

Interactive Model-Based Matching Retrieval

L. Cinque, S. Levialdi, A. Malizia, R. Mancini

Dipartimento di Scienze dell'Informazione Università "La Sapienza" di Roma
Via Salaria 113, 00198 Roma, Italy

Abstract. We present a system for image retrieval based on the query-by-sketch technique. In our system the user formulates a query by drawing a sketch of the shape of the object (the "model") he is looking for and requests all images, stored in a database, containing such similar object. The system evaluates the similarity between the shapes in the database and the rough sketch by applying, sequentially, three similarity criteria based on the shape description. The system tries to emulate some typical human features like the power of visual communication, the fast cognitive feedback loop and the relatively small memory load for a natural and inexpensive human-computer interaction.

Keywords- visual communication, human-computer interaction, shape description, query-by-sketch, image retrieval, model-based matching.

1 Introduction

Most image retrieval work has focused on particular techniques exploiting image contents [1] for matching whole images or shapes extracted from such images. The four main matching features are: shape, color, texture, and spatial and/or temporal relations. In shape matching, most of the 2D contour matching techniques developed for computer vision apply, although several researchers have particularly targeted their work towards image retrieval [2],[3]. Grosky [4] developed an algorithm for matching strings of chain codes describing the contours of picture objects. Grosky and Mehrotra [5],[6] used index trees to access a database of 2D contour models of industrial parts. Califano and Mohan [7] developed a related indexing method that uses multi-dimensional global invariants of tuples of local interest features as indexes that vote for object models in the database. This techniques also has some relationship to geometric hashing [8]. Del Bimbo [9] retrieves images containing specific 2D shapes by using an elastic matching technique. The user sketches the shape and the system returns these images that contain similar shapes in ranked order.

Shape matching is also part of QBIC (Query by Image Content) [10] and is based on classical shape measures such as area, circularity, eccentricity, major axis orientation, and algebraic moment invariants.

In this paper, we present an interactive model-based matching system for image retrieval. The user draws a sketch of a shape (representing a "model"),

and then the system, by applying three different successive similarity criteria, each criterion on a set of object shapes, reduces the search space as a result of the application of the previous one, finding the “candidate” shapes which are nearest to the drawn model. Such candidates are visualised for the user as query results. The three employed similarity criteria are based on the M.E.R. (Minimum Enclosing Rectangle), the Signature and the Distance Curve of the considered model shape. The use of three different similarity features in sequence does not allow the formulation of a query enriched by boolean operators, but enables the system to reduce, at each step, the set of shapes provided as a result. The main consequence of such a reduction is the increase of precision at the expense of recall [11]. Moreover, the system becomes very easy to use from the user’s point of view; in fact, not only the system functionalities are immediate and easy to learn simply by navigating through the system, but also the user can feel in control of his dialogue along the interaction. This is due to the presence of different recovery functionalities, the progressive querying technique [12] and of two browsers, one over a portion of the database, and the other along the query history, allowing the query-by-browsing technique.

This paper is organised as follows. In Section 2 we present an overview of our system architecture, while in Section 3 the adopted shape description is provided. A retrieval example is discussed in Section 4, which also includes the matching algorithm and the interactivity aspects of the system. Finally, Section 5 highlights the main results.

2 System Architecture Overview

The system architecture, shown in Fig. 1, includes four main components: the interface, the matching engine, and two databases, one for real images and the other for the objects’ shapes extracted from the stored images. The query may be formulated by simply drawing a sketch of the object’s shape the user is interested in retrieving.

As shown in Fig. 1, the interface contains three modules, VisEd, SQD and the query Results windows.

VisEd [13] is a visual editor that allows users to draw a sketch of the desired object’s shape by using a mouse (the use of a graphical pad is also developed).

The SQD (Sketch Query by Dialog) is a module which imports the sketch from the VisEd module using a shared memory communication technique and the query, formulated as the drawn shape, is processed. SQD allows the user to browse the shapes stored in the database. Moreover, the system stores the history of the performed query, so that the user can also browse the past query by following the principle of the progressive querying. The last interface module, is given by the query Results windows. After a query has been performed, the set of the retrieved shapes is visualised and, by clicking on one of them, the user can load the selected shape into the VisEd module if any editing is needed, or into the SQD module to submit a new query.

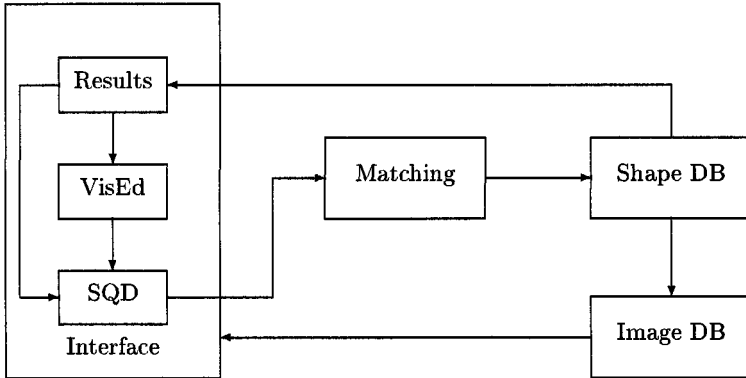


Fig. 1. Relationship of system components

The second architectural component is the matching one, which links together two different phases: indexing and retrieval. In the first phase, after a shape is drawn and the query is submitted, the shape features are computed. In the second phase, a comparison with the features of shapes stored in the database is performed. Such comparison is done following a given similarity criterion, and all the shapes which are similar to the drawn one, are visualised in the query window results.

The third and fourth components are two databases, one of real images, and the other of object shapes. Such shapes have been extracted from the objects present in the images stored in the database.

3 Shape Description

Different shape boundary features have been extracted and used to form a feature index for the database of object shapes. Such features are: 1) the Minimum Enclosing Rectangle (M.E.R.), 2) the Signature, and 3) the Distance Curve.

The Minimum Enclosing Rectangle of an object is defined as the smallest rectangle containing the given object. The M.E.R. is a feature which seems interesting as the first one for matching, since it gives an idea of the object elongation, thus we use the width/height ratio of it for a first classification of objects in the database.

In order to extract this feature we have to compute a main axis across the body-centre of the object minimizing the sum of the distances between the contour points and the axis itself. We next compute an orthogonal axis perpendicular to the previous one across the body-centre. At this point the object is rotated until the main axis is parallel to the y-axis of the coordinate system. In this way, we obtain a new coordinate system for the points of the object and is now easy to compute the width and height of the M.E.R., as the difference between the maximum and the minimum ordinates and maximum and minimum abscissae.

The second feature is the Signature, which is a polar representation of an object contour [14]. We sample, at fixed angles, the distance between points of the contour and the body-centre. This kind of angular sample is a fixed length representation, thus allowing to manipulate and efficiently match contours with small amounts of memory. In our work, the signatures are a sorted sequence of values which are lengths of N vector rays, computed for the same angles, projected from the body-centre of each object to its contour points. The assumption that the contours we examine are closed implies the periodicity of the signature. The properties of a contour Signature are its rotation, scaling and translation invariance. Moreover, the Signature is also a fixed length representation which is an important property to perform matching between images.

The third and last employed feature used to retrieve object shapes from our image database is the Distance Curve [15]. We compute it by finding the distances between the contour of the object and its convex hull [16]. For every shape, stored in the shape database, the features discussed above are computed, and inserted into a structured index mechanism with a pointer to the stored shape.

We use these features in sequence. By applying the M.E.R., we reduce the range of shapes to be next analysed, including only those shapes with the width/height ratio similar to the one of the drawn sketch. At this point, a comparison between the signature of the sketch and those of the shapes obtained by applying the M.E.R. is performed. Finally, if the user decides that the query should be further restricted, the third feature can be employed.

The application of these features in sequence has an interesting consequence, since such an application reduces, at each time, the research space. Moreover, the Distance Curve is the most restrictive feature, among those used, and particularly useful when the objects have concavities.

4 A Retrieval Example

We will now describe our application to retrieve an image from the database. In this example we consider two different databases: 1) containing over 100 gray level images with objects, and 2) containing over 200 hand-drawn shapes and hand extracted shapes.

Initially, the user can see two windows on the screen: VisEd (Visual Editor) and SQD (Sketch Query by Dialog).

In the VisEd window the user can draw a rough sketch of the object to retrieve the most similar in the database. In the example of Fig.2, the user wants to retrieve a knife; he draws a sketch of a knife in the VisEd window using a mouse, and, by clicking onto the *Query* button, the shape of the object is sent to the SQD engine.

After this step, the (resampled) shape of the object is visualized in the Query Space window of the SQD (left-hand side frame), where the user can perform the real query, while in the Data Base Space (right-hand side frame), the user can browse all the object contained in the database. The query may be wide or

restricted; in the first case, only two similarity criteria are applied, the M.E.R. and the Signature respectively, while in the second case, also the Distance Curve is applied. After selected the kind of query which has to be processed (if wide or restricted), the matching phase starts. The results of the application of the matching procedure are visualised in a results window, in the left-hand side of Fig. 2.

After the results are visualised, the user may decide whether to modify the model sketch or not. If he wants to proceed along this way, due to the interactivity of our application, he can modify his shape by clicking onto the *Edit* button of the VisEd window and redraw or edit his object. After the user has edited his sketch, he can then choose to iterate the query.

If the user clicks in the Results Window on the knife he was looking for, such knife is automatically displayed in both the VisEd window and in the Database Space of the SQD. Moreover by using the browser on the left-hand frame of SQD, the user can see all the past formulated queries and, eventually, re-use or refine them.

Finally, in order to visualize the image containing such found shape (a knife in our case), a *ShowPhoto* button in the SQD module may be clicked. Fig. 3 shows such an image.

5 Conclusions

In this paper, some image retrieval issues have been discussed and a system based on the query-by-sketch technique has been proposed. In our system the user formulates a query by drawing a sketch of the shape of the object (the model) he is looking for and requests all images, stored in a database, containing such similar object.

Sketches corresponding to all images matching the query are retrieved and displayed. The comparison between the shapes in the database and the drawn sketch is done by applying, sequentially, three similarity criteria based on the M.E.R., the Signature and the Distance Curve.

The similarity criteria, being sequentially applied, reduce each time the search space and increase precision versus recall.

We have shown the main system's functionalities and characteristics by means of a retrieval example. The system allows the user to reach a high interaction level and, since the working environment is so natural, he does not distinguishes VisEd and SQD as two different environments (one for editing and the other for queryin), but he simply moves between them in a very easy and natural way, as we expected in the design phase.

The user can apply the progressive querying technique, for which he can modify or refine a formulated query, but he can also deploy the query-by-browsing technique, with which he can browse the data stored in the database and, if needed, use them to formulate a new query. Moreover, there is also a history mechanism which allows the user to browse along the past formulated query.

The system tries to emulate some typical human features like the power of visual communication, the fast cognitive feedback loop and the relatively small memory load for a natural and inexpensive human-computer interaction. Our system may be considered as a first step in the direction of easy-to-use and easy-to-learn interactive image database systems provided with comfortable and effective visual environments

References

1. V.N. Guidivada and V.V. Raghavan, "Content-Based Image Retrieval Systems", *IEEE Computer*, **22**, N. 12, pp 18-22, 1995.
2. V.E. Ogle and Michael Stonebraker, "Chabot: Retrieval from a Relational Database of Images", *IEEE Computer*, **22**, N. 12, pp 40-48, 1995.
3. R.K. Srihari, "Automatic Indexing and Content-Based Retrieval of Captioned Images", *IEEE Computer*, **22**, pp 49-56, 1995.
4. W. I. Grosky, Y. Lu, "Iconic indexing using generalized pattern matching techniques", *Comp. Vision Graphics and Image Processing*, **35**, pp. 383-403, 1986.
5. W. I. Grosky, R. Mehrotra, "Index-based object recognition in pictorial data management", *Comp. Vision Graphics and Image Processing*, **52**, pp. 416-436, 1990.
6. W.R. Mehrotra, F. K. Kung, I. Grosky, "Industrial part recognition using a component-index", *Image and Vision Computing*, **3**, pp. 225-231, 1990.
7. A. Califano, R. Mohan "Multidimensional indexing for recognizing visual shape", *IEEE Trans. on Pattern Analysis Machine Intell.*, **4**, pp. 373-392, 1994.
8. Y. Lamdan, H. J. Wolfson "Geometric hashing: a general and efficient model-based recognition scheme", *Proc. of 2nd Intern. Conf. on Computer Vision*, pp 238-249, (1988).
9. A. Del Bimbo, P. Pala, S. Santini, "Visual image retrieval by elastic deformation of object sketches", *IEEE Symposium on Visual Languages.*, **4**, pp. 216-223, 1994.
10. M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steel and P. Yanker, "Query by Image and Video Content: The QBIC System", *IEEE Computer*, **22**, N. 12, pp 23-32, 1995.
11. P.D. Bruza, *Stratified Information Disclosure: a Synthesis Between Hypermedia and Information Retrieval*, PhD Thesis ISBN 90-9005760-9, Katholieke Universiteit Nijmegen, NL, 1993.
12. S.K. Chang, M.F Costabile and S. Levialdi, "Reality Bites - Progressive Querying and Result Visualization in logical and VR Spaces", *IEEE Symposium on Visual Languages*, pp. 100-109, 1994.
13. R. Mancini, "Interacting with a Visual Editor", *Proc. of the Int. Workshop on Advanced Visual Interfaces AVI'96*, ACM Press, T. Catarci, M.F. Costabile, S. Levialdi and G. Santucci Eds., pp. 125-131, 1996.
14. R.C. Gonzales, R.E. Woods, "Digital Image Processing", Addison-Wesley, Reading MA, 1992.
15. S.Liu-Yu, M.Thonnat, "Description of Object Shapes by Apparent Boundary and Convex Hull", *Pattern Recognition*, Vol.**26**, No. 1, pp. 95-107, 1993.
16. F.P. Preparata "Computational Geometry: an introduction", Springer-Verlag, 1988.

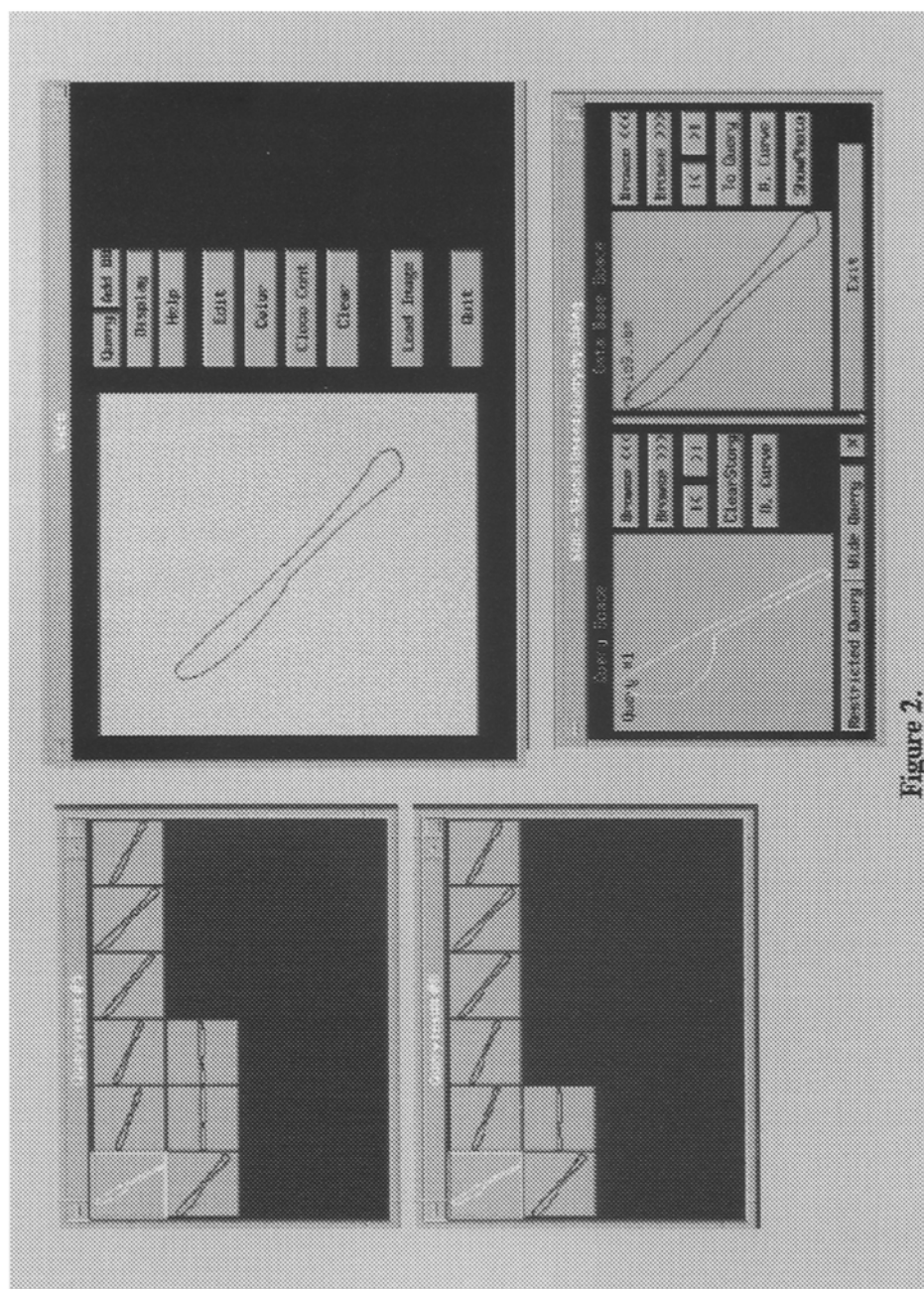


Figure 2.

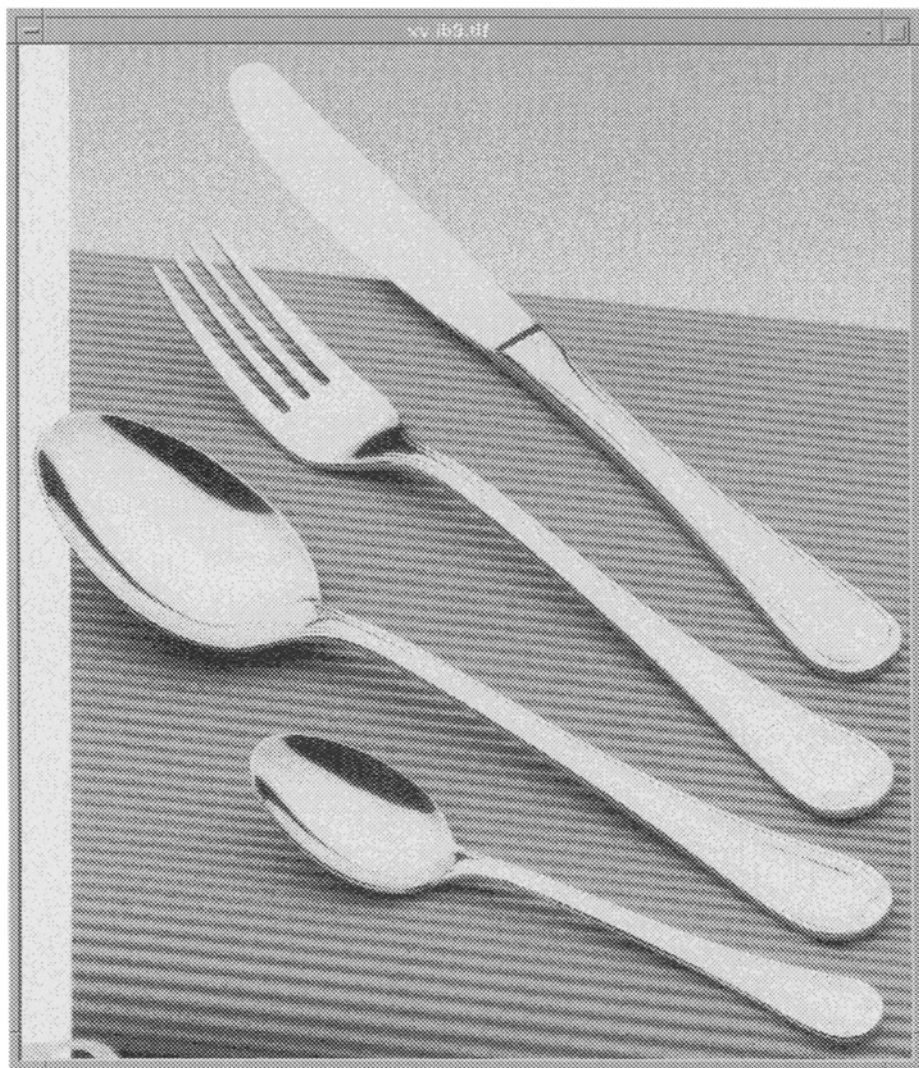


Figure 3.