A Region-Similarity-Based Image Retrieval System

J.F. Omhover **M. Detyniecki** **B. Bouchon-Meunier**

LIP6 – Pole IA
8, rue du Capitaine Scott
75015 Paris, France
contact: jean-francois.omhover@lip6.fr

Abstract

In this document, we present an image retrieval system that is based on a segmented representation of the visual content. This representation leads to a comparison of the image content that is more "semantic" than a classical global comparison. The system compares regions using fuzzy similarity measures that have been show to be psychologically intuitive and easy to aggregate. We then exploit the aggregation between regional similarity measures to let the user build four different types of original visual requests. Our system can handle these requests easily, thanks to its open architecture that let the expert user modify its parameters and compose new aggregation operators on them.

Keywords: Image Retrieval, Segmentation, Fuzzy Similarity Measures, Aggregation.

1 Introduction

In the last decade, there has been an increasing interest in image retrieval systems. The idea is to let some user query an image database for a specific image content. Unfortunately, this common goal has not yet been achieved [5].

To avoid the use of textual annotations, subjective and highly consumming in terms of human work charge, the community has focused on automatical extraction and comparison of image features. The query is then expressed not as some textual description, but as some image, chosen by the user, which is compared to the entries of the database. As a response, the user gets a list of images which are found most similar to his request.

Many different feature spaces and many systems have been developed to compute the similarities involved in those queries, among them we can point out QBIC [6] and VisualSEEk [12].

At the same time, in the last few years, segmentation has been used as a preprocessing tool for image indexation. Images can be segmented in regions, roughly corresponding to objets in the image. Thus reflecting the objective content of an image, measures were developed to compare regions one to the other, and build search engines based on these measures.

For the moment, only a few image retrieval systems follow this approach. For instance, Blobworld [1] proposes its segmentation tool to the user, who specifies the request by selecting a set of regions of interest. The retrieval then consists in comparing pairs of regions. The returned images are those for which the regions obtained the best scores in the one-to-one region similarity computation. Simplicity [7, 2], another system, extends this to image-to-image comparison. It compares images globally by aggregating the region-to-region similarities. Unfortunately, its scheme does not include the spatial configuration of the regions, though it tries to match regions of the request to regions of each image.

In this paper we introduce a system called "STRICT". Its design is based on a formulation of fuzzy similarity measures that is derived from psychological considerations, and that leads to an intuitive formulation of the similarity. These measures can be aggregated to form composite regional queries: that is the possibility to specify a set of regions, and to retrieve images containing these regions in an approximate spatial configuration.
Our system has been implemented so that it is open to online modifications. Its interface lets the user query the system by composing an aggregation operator to support different type of queries. We propose four types of specific aggregated queries that answer four different needs of visual comparison.
First, in section 2, we present the regional features and the similarity measures we use to compare the regional content and to retrieve the images. In section 3, we introduce the architecture of our system and the different type of requests it can handle. In the last section 4, we illustrate our approach with some experimental requests.

2 Segmented Image Index

From an image, a segmentation algorithm extracts regions complying to a given homogeneity criterium. This criteria is usually based on colors: the algorithm tends to isolate regions of connected pixels which present similar colors. Those regions roughly correspond to the objects present in each image. A similarity measure based on those regions can reflect an objective similarity of the content.
In order to build a semantic similarity between images, we developed a similarity measure for regions based on the similarities of color, shape, and position. First we briefly present the segmentation algorithm used to extract objects from images (section 2.1), then we introduce the definition of fuzzy similarity measures (section 2.2), and finally, we explain which region features we chose to extract and how these features are compared using fuzzy similarities (section 2.3).

Figure 1: Segmentation in regions

2.1 Segmentation

For the segmentation of an image in regions, many different approaches exist [3]. They have been developed for thirty years and applied to various application fields. They all aim at building a crisp partition of the image, based on vectors computed from the pixels: color, texture coefficients, edge orientation.
Our algorithm [4] provides good results (see figure 1) in a very short computation time. It follows the merge approach: each pixel is first considered as an isolated region, then fusions are operated to merge connected pixels of similar colors, until there is no more possible fusion. This occurs when all the connected regions are dissimilar enough.

2.2 Fuzzy Similarity Measures

A definition of fuzzy similarity measures [9] has been derived from Tversky's contrast model, a psychological framework for similarity measurement. As shown in [11], these measures provide an intuitive measurement of similarity. They are also independent of the scale of the fuzzy sets. In this scheme, the similarity between two fuzzy subsets A, B of feature space F can be calculated by a function S, with three variables:

- $M(A \cap B)$ the area of the fuzzy intersection of A and B, which measures *the common features* of fuzzy sets A an B.
- $M(A-B)$, the area of the fuzzy difference of A by B, which measures *the features that are only present in* A.
- $M(B-A)$, the area of the fuzzy difference of B by A, which measures *the features that are only present in* B.

These areas correspond to the following sums, caculated from the membership functions μ_A, μ_B of A, B:

$M(A \cap B) = \sum_{x \in F} min(\mu_A(x), \mu_B(x))$,

$M(A - B) = \sum_{x \in F} max(\mu_A(x) - \mu_B(x), 0)$,

$M(B - A) = \sum_{x \in F} max(\mu_B(x) - \mu_A(x), 0)$

This expression of fuzzy similarity measures falls within Tversky's contrast model [14]. Based on this general framework, different particular measures were derived. We chose to implement only four of them as region similarity measures: Jaccard, Dice, Ochiai and Fermi-Dirac similarity measures [10]. With $X = M(A \cap B)$, $Y = M(A - B)$, $Z = M(B - A)$, we have:

$$\begin{aligned}
S_{jaccard}(X, Y, Z) &= \tfrac{X}{X+Y+Z} \\
S_{dice}(X, Y, Z) &= \tfrac{2X}{2X+Y+Z} \\
S_{ochiai}(X, Y, Z) &= \tfrac{X}{\sqrt{X+Y}\sqrt{X+Z}} \\
S_{fermidirac}(X, Y, Z) &= \frac{F_{FD}(\phi) - F_{FD}(\frac{\pi}{2})}{F_{FD}(0) - F_{FD}(\frac{\pi}{2})}
\end{aligned}$$

with $F_{FD}(\phi) = \frac{1}{1+\exp^{\frac{(\phi-\phi_0)}{\Gamma}}}$,

$\phi = \arctan(\frac{Y+Z}{X})$.

Γ and ϕ_0 being parameters balancing the selectivity of Fermi-Dirac measure [10]. Considering different measures enable us to investigate their properties. In particular, we show in [8] that the ranking in the list of the results provided by an information retrieval system using these measures is conserved between some of them.

2.3 Region Fuzzy Similarity using color, shape and position

The resemblance between regions is more elaborated than the global one. We have to consider several aspects of comparison, that are color, shape and position. For each of these comparisons, a region is represented as a vector. Because these three vectors belong to different spaces, and cannot be related one to the other, the similarity between pairs of region relies on three different measures: one measure for each of the vector pairs. Formally, for two regions $R(i)$ and $I(j)$, extracted respectively from the image request R and from an image database entry I, the region similarity S_{reg} can be written as an aggregation of three "sub-measures": $S_{reg|color}$ measuring color similarity based on histograms, $D_{reg|position}$ evaluating the proximity of the centers of the two regions, and $S_{reg|shape}$ measuring the shape similarity based on the minimum bounding rectangles (MBR) of $R(i)$ and $I(j)$.
The aggregation operator used to combine these three measures is the average:

$$\begin{aligned}
S_{reg}(R(i), I(j)) = \quad & \lambda_c . S_{reg|color}(R(i), I(j)) \\
& + \lambda_p . D_{reg|position}(R(i), I(j)) \\
& + \lambda_s . S_{reg|shape}(R(i), I(j))
\end{aligned}$$

$$\text{with } \lambda_c + \lambda_s + \lambda_p = 1$$

The three parameters λ_c,λ_s,λ_p are chosen to balance the influence of each of the three features spaces on the region similarity. They have been set to $\lambda_c = 0.6$, $\lambda_s = 0.2$, $\lambda_p = 0.2$ for our experiments.

2.3.1 Fuzzy Region Color Similarity

The region color similarity $S_{reg|color}$ is based on the histograms of the regions [13]. We extract the color distribution of each region from each image. To compare the regional color histograms, we use the measures presented in section 2.2. For two regions $R(i)$, $I(j)$, extracted respectively from R, I, we compare their respective histograms $H_{R(i)}$, $H_{I(j)}$ by computing:

$$\begin{aligned}X &= Area(H_{R(i)} \cap H_{I(j)})\\ Y &= Area(H_{R(i)} - H_{I(j)})\\ Z &= Area(H_{I(j)} - H_{R(i)})\\ S_{reg|color}(R(i), I(j)) &= S(X, Y, Z)\end{aligned}$$

S being one of the Jaccard, Dice, Ochiai, or Fermi-dirac resemblance measures.

2.3.2 Fuzzy Proximity Measure

The proximity measure $D_{reg|position}$ is not a resemblance measure. It simply evaluates the degree of proximity of the center of gravity of two regions. We define a fuzzy set which represents the set of distance values considered as "near" (see figure 2). $D_{reg|position}(R(i), I(j))$ simply returns the membership of this set at $d = \|C_{R(i)} - C_{I(j)}\|$, that is the normalized distance between the centers $C_{R(i)}$,$C_{I(j)}$ of both regions $R(i), I(j)$.

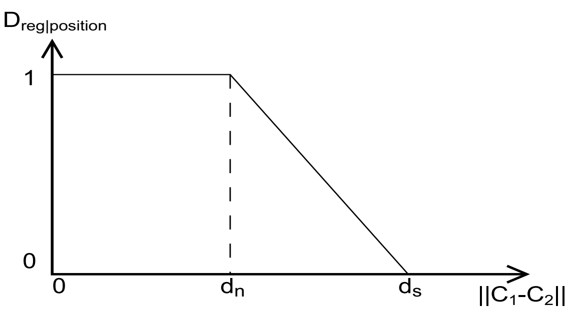

Figure 2: $D_{reg|position}$ from the distance between the centers of two regions

The parameters d_n and d_s specify the kernel and support of the fuzzy set D_{near}. They can be modified by the user of our system to widen or tighten the selectivity of this "proximity set" (see section 3.1), but they are set by default to $d_n = 0.2$ and $d_s = 0.4$.

2.3.3 Fuzzy Shape Similarity

The similarity of the shapes of two regions relies on the comparison of their minimum bounding rectangles (MBR). These are rough representation of shape, but they are easy to implement and compare. The MBR of a region is the rectangle that minimally covers the region and the borders of which are parallel to the borders of the image.

To compare the two shapes of regions $R(i), I(j)$, we center their minimum bounding rectangles $MBR_{R(i)}$,$MBR_{I(j)}$ on a common point, and we compute the common surface ($X = M(MBR_{R(i)} \cap MBR_{I(j)})$) and their different surfaces ($Y = M(MBR_{R(i)} - MBR_{I(j)})$,$Z = M(MBR_{R(i)} - MBR_{I(j)})$). A resemblance measure is then used to compute their similarity from these three areas.

At the time we prepare this article, the MBR comparison is the only similarity we have implemented for shape comparison. It is a trivial scheme that lacks of accurate representativity. We are currently working on another shape representation, based on shape invariants.

2.4 Region to Image comparison

To let the user perform an image retrieval by a single region, we build a global measure, based on the request of a specified region. Then we compare this request to all the regions of each image, and return a single score. We choose to take into account the maximum of the region-to-region comparisons between the request and each image of the database: this score indicates if the region specified in the request is roughly present in the image.

$$S_{reg}(R(i), I) = \mathrm{Max}_{I(j)\in I}(S_{reg}(R(i), I(j)))$$

3 STRICT Image Retrieval Platform

In order to implement, test and validate our approach, we propose a platform, which is evolutive enough to let us insert new methods, modify their parameters, and run them in parallel for benchmarking. This is the system STRICT (System for Testing Retrieval by Image ContenT).

Its main assets are its dynamic web interface, and its open computation engine. These two parts form a powerful tool for image retrieval testing and fuzzy similarity experimentations. Here we first present the architecture of our system (3.1), then explore the use of aggregation operators to form composite queries (3.2)

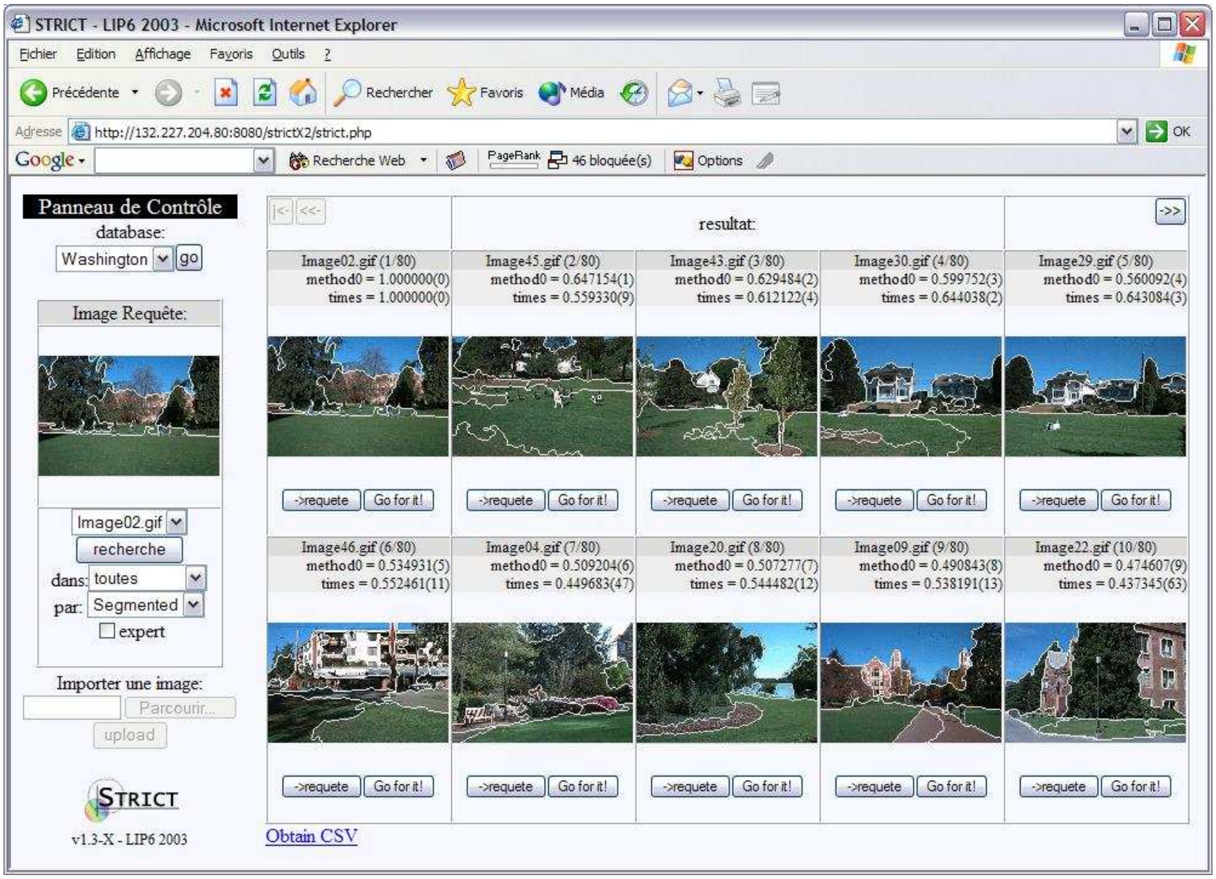

Figure 3: STRICT Interface, browsing a request's result

to look for a group of regions in an image database.

3.1 Architecture

STRICT is composed of two distinct servers: the indexation server (developed from scratch), that keeps the signature database and computes similarities on demand, and an HTTP server that runs the interface, serves the images to the user and presents the results of his requests.

The invisible part of the system is the indexation server. It maintains a database of vectors extracted from images. These vectors are the signature of each image. Each measure of comparison between images is based on a different vector (see section 2 for the description of our vectors and their corresponding measures). Different vector spaces and their similarity measures can be maintained by the indexation server. These different similarity measures can be runned simultaneously.

The indexation server receives the requests from the interface, computes the similarities, and returns a list of the most similar images.

The web interface is the visible part of the system. Its dynamic format enables the expert user to browse the image database, formulate a request and compare the results from different similarity measures (see figure 3).

Our similarity measures have parameters. The parameters type and values are listed within the indexation server: it registers all available measures and their corresponding parameters. When a user opens a new connection to the interface, the description of each measure is sent to the web server. From this list, the web server generates a form for the user that allows him to modify parameters.

A request can then be composed of a list of measures with different parameters. Each measure of the request is then runned against every image in the database. When the resulting list of images is displayed, each image is marked by the intensity of each measure specified. This feature enables the user to compare the effects of its parameters on the retrieval.

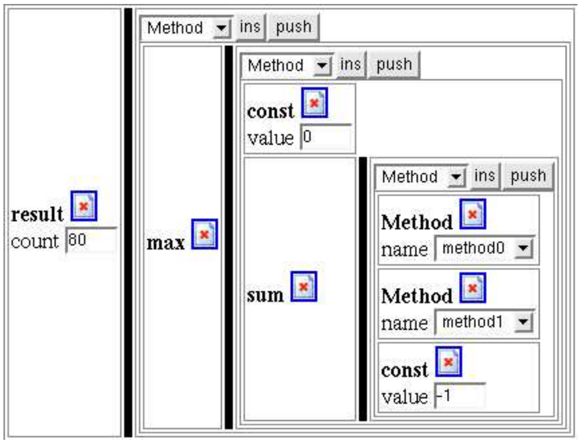

Figure 4: Lukasiewicz t-norm built by simple clicks on STRICT web interface

Furthermore, the system can aggregate these results on-the-fly. The user can formulate an aggretation operator with basic mathematical operations (see figure 4). Using the minimum, the maximum , the addition and the product, he can define his own operator, specify its entries within the available similarity measures, and then test this newly aggregated similarity against the images. The basic operators programmed in this application let us build most of the known aggregation operators. The utility of this scheme is shown practically in the following section.

3.2 Possible Requests

STRICT can compute global and regional similarities (see section 2). Those similarities can be aggregated on-the-fly under user's command. This feature let the user build different kinds of requests. This section points out four schemes, where aggregation is involved.

- From a single request image R, the user can compare the image entries using *multiple features* (color , texture..). STRICT then agregates the scores resulting from the different similarity measures. The final score $f(I)$ of the image I is then expressed as:

$$f(I) = Agg(S_1(R, I), ..., S_n(R, I))$$

- The user can also point out *multiple image* requests. The idea is to retrieve images, which are similar to all the requests (some of them might be more important than others). With $R_1, ..., R_n$ a set of request images, STRICT will compute the global similarity between each image I in the database and each request R_i. It will then aggregate these similarity degrees to provide a final score $f(I)$ for each I.

$$f(I) = Agg(S(R_1, I), ..., S(R_n, I))$$

  Here, the aggregation can also reflect the importance given by the user to each image request, by an optional weighting.

- From a *single image request* R, the user can select *different regions* $R(i), ..., R(j)$ that he wants to look for in the database. The final score $f(I)$ is then expressed as:

$$f(I) = Agg(S_{reg.}(R(i), I), ..., S_{reg.}(R(j), I))$$

  Attention must be paid to the fact that each score $S_{reg.}(R(i), I)$ is a global one (see 2): the aggregated score does not indicate if the set of regions in the image I is formed of distinct regions, they just individually look like one of the regions $R(i)$.

- Obviously, the previous scheme can also be applied to a request composed of regions extracted from different images. $f(I)$ then takes the following form:

$$f(I) = Agg(S_{reg.}(R_1(i_1), I), ..., S_{reg.}(R_n(i_n), I))$$

The dynamic interface of our system is a practical tool to test all of these four schemes. With simple actions, an expert user can simulate one of these four requests, specify its parameters, and compose an aggregator to fulfill its needs. No current CBIR system offers

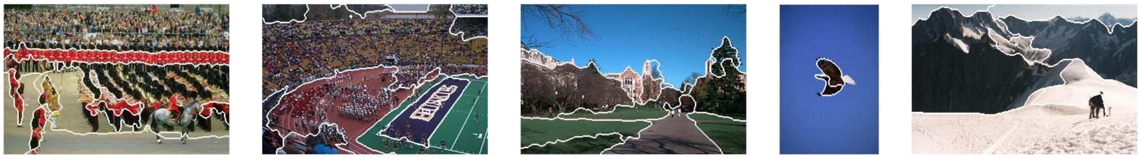

image1 image2 image3 image4 image5

Figure 5: Images used as requests to evaluate the efficiency of region similarity

all of these features. Though many similarity measures in our system are not yet fully developed, we have built a structure that let us implement, experiment and compare different methods of the image retrieval field.

4 Preliminary Experiments

As our system has been recently developed, we have not proceeded yet to its full evaluation. In this section, we only propose a preliminary array based on five visual requests (see figure 5): each time, 2 or 3 regions of interest were selected, then used as a query. We choose to aggregate regional similarity measures (based on color, position and shape) using an averaging operator. The effectiveness of this aggregated measure is compared to the effectiveness of the global histogram similarity (one global histogram is extracted and compared to another histogram by a fuzzy similarity measure). The system retrieves the 80 best results for both methods among a database of 7700 images. This operation takes about 1.2 seconds. We evaluate the effectiveness of both methods by measuring two classical values:

- the rank RK of the first non pertinent image in the list.
- the proportion P of pertinent images in the first 10 results (extended to 20 or 30 when the first non-pertinent image was not present in the 10 first). The number of 10 results was chosen only to reflect the satisfaction of the first page of the results returned.

For a method to be efficient, it has to maximise the rank of the first non pertinent image, and also maximise the proportion of pertinent images in the first results.

On table 1, we observe that our method based on region matching gets better results than the global histogram comparison. Obviously, though the present preliminary evaluation shows encouraging results, a larger benchmarking should be runned to prove the effectiveness of our system.

Table 1: Evaluation Results

	Segmented		**Global Histo.**	
	P	RK	P	RK
image1	19/20	19	7/20	5
image2	22/30	19	24/30	21
image3	4/10	5	4/10	3
image4	5/10	3	1/10	2
image5	4/10	4	4/10	2
average results				
	0.59	10	0.41	6.6

5 Conclusion

Our image retrieval system uses a regional representation of images, built using a segmentation algorithm. This solution, currently investigated in the community, has been implemented in STRICT to let the user specify his visual interest in the request. For the expert, our system proposes features that enables him to modify parameters, build complex requests, and compare their results in a dynamic and easy to use interface.
We realize that the choice of the parameters and the operators used by the system to build complex queries may be difficult for the non-expert users. STRICT is in fact a CBIR experimentation platform that requires from users to know what they are doing, and for

what purpose. A simpler version of the interface system is being developed to propose a ready-to-use image retrieval tool, including pre-chosen parameters and aggregation operators.
The efficiency of regional similarities has been tested against a classical global color measure, we led some preliminary experiments in which we got encouraging results. Although it still needs development, our system proposes a scheme relying on fuzzy resemblance measures and aggregation operators that reflect an intuitive similarity between visual objects.

References

[1] C. Carson, M. Thomas, S. Belongie, J.M. Hellerstein, and J. Malik. Blobworld: A system for region-based image indexing and retrieval. In *Third Int. Conf. on Visual Information Systems*, Amsterdam, 1999.

[2] Y. Chen and J.Z. Wang. A region-based fuzzy feature matching approach to content-based image indexing and retrieval. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(9), September 2002.

[3] H.D. Cheng, X.H. Jiang, Y. Sun, and J. Wang. Color image segmentation: advances and prospects. *Pattern Recognition*, 34:2259–2281, 2001.

[4] G. Durand and P. Faudemay. A fast region-merging segmentation algorithm for video analysis and indexing. In *CIR 98 Symposion and Workshop on Image Retrieval*, Newcastle, 1998.

[5] J.P. Eakins. Towards intelligent image retrieval. *Pattern Recognition*, 35:3–14, 2002.

[6] M.D. 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. *Computer*, 28(9):23–32, September 1995.

[7] J. Li, J.Z. Wang, and G. Wiederhold. Irm: Integrated region matching for image retrieval. In *8th ACM Int. Conf. on Multimedia*, pages 147–156, 2000.

[8] J.F. Omhover, M. Detyniecki, and M. Rifqi. Image retrieval using fuzzy similarity : Measure equivalence based on invariance in ranking. In *IEEE International Conference on Fuzzy Systems Fuzz-IEEE*, Budapest, Hungary, July 2004.

[9] M. Rifqi. *Mesures de comparaison, typicalit et classification d'objets flous : thorie et pratique*. PhD thesis, Universit Paris 6, Dcembre 1996.

[10] M. Rifqi, M. Detyniecki, and B. Bouchon-Meunier. Discrimination power of measures of resemblance. In *IFSA'03*, Istanbul, Turkey, 2003.

[11] S. Santini and R. Jain. Similarity measures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(9):871–883, 1999.

[12] J.R. Smith and S.F. Chang. Visualseek: a fully automated content-based image query system. In *ACM Multimedia '96*, November, 1996.

[13] M.J. Swain and D.H. Ballard. Color indexing. *International Journal of Computer Vision*, 7(1):11–32, 1991.

[14] A. Tversky. Features of similarity. *Psychological Review*, 84:327–352, 1977.