Semi-Supervised Collaborative Image Retrieval using Relevance Feedback

Ahire Bhavna A[1], Chaudhari Sayali R[2], Kamble Priya V[3], Ahire Rohini S[4]

[1,2,3,4] Student, Department of Computer Engineering, R. H. Sapat College of Engineering
College Road, Vidya Nagar, Nashik, Maharashtra, India
Prof. R.R. Chakre, Deparment of Computer Engineering. of R. H. Sapat College of Engineering, Nashik, Maharashtra, India

---

Abstract - *In Digital Image Processing, Content-Based Image Retrieval has gained much popularity. However, the accuracy of CBIR system is low. To improve the performance of CBIR system, Relevance feedback system can be used. In relevance feedback system the user refines the search results, progressively by marking images in the results as "relevant", "irrelevant", or "neutral" to the search query and then repeating the search with the new information. In many cases, there may be a large number of images to a label. Most of the times user would not like to label a large number of images. In this paper, a semisupervised method is used. This means the user needs to label only a few most informative images. These labeled images are then used as a training set for SVM classifier. The images in the database are resorted based on a new similarity metric. If the user is satisfied with the results, Relevance feedback is nolonger required and the system gives the final results. These results are most semantically relevant to the query image. Thus the image retrieval process is ended. Otherwise, Relevance Feedback will be performed iteratively.*

Keywords—Content Based Image Retrieval (CBIR), Collabora- tive Image Retrieval, Semantic Gap, Relevance Feedback (RF), Feature Selection, Binary Classifier, Precision, Recall.

1. INTRODUCTION

In Digital Image Processing, Content Based Image Retrieval (CBIR) [1]-[3] has gained much popularity. In Content-based Image Retrieval visual contents are used to search images from large scale image databases according to user interest. But one of the reasons for poor performance of CBIR is the gap between low level and high level features. Also users viewpoint for same image may be different at different times. Also in CBIR, images with same visual features but different semantics may be considered as identical. This state of problem is called as 'semantic gap'. Now, to solve these problem, relevance feedback [4] can be used. In relevance feedback system the user progressively refines the search results by marking images in the results as "relevant", "irrelevant", or "neutral" to the search query and then repeating the search with the new information. In many cases there may be large number of images to label which becomes hectic for the user. In this respect, semi supervised method [5] is used in which, the user needs to label only few most informative images. These labeled images are then used as training set for SVM classifier [6][7]. Then images in database are resorted based on new similarity metric. If the user is satisfied with the results, Relevance feedback is no longer required and the system gives the final results. These results are most semantically relevant to the query image and image retrieval process ends. Otherwise, Relevance Feedback will be performed iteratively. This improves the performance of CBIR system.

1.1 Content Base Image Retrival

Content-based image retrieval (CBIR), also known as Query by image content (QBIC) and Content-Based Visual Information Retrieval (CBVIR) is the application of computer vision techniques to the image retrieval problem. Thus it is the problem of searching for digital images in large databases. Content-based image retrieval contrasts with traditional concept-based approaches. "Content-based" means that on searching, the contents of the image are analysed rather than the metadata such as the keywords, descriptions or tags associated with the image according to process as shown in Figure 1.

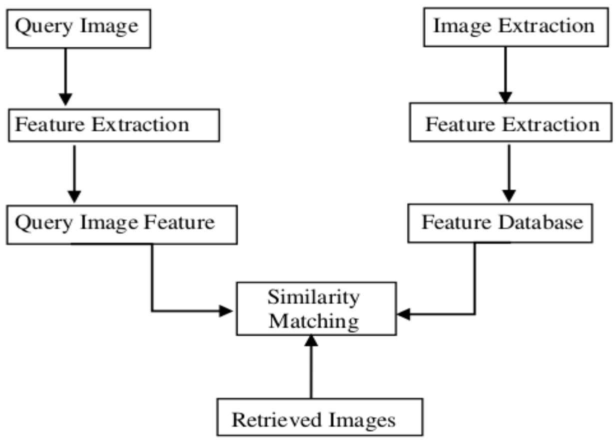

Fig-1: A general Overview of CBIR System.

The term "content" in this context might refers to shapes, colours, textures, or any other information that can be derived from the image itself. The searches that rely purely on metadata are dependent on annotation quality and completeness. Hence CBIR is desirable. Making humans manually annotate images by entering keywords or metadata in a large database can be time consuming and may not capture the keywords required to describe the image. The evaluation of the effectiveness of keyword image search is subjective and has not been well-defined. In the similar context, CBIR systems have similar challenges in defining success. Two main problems faced by CBIR systems are:

- Production of low level image features that accurately describe human visual perception.
- Computational complexity: The high dimensional feature vector gives better information about the image content. It increases the computational complexity when working with high dimensional vectors. Thus CBIR suffers with curse of dimensionality.
- Semantic gap: One of the reasons for poor performance of CBIR is the gap between low level and high level features. Also users viewpoint for same image may be different at different times. Also in CBIR, images with same visual features but different semantics may be considered as identical. This problem is called as semantic gap. The visual features of the litle girl's image and the dog's image are very similar, but their semantic meanings are totally different as shown in Figure 1.These problems are come under semantic gap.

Fig-2: Examples of a semantic gap.

1.2 Related Work

To describe our method clearly, let us first review two areas of research that are closely related to our work in this paper, i.e., 1) CIR, and 2) OED.

A.Review of CIR To reduce the labeling efforts of the user in image retrieval, a variety of research work has been conducted regarding the paradigm of CIR [8], [09], [10]–[12]. Some of the studies have attempted to address the challenges encountered by conventional RF by resorting to the user historical feedback log data or the large-scale web data,. Hoi et al. proposed a log-based RF scheme with SVM by engaging the user historical feedback log data in a conventional online RF task. In [13], a textual query based personal image retrieval system was proposed, which can significantly alleviate the labeling efforts of the user in RF by leveraging millions of loosely labeled web images. Active learning is well-known for getting the necessary information by labeling as few samples as possible. SVMactive is one of the most popular techniques in this category for CIR, which has attracted much attention during the last decade [13], [14]–[17]. To alleviate the small-sized training data problem, Wang et al. proposed to modify SVMactive with a transductive SVM by engaging unlabeled samples in the database. Hoi et al. combined some semi-supervised learning techniques with the traditional SVMactive, which can also effectively exploit the information of unlabeled samples. Despite the vast research work, SVMactive methods always require an initial optimal hyperplane to identify the most informative samples. However, this optimal hyperplane will not always be accurate with insufficient and inexactly labeled feedback samples, which is always the case in image retrieval. Besides the aforementioned methods, some other research efforts have also been devoted to CIR [18], [19]. In [20], a batch model active learning framework was proposed to employ the Fisher information matrix as an ambiguous measurement to select the most informative samples, which is fundamentally based on a probabilistic framework of the kernel logistic regression model. He employed an experimental design criterion to identify one sample after another with a greedy strategy, which does not have a clear interpretation to the selected samples and is usually not the optimal solution to select multiple informative samples for the user to label.

1.3 Relevance Feedback

A question that naturally emerges is, what can we do to deal with these problems? The answer is introducing the users to the process, having them interacting and telling what is ultimately relevant to the images being retrieved andanalyzed. Therefore, by gathering the user's indications, algorithms can be developed to change the placement of the query, or to change the similarity function employed inorder to better comply with the user's expectations. The approach that asksto the user to set the relevance of the images to a given query and to reprocess it based on the user's feedback is called relevance feedback (RF) and it been proven to be quite effective in bridging the semantic gap

The conventional process of RF is as follows:

- From the retrieved images, the user labels a number of relevant samples as positive feedbacks, and a number of irrelevant samples as negative feedbacks.

- The CBIR system then refines its retrieval procedure based on these labeled feedback samples to improve retrieval performance.

1. Subspace learning based methods [21][22][23] define a- class problem and find a subspace within which to separate the one positive class from the unknown number of negative classes. Few of the methods come under this category are: biased discriminant analysis or BDA, the direct kernel biased discriminant analysis (DKBDA),marginal biased analysis (MBA) [24].

2. Support vector machine (SVM) based methods [25][26] either estimate the density of positive instances or regard Relevance Feedback as a classification problem with the positive and negative samples as training sets. SVM active-learning selects the samples near the SVM boundary and queries the user for labels. After training, the points near the SVM boundary are regarded as the most informative images while the most- positive images are the farthest from the boundary on the positive side.

3.Random sampling-based methods apply statistical sampling techniques to reduce particular problems in Relevance Feedback which occurs in previous two methods.

For example, the asymmetric bagging random subspace scheme [27].

4.Feature selection-based methods adjust weights associated with various dimensions of the feature space to enhance the importance of those dimensions that help in retrieving the relevant images and to reduce the importance of those dimensions that hinder the retrieval performance. Alternatively, features can be selected by the boosting tech unique, e.g., AdaBoost,[18],in which a strong classifier can be obtained as a weighted sum of weak classifiers along different feature dimensions.

2. IMPLEMENTATION DETAILS

Relevance Feedback (RF) is one of the most powerful techniques to bridge the semantic gap by letting the user label semantically relevant and un-relevant images, which are positive and negative feedback samples respectively. One- class support vector machine (SVM) can calculate approxi mately the density of positive feedback samples. Concerning the positive and negative feedback samples as two different classes, Relevance Feedback can be considered an online binary classification problem. This is the reason for finding better classifier, which can classify the images in the database based on user feedback. Two-class Support Vector Machine was widely used to build the Relevance Feedback schemes due to its superior generalization ability. With the observation that all positive samples are alike and each negative sample is negative in its own way, Relevance Feedback was formulated as a biased subspace learning problem, where there i s an unknown number of classes, but the user is concerned only about the positive one.

A. Support Vector Machine (SVM)

Support vector machine (SVM) active learning can select ambiguous samples as the most informative ones for the user to label with the help of the optimal hyperplane of SVM, and thus alleviate the labeling efforts of conventional Relevance Feedback. To explain the mechanism of SVMactive, a simple example of a toy is illustrated in Figure 3.

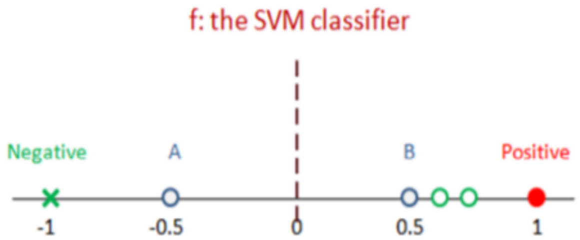

Fig-3: SVM Classifier

There are two labeled samples (i.e., the red solid dot for the positive feedback sample while the green cross point for the negative feedback sample) and several unlabeled ones (i.e. open circles). The six samples distribute along a line and the coordinates on the horizontal axis indicate their positions. By using the SVM, the optimal hyperplane of the classifier f, which separates the two labeled feedback samples with a maximum margin, crosses position 0 as shown in **Figure 3** with the dashed line. According to the most ambiguous criterion, i.e., the samples closest to f have the maximum ambiguity, we can get that A and B have the maximum and identical ambiguity because they have the same distance, i.e., 0.5 for both, to the optimal hyperplane. Therefore, A and B should be identified by the user and used as the training data in Relevance Feedback. If we can choose only one sample for labeling, it is more reasonable to label B than A since more unlabeled samples are distributed around B and thus B is more effective than A to represent the distribution of unlabeled samples in the database. However, SVMactive can only select the ambiguous samples for the user to label although labeling representative ones may bring more useful information for achieving much better performance. Moreover, the optimal hyperplane of SVM is always unstable with small sized training data i.e. this hyper plane is always sensitive when the size of the training data is small. Generally, in Relevance Feedback, the user would only label a small number of samples and cannot label each sample accurately all the time. Therefore, the optimal hyperplane of SVM cannot always be accurate with insufficient and inexactly labeled feedback samples.

B. Mathematical Model

A user gives the query in the form of image Iq. The system retrieves Top K images from the image database. On retrieved images, a user will give feedback as positive and negative image samples. These image samples are then used for th e feature will be further given classifier for classification. The proposed system S is defined as follows:

$S = \{I, Iq, F\,D, OI, RF, FS, F, ORF\}$

where, $I = \{I1, I2, \dots \dots \dots IN\}$

I = set of images in a database.

N = number of images.

Iq = Query image.

FD= $\{ 1, FD2, \dots F\,DN \}$

FD = set of vectors in the database

N = Number of images where FDi ={ FDi1, FDi2, ...FDid} is a set of d features associated with each feature vector.

OI = { OI1, OI2, ...OIk}

OI = set of retrieved images as output

ORF = set of positive and negative labeled samples given by user on retrieved images

FS = { FS1, FS2, ..FSM}

FS = set of features selected from the feature database

F= { F1, F2, ...}

Where F is a set of functions. The functionality of this system is to output top K images, which are relevant to the query image given by the user.

C. Process Block Diagram:

The block diagram for proposed system is shown in Figure4Relevance Feedback approach consists of different stages

- **Retrieval:** These are the retrieved images which are relevant to the query image provided by the user.
- **Relevance Feedback:** Now the user will ask to label the images as relevant or not relevant to positive and negative feedback samples.
- **Feature Selection:** The features which are most

  dominating are selected from the relevance of

  positive images.
- **Binary Classifier:** This feedback given to the

  classifier as a training data for classifying the images in the database into two classes as positive and negative.
- **Re-ranking:** After classification, the images in the database are ranked again.

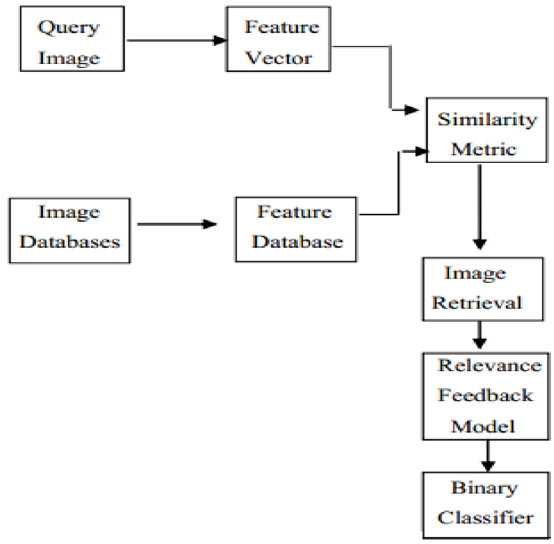

Fig-4: System Architecture Diagram

3. RESULTS

A. Experimental Setup

In order to assess the performance of the proposed method, an image set containing 1000 images from the Corel database of natural jpg images is used. Initially, all the images in the database are used once as queries. In each Relevance Feedback round, at most 3 relevant images are to be selected. These images are used in combination with the examples provided in the previous Relevance Feedback rounds to select a number of important features(K) and, then, to train a new SVM classifier in the resulting lower dimensional feature space. Based on this new classifier, the ranking of the database images is updated. For the initial ranking, when no feedback examples have been provided yet and, hence, neither feature selection nor classifier training can be employed, the Euclidean distance in the initial feature space is used.

B. Results

The precision and recall will be computed to evaluate the performance of retrieval system.

Precision= The Number of relevant images retrieved / Total Number of relevant images.

Recall= Number of relevant images retrieved/Total Number of images retrieved.

Retrieved Images	Curvelet Exiting System(CBIR)		Proposed System	
	Precision	Recall	Precision	Recall
10	1	0.1	1	0.1
20	1	0.2	1	0.2
30	0.83334	0.25	0.9334	0.28
40	0.8	0.32	0.875	0.35
50	0.72	0.36	0.76	0.38
60	0.65	0.39	0.6778	0.42
80	0.6	0.41	0.65	0.43
90	0.55	0.43	0.611	0.45
100	0	0	0	0

Table no.-1: Comparison Results For CBIR System And Proposed System

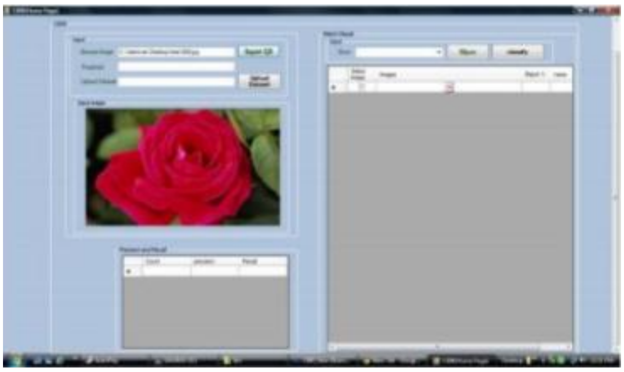

Fig-5: Query Image

Figure 5 shows given images is input image which is given to system for finding similar images.

Fig-6: CBIR Result

Figure 6 shows the result of CBIR system but this system gives an inaccurate result.

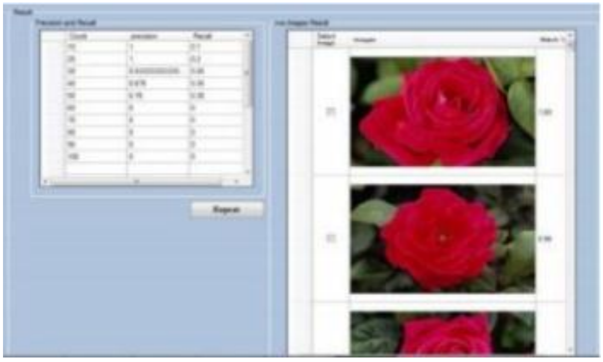

Fig-7: Improve Result

Figure 7 shows, using improve system user get most relevant images according to user requirement. Also for a given query, precision is calculated at different recall values considering entire database and is displayed.

4. CONCLUSIONS

CBIR system is presented. This approach uses binary classifiers to distinguish between the classes of relevant and irrelevant images, along with an SVM-based feature selection technique. As compared to existing systems, proposed system may give the better retrieval results. The precision and recall will be computed to evaluate the performance of retrieval system.

Finally, Outcomes of this system are:

1. When a query image is given to the system all images in the ranked as per their relevance to query image and top K images are retrieved.

2. Worst, moderate and best case queries are selected to study experimentally the effect of Relevance Feedback on system performance.

REFERENCES

[1]A. Smeulders, M. Worring, S. Santini, A. Gupta, and R.Jain, "Content-based image retrieval at the end of the early years," IEEE Trans. Pattern Anal. Mach. Intell. vol. 22, no. 12, pp. 1349–1380, Dec. 2000.

[2] Y. Rui, T. Huang, and S. Chang, "Image retrieval: Current techniques, promising directions, and open issues," J. Vis. Commun. Image Represent., vol. 10, no. 1, pp. 39–62, 1999.

[3] R. Datta, D. Joshi, J. Li, and J. Wang, "Image retrieval:Ideas, influences, and trends of the new age," ACM Comput. Surveys, vol. 40, no. 2, pp. 1– 60, May 2008.

[4] X. Zhou and T. Huang, Relevance feedback in image retrieval: A comprehensive review, Multimedia Syst., vol. 8, no. 6, pp. 536544, Apr. 2003.

[5] C. H. Hoi and M. R. Lyu, A semi-supervised active learning framework for image retrieval, in Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit., 2005, pp. 302309.

[6] Lei Zhang; Fuzong Lin; Bo Zhang; Support Vector Machine Learning For Image Retrieval, in Image Processing, 2001. Proceedings. 2001 International Conference, pp 721 - 724 vol.2, Oct 2001.

[7] K. Bennett and A. Demiriz, Semi-supervised support vector machines, Adv. Neural Inform. Process. Syst. vol. 12, no. 1, pp. 368374, 1998.

[8] C. Hoi, M. Lyu, and R. Jin, "A unified log-based relevance feedback scheme for image retrieval," IEEE Trans. Knowl. Data Eng., vol. 18, no. 4, pp. 509–524, Apr. 2006.

[9] Y.Liu, D. Xu, I. Tsang, and J. Luo, "Textual query of personal photos facilitated by large-scale web data," IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 5, pp. 1022–1036, May 2011

[10] L. Si, R. Jin, C. Hoi, and M. Lyu, "Collaborative image retrieval via regularized metric learning," Multimedia Syst., vol. 12, no. 1, pp. 34–44, 2006

[11] C. Hoi, W. Liu, and S. Chang, "Semi-supervised distance metric learning for collaborative image retrieval and clustering," ACM Trans. Multimedia Comput. Commun.,Appl., vol. 6, no. 3, pp. 1–26, 2010.

[12] L. Zhang, L. Wang, and W. Lin, "Conjunctive patches subspace learning with side information for collaborative image retrieval," IEEE Trans. Image Process., vol. 21, no. 8, pp. 3707–3720, Aug. 2012.

[13] S. Tong and E. Chang. (2001). "Support vector machine active learning for image retrieval," in Proc. 9th ACM Int. Conf. Multimedia, pp. 107– 118 [Online]. Available:http://doi.acm.org.ezlibproxy1.ntu.edu.sg/10.1145/500141.500159

[14] L. Wang, K. Chan, and Z. Zhang, "Bootstrapping SVM active learning by incorporating unlabelled images for image retrieval," in Proc. IEEE Conf. Comput. Vision Pattern Recognit., 2003, pp. 629–634.

[15] C. H. Hoi and M. R. Lyu, "A semi-supervised active learning framework for image retrieval," in Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit., 2005, pp. 302–309.

[16] C. Dagli, S. Rajaram, and T. Huang, "Leveraging active learning for relevance feedback using an information theoretic diversity measure," Image Video Retrieval, Lecture Notes in Comput. Sci., vol. 4071, no.1, pp. 123–132, 2006.

[17] C. H. Hoi, R. Jin, J. Zhu, and M. R. Lyu, "Semisupervised SVM batch mode active learning with applications to image retrieval," ACM Trans. Inform. Syst., vol. 27, no. 3, pp. 16:1–16:29, May 2009.

[18] M. Belkin, P. Niyogi, and V. Sindhwani, "Manifold Regularization: A geometric framework for learning from labeled and unlabeled examples," J. Mach. Learn. Res., vol. 7, pp. 2399–2434, Dec. 2006

[19] C. Hoi, R. Jin, and M. Lyu, "Batch mode active learning with applications to text categorization and image retrieval," IEEE Trans. Knowl. Data Eng., vol. 21, no. 9, pp. 1233–1248, Sep. 2009.

[20] X. Tian, D. Tao, X. Hua, and X. Wu, "Active reranking for web image search," IEEE Trans. Image Process., vol. 19, no. 3, pp. 805–820, Mar. 2010.

[21] Y. Fu and T.-S. Huang, Image classification using correlation tensor analysis, IEEE Trans. Image Process.,vol. 17, no. 2, pp. 226234, Feb. 2008

[22] D. Tao, X. Li, Wu, and S. J. Maybank, Geometric meanfor subspace selection, IEEE Trans. Pattern Anal. Mach.Intell., vol. 31, no. 2, pp. 260274, Feb. 2009.

[23] M. Sugiyama, Dimensionality reduction of multimodal labeled data by local Fisher discriminant analysis, J. Mach. Learn. Res., vol. 8, pp. 1,0271,061, May 2005.

[24] Tao, Dacheng; Tang, Xiaoou; Li, Xuelong and Rui, Yong . Direct kernel biased discriminant analysis: a new content-based image retrieval relevance feedback algorithm. IEEE Transactions on Multimedia 8 (4), pp. 716-727. ISSN 1520-9210. 2006.

[25] S. Tong and E. Chang, Support vector machine activelearning for image retrieval, in Proc. ACM Int. Conf.Multimedia, pp. 107 118, 2001

[26]. Lei Zhang; Fuzong Lin; Bo Zhang; "Support Vector Machine Learning For Image Retrieval," in Image Processing, 2001. Proceedings. 2001 International Conference, pp 721 - 724 vol.2, Oct 2001.

[27]. D. Tao, X. Tang, X. Li, and X. Wu, "Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval," IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 7, pp. 1088–1099,