Ihara Coefficients: A Flexible Tool for Higher Order Learning

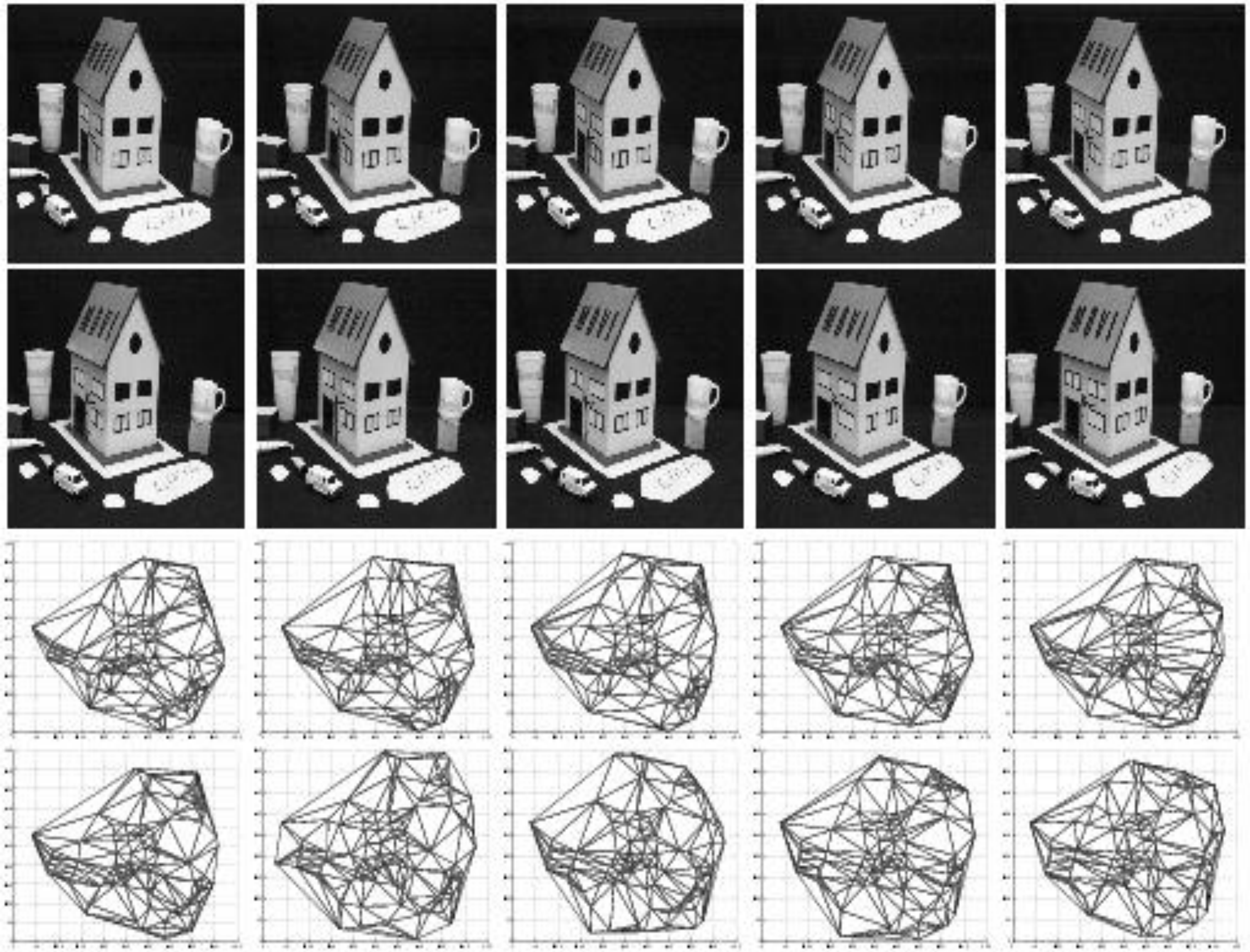
Peng Ren[1], Tatjana Aleksić[2], Richard C. Wilson[1], and Edwin R. Hancock[1]

[1] Department of Computer Science, The University of York,
York, YO10 5DD, UK
[2] University of Kragujevac, Faculty of Science, 34000 Kragujevac, Serbia

THE ROYAL SOCIETY

SIMBAD

Structural Variations



Problem studied

- How can we find efficient means of characterising graph structure which does not involve exhaustive search? Enumerate properties of graph structure without explicit search, e.g. count cycles, path length frequencies, etc..

- Can we analyse the structure of sets of graphs without solving the graph-matching problem? Inexact graph matching is computational bottleneck for most problems involving graphs.

- Past: Explored how diffusion processes based on heat equation can be used for this purpose..

Characterising graphs

- Topological: e.g. average degree, degree distribution, edge-density, diameter, cycle frequencies etc.

- Spectral: use eigenvalues of adjacency matrix or Laplacian.

- Algebraic: co-efficients of characteristic polynomial.

Prior work

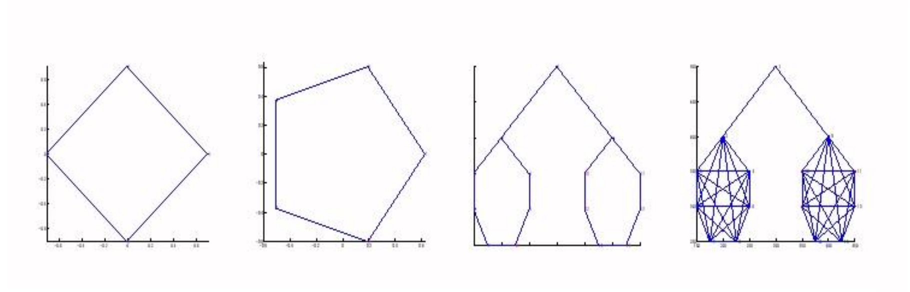
- Heat kernel trace provides a means of characterising graph structure (Xiao, Wilson, Hancock – PR 2010).

- Moments of heat-kernel trace are zeta functions (determined by product of non-zero Laplacian eigenvalues).

- Derivative of zeta-function at origin linked to number of spanning trees in graph.

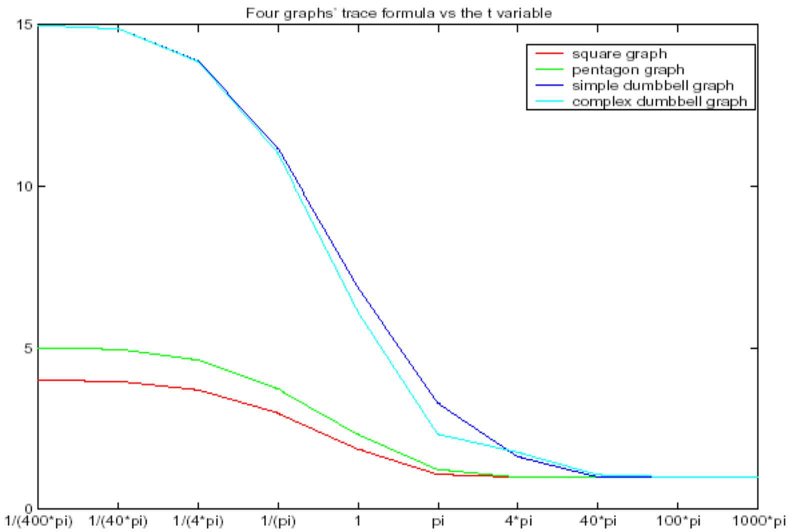
Heat Kernel Trace

$$Tr[h_t] = \sum_i \exp[-\lambda_i t]$$



Shape of heat-kernel distinguishes graphs…can we characterise its shape using moments

Trace



Time (t)->

Rosenberg Zeta function

- Definition of zeta function

$$\boxed{\varsigma(s)=\sum_{\lambda_k\neq 0}(\lambda_k)^{-s}}$$

Heat-kernel moments

- Mellin transform

$$\lambda_i^{-s} = \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} \exp[-\lambda_i t] dt$$

$$\Gamma(s) = \int_0^\infty t^{s-1} \exp[-t] dt$$

- Trace and number of connected components

$$Tr[h_t] = C + \sum_{\lambda_i \neq 0} \exp[-\lambda_i t]$$

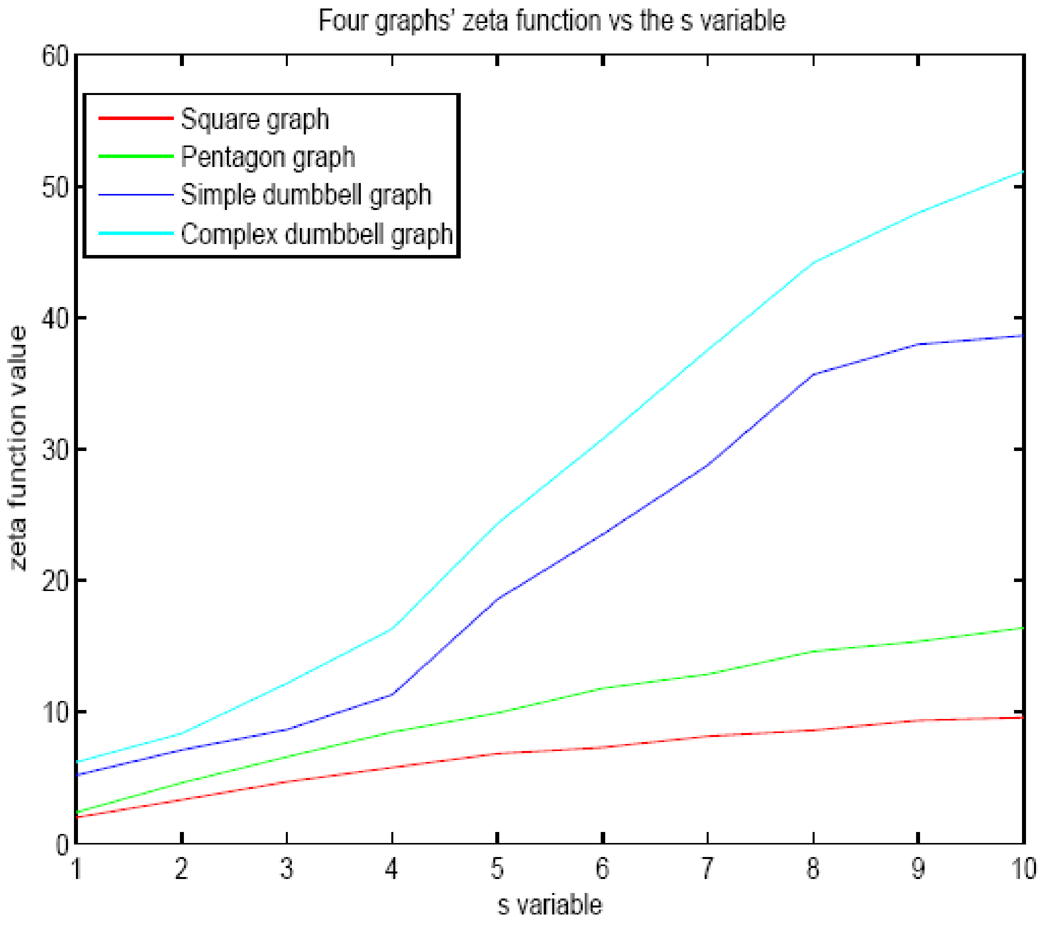
C is multiplicity of zero eigenvalue or number of connected components in graph.

- Zeta function

$$\varsigma(s) = \sum_{\lambda_i \neq 0} \lambda_i^{-s} = \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} \left[Tr[h_t] - C \right] dt$$

Zeta-function is related to moments of heat-kernel trace.

Zeta-function behavior



Deeper insights

- What more can zeta functions tell us about graph-structure?

- Can they be use to probe structure in a deeper way.

- How are they linked to graph spectra?

Zeta functions

- Used in number theory to characterise distribution of prime numbers.
- Can be extended to graphs by replacing notion of prime number with that of a prime cycle.

Aims

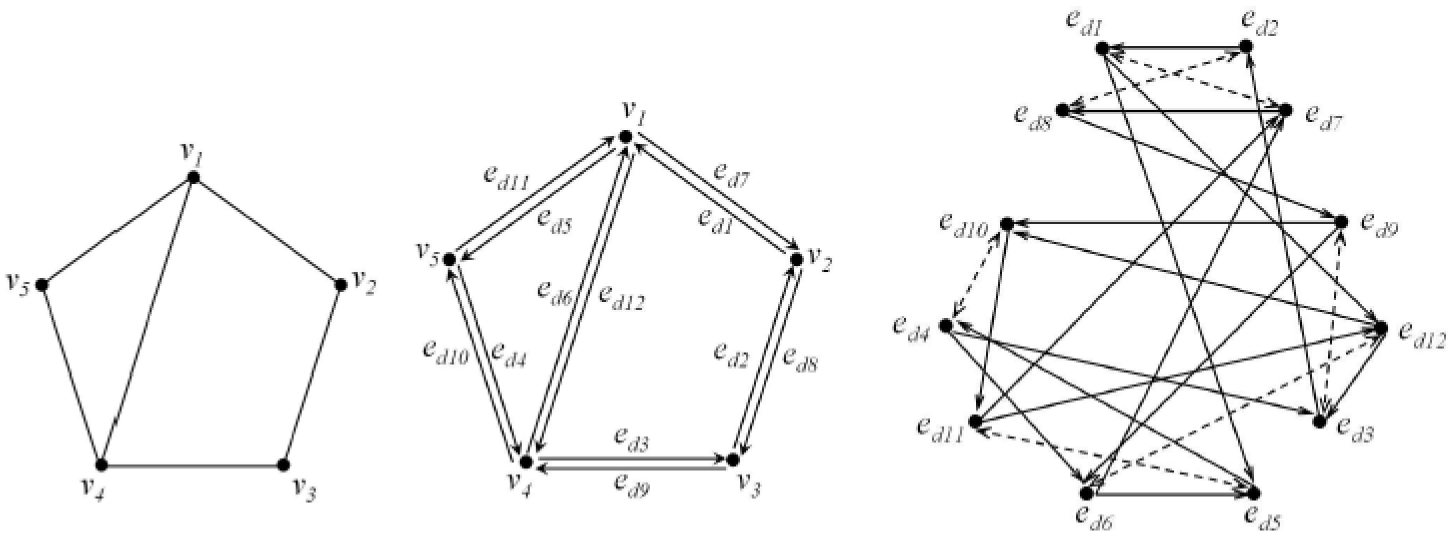
Ihara zeta function characterises graph in a manner that captures topological, spectral and algebraic properties.

Can easily be appled to graphs, weighted graphs and hypergraphs!

Ihara Zeta function

- Determined by distribution of prime cycles.
- Transform graph to oriented line graph (OLG) with edges as nodes and edges indicating incidence at a common vertex.
- Zeta function is reciprocal of characteristic polynomial for OLG adjacency matrix.
- Coefficients of polynomial determined by eigenvalues of OLG adjacency matrix.
- Coefficients linked to topological quantities such as cycle frequencies, number of spanning trees.

Oriented Line Graph



(a) Original Graph (b) Digraph (c) Oriented Line Graph

Ihara Zeta Function

- Ihara Zeta Function for a graph *G(V,E)*
  - Defined over prime cycles of graph

$$Z_G(u) = \prod_{p \in P} \left(1 - u^{|p|}\right)^{-1}$$

  - Rational expression in terms of characteristic polynomial of oriented line-graph

$$Z_G(u) = \left(1 - u^2\right)^{\chi(G)} \det\left(\boldsymbol{I}_{|V(G)|} - u\boldsymbol{A} + u^2\boldsymbol{Q}\right)^{-1}$$

A is adjacency matrix of line digraph

Q =D-I (degree matrix minus identity

Characteristic Polynomials from IZF

- Perron-Frobenius operator is the adjacency matrix $\boldsymbol{T}_H$ of the oriented line graph
- Determinant Expression of IZF

$$\zeta_H(u) = \det(\boldsymbol{I}_H - u\boldsymbol{T}_H)^{-1}$$
$$= (c_0 + c_1 u + \cdots + c_{M-1}u^{M-1} + c_M u^M)^{-1}$$

  – Each coefficient,i.e. Ihara coefficient, can be derived from the elementary symmetric polynomials of the eigenvalue set $\{\lambda_1, \lambda_2, \lambda_3 \ \ldots \}$

$$c_r = (-1)^r \sum_{k_1 < k_2 < \ \ldots \ < k_r} \lambda_{k_1} \lambda_{k_2} \ldots \lambda_{k_r}$$

- Pattern Vector in terms of $\vec{v} = [c_{r1} \ c_{r2} \ \ldots \ c_{rN}]^T$

Analysis of determinant

- From matrix logs

$$\zeta(s) = \frac{1}{\det[I - Ts]} = \exp[\sum_{k>1} Tr[T^k] \frac{s^k}{k}]$$

- Tr[T^k] is symmetric polynomial of eigenvalues of T

$$Tr[T^1] = \lambda_1 + \ldots\ldots\ldots + \lambda_N$$

$$Tr[T^2] = \lambda_1^2 + \lambda_1\lambda_2 + \ldots\lambda_N^2$$

$$\ldots\ldots$$

$$Tr[T^N] = \lambda_1\lambda_2\ldots\ldots\ldots\ldots\ldots\lambda_N$$

Points of contact

- Lifting cospectrality: Emms, Hancock, Severini and Wilson showed that positive support of T-cubed can lift copspectrality of strongly regular graphs and trees (see J.Comb07 and Pattern Recogonition08).

- Spectral polynomials: Wilson, Hancock and Luo have shown how to cluster graphs using symmetric polynomials on Laplacian (PAMI05).

Analysis of determinant

- Project-out symmetric polynomials by taking r-th derivative of zeta function at origin.

- From theory of multinomials

$$\zeta^{(r)}(s) = \exp[g[s]] \sum_{k_1,\ldots k_r} \frac{k!}{k_1!\ldots k_r!} \frac{g^{(1)}(s)^{k_1}}{1!} \ldots \frac{g^{(r)}(s)^{k_r}}{r!}$$

$$g(s) = \sum_{k>1} Tr[T^k] \frac{s^k}{k}$$

Symmetric polynomials

- Derivatives

$$g^{(r)}(s) = \sum_{k>r} Tr[T^k] \frac{(k-1)!}{(k-r)!} s^{k-r}$$

- At origin

$$g^{(r)}(0) = \sum Tr[T^k] \frac{(k-1)!}{(k-r)!}$$

Distribution of prime cycles

- Frequency distribution for cycles of length l

$$s\frac{d}{ds}\ln\zeta(s) = \sum_{l} N_l s^l$$

- Cycle frequencies

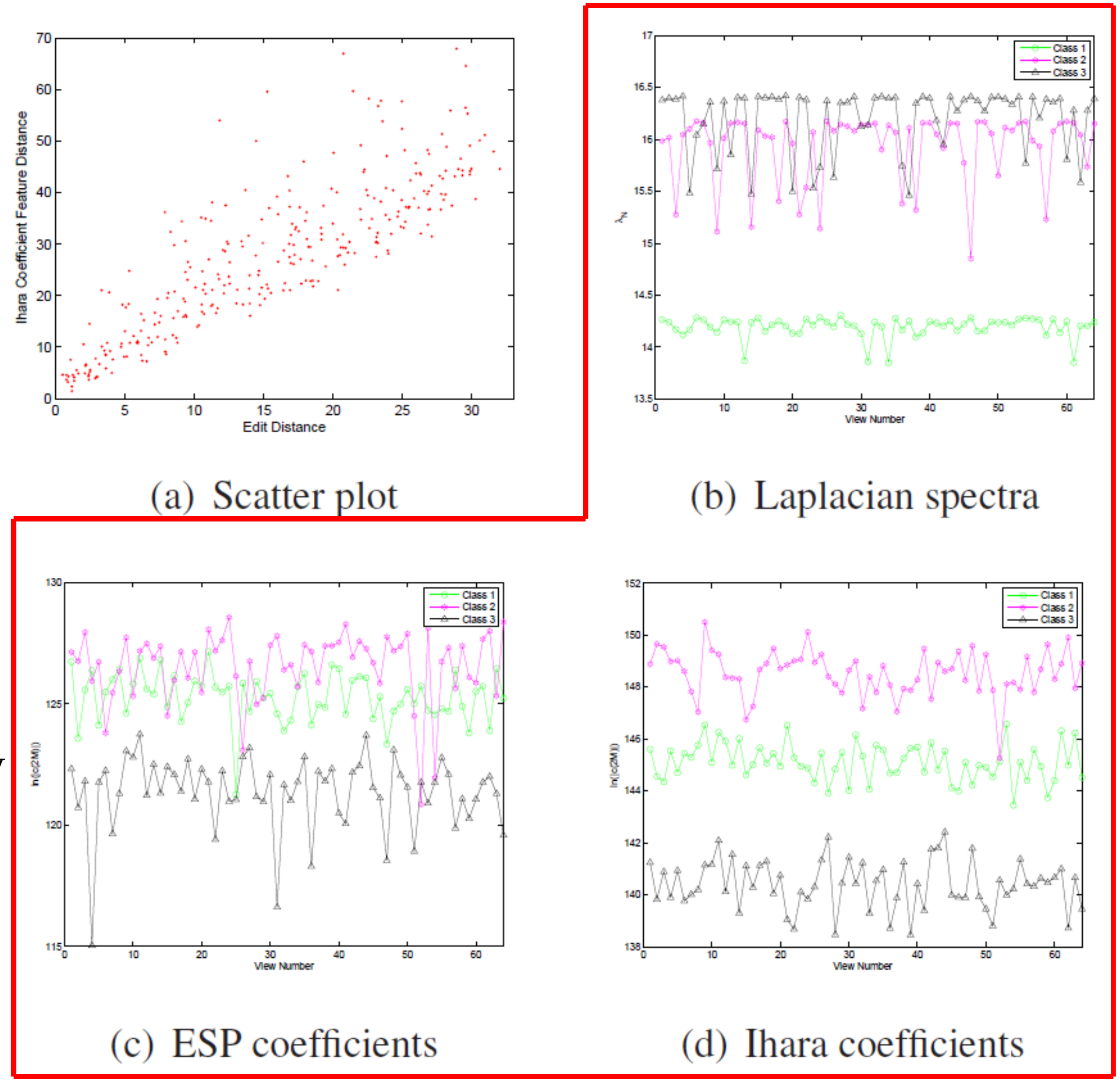
$$N_l = \frac{1}{(l-1)!}\frac{d^l}{ds^l}\ln\zeta(s)\Big|_{s=0} = Tr[T^l]$$

Experiments: Edge-weighted Graphs

Feature Distance

& Edit Distance

Three Classes of Randomly Generated Graphs



(a) Scatter plot

(b) Laplacian spectra

(c) ESP coefficients

(d) Ihara coefficients

Hypergraphs

- Hypergraphs acquiring interest as a means of representing patterns involving higher-order relations.
- Compact means of characterising hypergraphs is required so that they may be used in pattern recognition and machine learning tasks (e.g clustered, classified, similarity measures).
- No clearly accepted way of doing this.

Literature Review

- Matrix Representations for Hypergraphs
  - Chung defines Laplacian Matrix for K-regular Hypergraphs
  - Agarwal *et al.* have reviewed alternative graph representations of hypergraph and explored their relationships.
    - Graph representations of a hypergraph is needed
      - Star Expansion
      - Clique Expansion
    - The Laplacian matrix of the associated graph is regarded as that of the resulting hypergraph

Hypergraph Laplacian

- Alternative definitions of the hypergraph Lapliacian derived from different graph representations of a hypergraph.
- For a hypergraph with incidence matrix ***H***
  - Definition 1[Ren *et al*. SSPR 2008]

$$\boldsymbol{A}_H = \boldsymbol{H}\boldsymbol{H}^T - \boldsymbol{D}_v \quad \boldsymbol{L}_H = \boldsymbol{D}_v - \boldsymbol{A}_H = 2\boldsymbol{D}_v - \boldsymbol{H}\boldsymbol{H}^T$$

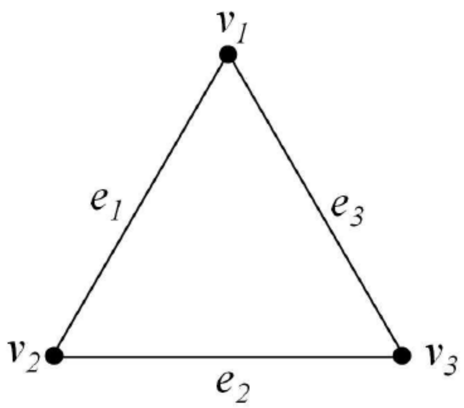
  - Definition 2 [Zhou *et al*. ICML 2005]

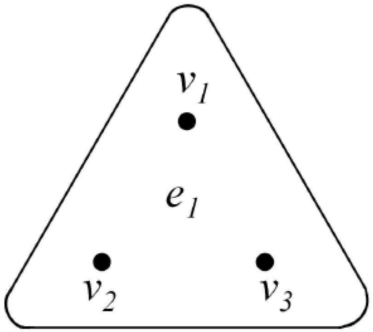
$$\hat{\boldsymbol{L}}_H = \boldsymbol{I} - \boldsymbol{D}_v^{-1/2}\boldsymbol{H}\boldsymbol{D}_e\boldsymbol{H}^T\boldsymbol{D}_v^{-1/2}$$

$\boldsymbol{D}_v$ is the diagonal vertex degree matrix $\boldsymbol{D}_e$ is the diagonal vertex degree

Examples of Hypergraph Laplacian

- An example





Fro Definition 1

$$\boldsymbol{A}_H = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \qquad \boldsymbol{L}_H = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

For Definition 2

$$\hat{\boldsymbol{L}}_{H1} = \begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{pmatrix} \qquad \hat{\boldsymbol{L}}_{H2} = \begin{pmatrix} 1/2 & -1/4 & -1/4 \\ -1/4 & 1/2 & -1/4 \\ -1/4 & -1/4 & 1/2 \end{pmatrix}$$

$$\hat{\boldsymbol{L}}_{H2} = \tfrac{3}{4}\hat{\boldsymbol{L}}_{H1}$$

Deficiencies of Hypergraph Laplacain

- Origins
  - The Laplacian matrix only records the adjacency relationships between pairs of nodes and neglects the cardinalities of the hyperedges. This results in information loss when relational orders of varying degree are present
- Possible Solution
  - Explore representations that are capable of distinguishing hypergraphs with the same pairwise connectivity between the same set of vertices, but with different relational orders.

Deficiencies of Hypergraph Laplacian

- Origins
  - The Laplacian matrix only records the adjacency relationships between pairs of nodes and neglects the cardinalities of the hyperedges. This results in information loss when relational orders of varying degree are present
- Possible Solution
  - Explore representations that are capable of distinguishing hypergraphs with the same pairwise connectivity between the same set of vertices, but with different relational orders.

Ihara Zeta Function for Hypergraph

- Ihara Zeta Function for a Graph *G(V,E)*
  - Definition $Z_G(u) = \prod_{p \in P} \left(1 - u^{|p|}\right)^{-1}$
  - Rational Expression

$$Z_G(u) = \left(1 - u^2\right)^{\chi(G)} \det\left(\boldsymbol{I}_{|V(G)|} - u\boldsymbol{A} + u^2\boldsymbol{Q}\right)^{-1}$$

- Ihara Zeta Function for a Hypergraph *H(V,E$_H$)*
  - Definition $\zeta_H(u) = \prod_{p \in P_H} \left(1 - u^{|p|}\right)^{-1}$
  - Rational Expression

$$\zeta_H(u) = (1 - u)^{\chi(BG)} \det\left(\boldsymbol{I}_{|V(H)|+|E_H(H)|} - \sqrt{u}\boldsymbol{A}_{BG} + u\boldsymbol{Q}_{BG}\right)^{-1}$$

$$\boldsymbol{A}_{BG} = \begin{bmatrix} \boldsymbol{0}_{|V(H)|\times|E_H(H)|} & \boldsymbol{H}^T \\ \boldsymbol{H} & \boldsymbol{0}_{|E_H(H)|\times|V(H)|} \end{bmatrix}$$

Adjacency matrix of the associated bipartite graph

Hypergraph Transformation

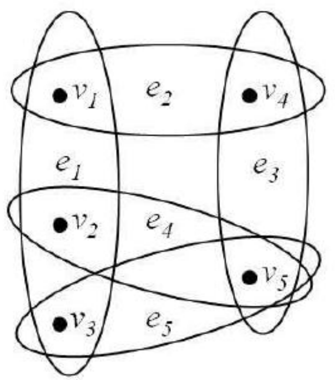


Fig. 1. Hypergraph

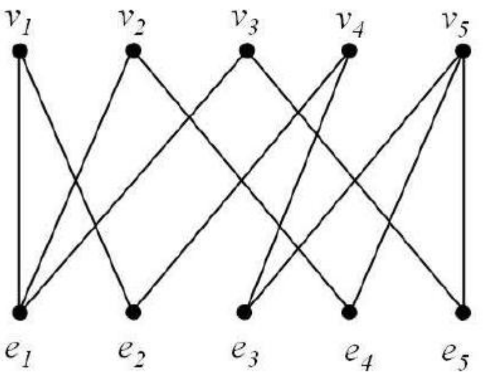


Fig. 2. Bipartite Graph

Determinant Expression of IZF for A Hypergraph

- Oriented Line Graph
  - Clique Graph from The Original Hypergraph

Hypergraph Transformation

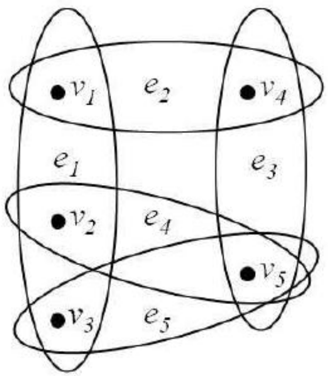


Fig. 1. Hypergraph

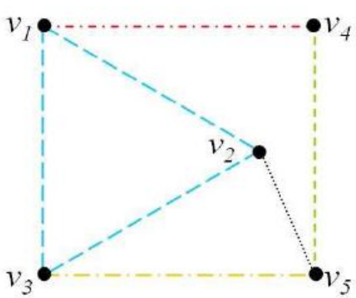


Fig. 3. Clique

Determinant Expression of IZF for A Hypergraph

- Oriented Line Graph
  - Clique Graph from The Original Hypergraph
  - Symmetric Digraph from The Clique Graph

Hypergraph Transformation

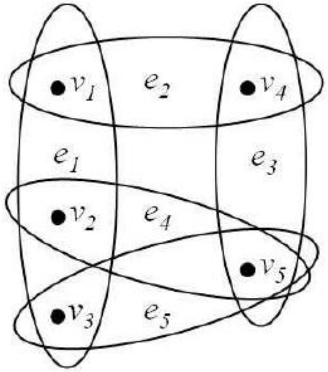


Fig. 1. Hypergraph

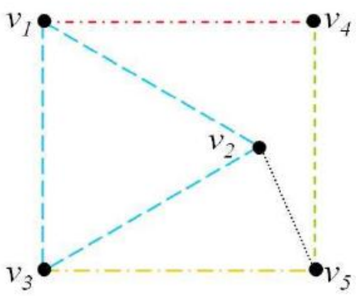


Fig. 3. Clique

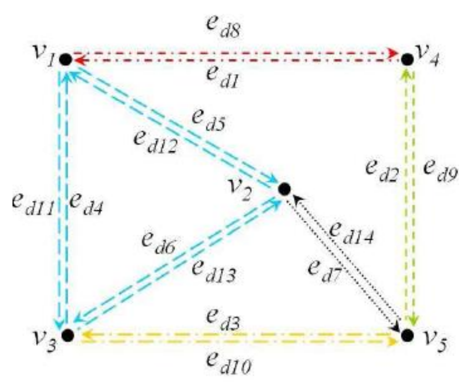


Fig. 4. Di-clique

Determinant Expression of IZF for A Hypergraph

- Oriented Line Graph
  - Clique Graph from The Original Hypergraph
  - Symmetric Digraph from The Clique Graph
  - Oriented Line Graph from The Symmetric Digraph According the Follow Rules

$$\begin{cases} V_{ol} &= E_d(DGH) \\ E_{ol} &= \{(e_d(u,v), e_d(v,w)) \in E_d \times E_d \,;\, u \cup w \not\subset E_H\}. \end{cases}$$

Hypergraph Transformation

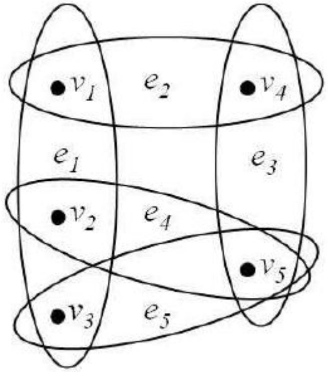


Fig. 1. Hypergraph

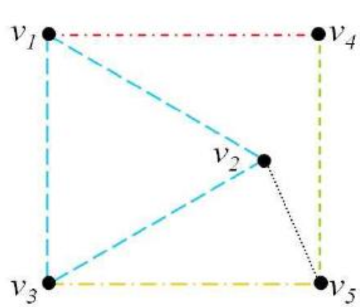


Fig. 3. Clique

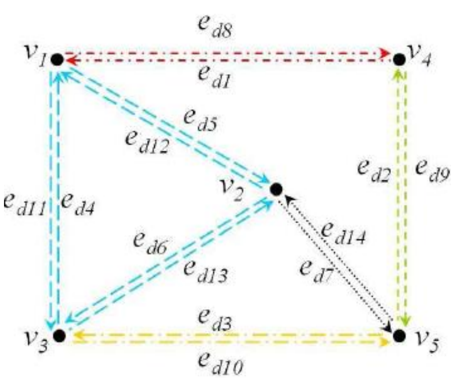


Fig. 4. Di-clique

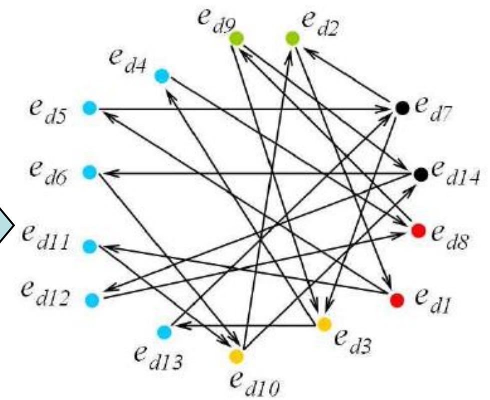


Fig. 5. Oriented Line Graph

Characteristic Polynomials from IZF

- Perron-Frobenius operator is the adjacency matrix $\boldsymbol{T}_H$ of the oriented line graph
- Determinant Expression of IZF

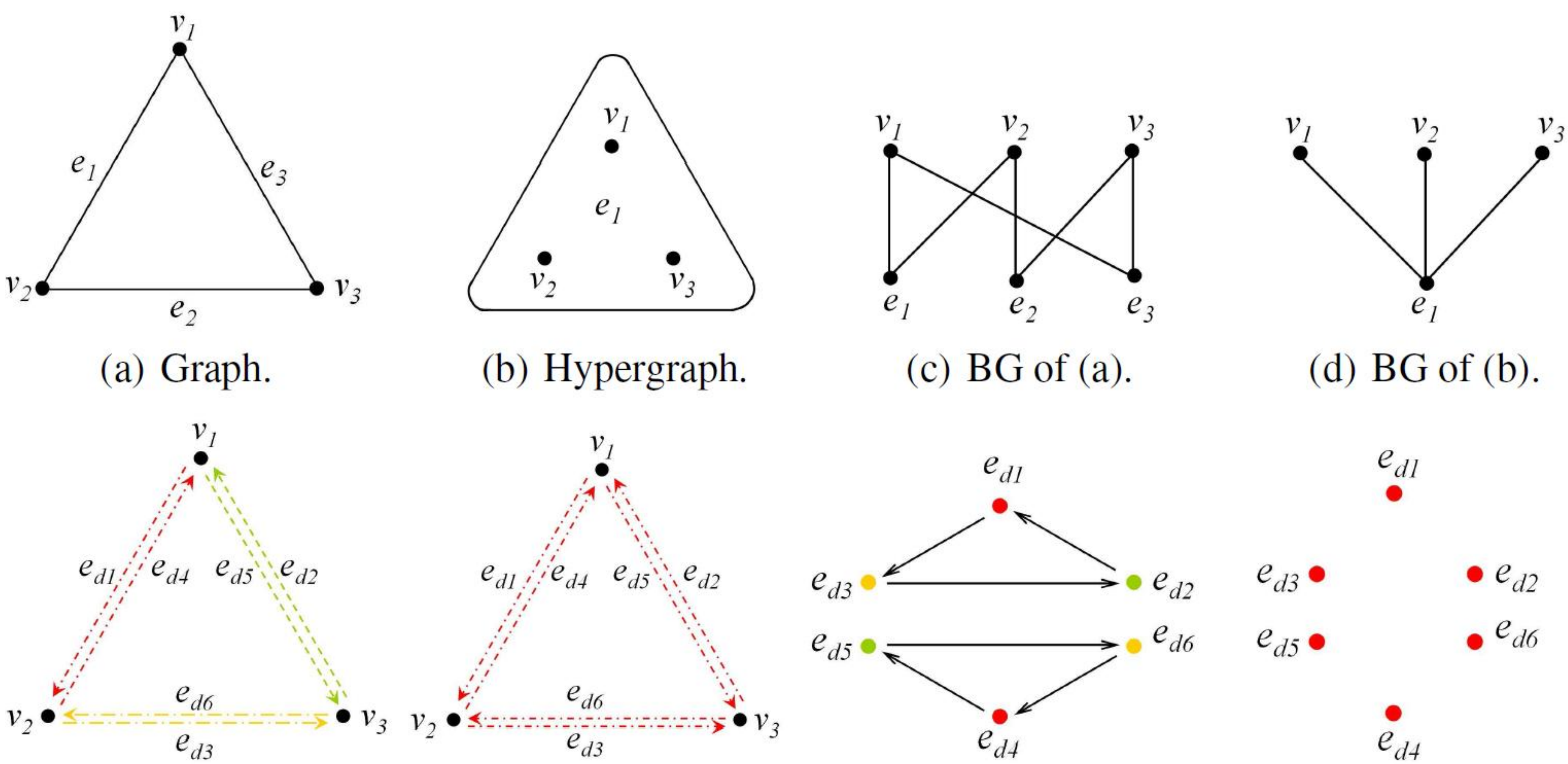
$$\zeta_H(u) = \det(\boldsymbol{I}_H - u\boldsymbol{T}_H)^{-1}$$
$$= (c_0 + c_1 u + \cdots + c_{M-1}u^{M-1} + c_M u^M)^{-1}$$

  - Each coefficient,i.e. Ihara coefficiet, can be derived from the elementary symmetric polynomials of the eigenvalue set $\{\lambda_1, \lambda_2, \lambda_3 \ \ldots \}$

$$c_r = (-1)^r \sum_{k_1<k_2<\ \ldots\ <k_r} \lambda_{k_1}\lambda_{k_2} \ldots \lambda_{k_r}$$

- Pattern Vector in terms of $\vec{v} = [c_{r1}\ c_{r2}\ \ldots\ c_{rN}]^T$

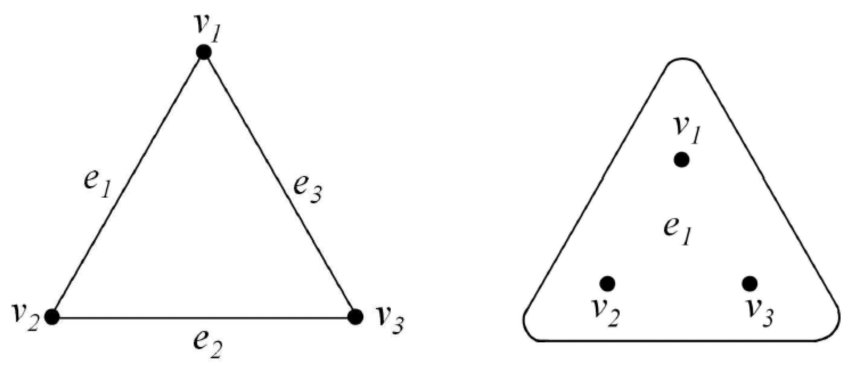
Ihara coefficients for distinguishing the previous example…



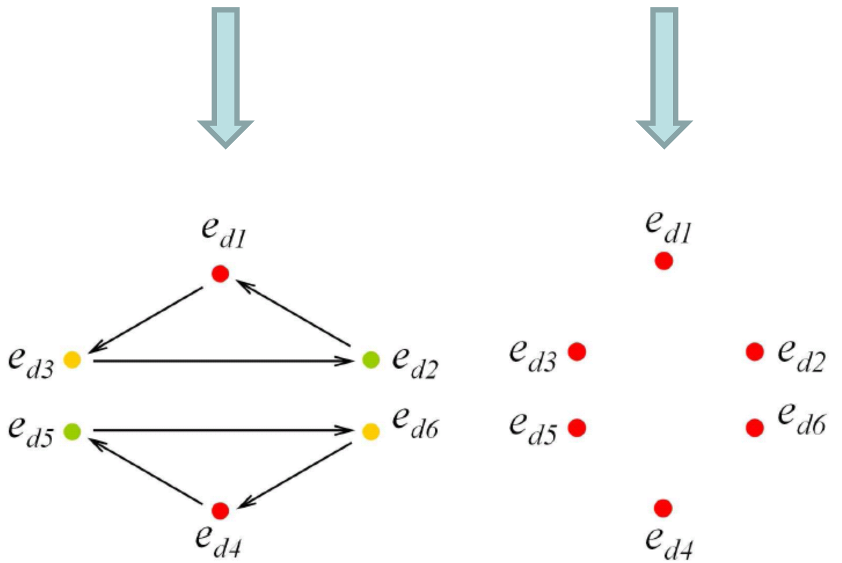
(a) Graph. (b) Hypergraph. (c) BG of (a). (d) BG of (b).

(e) Digraph of (a). (f) Digraph of (b). (g) OLG of (a). (h) OLG of (b).

Cont.



(a) Graph. (b) Hypergraph.



(g) OLG of (a). (h) OLG of (b).

Perron-Frobenius Operators for the hypergraph in Figs. (a) and (b)

$$\boldsymbol{T}_{Ha} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{T}_{Hb} = 0.$$

Coefficient Numerical Computation

- The eigen-decomposition on T_H tends to be computationally expensive
- Computation based on the bipartite graph rather than the oriented line graph would be computationally economical $\zeta_H^{-1}(u) = Z_{BG}^{-1}(\sqrt{u}) = \det(\boldsymbol{I}_{BG} - \sqrt{u}\boldsymbol{T}_{BG})$

where $Z_{BG}^{-1}(u) = \prod_{p \in P_{BG}} \left(1 - u^{|p|}\right)^{-1} = \left(1 - u^{|p_1|}\right)\left(1 - u^{|p_2|}\right)\left(1 - u^{|p_3|}\right)\cdots$

is a polynomial with the odd coefficients equal to zeros, that is:

$$
\begin{aligned}
Z_{BG}^{-1}(u) &= \det(\boldsymbol{I}_{BG} - u\boldsymbol{T}_{BG}) \\
&= \tilde{c}_0 + \tilde{c}_1 u + \tilde{c}_2 u^2 + \tilde{c}_3 u^3 + \tilde{c}_4 u^4 + \tilde{c}_5 u^5 + \tilde{c}_6 u^6 + \cdots \\
&= \tilde{c}_0 + \tilde{c}_2 u^2 + \tilde{c}_4 u^4 + \tilde{c}_6 u^6 + \cdots .
\end{aligned}
$$

So $\zeta_H^{-1}(u) = Z_{BG}^{-1}(\sqrt{u}) = \det(\boldsymbol{I}_{BG} - \sqrt{u}\boldsymbol{T}_{BG}) = \left(1 - (\sqrt{u})^{|p_1|}\right)\left(1 - (\sqrt{u})^{|p_2|}\right)\left(1 - (\sqrt{u})^{|p_3|}\right)\cdots$

$$
\begin{aligned}
&= \tilde{c}_0 + 0\sqrt{u} + \tilde{c}_2(\sqrt{u})^2 + 0(\sqrt{u})^3 + \tilde{c}_4(\sqrt{u})^4 + 0(\sqrt{u})^5 + \tilde{c}_6(\sqrt{u})^6 + \cdots \\
&= \tilde{c}_0 + \tilde{c}_2 u + \tilde{c}_4 u^2 + \tilde{c}_6 u^3 + \cdots = c_0 + c_1 u + c_2 u^2 + c_3 u^3 + \cdots .
\end{aligned}
$$

Thus the Ihara coefficients of hypergraph can be efficiently obtained by selecting even-indexed Ihara coefficients of the associated bipartite graph.

Experimental Evaluation

- Hypergraph Representation of Objects

$$h(i,j) = \begin{cases} 1 & \text{if } \|\boldsymbol{c}(v_i) - \boldsymbol{c}(v_j)\| \leq \Phi_{j1} \text{ and if } \mid I(v_i) - I(v_j) \mid \leq \Phi_{j2} \\ 0 & \text{otherwise.} \end{cases}$$

Hyperedges capture spatial proximity and grey-scale similarity.

- Dataset
  - Model house image sequences
  - COIL Dataset
- Methods for comparison
  - Hypergraph Normalized Laplacian [Zhou *et al*. ICML 2005]
  - Hypergraph Laplacian [Ren *et al*. SSPR 2008]

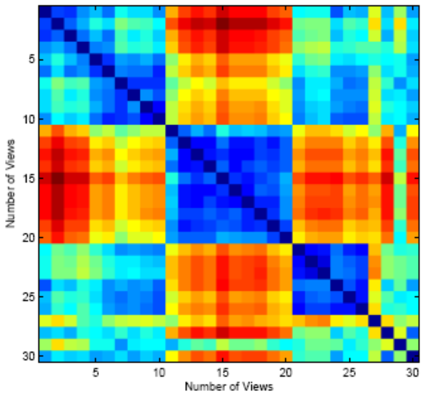
Real world data

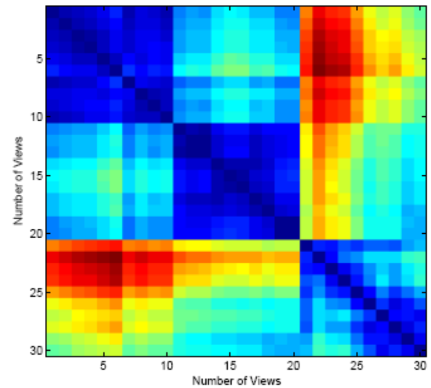
Multiple views of each object as camera pans.

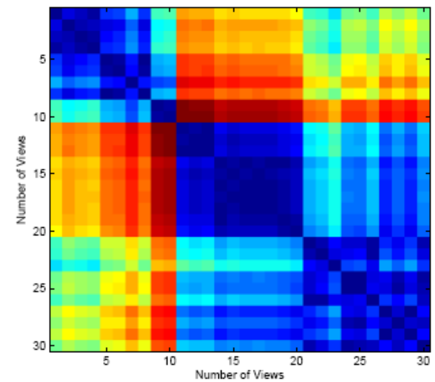
Houses

Experimental Evaluation

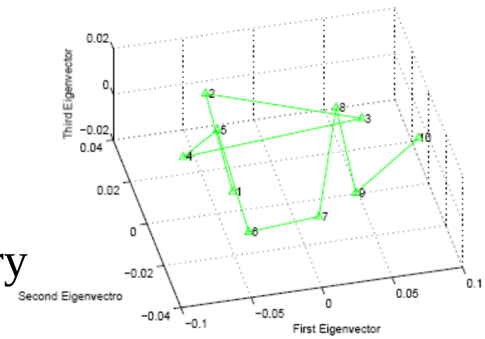
Distance Map

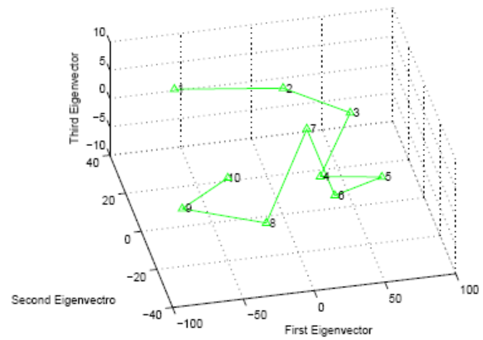


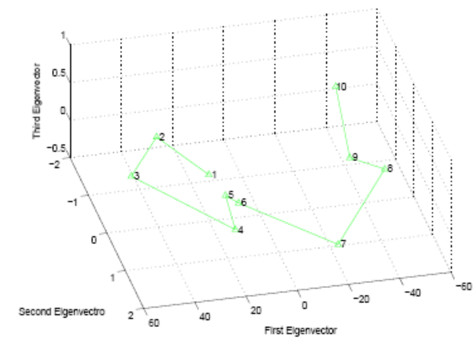




Within-class Trajectory



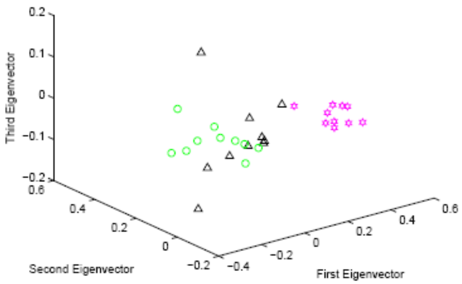




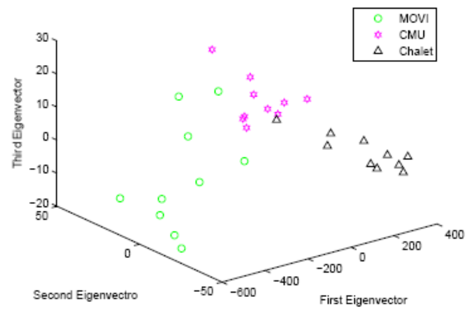
Normalized Laplacian Spectra Laplacian Spectra Ihara coefficients

Clustering

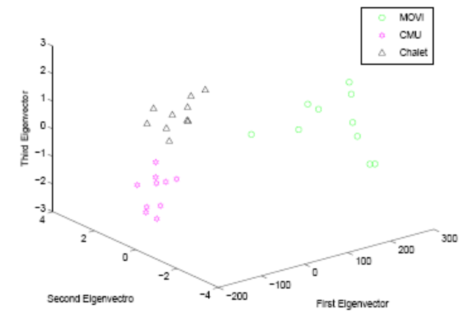
Clustering on Model Houses



Normlized Laplacian Spectra



Laplacian Spectra

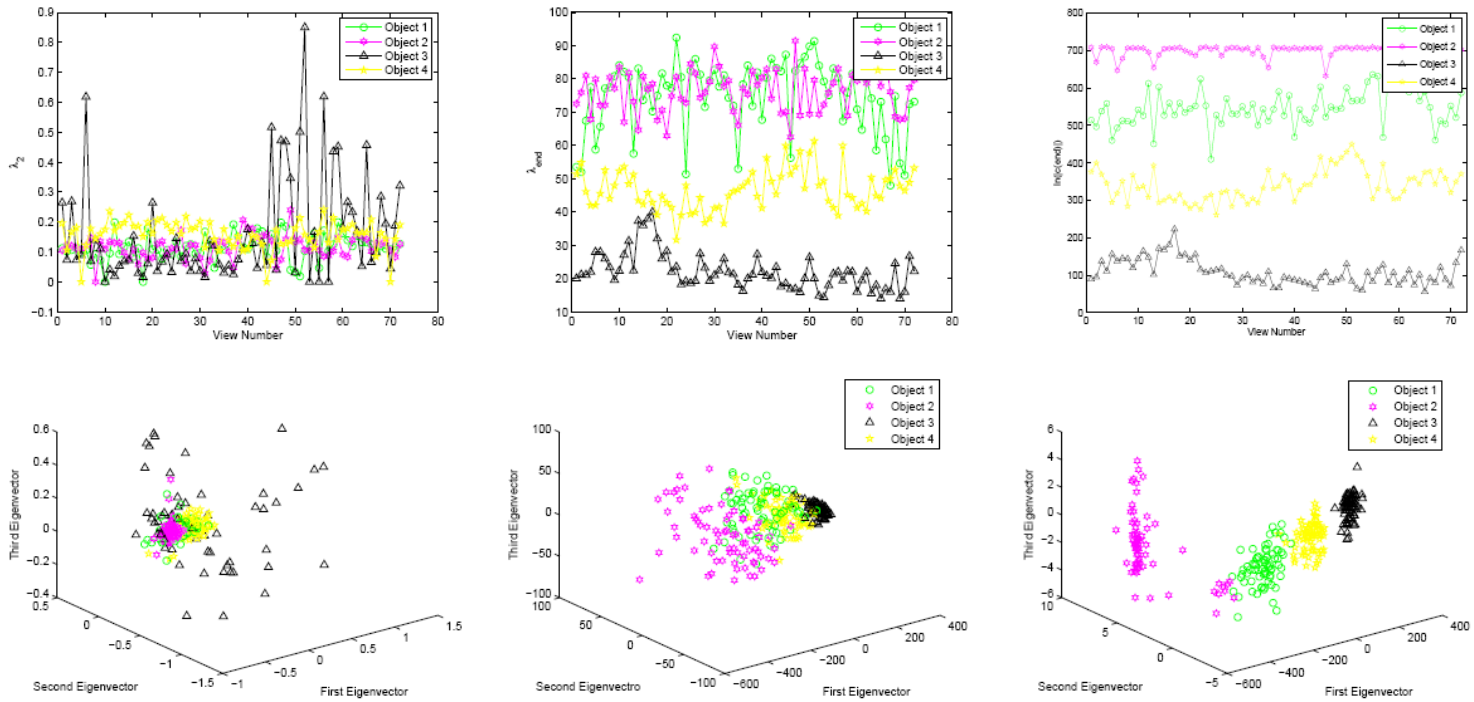


Ihara coefficients

COIL

COIL coefficient Plot

Clustering on COIL objects



Zou Laplacian Spectra | Peng Laplacian Spectra | Ihara coefficients

Pattern Vector	Number of Object Classes			
	5	6	7	8
Truncated Normalized Laplacian Spectra	0.7323	0.7074	0.7650	0.8030
Truncated Laplacian Spectra	0.8574	0.8564	0.8454	0.8449
Ihara Coefficients	0.9355	0.8859	0.8716	0.8812

Table 1: Rand Indices

Conclusion

- Ihara coefficients provide a flexible tool for both characterizing pairwise structures and higher order structures.
- Ihara coefficients capable of distinguishing structures with the same pairwise connectivity but different relational orders.
- Propose an efficient for computing Ihara zeta function