Shape Classification Based on Skeleton Path Similarity

Xingwei Yang[1], Xiang Bai[2], Deguang Yu[1], and Longin Jan Latecki[1]

[1] Dept. of Computer and Information Sciences, Temple University, 1805 North Broad Street, Philadelphia, PA 19122, USA
{xingwei.yang,deguang,latecki}@temple.edu
[2] Dept. of Electronics and Information Engineering, HuaZhong University of Science and Technology, Wuhan, Hubei 430074, China
xiang.bai@gmail.com

Abstract. Most of the traditional methods for shape classification are based on contour. They often encounter difficulties when dealing with classes that have large nonlinear variability, especially when the variability is structural or due to articulation. It is well-known that shape representation based on skeletons is superior to contour based representation in such situations. However, approaches to shape similarity based on skeletons suffer from the instability of skeletons and matching of skeleton graphs is still an open problem. Using a skeleton pruning method, we are able to obtain stable pruned skeletons even in the presence of significant contour distortions. In contrast to most existing methods, it does not require converting of skeleton graphs to trees and it does not require any graph editing. We represent each shape as set of shortest paths in the skeleton between pairs of skeleton endpoints. Shape classification is done with Bayesian classifier. We present excellent classification results for complete shape.

Keywords: Skeleton pruning, skeleton path, Bayesian Classification.

1 Introduction

An important goal in image analysis is to classify and recognize objects. They can be characterized in several ways, using color, texture, shape, movement, and location. Shape, as a significant factor of objects, is an important research direction in image classification and recognition. Shape of planar objects can be described based on their contours or on skeletons.

When utilizing contours in classification and recognition, shape classes that have a large nonlinear variability of global shape, due to structural variation, articulation, or other factors, present a challenge for several existing shape recognition approaches. Approaches that match the target shape to stored example shapes require a large number of stored examples to capture the range of variability [1]. Furthermore, existing example-based and model-based approaches cannot handle object classes that have different parts or numbers of parts without splitting the class into separate subclasses. This type of structural variation can be handled by approaches that

A.L. Yuille et al. (Eds.): EMMCVPR 2007, LNCS 4679, pp. 375–386, 2007.

represent part relationships explicitly and match shapes syntactically; however, these structural approaches are computationally expensive [2].

On the other hand, skeleton (or medial axis), which integrates geometrical and topological features of the object, is an important shape descriptor for object recognition [4]. Shape similarity based on skeleton matching usually performs better than contour or other shape descriptors in the presence of partial occlusion and articulation of parts [5][6][7][1]. However, it is a challenging task to automatically recognize the objects using their skeletons due to skeleton sensitivity to boundary deformation [8]. Usually the skeleton branches have to be pruned for recognition [8][9][10][11]. Moreover, another major restriction of recognition methods based on skeleton is a complex structure of obtained tree or graph representations of the skeletons. Graph edit operations are applied to the tree or graph structures, such as merge and cut operations [12][13][14][15][16], in the course of the matching process. Probably the most important challenge for skeleton similarity is the fact that the topological structure of skeleton trees or graphs of similar object may be completely different. Besides, some methods [21] have focused on utilizing geometry measures to gauge the similarity of 2D shapes by comparing their skeletons. This fact is illustrated in Fig. 1. Although the skeletons of the two horses (a) and (b) are similar, their skeleton graphs (c) and (d) are very different. This example illustrates the difficulties faced by approaches based on graph edit operations in the context of skeleton matching. To match skeleton graphs or skeleton trees like the ones shown in Fig. 1, some nontrivial edit operations (cut, merge, et al.) are inevitable.

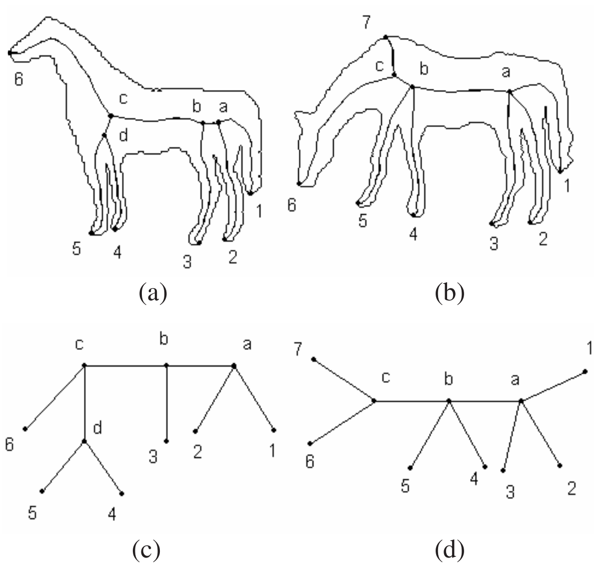


Fig. 1. Visually similar shapes in (a) and (b) have very different skeleton graphs in (c) and (d)

On the other hand, skeleton graphs of different objects may have the same topology as shown in Fig. 2. The skeletons of the brush in Fig. 2(a) and the pliers in Fig. 2(b) have the same topology as shown in Fig. 2(c).

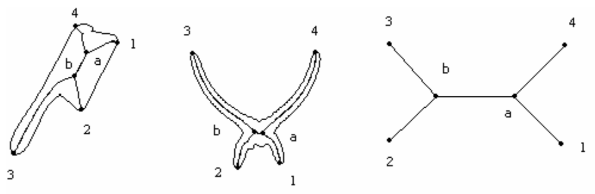


Fig. 2. Dissimilar shapes in (a) and (b) can have the same skeleton graphs (c)

The proposed method combines Bayesian classifier and a novel skeleton representation that overcomes the above limitations. This paper utilizes a three-level statistical framework including distinct models for dataset, class, and part. Bayesian inference is used to perform classification within this framework. Based on Bayes rule, the posterior probabilities of classes can be computed by the difference between skeletons of query shape and the shape in dataset. In the proposed framework, it can work well to classify complete shapes.

In section 2, the background of the related method will be discussed. The way to obtain and represent skeletons is introduced in section 3 and 4. The Bayesian framework is given in the section 5. In section 6, experimental results and analysis on two different datasets have been given. At last, conclusion and future work are drawn out.

2 Background

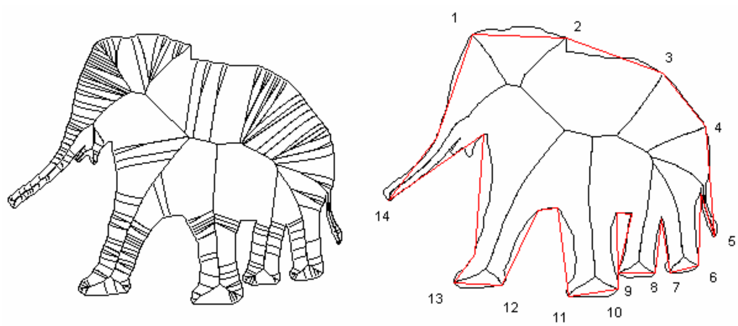
This section briefly introduces some recent methods developed for shape matching, including classification, detection, and retrieval. A number of approaches are based on the contour. Belongie et al. [1] proposed the concept of 'shape context', which are log-polar histograms among different points on the shape. Through finding the correspondence between points on different shapes, this approach can get the similarity between the shapes. Some methods used boosting to classify objects. Bar-Hillel, et al. [17] designed a classifier based on a part-based, generative object model. The approach given by Opelt, et al. [18] developed a novel learning algorithm which uses Adaboost to learn the shape features. Besides the learning algorithm, Gorelick et al. [19] used the Poisson Equation to extract various shape properties for shape classification. Tu and Yuille presented an algorithm for shape matching and recognition based on a generative model for how one shape can be generated by the other [26]. Sun and Super [3] used distribution of contour parts in known object classes to classify shapes with Bayesian classifier. Their classification works only for complete query shapes.

In contrast to the methods based on contour, many researchers have worked on the approaches based on skeleton. Zhu et al. matched the skeleton graphs of objects using a branch-bounding method that was limited to motionless objects [12]. Shock graph was a kind of ARG proposed by Siddiqi and Dickinson et al [24][25]., which was based on the shock Grammar. The distance between subgraphs was measured by comparing the eigenvalues of their adjacency matrices. Besides the methods for shape similarity based on skeleton, a few approaches implemented the skeleton in classification. Sebastian. et al [23] discussed an indexing technique on shock graph, Shokoufandeh, A et al [22]describe a framework for indexing such representations that embeds the topological structure of a directed acyclic graph (DAG) into a low-dimensional vector space The main reason for this is that the past methods have high complexity. The proposed method defines a novel approach to classify the shape. The main difference between the proposed method and other methods is it utilizes the skeleton path into the Bayesian framework, which has never been discussed before. The results are very promising and the complexity of the proposed method is much lower than current methods, such as shock graph [2].

3 Obtain Skeleton by Skeleton Pruning

Any topology preserving method can be used to compute skeletons. We used the method by Choi et al. [9].

The limitation of the skeleton is that it is sensitive to the boundary deformation and the noise. Therefore, it is difficult to obtain the ideal skeletons to recognize the objects. In order to solve this problem, this method utilizes skeleton pruning introduced in [10] to improve the skeleton. First, Discrete Curve Evolution (DCE) simplifies the polygon. Then the skeleton is pruned so that only branches ending at the DCE vertices remain. For example, in Fig. 3 (a) the skeleton contains a lot of noise. In Fig. 3 (b) all the endpoints (denoted by 1, 2, …, 14) of the elephant's skeleton are vertices of the DCE simplified polygon (in red). The pruned skeleton is guaranteed to preserve the topology of the shape and it is robust to noise and boundary deformation [10]. Moreover, the skeleton endpoints are guaranteed to lie on the object contour.



(a) (b)

Fig. 3. (a) The original input skeleton. (b) The skeleton pruned with contour partitioning [10].

4 Shape Representation with Skeleton Paths

The endpoint in the skeleton graph is called an end node, and the junction point in the skeleton graph is called a junction node. The shortest path between a pair of end nodes on a skeleton graph is called a skeleton path. We show a few example skeleton paths in Fig 4.

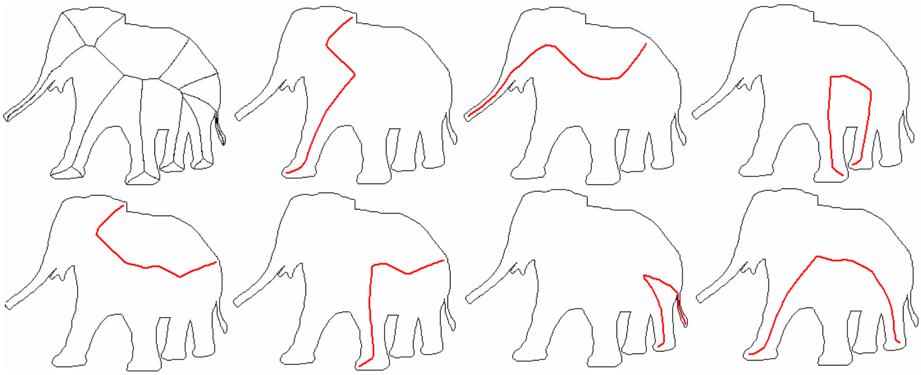


Fig. 4. The elephant's skeleton and the shortest paths (in red) between the pairs of endpoints

The shortest paths between every pair of skeleton endpoints are represented as sequences of radii of the maximal disks at corresponding skeleton points.

Suppose there are N end nodes in the skeleton graph G to be matched, and let v_i (i = 1, 2, …, N) denote the ith end node along the shape contour in the clockwise direction. Let sp(m, n) denote the skeleton path from v_m to v_n. We sample sp(m, n) with M equidistant points, which are all skeleton points. Let $R_{m,n}(t)$ denote the radius of the maximal disk at the skeleton point with index t of in sp(m, n). Let $R_{m,n}$ denote a vector of the radii of the maximal disks centered at the M sample skeleton points on sp(m, n):

$$R_{m,n}=(R_{m,n}(t))_{t=1,\dots,M}=(r_1,r_2,\dots,r_M) \tag{1}$$

In this paper, the radius $R_{m,n}(s)$ is approximated with the values of the distance transform DT(s) at each skeleton point s. Suppose there are N_0 pixels in the original shape S. To make the proposed method invariant to the scale, we normalize $R_{m,n}(s)$ in the following way:

$$R_{m,n}(s)=\frac{DT(s)}{\frac{1}{N_0}\sum_{i=1}^{N_0}DT(s_i)} \tag{2}$$

where s_i (i=1, 2, …, N_0) varies over all N_0 pixels in the shape.

The shape dissimilarity between two skeleton paths is called a path distance. If R and R' denote the vectors of radii of two skeleton paths sp and sp' respectively, the path distance pd between sp and sp' is:

$$pd(R,R')=\sum_{i=1}^{M}\frac{(r_i-r_i^{'})^2}{\left|r_i+r_i^{'}\right|} \tag{3}$$

5 Bayesian Classification

Compared to the method in [3], which uses contour segments and Bayesian classification to perform a recognition task, our method uses paths instead of contour segments. The basic idea is very simple, similar shape should have similar paths. Therefore, the difference of paths between similar shapes should be small. This Bayesian framework can obtain the classification by summing the difference of query shape's skeleton paths to the all of the shapes' skeleton paths in the same class. The smaller the difference is, the more possible the query shape belongs to the class.

Given a shape ω' that should be classified by Bayesian Classifier, we build the skeleton graph G(ω') of ω' and input G(ω')as the query. For a skeleton graph G(ω'), if the number of end nodes is n, the corresponding number of paths is n(n-1) compared to the number of parts n! in [3]. Then, the Bayesian Classifier computes the posterior probability of all classes for each path sp'∈G(ω'). By accumulating the posterior probability of all of the paths of G(ω'), the system automatically yields the ranking of class hypothesis.

If two different paths have small pd value, the value of probability should be large. Otherwise, it should be small. Therefore, we use Gaussian distribution to compute the probability p:

$$p(sp^{'} \mid sp) = \frac{1}{\sqrt{2\pi\alpha}} \exp(-\frac{pd(sp^{'}, sp)^2}{2\alpha}) \tag{4}$$

For different datasets, the α should be different. In our experiments, for the dataset of Aslan and Tari [20], α=0.15 and α=0.05 for Kimia dataset [2].

The class-conditional probability for observing sp' given that ω' belongs to class c_i is:

$$p(sp^{'} \mid c_i) = \sum_{sp \in G(c_i)} p(sp^{'} \mid sp) p(sp \mid c_i) \tag{5}$$

We assume that all paths within a class path set are equiprobable, therefore

$$p(sp \mid c_i) = 1/\left|G(c_i)\right| \tag{6}$$

c_i is one of the M classes.

The posterior probability of a class given that path sp'∈G(ω') is determined by Bayes rule:

$$p(c_i \mid sp^{'}) = \frac{p(sp^{'} \mid c_i) p(c_i)}{p(sp^{'})} \tag{7}$$

Similar to the above assumption, $p(c_i)=1/M$. The probability of sp' is equal to

$$p(sp^{'}) = \sum_{i=1}^{M} p(sp^{'} \mid c_i) p(c_i) \tag{8}$$

Through the above formulas, we can get the posterior probability of all paths of G(ω'). By summing the posterior probabilities of a class over the set of paths in the input shape, we obtain the probability that the input shape belongs to a given class. Obviously, the biggest one, Cm, is the class that input shape belongs to.

$$C_m = \underset{i=1,\dots,M}{\arg\max} \sum_{sp' \in G(\omega')} p(c_i \mid sp') \tag{9}$$

6 Experiments

In this section, we evaluate the performance of the proposed method based on the dataset of Aslan and Tari [20]. We selected this dataset due to large variations of shapes in the same classes. As shown in Fig. 5, Aslan and Tari dataset includes 14 classes of articulated shapes with 4 shapes in each class. We use each shape in this dataset as a query, and show the classification result of our system in Fig. 6. We used leave one out classification, i.e., the query shape was excluded from its class.

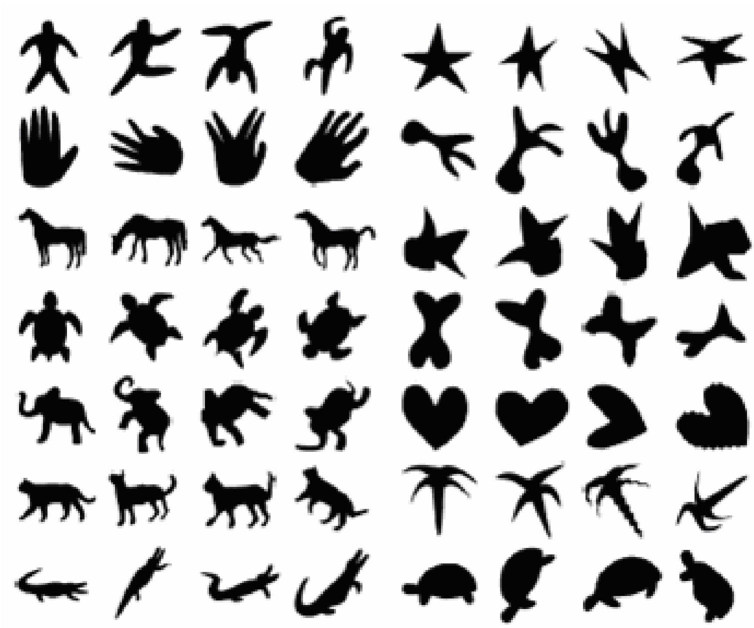


Fig. 5. Aslan and Tari dataset [20] with 56 shapes

The table in Fig. 6 is composed of 14 rows and 9 columns. The first column of the table represents the class of each row. For each row, there are four experimental results which belong to the same class. Each experimental result has two elements. The first one is the query shape and the second one is the classification result of our system. If the result is correct, it should be the equal to the first column of the row. The red numbers mark the wrong classes assigned to query objects. Since there is only one error in 56 classification results, the classification accuracy in percentage by this measure is 98.2%. In fact, the only error is reasonable. Even a human can misclassify it. The query shape is very similar to star, the class 8. Therefore, in some sense, we can conclude that all of our results are correct.

class	query	result	query	result	query	result	query	result
1		1		1		11		9
2		2		2		2		2
3		3		3		6		3
4		4		4		4		4
5		5		5		5		5
6		6		6		6		6
7		7		7		7		7
8		8		8		8		8
9		9		9		9		9
10		10		10		10		10
11		11		11		11		11
12		12		12		12		12
13		13		8		13		13
14		14		14		14		14

Fig. 6. Results of the proposed method on Aslan and Tari dataset [20]. Since each class is composed of 4 shapes, the class of query and the result should be the same. Red numbers mark the results where this is not the case.

class	query	result	query	result	query	result	query	result
1		1		1		11		9
2		2		2		2		2
3		3		3		6		3
4		4		4		4		4
5		5		5		5		5
6		6		6		6		6
7		7		7		7		7
8		8		8		8		8
9		9		9		9		9
10		10		10		10		10
11		11		11		11		11
12		12		12		12		12
13		13		8		13		13
14		14		14		14		14

Fig. 7. Results of the Sun and Super's method on Aslan and Tari dataset [20]. Since each class is composed of 4 shapes, the class of query and the result should be the same. Red numbers mark the results where this is not the case.

We compared our method to the method presented by Sun and Super in [3], their method uses the same Bayesian classifier but is based on contour parts. As shown in Fig.7, their method yields 4 wrong results for 56 query shapes, so the accuracy is only 92.8%.

Moreover, the classification time for all 56 shapes with the proposed method takes only 5 minutes on the PC with 1.5 GHZ CPU and 512M RAM. However, Sun and Super's method takes 13 minutes on the same computer.

We also apply the proposed method to Kimia dataset [23] as shown in Fig. 8, which includes 18 classes, and each class is consisted of 12 shapes. In each experiment, we remove the query shape from the dataset Therefore there are 215 shapes in dataset and one query shape. Since there are only 12 errors in 216 classification results, the classification accuracy in percentage is 94.4%, which is comparable to Sun and Super's result [3]. Though the accuracy of Sebastian et al [23] on the dataset is 100 percent which is better than the proposed method, the proposed method is still promising. The classification time for all 216 shapes with the proposed method takes only nearly 25 minutes on the PC with 1.5 GHZ CPU and 512M RAM.

Fig. 8. Eighteen classes in Kimia dataset [23]

In Fig.9, we just give out 2 correct experimental results for each class and the last four images are chosen from the 12 error classifications, the wrong classification results are in red. The reason for the first wrong classification is the skeleton of head of the bird is similar to the same part of camel. For the glass, the skeleton is similar to the end of the bone. Moreover, the turtle is misclassified to elephant, as the tail part is like the same part of elephant. The Misk is like the brick in some sense, therefore the misclassification is reasonable.

Moreover, based on the classification results, the proposed method is rotation and scale invariant. In the experiment of the first dataset, Aslan and Tari dataset [21], the shapes have been rotated but the results are still correct. For the dataset of kimia [23], the size of the shape in the same class is different from each other. The proposed method can still obtain over 94 percent accuracy.

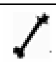
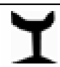
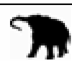
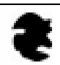
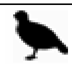
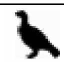
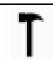
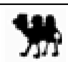
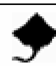
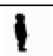
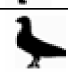
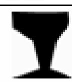
Query	Result	Query	Result	Query	Result	Query	Result
	Bone		Bone		Classic car		Classic car
	Glass		Glass		elephant		elephant
	heart		heart		face		face
	Misk		Misk		fork		fork
	Bird		Bird		fountain		fountain
	Brick		Brick		hammer		hammer
	camel		camel		key		key
	car		car		ray		ray
	child		child		turtle		turtle
	camel		bone		elephant		brick

Fig. 9. Part of the classification results on Kimia dataset

7 Conclusions

In this paper, we propose a novel method to classify the whole shape that is based on statistics of dissimilarities between shortest skeleton paths. Compared to the shock graph, we use the radius distance instead of the topology of skeleton to measure the similarity between two shapes. As the proposed method need not find corresponding between different skeleton paths, It avoids complex discussion on finding corresponding between two skeletons. Moreover, the result of two different datasets demonstrated that skeleton paths are very efficient shape representation for classification. However, as the probability of two paths is calculated based on the radius difference between two paths, if one shape's radiuses are totally different from other shapes in its class, the system may misclassify it to other classes. It is the drawback of the proposed method compared to the shock graph. In the future, our work will focus on solving this kind of problems and implementing the classification method in the part classification.

References

1. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE Trans. PAMI 24(4), 509–522 (2002)
2. Sebastian, T., Klein, P., Kimia, B.: Shock-Based Indexing into Large Shape Databases. In: Tistarelli, M., Bigun, J., Jain, A.K. (eds.) ECCV 2002. LNCS, vol. 2359, pp. 731–746. Springer, Heidelberg (2002)
3. Sun, K.B., Super, B.J.: Classification of Contour Shapes Using Class Segment Sets. In: CVPR, pp. 727–733 (2005)
4. Blum, H.: Biological Shape and Visual Science. J. Theoretical Biology 38, 205–287 (1973)
5. Sebastian, T.B., Kimia, B.B.: Curves vs skeletons in object recognition. Signal Processing 85, 247–263 (2005)
6. Basri, R., Costa, L., Geiger, D., Jacobs, D.: Determining the Similarity of Deformable Shapes. Vision Research 38, 2365–2385 (1998)
7. Huttenlocher, D.P., Klanderman, G.A., Rucklidge, W.J.: Comparing Images Using the Hausdorff Distance. IEEE Trans. Pattern Analysis and Machine Intelligence 15(9), 850–863 (1993)
8. Shaked, D., Bruckstein, A.M.: Pruning Medial Axes. Computer Vision and Image Understanding 69(2), 156–169 (1998)
9. Choi, W.-P., Lam, K.-M., Siu, W.-C.: Extraction of the Euclidean skeleton based on a connectivity criterion. Pattern Recognition 36(3), 721–729 (2003)
10. Bai, X., Latecki, L.J., Liu, W.-Y.: Skeleton Pruning by Contour Partitioning with Discrete Curve Evolution. IEEE Trans. Pattern Analysis and Machine Intelligence 29(3), 449–462 (2007)
11. Ogniewicz, R.L., Kubler, O.: Hierarchic voronoi skeletons. Pattern Recognition 28, 343–359 (1995)
12. Zhu, S.C., Yuille, A.L.: FORMS: A flexible object recognition and modeling system. Int. J. Computer Vision (IJCV) 20(3), 187–212 (1996)
13. Liu, T., Geiger, D.: Approximate Tree Matching and Shape Similarity. In: Proc. Int. Conf. Computer Vision (ICCV), pp. 456–462 (1999)
14. Geiger, D., Liu, T., Kohn, R.V.: Representation and self-similarity of shapes. IEEE Trans. Pattern Anal. Mach. Intell. 25(1) (2003)
15. Di Ruberto, C.: Recognition of shapes by attributed skeletal graphs. Pattern Recognition 37(1), 21–31 (2004)
16. Pelillo, M., Siddiqi, K., Zucker, S.W.: Matching hierarchical structures using association graphs. IEEE Trans. Pattern Anal. Mach. Intell. 21(11), 1105–1120 (1999)
17. BarHill, A., Hertz, T., Weinshall, D.: Object class recognition by boosting a part based model. CVPR (2005)
18. Fussenegger, M., Opelt, A., Pinz, A., Auer, P.: Object Recognition Using Segmentation for Feature Detection. In: Proc. Int. Conf. Pattern Recognition (ICPR), pp. 41–44 (2004)

19. Gorelick, L., Galun, M., Sharon, E., Basri, R., Brandt, A.: Shape Representation and Classification Using the Poisson Equation. IEEE Trans. Pattern Analysis and Machine Intelligence 28(12), 1991–2005 (2006)
20. Aslan, C., Tari, S.: An Axis Based Representation for Recognition. In: ICCV, pp. 1339–1346 (2005)
21. Torsello, A., Hancock, E.R.: A skeletal measure of 2D shape similarity. Computer Vision and Image Understanding 95(1), 1–29 (2004)
22. Shokoufandeh, A., Macrini, D., Dickinson, S., Siddiqi, K., Zucker, S.W.: Indexing hierarchical structures using graph spectra. IEEE Trans. Pattern Analysis and Machine Intelligence 27(7), 1125–1140 (2005)
23. Sebastian, T.B., Klein, P.N., Kimia, B.B.: Recognition of shapes by editing their shock graphs. IEEE Trans. Pattern Analysis and Machine Intelligence 26(5), 550–571 (2004)
24. Siddiqi, K., Shokoufandeh, A., Dickenson, S.J., Zucker, S.W.: Shock graphs and shape matching. ICCV, 222–229 (1998)
25. Siddiqi, K., Shokoufandeh, A., Dickinson, S., Zucker, S.: Shock graphs and shape matching. International Journal of Computer Vision 30, 1–24 (1999)
26. Tu, Z., Yuille, A.L.: Shape Matching and Recognition Using Generative Models and Informative Features. ECCV, 195–209 (2004)