RELEVANCE FEEDBACK WITH MAX-MIN POSTERIOR PSEUDO-PROBABILITY FOR IMAGE RETRIEVAL

Yuan Deng, Xiabi Liu* and Yunde Jia
School of Computer Science and Technology, Beijing Institute of Technology, Beijing, P. R. China

Keywords: Content-based image retrieval, relevance feedback, discriminative training, Gaussian mixture models, max-min posterior pseudo-probabilities.

Abstract: This paper proposes a new relevance feedback method for image retrieval based on max-min posterior pseudo-probabilities (MMP) framework. We assume that the feature vectors extracted from the relevant images be of the distribution of Gaussian mixture model (GMM). The corresponding posterior pseudo-probability function is used to classify images into two categories: relevant to the user intention and irrelevant. The images relevant to the user intention are returned as the retrieval results which are then labelled as true of false by the user. We further apply MMP training criterion to update the parameter set of the posterior pseudo-probability function from the labelled retrieval results. Subsequently, new retrieval results are returned. Our method of relevance feedback was tested on Corel database and the experimental results show the effectiveness of the proposed method.

1 INTRODUCTION

One of the main problems in content-based image retrieval (CBIR) is how to bridge the 'semantic gap' (Zhou and Huang, 2003). Relevance feedback (RF) technique was introduced to CBIR in mid 1990s, with the intention to bring user in the retrieval loop to reduce the gap and improve retrieval performance (Liu et al., 2007).

Currently, some researchers have regarded image retrieval as a supervised learning problem and some machine learning methods have been combined with RF to improve retrieval performance. Since many RF methods treat each image as a whole while the user only concerns a few parts of the image, some researchers transformed CBIR into a Multiple Instance Learning (MIL) problem to find those regions in which the user was interested (Chen et al., 2006). Support Vector Machine (SVM) itself has some drawbacks such as unstable for small-sized training set, biased optimal hyper plane for imbalanced sample set, overfitting, etc. To overcome those problems, some other methods have been applied, such as integrating bagging and random subspace (Tao et al., 2006), active learning (Cheng and Wang, 2007), Boosting (Yu et al., 2007), Biased Minimax Probability Machine (BMPM) (Peng and King, 2006), etc. Besides, due to the user's subjectivity and strict binary classification problem, fuzzy SVM (Rao et al., 2006) and Bayesian learning (Zhang and Zhang, 2006) were proposed to reduce misclassification and refine retrieval precision. Furthermore, Bayesian classifier combining with incremental learning was adapted to realize long term feedback (Goldmann et al., 2006).

In this paper, a new relevance feedback method for image retrieval based on max-min posterior pseudo-probabilities (MMP) (Liu et al., 2006) is proposed to learn user's intention during feedback. We assume that the feature vectors extracted from the relevant images should be of the distribution of Gaussian mixture model (GMM). The posterior pseudo-probability function for the relevant images is used as user intention model. According to the posterior pseudo-probabilities, the images in the database are classified into two categories: relevant to the user intention and irrelevant. The optimum parameter set of user intention model is learned from relevant and irrelevant images that user labelled during feedback using MMP criterion. Then the model obtained after learning is utilized to classify all images and return new results to user. Experiments on 5,000 Corel images show the effectiveness of our proposed method for improving retrieval performance.

* Corresponding Author

Deng Y., Liu X. and Jia Y. (2008).
RELEVANCE FEEDBACK WITH MAX-MIN POSTERIOR PSEUDO-PROBABILITY FOR IMAGE RETRIEVAL.
In *Proceedings of the Third International Conference on Computer Vision Theory and Applications*, pages 286-289
DOI: 10.5220/0001079602860289

2 IMAGE CLASSIFICATION ACCORDING TO USER'S INTENTION

2.1 Statistical Modelling of User's Intention

Each image is represented as an 80-D feature vector which consists of 9-D color moments and 71-D Gabor based texture features.

Relevant images that user labelled during feedback can reflect user's intention; therefore we can use them to describe user's intention. We assume that the feature vectors extracted from the relevant images should be of the distribution of Gaussian mixture model. Let $\mathbf{X}$ be the feature vector of the image, and ω be the relevant image. Let K be the number of Gaussian components, w_k, $\boldsymbol{\mu}_\mathbf{k}$ and $\boldsymbol{\Sigma}_\mathbf{k}$ be the weight, the mean, and the covariance matrix of the K-th Gaussian component respectively, $\sum_{k=1}^{K} w_k = 1$, $p(\mathbf{x}|\omega)$ be the class-conditional probability density function for $\mathbf{X}$:

$$P(x|\omega) = \sum_{k=1}^{K} w_k N(x|\mu_k, \Sigma_k) \tag{1}$$

where

$$N(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = (2\pi)^{-40} |\boldsymbol{\Sigma}_k|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1}(\mathbf{x}-\boldsymbol{\mu}_k)\right) \tag{2}$$

$\boldsymbol{\Sigma}_\mathbf{k}$ is further assumed to be diagonal for simplicity: $\boldsymbol{\Sigma}_\mathbf{k} = [\sigma_{kj}]_{j=1}^{80}$.

2.2 Image Classification using Posterior Pseudo-Probability

Posterior class probabilities are generally used to realize classification in classical Bayesian classifiers. Because it is not practicable to collect the representative examples of irrelevant images, posterior class probability is not adequate for the classification problem discussed here. We use posterior pseudo-probability to approximate $P(\omega|\mathbf{x})$ by embedding $p(\mathbf{x}|\omega)$ in a smooth, monotonically increasing function which takes value in $[0,1]$:

$$P(\omega|\mathbf{x}) \approx f(p(\mathbf{x}|\omega)) = 1 - \exp(-\lambda p(\mathbf{x}|\omega)) \tag{3}$$

where λ is a positive number.

For more details of posterior pseudo-probability, please refer to (Liu et al., 2006).

By substituting Eq. 1 into Eq. 3, we get user intention model:

$$f(\mathbf{X};\boldsymbol{\Lambda}) = 1 - \exp(-\lambda \sum_{k=1}^{K} w_k N(\mathbf{x}|\mu_k, \boldsymbol{\Sigma}_k)) \tag{4}$$

where $\boldsymbol{\Lambda}$ denotes the unknown parameter set:

$$\boldsymbol{\Lambda} = \{\lambda, w_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}, k = 1, \cdots, K \tag{5}$$

We can use the posterior pseudo-probability function to classify the images into two categories: relevant and irrelevant. We compute the values of the posterior pseudo-probabilities for all images in the database using Eq. 4 and sort those images in descending order according to their posterior pseudo-probabilities. Then top N images are returned as the retrieval results.

3 MMP LEARNING OF USER'S INTENTION

In order to classify images using Eq. 4, the unknown parameter set $\boldsymbol{\Lambda}$ must be determined. We use MMP criterion to learn those parameters from the training data. We collect the relevant and irrelevant images that user labelled as positive and negative examples respectively. MMP method is introduced below briefly.

The main idea of MMP method is to maximize the class separability by producing the posterior pseudo-probability function of each class to maximize the posterior probabilities for its positive examples, at the same time to minimize those for its negative examples. Let $\hat{\mathbf{x}}_i$ and $\bar{\mathbf{x}}_i$ be the feature vector of the i-th positive and negative example of the user intention model respectively. Let m and n be the number of positive and negative examples of user intention model respectively. Then the objective function of the MMP learning for user intention model is designed as:

$$F(\boldsymbol{\Lambda}) = \frac{n}{m+n}\sum_{i=1}^{m}[f(\hat{\mathbf{x}}_i;\boldsymbol{\Lambda})-1]^2 + \frac{m}{m+n}\sum_{i=1}^{n}[f(\bar{\mathbf{x}}_i;\boldsymbol{\Lambda})]^2 \tag{6}$$

It is obvious that $F(\boldsymbol{\Lambda}) = 0$ means the hundred-percent class separability: the less the value of $F(\boldsymbol{\Lambda})$ is, the more class separability is. Consequently, we can obtain the optimum parameter set $\boldsymbol{\Lambda}^*$ of user intention model by minimizing $F(\boldsymbol{\Lambda})$:

$$\mathbf{\Lambda}^* = \arg\min_{\mathbf{\Lambda}} F(\mathbf{\Lambda}) \tag{7}$$

The optimum parameter set of user intention model is updated iteratively through the gradient descent method until convergence or a prefixed maximum number of iteration is reached.

For more details of the MMP criterion, please also refer to (Liu et al., 2006).

4 EXPERIMENTS AND DISCUSSIONS

Relevance feedback experiments for querying by concept and querying by example (QBE) on 5,000 Corel images were taken to evaluate our proposed method. Those images are divided into 50 categories, such as African people, beach, buildings, etc. Each category includes 100 images. We also compared our method with other approaches.

4.1 Relevance Feedback Experiment for Querying by Concept

In this experiment, concept refers to the "name" of image category. Therefore relevant images are those images that belong to the specified image category that user query. We assumed that the feature vectors extracted from images with the same image category be of the distribution of Gaussian mixture model. 50 concept models were trained with 2,500 images (50 images each category). Concept retrieval experiment was performed on the remaining 2,500 images. After user input the concept, we computed the posterior pseudo-probabilities of the corresponding concept model for 2,500 images. Then those images were sorted in descending order according to the value of the posterior pseudo-probability functions and the top 50 images were returned as the results. Please refer to (Deng et al., 2007) for more details about querying by concept. During feedback, top 50 images were labelled automatically as relevant to the concept or irrelevant and then used as the training data for user intention model to obtain the optimum parameters set using MMP criterion. Then 2,500 images were classified according to user intention model after learning.

P20 and P50 were used as the performance measure. Table 1 shows the experiment data.

Table 1: Average precision for top 20 and top 50 images.

Iteration times	P20	P50
#0	0.4220	0.3280
#1	0.5770	0.3752
#2	0.6120	0.4024
#3	0.6350	0.4052
#4	0.6550	0.4144
#5	0.6720	0.4288

Gosselin and Cord proposed a retrieval method which combined transductive SVM with active learning strategy (Gosselin and Cord, 2004). Their retrieval experiment was performed on 11 Corel image categories, nine iterations were carried out and 20 images were labelled each time. We did similar experiment on 50 Corel image categories. Table 2 shows the experiment data between our method and Gosselin and Cord's method, which are denoted as MMP and RETINAL respectively.

Table 2: Average precision for top 20 images after nine iterations in two methods.

P20	#9
RETINAL	0.61
MMP	0.8790

4.2 Relevance Feedback Experiment for Querying by Example

This experiment was designed to find images that were similar to the query image. Two images with the same image category are similar; therefore relevant images are those images from the same image category as the query image. We assumed that the difference between feature vectors of two images from the same category be of the distribution of Gaussian mixture model. We randomly chose 20 images from each image category, or, 1,000 images in all, to train similarity model. QBE experiment was performed on the remaining 4,000 images. After the user input the query image, the system computed the posterior pseudo-probabilities for the query image and the target image in the database. Then those target images were sorted in descending order according to the value of the posterior pseudo-probability functions and the top 80 images were returned as the results. During feedback, top 80 images were labelled automatically as relevant to the query image or irrelevant and then used as the training set for user intention model to obtain the optimum parameters set using MMP criterion. Then 4,000 images were classified according to user intention model after learning. Table 3 shows the experiment data.

Table 3: Average precision for top 20 and top 50 images.

Iteration times	P20	P50
#0	0.4982	0.3878
#1	0.6026	0.4570
#2	0.6446	0.4853
#3	0.6694	0.5013
#4	0.6894	0.5107
#5	0.7028	0.5192

Rao et al. proposed a querying by example method based on Fuzzy SVM and performed experiment on 2,000 Corel images (Rao et al., 2006). Top 20 images were labelled each time. We did similar QBE experiment on 5,000 Corel images. Table 4 shows the experiment results between two methods at three iteration steps (#1, #5, and #10), Rao et al.'s method is denoted as Fuzzy SVM.

Table 4: Average precision for top 20 images in two methods at three iteration steps.

P20	#1	#5	#10
Fuzzy SVM	About 0.53	About 0.74	About 0.77
MMP	0.5730	0.6681	0.7172

5 CONCLUSIONS

In this paper, we have proposed a new relevance feedback method based on max-min posterior pseudo-probabilities framework for learning pattern classification. We assume that the feature vectors extracted from the relevant images be of the distribution of Gaussian mixture model. The corresponding posterior pseudo-probability function is used to determine whether the image is relevant to the user intention. In each feedback process, those images relevant to the user intention are returned as the retrieval results and then labeled as true or false by the user. According to labeled retrieval results, MMP training criterion is used to update the parameter set of posterior pseudo-probability function and subsequent retrieval results. We conducted concept retrieval and example retrieval experiments of relevance feedback on Corel database. After five iterations, P20 has been raised from 42.20% to 67.20% and from 49.82% to 70.28% respectively.

ACKNOWLEDGEMENTS

This research was partially supported by the 973 Program of China (2006CB303105) and BIT Excellent Young Scholars Research Fund (2006Y1202).

REFERENCES

Chen, S.C., Rubin, S.H., Shyu, M.L., Zhang, C., 2006. A dynamic user concept pattern learning framework for content-based image retrieval. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, 36(6): 772-783.

Cheng, J., Wang, K., 2007. Active learning for image retrieval with Co-SVM. *Pattern Recognition*, 40(1): 330-334.

Deng, Y., Liu, X., Jia, Y., 2007. Learning Semantic Concepts for Image Retrieval Using the Max-min Posterior Pseudo-Probabilities Method. In *ICME 2007, IEEE International Conference on Multimedia and Expo*.

Goldmann, L., Thiele, L., Sikora, T., 2006. Online image retrieval system using long term relevance feedback. In *CIVR 2006, International Conference on Image and Video Retrieval*.

Gosselin, P.H., Cord, M., 2004. RETIN AL: an active learning strategy for image category retrieval. In *ICIP 2004, International Conference on Image Processing*.

Liu, X., Jia, Y., Chen, X., Deng, Y., Fu, H., 2006. Max-Min Posterior Pseudo-Probabilities Estimation of Posterior Class Probabilities to Maximize Class Separability. Technical Report BIT-CS-TP-20060002, Beijing: Beijing Institute of Technology. (http://www.mcislab.org.cn/technicalreports/MMP.pdf)

Liu, Y., Zhang, D., Lu, G., Ma, W.Y., 2007. A Survey of Content-Based Image Retrieval with High-level Semantics. *Pattern Recognition*, 40(1): 262-282.

Peng, X., King, I., 2006. Imbalanced learning in relevance feedback with biased minimax probability machine for image retrieval tasks. In *ICONIP 2006, International Conference on Neural Information Processing*.

Rao, Y., Mundur, P., Yesha, Y., 2006. Fuzzy SVM Ensembles for Relevance Feedback in Image Retrieval. In *CIVR 2006, International Conference on Image and Video Retrieval*.

Tao, D., Tang, X., Li, X., Wu, X., 2006. Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(7): 1088-1099.

Yu, J., Lu, Y., Xu, Y., Sebe, N., Tian, Q., 2007. Integrating relevance feedback in boosting for content-based image retrieval. In *ICASSP 2007, IEEE International Conference on Acoustics, Speech, and Signal Processing*.

Zhang, R., Zhang, Z., 2006. BALAS: Empirical bayesian learning in the relevance feedback for image retrieval. *Image and Vision Computing*, 24(3): 211-223.

Zhou, X.S., Huang, T.S., 2003. Relevance feedback in image retrieval: A comprehensive review. *Multimedia Systems*, 8(6): 536-544.