A study of graph spectra for comparing graphs and trees

Richard C. Wilson*, Ping Zhu

Computer Science Department, University of York, Heslington, York YO10 5DD, UK

ARTICLE INFO

ABSTRACT

The spectrum of a graph has been widely used in graph theory to characterise the properties of a graph and extract information from its structure. It has also been employed as a graph representation for pattern matching since it is invariant to the labelling of the graph. There are, however, a number of potential drawbacks in using the spectrum as a representation of a graph. Firstly, more than one graph may share the same spectrum. It is well known, for example, that very few trees can be uniquely specified by their spectrum. Secondly, the spectrum may change dramatically with a small change structure.

There are a wide variety of graph matrix representations from which the spectrum can be extracted. Among these are the adjacency matrix, combinatorial Laplacian, normalised Laplacian and unsigned Laplacian. Spectra can also be derived from the heat kernel matrix and path length distribution matrix. The choice of matrix representation clearly has a large effect on the suitability of spectrum in a number of pattern recognition tasks.

In this paper we investigate the performance of the spectra as a graph representation in a variety of situations. Firstly, we investigate the cospectrality of the various matrix representations over large graph and tree sets, extending the work of previous authors. We then show that the Euclidean distance between spectra tracks the edit distance between graphs over a wide range of edit costs, and we analyse the accuracy of this relationship. We then use the spectra to both cluster and classify the graphs and demonstrate the effect of the graph matrix formulation on error rates. These results are produced using both synthetic graphs and trees and graphs derived from shape and image data.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

The spectrum of a graph has been widely used in graph theory to characterise the properties of a graph and extract information from its structure. It has also been employed as a graph representation for pattern matching tasks [1–3]. Its use has not gained wide acceptance as a representation for matching and comparison of graphs. There are two main reasons for this; firstly, more than one graph may share the same spectrum. Secondly, the spectrum may change dramatically with a small change structure. While these factors appear to count against the spectrum, they may or may not be important in practical graph matching problems.

Graph structures have been used to represent structural and relational arrangements of entities in many vision problems. Some of the earliest attempts to do so are due to Fischler and Elschlager [4], and Barrow and Burstall [5]. More recently, for example, shock graphs have been used to represent shape [6]. The key problem in utilising graph representations lies in measuring their structural similarity. This is a difficult problem because there is no explicit labelling of the parts, and typically correspondences must be established before similarity can be assessed. There are many methods in the literature which examine the problem of finding correspondences between graphs [7–9]. As an example, Sanfeliu and Fu [10] employed the concept of graph edit distance, giving separate edit costs for relabeling, insertion and deletion on both nodes and edges. A search is necessary to locate the set of operations which have minimal cost. More recently, Bunke [11,12] has established a relationship between the minimum graph edit distance and the size of the maximum common subgraph. The graph edit distance therefore provides a well-defined way of measuring the similarity of two graphs.

Spectral graph theory provides another approach to the problem of graph similarity. This approach is based on a branch of mathematics that is concerned with characterising the structural properties of graphs using the eigenvectors and eigenvalues of the adjacency matrix or the closely related Laplacian matrix (the degree matrix minus the adjacency matrix) [13]. One of the well-known successes of spectral graph theory in computer vision is the use of

* Corresponding author. Tel.: +44 1904 434726.
E-mail address: wilson@cs.york.ac.uk (R.C. Wilson).

eigenvector methods for grouping via pairwise clustering. Examples include Shi and Malik's [14] iterative normalised cut method which uses the Fiedler (i.e. second smallest) eigenvector for image segmentation and Sarkar and Boyer's use of the leading eigenvector of the weighted adjacency matrix [15]. Graph spectral methods have also been used for correspondence analysis. For example, Umeyama's method [16] allows the matching of two graphs of equal size by using the eigendecompositions of the adjacency matrices. Kosinov and Caelli [17] have used properties of the spectral decomposition to represent graphs and Shokoufandeh et al. [1] have used eigenvalues of shock graphs to index shapes. Wilson and Hancock have previously shown [18,19] how permutation invariant polynomials can be used to derive features which describe graphs and make full use of the available spectral information.

The spectrum of a graph (i.e. the set of eigenvalues) is generally considered to be too weak to be a useful tool for representing graphs, mainly due to the result of Schwenk [20] who showed that for trees at least, a sufficiently large tree nearly always has a partner with the same spectrum. Trees therefore cannot be uniquely defined by the spectrum. However, it is not known to what extent this is a problem in practice since Schwenk's work does not reveal how 'sufficiently large' actually it is. The situation for graphs is even less clear, as no similar result is known. Computational simulations by Haemers et al. [21] have shown that the fraction of cospectral graphs reaches 21% at 10 vertices (for the adjacency matrix) and is less for 11 vertices, which is the limit of their simulations. While far from conclusive, their results suggest that for small graphs, the fraction of cospectral graphs is small for some representations, and at least at 11 vertices, the trend is decreasing.

The graph spectrum is derived from a matrix representation of the graph, and is highly dependent on the form of the matrix. A number of alternative matrix representations for graphs have been proposed in the literature; these include the adjacency matrix, Laplacian and normalised Laplacian. The spectrum of all of these representations may be used to characterise the graph, and each may reveal different graph properties. Some of these representations may be more stable to perturbations in the graph. In this paper we analyse these matrices and quantify the effect the matrix representation has on the stability and representational power of the eigenvalues of the graph. We also examine the problem of cospectrality of graphs. In Section 2, we discuss the spectral decomposition of a graph. Section 3 describes the standard matrix representations of graphs. In Section 4, we investigate the cospectrality properties of these matrices with respect to trees and general graphs. In Section 5 we look at the relationship between graph spectra and the edit distance between graphs. Finally, Section 6 details the experiments aimed at measuring the utility of these representations in more practical situations such as clustering and classification.

2. Spectral decomposition of the representation matrix

The graphs under consideration here are undirected, unweighted graphs. While we do not consider weighted graphs here, these ideas are straightforwardly extended to such graphs. We denote a graph by $G = (V, E)$ where V is the set of nodes and $E \subseteq V \times V$ is the set of edges. The degree of a vertex u is the number of edges incident on the vertex u and is denoted d_u. A *matrix representation* of the graph is a $|V|$ by $|V|$ matrix X, such that an element X_{ij} of this matrix represents some property of the pair of vertices i and j. Diagonal elements X_{ii} encode information about the vertex i only. A simple example is the adjacency matrix A, where A_{ij} is 1 when there is an edge between i and j, and zero otherwise. We discuss specific representations in more detail in the next section.

The spectrum of the graph is obtained from the matrix representation using the eigendecomposition. Let $\mathbf{X}$ be the matrix

Fig. 1. A pair of graphs with the same adjacency matrix spectrum [22].

representation in question. Then the eigendecomposition is $\mathbf{X} = \Phi \Lambda \Phi^{\mathrm{T}}$ where $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{|V|})$ is the diagonal matrix with the ordered eigenvalues as elements (ordered in terms of magnitude, with the largest first) and $\Phi = (\phi_1 | \phi_2 | \ldots | \phi_{|V|})$ is the matrix with the ordered eigenvectors as columns. The spectrum is the set of eigenvalues

$$s = \{\lambda_1, \lambda_2, \ldots, \lambda_{|V|}\}$$

with

$$\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_{|V|}$$

One of the key problems with graph comparison is that the vertices are not labelled or ordered. As a result, the vertices may appear in different orders in two graphs under comparison, even if the graphs are isomorphic. Usually, correspondences between the vertices in the two graphs must be established before similarity can be measured. Clearly, the vertex ordering affects the matrix representation of the graph. If $\mathbf{P}$ is a permutation matrix which re-orders the vertices, then

$$\mathbf{L}' = \mathbf{PLP}^{\mathrm{T}}$$

represents the same graph as $\mathbf{L}$.

The spectrum is particularly useful as a graph representation because it is invariant under the similarity transform $\mathbf{PLP}^{\mathrm{T}}$. In other words, two isomorphic graphs will have the same spectrum. This is the motivation for using the spectrum as a graph feature. As noted earlier, the converse is not true, two non-isomorphic graphs may share the same spectrum. Fig. 1 shows an example of two such graphs which have the same adjacency matrix spectrum from [22]. However, the spectrum may be used as an approximate measure of graph similarity. One of the aims of this paper is to establish how useful the spectrum is as such a measure.

The spectral distance between graphs is simply the Euclidean distance between the spectra.

$$d_S(G_1, G_2) = \sqrt{\sum_i (s_i^{(1)} - s_i^{(2)})^2}$$

When the spectra are of different sizes, then the smaller may be padded with zero values (while maintaining the magnitude ordering). This is equivalent to adding disjoint vertices to the smaller graph to make both graphs the same cardinality.

3. Standard graph representations

In this section, we review the properties of some standard graph representations and their relationships with each other.

3.1. Adjacency matrix

The most basic matrix representation of a graph is using the adjacency matrix $\mathbf{A}$ for the graph. This matrix is given by

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

Clearly, if the graph is undirected, the matrix $\mathbf{A}$ is symmetric. As a consequence, the eigenvalues of $\mathbf{A}$ are real. These eigenvalues may be positive, negative or zero and the sum of the eigenvalues is zero. The eigenvalues may be ordered by their magnitude and collected into a vector which describes the graph spectrum.