2008 20th IEEE International Conference on Tools with Artificial Intelligence

Application of Relevance Feedback in Content Based Image Retrieval Using Gaussian Mixture Models

Apostolos Marakakis[1], Nikolaos Galatsanos[2], Aristidis Likas[2], and Andreas Stafylopatis[1]
[1]*School of Electrical and Computer Engineering,*
National Technical University of Athens, 15780 Athens, Greece
[2]*Department of Computer Science, University of Ioannina, 45110 Ioannina, Greece*
amara@central.ntua.gr, {galatsanos, arly}@cs.uoi.gr, andreas@cs.ntua.gr

Abstract

In this paper a relevance feedback (RF) approach for content based image retrieval (CBIR) is described and evaluated. The approach uses Gaussian Mixture (GM) models of the image features and a query that is updated in a probabilistic manner. This update reflects the preferences of the user and is based on the models of both positive and negative feedback images. Retrieval is based on a recently proposed distance measure between probability density functions (pdfs), which can be computed in closed form for GM models. The proposed approach takes advantage of the form of this distance measure and updates it very efficiently based on the models of the user specified relevant and irrelevant images. For evaluation purposes, comparative experimental results are presented that demonstrate the merits of the proposed methodology.

1. Introduction

The target of content-based image retrieval (CBIR) is to retrieve relevant images from an image database based on the similarity of their visual content with one or more query images. These query images are submitted by the user as examples of his/her preferences. Then, the CBIR system ranks the database images and displays the retrieved results ordered with respect to their similarity with the query images. Most CBIR systems, e.g. [1]–[8], model each image using a combination of low-level features, and then define a distance metric that is used to quantify the similarity between image models. Nevertheless, low-level image features cannot always capture the human perception of image similarity. In other words, it is difficult using only low-level image features to describe the semantic content of an image. This is known in the CBIR community as the *semantic gap* problem [11].

Relevance feedback (RF), has been proposed as a methodology to ameliorate this problem, e.g. [1]-[3] and [6]-[8]. RF attempts to insert the subjective human perception of image similarity into a CBIR system. Thus, RF is an interactive process that refines the distance between the query and the database images through interaction with the user and taking into account his/her preferences. To accomplish this, during a round of RF, users are required to rate the relevance of the retrieved images according to their preferences. Then, the retrieval system updates the matching criterion based on the user's feedback, e.g. [1]–[3], [6]–[8], [15] and [16].

In what concerns RF approaches proposed in literature, there is much work which has been done during the last years that can be classified in two main categories. The first category concerns learning-based methods, i.e. it includes the methods which are based on some learning model (usually SVMs) in order to train a classifier to distinguish between the positive and negative feedback examples, e.g. [28], [6], [26]. The main drawback of the learning-based approach is that for every feedback round a new classifier must be trained taking into account both the previously presented examples and the new ones presented in the last feedback round.

The second category of RF methods (model-based methods) includes those that attempt to model the statistical distribution of feedback examples in the feature space. These methods can be further divided in two subcategories.

The first subcategory includes methods that make the assumption that the feedback examples form one cluster in feature space. The cornerstone of such methods is MindReader [1]. Other methods that work under this assumption are presented in [12], [16], [34], [35]. The single cluster assumption which is made by these methods is usually very restrictive even for the set of positive examples. Moreover, the negative feedback examples cannot be taken into account because they naturally spread across different semantic categories, so it cannot be claimed that they form one cluster. Nevertheless, in

1082-3409/08 $25.00 © 2008 IEEE
DOI 10.1109/ICTAI.2008.110

[35] a solution to this problem is proposed based on a "two-step" retrieval approach. In the first step, only the positive examples are used to determine a reduced set of database images very similar to them. In the second step these images are re-ranked based on both the positive and the negative examples, in order for the final ranking to be produced.

The second subcategory of model-based RF techniques includes methods which assume that the feedback examples (either positive or negative) form more than one clusters, e.g. [17], [29], [30], [31], [32]. The methods which are based on Gaussian mixture models belong to this category.

Gaussian mixtures (GM) are a well-established methodology to model probability density functions (pdf). The advantages of this methodology such as adaptability to the data, modeling flexibility and robustness have made GM models attractive for a wide range of applications, e.g. [17], [18]. Furthermore, GM models have been employed for the CBIR problem, e.g. [4], [14] and [17]. The main challenge when using a GM model in CBIR, is to define a distance metric between GMs that separates well different models, and, in addition, it can be computed efficiently.

In [21] a distance measure between pdfs, called *C2 divergence*, has been introduced. This measure can be computed in closed form for GM models and can constitute the basis for an efficient use of GMs in CBIR. In this context, we propose a RF technique, which relies on a suitable and intuitive update of the query model using the relevance of the retrieved images. Moreover, an effective strategy is devised, that requires very few computations to incrementally update the distance metric after each RF round. A preliminary version of this work has been presented in [22]. In the present paper, we further elaborate on this approach and provide an in-depth experimental study for performance assessment of the proposed method. The experimental evaluation is based on comparative results on large-scale data sets using an enriched set of image features.

The rest of the paper is organized as follows: in Section 2 we describe GMs in the context of image modeling for CBIR Furthermore, we define the C2 divergence as a promising alternative distance metric for GMs, we analyze its properties and we present the proposed C2-based RF scheme. In Section 3 we provide the details and the results of the experiments. Finally, in Section 4 we present conclusions and directions for future research.

2. Relevance feedback based on Gaussian Mixture models in image retrieval

GM models have been used extensively in many data modeling applications. Using them for the CBIR problems allows us to bring to bear several powerful features of the GM modeling methodology, such as modeling flexibility and easy training, that make it attractive for a wide range of applications, e.g. [18], [19]. GM models have been used previously for CBIR, e.g. [4], [14], as probability density models of the features that are used to describe images. In this framework, each image is described as a bag of feature vectors which are computed locally (e.g. one feature vector for each pixel or each region of the image). This bag of feature vectors is subsequently used to train (in a maximum likelihood manner) a GM that models the probability density of the set of image features. A GM model for the image feature vectors $x \in R^d$ is defined as

$$p(x) = \sum_{j=1}^{K} \pi_j \phi(x \mid \theta_j) \tag{1}$$

$$\phi(x \mid \theta_j = [\mu_j, \Sigma_j]) = \frac{1}{\sqrt{(2\pi)^d |\Sigma_j|}} e^{-\frac{1}{2}(x-\mu_j)^T \Sigma_j^{-1} (x-\mu_j)} \tag{2}$$

where K is the number of Gaussian components in the model, $0 \leq \pi_j \leq 1$ the mixing probabilities with $\sum_{j=1}^{K} \pi_j = 1$, and $\phi(x \mid \theta_j)$ a Gaussian pdf with mean μ_j and covariance Σ_j.

In order to describe the similarity between images in this context, a distance metric must be defined. The Kullback-Leibler (KL) divergence [10] is the most commonly used distance metric between pdfs. However, this distance measure cannot be computed in closed form for GMs. Thus, one has to resort to time consuming random sampling Monte Carlo methods, which make its use impractical for CBIR. As far as RF is concerned, the situation is even worse, because RF is based on the online interaction with the user, which demands rapid distance update at every RF round. In order to overcome these difficulties, some alternatives have been proposed.

For example, in [14] the Earth Movers Distance (EMD) metric between GMs was proposed. This metric is based on considering the probability mass of one GM as piles of earth and of the other GM as holes in the ground and then finding the least work necessary to fill the holes with the earth in the piles. The EMD is an effective metric for CBIR, however it cannot be computed in closed form and requires the solution of a linear program each time that must be computed. Thus it is slow and cumbersome to use for RF were the query changes after each RF epoch.

2.1. The C2 divergence

In order to ameliorate these difficulties a new distance metric was proposed in [21]. This metric between two pdfs $p_1(x)$ and $p_2(x)$ is defined as

$$C2(p_1,p_2) = -\log\left(\frac{2S_{p_1p_2}}{S_{p_1p_1}+S_{p_2p_2}}\right) \quad (3)$$

with

$$S_{p_mp_l} = S_{ml} = \int p_m(x)p_l(x)dx \quad (4)$$

and it can be computed in closed form when $p_1(x)$ and $p_2(x)$ are GMs. In this case we have:

$$S_{ml} = \frac{1}{\sqrt{(2\pi)^d}}\sum_{i=1}^{K_m}\sum_{j=1}^{K_l}\frac{\pi_{mi}\pi_{lj}}{\sqrt{|C_{ml}(i,j)|e^{k_{ml}(i,j)}}} \quad (5)$$

where

$$C_{ml}(i,j) = \Sigma_{mi} + \Sigma_{lj} \quad (6)$$

$$k_{ml}(i,j) = (\mu_{mi}-\mu_{lj})^T C_{ml}^{-1}(i,j)(\mu_{mi}-\mu_{lj}) \quad (7)$$

π_{mi} is the mixing weight of the i-th Gaussian kernel of p_m, μ_{mi},Σ_{mi} are means and covariance matrices respectively of the kernels of the Gaussian mixture p_m, d is the dimension of the feature vector x and K_m is the number of Gaussian components in p_m.

For the C2 divergence the following properties hold:

1. $C2(p_1,p_2) = -\log\left[\frac{2\int p_1(x)p_2(x)dx}{\int p_1^2(x)dx+\int p_2^2(x)dx}\right] \geq 0$
2. $C2(p_1,p_2) = 0 \Leftrightarrow p_1(x) = p_2(x)$
3. $C2(p_1,p_2) = C2(p_2,p_1)$
4. The triangular inequality
   $C2(p_1,p_3) \leq C2(p_1,p_2) + C2(p_2,p_3)$
   does not hold, therefore C2 is not a metric as is also the case with the KL divergence.

In what concerns the relation between the $C2$ and the symmetric KL divergence

$$SKL(p_1,p_2) = \frac{1}{2}[KL(p_1 \| p_2) + KL(p_2 \| p_1)] \quad (8)$$

it can be proved that, for arbitrary pdfs p_1 and p_2, the difference between the SKL and the C2 is bounded by:

$$SKL(p_1,p_2) - C2(p_1,p_2) \leq \frac{1}{2}[\Delta_1(p_1)+\Delta_2(p_2)] + \log 2 \quad (9)$$

where

$$\Delta_i(p_i) = E[\log p_i]_{p_i} - \log E[p_i]_{p_i} \geq 0 \quad (10)$$

due to the Jensen inequality.

Furthermore, in what concerns the relation between the C2 and the norm L_2, it is straightforward to show that:

$$C2(f,g) = -\log\left(1 - \frac{L_2(f-g)}{L_2(f)+L_2(g)}\right) \quad (11)$$

where $L_2(f) = \int f^2(x)dx$ is the definition of the norm L_2 for real continuous functions. Given that $\log(x) \approx x-1$ for $x \approx 1$

$$C2(f,g) \approx \frac{L_2(f-g)}{L_2(f)+L_2(g)} \quad (12)$$

when $L_2(f-g) << L_2(f)+L_2(g)$

2.2. Relevance feedback based on C2

Assume we have a query q modeled as $q(x)$ (e.g. a GM model), and that the i-th database image is modeled by $i(x)$ for $i = 1,...,N$. The search based on this query requires the calculation of a $N \times 1$ table with the values $C2(q,i)$. Also assume that from the retrieved images the user decides that the images with models $r_m(x)$, $m = 1,2\ldots M$, are the most relevant and desires to update his/her query based on them. One simple and intuitive way to go about it is to generate a new query model given by

$$q'(x) = (1-\Lambda)q(x) + \sum_{m=1}^{M}\lambda_m r_m(x) \quad (13)$$

where $0 \le \lambda_m \le 1$, $\sum_{m=1}^{M} \lambda_m = \Lambda$, $0 \le \Lambda \le 1$, λ_m is the relevance assigned by the user to image r_m and $1-\Lambda$ is the weight of the contribution of the previous query to the formation of the new query. The attractive feature of the model in Eq. (13) is that relevance λ_m is consistent with the probabilistic framework that is used. It has a physical meaning; it is proportional to the relevance degree assigned by the user to image r_m and this defines a "composite model" that incorporates the preferences of the user.

User selected irrelevant images can also be incorporated, as negative feedback, into the RF process. We can thus define, in a way similar to Eq. (13), an updated query model $n'(x)$ for the irrelevant images, which we call negative query:

$$n'(x) = (1-\Lambda_n)n(x) + \sum_{m=1}^{M_n} \lambda_m^n r_m^n(x) \qquad (14)$$

where n, n' correspond to the previous and new negative query respectively, Λ_n, λ_m^n are analogous to the previously mentioned Λ, λ_m and r_m^n, $m = 1,2,\ldots,M_n$ are the negative examples. The negative query is initially "empty", contrary to the positive one which includes the initial query selected by the user.

Furthermore, it is desirable to efficiently compute the distances between the database image models $i(x)$, $i = 1,2\ldots N$, and the new query models $q'(x)$ and $n'(x)$. Taking into account Eq. (3)-(4) and (13), the updated distance measure for the new query q' is given by:

$$C2(q',i) = -\log\left(\frac{2S_{q'i}}{S_{q'q'} + S_{ii}}\right) \qquad (15)$$

$$S_{q'i} = (1-\Lambda)S_{qi} + \sum_{m=1}^{M} \lambda_m S_{r_m i} \qquad (16)$$

$$S_{q'q'} = (1-\Lambda)^2 S_{qq} + 2(1-\Lambda)\sum_{m=1}^{M} \lambda_m S_{qr_m} + \sum_{m=1}^{M}\sum_{m'=1}^{M} \lambda_m \lambda_{m'} S_{r_m r_{m'}} \qquad (17)$$

Similar equations hold in the case of the negative query distance update. To obtain these equations, the only thing to do is to replace $q, q', \lambda_m, \Lambda, M, r_m$ by $n, n', \lambda_m^n, \Lambda_n, M_n, r_m^n$, respectively, in Eq. (15)-(17).

From the above equations, it is obvious that the computation of the distance between $q'(x)$, $n'(x)$ and the database image models is very fast. This constitutes a notable advantage of our method. Indeed, computing the distance for the new query involves only rescaling operations. Actually the new query models in Eq. (13)-(14) do not need to be constructed. In order to update the distance between the new query and the database images only the computations of Eq. (15)-(17) are needed, and those computations only implicitly involve the new query model.

The best images to retrieve can be found by combining both positive and negative RF. This can be done by minimizing the following distance metric:

$$c(i) = a_{pos} C2(q,i) - (1-a_{pos})C2(n,i) \qquad (18)$$

with $0 \le a_{pos} \le 1$ being the relative weight given to the positive feedback. After computing the metric $c(i)$ for every database image, we can retrieve the images with the lowest values for this metric. Such images will be similar to the positive examples and dissimilar to the negative examples.

3. Experiments

In this work, we propose the use of GMs, in order to model the database images, along with the RF scheme based on the C2 divergence described in subsection 2.2. In order to demonstrate the merits of this method we conducted a number of experiments.

For a comparative evaluation of the proposed RF methodology we have implemented the method described in [35]. This method incorporates both positive and negative examples using a "two step" retrieval scheme. In the first step, only the positive examples are used in order to rank the database images. Then, only a relatively small number of the database images, which are very similar to the positive examples, are retained and re-ranked, in the second step which takes into account both the positive and the negative examples, to produce the final ranking from which the top images, placed near to the positive and far from the negative examples, are presented to the user. Both retrieval steps are based on Lagrange optimization in order to estimate the distance parameters that minimize the within class distance and maximize the between class distance of the positive and the negative examples.

3.1. Image databases and features

In order to test the validity of the proposed approach we used two image databases: The first image set (DB I) contains 3740 annotated 640x480 color images from the image database in [33]. These images have been classified into 17 semantic categories according to their content (e.g. airplanes, cars, birds, windows etc). Generally, we adopted the categorization specified by the database provider, except for a few cases where categories which are semantically very close to each other were merged. The second database (DB II) is a subset of the Corel image database. It contains 9923 images of size 256x384 that are classified into 42 semantic categories. Although the Corel database is professionally annotated and categorized, many images containing the same semantic content are distributed across different Corel categories. Hence, we decided to merge some Corel categories thus producing our own semantic categorization, which is considered as ground truth.

For each image pixel in the aforementioned databases a set of features has been extracted that are of three types:

1) *Position*: the *(x,y)*-coordinates of each pixel normalized by the image width for *x*-dimension and image height for *y*-dimension.
2) *Color*: we chose the CIE-Lab [14] color space as being approximately perceptually uniform, a property very useful for retrieval, thus 3 color features (L*, a*, b*) per pixel were used. In order to take the final values for these features, local Gaussian smoothing is performed according to the texture scale which is mentioned below.
3) *Texture*: as a scheme to extract texture features from the images we selected the one presented analytically in [4]. One measure of local texture scale and the anisotropy (*A*), the polarity (*P*) and the contrast (*C*) corresponding to this scale, are estimated for each pixel. Taking into account that the polarity and anisotropy values are meaningless in regions of low contrast, the selected texture features are $AC = A*C$, $PC = P*C$, and C.

Thus, finally, we have an 8-dimensional feature vector for each image pixel with

$$\begin{aligned} x &= (f_1, f_2, \ldots, f_8) \\ &= \left(x^{norm}, y^{norm}, (L*)^{smoothed}, (a*)^{smoothed}, (b*)^{smoothed}, AC, PC, C\right) \end{aligned} \tag{19}$$

$$x^{norm} = \frac{x}{x_{\max}},\ y^{norm} = \frac{y}{y_{\max}} \tag{20}$$

Prior to feature extraction, pixel sub-sampling was performed, using a spatially uniform grid. In DB I only the 15% of the image pixels was used for feature extraction and GM training. In DB II 50% of image pixels was used, because the images are of lower resolution. Prior to sub-sampling, the images were smoothed using a Gaussian kernel in order to avoid aliasing.

3.2. Implementation issues

Regarding our method, the parameters of the GM model of each image were estimated using a variation of the well-known EM algorithm, called the Greedy EM [23]. This algorithm avoids the problem of parameter initialization, which is critical for the normal EM. In all the experiments, we chose to use full covariance parameterization for the GM components. We also selected empirically to use 10 components per GM model. Before applying the EM algorithm, we normalize each feature to have zero mean and unit standard deviation, given the set of all feature vectors in the image.

The parameters λ_m, λ_m^n for the positive and negative examples are given equal values regardless of *m*, because for the simulations described analytically in the next subsection, the ground truth of the aforementioned pre-categorized databases is used in order to select the positive and negative examples. Thus, using the strict ground truth categorization, it is meaningless to assume different relevance degrees for each example. The weight of the previous positive (negative) query, $1-\Lambda$ ($1-\Lambda_n$), is given values proportional to the number of the positive (negative) examples given by the user until the current RF round.

In order to extract features appropriate for the method proposed in [35], we partitioned each database image in 9 sub-images dividing each of the *x* and *y* image axes in 3 equal width intervals, and for each sub-image we estimated a $4\times8\times8\left((L*)^{smoothed}-(a*)^{smoothed}-(b*)^{smoothed}\right)$ color histogram, and a $4\times4\times4\left(AC-PC-C\right)$ texture histogram. Thus, we get an image description of $I=18$ feature vectors, 9 256-dimensional color histogram vectors and 9 64-dimensional texture histogram vectors.

3.3. Simulations and results

In order to quantify the performance of the proposed RF system we have resorted to relevance feedback simulation. In this simulation, a percentage of the images of each database category are selected to form a query set. Each image in this set is used as initial query. The simulation scheme is similar to that proposed in [16]. The accuracy is measured as the ratio of relevant images (determined according to the ground truth) among the top *T* retrieved images. At each feedback step, at most K_p relevant images are selected from the top *M* retrieved images, as positive feedback. If the number of relevant images in the top M retrieved images is greater than K_p, then we select randomly K_p of them as feedback, else we select all the relevant retrieved images. In case we wish to provide the system with negative feedback we follow a similar procedure providing as feedback at most K_n of the non-relevant images retrieved in the first M retrieved images. We provide experiments with this methodology in both databases. In DB I each image was used once as the initial query whereas in DB II the 40% of the database images (a percentage sufficient from the statistical point of view) were used as initial queries. The average accuracy was computed in databases for all images in the set of initial queries and for each RF round. In both databases, for reasons of comparison we made several choices regarding the RF method and the RF parameters. The results are given in Figures 1, 2. The choices for the simulation parameters were: $T = 20,\ M = 150,\ K_p = 10,\ K_n = 10$. When negative feedback was provided, $a_{pos} = 0.65$. For the method proposed in [35], when negative feedback was provided, the $M\ (=150)$ top ranked images in the first step of retrieval were retained in the second step of retrieval. In Figures 1, 2 with "*GMM*" we denote our method and with "*Meth[35]*" the method proposed in [35]. Furthermore, "p" indicates the use of positive feedback and "n" indicates the use of negative feedback.

As it can be easily observed from Figures 1, 2, when compared with the method proposed in [35], our method almost always results in a performance improvement after the initial (one or two) RF rounds. When negative feedback is used, our method always results in significantly higher levels of accuracy. This is an assessment of the advantages of the proposed RF scheme, which easily, efficiently and intuitively incorporates both the positive and negative examples, and justifies the use of GMs for image modeling.

To provide an estimate of the retrieval time of the proposed method, it can be noted that for DB I and for the previously described simulation scheme (with both positive and negative feedback and with the simulation parameters having values as specified above) it takes about 166 sec of computation time (on a 3 GHz PC using Matlab) to execute for all 3740 images presented as initial queries, the initial retrieval plus six epochs of RF. This means that the average retrieval time per image query with 6 rounds of RF is 0.044 sec. On the other hand using the same simulation scenario, the method proposed in [35] takes about 19000 sec for all 3740 images used as initial queries which means an average of about 5 sec per query. This significant difference in time complexity, demonstrates the noticeable ability of our RF scheme to update rapidly the distances between the database images and the new query and to provide instantaneous response to the user.

4. Conclusions

This work focuses on the evaluation of a probabilistic framework for relevance feedback based on GM models. The main advantages of the proposed methodology are accuracy as indicated by our simulation study results, speed of implementation and flexibility. Incorporation of both positive and negative feedback examples is performed in a very intuitive manner which, in combination with the simple and easy to update form of the ranking measure *C2*, allows for real time evaluation of the image ranking criterion. In this way, fast retrieval is achieved after user feedback has been provided.

Furthermore, we compared our method with the method proposed in [35] which incorporates both positive and negative feedback examples. The performance of our method is clearly superior, especially when both positive and negative feedback is used.

In the future we intend to provide the user with the possibility to determine explicitly the degree of relevance of the feedback examples, by implementing a sophisticated interactive system based on GMs and using our RF scheme. In addition, we aim to generalize our RF scheme to support region-based similarity and retrieval. Furthermore, we aim to attempt to apply techniques for determining automatically the appropriate number of kernels of each mixture model. Moreover, we plan to test the performance of our method using more sophisticated image features. Finally, we would like to test the scalability of the proposed method using even larger image databases.

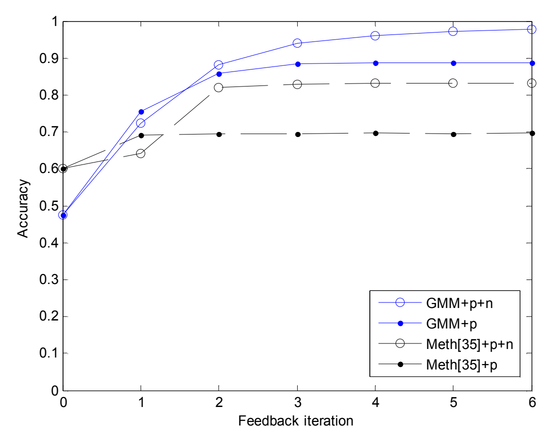

Figure 1. Average accuracy in scope T = 20 during different rounds of relevance feedback (DB I)

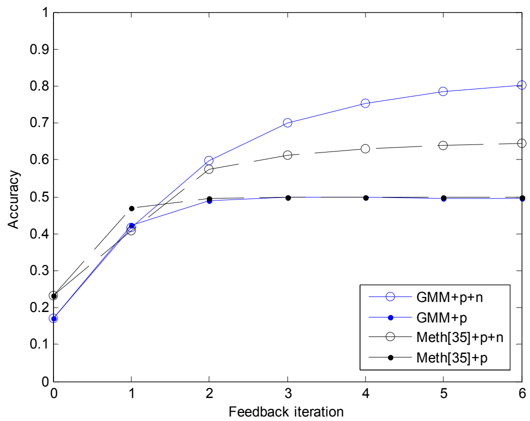

Figure 2. Average accuracy in scope T = 20 during different rounds of relevance feedback (DB II)

Acknowledgement. This work was partly supported by Public Funds under the PENED 2003 Project. The Project is co-funded by the European Social Fund (80%) and National Resources (20%) from the Hellenic Ministry of Development - General Secretariat for Research and Technology.

5. References

[1] Y. Ishikawa, R. Subramanya, and C. Faloutsos, “MindReader: Querying databases through multiple examples”, *Proc. VLDB Conf.*, 1998.

[2] I. J. Cox, M. L. Miller, T. P.Minka, T. V. Papathomas, and P. N. Yianilos, “The Bayesian image retrieval system, PicHunter: Theory, implementation, and psychophysical experiments”, *IEEE Trans. Image Processing.*, vol. 9, no. 1, pp. 20–37, Jan. 2000

[3] Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra, “Relevance feedback: A power tool for interactive content-based image retrieval”, *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 5, pp. 644–655, Sep. 1998.

[4] C. Carson, S. Belongie, H. Greenspan, and J. Malik, “Blobworld: Image segmentation using expectation-maximization and its application to image querying”, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 8, pp. 1026–1038, Aug. 2002.

[5] Y. Chen and J. Z. Wang, “A region-based fuzzy feature matching approach to content-based image retrieval”, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 9, pp. 1252–1267, Sep. 2002.

[6] G. D. Guo, A. K. Jain, W. Y. Ma, and H. J. Zhang, “Learning similarity measure for natural image retrieval with relevance feedback”, *IEEE Trans. Neural Networks.*, vol. 13, no. 4, pp. 811–820, Jul. 2002.

[7] G. Aggarwal, T. V. Ashwin, and S. Ghosal, “An image retrieval system with automatic query modification”, *IEEE Trans. Multimedia*, vol. 4, pp. 201–214, Jun. 2002

[8] F. Jing, M. Li, H. J. Zhang, and B. Zhang, “An efficient and effective region-based image retrieval framework”, *IEEE Trans. Image Processing*, vol. 13, no. 5, pp. 699–709, May 2004.

[9] B. S. Manjunath, J. R. Ohm, V. V. Vasudevan, and A. Yamada, “Color and texture descriptors”, *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 6, pp. 703–715, Jun. 2001.

[10] M. N. Do and M. Vetterli, “Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance”, *IEEE Trans. Image Processing.*, vol. 11, no. 2, pp. 146–158, Feb. 2002.

[11] Y. Rui, T. S. Huang, and S. F. Chang, “Image retrieval: Current techniques, promising directions, and open issues”, *J. Vis. Commun. Image Represen.*, vol. 10, pp. 39–62, Mar. 1999.

[12] Z. Su, H. Zhang, S. Li, and S. Ma, “Relevance feedback in content-based image retrieval: Bayesian framework, feature subspaces, and progressive learning”, *IEEE Trans. Image Processing*, vol. 12, no. 8, pp. 924–937, Aug. 2003.

[13] X. He, O. King, W. Y. Ma, M. Li, and H. J. Zhang, “Learning a semantic space from user’s relevance feedback for image retrieval”, *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 1, pp. 39–48, Jan. 2003.

[14] H. Greenspan, G. Dvir, and Y. Rubner, “Context-dependent segmentation and matching in image databases”, *Computer Vision and Image Understanding*, vol. 93, pp. 86–109, 2004.

[15] El. Naqa, Y. Yang, and N. Galatsanos and M. Wernick. “A Similarity Learning Approach to Content Based Image Retrieval: Application to Digital Mammography”, *IEEE Transactions on Medical Imaging*, vol. 23 , no. 10, pp. 1233-1244, Oct. 2004.

[16] C.T. Hsu, and C. Y. Li, “Relevance Feedback Using Generalized Bayesian Framework With Region-Based Optimization Learning”, *IEEE Trans. on Image Processing.*, vol. 14, no. 10, pp. 1617-1631, October 2005.

[17] F. Qian, M. Li, L. Zhang, H. J. Zhang, and B. Zhang, “Gaussian mixture model for relevance feedback in image retrieval”, *Proc. IEEE ICME*, Aug. 2002.

[18] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford Univ. Press Inc., New York, 1995.

[19] G. M. McLachlan and D. Peel, *Finite Mixture Models*, John Wiley & Sons Inc., New York, 2001.

[20] A. D. Bimbo, *Visual Information Retrieval*, San Mateo, CA: Morgan Kaufmann, 1999.

[21] G. Sfikas, C. Constantinopoulos, A. Likas and N. P. Galatsanos, “An analytic distance metric for Gaussian mixture models with application in image retrieval”, *Proc. International Conference Artificial Neural Networks (ICANN 2005)*, Warsaw, Poland, 2005.

[22] A. Marakakis, N. Galatsanos, A. Likas and A. Stafylopatis, “A Relevance Feedback Approach for Content Based Image Retrieval Using Gaussian Mixture Models”, *Proc. International Conference Artificial Neural Networks* (*ICANN 2006*), Athens, Greece, September 2006, LNCS 4132, pp. 84-93, Springer-Verlag.

[23] N. Vlassis, A. Likas, “A greedy EM algorithm for Gaussian mixture learning”, *Neural Processing Letters*, vol. 15, pp. 77-87, 2002.

[24] Jason V. Davis, Inderjit Dhillon, “Differential Entropic Clustering of Multivariate Gaussians”, *NIPS*, 2006

[25] N. Vasconcelos, A. Lippman, "Bayesian relevance feedback for content-based image retrieval", *Content-based Access of Image and Video Libraries,* 2000.
[26] S. Tong, E. Chang, "Support vector machine active learning for image retrieval", *ACM Multimedia,* 2001.
[27] N. Vasconcelos, "On the Efficient Evaluation of Probabilistic Similarity Functions for Image Retrieval", *IEEE Transactions on Information Theory,* vol. 50(7), pp. 1482-1496, July 2004.
[28] F. Jing, M. Li, H-J. Zhang, and B. Zhang, "Relevance Feedback in Region-Based Image Retrieval", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 5, May 2004.
[29] M. L. Kherfi and D. Ziou, "Relevance Feedback for CBIR: A New Approach Based on Probabilistic Feature Weighting With Positive and Negative Examples", *IEEE Transactions on Image Processing*, vol. 15, no. 4, April 2006.
[30] 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, July 2002.
[31] N. Vasconcelos and A. Lippman, "Learning over multiple temporal scales in image databases", *European Conference in Computer Vision,* Dublin, Ireland, 2000.
[32] N. Vasconcelos and A. Lippman, "Learning from user feedback in image retrieval systems", *Advances in Neural Information Processing Systems*, 1999.
[33] Microsoft Research Cambridge Object Recognition Image Database, version 1.0. http://research.microsoft.com/research/downloads/Details/b94de342-60dc-45d0-830b-9f6eff91b301/Details.aspx
[34] Rui, Y., Huang, T.S., "Optimizing learning in image retrieval", *IEEE International Conference on Computer Vision and Pattern Recognition*, Hilton Head, SC, USA, 2000.
[35] M. L. Kherfi, D. Ziou and A. Bernardi, "Combining positive and negative examples in relevance feedback for content-based image retrieval", *Journal of Visual Communication and Image Representation*, vol. 14, pp. 428-457, December 2003.
[36] Ritendra Datta, Jia Li, James Ze Wang, "Content-based image retrieval: approaches and trends of the new age", *Multimedia Information Retrieval*, pp. 253-262, 2005.