Trace Formula Analysis of Graphs

Bai Xiao and Edwin R. Hancock

Department of Computer Science,
University of York, York Y01 5DD, UK

Abstract. In this paper, we explore how the trace of the heat kernel can be used to characterise graphs for the purposes of measuring similarity and clustering. The heat-kernel is the solution of the heat-equation and may be computed by exponentiating the Laplacian eigensystem with time. To characterise the shape of the heat-kernel trace we use the zeta-function, which is found by exponentiating and summing the reciprocals of the Laplacian eigenvalues. From the Mellin transform, it follows that the zeta-function is the moment generating function of the heat-kernel trace. We explore the use of the heat-kernel moments as a means of characterising graph structure for the purposes of clustering. Experiments with the COIL and Oxford-Caltech databases reveal the effectiveness of the representation.

1 Introduction

The Laplacian spectrum of a graph has found widespread use in computer vision for a number of applications including segmentation [2] and routing [3], graph clustering[10] and graph indexing[4]. For instance, the Fiedler vector [1] [12] [13], i.e. the eigenvector associated with the smallest non-zero eigenvalue, can be used to perform pairwise clustering of data. The Laplacian eigenvalues may be used to characterise graphs for the purposes of clustering. Several authors have explored the use of the Laplacian and related operators to map data to manifolds in a low dimensional space [9] [15] [16] [17] [5]. These methods share the feature of using the spectrum of the Laplacian matrix to map data specified in terms of a proximity matrix to a vector space. For instance in the Laplacian eigenmap [15], the mapping is determined by the raw Laplacian spectrum. The diffusion map [5] of Lafon and Coifman constructs the mapping by raising the Laplacian eigensystem to a negative integer power. This mapping is shown to preserve the distances between nodes under a random walk, or diffusion, on the graph. In the heat-kernel embedding of Lebanon and Lafferty [6], the embedding is based on the heat-kernel and this is found by exponentiating the Laplacian eigensystem.

The aim in this paper is to explore whether the trace of the heat-kernel [7] can be used for the purposes of characterising the properties of graphs. The trace of the heat kernel is found by summing a series of terms, each of which is the result of exponentiating a Laplacian eigenvalue with time. As a result the heat-kernel trace is a function whose parameters are the Laplacian eigenvalues and whose argument is time. Our aim in this paper is to explore whether the shape of this function can be used to characterise the corresponding graph. There are several ways in which this can be done. In spectral geometry, the heat kernel trace has been used to characterise the differential geometry

of manifolds [7] [8]. Here the spectrum of the Laplace-Beltrami operator is used to construct a trace-function. This function can be expanded as a polynomial series in time, and the co-efficients of the series can be related to the Ricci curvature tensor of the manifold. Unfortunately, the relationships between the elements of the Ricci curvature tensor and the co-efficients are difficult to determine, and are only tabulated up to third order [8]. For large graphs, the Laplacian can be viewed as a discrete approximation of the Laplace-Beltrami operator and this analysis can be carried over from manifolds to graphs [14].

However, in this paper we deal with rather small graphs and take a different approach. Our idea is to measure the shape of the heat-kernel trace by taking moments with respect to time. Using the Mellin transform it is straightforward to show that the moment generating function is related to the zeta function of the graph. The zeta function is a series found by exponentiating and summing the reciprocals of the non-zero eigenvalues of the Laplacian. We construct a feature-vector whose components are the values of the zeta-function with integer argument.

Experiments with real world data taken from the COIL and Caltech-Oxford databases reveal that the zeta-function provides useful features for clustering graphs and to outperform the Laplacian spectrum.

2   The Laplacian Eigensystem and the Heat-Kernel

To commence, suppose that the graph under study is denoted by $G = (V, E)$ where V is the set of nodes and $E \subseteq V \times V$ is the set of edges. Since we wish to adopt a graph-spectral approach we introduce the adjacency matrix A for the graph where the elements are

$$A(u, v) = \begin{cases} 1 \text{ if } u, v \in E \\ 0 \text{ otherwise} \end{cases} \tag{1}$$

We also construct the diagonal degree matrix D, whose elements are given by $D(u, u) = \sum_{v \in V} A(u, v)$. From the degree matrix and the adjacency matrix we construct the Laplacian matrix $L = D - A$, i.e. the degree matrix minus the adjacency matrix. The normalised Laplacian is given by $\hat{L} = D^{-\frac{1}{2}} L D^{-\frac{1}{2}}$. The spectral decomposition of the normalised Laplacian matrix is $\hat{L} = \Phi \Lambda \Phi^T$, where $\Lambda = diag(\lambda_1, \lambda_2, ..., \lambda_{|V|})(0 = \lambda_1 < \lambda_2 < ... < \lambda_{|V|})$ is the diagonal matrix with the ordered eigenvalues as elements and $\Phi = (\phi_1|\phi_2|....|\phi_{|V|})$ is the matrix with the ordered eigenvectors as columns. Since $\hat{L}$ is symmetric and positive semi-definite, the eigenvalues of the normalised Laplacian are all positive. The eigenvector associated with the smallest non-zero eigenvector is referred to as the Fiedler-vector.

We are interested in the heat equation associated with the Laplacian, i.e.

$$\frac{\partial h_t}{\partial t} = -\hat{L} h_t \tag{2}$$

where h_t is the heat kernel and t is time. The heat kernel can hence be viewed as describing diffusion across the edges of the graph with time. The rate of flow is determined by

the Laplacian of the graph. The solution to the heat equation is found by exponentiating the Laplacian eigenspectrum, i.e.

$$h_t = \sum_{i=1}^{|V|} \exp[-\lambda_i t]\phi_i\phi_i^T = \Phi \exp[-t\Lambda]\Phi^T \tag{3}$$

The heat kernel is a $|V| \times |V|$ matrix, and for the nodes u and v of the graph G the resulting element is

$$h_t(u, v) = \sum_{i=1}^{|V|} \exp[-\lambda_i t]\phi_i(u)\phi_i(v) \tag{4}$$

When t tends to zero, then $h_t \simeq I - \hat{L}t$, i.e. the kernel depends on the local connectivity structure or topology of the graph. If, on the other hand, t is large, then $h_t \simeq \exp[-t\lambda_2]\phi_2\phi_2^T$, where λ_2 is the smallest non-zero eigenvalue and ϕ_2 is the associated eigenvector, i.e. the Fiedler vector. Hence, the large time behavior is governed by the global structure of the graph.

The trace of the heat kernel is

$$Z(t) = Tr[h_t] = \sum_{i=1}^{|V|} \exp[-\lambda_i t] \tag{5}$$

To provide an illustration of the potential utility of the trace-formula, in Figure 1 we show four small graphs with rather different topologies. Figure 2 shows the trace of the heat kernel as a function of t for the different graphs. From the plot it is clear that the curves have a distinct shape and could form the basis of a useful representation to distinguish graphs. For instance, the more "dumbbell" shaped the graph the more strongly peaked the trace of the heat-kernel at the origin. This is due to the fact the spectral gap, i.e. the size of λ_2, determines the rate of decay of the trace with time, and this in turn is a measure of the degree of separation of the graph into strongly connected subgraphs or "clusters".

3 Zeta-Function and Heat-Kernel Trace Moments

The aim in this paper is to use the shape of the heat-kernel trace function as a means of characterising graph-structure. Our characterisation is found by taking moments of trace-function over time.

To commence our development, we consider the zeta function associated with the Laplacian eigenvalues. The zeta function is given by

$$\zeta(s) = \sum_{\lambda_i \neq 0} \lambda_i^{-s} \tag{6}$$

In other words, it is the result of exponentiating and summing the reciprocal of the non-zero Laplacian eigenvalues.
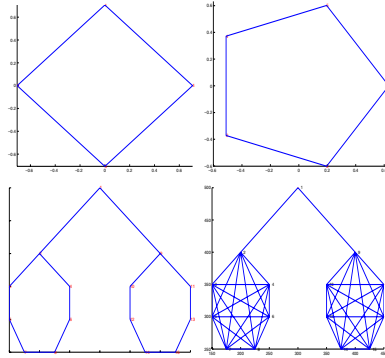
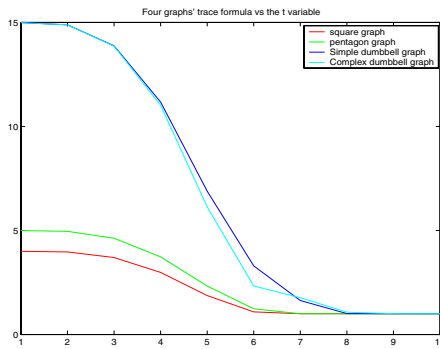**Fig. 1.** Four graphs used for heat-kernel trace analysis

Fig. 2. Heat kernel trace as a function of t for four simple graphs

To establish the link between the zeta function and the trace of the heat-kernel we make use of the Mellin transform

$$\lambda_i^{-s} = \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} \exp[-\lambda_i t] dt \tag{7}$$

where

$$\Gamma(s) = \int_0^\infty t^{s-1} \exp[-t] dt \tag{8}$$

Hence, we can write the zeta function as a moment generating function

$$\zeta(s) = \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} \sum_{\lambda_i \neq 0} \exp[-\lambda_i t] dt \tag{9}$$

The sum of exponentials inside the integral is clearly linked to the trace of the heat-kernel. To show this we make use of the fact that

$$Tr[h_t] = C + \sum_{\lambda_i \neq 0} \exp[-\lambda_i t] \tag{10}$$

where C is the multiplicity of the zero eigenvalue of the Laplacian, or the number of connected components of the graph. Substituting this result back into the Mellin transform, we have

$$\zeta(s) = \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} \left\{ Tr[h_t] - C \right\} dt \tag{11}$$

As a result the zeta function is related to the moments of the heat-kernel trace. It is hence a way of characterising the shape of the heat kernel trace.

4   Experiments

We have experimented with the zeta-function characterisation of the heat-kernel trace. The data used for our study furnished by two data-bases used widely in the object recognition literature, namely the COIL data-base and Oxford-Caltech data-base. For the Coil data-base, we extract the feature points using the method of Harris and Stephens [18]. We have extracted graphs from the images by computing the Voronoi tessellations of the feature-points, and constructing the region adjacency graph, i.e. the Delaunay triangulation, of the Voronoi regions. Figure 3 shows some examples images with the extracted Delaunay graph overlayed for each of the four objects studied. For the Caltech-Oxford data-base, in Figure 4, we use Gestalt relation graphs between line-segments. For each image we extract line-segments using the Canny edge detector and contour polygonalisation. We treat each line-segment as a node in the relation graph. The weights between each pair of nodes are from the relative distance and relative angles attributes between the line-segments. The weighted links between the line segments capture the regular Gestalt-inspired relationships of proximity, parallelism, closure, and continuity [19]. The graphs used in our study are undirected and unattributed.

Both data-sets contain multiple images of either objects of the same class or views of the same object in different poses with respect to the camera. Example images from the data-sets are shown in Figures 3 and 4.

We commence by illustrating the behavior of the zeta-function for the images of objects from COIL data-base. From left-to-right and top-to-bottom in Figure 5 we show
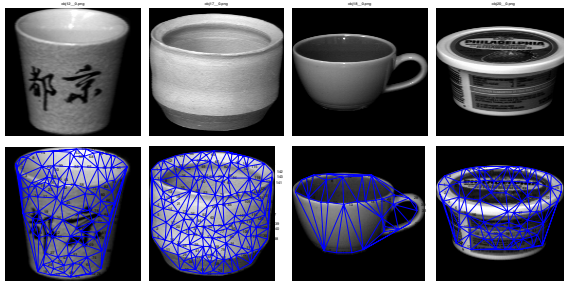

Fig. 3. Example images of four objects from the COIL data-base with their Delaunay graphs overlayed

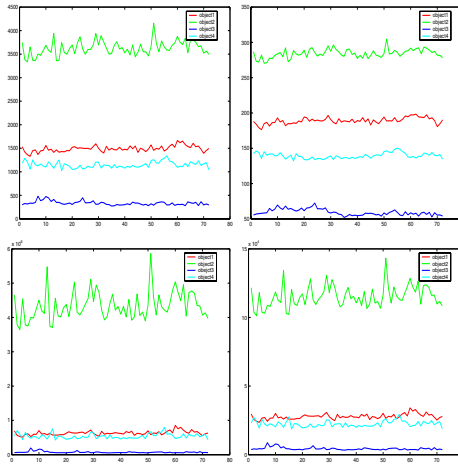**Fig. 4.** Example images from the Oxford-Caltech Database

Fig. 5. Zeta function variation with view number

the values of $\zeta(1)$, $\zeta(2)$, $\zeta(3)$ and $\zeta(4)$ as a function of view-number for the four objects. The different curves in the four plots correspond to the different objects. The main feature to note is that the curves for the different objects are well separated, and that the individual values of the zeta-function do not vary significantly with view number. Moreover the fluctuations in the values are generally smaller than the differences between different objects. This feature is shown more clearly in Figure 6. Here we show the average value of the zeta-function moments as a function of the moment order. The different curves are for different objects. The error-bars show the standard deviation of the moment over the different views (instances) of the same object. The left-hand plot is for the COIL data and the right-hand plot for the Oxford-Caltech data. The moments do not overlap for the different objects, and the best separation is achieved for moments of intermediate order.
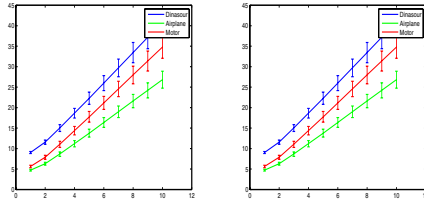
Fig. 6. Zeta function moments as a function or order for the different objects (COIL left and Oxford-Caltech right)
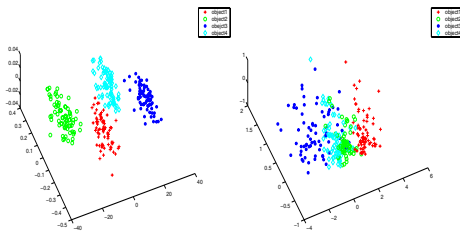

Fig. 7. Zeta-function and Spectral Clustering for the COIL database
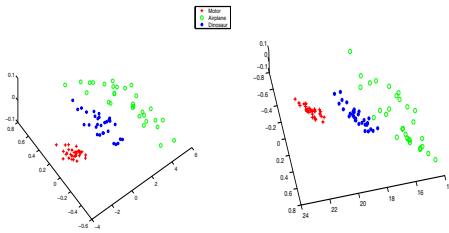

Fig. 8. Zeta-function and Spectral Clustering for the Oxford-Caltech database

In the left-hand panel of Figure 7 we show the result of performing principal components analysis on a feature-vector $f = (\zeta(1), \zeta(2)....., \zeta(10))^T$ which has as its components the zeta-function evaluated at the integers 1 to 10. Here we project the graphs onto the eigenspace spanned by the first three eigenvectors of the feature-vector covariance matrix. The different objects are denoted by points of a different color. The different objects are well separated in the eigenspace. For comparison the right-hand panel in Figure 7 we show the result of repeating this analysis on a vector of leading eigenvalues of the Laplacian matrix $f_\Lambda = (\lambda_1, \lambda_2...., \lambda_{10})^T$. In the eigenspace, the objects are severely overlapped, and the clustering is poorer. In Figure 8 we repeat the analysis of the zeta-function and Laplacian spectrum for the objects from the Caltech-Oxford database. Again, the best clusters are obtained using the zeta-function moments.

5   Conclusions

In this paper we have explored the use of the zeta-function as a means of characterising the shape of the heat-kernel trace for the purposes of graph-clustering. Using the Mellin transform, we have shown that the zeta-function is linked to the moment generating function of the heat-kernel trace. We have experimentally explored the use of the zeta-function as a means of characterising graphs for the purposes of clustering. The method works well on the COIL and Caltech-Oxford data-bases.

References

1. F.R.K.Chung. Spectral Graph Theory *American Mathematical Society*, 1997.
2. J.Shi and J.Malik. Normalized cuts and image segmentation. *IEEE PAMI*, 22:888–905, 2000.
3. J.E.Atkins, E.G.Bowman and B.Hendrickson. A spectral algorithm for seriation and the consecutive ones problem. *SIAM J. Comput.*, 28:297–310, 1998.
4. A.Shokoufandeh, S.Dickinson, K.Siddiqi, and S.Zucker. Indexing using a spectral encoding of topological structure. *IEEE Conf. on Computer Vision and Pattern Recognition*, 491–497, 1999.
5. R.R.Coifman and S.Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 2004.
6. J.Lafferty and G.Lebanon. Diffusion kernels on statistical manifolds. *Technical Reports CMU-CS-04-101*, 2004.
7. S.T.Yau and R.M.Schoen. Differential geometry. *Science Publication*, 1988.
8. P.B.Gilkey. Invariance theory, the heat equation, and the atiyah-singer index theorem. *Perish Inc.*, 1984.
9. Sam Roweis and Lawrence Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326 ,Dec.22, 2000.
10. Bin Luo, Richard C. Wilson and Edwin R. Hancock. Spectral embedding of graphs. Pattern Recognition, 36:2213–2230, 2003.
11. S.Rosenberg. The laplacian on a Riemannian manifold. *Cambridge University Press*, 2002.
12. B.Mohar Laplace eigenvalues of graphs - a Survey. *Discrete Math.*, 109:171–183, 1992.
13. L.Lovaz Random Walks on Graphs: A Survey. *Combinatorics, Paul Erds is eighty*, 2:353–397, 1996.
14. M.Hein, J.Y.Audibert and U.von Luxburg From graphs to manifolds – Weak and strong pointwise consistency of graph laplacian *18th Annual Conference on Learning Theory*, 470–485, 2005.
15. M.Belkin and P. Niyogi, "Laplacian Eigenmaps for Dimensionality Reduction and Data Representation", Neural Computation, **15**, pp. 1373–1396, 2003.
16. J.B. Tenenbaum, V.D. Silva and J.C.Langford, "A global geometric framework for non-linear dimensionality reduction", *Science*, **290**, pp. 586–591, 2000.
17. X.He and P. Niyogi, "Locality preserving projections", *NIPS03*.
18. C.G.Harris, and M.J.Stephens, "A Combined Corner and Edge Detector", *Fourth Alvey Vision Conference*, pp. 147–151, 1994.
19. S.Sarkar, and K.L.Boyer, "Quantitative measures of change based on feature organization: Eigenvalues and eigenvectors. Computer Vision and Image Understanding", *Computer Vision and Image Understanding*, pp. 110–136, 1998.