BBM 413
Fundamentals of Image Processing

Erkut Erdem
Dept. of Computer Engineering
Hacettepe University

Image Formation and the Digital Camera

Acknowledgement: The course slides are mostly adapted from the slides prepared by 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
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?
  – brightness
  – color

Figures: Francis Crick, The Astonishing Hypothesis,
Image Formation

Three Dimensional World → Two Dimensional Image Space

• What is measured in an image location?
  - brightness
  - color

viewpoint
<< illumination conditions
surface properties (local geometry and local material properties)

Figures: Francis Crick, The Astonishing Hypothesis,
Image Formation

Images cannot exist without light!

Why is there no image on a white piece of paper?

It receives light from all directions

From Photography, London et al.
Pinhole

A pinhole projects all rays through a common center of projection.

From Photography, London et al.
Pinhole
Pinhole Size?

Photograph made with small pinhole

Small pinhole—sharp but hard to collect enough light

Photograph made with larger pinhole

Larger pinhole—Blur

From Photography, London et al.
Pinhole Size

small hole => sharp, but don’t collect enough light (noise)
Solution: light refraction!
Lenses

• gather more light!

• But need to be focused

To make this picture, the lens of a camera was replaced with a thin metal disk pierced by a tiny pinhole, equivalent in size to an aperture of f/182. Only a few rays of light from each point on the subject got through the tiny opening, producing a soft but acceptably clear photograph. Because of the small size of the pinhole, the exposure had to be 6 sec long.

This time, using a simple convex lens with an f/16 aperture, the scene appeared sharper than the one taken with the smaller pinhole, and the exposure time was much shorter, only 1/100 sec. The lens opening was much bigger than the pinhole, letting in far more light, but it focused the rays from each point on the subject precisely so that they were sharp on the film.

From Photography, London et al.
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

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

Figure 2.19
A lens is focused at a single depth

Object away from focus depth

Rays emerge from multiple object points (circle of confusion) => the captured image is blurred
A lens is focused at a single depth

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

[Diagram of aperture sizes: Full aperture, Medium aperture, Stopped down]
Main effect of aperture

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

From Photography, London et al.
Depth of field

Image of object in focus - sharp (all rays hitting a single sensor point emerge from a single point on the object)
Depth of field

Image of object in focus- sharp (all rays hitting a single sensor point emerge from a single point on the object)

Image of an object away from focus depth- blurred (rays hitting a single sensor point emerge from multiple points on the object)
Depth of field

• We allow for some tolerance

Max acceptable circle of confusion

Depth of focus

Depth of field

Point in focus

Object with texture
Depth of Field
Depth of Field

Portait

Large Aperture
f/2.8

Small Aperture
f/5.6

Shallow Depth of Field

Landscape

Large Depth of Field

http://photographertips.net
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. 1/500\textsuperscript{th} sec) = less light

• Slower shutter (e.g. 1/30\textsuperscript{th} 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
Shutter speed

Short exposure - dark  medium exposure  long exposure - saturation
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
Main effect of slower shutter speed

• For dynamic scenes, the shutter speed also determines the amount of *motion blur* in the resulting picture.

• Scene motion

From Photography, London et al.
Effect of Shutter Speed

- Freezing motion

Walking people | Running people | Car | Fast train
---|---|---|---
1/125 | 1/250 | 1/500 | 1/1000
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
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 $f(x,y)$
Image Representation

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

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.92</td>
<td>0.93</td>
<td>0.94</td>
<td>0.97</td>
<td>0.62</td>
<td>0.37</td>
<td>0.85</td>
<td>0.97</td>
<td>0.93</td>
<td>0.92</td>
</tr>
<tr>
<td>0.95</td>
<td>0.89</td>
<td>0.82</td>
<td>0.89</td>
<td>0.56</td>
<td>0.31</td>
<td>0.75</td>
<td>0.92</td>
<td>0.81</td>
<td>0.95</td>
</tr>
<tr>
<td>0.89</td>
<td>0.72</td>
<td>0.51</td>
<td>0.55</td>
<td>0.51</td>
<td>0.42</td>
<td>0.57</td>
<td>0.41</td>
<td>0.49</td>
<td>0.91</td>
</tr>
<tr>
<td>0.96</td>
<td>0.95</td>
<td>0.88</td>
<td>0.94</td>
<td>0.56</td>
<td>0.46</td>
<td>0.91</td>
<td>0.87</td>
<td>0.90</td>
<td>0.97</td>
</tr>
<tr>
<td>0.71</td>
<td>0.81</td>
<td>0.81</td>
<td>0.87</td>
<td>0.57</td>
<td>0.37</td>
<td>0.80</td>
<td>0.88</td>
<td>0.89</td>
<td>0.79</td>
</tr>
<tr>
<td>0.49</td>
<td>0.62</td>
<td>0.60</td>
<td>0.58</td>
<td>0.50</td>
<td>0.60</td>
<td>0.58</td>
<td>0.50</td>
<td>0.61</td>
<td>0.45</td>
</tr>
<tr>
<td>0.86</td>
<td>0.84</td>
<td>0.74</td>
<td>0.58</td>
<td>0.51</td>
<td>0.39</td>
<td>0.73</td>
<td>0.92</td>
<td>0.91</td>
<td>0.49</td>
</tr>
<tr>
<td>0.96</td>
<td>0.67</td>
<td>0.54</td>
<td>0.85</td>
<td>0.48</td>
<td>0.37</td>
<td>0.88</td>
<td>0.90</td>
<td>0.94</td>
<td>0.82</td>
</tr>
<tr>
<td>0.69</td>
<td>0.49</td>
<td>0.56</td>
<td>0.66</td>
<td>0.43</td>
<td>0.42</td>
<td>0.77</td>
<td>0.73</td>
<td>0.71</td>
<td>0.90</td>
</tr>
<tr>
<td>0.79</td>
<td>0.73</td>
<td>0.90</td>
<td>0.67</td>
<td>0.33</td>
<td>0.61</td>
<td>0.69</td>
<td>0.79</td>
<td>0.73</td>
<td>0.93</td>
</tr>
<tr>
<td>0.91</td>
<td>0.94</td>
<td>0.89</td>
<td>0.49</td>
<td>0.41</td>
<td>0.78</td>
<td>0.78</td>
<td>0.77</td>
<td>0.89</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Figure: M. J. Black
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
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 differently 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.

\[
\begin{array}{c|c|c}
\text{Blue} & \text{Green} & \text{Gray} \\
\hline
\text{Same luminance?}
\end{array}
\]
Converting pixel precision

- Up is easy; down loses information—be careful

[photo: Philip Greenspun]
Next class

• Color perception
• Color spaces