The Generalized Haar-Walsh Transform

Jeff Irion with Dr. Naoki Saito
Department of Mathematics, University of California, Davis

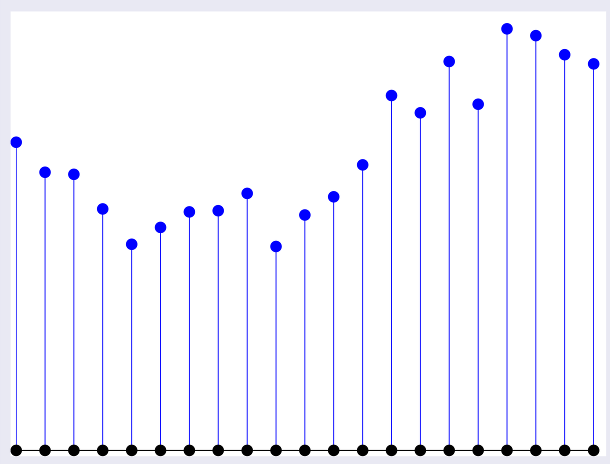
MMDS Workshop, Berkeley, CA, June 17-20, 2014

Aims & Objectives

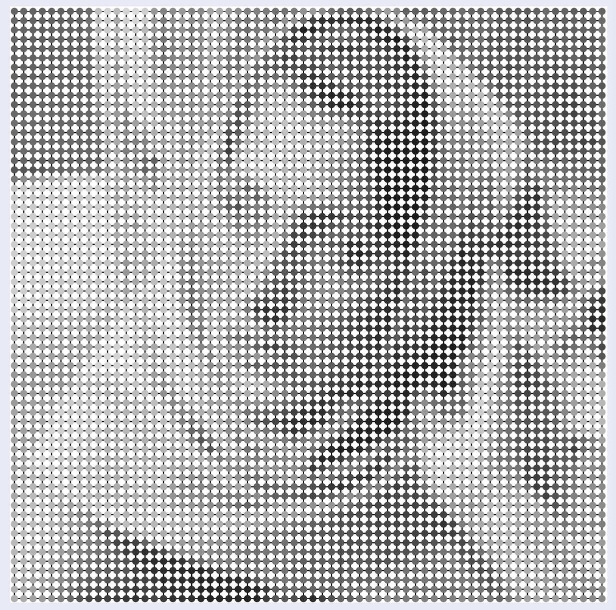
1. Develop an overcomplete multiscale transform for signals on graphs
2. Develop a corresponding best-basis search algorithm
3. Investigate usefulness for approximation, denoising, classification, function estimation, etc.

Motivation

The regular signals analyzed in classical signal processing can be viewed as signals on graphs with very simple structures. For example:

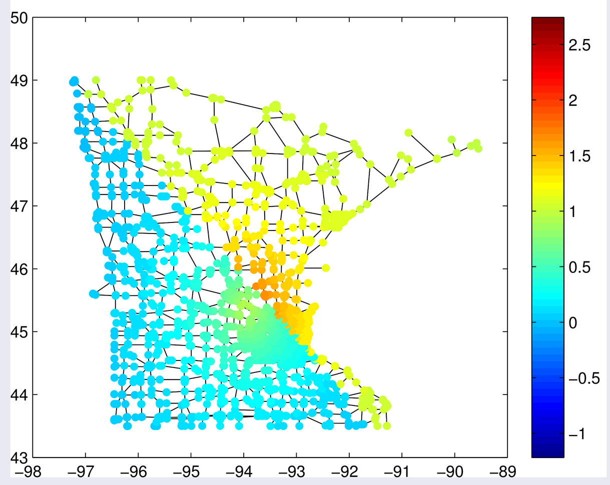


1a. A portion of a 1D signal

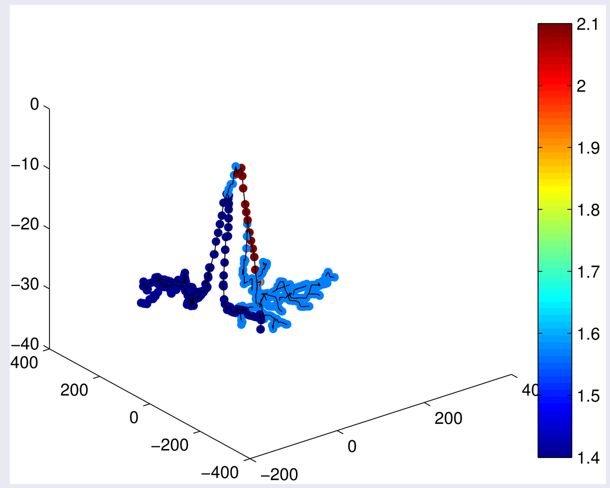


1b. Barbara

A current aim in signal processing is to extend the tools developed for regular signals to signals on graphs; e.g.,



2a. A mutilated Gaussian signal on the MN road network



2b. Thickness data on a dendritic tree

Problem Setup

- a graph G with vertices $V = \{v_1, \dots, v_N\}$
- an edge weight matrix W
- a data vector $\mathbf{f} \in \mathbb{R}^N$, where f_i corresponds to v_i

Notation & Recursive Partitioning

A preliminary step for the GHWT is to use recursive bisection to generate a recursive partitioning of the graph (see [1] for an example). Our notation is:

- $j \in [0, j_{\max}]$ denotes **levels**, where $j = 0$ is the coarsest level (the only region is the entire graph) and $j = j_{\max}$ is the finest level (each region is a single node)
  - $k \in [0, K^j)$ indexes the **regions** on level j

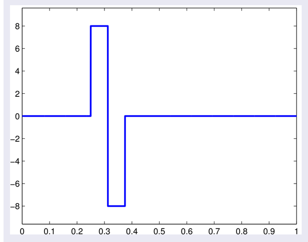
In addition, we let

- N_k^j denote the number of nodes in region k on level j
- If $N_k^j > 1$ then we let k' and $k'+1$ denote the indices of the children regions on level $j+1$

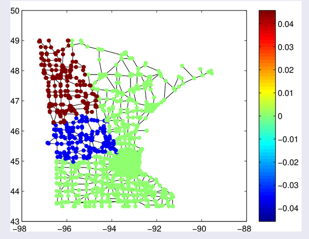
Basis Vector & Coefficient Notation

GHWT basis vectors and coefficients are written as $\psi_{k,\ell}^j$ and $d_{k,\ell}^j$, respectively, where j and k correspond to level and region and ℓ is the **tag**.

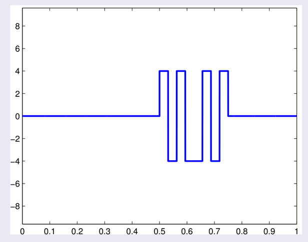
- $\ell = 0 \Rightarrow$ **scaling coefficient/basis vector**
- $\ell = 1 \Rightarrow$ **Haar-like coefficient/basis vector**
- $\ell \geq 2 \Rightarrow$ **Walsh-like coefficient/basis vector**



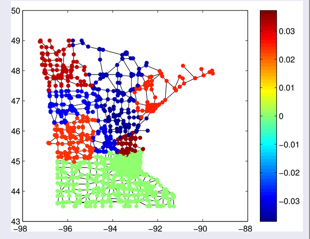
3a. Haar function



3b. Haar-like vector $\psi_{0,1}^2$



3c. Haar-Walsh wavelet packet



3d. Walsh-like vector $\psi_{0,5}^1$

Generalized Haar-Walsh Transform (GHWT)

Step 1: Start at level $j = j_{\max}$. As each region is a single node, define an orthonormal basis consisting of kronecker deltas. Obtain expansion coefficients $d_{k,0}^{j_{\max}}$ by reordering the signal.

Step 2: **For** $j = j_{\max}, \dots, 1$: use the coefficients on level j to compute coefficients on level $j-1$ as follows.
For $k = 0, \dots, K^{j-1} - 1$:

Step 2a: Compute the **scaling coefficient**.

Case 1: **If** $N_k^{j-1} = 1$, **then** set

$$d_{k,0}^{j-1} := d_{k',0}^{j}. \qquad (1)$$

Case 2: **If** $N_k^{j-1} > 1$, **then** compute

$$d_{k,0}^{j-1} := \frac{\sqrt{N_{k'}^j}\, d_{k',0}^j + \sqrt{N_{k'+1}^j}\, d_{k'+1,0}^j}{\sqrt{N_k^{j-1}}}. \qquad (2)$$

Step 2b: **If** $N_k^{j-1} \geq 2$, **then** compute the **Haar-like coefficient** as

$$d_{k,0}^{j-1} := \frac{\sqrt{N_{k'+1}^j}\, d_{k',0}^j - \sqrt{N_{k'}^j}\, d_{k'+1,0}^j}{\sqrt{N_k^{j-1}}}. \qquad (3)$$

Step 2c: **If** $N_k^{j-1} \geq 3$, **then** compute the **Walsh-like coefficients**. **For** $\ell = 1, \dots, 2^{j_{\max}-j} - 1$:

Case 1: **If** neither subregion has a coefficient with tag ℓ, **then** do nothing.

Case 2: **If** (without loss of generality) only subregion k' has a coefficient with tag ℓ, **then** set

$$d_{k,2\ell}^{j-1} := d_{k',\ell}^{j}. \qquad (4)$$

Case 3: **If** both subregions have coefficients with tag ℓ, **then** compute

$$d_{k,2\ell}^{j-1} := \left(d_{k',\ell}^j + d_{k'+1,\ell}^j\right) \Big/ \sqrt{2} \qquad (5)$$

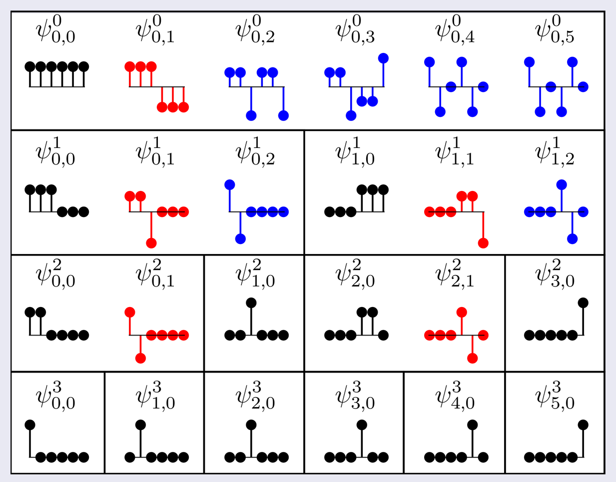
$$d_{k,2\ell+1}^{j-1} := \left(d_{k',\ell}^j - d_{k'+1,\ell}^j\right) \Big/ \sqrt{2}. \qquad (6)$$

GHWT Output: TWO Dictionaries

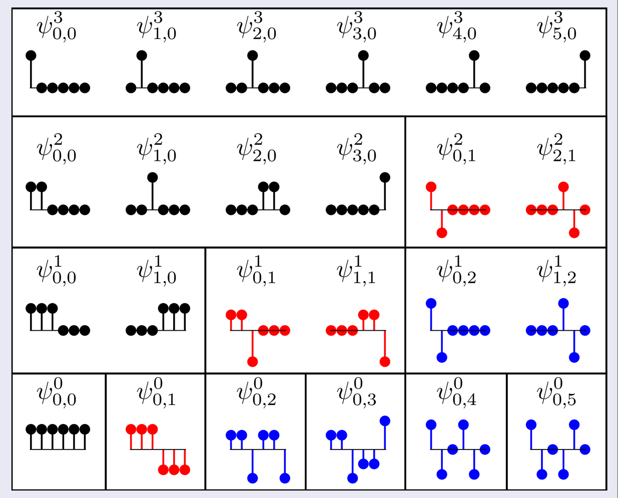
The GHWT expansion coefficients and basis vectors can be grouped in two ways:

1. **coarse-to-fine dictionary** – grouping is by region; i.e., the two child regions are beneath their parent region
2. **fine-to-coarse dictionary** – grouping is by tag, exploiting the fact that the vectors/coefficients with tag ℓ are used to generate those with tags 2ℓ and $2\ell+1$ on the next coarser level

Using the path graph of length 6 as an example, we have:



4a. Coarse-to-fine dictionary



4b. Fine-to-coarse dictionary

Considering both dictionaries affords more options for choosing a basis.

Best-Basis Algorithm

We generalized the best-basis algorithm [3] to our transform.

Step 1: Specify a cost functional $\mathcal{J}$ (e.g., ℓ^p quasi-norm with $0 < p < 1$)

Step 2: For the coarse-to-fine dictionary, intialize the best basis as the bottom level. Proceed upwards, using $\mathcal{J}$ to compare the cost of each block of coefficients to the blocks beneath it and updating the best basis as necessary. The result is the coarse-to-fine best basis.

Step 3: Repeat for the fine-to-coarse dictionary.

Step 4: Use $\mathcal{J}$ to compare the coarse-to-fine and fine-to-coarse best bases. The result is the overall best basis.

Observations

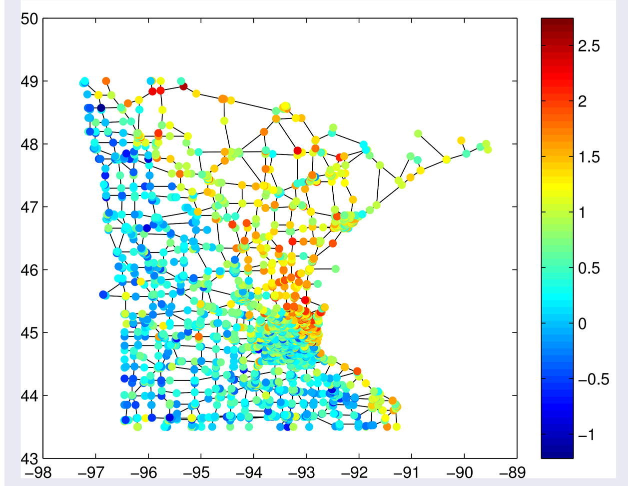
- When performed on an unweighted dyadic path graph (partitioned dyadically), the GHWT corresponds exactly to the Haar-Walsh wavelet packet transform
- The generalized Haar basis is a choosable basis from the fine-to-coarse dictionary
- Given a recursive partitioning with $O(\log N)$ levels, the computational cost of the GHWT is $O(N \log N)$
  - Recursive partitioning via Fiedler vectors costs from $O(N \log N)$ to $O(N^2)$

	N	$j_{\max}$	GHWT Run Time
MN Road Network	2,636	14	0.11 s
Facebook Dataset	4,039	26	0.62 s
Brain Mesh Dataset	127,083	20	4.29 s

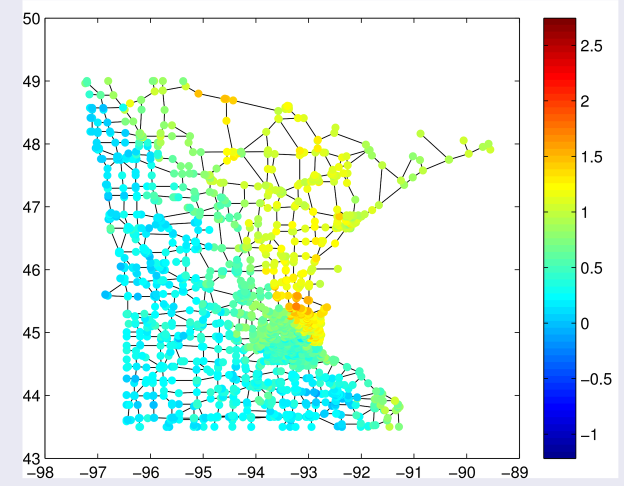
(Experiments performed on a personal laptop.)

Denoising Experiment

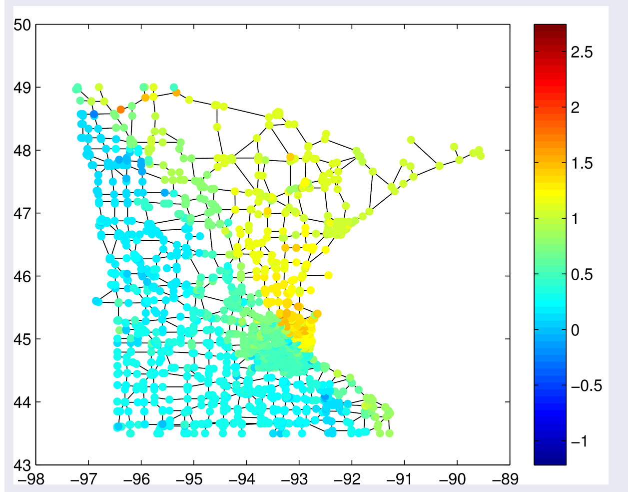
1. Add Gaussian noise to mutilated Gaussian on the MN road network (Fig. 2a) to yield a signal with SNR 5.00 dB (Fig. 5a)
2. Perform the GHWT and consider the fine-to-coarse best-basis ($\mathcal{J}$ is the ℓ^p quasi-norm with $p = 0.1$), Haar basis, and GHWT coarse-to-fine level $j = 6$ basis
3. Denoise via soft-thresholding
   - Haar & best basis: threshold chosen manually using original signal
   - Level 6: threshold chosen using elbow selection algorithm (original signal not used)



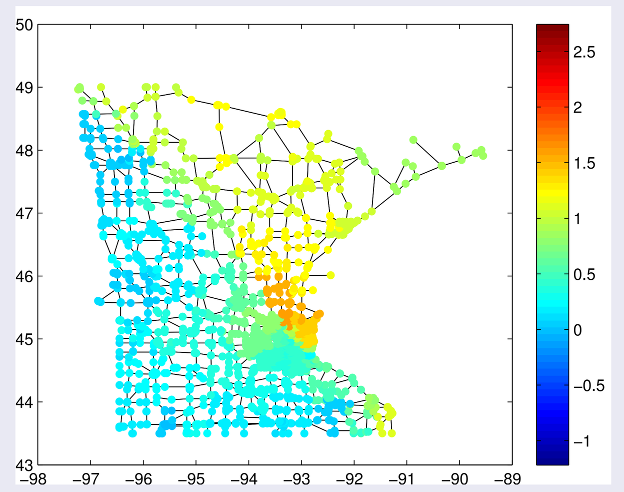
5a. Noisy signal, SNR = 5.00 dB



5b. GHWT fine-to-coarse best basis, SNR = 11.59 dB



5c. Haar basis, SNR = 12.83 dB



5d. GHWT coarse-to-fine level $j = 6$, SNR = 13.46 dB

- GHWT level 6 has the best localization for this denoising task; Haar is too localized and best basis is too global
- Most retained coefficients for GHWT level 6 are scaling coefficients, not Haar-like (as in the Haar basis) or higher frequency Walsh-like (as in the best basis) coefficients

Future Work

- Investigate usefulness for classification and function estimation
- Allow for partitions to have "soft" boundaries
- Investigate different graph partitioning methods

Acknowledgements

Funding by

- NSF VIGRE DMS-0636297
- NDSEG Fellowship, 32 CFR 168a
- ONR grant N00014-12-1-0177

References

1. J. Irion, N. Saito: "The Generalized Haar-Walsh Transform," *Proceedings of 2014 IEEE Workshop on Statistical Signal Processing*, pp. 488–491, 2014.
2. J. Irion, N. Saito: "Hierarchical Graph Laplacian Eigen Transforms," *Japan SIAM Letters*, vol. 6, pp. 21–24, 2014.
3. R. Coifman, M. Wickerhauser: "Entropy-based algorithms for best basis selection," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 713–718, 1992.