A Survey of Document Search Using Images

Pradnya More [1], Prof. Pradeep Patil [2]

[1] *Department of Computer Engineering, VPCOE, Baramati*
[2] *Department of Computer Engineering (PG), VPCOE, Baramati*

Abstract — *Web search engine is software system that is designed to search for information on WWW. When user gives a query into search engine, it provides to the user, a list of best matching web pages according to its criteria, generally with a short summary consisting the documents title and part of the text. A picture convey information more efficiently and effectively comparative to words. Traditional text based search engine is based on textual similarity. Web search engines do not use the pictures in the web pages to find relevant documents for a given user query. Instead of images, they operate by computing a measure of consistency between only the text portion of each page and the keywords provided by the user. This paper presents survey of various techniques which are used for re-ranking. It introduces the information about existing approaches towards the re-ranking techniques.*

Keywords- *document ranking; image search; search engines; web pages*

I. INTRODUCTION

Web has turned into the largest information repository over the world. Due to its huge volume, users easily feel lost and difficult to achieve what they need from the repository. In consequence, research on Web search has attracted more and more attention. Usually, each Web page is considered as a text document and Web is treated as a huge document collection. Thus, techniques from conventional text retrieval, such as Vector Space Model (VSM) [1] and TFIDF [2], were applied to search for documents with text contents similar to the query. Later, considering that Web is a hyperlinked environment and the links between Web pages encode the importance of the pages to some extent, Link information is used in many approaches, such as PAGERANK and HITS [3], as a helpful evidence to make the pages which are relevant to the query stand out. It is interesting that in addition to link and text, other modalities of information such as image, video and audio, also tell about rich and important information about contents of Web pages. In fact, Web is a multimedia environment. In the traditional Web search, the search process can be described as follows: first, a search query is produced by a user; second, the search result consisting of text snippets of the retrieved Web pages is returned to the user for browsing; third, if the search result is satisfactory to the user then the search is completed; if not, which is the general case, the user may reformulate the query through labeling relevant/irrelevant pages or directly revaluate the query terms manually. "A picture is worth a thousand words". Despite this old saying, current web search engines ignore the pictures in web pages and retrieve documents merely by comparing the query keywords with the text in the documents [4]. Of course this text involve the words in image captions and markup tags, but does not look at the pixels themselves.

A picture convey information more efficiently and effectively comparative to words. Traditional text based search engine is based on textual similarity. Web search engines do not use the pictures in the web pages to find relevant documents for a given user query. Instead of images, they operate by computing a measure of agreement between the keywords given by the user and only the text portion of each page. If a web page is determined to be relevant to a query in multiple information modalities simultaneously, then the page has high chance to be really relevant than other page which is judged to be relevant in only the text mode. For example, when a user searches for the animal tiger, a web page consisting a tiger photo as well as the word "tiger" would have a more chances to satisfy the query than a page including the word "tiger" without image of a tiger. In other words, related images in web pages are treated as additional evidence in judging the relevance of those pages in the search process. The relevance of web pages in the search process is determined by using related images in the web pages.

Web search engine is software system that is designed to search for information on WWW. Search engine is a program that searches for items in a database which is related to keywords or characters specified by the user, used for finding particular sites on the World Wide Web. Search engines first create an index of all the web documents and store it on the server. After the submission of a query by the user, the query is provided to the index, then it returns the documents containing the words in the query. Then, the returned documents are sent to a ranking function which allow a rank to each document and the top-k documents are returned to the user. Fig.1 shows the working of a typical search engine.
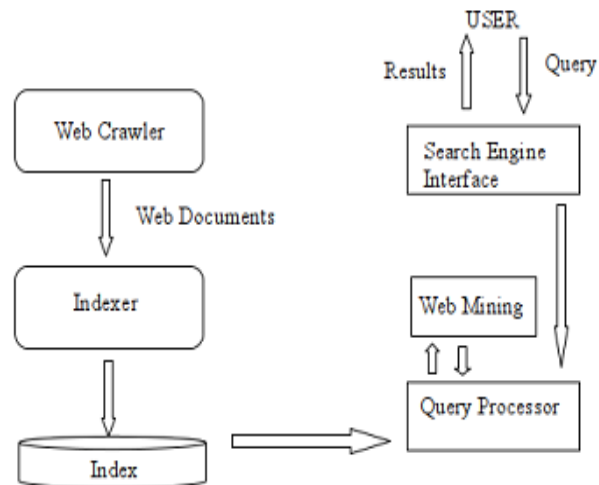
Figure 1.Working of search engine

II. RELATED WORK

Zhou and Dai [5], who were the first to show that content extracted from the pictures of Internet pages can be utilized to improve Web document search. Their system relies on an unsupervised method to discover a visual representation of the query from the images of Web pages retrieved via text search. The visual model of the query is computed through an iterative technique for density estimation aimed at finding the region of the visual feature space that contains the highest concentration of image examples associated to the query. These image examples are then averaged to form a single prototypical representation of the query. Then, an image-based rank of candidate Web pages is computed by calculating the distance between the pictures in the page and the visual prototype of the query. This image-based rank is fused with a conventional keyword-based rank to form the final sorted list of documents. To summarize the state of the art the recent paper of Schroff et al. [6] serves as an adequate exemplar. This work unites text, metadata and visual features in order to achieve a completely automatic ranking of the images related to a given query. Their approach begins from Web pages, recovered by text search for the query. Then images in the pages are re-ranked using both text and metadata features, and finally a form of pseudo-relevance feedback (PRF) is used: a classifier is trained to predict high rankings, and re-rank the image list. J. Krapac et al. [7] proposed an image re-ranking technique, based on textual and visual features, that does not require learning a separate model for every query. Features for each image are query dependent. The system proposed by Yeh et al. [8] is another example of multimedia search. However, their method needs additional user input, in the form of an image accompanying the text query. Barnard and Johnson [9] address the problem of word sense disambiguation in the context of words in image captions. For word sense disambiguation approach combining image information and text is more useful. Traditional approach is as shown in Figure 2.
Xue et al. [10] proposed to present image snippets along with text snippets to the user such that it is more easier and more accurate for the user to identify the Web pages he or she expects and to reformulate the initial query. An image snippet of a Web page is one of the images in the page that is both representative of the theme of the page and at the same time closely related to the query. Woodruff et al. [11] proposed to use an enhanced thumbnail to help the user quickly find the Web pages he or she expects. This improved thumbnail is an image, where the whole Web page is resized to fit into this image and important text is highlighted in order to be readable for the user.
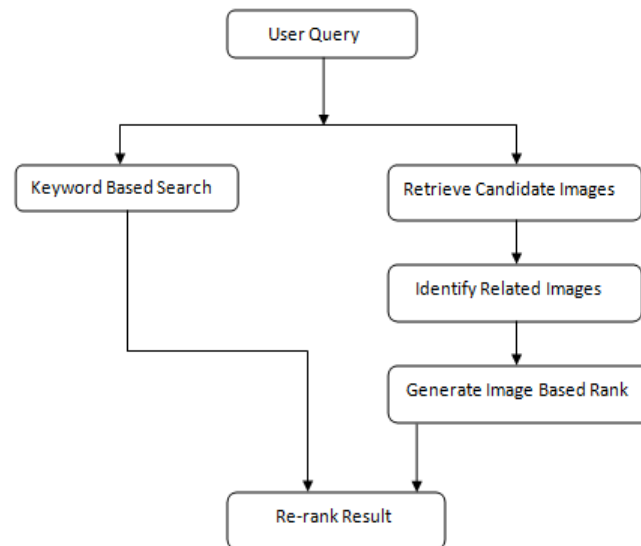
Figure 2. Traditional approach

III. RE-RANKING STRATEGIES

Ranking is sorting the documents according to some criterion. So that the best results appear first in the result list displayed to the user. Ranking criteria is in terms of relevance of documents with respect to an information need expressed in the query. Re-ranking is nothing but rank again for better result. There are different re-ranking techniques. The techniques are as follows:

1. Supervised Re-ranking
2. Visual Re-ranking
3. Bayesian visual re-ranking
4. Multimedia Search with Pseudo-relevance Feedback

L. Yang et al. [12] proposed a supervised re-ranking for web image search. In this method the learning-to-re rank model is used, which is implemented using the ranking SVM algorithm and 11 lightweight re-ranking. Due to unified re-ranking model is learned that can be applied to all queries, supervised learning is introduced. This approach does not go through scalability problems. In other words, a query-independent re-ranking model will be learned for all queries using query-dependent re-ranking features. The query-dependent re-ranking feature extraction is difficult because the visual documents have different representation.

In [13] Y. Jing et al. presented Visual Rank algorithm, a simple method take in the advances made in utilizing link and network analysis for Web document search into image search. As user created hyperlinks can be replaced with visual hyperlinks, so visual rank deviates from crucial source of information which makes Page Rank successful: the large number of manually created links on a diverse set of pages. Visual hyperlinks are automatically inferred. There are two techniques through which human coded information is re-captured. First, by making Visual Rank query dependent (by selecting the initial set of images from search engine answers), human knowledge, in terms of linking relevant images to Web pages, is directly introduced into the system. Second, common features between two images are captured to generate image similarity graph. The images capture the common themes such selected images which possess higher relevancy.

In [14] X. Tian.et al. presented Bayesian visual re-ranking which uses the visual and textual information from the probabilistic viewpoint and makes visual re-ranking more effective. In this framework to solve the problems existing in pair wise regularizers and point-wise ranking distance, a local learning-based visual consistency regularizer and a pair-wise ranking distance are proposed. The Bayesian framework solves the problem, i.e., maximizing the ranking score

consistency among visually similar samples while minimizing the ranking distance, which represents the dissimilarity between the objective ranking list and the initial text based.

R.Yan et al. [16] proposed a Multimedia Search with Pseudo-relevance Feedback. In this method an algorithm is used which is mainly focused on video retrieval that fuses the decisions of multiple retrieval agents in both text and image techniques. This emphasizes the successful use of negative pseudo-relevance feedback to get better image retrieval performance. Although it does not work out all issues in video information retrieval, the results are hopeful, indicating that pseudo-relevance feedback shows great promise for multimedia retrieval with very diverse results.

IV. SUMMERY TABLE OF RE- RANKING STRATEGIES

Table 1. Re-ranking strategies

Sr. No	Author	Methods	Advantage	Disadvantage
1	L.Yang & Alan Hanj-alic [12]	Supervised Re-ranking for Web Image Search	The re-ranking performance of individual query is related to the characteristics of the initial text-based search result. The queries for which the relevant images in the initial result will benefit more from re-ranking.	The performance is even degraded in some cases. Since the result diversity is also an important objective so that more informative search result can be provided to users. It does not consider the diversity of results.
2	Y. Jing and S. Baluja [13]	Visual Rank	The Visual Rank algorithm presents a simple technique to incorporate the advances made in using link and network analysis for Web document search into image search	The relationship between image similarity and likelihood for transition is not fully explored. No linkages between images that are labelled and those that need labels.
3	X. Tian et al. [14]	Bayesian Visual Re-ranking	It examines the re-ranking problem from the probabilistic perspective and derives an optimal re-ranking function based on Bayesian analysis.	There are some limitations of this method. The first one is to speed-up the computation of hinge re-ranking via working set selection approach. To better represent the visual consistency, the semantic similarity which can be learned using distance metric learning will be incorporated into the re-ranking objective function.
4	Rong Yan et al. [15]	Multimedia Search with Pseudo-relevance Feedback	Pseudo relevance Feedback is an interactive technique for query reformulation. It has been shown successfully used in Content Based Image Retrieval.	It Need user's involvement to provide relevance information. This imposes a burden to users. Most of RF algorithms focus only on positive examples.

V. CONCLUSION

This survey presents various methods used for re-ranking of web images. In this each method is significantly efficient in ranking of images. This paper shows the merits and demerits of each method. The efficiency of the surveyed method can be measured in terms of computational time and retrieval accuracy. The benefits of each method can be taken into account and further these techniques can be enhanced for large scale web image re-ranking mechanism efficiently.

REFERENCES

[1] V.V. Raghavan and S.K.M. Wong, "A critical analysis of vector space model for information retrieval," Journal of the American Society for Information Science, 37(5):279–287, 1986.

[2] G. Salton and C. Buckley,"Term-weighting approaches in automatic text retrieval," Information Processing and Management, 24(5):513–523, 1988.

[3] J. M. Kleinberg," Authoritative sources in a hyperlinked environment," Journal of the ACM , 46(5):604–632,1999.

[4] J.S. Sergio Rodriguez-Vaamonde, Lorenzo Torresan, Andrew W. Fitzgibbon "What Can Pictures Tell Us About Web Pages? Improving Document Search Using Images," IEEE transaction on Pattern Analysis And Machine Intelligence, Vol. 37, No. 6, June 2015.

[5] Z.-H. Zhou and H.-B. Dai, "Exploiting image contents in web search." In Proc. Int. J. Conf. Artif. Intell. , pp. 2922–2927, 2007.

[6] F.Schroff, A.Criminisi, and A.Zisserman, "Harvesting image databases from the web," IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 4, pp. 754–766, Apr. 2011.

[7] J. Krapac, M. Allan, J. J. Verbeek and F. Jurie, "Improving web image search results using query-relative classifiers," in Proc. IEEE Conf. Comput. Vis. Pattern Recognition., pp. 1094–1101, 2010.

[8] T.Yeh, J. J. Lee, and T.Darrell, "Photo-based question answering," in Proc. 16th ACM Int. Conf. Multimedia , ser. MM '08, New York, NY, USA: ACM, pp. 389–398, 2008.

[9] K.Barnard and M.Johnson, "Word sense disambiguation with pictures," Artif. Intell., vol. 167, no. 1, pp. 13–30, 2005.

[10] X.-B. Xue, Z.-H. Zhou, and Z. Zhang," Improve Web search using image snippets," In Proceeding of the 21st National Conference on Artificial Intelligence ,pages 1431–1436, Boston, WA, 2006.

[11] A. Woodruff, A. Faulring, R. Rosenholtz, J. orrison, and P. Pirolli, "Using thumbnails to search the Web," In Proceeding of the SIGCHI Conference on Human Factors in Computing Systems, pages 198–205,Seatle, WA, 2001.

[12] L.Yang and A.Hanjalic, "Supervised Re-Ranking for Web Image Search," Proc. ACM Int'l Conf. Multimedia, 2010.

[13] Y. Jing and S. Baluja, "Visual Rank: Applying Page Rank to Large-Scale Image Search," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 30.no. 11, pp. 1877-1890, Nov. 2008.

[14] X. Tian, L.Yang, J. Wang, X. Wu, and X. Hua, "Bayesian Visual Re-ranking," IEEE Trans. Multimedia, vol. 13, no. 4, pp. 639-652, Aug. 2011.

[15] R. Yan, A. Hauptmann, and R. Jin, "Multimedia search with pseudo relevance feedback, Image Video Retrieval", pp. 649-654, 2003.