Proc. 27 th Conf. of the Union of Bulgarian Math., Pleven, Bulgaria 1998 (63-71)

GENERAL IMAGE RETRIEVAL MODEL

Peter L. Stanchev
Institute of Mathematics and Computer Science, Bulgarian Academy of Sciences
Acad. G. Bonchev St. 8, 1113 Sofia, Bulgaria, phone: 359-2-713 3814, fax: 359-2-971 3649
e-mail: stanchev@bgearn.acad.bg

In this paper we analyse the existing approaches to image retrieval and we propose a General Image Retrieval (GIR) model. The model establishes a taxonomy based on the systematisation on the existing approaches. In it the queries for image retrieval can be specified visually and interactively. The query mechanism is based on the calculation of similarities between the image query representation and the images in the image database. The GIR model provides numerous retrieval mechanisms such as: retrieval by attribute values, shape, colour, texture, visual examples, spatial similarity, and browsing and is general enough to serve as an engine to various image applications. It is designed to be used as a part of an Image Database System. An example for applying the GIR model to a plant image database is given.

Keywords: Image, Image Database, Image Retrieval, Content-based Image Retrieval, Image Retrieval Model

1. INTRODUCTION

With the increasing the role of the image databases the problem of image retrieval becomes very essential. The leading commercial image database systems are the SQL Multimedia offered by Digital Equipment Corporation, Query By Image Content (QBIC) research project designed by IBM Corporation and Kodak Photo CD System introduced by Kodak. The most significant achievement in them is the efficiency of image retrieval.

There are several approaches for image retrieval. The first approach is text-based. The image is described as a set of key words or free text. The queries are based on exact or probabilistic match of query text. The alternative way allows the retrieval query to be based on the visual content of an image such as: image patterns, colours, textures, shapes of image object and location information. The content-based image retrieval can be characterised by the ability to retrieve relevant images to the user defined image query, based on the semantic content of the images. The queries that use the content-based indexing are based on the similarities of image descriptions (all obtained images as a result of the query have a description within some preliminary defined distance ε from the query defined image).

In this paper we propose a General Image Retrieval (GIR) model. The GIR model is based on the proposed General Image Data model [1] and General Image Database model [2]. The GIR model incorporates flexible tools for queries and browsing using traditional database tools and visual information. The query contains paintings, sketches, example images, icons, sample textures or colours. The proposed GIR model language is interactive**.** The user screen is divided into three sub-widows for: preparing the image query, giving query parameters and showing the obtained result images from the image database. The GIR model allows retrieval by attribute values, colour similarity, shape similarity, texture similarity, pictorial example similarity, and spatial constrains and browsing. The retrieval by shape similarity provides the user with tools for drawing image-object with similar shape. The retrieval by spatial constrain provides the user with a querying mechanism that is based on image-object icons, sketches and pictures, composed to represent spatial relationships among the image-objects. The retrieval by colour, or texture similarity allows the user to specify a given colour or texture, picking the colour from sample image or the colour palette or drawing texture or choosing some from sample texture library. When retrieval is based on a pictorial example similarities the user finds or composes an example image.

An example for applying the GIR model for retrieval from a plant image database is given.

2. IMAGE RETRRIEVAL MODELS

Some of the valuable proposals for image retrieval models are: QBIC System [3], AIR framework [4], PICQUERY [5], AMSTERDAM system [6].

The ***QBIC*** (query by image content) system was developed to explore the content-based retrieval methods. It allows queries based on colour and texture patterns or sketches and drawings of an object shape. The indexing is dome automatically. R trees are used for storing the multidimensional index.

The ***AIR*** framework suggested five classes of retrieval: by spatial constrains, shape similarity, semantic attributes, object attributes and browsing. The queering is done by window-based graphical interface.

The ***PICQUERY*** language has been designed with a flavour similar to the QBE as highly nonprocedural and conversational language for image database management. It includes the following sections: image manipulation operations such as: rotation, zooming, colour transformation, etc.; pattern recognition operations such as edge detection, thresholding, contour drawing, similarity retrieval, texture measure, clustering, segmentation, etc.; spatial or geometric operations such as: distance operations, length, area, intersection, union, etc.; function operation such as: statistical functions, user defined functions and input/output operations.

The image query language in the ***AMSTERDAM*** system is a combination of retrieval by visual example image and textual description of the image content. The resemblance degree between example image and the images in the database is defined as the recognition degree in fuzzy Hit-or-Miss transformation from the mathematical morphology. The textual retrieval is given in a form similar to the conventional database manipulation language.

3. THE GENERAL IMAGE RETRIEVAL MODEL

The GIR model is unique in the sense of its comprehensive coverage of the image features. The main characteristics of the proposed GIR models could be summarised as follow:

a.) the images are searched by their general image description model representation [1];

b.) the model is based on similarity retrieval;

c.) the model language is a sophisticated window-based graphical interface;

d.) the user interface supports the visual expression of a query and allows query refinement and manipulation of the results;

e.) the model supports retrieval by attribute values, colour, shape, texture, pictorial example similarity and spatial constrains;

f.) the query for each retrieval class is based on schemes that are both natural and efficient for specifying query in the corresponding class;

g.) the model is used as the image manipulation language in the General Image Database model [2];

h.) the model has a reasonable generality to deal with a wide variety of image applications.

The proposed GIR model is described schematically. In every particular realisation the proposed tools can by implemented differently. The main window contains 3 sub-windows for formulation of the query, for retrieval parameters, and for the query results. A schematic view of the retrieval interface is given in Figure 1.

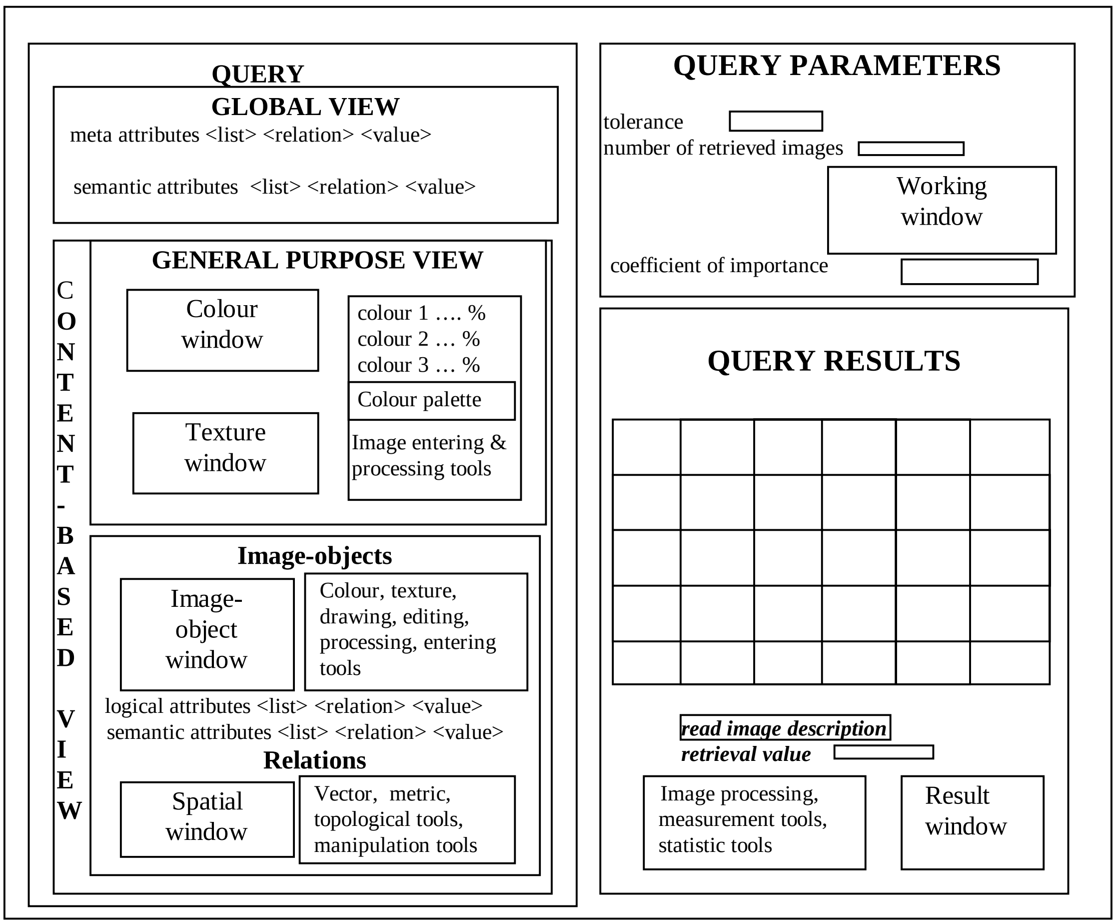


Figure 1. Schematic view of the interface for the GIR model

The image query contains description on the desired image or a part of it. During the retrieval the tolerance between the described image the result subset is given. In this place the user can also give the coefficient of importance of the obtained images in order to change the importance of the different parameters in the image description. Image processing tools, measurement tools, and statistical tools are also available. The available retrieval methods are:

a.) retrieval by attribute values

The retrieval by attribute values is similar to the retrieval in a conventional database systems. The query is formulated using meta, logical and semantic attributes. The user employs a list of a predefined attributes and a list of appropriate relations.

b.) retrieval by colour similarity

For retrieving an image or an image-object hued similarly to a given pattern. The following methods are used :

- pick the template from predefined images (such as a sunny sky, a clear sea, exc.) and query for the images having similar colour distribution;
- make an own template by using parts of stored images and/or use drawing tools;
- use a colour picker to specify from a colour palette the percentage of the desired colours in the result subset (in this way a query can be formed as follows: "Show me all images with colour distribution: 40% yellow and 60% blue").

c.) retrieval by shape similarity

For specifying an image containing image objects with the desired shape the user can:

- draw an approximation of the image-object shape;
- copy the shape for the query from a "shape" gallery.

d.) retrieval by texture similarity

For specifying an image or an image-object with some desired texture the user can:

- use a library of predefined image texture templates (such libraries exist in the drawing and image processing packages such as Corel Draw, Photo Stiller, etc.);
- make their own template by composing it from predefined images.

e.) retrieval by pictorial example similarity

For specifying an example image the user can:

- specify an image from the image database;
- make their own template.

f.) retrieval by spatial constrains

Retrieval by spatial constraints facilitates a class of queries that are based on the retrieval of the 2-D arrangement of the objects in the image. The query is composed by placing icons (assuming there is a predefined icon for every type of an image-object in the application domain) or user defined shapes on a plane. As well there are tools for scaling and rotating the icons and the shapes. The result subset contains the images that contain objects arranged in a manner similar to the way shown in the query.

g.) browsing

In the browsing process thumbnail images (64 by 64 pixels) are used.

4. SIMILARITY CALCULATION

Let a query be converted through the general image data model in an image description $Q(q_1, q_2, \ldots, q_n)$ and an image in the image database has the description $I(x_1, x_2, \ldots, x_n)$. Then the retrieval value (RV) between Q and I is defined as:

$RV_Q(I) = \Sigma_{i=1,\ldots,n} (w_i * sim(q_i, x_i)),$

where w_i ($i = 1,2, \ldots, n$) is the weight specifying the importance of the i^{th} parameter in the image description and $sim(q_i, x_i)$ is the similarity between the i^{th} parameter of the query image and database image and is calculated in the following way:

a.) if q_i, x_i ***are symbol values*** then:

$sim(q_i, x_i) = 1$, if q_i is equal to x_i

$= 0$, if q_i is different from x_i.

b.) if q_i, x_i ***are numerical*** values then:

$sim(q_i, x_i) = 1 - |q_i - x_i| /$ *(maximum value of* x_i *in the image database).*

c.) if q_i, x_i ***are linguistic*** values, then:

first: q_i, x_i are converted to fuzzy numbers, represented as $q_i(a, b, \alpha, \beta)$ and $x_i(c, d, \gamma, \delta)$ chosen from a basic set of fuzzy numbers. (The fuzzy number $q_i(a, b, \alpha, \beta)$, where a, b is the closed interval in which the membership function is equal to 1, α is the left bound and β the right bound is shown in Figure 2);

second: $sim(q_i, x_i) = ((a+c)/2, (b+d)/2, (\alpha+\gamma)/2, (\beta+\delta)/2)$;

third: $sim(q_i, x_i)$ is approximate with the nearest fuzzy number from the basic set;

fourth: the obtained fuzzy number is converted back to linguistic value.

More details about this procedure are given in [7].

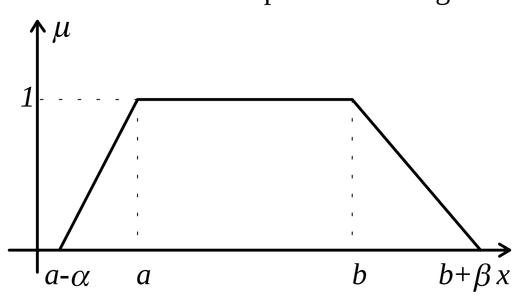


Figure 2. Characterisation of the fuzzy value $q_i = (a, b, \alpha, \beta)$

d.) if q_i, x_i ***are histograms***, then:

$sim(q_i, x_i) = (\Sigma_{l=1,\ldots,N} \min(q_{il}, x_{il})) / (\Sigma_{l=1,\ldots,N} q_{il}),$

where N is the number of the bins in the both histograms (we assume that the histograms have an equal number of bins) and q_{il} is the value of q_i in bin number l.

e.) if q_i, x_i ***are attribute relational graphs (ARG)***, then:

$sim(q_i, x_i)$ is the distance between the two graphs, calculated as the cost to find a sequence of transformation for producing one ARG from the other.

f.) if q_i, x_i ***are pictures*** and their pixels arrays are denoted as $Q(l, m)$ *and* $I(l,m)$, $1 \le l \le L$, $1 \le m \le M$, where L and M are the vertical and horizontal dimensions of the two pictures which coincide. Then

$sim(Q(l, m), I(l, m)) = \Sigma_{(l=1,\ldots,L;\, m=1,\ldots,M)} f(l, m) * I'(l, m),$

where $f(l, m)$ $(1 \le l \le L, 1 \le m \le M)$ is the normalised query image $Q(l, m)$. (The pixels' values in $f(l, m)$ are positive and their sum is equal to1). $f(l, m)$ is consider as a structuring element, in what is called a Hit-or-Miss operation in mathematics morphology and *I'* is the image *I* converted as a binary image, with values "+1" for the image objects and "-1" for the background pixels (for this purpose a threshold operation is used).

This procedure uses the fuzzy mathematical morphology and is described in [6].

g.) if q_i, x_i ***are interval values*** denoted with [*a*, *b*] and [*c*, *d*], then:
$sim(q_i, x_i) = (sim(a,c) + sim(b,d))/2;$

h.) if q_i, x_i ***are spatial representations,*** characterised with $\Theta\Re$ strings: $\Theta\Re_Q$ and $\Theta\Re_I$, then:
$sim(\Theta\Re_Q, \Theta\Re_I) = (Object\text{-}Factor + 2\,(Spatial\text{-}Factor) + Scale\text{-}Factor)/4,$
where the *Object-Factor* measures the similarity between the query image-object and the database image-object. Then to calculate the *Spatial-Factor* the similarity between the nearest left image-object of the query and the nearest left image-object in the database is calculated. The same is done with the right object and the mean of the two calculation yields the Spatial-Factor. The *Scale-Factor* is calculating by measuring the difference between the centres of the query image object and the database image object and normalising it in [0..1].

5. INDEXING THE IMAGE DESCRIPTION

The major problem when dealing with large image databases is the efficiency of the image retrieval. This efficiency depends on the used schema for indexing.

An image features could be represented as a feature vector that corresponds to a point in a multi-dimensional feature space. There are three popular approaches for multi-dimensional indexing: R-trees (particularly R^* trees), linear quadtrees, and grid files. Except them the following techniques for increasing the efficiency of the image indexing can be used:

a.) ***Technique for minimising the number of dimensions for the multi-dimensional vector which represents the logical view of the image***

Let the vector be $X(x_0, x_1, \ldots, x_{n-1})$. The Discrete Fourier Transform X_f of a X is defined as a sequence X_f of n complex numbers:

$$X_f = \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} x_f e^{\frac{-2\pi . f . t\sqrt{-1}}{n}}$$

$f = 0, 1, \ldots n-1.$

The Parseval theorem gives that:
$\Sigma_{t=0,\ldots,n-1}|x_t|^2 = \Sigma_{f=0,\ldots,n-1}|X_f|^2.$

Let a query is described with a sequence of real numbers q_i $(i = 0, \ldots, n-1)$ and an image from the database with a sequence of real numbers d_i $(i = 0, \ldots, n-1)$. Let's consider the two images being ε close if :

$(\Sigma_{t=0,\ldots,n-1}|q_i - d_i|^2)^{1/2} \leq \varepsilon^2$, but using the Parseval theorem
$\Sigma_{f=0,\ldots,n-1}|Q_f - D_f|^2 = (\Sigma_{t=0,\ldots,n-1}|q_i - d_i|^2) \leq \varepsilon^2$. Keeping only the first $fc < n$ coefficients, we have:

$\Sigma_{f=0,\ldots,fc}|Q_f - D_f|^2 \leq \Sigma_{f=0,\ldots,n-1}|Q_f - D_f|^2 = (\Sigma_{t=0,\ldots,n-1}|q_i - d_i|^2) \leq \varepsilon^2$.

In other words if we keep only *fc* elements of the complex vector the result subset will contain all the images from the database in ε vicinity of the query image, but as there will be spurious images also which are further then ε from the query image. In [8] an experiment was conducted. It shows that the optimal choice for *fc* is 3.

b.) Using self-organising map

In this method the indexing mechanism is based on the use of neural networks.

c.) *Using unique characteristics*

In the case that all possible image-objects that can be contained in the images in the database are predefined (this is the case for industrial object recognition) only the specific for the image-object attributes can be kept in the database [9].

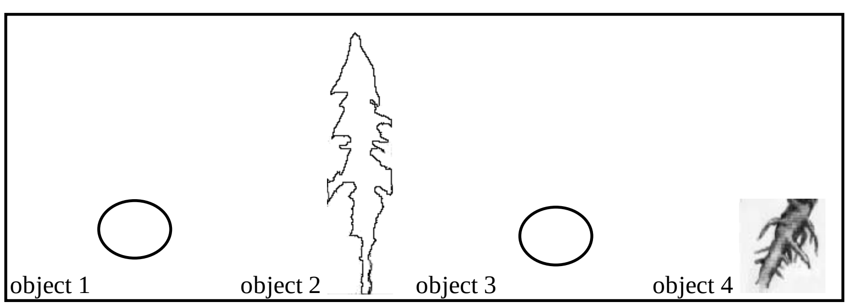
6. AN EXAMPLE FOR APPLYING THE GIR MODEL

Let's assume the user likes to retrieve at most three images, with the following general image data description:

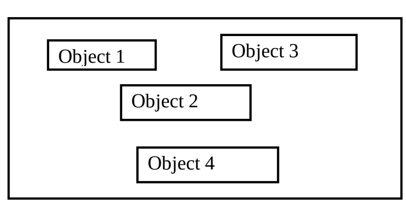
GLOBAL VIEW

ATTRIBUTE TYPE	< ATTRIBUTE LIST>	<RELATION>	<VALUE>
meta attributes:	source	MATCH	book102
	blossoming period	IN_INTERVAL	[April, May]
semantic attributes:	use-in-medicine	>=	small

GENERAL PURPOSE VIEW
OBJECTS:



RELATIONS:



The result set is:

7. CONCLUSIONS

The major achievements of the proposed GIR model could be summarised as follows:

(1) its generality. The model uses the main techniques from the existing image retrieval models and it is applicable to a wide variety of image collections;

(2) its practicality. The model can be used as an image manipulation language in an image database system;

(3) richness of the used methods for image retrieval, image indexing, and similarity measure. All main techniques for image querying are summarised in the proposed model and the model allows diverse ways for image retrieval. The model supports a variety of similarity measures and metrics. It is flexible in the methods it uses for image retrieval to achieve top efficiency in indexing;

(4) its flexibility. The model could be customised when used with a specific application.

The presented model could be extended for distributed image database systems and multi-media database containing text, video and speech signals.

At present software realisation of the model for Windows 95 is considered.

8. ACKNOWLEDGEMENT

This work is partially supported by a project I 531 of the National Foundation for Science Research of Bulgaria.

9. REFERENCES

[1] P. Stanchev, General Image Data Model, Paper presented in IEEE Multimedia, 1997.

[2] P. Stanchev, General Image Database Model, Paper presented in IEEE Multimedia, 1997.

[3] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele and P. Yanker, Query by Image and Video Content: The QBIC System, IEEE Computer, Vol. 28, No. 9, 1995, pp. 23-32.

[4] V. Gudivada, V. Raghavan and K. Vanapipat, A Unified Approach to Data Modelling and Retrieval for a Class of Image Database Applications, IEEE Transactions on Data and Knowledge Engineering, 1994.

[5] T. Joseph and A. Cardenas, PICQUERY: a High Level Query Language for Pictorial Database Management, IEEE Transaction on Software Engineering, Vol.14, No. 5, 1988, pp. 630-638.

[6] P. Stanchev, A. Smeulders and F. Groen, An Approach to Image Indexing of Documents, In E. Knuth and L. Wegner (Eds.), Visual Database Systems II, North Holland, 1992, pp. 63-77.

[7] P. Stanchev, Medical Knowledge Acquisition for Expert Consultation System, in M. Chytel and R. Engelbrecht (eds.), Medical Expert Systems, Sigma press, London 1988, pp. 153-163.

[8] R. Agrawal, C. Faloutsos and A. Swami, Efficient Similarity Search in Sequence Databases, FODO Conferense , Evanston, Illinois, 1993.

[9] P. Stanchev and V. Vutov, A Model-Based Technique with a New Indexing Mechanism for Industrial Object Recognition, IEEE/CHMT, IEMT, 1990, pp. 56-60.