Asymmetric Bayesian Learning for Image Retrieval with Relevance Feedback

Jun Wu* and Mingyu Lu

School of Information Science & Technology, Dalian Maritime University,
Dalian 116026, China
wujunas8@gmail.com

Abstract. Bayesian learning (BL) based relevance feedback (RF) schemes plays a key role for boosting image retrieval performance. However, traditional BL based RF schemes are often challenged by the small example problem and asymmetrical training example problem. This paper presents a novel scheme that embeds the query point movement (QPM) technique into the Bayesian framework for improving RF performance. In particular, we use an asymmetric learning methodology to determine the parameters of Bayesian learner, thus termed as asymmetric Bayesian learning. For one thing, QPM is applied to estimate the distribution of the relevant class by exploiting labeled positive and negative examples. For another, a semi-supervised learning mechanism is used to tackle the scarcity of negative examples. Concretely, a random subset of the unlabeled images is selected as the candidate negative examples, of which the problematic data are then eliminated by using QPM. Then, the cleaned unlabeled images are regarded as additional negative examples which are helpful to estimate the distribution of the irrelevant class. Experimental results show that the proposed scheme is more effective than some existing approaches.

Keywords: Asymmetric learning, Bayesian, image retrieval, relevance feedback.

1 Introduction

Relevance feedback (RF), as a powerful tool for bridging the gap between the high-level semantic concepts and the low-level visual features, has been extensively studied in content-based image retrieval (CBIR) [1]. RF focuses on the interactions between the user and the search engine by letting the user provide feedback regarding the retrieval results, i.e. the user has the option of labeling a few images returned as either positive or negative in terms of whether they are relevant to the query concept or not. From this feedback loop, the engine is refined and improved results are returned to the user. Early RF schemes are heuristic, which aims to improve the query vector or similarity measure function. However, these methods make strong assumption that the target class has an elliptical shape in the feature space, but it is not hold true in the semantically relevant image retrieval. Later on, researchers began to consider RF as a statistical learning problem, which attempts to train a learner to classify

* Corresponding author.

S. Boll et al. (Eds.): MMM 2010, LNCS 5916, pp. 650–655, 2010.

the images in the database as two classes, i.e. relevant (positive) class and irrelevant (negative) class, in terms of whether they are semantically relevant to the query or not. Support vector machine (SVM) has good performance for pattern classification by maximizing the margin of classification hyperplane and thus has been widely used to design RF schemes [2]. However, training a SVM learner is a very time-consuming process, which is inconsistent with the real time requirement of RF. In contrast, Bayesian learner is very easy to construct, not needing any complicated iterative parameter estimation methods [3].

Duan et al. [4] proposed an adaptive Bayesian RF algorithm, termed as Rich get Richer (RGR), which aims at emphasizing the more promising images and deemphasizing the less promising one by assigning high probabilities to the images similar to the query. However, this method often suffers from the small example problem. To address this problem, Yin et al [5] proposed a hybrid RF method by combining BL and query point movement (QPM) technique. But they ignore another characteristic indwelled in RF, i.e. the asymmetrical distribution between the positive and negative examples. Zhang et al [6] proposed a stretching Bayesian method which assumed that each negative example represents a unique irrelevant semantic class and the unlabeled examples near to the concerned negative example are regarded as additional negative examples. The examples close to the concerned negative example should have a strong chance to belong to the same semantic class, and thus a few irrelevant classes that contain the observed negative examples are emphasized. However, there are a lot of irrelevant classes existing in database and most of them are ignored.

In view of above discussion, an asymmetric Bayesian learning (ABL) scheme is developed in this paper, which investigates three special problems in RF, i.e. real time requirement, small example problem, and asymmetric training example problem [1]. First, we construct a very simple learner based on Bayesian inference so as to avoid a complicated learning process. Moreover, an asymmetric learning strategy is applied to estimate the distribution of the positive and negative class. Finally, a novel semi-supervised learning mechanism is presented for tackling the scarcity of negative examples.

2 The Proposed Scheme

Given the query, we apply Bayesian theory to determine the degree that an image in database is classified as a positive or a negative one according to the prior history of feedbacks provided by the user. Since the probability over the whole database is updated after each feedback, the CBIR system, therefore, able to retrieve as many as positive images and reject as many as negative images from being retrieved. Let P denote the positive example set while N denotes the negative example set, and x denotes a random image in database. We use $\mathrm{p}(\ )$ to denote a probability. Based on Bayesian inference, the following equations hold:

$$\mathrm{p}(P|x) = \mathrm{p}(x|P)\,\mathrm{p}(P)/\mathrm{p}(x) \tag{1}$$

$$\mathrm{p}(N|x) = \mathrm{p}(x|N)\,\mathrm{p}(N)/\mathrm{p}(x) \tag{2}$$

Then, the CBIR system can judge whether x is relevant to the query using the leaner:

$$L(x,P,N)=\frac{\mathrm{p}(P|x)}{\mathrm{p}(N|x)}=\frac{\mathrm{p}(x|P)\mathrm{p}(P)}{\mathrm{p}(x|N)\mathrm{p}(N)}\approx\xi\frac{\mathrm{p}(x|P)}{\mathrm{p}(x|N)}\propto\frac{\mathrm{p}(x|P)}{\mathrm{p}(x|N)} \tag{3}$$

Here, the response of the learner describes the relevancy confidence of image x to the query, and thus the learner could produces a rank of images according to how confident it believes the images are relevant to the query. Since the number of the positive images is much less than that of the negative images in the database, $\mathrm{p}(P)/\mathrm{p}(N)$ is treated as a small constant ξ and thus the learner is further simplified as $\mathrm{p}(x|P)/\mathrm{p}(x|N)$. The class-conditional probability density function $\mathrm{p}(x|P)$ and $\mathrm{p}(x|N)$ can be approximated by using Gaussian kernels. To simplify the following description, we use c_i (i=1, 2) denotes the class label ($c_1=P$, $c_2=N$). We assume that each feature dimension of all examples belonged to c_i class satisfies Gaussian distribution.

$$\mathrm{p}(x_k|c_i)=\frac{1}{\sqrt{2\pi}\sigma_k^{(c_i)}}exp\left[-\frac{1}{2}\left(\frac{x_k-\mu_k^{(c_i)}}{\sigma_k^{(c_i)}}\right)^2\right] \tag{4}$$

where x_k is the kth dimension of the feature vector of an image, $\mu_k^{(c_i)}$ and $\sigma_k^{(c_i)}$ are the mean value and the standard deviation of the kth dimension of all examples belonged to c_i class, respectively. Finally, the $\mathrm{p}(x|c_i)$ can be determined by using equation (5):

$$\mathrm{p}(x|c_i)=\prod_{k=1}^{T}\mathrm{p}(x_k|c_i) \tag{5}$$

where T is the number of dimensions of the feature space. Generally, $\mu_k^{(c_i)}$ and $\sigma_k^{(c_i)}$ can be estimated depending upon user labeled images. From Eq. (3)-(5), it can be seen that four parameters of the constructed learner are needed to determine: $\boldsymbol{\mu}^{(P)}=\{\mu_k^{(P)}\}$, $\boldsymbol{\sigma}^{(P)}=\{\sigma_k^{(P)}\}$, $\boldsymbol{\mu}^{(N)}=\{\mu_k^{(N)}\}$, $\boldsymbol{\sigma}^{(N)}=\{\sigma_k^{(N)}\}$, k=1…T.

Most BL based RF schemes directly estimate $\boldsymbol{\mu}^{(P)}$ and $\boldsymbol{\sigma}^{(P)}$ for relevant class using the observed positive examples. However, the positive examples labeled in RF may not be the most representative examples in the potential target class. Hence, the user has to repeat many rounds of feedback to achieve desirable results. Inspired by [5], ABL, by using QPM, attempts to mine a potentially better pattern for representing the target class. But the strategy used in ABL is different from that used in [5] which apply QPM to estimate the parameters for the relevant and irrelevant classes in the same manner. But ABL uses QPM for relevant and irrelevant classes with different purposes.

QPM aims to reformulate the query vector through user's feedback so as to move the query point to a region involving more positive examples in the feature space. Let Q denote the original query, the reformulated query, denoted as Q^*, can be computed by:

$$Q^*=\alpha Q+\beta\sum_{y_k\in P}\frac{y_k}{|P|}-\gamma\sum_{y_k\in N}\frac{y_k}{|N|} \tag{6}$$

where $|\bullet|$ denotes the size of a set, and α , β , and γ are constants used for controlling the relative contribution of each component. Our experiments show that the ABL is not sensitive to the setting of these parameters. Empirically, we set the values of $\alpha = 0.3$, $\beta = 0.6$, and $\gamma = 0.3$. In some sense, Q^* represents the mass centroid of the all possible positive examples, and thus it is reliable to assume that Q^* is also a rational estimate for the mean vector of the assumed Gaussian density of the positive examples. Hence, we set $\boldsymbol{\mu}^{(P)} = Q^*$. Then, based on $\boldsymbol{\mu}^{(P)}$ and P, $\boldsymbol{\sigma}^{(P)}$ can be estimated.

Unlike positive examples, each negative example is 'negative in its own way' [1] and the small number of labeled negative examples can hardly be representative the entire irrelevant class. ABL applies a semi-supervised learning mechanism to overcome the scarcity of negative examples. Our approach is based on a fact that, for any given query, negative examples make up an extremely large proportion of the existing database. So a random subset of the unlabeled image set can be selected as the additional negative examples. Furthermore, to improve the data quality, a QPM-based data cleaning method is applied to eliminate the possible positive examples in the selected unlabeled images.

(1) Collecting unlabeled images.
A random subset of the unlabeled images, denoted as N_U, is generated by using random sampling.

$$N_U = Sampling(U) \text{ with } |N_U| = fix(\sigma_S \cdot |U|) \tag{7}$$

where U denotes the unlabeled image set, $Sampling(\bullet)$ denotes random sampling from a certain set, $fix(\bullet)$ denotes the mantissa rounding operator, and $\sigma_S \in [0,1]$ is the sampling scale which is used for controlling the number of examples sampled from U.

(2) Cleaning the selected unlabeled images.
Since Q^* represents the mass centroid of all possible positive examples, the examples close to Q^* should have a strong chance to be positive. Depending upon this assumption, ABL tries to remove the k most "similar" examples to Q^* from N_U. A simplified version of Radial Basis Function is applied to measured the similarity between Q^* and the examples in N_U. The cleaned image set, denoted as N^*, can be generated by

$$N^* = N_U - \left\{ \mathrm{x}_i \left| \underset{\forall \mathrm{x}_i \in N_U}{argmax}\ exp\left(-\left\|\mathrm{x}_i - Q^*\right\|_2^2\right)\right.\right\}_{i=1}^{k} \text{ with } k = fix(\sigma_C \cdot |N_U|) \tag{8}$$

where $\sigma_C = \eta \cdot \sigma_S \in [0,1]$ is the cleaning scale used for controlling the number of examples removed from N_U, and η is a constant used for adjusting the relationship between σ_C and σ_S . To eliminate the 'bad' examples as much as possible σ_C may be slightly larger than σ_S because the candidate negative examples are readily available in the database. Empirically, η is set to 1.3. Finally, $\boldsymbol{\mu}^{(N)}$ and $\boldsymbol{\sigma}^{(N)}$ are estimated depending upon $N \cup N^*$.

Essentially, our approach could be regarded as a type of active semi-supervised learning algorithm. In the absence of the teacher, our approach just discards the problematic data after identification instead of asking the teacher for labels as in the standard active learning scenario.

3 Experiments

To demonstrate the effectiveness of the proposed ABL, we compare it with SVM active learning (SVM-AL) [2] and Rich get Richer (RGR) method [4]. 3000 images selected from COREL dataset are used to form the testing image database.

At the beginning of retrieval, the images in the database are ranked according to their Euclidean distances to the query. After user feedback, three learning methods are then used to rerank the images in the database. In each round of RF, the user labels 20 images for the system.

Sampling is the most important step for ABL. Selecting a small number of unlabeled examples might make the improvement trivial, while selecting a large number of unlabeled examples might include non-informative or even poor examples into the training set. To select optimal values of σ_S, various feasible values of each parameter are tested. After 60 experiments, the parameter with the best performance among those experiments is $\sigma_S = 0.1$.

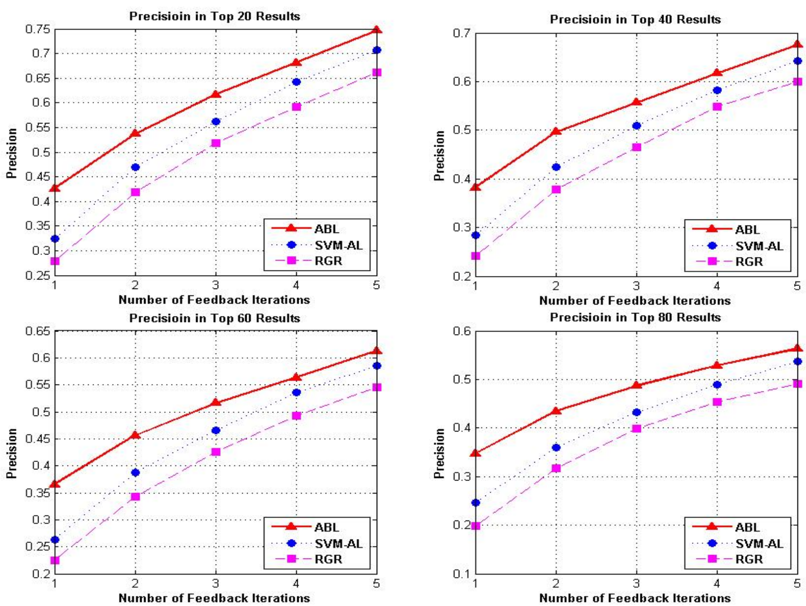

Fig. 1. Performance of the proposed algorithm compared with some existing algorithms

Fig. 1 shows the average precision at the top20, top 40, top 60, and top 80 retrieval results of the three methods. As can be seen, our ABL outperforms the other two methods, especially at the first round of relevance feedback. By iteratively adding the user's feedbacks, the performance difference between ABL and the other two methods gets smaller. Since the number of the labeled images is very limited after the first round of feedback, SVM-AL and RGR can hardly show meaningful results, yet ABL can outperform them obviously since the unlabeled examples are used by it. As the number of feedbacks increases, the performance of SVM-AL and RGR becomes much better. But our ABL continues to perform the best. The above observations show that the proposed asymmetric learning mechanism is effective and ABL method can improve the retrieval performance significantly by using only a few rounds of feedbacks.

4 Conclusions

In this paper, a novel asymmetric Bayesian leaning (ABL) based RF algorithm is presented. There are several key elements in our scheme:

(1) To satisfy the real-time requirement in RF, a very simple learner is constructed based on Bayesian inference. (2) Asymmetric learning strategy is applied to tackle the distribution imbalance between the positive and negative examples. (3) Semi-supervised learning mechanism is introduced in our scheme, and the unlabeled data is helpful to improve the generalization capability of learner. Experimental results illustrate the effectiveness of the proposed algorithm.

Acknowledgments. This research was supported by Natural Science Foundation of China (No.60773084, No. 60603023, No. 60973067) and Doctoral Fund of Ministry of Education of China (No. 20070151009).

References

1. Zhou, X., Huang, T.S.: Relevance Feedback in Image Retrieval: A Comprehensive Review. ACM Multimedia Syst. J. 8, 536–544 (2003)
2. Tong, S., Chang, E.: Support Vector Machine Active Learning for Image Retrieval. In: Proc. ACM Int. Conf. on Multimedia, pp. 107–118. ACM Press, Ottawa (2001)
3. Wu, X.D., Kumar, V., Quinlan, J.R., et al.: Top 10 Algorithms in Data Mining. Knowledge Information Systems 14, 1–37 (2008)
4. Duan, L., Gao, W., Zeng, W., et al.: Adaptive Relevance Feedback Based on Bayesian Inference for Image Retrieval. Signal Processing 85(2), 395–399 (2005)
5. Yin, P., Bhanu, B., Chang, K., et al.: Integrating Relevance Feedback Techniques for Image Retrieval using Reinforcement Learning. IEEE Trans. on Pattern Analysis and Machine Intelligence 27(10), 1536–1551 (2005)
6. Zhang, R., Zhang, Z.: BALAS: Empirical Bayesian Learning in Relevance Feedback for Image Retrieval. Image and Vision Computing 24, 211–223 (2006)