Combined 2D shape characterization for image databases *

J. M. Fuertes†, M. Lucena†, N. Pérez de la Blanca‡, J. Chamorro-Martínez‡
†Departamento de Informática. Escuela Politécnica Superior de Jaén
Universidad de Jaén - SPAIN
‡Departamento de Ciencias de la Computación e I.A. ETSI Informática
Universidad de Granada - SPAIN

Abstract

In this paper we present a Content-Based Images Retrieval (CBIR) system which uses characteristics of a region occupied by a shape and the outer contours of the same. The CBIR system is a two-stage image retrieval system: it first extracts a subset of images which contains those shapes most similar to those contained in the query shape, based on the use of morphological characteristics of the objects considered. It then indexes an extracted subset of shapes from the database in view of the similarity of the deformation of the outer contours of these shapes. For this first stage, we propose a geometric method of two-dimensional shape recognition based on the use of characterizations of those curves given by the morphological operators of erosion, closure, aperture and dilation with structural elements of increasing size. The second stage of this system uses an invariant characterization of shapes based on a deformable model. In the experiments, we have illustrated that the CBIR system can accurately retrieve the most similar images from the database by using scaled, rotated, small deformed or mirrored sketched query images.

Keywords: Image database, Morphological operators, Deformable models, Similarity measure, Image retrieval.

1 Introduction

In this paper we propose a system of indexation and query based on shape matching. Traditional shape description approaches include the extraction of a set of shapes and their representation as points of a multidimensional space. In related bibliography, we found description and retrieval systems based on shape. Representations using approaches to polygonal shapes and matching via polygonal vertexes [16]; shape matching using relaxation techniques in order to establish combinations of pairs of angles among shapes [3]; use of chains to represent and describe shapes [2][6]; comparing images using the Hausdorff distance [7]; and use of invariant moments, to represent characteristics of a shape with regard to transformations of translation, rotation and scale [18][10][14].

*This research was partially sponsored by the Dirección General de Enseñanza Superior CICYT, under grant PB98-1374

This paper focuses on describing a method of image extraction based on the similarity between characteristics of the shape of images under consideration. In order to match objects contained in the database, we have proposed a 2D shape retrieval system based on a combination of morphological filters and deformable models.

This approach is made up of two well-defined steps: first, from the full database, the system extracts a subset of those objects which are most similar to the query using the invariant characterization of finite sets based on the mathematical morphology given in Section 2.1. A similarity threshold is defined in order to establish a rejection-class from each database query. In a second step, the system reorders those shapes retrieved using the criteria of similarity proposed by the deformable model described in Section 2.2. This approach matches shapes according to the amount of necessary effort to deform an object and transform it into another one. References found in the bibliography related to the idea of object description using deformable models are for example [19][1][15].

2 Combining Morphological Filters and Deformable Models

According to results obtained in Section 3, the characteristics showed by both techniques suggest, in fact, that the two techniques should not be understood as alternatives to be used, but as complementary techniques. Therefore, in order to match two shapes, both the characterization of the area of the shape, and the deformation energy of its outer contour and the data, must be taken into account to index the result of the query.

Establishing similarity/matching between an input shape and the shapes contained in a database will be carried out in two different stages.

1. Using a given shape, as an input to the recognition system, we intend to index the shapes contained in the database comparing its similarity to the query shape. Thus, we will extract the $n-$first most similar shapes to the input prototype. The first extract will be carried out by using the morphological filtering technique shown in Section 2.1. In this way, we characterize the similarity by using the information supplied by the area filled by the shape.

2. Once we have the subset of the database shapes most similar to the one given in the input, a similarity measure is again established between the input shape and the subset of shapes retrieved in the first phase. To calculate the second similarity measure, a deformable model technique is used, as shown in Section 2.2. Thus, in this second step we characterize the similarity by using the information supplied by the shape contour.

The proposed approach involves linking both techniques. The first one works as a database consultation filter to define a subset of shapes that the second one will order according to the outer contour.

2.1 Morphological Filters

Using morphological models in shape recognition finds its development in the theory of random closed sets [11]. One image (background white) with a shape (foreground black) can be modelized as a point process or a random set; carrying out the point process is a random set. In this way we can define the random variable $N(E)$ as the sum of the number of points of set E.

A random closed set is characterized by $Q(T) = P(A \cap T = \varnothing)\ \forall T \in \mathcal{T}_0$, where $\mathcal{T}_0$ is some 'basis' for $\mathcal{T}$, closed under finite unions.

The most obvious characteristic of a random set A is the probability that point x belongs to the set: $p = P(x \in A) = 1 - Q(\{x\})$, for any such point x.

This value can be calculated by: $\hat{p} = \frac{\text{area}(A \cap W)}{\text{area}(W)}$, whereas the area is accounted for by means of counting the points in observations A and $A \cap W$ over a window W. The calculation given in [17] and [11] may be applied to define necessary operators. Given two sets A and B we determine: $\check{A} = \{-x \mid x \in A\}$, $A \oplus B = \{x + y \mid x \in A, y \in B\}$ and $A \ominus B = (A^c \oplus B)^c = \{x \mid B + x \subset A\}$, then we define:

$$e_T = \frac{\text{area}\{(A \cap W) \ominus \check{T}\}}{\text{area}\{W \ominus \check{T}\}} \qquad d_T = \frac{\text{area}\{(A \oplus \check{T}) \cap (W \ominus \check{T})\}}{\text{area}\{W \ominus \check{T}\}} \tag{1}$$

whereas the terms of the numerator of both expressions, using the Serra notation [17], denote *erosion* and *dilation*. The dual operations (o_T and c_T) denote *opening* and *closing* over disc T are given by the following, respectively: $(A \ominus \check{T}) \oplus T$ and $(A \oplus \check{T}) \ominus T$.

The curves characterize the objects of the image according to a variation in the relationship of surfaces existing between what would be the area of the object originally taken into account and that which is obtained by a morphological filter for each of the operators, when the radius of disc T varies.

The question of shape recognition based on a set of characteristics may be redefined as the establishment of a similarity measure between an object considered and the objects in the database. This object matching procedure follows from two different stages which are described now in detail.

2.1.1 Defining a measure of distance between the different characteristic vectors

The result of defining a shape recognition system is a vector of values representing the response from the shape of four morphological filters in a predetermined number of different diameters of T of equal distance.

The similarity of two shapes shall be determined by the Euclidean distance between the vector of characteristics X of the input shape and vectors Z_i of those shapes contained in the database:

$$d(X, Z_i) = \sqrt{(X - Z_i)^T (X - Z_i)} \tag{2}$$

The measure of distance given in the above equation (2) allows for a mechanism of indexation of the different shapes considered in the database.

2.1.2 Extension of the Approach to Non-isotropic Elements

This paper has included another experiment via the consideration of non-isotropic structural elements such as rectangulars positioned on their respective centers of gravity.

If we want to preserve an invariance to transformations of rotation, a study would have to be carried out on the morphological curves bearing in mind that the structural element rotates 360°.

Because this procedure considerably increases the computational cost of the process, we use a discretization of the first two quadrants. Since the structural elements are positioned on their center of gravity, discretization will affect the third and fourth quadrants equally.

The curves do not differ in the same sequence, but the groups of morphological curves have been shifted, so the proposed solution is:

1. Determine the displacement produced by the groups of curves.
2. Apply a measure of distance between the new vectors obtained whose elements have been previously corrected by possible displacement due to transformations of rotation.

2.2 Deformable Models

With respect to shape contour we can define a property of deformation, between the object in question and the analyzed correspondent, which allows us to match objects in the database. This leads to the use of a model of object contour deformation.

Different ways of modeling deformations are described in the bibliography [9] [5] [8] [4].

We will have to define some type of measure which allows us to evaluate the *effort* or *energy of deformation* which we have to apply to a given contour in order to adapt it to another.

In this paper we have used the deformable model given by Nastar in [12] [13] which is decribed briefly below.

2.2.1 Analytical Calculation of the Proper Vibration Modes

Nastar ([12]) proposes a calculation to determine the analytical expression of the modes and, thus, considerably reduce the costs of its computational calculation.

In the case of a closed curve, the eigenvalues ω^2 can be defined by the dispersion equation:

$$\omega^2 = \frac{4K}{M} \sin^2 \left(\frac{p\pi}{n} \right) \tag{3}$$

We should point out the fact that if N is odd, $\omega_0^2 = 0$ is the only simple eigenvalues and the rest are all double eigenvalues. If N is even, $\omega_0^2 = 0$ and $\omega_{n/2}^2 = 4K/M$ are the only simple eigenvalues, the others will be double eigenvalues.

For the simple eigenvalues, the expression of the associated eigenvector is given by :

$$\phi = \left[\cos\frac{2p\pi}{N}, \ldots, \cos\frac{2p\pi n}{N}, \ldots, \cos\frac{2p\pi N}{N}\right]^T \tag{4}$$

where $p \in \mathcal{B}(N)$, the *first Brillouin zone* [12].

$$\mathcal{B}(N) = \begin{cases} \{-\frac{N-1}{2}, \ldots, \frac{N}{2}\} \text{ if } N \text{ is even} \\ \{-\frac{N-1}{2}, \ldots, \frac{N-1}{2}\} \text{ if } N \text{ is odd} \end{cases} \tag{5}$$

For double eigenvalues, we obtain a couple of orthogonal eigenvectors given by:

$$\begin{cases} \phi = \left[\cos\frac{2p\pi}{N}, \ldots, \cos\frac{2p\pi n}{N}, \ldots, \cos\frac{2p\pi N}{N}\right]^T \\ \\ \phi' = \left[\sin\frac{2p\pi}{N}, \ldots, \sin\frac{2p\pi n}{N}, \ldots, \sin\frac{2p\pi N}{N}\right]^T \end{cases} \tag{6}$$

2.2.2 Deformation Spectrum and its Application

As seen in the beginning of Section 2.2, the idea is to define an object of a deformable model in order to adapt it to another, thus allowing us to measure the *energy used in the process as object similarity.*

This representation will be called modal spectrum [12]. Basically, it describes which modes and in which quantity the different modes are excited in order to alter a shape until it matches another. Thus we can obtain an indicator of the quantity of *deformation energy* needed for the process:

$$E_{deformation} = \frac{1}{2}\mathbf{U}^T\mathbf{K}\mathbf{U} = \frac{1}{2}\sum_{i=1}^{N}\omega_i^2\tilde{u}_i^2$$

Two deformations will be *similar* when their corresponding displacement fields $\mathbf{U}_1$ and $\mathbf{U}_2$ are similar, except the rigid transformations, which contain zero deformation energy. In other words: *two deformations will be similar when their corresponding deformation spectrum are similar.*

2.2.3 Similarity Measure Between Two Given Deformations

Since defining a vector with characteristics that allows us to determine the deformation of a given shape is necessary, we will have to define a shape prototype, common to all those shapes chosen, against which deformation can be calculated. The circular shape of a r radius selected *a priori* has been chosen as a prototype.

Using the deformation spectrum of a shape given to another circular one, we will have to be done to the other shapes shown, under the same conditions of r radius and number of points of sub-sampled contours. In other words, all deformation spectra of the shapes shown will be obtained. Then, we would only have to establish a distance between the deformation spectrum of the shape given as input and of the shapes contained in the database to establish a similarity rate. As distance, we can use any Euclidean rule as:

$$d(D_1, D_2) = \frac{1}{p}\sqrt{\sum_{i=1}^{p}(\tilde{u}_i(D_1) - \tilde{u}_i(D_2))^2} \tag{7}$$

where D_1 and D_2 are the deformation labels for each of the modes shown.

3 Experimental Results

In order to show the results obtained with this 2D shape recognition technique, we have used a database containing 161 2D simple shape images, which include: squares, rectangles, triangles, ellipses, pentagons, hexagons, characters, animals, airplanes, cars and silhouettes.

We could provide a list, *a priori*, of 20 different shape families, however, since the objects within those families have been deformed by using transformations such as translation, rotation, scaling, slicing or contour deformation, the use we have made of the family concept in this paper, has been to chose a representative of each of these in order to show the first 18 images retrieved in each experiment.

Two different types of experiments were carried out with the technique proposed in Section 2.1:*Use of an isotropic structural element* and *Use of non-isotropic structural elements.*

For the technique proposed in Section 2.2, experiments carried out have modified two elements:*The number of modes considered in the modal deformation space* and *the number of subsampled points considered in the outer contours of those query shapes.*

With this experimental selection of parameters, in the morphological filter phase, in those experiments carried out, we have used a rectangular structural element, positioned on its center of gravity with a maximum size of 2 pixels wide by 32 pixels high, considering 16 values of the morphological curves and 8 rotation angles for the structural element (in this way we cover 180°). This step rejects a 75% of the shapes. In the second stage, we have fixed a maximum prototype size equivalent to a circular shape with diameter of 128 pixels. Regarding the number of points of the contour shown, 128 points have been sub-sampled. Both options have been carried out.

Figure 1, show the results obtained in the different experiments in some of the families considered.

In those figures we can see how the techniques proposed in Sections 2.1 and 2.2 could be used in a independent way as 2-D shape characterization mechanisms, but such a characterization could to lead the query into errors in the 2-D shape recovery process as it is showed in step 1:Morphological Filter in Figure 1.

4 Conclusions

This paper has dealt with the question of indexing shapes contained in a database beginning with the concept of similarity. For this purpose, several discriminating characteristics have

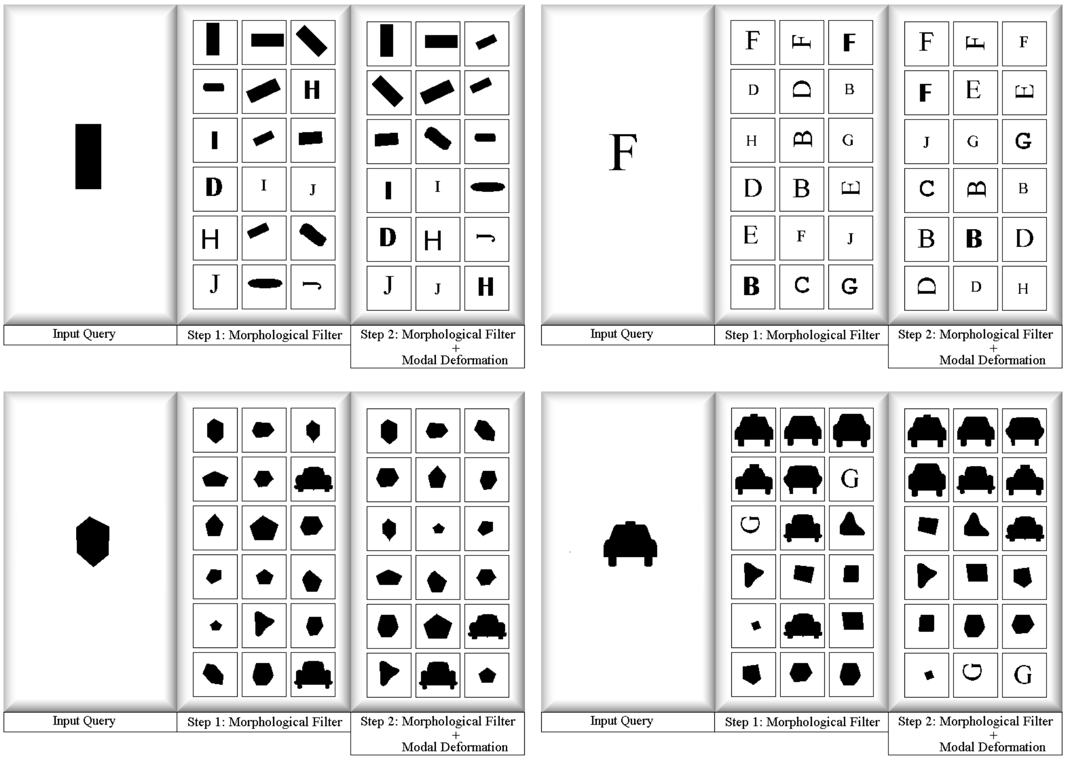

Figure 1: Indexing of the first 18 shapes (by rows) retrieved by using the proposed recognition technique.

been described within the different shapes which allow us to calculate their similarity, from a shape recognition process.

In some cases, the shapes are grouped in families according to a simple perceptual concept. To what extent can't a rectangle with rounded corners be considered an elliptic shape?

The proposed technique makes use of the morphological scheme in order to extract a subset of those most similar images from the database. The energy measure of the deformable model is then applied to index the resulting subset.

Experimental results have been obtained for queries made for several shapes varying the used structural element and the number of modes and subsampled points. Experiments indicate how indexing generated by the proposed technique gives very good results for the input prototype shape.

References

[1] F. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.

[2] G. Cortelazzo, G. A. Mian, G. Vezzi, and P. Zamperoni. Trademark shapes description by string-matching techniques. *Pattern Recognition*, 8(27):1005–1018, 1994.

[3] L. S. Davis. Shape matching using relaxation techniques. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1:60–72, 1979.

[4] M. P. Dubuisson, S. Lakshmanan, and A. K. Jain. Vehicle segmentation and classification using deformable templates. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(3):293–308, 1996.

[5] U. Grenander. Pattern synthesis. In *Lectures in Pattern Theory vol. 1*, volume 18 of *Applied Mathematical Sciences*. Springer-Verlag, 1976.

[6] P. W. Huang and Y. R. Jean. Using 2d c^+-strings as spatial knowledge representation for image database systems. *Pattern Recognition*, 9(27):1249–1257, 1994.

[7] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge. Comparing images using the hausdorff distance. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, (15):850–863, 1993.

[8] A. K. Jain and S. Lakshmanan. Object matching using deformable templates. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(3):267–278, 1996.

[9] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: active contour models. *Int. Journal of Computer Vision*, 1(4):321–331, 1987.

[10] A. Khotanzad and Y. H. Hong. Invariant image recognition by zernike moments. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, (12):489–497, 1990.

[11] G. Matheron. *Random sets and integral geometry.* Wiley, New York, 1975.

[12] C. Nastar. *Modèles physiques déformables et modes vibratoires pour l'analyse du mouvement non-rigide dans les images multidimensionelles.* PhD thesis, INRIA, 1994.

[13] C. Nastar and N. Ayache. Frequency-based nonrigid motion analysis: application to four dimensional medical images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(11):1067–1079, November 1996.

[14] W. Niblack, W. Barber, M. Equitz, M. Flickner, E. Glasman, D. Petkovic, and P. Yanker. The qbic project: querying images by content using color, texture and shape. In *Proc. SPIE Conf. on Storage and Retrieval of Image and Video Databses*, 1993.

[15] A. Pentland and S. Sclaroff. Closed-form solutions for physically-based shape modeling and recognition. *IEEE Trans. on Patterns Analysis and Machine Intelligence*, 13(7):715–729, July 1991.

[16] R. Schettini. A segmentation algorithm for color images. *Pattern Recognition Letters*, 14:499–506, 1993.

[17] J. Serra. *Image analysis and mathematical morphology, vol 2. Theoretical advance.* Academic Press, 1988.

[18] C. Teh and R. T. Chin. On image analysis by method of moments. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 10:496–513, 1988.

[19] D. Terzopoulos, A. Witkin, and M. Kass. Constraints on deformable models: Recovering 3d shape and nonrigid motion. *Artificial Intelligence*, 36:91–123, 1988.