

Logo Recognition Using Geometric Invariants

David S. Doermann, Ehud Rivlin and Isaac Weiss
Document Processing Group, Center for Automation Research
University of Maryland, College Park 20742
email: doermann@cfar.umd.edu, ehud@cfar.umd.edu, weiss@cfar.umd.edu

Abstract

The problem of logo recognition is of great interest in the document domain, especially for databases, because of its potential for identifying the source of the document and its generality as a recognition problem. By recognizing the logo we obtain semantic information about the document, which may be useful in deciding whether or not to analyze the textual components.

In this paper, we present a multi-level staged approach to logo recognition which uses global invariants to prune the database and local affine invariants to obtain a more refined match. We obtain an invariant signature which can be used for matching under a variety of transformations. We provide a method of computing Euclidean invariants, and show how to extend them to capture similarity, affine and projective invariants when necessary. We implement feature detection, feature extraction and local invariant algorithms and successfully demonstrate the approach on a small database.

1 Introduction

Logos are typically characterized as mixed text and graphic symbols which, when recognized, trigger an association of the object to which they are attached with a given group or organization. In the document domain, two analogous logo recognition tasks are of interest. First, given a document which contains a logo, classify the logo as one of a finite set of known logos or conclude that the logo is not present in the database. Second, given a representative logo (known or unknown), index into a database of documents and extract all documents which contain that logo. Several examples of logos are shown in Figure 1.

An approach to logo recognition can be formulated based on the use of *geometric invariants* to speed up searching and matching queries in databases. Invariants are quantities, computed from an image, which

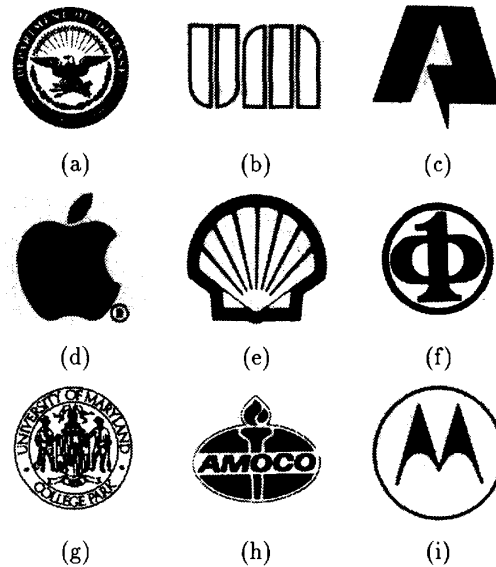


Figure 1: Logos

remain unchanged during changes in geometric conditions such as the point of view. These quantities, called *signatures* of the object, can be stored in the database in addition to the logos themselves. In the case of a database of documents, the invariant signatures can be computed for all logo-like regions. When a query is processed, its invariant signature can be matched with the stored signatures rather than matching the individual logo images. In this way we obtain a match even if the stored record and the query represent logos taken from different points of view, under scaling, rotation, or other geometric transformations. One of the advantages of signatures is that they can be indexed easily using current database technology (e.g., B-trees). We would like to compute a single signature for each logo that captures the features of interest in

the logo. We can then use one-dimensional indexing techniques to speed up access into the database.

This paper focuses on the various issues involved in the application of geometric invariants in a logo recognition process. In view of the diversity of possible document styles, we must deal with logos under a variety of transformations including translation, scaling, rotation and possibly skew. Rotation is less probable than the other transformations since a logo is likely to be oriented consistently with the text on the page. Transformations which distort the original shape of a logo should be recognized as different logos. In more general domains, such as a warehouse with a vision system which identifies logos on boxes, a logo may undergo perspective transformations.

Three main issues must be addressed independently of logo segmentation:

- partial occlusion of the logo or non-logo components in the logo region,
- affine transformation of the logo, and
- indexing into a very large database of logos

Section 2 overviews logo detection, Section 3 describes our approach to logo recognition and Section 4 presents some experimental results.

2 Logo Detection

Before recognition can be attempted, logos must be detected in the document. A logo can be described as a mixed text/graphic symbol which has the intent of associating an object (in this case a document) with a certain group or organization. Although graphic symbols are logos only if this association is intended, we can isolate regions which are “logo-like” prior to logo recognition. Page segmentation algorithms can delineate regions as logo candidates and generate a logo hypothesis (*logo detection*). We can distinguish a logo from standard textual components by its size and shape. Since the purpose of a logo is to provide a symbolic link to an organization, it is typically a perceptually salient marking, larger than individual text components. In addition, logos are typically confined to a compact region and are in an isolated portion of the document, not embedded in the text structure as other graphic figures. Thus we can often separate a logo candidate from other graphic components by its position on the page and by using knowledge of the document type (memo or letter vs. report), if this is known.

3 Logo Recognition

In this section we describe our approach to classification and verification. We believe that relying on a complete and accurate segmentation for all situations is not possible. We are approaching this problem under the assumption that although a logo hypothesis has been generated, portions may be missing, regions may be corrupted, or additional non-logo components may be present. “Partial match” and “not present in the database” are valid interpretations of the hypothesized logo region.

Although global, such invariants are useful for eliminating large numbers of candidates from the database. Our approach consists of three stages. In the first stage we use global shape classification to prune the database. In the second stage we use local geometric invariants [3] to generate a hypothesis. The third stage attempts to provide verification of the match. We briefly outline these stages in the following subsections.

3.1 Global Shape Classification

A typical database of logos may have thousands of entries. The problem of indexing into a large database is addressed by first attempting to partition the data into classes based on global shape properties and refining the interpretation of the elements using a decision tree. By doing so, we may attempt the matching process with a reasonable number of possible candidates.

We assume that the logo database consists of a graphic depiction of the logo, as well as slots for different attributes such as a priori information about orientation, textual components, landmark features, etc. Global invariants or approximate shape descriptors can be computed and used to prune the database. For example let us define the (i, j) th moment of a logo image I by

$$m_{ij} = \sum_x \sum_y x^i y^j I(x, y)$$

Moments can be given a physical interpretation by regarding gray level as mass, i.e. regarding I as composed of a set of point masses located at the point (x, y) . Thus m_{00} is the total mass of I , and m_{20} and m_{02} are the moments of inertia of I around the x and y axes respectively. The moment of inertia of I around the origin m_0 is

$$\sum \sum (x^2 + y^2) I(x, y) = m_{20} + m_{02}$$

It is easily verified that m_0 is invariant under rotation of I about the origin. Moreover, if I is rescaled, say by

a factor of s , then it is not hard to see that m_0 is multiplied by s^4 . Thus we can normalize I with respect to scale by requiring m_0 to be a constant. Alternatively, a ratio of two moments that have the same value of $i+j$ (m_{10}/m_{01}) is invariant to scale.

Other possible shape descriptors may include roundedness, elongation, etc. A survey of such descriptors can be found in Rosenfeld and Kak [2].

For the second stage of logo recognition we use local geometric invariants. The main problems for recognition are the extraction of reliable, meaningful features and the matching these features to the database. Since we are dealing with a small subset of possible logos, it is feasible to use the database to guide feature extraction. For large databases, under the types of transformations we are allowing, this would not be possible.

3.2 Computing Geometric Invariants

We are exploring a new method in which a shape descriptor which remains unchanged under affine and projective transformation can be extracted from a partial contour representation of a logo. The resulting description is an affine invariant signature of the image which can be matched against the signatures extracted from a database of known logos.

Our method of obtaining local projective and affine invariants is described and illustrated in [1]. Being local, the invariants are much less sensitive to occlusion than are global invariants. Computation of the invariants is based on a normalization method. This consists of defining a canonical coordinate system in terms of intrinsic properties of the shape. Since this canonical system is independent of the original coordinate system, it is invariant and all quantities measured in it are invariant.

Our invariants are computed as follows:

– For each pixel in the logo image:

- Define a window around the pixel and fit an implicit polynomial curve to it, using a quartic:

$$\begin{aligned} f(x, y) &= a_0 + a_1x + a_2y + a_3x^2 + a_4xy \\ &\quad + a_5y^2 + a_6x^3 + a_7x^2y + a_8xy^2 \\ &\quad + a_9y^3 + a_{10}x^4 + a_{11}x^3y \\ &\quad + a_{12}x^2y^2 + a_{13}xy^3 + a_{14}y^4 \\ &= 0 \end{aligned}$$

Once the curve order and window size have been chosen, the fitting itself can be done by standard methods. Simple least square fitting is ill conditioned because of the relatively large number of

unknowns. The SVD (Singular Value Decomposition) method is very successful in overcoming this problem and we obtain a reliable fit.

The following stages are performed analytically.

- Derive a canonical, intrinsic coordinate system based invariantly on the properties of the shape itself, independent of the given coordinate system. By doing so we eliminate all the unknown quantities of the original system (e.g. the viewpoint). To accomplish this: define an “auxiliary curve” which osculates the original fitted curve with a known order of contact. The canonical system is defined so that in it the osculating curve has a particularly simple, predetermined form.
- Transform the original fitted curve to this new system. Since the system is canonical, all shape descriptors defined in it are independent of the original coordinate system and are therefore invariants. Pick two invariants that are independent of the window size or the order of the fitted curve, and depend only on the shape itself.

– Plot one invariant against the other to obtain an invariant signature curve. This is based on the completeness property discussed above.

– If an invariant fit is needed, we repeat the previous steps, i.e. recompute the curve fitting in the new canonical system, and iterate until convergence.

3.3 Indexing and Verification

The matching of logo signatures is done for candidate logos remaining in the database. One strategy which can be undertaken is to qualitatively match the pointwise description of the original curve with those points in the signature of the unknown logo. To do this, a distance is computed from the candidate logo points to the original invariant curve. If a sufficient number of points are within a bounded distance, the logo is retained for further inspection. Ideally, a single logo hypotheses will remain after matching, but more likely, a set of logo hypothesis will remain and can be ranked.

Verification is an essential step after any matching method since only a subset of all of the features has been used to compute the invariants, and may be distorted due to noise. One standard verification procedure is to solve for the transformation and then correlate the model image with the logo. Another type of verification involves attempting to identify the

known textual components and to recognize them in the image. Certain classes of logos are very similar graphically, and differ only in their textual components. (For example, many universities with satellite campuses use a similar graphic, and change only the city name.) These types of similarities can be detected a priori in the database, and verification of the textual information can then be sought.

4 Experiments

We have implemented several portions of a complete system to demonstrate the feasibility of our approach to logo recognition. Figure 1 shows a subset of a logo database used for experiments and Figure 2 contains an example of the processed logos. Figures 2a and 2b show a pair of original logos (Shell and Apple), and 2c shows the Shell logo transformed by rotation, translation and scaling. Figures 2d-f show the results of preliminary processing which includes edge detection, edge linking and the selection of dominant contours. Figures 2g-i show the local invariant signature of each logo.

We can easily see the difference in the signatures of the Apple and the Shell logos. Although the signatures have some similarity, presumably due to their "roundness", they are different enough to distinguish between the two logos. The signatures of the two Shell logos are almost identical.

5 Conclusions

In this paper, we have presented a multi-level staged approach to logo recognition which uses global invariants to prune the logo database and local affine invariants to obtain a more refined match. Our approach allows us to obtain a signature which can be used for recognition under a variety of transformations without having to solve for the transformation explicitly. We have developed a method of computing shape invariants, and have shown how to extend it to capture similarity, affine and projective invariants when necessary. We have implemented feature detection, feature extraction and local invariant algorithms and successfully demonstrated the approach on a small database of logos.

The use of geometric invariants and qualitative matching shows promise for logo recognition and other document understanding problems including forms recognition, document class identification, and character recognition.

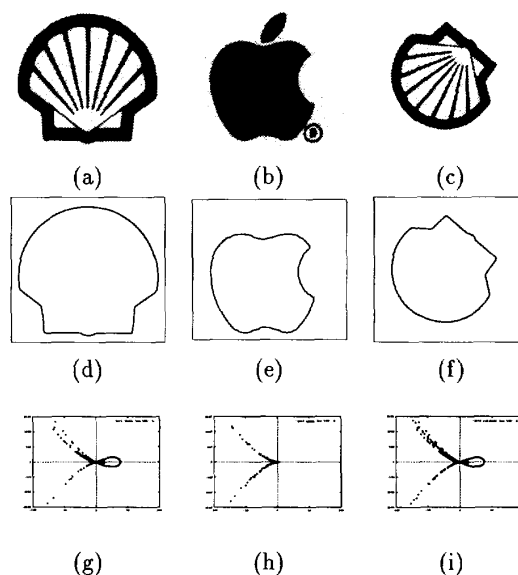


Figure 2: Two logo images are shown in a and b, along with a representative contours (d and e) and their invariant signatures (g and h). A translated, rotated and scaled instance of the shell logo is shown in c along with its contour (f) and signature (g).

References

- [1] E. Rivlin and I. Weiss. Local invariants for recognition. Technical Report CS-TR-2977, University of Maryland, College Park, MD, October 1992.
- [2] A. Rosenfeld and A. C. Kak. *Digital Picture Processing*. Academic Press, New York, 2nd edition, 1982.
- [3] I. Weiss. Projective invariants of shapes. In *Proc DARPA IUW*, pages 1125-1134, Cambridge, MA, April 1988.