© 2017 IJSRST | Volume 3 | Issue 6 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X
Themed Section: Science and Technology

An Enhanced Approach for Image Search Based on Attribute-Assisted Re-Ranking Model

Renuka S. Deshpande, Swati A. Patil

Computer Science and Engineering, GHRIEM Jalgaon Maharashtra, India

ABSTRACT

Image search re-ranking is an effective approach to refine the text-based image search results. Most existing re-ranking approaches are based on low-level visual features. Here, semantic attributes are exploited for image search re-ranking. In the existing system, image features are extracted through K-means clustering and then the histogram is generated. The extracted features were then used for attribute learning using Support Vector Machine (SVM) classifier. After that the attribute assisted hyper-graph is constructed to re-rank the images. In the proposed system Keyword expansion using word Net and user keywords are used where word net gives the synonyms for the query and user keywords gives the similar meaning words for the improvement in the search result. Image features are extracted through discrete cosine transform and perceptual hashing. Then histogram is generated for all the images. Attribute learning is carried out using multi SVM. The new image re-ranking agenda focuses on the semantic signatures associated with the images derived using a trained multi class classifier. A query image is selected and then the most similar three images to the query image are displayed. Then these most similar images are further compared with the remaining images and are arranged in order according to their visual appearance and by comparing their hash values using Euclidean distance. The experiments are carried out on multiple data sets.

Keywords: Image Search, Image Re-Ranking, Semantic Signature, Keyword Expansion.

I. INTRODUCTION

The image Re-ranking is an efficient technique to improve the results of web based image search [1]. It has been implemented by existing industrial search engines like Google, Bing and so many. For a given query keyword, search engine re-ranks the group of images based on the query. The user is prompt to pick a particular image from the group, the rest of the images are re-ranked based on the user selected image.

Image search is a widely growing feature of search engines namely `Google', `Yahoo', `Bing', etc. For a given text query, the search engine has to search for millions of images for retrieving the relevant images as early as possible. Generally the search engines are mainly based on using text meta-data namely keywords, tags, and/or text descriptions close to the images. The retrieval of images is typically grouped with undesirable non-relevant images as the meta-data do not constantly be in associate to the visual term of the images, Then also the retrieved images encloses enough relevant images which they are prepared for users that are in common more noticed by precision than recall. Then by re-ranking the original set of retrieved images the precision can be enhanced. Recently user's efforts are reduced by online image re-ranking method in which efforts are limited to just one-click feedback which is a useful way to improve search results.

Image retrieval has attracted significant attention in both academia and industry with the dramatic increase of online images. Many image search engines such as Google and Bing rely on matching textual information of the images against queries given by users. Text-based image retrieval suffers from difficulties and unreliability's that are caused mainly by the incapability of the associated text data to appropriately describe the content of image. Recently, to refine text-based search results by exploiting the visual information contained in the images visual re-ranking has been proposed. The existing visual re-ranking methods can be categorized into three categories as clustering based, classification based and graph based methods. The clustering based re-

ranking methods shows from the key observation that a wealth of visual characteristics can be shared by relevant images. With intellectual clustering algorithms like mean-shift, K-means, and K- medoids, initial search results from text-based retrieval can be group by visual convenience. However, the performance of the clustering-based methods is not guaranteed for queries that return highly diverse results or without clear visual patterns,. In the classification based methods, visual re-ranking is formulated as binary classification problem which aims to identify whether each search result is relevant or not. Pseudo Relevance Feedback (PRF)[2] is applied to select training images to learn a classifier or a ranking model. However, in many real life scenarios, representative examples obtained via PRF for the training images in dataset are very noisy and may not be adequate and useful for constructing effective classifiers. Graph based methods have been proposed recently and received increasing attention to be more effective. A collection of nodes and edges are used to represent multimedia entities in top ranks and their visual relationship. The local patterns or salient features discovered using graph analysis is very powerful and important to improve the effectiveness of rank lists. Yet, the re-ranking algorithms mentioned above are purely based on low-level visual features which generally do not consider any semantic relationship among initial ranked list. The high level semantic concepts which are crucial and hard to capture property of images could deliver more clearly semantic messages between various nodes in the graph. Thus, we propose to exploit stronger semantic relationship in the graph for image search re-ranking.

Alternatively, semantic attributes have received tremendous attention, where their effectiveness is demonstrated in broad and large applications like face verification, object recognition, fine-grained visual categorization and image search uses semantic attributes. Semantic attributes could be shape, color, texture, material, or part of objects, such as square, orange, mental, wheel and leg etc. An attribute has semantic meaning as opposed to low-level visual features as a kind of intermediate-level descriptor,. Thus attributes are expected to narrow down the semantic gap between low-level visual features and high-level semantic meanings. Due to spontaneous interpretation and cross-category, generalization property attribute-based image representation has also shown great promises for discriminative and descriptive ability. They describe image regions that are common within an object category but uncommon outside of it. Hence, attribute-based visual descriptor has achieved good performance in assisting the task of image classification. Besides that, an attribute is potentially any visual property that human beings can specifically communicate or understand, even if it does not correspond to a traditionally defined object part. Furthermore, the type of the most effective features should vary across queries. For example, color features will be useful for queries that are related to color distribution, such as sunset, sunrise and beach, Also queries like building and street, edge and texture features will be more successful. It can be understood that semantic attribute could also be viewed as a description or modality of image data. The multi modal [3] features can guarantee and assure that the useful features for different queries are contained. So, all these superiorities drive us to exploit semantic attributes for image representation in the task of image search re-ranking.

II. LITERATURE SURVEY

In the past, there was lot of research done on the image retrieval. The huge number of available images on the web makes image retrieval a challenging task. There are various techniques used by image retrieval to retrieve the most relevant image. The quality of results generated by the image retrieval system still not meets the user requirement. Due to lack of visual features and semantic signature quality results reduces drastically. Hence to overcome this drawback visual query expansion and keyword expansion is necessary.

2.1 Background

A number of techniques have been put forward for better performance of web image re-ranking in the recent days. Web image re-ranking, is a mixed process of both text based image result and visual features to obtained good quality performance in image search. The process can be explained by the way in which text query is given by the user and the search is done. It returns few unequal images which are not related to query and not useful to user. After that user selects the query image and again the search is done. Here, we get the most relevant image

based result, after that re-ranking is done. Image search re-ranking can be classified into two methods, supervised and unsupervised re-ranking. The unsupervised re-ranking method do not rely on user labeling of relevant images but require previous assumptions on how to utilize the information contained in underlying text based result for re-ranking. The challenge of supervised re-ranking method is to design query-independent re-ranking models based on query dependent re-ranking features.

Major web image search engines have adopted the strategy. A query keyword as a input is given by a user and lots of images relevant and similar to the query keyword are retrieved by the search engine according to a stored image database. After that user select a query image which observes its search objective from the set, the remaining images in the set are re-ranked based on their visual similarities with the query image given by user. The text-image index file and visual features of images are previously calculated offline and are stored, and then the web image collection is dynamically upgraded. If the visual features are not selected and only the similarity scores of images are stored then whenever a new image is added into the collection, similarities are computed with existing images, and then the visual features of images need be computed again.

2.2 Image Retrieval

Numerous search engines are available nowadays such as Bing, Google, Cydral, Yahoo, etc. which fulfills image search queries. Users express their needs in search engines, and related images are displayed as output. For improvement in the search results two methods are used they are image annotation and web image search re-ranking. The ever-growing large amount of digital images on the Internet, retrieving relevant images from a large set of databases has become an important research study. Over the past years, many image retrieval systems have been developed [4]; those are Text-based image retrieval, Content-based image retrieval and Hybrid approach.

2.2.1 Text-Based Approach

Text-based search technique is more effective and efficient in document and for image search. User's gives text query as an input. These text-based queries can be formulated as free-text and it can be compared to such text descriptors as description, subjects, title or the text surrounding an embedded image, using the text retrieval techniques. The main problem and difficulty of web image search is the variance between the image content and the web page text. The TBIR has been widely used in popular image search engines e.g. Google, Bing and Yahoo!. A user gives an input as a textual query to the retrieval system, then the system returns the ranked images whose adjacent texts contain the given query keyword, and the ranking score is obtained according to some similarity and corresponding measurements between the query keyword of user and the textual features of relevant images. Text-based search techniques have been established to perform well in textual documents; they often result in variance when applied to the query image search as the metadata cannot represent the semantic content of images.

2.2.2 Content-Based Approach

The majority of search engine works on Text Based Approaches but there exist alternative. In Content based image retrieval (CBIR) the visual features are extracted, such as color, texture and shape of images automatically and detect the similarity between images by distances in the features.

Implementation is simple and retrieval is fast here. Appropriate feature representation and a similarity measure for ranking images, given a query, are necessary here. Most of the Content based image retrieval systems performs feature extraction through images as a preprocessing step. It shows the procedure of CBIR where Features may include both text based features such as key words, annotations, tags and visual features like color, texture, shape, faces. Content based image retrieval that requires a user to submit a query image, and return images that are similar and relevant in content .Google is one of the search engines that works on this image re-ranking. The extracted visual information through this is natural and objective, but completely ignores the role of human understanding or knowledge in the interpretation process.

2.2.3 Hybrid approach

Recent research combines both the visual content of images and the textual information of images obtained from the Web for the WWW image retrieval. These methods make use of the usage of the visual information for refinement of the initial text-based search result. Particularly, through user's relevance feedback, that is, the giving in of desired query images or visual content-based queries, the re-ranking for image search results can achieve significant and good performance improvement.

III. PROPOSED WORK

With reference to literature work, various web Mining techniques used image re-ranking to solve the problem of lack of image retrieval. Use of visual query expansion and semantic signature found very useful in solving issues related to image retrieval, semantic complexity, re-ranking issues. The proposed solution introduces a new approach by using semantic signature and visual query expansion for accurate image retrieval from the database for effective re-ranking.

3.1 Proposed Approach

The proposed approach describes the solution to image retrieval. An existing system consist drawback of lack of image retrieval and proposed solution overcomes this drawback by using visual query expansion and semantic signature. Accurate and most relevant image retrieval is possible only when some technique are said to exist. Visual query expansion is provided using hashing, feature extraction and keyword query expansion provided using synonyms. Visual query expansion and semantic signature is best for such task. Features are extracted at the time if images inserted into database. Keyword expansion and visual expansion act as a technique to retrieval of most relevant image and re-rank other images based on their Euclidean Distances. In this way most relevant image is retrieved which is hidden in the database. Use of semantic signature and visual query expansion as a web mining technique in various data mining techniques found effective for improving the precision of image retrieval.

3.2 Proposed System Architecture

Figure 1 shows architecture of the proposed system. There are two phases: Training phase and Testing phase. Features are extracted at the time of image insertion in database. Semantic signatures of any image queried by the users are calculated and stored in database at the Training phase. The majority of the job is done at the Training stage. At the Testing phase, the user receives re-ranked images that are calculated using semantic signatures at the Training phase. Instead of developing a universal concept dictionary a novel framework is proposed for web image search re-ranking. It learns different visual semantic spaces for different query keywords independently and automatically.

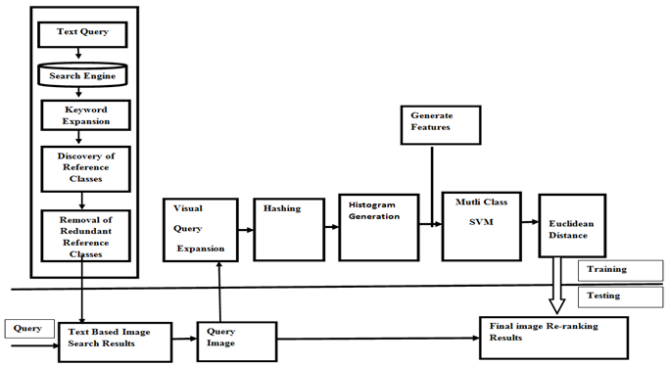

Figure 1: System Architecture Attribute-Assisted Re-ranking Model

3.3 Keyword Expansion

The target with keyword expansion [5] is to mark new keyword opportunities by discovering various synonyms that are both relevant and traffic driving. By adding additional relevant keyword terms to the original queries improves the performance of information retrieval systems.The query image to numerous positive visual examples from which new query specific visual and textual similarity metrics are obtained and learned are expanded from Expanded keywords for the given query. From the Methods which lengthened query keywords with their linguistically connected words such as synonyms. Word Net dictionary is been used here to differentiate word senses of query conditions. Whenever the sense of a query term is decided its synonyms. Keyword1 generate the synonyms which are given by user at the time of image insertion. Then keyword2 is use to find the synonyms through WorldNet. Based on

the keyword1 and keyword2 images are retrieved for the further process. The keywords provided by users tend to be short. They are unable describe the content of images accurately. The query keywords meanings may be more affluent than user's expectations. For example, the meanings of the word apple comprise apple fruit, apple computer, and apple iPod. The user may not have sufficient information on the textual description of target images.

Algorithm:

1. Procedure: Generate Synonyms of Keyword
Require: input query
2. Repeat
3. Find Synonyms of input keyword
Word-Net:-
Step 1: Find the Noun and Verb senses for the entered word from word-net
Step 2: For checking similarity between two words, Counts the noun and Verb senses,
Using

$$\text{Score} = A \cap B / \text{Min}(A, B)$$

4. until all synonyms of input keywords are not found
5. End procedure

3.4 Visual Query Expansion

The purpose of visual query expansion is to reach multiple positive example images to learn a visual similarity metric which is stronger and more specific to the query image. Visual Query Expansion develops an image re-ranking method, which only needs one click on the query image and thus positive examples have to be attained repetitively. The chosen image has the closest visual distance to the query instance and has reliable semantic meanings. As a result, they are used as added positive instances for visual query expansion.

The aim of visual query expansion is to obtain multiple positive images to learn a visual similarity metric which is stronger and more specific to the query image. One query image is not enough to capture the user's intention only. These similarity metrics react user intention at a finer level since every query image has different metrics. Visual query expansion also does not require users' feedback. Here, user select one image from the keyword based image result and based on that one image four images are retrieved using feature extraction in visual query expansion. The distance between those images are nearly similar. In this process, the image is firstly converted to RGB matrix then to Grey scale matrix. The grey scale matrix is further reduced to Gabor matrix.

3.4.1 Perceptual Hashing

Hashing is converting the image into gray scale and getting the hash value [6].In hashing firstly feature from image is extracted. Then the second stage compresses this feature vector to a final hash value. Perceptual hashing is a robust algorithm widely used for content identification. P-hash is reliable and fastest algorithm. The hash string is generated using the color of the images. Pixel color average is taken as mean value based on the average rate the hash string is formed. The image is scaled to defined small size. After the generation of the hash value, color pixel and its average color pixel value. With respect to the average of total color pixels and each color pixels of the image the Hash bit is generated. Then DCT is applied.

Algorithm: Perceptual Hashing

```
I←Input Image
I Convert Grey Scale(I)
I←Scale to a defined size of w and h
for← i  0 to w
for← j  0 to h
sum ← sum + colorpixel(i,j)
colors[] ←sum
avg ←sum/(w*h)
for each pixelval in the list of colors
if pixelval>avg then
hashbit ←1
else
hashbit ←0
loop ←loop+1
if not loop then
hash ←hash ++ hexadecimal(hashbit)
endif
```

3.4.3 Classifier of Reference Classes

Images are classify as per there feature for further process. Then one class SVM is adopted to improve the visual similarity. This takes the re-ranked image as input to the one-class SVM classifier and similarity to the

query image is used as output. User gives the class name at the time of images insertion in database. The classifiers show a lot of effect, a lot of training data is demanded for satisfactory performance as many parameters to be estimated. SVM is a supervised machine learning algorithm which can be used for issues related with classification or regression. An optimal boundary between the possible outputs is found by making use of a kernel trick technique transformations to transform the data.

SVM is abbreviated as Support Vector Machine used as an image processing classifier to classify images according to the visual content present in them. SVM is capable of providing a good performance in the domain of pattern recognition. The user can improve the selected image in this module with the image parameters. It is necessary to discover semantic features for each image. Semantic feature discovery is based on multiple modalities such as image contents and tags. Using semantic signature over reference classes user get the appropriate intention image. Reference classes give the different categories for image search to get the most accurate image.

Algorithm: MultiClass SVM [7]

for j=1 to (k-1) do

- For all samples from C1 to Cj classes, set labels to (+1)and all samples from Cj+1to Ck, set labels to (-1)
- Train jth binary SVM
- Classify the training samples
- if (j > 1), compute fuzzy scores p for all training samples _p classified as (+1) and define (j-1) thresholds by splitting the curve of sorted relevance scores into equally spaced intervals.
- if (j < k), compute fuzzy scores _n for all training samples _n classified as (-1) and define (k-j-1) thresholds by splitting the curve of sorted relevance scores into equally spaced intervals.

end for

Testing: Classification of a new sample zl

for j = 1 to (k-1) do

- Classify zl by jth model
- if (zl is classified as (+1))
  if (j = 1) $class_j$ $z_l \leftarrow c_1$ else use the defined thresholds to decide $class_j$ z_l
  else
  if (j = k-1) classj $z_l \leftarrow c_1$ else use the defined thresholds to decide classj z_l

end if
end for

3.4.4 Generate Histogram

An image histogram is a form of histogram that acts as a graphical representation of the tonal distribution in a digital image[1]. For Each tonal value the number of pixels are plotted. By looking at the histogram for a specific image an observer will be able to judge the entire tonal division at a glance.

Image histograms are present on many modern digital cameras. Photographers can use them as an support to show the distribution of tones captured, and whether image detail has been lost to blown-out highlights or blacked-out shadows. This is less useful when using a raw image format, as the dynamic range of the displayed image may only be an approximation to that in the raw file.

g(x,y) generate the graph for each image,

$$g(x,y)= \frac{f(x,y) - f,min}{fmax - fmin} * 2^{bpp}$$

Where f is pixel intensity, fmin is minimum pixel intensity and fmax is maximum pixel intensity, bpp be the bit per pixel.

3.4.5 Feature Extraction

In the feature extraction step, the two dimensional image is mapped to a one dimensional feature vector. Feature extraction plays an very important role in image processing. Features are use to classify and recognize the images. As a feature defines the behavior of the image

Generate Features:

Find Mean, deviation for retrieved image

Mean= $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x$

where n is the sample size and the x correspond to the observed valued.

Deviation We define the ***deviation*** to be

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x-\bar{x})^2$$

and the ***standard deviation*** to be

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x-\bar{x})^2}$$

3.5 Euclidean Distance

The Euclidean distance is the straight-line distance between two pixels. It is the distance between two points in Euclidean space. The two points in two dimensional Euclidean spaces where P with the coordinates (p1, p2) and Q with the coordinates (q1, q2). The line segment with the endpoints of P and Q form the hypotenuse of a right angled triangle. The distance between two point's P and Q can be given as the square root of the sum of squares of the differences between the coordinates of the points. The two-dimensional Euclidean geometry, the Euclidean distance between two points can be given as: a = (ax, ay) and b = (bx, by) is defined as:

$$\boldsymbol{d(a,b) = \sqrt{(bx-ax)2 + (by-ay)2}}$$

3.6 Re-ranking

At the online stage, after finding Euclidean distance the images are re-ranked. A process in which images are retrieved, arranged as per their features and user need is called image re-ranking. The re-ranking of the returned images based on metadata of image and text surrounding of image alone. Expand and follow the method proposed by using a set of textual features whose presence is great indication of the image content.

The main aim is to re-rank the searched images. Each feature is used as binary: "True" if it contains the query text and "False" if it doesn't contains the query text. To re-rank images for one fix class, it is not necessary employ the whole images for class.

IV. RESULT

The proposed system is tested on the image dataset which contains 2000 images with 15 classes along with their subclasses. The particular class contains the number of images with their feature extraction. Feature can be extracted or calculated at the time of image inserted in dataset. After image re-ranking Precision, Recall and F-measure is calculated.

Precision :- Precision is the basic measure used in evaluating search strategies. Precision means how many selected items are relevant. It is shown in Figure 4.1. Precision is calculated by using following formula:

$$\textbf{Precision} = \frac{\text{Number of Relevant Retrieved}}{\text{Number of Retrieved}}$$

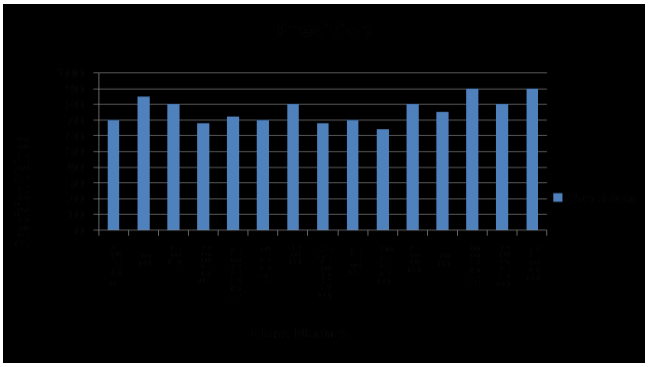

Figure 4.1: Bar Graph for Precision

Recall :-Recall is the ratio of the number of relevant records retrieved to the total number of relevant records in the database. It is usually expressed as a percentage. Recall means how many relevant items are selected. It is shown in Figure 4.2

Recall is calculated by using following formula:

$$\textbf{Recall} = \frac{\text{Number of Relevant Retrieved}}{\text{Total number of Relevant}}$$

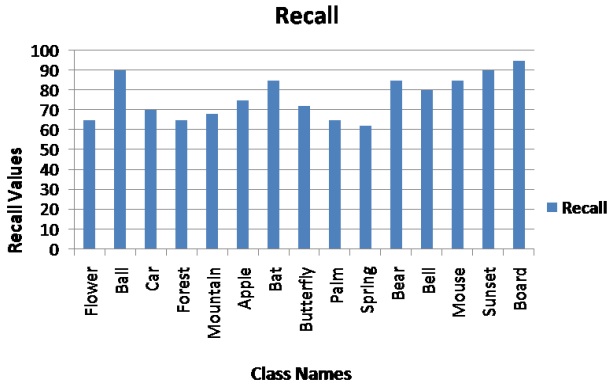

Figure 4.2: Bar Graph for Recall

F-measure:- F-measure is a Frequency measure of a tests accuracy. It considers both
the precision and recall of the test to compute the score. It is shown in Figure 4.3

F-measure is calculated from precision and recall by using following formula:

$$\textbf{F-measure} = \frac{2*\text{Precision}*\text{Recall}}{\text{Precision}+\text{recall}}$$

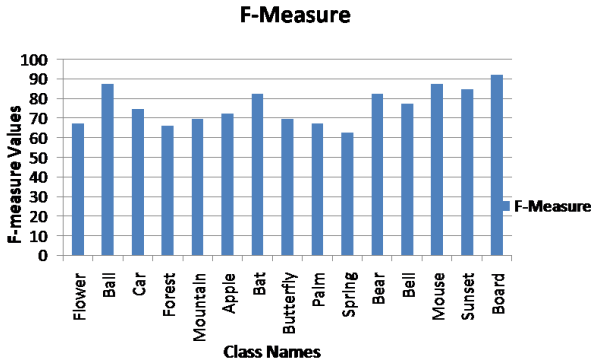

Figure 4.3: Bar Graph for F-measure

Precision and recall are the basic measures used in evaluating search strategies.

Table 1 shows the parameter calculation for classes.

Class Names	Precision	Recall	F-Measure
Flower	70	65	67.40741
Ball	85	90	87.42857
Car	80	70	74.66667
Forest	68	65	66.46617
Mountain	72	68	69.94286
Apple	70	75	72.41379
Bat	80	85	82.42424
Butterfly	68	72	69.94286
Palm	70	65	67.40741
Spring	64	62	62.98413
Bear	80	85	82.42424
Bell	75	80	77.41935
Mouse	90	85	87.42857
Sunset	80	90	84.70588
Board	90	95	92.43243

Table 1: Calculation of Parameters

V. CONCLUSION

Image search re-ranking has been studied for several years and various approaches have been developed recently to boost the performance of text-based image search engine for general queries. Semantic attributes are expected to narrow down the semantic gap between low-level visual features and high level semantic meanings. Motivated by that, the enhanced approach for image search based on attribute assisted re-ranking is proposed. Based on the classifiers for all the predefined attributes, each image is represented by an attribute feature consisting of the responses from these classifiers. Thus, by using some more features in the existing attribute assisted re-ranking model more accurate re-ranked images are obtained.

In Future Research, consideration of time complexity would be important point to focus. Also various algorithms can be merged with proposed solution to test and increase accuracy of image re-ranking. Proposed system can be applied on real world dataset to
test the accuracy of it.

VI. REFERENCES

[1]. Junjie Cai, Zheng-Jun Zha, Member, IEEE, Meng Wang, Shiliang Zhang, and Qi Tian, Senior Member, IEEE, "An Attribute-Assisted Reranking Model for Web Image Search", IEEE Transactions On Image Processing, Vol. 24, No. 1,January 2015.

[2]. R. Yan, A. Hauptmann, and R. Jin, “Multimedia search with pseudo relevance feedback”, in Proc. ACM Int. Conf. Image Video Retr., 2003, pp. 238247.
[3]. Meng Wang, Hao Li, Dacheng Tao, Ke Lu, and Xindong Wu ,Multimodal Graph-Based Reranking for Web Image Search, IEEE transactions On Image Processing,Vol. 21, No. 11, pp. 4649-4661,November 2012.
[4]. Suvarn V. Jadhav A. M. Bagade, Department of Information Technology, Pune Institute of Computer Technology, Pune,(India),"Comparative Analysis Of Reranking Techniques For Web Image Search", International Conference on Recent Innovations in Engineering and Management, Dhananjay Mahadik Group of Institutions (BIMAT) Kolhapur Maharashtra, 23rd March 2016.
[5]. Xiaogang Wang, Member, IEEE , Shi Qiu, Ke Liu and Xiaoou Tang, Fellow, IEEE, "Web Image Re-Ranking Using Query-Specific Semantic Signatures", IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 36, No. 4, April 2014.
[6]. Suresh Kumar Nagarajan, Shanmugam Saravanan ,Senior Assistant Professor, School of Computing Science and Engineering, VIT University, Vellore, India School of Computing Science and Engineering, VIT University, Vellore, India,” Content-based Medical Image Annotation and Retrieval using Perceptual Hashing Algorithm”, IOSR Journal of Engineering Apr. 2012, Vol. 2(4) pp: 814-818.
[7]. Hajer Fradi, Jean-Luc Dugelay, “A new multiclass svm algorithm and its application to crowd density Analysis using lbp features”, Eurecom Sophia Antipolis, France.