Recognition of Partially Occluded and Deformed Binary Objects

Ondřej Horáček, Jan Kamenický, and Jan Flusser [*]

Institute of Information Theory and Automation,
Academy of Sciences of the Czech Republic,
Pod vodárenskou věží 4, 182 08 Prague 8, Czech Republic
{horacek,kamenik,flusser}@utia.cas.cz

Abstract. A method dealing with recognition of partially occluded and affine transformed binary objects is presented. The method is designed for objects with smooth curved boundary. It divides an object into affine-invariant parts and uses modified radial vector for the parts description. Object recognition is performed via string matching in the space of radial vectors.

1 Introduction

Recognition of objects under partial occlusions and deformations caused by imaging geometry is one of the most difficult problems in computer vision. It is required always when analyzing 2-D images of a 3-D scene. Although many methods trying to solve this task have been published, it still remains open. Clearly, there is no universal algorithm which would be "optimal" in all cases. Different methods should be designed for different class of objects and for different groups of assumed deformations. In this paper we assume the objects are deformed by an unknown affine deformation. This assumption approximates real photos with a week perspective deformation.

We introduce a method developed for the recognition of smooth curved objects. First, the shape is divided into parts which are defined by means of inflection points of the object boundary. Then the shape of each part is described by a special kind of radial vector. Finally, the parameters of the affine deformation are estimated and classification is performed by string matching in the space of radial vectors. The performance of the method is demonstrated by experiments.

2 Overview of Current Methods

Current methods can be classified into two major categories. The methods of the first group divide the object into affine-invariant parts. Each part is described

by some kind of "standard" global invariants, and the whole object is then characterized by a string of vectors of invariants. Recognition under occlusion is performed by maximum substring matching. Since inflection points of the boundary are invariant to affine (and even projective) deformation of a shape, they become a popular tool for the definition of the affine-invariant parts. This approach was used by Ibrahim and Cohen [3], who described the object by area ratios of two neighboring parts. As a modification which does not use inflection points, concave residua of convex hull can be used. For polygon-like shapes, however, inflection points cannot be used. Instead, one can construct "cuts" defined by three or four neighboring vertices. Yang and Cohen [11] used area ratios of the cuts to construct affine invariants. Flusser [2] further developed their approach by finding more powerful invariant description of the cuts. Similar method was successfully tested for perspective projection by Rothwell et al. [6].

Lamdan [4] used mutual position of four "interesting" points for the recognition. To verify the received match, normalized concave areas were described by radial vector.

The methods of the second group are "intrinsically local" – they describe the boundary in every point by means of its small neighborhood. In that way they transform the boundary to so-called signature curve which is invariant to affine/projective transform. Typical representatives of this group are differential invariants. They were probably discovered by Wilczynski [10] and furthermore developed by Weiss [9], [8]. These invariants are based on derivatives of orders from four to eight. They have been experimentally proven to be extremely sensitive to inaccurate segmentation of the boundary, discretization errors and noise.

Mokhtarian and Abbasi [5] used inflection points themselves to characterize the boundary. They constructed so-called Curvature Scale Space and traced the position of inflection points on different levels of image pyramid. The trajectories of the inflection points then served as object descriptors. There have been also methods based on wavelet transform of the boundary. E.g., Tieng and Boles [7] introduced wavelet-based boundary representation, where affine invariance was achieved by enclosed area contour parameterization. However, the use of the wavelet-based methods in case of partial occlusions is questionable.

3   Definition of Affine-Invariant Parts

Both inflection points and central points of straight lines are affine invariant, i.e. the properties "to be an inflection point" and "to be a central point of a straight line" are preserved under arbitrary nonsingular affine transform. Thus, we use these points for the construction of affine-invariant parts. We connect each couple of neighboring cut points by a line. This line and the corresponding part of the object boundary form a convex region which may or may not lie inside the original object (in Fig. 1c). A sequence of such parts carry efficient information about the object.

Detection of inflection points of discrete curves has been discussed in numerous papers. Let us recall that, in the continuous domain, an inflection point is

defined by a constraint $\ddot{x}(t)\dot{y}(t) - \dot{x}(t)\ddot{y}(t) = 0$, where $x(t), y(t)$ is a parameterization of the curve and the dots denote derivatives with respect to t. When this definition is directly converted to the discrete domain, it becomes very sensitive to sampling and noise. Thus, we propose a new robust method of curvature estimation.

A circle with fixed radius is placed on each boundary point (in Fig. 1a). Ratio of the whole circle area to its area being inside the object serves for estimation of the curvature. When this ratio equals 0.5, the boundary has zero curvature and the corresponding point is either an inflection point or it lies inside a straight segment. We construct a curvature graph (in Fig. 1b), smooth it and define cut points as zero-crossing points and middle points of approximately zero-value segments. Furthermore, a request of sufficient part size is considered: the segment of the curvature graph between two cut points should have sum of values above some threshold, otherwise it is treated as a part of zero curvature.
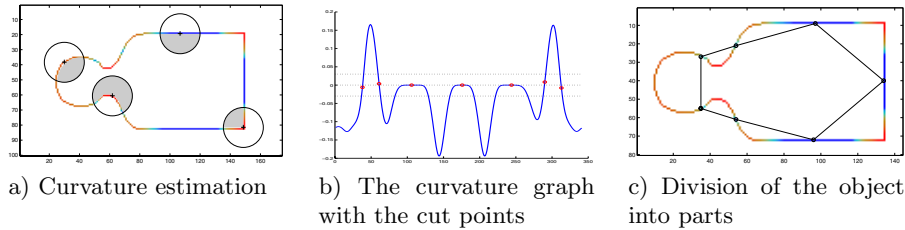

a) Curvature estimation   b) The curvature graph   c) Division of the object
                             with the cut points   into parts

Fig. 1. Definition of affine-invariant parts

4 Description of the Parts

The object is represented by the parts, construction of which has been described above. Describing the shape of the parts we get a description of the whole object which is robust to occlusion. Robustness to occlusion means that if some part of the object boundary is missing or changed, only few elements of the feature vector are changed. This is an important attribute. Note that traditional global methods, for instance description of the object by moment invariants or Fourier descriptors, do not have this property.

It would be possible to describe each part individually and eliminate the impact of the deformation by using proper affine invariants (moment invariants or Fourier descriptors for instance). In such a case, however, we do not employ important information that all the parts were deformed by *the same* transformation. Including this consistency information in the object description can significantly increase the recognition performance. Thus, we propose the following description of the parts by a modified radial vector, with included position of critical points. See complete demo object description in Fig. 2a.

The spokes of the radial vector come from the middle of the cutting line and they divide the part into subparts of equal area. For each part, they are constructed as follows.

1. Define the desired number n of the spokes (i.e. the length of the radial vector).
2. Go through the outer boundary of the part.
3. For each step calculate the area of the triangle between the neighbor boundary points and the midpoint of the cutting line.
4. If the cumulated area just steps over $k/(n-1)$ fraction of the total part area, the k-th spoke ends in the current boundary point.

The introduced modified radial vector divides the part invariantly under affine transformation. Note that a classical radial vector with constant-angle spokes distribution or constant-boundary length spokes distribution has not such a favorable property.
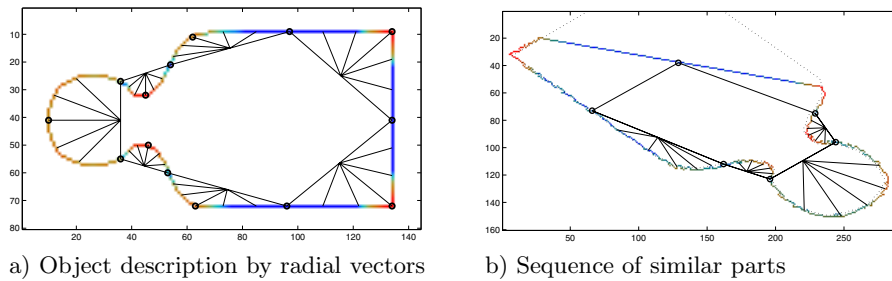

a) Object description by radial vectors   b) Sequence of similar parts

Fig. 2. Description and matching of the demo object

5   Matching

The image is classified by finding the longest and best matching part of the boundary (in Fig. 2b). This is realized by comparing a parts sequences of the image with a parts sequences of the database objects.

1. For each part of the database object and each part of the image do the following:
2. Take a sequence of parts starting from the current part. A length of this sequence is gradually incremented, it begins with the length of the previous longest successful match (or equals one in case of the first trial).
3. Recover the affine transformation T between the database and sensed objects. Use the least square fit of the control points of all parts involved in the sequence. Each part has three control points: two cut points and the middle-spoke end-point.
4. Transform the database radial vectors by transformation T.
5. Compare the transformed database radial vector sequence with the current one from the image. Similarity measure S is evaluated for this purpose – see below. If S is smaller than required similarity threshold, this sequences are considered not to match, and a start of next sequence is taken on the step 1.

6. The current sequence of radial vectors was successfully tested for a match in the previous step. If this sequence is longer than the previous best match, or is equal and the similarity level S is higher, select this match as the best one. Now try to make the matching sequence even longer – continue with the step 2.

Similarity measure S is introduced for suitable comparison of the radial vectors u, v. $S = 1$ only if $u = v$, S approaches zero for growing vector difference. The single similarity measure s_i of the i-th spoke lengths u_i, v_i is a Gaussian quantity of the $u_i - v_i$ difference

$$s_i = e^{-\frac{1}{\sigma_i^2}\left(\frac{u_i - v_i}{2}\right)^2}, \qquad \sigma_i = k_1 + k_2 \left|\frac{u_i + v_i}{2}\right|.$$

The Gaussian dispersion σ_i absolute component k_1 realizes a noise tolerance, the constant k_2 determines a tolerance relative to the value size.

We have the following requirements for combining single component s_i to overall similarity measure S: $S = s_i$ if all s_i are equal; $S = 0$ if some $s_i = 0$; S needs to be sensitive to all s_i; moreover, we require S to be 0.75 if all but one s_i equal 1 and one s_i equals 0.5. All these criteria are met for example by weighted average with weights w_i inversely proportional to s_i

$$S = \frac{\sum_{i=1}^{n} w_i \cdot s_i}{\sum_{i=1}^{n} w_i}, \qquad w_i = \frac{n-2}{s_i} - (n-3).$$

6   Experimental results

The proposed method was tested on a set of 24 binary objects, which had been previously segmented from color images. The objects were successively deformed by various affine transforms, their various regions were occluded and then the objects were matched against the database of the 24 originals. As a matching criterion which should be maximized we took the number of those parts (cuts) of the test object which match with the parts of the database object. This is in fact a well-known principle of string matching.

For illustration, two examples are shown in Fig. 3. On the left-hand side, one can see partially occluded and transformed objects. The corresponding database objects (which were successfully found in both cases) are shown on the right-hand side. The critical (inflection) points are highlighted, their connecting lines define the division into parts. The spokes of the corresponding radial vectors are drown inside the matched parts of the image.

The modified radial vector describes the boundary with a good precision, the tolerance to a shape perturbations is controlled by user-defined parameters/thresholds. This enables an optimization for various types of shapes. Interestingly, the boundary does not need to be a smooth curve with well-defined inflection points. The method finds critical points even on polygonal parts (see
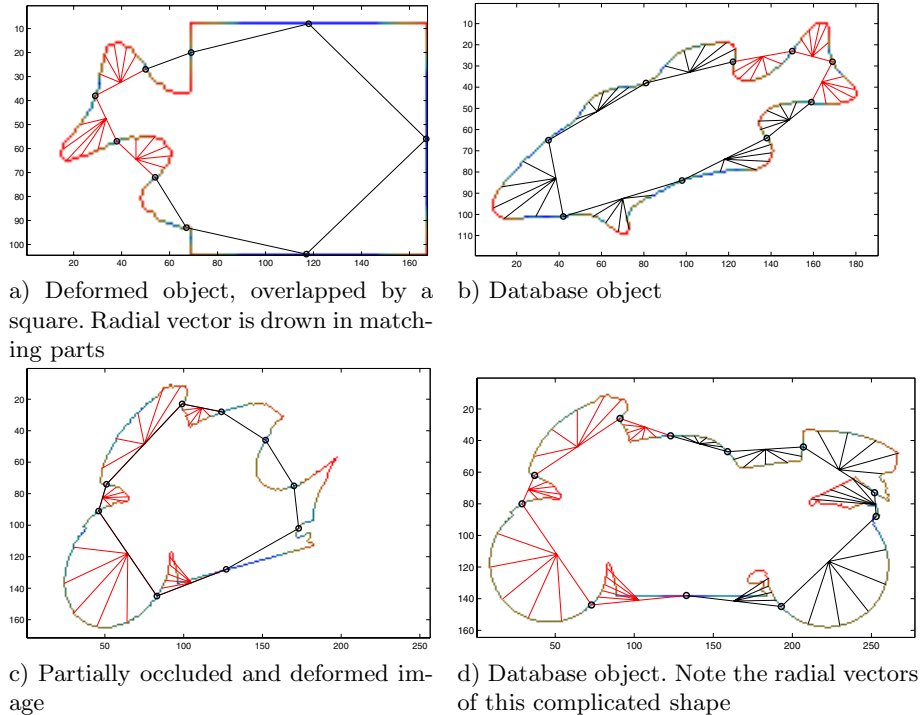
a) Deformed object, overlapped by a square. Radial vector is drown in matching parts

b) Database object

c) Partially occluded and deformed image

d) Database object. Note the radial vectors of this complicated shape

Fig. 3. Recognition samples and description of recognized database objects

Fig. 3a) and is able to construct radial vector even for non-convex parts (see Fig. 3d).

The most problematic part of the presented method is the critical point detection. Different positions of the critical points lead to different descriptions and of course affect the matching. Instability of the critical points can be caused by unsuitable shapes (without clear inflection points), affine transformation (affects the curvature), or occlusion (inflection points originally ignored can become significant). This situation is shown in Fig. 4. Although the database object (Fig. 4 top) was identified correctly in both cases (Fig. 4 bottom) one can see worse match on the right when overlaying the test and the database objects (the overlayed database object is drown by dotted line).

The results of our experiment depends on the object shapes, on the size of the occlusion, on the deformation, and other conditions. The summary is in Table 1. "Image area" denotes the size of the visible part of the test object (in per cent), "Constant scale of details" indicates whether or not the same thresholds were used for database and test objects when detecting inflection points, and "Transformation" means the significance of the deformation measured by skewing. The table itself shows the maximum number of matching parts over all database objects. In all instances where the maximum number of matching parts is greater than 2 the test objects were recognized correctly. One or two matching
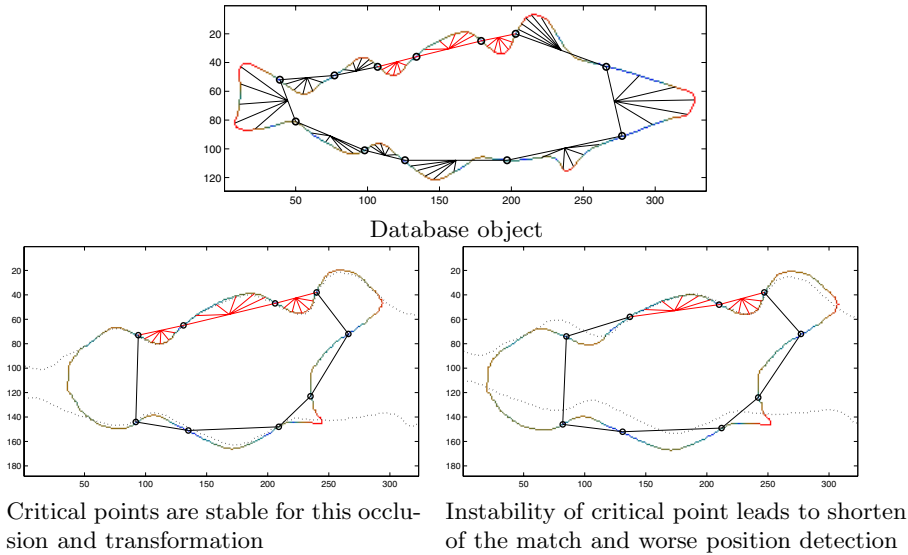
Database object

Critical points are stable for this occlusion and transformation

Instability of critical point leads to shorten of the match and worse position detection

Fig. 4. The impact of instability of the critical points

Table 1. Object recognition experiment. The number of matched boundary parts for various occlusion and transformation. All numbers greater than 2 lead to correct match.

Image area	100%	90%	50%	50%	100%	100%	50%
Constant scale of details	yes	yes	no	yes	yes	yes	yes
Transformation	none	none	none	none	medium	strong	medium
Image 1	12	7	4	3	8	5	3
Image 2	11	8	1	4	4	6	2
Image 3	11	8	3	4	4	5	4
Image 4	11	9	3	4	2	1	3
Image 5	7	4	2	2	7	3	2
Image 6	13	8	2	2	2	1	1
Image 7	10	7	2	6	8	1	4
Image 8	9	6	4	4	9	1	4

parts does not ensure unique correct match, the classification can be wrong in such cases. These bad situations were caused mostly by strong deformation or big occlusion which lead to instability of critical points.

7   Conclusion

We presented a method for recognition of partially occluded binary objects deformed by affine transformation. The method uses local affine-invariant description of object boundary by means inflection points and radial vectors. When working with digital boundary, the major limitation of the method is stability of

inflection points. As the experiments demonstrated, if the curve has "prominent" inflection points, they are usually very stable under affine transformation and the method works perfectly. On the other hand, in case of obscure boundary the inflection points may be detected at different positions depending on the particular transformation and/or occlusion and the recognition may fail.

Our experiment proved a good discrimination power of the method on a given test set. We discovered that if the maximum number of matched boundary parts between the unknown object and the database is greater than 2, it always indicated a correct match. Thus, this threshold can be recommended for prospective real experiments too.

References

1. Guggenheimer, H.W.: Differential Geometry. New York. McGraw-Hill (1963)
2. Flusser, J.: Affine Invariants of Convex Polygons. IEEE Transactions on Image Processing **11**(9) (September 2002)
3. Ibrahim Ali W.S., Cohen, F.S.: Registering Coronal Histological 2-D Sections of a Rat Brain with Coronal Sections of a 3-D Brain Atlas Using Geometric Curve Invariants and B-spline Representanion. IEEE Transaction on Medial Imaging **17**(6) (December 1998)
4. Lamdan, Y., Schwartz, J.T., Wolfson, H.J.: Object Recognition by Affine Invariant Matching. Computer Vision and Pattern Recognition, (June 1988) 335–344
5. Mokhtarian, F., Abbasi, S.: Shape Similarity Under Affine Transform. Pattern Recognition **35** (2002) 31–41
6. Rothwell, C.A., Zisserman, A., Forsyth, D.A., Mundy, J.L.: Fast Recognition Using Algebraic Invariants. Geometric Invariants in Computer Vision. Cambridge, MIT Press (1992) 398–407
7. Tieng, Q.M., Boles W.W.: An Application of Wavelet-based Affine-invariant Representation. Patern Recognition Letters **16** (1995) 1287–1296
8. Weiss, I.: Noise Resistant Invariants of Curves. Geometric Invariance in Computer Vision. Cambridge, MIT Press (1992) 135–156
9. Weiss, I.: Projective Invariants of Shapes. Proc. Image Understanding Workshop (1988) 1125–1134
10. Wilczynski, E.J.: Projective Differential Geometry of Curves and Ruled Surfaces. B. G. Teubner, Leipzig (1906)
11. Yang, Z., Cohen, F.S.: Image Registration and Object Recognition Using Affine Invariants and Convex Hulls. IEEE Transitional on Image Processing **8**(7) (June 1999)