Active Learning Applied To Image Retrieval

Neelam Longani[1]
M.E. (Electronics) Student,
Jawaharlal Nerhu engineering college, Aurangabad.

Prof. Faisal Shaikh[2]
Department of Electronics and Telecommunication,
Jawaharlal Nerhu engineering college, Aurangabad.

Abstract

Active learning in content-based image retrieval systems is a challenging task. For effective retrieval of visual information, statistical learning plays a pivotal role. Active learning methods have been considered with increased interest in the statistical learning community. This paper provides algorithms within a statistical framework to extend active learning for content-based image retrieval (CBIR). The classification framework is presented with experiments to compare several powerful classification techniques in this information retrieval context. Focusing on interactive methods, active learning strategy is then described. Active learning in such context optimizes the training data to get best classification with as few as possible user labeling. Experimental results on Corel database show the comparison of 3 classification strategies to evaluate active learning contribution for online CBIR and the improvement in retrieval performance by active learning approach.

1. INTRODUCTION

Content-based image retrieval has been a vigorous area of research for at least the last two decades. Contrary to the early systems, which focused on fully automatic strategies, recent approaches have introduced human-computer interaction [1], [2]. In this paper, we focus on the retrieval of concepts within a large image collection. We assume that a user is looking for a set of images, the query concept, within a database. The aim is to build a fast and efficient strategy to retrieve the query concept. Content-based image retrieval (CBIR) allow the users to submit a query by an image sample, namely, query_by_example, such that the database images that are "similar in content" to the query image are retrieved. Since end users, in general, do not know the make-up (kinds of images) of the image database and the content representation and search techniques used in the environment (what types of features and indexing methods are employed), it is hard for them to choose an appropriate query at the first trial. Therefore, the query formulation process is treated as a series of tentative trials until the target images are found. Relevance feedback (RF) is an interactive process which can fulfill the requirements of query formulation. It refines the retrievals to a particular query by utilizing the user's feedback on previously retrieved results. The principal idea behind the RF is as follows: The user initializes a query session by submitting an image. The system then compares the query

image to each image in the database and returns t images that are the nearest neighbors to the query. If the user is not satisfied with the retrieved result, he or she can activate an RF process by identifying which retrieved images are relevant and which are nonrelevant. The system then updates the relevance information, such as the reformulated query vector, feature weights, and prior probabilities of relevance, to include as many user-desired images as possible in the next retrieved result. The process is repeated until the user is satisfied or the results cannot be further improved.

To achieve the relevance feedback process, the first strategy focuses on the query concept updating. The aim of this strategy is to refine the query according to the user labelling. A simple approach, called query modification [2], computes a new query by averaging the feature vectors of relevant images. Another approach, the query reweighting, consists of computing a new similarity function between the query and any picture in the database. A usual heuristic is to weigh the axes of the feature space [3]. In order to perform a better refinement of the similarity function, optimization-based techniques can be used. They are based on a mathematical criterion for computing the reweighting, for instance Bayes error [4], or average quadratic error [5], [6]. Although these techniques are efficient for target search and monomodal concept retrieval, they hardly track complex image concepts. Performing an estimation of the query concept can be seen as a statistical learning problem, and more precisely as a binary classification task between the relevant and irrelevant classes [7],[10]. In image retrieval, many techniques based on statistical learning have been proposed, as for instance Bayes classification [8], k-Nearest Neighbors [9], support vector machines [10]. In order to deal with complex and multimodal concepts, we have adopted a statistical learning approach. Additionally, the possibility to work with kernel functions is decisive.

However, a lot of these learning strategies consider the CBIR process as a classical classification problem, without any adaptations to the characteristics of this context. For instance, some discriminative classifiers exclusively return binary labels when real values are necessary for CBIR ranking purposes. Furthermore, the system has to handle classification with few training data, especially at the beginning of the search, where the query concept has to be estimated in a database of thousands of images with only a few examples. Active learning strategies have been proposed to handle this type of problem. To deal

with few training data, active learning methods optimize training data to compensate its scarcity.

In this paper, we focus on statistical learning techniques for interactive image retrieval. We propose some discriminative classification methods for comparison. We propose a scheme to show that active learning [5] may be helpful to carry out an efficient relevance feedback strategy. Active learning strategies offer a natural framework for interactive image retrieval.

In this scope, we first present the binary classification framework to represent complex class distributions (Section II). Powerful classification methods and well-defined kernels for global image signatures are evaluated on real database experiments. Secondly, we present active learning to interact with the user (Section III). In Section IV, the experimental results on Corel database show the efficacy of our approach. Section V provides conclusions of the paper.

2. BINARY CLASSIFICATION FRAMEWORK FOR CBIR

In the classification framework for CBIR, estimation of retrieving image concepts is modelled as a binary classification problem: the relevant class (1), the set of images in the searched concept, and the irrelevant class (-1), composed by the remaining database. In this section, three generative and discriminative learning methods, Bayes, kNN and SVM are presented. They have been selected for their known classification performances in pattern recognition and CBIR context.

For the convenience of the readers, Table I provides the notations of the variables which will appear in this paper frequently.

2.1. Bayes Classifiers

Bayes classifier is tremendously appealing because of its simplicity, elegance, and robustness. It is one of the oldest formal classification algorithms, and yet even in its simplest form it is often surprisingly effective. Bayes classifier is used in text retrieval systems. Since ten years,CBIR community is transposing them to image retrieval [13],[14]. They are interesting since they are directly "relevance oriented," i.e., no modifications are required to get the probability of an image to be in the concept. However, since Gaussian-based models are generally used, they focus on estimating the center of each class, and then are less accurate near the boundary than discriminative methods, which is an important aspect for active learning. Furthermore, because of the unbalance of training data, the tuning of the irrelevant class is not trivial.

Table 1:-Notations Of Key Variables Used In The Paper

VARIABLES	NOTATIONS
$\{x_i\}_{i=1,n}$	n image indexes of the database
$A_y=\{(x_i,y_i)_{i=1,n} \quad y_i \neq 0\}$	Training set expressed from user label retrieval session.($y_i = 1$ if the image x_i is labelled as relevant,$y_i = -1$ if the image x_i is labelled as irrelevant (otherwise $y_i = 0$).
$f(x)$	Relevance function
$k(x_i,x_j)$	Kernel function between images x_i and x_j
X ----- {-1,1}	X labels images as -1 or 1
I^+ and I^-	Indexes of labelled and unlabelled images resp.
jA_y	Cost function

Bayes binary classifiers use the class-conditional likelihood associated with class c $p(x|c)$ to compute the mapping function $g(x)$ of an input vector x:

$$g(x) = \begin{array}{c} argmax \\ c \in (-1,1) \end{array} p(x|c)p(c) \dots\dots(1)$$

Because we have no prior assumption on the size of a class, we assume that $P(1) = P(-1) = ½$. Once $g(x)$ is computed, the relevance function $f(x)$ may be expressed as following

$$f(x) = p\big((x|c)g(x)\big)\dots\dots(2)$$

The Bayes Classification Algorithm:
- Given training data set Ay.
- Estimate the probabilistic parameters.
- Evaluate history's likelihood P $(x|c)$.Here p(c) represents the a priori probability.
- Estimate the a posteriori probability g(x).

2.2. kNN:k-nearest neighbours

This classification method has been used successfully in image processing and pattern recognition. k-nearest neighbour (kNN) classification [9], finds a group of k objects in the training set that are closest to the test object, and bases the assignment of a label on the predominance of

a particular class in this neighbourhood. There are three key elements of this approach: a set of labelled objects, e.g., a set of stored records, a distance or similarity metric to compute distance between objects, and the value of k, the number of nearest neighbors. To classify an unlabelled object, the distance of this object to the labelled objects is computed, its k-nearest neighbours are identified, and the class labels of these nearest neighbors are then used to determine the class label of the object.

Given a test point, a predefined similarity metric is used to find the k most similar points from the train set. For each class y_i, we sum the similarity of the neighbors of the same class. Then, the class y_i with the highest score is assigned to the data point x' by the k nearest neighbours algorithm. The relevance function $f(x_i')$ of the point x_i' is given

$$f(x_i') = \underset{y_j \in (-1,1)}{argmax} \sum_{i=1}^{k} \delta\left(y_j, f(x_i)\right) sim(x_i, x_i') \ldots\ldots (3)$$

k-nearest neighbour classification algorithm:

- Input: A_y, the set of training objects and test object .
- Process: Compute the distance between test object and every training object.Select, the set of k closest training objects to the test object.
- Output: The relevance function $f(x_i')$ for the point x_i' .

2.3. Support Vector Machines

In today's machine learning applications, support vector machines (SVM) [10] are considered a must try—it offers one of the most robust and accurate methods among all well-known algorithms. It has a sound theoretical foundation, requires only a dozen examples for training, and is insensitive to the number of dimensions. In addition, efficient methods for training SVM are also being developed at a fast pace. The aim of SVM classification method is to find the best hyperplane separating relevant and irrelevant vectors maximizing the size of the margin (between both classes). Initial method assumes that relevant and irrelevant vectors are linearly separable. To overcome this problem, kernels k(:, :) have been introduced. It allows to deal with non-linear spaces. Moreover, a soft margin may be used, in order to tolerate noisy configuration [11]. The resulting optimization problem may be expressed as follows:

$$\alpha^* = \underset{\alpha}{argmax} \sum_{i=1}^{N} \alpha_i \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j - K(x_i, x_j)$$

$$\text{With} \begin{cases} \sum_{i=1}^{N} \alpha_i y_i = 0 \\ \forall i \in [1,N] \quad 0 \le \alpha_i \le C \end{cases} \ldots\ldots (4)$$

Thanks to the optimal value α*, the distance between a vector x and the separating hyperplane is used to evaluate how relevant is x:

$$f(x) = \sum_{i=1}^{N} y_i \alpha_i^* \, K(x, x_i) + b \ldots\ldots (5)$$

where b is computed using KKT conditions [12].

3. ACTIVE LEARNING FOR CBIR

Traditional supervised learning methods set their parameters using whatever training data is given to them. By contrast, *active learning* is a framework in which the learner has the freedom to select which data points are added to its training set. An active learner may begin with a very small number of labeled examples, carefully select a few additional examples for which it requests labels, learn from the result of that request, and then using its newly-gained knowledge, carefully choose which examples to request next. In this way the active learner aims to reach high performance using as few labeled examples as possible. Thus active learning can be invaluable in the common case in which there are limited resources for labeling data, and obtaining these labels is time-consuming or difficult. Fig. 1 illustrates our system framework for active learning .
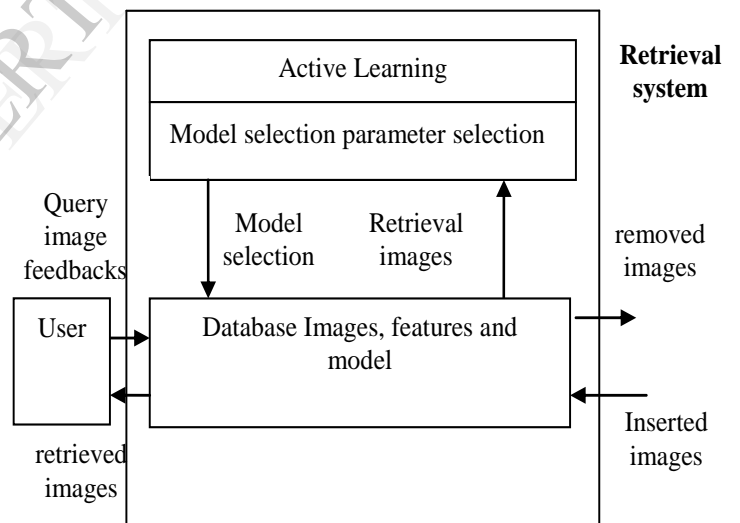
Fig. 1.System diagram with active learning for image databases

When the learner can only choose new data in a pool of unlabelled data, it is called *pool-based* active learning framework [14]. In CBIR, the whole set of images is available anytime. These data will be considered as the pool of unlabeled data during the selective sampling process.

3.1. Optimal active learning scheme

Starting from the image signatures$\{x_i\}_{1,n}$, the training set,indexes of labelled and unlabelled images I^+ and I^- resp. and the relevance function defined in table 1, the

active learning aims at selecting the unlabeled data x_i, that will enhance the most the relevance function $f(x)$ trained with the label $s(x_{i*})$ added to A_y. To formalize this selection process as a minimization problem, a cost function is jA_y. According to any active learning scheme, the selected image is x_{i*} minimizing $jA_y(x)$ over the pool of unlabeled images.

$$i^* = \begin{matrix} argmin \\ i \in I^- \end{matrix} \left(jA_y(x_i) \right) \dots \dots (6)$$

3.2. Active learning method

Two different strategies are usually considered for active learning: the uncertainty-based sampling, which selects the images for which the relevance function is the most uncertain and the error reduction based sampling estimation, which aims at minimizing the generalization error of the classifier. According to this distinction, we have selected two popular active learning strategies for presentation and comparison.

3.2.1. Uncertainty-based sampling

Uncertainty sampling methods iteratively request class labels for training instances whose classes are uncertain despite the previous labelled instances. These methods can greatly reduce the number of instances that an expert need label. In our context of binary classification, the learner of the relevance function has to classify data as relevant or irrelevant. Any data in the pool of unlabelled samples may be evaluated by the learner. Some are definitively relevant, others irrelevant, but some may be more difficult to classify.

To achieve this strategy, a first approach proposed by Cohn [15] uses several classifiers with the same training set, and selects samples whose classifications are the most contradictory. Another solution consists of computing a probabilistic output for each sample, and selecting the unlabelled samples with the probabilities closest to 0.5 [16]. Similar strategies have also been proposed with SVM classifier [17], with a theoretical justification [19], and with nearest neighbour classifier [18].

In any case, a relevance function $f(x)$ is trained. This function may be adapted from a distribution, a membership to a class (distance to the hyperplane for SVM), or a utility function. Using this relevance function, uncertain data x will be close to 0:

$$f(x) \sim 0$$

The solution to the minimization problem in (6) is

$$i^* = \begin{matrix} argmin \\ i \in I^- \end{matrix} (|f(x_i)|) \dots \dots (7)$$

The efficiency of these methods depends on the accuracy of the relevance function estimation close to the boundary between relevant and irrelevant classes.

4. EXPERIMENTS AND RESULTS

4.1. Database and Feature distributions

Experiments are carried out on the COREL photo database which is very interesting because the database is quite large, with many kinds of categories. It consists more than 10,000 pictures. To experiment for statistical evaluation, we randomly selected 10 of the COREL folders with 20 images in each folder. To perform evaluation, we built from the database 10 categories of different complexities ans sizes.

We use histogram of 18 colors and 48 textures for each image. RGB space is used for color and gabor filters in (4 scale factors x 6 orientations) x 2 = 48 different scales and orientations.

4.2. Experimental results for comparative methods

We evaluate 3 discriminative classification methods – Bayes, k-NN and SVM.

As illustrated from fig 2, the SVM technique performs slightly better than kNN and Bayes, they have a strong mathematical framework and efficient algorithmic implementation.

The kernel function used for training SVM is a Gaussian Radial Basis Function (rbf). Rbf kernel is given as

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2} \right) \dots \dots (8)$$

The output of kernel is depended on Euclidean distance of support vector and testing data point. The support vector will be the center of rbf and will determine the area of influence this support vector has over the data space. Hence we are using SVM as a default method for active learning.

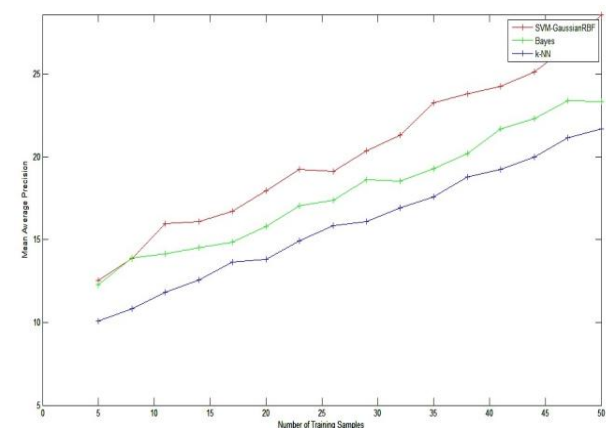The corel database is huge and simulated concept are more difficult.

Fig 2. Mean average precision for classification on the corel photo database

The training set remains too small to allow the above mentioned three classifier to efficiently learn the query concept. Hence to address this problem active learning is considered.

4.3. Experimental results for active learning

Active learning becomes more important when size of the database grows. Fig 3 shows the interest of selection step.The images are represented by 2-D feature vector. SVM classifier with rbf kernel function is computed for binary data classification. As can we see from fig 3, the red '+' are images the user is looking for, and the green '*' are the images the user is not interested in.
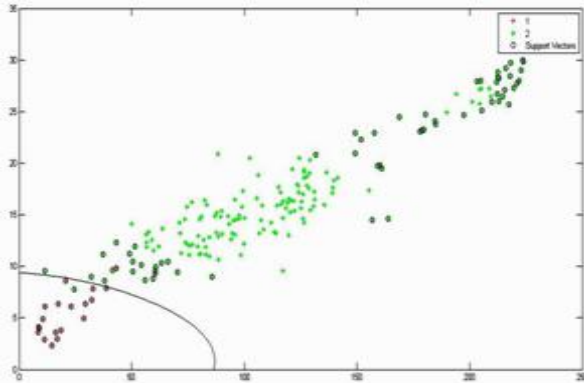
Fig 3 Active learning illustration. SVM training phase diagram for query concept (dinosaur)

The active learning selection working on uncertainty is proposed. In this context, uncertainty sampling selects the image that the user is most uncertain about. The uncertainty for each image depends on ratio between the distances to the closest labelled neighbours of different classes.
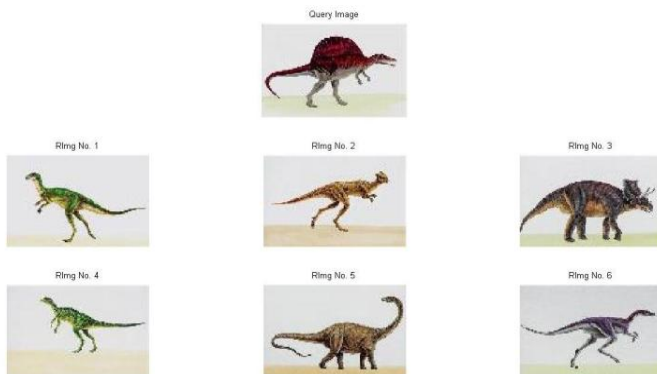
Fig 4 Program output of query concept

Feedback process is illustrated in Fig 4, the upper most centered image is query image. After the query image is submitted, the relevant retrieved images are acquired in 2 feedback stages.

5. CONCLUSION

In this paper, an efficient active learning strategy for interactive content-based image retrieval is introduced.

Our algorithm selects the most difficult images to classify, without using explicit boundary between relevant and irrelevant images.

We experimentally compare three well-known classification methods (Bayes, kNN and SVM) adapted to CBIR context. As compared to SVM, both kNN and Bayes are very simple and well understood. SVM is however, more appealing therotically and in practice, its strength is its power to address non linear classification task. SVM gives the best results on COREL database. After analysing the limitations of these methods to CBIR context, we introduce the active learning scheme. Experiments show that active learning is improving performances of image retrieval process.

REFERENCES

[1] R. Veltkamp, "Content-based image retrieval system: A survey", *Tech. Rep., Univ.Utrecht, Utrecht, The Netherlands, 2002.*

[2] Y. Rui, T. Huang, S. Mehrotra, and M. Ortega, "A relevance feedback architecture for content-based multimedia information retrieval systems", *in Proc. IEEE Workshop Content-Based Access of Image and Video Libraries*, 1997, pp. 92–89.

[3] S. Aksoy, R. Haralick, F. Cheikh, and M. Gabbouj, "A weighted distance approach to relevance feedback", *in Proc. IAPR Int. Conf. Pattern Recognition, Barcelona, Spain, Sep. 3–8, 2000, vol. IV, pp.812–815.*

[4] 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 Transactions on Image Processing*, 9(1):20–37, 2000.

[5] N. Doulamis and A. Doulamis, "A recursive optimal relevance feedback scheme for CBIR", *presented at the Int.Conf. Image Processing, Thessaloniki, Greece, Oct. 2001.*

[6] J. Fournier, M. Cord, and S. Philipp-Foliguet, "Back-propagation algorithm for relevance feedback in image retrieval", *in Proc. Int. Conf. Image Processing, Thessaloniki, Greece, Oct. 2001, vol. 1, pp.686–689.*

[7] D. Michie, D. J. Spiegelhalter, and C. C. Taylor, editors,"Machine Learning, Neural and Statistical Classification",*Ellis Horwood, 1994.*

[8] N. Vasconcelos, "Bayesian models for visual information retrieval", *Ph.D. dissertation, Mass. Inst. Technol., Cambridge, 2000.*

[9] T. Hastie, R. Tibshirani, and J. Friedman, "*The Element of Statistical Learning*", *Springer, 2001*

[10] O. Chapelle, P. Haffner, and V. Vapnik, "Svms for histogram based image classification," *IEEE Trans. Neural Netw.*, vol. 10, pp. 1055–1064, 1999.

[11] K. Veropoulos,"Controlling the sensivity of support vector machines",*In International Joint Conference on Artificial Intelligence (IJCAI99), Stockholm, Sweden, 1999.*

[12] C. Campbell,"Algorithmic approaches to training support vector machines",*A survey. In ESANN, pages 27–36. D-Facto Publications, 2000.*

[13] N. Vasconcelos,"Bayesian models for visual information retrieval", *PhD thesis, Massachusetts Institute of Technology, 2000.*

[14] N. Vasconcelos and M. Kunt,"Content-based retrieval from image databases: current solutions and future directions", *In International Conference in Image Processing (ICIP'01),volume 3, pages 6–9, Thessaloniki, Greece, October 2001.*

[15] D. Cohn, "Active learning with statistical models," *J. Artif. Intell. Res., vol. 4, pp. 129–145, 1996.*

[16] D. D. Lewis and J. Catlett, W. Cohen, H. Hirsh, and editors, Eds., "Heterogeneous uncertainly sampling for supervised learning,"*in Proc. Int. Conf. Machine Learning., Jul. 1994, pp. 148–56.*

[17] J. M. Park, "Online learning by active sampling using orthogonal decision support vectors", *in Proc. IEEE Workshop Neural Networks for Signal Processing, Dec. 2000, vol. 77, pp. 263–283.*

[18] M. Lindenbaum, S. Markovitch, and D. Rusakov, "Selective sampling for nearest neighbor classifiers", *Mach. Learn., vol. 54, no. 2, pp. 125–152, Feb. 2004.*

[19] S. Tong and D. Koller, "Support vector machine active learning with application to text classification", *J. Mach. Learn. Res., vol. 2, pp. 45–66, Nov. 2001.*

[20] N. Roy and A. McCallum, "Toward optimal active learning through sampling estimation of error reduction", *presented at the Int. Conf. Machine Learning, 2001.*