

# LEARNING IN CONTENT BASED IMAGE RETRIEVAL : A REVIEW

Renuka Devi S M

Assoc. Prof. ECE Department, G.Narayanamma Institute of Technology and Science, HYD-08 T.S. India

## Abstract

*Relevance feedback in Content Based Image Retrieval is an interactive process where the user provides feedback on the system-retrieved images to bridge the gap between user semantics at high level and machine extracted low level features of images. RF exploits Machine Learning and Pattern Recognition techniques for Short Term Learning and Long Term Learning to provide improved performance in retrieval. Intra query and across query learning have received enormous attention over the past decade. This paper first categorizes the various learning techniques and discusses the intuition behind each of these techniques. State-of-art learning techniques ranging from Feature Relevance learning to manifold learning in STL and Latent Semantic Analysis used in text processing to recent kernel semantic learning in LTL are discussed.*

**Keywords:** Relevance Feedback, Short Term Learning, Long Term Learning, Semantic Gap, High Level Features.

\*\*\*

## 1. INTRODUCTION

With the rapid advances in computer technology, and reduced cost of data capturing and storage devices, the volume of digital image data in digital libraries and on internet is growing. As a result, finding images of interest has become more and more like finding a needle in a haystack. Therefore this has raised the need for effective and efficient query formulation and retrieving of images from the database which is termed as Content-Based Image Retrieval (CBIR). CBIR is the research field that attempts to address this issue of finding the images of interest by analyzing and comparing the content of all images in a collection. A comprehensive survey of CBIR field is found in the review paper by S. Antani et al. [8]. The author has cited the URL's of various CBIR systems existed prior to his review. The first commercial CBIR system available to the public was QBIC [18], that emphasized on image and video retrieval which is mostly based on color structure. General purpose CBIR systems, which were introduced later, employed more sophisticated feature vectors, based on color, texture and shape visual descriptors which provided a possibility to use relevance feedback, that allows a better focus on the target of the search. Table 1 briefs some of the CBIR systems along with core technique used and the feature of having incremental learning. In [6], a high-level classification of various image retrieval systems is done based on three criteria viz., segmentation, relevance feedback, and query modification. In [38], a search for an image in CBIR is broadly classified into three categories: target search, category search and web search.

CBIR mainly involves the process of indexing and searching. In the indexing stage each image is stored as attributes of low level features such as color, shape, texture, spatial relationship and annotations. The search stage involves the use of similarity measures to retrieve semantically relevant images. Since the low level features [15] hardly capture the semantic content like house, garden, swimming, etc. of an image, while the humans perceive in

terms of the high level semantic and abstract concepts like happy, religious place etc.. from an image, the performance of CBIR is limited. This gap between the machine perception and human perception is known as semantic gap. Attempts to reduce this semantic gap have lead researches to explore various learning techniques. Learning refers to improvement in performance of a system over current iteration or across query sessions. Various approaches to bridge the semantic gap, ranging from feature learning to generative models and discriminative models to manifold distance learning, are attempted by researchers over the period. There exists two types of learning: Short term learning (STL) or intra query or in-session learning and Long Term learning (LTL) or inter query learning. Various review papers on learning such as Datta et al [14], Lew et al. [33] and Jing Li et al [34] provide a good insight into the more recent advances in the entire field of multimedia information retrieval and in particular learning in CBIR.

**Table 1:** Some of the CBIR systems with core technique used

CBIR system	Core Technique	Incremental
PicHunter [13]	Bayesian learning	No
Viper [39]	Building and searching inverted files	Yes
VisualSEEK [40]	Distance metric includes that best match color, location, size, and area of regions	No
Blobworld [12]	RBIR and Quadratic distance between histograms	No
Photobook [17]	comparing eigen image feature basis set	No
Accio [41]	Multiple Instance Learning	
SIMPLICITY [42]	Integrated Region Matching	No
FISH [43]	Feature weighing	Yes

## 2. FOCUS OF THIS SURVEY

This paper gives exhaustive survey of Short Term and Long Term Learning methods used for improved performance in image retrieval using RF. In particular the paper concentrates on RF techniques using only global

features and not regional features, features obtained after segmentation. The intuition in various recent techniques, present in the area of learning in CBIR is discussed. Section 3, describes the Semantic bridge, understanding of Semantic Concept, Section 4 of ways to give relevance feedback section 5 of features and dissimilarity measures used in CBIR section 6 of general architecture of CBIR with STL and LTL. Section 7 describes the various state of art short term and long term learning techniques. Section 8 lists the comparison of few prominent learning techniques in terms of performance, database and section 9 discusses the scope of research direction in learning techniques. Finally Section 10 gives conclusions.

### 3. BRIDGING SEMANTIC GAP

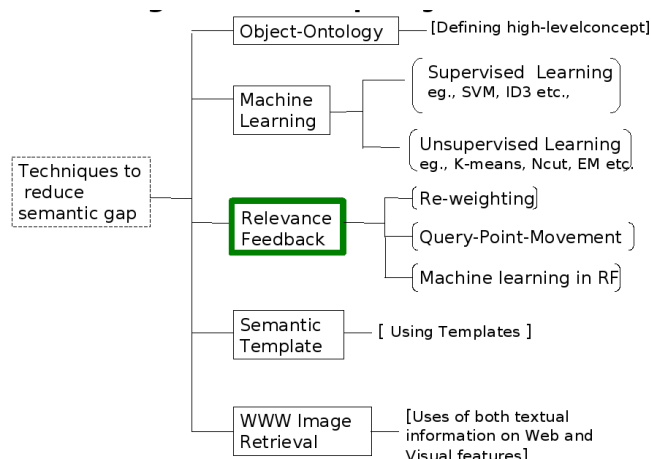


Figure 1: Methods to Reduce Semantic gap[36]

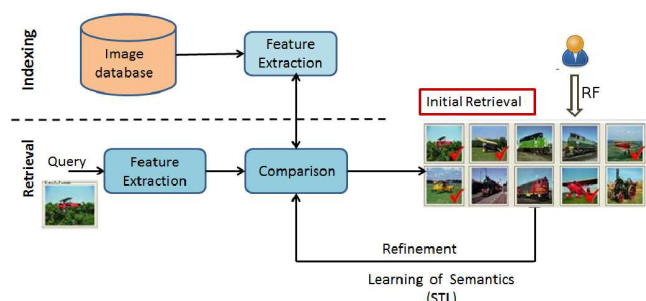


Figure 2: Components of CBIR, along with Relevance Feedback

Ying Liu et al. [36], have painted a detailed picture in this field, learning in image retrieval, and discussed how to successfully use the enormous potential of Pattern recognition and Machine Learning. Figure 1 gives the various approaches to reduce the Semantic gap. In [36], the author has classified the approaches to reduce the Semantic gap into five categories. In each of these approaches, different learning strategies such as offline unsupervised learning, online active learning, semi-supervised learning etc., can be considered to improve the retrieval efficiency of the system. RF is one approach that involves online learning of Semantics of the user. Following are the limitations faced by RF.

1. **Incapability of capturing semantics:** This arises as the amount of the training data is very low when compared to the number of dimensions of the description space.
2. **Scarcity and imbalance of feedback examples:** Usually much fewer positive examples exist than negative examples. The learner must have a low sensitivity to this imbalance in the training set.
3. **High dimensionality of signatures:** Usually in order to capture the semantics of images, high dimensional signatures are used. This becomes a limitation for fastness, and singularity issues.
4. **Need measures that capture nonlinear distribution of relevant images:** Requires generative or discriminative methods, that can capture complex shape distribution of positive labelled images even with small labelled examples for online learning needs.
5. **Need fast scoring measures:** To maintain interactivity, both learning from the labelled examples and scoring of the unlabelled database images, according to the criteria selected must be very fast. Thus computation cost is also very important criterion in the choice of a learning method.

**Query concept :** Semantic is hidden information of a set of images. The term Semantic concept is illustrated in the figure 3.

### 4. WAYS TO GIVE RELEVANCE FEEDBACK:

Most retrieval systems give the user the opportunity to give only positive feedback, positive and negative feedback or positive, neutral and negative feedback or group similar images. In some of the systems the user can give more accurate feedback using four or five levels[30]. Peng et al[8] has shown that with five relevance levels, namely, highly relevant, relevant, no opinion, nonrelevant, and highly nonrelevant, user's subjectivity is difficult to discriminate among nonrelevant and highly nonrelevant retrieved images. Hence Peng et al system with 4 levels excellent, fair, don't care, and bad gives better performance than using five levels.



Figure 3: An example to illustrate the Semantic concept as determined by group of images. When (a,c) are relevant the Query concept is 'yellow flower', When (b, c) are relevant the Query concept is 'rose flower', and When (a, b, c) are relevant the Query concept is 'flower.'

### 5. FEATURES AND SIMILARITY MEASURES:

Visual or low level features for an image can be obtained at global level, or block dividing the image[9] or obtaining regional features by segmentation[31]. The type of the features used, decide the similarity measures. Table 2 lists some of the similarity measures used in CBIR. It is generally accepted that the current state of the art image analysis techniques are not able to capture all meanings that

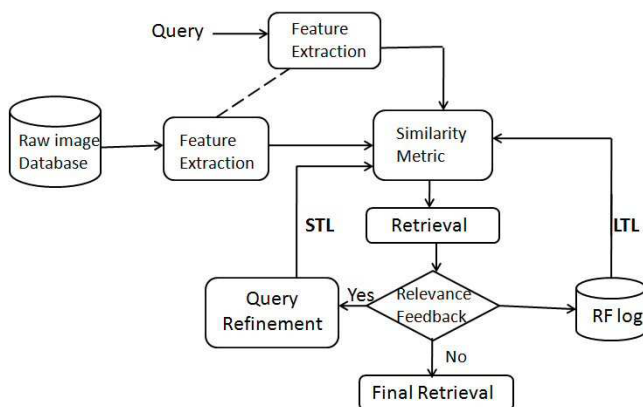
an image contains. To enhance the retrieval process, it is worthwhile to combine image low-level features with other sources of knowledge like text annotations, used to derive keywords from wordnet which relate to the high-level semantics of an image[17].

**Table 2:** Similarity Measures used over the features Intersection

Geometric Measures	Information theoretic Measures	Statistic Measures
1.Minkowski Family 2.Cosine Function based 3.Canberra 4.Squared Chord 5.Partial Histogram	1. Kullback-Leibler Divergence 2.Jeffrey Divergence	1. Chi-square Statistics 2.Cramer/von Mises Type 3.Kolmogorov - Smirnov 4.Pearson's Correlation Coefficient

**6. General CBIR system with STL and LTL:**

Figure 4 shows the general architecture of CBIR system with STL and LTL. When a query is submitted, the visual descriptors of the query is compared with that of database image descriptors (computed offline) by a any of similarity function or distance metric given in table 2. The top few images are retrieved and shown to user for RF. The RF information is used to fine tune the results by any of the learning techniques of STL or LTL discussed below. The RF log can be accumulated and used to improve the retrieval over the next query sessions. The use of RF log across the query sessions is termed LTL.



**Figure 4:** CBIR architecture with provision of STL and LTL

**7. STATEOF ART LEARNING TECHNIQUES**

**7.1 Short term learning:** The idea of Relevance feedback was formulated by J. J. Rocchoi in text retrieval in 1971 and this was adapted to image retrieval in 1997 in the work of Multimedia Analysis and Retrieval Systems (MARS) and MindReader (MR) approach[30] as Query Vector Movement (QVM). In QVM, the modified query moves towards P and away from N images of the user feedback.

**7.1.1 Feature weighing:** The idea of Rui et al.[45] was to give importance to those features that have high consistency(low variance) over relevant set of images marked by the user. Rui et al in [45], discussed the importance of a feature based on inter-weight and intra-weight of the feature. Feature reweighing has evolved in various dimensions.

The simplest weight being the inverse of standard deviation along individual dimensions as used in MARS[45]. This is similar to weighted Euclidean. Later on researchers began to think of generalized Mahalanobis distance metric[30], referred to MR approach. But this has the Small Sample Problem(SSP), when the dimension of signature is greater than number of P. This was resolved by use of Penrose pseudo-inverse method. Another approach to solve the singularity issue is by 'linear shrinkage estimator' i.e., addition of regularization term to covariance matrix. These weights are computed by least square estimate of the conditional density function over the dimensions at which the feature is relevant. This is referred to Probabilistic Feature Relevance Learning(PFRL).

S. N. Prasad et al [44] proposed that the weight of a feature can be decided by the ratio of separation of means of P and N to the sum of standard deviations of P and N respectively. Here the weight is decided by considering P and N to be clusters. But this fails when the distributions of these lusters are multimodal. This is similar to Fishers Discriminant Analysis(FDA)[11, 16]. This refers to entropy based disimilarity between the P and N disributions. The intuition was, for a feature to be weighed for a set of positive samples, it must cluster these samples, i.e., the entropy of the distribution should be small. This approach gives better performance than MARS. MacArthur et al. [37] used decision tree(DT) and viewed P and N from RF as a two class problem. Here DT is trained by P and N feedback images and later on in testing stage unlabelled database images are labelled by DT. Further research moved to exploring the discriminant subspace and using kernels to obtain the hyperplane that separates P and N clusters.

Probabilistic Approaches Bayesian estimation methods have been used in the probabilistic approaches to relevance feedback. Cox et al.[13], used Bayesian learning to incorporate user feedbacks to update the probability distribution of all the images in the database. Here the feedback examples are considered as the sequence of independent queries and they try to minimize the retrieval error by Bayesian rules. Bayesian framework to estimate the posterior probability p(R/Y), i.e., a database image Y is relevant to the query given the prior p(R) i.e., number of relevant images, and the conditional probability p(Y/R) is obtained by approximating parametric model such as Gaussian kernel using the feature vectors of R that are identified as relevant. Computing p(N/Y) similarly, the system can judge whether a database image Y is relevant to the query using the classifier that computes

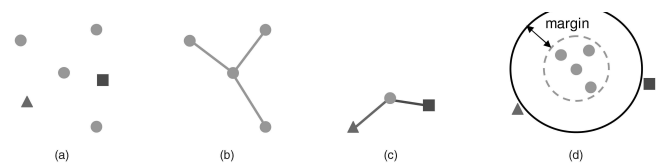
$$J = \frac{p(R/Y)}{p(N/Y)}$$

In [32] the author proposed a new image retrieval method termed Adaptive clustering which involves a two step process: classification and cluster merging. The classification is performed by Bayesian classifier and cluster-merging is done if the two cluster mean vectors follow the statistical Hotelling T2 statistics with significance level. The top  $k$  images for display are based on the aggregate distance of database image to each of the clusters.

**7.1.2 Discriminant Models:** In almost all of the above methods, P and N are assumed to come from uni-modal distribution and treated RF as two class problem. This formulation is not true in RF, as all P are alike is true, while each of the N is said to be negative in its own way[29]. Discriminant model identifies a feature space that is the combination of informative (compactness) and discriminant learning. Compactness refers to minimizing within class variance while discrimination refers to maximizing the separation between classes. LDA compacts both P and N while BDA compacts only positive examples keeping away negative points. Another kernel method was suggested, kernel Biased Discriminant Analysis (kBDA). Kernel BDA does the same, but in the higher dimensional feature space associated to the kernel rather than in the original feature space. In [22] 1-class SVMs train the hypersphere by P, which is an estimate of the support of distribution of positive examples.

**7.1.3 Generative Models:** The features of relevant images form an islands or distinct clusters in the low-level feature space. Each such local grouping of P are enclosed in hypersphere called cover[16] or hyperellipsoid called Connected Component[1]. This distribution of features can be well modeled by Gaussian Mixture Model. This has given improved learning over MARS and MR approach.

**7.1.4 Manifold learning:** The more recent approaches use Distance Metric Learning [38]. Many manifold learning algorithms were developed in recent years and this area of Manifold learning is based on the concept of Riemannian normal coordinates. Here the task is to map a set of high dimensional data points, with a manifold like structure to a low-dimensional space. This mapping perverts the information of direction and distance from a specific point on a manifold to any other nearby points. X. He et al,[25] proposed a novel semisupervised method of dimensionality reduction called Maximum Margin Projection(MMP). MMP is designed in preserving the local manifold structure, and aims to obtain large margin subspace[25], by using a Mahalanobis metric to learn in such a way to keep each of the target neighbours(P) closer than other inputs with different class labels (imposters or N) by a large margin as shown in figure 4.1. Unsupervised distance metric aims to construct a low-dimensional manifold, where geometric relationships between most of the observed data are largely preserved. This group of methods can be divided into nonlinear and linear methods. The main disadvantage of PCA and LDA is that both of them are based on Euclidean structure.



**Figure 5:** (a) Feature vectors with P (dots) and N imposters (square and triangle). (b) The within-class graph connecting P nearby points. (c) The between-class graph connecting nearby points with different labels. (d) After MMP, the margin between different classes is maximized.[25]

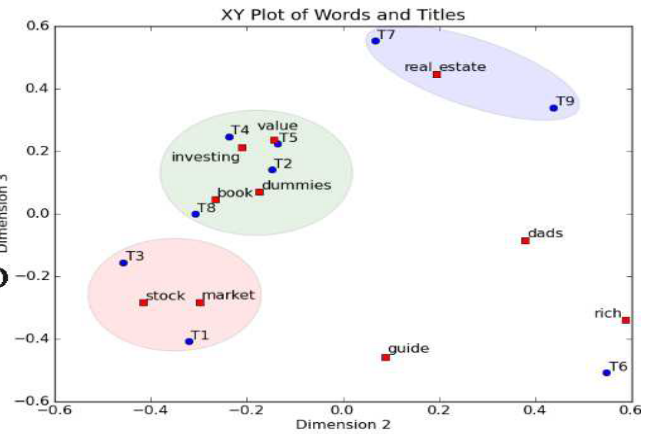
## 7.2. Long Term Learning

Long Term Learning is basically memorizing the RF information taken from earlier users interaction during each query session, and this RF log includes hidden semantic experiences that helps improved retrieval in subsequent query sessions. The semantic information or the co-occurrence association information of images in the RF log helps in pecculating information of RF to unlabelled images by means of clustering. In the review by Jing Li et al[34], the author categorizes the LTL methods based on core techniques used. This paper classifies earlier work on LTL into four categories.

**7.2.1 Latent Semantic Analysis:** LSA is directly borrowed from text retrieval. LSA, pLSA and Latent Dirichlet Allocation (LDA) are popular techniques in this direction. The fundamental difficulty arises when we compare low level features( words) to find semantically relevant images (documents) because what we really want to compare is meaning or concepts between words. LSA attempts to solve this problem by mapping both words and documents into a concept space and doing the comparison in this space. The term-document matrix is equivalent to RF log information of images-latent concept matrix in the intra query learning. Figure 6 illustrates the example term-document matrix, mapped to concept space of reduced dimension[4]. The pros of LSA are mainly the concept space has fewer dimensions compared to the original matrix that helps to contain most of the information and least noise. This allows to obtain concept space mapping by algebraic approach LSA assumes that the concept space to be of Gaussian distribution and Frobenius norm may not fit all problems. In [27] after the concept space is obtained using RF log, the similar images to the query is obtained by the most informative terms that have high gain using decision trees. The Performance of LSA at varying levels of RF log(documents) was compared with MARS and PFRL. It has been seen that LSA performed better after one-fourth of the size of the database is present as documents in the RF log. X.He et al. [26] combines the low level features with semantics derived by LSA.

	T1	T2	T3	T4	T5	T6	T7	T8	T9
BOOK			1	1					
dada						1			1
DUMIES		1							
ESTATE							1		1
GUIDE	1					1			
INVESTIN G	1	1	1	1		1	1	1	1
MARKET	1		1					1	
REAL						2			1
sTOCK	1		1					1	1
Value			1	1					

SVD



Images and documents are mapped to Concept space

Figure 6: Mapping of term-document matrix to Concept Space[4]

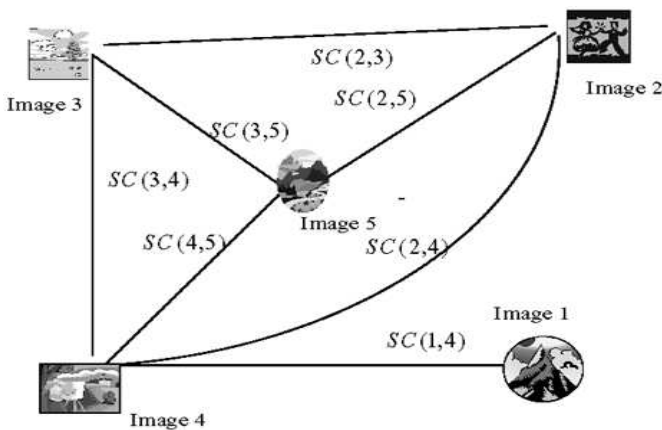


Figure 7: Semantic Network build over the images

**7.2.2 Statistical Correlation :** These category of methods relies more on co-occurrence information of images in RF. This information is exploited by using Association rule mining or collaborative filtering[31] or building a semantic link on the database of images[35, 24]. Figure 7 shows a semantic network with co-occurrence information over the link of two images in database. Here the author exploits the network in driving the links to images that are not seen by the user. P.Y.Yin et al. has used modified apriori algorithm(soft apriori) in deriving correlation rules among the RF images. S. Hoi et al [28] extracts correlation between two images in a db using RF log. Here the author overcomes the imbalance training of SVM by increasing the training images with the help of RF log using both P and N images of current iteration as seeds. Here the author uses score based SVM to score the unlabelled images in the database.

**7.2.3 Semantic Clustering:** This includes the work related to Semi-Supervised and unsupervised clustering of database. In [22], retrieval is based on exploiting both intra and interquery learning. The P from RF for each query is used to train one-class SVMs, and this results in High-level-concept clusters or concept database which forms RF log. For the retrieval set constituting the intra query learning, some k-Fuzzy memberships are estimated for k-nearest neighbor clusters. The reason for finding k memberships for a query being that each image constitutes many(k) concepts. The

membership is the relative distance between the query and the centers of all hyperspheres into which query falls. Most of the works in this category are from B. Bhanu et al. In [10], the concept space was visualised by Semi-Supervised Fuzzy C-means clustering (SSFCM). Here two matrices corresponding to P and N are used ,to keep track of feedback information of an image as positive or negative to a cluster or a concept. This forms the log, and the optimization function of SSFCM uses this log for clustering. The main idea here is in assigning labels to unseen images using RF history of groups. If the negative image in the current relevance feedback session is assigned as positive image in RF history, then all the other positive images in that group in association with this image are considered negative, similarly the positively associated images with the current positive images are extracted by the RF history. The performance plots show that 60% retrieval results is obtained with RF history equal to number of database images. In [19] the author does clustering on block generated features of database images and later, a membership is associated based on the position of a feature in the cluster.

**7.2.4 Feature Vector Based:** The fourth includes the learning of feature relevancy. In [21] the pooled variance over each iteration form the feature weights and hence these weights of the query corresponding to a session are learned. These weights form the model parameters corresponding to a query and are used across the sessions for retrieval. Here the author claims that the use of PCA to reduce the dimensionality, gives much advantage of having reduced features at the cost of minimal PR performance. B. Bhanu et al [8] proposed a novel technique that integrates STL and LTL experiences. Here every image in the database consists of two parts: low-level-features and virtual features. The virtual feature of an image records all the relevant index concepts the image has and the significance related to this image. The index of the concept is nothing but the system counter of a query and the significance is relevance level to the concept. At the end of each session, transform operations are defined to summarize the session’s relevance of an image to a concept. Here LTL is based on the fact that recent high relevant index concepts get more weight than the

old concepts of less relevance learned. The author claims maximal length of the VF for a database image can be 200, which is a reasonable estimate for the maximal number of concepts contained in an image. Gosselin and Cord[23] have proposed a kernel based similarity learning strategy.

## 8. COMPARISON AND EXPERIMENTATION PARAMETERS OF LEARNING TECHNIQUES.

The performance of the CBIR system using the given set of features and the database will greatly depend on the learning technique used. An optimal LTL model should be more of exploring in initial stages and later on after many in-session learning experiences gained, the exploitation of this log need to be more. The balance between exploration Vs exploitation is important at different stages of the system. The parameters that decide the PR performance include database size, the queries used, the learning technique, the

number of retrieved images termed *Scope*, number of classes and their complexity (amount of overlapping in feature space). Four of the recent works, each from STL and LTL are taken for comparison in table 3. PR refers to approximate precision at the second iteration, where in first retrieval is considered to be zeroth iteration. Most of the time the results are shown with *automatic RF* by the machine using ground truth, unless otherwise mentioned rarely manual RF is used. In addition to PR vs iteration, PR vs Recall and PR Vs scope, the performance for varying amount of noise that indicates the users mistakes in RF due to mislabelling and the amount of RF log in LTL are also considered for performance elevation of the system. The freely downloadable image databases are wang corel[3], corel photo gallery[2], caltech photos[5], MIRFLICKR database[1] with features, or images can also be downloaded from photo sharing sites using keyword search and the Application Program Interface(API).

Table 3. Comparison of Various STL and LTL methods

S. No	Method	Image DB. size	Scope	No. of classes	Precision (%)	Remarks
1.	Long Term Semi Supervised EM[7]	1,200	20	12	98	LTL with no mislabeling or noise in RF and RF log of 400 sessions, and PR over 1200 queries
2.	Image Relevance Association Rule Mining[67]	3,400	10	22	80	3,400 10 22 80 LTL (soft QVM with manual RF and association rules of 864,830 generated by 50 queries and average PR for 44 queries.
3.	Learning semantic space by SVD[26]	10,000	100	79	70	LTL, SVD used with 300 query experiences
4.	LTL SVM based on correlation of RF log[28]	50,000	20	50	56	100 RF log sessions
5.	Probabilistic approach[39]	10,000(www + corel + calPhotos)	20	10	62	STL manual RF with 6 users
6.	On-line learning [21]	1000	100	100	65	STL and using variance pooling approach
7.	Adaptive clustering[32]	30,000	100	30	28	LTL and using variance pooling approach

## 9. SCOPE FOR THE FUTUREWORK

The performance of a CBIR system using only visual features is limited. Hence learning technique that uses annotations and then using the ontology aspect in retrieving will result in better performance. More recent work is in the direction of distance metric learning by semi-supervised means, that use manifold approach like LLE or MMP or by the use of kernel methods.

## 10. CONCLUSION

This paper touches firstly on the methods to reduce semantic gap, and the difficulties faced by RF is discussed. More recent work in this area to reduce the gap is to use text labels in the framework of ontologies using Wordnet. Secondly

most of the existing learning methods in the area of Long term learning and Short term learning are first categorized based on the core technique used and then intuition in each of the methods is discussed. The performance of LTL depends on the RF log and the extent to which the knowledge is exploited from the log to label or score the unlabeled images.

## REFERENCES

- [1] <http://press.liacs.nl/mirflickr/>.
- [2] <http://sites.google.com/site/dctresearch/home/contentbase-d-image-retrieval>.
- [3] <http://wang.ist.psu.edu/jwang/test1.tar>.

- [4] [http://www.puffinwarellc.com/index.php/news\\_and\\_articles/articles/33-latent-semantic-analysis/tutorial.html](http://www.puffinwarellc.com/index.php/news_and_articles/articles/33-latent-semantic-analysis/tutorial.html).
- [5] <http://www.vision.caltech.edu/imagetdatasets/caltech101/>.
- [6] Gaurav Aggarwal, T. V. Ashwin, and Sugata Ghosal. An image retrieval system with automatic query modification *IEEE Transactions on Multimedia*, 4(2):201–214, 2002.
- [7] 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.
- [8] Peng-Yeng Yin, B. Bhanu, Kuang-Cheng Chang, and Anlei Dong. Long-term cross-session relevance feedback using virtual features. *IEEE Transactions on Knowledge and Data Engineering*, 20(3):352–368, march 2008.
- [9] S. M. Renuka Devi and Chakravarthy Bhagvati. Connected component in feature space to capture high level semantics in CBIR. In *Proceedings of the Fourth Annual ACM Bangalore Conference 2011*, New York, NY, USA.
- [10] 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.
- [11] C.M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.
- [12] C. Carson, M. Thomas, S. Belongie, J. M. Hellerstein, and J. Malik. Blobworld: a system for region-based image indexing and retrieval (long version). Technical report, EECS Department, University of California Berkeley, 1999.
- [13] I. J. Cox, M. L. Miller, T.P. Minka, Papathomas T.V., and Yianilos P.N. The bayesian image retrieval system, pichunter: theory, implementation, and psychophysical experiments. *IEEE Transactions on Image Processing*, 9(1):20–37, January 2000.
- [14] R. Datta, D. Joshi, J. Li, and J. Z. Wang. Image retrieval: Ideas, influences, and trends of the new age. *ACM Computing Surveys.*, 40(2):1–60, 2008.
- [15] Thomas Deselaers, Daniel Keysers, and Hermann Ney. Features for image retrieval: an experimental comparison. *Inf. Retr.*, 11, April 2008.
- [16] 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. International Conference on Multimedia and Expo, November 2002.
- [17] A. Pentland, R. W. Picard, and S. Sclaroff. Photobook: Content-based manipulation of image databases.
- [18] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker. Query by image and video content: the qbic system. *Computer*, 28(9):23–32, September 1995.
- [19] J. Fournier and M. Cord. Long-term similarity learning in content-based image retrieval. In *Proc. International Conference on Image Processing*, volume 1, pages I–441–I–444, 2002.
- [20] Keinosuke Fukunaga. *Introduction to Pattern Recognition: A MATLAB Approach*. Academic Press, Second edition, 1990.
- [21] L. Goldmann, L. Thiele, and T. Sikora. Online image retrieval system using long term relevance feedback. In *Image and Video Retrieval*, LNCS, pages 422–431. 2006.
- [22] I. Gondra, D. R. Heisterkamp, and J. Peng. Improving image retrieval performance by inter-query learning with one-class support vector machines. *Neural Computing Applications*, 13:130–139, 2004.
- [23] P. H. Gosselin, M. Cord, and S. P. Foliguet. Combining visual dictionary, kernel-based similarity and learning strategy for image category retrieval. *Comput. Vis. Imag Underst.*, 110:403–417, June 2008.
- [24] 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.
- [25] X. He, Deng Cai, and Jiawei Han. Learning a maximum margin subspace for image retrieval. *IEEE Trans. on Knowl. and Data Eng.*, 20:189–201, February 2008.
- [26] 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.
- [27] D. R. Heisterkamp. Building a latent semantic index of an image database from patterns of relevance feedback. In *ICPR (4)*, pages 134–137, 2002.
- [28] 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.
- [29] T. S. Huang and X. S. Zhou. Image retrieval with relevance feedback: From heuristic weight adjustment to optimal learning methods. In *Proceeding of International Conference on Image Processing*, 2001.
- [30] Y. Ishikawa, R. Subramanyay, and C. Faloutsos. Mindreader: Querying databases through multiple examples. *Proc. 24rd International Conference on Very Large Data Bases*, 4071:422–431, 1998.
- [31] X. Zhou, Q. Zhang, L. Zhang, Li Liu, and B. Shi. An image retrieval method based on collaborative filtering. In *Intelligent Data Engineering and Automated Learning*, volume 2690 of *Lecture Notes in Computer Science*, pages 1024–1031. 2003.
- [32] D. Kim and C. Chung. Qcluster:relevance feedback using adaptive clustering for content-based image retrieval. In *Proc. 2003 ACM SIGMOD*, SIGMOD '03, pages 599–610.
- [33] Michael S. Lew, Nicu Sebe, Chabane Djeraba, and Ramesh Jain. Content-based multimedia information retrieval: State of the art and challenges. *ACM Trans. Multimedia Comput. Commun. Appl.*, 2:1–19, February 2006.
- [34] J. Li and N. M. Allinson. Long-term learning in content-based image retrieval. *International Journal of Imaging Systems and Technology*, 18:, 2008.
- [35] M. Li, Z. Chen, L. Wenying, and H. Zhang. A statistical correlation model for image retrieval. In *Proc. of the ACM workshops on Multimedia: multimedia information retrieval*, MULTIMEDIA '01, pages 42–45, 2001.
- [36] Y. Liu, D. Zhang, G. Lu, and W. Y. Ma. A survey of content-based image retrieval with high-level semantics. *Pattern Recogn.*, 40, January 2007.

- [37] S. D. MacArthur, C. E. Brodley, and Chi Ren Shyu. Relevance feedback decision trees in content-based image retrieval. In *Proc. IEEE Workshop on Content-based Access of Image and Video Libraries*, pages 68–72, 2000.
- [38] Liu Yang and Advisor Rong Jin. Contents distance metric learning: A comprehensive survey.
- [39] David McG. Squire, Wolfgang Müller, Henning Müller, and Thierry Pun. Content-based query of image databases: inspirations from text retrieval. *Pattern Recogn. Lett.*, 21:1193–1198, December 2000
- [40] J. R. Smith and S. Chang. Visualeek: a fully automated content-based image query system. In *Proceedings of the fourth ACM international conference on Multimedia, MULTIMEDIA '96*, pages 87–98. ACM.
- [41] R. Rahmani, S. A. Goldman, H. Zhang, J. Krettek, and J.E. Fritts. Localized content based image retrieval. In *Proceedings of the 7th ACM SIGMM international workshop on Multimedia Information retrieval, MIR '05*, pages 227–236, 2005.
- [42] James Z. Wang, Jia Li, and Gio Wiederhold. Simplicity: Semantics-sensitive integrated matching for picture libraries. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23(9): 947–963, September 2001.
- [43] P. Tandon, P. Nigam, V. Pudi, and C. V. Jawahar. Fish: A practical system for fast interactive image search in huge databases. *Proceedings of the 7th ACM CIVR '08*, July 2008.
- [44] S. N. Prasad, V. Shiv, N. Prasad, A. G. Faheema, and S. Rakshit. Feature selection in example-based image retrieval systems. In *Proceeding of Indian Conference on Vision Graphics and Image Processing*, 2002.
- [45] Y. Rui, T. S. Huang, and S. Mehrotra. Content-based image retrieval with relevance feedback in mars. In *Proc. IEEE ICIP*, 2:815–819, 1997.