Enhancing efficency of image retrieval using reranking methods

Nishant N. Chavan
Computer Engineering Department,
SCOE, Kharghar
Mumbai, Maharashtra,India
e-mail: nishant.chavann@gmail.com

Deepa Parasar
Computer Engineering Department,
SCOE, Kharghar,
Mumbai, Maharashtra,India
e-mail: deepaparasar@gmail.com

Abstract—Fast development in web attracted user to mine data also provide opportunity of research. Search engine and reranking methods are evolved as most reliable platform to obtain information relevant to user needs. Visual search reranking is performed on visual information. To remove ambiguity caused by text based approach various reranking methods like pairwise learning, visual saliency and consistency are used also reranking methods are improved by mapping user intension. To capture user intention methods like active reranking and intent search has been surveyed. Pairwise learning has been formed to produce optimal documents pair. It deals with ambiguity problem with EP reranking and DP reranking. Visual saliency and consistency technique has been focuses on two methods- visual saliency which is generally used to capture visually salient features that is relevant to user query image and visual consistency which is to select most frequently occurred images. Due to ambiguity in query, image reranking has been become inefficient. Active reranking is method used to capture intention. This method involves approaches like active sample selection strategy and dimension reduction algorithm. Another method contributed for capturing user intention is IntentSearch. It consists of four main steps- adaptive similarity, keyword expansion, image pool expansion and visual query expansion. First two methods are efficiently producing reranked images. Sometimes image search reranking doesn't prove fruitful to capture user intention so aforementioned two methods have been focused to enhance performance.

Keywords- *visual search; active reranking; intent search; visual saleincy ;pair wise reranking;*

*****

I. INTRODUCTION

Due to explosive growth in web 2.0, large volume of images being transmitted and searched on Internet. Image searching on the search engine have got increasing importance in our daily life. The tremendous development in Web technology has been led with immense scope of research activities in visual search. Thus, Exponential growth of social media and online multimedia data such as text, image, audio and video retrieval have been become an active and challenging area for research.

Although image search has become famous feature in many search engine including Google, Bing, Yahoo!, etc., the majority of image search use very little image information. Most traditional methodologies for retrieving objects use keyword as query. In this user mention their keyword in search engine then search engine make use of spider program that crawl through entire database to populate thousands of related images ranked based on textual information associated with it. Textual queries have been used to obtain user expected photos. This approach is referred as "text based information retrieval" (TBIR).Text based search was followed by most of commercial search engines because of easy understanding and processing of text. However above mentioned approach suffers from drawback that textual information or metadata associated with images is often mismatched with their actual content, which in turn producing result that is not as per expectation of user. To solve this problem generally image annotation is used to categorize image with respect to set of high level semantic concepts [1].Results generated by text based search was ambiguous. This has been led development of notion so called content based information retrieval (CBIR) that is more prominent. Content based image searching has attracted attentions of many multimedia and computer vision community researchers. In content based information retrieval, in order to retrieve expected result sample image as query has been provided by user to search engine. However major limitation in CBIR is semantic gap between low level visual features and the high-level semantic concepts. To bridge the semantic gap, methods developed to map user intentions [1].

A common practice to improve search performance is to rerank(refine) the visual documents returned from a search engine using a larger and richer set of parameters. Search reranking has become general way to boost information retrieval. Image reranking has been opted by many today's commercial search engine. In image reranking technique, search engines retrieves pool of images based on textual information after providing query as keyword. Further images are reranked based on their visual similarity. Many traditional methods have used textual information associated with image rather than visual attributes to rerank pool of images retrieved. Those methods suffer from many limitations such as ambiguous results. So, nowadays these techniques are discontinued by many commercial search engines and adapting to visual search reranking methods. Here four techniques have been surveyed that are used for reranking. The techniques are reranking using pairwise learning[2], reranking using integrating visual saliency and consistency[3],active reranking[4], Intent Search to map user

intention[5]. This paper surveys & discusses above mentioned techniques in detail in further section.

II. VISUAL SEARCH USING PAIR WISE RERANKING :

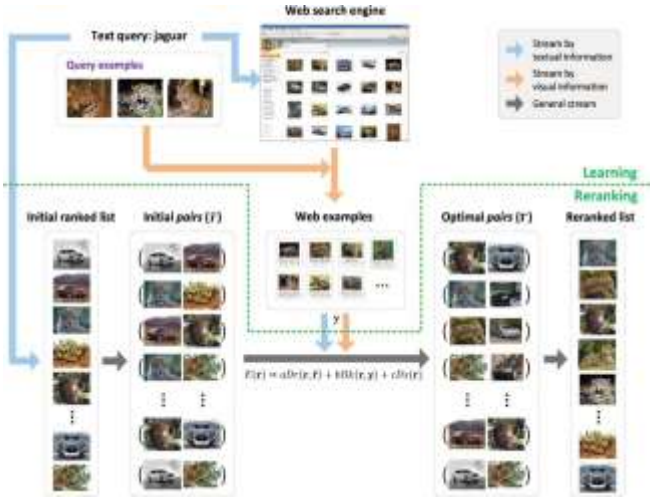

Figure 1: Flowchart of pair wise reranking [2]

On basis of external knowledge used[6], research on visual search has proceeded along three dimensions i.e. self-reranking[7] that does not make use of external knowledge, example-reranking that is based on user given query example and crowd-reranking which exploits online crowdsourcing knowledge[6].

From initial ranked list for reranking has been seeking consent from self-reranking. Fergus et al. presented probabilistic latent semantic analysis (pLSA) in which visual clustering has been executed on initial returned images and visual object categories have been discovered; finally reranking of images have been done on basis of distance from discovered category. Recurrent patterns have been found in Self-reranking, also refer as unsupervised-reranking [8].Whereas in second dimension i.e. example-reranking is referred as supervised re-ranking [8] in which accuracy of search has been enhanced by supplying some suggestions. Also reranking models have been prepared using query examples. Liu et al. focuses on notions like relevancy and irrelevancy between query examples for given query, later optimal set of documents pair has been discovered by means of an information theory. Ultimate reranking has been done by optimal pair set [8]. In third dimension i.e. crowd-reranking, crowdsourcing knowledge has been utilized obtained from public social network [6].

Many researchers have given plenty of emphasis on "query concept mapping". Kennedy et al. focuses on mutual information measurement for ranking [9]. While Wang et al. combines the text and visual information extracted from the text query and visual examples, respectively [10].In reality, length of text query are short and also suffers from ambiguity.

Pair-wise learn suggest that search engine has been supplied with text query that in turn returns visual documents along with associated text. At same time, web results with similarity are filtered using query example. Query example and online crowd-sourcing knowledge has been utilized to obtain more strong relatedness. Based on this optimal pair set has been found by theoretically formulating visual reranking as optimization problem. While formulating ranking distance, knowledge distance and smooth distance have been taken into account. Finally, round robin ranking [11] also refers as voting is used. In round robin, first document of each pair is given with ordinal score; while second document of each pair is given 0.All score given to same documents are added. Documents are reranked in descending order based on sum of score. Finally reranked list is retrieved.

III. RERANKING BASED ON VISUAL SALIENCY AND CONSISTENCY :

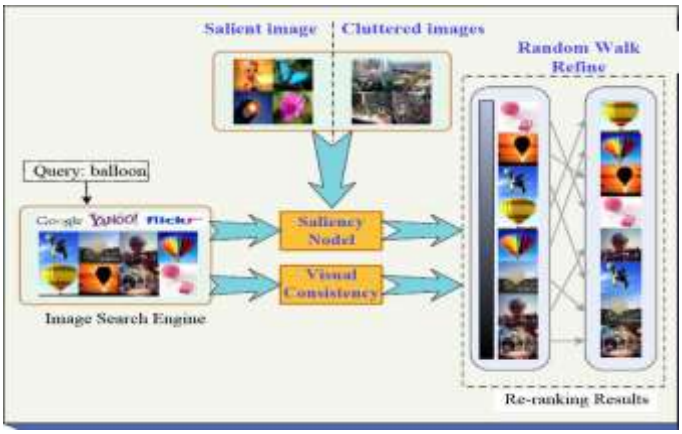

Figure 2: reranking based on visual saliency and consistency [3]

In framework of visual saliency and consistency based reranking, query is submitted to search engine then results for submitted query has been returned by search engine which are found to be disappointing for user. This reranking method refines resulting images such that appropriate images are displayed at the beginning and inappropriate images are displayed at end.

Results returned have been scrutinized by saliency model also referred as visual attention model. Relevance score is given to images depending on their saliencies. It has been observed that thumbnails are identified as most rigorously used for salient object, distinctive regions while user searching for images. Furthermore, these visually salient images in the front pages are often more relevant to the user's query. Therefore, these images should have high ranks in the refined result [3]. Concept of visual consistency has been proposed that calculates and elaborate on similarities among images. It enables filtration of inappropriate images by focusing more on frequently occurred images and discarding less frequently

occurred images. Final image retrieval is done by random walk based method which integrates two mechanisms.

IV. ACTIVE RERANKING :

Primary problem of capturing user intention has been focused in this method. Visual information is inadequate to map user's intention. In traditional methods, one fix label has been assign to each sample so an image significant for one user can become inappropriate to other user. The semantic space is user-driven, according to their different intentions but with identical query keywords. Secondly, user has been asked to label lots of images. This method has been targeting the user-driven intention from two aspects: collecting labeling information from users to obtain the specified semantic space, and localizing the visual characteristics of the user's intention in this specific semantic space [4].

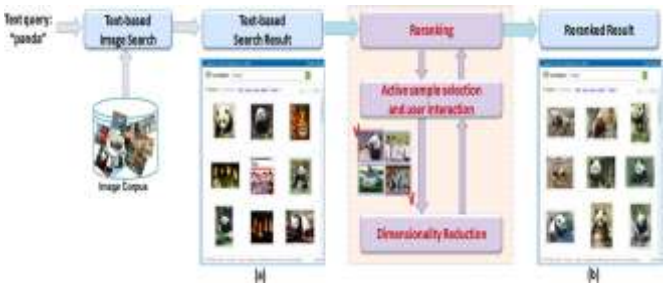

Figure 3: Framework for active reranking illustrated with the query "panda". When the query is submitted, the text-based image search engine returns a coarse result (a). Then the active reranking process is adopted to obtain a more satisfactory result (b), by learning the user's intention [4].

In active reranking, Structural information (SInfo) has been projected to gather labeling information from users. This method dynamically selects most useful query images. SInfo sample selection strategy covers two aspects: ambiguity and representiveness simultaneously. Ambiguity denotes uncertainty whether an image is relevant or not the user's intention. While representiveness is another part. Labeling a representative sample is more useful than labeling isolated one [4]. Local global discriminative (LGD) dimension reduction algorithm has been developed to restrict visual characteristics of user's intention. For reranking pool of images Bayesian reranking [12] has been used.

V. INTENT SEARCH

Visual content of target image cannot be easily depicted using keywords. This ambiguity is solved using capturing user search intention. It is also important to add visual information to image search. To make user interaction simpler bare minimum one click is used. To capture user intention for accurate image retrieval four step has been proposed.

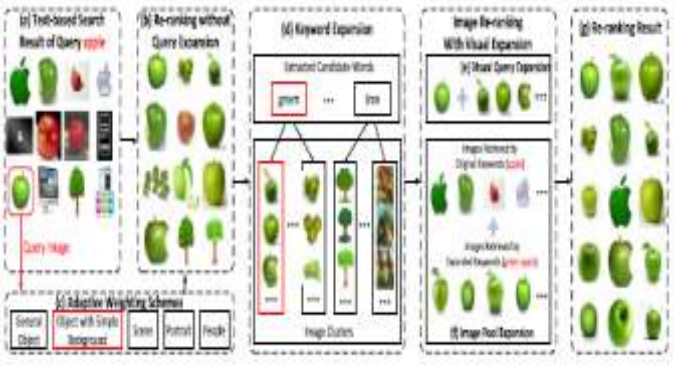

Figure 4: system architecture of Intent Search [5]

Initially, user tender text based keyword as query. A pool of images is retrieved using text based search. User is asked to choose query image from image pool. Here adaptive similarity is focused where image has been classified according to one of predefined adaptive weight category. For better reranking of text based search result, a specific pretrained weight schema is used to combine visual features adapting to this kind of images [5]. Initial result of this is not satisfactory so in following step keyword expansion has been done. Important keywords can be missed due to short user input and less knowledge of textual description of target image. To overcome this issue, with help of query image keyword has been expanded. Keyword expansion is novel approach has been introduced here that helps in better way to capture user intention using textual depiction and visual content. Thesaurus based methods [13][14] expanded query keyword with their linguistically related words such as, synonyms and hypernyms. Clustering [15] based approach is also suggested. Some methods [16][17] have been suggested tag suggestions or annotations depends upon visual content for input images. These methods have been designed as options to expand keyword. Method used here has been provided extra image clusters which is utilized for visual expansion. The purpose of visual query expansion is to obtain multiple positive example images which are more robust and more specific to query image so, one class SVM [18] has been adopted to refine visual similarity. All mentioned steps are automated with only one click [5].

VI. CONCLUSION AND FUTURE SCOPE

To summarize, four methods of visual search reranking has been surveyed. Image from search engine like Google, Bing is retrieved using concepts like TBIR AND CBIR but there after to enhance result re-ranking image is firstly introduced. These methods are derived to overcome problems of older method as well as to enhance performance by retrieving accurate result which is expected by user. First method is novel optimization based technique focuses ambiguity problem. This method generates optimal pair set by considering optimization problem. DP-reranking and EP-reranking are explored extensively to achieve relevant relation of each document pair.

Round robin is finally used to obtain rerank list. Second method is based on visual saliency and consistency that suggests integration of visual saliency and visual consistency for visual search re-ranking to improve user experience. Here only non-ambiguous notions are tested also perceptual visual quality parameters are not taken into consideration for re-ranking. Third method is active reranking which reduces effort of user by making user to label few images. This approach is vey proficient to map user intention which is further useful in reranking .Sample selection strategy and dimension reduction algorithm is very effective as they deliver performance over real web image search. Last method is intent search which shows significant improvement. User is contributing with one click which is very less efforts as compared to traditional methods. One deficiency has been observed that retrieved images may contain duplicate images.

In future scope, many researchers have been classified visual search reranking into two broad categories: visual pattern mining and multi modality fusion. In existing systems, initial stage of retrieval has been focusing more on single attribute of visual document.

To increase system performance, multiple parameters of visual document can be taken into account that will initially rerank images based on more than one visual attribute. Initial refinement in visual search will result in this process. Thereafter, single click on refined result will be made to capture expected result for user. A novel hybrid approach will receive advantage of both systems.

REFERENCES

[1] Yiming Liu, Dong Xu,Ivor Wai-Hung Tsang, and Jiebo Luo Textual Query of Personal Photos Facilitated by Large-Scale Web Data" IEEE PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 33, NO. 5, MAY 2011.

[2] Yuan Liu and Tao Mei, "Optimizing Visual Search Reranking via Pairwise Learning" IEEE MULTIMEDIA, VOL. 13, NO. 2, APRIL 2011, pp. 280-291.

[3] Jun Huang, Xiaokang Yang, Xiangzhong Fang, Weiyao Lin, and Rui Zhang, "Integrating Visual Saliency and Consistency for Re-Ranking Image Search Results", IEEE MULTIMEDIA, VOL. 13, NO. 4, AUGUST 2011,pp. 653-661

[4] Xinmei Tian, Dacheng Tao, Xian-Sheng Hua, Xiuqing Wu, "Active Reranking for Web Image Search" IEEE IMAGE PROCESSING, VOL. 19, NO. 3, MARCH 2010, pp.805-820

[5] Xiaoou Tang, Liu, Jingyu Cui, Fang Wen, Xiaogang Wang, "IntentSearch: Capturing User Intention for One-Click Internet Image Search", IEEE PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 34, NO. 7, JULY 2012, pp. 1342-1353.

[6] Y. Liu, T. Mei, and X.-S. Hua, "Crowdreranking: Exploring multiple search engines for visual search reranking," in *Proc. ACM Special Interest Group on Information Retrieval*, 2009, pp. 500–507.

[7] R. Fergus, L. Fei-Fei, P. Perona, and A. Zisserman, "Learning object categories from google's image search," in *Proc. IEEE Int. Conf.Comput. Vis.*, Oct. 2005, pp. 1816–1823.

[8] Y. Liu, T. Mei, X. Wu, and X.-S. Hua, "Optimizing video search reranking via minimum incremental information loss," in *Proc. ACM Int. Workshop Multimedia Information Retrieval*, 2008, pp. 253–259.

[9] L. Kennedy and S.-F. Chang, "A reranking approach for context-based concept fusion in video indexing and retrieval," in *Proc. ACMInt. Conf.Image and Video Retrieval*, 2007

[10] D. Wang, X. Li, J. Li, and B. Zhang, "The importance of query-concept- mapping for automatic video retrieval," in *Proc. ACMMultimedia*, 2007.

[11] Y. Liu, T. Mei, J. Tang, X.Wu, and X.-S. Hua, "Graph-based pairwise learning to rank for video search," in *Proc. Int. Multi-Media Modelling Conf.*, 2009.

[12] X. Tian, L. Yang, J.Wang, Y. Yang, X. Wu, and X.-S. Hua, "Bayesian video search reranking," in *Proc. ACM Int. Conf. Multimedia*, 2008, pp. 131–140.

[13] S. Liu, F. Liu, C. Yu, and W. Meng, "An Effective Approach to Document Retrieval via Utilizing WordNet and Recognizing Phrases," Proc. 27th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval, 2004.

[14] S. Kim, H. Seo, and H. Rim, "Information Retrieval Using Word Senses: Root Sense Tagging Approach," Proc. 27th Ann. Int'l ACM /SIGIR Conf. Research and Development in Information Retrieval, 2004.

[15] K. Sparck Jones, Automatic Keyword Classification for Information Retrieval. Archon Books, 1971.

[16] L. Wu, L. Yang, N. Yu, and X. Hua, "Learning to Tag," Proc. Int'l Conf. World Wide Web, 2009.

[17] C. Wang, F. Jing, L. Zhang, and H. Zhang, "Scalable Search-Based Image Annotation of Personal Images," Proc. Eighth ACM Int'l Workshop Multimedia Information Retrieval, 2006.

[18] Y. Chen, X. Zhou, and T. Huang, "One-Class SVM for Learning in Image Retrieval," Proc. IEEE Int'l Conf. Image Processing, 2001.