SELECTION OF CHARACTERISTICS FOCUSED ON STORAGE AND RETRIEVAL OF DATABASE IMAGES

Piedad Soriano Montiel, Esther de Ves Cuenca

Department of Computer Science, Universidad de Valencia

1. INTRODUCTION

In Spain, and particularly in the Valencian Community, there is a great tradition in the tile sector that goes back several centuries. It is not surprising therefore, that at present a great number of images can be found of the tiles designed up till now. The problem of having this enormous quantity is mainly of being able to find among all of these, the one that most approaches what a certain client is looking for. Another problem is how to store all the information on each tile in a suitable way. In this work it has been attempted to address this problem, using the technology and search techniques available at the moment, since one of the fields in artificial vision that has undergone profound development in the last decade has been the study of techniques for accessing large databases of graphic documents and images through key images. The new advances in telecommunications, calculation and multimedia technologies today allow accessing large quantities of digital/visual information, in the so-called Image Databases or Image Libraries.

The developed application consists of an image database storage and retrieval system for each tile image, and with these tile images, of a group of characteristics that will be discussed below, which describe each image.

Thus, the client that seeks a certain type of tile, but does not have concrete data to identify this particular tile can search the whole range of tiles, without having to do so manually, through all the catalogues that his company has. This provides great savings in time, as well as a greater satisfaction for the client, since it saves him considerable work and time.

Colour, shape, texture and movement are the basic characteristics that allow us to perceive, describe and recognise objects, and hence interact with our surroundings. In particular, the perception of textures plays a very important role in the human visual system. It is used to detect and distinguish objects, infer information about the orientation of surfaces and their perspective, and to determine the information on the shape of objects in three-dimensional scenes. Today textures, due to their aesthetic properties, play a very important part in design for consumers and in the marketing, sales and exchange of products. As neither movement nor shape are typical characteristics of tiles, we have focussed on analysing the colour and texture that make up these images.

Therefore, the attributes stored jointly with the images in the database correspond to colour and texture descriptors, extracted from the images contained in the database. For this, we will need to carefully select these descriptors, which will represent a certain image.

The retrieval, i.e., the search for images in a tile database, will be carried out starting with inquiries on the colour and the texture that make up this image. It will also be necessary to establish an indication of the existing similarity between the colour and texture of the inquiry image and the rest of the images in the database in a consultation. This will be termed the image fitting module, which enables selecting the most similar images from the database according to this measure of similarity.

A further characteristic is that the implemented application will work in the Net, i.e. that the database will be stored in a central computer, and the consultation made by a given client will take place from a remote computer. For this we have prepared a user interface in HTML, that is, it will be a Web site. This makes it a very simple application to use, while it is only necessary to store the database in a single computer, and not in each computer where the consultation is made, allowing low cost connection of as many computers as desired.

This work has been organised in the following way: section 2 describes the general structure of the system, section 3 continues with an explanation of each method used for obtaining the colour and texture characteristic, followed by a discussion of the measurement of similarity among the descriptors used and finally, the experiments are described which were carried out to evaluate the system.

2. SYSTEM STRUCTURE

Having described the problem, a schematic illustration will summarise the system's overall structure.

It shows the development of a complete scheme of information storage and retrieval in an image database, where attributes as important as colour or texture are the elements to be managed in the interaction process with the database:

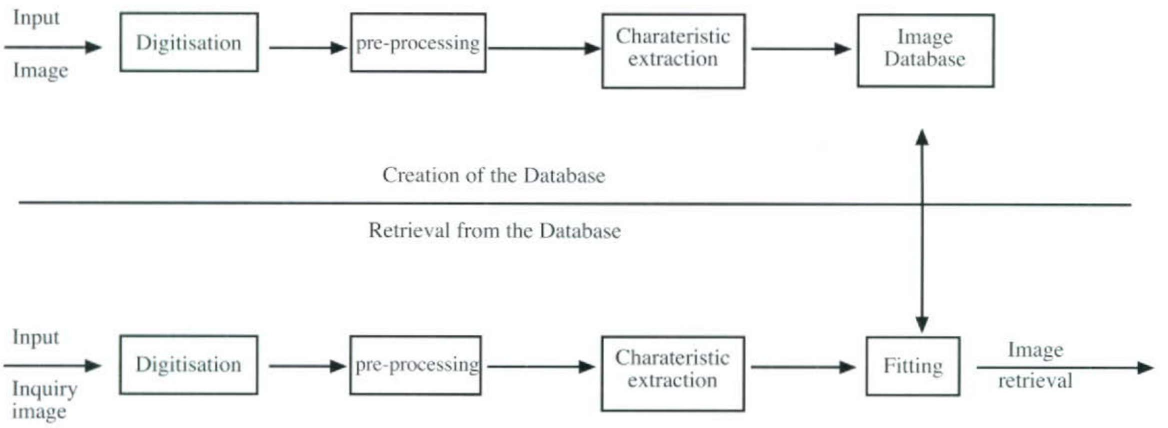

Fig. 2.1 Structure of the system.

As Figure 2.1 shows, up to the characteristic extraction module the scheme is identical for the creation (or insertion) part of the database, and for the part where the user makes the inquiry. Let us look at each of the modules in detail:

1. Digitisation: in this part, a digital photograph would be made of the tile, which would be entered into the system.
2. Pre-processing: after digitising, generic image pre-processing can be carried out depending on the type of image involved. In our particular case, the applied pre-processing consists of transformation of the colour space.
3. Characteristic extraction: here it will be attempted to establish the parameters that characterise the image, based on the pre-processed image.

4.1 Image database: finally, in the database creation part, these parameters are inserted in the database together with the image.

4.2 Fitting: in this module, when two images are compared, one corresponding to the inquiry image and the other corresponding to that in the database, it will be necessary to find a way of calculating how different or similar those two images are, and this we have called the fitting module.

5. Output: output only makes sense in the part in which an inquiry is made to the database, therefore as output we will find those images that meet certain conditions, result images.

In the selection process of the images that meet the conditions expressed in the inquiry made to the database, it is necessary to use some type of similarity operator (corresponding to the fitting module). This similarity will enable comparing the extracted values of the original image and those given by the image in the inquiry.

3. COLOUR AND TEXTURE DESCRIPTORS

The attributes that are stored in the database together with the image correspond to colour and texture characteristics. Good selection of these characteristics is very important, because appropriate operation of the whole system depend on this.

3.1 COLOUR DESCRIPTORS

Due to the way a human being sees the world that surrounds him, it is necessary to pre-process the image before beginning to work with it. This is because images are saved pixel by pixel, each one with a given **RGB** colour. Human beings do not see things in this way; we see the surrounding world in terms of saturation, light and contrast.

To adapt this as closely as possible to the way in which the human eye perceives colours, perceptual colour spaces have been developed in the past, one of which, the one that we have chosen, is the system of HSI colour co-ordinates, which transforms each RGB pixel of the image into its corresponding HSI (Hue, Saturation, Intensity).

After transforming the image, we can now analyse it to obtain its descriptors. The descriptors are set out below:

- **HISTOGRAM METHOD**

In image processing, the histogram[1] of an image is a bar diagram that represents the number of pixels that have a certain level of grey. In our case it will refer to the number of pixels that have a certain value of red, a certain value of blue, and a certain value of green, and also for the H and S of the image, which will be measured separately, i.e., 5 histograms will be made, one for each colour component.

From each histogram, we would keep the maximum value, i.e., the maximum for red, the maximum for green and the maximum for blue. But as it is suggested in[1], when distance is calculated for colour, this is done by looking at all the colour values, i.e., all the "bars" of the histograms of the images to be compared are compared. However, this would be too exhaustive for this application, particularly because a tile is not expected to exhibit a great variety of colours. At most, it would be expected to have three or four predominant colours. Therefore, we will keep the three predominant values of each histogram.

Otherwise, if we just stored a maximum, it would be insufficient. Consider an image whose background is white, and the texture red for example. The background could predominate, so that we would keep the image as white, while for us, as human beings the tile would be represented by the red and not the white colour.

- **3D HISTOGRAM METHOD**

Here we make a combined histogram of the three colour bands, i.e., we obtain a histogram in three dimensions. Because the range of colours oscillates between 0 and 255, a three-dimensional matrix measuring 256*256*256 would be found, which makes it excessively large and unmanageable. Therefore the range of 256 colour levels is divided by a certain number of colour levels, which will be chosen when the tests on the system are carried out, and which we will initially set at 64, with the possibility of changing this if necessary. This enables reducing the size of the resulting matrix considerably.

After constructing the histogram, we keep the maximum values, just as we did in the previous method, i.e., the three predominant colours in the image.

• HISTOGRAM METHOD FOR H AND S

The same is done as in the 3D histogram, but instead of using the R, G and B bands, the H and S bands are used.

3.2 TEXTURE DESCRIPTORS

Multiple definitions exist in the literature for the term texture. In this paper we have adopted the definition given by Harlick and Shapiro[4]:

"A texture is defined by the uniformity, density, thickness, roughness, regularity, intensity and directionality of discreet measurements of the tone and of its spatial relationships."

In this definition it is attempted to define texture based on certain characteristics. These characteristics will be the ones used for storage in the database. We shall now go on to describe them.

• GRADIENT METHOD

Starting with an image, the images, module and phase of the image gradient are calculated[6]. The average and variance of the values stored in these images will be used as texture descriptors. The attribute obtained by this method is an indicator of the uniformity of the image. This method is only applied to intensity band I.

• METHOD OF LOCAL EXTREMES

This method attempts to give a measure of the thickness and density of the texture. For this, the number of *local extremes per unit area* of the image is studied, which represents a texture. The algorithm that allows calculating the number of local extremes works in the following way: a given image is scanned by rows (and columns), marking all those pixels that are local maximums or minimums.

After finding the local maximums and minimums, the characteristic 'texture thickness' is then calculated, which is defined as:

A pixel of the image is calculated as the density of local extremes of the row (or column) of the image function in a window of size nxn centred on this pixel.

• FOURIER TRANSFORM METHOD

The Fourier transform[3] is one of the most widely used methods in image processing. The spectrum of the image is calculated, divided in rings and sectors. The area that each ring and each sector occupies is found. This method will be applied to the I band of the image. This represents the energy of the image in each of the rings or sectors.

The quantity of energy per ring gives a measure of the thickness or fineness of the texture, while the energy of each sector gives a measure of the directionality of the texture.

• **STATISTICAL METHODS (CO-OCCURRENCES)**

Co-occurrences[3], in a general way, can be specified in a matrix of relative frequencies $P(i,j;d,\theta)$ with which two neighbouring elements of texture, separated by a distance d and an orientation θ, occur in the image, one with property i and the other with property j.

A matrix P_d is constructed of size (GXG), where G is the number of levels of grey in the image, of co-occurrences of levels of grey for a vector of displacement $d=(d_x,d_y)$. The element (i, j) of P_d is the number of occurrences of the pair of levels of grey i and j that are at a distance d. Formally:

$$P_d(i,j) = |\{((r,s),(t,v)): I(r,s)=i,\ I(t,v)=j\}|$$

where$\{(r,s),(t,v)\} \in Nx$, $(t,v)=(r+d_x, s+d_v)$ y and the module signifies cardinality. Based on the calculated co-occurrence matrix, the characteristics of energy, entropy, contrast and homogeneity can be found.

• **MORPHOLOGICAL CHARACTERISTICS:**

A powerful method for characterising textures consists of carrying out several types of reducing and expanding operations of regions and analysing the results. The operation that we have used is expansion (basic operation in mathematical morphology[5], which yields another image as a result). A series of openings are performed on the original image that represents the texture. To carry out this operation, one needs what is known in mathematical morphology as a structuring element. In this work a circle of radius $\mathbf{h}_{min,}$ which will increase to $\mathbf{h}_{max}$ in each opening made, has been selected as structuring element. Once the opening is applied to the image, the area of the result image is calculated divided by the area of the original image.

$$v(h) = \frac{A(I_0,0)}{A(I)}$$

This value v(h) is calculated for different radii of the structuring element, yielding a probability distribution that can be used as a descriptor of the image.

• **PATTERN SPECTRUM:**

In this method the differences are established of the previous characteristics. This gives us the probability density:$PE(h) = v(h) - v(-1)$

The characteristics found with all the methods described in this section will be the attributes that are stored in the Database together with each.

3.3 SIMILARITY MEASUREMENTS

It is necessary to have some method that allows comparison of two images of the database or of an image and a model when making a consultation. In the literature multiple measures[2] have been advanced to quantify how similar the characteristics of two specific images are. Out of these we have selected two types of distance measurements: the Euclidean and the Mahalanobis distance.

The Euclidean distance, which is very well known, is calculated from:

$$D^2 = \sum_{\forall i} \left(C_{i,1} - C_{i,2}\right)^2$$

where $C_{i,1}$ is characteristic i of image one, and $C_{i,2}$ is characteristic i of image two.

This gives us a measure of the distance between the inquiry image and the database images. Therefore after calculating the distances between the inquiry image and the database images, we will keep the database image(s) whose distance is the smallest. If we set a threshold, we can obtain several result images.

The Mahalanobis distance is calculated from:

$$D^2 = V * \sum^{-1} * V^T$$

where$\sum^{-1}$ is the inverse of the matrix of co-variances of the values of various descriptors, and V is the vector formed by the difference between the values of the inquiry image characteristics and the values of the characteristics of each image stored in the database. With this it is attempted to weight each characteristic, so that some do not have a greater influence than others in the result, owing to the order of their numeric value.

4. TESTS AND RESULTS

To verify which results are acceptable, a series of tests was carried out for each case:

- Consultation only by colour
- Consultation only by texture

4.1 CONSULTATION ONLY BY COLOUR

To conduct the tests on colour, the characteristics corresponding to the 3D and HSI histograms were used. The experiment was performed on a group of 20 individuals, whose is evaluation was requested of the result of a consultation on a specified colour.

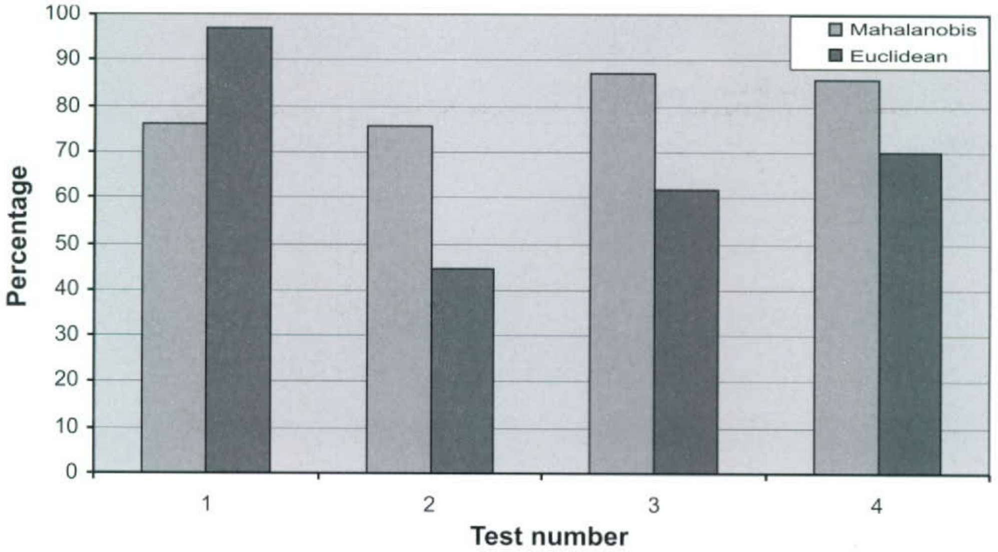

Fig. 4.1.1: Results of the evaluations on four different consultations on colour.

Figure 4.1.1 shows the results found on making the consultation on four different colours. In the tests the users were shown the colours on which the consultation was carried out and the result of these consultations, and the user had to select how many images had the inquiry colour as predominant colour.

It is necessary to keep in mind that not all people see colours in the same way, so that the results are very subjective. The diagram shows us, therefore, the percentage of satisfaction in the results for each colour used in the tests, for both the Mahalanobis and the Euclidean distance.

Next we present an example of a colour consultation:

Fig. 4.1.2: Consultation colour.

And the images obtained as a result of this consultation are:

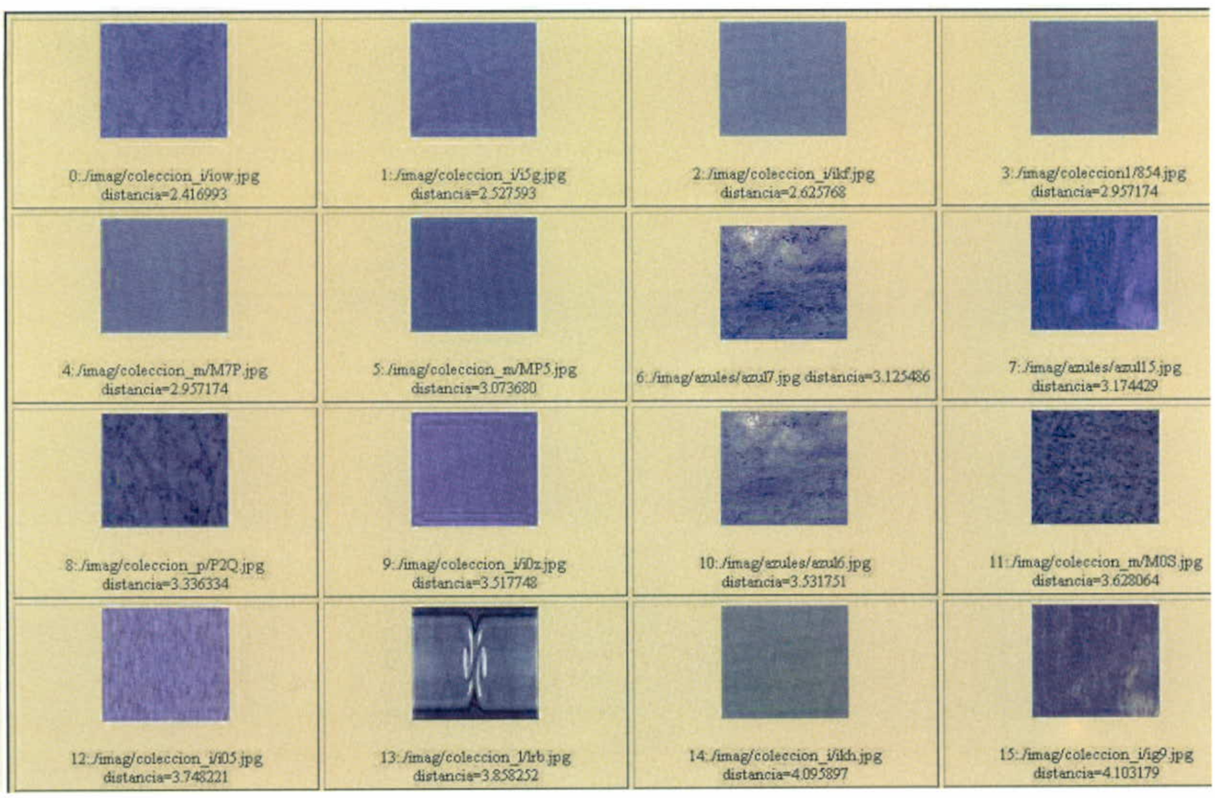

Fig. 4.1.3: Result of the colour consultation.

As can be observed, all the images have the colour specified in the inquiry image.

4.2 CONSULTATION ONLY BY TEXTURE

In this case, because we have implemented several descriptors, these will be analysed independently, and we will thus be able to verify which work best. The tests were conducted by "non-expert" users, i.e., people foreign to the work. The experiments consisted of evaluating the result images obtained on carrying out a consultation only using the texture descriptors. Each individual had to select the ones that he found significantly similar to the inquiry image from the result images.

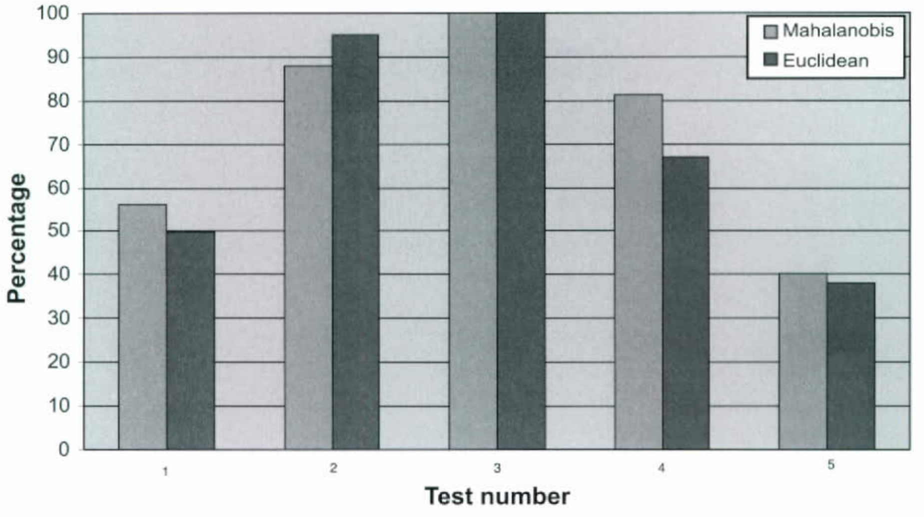

Fig. 4.2.1: Results of the evaluations on the Fourier descriptor for five different consultations.

Figure 4.2.1 shows the results obtained making the consultation on five different textures only using the Fourier descriptor in the consultation. As we can see, in general, Mahalanobis works better for Fourier, although the difference is not very great between both distances.

These results do not allow deciding with certainty which distance works better, since the results found are generally good and quite similar.

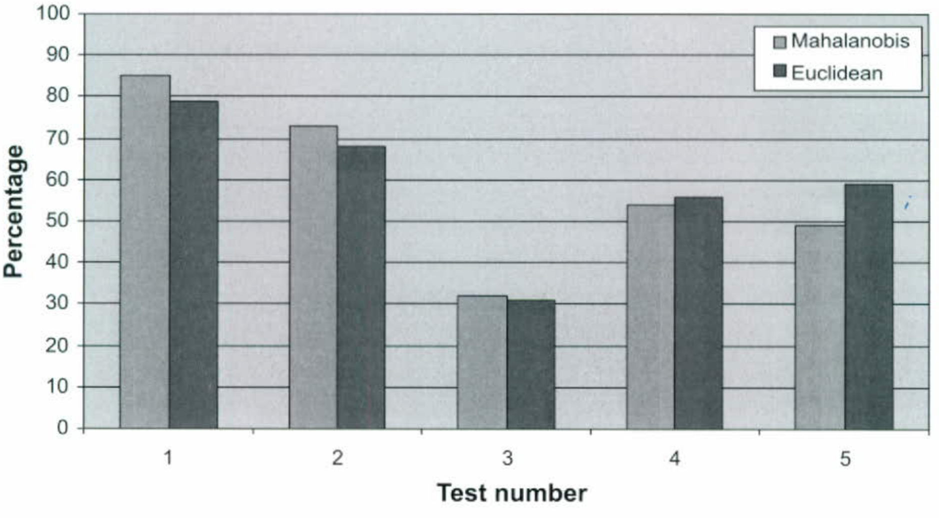

Fig. 4.2.2: Results of the evaluations on the Gradient descriptor for five different consultations.

Figure 4.2.2 shows the results obtained on carrying out the consultation on five different textures only using the Gradient descriptor in the consultation. Unlike the previous case, we cannot establish which of the two distances works better, while we cannot we discard this characteristic categorically either.

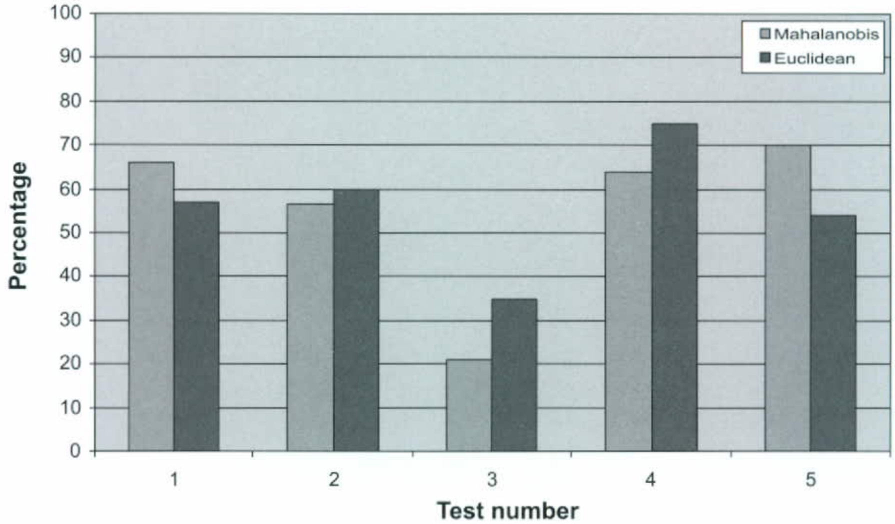

Fig. 4.2.3: Results of the evaluations on the Thickness descriptor for five different consultations.

Figure 4.2.3 shows the results obtained on conducting the consultation on five different textures only using the Thickness descriptor in the consultation. Here we can observe that the results are generally good, so that this characteristic cannot be discarded.

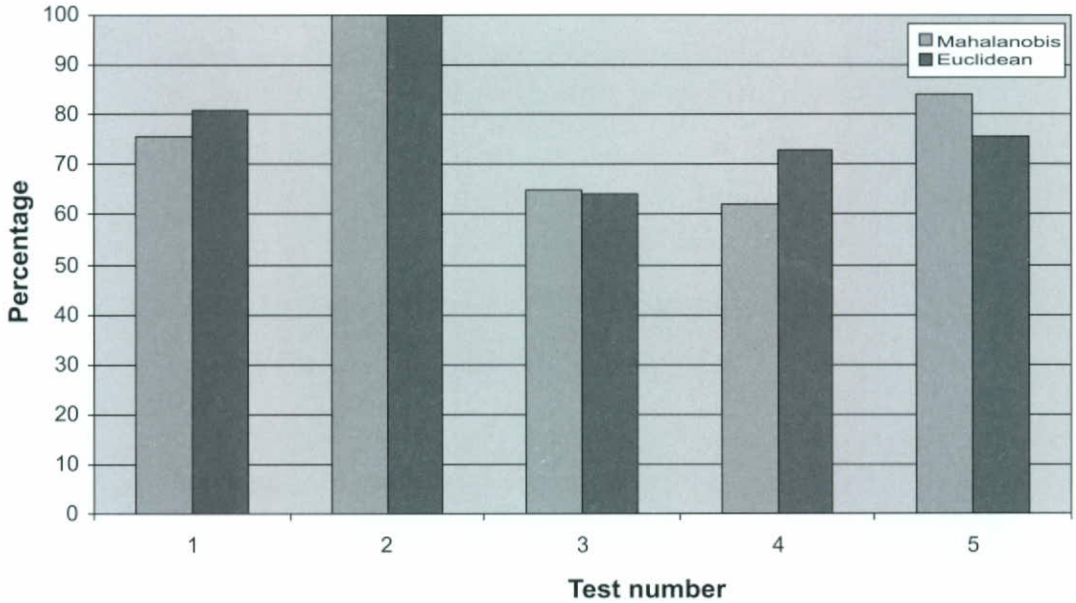

Fig. 4.2.4: Results of the evaluations on the Grain size descriptor for five different consultations.

Figure 4.2.4 shows the results obtained on making the consultation on five different textures only using the Grain size descriptor in the consultation. With these results we cannot assure which type of distance works better, although it can be assured that this texture characteristic will not be discarded, since the results are very good.

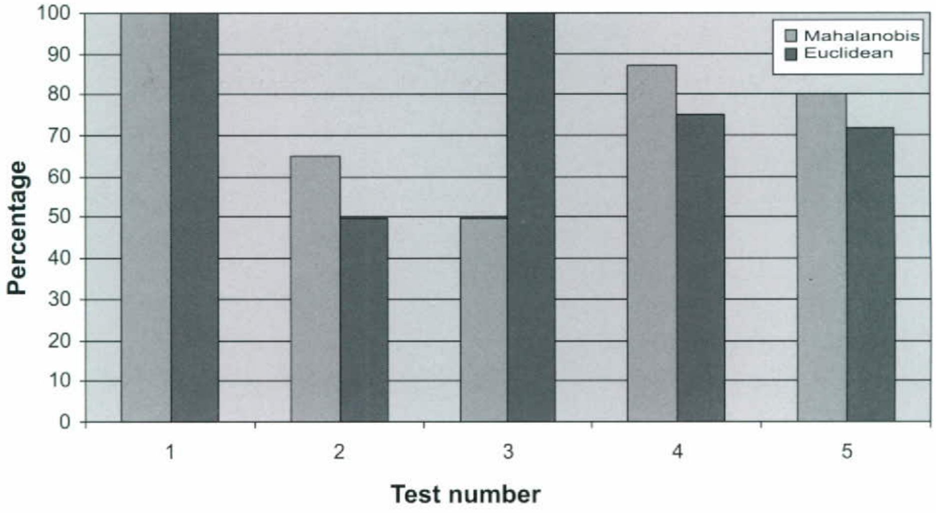

Fig. 4.2.5: Results of the evaluations on the Co-occurrences descriptor for five different consultations.

Figure 4.2.5 shows the results obtained on performing the consultation on five different textures only using the Co-occurrences descriptor in the consultation. As in the previous cases the elimination of this texture characteristic cannot be assured, nor can it be assured which distance works better.

Now we shall do the consultation for all the texture characteristics to see the results obtained:

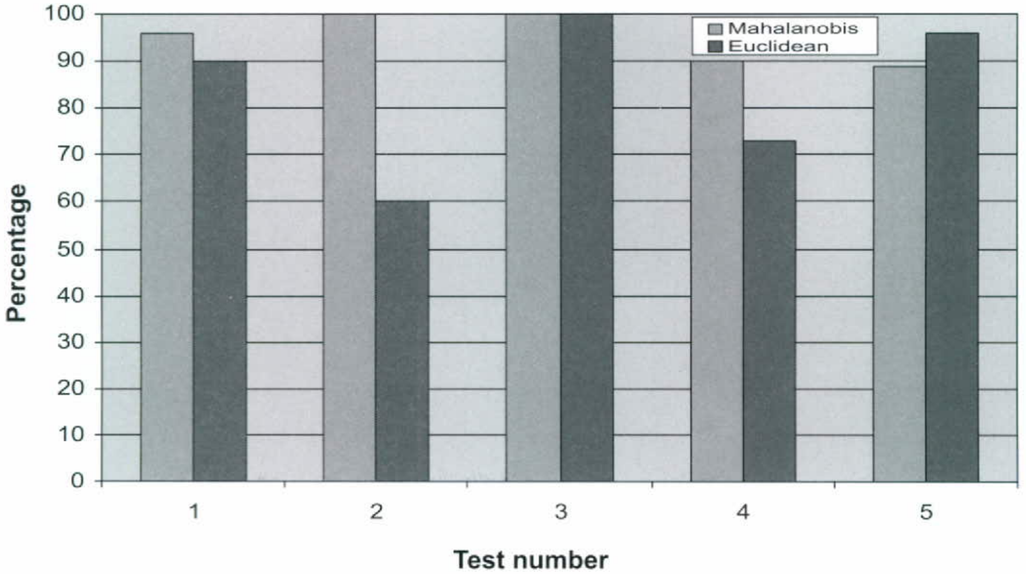

Fig. 4.2.6: Results of the evaluations on all the descriptors for five different consultations.

Figure 4.2.6 shows the results obtained on carrying out the consultation on five different textures using all the descriptors in the consultation. As we expected, the results for the consultation with all the characteristics work much better, in every case exceeding 90% of the images found, as well as establishing that the results are better with the Mahalanobis distance.

Next an example of the result is shown that would be obtained if we looked for the image that appears in Figure 4.2.7:

Fig. 4.2.7 Inquiry texture.

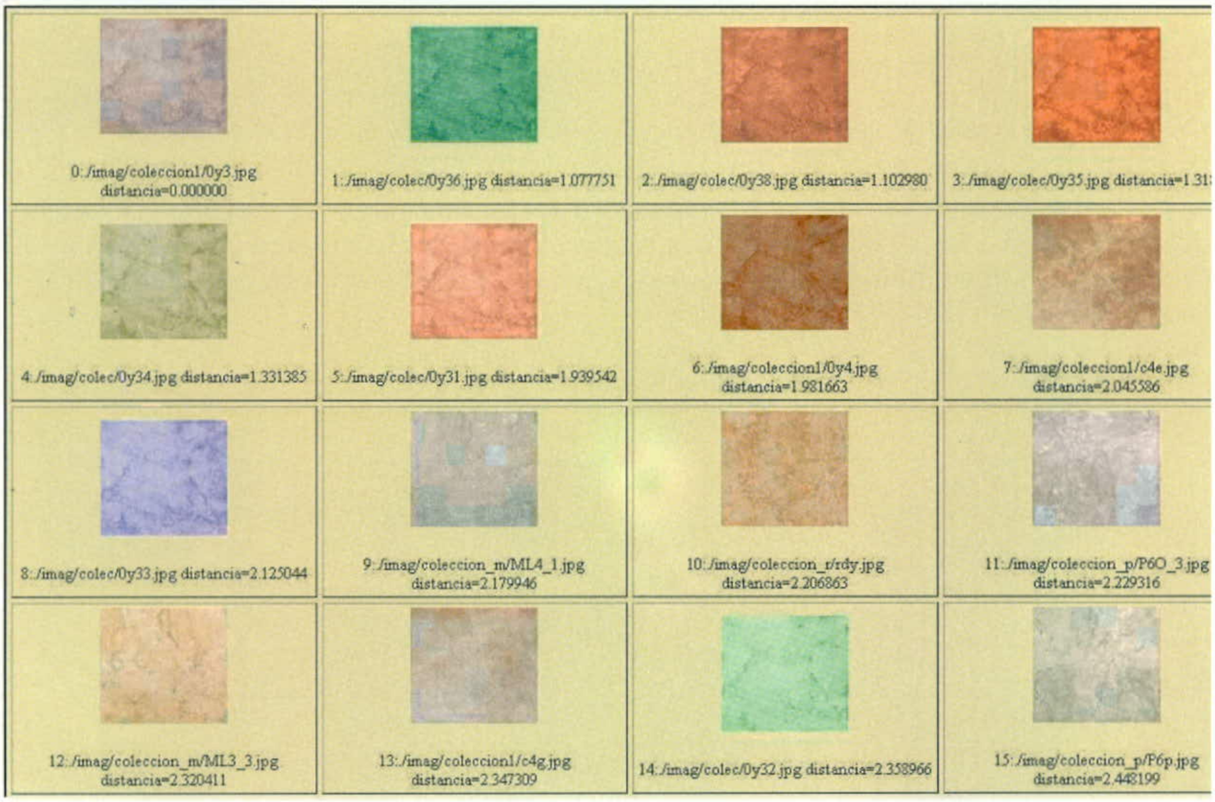

Fig. 4.2.8 Result of the consultation.

4.3 CONCLUSIONS

The experimental results obtained do not allow us to discard any of the characteristics described in this work since all contribute in some way to good operation of the system. There are characteristics that perform better for certain textures than others. It is therefore advisable to use all the descriptors in the same consultation, thus achieving 90% satisfactory results.

It can also be observed that, as expected, the results of the consultation for the Mahalanobis distance are considerably better than for the Euclidean, since the Mahalanobis distance, as remarked at the beginning, equally takes into account all the characteristics, while the Euclidean does not. Other distances could also have been tried, since the literature contains different methods.

With regard to colour, as can be observed, tests have not been carried out for the RGB Histogram characteristic, since poor results are expected for this, which do in fact occur. This is because of how the histogram is made, since it collects the maximum of each colour component independently, without taking into account the other components.

REFERENCES

[1] "Automated image retrieval using color and texture", Columbia University technical report TR# 414-95-20, July, 1995.

[2] "Similarity Measures", Simone Santini. And Ramesh Jain, IEEE transactions on pattern analysis and machine intelligence, Vol.21, nº 9, September 1999.

[3] "Texture Analysis" Handbook of Pattern Recognition and computer vision Eds. C.H. Chen, L.F. Pau and P.S.P. Wang.

[4] Textural Features for Content-Based Image Database Retrieval http://isl.wtc.washington.edu/~aksoy/thesis.shtml

[5] Procesado morfológico http://wgpi.tsc.uvigo.es/libro/procesim/node9.htm

[6] "Evolving descriptors for texture segmentation", Christophe Jacquelin, André Aurengo, Gilles Hejblum. Pattern Recognition, Vol. 30, Nº 7. pp 1069-1079.