Feature Extraction

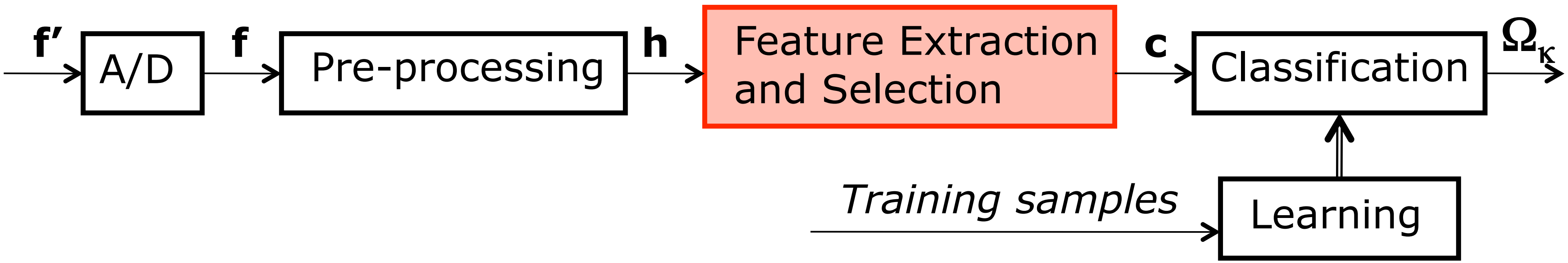
Spectrogram, Walsh Transform, Haar Transform

Dr. Elli Angelopoulou

Lehrstuhl für Mustererkennung (Informatik 5)

Friedrich-Alexander-Universität Erlangen-Nürnberg

Pattern Recognition Pipeline



- One common method for heuristic feature extraction is the projection of a signal $\vec{h}$ or $\vec{f}$ on a set of orthogonal basis vectors (functions), $\Phi = \left[\vec{\varphi}_1, \vec{\varphi}_2, \ldots, \vec{\varphi}_M\right]$

$$\vec{c} = \Phi^T \vec{f}$$

Speech Processing and Fourier Transform

- In speech processing we often want to analyze the sound of individual vowels or consonants or syllables.
- We want to analyze the sound signal in frames that last 10-20msec.
- Goal: compute the Fourier transform for each frame.
- How?
- Overlap the sound signal with a function that turns everything outside the frame of interest into 0.

Short Time Fourier Transform

- The idea of ignoring the signal (turning it to zero) for values outside a small time window has a broader application outside speech processing.
- It is known as the Short Time Fourier Transform.
- **Short Time Fourier Transform**: apply a windowing function to each frame before applying the Fourier transform.

$$F(\tau,\omega) = \int_{-\infty}^{\infty} f(t)w(t-\tau)e^{-j\omega t}dt$$

- Compared to the Fourier transform

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}dt$$

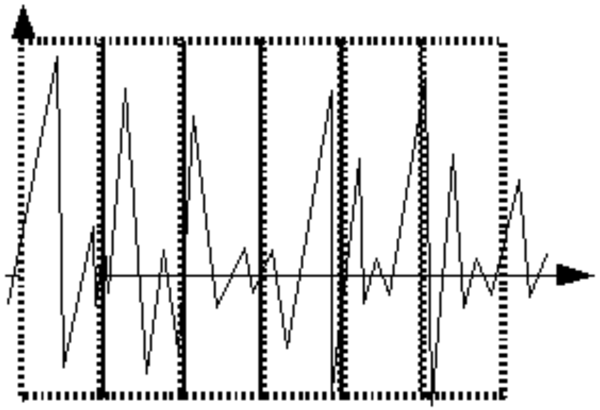
Short Time Fourier Transform - continued

- Short Time Fourier Transform:

$$F(\tau,\omega) = \int_{-\infty}^{\infty} f(t)w(t-\tau)e^{-j\omega t}dt$$

where $w(t)$ is the windowing function.

- It is used in determining the sinusoidal frequency and phase content of local sections of a signal as it changes over time.



Spectrogram

- In speech processing we use a special feature based on the Short Time Fourier Transform, called the Spectrogram:

$$\text{Spectrogram}\{f(t)\} = \left|F(\tau,\omega)\right|^2$$

- Spectrograms are used in:
  - identifying phonetic sounds
  - analyzing the cries of animals
  - analyzing music, sonar/radar signals, speech processing, etc.
- A spectrogram is also called a spectral waterfall, sonogram, voiceprint, or voicegram.
- The instrument that generates a spectrogram is called a sonograph.

Windowing Functions

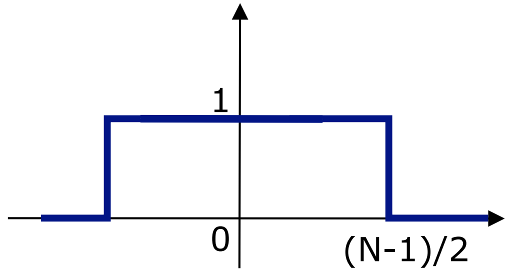
- One can use different windowing functions.
- Let N be the width of the window and $0 \le n \le N-1$.
- Then the time-shifted windowing functions are of the form:

$$w(n) = w_0\left(n - \frac{N-1}{2}\right)$$

where $w_0(t)$ is maximum at $t = 0$.

- Typically N is a power of 2, i.e. $N = m^2$.
- The simplest windowing function is a rectangle window:

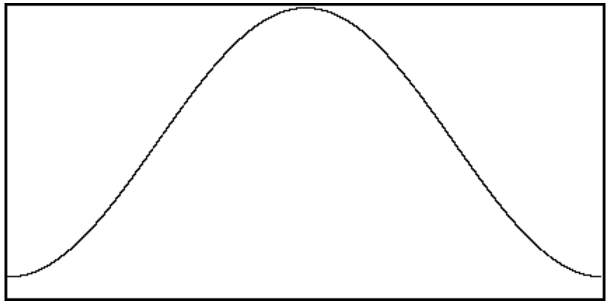
$$w(n) = 1$$



Windowing Functions - continued

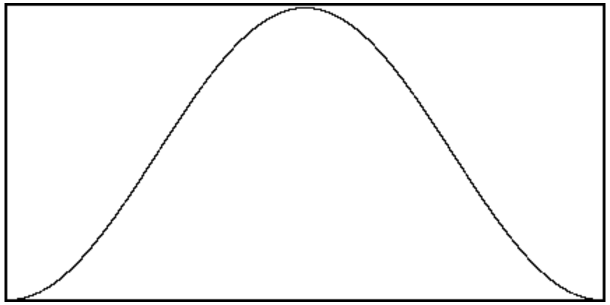
- A well-known windowing function is the **Hamming window**, which is a "raised cosine" proposed by Hamming (raised because it is not zero at the limits).
- It is defined as:

$$w(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right)$$



- Another widely-used windowing function is the **Hann window**:

$$w(n) = 0.5\left(1 - \cos\left(\frac{2\pi n}{N-1}\right)\right)$$



Plots courtesy of http://ecmc.rochester.edu/ecmc/docs/csound-manual-5.06.0/MiscWindows.html

Features based on Fourier Transform - review

- Recall that in the Fourier Transform we use sinusoidal functions for our signal decomposition:

$$e^{2\pi j\omega x} = \cos(2\pi\omega x) + j\sin(2\pi\omega x)$$

- When using the Fourier basis functions as an orthogonal basis, we used the following subset of the sinusoidal functions:

$$e^{-2\pi j\frac{\nu}{M}x} = \cos\left(-2\pi\frac{\nu}{M}x\right) + j\sin\left(-2\pi\frac{\nu}{M}x\right)$$

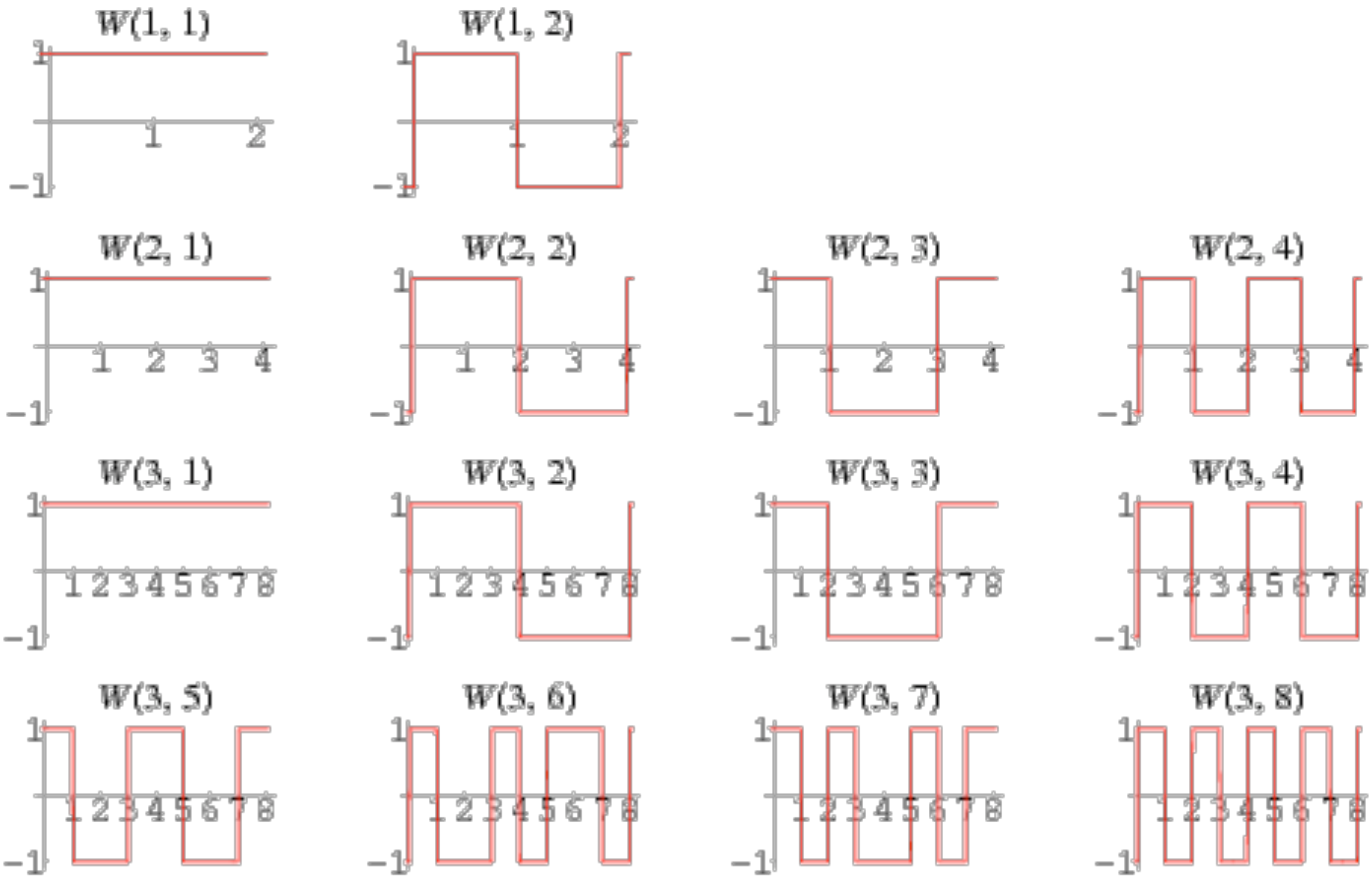
- The problems with such sinusoidal functions is that they are computationally expensive.

Walsh Functions

- Instead one can use a rectangular waveform with a magnitude range [-1, 1].
- One such type of function is the Walsh functions.
- The Walsh function can be thought of as a discrete version of sinus and cosinus functions.
- The frequency of sinusoidal function corresponds to the sequence of the Walsh function transitions.
- The Walsh functions are defined in the interval

$$-\tfrac{1}{2} \leq x \leq \tfrac{1}{2}$$

Walsh Function Plots



Plots courtesy of Wolfram Mathworld http://mathworld.wolfram.com/WalshFunction.html

Definition of Walsh Functions

The continuous Walsh functions are recursively defined:

$$w(x,0) = \begin{cases} 1 & \text{for } -\tfrac{1}{2} \le x \le \tfrac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$

$$w(x,2k+p) = -1^{\left\lfloor \frac{k}{2} \right\rfloor + p} \left(w\left(2\left(x+\frac{1}{4}\right),k\right) + (-1)^{k+p}\, w\left(2\left(x-\frac{1}{4}\right),k\right)\right)$$

for $k = 0,1,2,\ldots$ and $p = 0,1$.

The Walsh functions are orthonormal:

$$\int_{-1/2}^{1/2} w(x,k) \bullet w(x,n) = \begin{cases} 0 & \text{if } k \ne n \\ 1 & \text{if } k = n \end{cases}$$

Hadamard Matrix

- The orthogonal Walsh functions are the basis functions used in the Walsh-Hadamard transform.
- In the Walsh-Hadamard transform the key component is the Hadamard matrix, where the rows of the matrix are the Walsh functions.
- The Hadamard matrix is defined recursively:

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$H_M = H_2 \otimes H_{M/2}$$

$$= H_2 \otimes H_2 \otimes \ldots \otimes H_2 \qquad \text{q factors}$$

where H_M is an MxM Hadamard matrix and $M=2^q$.

Kronecker Product

- In the Hadamard matrix defintion, the operand $\otimes$ denotes the Kronecker product.
- Given an MxM matrix A and and mxm matrix B, their Kronecker product is an Mm x Mm matrix constructed as follows:

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & a_{13}B & \cdots & a_{1M}B \\ a_{21}B & a_{22}B & a_{23}B & \cdots & a_{2M}B \\ a_{31}B & a_{32}B & a_{33}B & \cdots & a_{3M}B \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{M1}B & a_{M2}B & a_{M3}B & \cdots & a_{MM}B \end{bmatrix}$$

Example Hadamard Matrix

- Consider the H_8 matrix: $H_8 = H_2 \otimes H_4 = H_2 \otimes H_2 \otimes H_2$

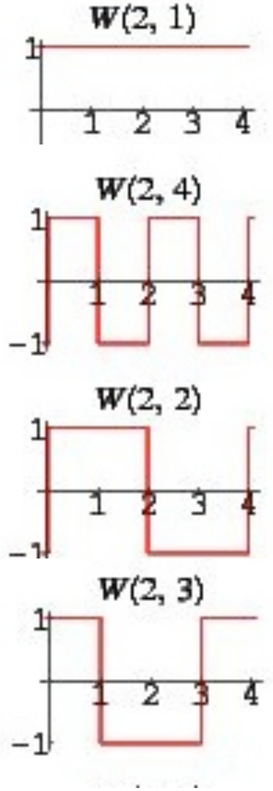
$$H_8 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \otimes H_2$$

$$H_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix}$$

More on the Hadamard Matrix

- The Hadamard matrix is simply just one way of arranging the Walsh functions.
- Consider for example the H_4 matrix.

$$H_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$



Ordering of Walsh Functions

Walsh functions can be ordered in many different ways.

in sequence | Hadamard | dyadic or Paley

(a) (b) (c)

Image adapted from S. Wolfram, http://mathworld.wolfram.com/WalshFunction.html

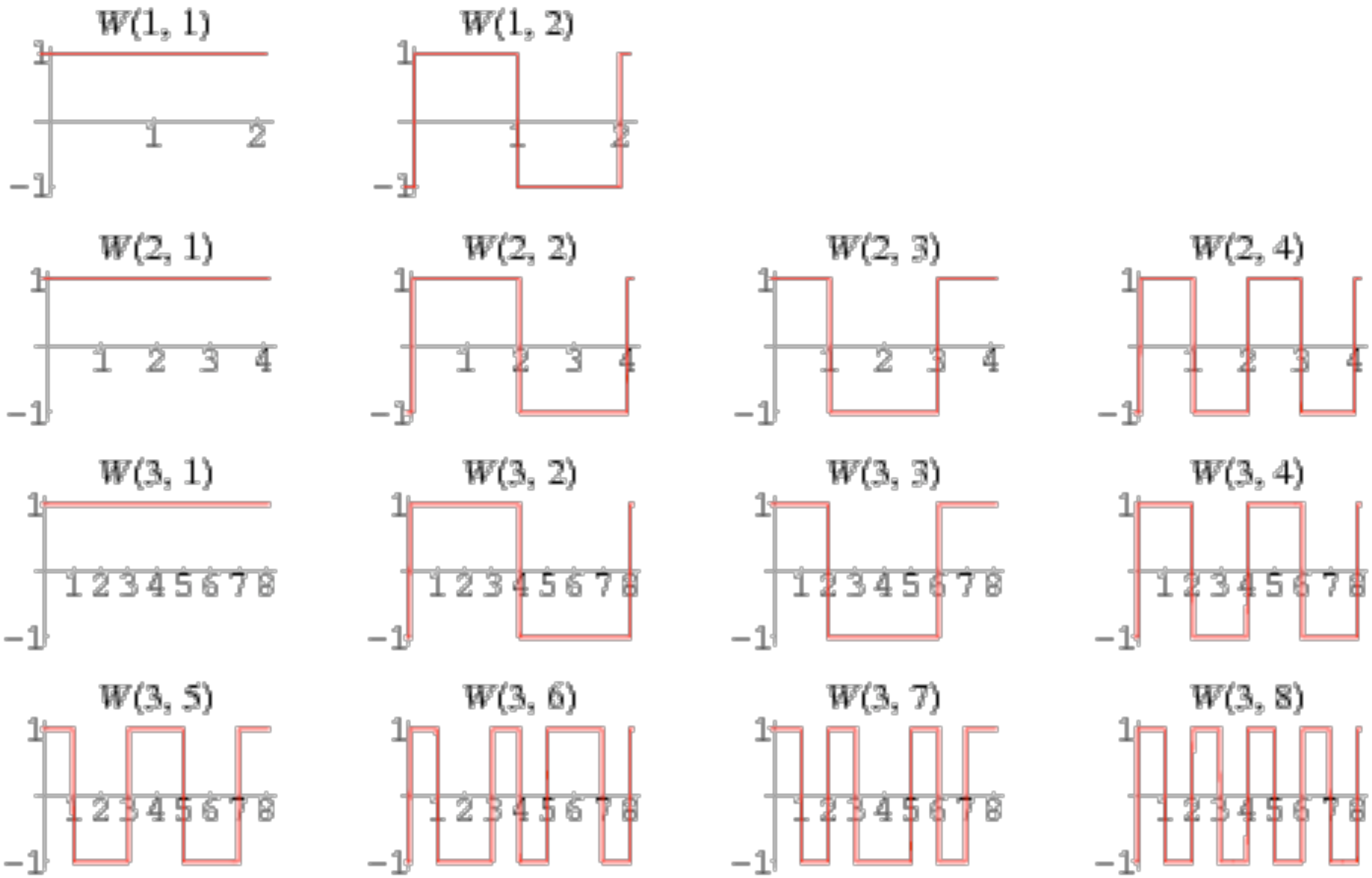
Walsh-Hadamard Transform

- We can then use the Hadamard matrix for computing an *M*-dimensional feature vector $\vec{c}$ as follows:

$$\vec{c} = H_M \vec{f}$$

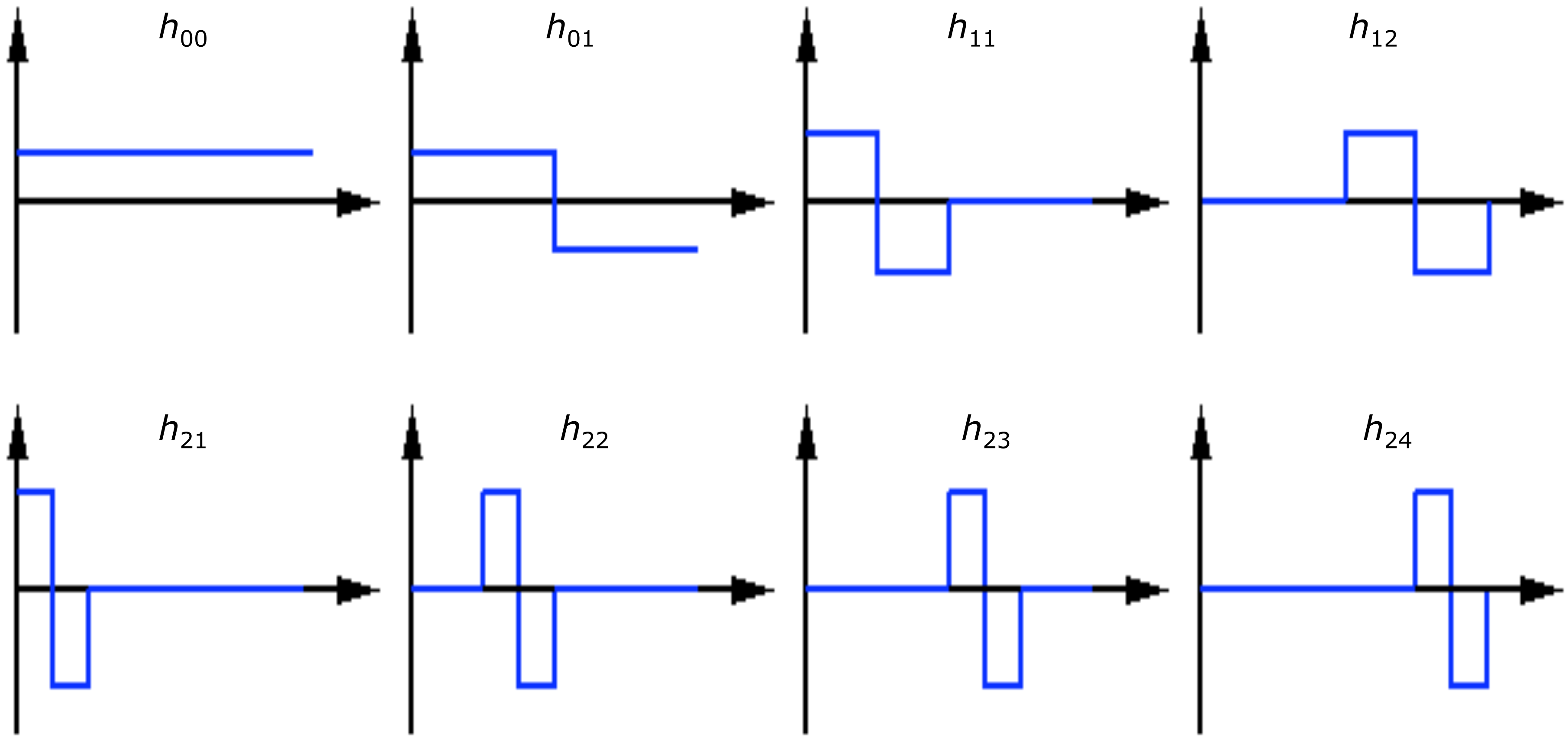
- This is known the Walsh-Hadamard Transform (WHT).
- Attributes of the WHT:
  - It only involves additions and subtractions of real numbers.
  - The results are real numbers.
  - There exists a divide-and-conquer implementation which decreases the M^2 additions and subtractions to MlogM additions/subtractions.

Walsh Function Plots - revisited



Plots courtesy of Wolfram Mathworld http://mathworld.wolfram.com/WalshFunction.html

Haar Functions



Plots courtesy of Ruye Wang http://fourier.eng.hmc.edu/e161/lectures/Haar/index.html

Definition of Haar Functions

- The collection of Haar functions is somewhat more intuitively constructed.
- The Haar functions $h_k(x) = h_{pq}(x)$ are recursively defined.
- For $k>0$ the Haar function always contains a single square wave where p specifies the magnitude and width of the shape (the narrower the wave, the taller it is) and q specifies its position
- The order of the function, k, is uniquely decomposed into 2 integers p and q.

Definition of Haar Functions - continued

- *p* and *q* are uniquely determined so that:
- ✓ 2^p is the largest power contained in k and
- ✓ *q* is the remainder
- The Haar functions are defined for the interval $0 \le x \le 1$ and for the following indices:

$$k = 0,1,2,\cdots,M-1 \quad \text{where} \quad M = 2^n$$

$$k = 2^p + q - 1$$

$$0 \le p \le n-1$$

$$q = \begin{cases} 0,1 & \text{for } p = 0 \\ 1 < q < 2^p & \text{for } p \ne 0 \end{cases}$$

Definition of Haar Functions - continued

- The Haar functions are then recursively defined as:

$$h_{00}(x) = \frac{1}{\sqrt{M}}$$

$$h_{pq}(x) = \frac{1}{\sqrt{M}} \begin{cases} 2^{p/2} & \text{for } \frac{q-1}{2^p} \leq x < \frac{q-0.5}{2^p} \\ -2^{p/2} & \text{for } \frac{q-0.5}{2^p} \leq x < \frac{q}{2^p} \\ 0 & \text{for other values of } x \text{ in } [0,1] \end{cases}$$

- The parameter M controls how fine our decomposition will be (how many basis functions we will use).

Example Haar Transformation Matrix

- For M=4, we get the Haar transformation matrix

$$Har_4 = \begin{bmatrix} h_{00} \\ h_{01} \\ h_{11} \\ h_{12} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix}$$

- Note that $Har_4^{-1} = Har_4^T$ which means that the matrix is orthogonal.
- This implies that the Haar basis functions are orthogonal to each other.

Another Haar Transformation Matrix

- For M=8, we get the Haar transformation matrix

$$Har_8 = \begin{bmatrix} h_{00} \\ h_{01} \\ h_{11} \\ h_{12} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{24} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{bmatrix}$$

- To create an M-dimensional feature vector $\vec{c}$ based on the Haar basis function, we compute:

$$\vec{c} = Har_M \vec{f}$$