Survey on Methodologies Used for Web Image Search

Dipalee N. Gaikwad
Department of Computer Engineering
Dr. D.Y.Patil College of Engineering Ambi
Savitribai Phule Pune University

Sandeep U. Kadam
Department of Computer Engineering
Dr. D.Y.Patil College of Engineering Ambi
Savitribai Phule Pune University

ABSTRACT
The image search has becomes an important feature of multimedia. It plays an important role in daily life. Some image search query results are satisfactory and some are unsatisfactory. The web image search mostly depends on the surrounding text of the image. It is difficult to understand the user intention only by query keywords and this leads to irrelevant image search results. In this paper the methods developed by different researchers in the area of web image search are reviewed. These methods vary from textual information search to user feedback. Also some methods are depend on the visual similarities between the images. To improve the result of web image search, strategies like keyword expansion, active re-ranking is also used. This paper focuses on the methods introduced by different authors for development in the area of web image search.

General Terms
Image Search, User Intention.

Keywords
Co-click Activities, Keyword Expansion, Relevance Feedback, Query Difficulty Prediction, Re-ranking.

1. INTRODUCTION
The aim of the image search is to retrieve the relevant image with respect to user query from a large image database. So identifying the accurate image with user intention is the most challenging task. The traditional search engine works by taking query as input, if any image contains that query keyword in its surrounding text then that image is retrieved as result. If image contains irrelevant surrounding information and if that keyword is found into that information then this image is also displayed even it is not related with user intention. e.g. if user is searching for fruit "apple" then the result of the image search contains images of fruit apple, apple i-phone, apple laptop.

To improve the results of the search engine lots of work is going on. This paper focuses on different methods used in web image search. These methods include maintaining history of co-click activities [2], implicit guidance of user [9], keyword expansion [6], Query specific semantic signatures [1], Active Re-ranking [3], prototype based image search [4] and query difficult prediction [10].

These all methods are developed by different researchers to improve effectiveness and efficiency of the web image search. Each of these methods are described in following sections. To refine the web image search and predicting a good result is the aim of each of this method.

Traditional search engine takes a query as a input from the user. Keywords in the query are taken and if the keyword is matched with surrounding text of the image then that images are retrieved as a result. There are number of images retrieved if they are having same text into their surrounding information as a query keyword text. So the number of images are retrieved as a result. Fig. 1 shows the working of traditional web image search engine.

2. LITERATURE SURVEY
The web image search refers to finding the accurate images related to the user query. The development in such image retrieval area is still going on. Different Researchers are working on the methods used to improve the performance of web search engine. Finding co-click history is best suited if two images are co-clicked frequently then they are more similar than third image co-clicked less often[2]. Difficulty with only textual search is that if the external text of image is ambiguous or not related with the query then this becomes a limitation for textual search[9].Logs of search engine is used to refine the image search results [8]. To make a query more descriptive keyword expansion is used [6].
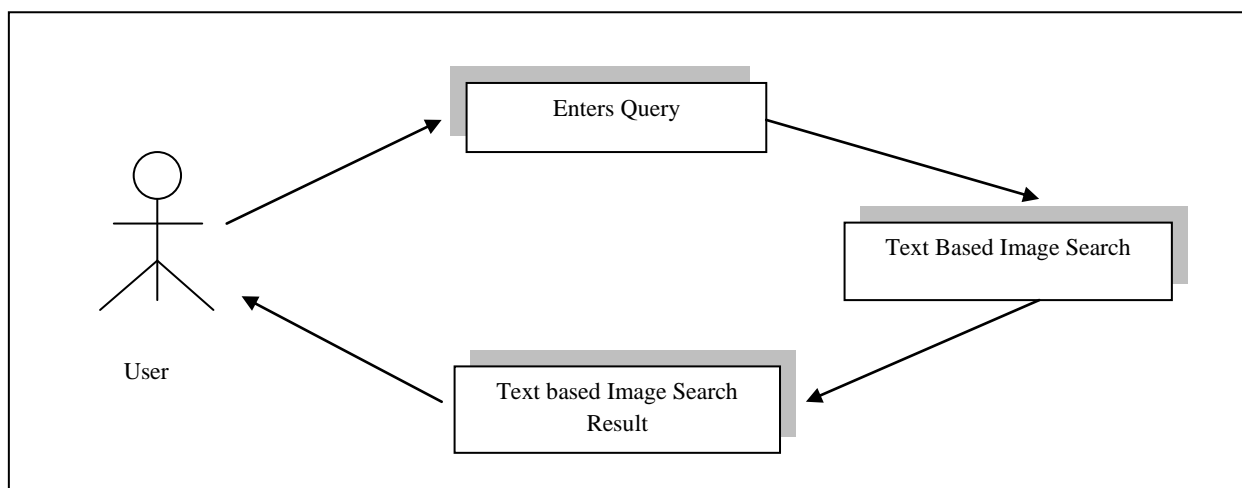

Fig. 1: User interacting with traditional web image search engine

By capturing user intension active re-ranking is performed [3] to produce best result. These are the different techniques developed in the area of web image search by different research persons described in following section.

3. METHODS USED

3.1 Using Textual Information [9]

As like a traditional method the image search is performed using text keywords in the query. If the keywords in the query appear in the surrounding text of the image then that image is retrieved as a resulting image. There is some work related with textual information is going on which uses text including filename of the image, URL of the image and description or the caption of the image as a surrounding text parameter. If the text is found in these surrounding parameters of the image then that image is displayed as a result image. The textual search is used as a input for visual similarity based search by grouping the images which is having same textual tag.

But some difficulty with textual search is that if the external text is ambiguous or not related with the image then this becomes a limitation for textual search. e.g. if user wants to search for the image "sky" and if there is one image of sky but having caption as a "blue" then this image will not be retrieved in the result even it contains the visual scene which user wants. If the surrounding text is ambiguous then also results generated may be poor.

3.2 Using Co-Click History [2]

When the user initiates an image search by typing the URL of a search engine a session is started and when user leaves or don't perform any activity long time the session is ended. By recording co-click images the decision is made whether two images are similar or not. When session starts with image search usually retrieve lots of images user may click one or more images. By recording click activities of users helps in image search. If image A and image B are both clicked by user then it is said that they are co-clicked. If two images are co-clicked frequently then they are more similar than third image co-clicked less often.

When session starts the users have some image in their mind what they are looking for. So user may compare retrieved image with the target image in their mind. The images similar to target images are only clicked and remaining is ignored. But sometimes user is interested in viewing the images generated by search query which is not related with the target image. In such cases also using co-click history is best suited because if two images are co-clicked frequently then they are more similar than third image co-clicked less often. Finding frequently co-clicked images the search is refined.

3.3 Relevance Feedback [8]

To refine an image search "logs" of search engine are used. These logs are used as a relevance feedback signals to refine image distance function. This approach is taken by Schultz [11], which is similar to soft margin SVM trained with relative comparison of the images. e.g. if Xi is more similar to Xj than Xi to Xk. Then only one method is learned and used to compare with all images in database in these strategies.

There are two ways [8] to increase the learning approaches. One is permit related text queries to share the learned distance function to improve query specific distance function and second is increase the ranking accuracy in some query groups more than others using query specific distance function which is able to which automatically picks up query group that is suitable and beneficial.

3.4 Keyword Expansion [6]

When user enters query for image search the resulting images are displayed by extracting surrounding information of the image. But if user query is not sufficient to express the intention of the user because of lack of knowledge about giving specific word for image search then the search performance also decreases.

To recover from this difficulty extra information related with the user query needs to be captured. Extra information is captured by using keyword expansion which makes query more descriptive. Existing methods of keyword expansion provides synonyms or related words which are frequently occurred with query word. e.g. Google image search provides related searches approach to find related images which are having query related keywords.

3.5 Query Specific Semantic Signatures [1]

Xiaogang Wang and Xiaoou Tang introduced a large scale benchmark database. Which includes 120000 images retrieved by the Bing image search using 120 keywords. Experiments on this database shows 25-40% improvement on re-ranking precision with 70 times speedup compared with state of the art methods. A user once find candidate image of target image the re-ranking function is used by choosing that candidate image as a query image. One of the approach described by xiao gang and xiaoou Tang [1] which has offline and online parts as follows.

3.5.1 The offline stage has following steps:

i. Set of related keywords with query keyword using keyword expansions are generated automatically. e.g. if query keyword is "apple" then expansion is "red apple", "apple i-watch".

ii. These expansions are used as a reference classes for query keyword. The training tuples are automatically obtained using keyword expansion. Obtained training examples of a reference class and to retrieve images by search engine based on textual information again.

iii. Images retrieved by keyword expansion are more correct than those retrieved by original keyword "apple".

iv. The lier images are removed automatically and top images are used as training examples of the reference class.

v. By calculating similarities between images and reference classes the semantic signature of an image is fetched using trained multiclass classifier.

vi. To extract one semantic signature for an image the features like shape, texture, color are grouped to train single classifier.

3.5.2 The online stage has following steps:

vii. With reference to query keyword the images are retrieved.

viii. The images retrieved are associated with query keyword reference to the word image index file. The word-image index file contains image associated with many query keywords.

ix. When user chooses a query, image semantic signatures are used for finding image similarities for re-ranking.

3.6 Active Re-ranking [3]

When the query term is unambiguous the method re-ranking is also fail to predict user search. To make prediction correctly about the user's query image intention the sample selection strategy is used to decrease the clients labelling efforts. The only text based search is however problematic. If user wants to search for "apple" and if any image contains name apple within its surrounding text then this image is also displayed even if it is irrelevant to the user's query image.

To display result of correct images to the user "active re-ranking" is used. In this method the user intention is captured and used for re-ranking the images. To improve the

performance of search, labelling information is collected from user and new method is proposed to actively select more informative query images through structural information. Few images are labelled by user in active re-ranking.

Author proposed active re-ranking [3] in which user first puts query e.g. "panda". Panda is an animal but there are some persons having name panda. The images having text panda into their surrounding are displayed. So the images retrieved are animals or persons. The user needs to label those images. After knowing the user intention by labeling information and discriminative submanifold the re-ranking process is conducted.
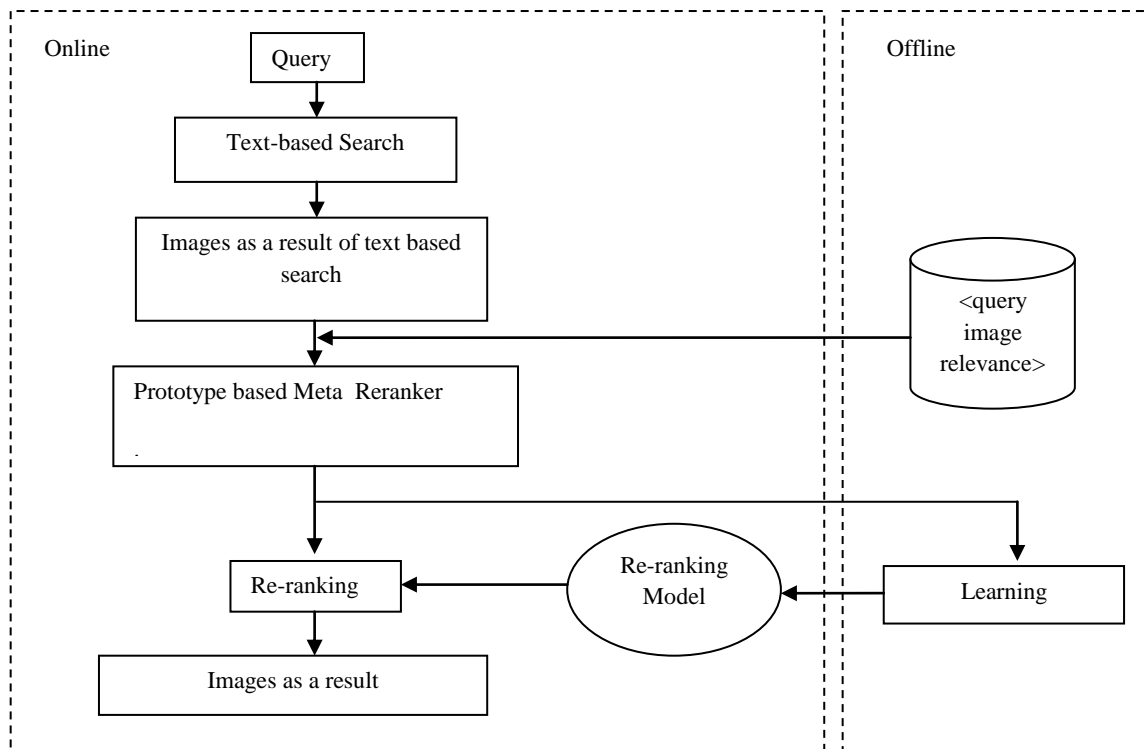

Fig. 2 Overview of the prototype based visual re-ranking framework proposed by Linjun Yang and Alan Hanjalic[4]

There are two methods through which the user intention is captured i.e. active sample selection strategy and the dimension reduction algorithms.

Prototype Based Image Search Re-ranking [4]

Strategies used for image search re-ranking can be divided into two types i.e. supervised and unsupervised. The unsupervised method is not depend upon user labeling. Author proposed prototype based re-ranking method. This method constructs meta re-rankers with respect to visual prototypes presenting the textual query and design the weights of linear model to group the result of individual meta-rankers and produce the re-ranking score of a given image taken from the initial text based search result.

The Framework developed by Linjun Yang and Alan Hanjalic [4] is shown in Fig. 2. Which has online and offline steps. When user enters a query initial search is performed using text based search. Then visual prototypes are generated and for every prototype meta reranker is built.

3.7 Query Difficulty Prediction [10]

The method query difficulty prediction [10] used to predict the quality of results generated by search without user's feedback and judgment. Query difficulty Prediction is easy in text search because query and document both are in textual format whereas in image search query is text and images are visual. This proposed method automatically predict the query difficulty for any query through machine learning approach.

Machine learning approach is designed by analyzing visual distribution characteristic of training queries then in learning process the relationship between features derived by author and query difficulty is mined. Based upon this the query difficulty model is built. This method is very effective and requires minimal search engine selection and image result merging.

Many QDP methods just predict the value for each query to show the difficulty but the method proposed by author is more successful than traditional method because the output of proposed model is the estimation of real performance rather than only indicator.

3.8 Joint Optimization [7]

For large scale images and indexing the bag of words (BOW) model has been known. In BOW search process depends on voting process. i.e. if image text describes the actual content of image. The voting score exactly reflect the image similarity. In multimedia search area the bag of word model becomes very popular because of its simplicity and effectiveness [12],[13],[14]. The technique used is to represent an image as a cluster of visual words irrespective of any order and organize all visual words into inverted table for efficient image search. Performance achieved in this scheme is outstanding , also it is tested on a very large image dataset to demonstrate scalability.

3.9 Confidentiality Preserving Image Search [5]

While performing image search the confidentiality of image is also taken into consideration. Also there are some methods used to preserve confidentiality during image search. For secure transmission the data is transmitted in encrypted format and it is decrypted at server side to operate on plain text. This makes user's private information vulnerable.

So the user's personal albums can be viewed by system administrator. By storing data in encrypted format at server side makes task difficult to server for handling client request and process encrypted data. There are technologies developed to work on encrypted databases that protect user's private data without any difficulties on server.

There are two types described i.e. homomorphic encryption and cryptography protocols and the second type is the randomization techniques for visual features and search indexes for preserving image search confidentiality.

These are the different methods used and developed with respect to web image search.

4. CONCLUSION

Web image search plays an important role in the area of multimedia. This paper focused on the methods developed by different researcher. The methods described are from oldest text based keyword search to the earlier prototype based image re-ranking. To remove unreliability and to increase the performance of web search engine these methods are applied. e.g. To refine query entered by user keyword expansion is used, for re-ranking the images actively user labeling is used. The development in the area of image search is still going on. But the methods described by different researches are very successful for faster and accurate results of the image search. The future work for image search is improved in the area of re-ranking using duplication detection. Using the combination of textual and visual features search is explored.

5. REFERENCES

[1] 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

[2] Yushi Jing, Michele Covell, David Tsai, and James M. Rehg, Member, IEEE, Learning Query-Specific Distance Functions for Large-Scale Web Image Search, IEEE Transactions On Multimedia, Vol. 15, No. 8, December 2013

[3] Xinmei Tian, Dacheng Tao, Member, IEEE, Xian-Sheng Hua, Member, IEEE, and Xiuqing Wu, Active Reranking For Web Image Search, IEEE Transactions On Image Processing, Vol. 19, No. 3, March 2010

[4] Linjun Yang, Member, IEEE, and Alan Hanjalic, Senior Member, IEEE, Prototype-Based Image Search Reranking , IEEE Transactions On Multimedia, Vol. 14, No. 3, June 2012 871

[5] WENJUN LU1, AVINASH L. VARNA2, (Member, IEEE), AND MIN WU3, (Fellow, IEEE), Confidentiality- Preserving Image Search: A Comparative Study Between Homomorphic Encryption and Distance-Preserving Randomization, Date of publication February 20, 2014, Digital Object Identifier 10.1109/ACCESS.2014.2307057

[6] Xiaoou Tang, Fellow, IEEE, Ke Liu, Jingyu Cui, Student Member, IEEE, Fang Wen, Member, IEEE, and xiaogang Wang, Member, IEEE, IntentSearch: Capturing User Intention for One-Click Internet Image Search , IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 34, NO. 7, JULY 2012

[7] Shikui Wei, Dong Xu, Xuelong Li, Fellow, IEEE, and Yao Zhao, Senior Member, IEEE, "Joint Optimization. Toward Effective and Efficient Image Search" IEEE Transactions On Cybernetics, Vol. 43, No. 6 December 2013.

[8] Yongdong Zhang, Senior Member, IEEE, Xiaopeng Yang, and Tao Mei, Senior Member, IEEE, Image Search. Re-ranking with Query-dependent Click-based Relevance Feedback, IEEE TRANSACTIONS ON IMAGE PROCESSING, 1057-7149 (c) 2013 IEEE

[9] Zheng Lu, Xiaokang Yang, Senior Member, IEEE, Weiyao Lin, Hongyuan Zha, and Xiaolin Chen, Inferring User. Image-Search Goals Under theImplicit Guidance of Users, IEEE Transactions On Circuits And Systems For Video Technology, Vol. 24, NO. 3, MARCH 2014.

[10] Xinmei Tian, Yijuan Lu, Member, IEEE, and Linjun Yang, Member, IEEE , Query Difficulty Prediction for Web Image Search, Transactions On Multimedia,Vol. 14,No. 4, 2012 951

[11] M. Schultz and T. Joachims, "Learning a distance metric from relative comparisons," in Proc. 16th Conf. Advances. in Neural Inf. Process. Syst. (NIPS).

[12] J. Sivic and A. Zisserman, "Video google: A text retrieval approach to object matching in videos," in Proc. IEEE Int. Conf. Comput. Vision, vol. 2. Oct. 2003, pp. 1470–1477.

[13] S. Wei, Y. Zhao, C. Zhu, C. Xu, and Z. Zhu, "Frame fusion for video copy detection," IEEE Trans. Circuits Syst. Video Technol., vol. 21, no. 1,pp. 15–28, Jan. 2011

[14] E. Nowak, F. Jurie, and B. Triggs, "Sampling strategies for bag-offeatures image classification," in Proc. Eur Conf. Comput. Vision, vol.3954.2006, pp.490–5