BBM 413
Fundamentals of Image Processing

Erkut Erdem
Dept. of Computer Engineering
Hacettepe University

Introduction
Today

• What is image processing?
• Image formation
• Digital images
Today

• But, first logistics..
  – About the class
  – Organization of this course
About this course

• This course is an advanced level undergraduate course about the fundamentals of image processing.

• Requirements
  – Programming skills
  – Good math background (Calculus, Linear Algebra, Statistical Methods)
  – Little or no prior knowledge of image processing techniques

• BBM 415 Introduction to Programming Practicum
  – The students will gain hand-on experience via a set of programming assignments.
About this course (cont’d.)

• **Goals of the course:**
  – to provide an introduction to students who wish to specialize in interrelated disciplines like image processing, computer vision and computational photography

• **Skills to develop:**
  – a foundational understanding and knowledge of concepts that underlie image processing

• **What is image processing?**
  – What does image processing deal with?
  – Computational analysis of low and mid-level vision
BBM 413-415 Team

Instructor
Erkut ERDEM
erkut@cs.hacettepe.edu.tr

TA
Özge Yalcinkaya
ozge@cs.hacettepe.edu.tr
Textbooks and Reference Material

• Lecture notes and handouts
• Papers and journal articles
Communication

• The course webpage will be updated regularly throughout the semester with lecture notes, programming and reading assignments and important deadlines.

http://web.cs.hacettepe.edu.tr/~erkut/bbm413.f18
Getting Help

• **Office hours**
  – Wednesdays, 4-6 pm

• **BBM 415 Image Processing Practicum**
  – Course related recitations, practice with example codes, etc.

• **Communication**
  – Announcements and course related discussions through https://piazza.com/hacettepe.edu.tr/fall2018/bbm413
BBM 415 Image Processing Practicum

• **Programming assignments (PAs) - 20% each**
  – Four programming assignments throughout the semester.
  – Each assignment has a well-defined goal such as solving a specific problem.
  – You must work alone on all assignments stated unless otherwise.

• **A set of quizzes – 20%**
  – Lowest 2 quiz grades will be dropped

• **Important Dates (Tentative)**
  – PA 1 due: November 2\textsuperscript{nd}
  – PA 2 due: November 23\textsuperscript{rd}
  – PA 3 due: December 14\textsuperscript{th}
  – PA 4 due: December 28\textsuperscript{th}
Policies

• **Work groups**
  – You must work alone on all assignments stated unless otherwise

• **Submission**
  – Assignments due at 23:59 on Friday evenings
  – Electronic submissions (no exceptions!)
  – Submission details will be announced soon.

• **Lateness penalties**
  – Get penalized **10% per day**
  – No late submission later than **3 days after due date**
Course work and grading

• **Course project (25%)**
  – done in groups of 2-3 students
  – [https://web.cs.hacettepe.edu.tr/~erkut/bbm413.f18/project/project.html](https://web.cs.hacettepe.edu.tr/~erkut/bbm413.f18/project/project.html)

• **Midterm exam (30%)**
  – Closed book and notes including reading assignments
  – In class on December 4th

• **Final exam (40%)**
  – Closed book and notes
  – To be scheduled by Registrar

• **Class participation (5%)**
Course Overview

– Introduction, and Image formation (1 week)
– Color and Point operations (1 week)
– Spatial filtering (1 week)
– Frequency Domain Techniques (2 weeks)
– Image pyramids and wavelets (1 week)
– Gradients, edges, contours (1 week)
– Image smoothing (1 week)
– Image segmentation (1 week)
– Deep learning basics (1 week)
– Convolutional neural networks and their applications (1 week)
Today

• What is image processing?
  – What does it mean, to see?
  – Vision as a computational problem
  – Sample image processing problems

• Image formation

• Digital images
Image Processing

- Signal Processing
- Comp. Photography
- Computer Vision
- Graphics
- Machine Learning
- Statistics
- Applied Math

Credit: P. Milanfar
What does it mean, to see?

• “The plain man’s answer (and Aristotle’s, too) would be, to know what is where by looking. In other words, vision is the process of discovering from images what is present in the world, and where it is.” David Marr, Vision, 1982

• Our brain is able to use an image as an input, and interpret it in terms of objects and scene structures.
What does Salvador Dali’s Study for the Dream Sequence in Spellbound (1945) say about our visual perception?

We see a two dimensional image

But, we perceive depth information

light reflected on the retina

converging lines

shadows of the eye
Why does vision appear easy to humans?

• Our brains are specialized to do vision.
• Nearly half of the cortex in a human brain is devoted to doing vision (cf. motor control ~20-30%, language ~10-20%)

• “Vision has evolved to convert the ill-posed problems into solvable ones by adding premises: assumptions about how the world we evolved in is, on average, put together”
  Steven Pinker, How the Mind Works, 1997

• Gestalt Theory (Laws of Visual Perception),
  Max Wertheimer, 1912
Why does vision appear easy to humans?

When a user takes a photo, the app should check whether they're in a national park...

Sure, easy GIS lookup. Gimme a few hours.

...and check whether the photo is of a bird.

I'll need a research team and five years.

In CS, it can be hard to explain the difference between the easy and the virtually impossible.

http://xkcd.com/1425/
Computer Vision

• “Vision is a process that produces from images of the external world a description that is useful to the viewer and not cluttered with irrelevant information”
  ~David Marr

• The goal of Computer Vision:
  To develop artificial machine vision systems that make inferences related to the scene being viewed through the images acquired with digital cameras.

Things that are easy for us are difficult for computers and viceversa ~ Marvin Minsky
THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".
Origins of computer vision


(a) Original picture.

(b) Differentiated picture.

(c) Line drawing.

(d) Rotated view.

Slide credit: S. Lazebnik
Vision: a very difficult computational problem, at several levels of understanding

• Vision as an information processing task [David Marr, 1982]

• Three levels of understanding:
  1. Computational theory
     – What is computed? Why it is computed?
  2. Representation and Algorithm
     – How it is computed?
     – Input, Output, Transformation
  3. Physical Realization
     – Hardware
Reading Assignment #1

Visual Modules and the Information Flow

- Visual perception as a data-driven, bottom-up process (traditional view since D. Marr)
- Unidirectional information flow
- Simple low-level cues $\Rightarrow$ Complex abstract perceptual units
Vision modules can be categorized into three groups according to their functionality:

- Low-level vision: filtering out irrelevant image data
- Mid-level vision: grouping pixels or boundary fragments together
- High-level vision: complex cognitive processes
**Fundamentals of Image Processing**

- What is a digital image, how it is formed?
- How images are represented in computers?
- Why we process images?
- How we process images?
Image Formation

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

Figures: Francis Crick, The Astonishing Hypothesis, 1995
Image Formation

- Discretization
  - in image space - sampling
  - In image brightness - quantization

Image Representation

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

Figure: M. J. Black
**Image Representation**

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

![Pixel Grid](image)

<table>
<thead>
<tr>
<th>0.92</th>
<th>0.93</th>
<th>0.94</th>
<th>0.97</th>
<th>0.62</th>
<th>0.37</th>
<th>0.85</th>
<th>0.97</th>
<th>0.93</th>
<th>0.92</th>
<th>0.99</th>
</tr>
</thead>
<tbody>
<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>
<td>0.91</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>
<td>0.92</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>
<td>0.95</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>
<td>0.85</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>
<td>0.33</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>
<td>0.74</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>
<td>0.93</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>
<td>0.99</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>
<td>0.97</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>
<td>0.93</td>
</tr>
</tbody>
</table>

*Figure: M. J. Black*
• Two types of receptor cells in retina:
  • Cone Receptor cells: 6-7 million → function in bright light, color sensitive, fine detail
  • Rod receptor cells: 75-150 million → function in dim light, color insensitive, coarse detail

Figure: Francis Crick, The Astonishing Hypothesis, 1995
Hierarchy of Visual Areas

- There are many different neural connections between different visual areas.

Visual Modules and the Information Flow

• Vision modules can be categorized into three groups according to their functionality:
  – Low-level vision: filtering out irrelevant image data
  – Mid-level vision: grouping pixels or boundary fragments together
  – High-level vision: complex cognitive processes

Subject matter of this course
Image Filtering

• Instagram
Image Filtering

- Filtering out the irrelevant information

\[ f(x) = u(x) + n(x) \]

- Image denoising, image sharpening, image smoothing, image deblurring, etc.

- Edge detection

- Required for many other image manipulation tasks
Edge Detection

- Edges: abrupt changes in the intensity
  - Uniformity of intensity or color
- Edges to object boundaries
Image Filtering

- **Difficulty:** Some of the irrelevant image information have characteristics similar to those of important image features.
Image Smoothing - A Little Bit of History

- Gaussian Filtering / linear diffusion
  - the most widely used method

\[
\frac{\partial u}{\partial t} = \nabla \cdot (\nabla u) = \nabla^2 u
\]

- mid 80’s – unified formulations
  - methods that combine smoothing and edge detection
  - Geman & Geman’84, Blake & Zisserman’87, Mumford & Shah’89, Perona & Malik’90
Image Denoising

- Images are corrupted with 70% salt-and-pepper noise
Non-local Means Denoising

Figure 1. Scheme of NL-means strategy. Similar pixel neighborhoods give a large weight, $w(p,q_1)$ and $w(p,q_2)$, while much different neighborhoods give a small weight $w(p,q_3)$.

Preserve fine image details and texture during denoising

A. Buades, B. Coll, J. M. Morel, A non-local algorithm for image denoising, CVPR, 2005
Context-Guided Smoothing

• Use local image context to steer filtering

Preserve main image structures during filtering

E. Erdem and S. Tari, Mumford-Shah Regularizer with Contextual Feedback, JMIV, 2009
Structure-Preserving Smoothing

L. Karacan, E. Erdem and A. Erdem, Structure Preserving Image Smoothing via Region Covariances, TOG, 2013
Structure-Preserving Smoothing

L. Karacan, E. Erdem and A. Erdem, Structure Preserving Image Smoothing via Region Covariances, TOG, 2013
Image Abstraction

L. Karacan, E. Erdem and A. Erdem, Structure Preserving Image Smoothing via Region Covariances, TOG, 2013
Detail Enhancement

L. Karacan, E. Erdem and A. Erdem, Structure Preserving Image Smoothing via Region Covariances, TOG, 2013
Artistic Stylizations

Abstract
This paper presents a fast deblurring method that produces a deblurring result from a single image of moderate size in a few seconds. We accelerate both latent image estimation and kernel estimation in an iterative deblurring process by introducing a novel prediction step and working with image derivatives rather than pixel values. In the prediction step, we use simple image processing techniques to predict strong edges from an estimated latent image, which will be solely used for kernel estimation. With this approach, a computationally efficient Gaussian prior becomes sufficient for deconvolution to estimate the latent image, as small deconvolution artifacts can be suppressed in the prediction. For kernel estimation, we formulate the optimization function using image derivatives, and accelerate the numerical process by reducing the number of Fourier transforms needed for a conjugate gradient method. We also show that the formulation results in a smaller condition number of the numerical system than the use of pixel values, which gives faster convergence. Experimental results demonstrate that our method runs an order of magnitude faster than previous work, while the deblurring quality is comparable. GPU implementation facilitates further speed-up, making our method fast enough for practical use.

1 Introduction
A motion blur is a common artifact that produces disappointing blurry images with inevitable information loss. It is caused by the nature of imaging sensors that accumulate incoming lights for an amount of time to produce an image. During exposure, if the camera sensor moves, a motion blurred image will be obtained. If a motion blur is shift-invariant, it can be modeled as the convolution of a latent image with a motion blur kernel, where the kernel describes the trace of a sensor. Then, removing a motion blur from an image becomes a deconvolution operation. In non-blind deconvolution, the motion blur kernel is given and the problem is to recover the latent image from a blurry version using the kernel. In blind deconvolution, the kernel is unknown and the recovery of the latent image becomes more challenging. In this paper, we solve the blind deconvolution problem of a single image, where both blur kernel and latent image are estimated from an input blurred image.

Single-image blind deconvolution is an ill-posed problem because the number of unknowns exceeds the number of observed data. Early approaches imposed constraints on motion blur kernels and used parameterized forms for the kernels [Chen et al. 1996; Chan and Wong 1998; Yitzhaky et al. 1998; Rav-Acha and Peleg 2005]. Recently, several methods were proposed to handle a more general motion blur given a single image [Fergus et al. 2006; Jia 2007; Shan et al. 2008]. While these methods can produce excellent deblurring results, they necessitate intensive computation. It usually takes more than several minutes for the methods to deblur a single image of moderate size.

Most blind deconvolution methods take an iterative process that alternatingly optimizes the motion blur kernel and the latent image. In the process, the blur kernel is obtained from the estimated latent image and the given blurred image. The kernel is then used to estimate the latent image by applying non-blind deconvolution to the given blurred image. The new estimated latent image is used for kernel estimation in the next iteration. The intensive computation of previous methods stems from the complicated methods used for kernel estimation and latent image estimation. Optimization involving large matrices and vectors is needed for kernel estimation, and sophisticated optimization techniques are necessary to handle non-blind deconvolution with non-linear priors.

This paper presents a fast blind deconvolution method that produces a deblurring result from a single image in only a few seconds. The high speed of our method is enabled by accelerating both kernel estimation and latent image estimation steps in the iterative deblurring process. Our method produces motion deblurring results with...
Image deblurring

Figure 5. Test results on the GOPRO dataset. From top to bottom: Blurry images, results of Sun et al. \cite{26}, results of Kim and Lee \cite{15}, and results of the proposed method.
Image superresolution

input

4x output

W.-S. Lai, J.-B. Huang, N. Ahuja and M.-H. Yang,
Deep Laplacian Pyramid Networks for Fast and Accurate Super-Resolution, CVPR 2017
Image superresolution

input

4x original

W.-S. Lai, J.-B. Huang, N. Ahuja and M.-H. Yang,
Deep Laplacian Pyramid Networks for Fast and Accurate Super-Resolution, CVPR 2017
Image superresolution

input

4x output

W.-S. Lai, J.-B. Huang, N. Ahuja and M.-H. Yang,
Deep Laplacian Pyramid Networks for Fast and Accurate Super-Resolution, CVPR 2017
Image superresolution

input

4x original

W.-S. Lai, J.-B. Huang, N. Ahuja and M.-H. Yang,
Deep Laplacian Pyramid Networks for Fast and Accurate Super-Resolution, CVPR 2017
Image Segmentation

- Partition an image into meaningful regions that are likely to correspond to objects exist in the image

Grouping of pixels according to what criteria?

high-level object specific knowledge matters!

Figures: A. Erdem
Normalized Cuts

- A graph-theoretic formulation for segmentation

Normalized Cuts
From contours to regions

- **State-of-the-art:** gPb-owt-ucm segmentation algorithm
From contours to regions

• **State-of-the-art:** gPb-owt-ucm segmentation algorithm

Prior-Shape Guided Segmentation

• Incorporate prior shape information into the segmentation process

E. Erdem, S. Tari, and L. Vese, Segmentation Using The Edge Strength Function as a Shape Prior within a Local Deformation Model, ICIP 2009
Image Inpainting

- Reconstructing lost or deteriorated parts of images

What do these examples demonstrate?
Image Resizing

- Resize an image to arbitrary aspect ratios
Image Retargetting

• automatically resize an image to arbitrary aspect ratios while preserving **important image features**

How we define the importance?

S. Avidan and A. Shamir, Seam Carving for Content-Aware Image Resizing, SIGGRAPH, 2007
Image Retargeting

S. Avidan and A. Shamir, Seam Carving for Content-Aware Image Resizing, SIGGRAPH, 2007
Image Retargeting

L. Karacan, E. Erdem and A. Erdem, Structure Preserving Image Smoothing via Region Covariances, TOG, 2013
J. Johnson, A. Alahi and L. Fei Fei, Perceptual losses for real-time style transfer and super-resolution, ECCV 2016
Style Transfer

J. Johnson, A. Alahi and L. Fei Fei, Perceptual losses for real-time style transfer and super-resolution, ECCV 2016
J. Johnson, A. Alahi and L. Fei Fei, Perceptual losses for real-time style transfer and super-resolution, ECCV 2016
J. Johnson, A. Alahi and L. Fei Fei, Perceptual losses for real-time style transfer and super-resolution, ECCV 2016
Style Transfer

J. Johnson, A. Alahi and L. Fei Fei, Perceptual losses for real-time style transfer and super-resolution, ECCV 2016
Style Transfer

J. Johnson, A. Alahi and L. Fei Fei, Perceptual losses for real-time style transfer and super-resolution, ECCV 2016
Style Transfer

F. Luan, S. Paris, E. Shechtman and K. Bala, Deep Photo Style Transfer, CVPR 2017
Style Transfer

F. Luan, S. Paris, E. Shechtman and K. Bala, Deep Photo Style Transfer, CVPR 2017
Style Transfer

F. Luan, S. Paris, E. Shechtman and K. Bala, Deep Photo Style Transfer, CVPR 2017
Today

• What is image processing?
• Image formation
• Digital images
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

<< viewpoint
illumination conditions
local geometry
local material properties

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,
A photon’s life choices

• Absorption
• Diffusion
• Reflection
• Transparency
• Refraction
• Fluorescence
• Subsurface scattering
• Phosphorescence
• Interreflection
A photon’s life choices

- Absorption
- Diffusion
- Reflection
- Transparency
- Refraction
- Fluorescence
- Subsurface scattering
- Phosphorescence
- Interreflection
A photon’s life choices

- Absorption
- **Diffuse Reflection**
- Reflection
- Transparency
- Refraction
- Fluorescence
- Subsurface scattering
- Phosphorescence
- Interreflection
A photon’s life choices

- Absorption
- Diffusion
- **Specular Reflection**
- Transparency
- Refraction
- Fluorescence
- Subsurface scattering
- Phosphorescence
- Interreflection
A photon’s life choices

- Absorption
- Diffusion
- Reflection
- **Transparency**
- Refraction
- Fluorescence
- Subsurface scattering
- Phosphorescence
- Interreflection
A photon’s life choices

• Absorption
• Diffusion
• Reflection
• Transparency
• **Refraction**
• Fluorescence
• Subsurface scattering
• Phosphorescence
• Interreflection
A photon’s life choices

- Absorption
- Diffusion
- Reflection
- Transparency
- Refraction
- **Fluorescence**
- Subsurface scattering
- Phosphorescence
- Interreflection
A photon’s life choices

- Absorption
- Diffusion
- Reflection
- Transparency
- Refraction
- Fluorescence
- **Subsurface scattering**
- Phosphorescence
- Interreflection
A photon’s life choices

- Absorption
- Diffusion
- Reflection
- Transparency
- Refraction
- Fluorescence
- Subsurface scattering
- Phosphorescence
- Interreflection
A photon’s life choices

• Absorption
• Diffusion
• Reflection
• Transparency
• Refraction
• Fluorescence
• Subsurface scattering
• Phosphorescence
• **Interreflection**
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.
• 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

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

From Photography, London et al.
Pinhole camera

• Add a barrier to block off most of the rays
  – This reduces blurring
  – The opening is known as the **aperture**
  – How does this transform the image?
Camera Obscura

- Basic principle known to Mozi (470-390 BC), Aristotle (384-322 BC)
- Drawing aid for artists: described by Leonardo da Vinci (1452-1519)
Camera Obscura
Camera Obscura

- The first camera
  - How does the aperture size affect the image?
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.
small hole => sharp, but doesn’t collect enough light (noise)
larger hole => easy to collect enough light, but blur occurs
Pinhole Size

- Why not make the aperture as small as possible?
  - Less light gets through
  - Diffraction effects...
Solution: light refraction!

From Photography, London et al.
• 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.
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)
- Real cameras use many lenses together (to correct for aberrations)
Thin lenses

- Thin lens equation: \( \frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f} \)

  - Any object point satisfying this equation is in focus
A lens is focused at a single depth

\[ \frac{1}{z_0} + \frac{1}{z_i} = \frac{1}{f} \]

\( z_0 \): distance to the (focused) object
\( z_i \): distance behind the lens at which the image is formed
\( f \): focal length

Object at focus depth

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

\[
\frac{1}{z_o} + \frac{1}{z_i} = \frac{1}{f}
\]
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.
• Changing the aperture size affects depth of field
  – A smaller aperture increases the range in which the object is approximately in focus

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
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
Depth of Field
Depth of Field

Portraits:
- Shallow Depth of Field
- Large Depth of Field

Landsacpes:
- Large Aperture (f/2.8)
- Small Aperture (f/5.6)

http://photographertips.net
Field of View (Zoom, focal length)

From London and Upton
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
Today

• What is image processing?
• Image formation
• Digital images
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)
  - CMOS

Digital camera

• Color typically captured using color mosaic
• Demosaicing
FIGURE 2.17 (a) Continuous image projected onto a sensor array. (b) Result of image sampling and quantization.
Issues with digital cameras

• Noise
  – big difference between consumer vs. SLR-style cameras
  – low light is where you most notice noise

• Compression
  – creates artifacts except in uncompressed formats (tiff, raw)

• Color
  – color fringing artifacts from Bayer patterns

• Blooming
  – charge overflowing into neighboring pixels

• In-camera processing
  – oversharpening can produce halos

• Interlaced vs. progressive scan video
  – even/odd rows from different exposures

• Are more megapixels better?
  – requires higher quality lens
  – noise issues

• Stabilization
  – compensate for camera shake (mechanical vs. electronic)

More info online, e.g.,
http://www.dpreview.com/
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.
Image Representation

• Discretization
  – in image space - sampling
  – In image brightness - quantization

Digital image: 2D discrete function $f$

Pixel: Smallest element of an image $f(x,y)$

Figure: M. J. Black
Image Representation

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

Figure: M. J. Black
Datatypes for raster images

- Bitmaps: boolean per pixel (1 bpp): \( I : \mathbb{R}^2 \to \{0, 1\} \)
  - interp. = black and white; e.g. fax
- Grayscale: integer per pixel: \( I : \mathbb{R}^2 \to [0, 1] \)
  - 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: \( I : \mathbb{R}^2 \to [0, 1]^3 \)
  - 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.

  Same luminance?
Converting pixel precision

- Up is easy; down loses information—be careful
Next week

• Color
• Point operations