Biased Minimax Probability Machine Active Learning for Relevance Feedback in Content-Based Image Retrieval

Xiang Peng and Irwin King

Department of Computer Science and Engineering
The Chinese University of Hong Kong
Shatin, N.T., Hong Kong
{xpeng, king}@cse.cuhk.edu.hk

Abstract. In this paper we apply Biased Minimax Probability Machine (BMPM) to address the problem of relevance feedback in Content-based Image Retrieval (CBIR). In our proposed methodology we treat relevance feedback task in CBIR as an imbalanced learning task which is more reasonable than traditional methods since the negative instances largely outnumber the positive instances. Furthermore we incorporate active learning in order to improve the framework performance, i.e., try to reduce the number of iterations used to achieve the optimal boundary between relevant and irrelevant images. Different from previous works, this model builds up a biased classifier and achieves the optimal boundary using fewer iterations. Experiments are performed to evaluate the efficiency of our method, and promising experimental results are obtained.

1 Introduction

Content-based Image Retrieval (CBIR) has attracted a lot of research interests in the past decade [9]. For Content-based Image Retrieval, i.e., searching in image database based on their content, the focus was on Query By Example (QBE). A representative CBIR system contains four major parts: image representation, high-dimensional image indexing, similarity measurement between images and system design [12]. At the early stage of CBIR research, scientists focused on the feature extraction for the best representation of the content of images. However these features are often low-level features. Therefore two semantically similar objects may locate far from each other in the feature space, while two absolutely different images may lie close to each other [12]. This is known as the problem of semantic gap between low-level features and high-level concepts and the subjectivity of human perception [2]. Although many features have been investigated for some CBIR systems, and some of them demonstrated good performance, the problem has been the major encumbrance to more successful CBIR systems.

Relevance feedback has been shown to be a powerful tool to address the problem of the semantic gap and the subjectivity of human perception problems in CBIR [2]. Widely used in text retrieval, relevance feedback was first introduced by Rui et al. [8] as an iterative tool in CBIR. Since then it becomes a

E. Corchado et al. (Eds.): IDEAL 2006, LNCS 4224, pp. 953–960, 2006.

major research topic in this area. Recently, researchers proposed a number of classification techniques to attack relevance feedback tasks, in which SVM-based techniques are considered as the most promising and effective techniques [2]. The major SVM technique treats the relevance feedback problem as a strict binary classification problem. However, these methods do not consider the imbalanced dataset problem, which means the number of irrelevant images are significantly larger than the relevant images. This imbalanced dataset problem would lead the positive data (relevant images) be overwhelmed by the negative data (irrelevant images). Furthermore, how to reduce the number of iterations in order to achieve the optimal boundary in this learning task is also a critical problem for image retrieval from large datasets.

In this paper, we propose a relevance feedback technique to incorporate both Biased Minimax Probability Machine and Active Learning to attack these two problems, which can better model the relevance feedback problem and reduce the number of iterations in the learning interaction.

The rest of the paper is organized as follows. In Section II, we review some previous work on relevance feedback. In Section III we first provide an introduction for Active Learning and Biased Minimax Probability Machine (BMPM), then we formulate the relevance feedback technique employing BMPM and Active Learning. Furthermore we show the advantages compared with conventional techniques. Experiments, performance evaluations are given in Section IV. Finally, Section V concludes our work and shows some directions for future work.

2 Related Work

In text retrieval, relevance feedback was used early on and had proven to improve results significantly. The adoption of relevance feedback in CBIR is more recent, and it has evolved to incorporate various machine learning techniques into applications recently. In [7], Decision Tree was employed to model the relevance feedback task. In [1], Bayesian learning was conducted to attack the problem of relevance feedback. Apart from these, many other conventional machine learning methods were also proposed, e.g., Self-organizing Map [5], Artificial Neural Network [9], etc. Furthermore, many state-of-the-art classification algorithms were suggested to model and solve the relevance feedback problem, e.g., Nearest Neighborhood classifier [9] and Support Vector Machine (SVM) [11], etc. Among these techniques, SVM-based techniques are the most effective techniques to address the relevance feedback task in CBIR.

However, conventional relevance feedback techniques by SVMs or other learning models are based on strict binary classification tasks. In other words, they do not consider the imbalanced dataset problem in relevance feedback. Moreover, these techniques always consume a number of iterations to obtain an optimal boundary which is not suitable for image retrieval from large datasets. In order to address this imbalance classification task and make relevance feedback more efficient, we propose the Biased Minimax Probability Machine Active Learning to construct the relevance feedback technique in CBIR.

3 Relevance Feedback by Active Learning Using BMPM

In this section, we introduce the concepts of Active Learning and Biased Minimax Probability Machine. We then present and formulate our proposed Biased MPM methodology with Active Learning, applying to relevance feedback.

3.1 Active Learning

In supervised learning, often the most time-consuming and costly process in designing classifiers is object labelling when we face large scale learning tasks. Instead of randomly picking instances to be manually labelled for training dataset, active learning is a novel mechanism for selecting unlabelled objects based on the result of past labelled objects. Under this framework, the learner could construct a classifier as quickly as possible while active learning method just provides fewer optimal data.

Based on the different criterion for optimal data, there are three main types of active learning method: "Most Information", "Minimizing the Expected Error" and "Farthest First" active learning methodologies. In each iteration of learning, the examples with highest classification uncertainty is chosen for manual labelling [12]. Then the classification model is retrained with additional labelled example. The key issue in active learning for relevance feedback is how to measure the information associated with an unlabelled images. In [6], various of distinct classifier models were first generated. Then, the classification uncertainty of a test image is measured by the amount of disagreement among the test images. Another batch of methodologies measure the information associated with a test example by how far the example is away from the classification boundary. One of the most promising approaches within this group is the SVM active learning developed by Tong and Chang [11].

Let O_i, $i = 1, 2, ..., N$ be the objects in the database, and $\mathbf{x}, \mathbf{y}$ be the two classes we want to perform classification. For each object O_i, we define probability P_{ic} to be the probability that this object belongs to a particular class $\mathbf{x}$ or $\mathbf{y}$. Furthermore we define P_{ix}=1 if the object O_i has been labelled to class $\mathbf{x}$, and P_{ix}=0 if it has been classified to class $\mathbf{y}$. P_{iy} is defined likewise. If the object has not been labelled, P_i is estimated by its nearest neighborhood. In order to derive the expected information gain when we label a certain object, we define an uncertainty measurement as follows:

$$G_i = \Phi(P_{ix}, P_{iy}), \qquad i = 1, 2, ..., N \tag{1}$$

where G_i is the information measurement and $\Phi(\cdot)$ is a function on the class probabilities of object O_i. Moreover, we use the entropy formulation to define the information measurement as

$$G_i = \Phi(P_{ix}, P_{iy}) = -P_{ix}\mathbf{log}P_{ix} - P_{iy}\mathbf{log}P_{iy} \tag{2}$$

3.2 Biased Minimax Probability Machine

We assume two random vectors $\mathbf{x}$ and $\mathbf{y}$ represent two classes of data with mean and covariance matrices as $\{\overline{x}, \Sigma x\}$ and $\{\overline{y}, \Sigma y\}$, respectively in a two-category classification task, where $\mathbf{x}, \mathbf{y}, \overline{x}, \overline{y} \in R^n$, and $\Sigma x, \Sigma y \in R^{n \times n}$. We also use $\mathbf{x}$ and $\mathbf{y}$ to represent the corresponding class of the $\mathbf{x}$ data and the $\mathbf{y}$ data respectively.[1]

With given reliable $\{\overline{x}, \Sigma x\}$, $\{\overline{y}, \Sigma y\}$ for two classes of data, we try to find a hyperplane $\mathbf{a}^T\mathbf{z} = b$ ($\mathbf{a} \neq 0, \mathbf{z} \in R^n, b \in R$, here the superscript T denotes the transpose) with $\mathbf{a}^T\mathbf{z} > b$ being considered as class $\mathbf{x}$ and $\mathbf{a}^T\mathbf{z} < b$ being judged as class $\mathbf{y}$ to separate the important class of data $\mathbf{x}$ with a maximal probability while keeping the accuracy of less important class of data $\mathbf{y}$ acceptable. We formulate this objective as follows:

$$
\begin{aligned}
\max_{\alpha,\beta,b,\mathbf{a}\neq\mathbf{0}} \quad & \alpha \\
s.t. \quad & \inf_{\mathbf{x}\sim(\bar{\mathbf{x}},\Sigma_{\mathbf{x}})} \mathbf{Pr}\{\mathbf{a}^T\mathbf{x} \geq b\} \geq \alpha, \\
& \inf_{\mathbf{y}\sim(\bar{\mathbf{y}},\Sigma_{\mathbf{y}})} \mathbf{Pr}\{\mathbf{a}^T\mathbf{y} \leq b\} \geq \beta, \\
& \beta \geq \beta_0,
\end{aligned}
\tag{3}
$$

where α represents the lower bound of the accuracy for the classification, or the worst-case accuracy of future data points $\mathbf{x}$; likewise β. The parameter β_0 is a pre-specified positive constant, which represents an acceptable accuracy level for the less important class $\mathbf{y}$.

The above formulation is derived from MPM, which requires the probabilities of correct classification for both classes to be an equal value α. Through this formulation, the BMPM model can handle the biased classification in a direct way. This model provides a different treatment on different classes, i.e., the hyperplane $\mathbf{a}_*^T\mathbf{z} = b_*$ given by the solution of this optimization problem will favor the classification of the important class $\mathbf{x}$ over the less important class $\mathbf{y}$.

Given the reliable mean and covariance matrices, the derived decision hyperplane is directly associated with two real accuracy indicators of classification of the future data, i.e., α and β, for each class. Furthermore with no assumption on the data distribution, the derived hyperplane seems to be more general and valid than generative classifiers.

3.3 Proposed Framework

Here we describe how to formulate the relevance feedback algorithm by employing the BMPM technique and Active Learning. Applying BMPM-based techniques in relevance feedback is similar to the classification task. However, the relevance feedback needs to construct an iterative function to produce the retrieval results. The following is our proposed methodology for image retrieval tasks in CBIR:

[1] The reader may refer to [3] for a more detailed and complete description.

Strategy 1. $BMPM_{active}$ loop summary

1: Randomly pick n_0 images from the pool and check their labels
2: Learn a BMPM on the current images whose labels are known
3: Select m images from the dataset based on the criterion of Eq. 2 with the highest value
4: Loop till local optimal boundary achieved or get to maximum number of iterations

After the iterations of relevance feedback have been performed, $BMPM_{active}$ returns the $Top - k$ most relevant images and learn a final BMPM based on the label known images.

Strategy 2. $BMPM_{active}$ final output

1: Learn a final BMPM from the labeled images
2: This decision line maybe the a local optimal one for the whole image dataset.
3: The final BMPM boundary separates relevant images from irrelevant ones.

When we train and engage BMPM in the classification task, the choice of parameters is very direct, for example a typical settings could be n_0=10, m=10, and k=50. Users can also set them empirically by experience.

4 Performance Evaluation

In this section, we will show some experimental results and we compare the performance of two different algorithms for relevance feedback: SVMs and our proposed $BMPM_{active}$. Both of them are based on Radial Basis Function Kernel. The experiments are evaluated on two real world image datasets: a two-category and a ten-category image dataset. These image datasets were collected from COREL Image CDs. All our works are done on a 3.2GHz machine with Intel Pentium 4 processor and having 1Gb RAM.

4.1 Experiment Setup

COREL Image Datasets. The real-world images are chosen from the COREL image CDs. We organize the datasets which contain various images with different semantic meanings, such as *bird*, *pyramid*, *model*, *autumn*, *dog* and *glacier*, etc.

(A) *Two-Bird set.* The 180 images in this dataset belong to two groups - *bird* which contains 80 images, and *pyramid* which consists of 100 images. And we assume the category of *bird* is relevant class.

(B) *Ten-Dog set.* The 480 images in this dataset fall into ten categories - *dog*, *autumn*, *bird*, *pyramid*, *Berlin*, *model*, *church*, *wave*, *tiger*, *Kenya*. In this set we assign the class of *dog* to be the user wanted group and it contains 80 images while the other categories have 100 images each belonging to the irrelevant classes.

Fig. 1. Example Images from COREL Image Database

Image Representation. For the real-world image retrieval, the image representation is an important step for evaluating the relevance feedback algorithms. We extract three different features to represent the images: color, shape and texture.

The color feature employed is the color histograms since it is closer to human natural perception and widely used in image retrieval. We quantized the number of pixels into 10 bins for each color channel (H, S, and V) respectively. Thus we could get a 30-dimensional color histogram.

We use edge direction histogram as shape feature to represent an image [4]. We first calculate the edge images by Canny edge detector and obtain the edge direction histogram by quantize it into 15 bins of 20 degrees. Therefore a 15-dimensional edge direction histogram is generated as the edge feature.

Texture is an important cue for image feature extraction. We apply the wavelet-based texture in our experiments [**?**]. Gabor Wavelet Decomposition is first performed and we compute the features for each Gabor filter output afterwards. Following this approach we use a 16-dimensional vector to describe the texture information for each image.

4.2 Experimental Results

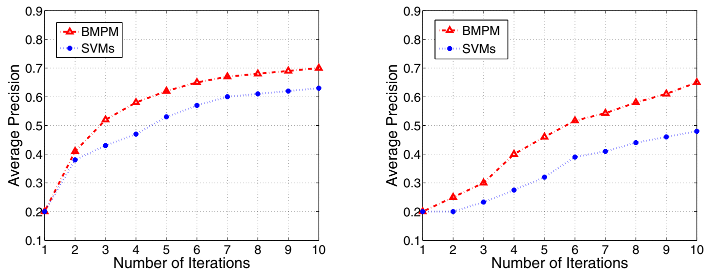
In the following, we present the experimental results by this algorithm on real-world images. The metric of evaluation is Recall vs. Precision for each query, and the Average Precision which is defined as the average ratio of the number of relevant images of the returned images over the total number of the returned images.

The System performances are defined as precision **Pre** and recall **Rec** during the retrieval progress as,

$$\begin{array}{l}\mathbf{Pre} = \frac{r}{\mathbf{A}},\\ \mathbf{Rec} = \frac{r}{\mathbf{R}}.\end{array} \tag{4}$$

where **A** is the number of images returned, r is the number of relevant images retrieved, and **R** is the total number of relevant images in the pool. In general, recall increases as more images are retrieved while precision decreases.

Since we define n_0=10, m=10, and k=50 in the experiments, two positive examples and eight negative examples are randomly picked from the dataset for the first iteration, then SVMs and $BMPM_{active}$ are applied with the same start point. For the iterations afterward, both methods select 10 image based on their



(a) Evaluation on Two-Bird Set **(b)** Evaluation on Ten-Dog Set

Fig. 2. Experimental results over COREL Images Dataset

own strategies. For SVM-based method in our evaluation we select images closest to the boundary from the dataset. In the iterative procedure, the precision and recall are recorded, and the maximum loop used to obtain the average precision is set to be 10 iterations for both methods.

Fig. 2 shows the evaluation results on the Two-Bird dataset and Ten-Dog dataset. From the results on the image sets, we can see that our proposed framework outperforms the other approach.

The following table shows the retrieval results after certain number of iterations by these two methods. We could get a more detailed comparison between these two methods. Here k could be 10, 20, 50, and 100 respectively since we define 10 returns for every iteration. In the right sub-table we notice that when BMPM return most of the relevant images from the pool within 7 iterations while for SVMs it takes more than 10 iterations. From this point we could say BMPM achieves the optimal decision line much earlier than SVMs.

Table 1. Number of relevant images in Top-k Returns

Two-Bird	No. of Iterations					Ten-Dog	No. of Iterations				
Dataset	1	2	5	7	10	Dataset	1	2	5	7	10
BMPM	2	8	31	47	70	BMPM	2	5	23	38	65
SVM	2	7	26	42	63	SVM	2	4	16	28	48

5 Conclusion

In this paper, we address the problem of biased classification needed by the relevance feedback in CBIR and present a novel learning tool, Biased Minimax Probability Machine Active Learning, to treat this problem more precisely. In contrast to the traditional methods, the BMPM provides a more elegant way to handle biased classification tasks. We evaluate the performance of the BMPM based on the COREL image dataset and obtain promising retrieval results.

Acknowledgment

The work described in this paper is supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CUHK4235/04E) and is affiliated with the Microsoft-CUHK Joint Laboratory for Human-centric Computing and Interface Technologies.

References

1. I. Cox, M. Miller, T. Minka, and P. Yianilos. An optimized interaction strategy for bayesian relevance feedback. In *Proceedings of 1998 IEEE Conference on Computer Vision and Pattern Recognition*, pages 553–558, Santa Barbara, USA, 1998.
2. C. H. Hoi, C. H. Chan, K. Huang, M. R Lyu, and I. King. Biased support vector machine for relevance feedback in image retrieval. In *Proceedings of International Joint Conference on Neural Networks*, pages 3189–3194, Budapest, Hungary, 2004.
3. K. Huang, H. Yang, I. King, and M. R Lyu. Learning classifiers from imbalanced data based on biased minimax probability machine. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 558–563, 2004.
4. A. Jain and A. Vailaya. Shape-based retrieval: a case study with trademark image database. *Pattern Recognition*, 9:1369–1390, 1998.
5. J. Laaksonen, M. Koskela, and E. Oja. PicSOM: Self-organizing maps for content-based image retrieval. In *Proceedings of International Joint Conference on Neural Networks, Washington, D.C., USA, July 10–16*, 1999.
6. D. Lewis and W. Gale. A sequential algorithm for training text classifiers. In *Proceedings of ACM International Conference on Research and Development in Information Retrieval*, pages 3–12, Dublin, IE, 1994.
7. S. MacArthur, C. Brodley, and C. Shyu. Relevance feedback decision trees in content-based image retrieval. In *IEEE Workshop on Content-based Access of Image and Video Libraries*, pages 68–72, 2000.
8. Y. Rui, T. Huang, M. Ortega, and S. Mehrotra. Relevance feedback: A power tool for interactive content–based image retrieval. *IEEE Transactions on Circuits and Systems for Video Technology*, 8(5):644–655, 1998.
9. A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, 2000.
10. J. Smith and S. F. Chang. Automated image retrieval using color and texture. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, November 1996.
11. S. Tong and E. Chang. Support vector machine active learning for image retrieval. In *Proceedings of ACM International Conference on Multimedia*, pages 107–118, 2001.
12. C. Zhang and T. Chen. An active learning framework for content-based information retrieval. *IEEE Transactions on Multimedia*, 4(2):260–268, June 2002.