Wavelets and Multiscale Edge Detection

Brody Dylan Johnson

Saint Louis University

Abstract:

In 1992, Mallat and Zhong published a paper presenting a numerical technique for the characterization of one- and two-dimensional discrete signals in terms of their multiscale edges [2]. With the appropriate choice of wavelet, the locations of edges correspond to modulus maxima of the continuous wavelet transform at a given scale. In this talk, we will explore the fundamentals of the Mallat-Zhong approach.

Overview:

- 1-D Edge Detection and Signal Characterization
  - smoothing functions and "wavelet derivatives"
  - stability of continuous wavelet transform
  - practical considerations
  - example
- 2-D Edge Detection
  - Canny edge detector
  - examples

The smoothing function:

- We say $\theta(x)$ is a **smoothing function** if $\theta \in C^2(\mathbb{R})$, has a fast decay (so that $\hat{\theta}$ is C^2), and $\int_{\mathbb{R}} \theta(x) = 1$. Under these assumptions, $\theta \in L^p(\mathbb{R})$, $1 \leq p \leq \infty$.
- Prototypical example: the Gaussian, $\theta(x) = \frac{1}{\sqrt{\pi}} e^{-x^2}$.
- At scale $s > 0$, we have a dilated version of the smoothing function, $\theta_s(x) := \frac{1}{s}\theta(\frac{x}{s})$, which also satisfies $\int_{\mathbb{R}} \theta_s(x) = 1$.
- For $f \in L^2(\mathbb{R})$, the convolution $(f * \theta_s)(x)$ is a smoothed version of f (twice-differentiable) at the scale $s > 0$. Moreover,

$$\lim_{s \to 0} (f * \theta_s)(x) = f(x) \quad \text{a.e.}.$$

- Interpretation: $(f * \theta_s)$ removes variation from f that occurs at resolutions finer than s.

The smoothing function:

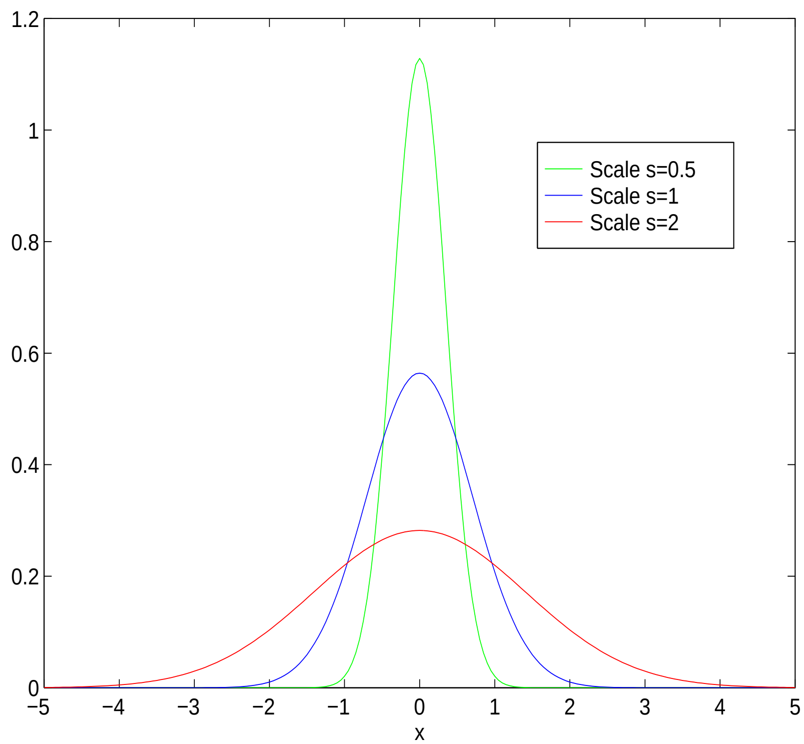


Figure 1: Various dilations of the smoothing function $\theta = \frac{1}{\sqrt{\pi}} e^{-x^2}$.

The Fourier transform:

- The Fourier transform of $f \in L^1 \cap L^2(\mathbb{R})$ is defined by

$$\mathcal{F}f(\xi) = \hat{f}(\xi) = \int_{\mathbb{R}} f(x) e^{-2\pi i x \cdot \xi} dx.$$

- Relevant properties of the Fourier transform:

1. $(\mathcal{F}f'(x))(\xi) = (2\pi i \xi)\hat{f}(\xi)$
2. $(\mathcal{F}xf(x))(\xi) = \frac{i}{2\pi}\hat{f}'(\xi)$
3. $\hat{f}(0) = \displaystyle\int_{\mathbb{R}} f(x) dx$

- The Parseval formula for $f, g \in L^1 \cap L^2(\mathbb{R})$:

$$\langle f, g \rangle = \int_{\mathbb{R}} f(x)\overline{g(x)} dx = \int_{\mathbb{R}} \hat{f}(\xi)\overline{\hat{g}(\xi)} d\xi = \langle \hat{f}, \hat{g} \rangle.$$

The wavelets:

- Given a smoothing function θ as above, define

$$\psi^a(x) = \frac{d\theta}{dx}(x) \quad \& \quad \psi^b(x) = \frac{d^2\theta}{dx^2}(x).$$

- ψ^a and ψ^b are wavelets in the sense that

$$\int_{\mathbb{R}} \psi^a(x)dx = \int_{\mathbb{R}} \psi^b(x)dx = 0.$$

This is because $\hat{\psi}^a(\xi) = (2\pi i\xi)\hat{\theta}(\xi)$, $\hat{\psi}^b(\xi) = (2\pi i\xi)^2\hat{\theta}(\xi)$, and $\hat{\theta}(0) = 1$, implying $\hat{\psi}^a(0) = \hat{\psi}^b(0) = 0$.

The wavelets:

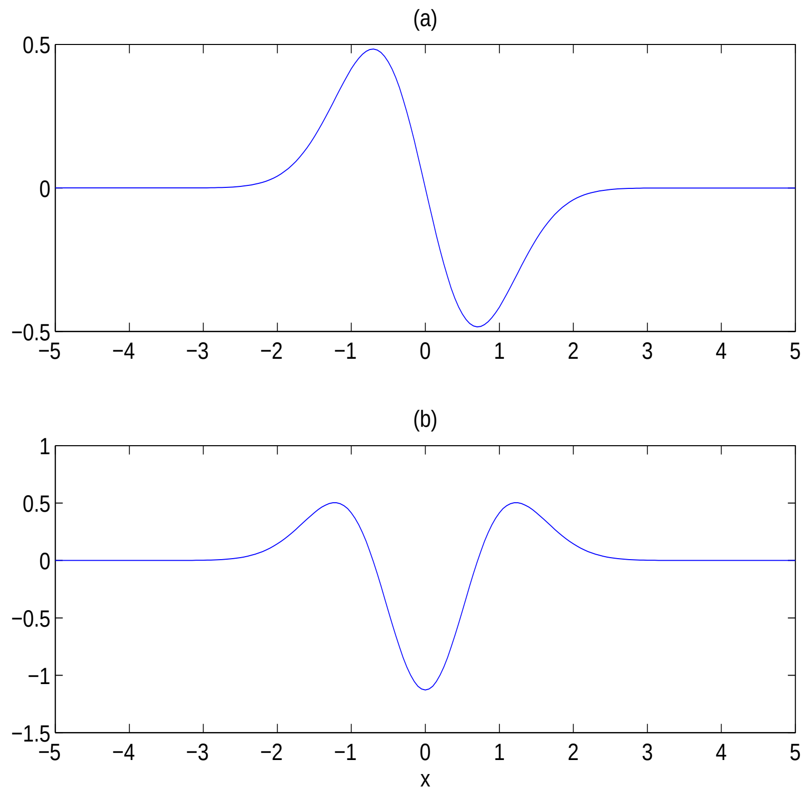


Figure 2: The wavelets: (a) ψ^a and (b) ψ^b associated with the smoothing function $\theta = \frac{1}{\sqrt{\pi}} e^{-x^2}$. The wavelet ψ^b is often referred to as the Mexican hat function.

Continuous wavelet transform:

- The continuous wavelet transforms defined by ψ^a and ψ^b, respectively, are

$$W_s^a f(x) = (f * \psi_s^a)(x) = s\frac{d}{dx}(f * \theta_s)(x)$$

and

$$W_s^b f(x) = (f * \psi_s^b)(x) = s^2\frac{d^2}{dx^2}(f * \theta_s)(x).$$

- $W_s^a f$ measures the derivative of the smoothed version of a signal f at scale s, while $W_s^b f$ measures the second derivative.

- Wavelets work by translation and dilation:

$$W_s^a f(x) = \int_{\mathbb{R}} f(y)\psi_s^a(x-y)dy = \langle f, T_x\tilde{\psi}_s^a\rangle,$$

i.e., $W_s^a f(x)$ is an inner product with a translation and dilation of $\tilde{\psi}_s^a$. (The involution of f is $\tilde{f}$, given by $\tilde{f}(x) = \overline{f(-x)}$.)

Defining edges:

- An edge should correspond to a point where $f(x)$ undergoes rapid variation, i.e., maxima of $f'(x)$. We cannot investigate $f'(x)$ directly, but we can instead study $W_s^a f(x)$.
- Loosely speaking, we will say that $f(x)$ has an edge at $x = a$ if $W_s f(x)$ has a local maxima at $x = a$. ($x = a$ should remain a local maxima as $s \to 0$)
- The local extrema of $W_s^a f(x)$ correspond to the zero crossings of $W_s^b f(x)$ and the inflection points of $(f * \theta_s)(x)$.
- Thus, W_s^a and W_s^b can each be used to locate eges, but the zero crossings of $W_s^b f$ fail to separate between the local maxima and minima of f. The minima of $W_s^b f$ correspond to points of smooth variation of f and will not give rise to edges.

Achieving a stable representation:

- Mallat and Zhong want to use the modulus maxima of $W_s^a f$ to reconstruct f, but it is not even obvious that one can reconstruct f from $W_s^a f$.

- Instead of consdering all scales $s > 0$ we will consider only dyadic scales 2^j, $j \in \mathbb{Z}$.

- Assume that ψ satisfies a Calderón inequality:

$$A \leq \sum_{j \in \mathbb{Z}} |\hat{\psi}(2^j \xi)|^2 \leq B \quad \text{a.e. } \xi \in \mathbb{R}.$$

- Define the Dyadic Wavelet Transform: $W^\psi : L^2(\mathbb{R}) \to L^2(\mathbb{Z}, \mathbb{R})$, $f \mapsto \{W_{2^j}^\psi f\}_{j \in \mathbb{Z}}$, where

$$W_{2^j} f := W_{2^j}^\psi f = (f * \psi_{2^j})(x).$$

Completeness of the wavelet transform:

Claim: $A\|f\|^2 \le \sum_{j\in\mathbb{Z}} \|W_{2^j} f\|^2 \le B\|f\|^2$.

Proof: Observe that

$$\begin{aligned}
\sum_{j\in\mathbb{Z}} \|W_{2^j} f\|^2 &= \sum_{j\in\mathbb{Z}} \int_{\mathbb{R}} |W_{2^j} f(x)|^2 \, dx \\
\text{(Parseval)} \qquad &= \sum_{j\in\mathbb{Z}} \int_{\mathbb{R}} |\hat{f}(\xi)|^2 |\hat{\psi}_{2^j}(\xi)|^2 \, d\xi \\
&= \sum_{j\in\mathbb{Z}} \int_{\mathbb{R}} |\hat{f}(\xi)|^2 |\hat{\psi}(2^j \xi)|^2 \, d\xi \\
&= \int_{\mathbb{R}} |\hat{f}(\xi)|^2 \Big(\sum_{j\in\mathbb{Z}} |\hat{\psi}(2^j \xi)|^2 \Big) \, d\xi.
\end{aligned}$$

Reconstruction:

- Suppose we find $\chi(x)$ so that

$$\sum_{j\in\mathbb{Z}} \hat{\psi}(2^j\xi)\hat{\chi}(2^j\xi) = 1,$$

then we can recover f from $W_{2^j} f$ via

$$f(x) = \sum_{j\in\mathbb{Z}} \big(W_{2^j} f * \chi_{2^j}\big)(x).$$

- This follows from the Fourier transform:

$$\sum_{j\in\mathbb{Z}} \hat{f}(\xi)\,\hat{\psi}(2^j\xi)\,\hat{\chi}(2^j\xi) = \hat{f}(\xi)\sum_{j\in\mathbb{Z}} \hat{\psi}(2^j\xi)\,\hat{\chi}(2^j\xi) = \hat{f}(\xi).$$

- Reconstruction from modulus maxima is another story, however, which will be addressed briefly below.

Practical considerations:

- In practice one encounters discretely defined functions, not functions of a continuous variable. Hence, we need a discrete version of the continuous wavelet transform.

- Let θ, ψ, and χ be refinable, i.e., there exists $m_0, m_1, m_2 \in L^\infty(\mathbb{T})$ such that

  $$\hat{\theta}(2\xi) = m_0(\xi)\hat{\theta}(\xi),\ \hat{\psi}(2\xi) = m_1(\xi)\hat{\theta}(\xi),\ \text{and}\ \hat{\chi}(2\xi) = m_2(\xi)\hat{\theta}(\xi)$$

  with the additional assumption (perfect reconstruction condition) that

  $$|m_0(\xi)|^2 + \overline{m_1(\xi)}m_2(\xi) = 1.$$

- We now replace the continuous wavelet transform with a discrete wavelet transform known as the à trous algorithm.

The à trous algorithm:

- The refinability of the smoothing function and wavelets provides useful relationships between the values of the wavelet transform across scales:

$$\langle f, T_k \theta_{2^{j-1}} \rangle = \sum_{\ell \in \mathbb{Z}} \alpha_\ell \langle f, T_{k+2^j \ell} \theta_{2^j} \rangle,$$

  and

$$\langle f, T_k \psi_{2^{j-1}} \rangle = \sum_{\ell \in \mathbb{Z}} \beta_\ell \langle f, T_{k+2^j \ell} \theta_{2^j} \rangle,$$

  where $m_0(\xi) = \sum_{\ell \in \mathbb{Z}} \alpha_\ell e^{-2\pi i \ell \xi}$ and $m_1(\xi) = \sum_{\ell \in \mathbb{Z}} \beta_\ell e^{-2\pi i \ell \xi}$.

- The à trous algorithm uses these relationships to compute $f * \theta_{2^{j-1}}$ and $W_{2^{j-1}} f$ from $f * \theta_{2^j}$. In practice a signal is interpreted as $f * \theta_{2^0}$ in this algorithm. Reconstruction is similar.

Reconstruction from modulus maxima:

- A frame for a Hilbert space $\mathbb{H}$ is a collection $\{x_j\}_{j\in\mathbb{J}}$ for which there exists $0 < A \leq B < \infty$ so that for each $x \in \mathbb{H}$

$$A\|x\|^2 \leq \sum_{j\in\mathbb{J}} |\langle x, x_j\rangle|^2 \leq B\|x\|^2.$$

- The reconstruction described above amounts to the existence of a dual frame $\{y_j\}_j$ which for each $x \in \mathbb{H}$ satisfies

$$x = \sum_{j\in\mathbb{J}} \langle x, x_j\rangle y_j.$$

- By considering only modulus maxima in reconstruction we are attempting to recover x using only $\{\langle x, x_{j_k}\rangle\}$ for some subsequence $\{j_k\} \subset \mathbb{J}$.

Reconstruction from modulus maxima:

- It has been shown that different functions can have the same modulus maxima (see [1] for references), but these signals tend to be very similar and for this reason fairly accurate reconstructions are possible using modulus maxima.

- A dual frame can no longer be used for reconstruction. Instead, the original function is recovered using the frame algorithm, which is an iterative algorithm for inverting the partial frame operator [1].

- The frame operator associated to $\{x_j\}$ is defined by

$$Sx = \sum_{j \in \mathbb{J}} \langle x, x_j \rangle x_j.$$

A natural signal

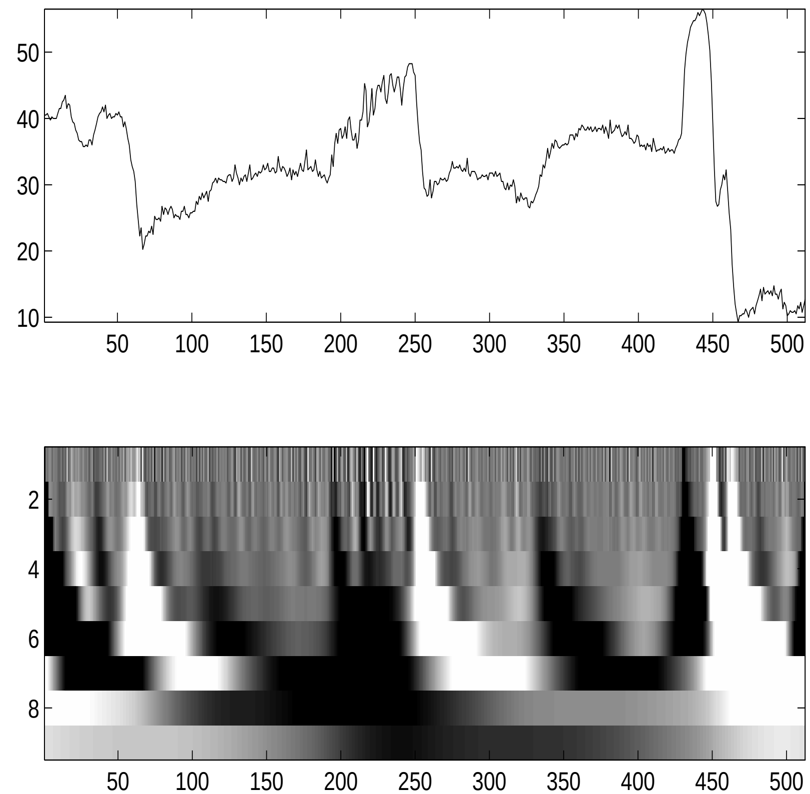


Figure 3: The 1-D continuous wavelet transform of a natural signal.

Smoothing across scales:

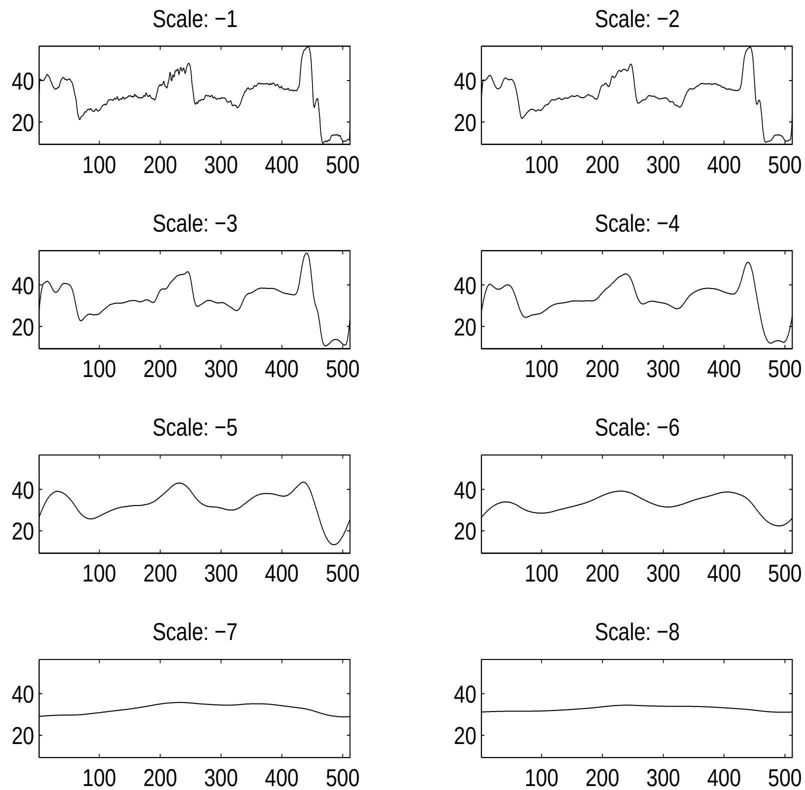


Figure 4: The smoothed versions of the signal at various scales.

Modulus of the wavelet transform:

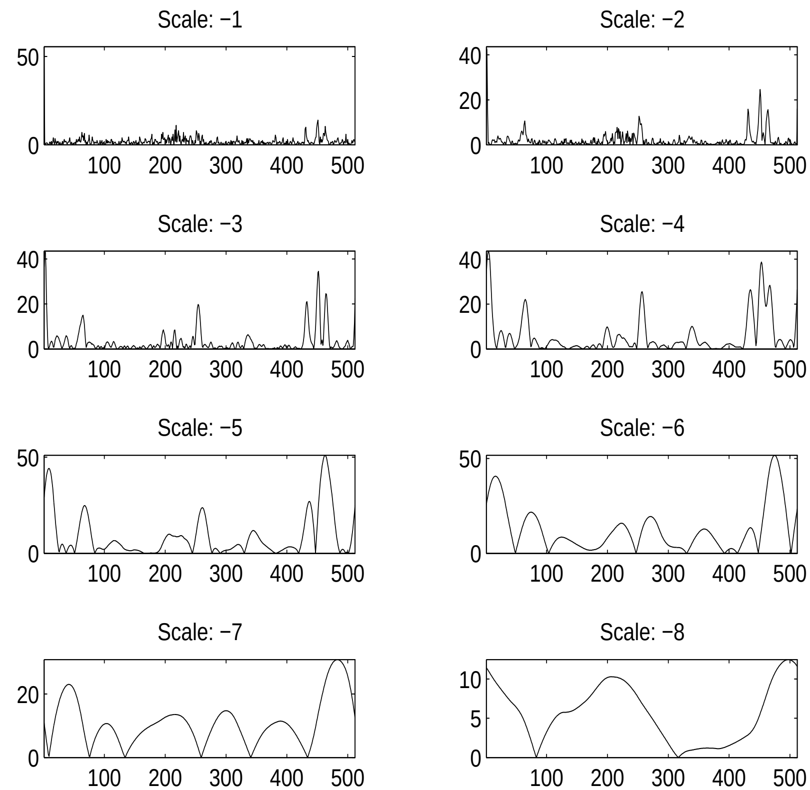


Figure 5: The modulus of the continuous wavelet transform at various scales.

Modulus maxima:

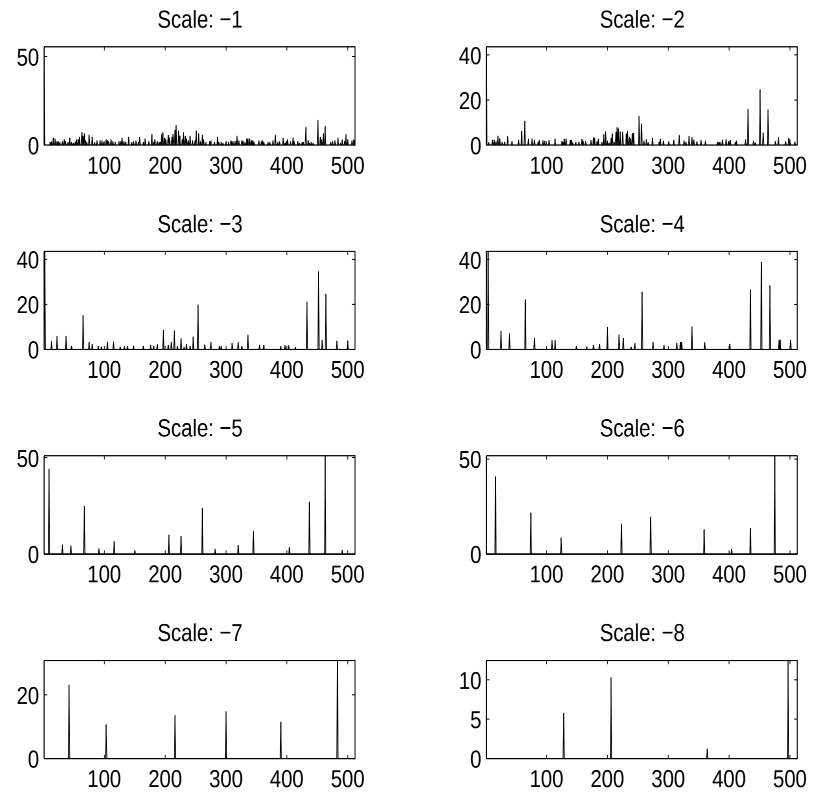


Figure 6: The modulus maxima of the continuous wavelet transform at various scales.

Reconstruction from modulus maxima:

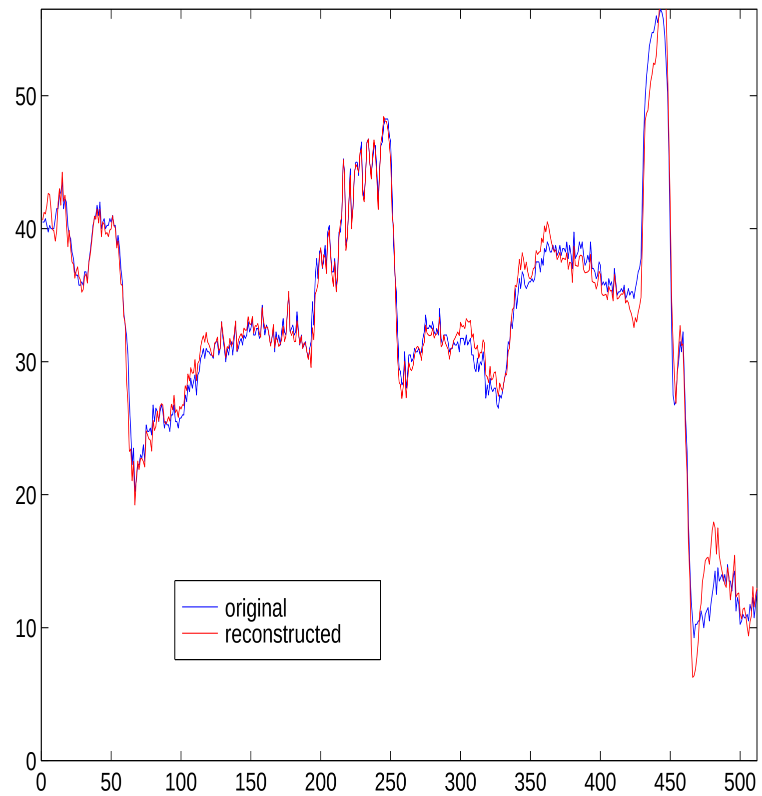


Figure 7: The comparison of the original signal and the signal reconstructed from the modulus maxima.

Two-dimensions:

- Smoothing function: $\tilde{\theta}(x,y) = \theta(x)\theta(y)$, where θ is a 1-D smoothing function.
- Define two wavelets: $\psi^1(x,y) := \frac{\partial}{\partial x}\tilde{\theta}(x,y) = \psi^a(x)\theta(y)$ and $\psi^1(x,y) := \frac{\partial}{\partial y}\tilde{\theta}(x,y) = \theta(x)\psi^a(y)$, where $\psi^a(x) = \frac{d}{dx}\theta(x)$.
- Let $\psi^1_s(x,y) = \frac{1}{s^2}\psi^1(\frac{x}{s},\frac{y}{s})$ and $\psi^2_s(x,y) = \frac{1}{s^2}\psi^2(\frac{x}{s},\frac{y}{s})$ and for $s = 2^j$, $j \in \mathbb{Z}$, consider the dyadic wavelet transforms:

$$W^1_s f(x,y) = (f * \psi^1_s)(x,y) \quad \text{and} \quad W^2_s f(x,y) = (f * \psi^2_s)(x,y).$$

- If $\chi^1(x,y)$ and $\chi^2(x,y)$ satisfy:

$$\sum_{j\in\mathbb{Z}} \hat{\psi}^1(\xi_1,\xi_2)\hat{\chi}^1(\xi_1,\xi_2) + \hat{\psi}^2(\xi_1,\xi_2)\hat{\chi}^2(\xi_1,\xi_2) = 1,$$

then the two-dimensional wavelet transform will allow reconstruction as above.

The Canny edge detector:

- Observe that

$$\begin{pmatrix} W_s^1 f(x,y) \\ W_s^2 f(x,y) \end{pmatrix} = s \begin{pmatrix} \frac{\partial}{\partial x}(f * \theta_s)(x,y) \\ \frac{\partial}{\partial y}(f * \theta_s)(x,y) \end{pmatrix} = s\nabla(f * \theta_s)(x,y).$$

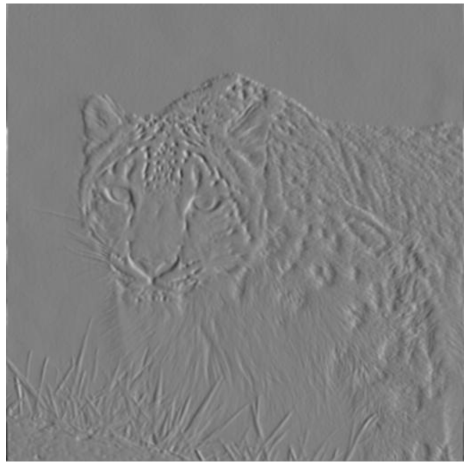
- The Canny algorithm defines (x_0, y_0) to belong to an edge if $\|\nabla f(x,y)\|$ is locally maximum at (x_0, y_0) in the direction of $\nabla f(x_0, y_0)$.

- According to [1], it remains an open problem as to whether or not such edges yield a complete and stable representation in two dimensions. The algorithm of [2] does provide numerical support for this hypothesis.

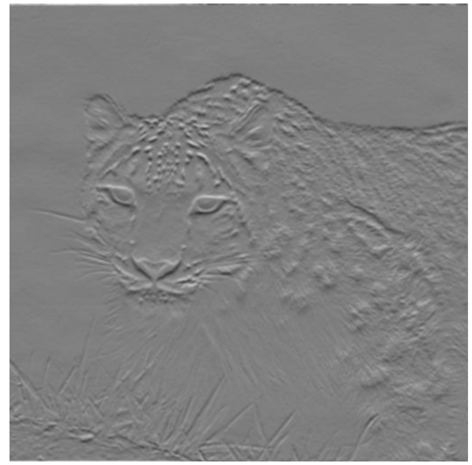
A natural image:

A snow leopard from the St. Louis Zoo.

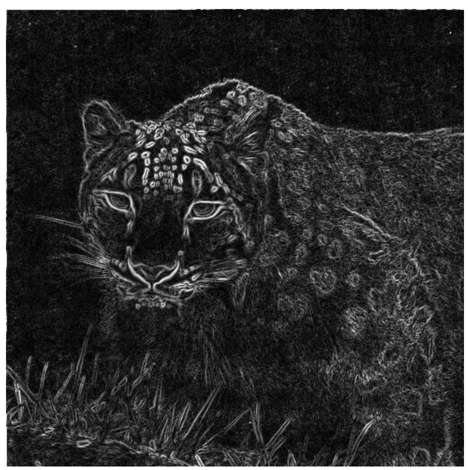
Scale $s = -1$:



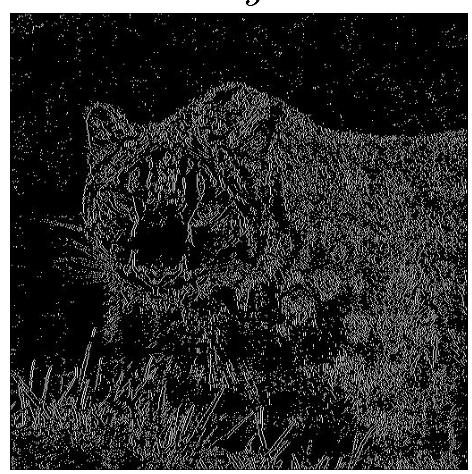
$\frac{\partial f}{\partial x}$



$\frac{\partial f}{\partial y}$

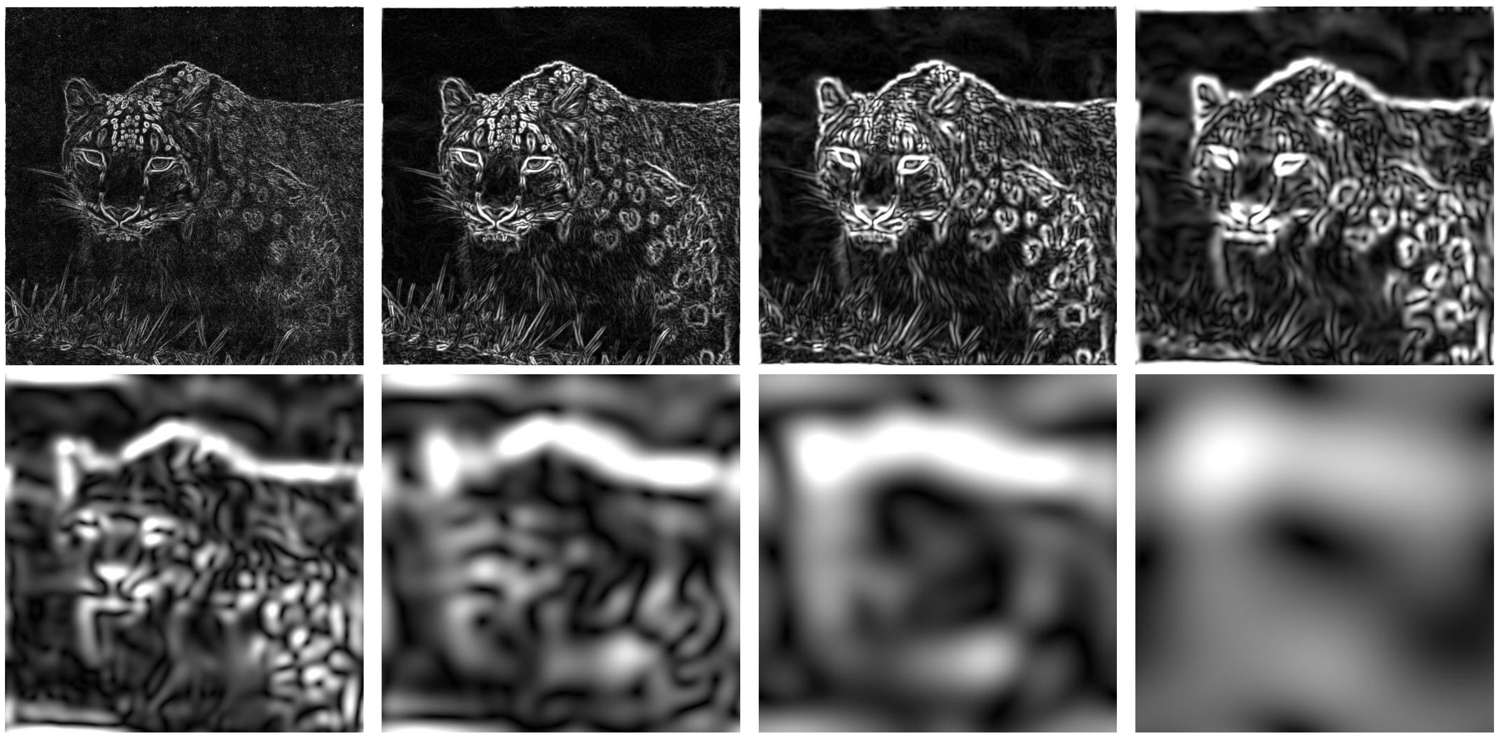


$\|\nabla f\|$



Modulus Maxima

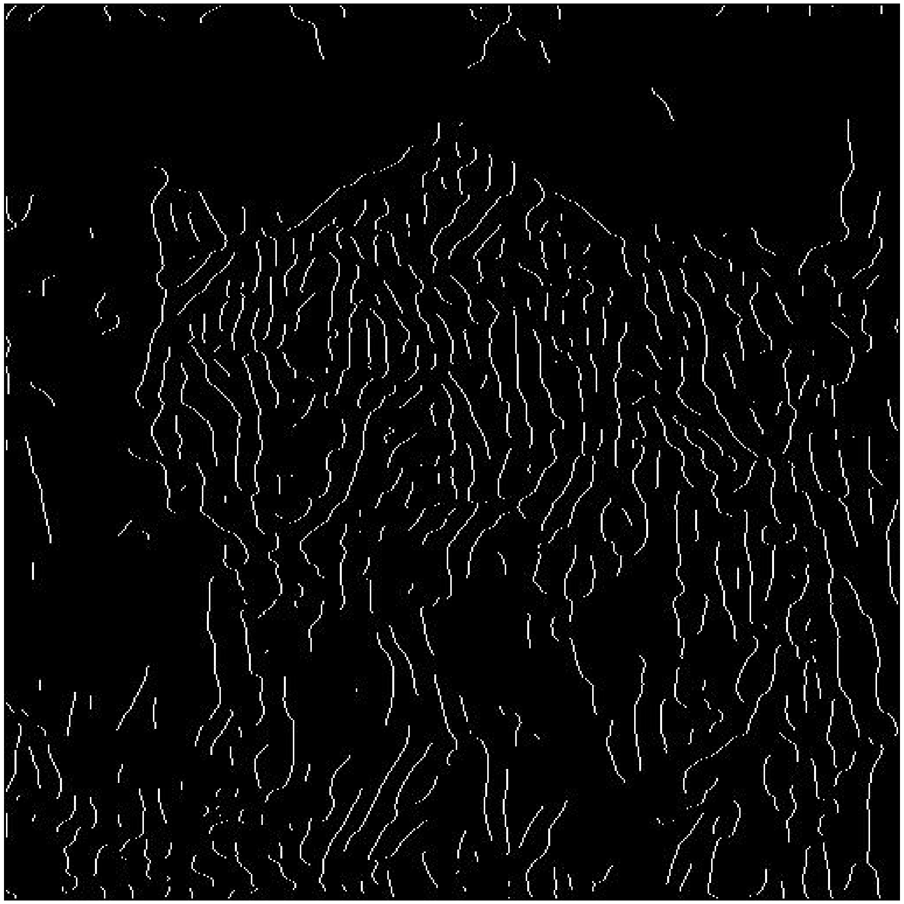
The modulus of the wavelet transform:



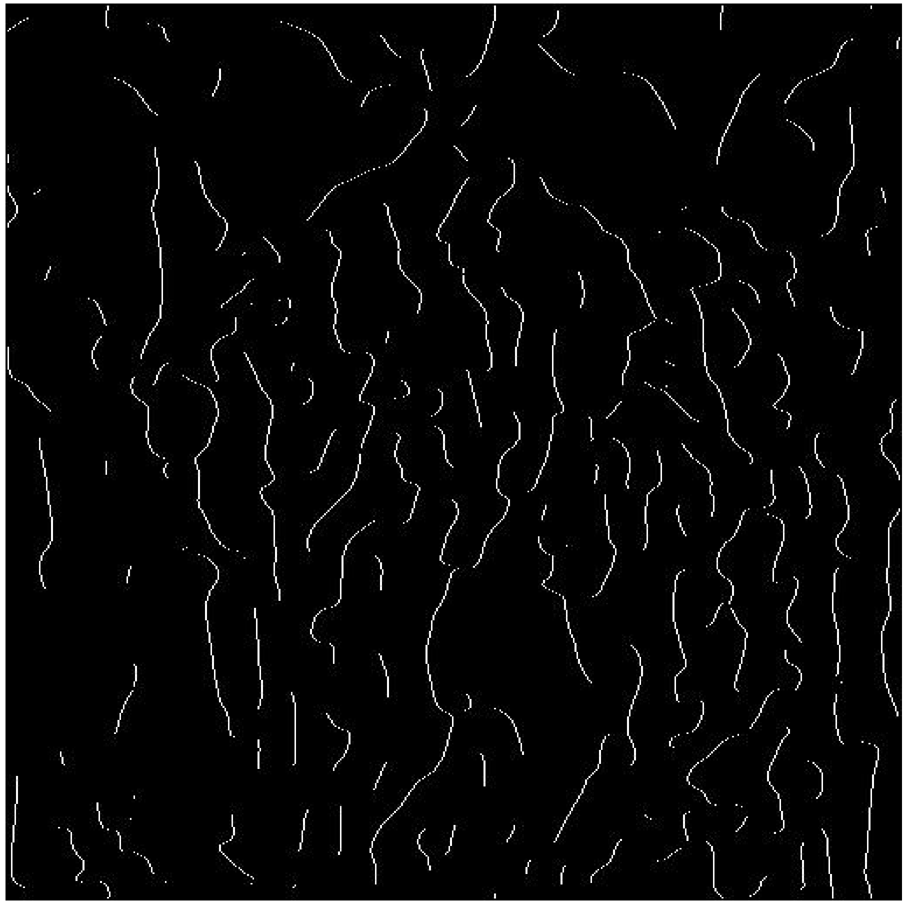
Modulus maxima: scale = -2

Modulus maxima: scale = -3

Modulus maxima: scale = -4



Modulus maxima: scale = -5

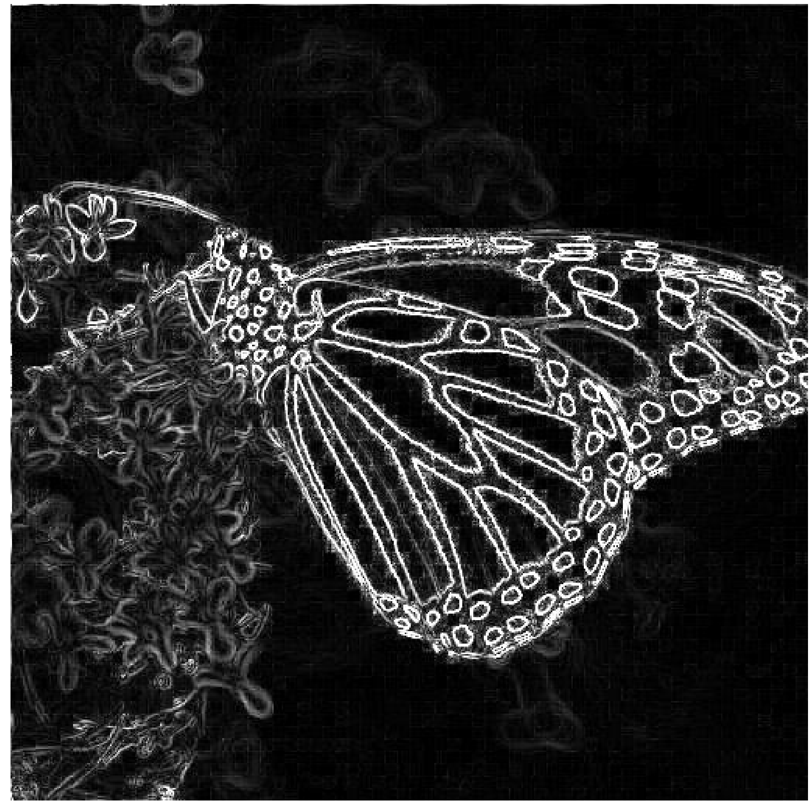


Reconstructed image:

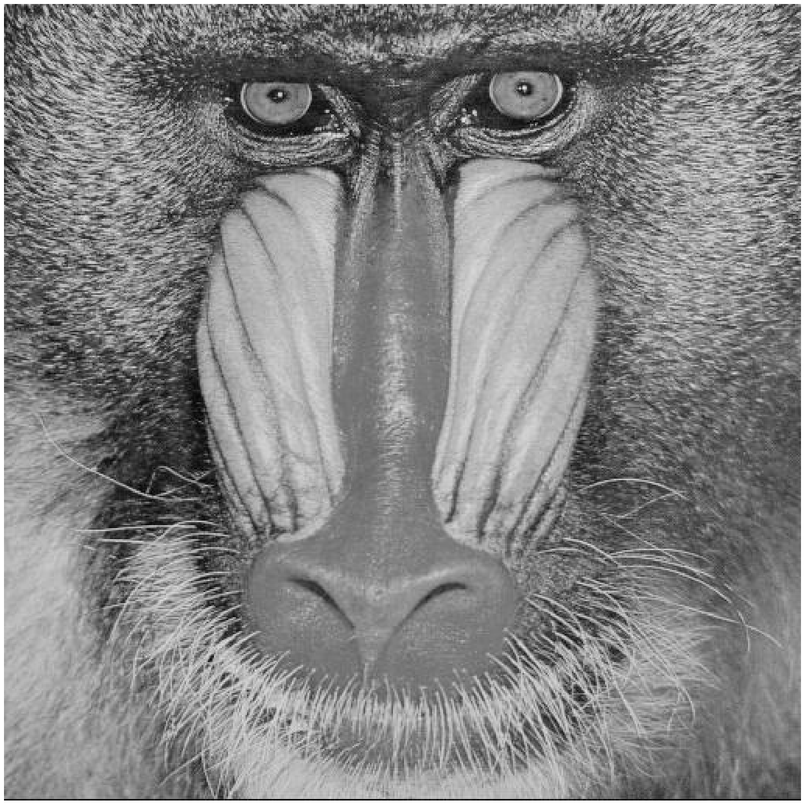
(Some "small" modulus maxima were ignored in reconstruction.)

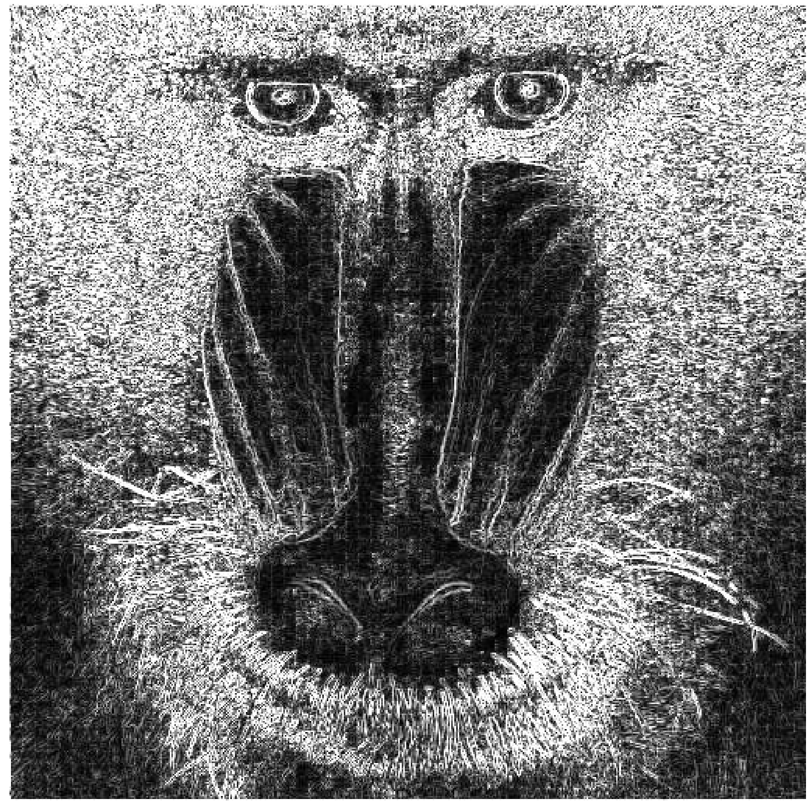
Original image:

Monarch example:

Mandrill example:





Lena example:

References

[1] Stéphane Mallat. *A wavelet tour of signal processing.* Associated Press, 1998.

[2] Stéphane Mallat and Sifen Zhong. Characterization of signals from multiscale edges. *IEEE Trans. Patt. Anal. and Mach. Intell.*, 14(7):710–732, 1992.