World Scientific
www.worldscientific.com

EFFECT OF HIGH CURVATURE POINT DELETION ON THE PERFORMANCE OF TWO CONTOUR BASED SHAPE RECOGNITION ALGORITHMS

ANARTA GHOSH* and NICOLAI PETKOV[†]

Institute of Mathematics and Computing Science
University of Groningen, P.O.Box. 800
9700 AV Groningen, The Netherlands
*anarta@cs.rug.nl
†petkov@cs.rug.nl

Psychophysical researches on the human visual system have shown that the points of high curvature on the contour of an object play an important role in the recognition process. Inspired by these studies we propose: (i) a novel algorithm to select points of high curvature on the contour of an object which can be used to construct a recognizable polygonal approximation, (ii) a test which evaluates the effect of deletion of contour segments containing such points on the performance of contour based object recognition algorithms. We use complete contour representations of objects as a reference (training) set. Incomplete contour representations of the same objects are used as a test set. The performance of an algorithm is reported using the recognition rate as a function of the percentage of contour retained. We consider two types of contour incompleteness obtained by deletion of contour segments of high or low curvature. We illustrate the test procedure using two shape recognition algorithms that deploy a shape context and a distance multiset as local shape descriptors. Both algorithms qualitatively mimic human visual perception in that the deletion of segments of high curvature has a stronger performance degradation effect than the deletion of other parts of the contour. This effect is more pronounced in the performance of the shape context method.

Keywords: Contour deletion; distance multiset; high curvature points; ICR test; performance evaluation; robustness to incompleteness; shape context; shape recognition.

1. Introduction

Psychologist Gollin[9] investigated the human ability to recognize objects from incomplete contour representations. The main objective of his study was to determine the influence of developmental characteristics, such as mental and chronological age and intelligence quotient. As subjects of his experiments he chose children of different age groups and a group of adults. In his experiments Gollin used sets of contour images with different degrees of incompleteness and addressed the following questions: (1) In order to be recognized, how complete the contours of common objects need to be? (2) How does training affect the recognition performance in

case of incomplete representations? Through his experiments he found that human ability to recognize objects with incomplete contours (a) depends on intelligence quotient and (b) is improved by training.

This aspect of recognition of objects with incomplete contours is also very important in the context of processing visual information using computers due to the fact that present contour detection algorithms[11,12] fail to extract a closed and neat contour of an object from a natural scene.[8,7]

Shape descriptor based object recognition methods have been evaluated and compared using various characteristics like recognition performance,[2,10,13] invariance, uniqueness and stability.[16] Marr and Nishihara[15] proposed three criteria for judging the effectiveness of a shape descriptor, viz. accessibility, scope and uniqueness, stability and sensitivity. Brady[4] put forward a set of criteria for representation of shape, viz. rich local support, smooth extension and propagation. Inspired by Gollin's study and keeping the aforementioned practical importance in mind, in our previous studies[8,7] we proposed the Incomplete Contour Representations (ICR) test to evaluate the robustness of contour based object recognition algorithms to incompleteness of extracted contours. In those studies, we did not take into consideration the fact that points along a contour can have different importance for human perception and possibly for computer algorithms. In the present paper we investigate whether the deletion of high curvature contour segments has a greater effect on the performance of two shape recognition algorithms than the deletion of low curvature segments. Our motivation originates in the fact that a similar question has been posed and answered regarding human visual perception.

In 1954, psychologist Attneave[1] proposed that perceptually important information along a visual contour is concentrated in the regions of high curvature rather than distributed uniformly. He took an image of a cat, selected points of high contour curvature and joined them with straight line segments. The resulting line drawing (now popularly known as *Attneave's cat*) is easily recognizable, proving that not much information is lost due to the linear approximation of the contour between the points of high curvature. We illustrate the essence of his considerations in Figs. 1(a) and 1(b). Attneave[1] also described the results of an experiment in which participants were asked to approximate a two-dimensional contour with a fixed number of points and to locate these points in the original shape. He found the points were selected more frequently in the regions of high curvature in the original contour. Moreover, experiments with partial contour deletion by Biederman[3] have shown that the deletion of high curvature contour segments creates greater difficulties in recognition than deletion of low curvature segments of comparable lengths when presentation time is short ($\sim$ 100 msec), c.f. Figs. 1(c) and 1(d).

In the current paper we report the results of an extension of our previous studies on the robustness of shape recognition algorithms to incompleteness of contours.[8,7] Similar to those studies we choose an idealized situation where: (a) complete contour representations of the objects to be recognized form the reference (training) set or "memory" of the system/algorithm, (b) incomplete contour representations of the
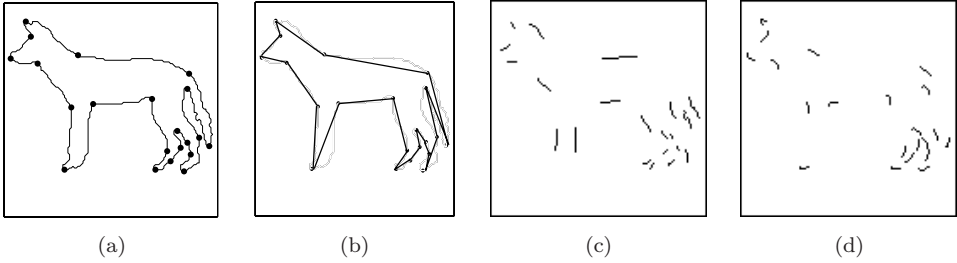
Fig. 1. (a) Contour image of a dog with high curvature points marked by dots. (b) Points of high curvature are joined by straight lines to construct a contour image similar to Attneave's cat. (c) Incomplete contour image in which the contour segments of high curvature are deleted. (d) Segments between the points of high curvature are deleted to construct this incomplete contour image. Though the percentages of contour deletion are the same in (c) and (d), it is yet easier to recognize the contour image in (d) as a dog due to the retainment of contour segments of high curvature. The shorter the presentation time, the more important is the presence of such contour segments for recognition by humans.[3]

same objects are derived from the aforementioned complete representations and are used as a test set, (c) the performance of the system/algorithm in recognizing the objects from these incomplete representations is evaluated.

In this study we use only one of the three ICR tests defined in our previous work,[8,7] viz. the segment-wise deletion test. In contrast to the previous studies where contour segments are deleted at random to construct an incomplete representation, in the current study contour segments are removed in two different ways by deleting segments of (i) high or (ii) low curvature [Figs. 1(c) and 1(d)].

We investigate the performance of two contour based shape recognition methods, which use a shape context[2] and a distance multiset[10] as shape descriptors. The two methods are briefly reviewed in Sec. 2. In Sec. 3, we describe the method used to obtain points of high curvature. The achieved results are presented in Sec. 4 and discussed in Sec. 5. A summary and conclusions are presented in Sec. 6.

2. Shape Recognition Methods

In the methods studied below, the recognition of objects is done by computing dissimilarity between the contour representations of two objects by using a point correspondence paradigm. Shape descriptors associated with the points are used to find the point correspondences. To maintain brevity and focus on the illustration of the proposed test we use simplified versions of two methods.

The **distance multiset**[10] of a point p in the contour of an object $\mathcal{O}$ of N points is formally defined as the following vector: $D_N^{\mathcal{O}}(p) = (ln(d_1(p)), ln(d_2(p)), \ldots, ln(d_{N-1}(p)))$ where $d_j(p)$ is the Euclidean distance between p and its jth nearest neighbor in $\mathcal{O}$. In this approach the shape of an object $\mathcal{O} \equiv \{p_1 \ldots p_N\}$ defined by a set of contour points is described by the collection of distance multisets of these points, $S_{\mathcal{O}}^{DM} \equiv \{D_N^{\mathcal{O}}(p)|p \in \mathcal{O}\}$. The cost $c(X, Y)$ of matching two distance

multisets $X = (x_1, x_2, \ldots, x_M)$ and $Y = (y_1, y_2, \ldots, y_N)$, $M \leq N$, is defined as the L_1 norm of the difference between X and a subvector of Y of M elements for which this norm is minimum.[8]

The cost $c_{i,j}^{DM}$ of matching a point p_i in an object $\mathcal{O}_1$ represented by M contour points to a point q_j in an object $\mathcal{O}_2$ represented by N contour points, $M \leq N$, is defined as the cost of matching the corresponding distance multisets: $c_{i,j}^{DM} \equiv c(D_N^{\mathcal{O}_1}(p_i), D_M^{\mathcal{O}_2}(q_j))$. The dissimilarity between two shapes is defined as the sum of the costs of pairwise point matching for the mapping from $\mathcal{O}_1$ into $\mathcal{O}_2$ for which that sum is minimum. The problem of finding the optimal point correspondences for which this sum is minimum can be solved using combinatorial optimization algorithms, e.g. the Hungarian algorithm.[2,10] In this study we do not do that for simplicity. Instead we compute an approximation to the optimal dissimilarity between the shapes $S_{\mathcal{O}_1}^{DM}$ and $S_{\mathcal{O}_2}^{DM}$ as follows: $d^{DM}(S_{\mathcal{O}_1}^{DM}, S_{\mathcal{O}_2}^{DM}) \equiv \sum_{i=1}^{M} \min\{c_{i,j}^{DM} | j = 1 \cdots N\}$.

The **shape context**[2] of a point p belonging to the contour of an object is a bi-variate histogram in a log-polar coordinate system that gives the distribution of contour points in the surroundings of p. Let an object $\mathcal{O}$ be represented by a set of contour points $\mathcal{O} \equiv \{p_1 \ldots p_N\}$. Formally, the authors of this method define the shape context of a point $p \in \mathcal{O}$ as a vector in the following way: $H_K^{\mathcal{O}}(p) = (h_1(p), h_2(p), \ldots, h_K(p))$, where $h_k(p) = \text{card}\{q \neq p | q \in \mathcal{O}, (q - p) \in \text{bin}(k)\}$ is the number of contour points in the kth bin — $\text{bin}(k)$ — and K is the total number of histogram bins. The bins are constructed by dividing the image plane into K partitions (in a log-polar coordinate system) with p as the origin. In this study we use five intervals for the log distance and 12 intervals for the polar angle, so $K = 60$. As suggested by the authors of that method,[2] we randomly choose 100 points (if available) from the contour of an object and calculate their shape contexts. The shape of the object is described using the set of shape contexts associated with the contour points, $S_{\mathcal{O}}^{SC} \equiv \{H_K^{\mathcal{O}}(p) | p \in \mathcal{O}\}$. The cost of matching a point p_i that belongs to the contour of an object $\mathcal{O}_1$ of M points, to a point q_j from the contour of an object $\mathcal{O}_2$ of N points is defined as follows: $c_{i,j}^{SC} \equiv \frac{1}{2} \sum_{k=1}^{K} \frac{[h_k(p_i) - h_k(q_j)]^2}{h_k(p_i) + h_k(q_j)}$, which yields an $M \times N$ cost matrix of point-wise dissimilarities. Similar to the distance multiset method we compute the dissimilarity between the shapes $S_{\mathcal{O}_1}^{SC}$ and $S_{\mathcal{O}_2}^{SC}$ of the objects in the following way: $d^{SC}(S_{\mathcal{O}_1}^{SC}, S_{\mathcal{O}_2}^{SC}) \equiv \sum_{i=1}^{M} \min\{c_{i,j}^{SC} | j = 1, \ldots, N\}$.

3. Extraction of High Curvature Points

We define as feature points, selected points of local curvature extrema on the contour of an object. We consider signed curvature where positive and negative values represent convexity and concavity, respectively, and use an algorithm for curvature computation proposed by Feldman and Sing.[5] Figures 2(a) and 2(b) illustrate the results of this computation. There are numerous local curvature extrema from which we select a subset as follows: for each contour point p_i which is a local
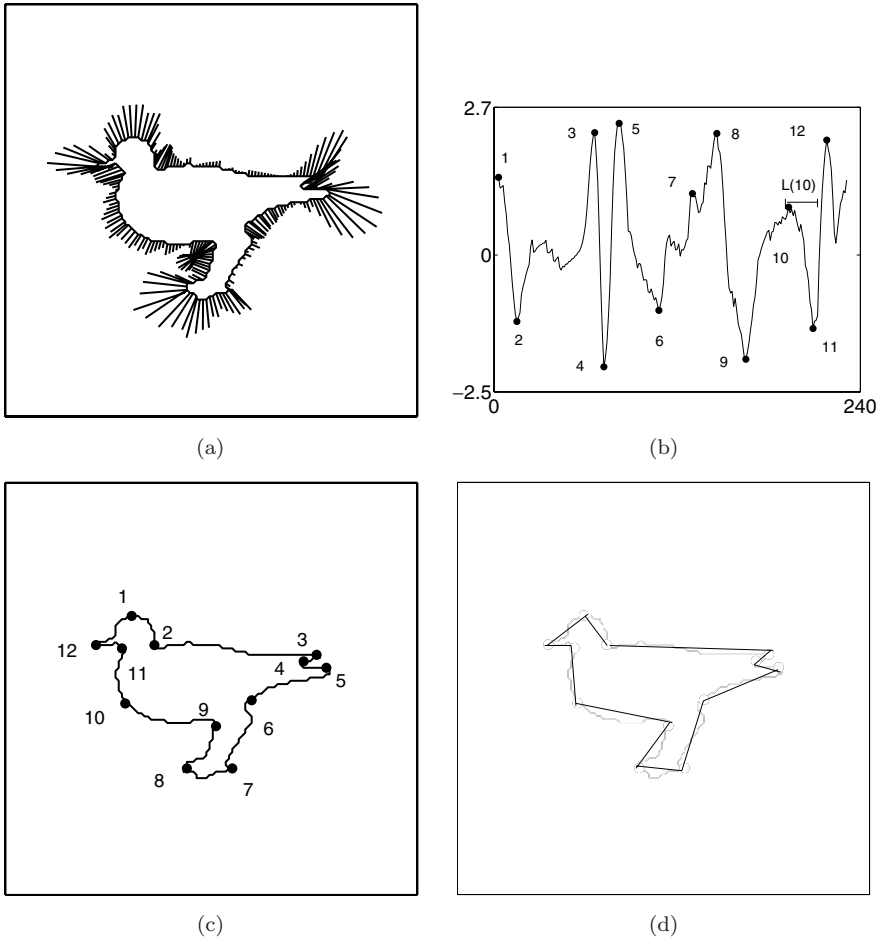
(a)

(b)

(c)

(d)

Fig. 2.   Extraction of feature points. (a) Needle diagram: the size of a protruding needle at a given contour point represents the magnitude of the curvature at that point. (b) Curvature κ plotted as a function of the length along the contour, starting at the top-left point and traversing the contour in clockwise direction. The dots denote the selected local extrema which we call the feature points. For feature point 10 the length $L(10)$ of the contour part within which that point is a curvature extremum (maximum in this case) is shown. (c) Feature points corresponding to the significant local extrema for a threshold $\Delta = 6$. (d) Polygonal approximation of the bird contour constructed by joining the feature points with straight line segments.

extremum of the curvature function κ, we compute a product $r(i) = |\kappa(i)|L(i)$ of the absolute value of the curvature $\kappa(i)$ at point p_i with the length $L(i)$, measured by the number of contour pixels, of the contour part from p_i to the nearest point along the contour that has the same curvature, Fig. 2(b). For the special case when the point is a global extremum $L(i)$ is chosen to be half the total number of pixels on the contour of the object. We select a contour point p_i which is a local curvature extremum as a feature point only if the aforementioned product $r(i)$ exceeds
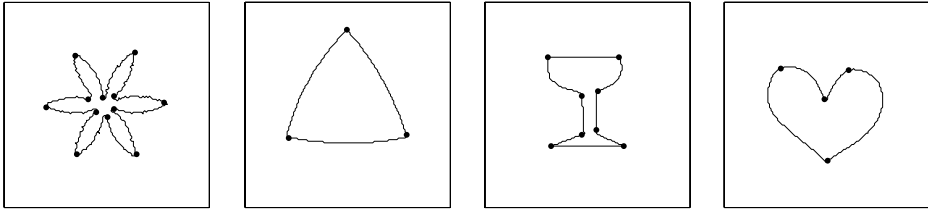
Fig. 3.   Examples of rescaled contour images obtained from the MPEG-7 silhouette database. These images are considered as complete representations that comprise the memory of the recognition system. The dots on the contour denote the selected feature points using the method described in Sec. 3.

a certain threshold Δ, $r(i) \geq \Delta$, Fig. 2(c). In our experiments with a subset of the MPEG-7 dataset, examples of which are shown in Fig. 3, we use a fixed threshold which is the same for all images in the dataset and is chosen in such a way that it results in at least three feature points per image. The actual number of obtained feature points varies from 3 to 21 across the dataset. The result of this selection looks similar to what may be expected from an Attneave-like[1] experiment with a human observer, but showing quantitative correlation with human perception is beyond the scope of this study.[a] Indeed, the obtained feature points allow a reasonable polygonal approximation, similar to Attneave's reconstruction. The minimum number (three) of feature points per object used to determine the threshold was chosen based on such considerations and reflects the simplicity of some of the objects in the used dataset, Fig. 3.

4. Experiments and Results

4.1. *Image set*

We choose silhouette images from the MPEG-7 database.[14] This dataset contains 1400 images divided into 70 classes, each of 20 similar objects (e.g. apple, bird, bat, etc). We choose one object from each class and extract the contours of the objects using Gabor filters.[11] The resulting 70 contour images are rescaled in such a way that the diameter (maximum Euclidean distance between contour pixels) is approximately the same (76 pixels) for all objects (Fig. 3). These 70 rescaled contour images are used as reference images or prototypes in our experiments and constitute the "memory" of the recognition system.

Next we construct two sets of images of objects with incomplete contours. One set is constructed by removing contour segments that include the feature points, Fig. 4. For brevity we refer to them as feature segments in the following. The other set is obtained by deleting contour segments with low curvature between the feature

---

[a]The problem of selection of points on the contour of an object has been discussed in a different context (curve partitioning) elsewhere.[6] For other methods we refer to the work by Wu[17] and references therein.
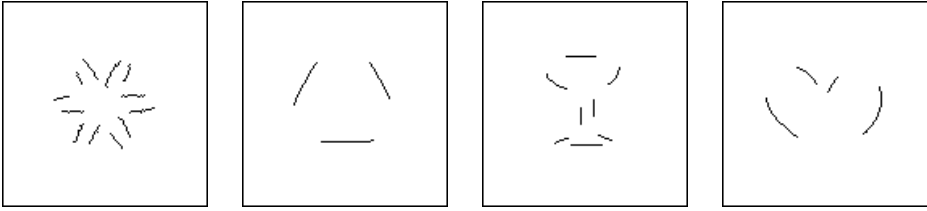
Fig. 4.   Incomplete contour representations in which feature segments, i.e. segments containing feature points are deleted, compare with Fig. 3. In these examples 40% of the contour is retained.
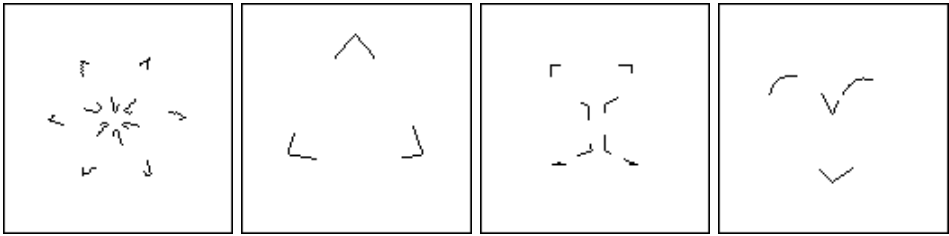

Fig. 5.   Incomplete contour representations in which midsegments, i.e. contour segments between the feature points, are deleted, compare with Fig. 3. 40% of the contour points are retained in these examples.

points, Fig. 5. Following Biederman,[3] we call such contour parts midsegments. The percentages of retained pixels in both types of incomplete contour representations are chosen in the following way: in steps of 5% from 5% to 85%, and 100%.

4.2. *Methodology*

A test image (incomplete contour representation of an object) obtained from one of the 70 reference images in one of the above mentioned ways is compared with all 70 reference images using a given shape comparison algorithm and a decision is taken about which reference image the degraded image is most similar to (nearest neighbor search). If the nearest neighbor is the reference image from which the degraded image was obtained, the recognition result is considered correct, otherwise incorrect. If the nearest neighbor is found to be not unique then the recognition is also considered incorrect. For each of the two types of incomplete contour representations and for each degree of contour image degradation, the corresponding 70 test images are compared with each of the 70 reference images and the percentage of correct recognition $P(c)$ is determined as a function of the percentage c of retained contour points.

4.3. *Results*

The results of our experiments are shown in Fig. 6. The recognition rate is a monotonic increasing function of the percentage of contour retainment for both
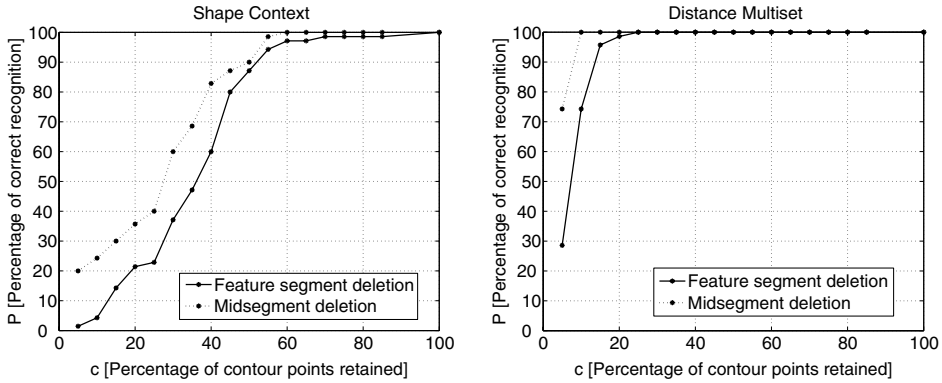
Fig. 6.   (Left panel) Results obtained for the shape context method — the effect of feature point deletion is significant. (Right panel) Results obtained with the distance multiset method — the effect of feature point deletion is substantial only for high degree of incompleteness.

algorithms and both types of deletion. For both algorithms the deletion of high curvature segments has a stronger performance degradation effect than the deletion of midsegments. This effect gets more pronounced at higher degree of contour incompleteness. The performance of the distance multiset method is appreciably better than that of the shape context method for any percentage of retained contour pixels and both types of deletion.

5. Discussion

In his experiments Biederman[3] found that the deletion of high curvature points (he called vertices) has a stronger degradation effect on the recognition performance of humans than the deletion of low curvature contour segments (he called midsegments) only when the presentation times were short (100 msec). For longer presentation times (750 msec) he found no such difference. To explain his findings, Biederman hypothesized a contour filling-in process: if sufficient time is given, midsegments are extrapolated until they meet, forming segments of high curvature which are then used to recognize an object. Contour segments of high curvature are thus essential for recognition of objects by the human visual system.

The results presented in Fig. 6 show that high curvature segments have a similar importance for the shape recognition algorithms studied here. However, we should note that these are statistical results obtained with many objects, showing that *on average* the deletion of high curvature segments has a stronger performance degradation effect than the deletion of midsegments. There are, however, objects in the dataset for which the opposite is true: the deletion of midsegments leads to false recognition while the deletion of an equal amount of high curvature segments does not have such an effect. Therefore, there is no simple relation between the design of the considered algorithms and the observed effect. Hence, in the following we only
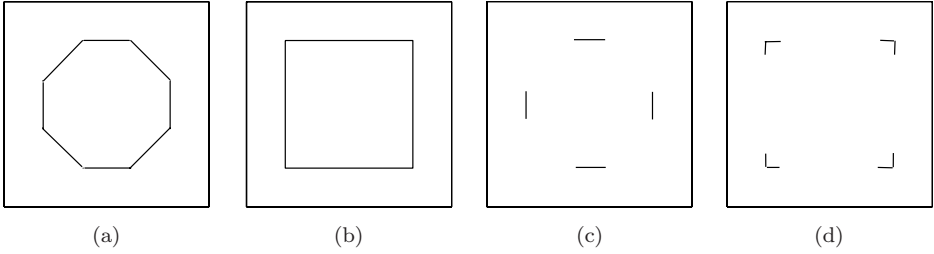
Fig. 7. Contour images of (a) an octagon and (b) a square. (c–d) Incomplete contour representations derived from the square by deleting contour segments with (c) high and (d) low curvatures. The set of contour points shown in (c) is a subset of the sets of contour points shown in (a) and (b). This will result in a misclassification of (c) by the distance multiset method. In contrast, the incomplete representation shown in (d) in which the high curvature contour segments are retained will not be misclassified because it has a unique nearest neighbor — the square.

present two examples to illustrate the possible ways via which the observed effect can arise in concrete cases.

Figures 7(a) and 7(b) show the complete representations of an octagon and a square and Figs. 7(c) and 7(d) present incomplete representations of the square obtained by deleting contour segments of high and low curvatures, respectively. The comparison of the incomplete representation shown in Figs. 7(c) with the complete representations presented in Figs. 7(a) and 7(b) using the distance multiset method will yield zero dissimilarity.[b] Since the nearest neighbor in the reference set is not unique, the incomplete contour image in Fig. 7(c) will be considered as misclassified in the proposed evaluation procedure. No such recognition error will be made by the distance multiset method if the incomplete representation shown in Fig. 7(d) is used: its comparison with Figs. 7(a) and 7(b) will yield nonzero and zero dissimilarity, respectively, and it will be classified as a square.

Figure 8 shows another example. In this case, the deletion of high curvature segments from the contours of an apple and a pocket watch results in incomplete circles that can be classified either as a pocket watch or an apple by the shape context method. The chance of correct classification is, however, as high as the chance of misclassification.

From an information theoretic point of view a point of high curvature value on the contour of an object carries more information than a point on a straight line segment because the curvature value encodes for a contour orientation change in the former versus no such change in the latter case.

The performance of the algorithms also depends on the number of feature points selected. As mentioned in Sec. 3, this number increases when the threshold Δ in the feature point extraction algorithm is decreased. In our experiments we found

---

[b]As shown elsewhere[8] the distance multiset method will yield a zero dissimilarity in the comparison of a shape defined by a set A of points with another shape that is defined by a subset C of the first set: $d^{DM}(S_C^{DM}, S_A^{DM}) = 0$ if $C \subset A$.
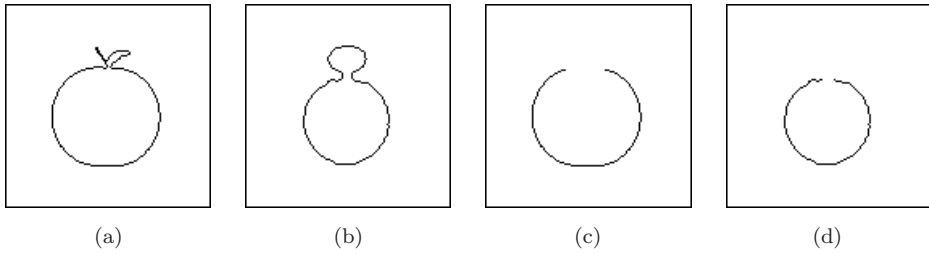
Fig. 8. Contour images of (a) an apple and (b) a pocket watch. (c–d) Incomplete contour representations derived from the contour images in (a) and (b) respectively by deleting contour segments with high curvature. Removal of high curvature extremities results in shapes that are indistinguishable.

that for a given percentage of contour retainment the performance of the algorithms improves with the number of contour parts present in the incomplete representations for both (feature segment and midsegment deletion) types of incompleteness. Since the number of contour parts in the incomplete representations is the same as the number of feature points selected, the higher the number of selected feature points the better is the performance of the algorithms. Hence, the results of the proposed test should be reported along with the range of the number of feature points used to construct the incomplete contour representations. In this study this number varies from 3 to 21.

6. Summary and Conclusion

In the present work we further elaborate on one of the ICR tests previously proposed[8,7] by distinguishing between deletion and retainment of contour segments containing points of high curvature. For illustration we tested two shape recognition methods based on the shape context and the distance multiset. In our experiments we found that both methods perform similar to the human visual system for short presentation time in that the deletion of segments of high curvature has a stronger degradation effect on the recognition performance than the deletion of low curvature segments. This phenomenon is more pronounced in the shape context method.

The main contributions of the research presented in this paper are: (A) a test procedure to evaluate the effect of deletion of segments containing high curvature feature points on the performance of contour based object recognition algorithms, (B) an algorithm to select high curvature feature points in the contour images of objects, and (C) characterization of the performance of the shape context and distance multiset shape recognition algorithms in the proposed test.

References

1. F. Attneave, Some informational aspects of visual perception, *Psychophys. Rev.* **61** (1954) 183–193.

2. S. Belongie, J. Malik and J. Puzicha, Shape matching and object recognition using shape contexts, *IEEE Trans. Patt. Anal. Mach. Intell.* **24**(4) (2002) 509–522.

3. I. Biederman, Recognition-by-components: a theory of human image understanding, *Psychophy. Rev.* **94**(2) (1987) 115–147.

4. M. Brady, Criteria for representations of shape, in *Human and Machine Vision*, eds. J. Beck, B. Hope and A. Rosenfeld (Academic Press, 1983), pp. 39–84.

5. J. Feldman and M. Singh, Information along contours and object boundaries, *Psychophy. Rev.* **112** (2005) 243–252.

6. M. A. Fischer and R. C. Boles, Perceptual organization and curve partitioning, *IEEE Trans. Patt. Anal. Mach. Intell.* **8**(1) (1986) 100–105.

7. A. Ghosh and N. Petkov, Incomplete contour representations and shape descriptors: ICR test studies, in *Proc. First Int. Symp. Brain, Vision and Artificial Intelligence BVAI 2005*, Naples, October 19–21, 2005, Lecture Notes in Computer Science, Vol. 3704 (Springer-Verlag, Berlin, Heidelberg, 2005), pp. 416–425.

8. A. Ghosh and N. Petkov, Robustness of shape descriptors to incomplete contour representations, *IEEE Trans. Patt. Anal. Mach. Intell.* **27**(11) (2005) 1793–1804.

9. E. S. Gollin, Developmental studies of visual recognition of incomplete objects, *Percep. Motor Skills.* **11** (1960) 289–298.

10. C. Grigorescu and N. Petkov, Distance sets for shape filters and shape recognition, *IEEE Trans. Imag. Process.* **12**(10) (2003) 1274–1286.

11. C. Grigorescu, N. Petkov and M. Westenberg, Contour detection based on nonclassical receptive field inhibition, *IEEE Trans. Imag. Process.* **12**(7) (2003) 729–739.

12. C. Grigorescu, N. Petkov and M. A. Westenberg, Contour and boundary detection improved by surround suppression of texture edges, *Imag. Vis. Comput.* **22**(8) (2004) 609–622.

13. A. C. Jalba, M. H. F. Wilkinson and J. B. T. M. Roerdink, Shape representation and recognition through morphological curvature scale spaces, *IEEE Trans. Imag. Process.* **15**(2) (2006) 331–341.

14. L. J. Latecki, R. Lakämper and U. Eckhardt, Shape descriptors for non-rigid shapes with single closed contour, in *Proc. IEEE Conf. Computer Vision and Pattern Recognition* (1998), pp. 424–429.

15. D. Marr and H. K. Nishihara, Representation and recognition of the spatial organization of three-dimensional shapes, in *Proc. Roy. Soc. London, B* **200** (1978) 269–294.

16. F. Mokhtarian and A. K. Mackworth, A theory of multiscale, curvature-based shape representation for planar curves, *IEEE Trans. Patt. Anal. Mach. Intell.* **14**(8) (1992) 789–805.

17. W. Y. Wu, An adaptive method for detecting dominant points, *Patt. Recogn.* **36**(10) (2003) 2231–2237.

Anarta Ghosh received his B.Sc. (1995) with honours in physics (1st class) from Calcutta University, India. He obtained the Master of Engineering (2000) degree in electrical engineering from the Indian Institute of Science, Bangalore. Presently he is working as a Ph.D. student in the Institute of Mathematics and Computing Science, University of Groningen, The Netherlands.

His research interests are in the area of cognitive computer vision, pattern recognition, machine learning, signal processing and multimedia systems.

Nicolai Petkov is Scientific Director (Head) of the Research Institute of Mathematics and Computing Science, University of Groningen, The Netherlands, where he also holds the Chair of Intelligent Systems. He is author of two books and 90 scientific publications.

His current research interests are in the area of computer simulations of the visual system, making links between computer vision, neurophysiology, psychophysics and arts.