

# IMAGE RETRIEVAL AND RE-RANKING TECHNIQUES - A SURVEY

Mayuri D. Joshi, Revati M. Deshmukh, Kalashree N.Hemke, Ashwini Bhake  
and Rakhi Wajgi

Computer Technology Department,  
Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India.

## **ABSTRACT**

*There is a huge amount of research work focusing on the searching, retrieval and re-ranking of images in the image database. The diverse and scattered work in this domain needs to be collected and organized for easy and quick reference.*

*Relating to the above context, this paper gives a brief overview of various image retrieval and re-ranking techniques. Starting with the introduction to existing system the paper proceeds through the core architecture of image harvesting and retrieval system to the different Re-ranking techniques. These techniques are discussed in terms of approaches, methodologies and findings and are listed in tabular form for quick review.*

## **KEYWORDS**

*Image Retrieval, Re-ranking, MI learning, Ontology, Multi-latent vector.*

## **1. INTRODUCTION**

Image retrieval is a key issue of user concern. Normal way of image retrieval is the text based image retrieval technique (TBIR)[12]. TBIR-needs rich semantic textual description of web images .This technique is popular but needs very specific description of the query which is tedious and not always possible.

Therefore generally the process of image search includes searching of image based on keyword typed. The process that occurs in the background is not so simple though.

When query is entered in the search box for searching the image, it is forwarded to the server that is connected to the internet. The server gets the URL's of the images based on the tagging of the textual word from the internet and sends them back to the client.

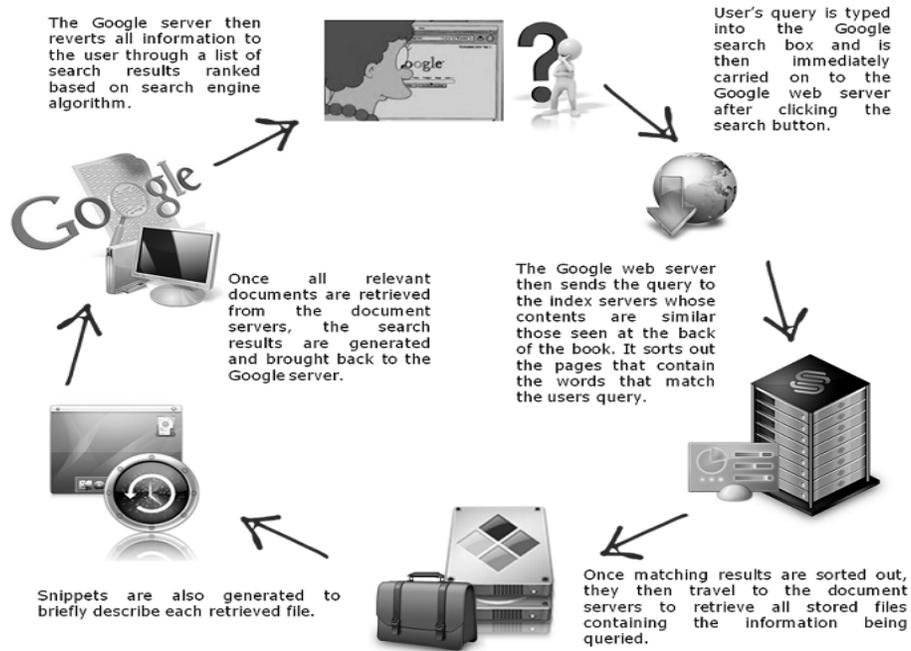


Figure 1. Working of Google search engine. [17]

The search engine thus navigates through the pages and collects the images. It gives the client the top ranked image which is the one with maximum number of hits from the user and a set of images. This is the technique of text based image retrieval system.

But it has certain drawbacks like images obtained are many a time duplicated, of low precision, and irrelevant. This scenario may occur due to sparse and noisy textual query. Due to this aspect user cannot be always sure of perfect images being obtained in available time. Many a times user has to surf many pages of images available to land at the perfect one. This possesses a great threat to the fast technology. Such problems surface when user needs large database of images. So due to these factors of complexity, "image harvesting and retrieval" is a topic which is gaining popularity in research sector.

What can be done in this respect is as follows-

1. Rerank the images obtained on client side and provide with top rank image.
2. Use highly efficient clustering algorithm to facilitate grouping of similar images and select perfect among them.
3. Use contents of image rather than url tagging to retrieve images from internet database
4. Use various concepts in combination to get an excellent image retrieval system.

The above mentioned factors are reviewed throughout this paper and different details and aspects are put forward for comparison. Each method has certain limitations but trade off between them surely evolves the best out of the available study.

## 2. LITERATURE SURVEY

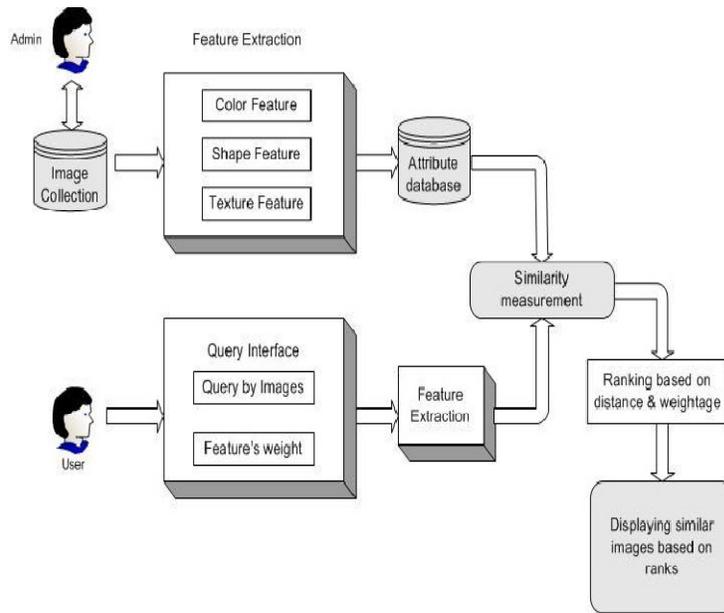


Figure 2. Architecture of image harvesting and re-ranking system [10]

From the architecture diagram (Fig. 2) [10] an overview is obtained. Each module observed in the figure is a complex module having own ways of implementation and understanding. Exclusive factors of Digital image are used.

The large image collection is subjected to feature extraction process where the attributes of the image both visual such as color, texture and shape and semantic such as intentional, clicks, labels etc. are extracted from the feature database using appropriate methods. The query image can be any of the popular formats. The query image is subjected to feature extraction process and query features are obtained. In similarity measurement process, the query's feature is compared with the features stored in feature database. The distance between the two features is calculated and weights are determined. The output images are then sorted and ranked, so that most similar images can be displayed to the user. This system is based on the following functionalities and features:

a) Extraction

(i) Visual features

If the entered query is "sunset", color should be the considered feature as color is the primary identifier. For "building" shape as a feature rather than color is appropriate. Whereas, for "snow" if color and shape is considered then differentiation between "snow" and "cotton" would become difficult for the system. Thus, texture will become the primary identifier for "snow" and not colour or shape.

(ii) Semantic features

Semantics is the actual intention of the user behind the query. This intention cannot be interpreted by the machine, resulting in the semantic gap. For instance, if the entered query is "ford", user may intend for a car or a person named "Ford". But system cannot interpret the intended semantic. Thus, to reduce the semantic gap, semantic feature need to be considered.

b) Distance calculation and similarity measurement:

This step calculates the difference between the images in terms of corresponding feature. Lesser the distance, more similar the images are. For example, if the entered query is "lake" and the selected feature is color. The images are plotted in feature space and distance between them is calculated. The images that lie closer in this space are considered to be more similar.

Given two feature vectors A and B such that

$$A = \begin{pmatrix} a_1 \\ a_2 \dots \\ a_n \end{pmatrix} \quad B = \begin{pmatrix} b_1 \\ b_2 \\ b_n \end{pmatrix}$$

Euclidean distance is given by:

$$\sqrt{\sum_{i=1}^n (a_i - b_i)^2}$$

City block is another approach for distance measurement. [5]

$$\sum_{i=1}^n |a_i - b_i|$$

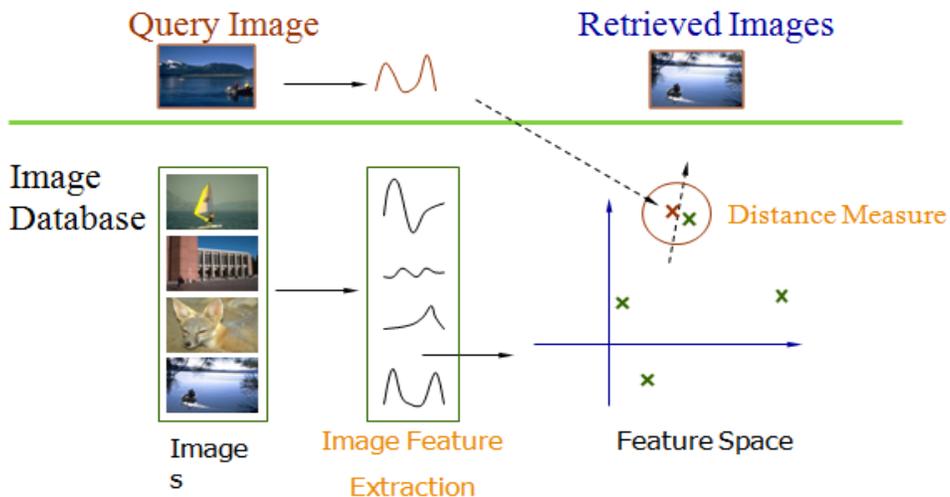


Figure 3. Distance calculation and measurement [18]

Feature extraction will be compulsorily followed by distance calculation and similarity measurement. As mentioned in [5], for CBIR implementation, image classification should be fast and efficient.

In this context if visual features are considered as features to be extracted then low level histogram representation is most efficient as histogram is a model of probability distribution of intensity levels of visual features. Also its generation is quick as well as easy for comparison.

If semantic features are considered satellite image retrieval system (SIRS) [8] is a good approach. Understanding of semantic features and their extraction require data and knowledge exchange. [8] proposes use of xml for data exchange and use of web ontology language for knowledge exchange. Semantic knowledge is described using rule based expert system, neural network, decision trees etc. In relation to this concept, ontology refers to expressing elements of domain as well as intended meaning of element. Query "ford" mentioned above is an example needing implementation of ontology.

c) The core architecture can be extended to Re-rank the images based on various parameters. The techniques for image retrieval and re-ranking may differ in feature extraction algorithms, score calculation methods, and score matching algorithms and re-ranking algorithms individually or in combination. This paper is a review work considering the above parameters through a detailed study of related domain specific features.

A simple and thinking friendly way to start with is Content based image retrieval (CBIR) technique [1].

## **2.1. Overview of CBIR**

This concept emphasises on use of visual content of image like colour, texture, shape etc. for image comparison and retrieval rather than textual query. In common words, visual feature of any image is anything that is seen or felt about that image. It includes any visual variation in the look of that image.

These contents are then extracted from images in the database and are described by multi-dimensional vectors. The feature vectors of the images in database form the feature database. To retrieve images, users provide the retrieval system with example images or sketched figures. The system then converts them into internal representation of feature vectors. The similarities /distances between the feature vectors of the query example or sketch provided and those of the images in the database are calculated and then retrieval is performed. Under this work various factors defining the concerned visual contents are described in details.

Retrieved images will need comparison based on various features. Comparison based on their appearance is one approach named as "appearance based image matching" [12]. It works using the basis of parts and shapes of image. But this concept is not widely in application because its time complexity is very high as each image retrieved from the database is matched with the desired image. So finally, clustering is found to be the solution for this problem.

Table 1. Visual Attributes of Image

Visual attribute	Factors under consideration
1.Colour	1.Colour space 2.Color Correlogram 3.Coherence vector 4.histogram[5] 5.colour moment
2.Texture	1.Tamura features 2.Wold Feature 3.Gabor filter feature
3.Shape	1.Moment Invariant 2.Turning Angles 3.Polynomial approximation 4.Fourier Descriptors

### 2.2. Bag based Image Re-ranking

Clustering means grouping similar images together and comparing or matching among clusters instead of individual images. This will reduce the concerned time complexity to a great extent. So cluster of similar images containing most of the relevant images is called positive bag and the bag containing least relevant images related to query is labelled as negative bag. This way of clustering is derived from the theory of Generalized Multi-instance learning (GMI) [12] and called as bag based image re-ranking. Diverse clustering algorithms are available with varying degree of success based on domain requirement. The task following bags formation is removal of irrelevant images and re-ranking the remainder. Iterative application of bag formation algorithm using weak bag annotation technique [12], yields bag more precise to the entered query. This is viewed through the following diagram.

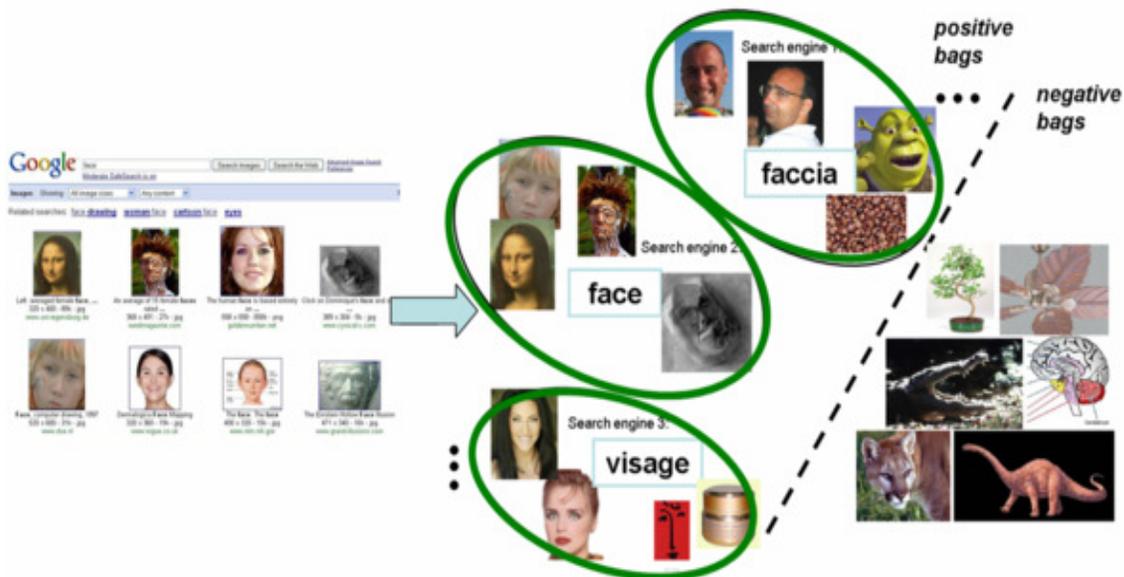


Figure 4. Labeling positive and negative bags for the query FACE .[17]

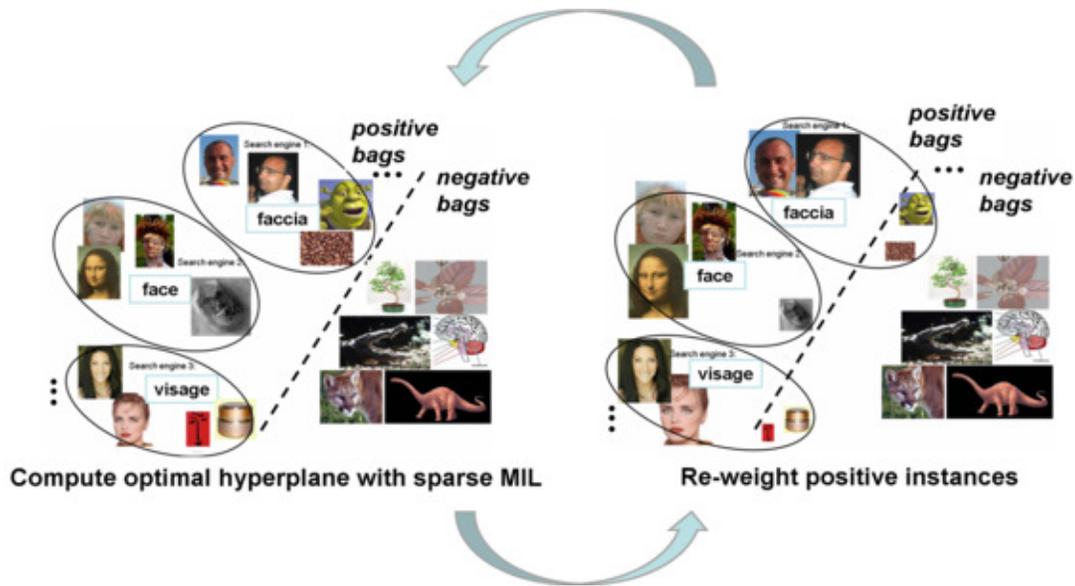


Figure 5. Iterative application of bag based algorithm for bag optimization.[17]

### 2.3. Assumption for Clustering and Re-ranking of Images

Some assumptions for clustering and re-ranking of images are mentioned. [13]

1. Pseudo-Relevance Feedback (PRF) assumption- The top-N images of initial result are regarded as pseudo-relevant.
2. clustering assumption - Visually similar images should be ranked nearby.

But these assumptions have following deficiencies-

1. They make visual similarity equal to similarity of relevance to query. This means similar looking images will not always be of same category.
2. They omit the fact that if two images are not similar, even then they can be equally relevant.

To cope up with these deficiencies, trend moves towards supervised re-ranking also called as active re- ranking [9].

### 2.4. Active Re-ranking

Active re-ranking is the re-ranking with user interactions. Figure [9] depicts the flow of active re-ranking technique for the query "panda". It involves active sample selection in which user labels the images as relevant or irrelevant. The images seen in the third module bearing tick-marks are the user labelled relevant images. This step is followed by dimension reduction [9] which localizes visual features. Iterative applications of above steps leads to proper result.

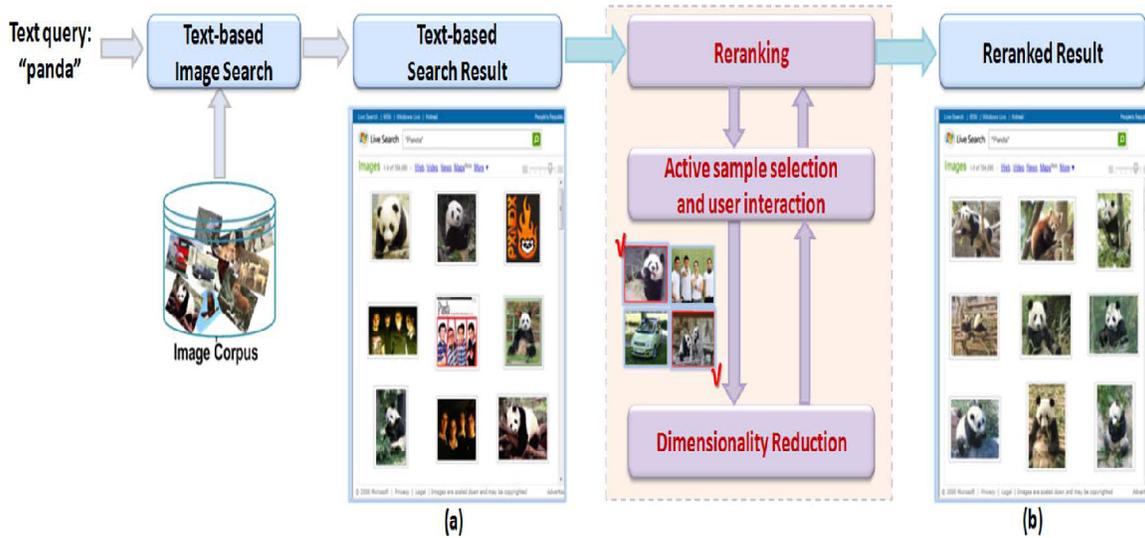


Figure 6. Framework for active re-ranking illustrated with the query “panda”. When the query is submitted, the text-based image search engine returns a coarse result (a). Then the active re-ranking process is adopted to obtain a more satisfactory result (b), by learning the user’s intention. [9]

The above explained techniques use single feature for re-ranking, but the type of most effective features vary across queries, as elaborated above under the topic extraction of visual features. Thus, employing multimodal features (color, texture, edge)[14] is a solution.

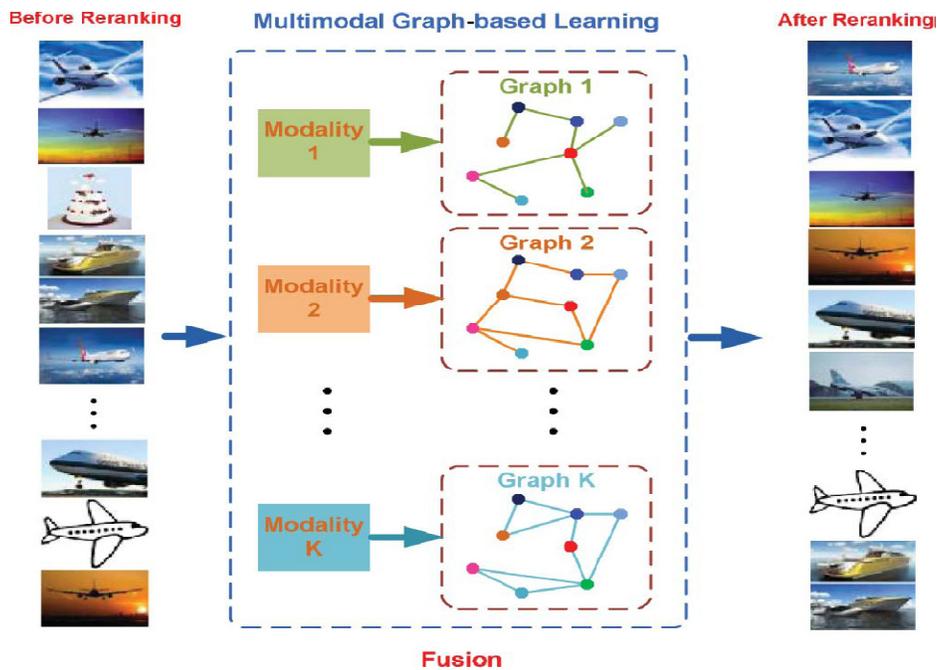


Figure7. Illustrates multimodal graph-based learning.[14]

In this approach, graphs are constructed each for one modality. Later the result of each modality is fused based on the relevance scores and based on it, the images are re-ranked.

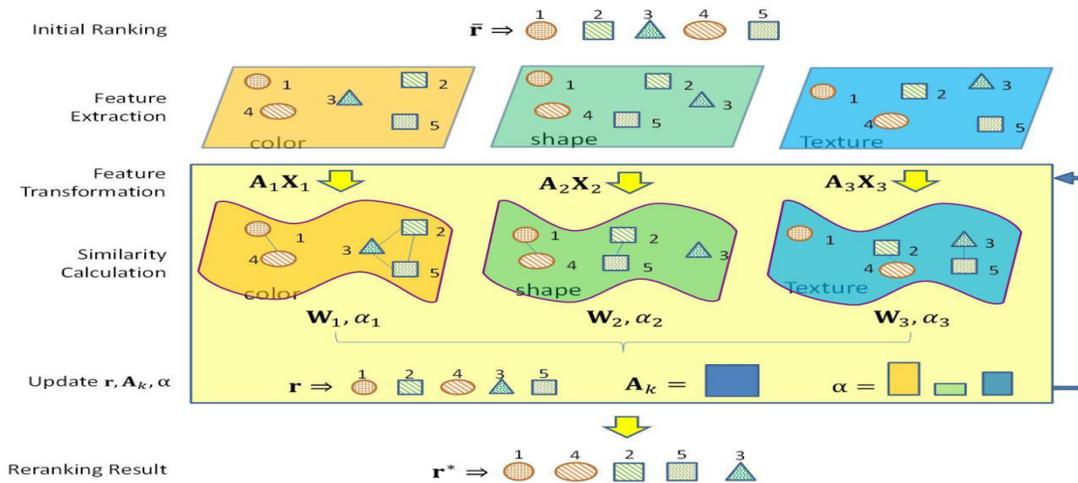


Figure 8. Fusion based on relevance scores, weight of modalities, distance metric.

$W_k$ -The similarity matrix of images for the  $k$ th modality.

$\alpha$  -The weight vector.

$A_k$ -The transformation matrix for the  $k$ th modality.

Multimodal fusion in combination with pattern mining forms a new re-ranking technique called circular re-ranking [15]. Circular re-ranking uses the mutual exchanges of information across multiple modalities for improving the search performance .

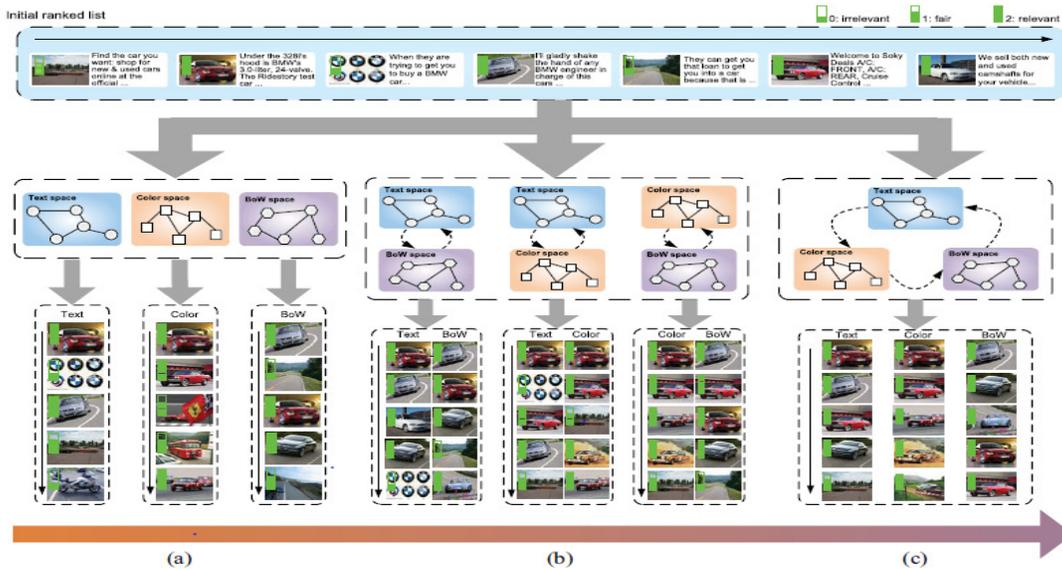


Figure 9. Circular re-ranking. [15]

Table 2. Summary Table

Year	Ref. No	Dataset Used	Methodology	Findings	Approach Used
1996	[2]	-	<p>1. Block oriented CBIR- Wavelet transforms used to extract image features.</p> <p>2. Feature vectors of images are constructed using two wavelet transforms. Content based image retrieval performed by comparing the feature vectors of the query image and the segments in database Images.</p>	<p>1. Introduced image segmentation concept and block-oriented image decomposition structure which is later used to support CBIR</p>	<p>model block oriented image decomposition structure i.e.</p> <ol style="list-style-type: none"> <li>1. Nona-tree decomposition</li> <li>2. quad-tree decomposition.</li> </ol>
April 1999	[3]	-	<p>1.Extraction of color feature</p> <p>2.order the obtained features</p> <p>3. calculate feature vector</p> <p>DSQ algorithm is used to achieve the above goal</p> <p>Application of DSQ is followed by dynamic matching for image Re-ranking.</p>	<p>Color based indexing is used as</p> <ol style="list-style-type: none"> <li>1. Color feature of image is less sensitive to noise and background complications.</li> <li>2.Colour compute image statistics independent of geometric variations</li> </ol>	<ol style="list-style-type: none"> <li>1. Dependent Scalar Quantization(DSQ)</li> <li>2. Dynamic matching</li> <li>3. Histogram intersection method.</li> <li>4. Distance method.</li> </ol>
2007	[4]	Animal images from Flickr.	<p>Partial grouping using BBC-</p> <ol style="list-style-type: none"> <li>1. Consider partial clusters using BBC based on result of text based search.</li> <li>2. Obtain cluster of relevant images based on relevance feedback.</li> <li>3. Images are re-ranked as per visual similarities.</li> </ol>	<ol style="list-style-type: none"> <li>1. Relevant and irrelevant images are less mixed in clusters formed by BBC</li> <li>2. BBC makes easier for user to label clusters.</li> <li>3. Select accurate cluster representatives without additional human labour.</li> </ol>	<ol style="list-style-type: none"> <li>1. Bregman Bubble Clustering(BBC) algorithm</li> </ol>

2009	[4]	Images from Google and Yahoo	One class classification- 1. Crawl images from Google and Yahoo. 2. Form a single class called target class containing irrelevant images. 3. Use kernel whitening and SVDD for detecting the relevant (outlier to target class) images. 4. Relevance feedback is used to improve performance.	1. Useful even with lack of clean data and contaminated unrelated data. 2. Useful in non-popular or non-typical category classification.	1. Kernel whitening 2. support vector data description 3. Relevance feedback
Mar 2010	[9]	Synthetic database	Active re-ranking- 1. Collect labelling information from user to obtain specified semantic space. 2. Localize the visual characteristics of the user intentions in space.	1. Use both ambiguity and representativeness . 2. Reduce user labelling efforts	1. Structural information based sample selection 2. Local global discriminative dimension reduction algorithm (LGD).
April 2011	[11]	Bing query logs	Click data based re-ranking- 1. Identify previously clicked images for the same query. 2. GPR is then trained to predict normalized clicks on each image. 3. Combining original and predicted click count re-rank images.	1. No need of user intervention. 2. Query independent method 3. Reduce label noise problem 4. Promote likely to be clicked images along with previously clicked images.	1. Gaussian Process Regression(GP R) 2. Click boosting as tie breaker.
Nov 2011	[12]	Flickr images with tags ,NUS-WIDE	Bag Based re-ranking- 1. Partition images into clusters using textual and visual features 2. Uses multi instance(GMI) framework 3. Treats each cluster as Bag and images as instances	1. MI learning problem 2. Weak bag annotation 3. Average precision for images	1. MI-SVM, GMI-SVM 2. K-means algorithm for clustering

Jun 2012	[13]	Web query (353 queries)	<p>Prototype Based re-ranking-</p> <ol style="list-style-type: none"> <li>1. Find relevance probability from rank position in initial search result.</li> <li>2. Generate visual prototypes.</li> <li>3. Meta ranker constructed using prototypes to calculate image score.</li> <li>4. Uses linear ranking.</li> </ol>	<p>Re-ranking model is query independent as learned model weights are related to initial text based rank position of any image and not image itself.</p>	<ol style="list-style-type: none"> <li>1. Visual modality</li> <li>2. Supervised and unsupervised image searching</li> <li>3. Clustering assumption used</li> </ol>
Nov 2012	[14]	MSRA-MM (version 1.0,2.0)	<p>Multi-modal re-ranking-</p> <ol style="list-style-type: none"> <li>1. Integrate learning of relevance score, weights of modality, distance matrix and its scaling into unified scheme.</li> </ol>	<ol style="list-style-type: none"> <li>1. More robust than using each individual modality.</li> <li>2. Better performance than existing approaches.</li> </ol>	<ol style="list-style-type: none"> <li>1. Multi modal graph based Re-ranking</li> <li>2. Use late fusion.</li> </ol>
April 2013	[15]	MSRA-MM	<p>Circular re-ranking method-</p> <p>Retrieved images are modelled as graphs in different feature spaces followed by-</p> <ol style="list-style-type: none"> <li>1. Random walks: Re-ranking the images by treating each feature space independently.</li> <li>2. Mutual reinforcement: Pair wise exchanging modality spaces.</li> <li>3. Circular Re-ranking: iteratively updating the image ranks by circular mutual reinforcement.</li> </ol>	<ol style="list-style-type: none"> <li>1. Addresses the issue of multi-modality Interaction in visual search by mutual reinforcement.</li> <li>2. In this way, the performance of the weak Modality is also benefited by learning from strong modalities.</li> </ol>	<ol style="list-style-type: none"> <li>1. Recurrent pattern mining- <ol style="list-style-type: none"> <li>a. Self</li> <li>b. Crowd</li> <li>c. Example based</li> </ol> </li> <li>2. Deriving fusion weights- <ol style="list-style-type: none"> <li>a. MAD</li> <li>b. Query class-dependent Fusion</li> </ol> </li> </ol>
Nov 2013	[16]	MSRA-MM	<p>Topical graph method[-</p> <p>Given a textual query-</p> <ol style="list-style-type: none"> <li>1. Initial Re-ranking list obtained by current search engine</li> <li>2. Sub-graph extracted from latent graph</li> <li>3. Finally optimal re-ranked list obtained</li> </ol>	<ol style="list-style-type: none"> <li>1. Offline part: Uses image collection to learn a latent space graph.</li> <li>2. Matrix factorization: Get global and local features</li> <li>3. Online part: For sub-graph extraction.</li> </ol>	<ol style="list-style-type: none"> <li>1. Re-ranking with multi-latent topical graph</li> <li>2. Uses latent semantic analysis and construct multi latent graph</li> </ol>

### 3. CONCLUSION

Basic thing reviewed from this survey of available image retrieval and re-ranking techniques is that the text-based image retrieval is not sufficient for obtaining precise images for a given query. Thus techniques based on CBIR are found to be more vibrant and are likely to be adopted for such applications. Most of the earlier techniques used only visual features and didn't capture users' intentions. To bridge this semantic gap, method like active re-ranking has been proposed. Multi-modal graph based and circular re-ranking techniques proposed in recent years capture more than one feature of image for more accurate re-ranking results. These methods do not always compete but can complement each other.

The domain of image harvesting, retrieval and re-ranking offers a vast scope for exploration as well as innovation. This survey will prove to be beneficial to gain overview of the work done in this field.

### REFERENCES

- [1] Venkat N.Gudivada, Vijay V. Raghavan "Content-Based Image Retrieval Systems" IEEE Transaction 0018-9162, 1995 .
- [2] Edward Remias, Gholamhosein Sheikholeslami, Aidong Zhang." Block-Oriented Image Decomposition and Retrieval in Image Database Systems". IEEE Transaction 0-8186-7469-5, 1996.
- [3] Soo-Chang Pei, Senior Member, IEEE, and Ching-Min Cheng." Extracting Color Features and Dynamic Matching for Image Data-Base Retrieval". IEEE Transactions On circuits and systems for video technology, VOL. 9, NO. 3, APRIL 1999.
- [4] Yang Hu, Nenghai Yu, Zhiwei Li, Mingjing Li. "Image Search Result Clustering And Re-ranking via PARTIAL GROUPING". IEEE transaction ,1-4244-1017-7/07, 2007.
- [5] Szabolcs Sergy'an, Budapest Tech, John von Neumann ,Faculty of Informatics." Color Histogram Features Based Image Classification in Content-Based Image Retrieval Systems".6th International IEEE Symposium on Applied Machine Intelligence and Informatics-2008.
- [6] Yihun Alemu, Jong-bin Koh, Muhammed Ikram, Dong-Kyoo Kim." Image Retrieval in Multimedia Databases: A Survey". Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing ,IEEE-2009.
- [7] Jie Xia, Yun Fu, Yijuan Lu, Qi Tian." REFINING IMAGE RETRIEVAL USING ONE-CLASS CLASSIFICATION". IEEE Transaction 978-1-4244-4291-1,2009
- [8] Jes'us M. Almendros-Jim'enez ,Jos'e A. Piedra and Manuel Cant'on." AN ONTOLOGY-BASED MODELING OF AN OCEAN SATELLITE IMAGE RETRIEVAL SYSTEM".IEEE transaction 978-1-4244-9566-5 ,2010.
- [9] Xinmei Tian, Dacheng Tao, Member, IEEE, Xian-Sheng Hua, Member, IEEE, and Xiuqing Wu." Active Re-ranking for Web Image Search". IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 19, NO. 3, MARCH 2010.
- [10] K.A. Shaheer Abubacker, L.K. Indumathi." Attribute Associated Image Retrieval and Similarity Re-ranking". Proceedings of the International Conference on Communication and Computational Intelligence – 2010, Kongu Engineering College, Perundurai, Erode, T.N.,India.27 – 29 December,2010.pp.235-240.
- [11] Vedit Jain, Manik Varma." Learning to Re-Rank: Query-Dependent Image Re-Ranking Using Click Data". ACM 978-1-4503-0632-4, April 2011.
- [12] Lixin Duan, Wen Li, Ivor Wai-Hung Tsang, and Dong Xu, Member, IEEE. "Improving Web Image Search by Bag-Based Re-ranking".IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 20, NO. 11, NOVEMBER 2011.
- [13] Linjun Yang, Member, IEEE, and Alan Hanjalic, Senior Member, IEEE." Prototype-Based Image Search Re-ranking".IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 14, NO. 3, JUNE 2012.

- [14] Meng Wang, Member, IEEE, Hao Li, Dacheng Tao, Senior Member, IEEE, Ke Lu, and Xindong Wu, Fellow, IEEE." Multimodal Graph-Based Re-ranking for Web Image Search".IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 21, NO. 11, NOVEMBER 2012
- [15] Ting Yao, Chong-Wah Ngo, Member, IEEE, and Tao Mei, Senior Member, IEEE." Circular Re-ranking for Visual Search" .IEEE Transaction on image processing, vol. 22, no. 4, April 2013.
- [16] Junge Shen, Tao Mei, Qi Tian, Xinbo Gao." Image Search Re-ranking with Multi-latent Topical Graph".IEEE transaction 978-1-4673-5762-3,Nov 2013.
- [17] J. Sivic and A. Zisserman. Video Google: A text retrieval approach to object matching in videos. In Proc. ICCV, 2003.

## AUTHORS

Mayuri D. Joshi is pursuing bachelor of engineering from YCCE, Nagpur. Currently, she is studying as final year student. Her area of interest includes Digital image processing, Operating systems and Data structures.



Revati M. Deshmukh is pursuing bachelor of engineering from YCCE, Nagpur. Currently, she is studying as final year student. Her area of interest include Digital image processing, DBMS and Software Engineering.



Kalashree M. Hemke is pursuing bachelor of engineering from YCCE, Nagpur. Currently, she is studying as final year student. Her area of interest includes Digital image processing, Operating systems and Data structures.



Ashwini S. Bhake is pursuing bachelor of engineering from YCCE, Nagpur. Currently, she is studying as final year student. Her area of interest includes Image Processing, Operating systems and Data structures.



Rakhi D. Wajgi received her Bachelor of Engineering degree from Pune University in 2004. She has completed her M.E. in Computer Science and Engineering from BITS Pilani, Rajasthan in 2008. She is an Asst. Professor in Yeshwantrao Chavan College of Engineering, Nagpur. She has around 6 Yrs of teaching experience. Currently she is pursuing her PhD in Gene Regulation from Nagpur University. Her area of research includes Data Structures, Operating Systems, Parallel Programming and Bioinformatics.

