Review – Frequency Domain Techniques

• The name “filter” is borrowed from frequency domain processing
• Accept or reject certain frequency components
• Fourier (1807):
  Periodic functions could be represented as a weighted sum of sines and cosines

Review - Fourier Transform

We want to understand the frequency \( w \) of our signal. So, let’s reparametrize the signal by \( w \) instead of \( x \):

\[
 f(x) \xrightarrow{\text{Fourier Transform}} F(w)
\]

For every \( w \) from 0 to \( \infty \), \( F(w) \) holds the amplitude \( A \) and phase \( \phi \) of the corresponding sine

\[
 A \sin(\omega x + \phi)
\]

• How can \( F \) hold both? Complex number trick!

\[
 F(\omega) = R(\omega) + iI(\omega)
\]

\[
 A = \pm \sqrt{R(\omega)^2 + I(\omega)^2}
\]

\[
 \phi = \tan^{-1} \frac{I(\omega)}{R(\omega)}
\]

We can always go back:

\[
 F(w) \xrightarrow{\text{Inverse Fourier Transform}} f(x)
\]

Review - The Discrete Fourier transform

• Forward transform

\[
 F[m,n] = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} f[k,l] e^{-i \left( \frac{km}{M} + \frac{ln}{N} \right)}
\]

• Inverse transform

\[
 f[k,l] = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} F[m,n] e^{i \left( \frac{km}{M} + \frac{ln}{N} \right)}
\]
**Review - The Discrete Fourier transform**

![Fourier spectrum diagram]

- Vertical orientation: 45°
- Horizontal orientation: 0°
- Low spatial frequencies: 0 to \( f_{max} \)
- High spatial frequencies: \( f_{max} \) to 0

**Log power spectrum**

Slide credit: B. Freeman and A. Torralba

---

**Review - The Convolution Theorem**

- The Fourier transform of the convolution of two functions is the product of their Fourier transforms:

  \[
  F[g * h] = F[g]F[h]
  \]

- The inverse Fourier transform of the product of two Fourier transforms is the convolution of the two inverse Fourier transforms:

  \[
  F^{-1}[gh] = F^{-1}[g] * F^{-1}[h]
  \]

- Convolution in spatial domain is equivalent to multiplication in frequency domain!

Slide credit: A. Efros

---

**Review - Filtering in frequency domain**

- Low-pass filtering:
  - \( F^{-1}[F[1] \cdot F[g]] \)
  - \( F^{-1}[F[1] \cdot F[g]] \)

- High-pass / band-pass filtering:
  - \( F^{-1}[F[1] \cdot F[g]] \)
  - \( F^{-1}[F[1] \cdot F[g]] \)

Slide credit: D. Hoiem

---

**Review - Low-pass, Band-pass, High-pass filters**

- Low-pass:
  - Filtered image
  - Log power spectrum

- High-pass / band-pass:
  - Filtered image
  - Log power spectrum

Slide credit: A. Efros
**Template matching**
- Goal: find in image

- Main challenge: What is a good similarity or distance measure between two patches?
  - Correlation
  - Zero-mean correlation
  - Sum Square Difference
  - Normalized Cross Correlation

**Matching with filters**
- Goal: find in image

- Method 0: filter the image with eye patch

\[ h[m,n] = \sum_{k,l} g[k,l] f[m+k,n+l] \]

- Method 1: filter the image with zero-mean eye

\[ h[m,n] = \sum_{k,l} (f[k,l] - \bar{f}) (g[m+k,n+l]) \]  

What went wrong?
response is stronger for higher intensity

- Method 2: SSD

\[ h[m,n] = \sum_{k,l} (g[k,l] - f[m+k,n+l])^2 \]
**Matching with filters**
- Goal: find $\varpi$ in image
- Method 2: SSD

$$h[m,n] = \sum_{k,l} (g[k,l] - f[m+k,n+l])^2$$

What’s the potential downside of SSD?
SSD sensitive to average intensity

**Matlab:**
```matlab
normxcorr2(template, im)
```

**Matching with filters**
- Goal: find $\varpi$ in image
- Method 3: Normalized cross-correlation

$$h[m,n] = \frac{\sum_{k,l} (g[k,l] - \bar{g})(f[m-k,n-l] - \bar{f}_{m,n})}{\left(\sum_{k,l} (g[k,l] - \bar{g})^2 \sum_{k,l} (f[m-k,n-l] - \bar{f}_{m,n})^2\right)^{0.5}}$$

**Matlab:**
```matlab
normxcorr2(template, im)
```
Q: What is the best method to use?

A: Depends
• SSD: faster, sensitive to overall intensity
• Normalized cross-correlation: slower, invariant to local average intensity and contrast

Q: What if we want to find larger or smaller eyes?

A: Image Pyramid

Image pyramids
Image information occurs over many different spatial scales.
• Gaussian pyramid
• Laplacian pyramid
• Wavelet/QMF pyramid
• Steerable pyramid

Image Pyramids
**Image pyramids**

Image information occurs at all spatial scales

- Gaussian pyramid
- Laplacian pyramid
- Wavelet/QMF pyramid
- Steerable pyramid

**The Gaussian pyramid**

- Smooth with Gaussians, because
  - A Gaussian * Gaussian = another Gaussian
- Gaussians are low pass filters, so representation is redundant.
- Gaussian pyramid creates versions of the input image at multiple resolutions.
- This is useful for analysis across different spatial scales, but doesn’t separate the image into different frequency bands.

**The computational advantage of pyramids**

[Burt and Adelson, 1983]
The Gaussian Pyramid

Convolution and subsampling as a matrix multiply (1D case)

\[ x_3 = G_2 x_2 \]

\[ G_1 = \begin{array}{cccccccccccc}
1 & 4 & 6 & 4 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 4 & 6 & 4 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 4 & 6 & 4 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 4 & 6 & 4 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 6 & 4 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 6 & 4 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 6 & 4 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 6 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 \\
\end{array} \]

\[ G_2 = \begin{array}{cccccccc}
1 & 4 & 6 & 4 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 4 & 6 & 4 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 4 & 6 & 4 \\
0 & 0 & 0 & 0 & 0 & 1 & 4 & 6 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
\end{array} \]

[Normalization constant of 1/16 omitted for visual clarity.]

Next pyramid level
The combined effect of the two pyramid levels

\[ x_3 = G_2 G_1 x_1 \]

\[ G_2 G_1 = \]

\[
\begin{array}{cccccccccccccccc}
1 & 4 & 10 & 20 & 31 & 40 & 44 & 40 & 31 & 20 & 10 & 4 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 4 & 10 & 20 & 31 & 40 & 44 & 40 & 31 & 20 & 10 & 4 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 10 & 20 & 31 & 40 & 44 & 40 & 30 & 16 & 4 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{array}
\]

Gaussian pyramids used for

- up- or down- sampling images.
- Multi-resolution image analysis
  - Look for an object over various spatial scales
  - Coarse-to-fine image processing: form blur estimate or the motion analysis on very low-resolution image, upsample and repeat. Often a successful strategy for avoiding local minima in complicated estimation tasks.

1D Gaussian pyramid matrix, for \([1 4 6 4 1]\) low-pass filter
**Template Matching with Image Pyramids**

Input: Image, Template
1. Match template at current scale
2. Downsample image
3. Repeat 1-2 until image is very small
4. Take responses above some threshold, perhaps with non-maxima suppression

---

**Coarse-to-fine Image Registration**

1. Compute Gaussian pyramid
2. Align with coarse pyramid
3. Successively align with finer pyramids
   - Search smaller range

Why is this faster?

Are we guaranteed to get the same result?

---

**Image pyramids**

Image information occurs at all spatial scales

- Gaussian pyramid
- Laplacian pyramid
- Wavelet/QMF pyramid
- Steerable pyramid

---

**The Laplacian Pyramid**

- Synthesis
  - Compute the difference between upsampled Gaussian pyramid level and Gaussian pyramid level.
  - Band pass filter - each level represents spatial frequencies (largely) unrepresented at other level.

- Laplacian pyramid provides an extra level of analysis as compared to Gaussian pyramid by breaking the image into different isotropic spatial frequency bands.
The Laplacian Pyramid

\[ x_1 \quad G_1 x_1 \quad F_1 G_1 x_1 \quad (I - F_1 G_1) x_1 \]

Slide credit: B. Freeman and A. Torralba
The Laplacian Pyramid

\[ x_1, \quad G_1 x_1 = x_2 \]

\[ F_1 G_1 x_1 \]

\[ (I - F_1 G_1) x_1 \]

The Laplacian Pyramid

\[ x_1, \quad G_1 x_1 = x_2 \quad \text{and} \quad x_2 \quad \text{and} \quad x_3 \]

\[ (I - F_1 G_1) x_1 \]

\[ F_1 G_1 x_1 \]

\[ (I - F_2 G_2) x_2 \]

\[ (I - F_1 G_1) x_3 \]

Upsampling

\[ y_2 = F_3 x_3 \]

Insert zeros between pixels, then apply a low-pass filter, \([1 \ 4 \ 6 \ 4 \ 1]\)

\[ F_3 = \begin{bmatrix} 6 & 1 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ 1 & 6 & 1 & 0 \\ 0 & 4 & 4 & 0 \\ 0 & 1 & 6 & 1 \\ 0 & 0 & 4 & 4 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 4 \end{bmatrix} \]
Showing, at full resolution, the information captured at each level of a Gaussian (top) and Laplacian (bottom) pyramid.

Slide credit: B. Freeman and A. Torralba

Laplacian pyramid reconstruction algorithm:
recover $x_1$ from $L_1$, $L_2$, $L_3$ and $x_4$

$G#$ is the blur-and-downsample operator at pyramid level #
$F#$ is the blur-and-upsample operator at pyramid level #

Laplacian pyramid elements:
$L_1 = (I - F_1 G_1) x_1$
$L_2 = (I - F_2 G_2) x_2$
$L_3 = (I - F_3 G_3) x_3$

$x_2 = G_1 x_1$
$x_3 = G_2 x_2$
$x_4 = G_3 x_3$

Reconstruction of original image ($x_1$) from Laplacian pyramid elements:
$x_3 = L_3 + F_3 x_4$
$x_2 = L_2 + F_2 x_3$
$x_1 = L_1 + F_1 x_2$

Slide credit: B. Freeman and A. Torralba
Laplacian pyramid applications

- Texture synthesis
- Image compression
- Noise removal

The Laplacian Pyramid as a Compact Image Code

Peter J. Burt, member, IEEE, and Edward H. Adelson

IEEE TRANSACTIONS ON COMMUNICATIONS, Vol. COM-31, No. 4, April 1983
Image blending

- Build Laplacian pyramid for both images: LA, LB
- Build Gaussian pyramid for mask: G
- Build a combined Laplacian pyramid:
  \[ L(j) = G(j) \cdot LA(j) + (1-G(j)) \cdot LB(j) \]
- Collapse L to obtain the blended image

Eulerian Video Magnification

- Video

Image pyramids

Image information occurs at all spatial scales

- Gaussian pyramid
- Laplacian pyramid
- Wavelet/QMF pyramid
- Steerable pyramid
Wavelet/QMF pyramid

- Subband coding
- Wavelet or QMF (quadrature mirror filter) pyramid provides some splitting of the spatial frequency bands according to orientation (although in a somewhat limited way).
- Image is decomposed into a set of band-limited components (subbands).
- Original image can be reconstructed without error by reassembling these subbands.

2D Haar transform

Basic elements: $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

Low pass

High pass vertical

High pass horizontal

High pass diagonal

Sketch of the Fourier transform

Horizontal low pass, vertical low-pass

Horizontal high pass, vertical low-pass

Horizontal low pass, vertical high-pass

Horizontal high pass, vertical high-pass

Pyramid cascade

Wavelet/QMF representation

Image pyramids

Image information occurs at all spatial scales

- Gaussian pyramid
- Laplacian pyramid
- Wavelet/QMF pyramid
- Steerable pyramid

Steerable Pyramid

The Steerable pyramid provides a clean separation of the image into different scales and orientations.

Images from: http://www.cis.upenn.edu/~eero/steerpyr.html

Steerable Pyramid

We may combine Steerability with Pyramids to get a Steerable Laplacian Pyramid as shown below.

Decomposition

Reconstruction

Images from: http://www.cis.upenn.edu/~eero/steerpyr.html
**Steerable Pyramid**

We may combine Steerability with Pyramids to get a Steerable Laplacian Pyramid as shown below.

Images from: http://www.cis.upenn.edu/~eero/steerpyr.html

Slide credit: B. Freeman and A. Torralba

Image pyramids

- Gaussian
- Laplacian
- Wavelet/QMF
- Steerable pyramid

There is also a high pass residual…

Slide credit: B. Freeman and A. Torralba
Image pyramids

- Gaussian
  Progressively blurred and subsampled versions of the image. Adds scale invariance to fixed-size algorithms.

- Laplacian
  Shows the information added in Gaussian pyramid at each spatial scale. Useful for noise reduction & coding.

- Wavelet/QMF
  Bandpassed representation, complete, but with aliasing and some non-oriented subbands.

- Steerable pyramid

Slide credit: B. Freeman and A. Torralba
Schematic pictures of each matrix transform

Shown for 1-d images

The matrices for 2-d images are the same idea, but more complicated, to account for vertical, as well as horizontal, neighbor relationships.

Fourier transform, or Wavelet transform, or Steerable pyramid transform

\[ \vec{F} = \vec{U} \vec{f} \]

\[ \text{Fourier transform} \]

\[ \text{Fourier bases are global: each transform coefficient depends on all pixel locations.} \]

Gaussian pyramid

\[ \text{Gaussian pyramid} \]

Overcomplete representation.
Low-pass filters, sampled appropriately for their blur.

Laplacian pyramid

\[ \text{Laplacian pyramid} \]

Overcomplete representation.
Transformed pixels represent bandpassed image information.
Wavelet (QMF) transform

Wavelet pyramid = Ortho-normal transform (like Fourier transform), but with localized basis functions.

Why use image pyramids?

- Handle real-world size variations with a constant-size vision algorithm.
- Remove noise
- Analyze texture
- Recognize objects
- Label image features
- Image priors can be specified naturally in terms of wavelet pyramids.

Phase-based Video Magnification

- Video