Region-Based Automated Relevance Feedback in Algae Image Retrieval

H. TABOUT, A. SOUISSI and A. SBIHI

Laboratory of Telecommunication Systems and Decision Engineering, Ibntofail University, Kenitra, Morocco

Summary
In this paper, we present a content based image retrieval system with automatic relevance feedback. The proposed technique is based on color interest points extraction and on the calculation of the local descriptors in interest regions. Global and local features are grouped to feed a two level Self Organizing Maps Network in order to discriminate between retrieved images without user interaction. An implemented prototype system has demonstrated a promising retrieval performance for a test database containing more than 1000 color algae images.

Key words:
Self Organizing Maps, Content Based Image Retrieval, Relevance Feedback, Interest points, algae images

1. Introduction

Recently, content based image retrieval (CBIR) methods have gained importance in the research area. The exponential growth of the quantity of multimedia data creates a need for effective and fast access. In this context, it becomes necessary to develop systems of indexing and automatic search to allow the exploration of these data. These systems rely on low-level representation of images in terms of their visual contents such as color, texture and shape in order to compare images. The comparison is usually performed using features extracted automatically from each image. These features should comply with the human perception. This requirement is a difficult challenge; the difficulty comes from the semantic gap between low level image representation and higher level concepts by which human understand images. To overcome the problems of the semantic gap and user subjectivity, interactive systems have been proposed. Relevance feedback originated from the text-based information retrieval is a powerful tool to improve retrieval performance [1]. The main idea of relevance feedback is to let user guide the system during the retrieval process. The user interacts with the system and judges the relevance of the retrieved images according to his subjective perception. With this provided relevance Information, the system dynamically learns the user's intention and gradually boosts its retrieval performance.

However at any iteration, the user labels images to be relevant or irrelevant; and in order to satisfy his interest need, the user should make a heavy work and more consuming time is required.
The present work is concerned with indexing and retrieving images by the contents. More precisely, we are interested in indexing and retrieving algae images. Algae are largely exploited in Morocco. These plants are widespread in the Atlantic coast and they are involved in the chemical, pharmaceutical and food industries. However, many species are deemed to be harmful [2]. The aim of this study is to retrieve the most similar algae images to a query image in order to identify and compare it with the prestored algae images. This is done by considering, on the one hand, the tools of analysis and image processing for the description of the contents of these images and on the other hand, the installation of a system of search and navigation in such a base. In this paper we introduce a novel region based automated relevance feedback method. This approach is based on an unsupervised learning in order to automate relevance feedback and then reduced human interaction. The basic idea of the proposed method comes from the assumption that the user judgment of relevance retrieval is related to the classification problem. Therefore, relevance identification is assured using a two level Self Organizing Maps, by incorporating both global and local features using interest points. For any iteration, feature vectors of retrieved images are implemented to train the neural network. This allows the system to distinguish between relevant and non relevant images automatically. Thus, the system, through the learning process, boosts his performance of retrieving without human involvement. But if the user wasn't satisfied, he could continue retrieving with a combined mode. This allows reducing human effort.

The remainder of the paper is organized as follows. The next section summarizes related works. Section 3 provides a brief review of the used indexing techniques. In section 4, the automated relevance feedback and the algorithm for learning similarity is described. In section 5, simulation results and evaluation are provided, and finally, concluding remarks are offered in section 6.

Manuscript received December 5, 2007
Manuscript revised December 20, 2007

2. Related work

In order to bridge the semantic gap and the subjectivity of the user, a various interactive CBIR methods have been proposed. These methods implement relevance feedback to enhance the performance of retrieval.

Among the frequently used approaches in relevance feedback is updating weights. It consists to reduce the weight of feature vector that has a high variation of the query image in the feature space and to increase the weight of feature vector that has a small variation. In [3], the weights embedded in the query object are dynamically updated to model the high-level concepts and perception subjectivity.

Other used technique is query movement. This method tries to find the ideal query point by moving toward the relevant images and away from the non relevant images. Thus, in each cycle of relevance feedback, the ideal query is determined through combinations of images judged to be relevant or non relevant by the user. Papers that have adopted this approach are [4], [5] and [6]. In [7] Porkaew & al. have used Rocchio's formula [8] to perform query shifting.

Many different techniques [9] have used SVM classifier to determine positive examples. This is done by dividing the database into relevant and non relevant classes.

In [10], the authors proposed a probability based approach which relies on Bayes rule to predict what the target is the user wants. This is done via a probability distribution over possible image targets, rather than by refining a query.

Another important issue is to use unsupervising learning. In Picsom [11], Laaksonen & al. have implemented a tree structured SOMs. The basic idea of this approach is to use several SOMs in parallel for retrieving relevant images from a database. These parallel SOMs have been trained with separate data sets obtained from the image data with different feature extraction techniques.

The major drawback of the aforementioned approaches is the laborious human effort. To alleviate this problem, an automatic machine interaction is proposed. In [12] and [13], the authors implemented a framework based on an unsupervised learning. This learning involves a novel neural network, Self Organizing Tree Maps (SOTM) to automate relevance feedback. However, these approaches did not take regional features into account.

In a previous work, we proposed in [14] a region based image retrieval using a Concurrent Self Organizing Maps (CSOM). This paper extends the earlier research work by introducing an automated relevance feedback. The proposed approach implements a novel neural network and uses interest points in order to take regional knowledge into consideration.

3. Feature Extraction and Image Retrieval

The automatic retrieval process consists of extracting features for a query image. The features of image database are precalculated and stored. Because of the exhaustive computational requirement, the calculation of descriptors is performed off-line. The system then compares the query image with each image database using these features. Similarity measurement, in the first retrieval, is carried out by Euclidean distance and 16 most similar images are returned and displayed by achieving the first retrieval stage. The retrieved images are then classified using a two layer Self-Organizing maps and a new query is reformulated. New features will be then extracted and this process will carry on for two subsequent cycles. An overview of the system is presented in figure 1.

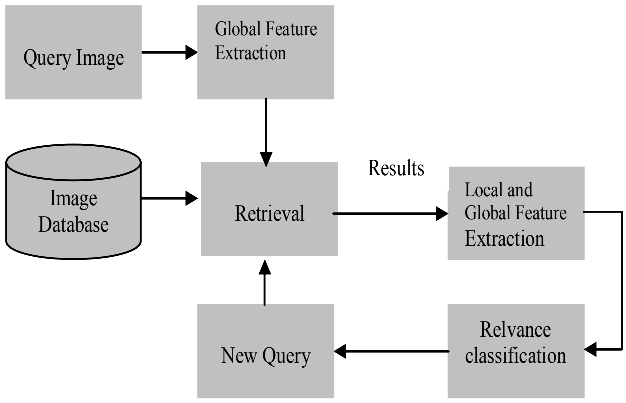


Fig. 1. Overview of the proposed CBIR System

To improve accuracy in automatic relevance feedback system, retrieval and relevance classification features should be different. In this paper four descriptor are used: Color Histogram, Color moments, Gabor Filters and Wavelet moments.

3.1 Color Histograms

Color histograms are largely used in content based image retrieval. It's a representation of the distribution of colors in an image. 48 components are generated in the HSV color space.

3.2 Color Moments

Proposed by Stricker and Orengo [15] to overcome the quantization effect of the histograms. In this approach, color features of images are represented by their color moments namely average, variance, and skewness in each color channel. The used moments have been proved to be efficient in representing color distribution [15]. 9 dimensional moments in HSV color space are extracted.

3.3 Gabor Filters

Frequently used [16], Gabor Filters are an effective method in texture feature extraction [17], [18]. A range of filters at different scales and orientations allows multichannel filtering of an image to extract frequency and orientation information, therefore capturing local texture features. A Gabor function is a complex sinusoidal modulated by a rotated Gaussian. Filtering an image *I(x, y)* with Gabor filters $g_{mn}(x,y)$ designed according to [18] results in its Gabor wavelets transform W_{mn}:

$$W_{mn} = \int I(x,y) g^{*}{}_{mn}(x - x_1, y - y_1) dx_1 dy_1 \qquad (1)$$

Where "*" is the complex conjugate.

In this work we build a texture signature vector by filtering the image with a set of orientation and scale filters and calculating the mean and standard deviation of the output in frequency domain. We found that a filter with 3 scales and 4 orientations gives a best performance. A 24 dimensional vector is obtained to characterize Gabor filters.

3.4 Wavelets Moments

To characterize local texture, we use Haar wavelets Transform. Haar wavelet is the simplest orthogonal wavelet with compact support, so is the fastest for execution [19]. After a three level Haar transform, we use mean and standard deviation to extract wavelet moments. 20 dimensional wavelets moments are extracted.

4 Automated Relevance Feedback

4.1 Interest points extraction

The second stage of our approach consists of extracting parts representing retrieved images. To have a significant representation for the class of interest, we should select different parts that are specific to the species and capture the various variations across the species in question. For automatically selecting such parts, we have opted to extract interest points from retrieved images.

Interest points are defined as points where important variations occur. Such points characterize corners, junctions and locations with significant texture changes [20].

To detect those points, we have chosen Harris color detector [21]. The Harris color extractor has been proved to be the most stable in regard to image rotation, illumination changes, viewpoint changes and noise [22].The extraction of interest points is done by the following formula:

$$P(x,y) = Det(M(x,y)) - k.trace?\, M(x,y)) \qquad (2)$$

With k=0.04 , *Det* and *trace* are the determinant and the trace of the matrix M respectively. M is defined as follows:

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$$

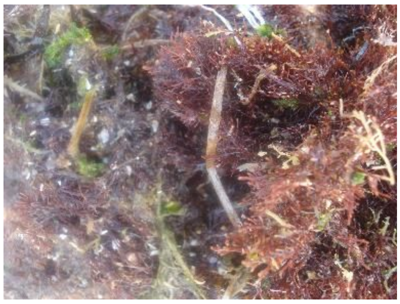
With:

$$M_{11}(x,y) = G_{\tilde{\sigma}} \otimes (r_x{}^2 + g_x{}^2 + b_x ? \qquad (3)$$

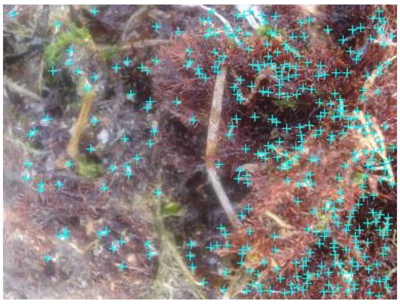
$$M_{12}(x,y) = G_{\tilde{\sigma}} \otimes (r_x r_y + g_x g_y + b_x b_y) \qquad (4)$$

$$M_{22}(x,y) = G_{\tilde{\sigma}} \otimes (r_y{}^2 + g_y{}^2 + b_y ? \qquad (5)$$

Where $G_{\tilde{\sigma}}$ is an isotropic Gaussian with variance $\tilde{\sigma}^2$ and (rx, gx, bx, ry, gy, by) represent the first order Gaussian derivatives in the RGB color space. These derivatives are implemented using a Gaussian with variance σ^2.



(a)



(b)

Fig. 2. Interest points extraction. (a). Gelidium image.
(b). Its interest points.

Figure 2 gives an example of the points of interests, identified on an underwater image of an area where the field of algae is largely dominated by gelidium sesquipedale.

The proposed method consists of applying Harris color detector on retrieved images, and extracting patches of size 13x13 pixels around each interest point.

As depicted in figure 2-b, several parts are perceptually very similar to each other. It's important then to abstract over these parts by merging similar patches. For this purpose and because of its simplicity, the K-means algorithm [23] is adopted to group similar parts into a few classes, each of which corresponds to a region of interest.

4.2 Images clustering using Self Organizing Maps

In this cycle, a neural network method has been introduced to achieve relevance feedback of the retrieved images without user interaction. In other terms, we implement an unsupervised learning network for the classification of relevance to assist the system. In this way, its retrieval performance can be enhanced automatically.
Region based representation of images is an effective way to improve accuracy in classification. Global features do not always represent salient objects seen in an image. Such features are computationally effective but provide rough representation of the image content. So, higher classification accuracy will be more effective with taking into account more precise information.

After Region features extraction, the image is represented by a number of regions of interest. Figure 1 shows that the image, in this stage, is represented by a two kind of features: global, which is the color histogram, and region-based features, are characterized by color and wavelet moments. This enables global and local image features to be integrated through a two-level Neural Network. Our method is based on the well-known Kohonen Self-organizing Feature Maps, or SOM [24]. The Kohonen technique creates a network that stores information in such a way that any topological relationships within the training set are maintained. SOMs provide a way of representing multidimensional data in much lower dimensional spaces. This process of reducing the dimensionality of vectors is essentially a data compression technique known as vector quantization.

Our purpose is to organize the retrieved images into a set of clusters such that similar images will fall in the same cluster. The SOM algorithm organizes a set of high dimensional vectors into a two dimensional map of neurons according to the similarities among the vectors. Applying the SOM algorithm to the retrieved images' feature vectors, allows us to perform a clustering process: similar images will belong to the same or neighboring neurons to build clusters.

Relevance identification is performed into two stages. First, we use The SOM's property of dimensionality reduction for dealing with region features. These features are processed by an unsupervised SOM in order to reduce the dimension of the input vector. To start the training process, SOM layer is trained by region inputs from all trained images. When training is carried out, each patch is associated with its best-matched neuron on the SOM.

The training SOM algorithm is summarized as follows:

Let $Xi = \{x_{in} / 1 \le n \le N\}$ a feature vector from the training vector set.

Step 1: Randomly select a training vector *Xi* .

Step 2: Find the winning node j with synaptic weight *wj* which is closest to *Xi* :

$$\|x_i - w_j\| = \min_{1 \le k \le J} \|x_i - w_k\|$$

Step 3: For every neuron l in the neighborhood of the winning node, update its synaptic weights by:

$$w_l(t+1) = w_l(t) + \alpha(t)(x_i(t) - w_l(t))$$

where $\alpha(t)$ is the learning rate which decreases with time, $0 < \alpha(t) < 1$.

Step4: Increase the time step t .

If $t \le T$, where T is the maximum training time,

Then halt the training process.

Else decrease the neighborhood size. go to Step 1.

Positions of the winnings nodes are combined with global features. The combined signatures of the images are then processed with the top level SOM. Therefore the bottom level SOM is used for region encoding while the top level SOM is used for relevance identification task. From another point of view, the retrieved images are encoded into a set of vectors. Each neuron in the SOM map is labeled by a list of images which are considered similar and are in the same cluster. Besides, images belonging to the same cluster as the query are deemed to be relevant and vice versa. Figure 3 demonstrates the processing of the automated relevance feedback.

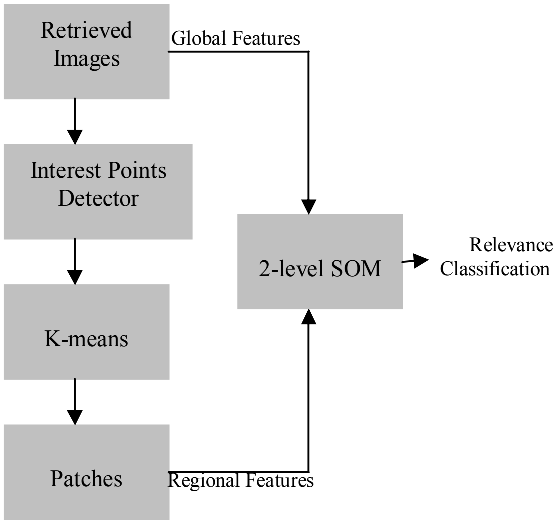


Fig. 3. Relevance Identification process

4.2 Similarity Learning

The positive and negative retrieved images are fed back to the similarity process, to enhance retrieval performance in the subsequent iteration. A nonlinear relevance feedback method using a Radial-Basis Function (RBF) is used [12] for image similarity. The used RBF is a single variable function with two controlling parameters: centres C and widths σ. The Gaussian shaped RBF is defined as:

$$f(x,c) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-c)^2}{2\sigma^2}\right) \tag{6}$$

The RBF process, uses the N_r relevant and N_n non relevant images' features vectors which are X_r and X_n respectively and updates these centers and widths thought relevance feedback at each iteration according to the following formulas:

$$c(t+1) = \bar{x}_r - \alpha_n(\bar{x}_n - c(t)) \tag{7}$$

$$\sigma = \exp(-\beta . Std(x_r)) \tag{8}$$

Where:

The parameter α_n is the negative weighting, β determines the overlapping factor, *Std* is the standard deviation function and $\bar{x}_r$, $\bar{x}_n$ are the mean of relevant and non relevant images' feature vectors respectively and they are defined as:

$$Std = \left(\frac{1}{N_r - 1}\sum_{i=1}(x_{ri} - \bar{x}_r)^2\right)^{1/2} \tag{9}$$

$$\bar{x}_r = \frac{1}{N_r}\sum_{1}^{N_r} x_{ri} \tag{10}$$

$$\bar{x}_n = \frac{1}{N_n}\sum_{1}^{N_n} x_{ni} \tag{11}$$

Centers in equation (7) denote the location of similar features and width in equation (8) determines the image feature deviation around the center.

The similarity measure between vector x and c is defined as:

$$s(x,c) = \sum_{i=1}^{p} f(x_i, c_i) \tag{12}$$

5. Discussion And Results

To validate the method we have described, we implemented an image retrieval system. We compare its retrieval performance, on the one hand, with an automated relevance Feedback with global descriptors where features are extracted over the whole image. On the other hand, with a non interactive technique and the combined relevance feedback method. Then we describe the experimental setup, the method of performance measurement, and the experimental results. The adopted method has been tested on an algae image database which a ground truth classification is available. The retrieved images are judged to be relevant if they are in the same class as the query. The algae image database contains more than 1000 algae images of various classes: Gelidium, Codium, Blidingia, … collected from Moroccan littoral.
Each image is indexed by 81dimensional feature vector used for retrieval. This vector includes HSV color

histogram and color moments for color descriptors, Gabor wavelet for texture.

For relevance identification, we use color histograms as global features, Color moments and wavelet moments as local features.

To obtain accurate precision in image retrieval, the learning process should exploit relevant patches in the image. First we changed the number of interest points per image by adjusting the threshold of color Harris detector. Results are presented in figure 4. The average performance of relevance identification increases when we take more interest points into account.

After, we extract a fixed number of interest points in the image. This number is picked out around 200 in order to find a good relationship between a good representation of the image and the processing time.

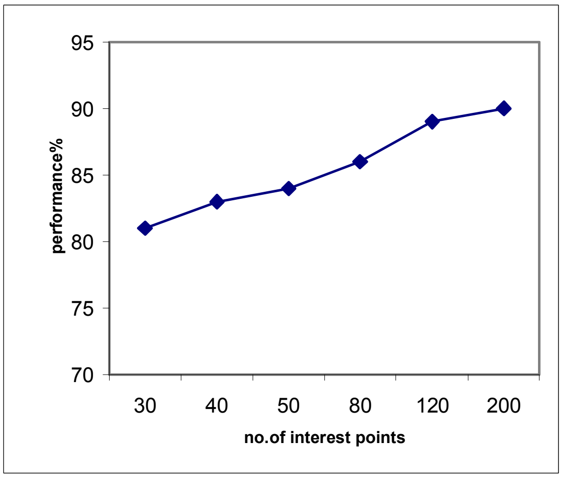
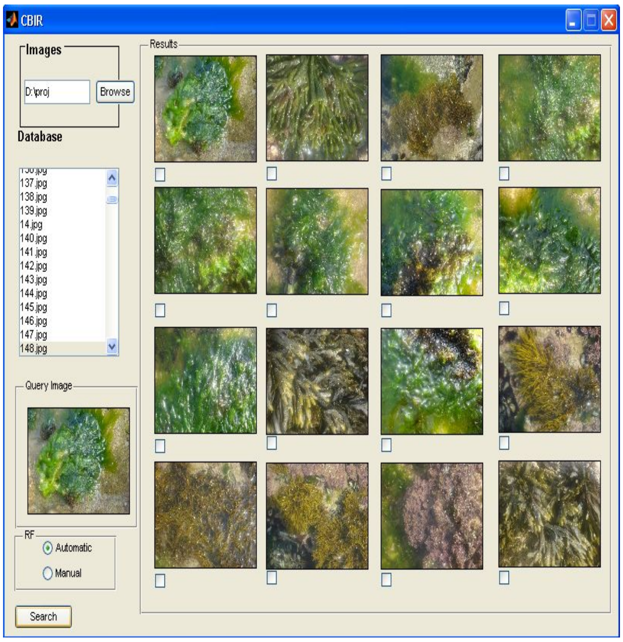


Figure 4: The performance of relevance identification as a function of the number of interest points

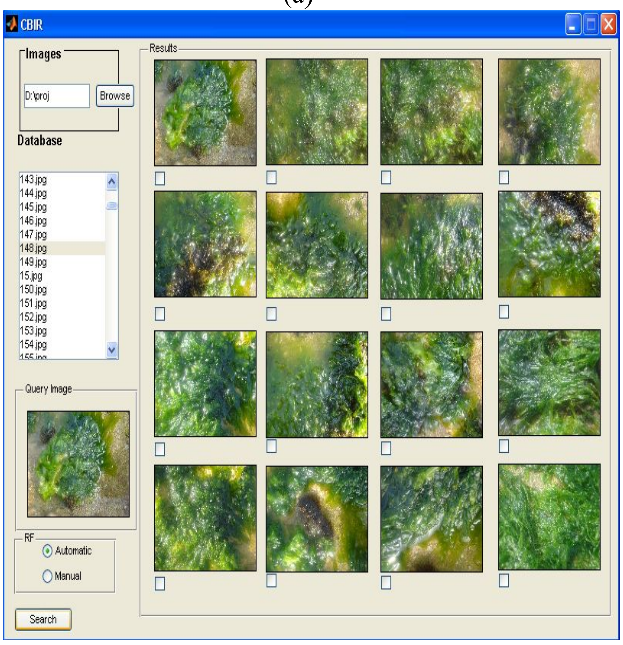
Two statistical metrics were used to evaluate system performance, based on two measures frequently used in information retrieval, namely *recall* and *precision*. Recall indicates the proportion of relevant images retrieved from the database when answering a query. For the sake of simplicity, we only computed the value of recall for the number of image that could fit into a display window of 16 images. Precision is the proportion of the retrieved images that are relevant to the query. A high value of precision therefore denotes that the top-ranked images are relevant.

Figure 5 illustrates retrieval examples with and without user relevance feedback. It shows that retrieval without user interaction has some difficulties to retrieve relevant images and demonstrates clearly that automated relevance feedback method outperforms the non-interactive method.

To show the performance of our approach, we have compared our approach with an automated relevance feedback with global descriptors [13]. The evaluation is done for 30 image queries. Results are summarized, in terms of precision and recall in figure 6.



(a)



(b)

Figure 5: Examples of retrieval results on the algae query image obtained by (a) non-relevance Feedback method,

the precision is: 50% and (b) the automated relevance Feedback method, the precision is:100%

In each experiment, one query image was randomly selected from the database and matched against the rest of the database. In the case of our approach, on average 75% of precision for a recall of 50%. This result shows that the proposed approach has better recall and precision performance than the other method. However, our method presents a slight increase in terms of consuming time.

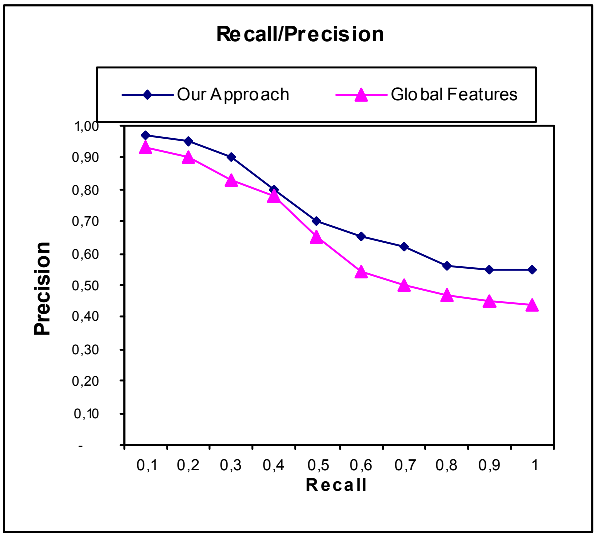


Figure 6: Precision and recall graph comparing two Automated Relevance Feedback.

Table 1: Average performances

Method	Recall	Precision at 100% recall
Combined RF	94.65%	81.35%
Automated RF	81.52%	67.18%
Without feedback	77.05%	46.08%

After automated relevance feedback retrieval, If the user wasn't satisfied, he will be invited to refine the results according to his interest need by choosing a combined relevance feedback mode. Table 1 exhibits the retrieval performances of the combined RF, Automated RF, and the non-relevance feedback method. Combined RF, presents, on average, 94.65% of the best 16 images displayed belong to the same class as the query, and the relevant images represents 81.45% of all the images required to be exhibited to not miss any suitable images. Furthermore, automated RF provides significant improvements compared to the non interactive method.

6. Conclusion

In this paper, we have introduced a novel framework for automated relevance feedback in algae image retrieval. We have proposed two level organizing maps that performs relevance classification of retrieved images. This allows avoiding at any iteration the human burden to label, relevant and non relevant images and to reduce the processing time. The use of interest points to extract regional features optimizes the learning process accuracy. Furthermore, it's showed that we may combine between the user subjectivity and the automatic process to minimize the number of iterations required to achieve a high retrieval performance.

The evaluation showed that the proposed approach gives good results according to our test image database.

References

[1] G. Salton and M.J. McGill, Introduction to Modern Information Retrieval. McGraw-Hill, 1983.

[2] R. Akallal , S.Alaoui, T.Givernaud et A.Mouradi « intoxications alimentaires associees aux efflorescences d'algues marines nocives » Revue Marocaine, 2001.

[3] Y. Rui, T.S. Huang, M. Ortega, and S. Mehrotra, "Relevance Feedback : A Power Tool in Interactive Content-Based Image Retrieval," IEEE Trans. Circuits and Systems for Video Technology,vol. 8, no. 5, pp. 644-655, Sept. 1998.

[4] H. Müller, W. Müller, D. M. Squire, S. Marchand-Maillet, and T. Pun,"Strategies for Positive and Negative Relevance Feedback in Image Retrieval," Comput. Vis. Group, Comput. Center, Univ. Geneva, Geneva,Switzerland, Tech. Rep., 2000.

[5] P. Muneesawang and L. Guan, "Multiresolution-histogram indexing for wavelet-compressed images and relevant feedback learning for image retrieval," in Proc. IEEE Int. Conf. on Image Processing, vol. 2, Vancouver, BC, Canada, 2000, pp. 526–529.

[6] S. Sclaroff, L. Taycher, and M. L. Cascia, "ImageRover : Acontent-based image browser for the world wide web," in Proc. IEEE Workshop on Content-Based Access of Image and Video Libraries, 1997, pp. 2–9.

[7] K. Porkaew; M. Ortega; S. Mehrota, "Query reformulation for content based multimedia retrieval in MARS" ICMCS'99, pp. 747 -751

[8] J. J. Rocchio Jr., "Relevance feedback in information retrieval,"in The Smart System—Experiments in Automatic Document Processing. Englewood Cliffs, NJ: Prentice-Hall, 1971, pp. 313–323.

[9] S. Tong and E. Chang, "Support vector machine active learning for image retrieval," presented at the ACM Multimedia Conf., Ottawa, ON, Canada, 2001.
[10] I. J. Cox, T. P. Minka, T. V. Papathomas, and P. N. Yianilos, "The Bayesian image retrieval system, PicHunter: Theory, implementation, and psychophysical experiments," IEEE Trans. Image Process., vol. 9, no. l, pp. 20–37, Jan. 2000.
[11] J. Laaksonen, M. Koskela, and E. Oja, "PicSOM: Self-organizing maps for content-based image retrieval," presented at the Int. Joint Conf. Neural Networks, Washington, DC, 1999.
[12] P. Muneesawang, and L. Guan, "Automatic machine interactions for content-based image retrieval using a self-organizing tree map architecture", IEEE Transactions on Neural Networks, vol. 13, no 4, pp. 821-834, 2002.
[13] K .Jarrah , P.Muneesawang, I.Lee and L.Guan "Minimizing Human-Machine Interactions in Automatic Image Retrieval " CCGEI may 2004.
[14] H.Tabout and A.Sbihi " Algae image indexing approach using color and texture features" SITA, November, 2006
[15] M. A. Stricker and M. Orengo. "Similarity of color images". In Proc. of the SPIE conference on the Storage and Retrieval for Image and Video Databases III, pages 381–392, 1995
[16] Q. Iqbal and J. Aggarwal. "Using structure in content-based image retrieval." In Int.Conf. Signal and Image Processing, pp. 129–133, Nassau, Bahamas, Oct. 1999.
[17] C.C Chen, "Filtering methods for texture discrimination" Pattern Recognition Letter 20(8) (1999) 783-90.
[18] B. Manjunath and W. Ma. "Texture features for browsing and retrieval of image data". IEEE Transactions on Pattern Analysis and Machine Intelligence, 18, August 1996.
[19] Q. Tian, N.Sebe, M.S Lew, E.Loupias, T.S. Huang "Image Retrieval using Wavelet-based Salient Points". Journal of Electronic Imaging, Vol. 10, No. 4, pp. 835-849, October, 2001.
[20] C.Harris et M.Stephens, "A combined corner and edge detector". Plessey Research Roke Manor, United Kingdom.1988.
[21] P. Montesinos, V. Gouet, and R. Deriche. "Differential invariants for color images". International Conference on Pattern Recognition, pages 838–840, 1998.
[22] V. Gouet and P. Montesinos, R. Deriche, D. Pelé, Evaluation de détecteurs de points d'intérêt pour la couleur, RFIA 2000.
[23] J.Macqueen, Some Methods of Classification and Analysis of multivariate observations. Proc. 5th Symp. Maths. Stat. Prob. 1976, pp.281-297
[24] T. Kohonen, "Self-Organizing Maps", 3rd ed. New York: Springer-Verlag, 2001, vol. 30, Springer Series in Information Sciences.

Hassan TABOUT is an engineer in a telecom company since 2002. He received the master and the Diploma of the higher Studies in Computer Engineering from the University Hassan Premier, Settat, Morocco in 1999 and 2001 respectively. Currently, he is a PhD student in the Telecommunication Systems and Decision Engineering Laboratory, Ibntofaïl University, Kenitra, Morocco. His research interests includes pattern recognition, content based image retrieval, interactive learning in multimedia systems.

Abdelmoghit SOUISSI is an engineer in a telecom company since 2001. He received the master and the Diploma of higher Studies in Computer Engineering from the University Hassan Premier, Settat, Morocco in 1999 and 2001 respectively. Currently, he is a PhD student in the Telecommunication Systems and Decision Engineering Laboratory, Ibntofaïl University, Kenitra, Morocco. His research interests have been oriented toward the pattern classification and content based image retrieval.

Abderrahmane SBIHI received the "License" of Science degree in physics from the University Mohammed V of Rabat, Morocco, in 1983 and the "Docteur ès Sciences" degree from the university Ibn Tofaïl of Kénitra, Morocco, in 1995. In 1985, he joined this last university where he worked as an Assistant Professor. From 1992 to 1995, he joined the University of Science and Technology of Lille, France, as a Guest Assistant Professor at the Institute of Technology and as a Research Assistant in the Automatic Control Center where he worked on the adaptation of numerical mathematical morphology to pattern classification problems. In 1995, he returned to the university Ibn Tofaïl where he is currently Professor and Head of the Telecommunication Systems and Decision Engineering Laboratory (LASTID). He is member of governing board of the IAPR since 2004. His current research interests include computer vision, pattern recognition and multidimensional data analysis.