Image Processing

6.815/6.865
Frédo Durand
A bunch of slides by Bill Freeman (MIT) & Alyosha Efros (CMU)
Message from G. Verghese

• Hi, Fredo. I just approved a petition for a student to consider your Computational Photography subject as an EC under 6.011, as the description and the prerequisites suggest this is quite justified. If you agree, you could make this possibility more widely known in your class, in case others find it helpful to count the subject that way.

• Thanks.

• George
Colorimetry Demystified
Jan J. Koenderink,

• **Wednesday 3pm in Kiva**

  • "Colorimetry" is the link between optics (physics) and psychology in that it describes the input to the brain. This renders it trivial perhaps, but also important as the foundation ("chapter zero") of any conceivable "Color Science". Unfortunately, the field has been, and still is, rife with misconceptions and formal errors. I will discuss the worst of these and show how to "do colorimetry right". Conceptually colorimetry is coarse grained spectroscopy, fully characterized through a "black subspace" of codimension three in the space of spectra. In applications one would like to "strip the black fluff" off spectra in order to retain just their causally effective "fundamental parts" ("Wyszecki's Hypothesis" of the 1950's). Cohen's method (of the 1970's) is believed to achieve just this, but fails because of an implicit assumption that seems hard to defend. One solution is to take Schopenhauer's notion (of 1816) of colors as "Parts of Daylight" seriously. This yields an intuitive and formally elegant system of considerable utility.
Image processing

• Filtering, Convolution, and our friend Joseph Fourier
What is an image?

- We can think of an image as a function, $f$, from $\mathbb{R}^2$ to $\mathbb{R}$:
  - $f(x, y)$ gives the intensity at position $(x, y)$
  - Realistically, we expect the image only to be defined over a rectangle, with a finite range:
    - $f: [a,b] \times [c,d] \rightarrow [0,1]$

- A color image is just three functions pasted together. We can write this as a “vector-valued” function:

$$f(x, y) = \begin{bmatrix} r(x, y) \\ g(x, y) \\ b(x, y) \end{bmatrix}$$
Images as functions
Image Processing

- **image filtering:** change *range* of image
  
  \[ f(x) = h(f(x)) \]

- **image warping:** change *domain* of image
  
  \[ g(x) = f(h(x)) \]
Image Processing

- image filtering: change range of image
  \[ g(x) = h(f(x)) \]

- image warping: change domain of image
  \[ g(x) = f(h(x)) \]
Point Processing

• The simplest kind of range transformations are these independent of position \(x,y\):

\[ g = t(f) \]

• This is called point processing.

• **Important:** every pixel for himself – spatial information completely ignored!
Negative

FIGURE 3.4
(a) Original digital mammogram.
(b) Negative image obtained using the negative transformation in Eq. (3.2-1).
(Courtesy of G.E. Medical Systems.)
Contrast Stretching

![Contrast Stretching Diagram]

**FIGURE 3.10**
Contrast stretching, (a) Form of transformation function. (b) A low-contrast image. (c) Result of contrast stretching. (d) Result of thresholding. (Original image courtesy of Dr. Roger Heady, Research School of Biological Sciences, Australian National University, Canberra, Australia.)
Image Histograms

histogram $H(f) = \# \text{ or } \% \text{ pixels with value } f$
(implies binning of the values)

cumulative histogram:
$C(f) = \sum_{f' \leq f} H(f)$
$= \# \text{ or } \% \text{ of pixels with value } \leq f$

FIGURE 3.15 Four basic image types: dark, light, low contrast, high contrast, and their corresponding histograms. (Original image courtesy of Dr. Roger Heady, Research School of Biological Sciences, Australian National University, Canberra, Australia.)
Histogram Equalization

• point transformation: $g(x) = t(f(x))$
• uniform across image (t does not depend on x)
• monotonic (preserve intensity ordering)
• so that histogram of g is uniform
  – perfect uniform only possible with continuous histogram
Qualitative Histogram equalization

- Qualitative
Derivation

• Normalized cumulative histogram $C$: there are $C(f)\%$ pixels equal or darker than $f$

• In an image in $[0, 1]$ with a flat histogram, what is the greyscale value $g$ so that $C(f)\%$ pixels are equal or darker than $f$?
  – $C(f)$ of course!

• Therefore, histogram equalization:
  – $g(x) = C(f(x))$
Questions?
Filtering

• So far we have looked at range-only and domain-only transformation

• But other transforms need to change the range according to the spatial neighborhood
  – Linear shift-invariant filtering in particular
Linear shift-invariant filtering

- Replace each pixel by a linear combination of its neighbors.
  - only depends on relative position of neighbors
- The prescription for the linear combination is called the “convolution kernel”.

<table>
<thead>
<tr>
<th>10</th>
<th>5</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>0</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Local image data  kernel  Modified image data (shown at one pixel)
Example of linear NON-shift invariant transformation?

- e.g. neutral-density graduated filter (darken high y, preserve small y)
  \[ J(x,y) = I(x,y) \times (1 - y/y_{\text{max}}) \]

- Formally, what does linear mean?
  - For two scalars a & b and two inputs x & y:
    \[ F(ax + by) = aF(x) + bF(y) \]

- What does shift invariant mean?
  - For a translation T:
    \[ F(T(x)) = T(F(x)) \]
  - If I blur a translated image, I get a translated blurred image
More formally: Convolution

\[ f[m, n] = I \otimes g = \sum_{k,l} I[m-k, n-l]g[k,l] \]
Convolution (warm-up slide)

original

coefficient

Pixel offset

1.0

?
Convolution (warm-up slide)

- Original image
- Coefficient value: 1.0
- Pixel offset value: 0
- Filtered image (no change)
Convolution

original

Pixel offset

coefficient

1.0

?
shift

original

shifted

Pixel offset

0

1.0

coefficient
Convolution

original

coefficient

0.3

Pixel offset

?
Blurring

original

Blurred (filter applied in both dimensions).
Blur examples

impulse

original

coefficient

Pixel offset

filtered

8

0.3

2.4
Blur examples

impulse

original

filtered

edge

original

filtered
Questions?
Convolution (warm-up slide)
Convolution (no change)
Convolution

original
(remember blurring)

original

Blurred (filter applied in both dimensions).
Sharpening

original

Sharpened original
Sharpening example

original

coefficient

Sharpened
(differences are accentuated; constant areas are left untouched).
Sharpening

before

after
Oriented filters

Gabor filters at different scales and spatial frequencies.

Top row shows anti-symmetric (or odd) filters, bottom row the symmetric (or even) filters.
Filtered images

Questions?
Studying convolutions

• Convolution is complicated
  – But at least it’s linear
    \[(f+kg) \otimes h = f \otimes h + k (g \otimes h)\]

• We want to find a better expression
  – Let’s study function whose behavior is simple under convolution
Blurring: convolution

This is an eigenvector
(output is the input multiplied by a constant)

Same shape, just reduced contrast!!!
Big Motivation for Fourier analysis

• Sine waves are eigenvectors of the convolution operator
Other motivation for Fourier analysis: sampling

- The sampling grid is a periodic structure
  - Fourier is pretty good at handling that
  - We saw that a sine wave has serious problems with sampling
- Sampling is a linear process
  - but not shift-invariant
Sampling Density

- If we’re lucky, sampling density is enough.
Sampling Density

- If we insufficiently sample the signal, it may be mistaken for something simpler during reconstruction (that's aliasing!)
Recap: motivation for sine waves

- Blurring sine waves is simple
  - You get the same sine wave, just scaled down
  - The sine functions are the eigenvectors of the convolution operator

- Sampling sine waves is interesting
  - Get another sine wave
  - Not necessarily the same one! (aliasing)

If we represent functions (or images) with a sum of sine waves, convolution and sampling are easy to study
Questions?
Fourier as change of basis

• Shuffle the data to reveal other information
• E.g., take average & difference: matrix

\[
\begin{bmatrix}
0.5 & 1 \\
0.5 & -1
\end{bmatrix}
\]
Fourier as change of basis

- Same thing with infinite-dimensional vectors
Question?
Fourier as a change of basis

• Discrete Fourier Transform: just a big matrix

http://www.reindeergraphics.com
To get some sense of what basis elements look like, we plot a basis element --- or rather, its real part --- as a function of \(x, y\) for some fixed \(u, v\). We get a function that is constant when \((ux + vy)\) is constant. The magnitude of the vector \((u, v)\) gives a frequency, and its direction gives an orientation. The function is a sinusoid with this frequency along the direction, and constant perpendicular to the direction.
Here $u$ and $v$ are larger than in the previous slide.
And larger still...
Question?
Other presentations of Fourier

• Start with Fourier series with periodic signal

• Heat equation
  – more or less special case of convolution
  – iterate -> exponential on eigenvalues
Motivations

• Insights & mathematical beauty
• Sampling rate and filtering bandwidth
• Computation bases
  – FFT: faster convolution
  – E.g. finite elements, fast filtering, heat equation, vibration modes
• Optics: wave nature of light & diffraction
Questions?
The Fourier Transform

- Defined for infinite, aperiodic signals
- Derived from the Fourier series by “extending the period of the signal to infinity”
- The Fourier transform is defined as
  \[ X(\omega) = \frac{1}{\sqrt{2\pi}} \int x(t)e^{-j\omega t} \, dt \]
- \( X(\omega) \) is called the spectrum of \( x(t) \)
- It contains the magnitude and phase of each complex exponential of frequency \( \omega \) in \( x(t) \)
The Fourier Transform

• The inverse Fourier transform is defined as
  \[ x(t) = \frac{1}{\sqrt{2\pi}} \int X(\omega) e^{j\omega t} d\omega \]

• Fourier transform pair
  \[ x(t) \xrightarrow{F\,\mathbb{R}} X(\omega) \]

• \( x(t) \) is called the *spatial domain* representation
• \( X(\omega) \) is called the *frequency domain* representation
Duality
Beware of differences

- Different definitions of Fourier transform
- We use
  \[ X(\omega) = \frac{1}{\sqrt{2\pi}} \int x(t)e^{-j\omega t} \, dt \]
- Other people might exclude normalization or include \(2\pi\) in the frequency
- \(X\) might take \(\omega\) or \(j\omega\) as argument
- Physicist use \(j\), mathematicians use \(i\)
Questions?
Phase

• Don’t forget the phase! Fourier transform results in complex numbers

• Can be seen as sum of sines and cosines

• Or modulus/phase
Phase is important!
Phase is important!

Figure 6.2  (a) The image shown in Figure 1.4; (b) magnitude of the two-dimensional Fourier transform of (a); (c) phase of the Fourier transform of (a); (d) picture whose Fourier transform has magnitude as in (b) and phase equal to zero; (e) picture whose Fourier transform has magnitude equal to 1 and phase as in (c); (f) picture whose Fourier transform has phase as in (c) and magnitude equal to that of the transform of the picture shown in (g).
Questions?
Low pass

black means 1, white means 0

http://www.reindeergraphics.com
High pass

http://www.reindeergraphics.com
Filtering in Fourier domain
Analysis of our simple filters

Original

Pixel offset

Coefficient

Filtered
(no change)

Spectrum: \( F(\omega) = 1 \)

(yes, I am now using the definition without \( 1/\sqrt{2\pi} \))
Analysis of our simple filters

spectrum:

\[ F(\omega) = e^{-2\pi j \omega \delta} \]
Analysis of our simple filters

Pixel offset

original

blurred

coefficient

0.3

Low-pass filter

spectrum:

\[ F(\omega) = \text{sinc}(\omega) = \frac{\sin(\omega)}{\omega} \]
Analysis of our simple filters

original

sharpened

spectrum:

\[ F(\omega) = 2 - \text{sinc}(\omega) \]
Convolution versus FFT

- 1-d FFT: $O(N\log N)$ computation time, where $N$ is number of samples.
- 2-d FFT: $2N(N\log N)$, where $N$ is number of pixels on a side
- Convolution: $K N^2$, where $K$ is number of samples in kernel
- Say $N=2^{10}$, $K=100$. 2-d FFT: $20 \times 2^{20}$, while convolution gives $100 \times 2^{20}$
Questions?
Sampling and aliasing
In photos too
More on Samples

• In signal processing, the process of mapping a continuous function to a discrete one is called *sampling*

• The process of mapping a continuous variable to a discrete one is called *quantization*

• To represent or render an image using a computer, we must both sample and quantize
  – Now we focus on the effects of sampling and how to fight them
Sampling in the Frequency Domain

original signal

sampling grid

sampled signal

Fourier Transform

multiplication

convolution

Fourier Transform

Fourier Transform

Fourier Transform
Reconstruction

- If we can extract a copy of the original signal from the frequency domain of the sampled signal, we can reconstruct the original signal!

- But there may be overlap between the copies.
Guaranteeing Proper Reconstruction

- Separate by removing high frequencies from the original signal (low pass pre-filtering)

- Separate by increasing the sampling density

- If we can't separate the copies, we will have overlapping frequency spectrum during reconstruction → aliasing.
Sampling Theorem

- When sampling a signal at discrete intervals, the sampling frequency must be greater than twice the highest frequency of the input signal in order to be able to reconstruct the original perfectly from the sampled version (Shannon, Nyquist, Whittaker, Kotelnikov)
Final project brainstorming

Fredo Durand
MIT EECS 6.815/6.865
Final project

- Groups of 1 or 2
- Proposal due soon (with pset 4)
- Deliverables: report + small presentation
Your ideas?
Some ideas

- Use CHDK to provide new features to Canon compact cameras
- Use flickr API to do something creative
- Explore different types of gradient reconstructions
- Improve time lapse
- Handle small parallax in panoramas
- Exploit flash/no-flash pairs
- Editing with images+depth (e.g. from stereo)
- Smart color to greyscale
- Face-aware image processing
- Sharpening out-of-focus images using other pictures from the sequences
- Application of morphing/warping
- Motion without movements and automatic illusions