
Proceedings Paper:
Aziz, Furqan, Hancock, Edwin R. orcid.org/0000-0003-4496-2028 and Wilson, Richard C. orcid.org/0000-0001-7265-3033 (2016) Graph Entropy from Closed Walk and Cycle Functionals. In: Structural, Syntactic, and Statistical Pattern Recognition - Joint IAPR International Workshop, S+SSPR 2016, Mérida, Mexico, November 29 - December 2, 2016, Proceedings. Lecture Notes in Computer Science . Springer , pp. 174-184.

https://doi.org/10.1007/978-3-319-49055-7_16

Graph Entropy from Closed Walk and Cycle Functionals

Furqan Aziz[1], Edwin R. Hancock[2], and Richard C. Wilson[2]

[1] Department of Computer Science, IM|Sciences, Peshawar, Pakistan.
furqan.aziz@imsciences.edu.pk
[2] Department of Computer Science, University of York, YO10 5GH, UK.
{edwin.hancock, richard.wilson}@york.ac.uk

Abstract. This paper presents an informational functional that can be used to characterise the entropy of a graph or network structure, using closed random walks and cycles. The work commences from Dehmer's information functional, that characterises networks at the vertex level, and extends this to structures which capture the correlation of vertices, using walk and cycle structures. The resulting entropies are applied to synthetic networks and to network time series. Here they prove effective in discriminating between different types of network structure, and detecting changes in the structure of networks with time.

Keywords: Graph entropy, Random walks , Ihara coefficients.

1 Introduction

The problem of determining the complexity of network structures is an elusive one, which has challenged graph-theorists for over five decades. Broadly speaking the are two approaches to the problem. According to randomness complexity, the aim is to probabilistically characterise the degree of disorganisation of a network, and Shannon entropy provides one convenient way of doing this. One of the earliest attempts at computing network entropy was proposed by Körner [3]. This involves computing the entropy of a vertex packing polytope, and is linked to the chromatic number of the graph. Another simple approach is to use Shannon entropy to compute the homogeneity of the vertex degree distribution. Statistical complexity, on the other hand aims to characterise network complexity beyond randomness. One of the shortcomings of randomness complexity is that it does not capture vertex correlations. To overcome this problem, statistical complexity allows more global structure to be probed. For instance, by using the logical or thermodynamic depth of a network, the details of inhomogeneous degree structure can be problem. One powerful techniques here is to use a variant of the Kologomorov-Chaitin [4, 5] complexity to measure how many operations are need to transform a graph into a canonical form (see [9] for a review of network complexity).

So although entropy based methods provide powerful tools to characterise the properties of a complex network, one of the challenges is to define the entropy in a manner that can capture the correlations or long-range interactions

between vertices. One way to do this is to adopt path or cycle-based methods or to use other substructures that allow networks to be decomposed into non-local primitives [7, 8]. In this way some of the strengths of both the randomness and statistical approaches to complexity can be combined. One approach that takes an important step in this direction is thermodynamic depth complexity [9]. Here a Birckhoff-vonNeumann polytope is fitted to the heat kernel of a graph. The polytopal expansion uses permutation matrices as a basis, and the Shannon entropy associated with the polytopal expansion coefficients can be used to provide a depth based characterisation of network structure as a function of the diffusion time. However, this approach is time consuming and does link directly to the topological sub-structures which go to form the global network structure.

Here we adopt a different approach, with the aim measuring the entropy associated with closed random walks and cycles in graphs. Our starting point, is the information functional introduced by Dehmer and his co-workers. This allows the entropy of a network or graph to be computed from a functional defined on its vertices. Here, on the other hand we extend this functional to be defined over closed random walks and cycles. The functional for closed random walks builds on Estrada's index [2], while that for cycles uses the coefficients of the Ihara-zeta function, which measure the frequencies of prime cycles in a graph. These two new informational functionals are applied to a variety of synthetic and real world data.

2 Graph Entropy and Information Functionals

In this section we briefly explain the general framework proposed by Dehmer to define graph entropy.

Definition 1 (Dehmer [1]) *Given a graph $G = (V, E)$, its entropy is defined as*

$$I_f(G) := -\sum_{i=1}^{|V|} \frac{f(v_i)}{\sum_{j=1}^{|V|} f(v_j)} log \frac{f(v_i)}{\sum_{j=1}^{|V|} f(v_j)} \tag{1}$$

where $f(v_i)$ is an arbitrary local vertex information functional. □

A number of information functionals can be defined that capture different local properties of the graph. For example, Dehmer has proposed the following definitions for the information functional that gauges the metrical properties of a graph.

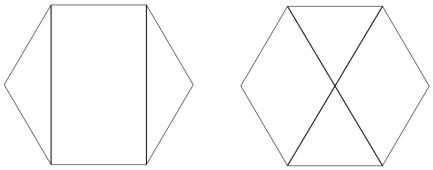
Definition 2 *Given a graph G, the local information functional is defined as*

$$f^V(v_i) := \alpha^{c_1|S_1(v_i,G)|+c_2|S_2(v_i,G)|+....+c_\rho|S_\rho(v_i,G)|}$$

where $|S_k(v_i, G)|$ represents the number of shortest paths of length k reachable from the node v_i and $c_1, c_2, ...c_\rho$ are real valued constants. □

The above information functional can easily be obtained by definition of the j-sphere [1].

One of the problem with entropy defined in this way is that it captures only the local properties of a graph, and it is therefore sensitive to the degree distribution of the graph. For example, consider the non-isomorphic graphs in the Figure 1. The above two graphs are structurally different as G_1 contains



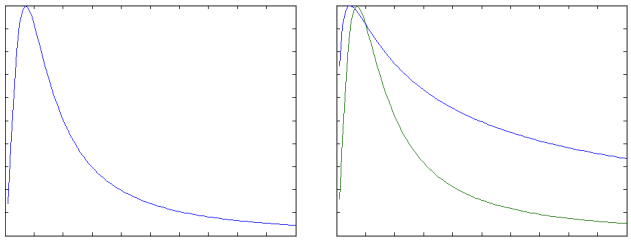
(a) G1 (b) G2

Fig. 1. Two non-isomorphic graphs

two triangles while G_2 does not contain any triangle. However it can be shown that, using Definition 2, the entropy for both the graphs is

$$I_f(G) = -\left[\frac{2\alpha^{16}}{2\alpha^{16}+4\alpha^{18}}\log\left(\frac{2\alpha^{16}}{2\alpha^{16}+4\alpha^{18}}\right)+\frac{4\alpha^{18}}{2\alpha^{16}+4\alpha^{18}}\log\left(\frac{4\alpha^{18}}{2\alpha^{16}+4\alpha^{18}}\right)\right]$$

Figure 2(a) plots the entropy of the two graphs as a function of α.



(a) Path Length (b) Random Walk

Fig. 2. Entropy computed from different information functionals

3 Substructure based approach for Graph Entropy

To overcome the problems associated with the methods discussed above, we use random walks on a graph to define graph entropies. Motivated by Esterada's Index (EI), we define the information functional based on closed random walks.

Definition 3 *Let $|R_k(v_i, G)|$ denotes the number of random walks in graph G of length k, starting from and ending at the node v_i. We define the information functional as*

$$f^V(v_i) := \alpha^{c_1|R_1(v_i,G)|+c_2|R_2(v_i,G)|+....+c_k|R_k(v_i,G)|}$$

Here $c_1, c_2, ...c_\rho$ are real valued constants. □

Note that if we put $c_n = \frac{1}{n!}$, then the value in the exponent becomes equal to Estrada's Index (EI) [2].

Using Definition 3 the entropy for graph G_1 of Figure 1 can be shown to be

$$I_f(G_1) = -\left[\frac{2\alpha^{1.5}}{2\alpha^{1.5}+4\alpha^{3.42}}\log\left(\frac{2\alpha^{1.5}}{2\alpha^{1.5}+4\alpha^{3.42}}\right)+\frac{4\alpha^{3.42}}{2\alpha^{1.5}+4\alpha^{3.42}}\log\left(\frac{4\alpha^{3.42}}{2\alpha^{1.5}+4\alpha^{3.42}}\right)\right]$$

while for G_2 it can be shown to be

$$I_f(G_2) = -\left[\frac{2\alpha^{1.8}}{2\alpha^{1.8}+4\alpha^{2.68}}\log\left(\frac{2\alpha^{1.8}}{2\alpha^{1.8}+4\alpha^{2.68}}\right)+\frac{4\alpha^{2.68}}{2\alpha^{1.8}+4\alpha^{2.68}}\log\left(\frac{4\alpha^{2.68}}{2\alpha^{1.8}+4\alpha^{2.68}}\right)\right]$$

Figure 2(b) plots the entropies as a function of α for both graphs, which suggests that random walks are a more powerful tool for discriminating between different graph structures.

Our goal in this paper is to define graph entropy using the substructures in the graph. To this end, we decompose the graph into substructures and then use the frequency of a particular substructure to compute the information functional. We now propose a general framework to define entropy.

Definition 4 *Given a graph G, we define the graph entropy as*

$$I_f(G) = -\sum_{k=1}^{n}\frac{f(S_k)}{\sum_{i=1}^{n}f(S_i)}log\frac{f(S_k)}{\sum_{i=1}^{n}f(S_i)} \tag{2}$$

where S_k represents the information functional computed from k^{th} substructure. □

There are a number of ways in which a graph can be decomposed and information functionals can be defined. Here we use the coefficients of Ihara zeta function. The reason for using Ihara coefficients is that they are related to the number of simple cycles in the graphs, and hence can be used to gauge the cyclic complexity of a graph.

The Ihara zeta function associated to a finite connected graph G is defined to be a function of $u \in \mathbb{C}$ with $|u|$ sufficiently small by [13]

$$\zeta_G(u) = \prod_{c\in[C]}\left(1-u^{l(c)}\right)^{-1} \tag{3}$$

The product is over equivalence classes of primitive closed backtrackless, tail-less cycles $c = (v_1, v_2, v_3, ..., v_r = v_1)$ of positive length r in G. Here $l(c)$ represents the length of the cycle.

The reciprocal of Ihara zeta function can also be written in the form of a determinant expression [14]

$$\zeta_G(u) = \frac{1}{\det(I-uT)} \tag{4}$$

where T, the Perron-Frobenius operator, is the adjacency matrix of the oriented line graph of the original graph, and I is the identity matrix of size $2m$. Here m is the number of edges in the original graph. The oriented line graph is constructed by first converting the graph into equivalent digraph, and then replacing each arc of the resulting digraph by a vertex. These vertices are connected if the head of one arc meets the tail of another while preventing backtracking, i.e., arcs corresponding to same edge are not connected.

Since the reciprocal of the Ihara zeta function can be written in the form of a determinant expression, therefore it can also express as a polynomial of degree at most 2. i.e.,

$$\zeta_G(u)^{-1} = \det(I - uT) = c_0 + c_1 u + c_2 u^2 + c_3 u^3 + ... + c_{2m} u^{2m} \tag{5}$$

where $c_0, c_1, ..., c_{2m}$ are called Ihara coefficients and are related to the frequencies of simple cycles in the graph. In particular, it can be shown that if G is a simple graph then $c_0 = 1$, $c_1 = 0$, $c_2 = 0$. Furthermore, the coefficients c_3, c_4 and c_5 are the negative of twice the number of triangles, squares, and pentagons in G respectively. The coefficient c_6 is the negative of the twice the number of hexagons in G plus four times the number of pairs of edge disjoint triangles plus twice the number of pairs of triangles with a common edge, while c_7 is the negative of the twice the number of heptagons in G plus four times the number of edge disjoint pairs of one triangle and one square plus twice the number of pairs of one triangle and one square that share a common edge [6, 7]. In [7], we have developed method that can be used to compute Ihara coefficients in a polynomial amount of time.

Definition 5 *Let c_i represents the i^{th} Ihara coefficient. We define the information functionals, $f(c_i)$, as*

$$f(i) := \alpha^{k_i c_i},$$

where k_i are constants. □

Note that the first three Ihara coefficients are constants [6, 7], and therefore we can ignore these coefficients in our computation. Also, since the coefficient beyond the first few coefficients contain redundant information [7], therefore we only retain few coefficients and discard the remainder.

4 Experiments

In this section we explore whether the proposed method can be used to distinguish between graphs that are structurally different. The first dataset comprises synthetically generated networks according to some known network models. Next we apply the proposed methods to three different types of graphs extracted from the COIL [20] dataset. The purpose here is to demonstrate the ability of entropy defined using Ihara coefficients to differentiate between md2 graphs (graphs where the degree of each node is at least 2) with different structures. Finally we investigate the use of the proposed method for the purpose of detecting crises and different stages in an evolving network.

4.1 Random Graphs

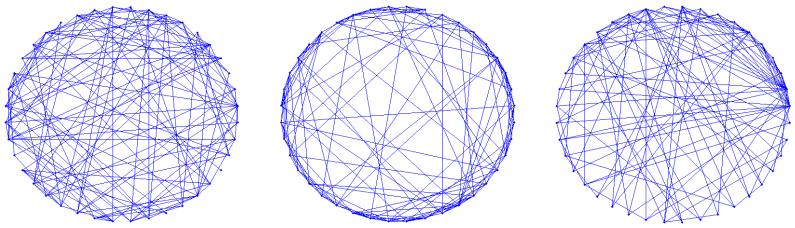
We commence by experimenting with the proposed method on synthetically generated graphs according to the following three models.

Erdős-Rényi model(ER) [10]: An ER graph $G(n, p)$ is constructed by connecting n vertices randomly with probability p. i.e., each edge is included in the graph with probability p independent from every other edge. These models are also called $r$andom networks.

Watts and Strogatz model(WS) [11]: A WS graph $G(n, k, p)$ is constructed in the following way. First construct a regular ring lattice, a graph with n vertices and each vertex is connected to k nearest vertices, k/2 on each side. Then for every vertex take every edge and rewire it with probability p. These models are also called $s$mall-world networks.

Barabási-Albert model(BA) [12]: A BA graph $G(n, n_0, m)$ is constructed by an initial fully connected graph with n_0 vertices. New vertices are added to the graph one at a time. Each new vertex is connected to m previous vertices with a probability that is proportional to the number of links that the existing nodes already have. These models are also called $s$cale-free networks.

Figure 6 shows an example of each of these models. We have generated 200



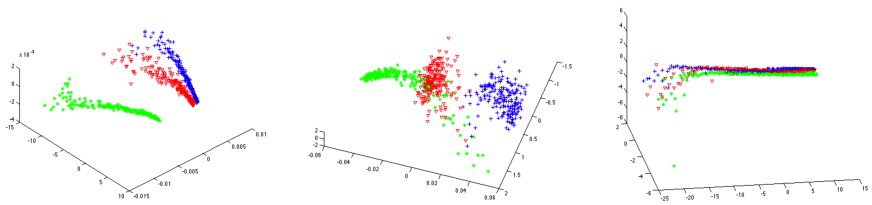
(a) random graphs (b) small-world graph (c) scale-free graph

Fig. 3. Graph models

graphs for each of these models with n = 21, 22,...,220. The parameters for these models were chosen in such a way that graphs with the same number of vertices have roughly the same number of edges. For ER graphs we choose $p = 15/n$, For WS graphs we choose $k = 16$ and $p = .25$, and for BA graphs we choose $n_0 = 9$ and $k = 8$.

Next we compute the entropy of each graph using both the random walks and Ihara coefficients. We have used the coefficients of Ihara zeta function and closed random walks as information functional to compute entropy. For the Ihara coefficients, we have selected the first six coefficients starting for c_3 to compute the information functional, and used Definition 4 to find entropy. To avoid scaling effects, the last three coefficients were multiplied with $1/|E|$. For random walks, we have selected random walk up to length 6 and used Definition 1 to find entropy. Each coefficient was multiplied with $1/k!$, where k represents the length of the random walk.

For each graph, we have generated a feature vector of length 100. The feature vector is constructed by choosing different values of α and computing information functional and the resulting entropy for each value of α. In our case we have put $\alpha = 0.1, 0.2, ..., 10$. This transforms each graph into a feature vector in a 100 dimensional feature space. To visualise the results, we have performed Principal Component Analysis *PCA* on the resulting feature vectors and embed the results in a three dimensional vector space. *PCA* is mathematically defined [17] as an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on. Figure 4(a) shows the resulting embedding on the first three principal components for feature vectors computed using random walks, while Figure 4(b) shows the resulting embedding on the first three principal components for feature vectors computed using Ihara coefficients. To compare the results, we have also used the local information functional defined by Dehmer [1] that is computed from path lengths. Figure 4(c) shows the resulting embedding. The resulting embedding shows that the



(a) Random Walks (b) Ihara Coefficients (c) Path Length

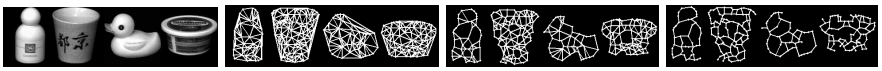
Fig. 4. *PCA* embedding of different methods

entropy computed from random walks gives best results. The Ihara coefficients on the other hand does not provide very good inter-class separation. This is due to the fact that graphs generated using random models have limited number of cycles. Figure also suggest that local information (path length) is not very helpful in distinguishing the different families of graphs.

4.2 Graphs extracted from COIL dataset

We now perform experiments on the graphs extracted from the images in the Columbia object image library (COIL) dataset [20]. This dataset contains views of 20 different 3D objects under controlled viewer and lighting condition. For each object in the database there are 72 equally spaced views. To establish a graph on the images of objects, we first extract feature points from the image. For this purpose, we use the Harris corner detector [19]. We then construct three different types of graphs using the selected feature points as vertices, i.e., Delaunay triangulation(DT) [15], Gabriel graph(GG) [18], and relative neighbourhood graph(RNG) [16]. Figure 5(a) shows some of the COIL objects, while Figure 5(b), Figure 5(c), and Figure 5(d) show the DT, GG and RNG extracted

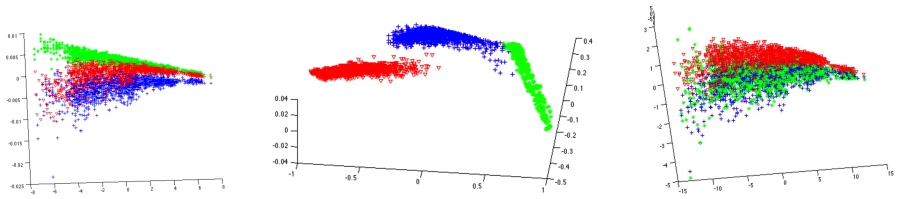
from the corresponding objects respectively. Next we used Dehmer's defini-



(a) COIL (b) DT (c) GG (d) RNG

Fig. 5. COIL objects and their extracted graphs.

tion to compute the entropy of each graph extracted from each of the 72 views of all the 20 objects. We have used the coefficients of Ihara zeta function and closed random walks as information functional to compute entropy. For the Ihara coefficients, we have selected the first six coefficients starting for c_3. To avoid the scaling effect, the last three coefficients were multiplied with $1/|E|$. For each graph, we have generated a signature of length 100 by putting $\alpha = 0.1, 0.2, ..., 10$. For random walk, we have selected random walk upto length 6. Each coefficient was multiplied with $1/k!$, where k is the length of the random walk. Finally, we have also used the local information functional defined by Dehmer to compute entropy. To compare the results, we have performed Principal Component Analysis *PCA*, on the signatures obtained by choosing different values of α for each of these methods and embed the results in a three dimensional vector space. Figure 4 compares the resulting embedding of the feature vectors on the first three principal components for all the three methods. It is clear from the figure that Ihara coefficients proves to be a powerful tool to distinguish graphs that exhibit a cyclic structure. On the other hand, the entropy computed from random walks and local paths is not very helpful in distinguishing these graphs.



(a) RW (b) Ihara (c) Path Lengths

Fig. 6. PCA embedding for COIL dataset

4.3 Time-varying Networks

In our last experiment we explore whether the proposed methods can be used as a tool for understanding the evolution of a complex network. For this purpose we choose publicly available New York Stock Exchange (NYSE) dataset[25]. This dataset consists of the daily prices of different stocks traded continuously on the New York Stock Exchange for a 25 year span from January 1986 to February 2011. A total of 347 stock were selected from this set. We construct a network over a time-window of 28 days. Here the stocks represent the nodes of the network. An edge is present, if the correlation value is above a threshold. We select

an empirical value of 0.85 as a threshold value. This was done under the assumption that, at any given time, a particular stock must interact with another stock. A new network is generated by sliding the window of 1 day and repeating the process. In this way a total of 5977 time-varying networks are generated. Since the networks generated in this way have very limited number of prime cycles, We have used the entropy defined using closed random walks. We next applied PCA on the resulting signatures. Figure 7 shows the values of the eigenvector with the highest variance. The above result suggests that the proposed method

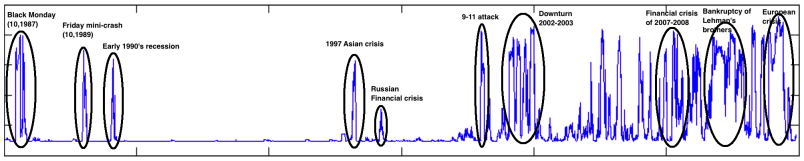


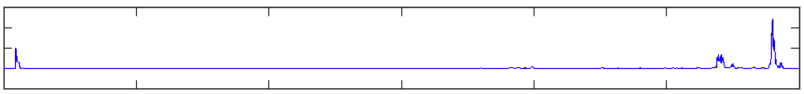
Fig. 7. Entropy computed from random walks on networks

is a very useful tool for detecting changes in a time evolving network. To compare these results, we have computed the entropy using path length as defined by Dehmer. Figure 8 shows the values of the first principal component, after applying PCA on the resulting signatures. This clearly suggests that entropies

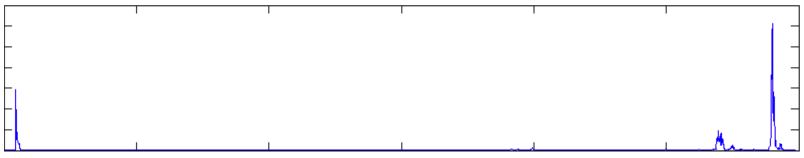
Fig. 8. Entropy computed from path lengths using j-sphere

defined using local structural properties are not very helpful to detect changes in a time evolving network.

To compare the results, we have also computed the von Neumann entropy and the Estrada index of the evolving networks. Figure 9 shows the results.



(a) VonNeumann Entropy



(b) Estrada Index

Fig. 9. VonNeumann Entropy and Estrada Index for NYSE

5 Conclusion

In this paper we have used closed random walks and simple cycles to define information functionals that can be used to define the entropy of a graph (or

a network). We have decomposed the graph into substructures and used the frequencies of these substructures to define entropy. We have also presented a generic framework that can be used to define graph entropy by decomposing the graph into substructures. Experiments were performed on numerous datasets including synthetic data, cyclic graphs, and time series data, and the results suggest that the proposed methods can be used to characterise the graphs (and networks) with very higher accuracy as compared to some of the other state-of-the-art methods.

References

1. Dehmer, M.: Information processing in complex networks: Graph entropy and information functionals, Applied Math. and Comp., Vol 201, 2008, Pages 82-94.
2. Estrada, E.: Characterization of 3D molecular structure. Chemical Physics Letters, Volume 319, Issues 5-6, 24 March 2000, Pages 713-718.
3. Körner, J.: Coding of an information source having ambiguous alphabet and the entropy of graphs. 6th Prague conference on information theory. 1973.
4. Kolmogorov, A.N.: Three approaches to the definition of the concept ?quantity of information? Probl. Peredachi Inf., 1965, Volume1, Issue1, Pages3?11
5. Chaitin, G. J. : On the Length of Programs for Computing Finite Binary Sequences Journal of the ACM, Volume 13 Issue 4, Oct. 1966.
6. Scott, G., Storm, C.: The coefficients of the Ihara zeta function, Involve - a journal of mathematics. Vol. 1 (2008), No. 2, 217?233.
7. Aziz, F., Wilson, R. C., Hancock, E. R.: Backtrackless Walks on a Graph. IEEE Trans. on Neural Netw. and Learning Systems, Volume 24, Issue 6, Pages 977-989.
8. Ren, P., Wilson, R.C., Hancock, E.R.: Graph Characterization via Ihara Coefficients. IEEE Transactions on Neural Networks, vol 22(2), 233 - 245, 29 November 2010.
9. Escolano, F., Hancock, E. R., Lozano, M.A.: Heat diffusion: Thermodynamic depth complexity of networks. Phys. Rev. E, vol 85(3), Mar 2012.
10. Erdõs, P., Rényi A.: On the evolution of random graphs. Publications of the Mathematical Institute of the Hungarian Academy of Sciences, 17-61 (1960).
11. Watts, D. J., Strogatz, S. H.: Collective dynamics of 'small-world' networks. Nature, 440-442 (1998).
12. Barabási A., and Albert R.: Emergence of Scaling in Random Networks. Science 509-512 (1999).
13. Bass, H.: The Ihara-Selberg zeta function of a tree lattice. Int. J. Math., 1992, 717-797.
14. Kotani, M., Sunada, T.: Zeta function of finite graphs. Journal of Mathematics, University of Tokyo 7(1), 2000, pages 7-25.
15. Delaunay, B.: Sur la sphre vide, Izvestia Akademii Nauk SSSR, Otdelenie Matematicheskikh i Estestvennykh Nauk, 1934, 793800.
16. Toussaint, G.T.: The relative neighbourhood graph of a finite planar set, Pattern Recognition, 1980, 261 - 268.
17. Jolliffe, I. T.: Principal Component Analysis, Springer-Verlag, New York, 1986.
18. Gabriel, K. R., Sokal R. R.: A New Statistical Approach to Geographic Variation Analysis, Systematic Zoology, 1969, 205-222.
19. Harris, C., Stephens, M.: A combined corner and edge detector, In Fourth Alvey Vision Conference, Manchester, UK, 1988, 147-151.
20. Nayar, S.K., Nene, S.A., Murase, H.: Columbia object image library (coil 100), Department of Comp. Science, Columbia University, Tech. Rep. CUCS-006-96,1996.