Advanced Multimedia

Images, cameras & color
Today: Images, Cameras & Color
How do we see the world?

• 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
• Pinhole model:
  – Captures **pencil of rays** – all rays through a single point
  – The point is called **Center of Projection (focal point)**
  – The image is formed on the **Image Plane**
Camera Obscura

Camera Obscura, Gemma Frisius, 1558

- The first camera
  - Known to Aristotle
Abelardo Morell

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
Dimensionality Reduction Machine (3D to 2D)

What have we lost?

- Angles
- Distances (lengths)
Funny things happen...

• Parallel lines converge at a vanishing point
  – Each direction in space has its own vanishing point
  – But parallels parallel to the image plane remain parallel
Funny things happen...

- Parallel lines converge at a vanishing point
  - Each direction in space has its own vanishing point
  - But parallels parallel to the image plane remain parallel

How do we construct the vanishing point/line?
Lengths can’t be trusted...
...but humans adopt!

Müller-Lyer Illusion

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

- Problem for architectural photography: converging verticals

Source: F. Durand
Perspective distortion

- Problem for architectural photography: converging verticals

  - Tilting the camera upwards results in converging verticals
  - Keeping the camera level, captures only the bottom portion of the building
  - Shifting the aperture upwards results in a picture of the entire subject

- Solution: view camera (aperture shifted w.r.t. film)

[Diagram showing perspective distortion and solutions]

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

Source: F. Durand
Perspective distortion

- Problem for architectural photography: converging verticals
- Result:

Source: F. Durand
Perspective distortion

• However, converging verticals work quite well for horror movies...
Pinhole camera model

• Pinhole model:
  – Captures *pencil of rays* – all rays through a single point
  – The point is called **Center of Projection (focal point)**
  – The image is formed on the **Image Plane**
Modeling projection

• The coordinate system
  – We will use the pin-hole model as an approximation
  – Put the optical center (Center Of Projection) at the origin
  – Put the image plane (Projection Plane) *in front* of the COP
    - Why?
Modeling projection

- **Projection equations**
  - Compute intersection with PP of ray from \((x, y, z)\) to COP
  - Derived using similar triangles

\[
(x, y, z) \rightarrow \left(-d \frac{x}{z}, -d \frac{y}{z}, -d\right)
\]

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

\[
(x, y, z) \rightarrow \left(-d \frac{x}{z}, -d \frac{y}{z}\right)
\]
Homogeneous coordinates

Trick: add one more coordinate:

\[
(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}
\]

homogeneous image coordinates

homogeneous scene coordinates

Converting \textit{from} homogeneous coordinates

\[
\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w) \quad \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)
\]
Perspective Projection

- Projection is a matrix multiply using homogeneous coordinates:

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

divide by third coordinate

This is known as **perspective projection**

- The matrix is the **projection matrix**
Building a real camera
Home-made pinhole camera

http://www.debevec.org/Pinhole/
Shrinking the aperture

- Why not make the aperture as small as possible?
  - Less light gets through
  - Diffraction effects...
Shrinking the aperture
The reason for lenses
Focus
Adding a lens

- A lens focuses light onto the film
Focus

When you turn the lens of a camera to focus it -- you're moving it closer or farther away from the film surface. As you move the lens, you can line up the focused real image of an object so it falls directly on the film surface.
• 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* $f$
• 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
Thin lens formula
Thin lens formula

Similar triangles everywhere!
Thin lens formula

Similar triangles everywhere!

\[ \frac{y'}{y} = \frac{D'}{D} \]
Thin lens formula

\[ \frac{y'}{y} = \frac{D'}{D} \]

\[ \frac{y'}{y} = \frac{D'-f}{f} \]

Similar triangles everywhere!
Thin lens formula

\[ \frac{1}{D'} + \frac{1}{D} = \frac{1}{f} \]

Any point satisfying the thin lens equation is in focus.
Varying Focus
Depth Of Field
Depth of Field

http://www.cambridgeincolour.com/tutorials/depth-of-field.htm
Aperture affects 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
Varying the aperture

Large aperture = small DOF

Small aperture = large DOF

Slide by Alyosha Efros
Nice Depth of Field effect
Field of View (Zoom)
Field of View (Zoom)

From London and Upton

Slide by Alyosha Efros
Field of View (Zoom)

From London and Upton
FOV depends on Focal Length

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

Smaller FOV = larger Focal Length
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

wide-angle  standard  telephoto

Source: F. Durand
Lens Flaws
Lens Flaws: Chromatic Aberration

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

Source: L Lazebnik
Lens flaws: Spherical aberration

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

Source: L. Lazebnik
Lens flaws: Vignetting

Source: L Laezenik
Digital camera

• 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)
Sampling and Quantization

FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.
CCD vs. CMOS

- **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.

- **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.

  [Diagram of CCD and CMOS]

  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.

  ![Diagram of CCD and CMOS](http://www.dalsa.com/shared/content/pdfs/CCD_vs_CMOS_Litwiller_2005.pdf)

  Source: A Efros
Issues with digital cameras

• 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

• Color artifacts
  • purple fringing from microlenses, artifacts from Bayer patterns
  • white balance

• More info online:
  – http://www.dpreview.com/

Slide by Steve Seitz
Historical context

- **Pinhole model:** Mozi (470-390 BCE), Aristotle (384-322 BCE)
- **Principles of optics (including lenses):** Alhacen (965-1039 CE)
- **Camera obscura:** Leonardo da Vinci (1452-1519), Johann Zahn (1631-1707)
- **First photo:** Joseph Nicéphore Niépce (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)
• 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
The Retina

Cross-section of eye

Cross section of retina

Ganglion axons
Ganglion cell layer
Bipolar cell layer
Receptor layer
Pigmented epithelium

Source: A Efros
Retina up-close

Source: A Efros
Two types of light-sensitive receptors

**Cones**
- cone-shaped
- less sensitive
- operate in high light
- color vision

**Rods**
- rod-shaped
- highly sensitive
- operate at night
- gray-scale vision

Source: A Efros
The famous sock-matching problem...

Source: A Efros
Distribution of Rods and Cones

Night Sky: why are there more stars off-center?

Source: A Efros
Electromagnetic Spectrum

Human Luminance Sensitivity Function

http://www.yorku.ca/eye/photopik.htm
Visible Light

Why do we see light of these wavelengths?

...because that's where the Sun radiates EM energy
Any patch of light can be completely described physically by its spectrum: the number of photons (per time unit) at each wavelength 400 - 700 nm.
The Physics of Light

Some examples of the spectra of light sources

A. Ruby Laser

B. Gallium Phosphide Crystal

C. Tungsten Lightbulb

D. Normal Daylight
The color of objects

- Colored light arriving at the camera involves two effects
  - The color of the light source
  - The color of the surface
  - Changes caused by different colored light sources can be large

Source: DA Forsyth
The Physics of Light

Some examples of the reflectance spectra of surfaces

- Red
- Yellow
- Blue
- Purple

Source: A Efros
There is no simple functional description for the perceived color of all lights under all viewing conditions, but …..

A helpful constraint:
Consider only physical spectra with normal distributions
The Psychophysical Correspondence

Mean ↔ Hue

# Photons

Wavelength

Source: A Efros
The Psychophysical Correspondence

Variance ↔ Saturation

# Photons

Wavelength

Source: A Efros
The Psychophysical Correspondence

Area $\leftrightarrow$ Brightness

# Photons

Wavelength

Source: A Efros
Physiology of Color Vision

Three kinds of cones:

- *Why are M and L cones so close?*
- *Why are there 3?*

Source: A Efros
Color Adaptation
Color Sensing in Camera (RGB)

- 3-chip vs. 1-chip: quality vs. cost

Why 3 colors?

http://www.cooldictionary.com/words/Bayer-filter.wiki
Color Sensing in Camera (RGB)

• 3-chip vs. 1-chip: quality vs. cost
• Why more green?

Why 3 colors?

http://www.cooldictionary.com/words/Bayer-filter.wikipedia

Slide by Steve Seitz
Color Sensing in Camera (RGB)

• 3-chip vs. 1-chip: quality vs. cost
• Why more green?

Why 3 colors?

http://www.cooldictionary.com/words/Bayer-filter.wikipedia
RGB color space

![RGB 24-bit color cube](image)

**FIGURE 6.8** RGB 24-bit color cube.

• RGB cube
  – Easy for devices
  – But not perceptual
  – Where do the grays live?
  – Where is hue and saturation?
HSV Cone

- Hue, Saturation, Value (Intensity/Brightness)
- Use rgb2hsv() and hsv2rgb() in Matlab