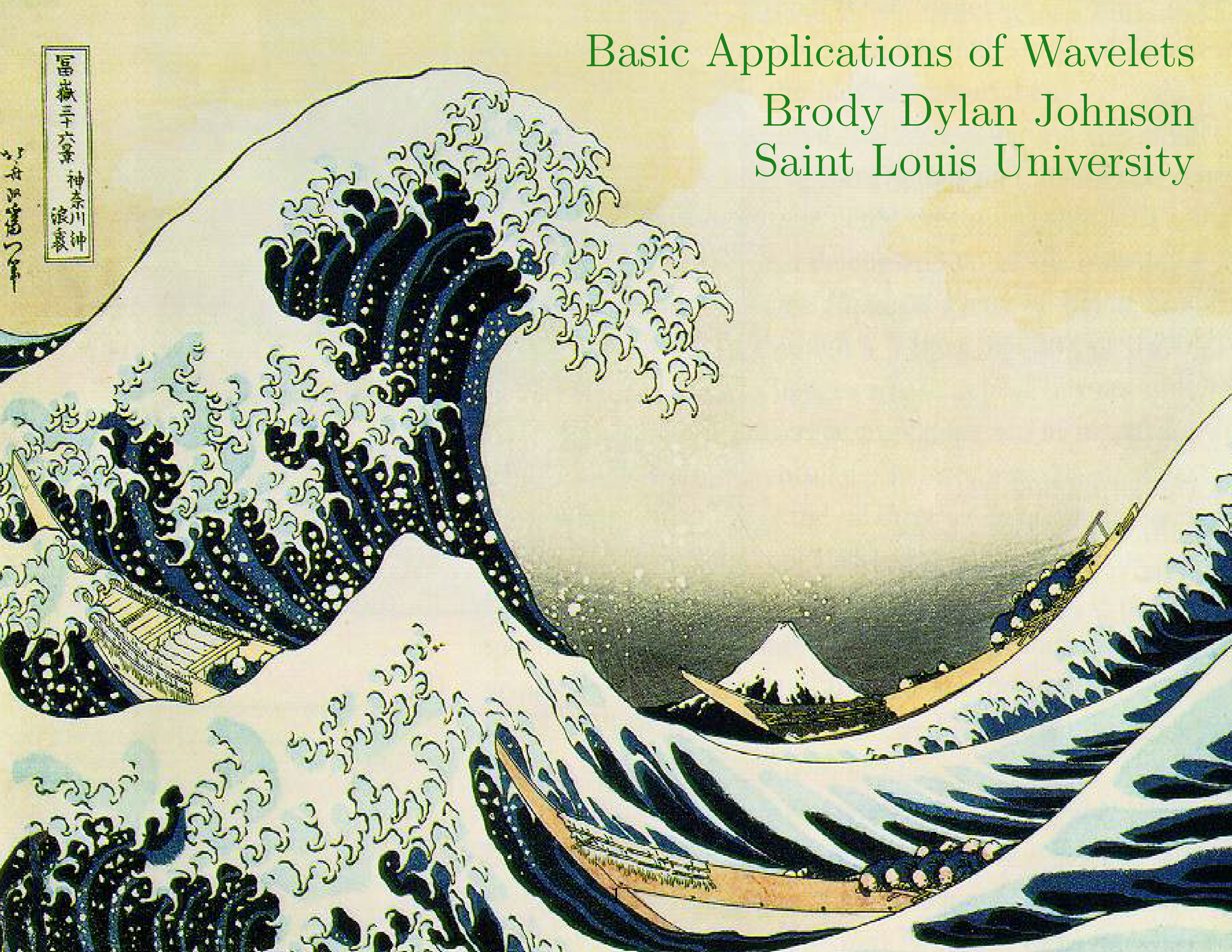


Basic Applications of Wavelets

Brody Dylan Johnson
Saint Louis University

Outline:

- Multiresolution Analysis & Wavelets
  - ⋆ Definitions & Basic properties
  - ⋆ Low-pass & high-pass filters
  - ⋆ Vanishing moments
  - ⋆ Discrete Wavelet Transform (DWT)
  - ⋆ Implementation on finite-dimensional signals
- Applications
  - ⋆ Compression
  - ⋆ Noise Reduction

Starting Point:

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

$$\hat{f}(\xi) = \int_{\mathbb{R}} f(x) \exp(-2\pi i x \xi)\, dx.$$

- Translation: $Tf(x) = f(x-1)$.
- Dilation: $Df(x) = \sqrt{2} f(2x)$.
- We say $\psi \in L^2(\mathbb{R})$ is an *orthonormal wavelet* if the collection

$$\{D^j T^k \psi \,:\, j, k \in \mathbb{Z}\}$$

is an orthonormal basis for $L^2(\mathbb{R})$.

Multiresolution Analysis: (1 of 3)

- The notion of a multiresolution analysis plays a major role in the application of wavelets.
- A collection of close subspaces of $L^2(\mathbb{R})$ is called a *multiresolution analysis (MRA)* if
  1. $f \in V_{j+1} \;\Leftrightarrow\; f(2\cdot) \in V_j$, $j \in \mathbb{Z}$;
  2. $V_j \subseteq V_{j+1}$, $j \in \mathbb{Z}$;
  3. $\overline{\cup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R})$;
  4. $\cap_{j\in\mathbb{Z}} V_j = \{0\}$;
  5. There exists $\varphi \in V_0$ (the *scaling function*) such that the collection $\{\varphi(\cdot - k) : k \in \mathbb{Z}\}$ is an orthonormal basis for V_0.

Multiresolution Analysis: (2 of 3)

- Define W_j, $j \in \mathbb{Z}$, to be the closed subspace of $L^2(\mathbb{R})$ such that

$$V_{j+1} = V_j \oplus W_j.$$

- One means for constructing an orthonormal wavelet is to find $\psi \in V_1$ such that $\{T^k \psi\}_{k \in \mathbb{Z}}$ is an ONB of W_0. This follows very naturally from the MRA properties.

- Because $V_0 \subseteq V_1$, there exists $\{\alpha_k\}_{k \in \mathbb{Z}} \in \ell^2(\mathbb{Z})$ such that

$$\varphi(x) = 2 \sum_{k \in \mathbb{Z}} \alpha_k \varphi(2x - k), \tag{1}$$

  which is equivalent to

$$\hat{\varphi}(2\xi) = m_0(\xi) \hat{\varphi}(\xi),$$

  where $m_0(\xi) := \sum_{k \in \mathbb{Z}} \alpha_k \exp\left(-2\pi i k \xi\right)$ is called the *low-pass filter*.

Multiresolution Analysis (3 of 3)

- It is well known that, given a scaling function φ, one obtains an orthonormal MRA wavelet ψ by defining

$$\hat{\psi}(2\xi) = \exp(2\pi i\xi)\overline{m_0(\xi + \frac{1}{2})}\hat{\varphi}(\xi).$$

  Let $m_1(\xi) = \exp(2\pi i\xi)\overline{m_0(\xi + \frac{1}{2})}$, which is called the *high-pass filter*.

- This definition is equivalent to writing

$$\psi(x) = 2\sum_{k\in\mathbb{Z}} \beta_k \varphi(2x - k), \tag{2}$$

  with $m_1(\xi) := \sum_{k\in\mathbb{Z}} \beta_k \exp(-2\pi i k\xi)$.

  Note: $\beta_k = (-1)^{k+1}\alpha_{-k-1}$, $k \in \mathbb{Z}$.

Vanishing Moments: (1 of 3)

- If φ is compactly supported then m_0 is a trigonometric polynomial, $\hat{\varphi}$ is continuous, and it follows that

$$m_0(0) = 1 \quad \text{and} \quad \hat{\varphi}(0) = 1.$$

  In other words, $\int_{\mathbb{R}} \varphi(x)\,dx = 1.$

- If φ is compactly supported, then so is ψ and, moreover, it follows that

$$m_1(0) = 0 \quad \text{and} \quad \hat{\psi}(0) = 0,$$

  which means $\int_{\mathbb{R}} \psi(x)\,dx = 0.$

- Obviously, if m_0 is a trigonometric polynomial then m_1 will be as well.

Vanishing Moments (2 of 3)

- If ψ is a compactly supported orthonormal wavelet then for $f \in L^2(\mathbb{R})$ one has

$$f = \sum_{j,k \in \mathbb{Z}} \left\langle f, D^j T^k \psi \right\rangle \, D^j T^k \psi.$$

  For smooth functions f it is desirable that this representation be sparse, i.e., that few of the inner products $\left\langle f, D^j T^k \psi \right\rangle$, $j, k \in \mathbb{Z}$, are essential in the above representation of f.

- A wavelet ψ is said to have vanishing moments of order $m \in \mathbb{N}$ if for $0 \leq n \leq m$,

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

Vanishing Moments (3 of 3)

- Suppose that ψ is an MRA wavelet associated with a compactly supported scaling function φ. Then ψ has vanishing moments of order m if and only if

$$\frac{d^n}{d\xi^n}\left[m_1(\xi)\right]\Big|_{\xi=0} = 0, \quad 0 \leq n \leq m. \tag{3}$$

- Equation (3) is equivalent to the following identity for the filter coefficients of m_1:

$$\sum_{k\in\mathbb{Z}} k^n \beta_k = 0, \quad 0 \leq n \leq m.$$

  This can be viewed as a discrete notion of vanishing moments of order m, which suggests that vanishing moments for ψ should provide sparse representations for "smooth" discrete signals.

Interscale Relationships:

- The MRA refinement equations (1) and (2) provide interscale relationships for the sequences of inner products,

$$\{\langle f, D^j T^k \varphi\rangle\}_{k\in\mathbb{Z}}, \ \{\langle f, D^j T^k \psi\rangle\}_{k\in\mathbb{Z}}, \ \& \ \{\langle f, D^{j+1} T^k \varphi\rangle\}_{k\in\mathbb{Z}}.$$

- Namely, for $f \in L^2(\mathbb{R})$ and each $j \in \mathbb{Z}$:

$$\langle f, D^j T^n \varphi\rangle = \sqrt{2}\sum_{k\in\mathbb{Z}} \overline{\alpha_{k-2n}} \langle f, D^{j+1} T^k \varphi\rangle, \tag{4}$$

and

$$\langle f, D^j T^n \psi\rangle = \sqrt{2}\sum_{k\in\mathbb{Z}} \overline{\beta_{k-2n}} \langle f, D^{j+1} T^k \varphi\rangle. \tag{5}$$

These are easily recognizable as convolutions in $\ell^2(\mathbb{Z})$.

Discrete Wavelet Transform (1 of 3):

- Because $\hat{\varphi}(0) = 1$ we can think of $\langle f, D^j T^k \varphi \rangle$ as a local average of f near $x = 2^{-j}k$ at scale j. (as $j \to \infty$ this average becomes "more" localized)

- Example: In the case of the Haar wavelet & MRA, we have $\varphi = \chi_{[0,1)}$ and $\langle f, D^j T^k \varphi \rangle$ is proportional to the average of f on $\left[2^{-j}k, 2^{-j}(k+1)\right)$.

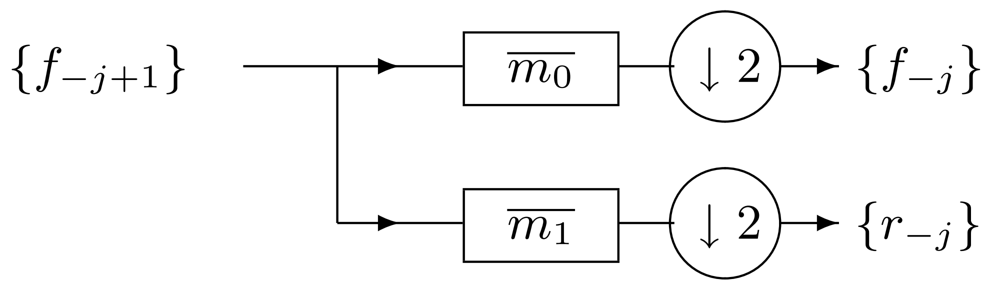
- Given a digital signal $\{f_k\}_{k \in \mathbb{Z}}$ one then interprets the values f_k, $k \in \mathbb{Z}$, as the local averages of a "continuous" function at scale $j = 0$. In other words, one defines

$$\langle f, T^k \varphi \rangle := f_k.$$

The interscale relationship is then used to decompose the signal at scales $j < 0$. This is the essence of discrete wavelet transform.

Discrete Wavelet Transform (2 of 3):

- At each scale j, the sequence f_{-j} is again decomposed and the DWT consists of all of the sequences r_{-j}, $j > 0$.

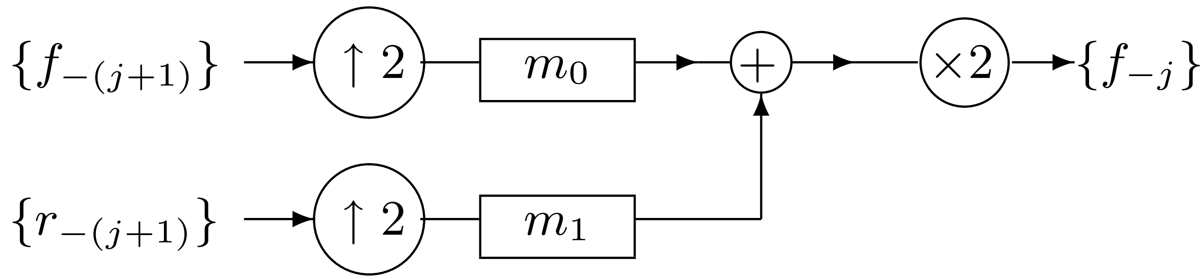


- In practice, a signal has finitely many samples, say $N = 2^M$. We then define the sequence f_0 by periodic extension. At each scale j the period is reduced by 2 and after M decomposition stages the signal is constant. The total size of all the required wavelet coefficients equals that of the original signal:

$$2^M = \underbrace{(2^{M-1} + 2^{M-2} + \cdots + 1)}_{\text{from } r_{-j},\ 1 \le j \le M} + \underbrace{1.}_{\text{from } f_{-M}}$$

Discrete Wavelet Transform (3 of 3):

- Our choice of m_1 allows for a similar reconstruction operation from the DWT:



- This discussion has been limited to one-dimensional signals; however, one can easily extend the above analysis to two-dimenional signals by considering separable products of the scaling function and wavelet or the corresponding filters:

$$\phi(x,y) = \varphi(x)\varphi(y) \qquad \psi_1(x,y) = \psi(x)\varphi(y)$$
$$\psi_2(x,y) = \varphi(x)\psi(y) \qquad \psi_3(x,y) = \psi(x)\psi(y).$$

An Example:

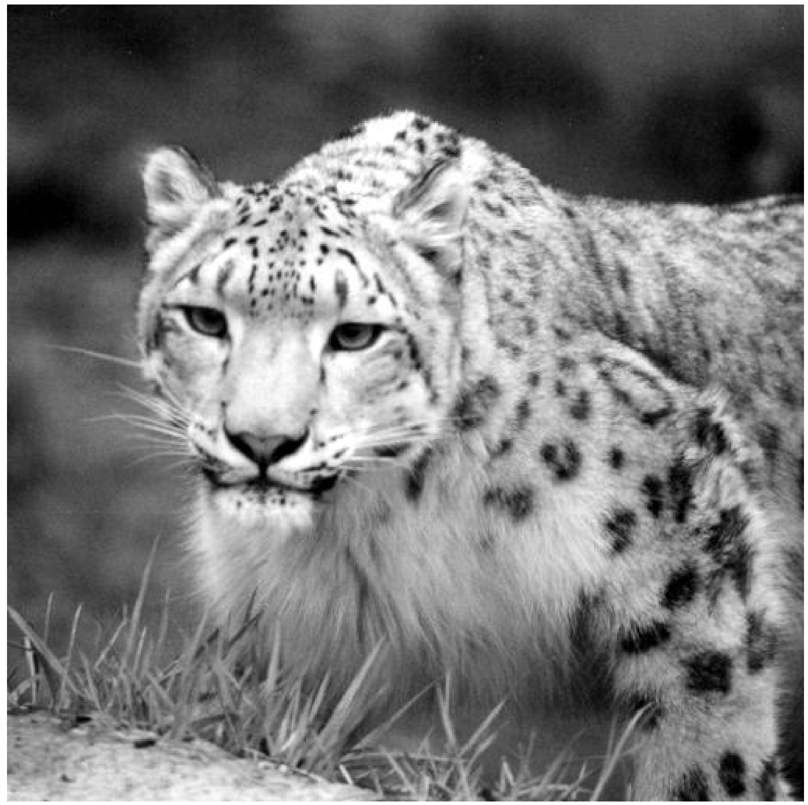


Figure 1: Original 512×512 snow leopard image.

An Example:

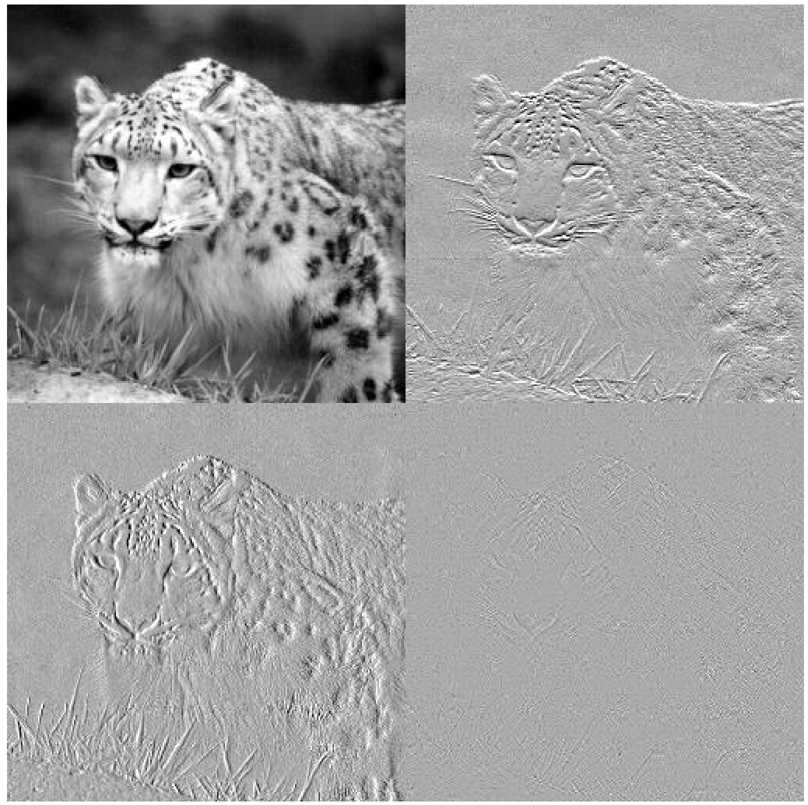


Figure 2: Scale $j = -1$: snow leopard image. (Haar filters)

An Example:

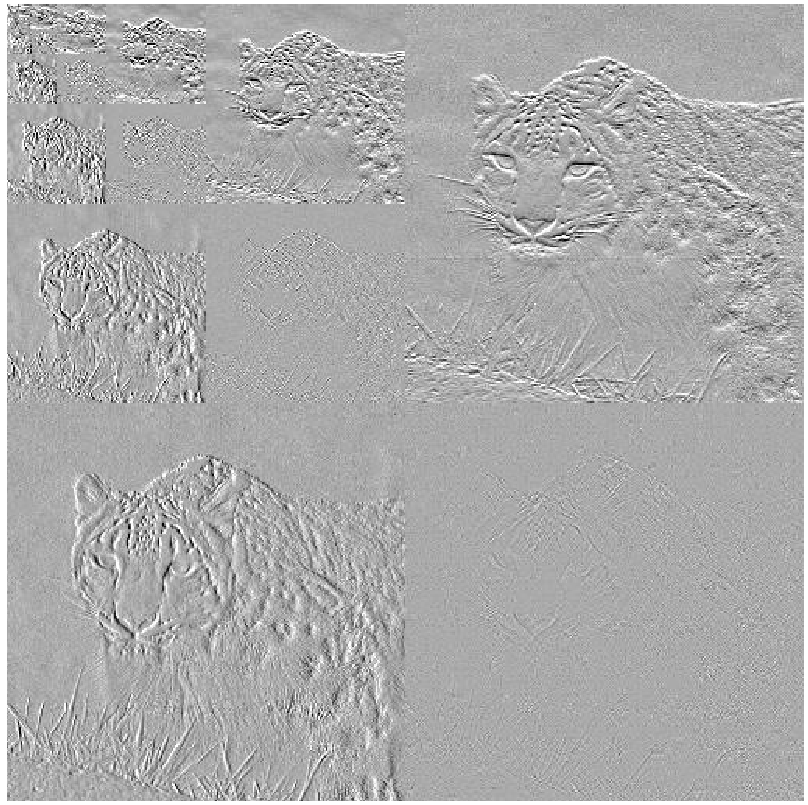


Figure 3: Full DWT decomposition of snow leopard image.

Image Compression:

- Looking at the DWT of the snow leopard image we see mostly gray, which corresponds to zero for the display of the DWT. This suggests that the number of coefficients $\left\langle f, D^j T^k \psi \right\rangle$ that are essential for reconstructing the original image is small.
- By ignoring the "small" coefficients we can compress the image. This is called *thresholding*.
- We will quantify the benefit and cost of thresholding by considering two quantities:

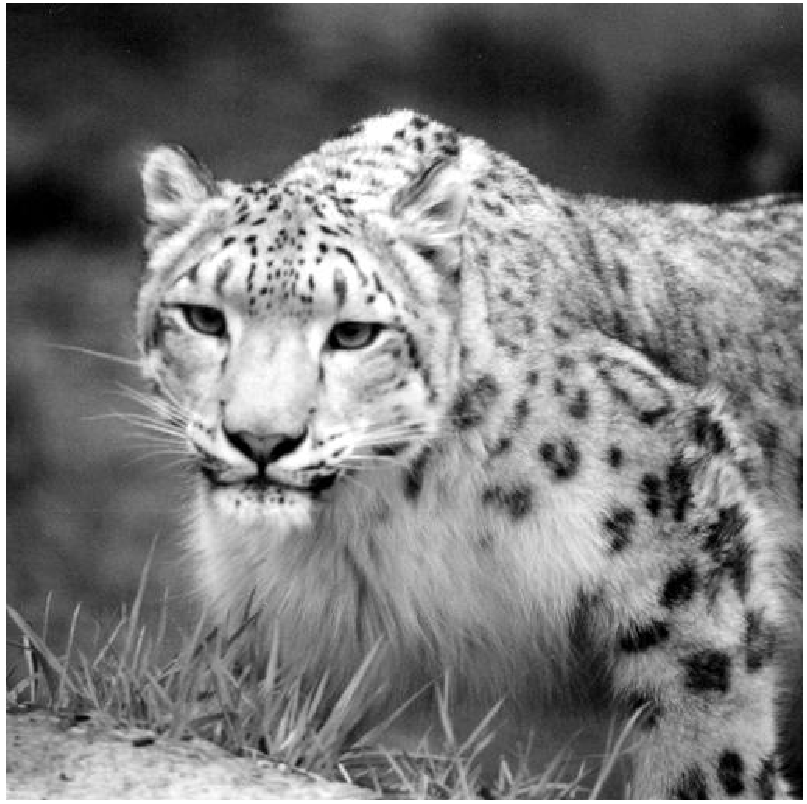
$$\text{Compression Factor} := \frac{\text{Total \# of pixels}}{\text{\# of coefficients} \geq \text{threshold}},$$

$$\text{Mean Squared-Error} := \frac{1}{N^2} \sum_{j,k=1}^{N} \left(\tilde{f}(j,k) - f(j,k) \right)^2 .$$

Image Compression Examples:

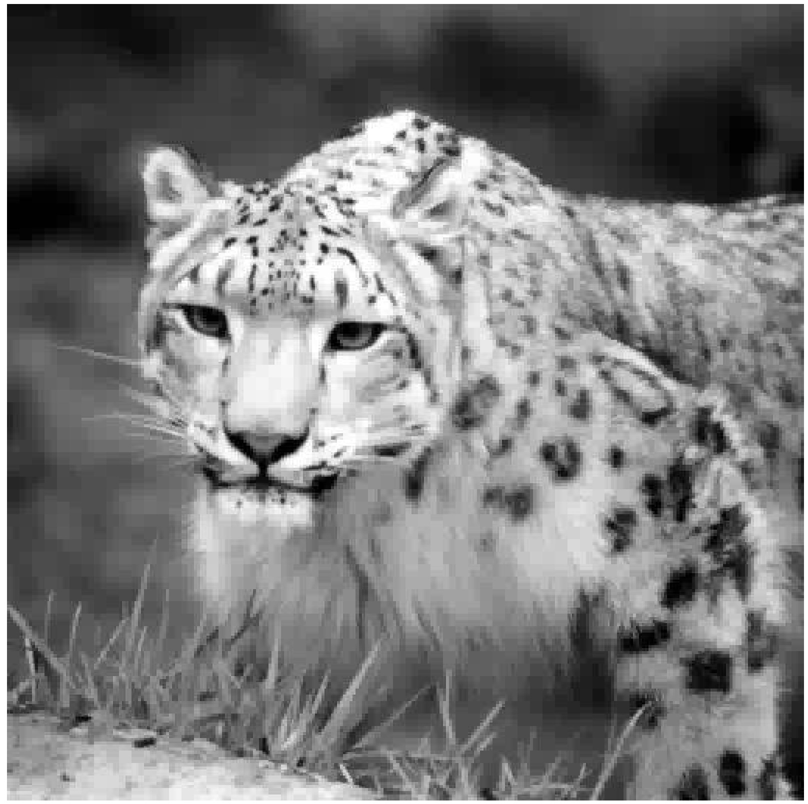
Haar filters, Threshold=5: C.F. ≈ 13.6 & M.S.E. ≈ 2.10.

Image Compression Examples:



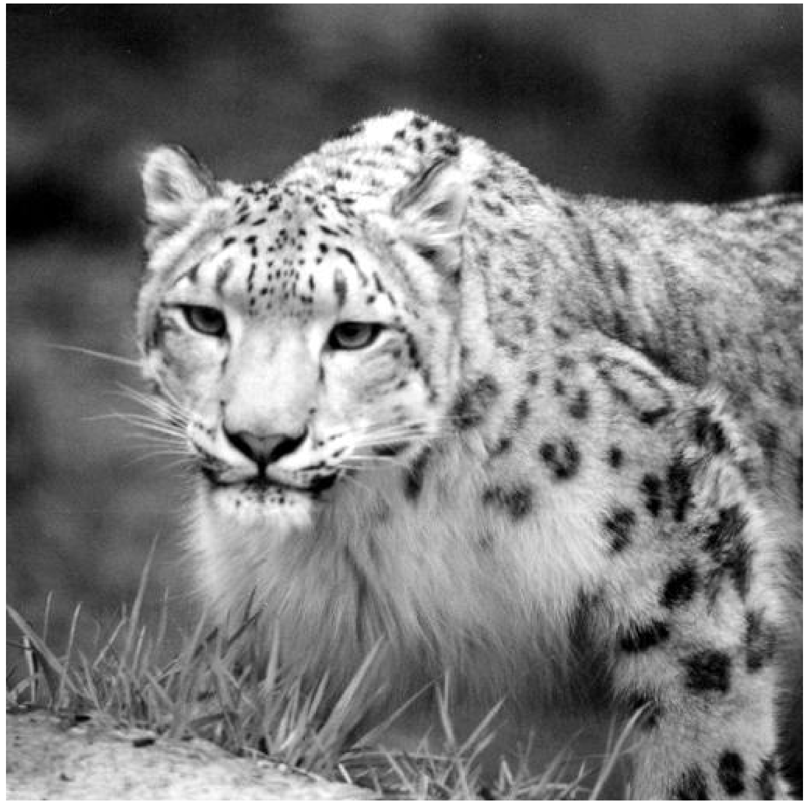
Original image.

Image Compression Examples:



Haar filters, Threshold=10: C.F. ≈ 27.0 & M.S.E. ≈ 4.53.

Image Compression Examples:

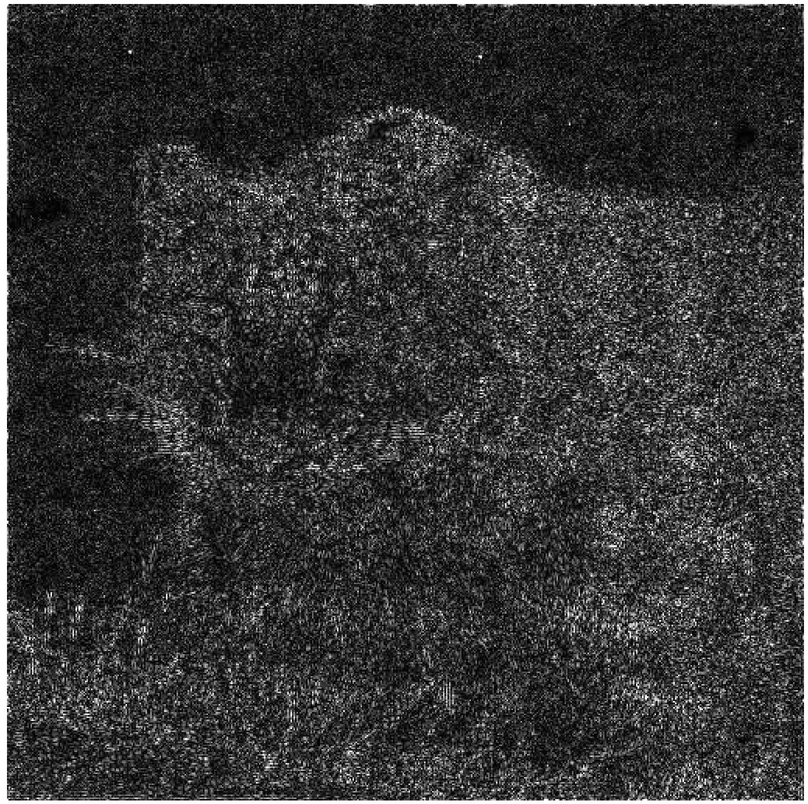


Original image.

Image Compression Examples:

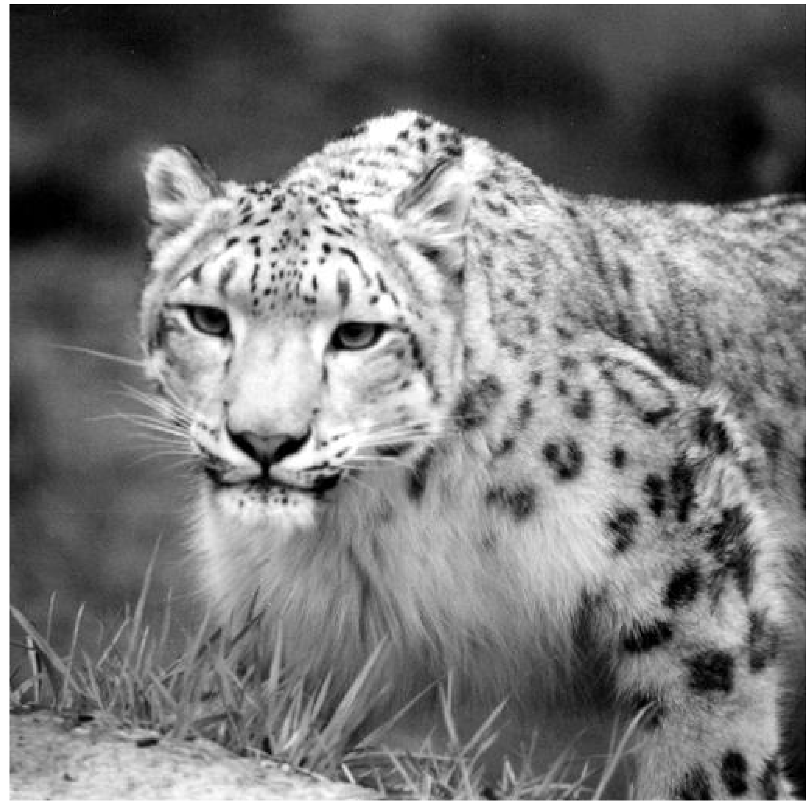
Daubechies filters ($n = 6$), Threshold=10: C.F. $\approx$ 33.8 & M.S.E. $\approx$ 2.66.

Image Compression Examples:



Absolute error $\times 10$: Daubechies filters ($n = 6$), Threshold=10.

Image Compression Examples:

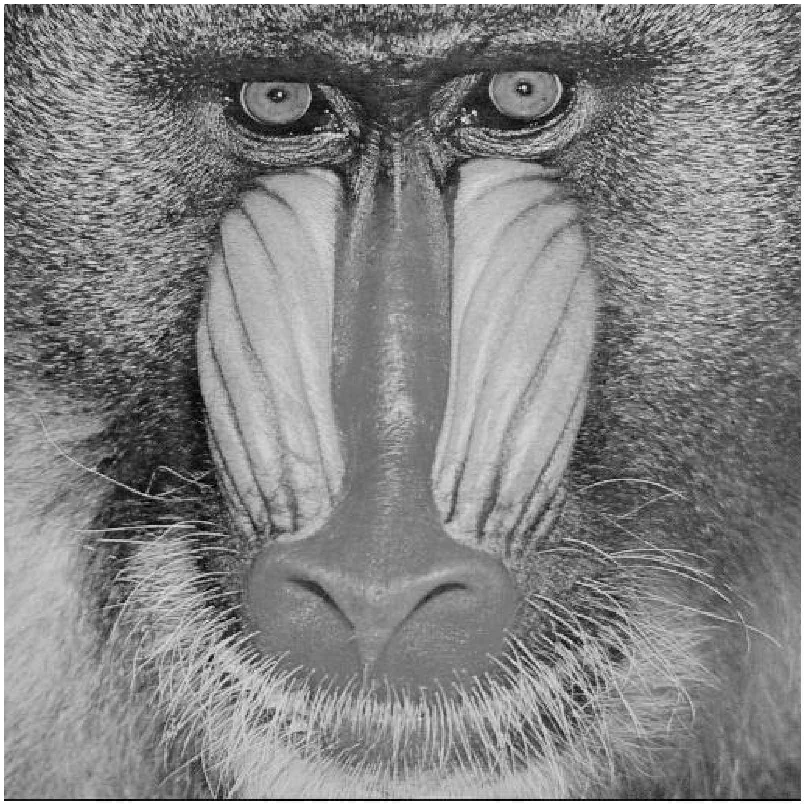
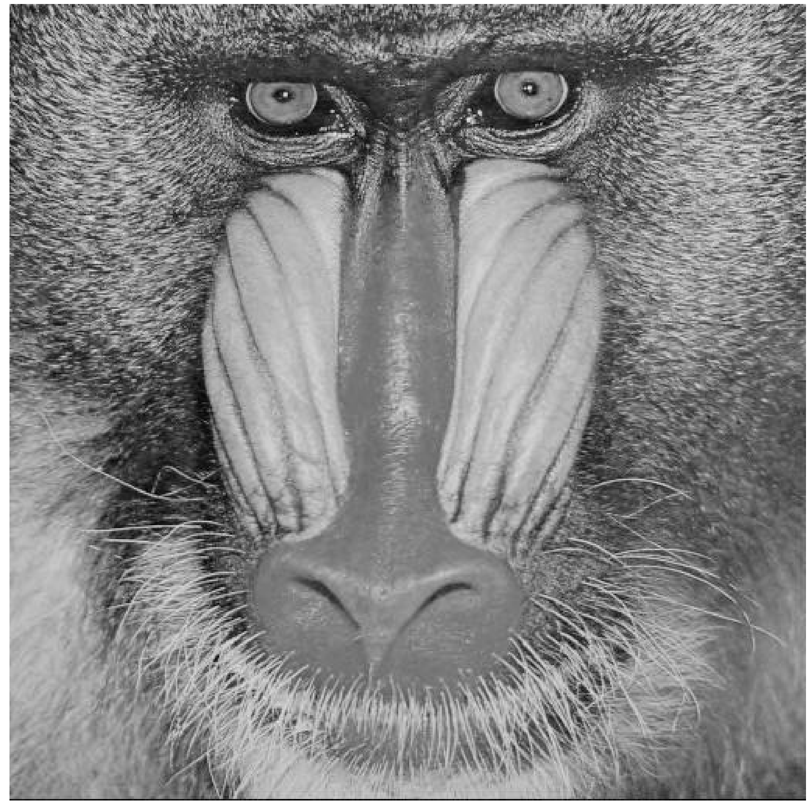
Daubechies filters ($n = 6$), Threshold=3: C.F. $\approx$ 12.3 & M.S.E. $\approx$ 0.91.

Image Compression Examples:

Daubechies filters ($n = 6$), Threshold=3: C.F. $\approx$ 17.0 & M.S.E. $\approx$ 0.83.

Image Compression Examples:

Daubechies filters ($n = 6$), Threshold=3: C.F. $\approx$ 5.1 & M.S.E. $\approx$ 1.15.

Image Denoising:

- Suppose that an image is corrupted by random noise.
- The noise will cause scattered coefficients in "smooth" areas to be large compared to neighboring coefficients in the image's DWT:

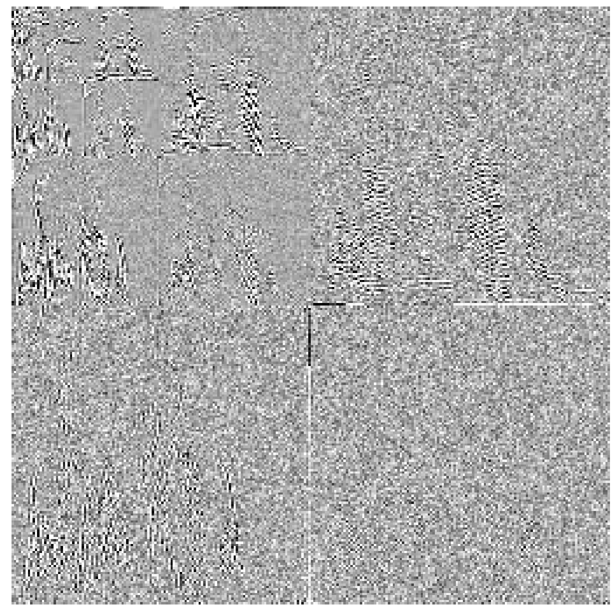
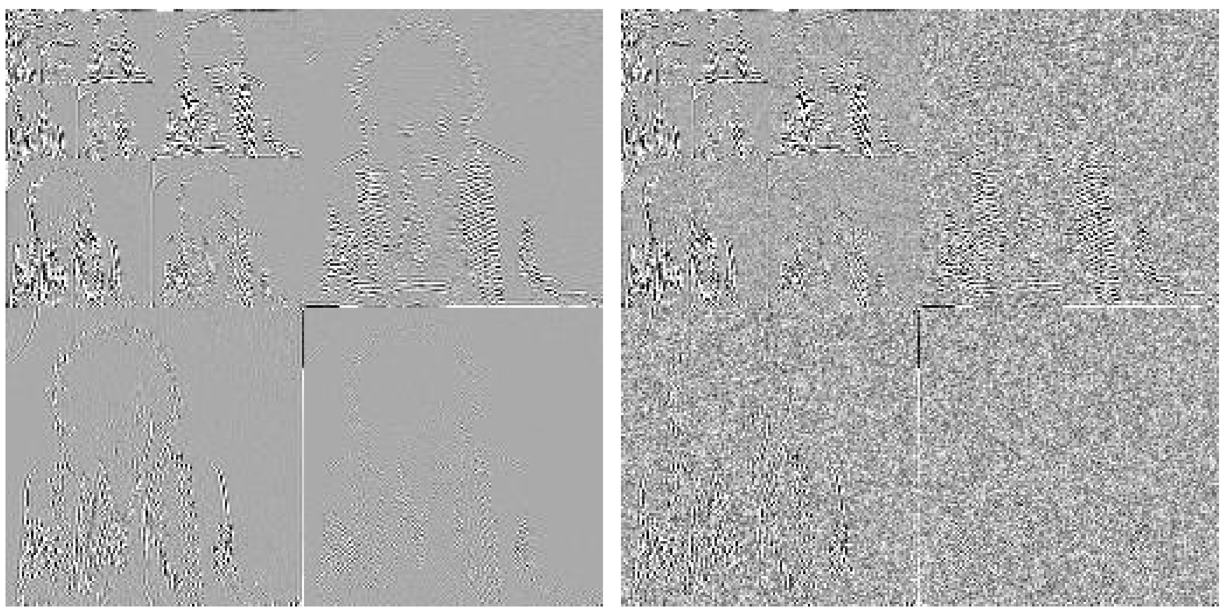


Image Denoising:

- Thresholding can reduce such noise, but at the cost of distortion due to the removal of other meaningful, but small components of the DWT. We will quantify the benefit of thresholding by considering M.S.E. values comparing the noisy and denoised images to the original.



Denoising Examples:

Daubechies filters, Threshold=11: M.S.E. $\approx$ 12.00 before denoising and M.S.E. $\approx$ 6.50 after denoising.

Denoising Examples:

Daubechies filters, Threshold=6: M.S.E. $\approx$ 11.9 before denoising and M.S.E. $\approx$ 3.72 after denoising. (à trous filtering)