principal Point (u0,v0), aspect ratio of the pixel ‘γ’ and the skew ‘s’ of the sensor pixel
- Intrinsic parameters can be estimated using camera calibration techniques

$$K = \begin{pmatrix} \gamma f & s & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

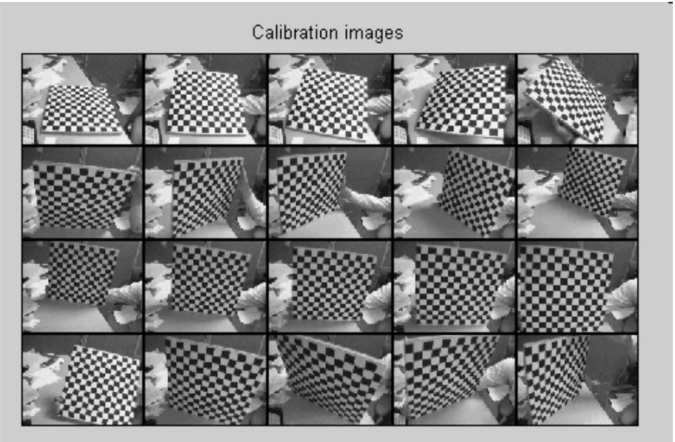


Ideal image sensor

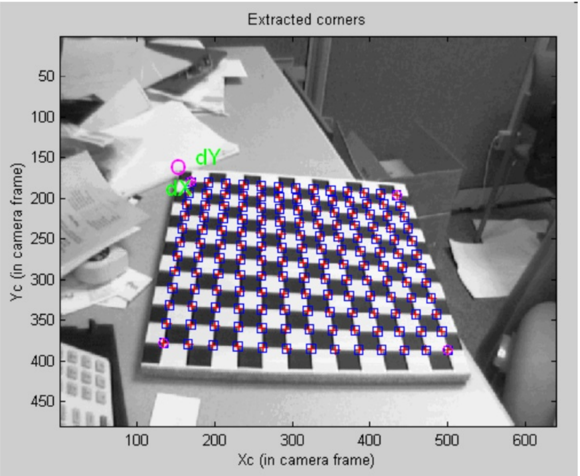


Sensor pixel with skew

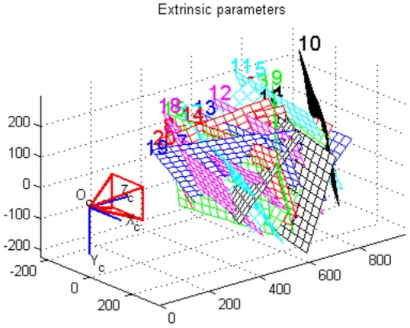
Camera Calibration Example



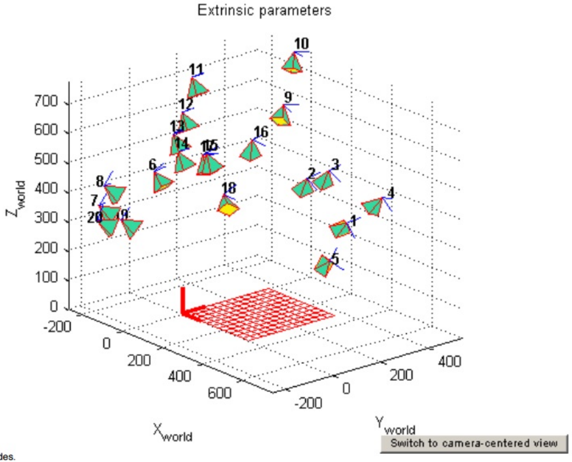
Camera Calibration Example



Camera Calibration Example

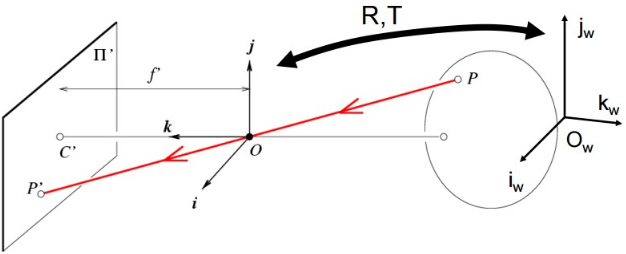


Camera Calibration Example



Source: S Savarese slides.

Extrinsic Parameters



$$\mathbf{x} = \mathbf{K}[\mathbf{R} \quad \mathbf{t}]\mathbf{X}$$

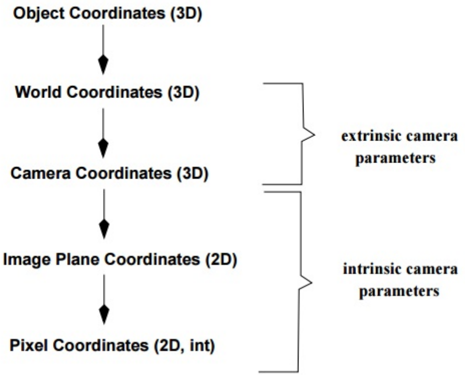
x: Image Coordinates: (u,v,1)
K: Intrinsic Matrix (3x3)
R: Rotation (3x3)
t: Translation (3x1)
X: World Coordinates: (X,Y,Z,1)

Rotation and Translation constitute extrinsic parameters

Camera Model Pipeline

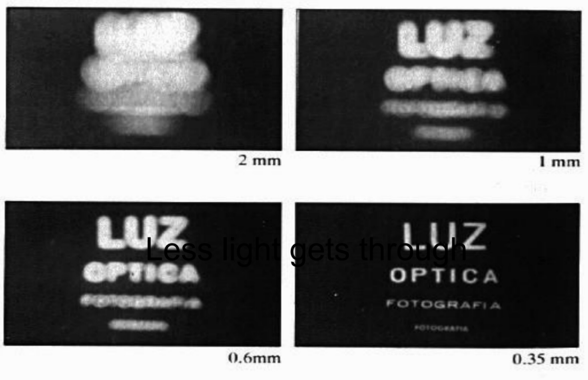
Extrinsic camera parameters: the parameters that define the *location* and *orientation* of the camera reference frame with respect to a known world reference frame.

Intrinsic camera parameters: the parameters necessary to link the pixel coordinates of an image point with the corresponding coordinates in the camera reference frame.



Pin Hole Aperture Size

What happens when we reduce the aperture size further?



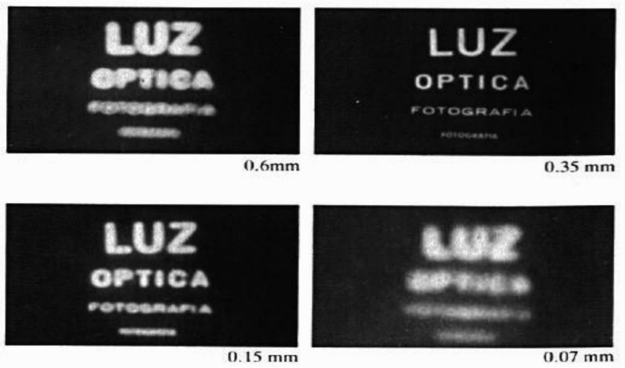
Less light gets through

Pin Hole Aperture Size

What happens when we reduce the aperture size further?

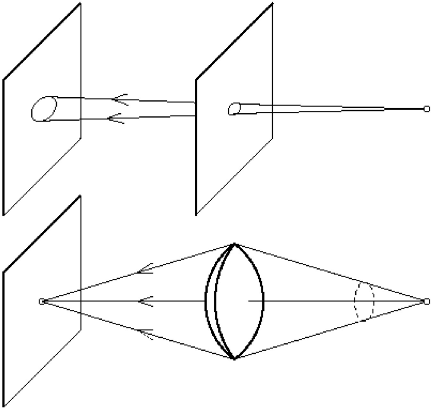
Less light gets through

and diffraction effects



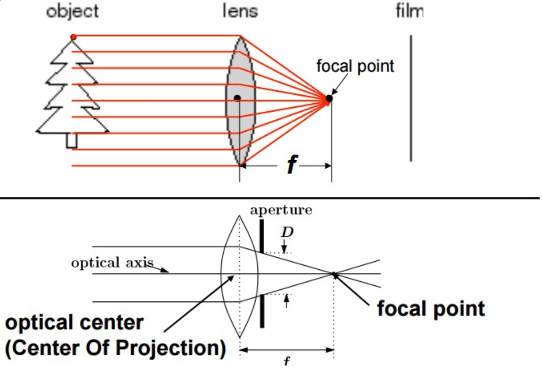
Adding lens

More light rays from a point in the scene get projected on the image plane

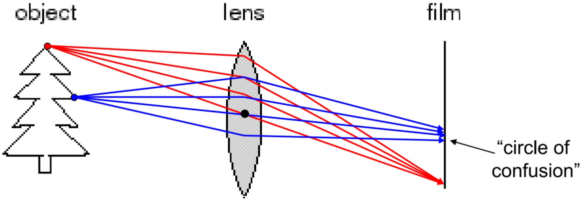


Adding lens

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



Adding lens

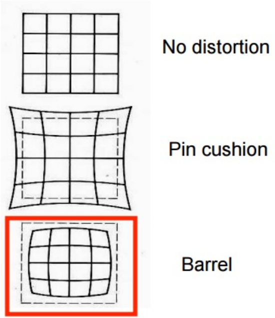


A lens focuses light onto the film

- There is a specific distance at which objects are "in focus"
  - other points project to a "circle of confusion" in the image
- Changing the shape of the lens changes this distance

Lens Distortions

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



o: S Savarese slides

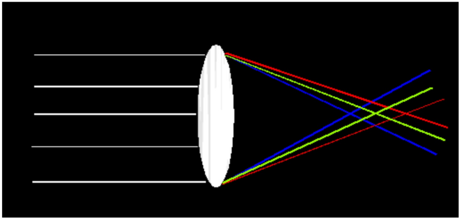
Lens Distortions

Chromatic Aberration

Dispersion: wavelength-dependent refractive index

- (enables prism to spread white light beam into rainbow)

Modifies ray-bending and lens focal length: $f(\lambda)$

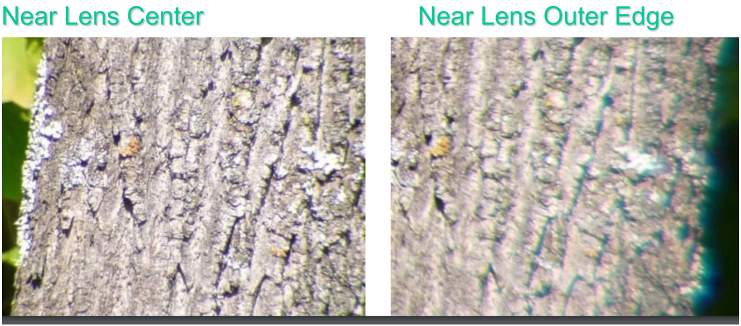


color fringes near edges of image

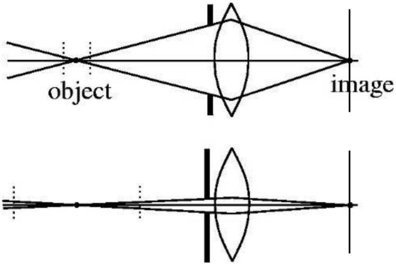
Corrections: add 'doublet' lens of flint glass, etc.

Lens Distortions

Chromatic Aberration



Depth of Field

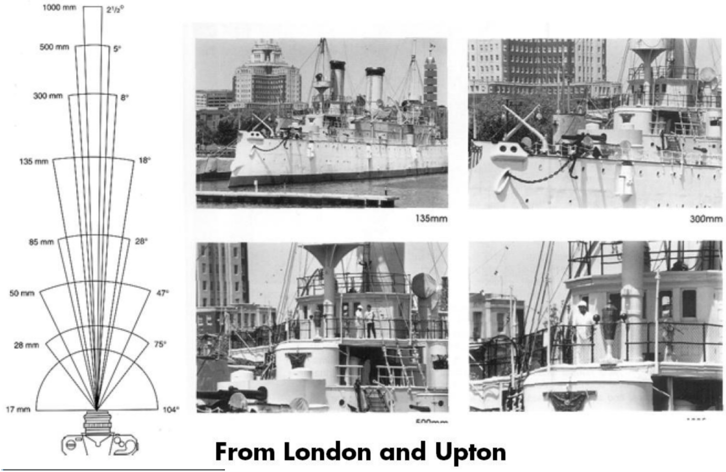


Changing the aperture size affects depth of field

- 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

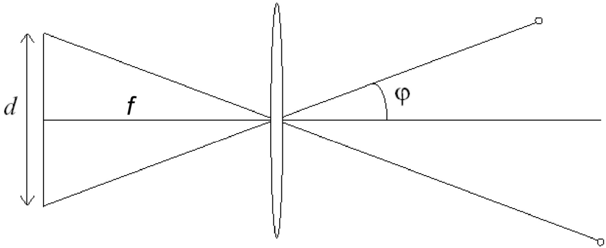
Depth of Field

Field of View (FOV)



From London and Upton

Field of View (FOV)

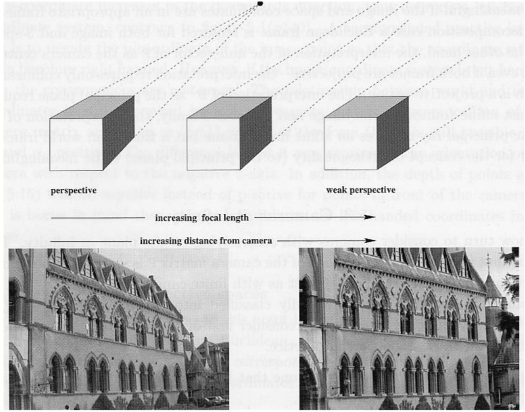


Size of field of view governed by size of the camera retina:

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

Smaller FOV = larger Focal Length

Small FOV - Weak Perspective



From Zisserman & Hartley

Field of View (FOV)

Which image is probably taken on a Smartphone and which one is taken on a professional Digital Camera?

Field of View (FOV)

Large FOV, small f
Camera close to car

Small FOV, large f
Camera far from the car

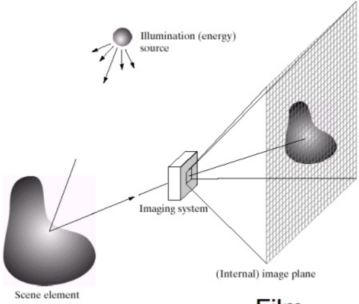
Photometric Imaging

- Three components to pixel brightness
  - Illumination and light sources

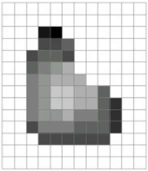
  - Surfaces and reflection

  - Camera response

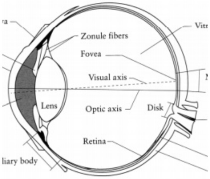
Digital Camera



Film



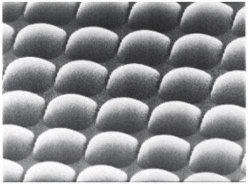
Digital Camera

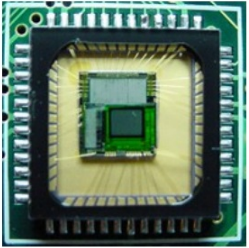


The Eye

Digital Camera

- Basic process:
  - photons hit a detector
  - the detector becomes charged
  - the charge is read out as brightness
- Sensor types:
  - CCD (charge-coupled device)
    - most common
    - high sensitivity
    - high power
    - cannot be individually addressed
    - blooming
  - CMOS
    - simple to fabricate (cheap)
    - lower sensitivity, lower power
    - can be individually addressed





Camera Sensor

Sampling and Quantization

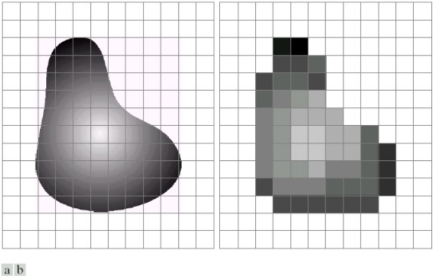
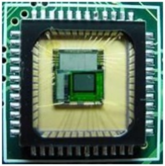


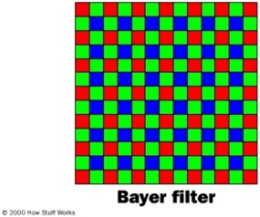
FIGURE 2.17 (a) Continuos image projected onto a sensor array. (b) Result of image sampling and quantization.

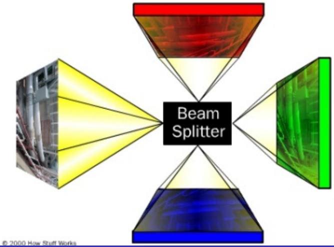


CMOS sensor

Color Cameras

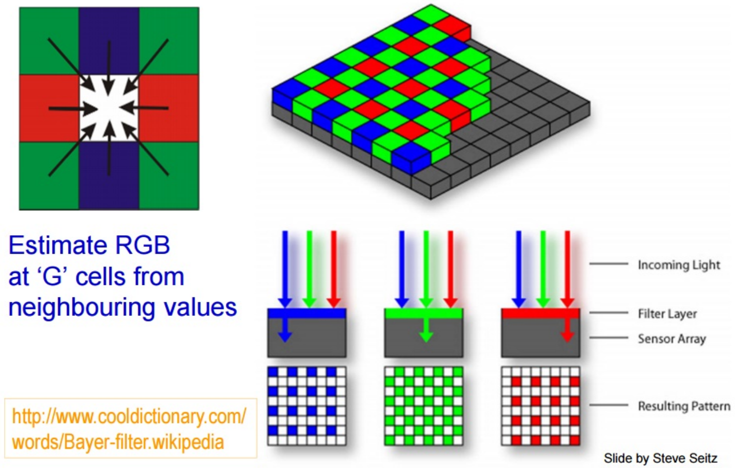
- 1 CCD cameras
  - A **Bayer** pattern is placed in front of the CCD
  - A **Demosaicing** process reads the pixels in a region and computes color and intensity
- 3 CCD camera use a beam splitter and 3 separate CCDs
  - higher color fidelity
  - needs lots of light
  - requires careful alignment of ccds





Bayer Filter

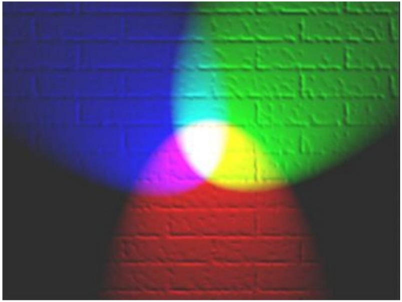
Sampling and Quantization for a Color Image



Shutter Mechanism

Rolling shutter example

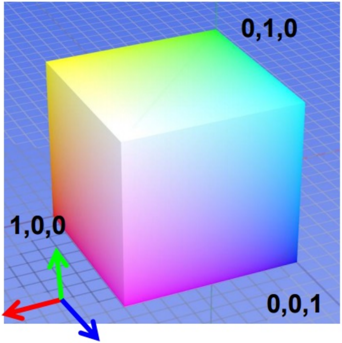
Color



Color Images - Three Channels

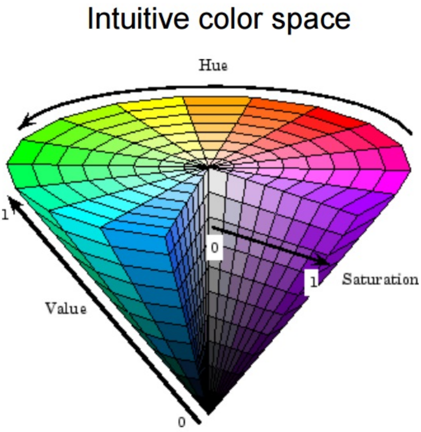
RGB Color Space

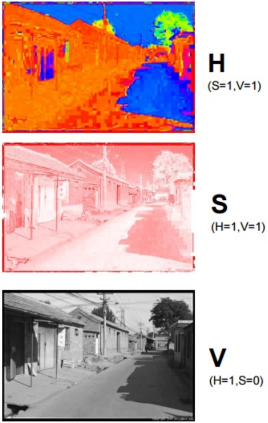
Default color space



Some drawbacks
- Strongly correlated channels
- Non-perceptual

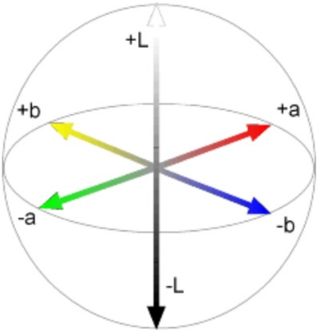
HSV Color Space

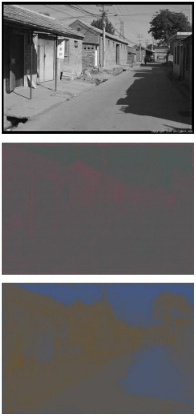




LAB Color Space

"Perceptually uniform"* color space





L
(a=0,b=0)

a
(L=65,b=0)

b
(L=65,a=0)

Luminance vs Chrominance

If you had to choose, would you rather go without luminance or chrominance?

Which of these two captures maximum information about a scene?

Information from Chrominance

Only color shown

Information from Luminance

Only intensity shown

Imaging in a Nut-Shell

Image Sciences in a nutshell

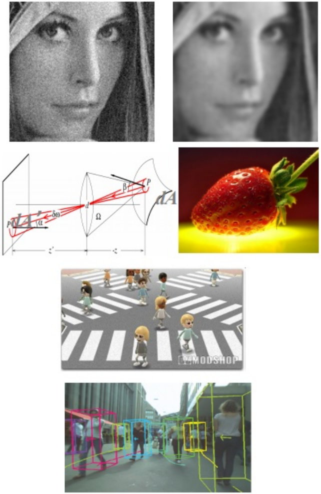


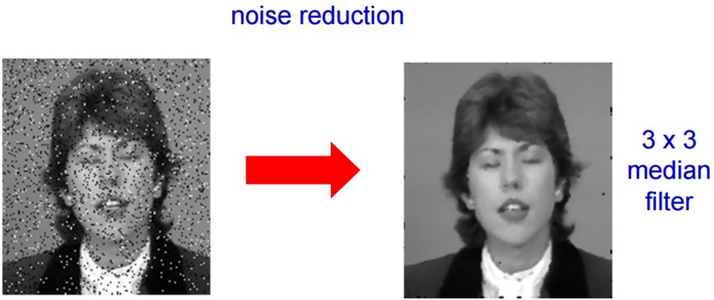
Image Processing
Image to Image

Imaging
Physics to Image

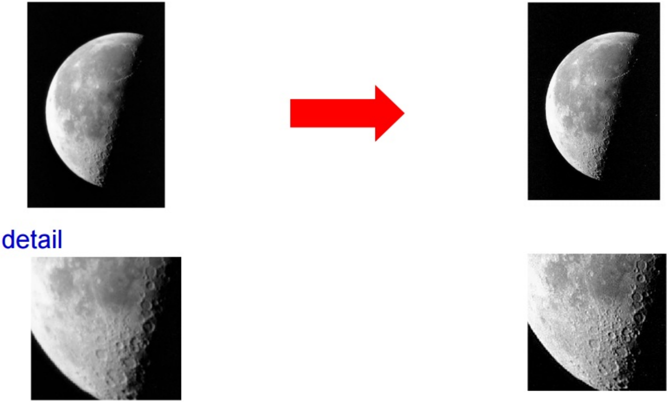
Computer Graphics
Models to Image

Computer Vision
Image to Models

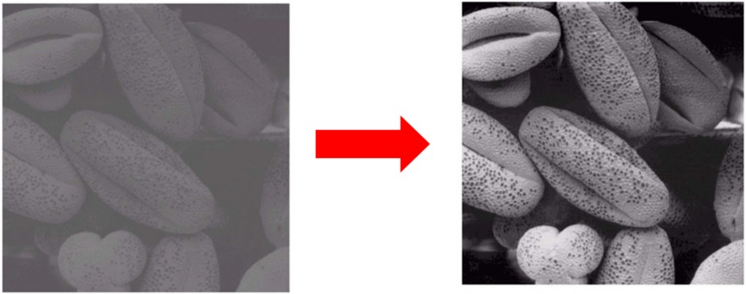
Example 1 - Salt and Pepper Noise Removal



Example 2 - Image Sharpening



Example 3 - Contrast Enhancement



Example 4 - De-blurring

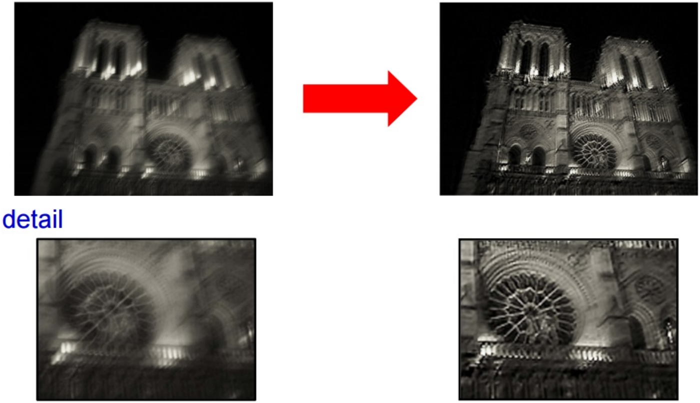


Image as a function of Intensity

- We can think of an **image** as a function, f ,from $\mathbb{R}^2 \rightarrow \mathbb{R}$:
  - $f(x,y)$ gives the **intensity** at position (x,y)
  - Realistically, we expect the image only to be defined over a rectangle, with a finite range:

$$f: [a,b] \times [c,d] \rightarrow [0,1]$$

- A color image is just three functions pasted together. We can write this as a "vector-valued" function:

$$f(x,y) = \begin{bmatrix} r(x,y) \\ g(x,y) \\ b(x,y) \end{bmatrix}$$

Discrete Images

- We usually work with **digital** (**discrete**) images:
  - **Sample** the 2D space on a regular grid
  - **Quantize** each sample (round to nearest integer)
- If our samples are Δ apart, we can write this as:

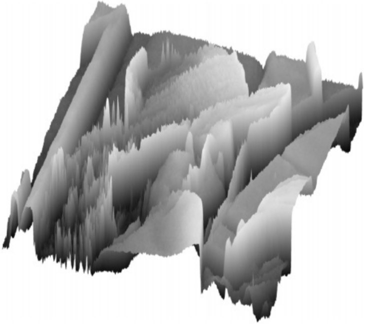
$$f[i,j] = \text{Quantize}\{f(i\Delta, j\Delta)\}$$

- The image can now be represented as a matrix of integer values

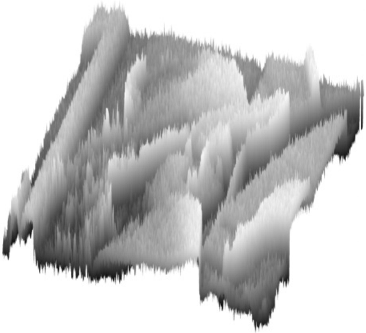
62	79	23	119	120	105	4	0
10	10	9	62	12	78	34	0
10	58	197	46	46	0	0	48
176	135	5	188	191	68	0	49
2	1	1	29	26	37	0	77
0	89	144	147	187	102	62	208
255	252	0	166	123	62	0	31
166	63	127	17	1	0	99	30

Source: Seitz and Szeliski Slides

Image Intensity Plot



Noise in Images



Key assumption: clean image is smooth

Noise in Images

Radhakrishna Dasari, Robert Collins, David Forsyth