

3D Scene Comparison using Topological Graphs

L. Paraboschi and S. Biasotti and B. Falcidieno

Istituto di Matematica Applicata e Tecnologie Informatiche - Sez. di Genova
Consiglio Nazionale delle Ricerche
{laurap,silvia,bianca}@ge.imati.cnr.it

Abstract

New technologies for shape acquisition and rendering of digital shapes have simplified the process of creating virtual scenes; nonetheless, shape annotation, recognition and manipulation of both the complete virtual scenes and even of subparts of them are still open problems.

In this paper we deal with the problem of comparing two (or more) object sets, where each model is represented by an attributed graph. We will define a new distance to estimate the possible similarities among the sets of graphs and will validate our work using the shape graph [BGSF06].

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Line and Curve Generation

1. Introduction

Object recognition is one of the main tasks of Computer Vision and Graphics, and graph theory has been trying to give insights for that purpose. Identifying an object with a skeleton, we obtain a *description* ([FS98]), which is a synthetic representation able to recover object topology and, if attributed, some of its geometric features (see [BGSF06]). Once every object has been represented as a graph, we need a tool able to evaluate the possible similarity among the elements: a distance measure is an effective solution to this problem. If the distance is a metric, it verifies well stated mathematical properties: identity discriminates an entity from another one, symmetry guarantees measure reciprocity, positivity says that we cannot obtain a distance equal to 0 as a cancellation, while triangle inequality helps ranking a database without calculating the distance between every couple of descriptions.

In this contribution we deal with the problem of comparing two or more scenes, that we define as sets of objects. Since every object is described by a graph we introduce a new distance between two sets of graphs that will be used to evaluate the scene similarity. Then we propose two different normalizations of this distance: the first one is more suitable to compare two scenes, while the second one is mainly suitable for shape retrieval.

The remainder of this paper is organized as follows: in the next section we provide an overview on the main approaches to the graph matching problem. In Section 3 we describe the method that inspired our work, proposing our extension and introducing two new distances between sets of graphs; finally, some simple examples are described in Section 4 highlighting experimental properties of the method.

2. Related work

In the last decades several approaches have been developed for graph matching, and especially in defining distances between them.

The matching problem is solvable in polynomial time when dealing with trees, a subclass of graphs that are maximally acyclic and minimally connected (see Theorem 1.5.1 [Die05]). In the large literature related to tree matching an interesting contribution has been proposed by Torsello et al. [THR05], where the authors present four novel distance measures (metrics) for attributed trees based on the notion of a maximum similarity subtree isomorphism. Since many problems deal with graphs instead of trees, a possible solution is to reduce the graph representation into an attributed tree obtained through editing operations (removing, adding or replacing a vertex or an edge). Unfortunately these tech-

niques discard a lot of information about the original structure of the object and the tree we get is non-unique.

As far as graph matching is concerned, several approaches are based on maximum common subgraph *MCS* and minimum common supergraph *mcs*, like those proposed by Bunke [BS98] and Fernández and Valiente [FV01]. In both cases the distances proposed are metrics between non-attributed graphs, therefore the attributes associated to graphs are not taken into account. Since the problem of finding the *MCS* or the *mcs* is NP-complete, some authors have proposed algorithms for approximating their computation (see [MSF05]), obviously providing a lower bound of the graph distance.

Recently, the need for having flexible graph matching frameworks that admit the mapping of many nodes into many others has led to the definition of the so-called “many-to-many” approach; actually, nowadays there are still a few works that face this problem. For example Demirci et al. [DSK*06] present an approximation algorithm for the many-to-many graph matching of two graphs. They begin by constructing a tree metric representation for graphs. Next, they embed them in a geometric space with low distortion using a spatial encoding of the graph vertices. Then they translate their problem into a many-to-many geometric points matching task, for which the Earth Mover’s Distance algorithm (see [RTG00]) is well suited.

Finally, a large family of methods is based on linear algebra techniques, exploiting the adjacency, incident or Laplacian matrix of a graph. For example, Shokoufandeh et al. [SMD*05] present a signature of the topology of a directed acyclic graph that provides an effective mechanism for indexing large databases of graphs. In this case the signature is based on a combination of the spectral properties of the graph underlying adjacency matrix. Another spectral method for graph matching is described by Robles-Kelly and Hancock [RKH01]: it makes use of a brushfire search procedure using the rank-order of the co-efficients of the leading eigenvector of the adjacency matrix. The search procedure starts from the node of the largest co-efficient and it proceeds via first-order neighbourhoods to assign correspondences on the basis of local rank order. Similarly, the method proposed by Wilson et al. [WHL05] analyses the spectrum of the Laplacian matrix of the graph; since we have identified this approach as the most appropriate for our purposes, it is briefly described in section 3.1. Another method has been proposed by Bapat [Bap05], where the author introduces the concept of a tree with attached graphs and then he defines a distance matrix for that original tree.

3. Comparing scenes

The problem of comparing two scenes can be faced comparing singularly their components; obviously this kind of approach needs a number of comparisons which is quadratic

in the total numbers of components. Especially when dealing with complex scenes, it could be useful and helpful having a method able to decide rapidly if the scenes are compatible, judging if they are overallly similar. Moreover, to simplify the information necessary for comparisons, we can imagine that every object is represented by a description, for example an attributed graph. In this last case, the comparison of two scenes is reduced to the subproblem of comparing two sets of graphs. We have identified a graph matching technique which is efficient, adaptable and easily extendible to our problem. Moreover, we propose the use of two new distances able to evaluate the similarity between the two original scenes.

3.1. Graph matching based on symmetric polynomials

Wilson et al. [WHL05] show how graphs can be converted into pattern vectors by using the spectral decomposition of the Laplacian matrix and basis sets of symmetric polynomials.

A graph $G \in \mathcal{G}$ is usually defined as a couple (V, E) , where \mathcal{G} is a set of graphs, V is the node set ($n = |V|$), and E the edge set. If a function $w : V(\text{or } E) \rightarrow \mathbf{R}$ is given, G is said to be attributed. Denoting L the Laplacian matrix of a graph G , let \mathbf{e}_i be an eigenvector (suppose $\|\mathbf{e}_i\| = 1 \forall i$) and λ_i be its eigenvalue. If we define $\Phi = (\Phi_{i,j})_{i,j=1,\dots,n} = (\sqrt{\lambda_1}\mathbf{e}_1, \dots, \sqrt{\lambda_n}\mathbf{e}_n)$, then $L = \Phi\Phi^t$, see [WHL05] for further details.

Graph topology is invariant for any permutation of node labels: in our context it means that, denoting P a permutation matrix, L and PLP^t have the same eigenvalues; therefore any permutation of the node labels does not alter the spectral representation of the graph. In particular,

$$\lambda_j = \sum_{i=1}^n \Phi_{ij}^2 . \quad (1)$$

Equation (1) is a symmetric polynomial in the components of eigenvectors \mathbf{e}_i . To measure the invariant features of the graphs, the authors propose to consider the set of elementary symmetric polynomials

$$S_j(v_1, \dots, v_n) = \sum_{i_1 < \dots < i_j} v_{i_1} v_{i_2} \dots v_{i_j} , \quad j = 1, \dots, n \quad (2)$$

A matrix $F = (f_{i,j})_{i,j=1,\dots,n}$, $f_{i,j} = \text{sign}(S_j(\Phi_{1,i}, \dots, \Phi_{n,i})) \ln(1 + |S_j(\Phi_{1,i}, \dots, \Phi_{n,i})|)$ is introduced to define the so-called **feature vector** of G :

$$\mathbf{B} = (f_{1,1}, \dots, f_{1,n}, \dots, f_{n,1}, \dots, f_{n,n})^t .$$

Then, given two graphs $G_1, G_2 \in \mathcal{G}$ whose feature vectors \mathbf{B}_i , $i = 1, 2$, are known, a metric on \mathcal{G} is given by

$$d(G_1, G_2) = \|\mathbf{B}_1 - \mathbf{B}_2\| . \quad (3)$$

3.2. Extension to sets of graphs

To extend the previous technique to graph set matching, we introduce in every set of graphs $\mathbf{G} = \{G_i \in \mathcal{G}\}_{i=1,\dots,m}$

($|V_i| = n_i$) a virtual node X without attributes, which is joined to a vertex of degree 1 for each component G_i of the set (see Figure 1). Since the Laplacian spectrum is invariant to node permutations, among all the nodes of degree 1 the choice of the vertex to which X is joined is not relevant for the extraction of the eigenvalues of L .

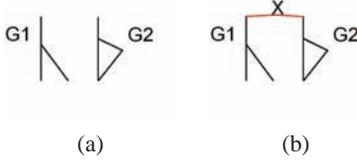


Figure 1: (a): two graph components, (b): the resulting scene graph

Therefore we obtain a single graph whose Laplacian matrix has the form shown in (4). In particular, if the set is composed by m graphs, we obtain

$$L = \begin{pmatrix} m & -1 & \mathbf{0} & \dots & -1 & \mathbf{0} \\ -1 & l_{1,1}^1 + 1 & l_{1,n_1}^1 & & & \\ \mathbf{0} & l_{n_1,1}^1 & l_{n_1,n_1}^1 & & & \\ \vdots & & & \ddots & & \\ -1 & & & & l_{1,1}^m + 1 & l_{1,n_m}^m \\ \mathbf{0} & & & & l_{n_m,1}^m & l_{n_m,n_m}^m \end{pmatrix}, \quad (4)$$

where L is the Laplacian matrix of a connected graph, specifically the *scene graph*, while $L(G_i) = L^i = (l_{k,j}^i)_{k,j=1,\dots,n_m}$ is the Laplacian matrix of the i -th graph. Note the element added to $l_{1,1}^i \forall i$: it stresses the existence of the virtual node joined with a component vertex, whose degree increases of 1.

When dealing with attributed graphs, it is possible to use an expression of L given by the authors [WHL05] which takes into account even node or edge attributes. In our case the attributed Laplacian matrix of the shape graph is modified according to the procedure described above.

3.3. A new pseudo-metric

In order to have a general idea about the possible similarity among the scene graphs, we introduce a new distance

$$D(G_1, G_2) := \left| \|\mathbf{B}_1\| - \|\mathbf{B}_2\| \right|, \quad (5)$$

and we call it **global distance**. D is a pseudo-metric: it satisfies positivity, symmetry and triangle inequality; identity is not verified ($D(G_1, G_2) = 0 \not\Rightarrow G_1 \simeq G_2$). We notice that D underestimates d in (3), in fact the following relation holds:

$$D(G_1, G_2) \leq d(G_1, G_2).$$

In our experiment we use a normalized version of (5):

$$D_N(G_1, G_2) = \frac{\left| \|\mathbf{B}_1\| - \|\mathbf{B}_2\| \right|}{\max(\|\mathbf{B}_1\|, \|\mathbf{B}_2\|)}. \quad (6)$$

D_N is a pseudo-semi-metric, since it satisfies neither identity nor the triangle inequality. It is appropriate if just two sets of graphs are compared, and it gives the proportional difference between them. When we compare two scenes, we first get their graphs and then estimate the distance in (6): the smaller D_N is, the more the two scenes are supposed similar.

Introducing $d_N(G_1, G_2) = \frac{\|\mathbf{B}_1 - \mathbf{B}_2\|}{2 \max(\|\mathbf{B}_1\|, \|\mathbf{B}_2\|)}$ as a normalization of the distance d , it follows that

$$D_N(G_1, G_2) \leq 2d_N(G_1, G_2).$$

We finally propose another possible normalization of D , which is more suitable to arrange a database with m sets of graphs. Let it be $\|\mathbf{B}_i\| = \max_{i=1,\dots,m} \|\mathbf{B}_i\|$; then

$$D_m(G_1, G_2) = \frac{\left| \|\mathbf{B}_1\| - \|\mathbf{B}_2\| \right|}{\|\mathbf{B}_i\|}.$$

We remark that D_m is a pseudo-metric, that is it does not satisfy identity.

4. Examples and discussions

We have evaluated our method using closed triangle meshes, such as, for examples, tables, chairs, cups, teddy bears and humans. The models are taken from various public databases: the AIM@SHAPE repository (<http://shapes.aim-at-shape.net>), the CAESAR Data Samples (<http://www.hec.afri.af.mil/HECP/Card1b.shtml#caesar>), and the McGill 3D Shape Benchmark (<http://www.cim.mcgill.ca/~shape/benchMark/>).

4.1. The graph representation

The construction of the graph used in our framework relies on the discretization of the Reeb graph theory defined by Bisotti [Bia04]. Given a shape represented by a regular triangle mesh M , we subdivide the co-domain $[f_{min}, f_{max}]$ of $f : M \rightarrow \mathbf{R}$ considering nv regular values of f , $f_i \in [f_{min}, f_{max}]$, $i = 1, \dots, nv$. The level sets of f corresponding to these values partition the mesh M into regions (see Figure 2(b)). Hence all points belonging to a region of a contour are identified and represented as nodes and edges of traditional graph (see Figure 2(c,d)).

Four different mapping functions f are considered in our framework, namely the distance from the barycenter, the height function (with respect to z axis), the integral geodesic distance, originally proposed by Hilaga et al. [HSKK01], and the minimum distance from a source point ([LV99]).

For every node $v \in V$, corresponding to a region R , it is possible to associate a property characterizing the region R

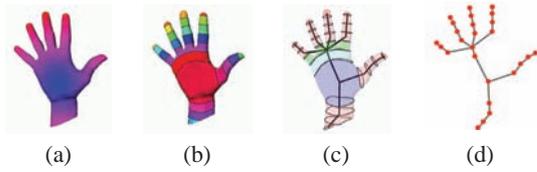


Figure 2: (a): Evaluation of the distance from the barycenter, (b): the mesh partition, (c-d): the skeleton.

or its boundary B_R ; in particular we use the minimum, maximum and average distance of the barycenter of R from the region vertices, the sum of the pseudo-cone lateral areas for each component of R , the sum of B_R lengths and the value of f in every vertex in V . Details on this representation can be found in the paper of Biasotti et al. [BGSF06].

4.2. Experimental results

Let us consider two scenes with a man and a teddy bear (Figure 3): they look really similar, in particular the man has open arms in both scenes. Since the objects are disposed similarly, the height function could be a good compromise between representation and computational efficiency.

	
Scene 1	Scene 2
0.0122	minimum distance
0.0053	maximum distance
0.0226	average distance
0.0091	sum of the pseudo-cone areas
0.0267	sum of B_R lengths
0.0462	value of f
0.0204	average distance

Figure 3: Scenes with a man and a teddy bear: distance values when using height function

The results in Figure 3 are consistent with our perception: in fact the biggest distance is 4.6%, which is a low value for D_N . On the contrary, if we chose the height function with respect to another axis (for example, x), we would not observe similarity between the two scenes: it means in fact that we must be very careful about the choice of f . Similarly, comparing some chairs analysed again through the height function we obtain the average distances in Figure 4. The underlined values refer to the fallen chair: as we can imagine, these distances are bigger than the others, because the used function is sensitive to the position of the object in the reference space.

Consider the scenes in Figure 5: in the first scene, there is a woman sitting on a chair in front of a table, on which

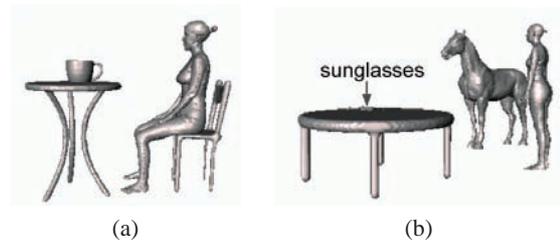


Figure 5: (a): scene 1, and (b): scene 2

there is a cup, while in the second one there are a table and a pair of spectacles on it, close to a woman and a horse. Let us analyze what we get when comparing them through their graphs (extracted through the barycentre distance function, every attribute):

0.8058 0.7864 0.7454 0.0074 0.7509 0.9519.

These are all high values (apart from the fourth), and they indicate that globally the two scenes are not similar; in fact, we do not generally see similarity between them. If we compare the table and the cup in (a) with the table and the spectacles in (b) we get, with respect, for example, to the fifth attribute (the sum of B_R lengths), $D_N = 0.3523$, which refers to non-similar subscenes. Instead, if we analyze together the table and the woman in (a) with the table and the horse in (b), we get

0.0936 0.0905 0.0956 0.0324 0.1125 0.0916.

These are low values, and they indicate a so-called “false positive”, since the subscenes are not semantically similar, but structural (see also Figure 8).

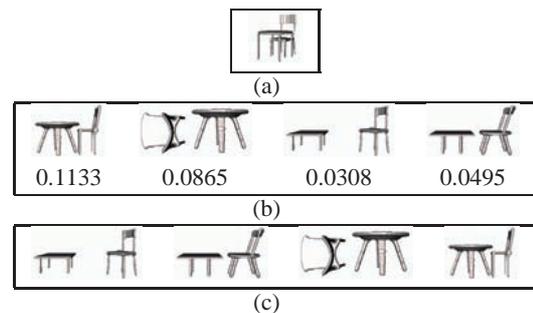


Figure 6: (a) Query, (b) the database and the distances from Q , and (c) the database ordered according to the distance

As a further example, we show a retrieval result (Figure 6), where the family of scenes with a chair and a table is re-ordered with respect to the query Q (attribute: average distance of the barycentre from the region vertices). Since the distances we obtain are those in Figure 6(b), then ordering the database with respect to D_N we get Figure 6(c). This

						
	0	0.1959	<u>0.1217</u>	0.1016	0.1293	0.2392
		0	<u>0.2536</u>	0.1057	0.0763	0.0526
			0	<u>0.1854</u>	<u>0.1969</u>	<u>0.2699</u>
				0	0.0344	0.1536
					0	0.1261
						0

Figure 4: Scenes with one chair

is a consistent result, because, for example, in the third and fourth scene there is the same table, which is quite different, being thicker than the other ones.

Let us now observe the four scenes in Figure 7: again, they are composed by the same kind of models, however we can expect a bigger similarity between the third and the fourth, in which the man has open arms. For every function we report the average distances : height (h), barycenter distance (b), the minimum distance from a source point (mds) and the integral geodetic distance (igd).

With respect to every function, the biggest similarity is between the third and the fourth scene (double-marked square); moreover, leaving igd results out, each distance remains below 13%. As established a posteriori, it seems to be a result of similarity. This example shows us that the integral geodetic distance cannot guarantee a reliable result, and it is not so suited for our purpose.

Finally, in Figure 8 we show the distances between a chair and a horse whose graphs are obtained through barycenter distance with respect to every attribute and then the average distance. This last value is 0.094, which could indicate similarity. Obviously a chair is not similar to a horse, while their graphs are comparable: it is an example of a so called “false positive”.

4.3. Properties

We now conclude describing some particular features of the method.

	
	0.0987 1 st attribute
	0.0982 2 nd attribute
	0.0907 3 rd attribute
	0.1516 4 th attribute
	0.0963 5 th attribute
	0.0282 6 th attribute
	0.0940 global average

Figure 8: A false positive result

Stability: as shown in Table 1, $\|\mathbf{B}\|$, and consequently D_N , is well-conditioned.

Considering for example a node attribute 0.000755: the table shows how the distance changes while adding noise to the weight. The smaller the noise, the lower the distance: the method is stable with respect to the attribute noise.

Moreover, there is also stability with respect to small model perturbations: if we consider one of our models, specifically a human body, and then we perturb it with increasing Laplacian smoothing (Table 2) or simplifying its triangulation (Figure 9) using the tool developed by Attene [Att06], we obtain the distance of the original model from those perturbed is always lower than 3%. It indicates strong similarity, as expected.

Robustness: an inherent numerical error appears in computing symmetric polynomials in (2), that is

	0	0.1028(h) 0.0478(b) 0.0571(mds) 0.2026(igd)	0.0232(h) 0.0817(b) 0.0774(mds) 0.3662(igd)	0.0177(h) 0.0977(b) 0.0903(mds) 0.4292(igd)
		0	0.1236(h) 0.0355(b) 0.0715(mds) 0.2120(igd)	0.1187(h) 0.0523(b) 0.0531(mds) 0.2894(igd)
			0	0.0111(h) 0.0175(b) 0.0413(mds) 0.1448(igd)
				0

Figure 7: Scenes with a man and a table

Table 1: Stability with respect to noise: G' is the graph G after the perturbation of an attribute

Attribute noise	Attribute	Distance $D_N(G, G')$
$+10^{-1}$	0.000755	$0.5541 \cdot 10^{-2}$
$+10^{-2}$	0.010755	$0.5367 \cdot 10^{-3}$
$+10^{-3}$	0.001755	$0.5357 \cdot 10^{-4}$
$+10^{-4}$	0.000855	$0.5357 \cdot 10^{-5}$
$+10^{-5}$	0.000765	$0.5357 \cdot 10^{-6}$
$+10^{-6}$	0.000756	$0.5357 \cdot 10^{-7}$

Table 2: Stability with respect to Laplacian smoothing: (a) original model, (b) distance from (a)

Laplacian smoothing	Distance D_N
1 iteration	0.0090
2 iterations	0.0095
3 iterations	0.0105
4 iterations	0.0110
5 iterations	0.0198
7 iterations	0.0204

$S_j(\Phi_{1,i}, \dots, \Phi_{n,i}) = \lambda_i^{\frac{1}{2}} \sum_{k_1 < \dots < k_j} V_{k_1,i} V_{k_2,i} \dots V_{k_j,i}$. Actually it is typical when working with Laplacian matrices that the most significant information on graphs are related to the first eigenvalues: a common practice is to eliminate some of the last eigenvalues. In our context, to guarantee coherent results, it is necessary discarding some of them: since the last ones are bigger than the previous (see Figure 10), a numerical error appears and distorts results. Since the growth of the matrix spectrum has been shown to be almost linear ([DBG*06]), eigenvalues can be automatically

		10%	34%	50%
	0	0.0160	0.0162	0.0161
10%		0	0.0003	0.0005
34%			0	0.0003
50%				0

Figure 9: Simplified meshes (10%, 34%, 50%): D_N is evaluated with respect to barycenter distance function. The results refer to the average distance

removed by analysing the increase of the spectrum and by discarding the values that diverge from linearity. For example, referring to Figure 10, it is appropriate to discard the last three eigenvalues.

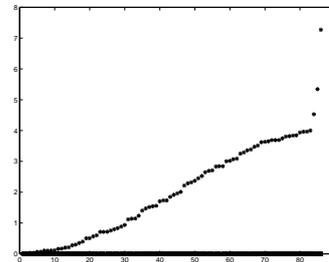


Figure 10: Spectrum of a Laplacian matrix

Acknowledgements

The authors thank the EU Newtwork of Excellence "AIM@SHAPE" and the project 4 at IMATI-GE/CNR.

References

- [Att06] ATTENE M.: *An interactive and user-friendly environment for remeshing surface triangulations*. Shape Modelling Group IMATI GE CNR, <http://www.ima.ge.cnr.it/ima/smg/resources.html>, July 2006.
- [Bap05] BAPAT R. B.: Distance matrix and Laplacian of a tree with attached graphs. *Linear Algebra and its Applications* 411 (2005), 295–308.
- [BGSF06] BIASOTTI S., GIORGI D., SPAGNUOLO M., FALCIDIENO B.: Size functions for 3D shape retrieval. In *EG SGP '06* (2006), pp. 239–242.
- [Bia04] BIASOTTI S.: *Computational Topology Methods for Shape Modelling Applications*. PhD thesis, Università degli Studi di Genova, May 2004.
- [BS98] BUNKE H., SHEARER K.: A graph distance metric based on the maximal common subgraph. *Pattern Recognition Letters* 19 (1998), 255–259.
- [DBG*06] DONG S., BREMER P.-T., GARLAND M., PASCUCCI V., HART J. C.: Spectral surface quadrangulation. *ACM Trans. Graph.* 25, 3 (2006), 1057–1066.
- [Die05] DIESTEL R.: *Graph Theory*, electronic version of the third edition ed. Springer - Verlag Heidelberg, New York, 2005.
- [DSK*06] DEMIRCI M. F., SHOKOUFANDEH A., KESELMAN Y., BRETZNER L., DICKINSON S.: Object recognition as many-to-many feature matching. *IJCV* 69, 2 (August 2006), 203–222.
- [FS98] FALCIDIENO B., SPAGNUOLO M.: A shape abstraction paradigm for modeling geometry and semantics. In *CGI '98* (1998), pp. 646–657.
- [FV01] FERNÁNDEZ M.-L., VALIENTE G.: A graph distance metric combining maximum common subgraph and minimum common supergraph. *Pattern Recognition Letters* 22 (2001), 753–758.
- [HSKK01] HILAGA M., SHINAGAWA Y., KOHMURA T., KUNII T. L.: Topology matching for fully automatic similarity estimation of 3D shapes. In *ACM Computer Graphics, (Proc. SIGGRAPH 2001)* (Los Angeles, CA, August 2001), ACM Press, pp. 203–212.
- [LV99] LAZARUS F., VERROUST A.: Level set diagrams of polyhedral objects. In *SMA '99: Proceedings of the 5th ACM Symposium on Solid Modeling and Applications 1999* (1999), Bronsvoort W., Anderson D., (Eds.), ACM Press, pp. 130–140.
- [MSF05] MARINI S., SPAGNUOLO M., FALCIDIENO B.: From exact to approximate maximum common subgraph. In *GbR 2005* (Poitiers (France), April 11-13 2005), LNCS, pp. 263–272.
- [RKH01] ROBLES-KELLY A., HANCOCK E. R.: Graph matching using adjacency matrix Markov chains. In *BMVC* (University of Manchester, Manchester, UK, September 2001), pp. 383–390.
- [RTG00] RUBNER Y., TOMASI C., GUIBAS L. J.: The earth mover's distance as a metric for image retrieval. *IJCV* 40 (2000), 99–121.
- [SMD*05] SHOKOUFANDEH A., MACRINI D., DICKINSON S., SIDDIQI K., ZUCKER S. W.: Indexing hierarchical structures using graph spectra. *IEEE PAMI* 27, 7 (2005), 1125–1140.
- [THRP05] TORSELLO A., HIDOVIC-ROWE D., PELILLO M.: Polynomial-time metrics for attributed trees. *IEEE PAMI* 27, 7 (2005), 1087–1099.
- [WHL05] WILSON R. C., HANCOCK E. R., LUO B.: Pattern vectors from algebraic graph theory. *IEEE PAMI* 27, 7 (2005), 1112–1124.