Extracting the Canonical Set of Closed Contours Using the Best-First Search Algorithm

Siniša Šegvić, Zoran Kalafatić and Slobodan Ribarić
Faculty of Electrical Engineering and Computing
Unska 3, 10000 Zagreb, Hrvatska
e-mail: sinisa.segvic@fer.hr; phone: +385 1 6129 935; fax: +385 1 6129 653

ABSTRACT

We address the problem of finding an intermediate-level representation of structured objects based on the output of an edge detector. The adopted approach for grouping edge elements is based on three principles of perceptual organization: proximity, continuation and closure. We are interested in a special subset of all closed contours in the image which is called canonical since it exhibits the desired properties of completeness, uniqueness, and minimality. Unfortunately, the elements of edge maps are often discontinuous where a single contour is perceived and vice versa, which imposes the need for devising a robust extracting procedure. Unlike the known approaches, the proposed procedure is formulated as the state-space search problem, which allows complex schemes for calculating compound saliency (the measure of perceptual importance) from the saliencies of individual junctions. The shortcoming of such an approach is the exponential worst-case complexity which we try to alleviate by several strategies for pruning the search tree. The procedure was tested on many synthetic images and a representative set of processing results is provided.

1. INTRODUCTION

The aim of techniques for perceptual grouping is to identify image parts which likely correspond to distinct structures of the world. Instead of using any specific knowledge about the analysed scene, these techniques perform grouping with respect to principles which are believed to be exploited by the human visual system [1]. These *Gestalt* principles include similarity, proximity, continuation, closure and symmetry, and are named after the German school of psychology whose members performed early research on perception. Unfortunately, implementations of these techniques are computationally very complex, which partly explains why they received relatively poor attention from the vision community until recently. In particular, considerable interest has been addressed to applying principles of proximity, continuation and closure for extracting intermediate features in both intensity [1, 2, 3, 4, 5] and range images [6, 7]. The closure principle has been applied for enhancing both edge and area segmentations, i.e. for disambiguation of the missing fragments of edge chains due to noise effects, as well as for finding regions in the image corresponding to distinct objects in the scene.

One of the first applications of the closure principle to object recognition was contributed by Jacobs [2] who investigated the problem of indexing of partially occluded convex objects in real images. The proposed approach relies on the list of corner points obtained as intersections of extracted straight line segments. Elder and Zucker [1] described a generic procedure for extracting arbitrary closed contours from real images. The procedure represents curves as sequences of disjoint line segments (these are also called super-segments [8]) and builds incidency matrix which for each super-segment pair provides the saliency of their concatenation. Thus, the image is represented as a graph, while the process of finding a closed contour is formulated and solved as the shortest-path problem. Williams and Thornber [4] provide a review of existing procedures for expressing saliency of an edge element in terms of its affinities to its neighbours which are based on principles of continuation and proximity. They also propose a novel procedure which takes advantage of the principle of closure and, according to their experiments, performs better than the others. An application of their measure to finding multiple simple objects in cluttered real images by segmenting out closed contours is proposed by Mahamud et al. [5].

The two latter approaches [4, 5] propose a complex model to enforce continuation and closure of the extracted contours, however, little attention is addressed towards dealing with situations in which multiple feasible closed contours pass through the same edge. Compared to [1], our approach trade-offs the computational simplicity in return for the enhanced capability of modelling the compound saliency of a closed contour. Additionally, we propose a strategy for extracting a special *canonical* configuration of multiple possibly neighbouring closed contours which is complete, minimal and unique. This issue is particularly important if the results of perceptual organization are employed for solving a higher level perceptual task such as establishing correspondences between neighbouring frames in a sequence. Our approach is similar to [3], in which the problem is also solved by the informed search; however, the procedure proposed in [3] comes short of exploiting capabilities offered by the search approach, and is unable to extract arbitrary closed contours.

The detailed description of the problem is given in Section 2, while Section 3 describes the procedure for extracting canonical subset of the closed contours, based on the output of the Canny edge detector. Finally, experimental results are provided in Section 4.

2. PROBLEM DESCRIPTION

Most methods which have been proposed for overcoming the difficulties tied with imperfect edge maps by perceptual organization use only local principles of proximity and continuation [4]. This is surprising since these local cues may be clearly incorrect in global context, while the closure as a global principle may disambiguate many of such situations. Consider for example the image in Fig.1a, and its enlarged central part in Fig.1b. Without global context, there is no way of knowing whether we have two black or two white touching objects in the image. The correct interpretation of the image can be obtained by formulating global context as a rule that the closed contours are more salient than the open ones.

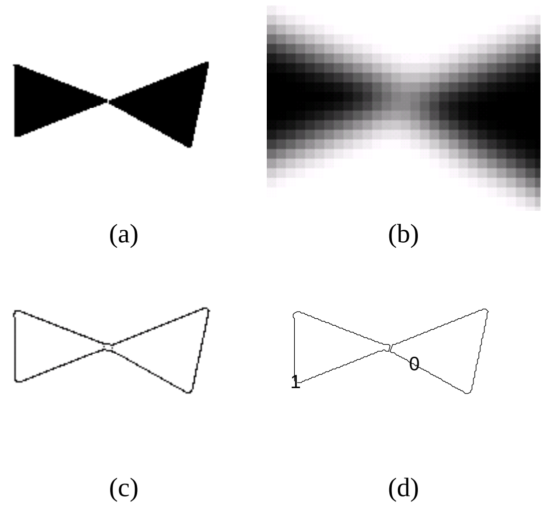

Figure 1: A synthetic test image (a), the enlarged central portion of the image smoothed with $\sigma = 2$ (b), the edge map provided by the Canny edge detector (c), and the extracted canonical set of 2 closed contours (d) labeled with “0” and “1”.

The method of grouping edge chains according to the closure principle may have difficulties when dealing with structured objects, which give rise to multiple feasible closed contours. For instance, consider the image in Fig.2a, featuring three adjacent homogeneous regions and 7 feasible closed contours. Which of these contours should be reported to a higher level perceptual procedure? A bad approach would be to allow assigning of each edge to exactly one contour since it would result in only one (which?) feature extracted from image in Fig.2a. The other non-viable approach would be to extract all feasible contours which, for our example, would result in 7 contours. The only solution left is to establish a criterion for choosing a representative subset of all feasible closed contours in the image, and consequently devise an algorithm which will enforce it.

Let $\mathcal{O}$ be the set of all feasible closed contours in the image. We define its *canonical* subset $\mathcal{C}_\mathcal{O}$ as the set of all elementary [3] contours $c_i \in \mathcal{O}$ for which the following holds: for each $c_j \in \mathcal{O}\backslash\{c_i\}$, either c_i does not encompass c_j, or c_i and c_j do not share a part of boundary. It can easily be shown that each image edge can be assigned either to 0, 1 or 2 elementary contours. We propose representing

$\mathcal{O}$ with $\mathcal{C}_\mathcal{O}$, since it exhibits the following favorable properties:

- **completeness**: each $c_k \in \mathcal{O} \setminus \mathcal{C}_\mathcal{O}$ can be obtained as a combination of two elements of $\mathcal{C}_\mathcal{O}$;
- **uniqueness**: there is only one canonical subset for each $\mathcal{O}$;
- **minimality**: the set obtained by removing any element of $\mathcal{C}_\mathcal{O}$ is not complete.

The listed properties make the canonical subset particularly suitable for intermediate representation of structured objects since the lack of redundant information makes the tasks of correspondence and indexing less complex. The algorithm for extracting the canonical set of closed contours from the input image is proposed in Section 3, while the experimental results are provided in Section 4.

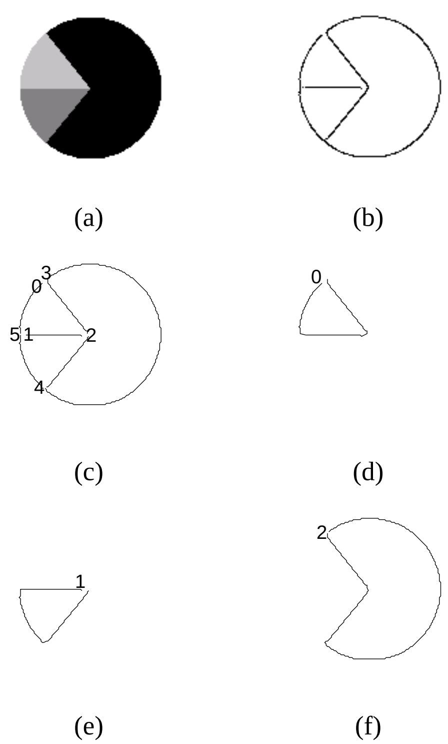

Figure 2: A synthetic test image (a), the edge map provided by the Canny edge detector (b), 6 edge chains broken by the criteria of endpoint proximity and high curvature (c), and the extracted canonical set of 3 closed contours (d,e,f).

3. THE PROPOSED PROCEDURE

In general, a procedure for extracting the canonical subset of closed contours $\mathcal{C}_\mathcal{O}$ from a given edge map consists of the following tasks: (1) obtaining a set of edge chains suitable for grouping, (2) determining which pairs of those chains can be linked into a common contour, and (3), extracting only those closed contours which are members of $\mathcal{C}_\mathcal{O}$. The former two tasks rely on local principles (proximity, continuation), while the third one organizes the obtained results with respect to the closure criterion.

3.1. Obtaining suitable edge chains

The task of this stage is to perform a partition of elements of the edge map into a set of edge chains, with respect to the following criteria:

- **saliency**: chains should be in accordance with local perceptual principles (proximity, continuation);
- **correctness**: chains should not span between more than two entities (objects or background);
- **minimality**: there should be as few chains as possible, when not in conflict with previous criteria, in order to avoid over-segmentation and lower the complexity of further processing.

While much work has been done regarding the first criterion [9, 10], it is not clear how to enforce the second one. A popular strategy is to make all chains short [8], which certainly solves the problem at hand, but badly affects the third criterion. We adopt a different approach, in which smoothly connected chains are broken only at locations for which there is a perceptual indication of a junction or the end of an object. Thus, intermediate edge chains are extracted following certain considerations from [10]. Consequently, chains are broken at points where radius of curvature is less than 5 pixels. Finally, for each chain endpoint, we consider the breaking of all other chains whose nearest point is closer than 3σ, where σ is the parameter of the Gaussian smoothing used by edge detector.

3.2. Building the incidency lists

Each edge chain provided by the previous processing step is first assigned the referential direction (see [4] for a discussion on the subject). Consequently, individual chains are checked for closure: this is possible if the noise is low and the contours produced by the linking procedure are not fragmented. On positive result, the closed contour is registered while the corresponding chain is not further considered. For each of the remaining chains, two incidency lists are created, one for each direction. Each of the created lists is filled with neighbours of the corresponding chain and direction by a simple thresholding over the saliency measure which is defined with respect to the principle of proximity. Besides the index and direction which identify the neighbouring chain, a single list node also contains the computed saliency measure, and the angle ϕ between the tangent at the last point of the chain corresponding to list and the tangent at the first point of the neighbouring chain. This angle is used in the latter phase, in which we reduce the count of neighbours in order to ensure computational feasibility of the next stage. The nodes are sorted with respect to ϕ; then, each node having a lower saliency than some other node differing in ϕ for less than 30° is suppressed; finally, only the strongest four neighbours according to saliency are left in the list.

3.3. Finding $\mathcal{C}_\mathcal{O}$ by the best-first search algorithm

Unlike the previous methods, we formulate the problem as the search problem over the space of all feasible unconnected contours. Each state s_i of the search space corresponds to a concatenation of nodes n_j obtained from incidency lists, together with several additional traits of the corresponding contour. Two states $s_1 = \{n_0, n_1, \ldots, n_m\}$ and $s_2 = \{n_0, n_1, \ldots, n_m, n_{new}\}$ are connected iff n_{new} can be found in the incidency list for node n_m, and $n_{new} \neq n_j, j = 1, 2, \ldots, m$. The search space therefore has the form of a tree, with branching factor bounded by the limit imposed on the count of nodes in incidency lists. The depth of the tree is determined by the longest salient concatenation of chains, whether open or closed. The first node is constructed with the chain and the corresponding direction for which a closed contour is sought, and assigned the saliency of 1. Goal states $s_g = \{n_0, n_1, \ldots, n_m, n_0\}$ are easily identified, by checking that the rotation index [1] of the contour is equal to one and testing the first and the last node for equality.

A nice property of such state tree is that closed contours which belong to $\mathcal{C}_\mathcal{O}$ are generated by the space-efficient best-first search [11], in which the order of opening of new nodes is imposed by the order of nodes in incidency lists. Since the nodes in incidency lists are ordered with respect to $\phi \in [-\pi, \pi]$, the search algorithm will always first explore the leftmost continuation of each chain and ensure that the first open goal-state will correspond to an elementary left-oriented contour. All such contours can be obtained by performing the search from each node, in both possible directions. For example, consider Fig.2c. Let the referential direction of an edge chain be defined as going from the endpoint near the chain label, towards the other endpoint. Then, the search from the horizontal edge chain labeled with #1 would find the contour $\{1, \overline{3}, 0\}$ for the referential direction of the edge #1, and the contour $\{\overline{1}, 5, \overline{2}\}$ for the reverse.

At first, the described formulation of the problem does not seem attractive, since the worst-case time complexity of depth-first search is $O(b^d)$, where b is the branching factor and d the depth of the search tree. Elder and Zucker [1] formulated the same problem as the shortest-path graph search, for which there are known solutions with polynomial complexity $O(n^2 log(n))$. However, their formulation has an important qualitative shortcoming which considerably restricts the freedom in modelling the perceptual processes. Namely, the shortest-path search is able to model only the quite trivial scheme of inferring compound saliency of the contour from individual saliences of chain junctions, which is to make the compound saliency equal to the sum (or the product) of individual saliencies. Consequently, it can't make dynamical decisions during the search, based on properties of the *pattern* of the individual saliencies. One consequence of such approach is the inability to filter out "double figure eights" from the extracted set of contours. Dynamic decisions are also useful in modelling at least two patterns of perceptually salient non-smooth chain concatenations, namely, apices and interruptions (see Fig.4 and Fig.3 for examples). The influence of these effects to the compound saliency can't be modelled faithfully by the sum model, since one or two interruptions or apices may be salient due to occlusions, but five or ten are certainly not. For instance, the sum model can not assign saliency of 0.9 to the first apex, 0.8 to the second one etc, since it must assign the same saliency to each chain junction, regardless of the pattern of other saliencies in the contour. These problems are easily solved in the state-space search formulation, since each state may be augmented with the following data about the corresponding contour: the compound saliency so far; the rotation index; the angular distance with respect to the first point of the first chain; counts of long interruptions and apices.

Nevertheless, the exponential complexity (although worst-case) does pose a significant problem which we try to alleviate by several strategies for pruning the search tree:

- for each chain, only one contour per direction is possible: a data structure is therefore kept with chains and directions which already have been assigned to a closed contour, and their appearance in an another contour is not permitted;
- contours with partial rotation indices larger than 1 and smaller than -1 are pruned;
- we limit the depth of the search to 20 nodes, or 10 if the angular distance of the contour is less than 0;
- for each chain only two apices and two long interruptions are permitted.

4. EXPERIMENTAL RESULTS

We have tested the proposed procedure on numerous synthetic images of increasing complexity. The processing results are shown in figures Fig.1, Fig.2 and Fig.3. The obtained canonical set for a complex image in Fig.3 contains 6 closed contours, only one of which is not perceptually salient. The contour labeled #2 corresponds to an almost closed part of the background between objects whose contours are labeled with #0, #3 and #5. While it could be deduced that the region in question should be attributed to background, the proposed procedure does not try to perform such reasoning.

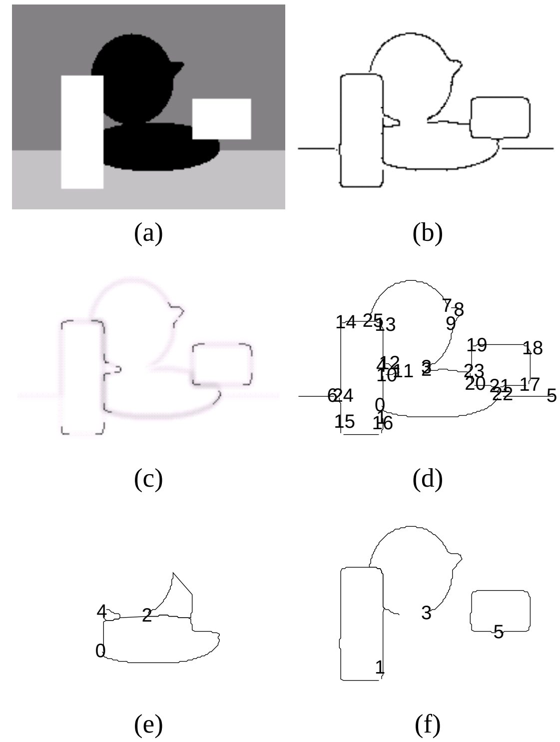

Figure 3: A complex synthetic test image (a), its edge map (b), the points of high curvature at which the chains are broken superimposed to the gradient image used by the edge detector (c), the final set of 30 edge chains (d), and finally, the obtained canonical set of 6 closed contours (e) and (f). In both contours #0 and #3, one large interruption is tolerated.

The obtained results confirm that the devised procedure can deal with problematic configurations such as apices and interruptions. These configurations are quite frequent in real images, which is illustrated by the image in Fig.4: due to noise and smoothing effects, the ends of both wings are not represented as edges at all. Such situation in which the contour "opens" when the object becomes too narrow occurs both in images of cars and airplanes.

5. CONCLUSION

We have presented a procedure for extracting canonical set of closed contours from an edge map. Providing such a representation is particularly important in the case of structured objects, since straightforward approaches may provide different (although correct) outputs for very similar input images. The proposed procedure allows more complex modelling of the perceptual grouping pro-

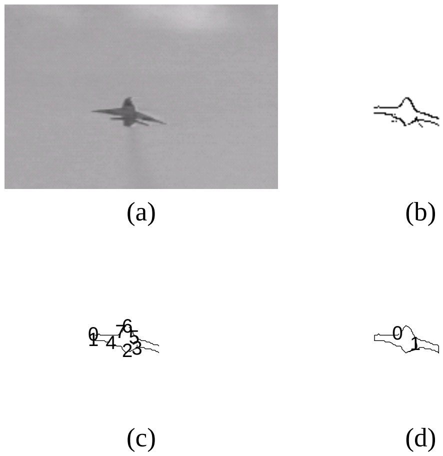

Figure 4: A simple real image (a), its edge map (b), the final set of 8 edge chains (c), and the obtained canonical set of 2 closed contours (d). In both contours, one apex is tolerated.

cess than some of the known approaches, at the cost of increased computational complexity. Several strategies for pruning the search tree have been devised in order to simplify the computations, however we still can not provide conclusive evidence about their efficacy. The obtained processing results on many synthetic images confirm the viability of the procedure in the context of ambiguous configurations such as apices and interruptions.

REFERENCES

[1] J. Elder and S. Zucker, "Computing contour closure," in *Proceedings of the European Conference on Computer Vision*, vol. 1, (Cambridge, England), pp. 399–412, Springer-Verlag, Apr. 1996.

[2] D. W. Jacobs, *Recognizing 3-D Objects Using 2-D Images*. PhD thesis, Massachussets Institute of Technology, Cambridge, Laboratory of Artificial Intelligence, 1992.

[3] M. Xie and M. Thonnat, "An algorithm for finding closed curves," *Pattern Recognition Letters*, vol. 13, pp. 73–81, Jan. 1992.

[4] L. Williams and K. Thornber, "A comparison of measures for detecting natural shapes in cluttered backgrounds," *International Journal of Computer Vision*, vol. 34, pp. 1–16, August 1999.

[5] S. Mahamud, K. Thornber, and L. Williams, "Segmentation of salient closed contours from real images," in *Proceedings of the International Conference on Computer Vision*, no. 7, (Corfu, Greece), pp. 891–897, Sept. 1999.

[6] B. Parvin and G. Medioni, "B-rep object description from multiple range views," *International Journal of Computer Vision*, vol. 20, no. 1/2, pp. 81–112, 1996.

[7] X. Jiang, "An adaptive contour closure algorithm and its experimental evaluation," *IEEE Transactions on Pattern recognition and Machine Intelligence*, vol. 22, pp. 1252–1265, November 2000.

[8] P. Havaldar, G. Medioni, and F. Stein, "Perceptual grouping for generic recognition," *International Journal of Computer Vision*, vol. 20, no. 1/2, pp. 59–80, 1996.

[9] M. Xie, "Edge linking by using causal neighborhood window," *Pattern Recognition Letters*, vol. 13, pp. 647–656, 1992.

[10] A. Farag and E. Delp, "Edge linking by sequential search," *Pattern Recognition*, vol. 28, pp. 611–633, May 1995.

[11] S. J. Russel and P. Norvig, *Artificial Intelligence: A Modern Approach*. Upper Saddle River, New Jersey: Prentice Hall, 1995.