BSB 663
Image Processing

Image Formation and the Digital Camera

Aykut Erdem
Dept. of Computer Engineering
Hacettepe University

Acknowledgement: The course slides are mostly adapted from the slides prepared by Erkut Erdem, Steve Marschner and Anat Levin

Today
- Image formation
- Display devices and digital camera
- Digital images

What is an image?
- A photographic print
- A photographic negative
- This projection screen
- Some numbers in RAM

Today
- Image formation
- Display devices and digital camera
- Digital images

What is an image?
- A photographic print
- A photographic negative
- This projection screen
- Some numbers in RAM
An image is:

- A 2D distribution of intensity or color
- A function defined on a two-dimensional plane

\[ I : \mathbb{R}^2 \rightarrow \ldots \]

- Note: no mention of pixels yet
- To process images, must:
  - obtain images—capture the scenes via hardware
  - represent images—encode them numerically

Image Formation

Three Dimensional World \[ \rightarrow \] Two Dimensional Image Space

- What is measured in an image location?
  - viewpoint
  - illumination conditions
  - local geometry
  - local material properties

Dimensionality Reduction Machine (3D to 2D)

3D world \[ \rightarrow \] 2D image

What have we lost?
- Angles
- Distances (lengths)
Funny things happen…

- Parallel lines in the world intersect in the image at a vanishing point.

Parallel lines aren't...

The images of parallel lines intersect at the horizon.

Vanishing points and lines

Photo from online Tate collection

Vanishing points and lines

Slide credit: Alyosha Efros

Slide credit: Alyosha Efros

Figure by David Forsyth

Slide credit: J. Hays

Slide credit: J. Hays
Lengths can’t be trusted...

...but humans adopt!

Müller-Lyer Illusion

We don’t make measurements in the image plane

How do we see the world?

- Let’s design a camera
  - Idea 1: put a piece of film in front of an object
  - Why is there no picture appearing on the paper?

http://www.michaelbach.de/ot/3dp_muehle/index.html
Light rays from many different parts of the scene strike the same point on the paper.

Pinhole Camera

- Light from a scene passes through the pinhole (a single point) and projects an inverted image on the image plane.

Pinhole

A pinhole projects all rays through a common center of projection.
Camera Obscura (darkened room)

The first camera
- known to Aristotle
- Depth of the room is the effective focal length

Camera Obscura, Gemma Frisius, 1558

Camera Obscura used for Tracing

Lens Based Camera Obscura, 1568

Vermeer and The Camera Obscura

Officer and Laughing Girl, 1657

http://www.essentialvermeer.com/camera_obscura/Officer_and_Laughing_Girl.html

Camera Obscura

from BBC's Genius of Photography
**Pinhole Size**

Small pinhole - sharp but hard to collect enough light.

Larger pinhole - blur.

**Cameras and Lenses**

Shrinking aperture size

- Rays are mixed up
- Why the aperture cannot be too small?
  - Less light passes through (dark images)
  - Diffraction effect

- To make this picture, the hole of a camera was replaced with a thin metal disk pierced by a tiny pinhole, equivalent to an aperture of F/16. Only a few rays of light from each point on the subject get through the very small opening, producing a soft but acceptable close-up photograph. Because of the small size of the pinhole, the exposure had to be 20 seconds long.

- The web opening was much bigger than the pinhole, allowing more light in. However, this exposed the scene from each pinhole to light, necessitating a much brighter image overall.

**The reason for Lenses**

- Gather more light!
- But need to be focused

Slide credit: Fei-Fei Li
A lens is focused at a single depth

\[
\frac{1}{z_o} + \frac{1}{z_i} = \frac{1}{f}
\]

- \(z_o\): distance to the (focused) object
- \(z_i\): distance behind the lens at which the image is formed
- \(f\): focal length

Object at focus depth

Rays emerge from multiple object points
(circle of confusion) => the captured image is blurred

All rays emerge from a single object point => The captured image is sharp

Aperture

- Diameter of the lens opening (controlled by diaphragm)
- Controls depth of field
- Expressed as a fraction of focal length, in f-number
  - f/2.0 on a 50mm means that the aperture is 25mm
  - f/2.0 on a 100mm means that the aperture is 50mm
- Disconcerting: small f number = big aperture
- What happens to the area of the aperture when going from f/2.0 to f/4.0?
- Typical f numbers are
  - f/2.0, f/2.8, f/4, f/5.6, f/8, f/11, f/16, f/22, f/32

Object away from focus depth

Rays emerge from multiple object points

Expressed as a fraction of focal length, in f-number
Main effect of aperture

- **Depth of field**: Allowable depth variation in the scene that limits the circle of confusion to a tolerable number

![Depth of field diagram](image)

From Photography, London et al.

Depth of field

- We allow for some tolerance

![Depth of field diagram](image)
Light Field / Plenoptic Cameras

- Lytro ($399, Sept 2012)
- Multiple, "synthetic apertures"
- Allows focusing after capture

“Achieved by inserting a micro-lens array between the sensor and main lens, creating a plenoptic camera. Each microlens measures not just the total amount of light deposited at that location, but how much light arrives along each ray.”
Field of View (Zoom) = Cropping

Field of View / Focal Length

FOV depends of Focal Length

Changing FOV – magnification constant
Changing FOV – magnification constant

- Hitchcock zoom (“Vertigo effect”)

Exposure

- **Exposure**: How much light falls on sensor
- Get the right amount of light to sensor/film
- Main parameters:
  - Shutter speed: How long sensor is exposed to light
  - Aperture (area of lens): How much light can pass through from the lens
  - Sensitivity: How much light is needed by the sensor
  - Lighting conditions

Shutter speed

- Controls how long the film/sensor is exposed, i.e. the amount of light reaching the sensor
- Pretty much linear effect on exposure
- Usually in fraction of a second:
  - 1/30, 1/60, 1/125, 1/250, 1/500
  - Get the pattern?
- Faster shutter (e.g. 1/500th sec) = less light
- Slower shutter (e.g. 1/30th sec) = more light
- On a normal lens, normal humans can hand-hold down to 1/60
  - In general, the rule of thumb says that the limit is the inverse of focal length, e.g. 1/500 for a 500mm lens

Exposure

- **Exposure**: How much light falls on sensor
- Get the right amount of light to sensor/film
- Main parameters:
  - Shutter speed: How long sensor is exposed to light
  - Aperture (area of lens): How much light can pass through from the lens
  - Sensitivity: How much light is needed by the sensor
  - Lighting conditions
Shutter speed

- Short exposure after contrast adjustment-noise
- Medium exposure
- Long exposure-saturation

Main effect of slower shutter speed

- For dynamic scenes, the shutter speed also determines the amount of motion blur in the resulting picture.
- Camera shake

Image taken with a tripod
Image taken with a hand held camera

Effect of Shutter Speed

- Freezing motion

Walking people
Running people
Car
Fast train

From Photography, London et al.
Today

• Image formation
• Display devices and digital camera
• Digital images

Representative display technologies

Computer displays
• Raster CRT display
• LCD display

Printers
• Laser printer
• Inkjet printer

Cathode ray tube

• First widely used electronic display
  – developed for TV in the 1920s–1930s

Raster CRT display

• Scan pattern fixed in display hardware
• Intensity modulated to produce image
• Originally for TV
  – (continuous analog signal)
• For computer, intensity determined by contents of framebuffer
**LCD flat panel or projection display**

- Principle: block or transmit light by twisting its polarization
- Intermediate intensity levels possible by partial twist
- Fundamentally raster technology
- Fixed format

**Raster display system**

- Screen image defined by a 2D array in RAM
  - for CRT, read out and convert to analog in sync with scan
- In most systems today, it’s in a separate memory
- The memory area that maps to the screen is called the frame buffer

**Color displays**

- Operating principle: humans are trichromatic
  - match any color with blend of three
  - therefore, problem reduces to producing 3 images and blending
- Additive color
  - blend images by sum
  - e.g. overlapping projection
  - e.g. unresolved dots
  - R, G, B make good primaries

**Color displays**

- CRT: phosphor dot pattern to produce finely interleaved color images
- LCD: interleaved R,G,B pixels
**Laser printer**

- Xerographic process
- Like a photocopier but with laser-scanned raster as source image
- Key characteristics
  - Image is binary
  - Resolution is high
  - Very small, isolated dots are not possible

**Inkjet printer**

- Liquid ink sprayed in small drops
  - Very small—measured in picoliters
- Head with many jets scans across paper
- Key characteristics:
  - Image is binary (drop or no drop; no partial drops)
  - Isolated dots are reproduced well

**Digital camera**

- A raster input device
- Image sensor contains 2D array of photosensors

**Digital camera**

- Color typically captured using color mosaic
- Demosaicing
Today
- Image formation
- Display devices and digital camera
- Digital images

Raster image representation
- All these devices suggest 2D arrays of numbers
- Big advantage: represent arbitrary images
  - approximate arbitrary functions with increasing resolution
  - works because memory is cheap (brute force approach!)

Meaning of a raster image
- Meaning of a given array is a function on 2D
- Define meaning of array = result of output device!
  - that is, piecewise constant for LCD, blurry for CRT
  - but: we don’t have just one output device
  - but: want to define images we can’t display (e.g. too big)
- Abstracting from device, problem is reconstruction
  - image is a sampled representation
  - pixel means “this is the intensity around here”
    • LCD: intensity is constant over square regions
    • CRT: intensity varies smoothly across pixel grid

Image Representation
- Discretization
  - in image space - sampling
  - In image brightness - quantization
**Image Representation**

- **Digital image**: 2D discrete function \( f \)
- **Pixel**: Smallest element of an image \( f(x,y) \)

**Datatypes for raster images**

- **Bitmaps**: boolean per pixel (1 bpp): \( I : \mathbb{R}^2 \rightarrow \{0,1\} \)
  - interp. = black and white; e.g. fax
- **Grayscale**: integer per pixel: \( I : \mathbb{R}^2 \rightarrow [0,1] \)
  - interp. = shades of gray; e.g. black-and-white print
  - precision: usually byte (8 bpp); sometimes 10, 12, or 16 bpp
- **Color**: 3 integers per pixel: \( I : \mathbb{R}^2 \rightarrow [0,1]^3 \)
  - interp. = full range of displayable color; e.g. color print
  - precision: usually byte \( \{3\} \) (24 bpp)
  - sometimes 16 (5+6+5) or 30 or 36 or 48 bpp
  - indexed color: a fading idea

**Image Representation**

- **Digital image**: 2D discrete function \( f \)
- **Pixel**: Smallest element of an image \( f(x,y) \)

**Datatypes for raster images**

- **Floating point**: \( I : \mathbb{R}^2 \rightarrow \mathbb{R}_+ \) or \( I : \mathbb{R}^2 \rightarrow \mathbb{R}_+^3 \)
  - more abstract, because no output device has infinite range
  - provides high dynamic range (HDR)
  - represent real scenes independent of display
  - becoming the standard intermediate format in graphics processors
- **Clipping and white point**
  - common to compute FP, then convert to integer
  - full range of values may not “fit” in display’s output range
  - simplest solution: choose a maximum value, scale so that value becomes full intensity (\(2^n-1\) in an \(n\)-bit integer image)
Intensity encoding in images

- What do the numbers in images (pixel values) mean?
  - they determine how bright that pixel is
  - bigger numbers are (usually) brighter

Datatypes for raster images

- For color or grayscale, sometimes add alpha channel
  - describes transparency of images

Storage requirements for images

- 1024x1024 image (1 megapixel)
  - bitmap: 128KB
  - grayscale 8bpp: 1MB
  - grayscale 16bpp: 2MB
  - color 24bpp: 3MB
  - floating-point HDR color: 12MB

Converting pixel formats

- Color to gray
  - could take one channel (blue, say)
    - leads to odd choices of gray value
  - combination of channels is better
    - but different colors contribute
derrently to lightness
    - which is lighter, full blue or full green?
    - good choice: gray = 0.2 R + 0.7 G + 0.1 B
    - more on this in color, later on

Same pixel values.

Same luminance!
Converting pixel precision

- Up is easy; down loses information—be careful

Next class

- Color perception
- Color spaces