Mathematics of Image and Data Analysis
Math 5467

Lecture 18: Wavelets and Image Classification

Instructor: Jeff Calder
Email: jcalder@umn.edu

http://www-users.math.umn.edu/~jwcalder/5467S21

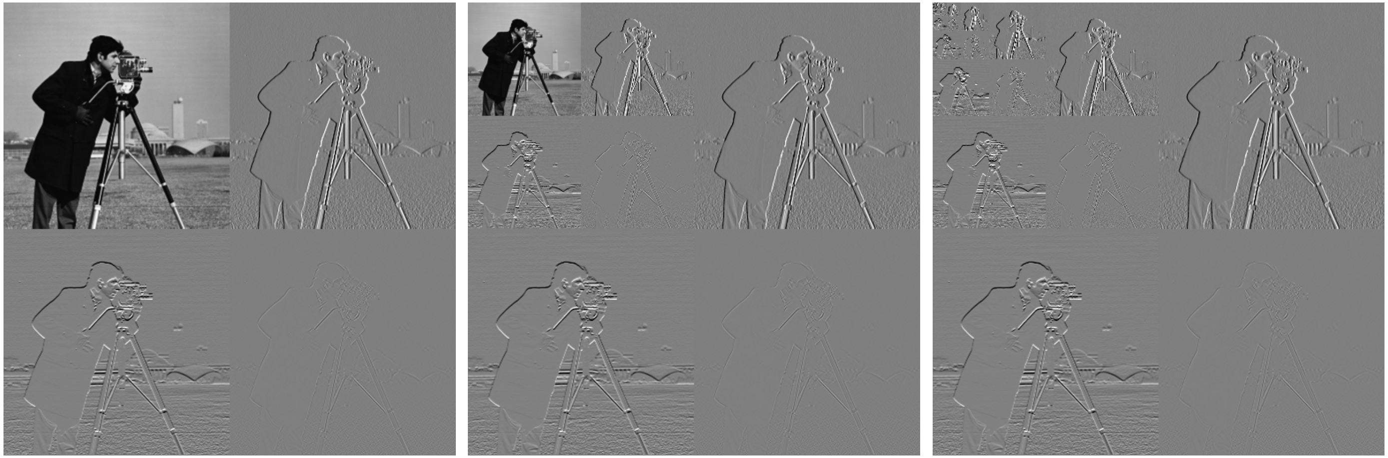
Last time

- The Wavelet Transform (1D and 2D Haar Wavelets)

Today

- Wavelets-based image classification
- General discrete wavelet transformations

Recall Haar Wavelet



(a) 1 Level (b) 2 Levels (c) 9 Levels

Figure 1: The Haar Wavelet Transformation of levels 1, 2, and 9 on the cameraman image. The approximation image is placed in the upper left coner, the horizontal detail in the lower left, the vertical detail in the upper right, and the diagonal detail in the lower right.

Wavelets on MNIST

(a) Zero

(b) One

Figure 2: One level Haar Wavelet Transform of MNIST digits.

Pooling detail features

Let $H(i,j)$ and $V(i,j)$ be horizontal + vertical detail.

$1 \le i,j \le 14$

(1) Take absolute values $|H(i,j)|$ and $|V(i,j)|$.

(2) Pooling: Sum absolute values of filter responses

$$v = \sum_{i=1}^{14} \sum_{j=1}^{14} |V(i,j)| \quad , \quad h = \sum_{i=1}^{14} \sum_{j=1}^{14} |H(i,j)|.$$

(*) Introduces Translation Invariance.

→ Pooled values v, h are unchanged under translations of the images.

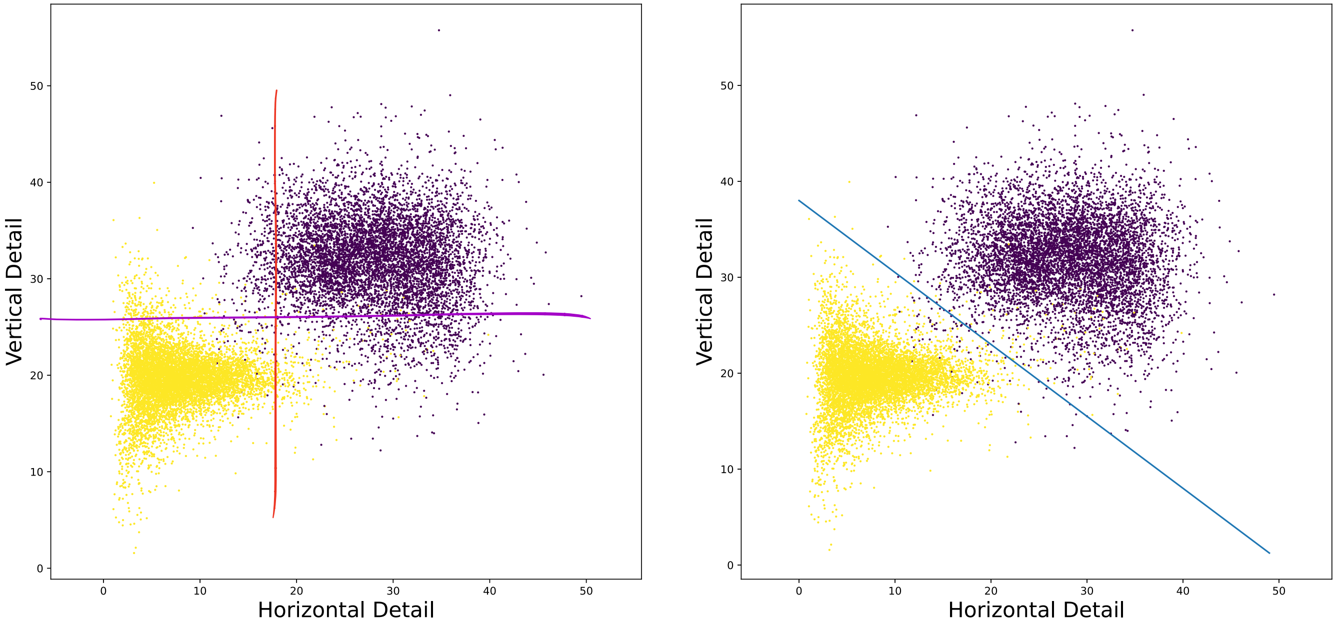
This gives 2 numbers, (h, v), for each MNIST digit. ↑ **Features**

- Think of as an embedding of MNIST into $\mathbb{R}^2$.

$(h, v) \longrightarrow$ Wavelet-Based Features.

Wavelet Features of MNIST Digits

MNIST 0 = purple, 1 = Yellow.



(a) Horizontal vs Vertical

(b) Decision boundary

Wavelet-based MNIST classification (.ipynb)

Connections to Convolutional Neural Networks

Horizontal detail $H = I * \psi_1$

Vertical detail $V = I * \psi_2$

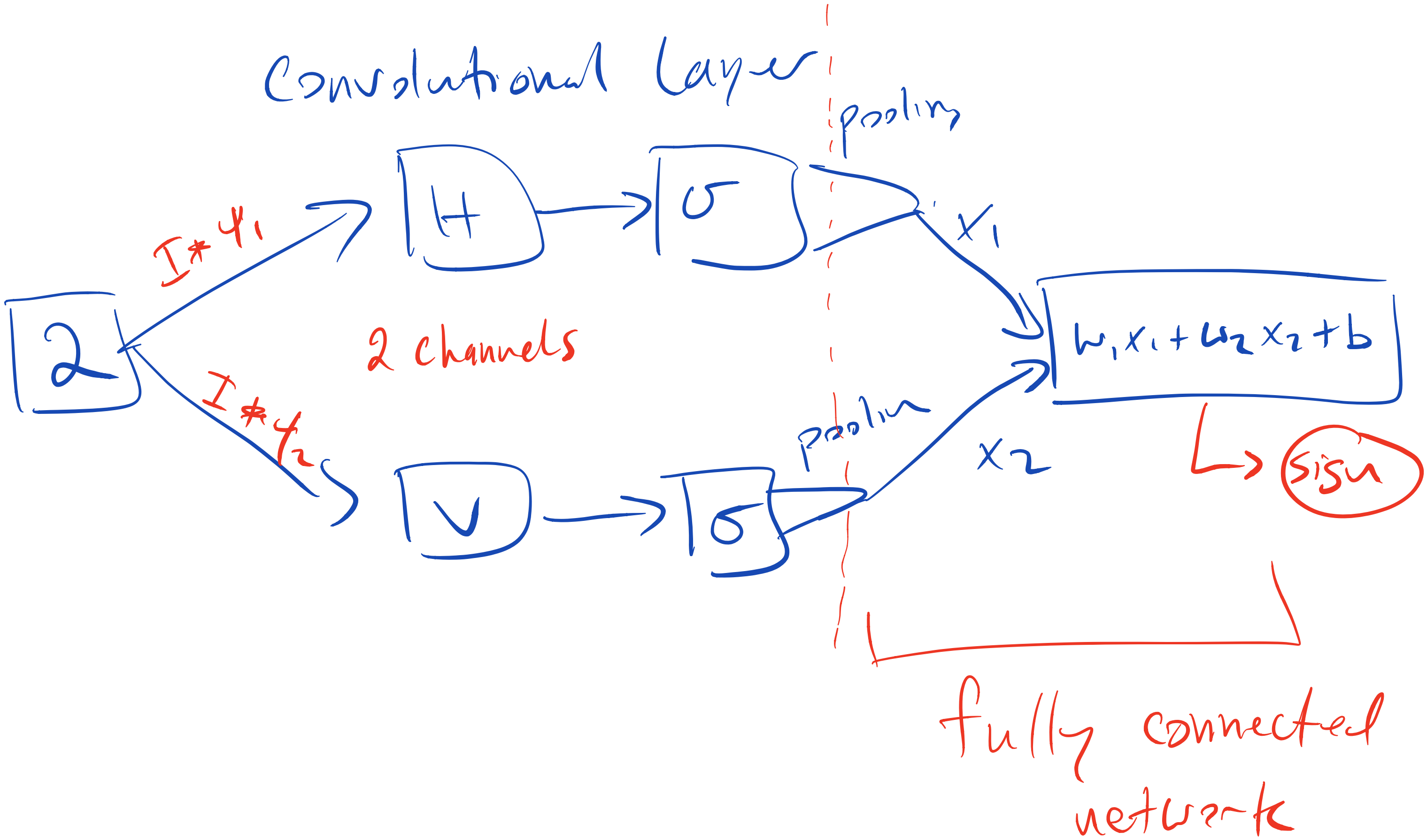
$$\psi_1 = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$\psi_2 = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}$$

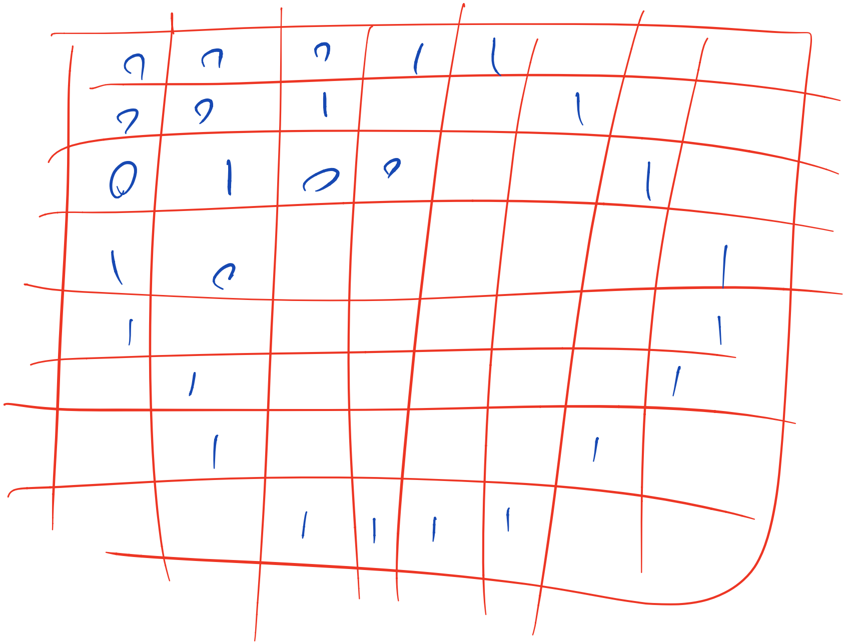
Classifier has the form

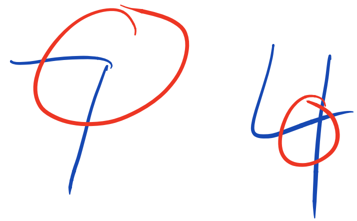
$$f = \text{sign}\Big(w_1 \cdot \text{pool}\big(\sigma(\underbrace{I * \psi_1}_{\text{Horizontal}})\big) + w_2 \cdot \text{pool}\big(\sigma(\overbrace{I * \psi_2}^{\text{vert.}})\big) + b \Big)$$

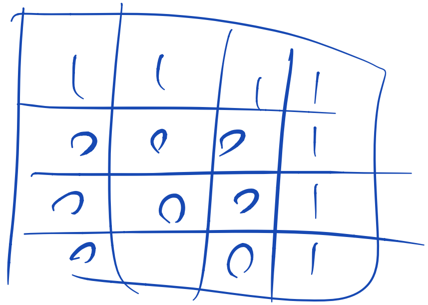
$$\sigma(t) = |t| \, , \quad \text{pool}(v) = \sum_i \sum_j v(i,j)$$


Convolutional layer
pooling
I * ψ1
H
σ
x1
2
2 channels
w1x1 + w2x2 + b
I * ψ2
pooling
x2
sigm
V
σ
fully connected
network

8x8







General Wavelets

Recall that the 1D Haar Wavelet Transformation on a signal of length $n = 8$ is represented by the matrix W defined by

$$W = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 & 1 & 1 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}. \tag{1}$$

The rows of the matrix W are the basis vectors for the Haar Wavelet Transformation.

The Haar Mother Wavelet

It turns out vectors can be written as rescaled and shifted versions of a *mother wavelet* ψ given by

$$\psi(t) = \begin{cases} 0, & \text{if } t < 0, \\ -1, & \text{if } 0 \leq t < 1, \\ 1, & \text{if } 1 < t < 2, \\ 0, & \text{if } t \geq 2. \end{cases} \tag{2}$$

Consider the rescaled wavelets $\psi_{j,k} \in L^2(\mathbb{Z}_8)$ given by

$$\psi_{j,k}(\ell) = \psi\left(2^{-j}\ell - k\right). \tag{3}$$

Exercise 1. Show The last 4 rows of W are exactly $\psi_{0,0}, \psi_{0,1}$, $\psi_{0,2}$ and $\psi_{0,3}$, the third and forth rows of W are $\psi_{1,0}$ and $\psi_{1,1}$, and the second row of W is $\psi_{2,0}$. △

Approximation coefficient

The first row—the approximation coefficient—is obtained by another function called the *scaling function*, which in this case is

$$\varphi(t) = \begin{cases} 0, & \text{if } t < 0, \\ 1, & \text{if } 0 \le t < 2, \\ 0, & \text{if } t \ge 2. \end{cases} \tag{4}$$

The first row of W is the rescaled scaling function $\varphi_{3,0}(\ell) = \varphi(2^{-3}\ell)$. We note that the mother wavelet ψ and the rescaling function φ are precisely the filters used to obtain the approximation and detail coefficients in the recursive definition of the Haar Wavelet Transformation given in previous lectures.

General definitions

Definition 2. For a general signal of length n that is a power of 2, the Haar Wavelet basis is given by the collection of functions $\psi_{j,k}$ with $j = 0, 1, 2, \ldots, \log_2(n) - 1$ and $k = 0, 1, 2, \ldots, 2^{\log_2(n)-1-j} - 1$, and the coarsest scale approximation coefficient $\varphi_{\log_2(n),0}$.

The Haar basis is an orthogonal basis.

Lemma 3. *For any* (j_1, k_1) *and* (j_2, k_2)*, not identically equal, we have*

$$\sum_{\ell \in \mathbb{Z}} \psi_{j_1,k_1}(\ell)\psi_{j_2,k_2}(\ell) = 0.$$

Exercise 4. Proof Lemma 3. △

It is also commont to choose a different normalization of the wavelets, given by

$$\psi_{j,k} = 2^{-\frac{j+1}{2}}\psi(2^{-j}\ell - k).$$

This ensures that $\|\psi_{j,k}\| = 1$ for all j, k, so the Haar Wavelet basis is an orthonormal basis.

Other wavelets

This general, and non-recursive, definition of the Haar Wavelet makes it possible to construct other types of wavelets by choosing other mother wavelets ψ. One common choice is the Ricker Wavelet

$$\psi(t) = \frac{2}{\sqrt{3\sigma}\pi^{1/4}} \left(1 - \left(\frac{t}{\sigma}\right)^2\right) e^{-\frac{t^2}{2\sigma^2}}. \tag{5}$$