Wavelets
Raghu Machiraju, The Ohio State University

Why Wavelets?

- We are generating and measuring larger datasets every year
- We can not store all the data we create (too much, too fast)
- We can not look at all the data (too busy, too hard)
- We need to develop techniques to store the data in better formats

Data Analysis

- Frequency spectrum correctly shows a spike at 10 Hz
- Spike not narrow - significant component at between 5 and 15 Hz.
- Leakage - discrete data acquisition does not stop at exactly the same phase in the sine wave as it started.

QuickFix

Windowing & Filtering
Image Example
- 8x8 Blocked Window (Cosine) Transform
- Each DCT basis waveform represents a fixed frequency in two orthogonal directions
- Frequency spacing in each direction is an integer multiple of a base frequency

Time Frequency Diagram

Windowing & Filtering
Windows - fixed in space and frequencies
Cannot resolve all features at all instants

Linear Scale Space
\[ \mu_r(x; \sigma) = G(\sigma, x) \otimes I(x) = \int G(\sigma, \tau) I(x - \tau) d\tau \]

Successive Smoothing
Keep 1 of 4 values from 2x2 blocks
This naive approach introduces aliasing
Sub-samples are bad representatives of area
Little spatial correlation

Sub-sampled Images
Image Pyramid

Average over a 2x2 block
- This is a rather straight forward approach
- This reduces aliasing and is a better representation
- However, this produces 11% expansion in the data

Image Pyramid - MIP MAP

Image Pyramid - Another Twist

Pyramids

10 2 8 6 4 2 4 0

Average Sum

Average Differences

This is the Haar Wavelet Transform

Ideally!

Create new signal G such that ||F-G|| = e

Wavelet Analysis

Wavelets can be used to detect features and to compare features
Wavelets can provide compressed representations
Wavelet Theory provides a unified framework for data processing

Why Wavelets? Because...

- We need to develop techniques to analyze data better through noise discrimination
- Wavelets can be used to detect features and to compare features
- Wavelets can provide compressed representations
Scale-Coherent Structures

- Coherent structure - frequencies at all scales
- Examples - edges, peaks, ridges
- Locate extent and assign saliency

Wavelets - Analysis

Wavelets - DeNoising

Wavelets - Compression

Wavelets - Compression
Yet Another Example

50%
7%

Information Rate Curve

- Energy Compaction - Few coefficients can efficiently represent functions
- The Curve should be as vertical as possible near 0 rate

Filter Bank Implementation

G: High Pass Filter
H: Low Pass Filter

Synthesis Bank

(The Other Half) Synthesis Filter Bank

Successive Approximations
Successive Details

Wavelet Representation

Coefficients

Lossey Compression

Lossey Compression

Image Example
Image Example

Average

Difference

Wavelet Transform

Approx. at res. j =

Boxes have same meaning!

Frequency Support

Image Example

How Does One Do This?

\[
\begin{align*}
\phi(x) &= \begin{cases} 
1 & 0 \leq x < 1 \\
0 & \text{otherwise}
\end{cases} \\
\psi(x) &= \begin{cases} 
-1 & 0 \leq x < \frac{1}{2} \\
1 & \frac{1}{2} \leq x < 1 \\
0 & \text{otherwise}
\end{cases}
\end{align*}
\]

Desired Goal:

\[
f(x) = A_2 + D_1
\]
Dilations

- Rescaling Operation \( t \rightarrow 2t \)
- Down Sampling, \( n \rightarrow 2n \)
- Halve function support; Double frequency content
- Octave division of spectrum- Gives rise to different scales and resolutions
- Mother wavelet! - basic function gives rise to differing versions \( \phi(x) = \frac{1}{\sqrt{2^j}} \phi(\frac{x}{2^j}) \)

Successive Approximations

![Successive Approximations Diagram]

Wavelet Decomposition

- Induced functional Space - \( W_j \)
- Related to \( V_j \) \( V_j = V_{j+1} \oplus W_{j+1} \)
- Space \( W_j \), is orthogonal to \( V_j \)
- Also \( W_{j+1} \subset V_j \)
- \( j \)-level wavelet decomposition -
  \( V_j = V_{j+1} \oplus W_{j+1} \oplus W_{j+1} \)
  \( V_Q = V_j \oplus W_j \oplus W_{j-1} \oplus W_{j-2} \oplus ... \oplus W_1 \)

Translations

- Covers space-frequency diagram
- Versions are \( \phi_j(x) = \frac{1}{\sqrt{2^j}} \phi(\frac{x}{2^j}) \)

Successive Differences

![Successive Differences Diagram]
Wavelet Expansion
- Wavelet expansion (Tiling: $j$: scale, $k$: translates), Synthesis
  $$f(x) = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{1}{2^j} \varphi_{j,k}(x) + \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{1}{2^j} \psi_{j,k,l}(x)$$
- Orthogonal transformation, coarsest level of resolution - $J$
- Smoothing function - $f$, Detail function - $y$
- Analysis: $a_{nj} = \sum f(x) \tilde{\varphi}_{j,k}(x)$, $d_{nj} = \sum f(x) \tilde{\psi}_{j,k}(x)$
- Commonly used wavelets are Haar, Daubechies and Coiflets

Scaling Functions
- Compact support
- Band limited - cut-off frequency
- Cannot achieve both
- DC value (or the average) is defined $\phi(x) dx = 1$
- Translates of $f$ are orthogonal $\phi(x) \phi(x-k) dx = \delta(k)$

Wavelet Functions
- Wavelet Equation - Haar System: $G$ Filter
  $$\psi(t) = 2g(0)\psi(2t) - 2g(1)\psi(2t-1)$$
  $$g(0) = 1/2, g(1) = -1/2$$

Orthogonal Filter Banks
- Alternating Flip $h_{2^n} = (-1)^n h_{N-2^n}$
- Not symmetric - $h$ is even length!
- Example $H = (h_0, h_1, h_2, h_3) \quad H^T = (h_3, h_2, h_1, h_0)$
  $$G = (h_3, -h_2, h_1, -h_0) \quad G^T = (-h_0, h_1, -h_2, h_3)$$
- Orthogonality conditions
  $$\sum h_n^2 = \delta(k)$$
  $$\sum h_n g(n-2k) = 0$$
  $$\|h\|^2 - \|h+\pi\|^2 = \|\pi\|^2 = 1$$
Examples

\[ h(0) + h(1) = \frac{1}{2} \]
\[ h(0) + h(1) + h(2) + h(3) = \frac{1}{4} \]

Haar

\[ \text{Approximation: Vanishing Moments Property} \]

- Function is smooth - Taylor Series expansion
  \[ f(x) = \sum_{p=0}^{\infty} \frac{f^{(p)}(0)}{p!} x^p \]

- Wavelets with \( m \) vanishing moments
  \[ h_{\text{maxflat}}(\omega) = \frac{1}{\sin(\omega/2)} \]

- Function with \( m \) derivatives can be accurately represented!

Example

\[ N=4 \]

Design of Compact Orthogonal Wavelets

- Compute scaling function
- Use Refinement Equation:
  \[ \Phi_\omega = \frac{1}{\pi} \left( \frac{1}{2} \right) \frac{1}{\omega} \mathcal{H}(\omega) \mathcal{P}(\omega) \]

- \( N \) vanishing moments property - \( H(\omega) \) has a zero of order \( N \) at \( \omega = p \)

- \( \mathcal{P}(y) \) is \( p^{th} \) order polynomial (Daubechies 1992)

- Maxflat filter

Example

\[ N=16 \]

Noise

- Uncorrelated Gaussian noise is correlated
- Region of correlation is small at coarse scale
- Smooth versions - no noise
- Orthogonal transform - uncorrelated
Denoising

- Statistical thresholding methods [Donohoe]
- Assuming Gaussian Noise
- Universal Threshold $\theta = 2\sigma_n\sqrt{\log(n)}$
- Smoothness guaranteed
  - Hard: $\mathcal{Y}_d(x) = \sigma \mathcal{Y}(x) \cdot (\mathcal{Y}(x) > \theta)$
  - Soft: $\mathcal{Y}_s(x) = \sigma \mathcal{Y}(x) \cdot (\mathcal{Y}(x) - \theta)$
- Works for additive noise since wavelet transform is linear $\mathcal{W}(x, h)[f + \eta \psi] = \mathcal{W}(x, h)[f] + \mathcal{W}(x, h)[\eta \psi]$

Bi-Orthogonal Filter Banks

- Analysis/synthesis different
- Aliasing - overlap in spectras
- Alias cancellation $H_1(a)G_0(a + b) + G_1(a)H_0(a + b) = 0$
- Distortion Free (phase shift $I$) $H_1(a)G_0(a + b) + G_1(a)H_0(a + b) = 2\sigma_\psi^2$\delta$
- Alternating Flip condition valid
- Can be odd length, symmetric

Discontinuity

Multi-scale Edges

- Mallat and Hwang
- Location - maxima (edges) of wavelet coefficients at all scales
- Maxima chains for each edge
- Ranking - compute Lipschitz coefficient at all points
- Representation - store maxima
- Reconstruction - approximate but works in practice
### Bi-Orthogonal Wavelets

- Governing equations
  \[ f(n) = (-1)^n h_1(N-n) \]
  \[ f'(n) = (-1)^n h_0(N-n) \]
  \[ \sum h(n) h_1(n+2k) = g(k) \]
- Spline Wavelets - Many choices of either \( H_0 \) or \( H_1 \)
- Choose \( H_0 \) as spline and solve equations to generate \( H_1 \)

### Bi-orthogonal: Lifting Scheme

- Lazy wavelet transform: split data in 2 parts
- Keep even part; predict (linear/cubic) odd part
- Lifting - update \( \lambda_{j+1} \) with \( \gamma_{j+1} \): Maintain properties (moments, avg.)
- Synthesis is just flip of analysis

### Summary

- Wavelets have good representation property
- They improve on image pyramid schemes
- Orthogonal and biorthogonal filter bank implementations are efficient
- Wavelets can filter signals
- They can efficiently denoise signals
- The presence of singularities can be detected from the magnitude of wavelet coefficients and their behavior across scales