Graph Signal Processing (GSP)

(or how I started seeing graphs everywhere)

Antonio Ortega

Signal and Image Processing Institute
Department of Electrical Engineering
University of Southern California
Los Angeles, California

Apr 14, 2017

Outline

Graphs Everywhere

GSP basics

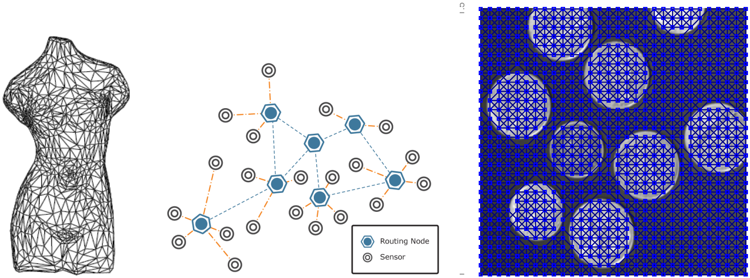
Contributions

Discussion

Graphs everywhere ...

Graphs provide a flexible model to represent many datasets:

- Examples in Euclidean domains

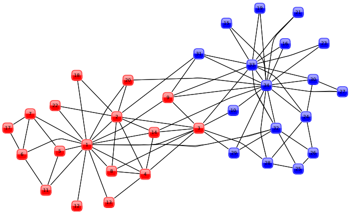


(a) (b) (c)

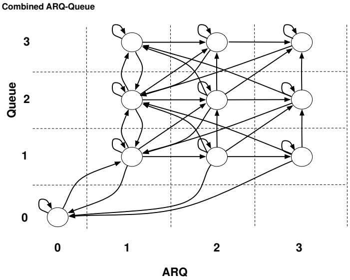
(a) Computer graphics[1] (b) Wireless sensor networks [2] (c) image - graphs

… and then some

- Examples in non-Euclidean settings



(a)



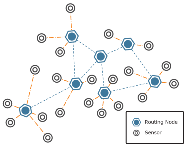
(b)

(a) Social Networks [3], (b) Finite State Machines(FSM)

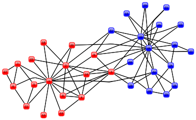
Graph Signal Processing

- Given a graph (fixed or learned from data)
- and given signals on the graph (set of scalars associated to vertices)
- define frequency, sampling, transforms, etc
- in order to solve problems such as compression, denoising, interpolation, etc
- Overview papers:
  [Shuman, Narang, Frossard, Ortega, Vandergheysnt, SPM'2013]
  [Sandryhaila and Moura 2013]

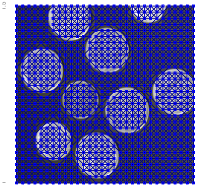
Examples



- Sensor network
  - Relative positions of sensors (kNN), temperature
  - Does temperature vary smoothly?



- Social network
  - Friendship relationship, age
  - Are friends of similar age?

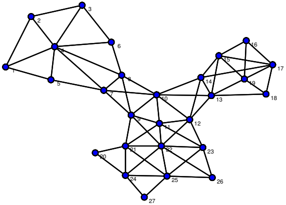


- Images
  - Pixel positions and similarity, pixel values
  - Discontinuities and smoothness

Outline

Graphs 101

- Graph : vertices (nodes) connected via some links (edges)
- Graph Signal: set of scalar/vector values defined on the vertices.



Graph-signal

Graph $G = (\mathcal{V}, E, w)$

Vertex Set $\mathcal{V} = \{v_1, v_2, ...\}$

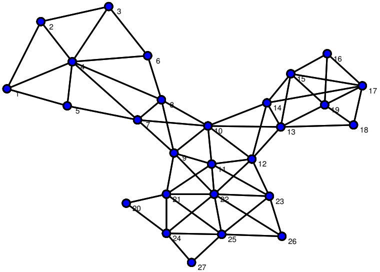
Edge Set $\mathbf{E} = \{(v_1, v_2), (v_1, v_3), ...\}$

Weighted edges w, sets of weights a_{ij}

Graph Signal $\mathbf{x} = \{x_1, x_2, ...\}$

Neighborhood, h-hop
$\mathcal{N}_h(i) = \{j \in \mathcal{V} : hop\_dist(i, j) \leq h\}$

Multiple algebraic representations

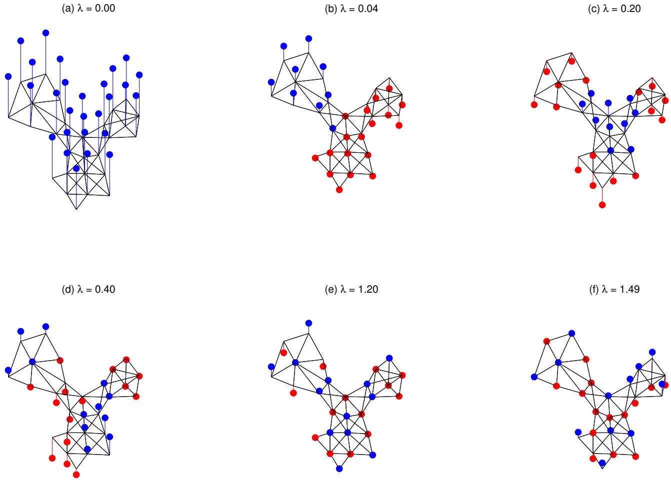


- Graph $G = (\mathcal{V}, E, w)$.
- Adjacency $\mathbf{A}$, a_{ij}, a_{ji} = weights of links between i and j (could be different if graph is directed.)
- Degree $\mathbf{D} = diag\{d_i\}$, in case of undirected graph.
- Various algebraic representations
  - Normalized adjacency $\frac{1}{|\lambda_{max}|}\mathbf{A}$
  - Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{A}$.
  - Symmetric normalized Laplacian $\mathcal{L} = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$
- Discussion:
  1. Undirected graphs easier to work with
  2. Some applications require directed graphs
  3. Graphs with self loops are useful

Graph Fourier Transform (GFT)

- Different results/insights for different choices of operator
- Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}'$
- Eigenvectors of $\mathbf{L}$: $\mathbf{U} = \{\mathbf{u}_k\}_{k=1:N}$
- Eigenvalues of $\mathbf{L}$: $diag\{\boldsymbol{\Lambda}\} = \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_N$
- **Eigen-pair system $\{(\lambda_k, \mathbf{u}_k)\}$ provides Fourier-like interpretation — Graph Fourier Transform (GFT)**

Eigenvectors of graph Laplacian

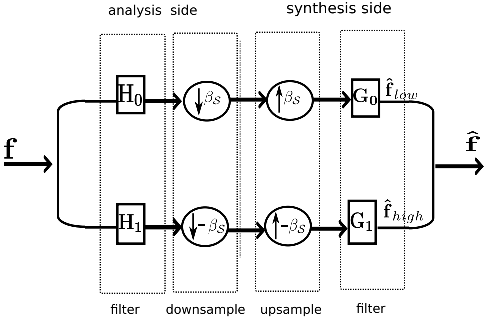


- Basic idea: increased variation **on the graph**, e.g., $\mathbf{f}^t\mathbf{L}\mathbf{f}$, as frequency increases

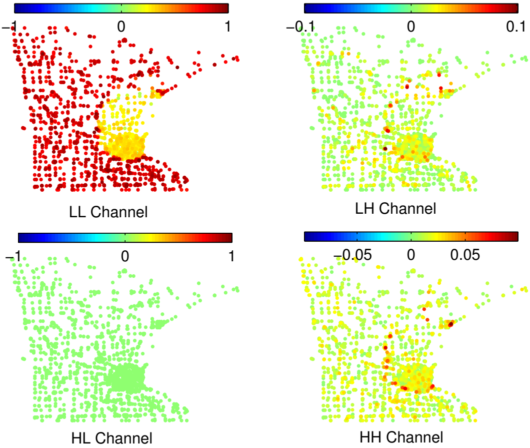
Outline

Graph Filterbank Designs

- Formulation of critically sampled graph filterbank design problem
- Design filters using spectral techniques [Hammond et al. 2009].
- Orthogonal (not compactly supported) [Narang and O. TSP'12]
- Bi-Orthogonal (compactly supported) [Narang and O. TSP'13]



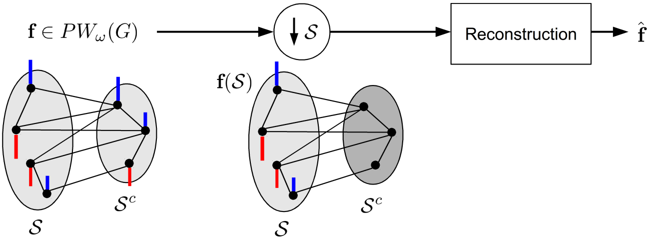
Example



Reconstructed graph-signals for each channel.

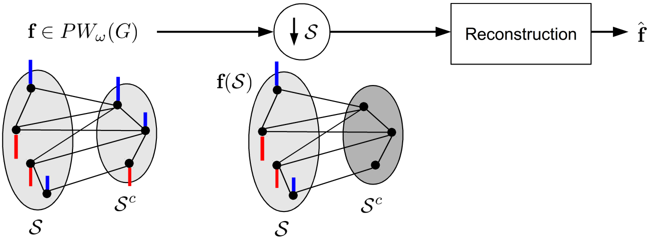
Graph Sampling?

- Measure a few nodes to estimate information throughout the graph
- Reconstruct signal in whole graph



Graph Sampling?

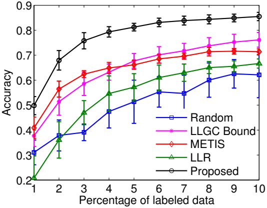
- Measure a few nodes to estimate information throughout the graph
- Reconstruct signal in whole graph



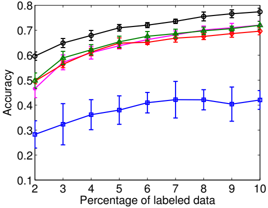
Questions:

- What properties enable recovery? Need to define frequency
- How to sample? No obvious regular sampling
- How to reconstruct? Filtering is needed

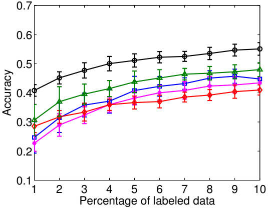
Results: Real Datasets



- USPS: handwritten digits
- $\mathbf{x}_i = 16 \times 16$ image
- number of classes $= 10$
- K-NN graph with $K = 10$
- $w_{ij} = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right)$



- ISOLET: spoken letters
- $\mathbf{x}_i \in \mathbb{R}^{617}$ speech features.
- number of classes $= 26$
- K-NN graph with $K = 10$
- $w_{ij} = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right)$



- Newsgroups: documents
- $\mathbf{x}_i \in \mathbb{R}^{3000}$ tf-idf of words
- number of classes $= 10$
- K-NN graph with $K = 10$
- $w_{ij} = \frac{\mathbf{x}_i^\top \mathbf{x}_j}{\|\mathbf{x}_i\| \|\mathbf{x}_j\|}$

Assumptions

- Given data matrix $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_N] \in \mathbb{R}^{n \times N}$.
- The k-th row of $\mathbf{X}$ (k-th variable) is attached to k-th vertex of the graph.
- Each $\mathbf{x}_i$ is a graph signal in an unknown graph.
  - Sensor network: each vertex is a sensor, signal is a measurement/time series
- Data model: attractive Gaussian Markov Random Field (a-GMRF) $\Leftrightarrow$ Gaussian with a Generalized Laplacian (GL) for precision matrix.
  - $\mathbf{Q} = \mathbf{P} + \mathbf{L}$ with $\mathbf{P}$ diagonal (self loop matrix) and $\mathbf{L}$ a combinatorial Laplacian.
  - $\mathbf{Q} = (q_{ij})$ and $q_{ij} \leq 0$ for all $i \neq j$
- Graph estimation: Max. Likelihood under aGMRF model.

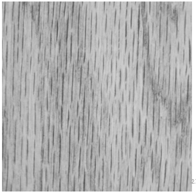
Use block coordinate descent to solve

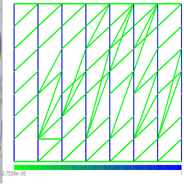
$$\min_{\mathbf{Q} \text{ is GL}} -\log\det(\mathbf{Q}) + \mathrm{tr}(\mathbf{QS}),$$

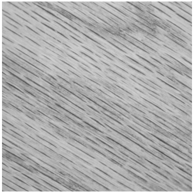
with $\mathbf{S} = \frac{1}{N}\mathbf{X}\mathbf{X}^T$.

Experiment: Texture graph

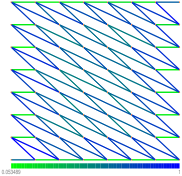
We consider *wood* textures from Brodatz dataset, with 0 and with 60 degree rotation. For each texture of Brodatz dataset, take 8×8 blocks, compute their covariance matrix $\mathbf{S}$ and solve GL estimation $rho = 0$.







wood060



Texture graphs using our GL estimation (only off diagonal elements of $\mathbf{Q}$). The graphs have $|E| = 130$ and $|E| = 117$ edges respectively

.

Outline

Discussion

- Research Question (Twitter Edition):

Discussion

- Research Question (Twitter Edition):
- GSP 4 CPS + IOT @USC-CCI ?

Discussion

- Research Question (Twitter Edition):
- GSP 4 CPS + IOT @USC-CCI ?

- Systems are large and irregular in space/time
  - Sampling and interpolation (sensors)

Discussion

- Research Question (Twitter Edition):
- GSP 4 CPS + IOT @USC-CCI ?

- Systems are large and irregular in space/time
  - Sampling and interpolation (sensors)
  - Variable topology communication networks

Discussion

- Research Question (Twitter Edition):
- GSP 4 CPS + IOT @USC-CCI ?

- Systems are large and irregular in space/time
  - Sampling and interpolation (sensors)
  - Variable topology communication networks
  - Data reduction to reduce complexity; multiresolution representations

Discussion

- Research Question (Twitter Edition):
- GSP 4 CPS + IOT @USC-CCI ?

- Systems are large and irregular in space/time
  - Sampling and interpolation (sensors)
  - Variable topology communication networks
  - Data reduction to reduce complexity; multiresolution representations
  - Control and optimization of large scale systems

References I

D. Shuman, S. K. Narang, P. Frossard, A. Ortega, P. Vandergheynst,
"Signal Processing on Graphs: Extending High-Dimensional Data Analysis to Networks and Other Irregular Data Domains"
Signal Processing Magazine, May 2013

A. Sandryhaila and J. Moura,
"Discrete Signal Processing on Graphs"
IEEE Transactions on Signal Processing, 2013