# Coarse-To-Fine Matching of Shapes using Disconnected Skeletons by Learning Class-Specific Boundary Deformations

Aykut Erdem[1] and Sibel Tari[2]

[1] Dipartimento di Informatica, Universitá Ca' Foscari di Venezia
Via Torino 155, Mestre, Venezia, 30172, Italy
erdem@dsi.unive.it

[2] Department of Computer Engineering, Middle East Technical University
Inonu Bulvari, 06531, Ankara, Turkey
stari@metu.edu.tr

**Abstract.** *Disconnected skeleton* [1] is a very coarse yet a very stable skeleton-based representation scheme for generic shape recognition in which recognition is performed mainly based on the structure of disconnection points of extracted branches, without explicitly using information about boundary details [2, 3]. However, sometimes sensitivity to boundary details may be required in order to achieve the goal of recognition. In this study, we first present a simple way to enrich disconnected skeletons with radius functions. Next, we attempt to resolve the conflicting goals of stability and sensitivity by proposing a coarse-to-fine shape matching algorithm. As the first step, two shapes are matched based on the structure of their disconnected skeletons, and following to that, the computed matching cost is re-evaluated by taking into account the similarity of boundary details in the light of class-specific boundary deformations which are learned from a given set of examples.

## 1 Introduction

There is a long history of research in computer vision on representing generic shape since shape information is a very strong visual clue in recognizing and classifying objects. A generic shape representation should be insensitive to not only geometric similarity transformations (*i.e.* translation, rotation, and scaling) but also visual transformations such as occlusion, deformation and articulation of parts. Since their introduction by Blum in [4], *local symmetry axis based* representations (commonly referred to as *shape skeletons*), have attracted and still attracts many scientists in the field, and became a superior alternative to boundary-based shape representations. These representation schemes naturally capture part structure by modeling any given shape via a set of axial curves, each of which explicitly represents some part of the shape. Once the relations among extracted shape primitives, *i.e.* the skeleton branches, are expressed in terms of a graph or a tree data structure (*e.g.* [5–7]), resulting shape descriptions are insensitive to articulations and occlusions.

A challenging issue regarding skeleton-based representations is the so-called *instability of skeletons* [8]. These representations are very sensitive to noise and/or small details on the shape boundary, and hence two visually very similar shapes might have structurally different skeleton descriptions. Hence, the success of any skeletonization method depends on how robust the final skeleton descriptions are in the presence of noise and shape features such as protrusions, indentations, necks, and concavities. As one might expect, this instability issue can also be passed over to the recognition framework, but in this case, the recognition algorithm should be devised in such a way that it includes a mechanism to handle possible structural changes (*e.g.* [5, 9–14]). A line of studies that focuses on solving the instability issue early in the representation level investigates the *abstraction of skeleton graphs*. This includes the methods which seek for a simplified graphical representation where the level of hierarchy is reduced to a certain extent (*e.g.* [5, 7, 15, 16]), the studies which try to come up with an abstract representation from a set of example skeletons (*e.g.* [5, 17, 18]), and more general *graph spectral* approaches (*e.g.* [19, 20]).

The method proposed in [1] is conceptually different than other approaches in the sense that the aim is obtaining the coarsest yet the most stable skeleton representations of shapes from scratch. The method depends on computing a special, excessively smooth distance surface where each skeleton extracted from this surface is in the form of a set of unconventionally *disconnected* and simple branches, *i.e.* the skeleton branches all terminate before reaching the unique shape center and no extra branching occurs on them. Hence, one can express disconnected skeletons in terms of *rooted attributed depth-1 trees*, whose nodes store some measurable properties, such as the location of the disconnection points, the length and the type (*positive* or *negative*, respectively identifying protrusions or indentations) of the branches [3] (Fig. 1).

Disconnected skeletons have been previously used for recognition in [2, 3] in which quite successful results are reported. Although the representation does not suffer from the instability of skeletons as a direct result of the disconnected nature of extracted branches, and that structure alone is an effective shape

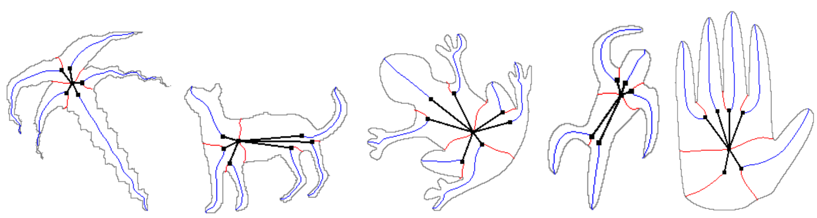


Fig. 1: Disconnected skeletons of some shapes and the corresponding tree representations. Note that each disconnection point (except the pruned major branches) gives rise to two different nodes in the tree, representing the positive and negative skeleton branches meeting at that disconnection point. However, for illustration purposes, only one node is drawn.

representation, as commented in [21], one might criticize the very coarseness of descriptions that they do not explicitly carry any information about boundary details. This issue is in fact about a philosophical choice of compromise between sensitivity and stability. Clearly, in distinguishing shapes, it might happen that the similarity of boundary details is more distinctive than the similarity of the structure of disconnection points (Fig. 6, 7).

In this study, we present a coarse-to-fine strategy to deal with such situations. The organization of the paper is as follows. In Section 2, we describe a way to obtain *radius functions* [4] (associated with the positive skeleton branches) in order to enrich the disconnected skeleton representation with information about shape boundary details. In Section 3, we utilize this extra information to enhance the class-specific knowledge utilized in the *category influenced matching* method proposed in [3] that boundary deformations in a shape category are additionally learned from examples. Following to that, in Section 4, we introduce a fine tuning step to the category influenced matching method, which then takes into account the similarity of boundary details. In Section 5, we present some matching results. Finally, in Section 6, we give a brief summary and provide some concluding remarks.

## 2 Obtaining Radius Functions

Disconnected skeleton of a shape is obtained from a special distance surface $\phi$, the level curves of which are the excessively smoothed versions of the initial shape boundary (Fig. 2(b)). The surface has a single extremum point capturing the center of a blob-like representation of the shape, and from that one can extract skeleton branches using the method in [22] in a straightforward way, without any need of a shock capturing scheme. As analyzed in detail in [2], this special surface is naively the limit case of the *edge strength function* $v$ [22] when the degree of regularization specified by the parameter $\rho$ tends to infinity (Fig. 2(c)-(e)). The excessive regularization employed in the formulation of $\phi$ makes it possible to obtain a very stable skeleton representation but this stability comes at the expense of losing information about boundary details. In contrast to Blum's skeletons, it is impossible to recover the distance from a skeleton point to the closest point on the shape boundary from the surface values.

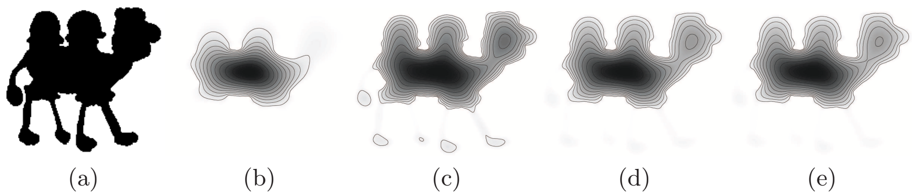


Fig. 2: (a) A camel shape. The level curves of the surfaces (b) $\phi$, (c) $v$, computed with $\rho = 16$, (d) $v$, computed with $\rho = 64$, (e) $v$, computed with $\rho = 256$.

In this study, we exploit the link between the surfaces $\phi$ and $v$, and in order to obtain the radius functions associated with the *positive* branches of disconnected skeletons (which are analogous to the Blum skeleton), we propose to benefit from a corresponding $v$ surface. Consider a ribbon-like section of a shape illustrated in Fig. 3 in which the dotted line shows the skeleton points representing that shape section. Assuming the 1D form of the edge strength function $v$, the diffusion process along a 1D slice (shown in red) is given by:

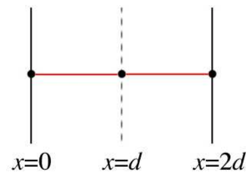



Fig. 3: An illustration of a ribbon-like section and its skeleton (the dotted line).

$$v_{xx}(x) - \frac{v(x)}{\rho^2} = 0 \; ; \quad 0 \leq x \leq 2d$$

with the boundary conditions $v(0) = 1,\ v(2d) = 1$.

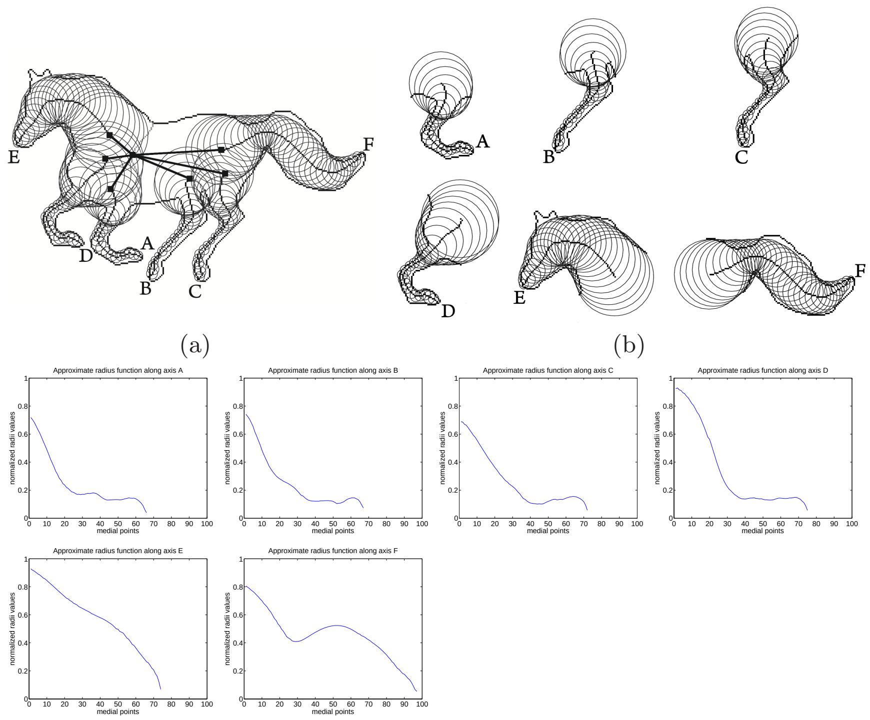
The explicit solution of this equation can be easily derived as:

$$v(x) = \left( \frac{1 - e^{2d/\rho}}{e^{-2d/\rho} - e^{2d/\rho}} \right) e^{-x/\rho} - \left( \frac{1 - e^{-2d/\rho}}{e^{-2d/\rho} - e^{2d/\rho}} \right) e^{x/\rho} \tag{1}$$

The value of $v$ on the skeleton point (the midpoint $x = d$) is equal to the hyperbolic cosine function $\frac{1}{cosh(d/\rho)}$, or equivalently, the distance from the skeleton point to the closest point on the boundary is given by $\rho cosh^{-1}(\frac{1}{v(d)})$. This explicit solution is certainly not valid for the 2D case as the interactions in the diffusion process are more complicated but it can be used as an approximation. Let $s$ be a skeleton point located at $(s_x, s_y)$ along a positive skeleton branch. Given a corresponding edge strength function $v$ computed with a sufficiently large value of $\rho$, the minimum distance from $s$ to the shape boundary, denoted by $r(s)$, can be approximated with:

$$r(s) = \rho cosh^{-1} \left( \frac{1}{v(s_x, s_y)} \right) \tag{2}$$

Fig. 4(a) shows the disconnected skeleton of a horse shape where the radius functions of the positive skeleton branches are approximately obtained from the edge strength function computed with $\rho = 256$ (the same value of $\rho$ is used in the experiments). The reconstructions of the shape sections associated with the positive skeleton branches are given separately in Fig. 4(b). Notice that small details on the shape boundary, *e.g.* the horse's ears, cannot be recovered completely since the perturbations on the shape boundary are ignored in disconnected skeleton representation. Moreover, the reconstructions might deviate from their true form at some locations, *e.g.* the skeleton points close to the leg joints, where a positive branch loses its ribbon-like structure of having slowly varying width. However, these approximate radius functions, when normalized with respect to the radius of maximum circle associated with the shape center, can be used as the descriptions of the most prominent boundary details (Fig. 4(c)).




(a) (b)

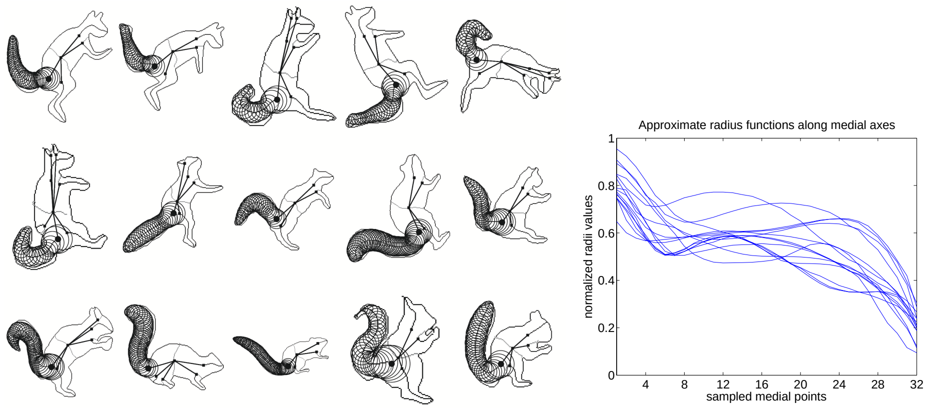
(c)

Fig. 4: (a) Disconnected skeleton of a horse shape and the radius functions obtained from the edge strength function computed with $\rho = 256$ (the maximal inscribed circles are drawn at every 3 consecutive skeleton points). (b) Shape sections associated with the positive skeleton branches. (c) Normalized radius functions associated with the branches A-F (from top left to bottom right).

## 3 Learning Boundary Deformations in a Shape Category

In the previous section, we developed a way to supply information about boundary details to disconnected skeletons. In this section, we extend our analysis and use the enriched skeleton descriptions to learn boundary deformations in a shape category from a given set of examples. It is noteworthy that the one-level hierarchy in the skeleton descriptions makes the learning process very practical since each positive skeleton branch simply corresponds to a major protrusion of the shape, and hence the correspondences among two disconnected skeletons can be found by a one-to-one matching.

Once the correspondence information is available, we follow the approach in [5], and model boundary deformations of a shape section in a category by forming a low-dimensional linear space from the corresponding radius functions. To be specific, we first uniformly sample equal number of points along matched positive branches (Fig. 5). The deformation space is then modeled by applying principal component analysis (PCA), where the first few principal components describe the representation space for possible deformations. In the experiments, our sample rate is 32 per each positive skeleton branch, and we use the first five principal components. Hence, each sampled radius functions are represented with a 5-dimensional vector.




(a) (b)

Fig. 5: An analysis of boundary deformations using approximated radius functions. (a) Equivalent shape sections of 15 `squirrel` shapes, each associated with a positive skeleton branch. (b) The corresponding set of uniformly sampled radius functions.

## 4 A Coarse-To-Fine Strategy to Incorporate Similarity of Boundary Details into Category-Influenced Matching

In [3], we presented a novel tree edit distance based shape matching method, named as *category influenced matching*, in which we used rooted attributed depth-1 trees to represent disconnected skeletons of shapes. The novelty in that work lies in the fact that the semantic roles of the shapes in comparison are distinguished as *query shape* or *database shape* (*i.e.* a member of a familiar shape category), and the knowledge about the category of the database shape is utilized as a context in the matching process in order to improve the performance. Such a context is defined by a *category tree*, which is a special tree union structure, nodes of which store basically the correspondence relations among the members of the same shape category, and some statistical information about observed skeleton attributes.

Here, we propose a fine tuning step to our category influenced matching method, in which the computed distance between the shapes in comparison is re-evaluated based on the similarity of their boundary details. Note that the process presented in Section 2 for learning class-specific boundary deformations can be easily integrated to the formation procedure of category trees. In that case, we additionally store the mean of the matched radius functions together with the reduced set of principle components in the nodes of the category tree. More formally, the overall algorithm can be summarized with the following two successive steps:

1. Let $\mathcal{T}_1$ be the shape tree of the query shape which is being compared with the shape tree of a database shape, denoted by $\mathcal{T}_2$, nodes of which is linked with a specific leaf node of the corresponding category tree. Compute an initial distance and the correspondences between $\mathcal{T}_1$ and $\mathcal{T}_2$ using category influenced matching method:

$$d\left(\mathcal{T}_1, \mathcal{T}_2\right) = \min_{\mathcal{S}} \left[ \sum_{u \in \Lambda} \texttt{rem}\left(u\right) + \sum_{v \in \Delta} \texttt{ins}\left(v\right) + \sum_{(u,v) \in \Omega} \texttt{ch}\left(u, v, \mathcal{B}\right) \right] \tag{3}$$

   where $\Lambda$ and $\Delta$ respectively denote the set of nodes removed from $\mathcal{T}_1$ and the set of nodes inserted to $\mathcal{T}_1$ from $\mathcal{T}_2$, and $\Omega$ denotes the set of matched nodes (See [3] for the details about the definition of cost functions associated with the edit operations `rem(ove)`, `ins(ert)` and `ch(ange)`).

2. Let $\mathcal{S}^* = (\Lambda^*, \Delta^*, \Omega^*)$ be the sequence of edit operations transforming $\mathcal{T}_1$ into $\mathcal{T}_2$ with the minimum cost. Re-calculate the distance between $\mathcal{T}_1$ and $\mathcal{T}_2$ according to Equation 4, in which $\Phi(u, v)$, appearing inside the extra term in front of the label change cost function, is the similarity between the radius functions associated with matched skeleton branches. Note that $\Phi(u, v)$ is calculated after projecting the corresponding uniformly sampled radius functions onto the related low-dimensional deformation space, as in Equation 5.

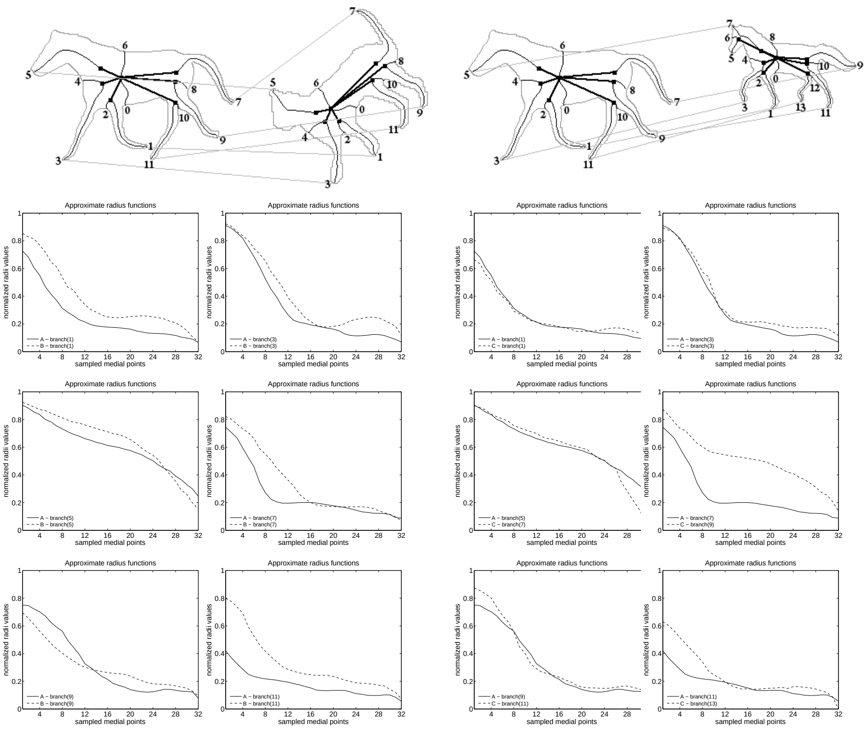
$$\widehat{d}\left(\mathcal{T}_1, \mathcal{T}_2\right) = \sum_{u \in \Lambda^*} \texttt{rem}\left(u\right) + \sum_{v \in \Delta^*} \texttt{ins}\left(v\right) + \sum_{(u,v) \in \Omega^*} \Big( \left(1 - \Phi(u, v)\right) \times \texttt{ch}\left(u, v, \mathcal{B}\right) \Big) \tag{4}$$

$$\Phi(u, v) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma^2}} exp\left( - \sum_{i=1}^{5} \dfrac{\left(\alpha_i - \beta_i\right)^2}{2\sigma^2} \right) & \text{if } u, v \text{ express \textit{positive} branches} \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

   where $\alpha$ and $\beta$ are to the vectors formed by projecting the radius functions associated with $u$ and $v$ onto related deformation space ($\sigma$ is taken as $\sigma = 0.4$ in the experiments).

## 5 Experimental Results

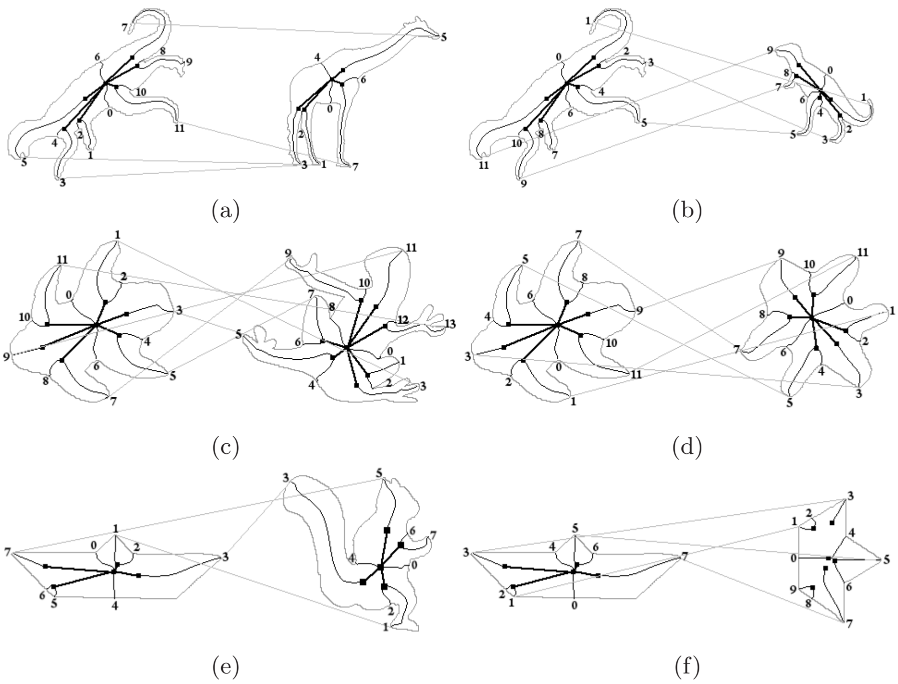
To demonstrate the effectiveness of the proposed approach, we test our method on the matching examples shown in Fig. 6, 7, in which the coarse structure of disconnected skeletons alone is not enough in distinguishing the shapes. In these examples, although part correspondences are correctly determined, the costs obtained with the category influenced matching method in [3] do not well reflect the perceptual dissimilarities[3]. On the other hand, when one examines the differences in the boundary details, it is clear that a more perfect decision can be made. For example, refer to Fig. 6. The pairs of radius functions associated with the matched branches is much similar in the case of matching of the horse shapes than the ones in the matching of the query horse shape with the cat shape. The only exception is the similarity of the horses' tails (Fig. 6(b), in the middle row and on the right) but note that these radius functions are compared in the corresponding deformation spaces that are learned from the given set of examples. In this regard, the proposed coarse-to-fine strategy can be used to refine the matching results.




(a) (b)

Fig. 6: Some matching results and the uniformly sampled radius functions of matched branches. The final matching costs are (a) 0.5800 (reduced from 0.7240), (b) 0.5368 (reduced from 0.7823). Note that the similarity of radius functions are actually computed in the related low-dimensional deformation spaces.

[3] In each experiment, the knowledge about the category of the database shape (the ones on the right) is defined by 15 examples of that category, randomly selected from the shape database given in [3].



(a) (b) (c) (d) (e) (f)

Fig. 7: Some other matching results. The final matching costs are (a) 1.1989 (reduced from 1.2904), (b) 0.9458 (reduced from 1.4936), (c) 1.9576 (reduced from 2.1879), (d) 1.8744 (reduced from 3.0387), (e) 0.8052 (reduced from 0.8105), (f) 0.6738 (reduced from 1.0875).

## 6 Summary and Conclusion

Despite its coarse structure, disconnected skeleton representation is a very stable and effective skeleton based representation. However, as the result of the excessive regularization employed in the extraction process, no information about boundary details is available in the skeleton descriptions. As articulated in [2], this is in fact a compromise between the opposing goals of stability and sensitivity. To enrich disconnected skeletons, we present a simple way to obtain the radius functions associated with the positive skeleton branches. This allows us to learn class-specific boundary deformations in a category when the correspondence relations among the members of the category is specified. This extra information is then incorporated into the category influenced matching method in [3] as a refinement step, in which the initial matching cost is re-evaluted by taking into account the similarity of radius functions of the matched positive branches. Our experiments show that this approach can be used to obtain perceptually more meaningful matching costs when the structure of disconnection points by themselves are not so distinctive in distinguishing shapes.

## References


1. Aslan, C., Tari, S.: An axis-based representation for recognition. In: ICCV 2005. Volume 2. (2005) 1339– 1346
2. Aslan, C., Erdem, A., Erdem, E., Tari, S.: Disconnected skeleton: Shape at its absolute scale. IEEE Trans. Pattern Anal. Mach. Intell. **30**(12) (2008) 2188–2203
3. Baseski, E., Erdem, A., Tari, S.: Dissimilarity between two skeletal trees in a context. Pattern Recognition **42**(3) (2009) 370–385
4. Blum, H.: Biological shape and visual science. In: Journal of Theoretical Biology. Volume 38. (1973) 205–287
5. Zhu, S.C., Yuille, A.L.: Forms: A flexible object recognition and modeling system. Int. J. Comput. Vision **20**(3) (1996) 187–212
6. Siddiqi, K., Kimia, B.B.: A shock grammar for recognition. In: CVPR. (1996) 507–513
7. Liu, T.L., Geiger, D., Kohn, R.V.: Representation and self-similarity of shapes. In: ICCV. (1998) 1129–1135
8. August, J., Siddiqi, K., Zucker, S.W.: Ligature instabilities in the perceptual organization of shape. Comput. Vis. Image Underst. **76**(3) (1999) 231–243
9. Sebastian, T.B., Klein, P.N., Kimia, B.B.: Recognition of shapes by editing shock graphs. In: ICCV. Volume 1. (2001) 755–762
10. Siddiqi, K., Shokoufandeh, A., Dickinson, S.J., Zucker, S.W.: Shock graphs and shape matching. Int. J. Comput. Vision **35**(1) (1999) 13–32
11. Pelillo, M., Siddiqi, K., Zucker, S.W.: Many-to-many matching of attributed trees using association graphs and game dynamics. In: Int. Workshop Visual Form. (2001) 583–593
12. Pelillo, M., Siddiqi, K., Zucker, S.W.: Matching hierarchical structures using association graphs. IEEE Trans. Pattern Anal. Mach. Intell. **21**(11) (1999) 1105–1120
13. Torsello, A., Hancock, E.R.: A skeletal measure of 2d shape similarity. CVIU **95**(1) (2004) 1–29
14. Liu, T., Geiger, D.: Approximate tree matching and shape similarity. In: ICCV. Volume 1. (1999) 456–462
15. Macrini, D., Siddiqi, K., Dickinson, S.: From skeletons to bone graphs: Medial abstraction for object recognition. In: CVPR. (2008)
16. Bai, X., Latecki, L.J.: Path similarity skeleton graph matching. IEEE Trans. Pattern Anal. Mach. Intell. **30**(7) (2008) 1282–1292
17. Torsello, A., Hancock, E.R.: Matching and embedding through edit-union of trees. In: ECCV. Volume 3. (2002) 822–836
18. Demirci, M.F., Shokoufandeh, A., Dickinson, S.J.: Skeletal shape abstraction from examples. IEEE Trans. Pattern Anal. Mach. Intell. (to appear (2009)
19. Shokoufandeh, A., Dickinson, S.J., Siddiqi, K., Zucker, S.W.: Indexing using a spectral encoding of topological structure. In: CVPR. (1999) 2491–2497
20. Demirci, M.F., van Leuken, R., Veltkamp, R.: Indexing through laplacian spectra. CVIU **110**(3) (2008) 312–325
21. Bai, X., Latecki, L.J., Liu, W.Y.: Skeleton pruning by contour partitioning with discrete curve evolution. IEEE Trans. Pattern Anal. Mach. Intell. **29**(3) (2007) 449–462
22. Tari, S., Shah, J., Pien, H.: Extraction of shape skeletons from grayscale images. CVIU **66**(2) (1997) 133–146