Invariant-based data model for image databases

Michael Kliot Ehud Rivlin

Computer Science Department
Technion - IIT, Haifa, 32000 Israel 32000

Abstract

We describe a new invariant-based data model for image databases under our approach for shape-based retrieval. The data model relies on contours description of the image shape, and emphasizes the use of invariants. Efficient indexing is based on geometric invariant features, while semi-local multi-valued invariant signatures are used for ranking the answers. The advantages of the proposed approach are its ability to retrieve images in situations in which part of the shape is missing (i.e., in case of occlusion or sketch queries), its ability to handle images distorted by different viewpoint transformations, and its ability to flexibly answer queries based on logical or shape descriptions (query by example), or on combination of both. The approach also handles sketch based queries.

We implemented our data model in an object oriented database system with a SQL-like user interface. The paper presents experimental results demonstrating the effectiveness of the proposed approach.

1 Background

Invariants play a significant role in pictorial databases systems, since there is a need for shape-based retrieval which is insensitive to viewpoint transformations. In pictorial databases, global geometric invariants such as moments and Fourier descriptors are extensively used (Faloutsos et al. [1], Chang et al. [2]). The main advantage of global invariants is their relative numerical stability. However, they are sensitive to occlusion. *Local invariants* are usable in case of occlusion or shapes overlapping. The most popular and well-studied kind of local invariants are *differential* invariants, i.e. invariants, based on derivatives of different order. The main disadvantage of this kind of invariants is their sensitivity to noise, resulting from their differential character. Hybrid, or semi-differential invariants try to combine the advantages of the global and local approaches. Moons et al. [3] present a general theory of local semi-differential invariants. Bruckstein et al. [4] use locally applied global geometric invariants. An overview on the use of geometric invariants for object recognition can be found in Weiss [5].

Content based retrieval of images is a subject of extensive research. Content is mostly defined in terms of color, texture and shape. We will concentrate on works dedicated to shape-based retrieval (or combination of shape with color and texture).

Several works use *global numerical attributes* for indexing and retrieval of shapes. Those features are usually invariant only under Euclidean or similarity transformations, limiting the system. In QBIC[1] [1] the shape features are based on a combination of global features: area, circularity, eccentricity, major axis orientation and a set of algebraic moments. The assumption is that the shapes are unoccluded. The allowed shape queries are based on example image. Chang et al. [2] focus on automatic extraction of low-level visual features such as texture, color, and shape, especially in compressed form. In determining the similarity between different shapes, they use higher-level attributes, such as area, orientation, aspect ratio, etc. as well as intermediate-level representations, such as Fourier descriptors, chain code, etc. In the work of this group, the input image is generally allowed to undergo transformations up to similarity, and occlusion is not handled.

Another form of shape description is based on *feature points*, detected from the image. Pentland et al. [6] present a shape model that is based on "interconnectedness" of shape features, e.g., edges, corners, or high-curvature points. Mehrotra and Gary [7] use for shape representation a collection of few adjacent interest points, like maximum local curvature boundary points or vertices of the shape boundary's polygonal approximation. Mokhtarian et. [8] use a Curvature Scale Space image, which is a multi-scale organization of the inflection points (or curvature zero-crossing points) of the contour. This approach is stable for minor occlusion, however, since curvature is not affine invariant, the stability for affine transformation is problematic.

Del Bimbo and Pala [9] present a method for elastic matching of user-drawn sketches. Jain and Vailaya [10] use a histogram of the edge directions as a shape feature. The method is limited to similarity transformation and unoccluded input image.

Our work attempts to overcome three main limitations of the existing approaches: a restriction of the allowed viewing transformation to similarity, a requirement for a whole, unoccluded input image and a restriction of the queries to be example-based only.

[1] Query by Image Content

0-8186-8821-1/98 $10.00

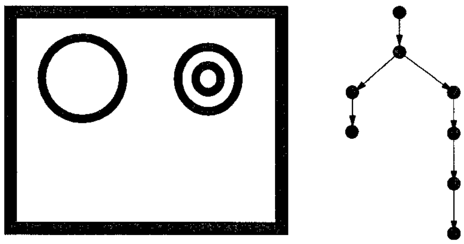

Figure 1: The image and its containment tree .

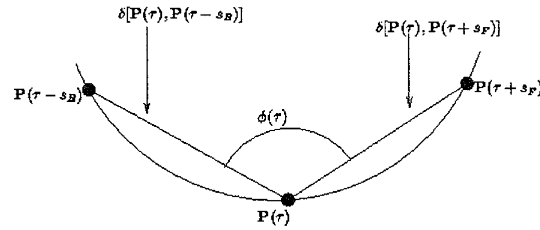

Figure 2: Local invariants for similarity transformation For a given s_F and s_B, the angle $\phi(\tau)$ as well as the ratio of lengths $\delta[\mathbf{P}(\tau), \mathbf{P}(\tau - s_B)]/\delta[\mathbf{P}(\tau), \mathbf{P}(\tau + s_F)]$ are similarity invariants.

2 Our approach

Selecting a characterizing set of features is an important step in designing an image database system. Our approach is based on representation of shape information as a set of contours. It emphasizes the invariance of the representation, allowing for variety of viewpoint transformation. In addition to similarity invariants, we exploit features which are invariant for wider range of transformations: affine and projective. The advantage of the exploited features is that if some shape information is missing, they can still be used for image description using only part of the contours. This allows to handle situations like partial occlusion and sketch-based retrieval.

As primitives for shape description we use geometric entities, such as circles, ellipses and straight lines. A number of entities can serve for both efficient indexing and query generation. We exploit the existence of these entities for filtering the database while searching for a candidate set of images answering a query. Features, on which the description is based, include quantities of the different geometric entities, and relations of their dimensions and positions. Another feature that we use reflects the relationship between the different contours. It exploits the property that being an inner contour of another is a projective invariant, given the same contour representation for the image and its transformed version. This property allows us to represent images exploiting the internal-external relations between contours. Each image is represented as a tree, called "containment tree ". The vertices of these trees are curves. For two curves, C_1 and C_2, the edge ($C_1 \rightarrow C_2$) exists if C_2 is inside C_1, and there is no C_3 such that C_2 is inside C_3 and C_3 is inside C_1. This representation can easily be obtained after extracting the curves. The example of the containment tree is presented on Fig. 1. The representation is unique, up to the order between the vertices on the same level. Thus tree-matching algorithms can be used to compare the representation of the input image with one of the library images. Algorithms for partial matching of containment trees are used for retrieval under partial occlusion and for sketch-based retrieval.

Since the image is regarded as a set of contours, we need an invariant representation of contours, which will be applicable in case of occlusion. Such representation is obtained via *local multivalued invariant signatures* ([4]). Invariant signature is a function of a curve, calculated pointwise, and invariant for a given set of transformations. That is, given the curve $\mathbf{P}(t)$ and its transformed version $\tilde{\mathbf{P}}(\tilde{t}(t)) = \mathbf{T}_\psi(\mathbf{P}(t))$, the values of the signature $\mathbf{S}$ at the corresponding points should be equal: $\mathbf{S}(\mathbf{P}(t)) = \mathbf{S}(\tilde{\mathbf{P}}(\tilde{t}(t)))$. Invariant signatures can be used for recognition of planar curves. The important property of invariant signatures is their applicability in situations where the input curve is partially occluded. Having an invariant signature, curve matching reduces to matching signatures. If the input curve is occluded, it should be matched to part of the library signature. The exploited approach combines the advantages of global and local methods, overcoming the ill-posed problem of finding derivatives, on which local invariants usually depend. The basic idea is to use invariant finite differences, with a scale parameter that determines the size of the differencing interval. On the first stage, the curve (the object boundary), given in arbitrary parameterization, should be reparameterized invariantly, using the lowest possible order of derivatives. It means that corresponding points should have the same parameter value, up to a constant cyclic shift, resulting from the fact that the curve initial point is unknown. Then, for each point on the curve, the invariant quantity based on locally applied global geometric invariants is calculated. Choosing a set of *locality parameters*, τ_i, $i = 1, \ldots, n$, in the invariant parameter space, we start from a point $\mathbf{P}(\tau_0)$ on the curve, for which we want to find invariants. We find the points $\{\mathbf{P}(\tau_0 + \tau_i)\}$, $i = 1, \ldots, n$. Using these points, we calculate any geometric invariant involving them. For example, under similarity transformation, the angle formed by $\mathbf{P}(\tau_0)$ and the two additional points $\mathbf{P}(\tau_0 + s_F)$ and $\mathbf{P}(\tau_0 - s_B)$ stays invariant (Fig. 2). The calculated invariants form a function of the curve parameter, called an *invariant signature*. After calculating the invariant signature, matching curves reduces to the matching of their signatures. If the locality parameter set is let to be free parameters rather than setting them in advance, we obtain a whole range of invariants at each point rather than a single value. The signature functions for curves

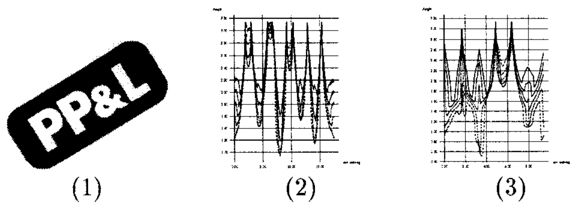

Figure 3: Input image and multivalued invariant signatures versus invariant parameter for some contours. (1) the input image, (2) the signature for the & symbol, (3) the signature for the external contour of P letter.

than a single value. The signature functions for curves become signature vectors or even continuum of values, i.e., surfaces or hyper-surfaces. Matching them will be less sensitive to peculiarities that may exist at some fixed pre-set value of the locality parameters.

In our work, local multivalued invariant signatures based on global geometric invariants have been implemented and used for the matching of corresponding contours of images. First, the contour is transferred to a parametric form by B-spline approximation. Next, the reparameterization and invariant signatures extraction is performed as described in [4]. To match two signatures, we map them to matrices having number of rows equal to the number of signatures in the multivalued signature, and the number of columns equal to the number of the samples . In order to estimate the unknown value of relative cyclic shift, we exploit reference points. Minimizing the difference between the matrices over all checked shift values provides a measure for the distance between the signatures. Detailed description of the matching procedure can be found in [11]. Figure 3 presents example image, with multivalued signatures calculated for some its contours. The obtained measure of distance between two curves allows us to rank images according to their distance from the input image. In this way, the candidate images are ranked.

The proposed approach is able to handle situations in which part of the shape information is missing. For example, if the input image is occluded, geometric entities and containment tree are used for pruning, and multivalued invariant signatures provide a measure of closeness to the input image. Sketch retrieval is treated in a similar manner.

3 Data model and query processing

We have implemented our approach in a heterogeneous database system having a SQL-like user interface augmented with sketch-based queries. The system is built on top of a commercial database system (Oracle), and can be activated from the Web. Each database entry includes contours representation of the image and geometric features, based on exploited entities. Some features reside in relational database (RDB), giving efficient performance and convenient

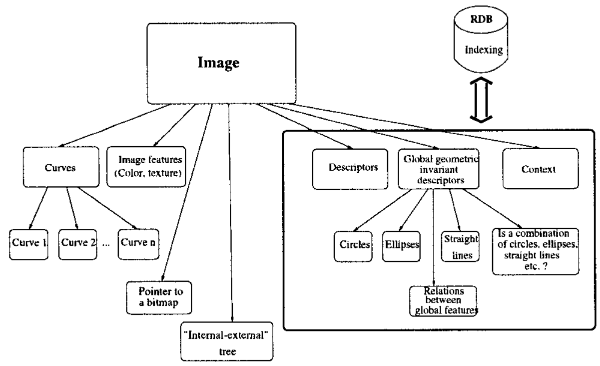

Figure 4: The image data model

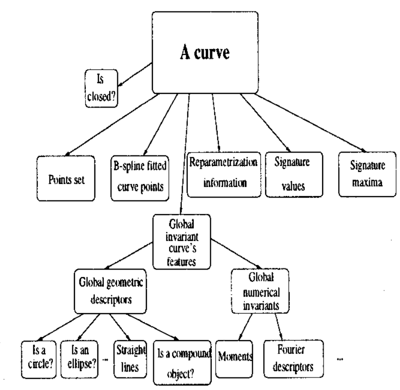

Figure 5: The curve data model

user interface. Figure 4 presents the data model. The model contains three types of fields:

1. Low-dimensional geometric descriptors and alpha-numerical fields:

   –*Descriptors* fields: store alpha-numerical descriptors of the image that the user assigns when the image is inserted (i.e data on a company, etc.).
   –*Context*: can include a verbal description of the image or its parts ("bird", "apple" etc.), characters appearing in the image etc.
   –*Global geometric invariant descriptors*

2. High-dimensional geometric and non-shape features:

   –*Image features* - color, texture, etc.
   –*Bitmap* field - pointer to a file.
   –*Containment tree* .

3. Curves:
   Each curve is a complex object represented according to the curve data model. For each curve the data model includes (see Fig. 5):

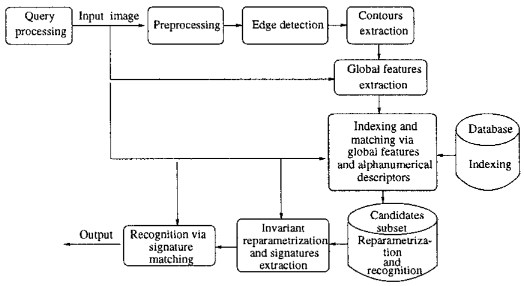

Figure 6: Query processing.

- An ordered set of points.
- Whether the curve is closed or not.
- The points of the B-spline which was fitted to the curve.
- The information about the invariant reparameterization (i.e. the values of the invariant parameter for each point).
- The invariant signature values - for each point, Num values are stored, where Num is a number of signatures (usually 1, higher for the multivalued case).
- Auxiliary information, i.e. signature maxima, global curve features, etc.

Our shape retrieval scheme consists of two main phases: filtering the database in order to drop the irrelevant images, and ranking the *candidates subset* according to the distance from the input image. For indexing we use geometric entities, such as circles, ellipses, etc. In our system we detect geometric entities from image contours. Circle detection is based on the fact that the characteristic ratio, which equals to the quotient of the area of the curve to the square of its perimeter, is known to be minimal $(1/4\pi)$ for a circle. The curve is detected as a circle if its characteristic ratio is close to $1/4\pi$ up to a predefined threshold. The detection of ellipses is based on the fact that any ellipse can be transformed to a circle by appropriate affine transformation. So, if we calculate affine invariant signature for an ellipse, it should be equal to that of a circle, which should be constant since circle is fully symmetrical image. Thus, we can calculate the affine-invariant signature described in Section 2, setting its parameters to some predefined values, and check the difference between the obtained signature and the apriori known constant value. The curve is detected as an ellipse if the difference value is close to zero. The number of these entities can serve for efficient filtering of the image collection. Final ranking within the set is based mainly on the semi-local multi-valued invariant signatures. The number of geometric entities in the image instances provides an accurate filtering. In order to avoid miss-matches, we allow the database image to have more entities (circles, ellipses) than the input image. In addition, the threshold used for entities detection in the input image are tighter than those for database images. Thus, correct database images are not pruned out.

The general scheme of query processing is presented on Fig. 6. After edge detection and curves extraction from the input image, geometric entities are detected. This global features extraction is used as a basis for the filtering process. Relational database is used in this stage in order to retrieve the candidates subset. Features used for indexing to the relational database and eventually for pruning the candidates set vary from the number of the curves and the geometric entities, their relative dimensions, etc. The relational database includes alpha-numerical data as well. The query may contain this kind of information, like the organization name (for the trademarks database) etc., and this part of the query will be processed as a regular query (further pruning the set). The candidates obtained at this stage go to a matching with the input image. The final set, ordered according to the similarity measure is given as an answer.

4 Experiments

In the experimental part of the paper we present a number of queries illustrating capabilities of the proposed approach and data model. The queries in SQL-like notation were run on combinations of different databases: SQUID[2] database of about 1100 marine images, and two databases from the University of Maryland[3] : a database of 100 trademark images and a databases of 300 "road sign" images.

First we present an example of a real image, photographed under partial occlusion. The extracted trademark part of the image served as an input to the query against the trademarks database. Figure 7 presents the input image and the retrieval results.

Figure 8 illustrates the filtering obtained by using invariant features. In Fig. 8(1) the user operated with the query

select title from Marines ordered by dist(mar100).

The results are ordered according to the distance measure based on invariant signatures. In the following example (Fig. 8(2)) the user operated with

select title from ALL ordered by dist(PD112ABU-tr).

The input image is a transformed version of one of the database images. The next example (Fig. 8(3)) illustrates retrieval from the combination of trademarks and "road signs" databases. The query is

select title from (Trademarks + Signs) ordered by dist(PD112ABU-tr).

Our data model allows queries based on the combination of the alpha-numerical and pictorial information. Figure 8(4) present the results of processing the query

select title from Trademarks where org_kind != "company" AND n_ellipses $\geq$ 2

[2] *www.ee.surrey.ac.uk/Research/VSSP/imagedb/demo.html*
[3] *ftp://documents.cfar.umd.edu/pub/contrib/databases/*

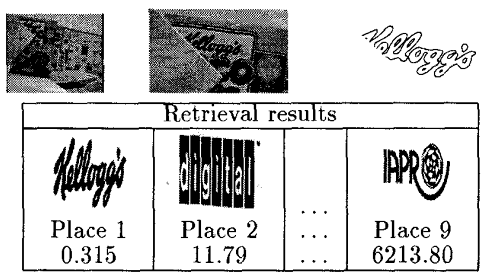

Figure 7: Example of real occluded scene. Above: the photographed scene, the region containing the trademark and the extracted contours. Below: retrieval results.

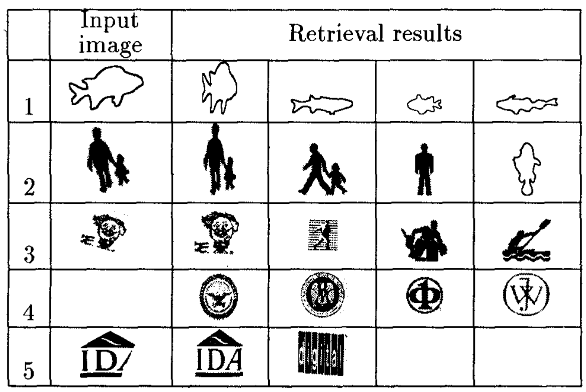

Figure 8: Results of queries

The next example (Fig. 8(5)) illustrates retrieval under partial occlusion. The retrieval is based on the partial matching of the containment trees .

select title from (Trademarks +Signs) ordered by OCCLUDED dist(logo54) .

This experiment demonstrates the discrimination power of containment tree - after the pruning stage, candidate set contains only two logos.

5 Summary

We have described a data model for image databases that relies on contours description of the image shape, and emphasizes the use of invariants. We have used geometric invariant features for efficient indexing, while semi-local multi-valued invariant signatures have been used for ranking the answers. Our approach retrieves images while part of the shape is missing, handles images distorted by different viewpoint transformations, and flexibly answers queries based on logical or alpha-numerical descriptions, shape (query by example), or combination of both.

We implemented the data model in an object oriented database system with a SQL-like user interface. We have presented examples of different queries, including queries based on combinations of alphanumerical and pictorial information. Our experimental results demonstrate the effectiveness of the approach.

References

[1] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovich, and W. Equitz, "Efficient and effective querying by image content", *Journal of Intelligent Information Systems* **3**, 1994, pp.231–262.

[2] S.-F. Chang and J. Smith, "Extracting multidimensional signal features for content-based visual query", in *SPIE Int. Symp. on Visual Communications and Image Processing*, 1995.

[3] T. Moons, E. Pauwels, L. V. Gool, and A. Oosterlinck, "Foundations of semi-differential invariants", *Int. Journal Of Comput. Vision* **14**, 1995, pp.25–47.

[4] A. Bruckstein, J. Holt, A. Netravali, and T. Richardson, "Invariant signatures for planar shape recognition under partial occlusion", *CVGIP:Image Understanding* **58**, 1993, pp.49–65.

[5] I. Weiss, "Geometric invariants and object recognition", *Int. Journal of Comput. Vision* **10**, 1993, pp.207–231.

[6] A. Pentland, R. Picard, and S. Sclaroff, "Photobook: Tools for content-base manipulation of image databases", in *Proceedings, SPIE Conf. on Storage and Retrieval of Image Databases II*, 1994, pp. 37–50.

[7] R. Mehrotra and J. Gary, "Similar-shape retrieval in shape data management", *IEEE Computer* **28**, 1995, pp.57–62.

[8] S. F. Mokhtarian, S. Abbasi, and J. Kittler, "Efficient and robust retrieval by shape content through curvature scale space", in *Proc. of International Workshop on Image DataBases and MultiMedia Search*, Amsterdam, The Netherlands, 1996, pp. 35–42.

[9] A. D. Bimbo and P. Pala, "Visual image retrieval by elastic matching of user sketches", *IEEE Trans. Patt. Anal. Mach. Intell.* **19**, 1997, pp.121–132.

[10] A. Jain and A. Vailaya, "Image retrieval using color and shape", *Pattern Recognition* **29**, 1996, pp.1233–1244.

[11] M. Kliot and E. Rivlin, "Invariant-based shape retrieval in pictorial databases", in *Proc. of ECCV*, Freiburg, 1998, pp. 491–507.