ISSN No. 0976-5697

Volume 5, No. 6, July-August 2014

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

MULTIMODAL GRAPH BASED IMAGE RETRIEVAL

Shweta Gonde
Department of Computer Science
University Institute of Technology-RGPV
Bhopal, India

Prof. Uday Chourasia
Department of Computer Science
University Institute of Technology-RGPV
Bhopal, India

Prof. Raju Barskar
Department of Computer Science
University Institute of Technology-RGPV
Bhopal, India

Abstract— Images are the increases day by day on the Internet. Retrieving relevant images from a large collection of database has become an important research topic. This paper focus on the reranking of images by utilizing the both the visual and textual features, so given a textual query in traditional image retrieval, relevant images are to be re-ranked using visual features after the initial text-based search. Here first query keywords are utilize for separating the dataset images into two group of relevant image and irrelevant image then all the images are ranked base on the image different modality of image features as the similar images need to be display closer. Using single modality is not effective as different image need different kind of feature for analysis and it was obtained in experimental that the proposed re-ranking approach has better performances than using single modality.

Keywords— Information Extraction, Text Analysis, Ontology, feature extraction, text categorization, clustering

I. INTRODUCTION

Due to the immense successful journey of text retrieval methodology which has been used for performing the task of searching the data of any type such as text data, or video, images and many more on world-wide-web named as internet.[3] The large number of systems are fully depended upon this system only, and the working procedure of this system can be defined as to find out the words associated with the images in their surrounding or exists in the form of name .as it can be seen that visual data relevancy cannot be judged easily and simple by making the use of text based retrieval process or approach, since this type of data information is generally too noisy to precisely provide the explanation about the visual content and also not available under the content of an image obtained as a result through the various search engines such as Google,Yahoo,Bing and many more as the functioning system follows the predefined sequence of steps in which it recovers rank images most of the times on the foundation of characterized data allied along the image into the précised web pages available, such as name of image and text surrounded to it. While text-based image ranking is often effective to do the searching for interrelated images, but the precision of the search results is limited by the dissimilarity amongst the true relevancy of an image and its relevance implicit from the associated textual descriptions. The prevailing procedures for image search reranking grieve from the treacheries of the assumptions underneath which are the text-based images search result. Various numbers of follow-on metaphors encompass more neither here nor there images, for the reason that of which the re-ranking conception arises to re-rank the retrieved images based on the text around the image and Meta data of data of the image and illustration quality of an image. The various integer of systems are discriminated for this re-ranking. The high ranked imaginings are worn as earsplitting data and a 'k' means algorithm for cataloging is learned to put right the ranking further. The most tremendous creativity of the overall method is in collecting text/metadata of look-alike and photographic geographies in order to have achievement of an instinctive ranking of the images.

II. RELATED WORK

In [1], Ben-Haim et al. take a subsection of images and fragments all images into blobs. At the time when clustered, densities of a blob cluster comes into the parameters of directly proportional to the relevancy of images in a cluster of the query. By making use of this idea, rests of the images are injected to applicable clusters and then the reranking of images is performed. In the same manner the main objective of the work done by Schroff et al. [9] is to practice characterized various image databases ingathered from the web. The task of re-ranking operation is accomplished for separating relevant and irrelevant, noisy results by making the usage of both of the textual and visual features in combination. In [7] a technique that uses user interference to re-rank the results is proposed. In their methodology, an insignificant subset of correct images is manually fabricated and after this the subset is then again used to filter the noise from the data. Here we have some work done exploiting multiple modalities for image search applications. Raguram and Lazebnik [12] implement joint clustering in a space built out of visualization feature and tag descriptors to treasure trove iconic extractions of abstract conceptions.
Wang and Forsyth [14] repossess entity images from the area of web pages by investigating the surrounding text together with the image itself. In [13], Schroff et.al. makes

use of the textual data that is present over there in the attachment with the web images so as to perform the task of re-ranking purposes before providing training to a SVM classifier with the help of the foundation of a visual features. The few numeral of image ranking methodologies have been encouraged by [17] in which they have shown that how to amend the page reranking [11] method with the visual domain. The methodology of Page Reranking is a type of procedure to rerank web pages according to their importance or on the foundation of priority. This technique performs his work by building up a graph representing the link structure of the web. The significance of a webpage is always presumed to be proportional to the number of hyperlinks pointing towards the particular page, i.e. the number of pages linked to it.The related work of [2] briefs a new approach for demonstrating multi-modal data sets, concentrating on the precise case of segmented images with the text associated to it. Book learning the joint distribution of image sections and words has many applications. It think through in detail forecasting words accompanying with whole images and working on a particular corresponding image regions. Auto annotation might also helped us to bring together and to have a permission of accessing a large collections of images in an abundance.
In our system the work is comprised into following sections as mentioned one by one.

A. Image Retrieval Problem

At this present age of computer system, virtually all spheres of human life including commerce, government, academics, hospitals, crime prevention, surveillance, engineering, architecture, journalism, fashion and graphic design, and historical research use images for well-organized services. A place where a large amount of images are stores and kept together is referred as image database. An image database is a system where image data is incorporated and stored [1]. Image data type take account of the rare images and information is extracted out of the images through the procedure of computerized or computer assisted image analysis. Let take some cases as a police, who maintains the database of criminals, crime scene and stolen items. If we talk about medical profession, the type of data categorized as X-ray reports and scanned images database are kept under the record files of diagnosis, monitoring and for the purpose of research also in the future can be used. If we talk about architectural and engineering design, the database that exists and stores the data in form of images or pictorial representation is for design projects, finished projects, and machine parts and specially for designing purpose. In the area of publishing and advertising, journalists maintains an image databases on paper file for different-different events and activities such as sports, buildings, personalities, national and international events, and advertisement of products are also does by them. In historical research, the databases management system is maintained for the storage of visual type of data for example like the pictorial representation of archives in areas that comes under the field of arts, sociology, and medicine. In a small size of database where collections of images are less over there, simple browsing can recognize an image. But, an image retrieval problem is the problem which is encountered and takes place when searching and retrieving of images that are relevant to a user's requested query from a database. To sort out this problem, text-based and content-based retrieval methodologies are the two techniques that have been implemented for searching and retrieval of an image in the database system.

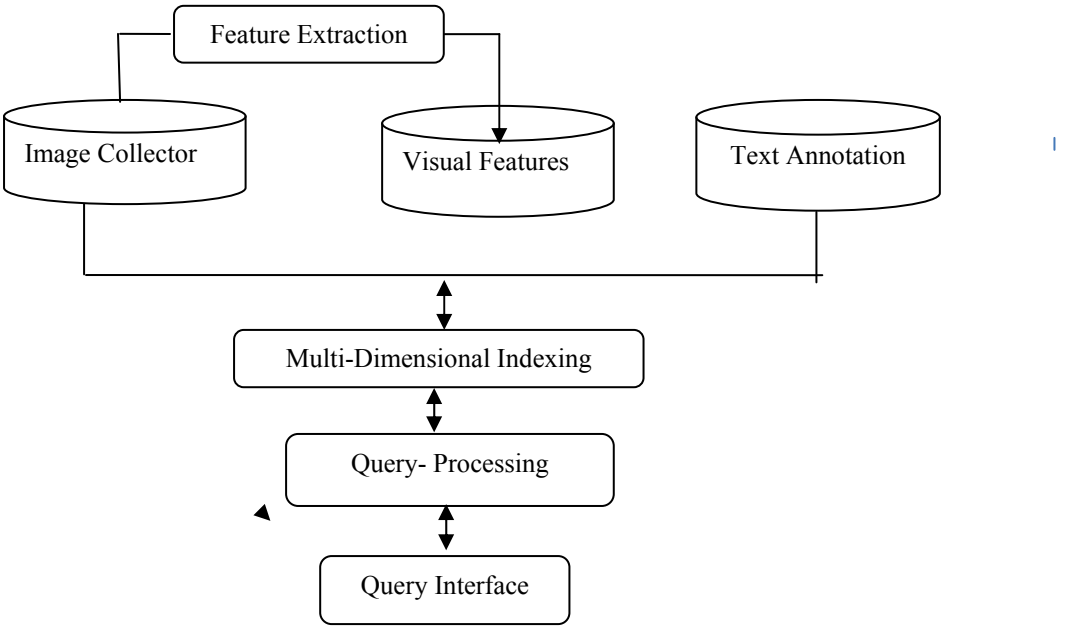


Fig. 1: Arcitecture of Image Retrieval System

B. Text Based and Content Based Image Retrieval

Under this section the task is performed by maintaining the list of those types of names which are used for specify someone or we can also say that to identify the individuality of something and for all these names an index is been maintained that contains keywords for particular word and also the subject headings, or certain codes for classification purpose, which in turn are used as words for passing the query while searching is done through search engines [2].retrieval methodology on the basis of textual data is non-homogeneous because of the reason that different-different users fire non identical keywords for explanation. This type of methodology is failed sometimes at one or the other level as it is not able to predict the actual thought if the user for some or many complex images of the database in appropriate manner. We can take an example of texture images which cannot be expressed into words. To extract the documented information about the images are quiet easy because of the already available techniques, but there is a need of human perception [16] were the image need to be described into the words for each and every image in a database. This thing is practically not possible in each and every case especially while dealing with big database, or the type of images that are of natural things by their nature example a picture captured through surveillance cameras. Not only are these but also to retrieve the images which are titled by the synonyms of the actual name also very difficult to catch. But if we talk about the image of type let us make it clear by taking an example suppose we want to search tiger in the category of animal keyword than here it is not so difficult to do the work, but till now we are facing the same scaling questions. [3]The content based image retrieval technique makes use of Meta data residing in the image to search and find out the images into the database system. This methodology when came into existence worked as to address the problems related to the text based retrieval method. Content based image retrieval works on the factor that it is a type of retrieval methodology in which semantically relevant images are retrieved from the database of an image type database where the images are automatically derived on the basis of predefined features. [4].the final goal of CBIR is to gain efficiency at the time of image indexing and retrieval, hence we are trying to reduce the help of human perception and intelligence in the working process of indexing. The digital computerized system must be capable of accessing the images from inside the database without any human postulation for any of the specific domain(example are: texture v/s non-texture or indoor v/s outdoor).other than the one of the most important task of CBIR system is to find out the comparison between the two in the basis of similarity measurement ,to extract the signature features of each and every image in the foundation of the pixel values and to establish the set of rules for the comparison of images. Now the image is compared with all other images in the database by taking or calculating the distance value among the group of images between their corresponding features.

C. Graph Based Reranking

Here we find that only a subset of text-based image search results contains relevance images and this subset usually forms a dense component in a fully-connected graph .Based on this surveillance, the functioning of this methodology is presented on the basis of similarities of the top-ranked images (based on the local descriptors) in a graph structure, where we have to find out the deepest component in the graph, and after than the ranking will be assigned to the images under the category of the heaviest component set while lower specifications re- ranks to the others. The Page Rank method is applied to perform the product image search and to design the Visual based Ranked algorithm for ranking. The Visual Rank employs the Random Walk perception to ranked images according to the visual hyperlinks among the variety of images. Intuitively, if a user is watching an image, and there are some other (visually similar) images also, then there is a probability that the user will barrier from this image to another which is in the feature with the existing one. This is corresponding to Page Rank where the prominence of a web page is commonly dignified based on the link structure.

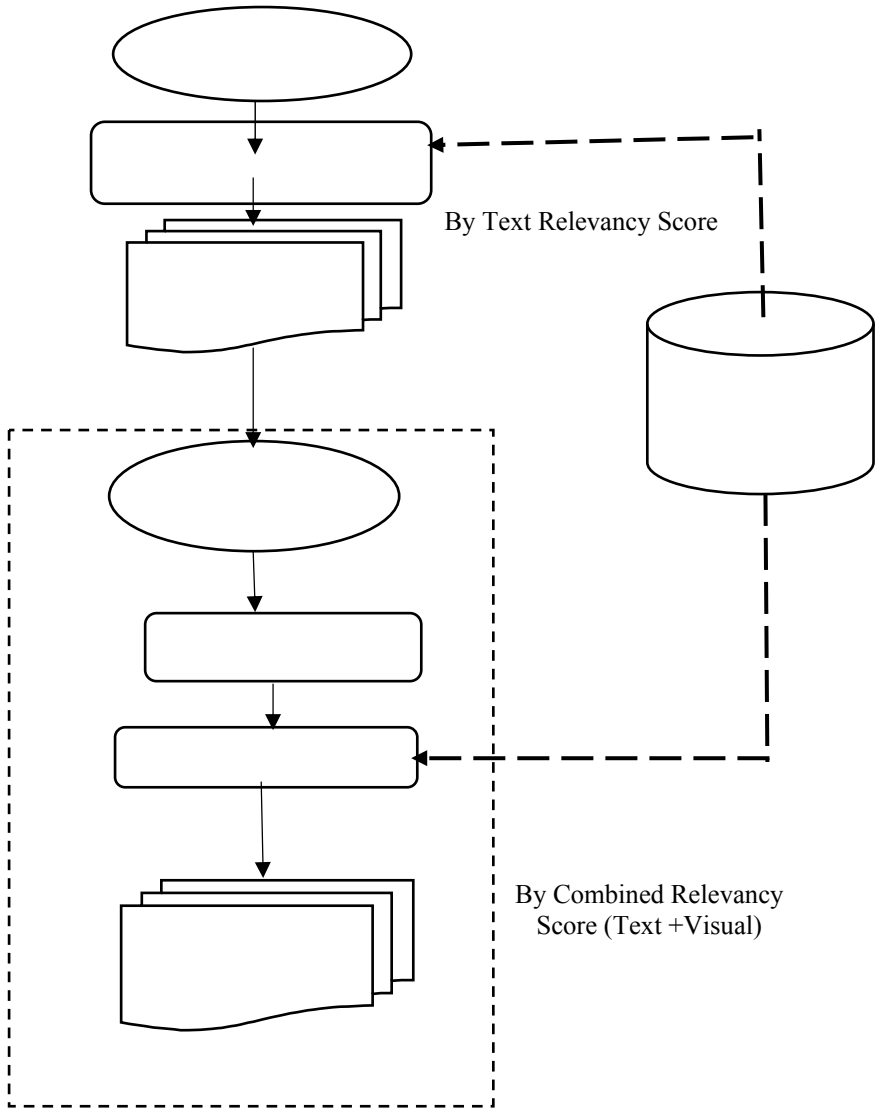


Figure 2: Working Process of TBIR and CBIR

III. BACKGROUND

This section is further divided into subsections named as:

a) Image Reranking
b) Role of Reranking
c) Modality
d) Multimodal Feature Extraction

A. ImageReranking

The term image Reranking can be defined as a particular ordering or sequencing of the number of images retrieved on the scale of some predefined measurements.[8] So it is an outcome in the form of reordering of the types of visual data when they are available there in a precise order and the set of rules or instructions applied for the reordering of visual documents can be briefed as the type of data related to that of the query or argument made by the user to fulfill the need of getting the information about a type of data can be classified as image data or visual data, as this type of procedure calmed into existence in order to overcome the limitations of only text based image retrieval. From another point of view, re-ranking can be viewed as a post-process of central part of a search. Figure 3 shows a commonly specifies the process for image search re-ranking. A visual type of document may exists in form of any type of data as it can be an image, a video clip, or a Web page containing images or videos. prearranged an initial ranked list of visual documents returned by any search come within reach of, visual search re-ranking improves search routine by reordering these visual documents on the foundation of the multimodal cue

Figure 3: A general process of Image Reranking

B. Role of Reranking

A representative visual search system comprises of several components, that inherits the query analysis, an index module, the uni-modal search (e.g., text, visual, and concept searches), and at last the re-ranking. Figure 4 shows a generic visual search framework. [5].Here basically the query which is passed through is the set of words itself (e.g., "different types of flowers") and/or probably a set of query examples (e.g., objects, images, or video key frames or clips) this is done through query analysis, the meaningful, useful and important keywords and the study of their meaning is obtained on the foundation of textual query. In the meantime, the mapping between the visual examples can be plotted to some relevant high-level concepts with the help of pre-trained classifiers for concept based search .Specifically, the scores which we have obtained as confidence scores through those classifiers can be treated as the weights for the analogous concepts (i.e., hidden text), and can also be used for further work which means can be used in a text-alike search (e.g., inverted index based on term and document frequency) or used as a feature vector in a concept space for penetrating via query-by-example (QBE). In addition to, a set of low-level visual features (global and local features) is taken out to represent these query examples for visual-based search[6].Now these multimodal queries are passed through into the singular search models, such as text, concept, and visual-based searches, respectively the confidence vectors from concept detectors or low-level feature vectors can be used in same way as the QBE or tf-idf scheme for searching. More all-inclusive extractions of content-based image and video retrieval can be found .Based on these initial search results, as well as some knowledge, a re-ranking module is functional to comprehensive the search results and reorder the initial document gradient to improve the search performance. We can observe from Figure 3 that re-ranking plays a vital role in the visual search framework to recover the preliminary search presentation.

ISSN No. 0976-5697

Volume 5, No. 6, July-August 2014

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

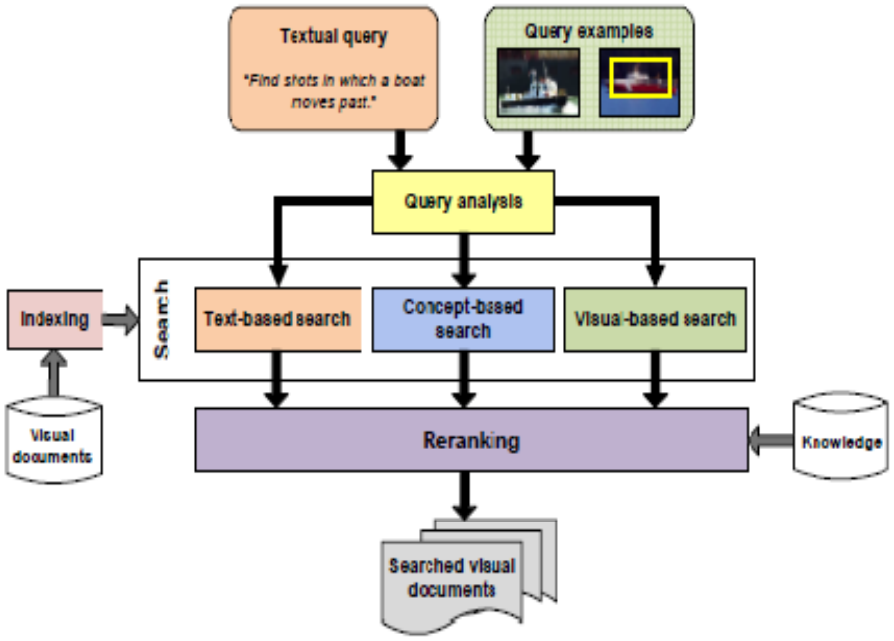


Figure 4: A general multimedia work framework

C. *Modality*

Modality can be defined as any feature or a combination of lots of features of a typical object which has been professed as an atomic information unit with reference to a provided application Now a modality is abbreviated as by the purpose of instead of by the original data, which permits us to model the existing retrieval system in more easier way and generic way. In agreement with the actual meaning of the word, modalities always stood up there as a basic building block of data management that is used to formula more multifaceted objects.

D. *Multimodal Feature Extraction*

Prevailing readings make known that the distances in the middle of samples become progressively more similar at the time when the dimension of adopted feature space is in elevation. This may bring together performance deprivation if we unswervingly apply high-dimensional features to distance based learning algorithms, such as the graph-based method adopted in this work. To make a solution and also to deal with this certain issue, as expected method that can be used over here is to interchange the high-dimensional learning assignment with the multiple low-dimensional learning responsibilities, i.e., which means we can, Separately apply the dissimilar modalities to learning algorithms and after then combining the results to have desired output [15]. A modality can be defined as or viewed as an explanation of the certain type of data such as image or video data, under which we will have some fixed type of properties resisting inside them, named as color, edge, texture, audio, and text. This type of functioning is generally named as the method known as “multimodal fusion” or “the study of multimodality”. From time to time it is also named “late fusion”, whereas the line of attack of by making the use of concatenated high-dimensional comprehensive feature vector is entitled as “early fusion” [18].

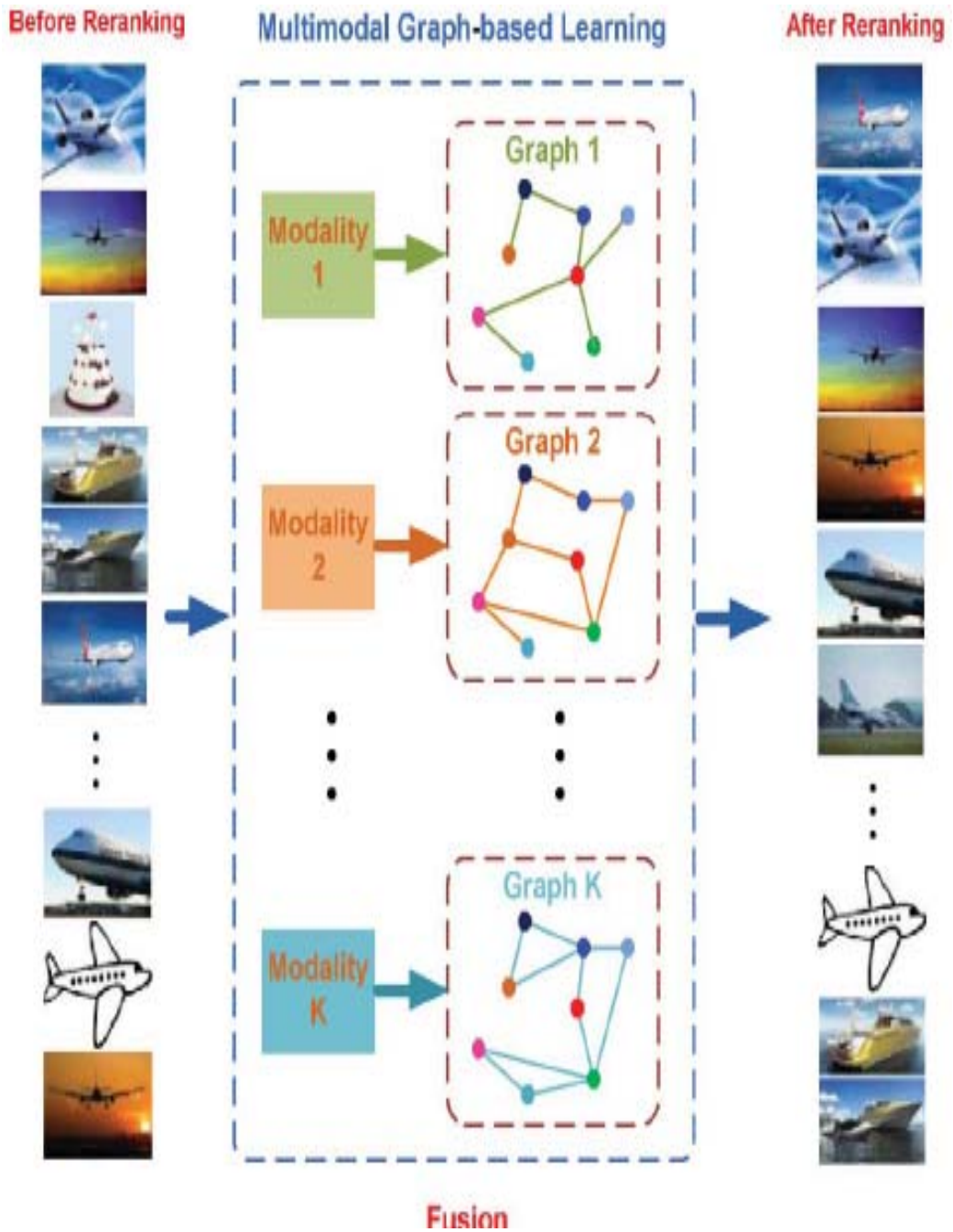


Figure 5:Multimodal Graph Based Reranking of images based on different visual features.

The contribution of this work is summarized as follows:

- A Textual and visual based learning approach for web image search reranking. It is able to integrate multiple modalities into a graph-based learning framework.
- This approach simultaneously learns the relevance scores, weights of modalities, and the distance metric and it's scaling for each modality.
- Although multiple modalities are involved for reranking but initial rank is decided by textual features which make this work efficient. Here it filter the images on the basis of the user query text, so algorithm work is limited to the images which are fetch by these text query.

There are 8 global features extracted, including

1) **225-dimensional Block-wise color moments**. Each image is split into 5-by-5 blocks, and 9-dimensional color moment features are extracted from each block.

2) **64-dimensional HSV color histogram**. Color histogram is an imperative goods for detecting and tracing as well as, not only this but also with the help of this feature. A 64-block size feature vector is generated in HSV color space for each image's acronym as Hue Saturation Value.

3) **144-dimensional Color autocorrelogram.** HSV color moments are quantized into 36 bins with 4 different pixels pair distances.

4) **256-dimensional RGB color histogram.** It is a three dimensional color representation of a particular image into which three matrices of two dimensional each is present having all the three colors red, green and blue. For finding out the mathematical value of an intensity calculation is done for gray scale format of each and every image which is again a two dimensional value ranging from 0 to 255. A 256-dimensional histogram feature vector is extracted in RGB color space.

5) **75-dimensional Edge distribution histogram.** Each image is divided into 5 blocks and 15-dimensional EDH features are extracted, as we have known that images are collection of intensity values of an image and if certain change occurs than a new change arises which takes place in form of edge feature. It is used for detection of different type of or various type of images such as buildings, scenes, roads and many more. Large number of algorithms have been developed to effectively figure out all the images out of the image or frame some of them are named as sobel, perwitt, canny etc. Out of all these mentioned "cannay" is best and very well suitable for edge detection problem which takes all the boundary values of an image

6) **128-dimensional Wavelet texture.** 128-dimensional features are extracted using the mean and standard deviation of the energy distribution of each sub-band at different levels. The image retrieved on the basis of text is disintegrated into four category images, named as low-low, low-high, high-low and high-high sub-bands[8]. Here first we have to calculate the energy level of each and every sub band one by one. This functionality come under the first level of disintegration. We make usage of low-low sub-band for keeping the further breakdown to be done continuously. The intention above and beyond the type of decomposition is to have the postulation which is defined as the energy of an image is always determined in the low-low band. Energy of all disintegrated images is calculated using energy level algorithm.

7) **7-dimensional Face features.** The features include the number of faces, the ratio of face areas and the position of the largest face region. For the purpose of face recognition the face prominent face features will allow us a pathway to the meaning full result which means we will have a high quality of performance while searching, but this single feature is not so relevant and meaning full as it doesn't set well with wrong side up indexing as this is of high-dimensional value and comprehensive and is not accessible in either computational or storage charge.

8) **Corner feature of the 256 dimensional image of gray format.** To provide stabilization in video frames in case of moving camera there is need of difference among the two frames which are pointed out by the corner feature in the image or frame. Hence by knowing the corner position of two frames one can find out how to resize the window in an original view. It is also used to find out the angles, distance between the objects of two dissimilar frames. Since it represent point in the image hence it is used to track the objective object.

For initial ranking of the images y need to be initialize as per the query so special feature which is text added with each image for fast arranging of different image as per the query of the user. Basis on the similar text of the text and the query y is ranked.

This can be understand as let the query be the 'tajmahal Agra mumtaj' ranking be given to the image with most similar text in query. Once the images is ranked as per y then it is utilize for fea

Similarity of the images for which graphs are generate.
A ranking score list, y = [y1, y2. . .yn], is a vector of ranking scores, which corresponds to a sample set X = {x1, x2, . . . , xn}. Reranking aims to obtain a new ranking score list by performing learning based on images' visual content.
Generally, graph-based reranking can be formulated as a regularization term R that makes the ranking scores of visually close similar images, and W is a similarity matrix in which Wijindicates the visual similarity of xi and xj.
Normalized graph Laplacianregularizer, i.e.

$$R(\mathbf{y}, \mathcal{X}) = \sum_{i,j} W_{ij} \left(\frac{y_i}{d_{ii}} - \frac{y_j}{d_{jj}} \right)^2 = \mathbf{y}^T \tilde{\mathbf{L}} \mathbf{y} \quad (1)$$

Where dii is the sum of the i-th row of W, and L = I−D−1/2WD−1/2 is named normalized graph Laplacian. For the loss term, usually it estimates the difference between two ranking lists. It can be defined based on either the relevance scores or ranking scores. There are several different choices, such as the squared difference and the hinge distance of relevance scores. More details and discussion on the distance of ranking lists for reranking can be found in [10]

$$R(\mathbf{y}, \mathcal{X}, a, \mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_K) = \sum_{k=1}^{K} \sum_{ij} a_k W_{k,ij} \left(\frac{y_i}{d_{k,ii}} - \frac{y_j}{d_{k,jj}} \right)^2 \quad (2)$$

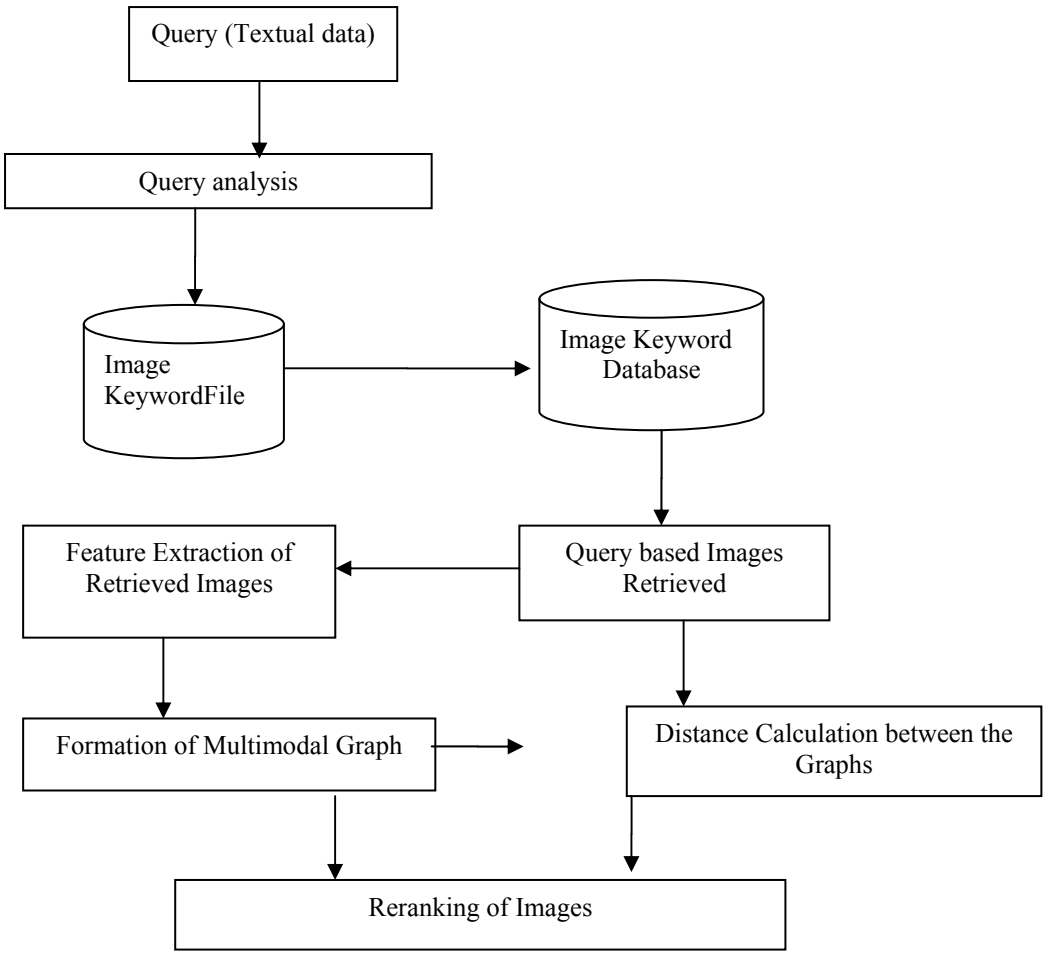
Where $W_{k,ij} = \exp\left(-\|\mathbf{A}_k(\mathbf{x}_{k,i} - \mathbf{x}_{k,j})\|^2\right)$ and $\alpha_K$is the weight for the k-th modality. The weights satisfy $0 \leq \alpha k \leq 1$ and

$\sum_{k=1}^{K} a_k = 1$. Wij is the Mahalanobis distance formula and A_K transformation matrix for the *k*th modality. After having the deep study of the following features one can easily say that the ranking of the images when the images are already there in their initial ranked stage than that process is named as re-ranking procedure and this is done so as to acquire the target in which we specify that the arrangement of images are done in a specified manner .Since, as we know that when user performs the search at that time the images should get automatically organized in a particular manner and also the procedure which has been followed in the this organization should be made on the basis of the fact that in the result it should be clear that those of the images which are closer to each other, that it should be decided by the features mentioned above and after this on the foundation of the distance value between the images one can easily form the graph for each particular feature.Now after the completion of all the steps the graph which we have now in the form of result will be utilized for the generation of new ranking pattern of the images even though there are different weight-age values for each and every modal graph.

Algorithm for Re-Ranking (Update Y parameter)

Keyword <- Query
Y <- Initialization (Keyword) // Filter image from dataset
F[n] <-Feature generation(Y)
W [M,Y,Y] = 0//Initialize weight matrix by zero of each modal
Loop m:M // M Represent number of visual features
Loop n:Y
Loop o:Y
ηt = 1
Loop k =1:T1 // T1 is random iteration
A (I+1) = A (I) k− ηt∂Q∂Ak\|Ak=A (I) k
If Q (A (I+1)) <Q (A (I) k)
W[m,n,o] <- Mahal_distance(F[n], F[o])
End Loop
End Loop
End Loop
Loop1: T2 // T2 is random iteration
Loop I = 1: M
A [I] = mean (sum (W [I]))
D [I] = diag (sum (W [I]))
L [I] =I-sqrt(D[I])*W[I]*sqrt(D[I]
Total = (total + alpha [I] * L [I])
End Loop
Y <- inverse (I + total)*Y //Update Y

Then ηt+1 = 2ηt;
Otherwise A (I+1) k= A (I) k, ηt+1 = ηt/2.
End Loop
End Loop



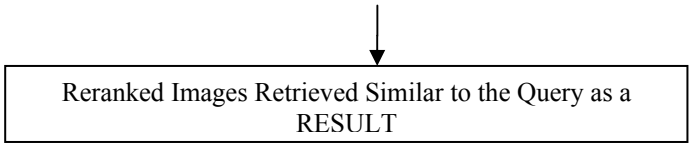


Figure 6: Working Process ofProposed Work.

IV. EXPERIMENT AND RESULT

In this section, first introduce experimental settings, and then present the experimental results that validate the effectiveness of the approach. The experiments actually contain two parts. This work is compare with other several existing methods that adopt all features.We adopt NDCG [16] as the performance evaluation measure.
The NDCG measure is computed as

$$NDCG@P = Z_P \sum_{i=1}^{P} \frac{2^{l(i)} - 1}{\log(i+1)}$$

Where Pis the considered depth, *l(i)* is the relevance level of the *i*-th image and *ZP* is a normalization constant that is chosen to let the optimal ranking's NDCG score to be 1.

A. Datasets and Preprocessing

In order to conduct the experiment an artificial dataset which is a collection of images from different category are utilize. As images are of different format so first it is necessary to make it in readable format for experiment tool MATLAB. Now these collections of images of different category are shown in table 1 for which one can make some important keyword collection for different images. In this way each image have one more feature to identify that is the keys of the images.

Table I. Dataset of Different category.

Sr No.	*Category*	*Number of Query*	*Examples*
1.	*Objects*	*50*	*sculpture, gemstones*
2.	*Places*	*90*	*Taj Mahal, Worcester, Trinity*
3.	*Animal*	*60*	*tiger, cheetah, wild dog, polar bear*
4.	*Person*	*45*	*Abdul Kalam Azad, Narendra Modi, Obama*
5.	*Cartoon*	*50*	*goofy, daisy duck, dragon tales, mickey, mini, daffy*

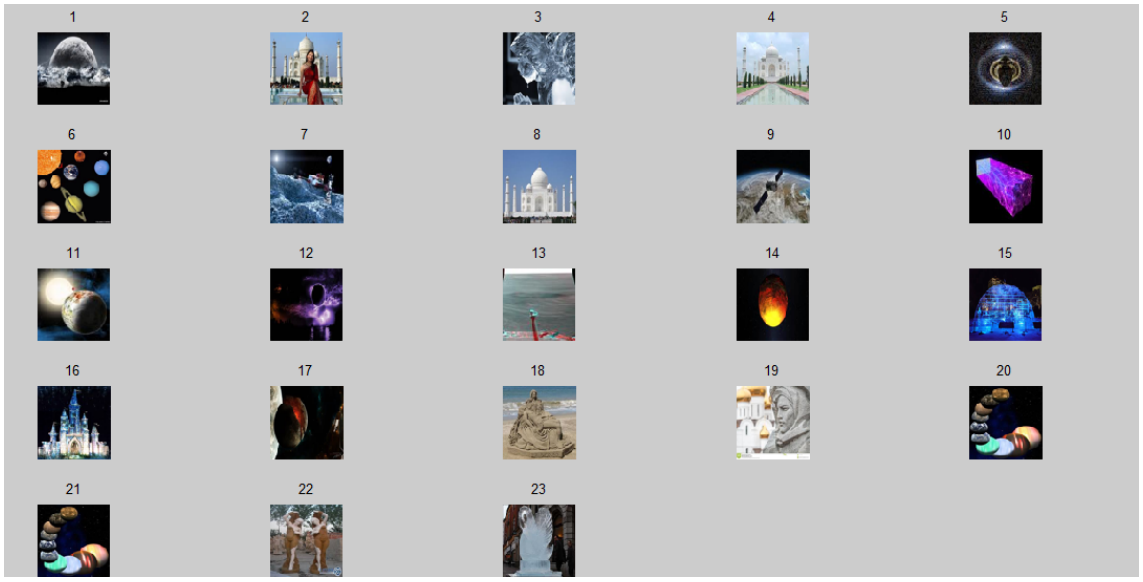
By entering the query and search the desired image it was obtained that they can be categorize into few levels such as relevant or not. It can be further categorize into most relevant, relevant, less relevant, and irrelevant.

Fig 7. Above image are generate from two category relevant and irrelevant for two query 'Taj Mahal', 'Barack Obama

Fig 7. Results obtained by proposed work and MGL [19] for Query of 'Barak Obama'.

Figure 8.1 Initial sequences of images for keyword "red apple" by proposed work

Figure 8.2: Reranked sequence of images for the query" red apple" by proposed work



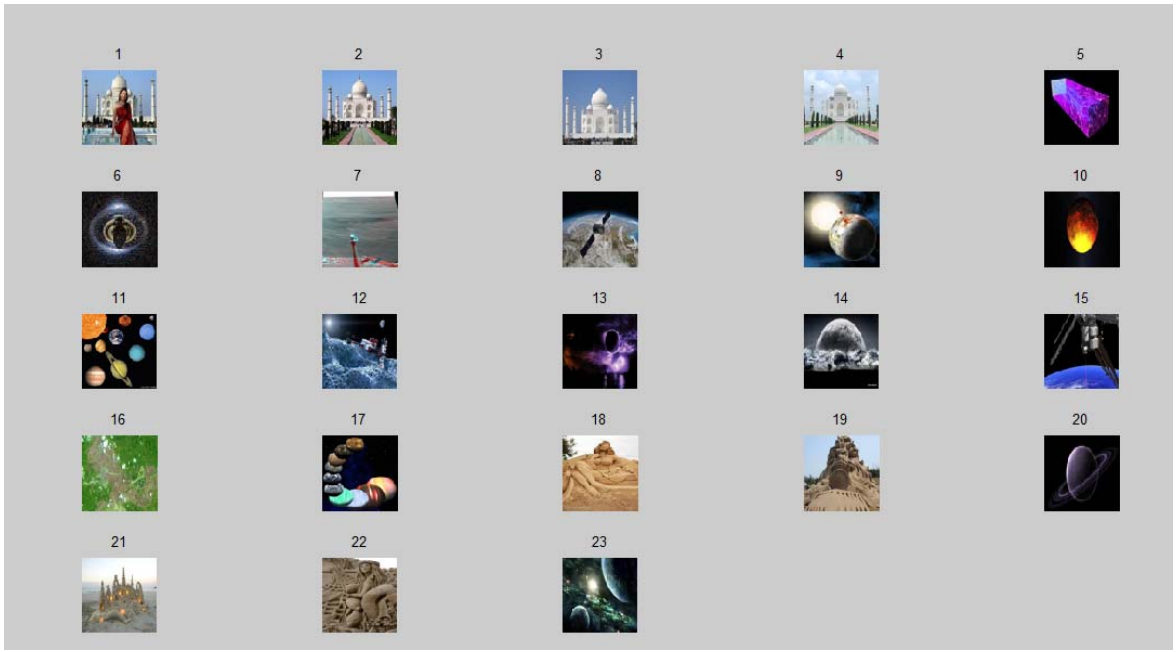
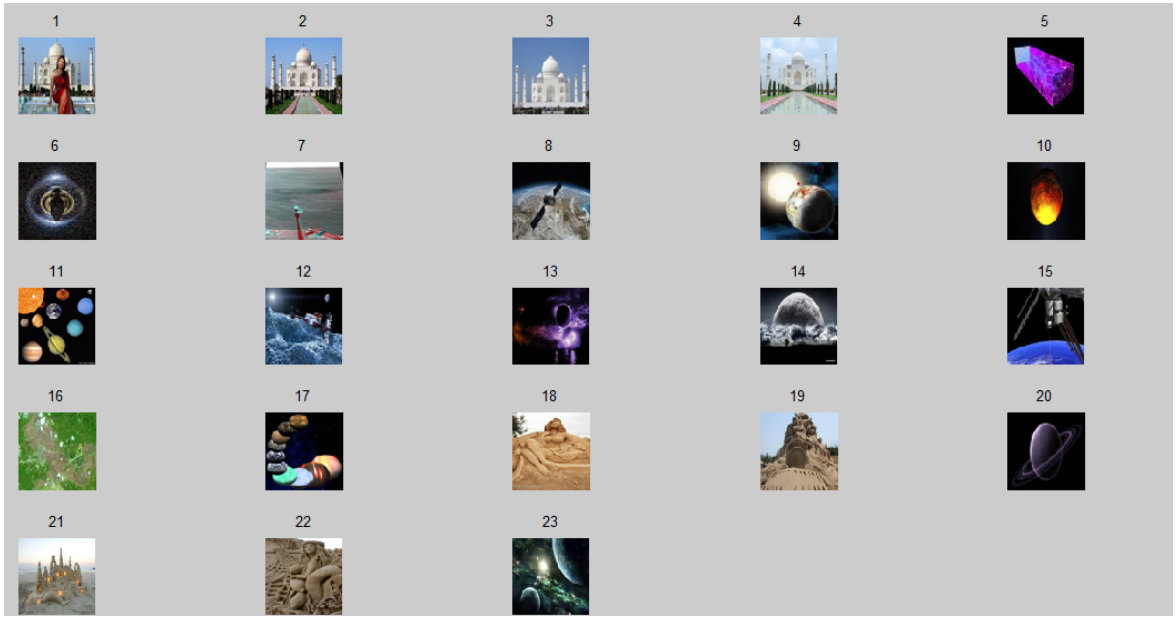


Figure 9.1: Images retrieved on the basis of keyword “a type of art” by our proposed work.



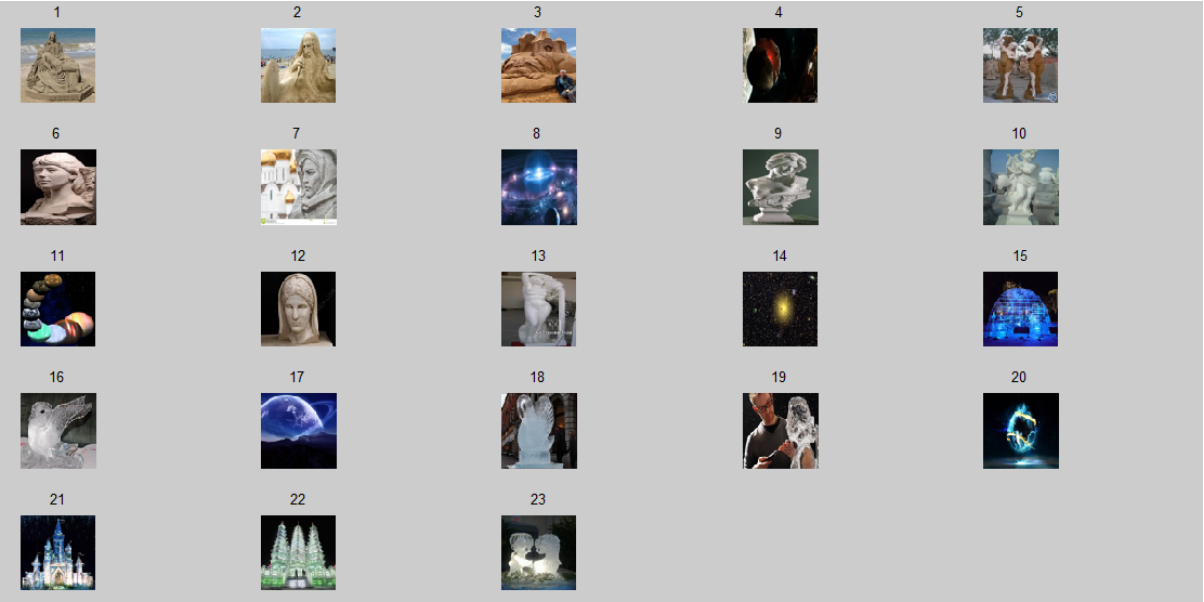


Figure 9.2: Reranked images on the basis of proposed work for the query "a type of art"

Table II. Data set used for calculation of mathematical values:

Sr No.	*Category*	*Number of Query*	*Examples*
1.	*Places*	*90*	*Taj Mahal, Worcester, Trinity*
2.	*Animal*	*60*	*tiger, cheetah, wild dog, polar bear*
3.	*Person*	*45*	*Abdul Kalam Azad, Narendra Modi, Obama*

Table III. Values of NDCG@10 by single modal and proposed method comparison

CATEGORY	*HSV color*	*histogram 36 bins*	*RGB Color*	*Edge*	*Wavelet*	*color moments*	*corner*	*Proposed*
Person	0.858	0.3972	0.7223	0.6749	0.608	0.67	0.608	1
Animal	0.5583	0.2903	0.0734	0.1585	0.8007	0.8007	0.8107	0.8207
Places	0.890	0.685	0.726	0.775	0.93	0.618	0.93	0.934

Table III. Values of NDCG@7 by single modal and proposed method comparison

CATEGORY	*HSV color*	*histogram 36 bins*	*RGB Color*	*Edge*	*Wavelet*	*color moments*	*corner*	*Proposed*
Person	0.725	0.725	0.7223	0.902	0.893	0.775	0.893	0.908
Animal	0.413	0.2798	0.096	0.198	1	1	1	1
scene	0.587	0.6067	0.3834	0.6583	0.606	0.6189	0.756	0.756

From the above table it is find that the including of the new feature of text has increase the efficiency of image re-ranking. In different categories of the images one can find that results are improved.

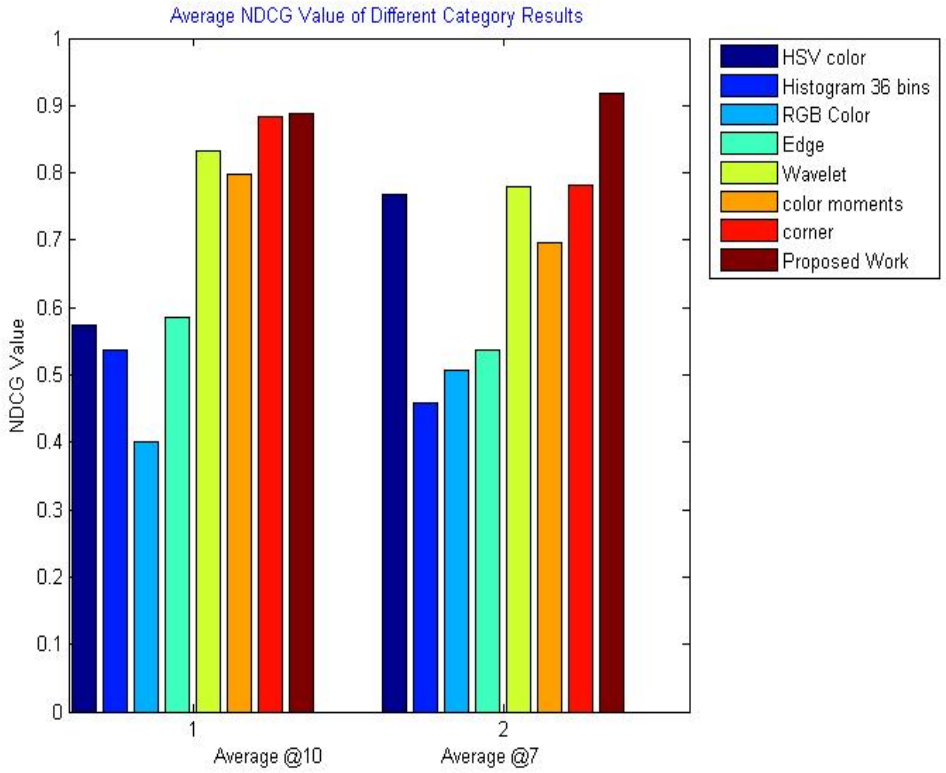


Fig 10: Average comparison of the Proposed other methods (1) denote NDCG@10 (2) denote NDCG@7

Table IV. Values of NDCG @10 for MGL and Proposed Work

CATEGORY	*MGL*	*PROPOSED WORK*
Person	*0.7827*	*1*
Animal	*0.7401*	*0.8207*
scene	*0.5779*	*0.934*

Table IV. Values of NDCG @7 for MGL and Proposed Work

CATEGORY	*MGL*	*PROPOSED WORK*
Person	**0.6166**	**0.908**
Animal	**0.4912**	**1**
scene	**0.4912**	**0.756**

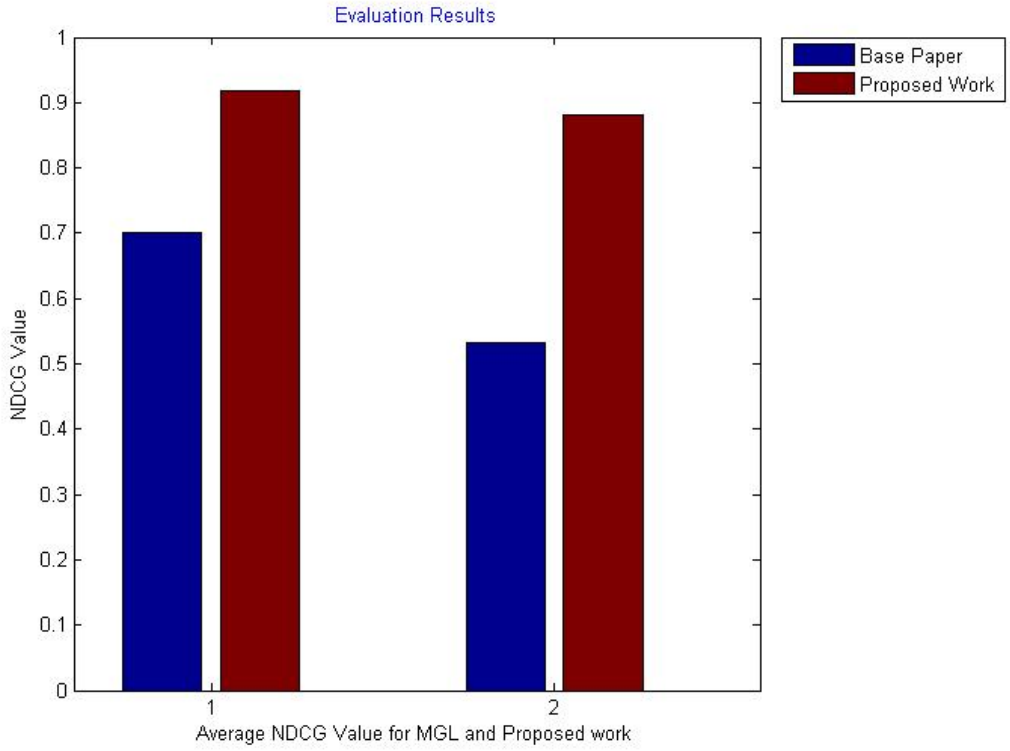


Fig 11: Average comparison of the MGL and Proposed method (1) denote NDCG@10 (2) denote NDCG@7

It is clear from the above bar graph that the values at different NDCG levels are better than the MGL individual feature method. So the utilization of both the textual as well as the visual features has increase the efficiency of the work then the previous.

V. CONCLUSION

Web image re-ranking has been widely used to reduce the user searching time on the internet; its success mainly depend on the accuracy of image features similarities. This paper present utilizing of the new text as well as visual features for ranking the image as both make the re-ranking process more powerful, which is shown in results. In future in order to improve the efficiency more features of images will be include. As to reduce the server time of making graph and re-ranking one filter need to be inserted into this so that it will filter relevant and irrelevant images at the initial stage.

REFRENCES

[1].Meng Wang, Hao Li, Dacheng Tao, Ke Lu, and Xindong Wu "Multimodal Graph-Based Reranking for Web Image Search" IEEE Transaction on image processing Vol. 21, NO. 11, November 2012.

[2]. Pushpanjali M. Chouragade, Prashant N.Chatur, "Visual Re-rank: An Approach for Image Retrieval from Large-scale Image Database" Volume 3, Issue 3, March 2013

[3].Shweta Gonde, UdayChourasia, RajuBarskar"A survey on web image search using reranking",International Journal of Engineering and Computer Science ISSN: 2319-7242 Volume 3 Issue 5 may, 2014.

[4]. M. Wang, X. S. Hua, R. Hong, J. Tang, G. Qi, and Y. Song, "Unified video annotation via multigraph learning," IEEE Trans. Circuits Syst. Video Technol., vol. 19, no. 5, pp. 733–746, May 2009.

[5]. Tao Mei, Microsoft Research Asia,Yong Rui, Microsoft Research Asia Shipeng Li, Microsoft Research Asia Qi Tian, University of Texas at San Antonio, "Multimedia Search Reranking: A Literature Survey",ACM Journal Name, Vol. 2, No. 3, Article 1, Publication date: May 2012.

[6]. N. Morsillo, C. Pal, and R. Nelson. "Mining the web for visual concepts", In 9th KDD Multimedia Data Mining workshop, 2008.

[7]. L. Page, S. Brin, R. Motwani, and T. Winograd. "The pagerank citation ranking: Bringing order to the web", Technical report, Stanford Digital Library Technologies Project, 1998.

[8]. F. Schro®, A. Criminisi, and A. Zisserman. "Harvesting image databases from the web", In Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference , Oct. 2007.

[9]. Y. Jing and S. Baluja. "Visualrank: Applying pagerank to large-scale image search" IEEE Trans. Pattern Anal. Mach. Intell. , 2008.

[10]. N. Ben-Haim, B. Babenko, and S. Belongie. Improving web based image search via content based clustering. In IEEE Xplore SLAM, New York City, NY, 2006.

[11]. D. Ozkan and P. Duygulu. A graph based approach for naming faces in news photos. In IEEE
Conference on Computer Vision and Pattern Recognition, New York City, NY, 2006.
[12]. K. Barnard, P. Duygulu, N. de Freitas, D. Forsyth, D. Blei, and M. Jordan. Matching words and pictures. J. Machine Learning Research, Feb 2003.
[13]. G. Iyengar, H. J. Nock, and C. Neti, "Discriminative model fusion forsemantic concept detection and annotation in video," in Proc. ACM Multimedia,pp. 255–258 in 2003.
[14]. R. Yan and A. Hauptmann, "The combination limit in multimedia retrieval," in Proc. ACM Multimedia,pp. 339–342 in 2003.
[15]. F. Schroff, A. Criminisi, and A. Zisserman. Improving web based image search via content based clustering. In International Conference on Computer Vision, Rio de Janeiro, Brazil, 2007.
[16]. X. Tian, L. Yang, J. Wang, Y. Yang, X. Wu, and X. S. Hua, "Bayesian video search reranking," in Proc. 16th ACM Int. Conf. Multimedia, pp. 131–140, 2008.
[17] Mayuri D. Joshi , Revati M. Deshmukh, Kalashree N.Hemke, Ashwini Bhake, "Image Retrieval using Re-Ranking Algorithms" International Journal Of Core Engineering & Management(IJCEM) Volume 1, Issue 1, April 2014
[18] A. Smeulders, "Content-Based Image Retrieval at the End of the Early Years," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349-1380, May. 2000.