# Image Retrieval by Content as an Anticipatory System

Irina Mocanu, Eugenia Kalisz
University Politehnica of Bucharest, Computer Science and Engineering Dept.
Splaiul Independentei 313, sector 6
060042, Bucharest, Romania
E-mail: irinam@cs.pub.ro, ekalisz@cs.pub.ro
http://prof.cs.pub.ro/~irinam , http://prof.cs.pub.ro/~ekalisz

**Abstract.**
Image retrieval by content represents a part of a multimedia database management system. The components of the image retrieval by content could be seen as anticipatory processes. Each component could anticipate its own evolutionary or functional process or the evolutionary or functional process of other component. This paper presents the anticipatory aspects of the components of an image retrieval by content system.
**Keywords**: Anticipation, image retrieval, image retrieval by content, performance evaluation, image query

## 1 Introduction

With the increasing role of the image databases the problem of image retrieval becomes an essential one. 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 using a set of key words or free text. The queries are based on exact or probabilistic match of query text.

Image retrieval by content boils down two problems [1]: (i) how to mathematically describe an image, and (ii) how to assess the similarity between a pair of images based on their abstracted descriptions. The first issue arises because the original representation of an image is an array of pixel values, which corresponds poorly to visual response and let alone semantic understanding of the image. Thus the image will be represented based on its features such as: colour, texture, shape, or combination of colour, texture and shape, spatial location information or semantics.

Images have many types of attribute which could be used for retrieval, including: (i) the presence of a particular combination of colour, texture or shape features (e.g. green stars); (ii) the presence or arrangement of specific types of object (e.g. chairs around a table); (iii) the depiction of a particular type of event (e.g. a football match); (iv) the presence of named individuals, locations, or events (e.g. the Queen greeting a crowd); (v) subjective emotions one might associate with the image (e.g. happiness); (vi) metadata such as who created the image, where and when.

**International Journal of Computing Anticipatory Systems, Volume 20, 2008**
**Edited by D. M. Dubois, CHAOS, Liège, Belgium, ISSN 1373-5411 ISBN 2-930396-07-5**

Each listed query type (with the exception of the last) represents a higher level of abstraction than its predecessor, and each is more difficult to answer without reference to some body of external knowledge. This leads naturally on to a classification of query types into three levels of increasing complexity [2]:

(1) comprises retrieval by *primitive* features such as colour, texture, shape or the spatial location of image elements;

(2) comprises retrieval by *derived* (or *logical*) features, involving some degree of logical inference about the identity of the objects depicted in the image;

(3) comprises retrieval by *abstract* attributes, involving a significant amount of high-level reasoning about the meaning and purpose of the objects or scenes depicted.

The most significant gap at present lies between queries of type 1 and 2. Queries of type 2 and 3 are referred as semantic image retrieval.

The content-based image retrieval can be characterized by the ability to retrieve relevant images starting from a user defined image query, based on the semantic content of the images. The search is usually based on similarity rather than on exact match, and the retrieved images are then ranked according to a similarity value - a metric distance between the image query and each image from the database. All the retrieved images have this calculated distance under a fixed threshold, $\varepsilon$. The retrieved images should also be similar to the query image from the human intuition point of view. Consequently, in the process of evaluating these methods a human expert will be needed. In this paper we present our point of view on the application of anticipation to image retrieval by content.

## 2 Structure of the Image Retrieval by Content

Figure 1 provides an overview of the multimedia image retrieval systems operation. Images from the database are preprocessed to extract features and are indexed according to these features. The extracted features are dependent on the method used for computing similarities between images. Each user query is processed and its main features are extracted. Image similarity computation is performed by using a metric distance applied to two sets of image features:

- features extracted from the database (or a subset of these features, if an index is used)
- features of the query.

Only images for which this distance is below a defined threshold are considered similar to the query image and will be retrieved and presented to the user. These images are presented in descending order in terms of similarity or in ascending order in terms of distance.

The structure of the image retrieval by content described in Figure 1 uses one method (criterion) for comparing images. There are situations in which the images obtained by a retrieval method must be filtered by one or several other methods. In this case the image retrieval process is a multi-step one.

Any new or enhanced image retrieval method should be evaluated before applying it in a working system. One way to evaluate an image retrieval method is to compare its results with those obtained by a human expert, who selects similar images from the database according to human intuition. The two sets of results – expert's and method's - will be compared for several images. The method's performance will be evaluated using precision and recall [3].

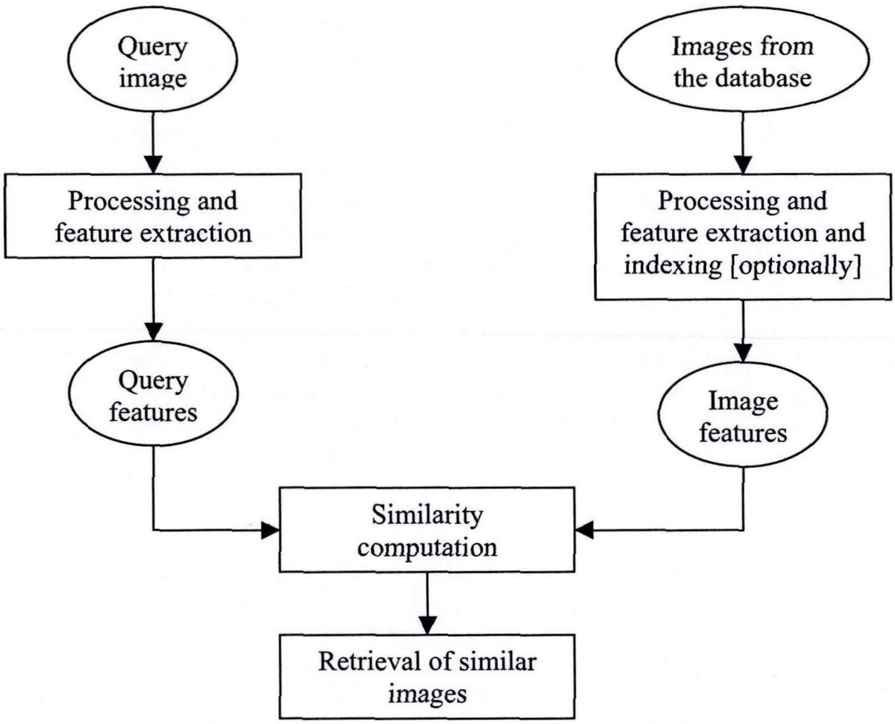



**Figure 1:** Structure of the image retrieval by content

*Precision* measures the retrieval accuracy. It is defined in [3] as the ratio between **nRr** - the number of relevant images retrieved and **nr** - the total number of retrieved images. *Recall* measures the capacity to retrieve relevant information items from the database. It is defined in [3] as the ratio **nRr / nR**, between the number of relevant images retrieved and the total number of relevant images in the database.

A method with high recall but low precision (Figure 2 a) will return a long list of images, many of which are irrelevant. On the other hand, high precision but low recall (Figure 2b) means that many images relevant to the query are not retrieved. A good retrieval method should balance the recall and precision. Ideally all retrieved images should be relevant: **nr = nR = nRr.**

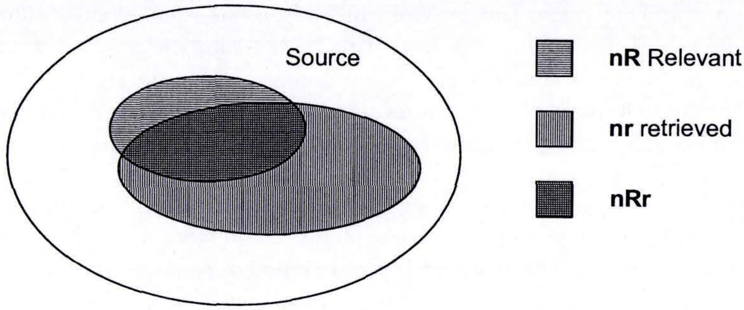



**Figure 2a:** High recall, low precision

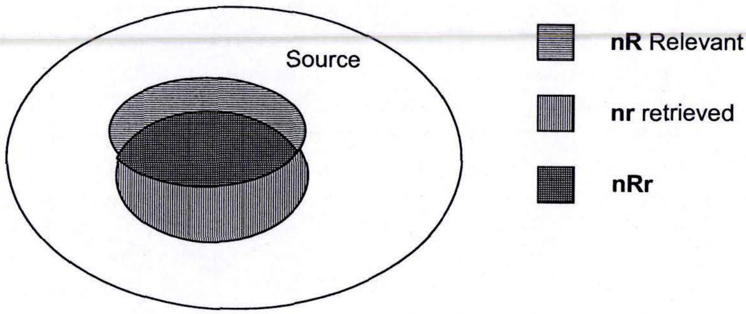



**Figure 2b:** High precision, low recall

To compare the performance of image retrieval methods, both recall and precision should be compared. One technique to do this is to determine precision values corresponding to recall values ranging from 0 to 1 (with step 0.1) and to plot a precision-recall graph for each method, as shown in Figure 3.

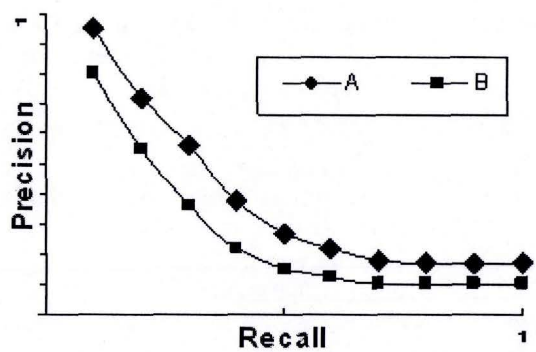



**Figure 3:** Precision-recall graphs for two methods A and B

A system with the graph further from the origin has higher performance. So the method with graph **B**, from Figure 3, performs better than the method with graph **A**. In [4], several methods for image retrieval by content based on shape were analyzed using this approach. **A** may be the centroid radii method [5] and **B** may be distance histograms method [6]. The performance of these two methods was analyzed using a synthetic image database constructed with images from SQUID database [7]. For each of these shapes, three similar shapes were generated: one scaled, one mirrored and one flipped. The database used consists of approximately 3,000 synthetic shapes with single contours, affected by noise. The query shape was selected from the database.

## 3 Image Retrieval by Content as an Anticipatory System

Let us consider that at time **t** an image query **IQ(t)** is issued, and the expected result is the list of relevant images **R(IQ(t))**, retrieved from the source **S** – database or index. The image retrieval method used **(M)** will search in the database (using or not an index) images which satisfy the similarity condition, and will build a list of images **r(IQ(t))**, relevant from its point of view. But, in most cases, this list represents only an anticipation of the final list $(\mathbf{R^*(S, M, IQ(t))}$, and requires additional processing. Consequently, the method acts as a *model generating the anticipation* of a next state in the retrieval process, and can be viewed as component of a *weak anticipatory system* [8].

### 3.1 Single Step Image Retrieval by Content

The image retrieval by content method M builds the list **r(IQ(t))** iteratively, analyzing each of the N source images.

For simplicity let us note:

$\mathbf{C_i}$ - the currently analyzed image

$\mathbf{d_i} = \mathrm{Dist}(C_i, IQ(t))$ – the distance current – query image

$\varepsilon$ - the threshold discriminating images similar to IQ(t), with $\mathbf{d_i} \le \varepsilon$

$\mathbf{r_i}$ – the current state, represented by the list of images similar to IQ(t), ordered in ascending order of **d**; the initial state $\mathbf{r_1}$ is an empty list and the final state, $\mathbf{r_{N+1}}$, represents the retrieved images list, anticipated by method M as relevant

$\mathbf{nr_i}$ – the length of $\mathbf{r_i}$ (the number of images retrieved so far)

where i = 1, 2,…, N.

The transition from one state to the next (Figure 4) is defined as

$$r_{i+1} = \mathbf{InsertSimilar}(r_i, C_i, IQ(t))$$

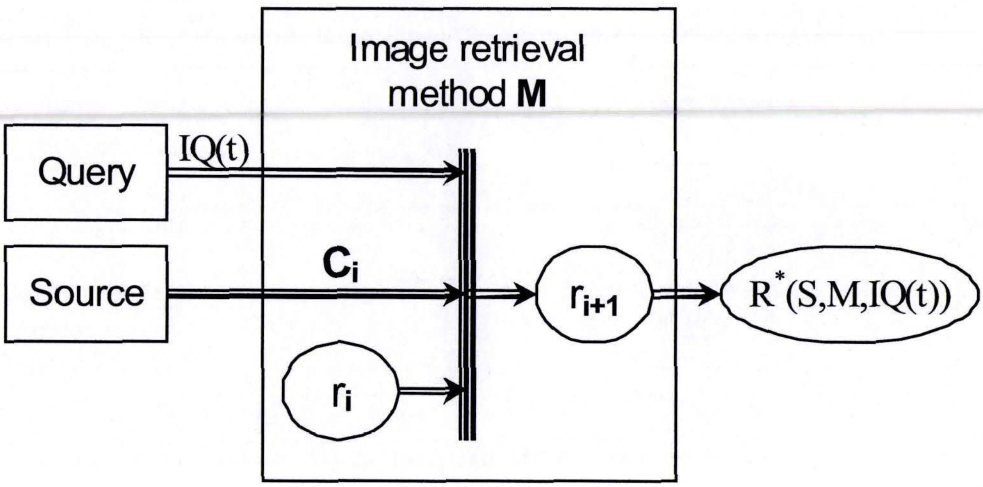
where **InsertSimilar** is the function that computes the distance $\mathbf{d_i}$, compares it to with the threshold $\varepsilon$ and, if lower or equal, inserts the current image in the ordered list $\mathbf{r_i}$, according to the value of $\mathbf{d_i}$.

Since the method M analyzes all the N source images, each image being considered as possibly relevant, it can be considered a strong anticipatory system, where the anticipated next state is

$$r^*_{i+1} = \mathbf{Insert}(r_i\, ,\, C_i)$$

and

$$r_{i+1} = \mathbf{InsertSimilar'}(r_i\, ,\, r^*_{i+1}\, ,\, IQ(t))$$




i = 1,2,…,N

**Figure 4:** Single criterion image retrieval by content

### 3.2 Multi-Step Image Retrieval by Content `

In our view a multi-step image retrieval by content system uses a collection of several image retrieval methods which may be applied in different ways to obtain different (or not) collections of retrieved images, similar to the current query image, IQ(t). For simplicity, in the following formulas we will not mention explicitly the parameter IQ(t).

Let us consider two image retrieval methods, M1 and M2. They can be used to obtain:

- all the images retrieved by either M1 or M2 ($M1 \cup M2$) - Figure 5a;
- only images retrieved by both M1 and M2 ($M1 \cap M2$) - Figure 5b.

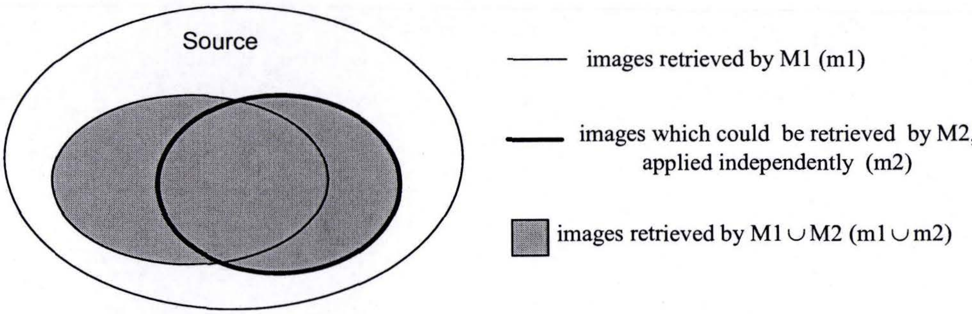



**Figure 5a:** Multi-step image retrieval by content for m1 ∪ m2

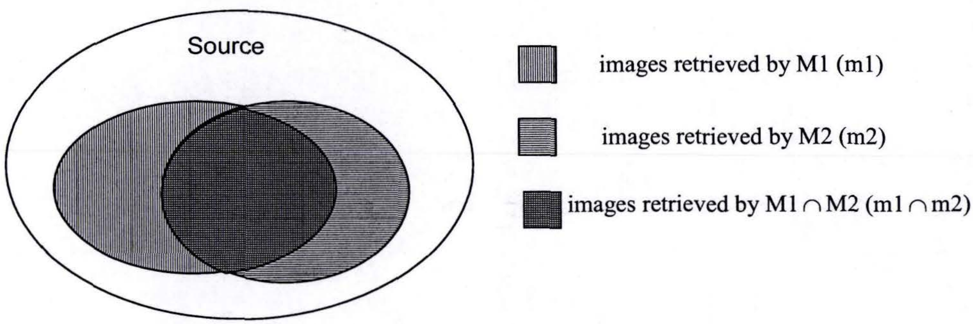



**Figure 5b:** Multiple step image retrieval by content for m1 ∩ m2

In multi-step image retrieval by content system several image lists are built. The simplest approach is to consider unordered lists, but a global distance measure could also be computed as follows:

$$\delta_i = \Sigma\ (Dist_{Mj}\ (C_i)\ /\ \varepsilon_{Mj})$$

where $C_i$ represents the current image, $Dist_{Mj}$ - the distance computed by method j and $\varepsilon_{Mj}$ - the specific threshold.
For retrieving images in m1 ∪ m2 (Figure 6a):

- ✓ M1 is applied to the source images from the database or from the index (if an index is used); the result of this process is the list $\mathbf{L_{M1}}$, of the images retrieved by M1 and ordered according to the distances computed by the specific function $\mathbf{Dist_{M1}}$;
- ✓ M2 is also applied on the source images, returning the list $\mathbf{L_{M2}}$;
- ✓ The two resulting lists are merged as follows

```
for each image Ci retrieved by M2
    if Ci exists in LM1
      then Modify(δi, Ci, M2)
      else Insert(LM2, Ci)
```

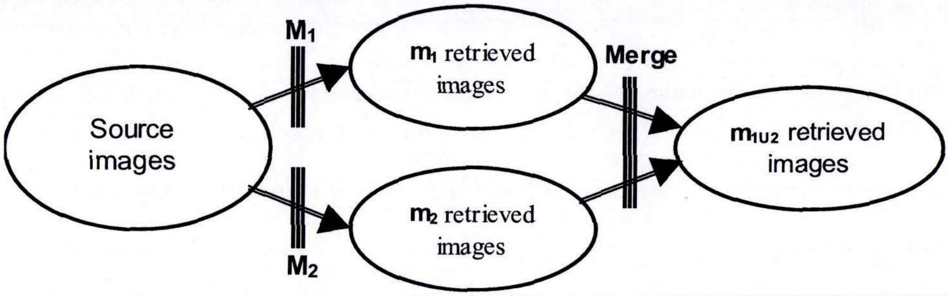



**Figure 6a:** Multiple step image retrieval by content for m1 $\cup$ m2

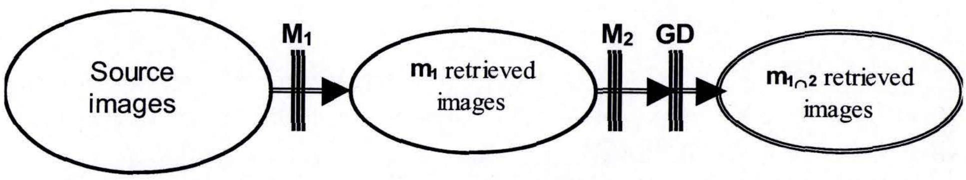



**Figure 6b:** Multiple step image retrieval by content for m1 $\cap$ m2

For retrieving images for m1 $\cap$ m2 (Figure 6b):

- ✓ M1 is applied to the source images from the database or from the index (if an index is used);
- ✓ M2 uses as source the images retrieved by method M1
- ✓ GD computes the global distance $\delta$ for each image retrieved by M2 and inserts it in the list accordingly.

The process of multiple step image retrieval by content can be seen as a weak anticipatory system, each method anticipating the list of images relevant for the current query image IQ.

## 4 Anticipation in Testing and Tuning an Image Retrieval Method

The evaluation of an image retrieval method requires to plot precision-recall graphs (as the one in Figure 3) for several databases and query images. Two types of evaluation (tests) can be performed:

- stability evaluation – a method is considered stable if the graphs obtained for the same types of queries, addressed to the same data base, are very close one to each other
- comparative evaluation, with reference graphs, obtained for a reference method, considered the best up to the evaluation model, and applied to the same data bases and query images.

In this process, at each evaluation step the anticipated graph is one situated in the region either close to (in the first case) or above (in the second case) the reference graphs. If the resulting graph does not satisfy the anticipation, a tuning of the method, by adjusting the specific threshold, or even a more deep method revision, should be considered.

As future research we intend to develop a general framework for adjusting the image retrieval method threshold, which will behave as anticipatory strong system.

## 5 Conclusions

Image retrieval by content (single or multi-step) and the evaluation of an image retrieval method have all anticipatory aspects. The paper presented the way these methods are used and outlined their anticipatory behavior, either weak or strong.

## References


[1] Y. Rui, H.S. Thomas. Image Retrieval: Current Techniques, Promising Directions, and Open Issues, Journal of Visual Communication and Image Representation, No. 1, March 1999, pp. 39-62.

[2] J. Eakins, M. Graham. Content-based Image Retrieval, JISC Technology Applications Programme Report, October 1999, http://www.jisc.ac.uk/uploaded_documents/jtap-039.doc.

[3] G. Lu. Multimedia Database Management Systems - Artech House Publishers, 1999.

[4] I. Mocanu. Optimizations Algorithms in Multimedia Database Systems, Ph.D. Thesis, University Politehnica of Bucharest, Computer Science and Engineering Dept. , October 2006.

[5] K. L. Tan, L. F. Thiang. Retrieving similar shapes efficiently – Multimedia Tools and Applications - Kluwer Academic Publishers, 2001.

[6] S. Fan. Shape Representation and Retrieval Using Distance Histograms - Technical Report TR 01-14, department of computer science, University of Alberta, Edmonton, Canada, 2001.

[7] SQUID Database - http://www.ee.surrey.ac.uk/Research/VSSP/imagedb, accessed june 2007

[8] D.M. Dubois. Review of Incursive, Hyperincursive and Anticipatory Systems - Foundation of Anticipation in Electromagnetism. Computing Anticipatory Systems: CASYS'99. Edited by Daniel M. Dubois, Published by The American Institute of Physics, AIP Conference Proceedings 517, 2000, pp. 3-30.