Filtering
Introduction to Computer Vision
CSE 152
Lecture 9

Announcements
• Homework 2 will be assigned today
  – Due Tuesday, May 12, 11:59 PM
• Midterm Exam on Thursday, May 7
• Reading:
  – Chapter 3 Image processing

Image Filtering
Input
Output

Filtering
Introduction to Computer Vision
CSE 152
Lecture 9

Example: Smoothing by Averaging

Linear Filters
• General process:
  – Form new image whose pixels are a weighted sum of original pixel values, using the same set of weights at each point.
• Properties
  – Output is a linear function of the input
  – Output is a shift-invariant function of the input (i.e. shift the input image two pixels to the left, the output is shifted two pixels to the left)
• Example: smoothing by averaging
  – form the average of pixels in a neighborhood
• Example: smoothing with a Gaussian
  – form a weighted average of pixels in a neighborhood
• Example: finding a derivative
  – form a difference of pixels in a neighborhood

What is image filtering?
• Modify the pixels in an image based on some function of a local neighborhood of the pixels.

From Bill Freeman
Linear functions

- Simplest: linear filtering.
- Replace each pixel by a linear combination of its neighbors.
- The prescription for the linear combination is called the "convolution kernel".

Note: Typically Kernel is relatively small in vision applications.

Convolution: \( R = K \ast I \)

Kernel size is \( m+1 \) by \( m+1 \)

\[
R(i, j) = \sum_{k=-m}^{m} \sum_{h=-m}^{m} K(h, k) I(i-h, j-k)
\]
Convolution: $R = K \ast I$

Kernel size is $m+1$ by $m+1$

$R(i, j) = \sum_{d=-m}^{m} \sum_{k=-m}^{m} K(h, k) I(i-h, j-k)$
Convolution: $R = K \ast I$

Kernel size is $m+1$ by $m+1$

$R(i, j) = \sum_{h=-m}^{m} \sum_{k=-m}^{m} K(h,k) I(i-h, j-k)$

Linear filtering (warm-up slide)

Original

Pixel offset

Filtered (no change)

Linear filtering

Original

Pixel offset

shift

Original

Pixel offset

Shifted one Pixel to the left

Linear filtering

Original

$\frac{1}{9}$

$111$

$111$

$111$
Blurring

Blurred (filter applied in both dimensions).

Practice with linear filters

Original

Sharpening example

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

Sharpening

before  after
Properties of Continuous Convolution (Holds for discrete too)
Let f, g, h be images and * denote convolution
\[ f * g(x, y) = \int \int f(x-u, y-v)g(u,v)dudv \]
• Commutative: \( f * g = g * f \)
• Associative: \( f * (g * h) = (f * g) * h \)
• Linear: for scalars \( a \) & \( b \) and images \( f, g, h \)
  \( (af+bg)*h=a(f*h)+b(g*h) \)
• Differentiation rule
  \[ \frac{\partial}{\partial x} (f * g) = \frac{\partial f}{\partial x} * g = f * \frac{\partial g}{\partial x} \]

Filtering to reduce noise
• Noise is what we’re not interested in.
  – We’ll discuss simple, low-level noise today: Light fluctuations; Sensor noise; Quantization effects; Finite precision
  – Not complex: shadows; extraneous objects.
• A pixel’s neighborhood contains information about its intensity.
• Averaging noise reduces its effect.

Additive noise
• \( I = S + N \). Noise doesn’t depend on signal.
• We’ll consider:
  \[ I_i = s_i + n_i \text{ with } E(n_i) = 0 \]
  \( s_i \) deterministic. \( n_i \) a random var.
  \( n_i, n_j \) independent for \( i \neq j \)
  \( n_i, n_j \) identically distributed

Averaging Filter
• Mask with positive entries, that sum 1.
• Replaces each pixel with an average of its neighborhood.
• If all weights are equal, it is called a Box filter.
  \[
  \begin{bmatrix}
  1 & 1 & 1 \\
  1 & 1 & 1 \\
  1 & 1 & 1 \\
  \end{bmatrix}
  \]
  \( \frac{1}{9} \)

Smoothing by Averaging
Kernel: 

Gaussian Noise: \( \text{sigma}=1 \)  
Gaussian Noise: \( \text{sigma}=16 \)
An Isotropic Gaussian

- The picture shows a smoothing kernel proportional to
  \[ \frac{e^{-\frac{x^2 + y^2}{2\sigma^2}}}{2\pi \sigma^2} \]

(which is a reasonable model of a circularly symmetric fuzzy blob)

Smoothing with a Gaussian

Kernel:

The effects of smoothing:
Each row shows smoothing with gaussians of different width, each column shows different realizations of an image of gaussian noise.

Gaussian Smoothing

original

\[ \sigma = 2.8 \]

\[ \sigma = 2 \]

\[ \sigma = 4 \]
Efficient Implementation

• Both, the Box filter and the Gaussian filter are separable:
  – First convolve each row with a 1D filter
  – Then convolve each column with a 1D filter.

Other Types of Noise

• Impulsive noise
  – randomly pick a pixel and randomly set to a value
  – saturated version is called salt and pepper noise

• Quantization effects
  – Often called noise although it is not statistical

• Unanticipated image structures
  – Also often called noise although it is a real repeatable signal.

Some other useful filtering techniques

• Median filter
• Anisotropic diffusion

Median filters: principle

Method:
1. rank-order neighbourhood intensities
2. take middle value

• non-linear filter
• no new grey levels emerge...

Median filters: Example for window size of 3

1.1.1.7.1.1.1
\downarrow
?1.1.1.1.1.?…

The advantage of this type of filter is that it eliminates spikes (salt & pepper noise).
Median filters: example
filters have width 5:

<table>
<thead>
<tr>
<th>INPUT</th>
<th>MEDIAN</th>
<th>MEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>[\cdots\cdots\cdots]</td>
<td>[\cdots\cdots\cdots]</td>
<td>[\cdots\cdots\cdots]</td>
</tr>
</tbody>
</table>

Median filters: analysis
median completely discards the spike,
linear filter always responds to all aspects
median filter preserves discontinuities,
linear filter produces rounding-off effects
Do not become all too optimistic

Median filters: images
3 x 3 median filter:
sharpens edges, destroys edge cusps and protrusions

Median filters: Gauss revisited
Comparison with Gaussian:
e.g. upper lip smoother, eye better preserved

Example of median
10 times 3 x 3 median
patchy effect
important details lost (e.g. ear-ring)

Fourier Transform
• 1-D transform (signal processing)
• 2-D transform (image processing)
• Consider 1-D
  Time domain ↔ Frequency Domain
  Real ↔ Complex
• Consider time domain signal to be expressed as
  weighted sum of sinusoid. A sinusoid \(\cos(ut+\phi)\) is
  characterized by its phase \(\phi\) and its frequency \(u\)
• The Fourier transform of the signal is a function
  giving the weights (and phase) as a function of
  frequency \(u\).
Fourier Transform

Discrete Fourier Transform (DFT) of \( I[x,y] \)

\[
F[u,v] = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} I[x,y] e^{-\frac{2\pi j}{N}(ux+vy)}
\]

Inverse DFT

\[
I[x,y] = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F[u,v] e^{\frac{2\pi j}{N}(ux+vy)}
\]

\( x,y \): spatial domain
\( u,v \): frequency domain
Implemented via the "Fast Fourier Transform" algorithm (FFT)

Fourier basis element

\( e^{-\frac{2\pi j}{N}(ux+vy)} \)

Transform is sum of orthogonal basis functions

Vector \((u,v)\)
- Magnitude gives frequency
- Direction gives orientation.

Using Fourier Representations

Dominant Orientation

Limitations: not useful for local segmentation

Phase and Magnitude

\[
\theta = \cos \theta + i \sin \theta
\]

- Fourier transform of a real function is complex
  - difficult to plot, visualize
  - instead, we can think of the phase and magnitude of the transform
- Phase is the phase of the complex transform
- Magnitude is the magnitude of the complex transform
- Curious fact
  - all natural images have about the same magnitude transform
  - hence, phase seems to matter, but magnitude largely doesn't
- Demonstration
  - Take two pictures, swap the phase transforms, compute the inverse - what does the result look like?
This is the magnitude transform of the cheetah pic.

This is the phase transform of the cheetah pic.

This is the magnitude transform of the zebra pic.

This is the phase transform of the zebra pic.
The Fourier Transform and Convolution

- If $H$ and $G$ are images, and $F(.)$ represents Fourier transform, then
  \[ F(H \ast G) = F(H)F(G). \]
- Thus, one way of thinking about the properties of a convolution is by thinking of how it modifies the frequencies of the image to which it is applied.
- In particular, if we look at the power spectrum, then we see that convolving image $H$ by $G$ attenuates frequencies where $G$ has low power, and amplifies those which have high power.
- This is referred to as the Convolution Theorem.

Various Fourier Transform Pairs

- Important facts
  - scale function down $\Leftrightarrow$ scale transform up
    i.e. high frequency = small details
  - The Fourier transform of a Gaussian is a Gaussian.
    
    compare to box function transform