Memory Learning using Cluster Assimilation approach in CBIR

Dr. Renuka Devi SM
Associate Professor
Department of Electronics & Communication Engineering
GNITS Hyderabad

Abstract *— A elementary problem in CBIR is Semantic gap, between the low-level features and high-level semantic as perceived by humans. Relevance Feedback is one solution to bridge this gap. Relevant images corresponding to a query form many clusters in feature space. This concept is used in framing this method of Cluster assimilation, memory learning in CBIR. The clusters of RF images associated to a query are labelled with query index. These clusters of many queries are merged based on some criterias like overlapping, nearness etc., to obtain overall merged clusters with associated tagged query indices. Experimental results illustrate the potential of proposed method of memory learning by cluster assimilation or integration in initial and as well over the iterations. The efficiency of our method is shown using standard Wang database of 1000 images.*

Keywords— Memory Learning, Inter-Query Learning, Relevance Feedback, Cluster Assimilation, Mahalanobis Distance, RF Log

I. Introduction

Ease of capturing and sharing images has led to huge collection of databases both for professional and personal use. Such data is usually accessible via the internet through photo sharing sites such as Picasa, Flickr, Panoramic and also as personal collections on local hard disks. This has led for the management of large database and hence Content Based Image Retrieval(CBIR) has emerged as active research area. CBIR mainly deals with querying for relevant images based on low-level features such as color, texture and shape, and also uses others features such as spatial information and annotations.

The holy grail of CBIR research is to bridge the semantic gap between the machine perceiving of image and the need of users. Relevance Feedback (RF) is one of the technique to bridge this gap. Here the user gives feedback on retrieved images, and the system will respond (by learning) to the users semantics in subsequent retrieval. Learning refers to improving the throughput performance of the system over iterations using retrieval of only current query or across several queries. Hence, accordingly termed as Short-Term Learning (STL) or Long Term Learning (LTL) respectively. The LTL is also called memory learning or inter-query learning.

Plenty of work exists in literature on LTL, on how well the RF log is exploited to bridge the gap faster. All the existing related work in LTL is based on summarizing the RF log for the concept information using techniques like latent semantic analysis [10, 11], pLSA [20] or generating the semantic associations using like apriori algorithm, Fuzzy SVM [12], or performing semantic clustering the data based on RF input using techniques like semi-supervised Fuzzy-C means (SSFCM) [2], semi-supervised EM (SSEM)[1] clustering. In all of these methods, semantic relations form the commonality and the bases for RF log in the LTL based CBIR system.

The main motivation of this proposed LTL by cluster merging is, for all the images belonging to the same conceptual category there is a considerable variance in low-level features to be understandable by a machine. Some of the sub-category definitions could be “a single yellow rose”, “single rose of any color”, “two roses” etc.,. Although all of these subcategories are conceptually same, the human brain is automatically processing and creating the picture of the concept. But the machine requires to learn this by earlier queries and use the learnt knowledge for subsequent queries. This is because of the reason that conceptual consistency does not imply descriptive low-level features consistency in color and texture across images within a category. These sub-categories of a concept that form a localized cluster in low-level features are attempted to be captured in a component or cluster and later use these clusters by tagging them with the query index. Each subcategory of a concept forms a localized cluster/island in feature space. Here we have proposed a novel method in the direction of building the semantic space from the RF log. Our method tries to give more coverage over the search space in contrast to Mahalanobis and is in-line with the fact that relevant images form islands[4] in the feature space. This approach believes in the fact that a query has set of relevant images and vice-versa a relevant image has set of queries. This belief is incorporated in this method in tagging the relevant image clusters(islands) of a query by a query index. After many such clusters corresponding to various queries are accumulated, these are thereby merged, based on some criteria that decides nearness and overlapping. The initial retrieval set in LTL consists of images retrieved from RF log based on tagged query indices and also from the STL via Mahalanobis distance over the database images. This proposed LTL by cluster merging approach is found to give better results in initial and as well over iterations compared to only STL by connected component(CC) approach[4]. We demonstrate the effectiveness of our method over a standard database of 1000 images corel Wang database. The proposed method has the advantage of storing just the merged cluster Information along with the tagged query indices instead of saving the entire RF log [12] or the semantic rules [28], that is required to develop semantic associations in accordance with the current query. Further our method of retrieval is fast where in it is assumed that merging the clusters happens when the system is idle. The contribution of this paper is mainly in two directions 1. This is a novel idea to enhance the search space for retrieval of similar semantic images over that of Mahalanobis distance metric.

In the remainder of the paper, Section 2 discusses the related work, and section 3 gives the architecture of the proposed system along with the details of the proposed approach. Section 4 is devoted to the promising experimental results and finally section 5 presents the conclusion and future work.

II. RELATED WORK

RF is an interactive process by which the system learns the high-level concepts of the user. The learning of semantics can be over a single query or across queries, which is accordingly STL or LTL. Previous research on STL is usually classified into QVM, where in refined query obtained moves towards relevant images and away from irrelevant images, FRE [25, 14] where the feature dimensions are weighed and finally Probability based methods, that estimates the likeness of an image being relevant given a set of relevant and non-relevant images[13]. Some of the other ways to achieve STL include discriminant methods like LDA,kBDA [6], and generative models like GMM [15], non-parametric models [13] and finally the more recent approach uses Manifold learning methods like LLE[18], Isomap etc.,. In our previous work [3] a novel method of integrating unsupervised clustering with RF was proposed, where in it was assumed that the clustered data forms an intermediate level of high level semantics.

All the existing research in LTL is based in either of the directions like using Latent Analysis to construct semantic space [10], building semantic network [8], developing association rules [28], doing semantic clustering or developing a kernel that reflects the semantic information of RF log [7] or learning a semantic mainfold [9, 19] using the more recent dimensionality reduction techniques.

III. PROPOSED SYSTEM

The block diagram of the proposed CBIR system is shown in the figure 1. As shown in block diagram, after a query is submitted the initial retrieval set presented to the user is formed by using the retrieval from Mahalanobis and that over RF log. Then over the subsequent iterations, the learning in based purely on STL by connected component(CC) approach [4]. The components that are collected in the process are tagged by the index corespounding to the query(called query index). The above two steps are repeated for few more queries and the resultant CCs obtained are merged using our cluster merging algorithm. Using these merged clusters RF log is updated. Each of the above three steps namely initial retrieval, STL by CC approach and finally updating RF log are discussed in next subsections. The following subsections deals with each of the steps in detail.

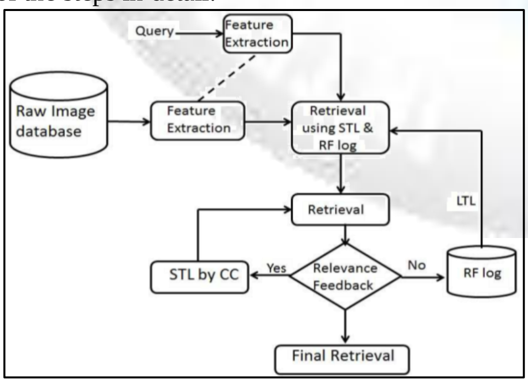

Fig. 1: CBIR system with inter-query learning

A. STL by CC and Indexing Clusters

In our previous works [5], we proposed that Connected components(CCs) in feature space is used to decide the number of components of Gaussian Mixture Model(GMM), and in [4] we proposed STL by CC, based on fitting hyperellipsoid to the CC of relevant images. The connected component is defined as the set of relevant images that are closer to each other than to any non-relevant image. This connected component was found to correspond to a localized highlevel semantic cluster after fitted by Gaussian. Further this approach of STL by CC requires less number of Gaussians to enclose the relevant images compared to Cover approach[15]. As shown in figure 3 queries Q1, Q2, Q3 and Q4 each result in three clusters using STL by CC approach. Thus a total of twelve clusters are formed. Our cluster merging algorithm reduces them to five clusters as shown by solid hyperellipsoid.

B. Cluster Merging and updation of RF log

The hyperellipsoids(clusters) are merged based on the algorithm discussed below. Initially compute the Hausdorff distance[16] between all possible pair of clusters. Hausdroff distance(HD) between two sets of points X = {x1, . . . xn} and Y = {y1, . . . ym} is

$$d_{HD}(X, Y) = \max\{d(X, Y), d(Y,X)\}$$

And d(X,Y) is where d(x, y) is the euclidean distance between two points x,y.

The cluster merging algorithm reduces large clusters to N merged clusters. The illustration of cluster merging is shown in figure 2.

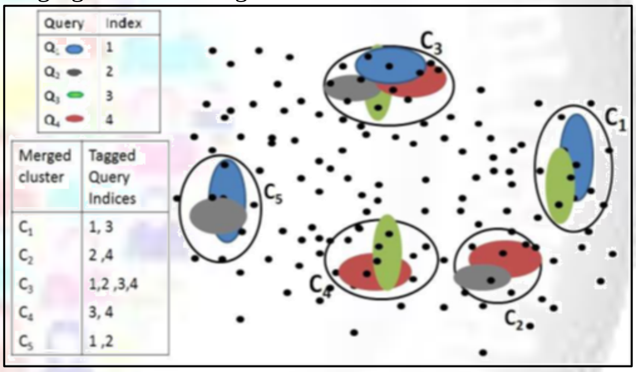

Fig. 2: Illustration of feature space queried for four queries and clusters for each query is shown by same color hyper-ellipsoids. After applying cluster merging algorithm five clusters are formed (solid curves). The query tags for each merged cluster are shown in table.

A merged cluster is defined by Gaussian parameters. The merged cluster so formed is tagged by query indices, if it has relevant image of a given query. As shown in table in figure 3 for example, merged cluster C1 has relevant features from clusters with query index 1 and 3, so tagged query indices of C1 is 1, 3. Similarly tagged indices are generated for other merged clusters.

C. Initial Retrieval of Images

The initial retrieval set to be shown to the user is formed by combination of two ranked list, one as given by RF log and other by Mahalanobis distance (STL) over the database images. Modified Borda Count is used to obtain final ranked list. If ‘k’ images are to be retrieved, the ranked list formed by STL is given weightage of α and the list formed using RF log receives (1−α) weight. We have used α = 0.8 in our experiments. This weight α increases as the system collects/learns more queries.

1) Algorithm 1: Cluster assimilation

- Input: Set of 'c' Clusters/hyperellipsoid
- Output: Set of N Merged clusters
- Step1: Compute Hausdorff Distance between each pair of clusters;
- Step 2: Let A merged(B,C) Where B, C are two clusters at Minimum Hausdroff Distance;
- Step 3: Compute e_A , e_B and e_C;
- Step 4: if $e_A > e_B + e_C$ then
- Step 5: Pick two clusters at next Minimum Hausdorff Distance and go to step 2;
- Step 6: else
- Step 7: Merge two clusters B, C and go to step 1;
- Step 8 Repeat steps 1 to 4 until no two clusters merge;

a) Retrieval using RF log

This approach retrieves top 'k images by searching over the selected merged clusters by using the procedure discussed below:

1) Find the merged cluster(Ct) for which the given query has maximum posterior probability.
2) Using the tagged indices of Ct, compute the Semantic Similarity (SS) of with each of the merged clusters (Cc). SS of Ct with Cc i.e., s(Ct,Cc) is defined as the ratio of the number of indices that are common to both Ct and Cc to the total number of indices in Ct.

$$s(Ct,Cc) \frac{|\text{matched indices of Cc with Ct}|}{|\text{ indices of Ct }|}$$

Where | A | denotes cardinality of set A. All the SS of Ct with other merged clusters forms a vector

$$S_Q = \{s(C_t,C_1), s(C_t,C_2) \ldots , s(C_t,C_c) \ldots , s(Ct,C_N)\}$$

Where some of the $s(C_t,Cc)$ can be zero.

3) Consider only those merged clusters having non-zero SS to the given query Q. Let p_Q, and p_I denote the vector of posterior probabilities fromed from merged clusters having non-zero SS, for given query(Q) and the database image(I) respectively.
4) Find the normalized dot-product between pQ and pI , where $I = \{I_1 \ldots I_M\}$ and M is the number of database images.
5) Rank the top 'k' database images with respect to Query Q, by using the measure of normalized dot-product. Thus the top 'k' images are obtained.

b) Retrieval using Short term learning

The top 'k' images is obtained by rank ordering the database images(I) using the Mahalanobis distance with respect to the query.

$$D(Q, I) = (I - \mu)^T \sum\nolimits^{-1} (I - \mu)$$

Where µ is mean and ∑ is covariance matrix formed by database images. This list is weighed by factor α, in estimating the combined retrieval set as discussed earlier.

IV. EXPERIMENTAL RESULTS

We have experimented with 1000 images Wang database [23] having 10 categories i.e., faces, beaches, monuments, buses, dragons, elephants, flowers, horses, mountains and food. Most of the research in image retrieval is based on automatic RF, but in real time CBIR systems human gives RF. Novelty in our experiments is that we subcategorized each of the broad category based on human similarity. For example all single elephants, two elephants and more than two elephants were grouped into three different categories, while categories like dragon do not have much variation in similarity and so were considered as single category. This was done to justify our motivation that a subcategory of a semantic concept forms an island, and also to depict the real time environment of manual RF. Automatic RF was done on this subcategories of images. The results and discussions show the efficiency of the proposed method of LTL by Cluster Merging.

A. Results on Cluster Merging

For cluster merging process, we have considered four images each from dragon, flower, elephants and food category and then clusters(or hyper-ellipsoids) obtained by STL with CCs were accumulated. The number of clusters formed for each query depends on complexity of query. These clusters so accumulated are around 90 clusters that is given to cluster merging algorithm. This algorithm stops merging when all pairs of intermediate level of merged clusters fail to meet the merging criteria. This has resulted in 19 clusters, of which one of them are shown in figure 3, along with their respective query indexes in figure 4 clarity. The remaining merged clusters are not shown here for want of space. The cluster merging processes is initiated on accumulated clusters from next ten queries and previous merged clusters in RF log. This forms the updated merged clusters of the RF log.

Fig. 3: Merged cluster formed as the result of Cluster merging algorithm

Fig. 4: Query index of above cluster

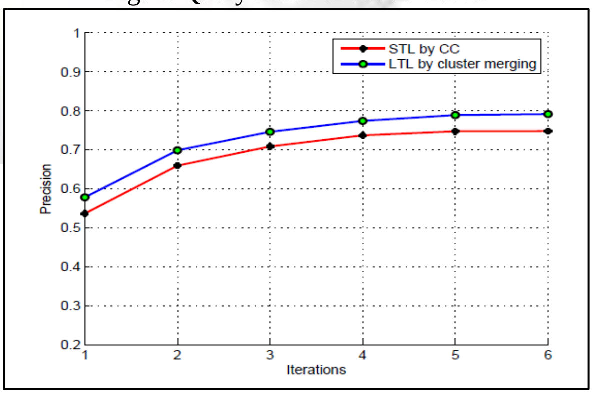

Fig. 5: Comparison of average Precision of proposed LTL by cluster merging over STL by CC

B. Precision and Recall of LTL

The Precision and Precision-Recall(PR) performances are the more widely used performance measures to compare two CBIR systems. Here we evaluate our proposed method of LTL by cluster merging with the STL by CC approach[4]. Precision defines the portion of relevant images retrieved in given number of retrieved images, and Recall is the performance that indicates the portion of relevant images with respect to total relevant images in database, and also the PR curve gives idea of how early the relevant images occur in the set of retrieved images. It is observed in figure 5 that LTL by cluster merging gives improved performance by around 5% averaged over 50 queries, over the Mahalanobis distance used in initial iteration of STL by CC. Once the relevant features are spotted in feature space in initial iteration by the proposed method of LTL by cluster merging, the precision improves further over iterations and is more than STL by CC as observed in figure 5.

The PR curve is plotted at iteration five for 90% recall averaged over 50 queries from different classes. As observed in figure 6, PR curve of proposed method remains at 100% till 35% recall, while that of STL by CC remains so till 22% recall only.

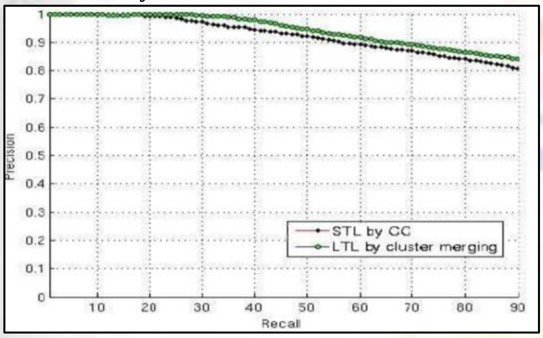

Fig. 6: Comparison of average Recall of proposed LTL by cluster merging over STLby CC.

V. Conclusions

Our proposed method gives an improvement in precision of 5% to 8% in retrieval over short term learning by Connected component approach. The improvement in initial retrieval indicates that this method tries to give more coverage over the search space in contrast to Mahalanobis distance metric. In future the retrieval over RF log can also be attempted with other approaches like dibbling[21]. Further we propose to experiment with larger database such as Flickr, and also think of weight given to STL and LTL be updated as the system learns over many users.

References

[1] Dong Anlei and Bhir Bhanu. Active concept learning in image databases. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 35(3):450– 466, June 2005.

[2] B. Bhanu and A. Dong. Concepts learning with fuzzy clustering and relevance feedback. In Proceedings of the Second International Workshop on Machine Learning and Data Mining in Pattern Recognition, MLDM ’01, pages 102–116, 2001.

[3] Renuka Devi and C. Bhagvati. Connected Components to capture High-level semantics in CBIR. Proceedings of the 4th Annual ACM Bangalore Compute Conference, 2011.

[4] Renuka Devi and C. Bhagvati. A Novel Method for Image Retrieval using Relevance Feedback and Unsupervised Clustering. Proceedings of the 4th Annual ACM Bangalore Compute Conference, 2011.

[5] Renuka Devi and C. Bhagvati. Connected Component Method to Find Components of GMM in Image Retrieval. Proceedings of the International Conference on Computational Intelligence and Communication Networks(CICN), pages 50–54, Nov. 2010.

[6] D.Tao, X.Tang, X. Li, and Y. Rui. Direct kernel biased discriminant analysis: a new content-based image retrieval relevance feedback algorithm. Multimedia, IEEE Transactions on, 8(4):716 –727, Aug 2006.

[7] P. H. Gosselin and M. Cord. Semantic kernel updating for content-based image retrieval. In Proceedings of the IEEE Sixth International Symposium on Multimedia Software Engineering, ISMSE ’04, pages 537–544.

[8] J. Han, K. N. Ngan, M. Li, and H. Zhang. A memory learning framework for effective image retrieval. IEEE Trans. Image Processing, 14:2005, 2005.

[9] J. He, M. Li, H. Zhang, H. Tong, and C. Zhang. Manifold-ranking based image retrieval. In Proc. 12th annual ACM international conference on Multimedia, MULTIMEDIA ’04, pages 9–16, 2004.

[10] X. He, O. King, W. Ma, M. Li, and H. Zhang. Learning a semantic space from user’s relevance feedback for image retrieval. IEEE Trans. Circuits Syst. Video Technology, 13(1):39–48, 2003.

[11] D. R. Heisterkamp. Building a latent semantic index of an image database from patterns of relevance feedback. In ICPR (4), pages 134–137, 2002.

[12] C. H. Hoi and L. Michael. A novel log-based relevance feedback technique in content-based image retrieval. In Proc. 12th annual ACM international conference on Multimedia, MULTIMEDIA ’04, pages 24–31, 2004.

[13] C. Meilhac and C. Nastar. Relevance feedback and category search in image databases. In ICMCS, Vol. 1, pages 512–517, 1999.

[14] J. Peng, B. Bhanu, and S. Qing. Probabilistic feature relevance learning for contentbased image retrieval. Computer Vision and Image Understanding, 75(1-2):150 – 164, 1999.

[15] F. Qian, M. Li, L. Zhang, H. J. Zhang, and B. Zhang. Gaussian mixture model for relevance feedback in image retrieval. In Proceedings of IEEE, pages 229–232.