

The Average Mixing Matrix Signature

Luca Rossi¹, Simone Severini², and Andrea Torsello^{3*}

¹ School of Engineering and Applied Science, Aston University, UK

² Department of Computer Science, University College London, UK

³ DAIS, Università Ca' Foscari Venezia, Italy

Abstract. Laplacian-based descriptors, such as the Heat Kernel Signature and the Wave Kernel Signature, allow one to embed the vertices of a graph onto a vectorial space, and have been successfully used to find the optimal matching between a pair of input graphs. While the HKS uses a heat diffusion process to probe the local structure of a graph, the WKS attempts to do the same through wave propagation. In this paper, we propose an alternative structural descriptor that is based on continuous-time quantum walks. More specifically, we characterise the structure of a graph using its average mixing matrix. The average mixing matrix is a doubly-stochastic matrix that encodes the time-averaged behaviour of a continuous-time quantum walk on the graph. We propose to use the rows of the average mixing matrix for increasing stopping times to develop a novel signature, the Average Mixing Matrix Signature (AMMS). We perform an extensive range of experiments and we show that the proposed signature is robust under structural perturbations of the original graphs and it outperforms both the HKS and WKS when used as a node descriptor in a graph matching task.

Keywords: Graph Characterisation, Structural Descriptor, Quantum Walks, Average Mixing Matrix

1 Introduction

Graph-based representations have been used with considerable success in computer vision in the abstraction and recognition of object shape and scene structure [8, 4, 13]. A fundamental problem in graph-based pattern recognition is that of recovering the set of correspondences (matching) between the vertices of two graphs. In computer vision, graph matching has been applied to a wide range of problems, from object categorisation [4, 1] to action recognition [14, 15]. More formally, in the graph matching problem the goal is to find a mapping between the nodes of two graphs such that the edge structure is preserved.

While there exists a wide range of methods to solve this problem, many graph matching algorithms greatly benefit from the use of local structural descriptors to maximise their performance. To this end, a structural descriptor or signature is assigned to each node of the graphs, effectively embedding the graphs nodes onto

* Authors listed alphabetically.

a vectorial space. Given an underlying correspondence between the node sets of the graphs, the assumption underpinning these approaches is that corresponding nodes will be mapped to points that are close in the signature space.

Two structural signatures that have proven to be particularly successful are the Heat Kernel Signature (HKS) [12] and the Wave Kernel Signature (WKS) [2]. Both signatures belong to the family of Laplacian-based spectral descriptors and were introduced for the analysis of non-rigid three-dimensional shapes. More specifically, they are similarly based on the idea of using the spectrum of the graph Laplacian (and more in general the Laplace-Beltrami operator on the shape surface) to characterise the points on the shape surface. While the Heat Kernel Signature is based on a heat diffusion process, the Wave Kernel Signature is based on the propagation of a wavefunction.

In this paper, we introduce a novel structural signature based on a quantum mechanical process taking place over the graph. In particular, we propose to probe the graph structure using a continuous-time quantum walk [6]. We make use of the average mixing matrix [5] to define a quantum analogue of the heat kernel, and we propose to take the rows of the average mixing matrix at increasing times t as our vertex descriptor. The average mixing matrix was introduced by Godsil, and it describes the time-averaged probability of a quantum walk starting from node x to visit node y at time t . The motivation behind our work is based on the fact that quantum walks have proven to be very successful in characterising graph structures [11, 10, 3, 9]. Moreover, using the average mixing matrix allows us to overcome the lack of convergence of quantum walks. We show that the proposed signature is robust to structural noise and outperforms state-of-the-art signatures in a graph matching task.

The remainder of the paper is organised as follows: Section 2 reviews the necessary quantum mechanical background and Section 3 introduces the proposed signature. The experimental evaluation is presented in Section 4 and Section 5 concludes the paper.

2 Quantum Walks and Average Mixing Matrix

In this section, we introduce the necessary quantum mechanical background. We start by reviewing the definition of continuous-time quantum walks, and then we show how to compute the average mixing matrix of a graph.

2.1 Continuous-Time Quantum Walks

The continuous-time quantum walk represents the quantum analogue of the continuous-time random walk [6]. Let $G = (V, E)$ denote an undirected graph over n nodes, where V is the vertex set and $E \subseteq V \times V$ is the edge set. In a continuous-time random walk, $\mathbf{p}(t) \in \mathbb{R}^n$ denotes the state of a walk at time t . The state vector evolves according to the heat equation

$$\mathbf{p}(t) = e^{-Lt} \mathbf{p}(0), \quad (1)$$

where the graph Laplacian L is the infinitesimal generator matrix of the underlying continuous-time Markov process.

Similarly to its classical counterpart, the state space of the continuous-time quantum walks is the vertex set of the graph. The classical state vector is replaced by a vector of complex amplitudes over V whose squared norm sums to unity, and as such the state of the system is not constrained to lie in a probability space, thus allowing interference to take place. The general state of the walk at time t is a complex linear combination of the basis states $|u\rangle$, i.e.,

$$|\psi(t)\rangle = \sum_{u \in V} \alpha_u(t) |u\rangle, \quad (2)$$

where the amplitude $\alpha_u(t) \in \mathbb{C}$ and $|\psi(t)\rangle \in \mathbb{C}^{|V|}$ are both complex. Moreover, we have that $\alpha_u(t)\alpha_u^*(t)$ gives the probability that at time t the walker is at the vertex u , and thus $\sum_{u \in V} \alpha_u(t)\alpha_u^*(t) = 1$ and $\alpha_u(t)\alpha_u^*(t) \in [0, 1]$, for all $u \in V$, $t \in \mathbb{R}^+$.

The evolution of the walk is governed by the Schrödinger equation

$$\frac{\partial}{\partial t} |\psi(t)\rangle = -iH |\psi(t)\rangle, \quad (3)$$

where we denote the time-independent Hamiltonian as H . Generally speaking, a continuous-time quantum walk is induced whenever the structure of the graphs is reflected by the (0,1) pattern of the Hamiltonian. For example, we could take the adjacency matrix or the Laplacian. In the following we assume $H = A$.

Given an initial state $|\psi(0)\rangle$, solving the Schrödinger equation gives the expression of the state vector at time t ,

$$|\psi(t)\rangle = U(t) |\psi(0)\rangle, \quad (4)$$

where $U(t) = e^{-iAt}$ is the unitary operator governing the temporal evolution of the quantum walk. Eq. 4 can be conveniently expressed in terms of the spectral decomposition of the adjacency matrix $A = \Phi A \Phi^\top$, i.e., $|\psi(t)\rangle = \Phi^\top e^{-iAt} \Phi |\psi(0)\rangle$, where Φ denotes the $n \times n$ matrix $\Phi = (\phi_1 | \phi_2 | \dots | \phi_j | \dots | \phi_n)$ with the ordered eigenvectors ϕ_j s of H as columns and $A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_j, \dots, \lambda_n)$ is the $n \times n$ diagonal matrix with the ordered eigenvalues λ_j of A as elements, and we have made use of the fact that $\exp[-iAt] = \Phi^\top \exp[-iAt] \Phi$.

2.2 Average Mixing Matrix

Given a graph $G = (V, E)$ with adjacency matrix A and its associated unitary operator $U(t)$, the behaviour of a continuous-time quantum walk at time t is captured by the mixing matrix [5]

$$M_G(t) = e^{iAt} \circ e^{-iAt} = U(-t) \circ U(t), \quad (5)$$

where $A \circ B$ denotes the Schur-Hadamard product of two matrices A and B .

Note that, while in the classical case the probability distribution induced by a random walk converges to a steady state, this does not happen in the quantum case. However, we will show that we can enforce convergence by taking a time-average even if $U(t)$ is norm-preserving. Let us define the average mixing matrix [5] at time T as

$$\widehat{M}_{G;T} = \frac{1}{T} \int_0^T M_G(t) dt. \quad (6)$$

The entry uv can be interpreted as the average probability that a walker is found at vertex v when starting the walk from vertex u . Let us define $P_\lambda = \sum_{k=1}^{\mu(\lambda)} \phi_{\lambda,k} \phi_{\lambda,k}^\top$ to be the projection operator on the subspace spanned by the $\mu(\lambda)$ eigenvectors $\phi_{\lambda,k}$ associated with the eigenvalue λ of A . Given this set of projectors, the unitary operator inducing the quantum walk can be rewritten as

$$U(t) = \sum_{\lambda} e^{-i\lambda t} P_\lambda \quad (7)$$

Given Eq. 7, we can rewrite the equation for the mixing matrix as

$$M_G(t) = \sum_{\lambda_1 \in A} \sum_{\lambda_2 \in A} e^{-i(\lambda_1 - \lambda_2)t} P_{\lambda_1} \circ P_{\lambda_2}, \quad (8)$$

and thus we can reformulate Eq. 6 as

$$\widehat{M}_{G;T} = \sum_{\lambda_1 \in A} \sum_{\lambda_2 \in A} P_{\lambda_1} \circ P_{\lambda_2} \frac{1}{T} \int_0^T e^{-i(\lambda_1 - \lambda_2)t} dt, \quad (9)$$

which has solution

$$\widehat{M}_{G;T} = \sum_{\lambda_1 \in A} \sum_{\lambda_2 \in A} P_{\lambda_1} \circ P_{\lambda_2} \frac{i(1 - e^{iT(\lambda_2 - \lambda_1)})}{T(\lambda_2 - \lambda_1)}. \quad (10)$$

In the limit $T \rightarrow \infty$, Eq. 10 becomes

$$\widehat{M}_{G;\infty} = \sum_{\lambda \in \tilde{A}} P_\lambda \circ P_\lambda \quad (11)$$

where \tilde{A} is the set of distinct eigenvalues of the adjacency matrix.

2.3 Properties of the Average Mixing Matrix

Given a graph $G = (V, E)$, let $\widehat{M}_{G;\infty}(i, j)$ denote the element of $\widehat{M}_{G;\infty}$ at the i th row and j th column. Recall that two nodes $u, v \in V$ are strongly cospectral if and only if $P_\lambda e_u = \pm P_\lambda e_v$ for all eigenvalues λ , where e_u is an element of the standard basis of $\mathbb{R}^{|V|}$. It can be shown that the following properties hold for the matrix $\widehat{M}_{G;\infty}$:

1. $\widehat{M}_{G;\infty}$ is doubly stochastic and rational, and that $\widehat{M}_{G;\infty}$ is positive semidefinite whenever G is connected [5];
2. Let G and H be cospectral, regular, and switching equivalent graphs. Then $\widehat{M}_{G;\infty}(j, j) = \widehat{M}_{H;\infty}(j, j)$, for every $j = 1, 2, \dots, n$;
3. Let G and H be cospectral, walk-regular graphs. Then $\widehat{M}_{G;\infty}(j, j) = \widehat{M}_{H;\infty}(j, j)$, for every $j = 1, 2, \dots, n$;
4. two vertices i and j are strongly cospectral if and only if the two corresponding rows of the average mixing matrix are the same, i.e., $\widehat{M}_{G;\infty}(i, k) = \widehat{M}_{G;\infty}(j, k)$, for each k [5];
5. suppose G is a regular and connected graph and \overline{G} is the complementary graph. Then $\widehat{M}_{G;\infty} = \widehat{M}_{\overline{G};\infty}$ if and only if \overline{G} is connected.

We omit the proof of 2, 3 and 5 for lack of space, while the proofs of 1 and 4 can be found in [5]. Note that 2, 3, 4 and 5 outline the conditions under which some diagonal element of the mixing matrix of can be duplicated, the rows of two mixing matrices are identical, and two mixing matrices have exactly the same elements, respectively. In other words, the above properties give an important indication of the limitations of the matrix $\widehat{M}_{G;\infty}$ to discriminate between different graphs. Although the same properties have not been proved for an arbitrary stopping time T , our experimental analysis in Section 4 suggests that some of these properties may still hold for $T < \infty$.

3 The Average Mixing Matrix Signature

Given a graph $G = (V, E)$ and a vertex $x \in V$, we define the Average Mixing Matrix Signature (AMMS) at x at time T as

$$AMMS(x, T) = \text{sort}(\widehat{M}_{G;T}(x, -)), \quad (12)$$

where $\widehat{M}_{G;T}(x, -)$ denotes the row of the average mixing matrix corresponding to x , and sort is a sorting function. Given a time interval $[T_{min}, T_{max}]$, the AMMS is then created by concatenating the sorted rows for every $T \in [T_{min}, T_{max}]$.

As Eq.12 shows, for each stopping time T , we decide to take the whole row of the average mixing matrix rather than just the diagonal element $\widehat{M}_{G;T}(x, x)$. Recall from Section 2 that under some particular conditions two graphs can have the same diagonal entries of $\widehat{M}_{G;\infty}$. Although the same has not been proved for an arbitrary T , we take the entire row in an attempt to maximise the discriminatory power of the signature. Finally, to ensure the permutational invariance of the signature, for each T we sort the elements of the average mixing matrix rows.

Note that, while there is no specific reason to take all the elements of the sorted row rather than simply the first k , in the experimental evaluation we show that the best performance is usually achieved for large values of k , so we propose to take the entire row. Finally, from Eq. 10 we observe that the computational complexity of the proposed signature is $O(|V|^4)$.

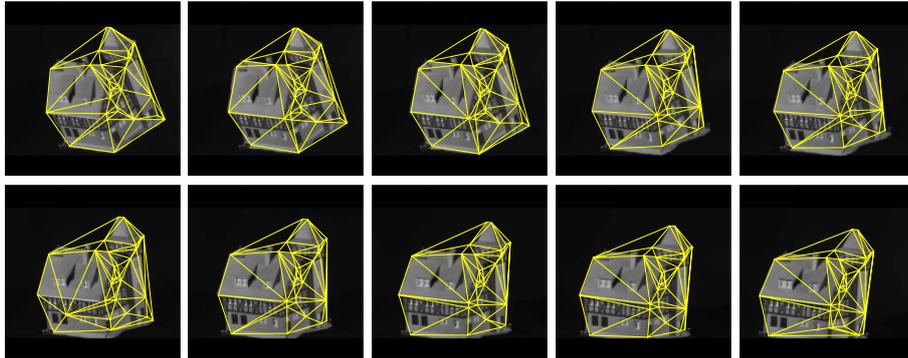


Fig. 1. CMU house sequence with the feature points and the Delaunay graphs superimposed.

4 Experimental Evaluation

To evaluate the proposed signature, we compare it on a graph matching problem against the HKS as well as the WKS. In the following, we refer to the percentage of correctly matched nodes as accuracy.

4.1 HKS and WKS

Heat Kernel Signature The HKS is based on the analysis of a heat diffusion process on the shape governed by the heat equation

$$-\frac{\partial u(x, t)}{\partial t} = \Delta u(x, t), \quad (13)$$

where Δ is the Laplace-Beltrami operator. Assuming at time $t = 0$ all the heat is concentrated at a point x , it can be shown that the amount of heat that remains at x at time t is

$$HKS(x, t) = \sum_{i=0}^{\infty} e^{-\lambda_i t} \phi_i(x)^2 \quad (14)$$

where λ_i and ϕ_i denote the i th eigenvalue and the i th eigenfunction of the Laplace-Beltrami operator, respectively. The HKS at a point x is then constructed by evaluating Eq. 14 for different $t \in [t_{min}, t_{max}]$. As suggested in [12], here we sample 100 points logarithmically over the time interval defined by $t_{min} = 4 \ln(10) / \lambda_{max}$ and $t_{max} = 4 \ln(10) / \lambda_{min}$, where λ_{min} and λ_{max} denote the minimum and maximum eigenvalue of the graph Laplacian, respectively.

Wave Kernel Signature The WKS is based on the analysis of wavefunction evolution on the shape surface governed by the Schrödinger equation

$$-\frac{\partial \psi(x, t)}{\partial t} = i \Delta \psi(x, t). \quad (15)$$

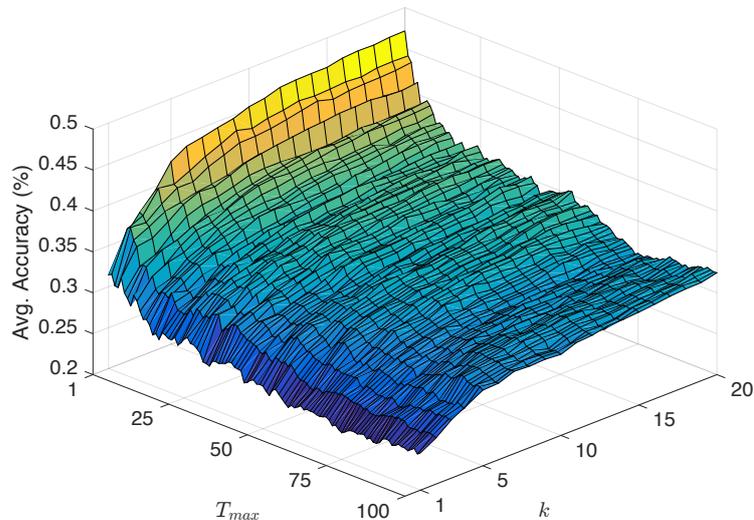


Fig. 2. Evaluation of the sensitivity to the parameters values. Here k denotes the number of row elements used to construct the signature and T_{max} is the maximum stopping time considered.

Although similar in the definition, note that the dynamics behind the WKS (oscillation) and the HKS (dissipation) are fundamentally different. The WKS evaluates the expected probability of measuring a quantum particle in the point x of the shape surface at any time, i.e.,

$$WKS(x, e) = \sum_{i=0}^{\infty} \phi_i(x)^2 f_e(\lambda_i)^2, \quad (16)$$

where λ_i and ϕ_i denote the i th eigenvalue and the i th eigenfunction of the Laplace-Beltrami operator, respectively, and f_e is a log-normal energy distribution that depends on a parameters e . The WKS for a point x is then constructed by evaluating Eq. 16 for different values of $e \in [e_{min}, e_{max}]$. Here we set all the parameters of the WKS as illustrated in [2].

4.2 Datasets

We use a dataset made of the images from the CMU house sequence, where each image is abstracted as a Delaunay graph over a set of corner feature points. All the resulting graphs are composed of 30 to 32 nodes. Fig. 1 shows the ten images with the feature points and the graphs superimposed.

In addition, we consider a synthetic datasets consisting of 10 Delaunay graphs over 20 nodes. This dataset is generated as follows. We create a first graph by generating 20 computing the Delaunay triangulation of 20 uniformly randomly scattered two-dimensional points. Then, we generate 9 additional graphs by slightly perturbing the original points and recomputing the Delaunay triangulation.

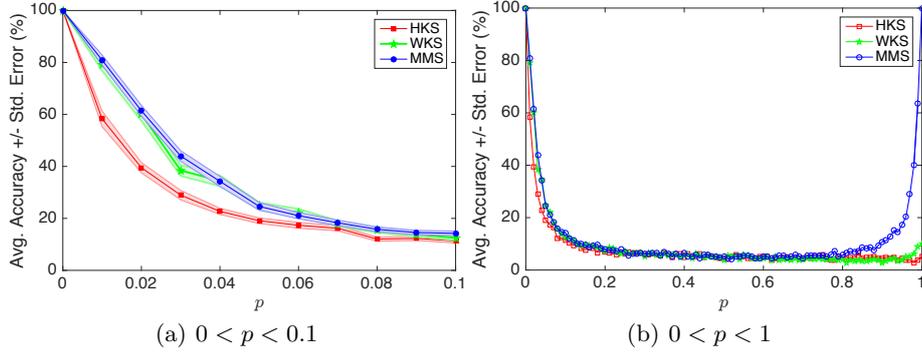


Fig. 3. Average accuracy for increasing levels of Erdős-Rényi noise. Here p denotes the probability of adding or removing an edge.

4.3 Graph Matching Experimental Setup

For a pair of input graphs, we compute the AMMS for each graph node. Here we decide to sample 10 points logarithmically over the time interval $[0.1, 1]$. The choice of $T_{max} = 1$ is experimentally motivated in the next subsection. In general, we observe that on all the datasets considered in this paper the best performance is achieved for low values of T_{max} . Given the nodes signatures, we then compute the matrix of pairwise Euclidean distances between the signatures. Finally, we cast the graph matching problem as an assignment problem which can be solved in polynomial time using the well-known Hungarian method [7].

4.4 Sensitivity to Parameters

We first investigate the sensitivity of the proposed signatures to the number k of row elements that we consider for each stopping time $T \in [T_{min}, T_{max}]$, as well the value of T_{max} . Starting from a seed graph G from the synthetic dataset, we generate 100 noisy copies G' by removing or adding an edge with probability $p = 0.03$. Then, we compute the matching for each pair (G, G') and each choice of the parameters k and T_{max} , and we plot the average accuracy in Fig. 2. Here we let $k = 1, 2, \dots, n$, where n is the number of nodes of G , and we let $T_{max} = 1, 2, \dots, 100$. Fig. 2 shows that the best accuracy is achieved for k larger than 10 and $T_{max} = 1$.

4.5 Robustness to Erdős-Rényi Structural Noise

In order to evaluate the robustness of the AMMS to structural noise, we perform the following experiment. Given a graph G from the synthetic dataset, we create a series of noisy versions of G , for increasing amounts of Erdős-Rényi noise. To this end, we flip an edge of the graph with probability p and obtain the noisy graph G' . For each value of p , we repeat the noise addition process 100 times to

Dataset	CMU House	Synthetic
HKS	30.2963 ± 3.8879	14 ± 1.3162
WKS	44.3704 ± 4.1975	14.5556 ± 1.7467
AMMS(<i>diag</i>)	43.3333 ± 3.8461	16.3333 ± 1.6314
AMMS(<i>row</i>)	52.5185 ± 3.8170	19.4444 ± 2.1095

Table 1. Average accuracy (\pm standard error).

obtain 100 noisy copies of G . Then, we compute the optimal matching using the proposed signature and we compare the results against those obtained using the HKS and the WKS. Fig. 3(a) shows the average accuracy for increasing levels of structural noise. Note that the AMMS and the WKS are significantly less susceptible to noise than the HKS. It is also interesting to observe that as p grows over 0.8, the number of correctly matched nodes for the AMMS start to increase (see Fig. 3(b)). We posit that this is evidence that the average mixing matrix of a graph G is equivalent or very similar to that of its complement \overline{G} also for arbitrary $T < \infty$. Note that for $T \rightarrow \infty$ this is true if G is regular and connected, and \overline{G} is connected as well. Although this is not the case for the graphs used in these experiments, the geometric nature of the Delaunay triangulation yields indeed fairly regular graphs.

4.6 Graph Matching Accuracy

Table 4.6 shows the average accuracy on the CMU house and the synthetic datasets. We consider two alternative versions of the proposed signatures: 1) AMMS(*diag*) is the signature where only the diagonal elements $\widehat{M}_{G;T}(x, x)$ are used; 2) AMMS(*row*) is the signature where all the rows elements are used. In both the datasets considered, our signature performs significantly better than both the HKS and the WKS. Moreover, the AMMS(*row*) always outperforms the AMMS(*diag*), highlighting the importance of using all the elements of the row and suggesting once again that the properties listed in Section 2 hold for an arbitrary T . Finally, we note that the AMMS(*diag*) always outperforms the HKS but not the WKS.

5 Conclusion

In this paper, we have introduced a novel structural signature for graphs. We probed the structure of the graph using continuous-time quantum walks and we proposed to consider the rows of the average mixing matrix for increasing stopping times T as the node descriptor. The experimental results show that this signature can outperform state-of-the-art signatures like the Heat Kernel Signature and the Wave Kernel Signature in a graph matching task. Future work should focus on the optimisation of the signature parameters for general graphs, as well as reducing the computational complexity associated with the Average

Mixing Matrix Signature. Recall that we observed that the best performance is usually achieved for low values of T_{max} . Therefore, one idea may be to reduce the computational complexity of the proposed signature by resorting to a linear approximation of its values.

Acknowledgments. Simone Severini was supported by the Royal Society and EPSRC.

References

1. Albarelli, A., Bergamasco, F., Rossi, L., Vascon, S., Torsello, A.: A stable graph-based representation for object recognition through high-order matching. In: Pattern Recognition (ICPR), 2012 21st International Conference on. pp. 3341–3344. IEEE (2012)
2. Aubry, M., Schlickewei, U., Cremers, D.: The wave kernel signature: A quantum mechanical approach to shape analysis. In: Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on. pp. 1626–1633. IEEE (2011)
3. Bai, L., Rossi, L., Torsello, A., Hancock, E.R.: A quantum jensen–shannon graph kernel for unattributed graphs. *Pattern Recognition* 48(2), 344–355 (2015)
4. Duchenne, O., Joulin, A., Ponce, J.: A graph-matching kernel for object categorization. In: Computer Vision (ICCV), 2011 IEEE International Conference on. pp. 1792–1799. IEEE (2011)
5. Godsil, C.: Average mixing of continuous quantum walks. *Journal of Combinatorial Theory, Series A* 120(7), 1649–1662 (2013)
6. Kempe, J.: Quantum random walks: an introductory overview. *Contemporary Physics* 44(4), 307–327 (2003)
7. Munkres, J.: Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial and Applied Mathematics* 5(1), 32–38 (1957)
8. Riesen, K., Bunke, H.: Graph classification and clustering based on vector space embedding. World Scientific Publishing Co., Inc. (2010)
9. Rossi, L., Torsello, A., Hancock, E.R.: Measuring graph similarity through continuous-time quantum walks and the quantum jensen–shannon divergence. *Physical Review E* 91(2), 022815 (2015)
10. Rossi, L., Torsello, A., Hancock, E.R., Wilson, R.C.: Characterizing graph symmetries through quantum jensen–shannon divergence. *Physical Review E* 88(3), 032806 (2013)
11. Suau, P., Hancock, E.R., Escolano, F.: Analysis of the schrödinger operator in the context of graph characterization. In: *Similarity-Based Pattern Recognition*, pp. 190–203. Springer (2013)
12. Sun, J., Ovsjanikov, M., Guibas, L.: A concise and provably informative multi-scale signature based on heat diffusion. In: *Computer graphics forum*. vol. 28, pp. 1383–1392. Wiley Online Library (2009)
13. Torsello, A., Rossi, L.: Supervised learning of graph structure. In: *Similarity-Based Pattern Recognition*, pp. 117–132. Springer (2011)
14. Yao, B., Fei-Fei, L.: Action recognition with exemplar based 2.5 d graph matching. In: *Computer Vision–ECCV 2012*, pp. 173–186. Springer (2012)
15. Yi, Y., Lin, M.: Human action recognition with graph-based multiple-instance learning. *Pattern Recognition* (2015)