Learning in CBIR

G.Ramya, D.Mounika, G.Swetha
CVR College of Engineering, Hyderabad, India

Abstract— Content Based Image retrieval is a task of retrieving images which are visually similar to the query image, from the database. Relevance feedback is commonly used to improve the performance of the CBIR system by incorporating user feedback iteratively. Main aim of our work is to infer a semantic space from the user's relevance feedback so that the system will improve the performance of retrieval. This is accomplished by using both short term and long term learning strategies. In short term learning through relevance feedback the query vector is modified according to the relevant images feature vectors. A long-term learning creates a semantic space from the feedbacks given by the user in short term learning. The idea is that after several rounds of feedback the user has a pool of images relevant to the particular query, aggregating these results we construct a semantic space with a concomitant improvement in the system performance .We use SVD to reduce the dimensionality of semantic space both for savings in storage and for improvement in retrieval performance.

Keywords—*Image , Content based,Relevance feedback.*

I. INTRODUCTION

Due to the rapid growth of number of digital images, there is an increasing demand for effective image management tools. Conventional content based image retrieval systems (CBIR) systems uses low level features automatically extracted from the images themselves to search for images relevant to a user's query. While there are research efforts to improve the performance by using different low-level features, and by modifying the similarity measures constructing from them, it is argued that, as unconstrained object recognition is still beyond the reach of current technology, these content based systems can at best capture only pre -attentive similarity, not the semantic similarity. The paper is organized as follows. Chapter II gives an overview of traditional CBIR systems .chapter III is about the literature survey. Algorithm is explained in chapter IV. Experimental results are given in chapter V. Conclusion and future work in chapter VI.

II. INTRODUCTION TO CBIR

Digital contents are being generated at a dramatically speed. The problem of locating a desired image in an enormous collection becomes very difficult. Therefore the need of an efficient method to retrieve digital images is recognized by the public. There are two approaches to image retrieval, Text Based approach and Content Based approach. The former solution is a more traditional approach which indexes images by using keywords. The keyword indexing of digital images is useful but requires a considerable level of effort and often limited for describing image content. The alternate approach, the Content-Based image retrieval indexes images by using the low level features of the digital images and the searching depends on features being automatically extracted from the image. Content Based Image Retrieval is the term used to describe the process of retrieving images from a database on the basis of the internal features of images. In CBIR, digital images are indexed by summarizing their visual contents through automatically extracted features such as texture, color and shape. There exist different ways to express the query. CBIR retrieves stored digital images from a collection by comparing features extracted from the images. The most common features used are mathematical measures of color, texture or shape. The CBIR system [1],[2],[3] identifies those stored images Whose feature values match those of the query most closely, and displays these found images to the user.

III. LITERATURE SURVEY

A. Relevance Feedback:

Relevance Feedback is commonly used to improve the performance of CBIR [6] system by allowing incorporation of user relevance feedback iteratively. The main idea behind Relevance feedback is as follows: the user submitting the query image to the system, it compares the query image to each image in the database and retrieve the images that are nearest to the query image. User marks some of the images as relevant and non relevant according to his/her information needs. If the user is not satisfied with the results, system updates the feature vector using relevant and non-relevant images to get better results in next iterations. The process is repeated until the user is satisfied or the results cannot be further improved. The RF techniques provide a way to bridge the gap between the machine subject in terms of low-level features and the human subject that is driven by high-level semantics.

B. Relevance Feedback Techniques:

Let the query image and a data base image be represented by feature vectors $X = (x_1; x_2; x_3;\dots ;x_d)$ and $Y = (y_1; y_2;\dots;y_d)$ respectively, where d is the number of selected features and x_i and y_i are the values of the i^{th} feature. The system derives the similarity between X and Y by computing the distance under the given dissimilarity Metric. The normalized [1] Euclidean metric

$$Dist(X, Y) = \sqrt{(\Sigma (X_i - Y_i)^2/d)}$$

is generally used for this purpose. The top t database images that are the nearest neighbors of the query are then returned to the user. If the user is not satisfied with the retrieved results, he or she can activate an iterative RF process until satisfied.

IV. APPROACH

1) Existing Procedure:

1. Extract the low level features of all the database images and also the query image. The low-level features we considered here are color, entropy, and contrast and occupancy ratio. In a

similar manner extract the features of the query image; it will be vector of dimension $1\times n$.

2. Compute the Manhattan distance between the database image feature vector and the query vector.

3. After computing the distance the top 20 images are displayed to the user to give the feedback .The images which are relevant to the query is marked as 1 and the images which are non-relevant are marked as 0.Thus we get a feedback vector for the query image which of size 20×1.We call the feedback matrix as F.

4. By using the feedback matrix F, modify the query vector by taking the average of the features of all the relevant images to that query.

5. Compute the dot product between the modified query vector and the feature vector of the database image which was given feedback. Set a threshold and if the value of the dot product exceeds that threshold keep the corresponding value as one else keep the corresponding value as zero. By this we will get a matrix H of dimension 20×1 .The matrix H is the machine generated feedback matrix.

6. Compare this machine generated matrix H with the feedback matrix F and update the query vector by a factor of α according to positive mistakes or negative mistakes.

7. Compute the dot product between the above modified query vector and the images in the database for which the feedback was given and again a generate a new H. Repeat steps 5 and 6 until the H matrix is stabilized.

8. Using the stabilized H matrix and a classifier dune H to the whole database and get the H matrix of size $m \times 1$. Repeat steps 1 to 8 for all k queries to get the matrix H of size $m \times k$.

NOTE: Short term learning is always used only for that query image

9. The above steps end the short term learning session. For storage and performance requirements SVD is computed on H matrix. The rank of the reduced matrix is equal to the number of classes of images in the database.

$H = [U \times S \times V^{'}]$

The B matrix which contains the low-level features is replaced by the product of $U \times S$. This will be considered as the semantic space.

10. For a new query, first extract the low level features and compute the Manhattan distance between original B and new query feature vector.

11. Get the feedback for top 20 images and mark relevant and non relevant images. Call this feedback matrix as F_n which consists of ones and zeroes. The images which is relevant to the query is given a one otherwise a zero.

12. Modify the query vector by taking the average features of the relevant images from the modified B.

13. Compute the dot product between the modified B and the modified new query vector .Set a threshold and if the value of the dot product exceeds that threshold keep the corresponding value as one else keep the corresponding value as zero. By this we will get a matrix H which consists of ones and zeroes.

14. Compare the matrix H with the feedback matrix F and update the query vector by a factor of α according to positive mistakes or negative mistakes.

15. Compute the dot product between the above modified query vector and the images in the database for which the feedback was given and again generate a new H.

16. Repeat the steps 14 and 15 until the H matrix is stabilized.

17. Thus the final results are obtained.

II) Modified Procedure:

1. Construct matrix B of size $m \times n$, matrix Q of size $1 \times n$ and matrix F of size $1 \times k$, where m corresponds to the number of images in the database and n corresponds to the number of features; K corresponds to the query images.

2. Compute the Manhattan distance between the matrix B and Q and get a resultant matrix of size $m \times 1$.

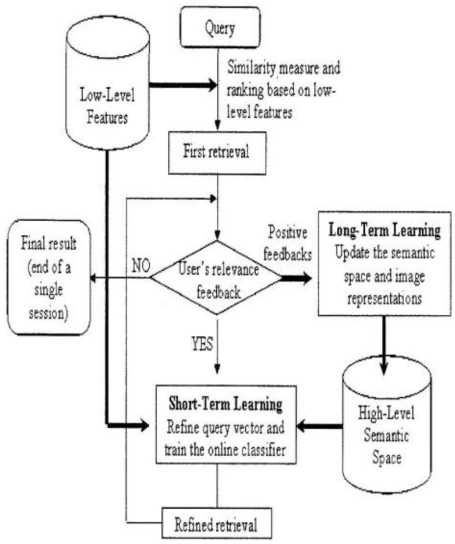

Fig1: Block diagram of proposed system

3. Rank order the values in the R matrix and display the top 20 images to the user.

4. From the displayed images the user marks relevant and non relevant images, apply Query Vector Modification (QVM) to modify the query vector.

5. Compute the dot product between the modified query vector and the B matrix and get resultant R matrix. Set a threshold and, if the value in the R matrix exceeds a threshold keep a 1 otherwise keep a 0 in this manner we get a matrix H of size $m \times 1$.

6. Compare the H matrix with F matrix and modify the query vector by a vector of alpha according to the positive or negative mistake by using mistake driven learning algorithm.

7. Repeat the steps 5 and 6 until the hamming distance between the F and H matrices is minimized. Repeat the steps 1- 6 for k queries to get H of dimension $m \times k$.

8. This ends the short term learning session and after this appends the matrix H to the matrix B. By appending the H matrix to the B matrix the low-level features are combined along with the relevant and non relevant information of an image to the particular query. Compute the svd of the matrix B augmented with H to reduce the space.

$[U, S, V] = svd(B|H)$

Now the B new will be the product of U × S.

9. For a new query extract the low-level features and compute the distance between the new query feature vector and the B matrix which contains the database image feature vectors, display the top 20 images to the user to give the feedback.

10. From the feedback given by the user pick the relevant images feature vectors and non relevant feature vectors from the matrix B_n (it is a modified B matrix which is considered as the semantic space.) ,and apply the QVM to obtain the modified query feature vector.

11. Compute the dot product between the modified query vector and the B_n, set a threshold and if the corresponding value exceeds that threshold keep a one otherwise keep a zero. This matrix is considered as the H matrix.

12. Compare this H matrix with the feedback matrix and modify the query vector according to the mistake driven algorithm.

13. Repeat the above step until the hamming distance between the H and F is minimized.

14. Display the images to the user which are having their values as one in H matrix. Thus the final results are obtained.

V. EXPERIMENTS AND RESULTS

Fig 2 is the query image. Given this query image first retrieval (fig 2) are the results which are obtained by computing the Manhattan distance between the query images and the database images.Table1 shows the short term learning output which is obtained by modifying the query vector by taking the average of the relevant image feature vectors and then updating it by a factor of alpha according to positive or negative mistakes as explained in the procedure.

Table 2 and 3 shows the long term learning outputs. Here the performance of these methods is same with respect to the query which is given in figure1

Fig 2: Query Image

Fig 2: First Retrieval

Table1: Short term learning output

Table 2: Long Term Learning Output

Table 3: Modified Learning Output

VI. CONCLUSION AND FUTURE WORK

A learning frame work was described which makes use of relevance feedback to enhance the performance of an image retrieval system from both short term learning and long term learning perspectives. The proposed long term learning scheme infers a semantic space from user's interactions. A method of updating the semantic space and guidelines for choosing the optimal dimensionality were also discussed. The learned semantic space supplements the low-level features in making image search result more satisfactory to the user. The semantic space is constructed by doing svd on H which only captures semantics arrived from relevance of an image. Hidden semantics of low level features are ignored. To extract meaningful semantics method is modified by constructing the semantic space which includes both low-level and high level features. It is seen that it will have more meaningful semantics since the facts of the images and the expert perceptions are coupled.

Currently, each image belongs to one and only one semantic category; further research could analyze how results change when images belong to multiple semantic categories. One can also use gradations in constructing the semantic space instead of ones and zeroes and see the performance.

VII. REFRENCES

[1]. XiaofeiHe, OliverKing, Wei-Ying Ma, MingjingLi, Hong-Jiang Zhang, "Learning a Semantic Space From Users Relevance Feedback for Image Retrieval." IEEE Transactions CSVT,Vol.13,No.1,pp.39-48,jan2003

[2]. M-Shan, J.Xiaojun Qi, "Short term and Long-term learning approach for Content based image retrieval". Proc IEEE Conference, volume: 2, pp: II-II, 2006

[3]. Praks P., Dvorsky J., Snasel V, "Latent Semantic Indexing for Image Retrieval System", Proc. of the SIAM Conference on Applied Linear Algebra, Williamsburg, July 15-19, 2003.

[4]. S. Deerwester, S. Dumais, G. Furnas, T. Landauer, R. Harshman, "Indexing by La-tent Semantic Analysis." Journal of the American Society for Information Science, 41(6):391-407,1990

[5]. Ritendra Datta and Dhiraj Joshi and Jia Li and James and Z. Wan, "Image retrieval: Ideas, influences, and trends of the new age ACM Computing Surveys," volume 39, pp .2007, 2006.

[6]. T.Yoshizawa and H.Schweitzer. "Longterm learning of semantic grouping from relevance feedback." proc of Intl workshop on MIR pages,pages 165- 172,NY,USA,2004.ACM

[7]. N.Littlestone "Learning quickly when irrelevant attributes abound : A new linear- threshold algorithm," Mach.Learn., vol.2,no.4,pp.285-318-1988.

[8]. I.Dagan,Y.Karov,and D.Roth,, "Mistake Driven learning in text categorization", in Proc.2ndConf.Empirical Methods in Natural Language Processing,1997,pp.55-63.

[9]. W.Y.Ma and B.S.Manjunath, Netra :" A toolbox for navigating large image databases", ACM Multimedia Syst.,vol.7,pp.184-198,1999.

[10]. YingLiu,DengshengZhang,GuojunLu,Wei-YngMa, "A survey of content based image retrieval with high-level semantics", Pattern Recognition,v.40 n.1,p.262- 282,2007.

[11]. Jiang, Wenchao and Wan, Baoping and Zhang, Qijun and Zhou, Yanhong, "Image Search by Latent Semantic Indexing Based on Multiple Feature Fusion" CISP 31 '08: Proceedings of the 2008 Congress on Image and Signal Processing, Vol. 2,2008.pp:515-519

[12]. Peter Praus,Pavel Praks,, "Automatic Retrieval with in water related pictures database using latent semantic indexing method" ISSN Vol 2 pp:19-28-2009.

[13]. Chun Che Fung and Kien-Ping Chung, "Establishing semantic Relationship in Inter-query Learning for Content Based Image Retrieval Systems" PAKDD, pp: 498-506,2007

[14]. BING WANG ,MEI-WU HE,SHUO WANG,MIAO WANG, "Image Retrieval using long term learning relevance feedback" Proceedings of ML and Cybernetics,Hongkong,pp:19-22,2007.