

Efficient Partial Shape Matching of Outer Contours

Michael Donoser, Hayko Riemenschneider and Horst Bischof

Institute for Computer Graphics and Vision, Graz University of Technology
{donoser,hayko,bischof@icg.tugraz.at}

Abstract. This paper introduces a novel efficient partial shape matching method named *IS-Match*. We use sampled points from the silhouette as a shape representation. The sampled points can be ordered which in turn allows to formulate the matching step as an order-preserving assignment problem. We propose an angle descriptor between shape chords combining the advantages of global and local shape description. An efficient integral image based implementation of the matching step is introduced which allows detecting partial matches an order of magnitude faster than comparable methods. We further show how the proposed algorithm is used to calculate a global optimal Pareto frontier to define a partial similarity measure between shapes. Shape retrieval experiments on standard shape databases like MPEG-7 prove that state-of-the-art results are achieved at reduced computational costs.

1 Introduction

Shape matching is a well investigated problem in computer vision and has versatile applications as e. g. in object detection [1–3] or image retrieval [4]. The most important part of designing a shape matcher is the choice of the shape representation which has a significant effect on the matching step. Shapes have for example been represented by curves [5], medial axes [6], shock structures [7] or sampled points [8].

In general current shape matching algorithms can be divided into two main categories: global and local approaches. Global matching methods compare the overall shapes of the input objects by defining a global matching cost and an optimization algorithm for finding the lowest cost match. One of the most popular methods for global shape matching is the shape context proposed by Belongie et al. [8]. Their algorithm uses randomly sampled points as shape representation and is based on a robust shape descriptor – the shape context – which allows to formulate the matching step as a correspondence problem. The shape context is the basis for different extensions as proposed by Ling and Jacobs [9] or Scott and Nowak [10].

While such global matching methods work well on most of the standard shape retrieval databases, they cannot handle strong articulation, part deformations or occlusions. For example, shape context is a global descriptor and local articulations influence the description of every sampled point. To reduce this effect

larger histogram bins are used further away from the point. Although this reduces the problem e. g. occlusions still lead to matching errors as it is illustrated in Figure 1 for the shape context based COPAP framework [10].

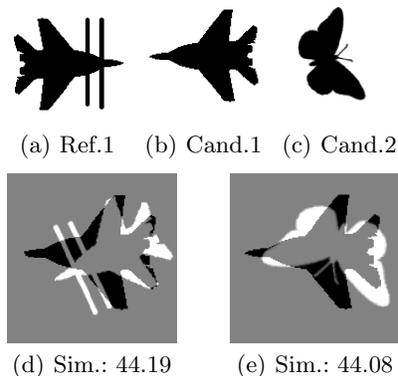


Fig. 1: Global shape matching methods like COPAP [10] are not able to handle occlusion or strong articulation, because of internally using global descriptors. The same similarity value for the partially occluded shape and a totally different one is returned.

These problems are handled well by purely local matching methods as e. g. proposed by Chen et al. [11], which allow accurately measuring local similarity, but in contrast fail to provide a strong global description for accurate shape alignment. In this work, we try to bridge the gap between the two worlds by combining their advantages.

We propose a novel shape matching method denoted as *IS-Match* (Integral Shape Match). We use sampled points along the silhouette as representation and exploit the ordering of the points to formulate matching as order-preserving assignment problem. We introduce a chord angle descriptor which combines local and global information and is invariant to similarity transformations. An integral image based matching algorithm detects partial matches with low computational complexity. The method returns a set of partial sub-matches and therefore also allows matching between differently articulated shapes.

The main contributions of this paper are: (1) a chord angle based descriptor combining local and global information invariant to similarity transformations (2) an efficient integral image based matching scheme where matching in practice takes only a few milliseconds and (3) the calculation of a global optimal Pareto frontier for measuring partial similarity.

The outline of the paper is as follows. Section 2 describes the partial shape matching concept named *IS-Match* in detail. Section 3 presents a comprehensive evaluation of the proposed algorithm for shape retrieval experiments on common databases like MPEG-7. All evaluations prove that state-of-the-art results are achieved at reduced computational costs.

2 Partial Shape Matching: *IS-Match*

Our shape matching algorithm named *IS-Match* takes two shapes as input and returns detected partial matches and a similarity score as result. Section 2.1 describes the shape representation used, which is a sequence of points sampled from the silhouette. Section 2.2 introduces a chord angle based descriptor invariant to similarity transformations. In Section 2.3 an efficient integral image based algorithm for matching the descriptor matrices to each other is outlined, which allows detecting subparts of the contours that possess high similarity with low computational complexity. Section 2.4 describes how a global optimal Pareto frontier is calculated and the corresponding Salukawdze distance is returned as measure for partial similarity. Finally, Section 2.5 analyzes the required computational complexity for single shape matches.

2.1 Shape Representation

The first step of our method is to represent the shapes by a sequence of points sampled from the contour. There are two different variants for point sampling: (a) sampling the same number of points from the contours or (b) equidistant sampling, i. e. fixing the contour length between sampled points. The type of sampling significantly influences the invariance properties of our method. Based on equidistant sampling occlusions as e. g. shown in Figure 1 can be handled but then only shapes at the same scale are correctly matched. By sampling the same number of points our method becomes invariant to similarity transformations, but strong occlusions cannot be handled anymore. In this paper we always use the equidistant sampling for the task of shape retrieval on single scale data sets. Nevertheless all subsequent parts of the method are defined in a manner independent of the sampling type. Therefore, we can switch the sampling type without requiring any modifications of the method. Please note that in other applications, as e. g. shape based tracking [12] or recognizing actions by matching to pose prototypes [13], equidistant sampling should be preferred.

Because the proposed shape matching method focuses on analyzing silhouettes (as required in the areas of segmentation, detection or tracking) the sampled points can be ordered in a sequence which is necessary for the subsequent descriptor calculation and the matching step. Thus, any input shape is represented by the sequence of points $P_1 \dots P_N$, where N is the number of sampled points.

2.2 Shape Descriptor

The descriptor constitutes the basis for matching a point P_i of the reference shape to a point Q_j of the candidate shape. We formulate the subsequent matching step presented in Section 2.3 as an order-preserving assignment problem. Therefore, the descriptor should exploit the available point ordering information. In comparison, the shape context descriptor loses all the ordering information due to the histogram binning and for that reason does not consider the influence of the local neighborhood on single point matches.

We propose a descriptor inspired by chord distributions. A chord is a line joining two points of a region boundary, and the distribution of their lengths and angles was used as shape descriptor before, as e. g. by Cootes et al. [14]. Our descriptor uses such chords, but instead of building distribution histograms, we use the relative orientations between specifically chosen chords.

Our descriptor is based on angles α_{ij} which describe the relative spatial arrangement of the sampled points. An angle α_{ij} is calculated between a chord $\overline{P_i P_j}$ from a reference point P_i to another sampled point P_j and a chord $\overline{P_j P_{j-\Delta}}$ from P_j to $P_{j-\Delta}$ by

$$\alpha_{ij} = \sphericalangle (\overline{P_i P_j}, \overline{P_j P_{j-\Delta}}), \quad (1)$$

where $\sphericalangle(\dots)$ denotes the angle between the two chords and $P_{j-\Delta}$ is the point that comes Δ positions before P_j in the sequence as is illustrated in Figure 2. Since angles are preserved by a similarity transformation, this descriptor is invariant to translation, rotation and scale.

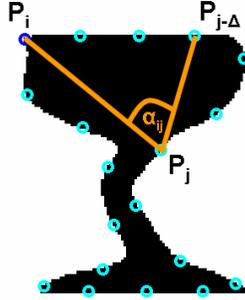


Fig. 2: Our shape descriptor is based on calculating N angles for each sampled point of the shape. In this case P_i is the reference point and the calculation of the angle α_{ij} to the point P_j with $\Delta = 3$ is shown.

In the same manner N different angles $\alpha_{i1} \dots \alpha_{iN}$ can be calculated for one selected reference point P_i . Additionally, each of the sampled points can be chosen as reference point and therefore a $N \times N$ matrix A defined as

$$A = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1N} \\ \vdots & \ddots & \vdots \\ \alpha_{N1} & \cdots & \alpha_{NN} \end{pmatrix} \quad (2)$$

can be used to redundantly describe the entire shape. Obviously, elements on the main diagonal $\alpha_{11}, \dots, \alpha_{NN}$ are always zero. This descriptor matrix is not symmetric because it considers relative orientations. Please note, that such a shape descriptor implicitly includes local information (close to the main diago-

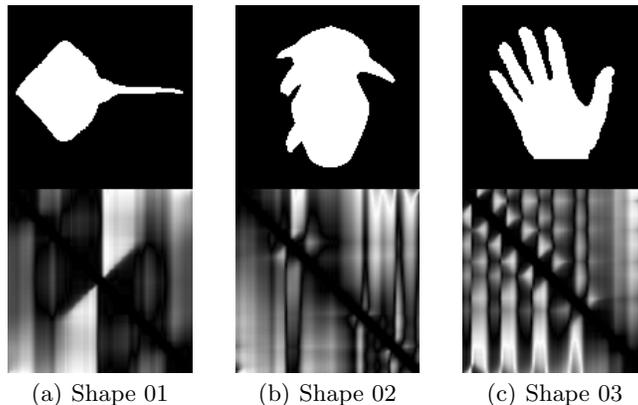


Fig. 3: Visualizations of distinct chord angle based shape descriptors. Bright areas indicate parts of the silhouettes which significantly deviate from straight lines.

nal) and global information (further away from the diagonal). Figure 3 shows different descriptor matrices for selected shapes.

The descriptor depends on which point is chosen as the first point of the sequence. For example the descriptor matrix A shown before changes to

$$A_{(k)} = \begin{pmatrix} \alpha_{kk} & \dots & \alpha_{k1} & \dots & \alpha_{k(k-1)} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \alpha_{1k} & \dots & \alpha_{11} & \dots & \alpha_{1(k-1)} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \alpha_{(k-1)k} & \dots & \alpha_{(k-1)1} & \dots & \alpha_{(k-1)(k-1)} \end{pmatrix} \quad (3)$$

if the k -th point is set as the first point of the sequence. Because only closed boundaries are considered, these two matrices $A_{(k)}$ and A are directly related by a circular shift. Matrix A can be obtained by shifting the $A_{(k)}$ matrix $k - 1$ rows up and $k - 1$ columns to the left. This is an important property for the efficient descriptor matching algorithm presented in the next section.

2.3 Matching Algorithm

To find a partial match between two given shape contours R_1 and R_2 the corresponding descriptor matrices A_1 with size $M \times M$ and A_2 with size $N \times N$ are compared. For notational simplicity we assume that $M \leq N$.

The aim of shape matching is to identify parts of the two shapes that are similar to each other. In terms of comparing the two descriptor matrices this equals to finding $r \times r$ sized blocks starting at the main diagonal elements $A_1(s, s)$ and $A_2(m, m)$ of the two descriptor matrices which yield a small angular difference value $D_\alpha(s, m, r)$ defined by

$$D_\alpha(s, m, r) = \frac{1}{r^2} \sum_{i=0}^{r-1} \sum_{j=0}^{r-1} [A_1(s+i, s+j) - A_2(m+i, m+j)]^2 \quad (4)$$

between them. This equation is valid due to the previously explained property that a different starting point just leads to a circular shift of the descriptor matrix as illustrated in Equation 3. To find such blocks all different matching possibilities and chain lengths have to be considered and the brute-force method becomes inefficient for larger number of points. Therefore, different authors as e.g. [15] proposed approximations where for example only every n -th point is considered as starting point.

We propose an algorithmic optimization to overcome the limitations of the brute-force approach, which is based on an adaption of the Summed-Area-Table (SAT) approach to calculate all the descriptor differences $D_\alpha(s, m, r)$. The SAT concept was originally proposed for texture mapping and brought back to the computer vision community by Viola and Jones [16] as integral image. The integral image concept allows to calculate rectangle image features like the sum of all pixel values for any scale and any location in constant time.

For calculating the similarity scores for all possible configuration triplets $\{s, m, r\}$ in the most efficient way N integral images $Int^1 \dots Int^N$ each of size $M \times M$ are built for N descriptor difference matrices M_D^n defined by

$$M_D^n = A_1(1 : M, 1 : M) - A_2(n : n + M - 1, n : n + M - 1). \quad (5)$$

The difference matrices M_D^n represent the N possibilities to match the point sequences onto each other. Based on these N integral images $Int^1 \dots Int^N$ the difference values $D_\alpha(s, m, r)$ can be calculated for every block of any size starting at any point on the diagonal in constant time.

As a final result all matching triplets $\{s, m, r\}$ which provide a difference value $D_\alpha(s, m, r)$ below a fixed threshold are returned. Obviously, the detected matches may overlap. Thus, the final result is obtained by merging the different matches providing a set of connected point correspondences. This ability of matching sub-fragments between two input shapes allows to handle articulations, as it is illustrated in Figure 4, where for two scissors four different partially matched sub-fragments are returned.

2.4 Shape Similarity

It is important to provide a reasonable similarity measure in addition to the identified matching point sequences, e.g. for tasks like shape retrieval. Commonly, a combination of descriptor difference, matched shape distances like the Procrustes distance and bending energy of an estimated transformation like a Thin Plate Spline is used. Since we focus on partial similarity evaluation we adapt a measure described by Bronstein et al. [17]. They proposed to use the Pareto-framework for quantitative interpretation of partial similarity. They define two quantities: partiality $\lambda(X', Y')$, which describes the length of the parts

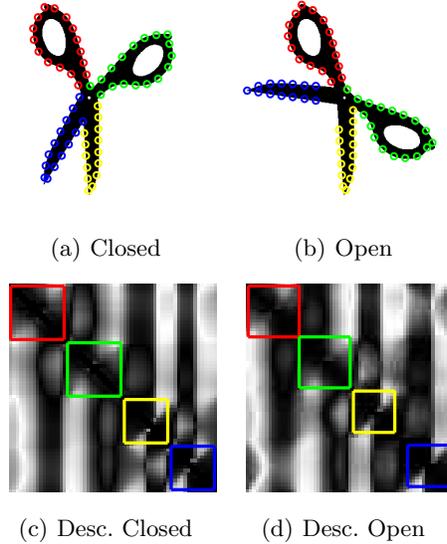


Fig. 4: Articulation invariance is handled by returning a set of partially matching boundary fragments. Corresponding fragments are depicted by the same color.

(the higher the value the smaller the part) and dissimilarity $\varepsilon(X', Y')$, which measures the dissimilarity between the parts, where X' and Y' are two contour parts of the shape. A pair $\Phi(X^*, Y^*) = (\lambda(X^*, Y^*), \varepsilon(X^*, Y^*))$ of partiality and dissimilarity values, fulfilling the criterion of lowest dissimilarity for the given partiality, defines a Pareto optimum. All Pareto optimums can be visualized as a curve, referred to as the set-valued Pareto frontier.

Since finding the Pareto frontier is a combinatorial problem in the discrete case, mostly rough approximations are used as final distance measure. Our matching algorithm automatically evaluates all possible matches for all possible lengths. Therefore, by focusing on the discretization defined by our point sampling, we can calculate a global optimal Pareto frontier, by returning the minimum descriptor difference for all partialities.

Finally, to get a single partial similarity value, the so-called Salukwadze distance is calculated based on the Pareto frontier by

$$d_s(X, Y) = \inf_{(X^*, Y^*)} |\Phi(X^*, Y^*)|_1, \quad (6)$$

where $|\dots|_1$ is the L1-norm of the vector. Therefore, $d_s(X, Y)$ measures the distance from the utopia $(0, 0)$ to the closest point on the Pareto frontier. The Salukwadze distance is then returned as the shape matching similarity score. Figure 5 illustrates the calculation of the global optimal Pareto frontier and Salukwadze distance.

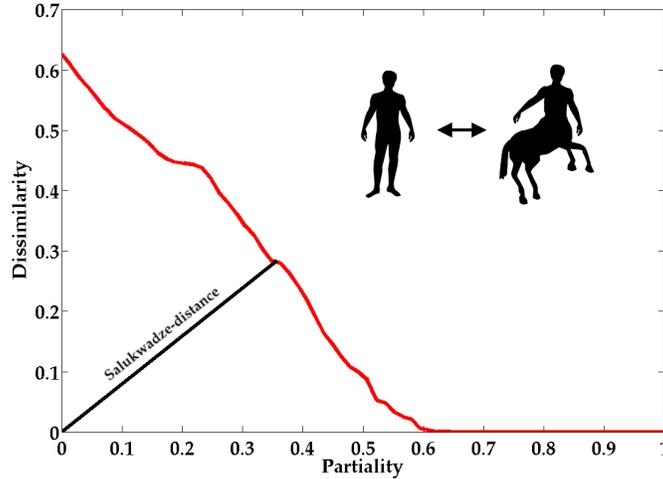


Fig. 5: *IS-Match* returns similarities for all possible matches and lengths which allows calculating a global optimal Pareto frontier. The Salukwadze distance is returned as partial similarity score.

2.5 Computational Complexity

An exhaustive search over all possible matches for all possible lengths has a complexity of $O(2^{n+m})$. Our proposed approach based on integral image analysis enables matching in $O(nm)$ time, where n and m are the number of sampled points on the two input shapes. We implemented our method in C, which enables shape matching on a desktop PC within milliseconds.

For comparison, Table 1 summarizes complexities and runtimes of current state-of-the-art shape matching methods. As it is shown in Section 3 only 30 sampled points are required to provide close to state-of-the-art shape retrieval results, which is possible within only 3 *ms*. Please note, that the runtimes may vary due to differences in implementations and machine configurations. But as can be seen in general *IS-Match* outperforms state-of-the-art concerning computational complexity and actual runtime. To the best of our knowledge this constitutes the fastest method for combinatorial matching of 2D shapes published so far.

3 Experimental Evaluation

To evaluate the overall quality of *IS-Match*, we first analyzed the influence of the number of sampled points and different parametrizations on the performance of shape retrieval on a common database in Section 3.1. The evaluation shows that only approximately 30 sampled points are required to achieve promising results, where a single match only requires 3 *ms* of computation time outperforming all

Table 1: Comparison of computational complexity and runtime in milliseconds for a single match. Please note, that as it is shown in Figure 6 our algorithm only requires 30 points to achieve competitive results on reference data sets.

Method	N	Complexity	Runtime
Felzenszwalb. [18]	100	$O(m^3k^3)$	500ms
Scott [10]	100	$O(mnl)$	450 ms
IDSC [9]	100	$O(m^2n)$	310ms
SC [8]	100	$O(m^2n)$	200ms
Schmidt [19]	200	$O(m^2\log(m))$	X
Brendel and Todorovic [5]	100	$O(nm)$	X
IS-Match	30	$O(nm)$	3ms

other shape matching algorithms by an order of magnitude. Section 3.2 shows results on the largest and currently most important benchmark for evaluating shape matching algorithms, the MPEG-7 database.

3.1 Performance Analysis

To evaluate the influence of the number of sampled points and different parameterizations we applied *IS-Match* for the task of shape retrieval on the common database of Sharvit et al. [20]. This database consists of 25 images of 6 different classes. Each shape of the database was matched against every other shape of the database and the global optimal Salukwadze distance as described in Section 2.4 was calculated for every comparison. Then for every reference image all the other shapes were ranked by increasing similarity value. To evaluate the retrieval performance the number of correct first-, second- and third ranked matches that belong to the right class was counted. In all the experiments Δ was set to 5, but experimental evaluations with different parameterizations revealed that changing Δ only has a small effect on shape retrieval performance.

Figure 6 illustrates the performance of our algorithm on this database, where the sum over all correct first-, second- and third ranked matches is shown. Therefore, the best achievable performance value is 75. We present results of *IS-Match* in dependence of the number of sampled points. As can be seen by sampling 30 points we achieve the highest score of 25/25, 25/25, 24/25 which represents state-of-the-art for this database as it is shown in Table 2.

3.2 Shape retrieval on MPEG-7 database

We further applied *IS-Match* to the MPEG-7 silhouette database [21] which is currently the most popular database for shape matching evaluation. The database consists of 70 shape categories, where each category is represented by 20 different images with high intra-class variability. The parametrization of our algorithm is based on the results shown in the previous section. The overall shape matching performance was evaluated by calculating the so-called bullseye

rating, in which each image is used as reference and compared to all of the other images. The mean percentage of correct images in the top 40 matches (the 40 images with the lowest shape similarity values) is taken as bullseye rating.

The measured bullseye rating for *IS-Match* was 84.79% and is compared to state-of-the-art algorithms in Table 4. As can be seen the score is close to the best ever achieved by Felzenszwalb et al. [18] of 87.70%. But please note that [18] uses a much more complex descriptor and requires about 500ms per match. Therefore, analyzing the entire database takes approximately 136 hours for [18], while with *IS-Match* all similarity scores are provided within a single hour (!).

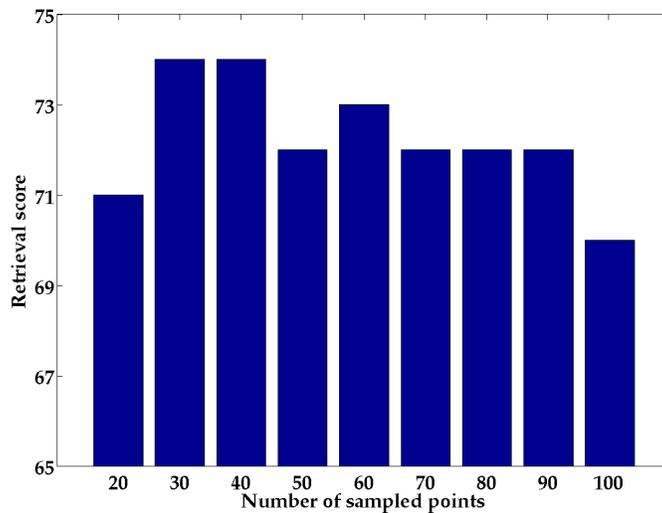


Fig. 6: Retrieval results in dependence of number of sampled points on database of [20] consisting of 25 shapes of 6 different classes. Maximum achievable score is 75.

Table 2: Comparison of retrieval rates on database of [20] consisting of 25 shapes of 6 different classes. The number of correct first-, second- and third ranked matches is shown.

Algorithm	Top 1	Top 2	Top 3	Sum
Sharvit et al. [20]	23/25	21/25	20/25	64
Belongie et al. [8]	25/25	24/25	22/25	71
Scott and Nowak [10]	25/25	24/25	23/25	72
Ling and Jacobs [9]	25/25	24/25	25/25	74
IS-Match	25/25	25/25	24/25	74

Table 3: Comparison of retrieval rates and estimated overall runtimes in hours (!) for calculating the full $N \times N$ similarity matrix on MPEG-7 database consisting of 1400 images showing 70 different classes.

Algorithm	Mokht. [22]	Belongie [8]	Scott [10]	Ling [23]	Felz. [18]	IS-Match
Score	75.44%	76.51%	82.46%	86.56%	87.70%	84.79%
Runtime	X	54 h	122 h	84 h	136 h	1 h

Table 4: Comparison of retrieval rates and estimated overall runtimes in hours (!) for calculating the full $N \times N$ similarity matrix on MPEG-7 database consisting of 1400 images showing 70 different classes.

Algorithm	Belongie [8]	Scott [10]	Ling [23]	Felz. [18]	Graph [22]	IS-Match
Score	76.51%	82.46%	86.56%	87.70%	88.59%	84.79%
Runtime	54 h	122 h	84 h	136 h	500 h	1 h

4 Conclusion

This paper introduced a partial shape matching method denoted as *IS-Match*. A chord angle based descriptor is presented which in combination with an efficient matching step allows detecting subparts of two shapes that possess high similarity. We proposed a fast integral image based implementation which enables matching two shapes within a few milliseconds. Shape retrieval experiments on common databases like the MPEG-7 silhouette database proved that promising results are achieved at reduced computational costs. Due to the efficiency of the proposed algorithm it is also suited for real-time applications as e.g. in action recognition by matching human silhouettes to reference prototypes or for tracking applications, which will be the focus of future work.

References

1. Opelt, A., Pinz, A., Zisserman, A.: A boundary-fragment-model for object detection. In: Proceedings of European Conference on Computer Vision (ECCV). Volume 2. (2006) 575–588
2. Felzenszwalb, P.: Representation and detection of deformable shapes. In: Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR). Volume 1. (2003) 102–108

3. Ferrari, V., Tuytelaars, T., Gool, L.V.: Object detection by contour segment networks. In: Proceedings of European Conference on Computer Vision (ECCV). Volume 3. (2006) 14–28
4. Mori, G., Belongie, S., Malik, H.: Shape contexts enable efficient retrieval of similar shapes. In: Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR). (2001) 723–730
5. Sebastian, T.B., Klein, P.N., Kimia, B.B.: On aligning curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* **25**(1) (2003) 116–125
6. Kuijper, A., Olsen, O.F.: Describing and matching 2D shapes by their points of mutual symmetry. In: Proceedings of European Conference on Computer Vision (ECCV). Volume 3. (2006) 213–225
7. Sebastian, T.B., Klein, P.N., Kimia, B.B.: Recognition of shapes by editing shock graphs. In: Proceedings of International Conference on Computer Vision (ICCV). Volume 1. (2001) 755–762
8. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. *IEEE Transactions of Pattern Analysis and Machine Intelligence (PAMI)* **24**(4) (2002) 509–522
9. Ling, H., Jacobs, D.W.: Using the inner-distance for classification of articulated shapes. In: Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR). Volume 2. (2005) 719–726
10. Scott, C., Nowak, R.: Robust contour matching via the order-preserving assignment problem. *IEEE Transactions on Image Processing* **15**(7) (2006) 1831–1838
11. Chen, L., Feris, R., Turk, M.: Efficient partial shape matching using smith-waterman algorithm. In: Proceedings of NORDIA Workshop at CVPR. (2008)
12. Donoser, M., Bischof, H.: Fast non-rigid object boundary tracking. In: Proceedings of British Machine Vision Conference (BMVC). (2008)
13. Weinland, D., Boyer, E.: Action recognition using exemplar-based embedding. In: Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR). (2008)
14. Cootes, T., Cooper, D., Taylor, C., Graham, J.: Trainable method of parametric shape description. *Journal of Image Vision Computing* **10**(5) (June 1992) 289–294
15. Osada, R., Funkhouser, T., Chazelle, B., Dobkin, D.: Shape distributions. *ACM Transactions on Graphics* **21** (2002) 807–832
16. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR). (2001) 511–518
17. Bronstein, A., Bronstein, M., Bruckstein, A., Kimmel, R.: Partial similarity of objects, or how to compare a centaur to a horse. *International Journal of Computer Vision* (2008)
18. Felzenszwalb, P.F., Schwartz, J.D.: Hierarchical matching of deformable shapes. In: Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR). (2007)
19. Schmidt, F.R., Farin, D., Cremers, D.: Fast matching of planar shapes in sub-cubic runtime. In: Proceedings of International Conference on Computer Vision (ICCV). (2007)
20. Sharvit, D., Chan, J., Tek, H., Kimia, B.: Symmetry-based indexing of image database. *Journal of Visual Communication and Image Representation* **9**(4) (1998)
21. Latecki, L.J., Lakämper, R., Eckhardt, U.: Shape descriptors for non-rigid shapes with a single closed contour. In: Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR). (2000) 424–429

22. Mokhtarian, F., Abbasi, S., Kittler, J.: Efficient and robust retrieval by shape content through curvature scale space. In: Proc. of International Workshop on Image Databases and Multimedia Search. (1996) 35–42
23. Ling, H., Okada, K.: An efficient earth movers distance algorithm for robust histogram comparison. IEEE Transaction on Pattern Analysis and Machine Intelligence (PAMI) **29** (2007) 840–853