Partial Highest Possible Edge Analysis for Interactive Image Accessibility

R.SRIDEVI
PG Scholar, Department of Computer Science
and Engineering
SCSVMV University, India

***Abstract*:** Relevance feedback is a technique that takes advantage of human-computer interaction to refine high level queries represented by low level features. Among RF schemes, the most popular technique is SVM based RF scheme. When SVM is used as a classifier in RF, there are two strategies. One strategy is to display the most positive images and use them as the training samples. The most-positive images are chosen as the ones farthest from the boundary on the positive side, plus those nearest from the boundary on the negative side if necessary. Another strategy is that most of SVM based RF scheme does not consider the unlabeled samples even though they are useful in constructing a good classifier. To overcome these drawbacks, in this paper we propose a biased maximum margin analysis (BMMA) and semi supervised BMMA (semiBMMA) for integrating the distinct properties of feedbacks and also to utilize the information of unlabeled samples. The BMMA differentiates positive from negative feedbacks, whereas semiBMMA takes into account the information of unlabeled samples by the introduction of Laplacian regularizer to BMMA. To validate the efficacy of the proposed approach, we test it on both synthesized data and real-world images. Promising results are achieved and this can significantly improve the performance of CBIR systems.

Keywords: *Content-based image retrieval (CBIR), relevance feedback (RF), support vector machines (SVM), graph embedding framework.*

Introduction

Content-based image retrieval (CBIR), as we see it today, is any technology that in principle helps organize digital picture archives by their visual content. By this definition, anything ranging from an image similarity function to a robust image annotation engine falls under the purview of CBIR. This characterization of CBIR as a field of study places it at a unique juncture within the scientific community. In the CBIR context, an image is represented by a set of low-level visual features, which are generally not effective and efficient in representing the image contents, and they also have no direct correlation with high-level semantic information. The gap between high-level information and low-level features is the fundamental difficulty that hinders the improvement of the image retrieval accuracy. Recently, a variety of solutions have been suggested that aim to bridge this semantic gap.

The relevance feedback [1] narrows the semantic gap by making use of user provided judgments which are the labels (relevant or non-relevant) on the retrieved images for a query. The retrieval performance improves as the user provides more and more feedback information to the CBIR system. Query vector modification (QVM) [2] and feature relevance learning [3] are the two widely used methods to integrate user feedback information into the CBIR system. Majority of the work uses relevance feedback to learn the relative importance of different features, with some tries to learn a feature weighting scheme either with [4] or without[5] considering correlations among feature components; while others either use a probabilistic scheme , or Self-Organizing Maps , or boosting technique , etc., to do so. A typical problem with CBIR system with relevance feedback is the relatively small number of training samples and the high dimension of the feature space. The system can only present the user with a few dozen of images to label (relevant or irrelevant).

Paper Type: Technical
Corresponding Author: R.SRIDEVI, email: sridevi.ragu@gmail.com

The interesting images to the user are only a very small portion of the large image database, in which most images remain unlabeled. Therefore, small sample learning methods are most promising for RF.

Two-class SVM is one of the popular small sample learning methods widely used in recent years and obtains the state-of-the-art performance in classification for its good generalization ability. Guo *et al*. developed a constraint similarity measure for image retrieval which learns a boundary that divides the images into two groups, and samples inside the boundary are ranked by their Euclidean distance to the query image. The SVM active learning method selects samples close to the boundary as the most informative samples for the user to label. It is almost impossible to estimate the real distribution of negative images in the database based on the relevant feedback. Nevertheless most of the SVM RF approaches ignore the basic difference between the two distinct groups of feedbacks, i.e all positive feedbacks share a similar concept (Fig1) while the negative feedbacks share a different concepts (Fig 2). Directly using SVM as an RF scheme damages the entire performance of CBIR systems. One problem is that different semantic concepts live in different subspace and it is the goal of RF schemes to figure out “which one”. Additionally it has another problem of incorporating the unlabeled samples into traditional SVM based RF schemes, even though they are useful in constructing a good classifier.

Fig1. Set of negative feedback in RF iteration ***Fig 2. Set of positive feedback in RF iteration***

To explore solutions to the above problems, in this paper we propose a technology, biased maximum margin analysis (BMMA) and semi supervised BMMA (SemiBMMA) for the traditional RF scheme, based on graph embedding framework [30]. The proposed scheme is mainly based on the following: 1) effectiveness of treating positive and negative samples differently; 2) the success of graph embedding in characterizing intrinsic geometric properties of the data set in high dimensional space;

The convenience of graph embedding framework in constructing semi supervised learning techniques.

2. RELATED PREVIOUS WORK

i) SVM RF FOR CBIR SYSTEMS

Many relevance feedback methods have been developed in recent years. They either adjust the weights of various features to adapt to the user’s preferences or estimate the density of the positive feedback examples. Regarding the positive samples and the negative samples as two difference groups and aiming at finding a classifier to identify these two groups from each other, relevance feedback in CBIR becomes a real-time classification problem. Among these classifiers, the Support Vector Machines (SVM) based relevance feedback (SVM RF) [6] has shown promising results owing to its good generalization ability. SVM has a very good performance for pattern classification problems by minimizing the Vapnik-Chervonenkis dimensions achieving a minimal structural risk. SVM active learning halves the image space each time in which the most positive samples are selected farthest from the classifier boundary on the positive side and the samples close to the boundary are deemed as the most informative ones for the user to label. Guo et al. [2] developed a constrained similarity measure (CSM) for image retrieval in which the SVM is also employed with AdaBoost. The CSM also learns a boundary that halves the images in the database into two groups and images inside the boundary are ranked by their Euclidean distances to the query. There are also some more kinds of SVM-based relevance feedback algorithms [3].

ii) GRAPH EMBEDDING FRAMEWORK

A unified view for understanding and explaining many popular algorithms such as PCA/LDA/LPP which may be used for linear techniques and ISOMAP/Laplacian Eigenmap/LLE which may be used for non-linear techniques. Graph embedding framework may be a platform for developing new dimension reduction algorithms.

Let the sample set be represented as

$$X = [x_1, \ldots x_N], x_i \in R^m$$

In this case often m is very large so there is a need to find the function F,intrinsic graph G and penalty graph $G^{p.}$

The function F can be represented as

$$F : x \mapsto y, y \in R^{m'}, m' << m$$

The intrinsic graph can be represented as

$$G = \{X, W\} \quad , W \in R^{N \times N}$$

The penalty graph can be represented as

$$G^p = \{X, W^p\} \quad , W^p \in R^{N \times N}$$

For a dimensionality reduction problem, direct graph embedding requires an intrinsic graph G , whereas a penalty graph G^P not a necessary input. The graph preserving criterion can be explained as

$$y^* = \arg\min_{y^T By = d} \sum_{i \neq j} \left\| y_i - y_j \right\|^2 W_{ij} = \arg\min_{y^T By = d} y^T L y$$
$$L = D - W, D_{ii} = \sum_{i \neq j} W_{ij}$$

Where tr(.)is the trace of an arbitrary square matrix, C is a constant, and B is the constraint matrix. The Laplacian matrix L can be explained as $L=D^P {}_{-}W^P$

The graph-embedding framework preserves the intrinsic property of the samples in two ways. For larger similarity between samples x_i and x_j, the distance between y_i and y_i should be smaller to minimize the objective function. Conversely, smaller similarity between x_i and x_j, should lead to larger distance between y_i and y_j . Hence, through the intrinsic graph G and the penalty graph G^P, the similarities and the differences among vertex pairs in graph can be preserved in the embedding.

3. CBIR SYSTEMS

In experiments, we use a subset of the Corel Photo Gallery as the test data to evaluate the performance of the proposed scheme. Given a query image by the user, the CBIR system is expected to feed back more semantically relevant images after each feedback iteration. However, during RF, the number of the relevant images is usually very small because of the semantic gap. At the same time, the user would not like to label a large number of samples. The user also expects to obtain more relevant images with only a few rounds of RF iterations. Keeping the size of labeled relevant images small and keeping the RF iterations few are two key issues in designing the image retrieval system. Therefore, we devise the following CBIR framework accordingly to evaluate the RF algorithms.

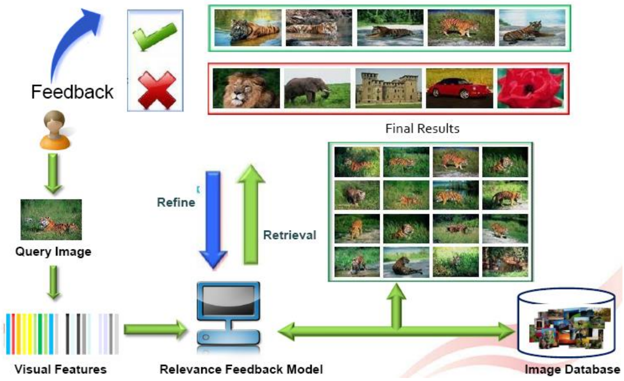

Fig 3 Framework of CBIR systems

Low-level image feature extraction is the basis of CBIR systems. To perform CBIR, image features can be either extracted from the entire image or from regions.

1. Image segmentation

Automatic image segmentation is a difficult task. A variety of techniques have been proposed in the past, such as curve evolution, energy diffusion, and graph partitioning. Many existing segmentation techniques work well for images that contain only homogeneous color regions, such as direct clustering methods in color space. These apply to retrieval systems working only with colors. Texture is an important feature in defining high-level concepts. As stated in [4] texture is the main difficulty in a segmentation method. Many texture segmentation algorithms require the estimation of texture model parameters which is a very difficult task. 'JSEG' segmentation overcomes these problems. Blobworld segmentation is another widely used segmentation algorithm. Some systems design their own segmentations in order to obtain the desired region features during segmentation, be it color, texture, or both.

The use of segmentation algorithm depends on the requirements of the system and the data set used. It is hard to judge which algorithm is the best.

2. Low-level image features

Many sophisticated feature extraction algorithms have been designed and good surveys are available. Here are the features with high-level semantics.

2.1 Color feature

Color feature is one of the most widely used features in image retrieval. Color spaces shown to be closer to human perception and used widely in RBIR include, RGB, LAB, LUV, HSV (HSL), YCrCb and the hue-min-max-difference (HMMD) . Common color features or descriptors in RBIR systems include, color-covariance matrix, color histogram, color moments, and color coherence vector. As the result, a set of viewpoint invariant color features have been computed.

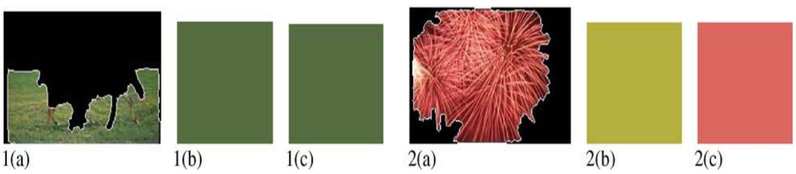

Fig. 4 Average color and dominant color: (a) original region; (b) average color; (c) dominant color.

2.2 Texture feature

Texture is not as well-defined as color features, some systems do not use texture features. However, texture provides important information in image classification as it describes the content of many real-world images such as fruit skin, clouds, trees, bricks, and fabric. Hence, texture is an important feature in defining high-level semantics for image retrieval purpose. Texture features commonly used in image retrieval systems include spectral features, such as features obtained using Gabor filtering [4] or wavelet transform, statistical features such as the six Tamura texture features and world features proposed by Liu et al.. Among the various texture features, Gabor features and wavelet features are widely used for image retrieval and have been reported to well match the results of human vision study. Gabor filtering and wavelet transform are originally designed for rectangular images. The Weber local descriptors (WLDs) are adopted as feature descriptors, which are mainly based on the human perception of pattern.

2.3 Shape feature

Shape is a fairly well-defined concept. Shape features of general applicability include aspect ratio, circularity, Fourier descriptors, moment invariants, consecutive boundary segments, etc.

Fig 5 Image representations

Shape features are important image features though they have not been widely used as color and texture features. Shape features have shown to be useful in many domain specific images such as man-made objects. For color images used in most papers, however, it is difficult to apply shape features compared to color and texture due to the inaccuracy of segmentation. Despite the difficulty, shape features are used in some systems and has shown potential benefit for RBIR.

4. MAXIMUM MARGIN FOR SVM RF IN CBIR

After reviewing all the approaches of image retrieving process there are some pros and cons. So there is a need of effective method to solve the problem. In this paper we have used biased maximum margin.

Algorithm

1. Train the images into the database so that the images are categorized based on the content.
2. Input the query image that needs to be reported as output.
3. Color extraction process-

i) Convert the image from RGB to HSV.

ii) Calculate vector point.

iii) Apply biased maximum margin and perform training and mapping.

iv) Finally get the feature vector.

4. Texture extraction process-
i) Convert the input query image from RGB to grayscale.
ii) Calculate vector point.
iii) Apply biased maximum margin and perform training and mapping.
iv) Finally get the feature vector.

5. Shape extraction process

i) Convert the image and perform removal of background.

ii) Apply Weber local descriptor
iii) Apply biased maximum margin and perform training and mapping.
iv) Finally get the feature vector.
6. All the feature vector go to a feature space.
7. Similarity computation
i) Query image and the database image are viewed.
ii) If both the images are found to be similar then the results are produced to the user.
iii) If both the images are mismatched then the process is repeated until the user satisfied output is produced.

The two groups of feedbacks have different properties for CBIR with the observation that "All positive are alike and each negative samples are negative in its own way". To utilize the information of unlabelled samples in the database this paper introduces Laplacian regularizer to BMMA which will lead to semiBMMA for SVM RF.

Then the remaining images in the database are projected onto this resultant semantic subspace, and a similarity measure is applied to sort the images based on new representations. The major criterion in SVM based RFs is that the distance to the hyperplane of the classifier to differentiate the query relevant samples from the query irrelevant samples? After all the images are projected into subspace, all the positive feedbacks are clustered together while the negative feedbacks are separated from positive feedbacks by a maximum margin.

Therefore, the resultant classifier seems to be much simpler and better than that in the original high-dimensional feature space.BMMA aims to learn a projection matrix α such as that in the projected space all positive samples have high local within class similarity but the samples with different labels have high between-class separability. Here two graphs are constructed to show the similarity and dissimilarity i.e Intrinsic graph G, to show the similarity between positive feedbacks and the Penalty graph G^P, to show the dissimilarity between positive feedbacks and negative feedbacks. $L^u = D^u - W^u$ Can be known as a Laplacian matrix. Hence, we call this term as a Laplacian regularizer. There are a lot of possible ways to choose a regularizer for the proposed BMMA but we have chosen Laplacian regularizer, which is largely inspired by emerging manifold learning community. For the penalty graph G^P , its similarity matrix represents geometric properties to be avoided and is used as a constraint matrix in the graph embedding framework.

The BMMA algorithm optimizes the objective function in a trace difference form,

$$\begin{aligned}\alpha^* &= \arg\max_{\alpha} 2tr[\alpha^T X(D^P - W^P)X^T\alpha] - 2tr[\alpha^T X(D-W)X^T\alpha] \\ &= \arg\max_{\alpha} tr(\alpha^T XBX^T\alpha) - tr(\alpha^T XLX^T\alpha) \\ &= \arg\max_{\alpha} tr[\alpha^T X(B-L)X^T\alpha]\end{aligned}$$

This objective function works in two ways, i.e trying to maximize $\mathrm{tr}(\alpha^T XBX^T\alpha)$ and at the same time minimize $\mathrm{tr}(\alpha^T XLX^T\alpha)$. The difference between BMMA and MMC are the definitions of the interclass separability and infraclass compactness. In MMC, both the interclass separability and intraclass com-pactness are defined as the same in LDA, which treats the two different classes equally, and MMC can see only the linear global Euclidean structure. In BMMA, the intraclass compactness is constructed by only considering one class and characterized by a sum of the distances between each positive sample and its k_1 nearest neighbors in the same class. It should be noted that previous methods that followed MMC cannot be directly used for the SVM RF in image retrieval because these methods treat samples in different classes equally.

The solution to this problem is trivial. Therefore an arbitrary scaling factor has to be removed in the projection,

$$\max_{\alpha} tr(\alpha^{T} X(B-L)X^{T}\alpha)$$

$$=\sum_{k=1}^{l} \alpha_k^{T} X(B-L)X^{T}\alpha_k$$

$$s.t.\ \ \alpha_k^{T}\alpha_k - 1 = 0,\ k = 1,2,\ldots,l$$

The other constraints such can be used but this also encounters “small size problem”. This maximum margin approach can be compared with traditional MFA with the same constraint but when analyzing, it gives the difference that maximum approach involves a constrained approach, whereas the traditional MFA solves an unconstrained optimization problem. In order to solve the mentioned problem, this paper introduces a Lagragian, i.e.,

$$L(\alpha_k,\lambda_k) = \sum_{k=1}^{l} \alpha_k^T X(B-L)X^T\alpha_k - \lambda_k\left(\alpha_k^T\alpha_k - 1\right)$$

Lagragian L should be maximized with respect to both λ_k and α_k. The condition is that, at the stationary point, the derivatives of respect to α_k must vanish, i.e.,

$$\frac{\partial L(\alpha_k,\lambda_k)}{\partial \alpha_k} = \left(X(B-L)X^T - \lambda_k I\right)$$
$$\alpha_k = 0, \quad k = 1,2,\ldots,l$$

And therefore, $X(B-L)X^T\alpha_k = \lambda_k\alpha_k, \qquad k = 1,2,\ldots,l$

Which means that they $\lambda_k' s$ are the Eigen values of $X(B-L)X^T$ and $\alpha_k' s$ are the corresponding Eigen vectors. Thus, we have

$$J(\alpha) = \sum_{k=1}^{l} \alpha_k^T X(B-L)X^T\alpha_k = \sum_{k=1}^{l} \lambda_k\alpha_k^T\alpha_k = \sum_{k=1}^{l} \lambda_k$$

Therefore, the objective function is maximized when α is composed of the largest Eigen vectors of $X(B-L)X^T$. This allows us to avoid the “small sample size” problem easily.

CONCLUSION

BDA and its kernal version map were first proposed to address the asymmetry between the positive and negative samples in interactive image retrieval. However, to use BDA “small size problem” and Gaussian assumption are the two major challenges. Different from the original BDA, our BMMA algorithm is a local discriminant analysis approach, which does not make any assumption on the distribution of the samples.

Since the graph-embedding technique is an effective way to capture the intrinsic geometry structure in the original feature space, we propose a way to incorporate the unlabeled samples based on the intrinsic graph, which is helpful in capturing the manifold structure of samples and alleviating the over fitting problem.

This scheme can preserve weak (probably correct) similarities between all unlabeled sample pairs and thus effectively integrate the similarity information of unlabeled samples into BMMA. After all the projection of images has been completed, the traditional SVM RF is executed on the new representations.

REFERENCES.

[1] M. R. Azimi-Sadjadi, J. Salazar and S. Srinivasan, “An adaptable image retrieval system with relevance feedback using kernel machines and selective sampling” *IEEE PAMI*, vol. 18, no.

7, pp. 1645-1659, July 2009.

[2] M. R. Widyanto, and T. Maftukhah, “Fuzzy relevance feedback in image retrieval for color feature using query vector modification method” *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol.14, No.1, pp, 34-38, Jan. 2010.

[3] B. Ko, J. Peng, and H. Byun,”Region based image retrieval using probabilistic feature relevance learning,” Journal of pattern Analysis and Applications, Vol.4, No.2, pp. 174-184, June 2001.

[4] Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra “Relevance Feedback: A Power Tool in Interactive Content-Based Image Retrieval”, IEEE Transactions on Circuits and Systems for Video Technology, Vol 8, No. 5, Sept., 644-655, 1998.

[5] Y. Rui, T. S. Huang, “Optimizing learning in image retrieval”, Proc. IEEE, 236-243.

[6] P. Hong, Q. Tian, and T. S. Huang, “Incorporate support vector machines to content-based image retrieval with relevant feedback,” in *Proc. IEEE ICIP*, Vancouver, BC, Canada, 2000, pp. 750–753.