Generalized Walsh bases and some applications to digital signal processing

Smooth and Non Smooth Harmonic Analysis Seminar
Iowa State University, June 3, 2018

Gabriel Picioroaga, University of South Dakota

Classic Walsh Basis

Rademacher functions:

$$\phi_n(x) = (-1)^k, \quad \frac{k}{2^n} \leq x < \frac{k+1}{2^n}, \quad k = 0, 1, \ldots, 2^n - 1$$

Orthogonal but not complete in $L^2[0,1]$.

Classic Walsh Basis

Rademacher functions:

$$\phi_n(x) = (-1)^k, \quad \frac{k}{2^n} \leq x < \frac{k+1}{2^n}, \quad k = 0, 1, \ldots, 2^n - 1$$

Orthogonal but not complete in $L^2[0,1]$*.*

Walsh functions: all possible finite products of Rademacher's.

Paley- ordered Walsh functions :

$$w_n = \phi_{n_1}\phi_{n_2} \cdot \cdots \cdot \phi_{n_p}, \quad n = 2^{n_1} + 2^{n_2} + \ldots 2^{n_p}$$

$(w_n)_n$ *is orthogonal and complete in* $L^2[0,1]$

Generalized Walsh Bases

Walsh, Fourier and more bases are captured by the following construction:

Unitary $N \times N$ matrix

$$A = \begin{pmatrix} \frac{1}{\sqrt{N}} & \frac{1}{\sqrt{N}} & \cdots & \frac{1}{\sqrt{N}} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,N-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N-1,0} & a_{N-1,1} & \cdots & a_{N-1,N-1} \end{pmatrix}$$

N filters m_i either polynomials of degree N with coefficients a_{ij} or

$$m_i(x) = \sqrt{N} \sum_{j=0}^{N-1} a_{ij} \chi_{[j/N,(j+1)/N)}(x), \quad x \in [0,1], i = 0, \ldots, N-1$$

Generalized Walsh Bases

Rademacher's played by

$$m_i(r^j(x)), \quad \textit{where } r(x) = Nx \textit{ mod } 1$$

Theorem

(Dutkay, P, Song)
The set $(w_n)_{n\geq 0}$ *forms an orthonormal basis in* $L^2[0,1]$.

$$w_n(x) = m_{i_0}(x)m_{i_1}(rx)m_{i_2}(r^2x)...m_{i_{p-1}}(r^{p-1}x)$$

where $n = i_0 + i_1 \cdot N + i_2N^2 + ... + i_{p-1}N^{p-1}$ *is the base N expansion of n*

Some properties

Theorem

(Dutkay, Harding, P)

- $\sum_{n=0}^{\infty} \langle f \, , \, w_n \rangle \, w_n$ *converges to* f *in* L^2

- *'Gap' sums* $\sum_{k=0}^{N^p} \langle f \, , \, w_k \rangle \, w_k$ *converge pointwise to* $f \in L^1$*, uniformly if* f *continuous*

- *if* f *piecewise constant the assignment*

$$A \longrightarrow (\langle f \, , \, w_n \rangle)_{n \geq 0}$$

is continuous.

- *'Filter' operators* $S_i : L^2 \to L^2$ *satisfy*

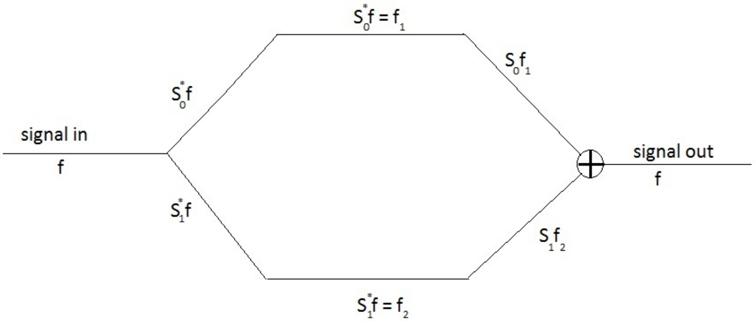
$$S_i f = m_i f(r)$$

$$S_i^* S_j = \delta_{i,j} I, \qquad \sum_{i=0}^{N-1} S_i S_i^* = I$$

Cuntz Relations

Definition

$$\text{Algebra } \mathcal{O}_N : \qquad S_i^* S_j = \delta_{i,j} I, \qquad \sum_{i=0}^{N-1} S_i S_i^* = I$$



Restricting to finite dimensions

Generalized Walsh transform $T : L^2 \to l^2$:

$$Tf = \langle f \, , \, w_n \rangle_{n \geq 0}$$

Discrete (and normalized) transform:
Signal space $\mathbf{R}^{N^p}$ *or* $\mathbf{C}^{N^p}$

$$Tf := A^{\otimes p} f$$

Expensive to implement $A \otimes B \otimes \ldots$ *but there is a 'fast' way for*

$$\overline{A}^{\otimes p} = \prod_{k=0}^{p-1} I_{N^k} \otimes \overline{A} \otimes I_{N^{p-1-k}}$$

Similar to FFT, fast Walsh: $O(M \log M)$ *operations,* $M = N^p$.

The Variance Criterion

$T : \mathbb{R}^n \to \mathbb{R}^n, \quad T^*T = I$

1. *Fix* $M < n$
2. $X = (x_1, x_2, \ldots, x_n),\ Y = TX = (y_1, y_2, \ldots, y_n)^t$
3. $cov(Y) := YY^t$
   Pick highest M *variances* $\sigma_{i_1,i_1}, \sigma_{i_2,i_2}, \ldots, \sigma_{i_M,i_M}$.
   Compression: $\tilde{Y} := (0, .., 0, y_{i_1}, 0, .., y_{i_2}, \ldots, 0, y_{i_M}, 0, ..0)$.
4. $\tilde{X} := T^{-1}(\tilde{Y})$ *approximates* X

Compare the (normalized) variance distributions:

$$k \to \frac{\sigma_{i_k,i_k}}{trace}$$

Compression Experiments

Two generalized Walsh based on 3×3 matrices

$$A := \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix}, \qquad B := \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -0.2 & -0.58 & 0.78 \\ -0.79 & 0.57 & 0.22 \end{pmatrix} \tag{1}$$

Compression Experiments

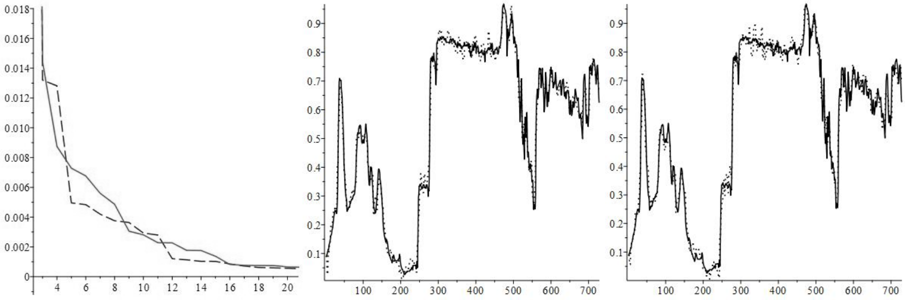


Figure: Variance distribution (left) for T_1-solid curve packs more energy than T_2 -dash. Approximation of a signal with T_1 (middle) and T_2(right), 90% components removed. Errors 0.84 and 0.95

Compression Experiments

Classic Walsh

$$A := \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \tag{2}$$

and a generalized Walsh based on 4×4*:*

$$B := \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix} \tag{3}$$

Compression Experiments

High variation signal, 2^8 length

$$X(i) := \begin{cases} \frac{i}{3i+1}, & \text{if } i|9 \\ \frac{i}{i+1}, & \text{otherwise} \end{cases}$$

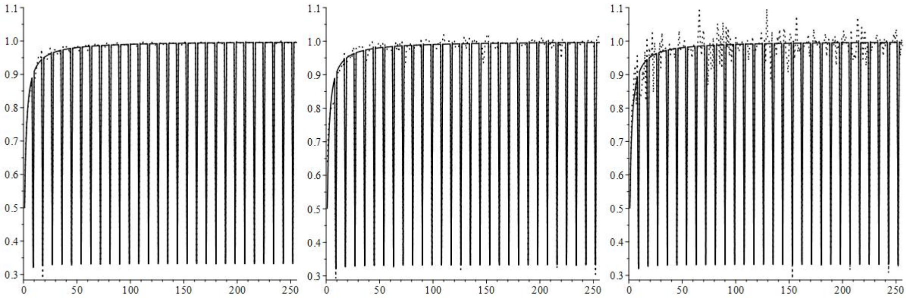


Figure: Approximation with DCT(left), classic Walsh(middle) and generalized 4 by 4 Walsh (right) with 55% components removed

Compression Experiments

Same signal, extended to 3^6 components. Apply now previous 3×3 generalized Walsh.

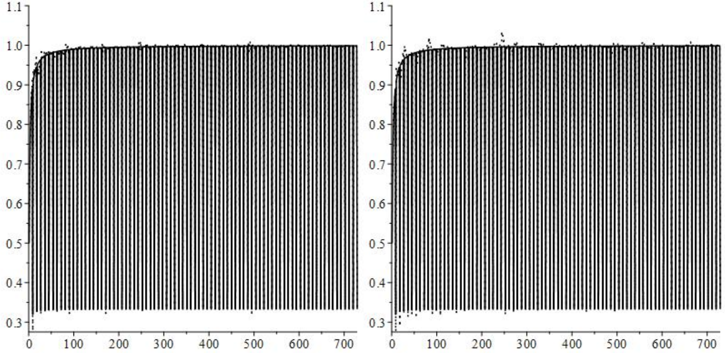


Figure: Same signal extended to dimension 3^6 with previous two generalized Walsh. 90% components removed. Recovery errors 0.02 and 0.08

Compression Experiments

Figure: original Lena pic, 256 by 256

Compression Experiments

Figure: classic Walsh Lena , 60% components removed, 256 by 256

Compression Experiments

Figure: generalized Walsh Lena, 60% components removed, 243 by 243

Compression Experiments

Figure: classic Walsh Lena , 88% components removed, 256 by 256

Compression Experiments

Figure: generalized Walsh Lena, 88% components removed, 243 by 243

Compression Experiments

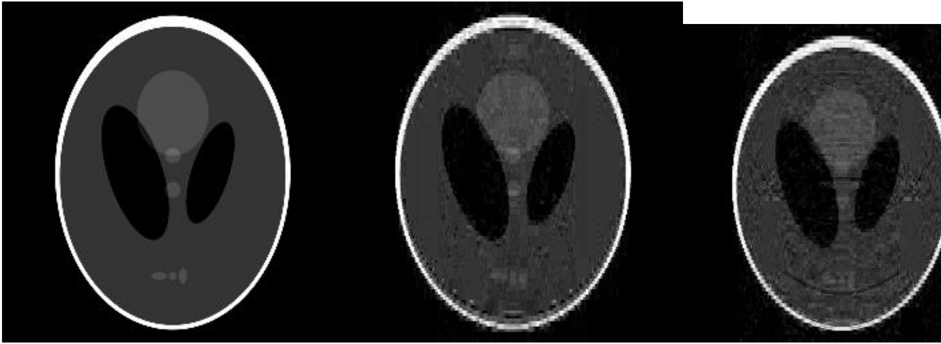


Figure: original phantom, classic Walsh, generalized Walsh, 88% compression

Compression Experiments

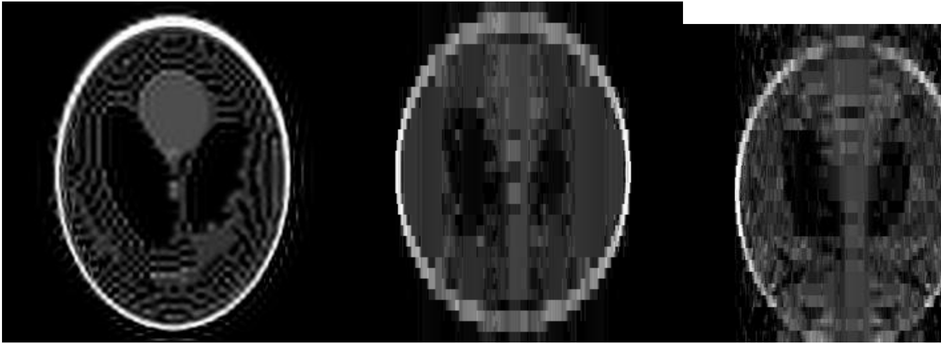
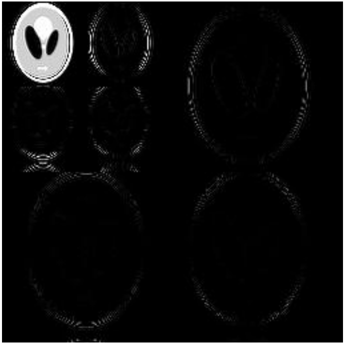


Figure: Wavelet, classic Walsh, generalized Walsh, 98% compression

Compression Experiments

The Discrete Wavelet transforms removes pixel values close to zero from this pic:



Uncertainty Principle

Uncertainty Principle for discrete signals (Donohoe, Stark '89)
$f \in \mathbf{C}^m$, $f \neq 0$, $\hat{f} \in \mathbf{C}^m$ *its discrete Fourier transform. Then*

$$|supp(f)| \cdot |supp(\hat{f})| \geq m \tag{4}$$

Theorem

(Dutkay, P, Silvestrov)
Let T *the generalized Walsh associated to* A, $f \in \mathbf{C}^{N^p}$, $f \neq 0$. *If* $\alpha > 0$ *satisfies* $max|a_{ij}| \leq (\frac{1}{\sqrt{N}})^{\alpha}$ *then :*
i)

$$|supp(f)| \cdot |supp(Tf)| \geq N^{p\alpha} \tag{5}$$

ii) $\alpha \leq 1$, *and* $= 1$ *if and only if* $\sqrt{N} \cdot A$ *is Hadamard matrix.*

Uncertainty Principle

What's the big deal?

Uncertainty Principle

What's the big deal?

We want to recover $f \in \mathbf{C}^{N^p}$ *and know :*
i) $N_f = |supp(f)|$.
ii) a subset $B \subset supp(Tf)$ *of observed 'frequencies' with which form the signal*

$$\tilde{f}(k) := \begin{cases} Tf(k), & \text{if } k \in B \\ 0, & \text{otherwise} \end{cases}$$

iii) the number of unobserved 'frequencies' $N_w = |B^c|$ *satisfies*

$$2N_f N_w < N^{p\alpha} \tag{6}$$

Then f can be uniquely reconstructed.

Uncertainty Principle

What's the big deal?

We want to recover $f \in \mathbf{C}^{N^p}$ *and know :*
i) $N_f = |supp(f)|$.
ii) a subset $B \subset supp(Tf)$ *of observed 'frequencies' with which form the signal*

$$\tilde{f}(k) := \begin{cases} Tf(k), & \text{if } k \in B \\ 0, & \text{otherwise} \end{cases}$$

iii) the number of unobserved 'frequencies' $N_w = |B^c|$ *satisfies*

$$2N_f N_w < N^{p\alpha} \tag{6}$$

Then f can be uniquely reconstructed.

Why does it work? Note also we don't know the missed frequency locations.

Uncertainty Principle

What's the big deal?

We want to recover $f \in \mathbf{C}^{N^p}$ *and know :*
i) $N_f = |supp(f)|$ *.*
ii) a subset $B \subset supp(Tf)$ *of observed 'frequencies' with which form the signal*

$$\tilde{f}(k) := \begin{cases} Tf(k), & \text{if } k \in B \\ 0, & \text{otherwise} \end{cases}$$

iii) the number of unobserved 'frequencies' $N_w = |B^c|$ *satisfies*

$$2N_f N_w < N^{p\alpha} \qquad (6)$$

Then f *can be uniquely reconstructed.*

Why does it work? Note also we don't know the missed frequency locations.

If $\tilde{f} = \tilde{g}$, $N_f = N_g$, *then signal* $h = f - g \neq 0$ *would violate the uncertainty principle.*

THANK YOU!