Planar Shape Databases with Affine Invariant Search

Startchik, Sergei; Milanese, Ruggero; Rauber, Christian; Pun, Thierry

How to cite

STARTCHIK, Sergei et al. Planar Shape Databases with Affine Invariant Search. In: Image Databases and Multi-Media Search. [s.l.] : World Scientific, 1998. p. 285–292. doi: 10.1142/9789812797988_0025

This publication URL: https://archive-ouverte.unige.ch//unige:121377
Publication DOI: 10.1142/9789812797988_0025

Planar shape databases with affine invariant search

S. Startchik, R. Milanese, C. Rauber, T. Pun

Dept. of Computer Science
University of Geneva, Switzerland.
E-mail: startchi@cui.unige.ch

1. Overview

Image databases are often used to archive and retrieve images containing man-made 3D objects usually taken from arbitrary viewpoints. These objects generally incorporate planar surfaces containing different kinds of highly curved patterns. It is often the case that the *form* of such patterns characterizes well the corresponding object. Besides classical retrieval by colour or texture, the database should therefore inevitably support queries based on geometrical shape.

In this paper we show how images in a database can be provided with additional information allowing viewpoint invariant shape query. Viewpoint invariance allows to compensate for affine or projective distortions, and is achieved by the usage of invariant reference frames. Those frames serve as a canonical coordinate system, common for all shapes in the database; in this system their similarity can be estimated and neighbourhood relationships can be established for subsequent search. The contributions of this work are twofold: first, the construction of such reference frames is done by combining different types of local invariant properties of the curve. Second, in the context of search operations, an important speed-up is introduced by means of moment-based indexing. Both strategies are tested on a database of trademarks. The database contains images of supermarket box packages with trademarks placed on their flat sides. Section 2 briefly describes the method for computing viewpoint-independent shape similarity. Section 3 presents the search mechanism, including indexing speed-up routine, and an evaluation of the results.

2. Viewpoint invariant similarity of shapes

In this section, the comparison of two shapes seen from two different viewpoints is considered, where each viewpoint transformation corresponds to either an affine or projective transformation of the image. For shapes that can be characterized in terms of points, lines or higher-order algebraic curves, comparison can be done using algebraic invariants for both affine and projective camera transformations [5, 7, 8]. For a broader class of shapes, general, non closed-form invariants can be obtained by exploiting local, differential properties of the shape [3, 6, 9]. In this latter situation, the approach consists in the construction of a reference frame based on invariant properties of the shape itself, and in studying the representation of the shape with respect to this frame. The selected frame becomes a canonical or unit coordinate system, in which the shape becomes an invariant signature.

Before detailing a particular curve representation, we concentrate on the construction of a reference frame which is based on projectively invariant properties of a shape such as straight and bitangent lines, cusp and inflection points [5, 7, 9]. In general, lines lead to more stable invariants. For this reason, we replace cusps and inflections by their tangent lines, which yields four types of lines that can be extracted from the shape. Whenever each of these lines intersect one another or a shape, the intersection point gives a stable invariant reference point (cf. figure 1). We pay attention to special cases when intersections of two pairs of lines of different types gives the same point (cf. figure 1.b), thereby removing its multiplicity.

To compensate for affine transformation, three reference points suffice, while for projective case, the number of necessary points is four. Taking the affine case as an example, the selection of triples of points amongst n invariant points of the curve gives C_n^3 possible combinations. By associating each point with a point on the curve, an ordering of points with respect to the curve is established; by only selecting successive triples of points, the number of possible coordinate frames thus reduces to n. For example, in figure 1.a four reference frames are obtained (P_1, P_2, P_3), (P_2, P_3, P_4), etc. Two curves are considered similar if in *one* of their n invariant reference frames they perfectly match. In one frame, e.g. given two triples, the measure of match is defined as the sum of closest distances from a number of sampled points on the first curve to the other curve. The use of the closest distance provides this similarity measure with the useful property of being independent of curve parametrization. Moreover, even if the first curve is partially occluded, but corresponds to a subset of the second curve, the measure of distance will remain small. The overall similarity measure between two curves Γ_1 and Γ_2 with n and m invariant points respectively is the maximum of $mn/2$ curve matches in all combinations of their invariant reference frames i and j: $\sup_{ij}\{M_{ij}(\Gamma_1, \Gamma_2)\}$, where M_{ij} is the match using reference frames i, j. This requires n signatures of the first curve in its n frames to be compared to the m signatures of the second curve in its m frames, leading to $m + n$ transformations of original curves and $mn/2$ comparisons. The cost of each transformation and comparison is heavily dependent on the curve representation that will be adopted.

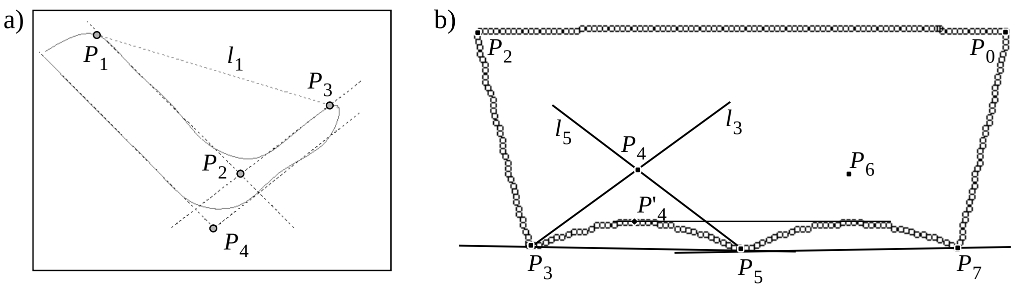

Figure 1. Finding invariant points associated with the curve and ordering them. In a) four straight lines and one bitangent l_1 were found for the L-shape curve, their intersection gives invariant points P_1, ..., P_4. In b), straight lines, bitangents and cusps are used. Points of contact of bitangents with the curve P_3, P_5, P_7 are cusps at the same time and should be considered only once. Point P_4 which is the intersection of tangents l_5 and l_3 at cusps P_5 and P_3. Such point is preferred over the internal bitangent point P'_4.

For image database retrieval, a large number of comparisons is required; a parametric representation of curves should thus be chosen so as to guarantee low-cost geometric transformations. This has restrained our attention to the class of parametric curves $p(t)= \sum_{i=1}^{N} C_i g_i(t)$, where C_i are the control points and $g_i(t)$ are the base functions [1, 3]. In this case, only the control points are affected by geometric transformations, under the condition that the relation between the curve and the control points is invariant to the applied transformation. In the case of projective transformation this nice property is available only for *rational* curves, i.e. those for which the base functions are of the form $g_i(t) = f_i(t)/\sum_{j=1}^{N} f_j(t)$. Another useful property of rational curves is their ability to model almost polygonal objects while preserving differentiability at “corners”. Two families of rational curves were selected and compared in this work: NURBS and RaG [2], respectively defined with polynomial and exponential base functions. The advantages of RaGs include easy multiscale analysis, control of smoothness, infinite differentiability and simplified fitting (minimization is global); NURBS, however, allow fast matching operations. Curve fitting in both cases is done using the Levenberg-Marquardt method with initial uniform knots placement. After extensive tests, NURBS have been preferred for efficiency considerations; RaGs will, however, be kept as a possible alternative.

3. Indexing and search in shape database

Shape-based queries in image databases consist in finding, given an input shape, a certain number of similar shapes from the database. Given an input shape containing n invariant points, the search in a database of N_s shapes will require $O(n\bar{n}N_s)$ comparisons, where $\bar{n}$ is the average number of invariant points for each shape in the database. In order to assess the discriminative power of the retrieval, experiments were performed with a database of $N_s = 20$ shapes. A pairwise distance comparison was done between all shapes in the database according to the similarity measure defined earlier; all distances are depicted in the distance matrix (figure 2.a). It can be seen that the values of diagonal elements are significantly lower that the others by at least an order of magnitude. This demonstrates strong discrimination capability as well as high precision of comparison. Asymmetry of the distance matrix originates from the use of closest distance in shape comparison.

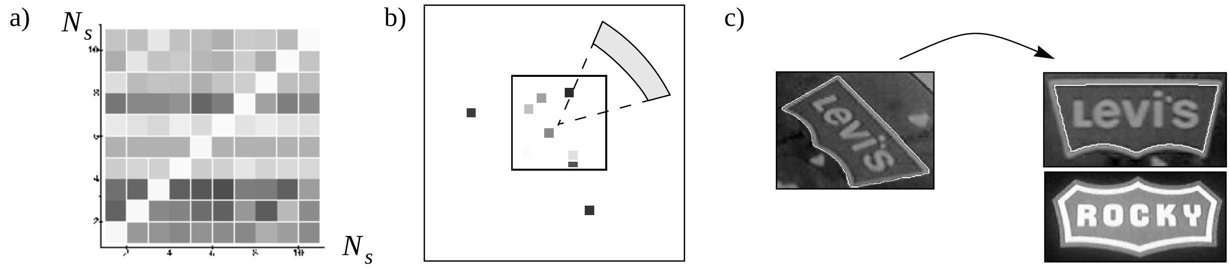

Figure 2. Distance matrix a) between two curves in the database (white corresponds to small distance). In b) moments associated with signatures of shapes stored in the database are shown in the unit frame. Examples of images retrieved according to the shape of "levis" logo c).

The number of comparisons is linear in the number of shapes and in the number of their invariant features. Since the time for the comparison between any two curves is 0.2s on a Sparc 10, this approach would be unfeasible for very large databases. A significant speed-up of the search operation is achieved by an indexing mechanism reducing the subset of shapes to compare. As signatures are represented in invariant frames, all Euclidean measures performed in this frame are also invariant; first-order moments of the signatures can therefore be used. Given a value of a moment computed from an input shape signature, a number of closest moments from the database can be obtained in about 0.01s. Considering a subset of k neighbours, this gives roughly $0.2k$ seconds for subsequent comparisons. The number k depends on the method of neighbourhood estimation which in turn can be fixed only after evaluation of the moments distribution in the database. This has to be done in order to assure the presence of correct matches in the subset of neighbours. Estimations show that noise, produced by viewpoint variations, entail variations in moment vector equivalent to 3% of range in orientation and 0.02% of range in magnitude. Using these values for construction of a ring-sector neighbourhood (cf. figure 2.b), we can guarantee the retrieval of correct matches and estimate the number k. For our database, it is equal to 7 moments in average, giving an acceptable retrieval time.

Figure 2.c) contains an example of a shape query (left) and two most similar images retrieved from the database. To conclude, database search require an efficient moment-based indexing process and a precise verification, based on direct comparison. Future improvements will allow to weight moments with curve arc length, in order to increase their stability to occlusion (direct comparison of curves already possess this property).

References

[1] F. S. Cohen, J.-Y. Wang, "Part I: Modeling image curves using invariant 3-D object curve models - a path to 3-D recognition and shape estimation from image contours", IEEE Trans. Patt. Anal. Mach. Intell., 1994, vol. 16, No. 1, p. 1-12.

[2] A. Goshtasby, "Design and Recovery of 2-D and 3-D shapes using rational gaussian curves and surfaces", Intern. J. Comput. Vis. 10:3, 1993, pp. 233-256.

[3] T.Moons, E.J. Pauwels, L.J. Van Gool, A. Oosterlinck, "Foundations of semi-differential invariants", Intern. J. Comput. Vis. 14, 1995, p. 25-47.

[4] E.J. Pauwels, T.Moons, L.J. Van Gool, P. Kempenaers, A. Oosterlinck, "Recognition of planar shapes under affine distortion", Intern. J. Comput. Vis. 14, 1995, p. 49-65.

[5] T. H. Reiss, "Recognizing planar objects using invariant image features", Lecture Notes in Computer Science, Springer-Verlag, 676, 1993.

[6] E. Rivlin, I. Weiss, "Local invariants for recognition", IEEE Trans. Patt. Anal. Mach. Intell., 1995, vol. 17, No. 3.

[7] J.L. Mundy and A. Zisserman (editors), "Geometric Invariance in Computer Vision", MIT Press, Cambridge Ma, 1992.

[8] S. Startchik, C. Rauber, T. Pun, "Recognition of planar objects over complex backgrounds using line invariants and relevance measures", Workshop on Geometric Modeling & Invariants for Computer Vision, Xian, China , 1995, p. 301-307.

[9] A. Zissermann, D.A. Forsyth, J.L. Mundy, C.A. Rothwell, "Recognizing general curved objects efficiently", In [7]