Image formation and cameras

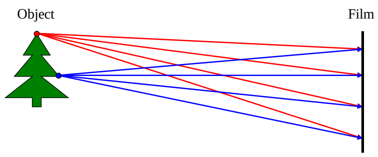
CSE P 576

Larry Zitnick (larryz@microsoft.com)

Many slides courtesy of Steve Seitz

Photography

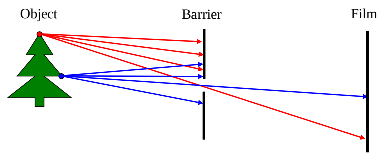
Image formation



Let's design a camera

- Idea 1: put a piece of film in front of an object
- Do we get a reasonable image?

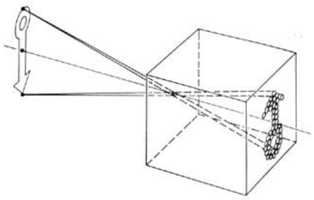
Pinhole camera



Add a barrier to block off most of the rays

- This reduces blurring
- The opening known as the **aperture**
- How does this transform the image?

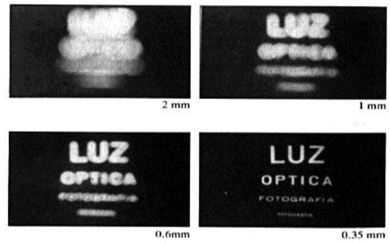
Camera Obscura



The first camera

- Known to Aristotle
- How does the aperture size affect the image?

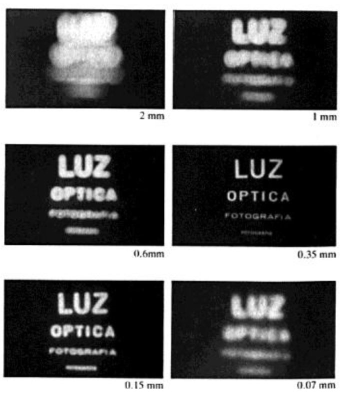
Shrinking the aperture



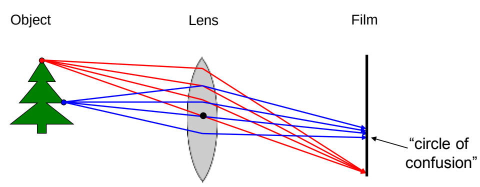
Why not make the aperture as small as possible?

- Less light gets through
- *Diffraction* effects...

Shrinking the aperture



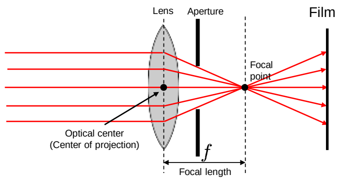
Adding a lens



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
- Changing the shape of the lens changes this distance

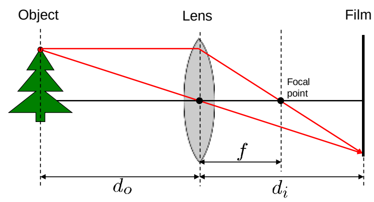
Lenses



A lens focuses parallel rays onto a single focal point

- focal point at a distance *f* beyond the plane of the lens
  - *f* is a function of the shape and index of refraction of the lens
- Aperture of diameter D restricts the range of rays
  - aperture may be on either side of the lens
- Lenses are typically spherical (easier to produce)

Thin lenses

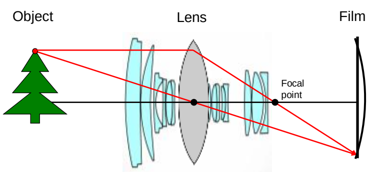


Thin lens equation: $$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}$$ Not quite right…

- Any object point satisfying this equation is in focus
- What is the shape of the focus region?
- How can we change the focus region?
- Thin lens applet: http://www.phy.ntnu.edu.tw/java/Lens/lens_e.html (by Fu-Kwun Hwang)

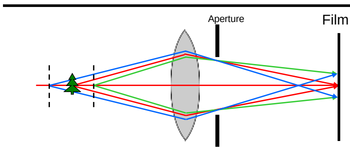
Thin lens assumption

The thin lens assumption assumes the lens has no thickness, but this isn't true…

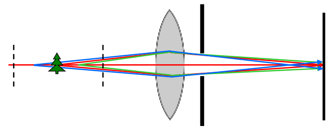


By adding more elements to the lens, the distance at which a scene is in focus can be made roughly planar.

Depth of field



f / 5.6



f / 32

Changing the aperture size affects depth of field

- A smaller aperture increases the range in which the object is approximately in focus

Flower images from Wikipedia http://en.wikipedia.org/wiki/Depth_of_field

Camera parameters

Focus – Shifts the depth that is in focus.

Focal length – Adjusts the zoom, i.e., wide angle or telephoto lens.

Aperture – Adjusts the depth of field and amount of light let into the sensor.

Exposure time – How long an image is exposed. The longer an image is exposed the more light, but could result in motion blur.

ISO – Adjusts the sensitivity of the “film”. Basically a gain function for digital cameras. Increasing ISO also increases noise.

Causes of noise

Shot noise – variation in the number of photons (low light situations.)

Readout noise – Noise added upon readout of pixel. In some cases can be subtracted out.

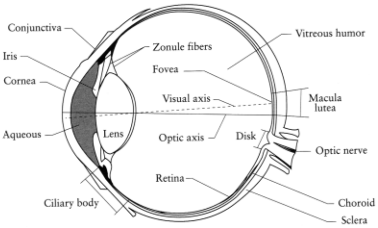
Dark noise – Noise caused by electrons thermally generated. Depends on the temperature of device.

Sport photography

Why do they have such big lenses?

Dirkus Maximus

The eye



The human eye is a camera

- **Iris** - colored annulus with radial muscles
- **Pupil** - the hole (aperture) whose size is controlled by the iris
- What’s the “film”?
  - photoreceptor cells (rods and cones) in the **retina**

Digital Cameras

Digital camera

CCD vs. CMOS

CCD

- Low-noise images
- Consume more power
- More and higher quality pixels

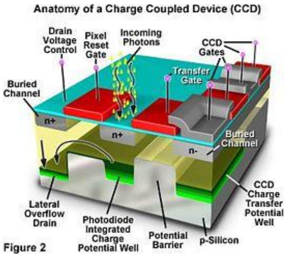
CMOS

- More noise (sensor area is smaller)
- Consume much less power
- Popular in camera phones
- Getting better all the time

http://electronics.howstuffworks.com/digital-camera.htm

Mega-pixels

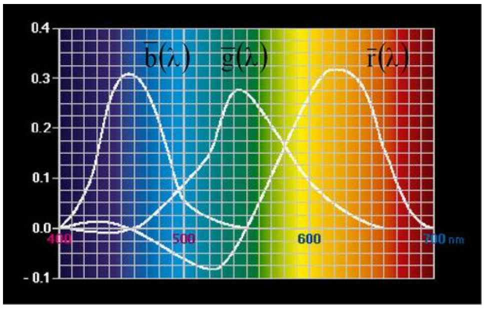
Are more mega-pixels better?



More mega-pixels require higher quality lens.

Colors

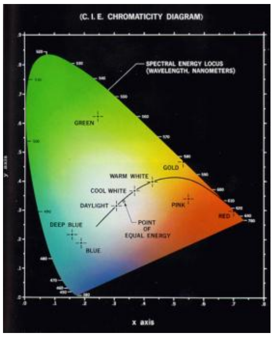
What colors do humans see?



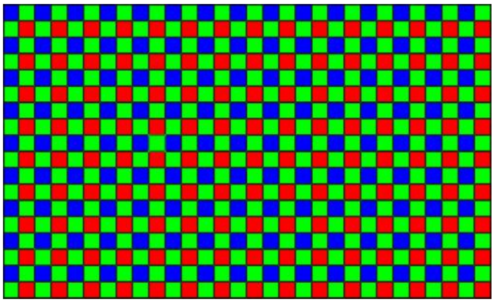
RGB tristimulus values, 1931 RGB CIE

Colors

Plot of all visible colors (Hue and saturation):



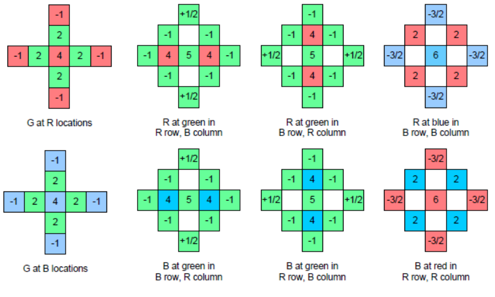
Bayer pattern



Some high end video cameras have 3 CCD chips.

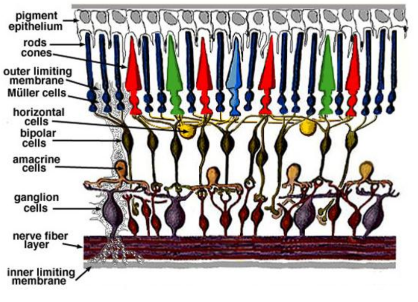
Demosaicing

How can we compute an R, G, and B value for every pixel?

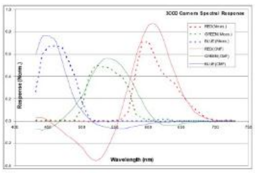


Malvar, H., He, L., Cutler, R.: High-quality linear interpolation for demosaicing of bayer-patterned color images. In: Procs. of ICASSP. (2004)

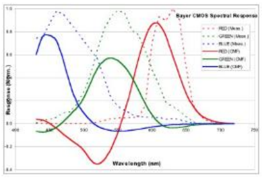
Human eye



Spectral response



3 chip CCD



Bayer CMOS

http://www.definitionmagazine.com/journal/2010/5/7/capturing-colour.html

Blooming

The buckets overflow…

Chromatic aberration

Different wavelengths have different refractive indices…

Interlacing

Some video cameras read even lines then odd…

Rolling shutter

Some cameras read out one line at a time:

Vignetting

The corners of images are darker than the middle:

Projection

Projection

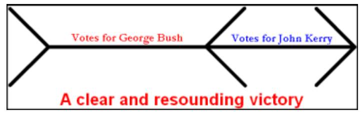
Readings
- Szeliski 2.1

Projection

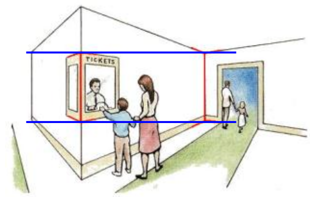
Readings

- Szeliski 2.1

Müller-Lyer Illusion

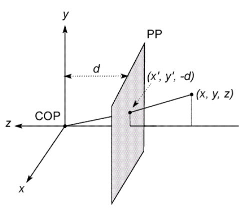


by Pravin Bhat



http://www.michaelbach.de/ot/sze_muelue/index.html

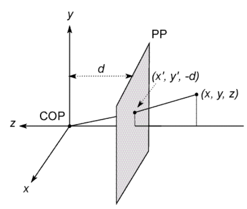
Modeling projection



The coordinate system

- We will use the pin-hole model as an approximation
- Put the optical center (**C**enter **Of P**rojection) at the origin
- Put the image plane (**P**rojection **P**lane) *in front* of the COP
  - Why?
- The camera looks down the *negative* z axis
  - we need this if we want right-handed-coordinates

Modeling projection



Projection equations

- Compute intersection with PP of ray from (x,y,z) to COP
- Derived using similar triangles (on board)

$$(x, y, z) \rightarrow (-d\frac{x}{z},\ -d\frac{y}{z},\ -d)$$

- We get the projection by throwing out the last coordinate:

$$(x, y, z) \rightarrow (-d\frac{x}{z},\ -d\frac{y}{z})$$

Homogeneous coordinates

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

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

Perspective Projection

Projection is a matrix multiply using homogeneous coordinates:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/d & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ -z/d \end{bmatrix} \Rightarrow (-d\frac{x}{z}, \ -d\frac{y}{z})$$

divide by third coordinate

This is known as **perspective projection**

- The matrix is the **projection matrix**

Perspective Projection

How does scaling the projection matrix change the transformation?

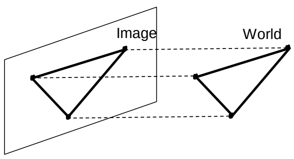
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/d & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ -z/d \end{bmatrix} \Rightarrow (-d\frac{x}{z}, \ -d\frac{y}{z})$$

$$\begin{bmatrix} -d & 0 & 0 & 0 \\ 0 & -d & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} -dx \\ -dy \\ z \end{bmatrix} \Rightarrow (-d\frac{x}{z}, \ -d\frac{y}{z})$$

Orthographic projection

Special case of perspective projection

- Distance from the COP to the PP is infinite



- Good approximation for telephoto optics
- Also called "parallel projection": (x, y, z) → (x, y)
- What's the projection matrix?

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

Telephoto lenses

Commonly used to make distant objects look closer than they really are.

Variants of orthographic projection

Scaled orthographic

- Also called “weak perspective”

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1/d \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1/d \end{bmatrix} \Rightarrow (dx, dy)$$

Affine projection

- Also called “paraperspective”

$$\begin{bmatrix} a & b & c & d \\ e & f & g & h \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

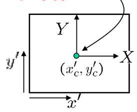
Camera parameters

A camera is described by several parameters

- Translation T of the optical center from the origin of world coords
- Rotation R of the image plane
- focal length f, principle point (x'_c, y'_c), pixel size (s_x, s_y)
- blue parameters are called “extrinsics,” red are “intrinsics”

Projection equation

$$\mathbf{x} = \begin{bmatrix} sx \\ sy \\ s \end{bmatrix} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \mathbf{\Pi X}$$



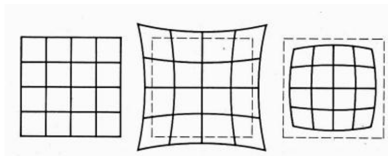
- The projection matrix models the cumulative effect of all parameters
- Useful to decompose into a series of operations

identity matrix

$$\mathbf{\Pi} = \underbrace{\begin{bmatrix} -fs_x & 0 & x'_c \\ 0 & -fs_y & y'_c \\ 0 & 0 & 1 \end{bmatrix}}_{\text{intrinsics}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\text{projection}} \underbrace{\begin{bmatrix} \mathbf{R}_{3x3} & \mathbf{0}_{3x1} \\ \mathbf{0}_{1x3} & 1 \end{bmatrix}}_{\text{rotation}} \underbrace{\begin{bmatrix} \mathbf{I}_{3x3} & \mathbf{T}_{3x1} \\ \mathbf{0}_{1x3} & 1 \end{bmatrix}}_{\text{translation}}$$

- The definitions of these parameters are **not** completely standardized
  - especially intrinsics—varies from one book to another

Distortion

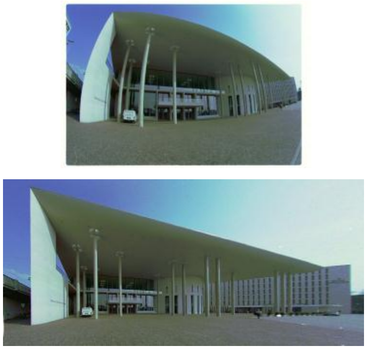


No distortion   Pin cushion   Barrel

Radial distortion of the image

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

Correcting radial distortion



from Helmut Dersch

Modeling distortion

Project $(\widehat{x}, \widehat{y}, \widehat{z})$ to "normalized" image coordinates

$$x'_n = \widehat{x}/\widehat{z}$$
$$y'_n = \widehat{y}/\widehat{z}$$

Apply radial distortion

$$r^2 = {x'_n}^2 + {y'_n}^2$$
$$x'_d = x'_n(1 + \kappa_1 r^2 + \kappa_2 r^4)$$
$$y'_d = y'_n(1 + \kappa_1 r^2 + \kappa_2 r^4)$$

Apply focal length translate image center

$$x' = f x'_d + x_c$$
$$y' = f y'_d + y_c$$

To model lens distortion

- Use above projection operation instead of standard projection matrix multiplication

Other types of projection

Lots of intriguing variants…

(I'll just mention a few fun ones)

360 degree field of view…

Basic approach

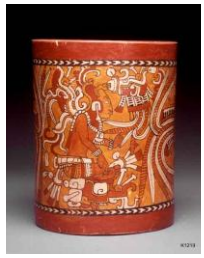
- Take a photo of a parabolic mirror with an orthographic lens (Nayar)
- Or buy one a lens from a variety of omnicam manufacturers…
  - See http://www.cis.upenn.edu/~kostas/omni.html

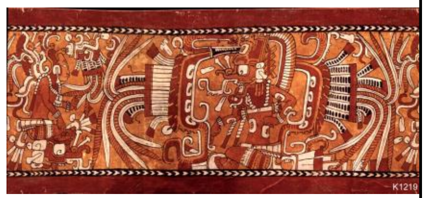
Tilt-shift

http://www.northlight-images.co.uk/article_pages/tilt_and_shift_ts-e.html

Titlt-shift images from Olivo Barbieri
and Photoshop imitations

Rotating sensor (or object)

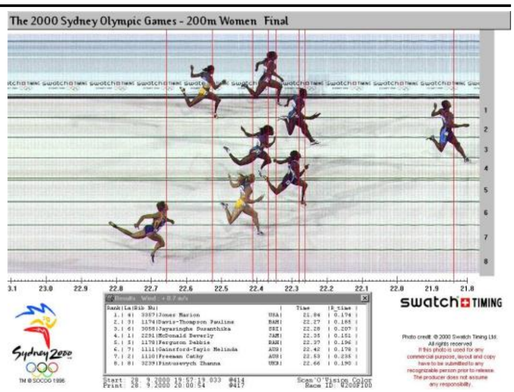




Rollout Photographs © Justin Kerr
http://research.famsi.org/kerrmaya.html

Also known as “cyclographs”, “peripheral images”

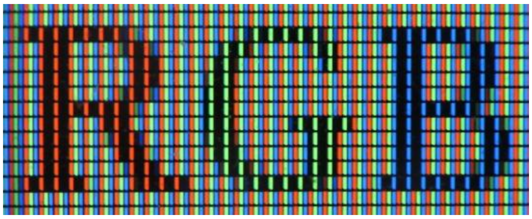
Photofinish



Displays

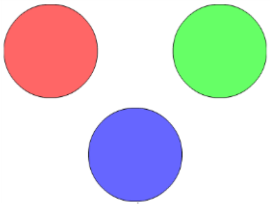
LCD

Monitors have an equal number of R, G, and B elements:



Displays

Mixing colors:



Displays

Most displays cannot generate the full spectrum of visible colors: