Shape Similarity Measure Based on Correspondence of Visual Parts

Longin Jan Latecki and Rolf Lakämper

Abstract—A cognitively motivated similarity measure is presented and its properties are analyzed with respect to retrieval of similar objects in image databases of silhouettes of 2D objects. To reduce influence of digitization noise, as well as segmentation errors, the shapes are simplified by a novel process of digital curve evolution. To compute our similarity measure, we first establish the best possible correspondence of visual parts (without explicitly computing the visual parts). Then, the similarity between corresponding parts is computed and aggregated. We applied our similarity measure to shape matching of object contours in various image databases and compared it to well-known approaches in the literature. The experimental results justify that our shape matching procedure gives an *intuitive* shape correspondence and is stable with respect to noise distortions.

Index Terms—Shape representation, shape similarity measure, visual parts, discrete curve evolution.

———————————— ✦ ————————————

1 INTRODUCTION

A shape similarity measure useful for shape-based retrieval in image databases should be in accord with our visual perception. This basic property leads to the following requirements:

1. A shape similarity measure should permit recognition of perceptually similar objects that are not mathematically identical.
2. It should abstract from distortions (e.g., digitization noise and segmentation errors).
3. It should respect significant visual parts of objects.
4. It should not depend on scale, orientation, and position of objects.

If we want to apply a shape similarity measure to distributed image databases, where the object classes are generally unknown a priori (e.g., in the Internet), it is necessary that:

5. A shape similarity measure is universal in the sense that it allows us to identify or distinguish objects of arbitrary shapes, i.e., no restrictions on shapes are assumed.

In this paper, we present a shape similarity measure that satisfies requirements 1 through 5. We demonstrate this by theoretical considerations, experimental results, and by comparison to the existing similarity measures.

The main contribution of this paper is presented in Sections 3 and 4, where our shape similarity measure is defined for object contours. Since contours of objects in digital images are distorted due to digitization noise and segmentation errors, it is desirable to neglect the distortions while, at the same time, preserving the perceptual appearance at the level sufficient for object recognition. Therefore, our similarity measure is applied to contours whose shape has been previously simplified by a discrete curve evolution. This allows us:

————————————————

- *The authors are with the Department of Applied Mathematics, University of Hamburg, Bundesstr. 55, 20146 Hamburg, Germany.*
  E-mail: {latecki, lakaemper}@math.uni-hamburg.de.

- to reduce influence of noise and
- to simplify the shape by removing *irrelevant* shape features without changing *relevant* shape features,

which contributes in a significant way to the fact that the similarity measure satisfies requirements 1 and 2. Observe that our discrete curve evolution is context sensitive since whether shape components are relevant or irrelevant cannot be decided without context. Our discrete curve evolution is presented shortly in Section 2 (more detailed presentations are given in [6], [7]).

In [6], we showed that significant visual parts become maximal convex arcs on an object contour simplified by the discrete curve evolution. Since we apply our similarity measure to contours simplified by the curve evolution and the similarity measure establishes the best possible correspondence of maximal convex/concave arcs, it follows that our similarity measure respects significant visual parts (requirement 3). Since requirements 1 and 3 are of a cognitive nature, they should be justified by cognitive experiments. We achieve this in Section 5 by relating our shape similarity measure to well-known measures that have been justified by cognitive experiments.

Requirements 4 and 5 are of pure mathematical nature and their satisfaction can be shown by simple arguments. The satisfaction of requirement 5 follows from the fact that we represent object boundaries as simple closed polygonal curves and that our shape similarity measure allows us to compare any two such curves. We simply obtain the polygonal curves from the boundary chain code (without any smoothing or other preprocessing) of segmented objects in digital images. Thus, every object contour in a digital image can be represented as a simple closed polygonal curve (with a possibly large number of vertices) without loss of information and without any additional computation.

Our approach to define a shape similarity measure is related to the one in Arkin et al. [1], where comparison of polygonal curves is based on L_2 distance of their turn angle representations (which we call tangent space representations). A more detailed comparison is given at the end of Section 4. The main difference is that our shape similarity measure is based on a subdivision of objects into parts of visual form. According to Siddiqi et al. [14], part-based representations allow for robust object recognition and play an important role in theories of object categorization and classification. There is also strong evidence for part-based representations in human vision, see, e.g., [14], [4]. Hoffman and Richards [3] provide strong evidence that contours are psychologically segmented at negative curvature minima.

2 DISCRETE CURVE EVOLUTION

Since contours of objects in digital images are distorted due to digitization noise and segmentation errors, it is desirable to neglect the distortions while at the same time preserving the perceptual appearance at the level sufficient for object recognition. An obvious way to neglect the distortions is to eliminate them by approximating the original contour with one that has a similar perceptual appearance. To achieve this, an appropriate approximation (or curve evolution) method is necessary. We achieve this through a novel method for evolution of polygonal curves.

Since any digital curve can be regarded as a polygon without loss of information (with possibly a large number of vertices), it is sufficient to study evolutions of polygonal shapes. The basic idea of the proposed evolution of polygons is very simple:

- In every evolution step, a pair of consecutive line segments s_1, s_2 is substituted with a single line segment joining the endpoints of $s_1 \cup s_2$.

The key property of this evolution is the order of the substitution. The substitution is done according to a relevance measure K given by:
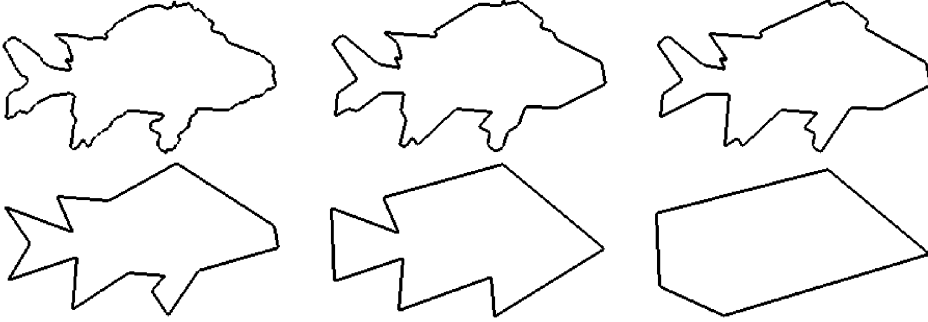
Fig. 1. A few stages of the proposed curve evolution. A distorted version of the contour in WWW page http://www.ee.surrey.ac.uk/Research/VSSP/imagedb/demo.html.

$$K(s_1, s_2) = \frac{\beta(s_1, s_2)l(s_1)l(s_2)}{l(s_1) + l(s_2)}, \qquad (1)$$

where $\beta(s_1, s_2)$ is the turn angle at the common vertex of segments s_1, s_2 and l is the length function normalized with respect to the total length of a polygonal curve C. The main property of this relevance measure is the following:

- The higher the value of $K(s_1, s_2)$, the larger is the contribution to the shape of the curve of arc $s_1 \cup s_2$.

A cognitive motivation of this property is given in [6]. A detailed description of our discrete curve evolution can be found in [7]. Online demonstrations can be viewed on our WWW site [8]. A few example stages of the discrete curve evolution are shown in Fig. 1.

Our curve evolution method does not require any control parameters to achieve the task of shape simplification, i.e., there are no parameters involved in the process of the discrete curve evolution. However, we clearly need a stop parameter, which is the number of iterations the evolution is performed. This parameter is automatically determined in accord with our visual perception by the procedure described at the end of Section 4.

3   SHAPE SIMILARITY MEASURE

In this section, we define our shape similarity measure. This measure is applied to contours which have been previously simplified by the discrete curve evolution. The appropriate evolution stage is selected for each shape and then the similarity is computed for the obtained instances of the shapes.

Our similarity measure profits from the decomposition into visual parts based on convex boundary arcs [6]. The key idea is to find the right correspondence of the visual parts. We assume that a single visual part (i.e., a convex arc) of one curve can correspond to a sequence of consecutive convex and concave arcs of the second curve, e.g., part number 0 of the top-left fish contour in Fig. 2. This

assumption is justified by the fact that a single visual part should match its noisy versions which can be composed of sequences of consecutive convex and concave arcs or by the fact that a visual part obtained at a higher stage of evolution should match the arc it originates from. Since maximal convex arcs determine visual parts, this assumption guarantees preservation of visual parts (without explicitly computing visual parts).

In this section, we assume that polygonal curves are simple, i.e., there are no self-intersections and they are closed. We also assume that we traverse polygonal curves in the counterclockwise direction.

Let $convconc(C)$ denote the set of all maximal convex or concave subarcs of a polygonal curve C. Then, the order of traversal induces the order of arcs in $convconc(C)$.

Since a simple one-to-one comparison of maximal convex/concave arcs of two polygonal curves is of little use, due to the fact that the curves may consist of a different number of such arcs and even similar shapes may have different small features, we join together maximal arcs to form *groups*:

A **group** g of curve C is a union of a (nonempty) consecutive sequence of arcs in $convconc(C)$. Thus, g is also a subarc of C. We denote $groups(C)$ as the set of all groups of C. We have $convconc(C) \subseteq groups(C)$. A grouping G for a curve C is an ordered set of consecutive groups $G = (g_0, \ldots, g_{n-1})$ for some $n \geq 0$ such that

- $g_i \cap g_{i+1(mod\ n)}$ is a single line segment for $i = 0, \ldots, n-1$.

Since any two consecutive groups intersect in exactly one line segment, the whole curve C is covered by G. We denote the set of all possible groupings G of a curve C as $\mathcal{G}(C)$. Fig. 2 shows example groupings of the given contours, where each group is assigned a different number.

Given two curves C_1, C_2, we say that groupings $G_1 \in \mathcal{G}(C_1)$ and $G_2 \in \mathcal{G}(C_2)$ **correspond** if there exists a bijection $f : G_1 \to G_2$ such that

1. f preserves order of groups and
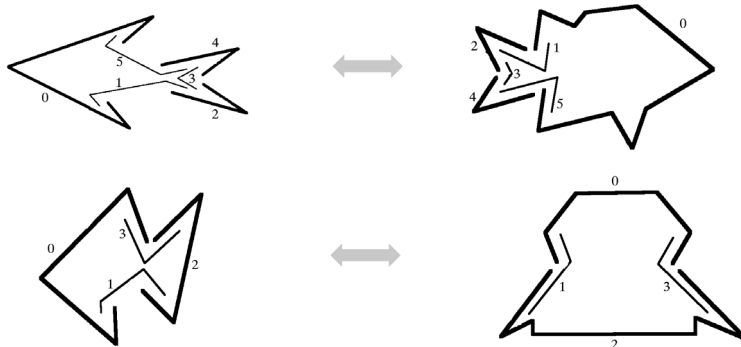2.   For all $x \in G_1$, $x \in convconc(C_1)$ or $f(x) \in convconc(C_2)$.

Fig. 2. The corresponding arcs are labeled by the same numbers.

We call the bijection f a correspondence between G_1 and G_2. We denote the set of all corresponding pairs (G_1, G_2) in $\mathcal{G}(C_1) \times \mathcal{G}(C_2)$ by $\mathcal{C}(C_1, C_2)$. Two example correspondences are shown in Fig. 2. The condition that any f is a bijection means that both curves are decomposed into the same amount of groups. Condition 2 means that at least one of corresponding groups $x \in G_1$ or $f(x) \in G_2$ is a maximal (convex or concave) arc. The reason is that we want to allow mappings between one-to-many maximal arcs or many-to-one maximal arcs, but never between many-to-many maximal arcs. Since maximal convex arcs determine visual parts, condition 2 guarantees preservation of visual parts (without explicitly computing visual parts). Condition 2 also implies that every maximal (convex or concave) arc in a higher stage of abstraction will match to the consecutive sequence of arcs it originates from.

A **similarity measure** for curves C_1, C_2 is defined as

$$S_c(C_1, C_2) = \min \left\{ \sum_{x \in G_1} S_a(x, f_{(G_1, G_2)}(x)) : (G_1, G_2) \in \mathcal{C}(C_1, C_2) \right\}, \tag{2}$$

where $f_{(G_1, G_2)}$ is the correspondence between G_1 and G_2 and S_a is a similarity measure for arcs that will be defined in the next section. To compute $S_c(C_1, C_2)$ means to find in the set $\mathcal{C}(C_1, C_2)$ of all corresponding groupings a pair of groupings for which the sum of the differences between the corresponding groups $S_a(x, f_{(G_1, G_2)}(x))$ is minimal. The task of computing the similarity measure defined in (2) can be formulated as a problem of computing the global minimum: Given a function

$$\mathcal{M}(X, Y) = \sum_{x \in X} S_a(x, f_{(X,Y)}(x))$$

that assigns a group matching value to every corresponding pair $(X, Y) \in \mathcal{C}(C_1, C_2)$ related by the correspondence $f_{(X,Y)}$, find a pair $(G_1, G_2) \in \mathcal{C}(C_1, C_2)$ for which $\mathcal{M}(G_1, G_2)$ is minimal, i.e., $\mathcal{M}(G_1, G_2) \leq \mathcal{M}(X, Y)$ for all elements $(X, Y) \in \mathcal{C}(C_1, C_2)$.

The similarity measure defined in (2) is computed using dynamic programming. Numerous experimental results show that it leads to intuitive arc correspondences, e.g., see Fig. 2. The experimental results are described in Section 5.

4 TANGENT SPACE REPRESENTATION

The goal of this section is to define the similarity measure S_a for arcs that is part of definition of our shape similarity measure in Section 3. As mentioned in the introduction, any digital curve C can be interpreted as a polygonal curve with a possibly large number of vertices without loss of information.

We assign to every polygonal curve a *tangent function*, which is a step function. We use the tangent function as a basis for the proposed similarity measure of simple polygonal arcs. Let C be a polygonal curve. We treat it as a function $C : [0, 1] \to \mathbb{R}^2$, i.e., the length of C is rescaled to 1. The **tangent function** of C (which is also called a *turning function*) is a multivalued function $T(C) : [0, 1] \to [0, 2\pi]$ defined by $T(C)(s) = C'_-(s)$ and $T(C)(s) = C'_+(s)$, where $C'_-(s)$ and $C'_+(s)$ are left and right derivatives of C. For example, see Fig. 3. Clearly, only if $C(s)$ is a vertex of the polygon $C'_-(s) \neq C'_+(s)$. The y-difference between two adjacent steps of the tangent function represents the turn angle of the corresponding pair of line segments.

Now, we define the similarity measure for arcs. Let c, d be simple polygonal arcs that are parts of closed curves C, D. We denote by $T(c), T(d)$ their tangent functions, uniformly scaled so that their projections on the x-axis $\pi_x(T(c))$ and $\pi_x(T(d))$ both have length one. The arc similarity measure is given by (e.g., see Fig. 3)

Fig. 3. Two polygonal curves, say c and d, their tangent functions $T(c), T(d)$, and the distance $D_{L_2}(T(c), T(d))$ of the tangent functions.

$$S_a(c, d) = \left(\int_0^1 (T(c)(s) - T(d)(s) + \theta_0)^2 \, ds \right) \max(l(c), l(d)) \max \left(\frac{l(c)}{l(d)}, \frac{l(d)}{l(c)} \right), \tag{3}$$

where l is the relative arclength of an arc with respect to the boundary length of the curve and θ_0 is defined below. The integral in (3) is weighted with the arc length penalized by the difference in length of the corresponding parts. For example, if $l(c) > l(d)$, then the scaling term is equal to $l(c) \frac{l(c)}{l(d)}$, where $l(c)$ scales the value of the integral by the relative arclength of arc c with respect to the length of curve C and $\frac{l(c)}{l(d)}$ is the penalty for the relative length difference of arcs c and d.

The constant θ_0 is a translation of $T(d)$ that minimizes the integral, i.e.,

$$\int_0^1 (T(c)(s) - T(d)(s) + \theta_0)^2 \, ds =$$
$$\inf_{\theta \in [0, 2\pi]} \int_0^1 (T(c)(s) - T(d)(s) + \theta)^2 \, ds.$$

The constant θ_0 exists and is given by Lemma 3 in Arkin et al. [1]. Observe that we apply measure (3) with the restriction that $c \in convconc(C)$ or $d \in convconc(D)$.

Now, we describe the procedure that determines the stop parameter, i.e., a stage at which the curve evolution halts. The evolved contours obtained at this stage are used as input to our shape similarity measure.

Let $P = P^0, \ldots, P^m$ be polygons obtained from a polygon P in the course of discrete curve evolution such that P^m is the first convex polygon.

For $i = m$ with a step -1 do:

The curve of an abstraction level P^i is segmented into maximal convex/concave parts. These parts are compared to their corresponding parts on the original polygon P^0, where the corresponding parts are the ones having the same endpoints. The comparison is done using the S_a-measure. If the comparison of a single part in P^i leads to a value higher than a given threshold s, the shape P^i is abstracted too much and the previous abstraction level P^{i-1} is taken.
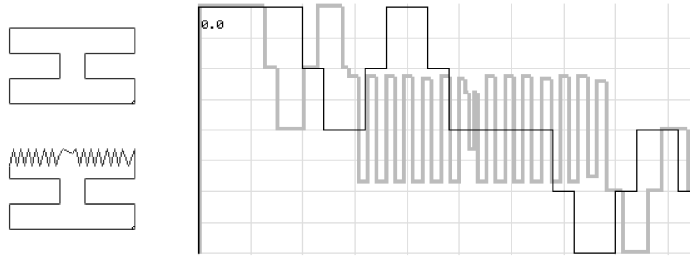
Fig. 4. A comparison of tangent functions of two contours based on Arkin et al. [1].

5   COMPARISON TO KNOWN SIMILARITY MEASURES

We concentrate on comparison to universal similarity measures that are translation, rotation, reflection, and scaling invariant. This excludes, for example, Hausdorff distance, which is universal but is not rotation, reflection, and scaling invariant.

As stated in Section 1, our approach to defining a shape similarity measure is related to the one in Arkin et al. [1], where L_2 distance of tangent functions is used for comparing polygonal shapes. The main difference of our approach is that we use L_2 distance for comparing tangent functions of parts of polygonal shapes, which makes our approach significantly more robust with respect to nonuniform distortions. This is illustrated by comparing the two shapes in Fig. 4. Since the tangent functions of both shapes are scaled to the same length, the local distortions make it impossible to overlay these functions in such a way that the corresponding parts are on top of each other. This results in a large similarity value of the similarity measure of Arkin et al. [1].

To compute our shape similarity measure, we first establish the best possible correspondence of the maximal convex/concave arcs. The corresponding maximal convex/concave arcs are labeled with the same numbers 0 to 3 in Fig. 5. The maximal convex arc 3 of the first shape correctly corresponds to the part 3 of the second shape (which is the part between lines 2 and 3 on the tangent function). In our approach, the comparison of tangent functions is done for each pair of corresponding parts separately. We scale each corresponding pair to the same length 1 and compute the distance of the local tangent functions obtained in this way. This correctly results in a small value of our similarity measure for the two shapes. This example also demonstrates that our similarity measure satisfies requirements 1, 2, and 3 described in Section 1. We want to stress

that the shape of the distorted object in this example was not simplified before the comparison.

Our shape similarity measure was throughly tested on the MPEG-7 Core Experiment CE-Shape-1 for shape desciptors. We report on this in [9]. Further, we compared our approach to retrieval of similar objects with a similarity measure based on curvature scale space in Mokhtarian et al. [10]. The curvature scale space representation is obtained by curve evolution guided by diffusion equation [11]. The similarity measure in [10] is applied to a database of marine animals in which every image contains one animal. We applied our similarity measure to the same database. The results, which can be viewed on our home page [8], are very similar, but not identical to the results in [10].

We compared the results of our approach with the approach presented in Siddiqi et al. [13], which is based on a hierarchical structure of shocks in skeletons of 2D objects. In this approach, object shape is represented as a graph of shocks. The similarity of objects is determined by a similarity measure of the graphs of shocks. Although the shape representation in [13] is not based on boundary curves, the results of our similarity measure are very similar to the results in [13]. These results can be viewed on our home page [8].

An interesting approach to establishing desirable properties of shape similarity measures is given in Basri et al. [2]. The desirable properties are illustrated and tested on three proposed similarity measures: spring model, linear model, and continuous deformation model. These models measure deformation energy needed to obtain one object from the other. The calculation of deformation energy is based on (best possible) correspondence of boundary points and local distortions of corresponding points as a function
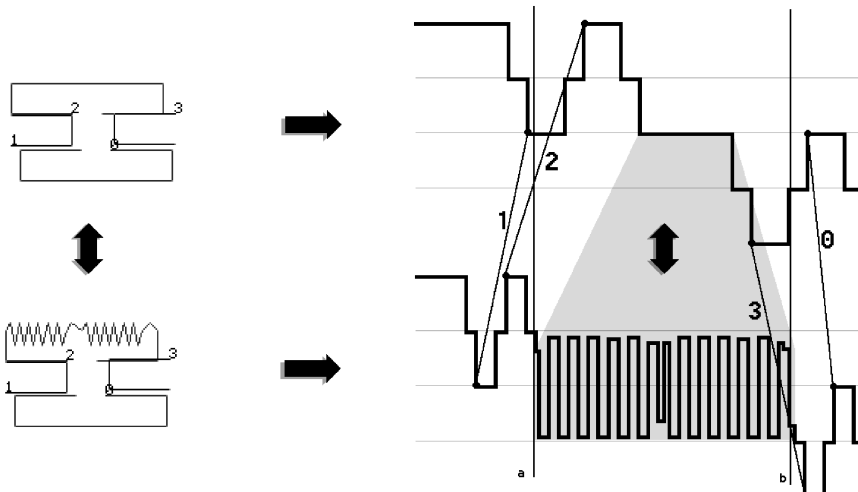

Fig. 5. The gray region shows the side of the first shape that should correspond to the distorted side of the second shape.

Fig. 6. The more vertices a regular polygon has, the more similar to a circle it is.

Fig. 7. Bending at a part boundary should imply fewer changes than bending in the middle of a part.

Fig. 8. Our similarity measure yields results in accord with our visual perception: $S_c((a),(b)) = 368$ and $S_c((a),(c)) = 518$.

of local curvature differences. Thus, the calculation of the three measures requires local computation of curvature.

We demonstrate that our measure yields desirable results in accord with experiments proposed in Basri et al. [2]. In Table 1 in [2], an experiment with regular polygons is demonstrated. The intuitive idea is that the more vertices a regular polygon has, the more similar to a circle it is. The results of our similarity measure on images from Table 1 are shown in Fig. 6. It can be easily observed that our measure yields the desirable results.

Basri et al. [2] further argue that similarity measures should be sensitive to structure of visual parts of objects. To check this property, they suggest that bending an object at a part boundary should imply fewer changes than bending in the middle of a part. This property of our measure is illustrated in Fig. 7.

The similarity measures in Basri et al. [2] are obtained as the integral of local distortions between corresponding contour points. The authors themselves point out a counterintuitive performance of their measures when applied to the objects like the ones in Fig. 8 ([2, Fig. 17]). The H-shaped contour (Fig. 8a) is compared to two different distortions of it. Although the shape (Fig. 8b) appears more similar to (Fig. 8a) than the shape (Fig. 8c), the amount of local distortion to obtain (Fig. 8b) and (Fig. 8c) from (Fig. 8a) is the same. Therefore, all three measures presented in [2] imply that shapes
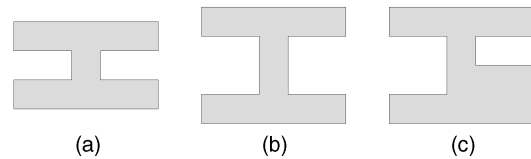
(Fig 8b) and (Fig. 8c) are equally similar to (Fig. 8a). Basri et al. argue that this counterintuitive performance is due to the fact that their measures are based on contour representation of shapes, but the performance of our measure clearly proves that this is not the case:

Our similarity measure is based on contour representation and gives similarity values in accord with visual perception. Our measure yields $S_c((a),(b)) = 368$ and $S_c((a),(c)) = 518$, i.e., Fig. 8b is more similar than Fig. 8c to Fig. 8a. The main difference is that our measure is not based on local properties, i.e., it is not based on correspondence of contour points and their local properties, but on correspondence of contour parts.

The approach described in Sclaroff [12] is based on distance to object prototypes representing classes of shapes. Shape similarity is computed in terms of the amount of strain energy needed to deform one object into another. Therefore, the above discussion of approaches based on deformation energy also applies to [12]. Additionally, the computation of the shape similarity in [12] requires establishing direct point correspondence and shape alignment, which is a highly nontrivial task. Sclaroff uses Hausdorff distance [5] to achieve this task in his experiments.

6 CONCLUSIONS

This paper presents novel techniques for shape-based object recognition, especially developed to match the discrete nature of digital image data. We developed a shape similarity measure that fulfills necessary requirements for cognitively motivated shape similarity measures. These requirements are proposed by the authors (see Section 1) and by Basri et al. [2] (see Section 5). We applied our measure to retrieval of similar objects in a database of object contours, see Fig. 9. The user query can be given either by a graphical sketch or by an example silhouette. Numerous experiments with various databases of real images and comparison to known universal shape similarity measures justify an *intuitive*
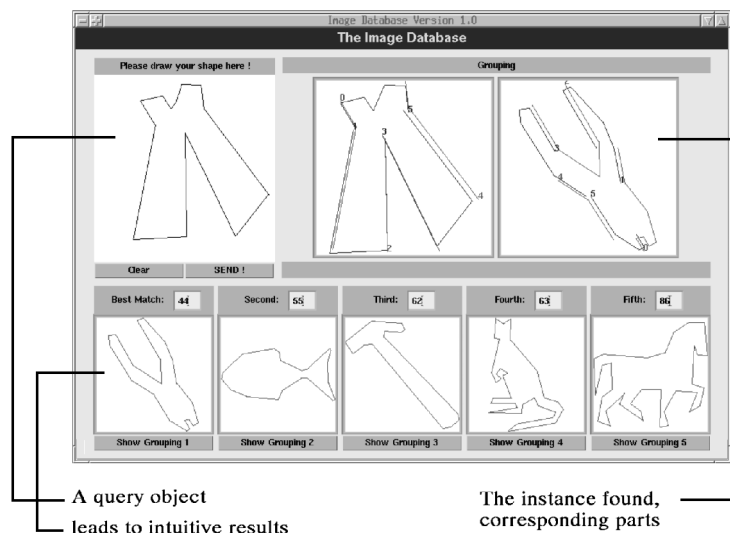

Fig. 9. Retrieval of similar objects based on our similarity measure.

shape correspondence and stability with respect to noise distortions of our shape similarity measure.

The main strength of our shape similarity measure is that it establishes the best possible correspondence of boundary parts that are visually significant. A discrete evolution method that is used as a prefilter for shape comparison is a basis for a shape decomposition into visual parts.

REFERENCES

[1]   M. Arkin, L.P. Chew, D.P. Huttenlocher, K. Kedem, and J.S.B. Mitchell, "An Efficiently Computable Metric for Comparing Polygonal Shapes," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 13, pp. 209-206, 1991.
[2]   R. Basri, L. Costa, D. Geiger, and D. Jacobs, "Determining the Similarity of Deformable Shapes," *Vision Research,* vol. 38, pp. 2,365-2,385, 1998.
[3]   D.D. Hoffman and W.A. Richards, "Parts of Recognition," *Cognition,* vol. 18, pp. 65-96, 1984.
[4]   D.D. Hoffman and M. Singh, "Salience of Visual Parts," *Cognition,* vol. 63, pp. 29-78, 1997.
[5]   D. Huttenlocher, G. Klanderman, and W. Rucklidge, "Comparing Images Using the Hausdorff Distance," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 15, pp. 850-863, 1993.
[6]   L.J. Latecki and R. Lakämper, "Convexity Rule for Shape Decomposition Based on Discrete Contour Evolution," *Computer Vision and Image Understanding,* vol. 73, pp. 441-454, 1999.
[7]   L.J. Latecki and R. Lakämper, "Polygon Evolution by Vertex Deletion," *Scale-Space Theories in Computer Vision Proc. Int'l Conf. Scale-Space '99,* M. Nielsen, P. Johansen, O.F. Olsen, and J. Weickert, eds., Sept. 1999.
[8]   L.J. Latecki, R. Lakämper, and U. Eckhardt   http://www.math.uni-hamburg.de/home/lakaemper/shape.
[9]   L.J. Latecki, R. Lakämper, and U. Eckhardt, "Shape Descriptors for Nonrigid Shapes with a Single Closed Contour," *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR),* June 2000.
[10] F. Mokhtarian, S. Abbasi, and J. Kittler, "Efficient and Robust Retrieval by Shape Content through Curvature Scale Space," *Image Databases and Multi-Media Search,* A.W.M. Smeulders and R. Jain, eds., pp. 51-58, Singapore: World Scientific, 1997.
[11] F. Mokhtarian and A.K. Mackworth, "A Theory of Multiscale, Curvature-Based Shape Representation for Planar Curves," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 14, pp. 789-805, 1992.
[12] S. Sclaroff, "Deformable Prototypes for Encoding Shape Categories in Image Databases," *Pattern Recognition,* vol. 30, pp. 627-641, 1997.
[13] K. Siddiqi, A. Shokoufandeh, S.J. Dickinson, and S.W. Zucker, "Shock Graphs and Shape Matching," *Int'l J. Computer Vision,* vol. 35, no. 1, pp. 13-32, 1999.
[14] K. Siddiqi, K.J. Tresness, and B.B. Kimia, "Parts of Visual Form: Psychophysical Aspects," *Perception,* vol. 25, pp. 399-424, 1996.