ISSN:0975-9646

Shekhar Maity et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 7 (3) , 2016, 1058-1060

Web Image Re-Ranking using Query-Specific Semantic Signatures

Shekhar Maity, Naveen Yerawar, Shivratan Pardeshi, Monalisa Singh

Department of Computer Engineering, Savitribai Phule Pune University
G.H.R.I.E.T., Wagholi, Pune, India.

***Abstract*- Image re-ranking is, the easiest way to improve the performance of web-based image search and it is used by the various search engine .Nowadays Google Image search, Microsoft live image search rely almost purely on surrounding text features. In the query based proposed system, keyword expansions help provide better results whereas in image counsel, re-ranking based on a priority of images accessed by other users provides more precise results. Here images are re-ranked by comparing their semantic signatures obtained from the visual semantic space specified by the query keyword.**

Image re-ranking mainly helps in reducing the time required for getting the desired result by providing the specific results related to the provided query.

***Keywords*--Image search, visual semantics, adaptive similarity, image re-ranking**.

I. INTRODUCTION

Today's commercial Internet image search engines use only for text information. Users type keywords in the hope of finding a certain type of images. The search engine returns thousands of images ranked by the text keywords. However, the returned Image results contain noise, disorganized, or irrelevant. Images are re-ranked based on the visual similarities. However, for web-scale commercial systems, the user input has to be limited to a minimum without online training.

Online image re-ranking which limits users' effort to just one-click feedback, is an effective way to improve search results and its interaction is simple enough. Many web-image re-ranking uses this one-click feedback. Its diagram is shown in Fig given below i.e. 1. Given a query keyword input by a user, a pool of images relevant to the query keyword is retrieved by the search engine according to a stored word-image index file. Usually, the size of the returned image pool is fixed, e.g. containing 1300 images. By asking the user to select a query image, which reflects the user's search intention, from the pool, the remaining images in the pool are re-ranked based on their visual similarities with the query image. Most of the existing works assume that there is one dominant cluster of images inside each image set returned by a query keyword.

In addition, these approaches cannot handle ambiguity inside a keyword query, since the assumption that images returned by querying one keyword are all from one class does not hold and the structure of the returned image set is much more complicated. For example, the query for "apple" can return images from 3 main classes (images that are semantically similar), such as Fruit apple, apple pie, and Apple digital products. Within each class, there are several sub-classes (there should be many similarities between images). Also, there are some images noise and neglect (hard to judge).

In this paper, we propose a framework and a search engine to get results in sorted form. A user will select the image by entering the query. We re-rank the return image according to their similarity with the image. The most challenging problem in this framework is how to find similarity. There are many

Features in vision, CBIR (content-based image retrieval), QBIC (Query by image content) and CBVIR (Content-based visual information retrieval). The query is first categories in one pre-define category. Inside each category, we have images which have similar in categories and query. By using this similarity measurement according to the query image, we improve overall retrieval performance.

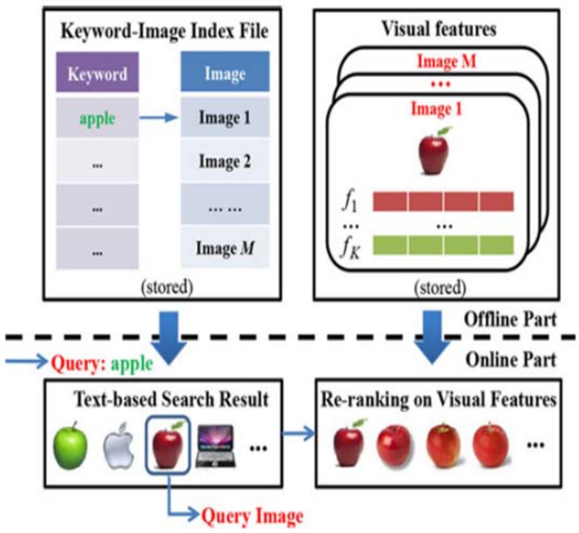

Fig.1.The conventional image re-ranking framework.

II. LITERATURE SURVEY

Even today we are more familiar with traditional web-based image search engine, in which we specify our query and then get results related to the query and choose the desired result. In some cases, we require to keep patience as we need to check more web pages for finding the desired result. This causes user to spend more time for finding the results, as he gets many categories of results altogether.

The more specific technique which we produce, it helps users to get the desired results in less amount of time without need of checking out for many webpages. We produce web-based image re-ranking search engine, which helps in finding the queries faster than that of traditional search engine and it overcomes the disadvantages of it. In this approach we form a cluster of all results depending on the query searched by the user and this cluster will only have images related to that cluster and it will avoid the other images of different cluster, which reduces a lot more time. User then can choose the desired image from the cluster.

There are many algorithms used for forming cluster of images, Cluster is collection of images that will only contain the images related to similar semantics or visual features. We choose K-means algorithm by comparing its features, advantages and disadvantages. Dbscan has disadvantages like it is not scalable and Datasets with varying densities are problematic. Hierarchical clustering have more time complexity than k-means, as the no of inputs increases hierarchical clustering gives worst performance. K-means algorithm is good for large datasets, It is sensitive for noise, sensitive for outliers and scalable .K-means give better performance than other two clustering algorithms hence we prefer k-means algorithm.

III. APPROACH OVERVIEW

The working approach is shown in Figure 2.The reference class of query keywords are automatically discovered after the keyword is searched. Suppose our query keyword is (e.g. "apple"), then set of most related keyword are automatically selected depending on both textual and visual information of the query. (Such as "red apple"," apple MacBook", and "apple IPhone") the keyword expansion (e.g. "red apple") is used to retrieve images by the search engine.

Result obtain is less diverse as compared to the result obtain in case we don't use concept of keyword expansion. This removes the unnecessary reference classes automatically and provides more relevant reference classes. Ex (when keyword is apple) then retrieved reference classes are (such as "apple laptop" and "apple MacBook") .Here redundant reference classes are removed for obtaining the better efficiency. For a every keyword, a multi-class classifier has sub relevant classes which has similar kind of visual features .This sub-relevant classes has similarities related to query keyword. When later user select its choice, that increase the re-ranking accuracy but it also increase storage and reduce the online matching efficiency due to the increased size of semantic signatures. An image may be related to many query keyword that are searched due to searched query may have similar kind of semantics features. Every image in the database is associated with a some relevant keywords. If an image has N relevant keywords, then it is said that it has N semantic signatures which are need to be computed and stored offline. At online stage, several images are retrieved by the search engine depending on the provided query. This pool of images are obtained based on main visual semantics related to the searched query by the user. Later user chooses an image from one of category and then all the images related to that category based on semantic signatures are re-ranked and view by the user.

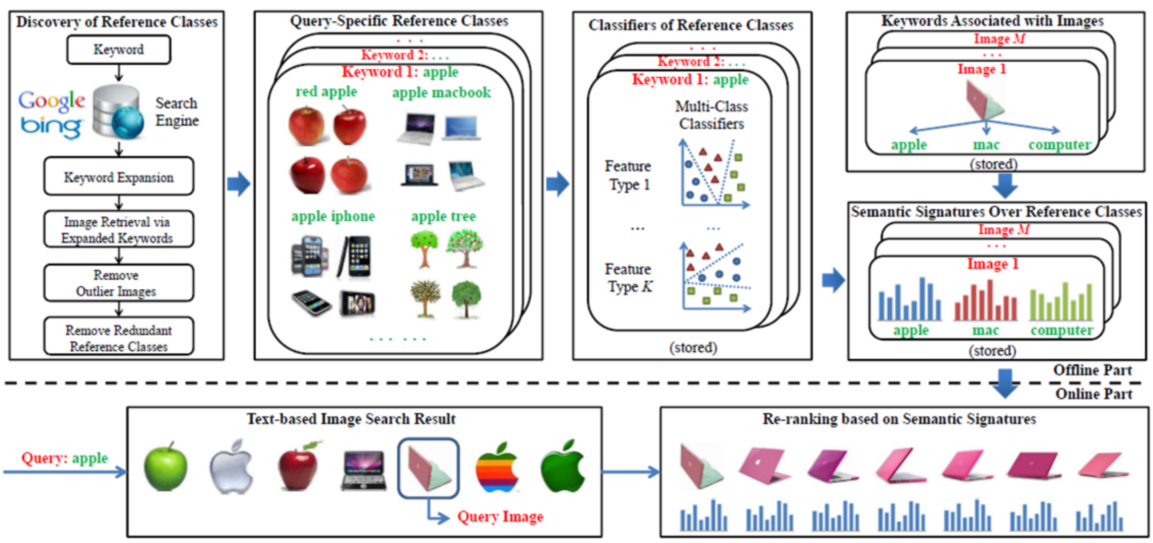

Fig. Diagram for image re-ranking framework.

IV. ARGUMENT ON COMPUTATIONAL COST AND STORAGE

The extra storage of classified and semantic signatures are even smaller than the storage of optical features of images. If we increase the query keywords or images in the database then the cost of the system increases accordingly. Image re-ranking approach is more efficient at the online stage, because the main computational cost of online image re-ranking is on comparing visual features or semantic signatures and the lengths of semantic signatures are much shorter than those of low-level visual features. For example, there are 1500 dimensions are used in visual features. Based on our experimental results, as an average each keyword has 25 reference classes. If only one classifier is trained combining all types of visual features, the semantic signatures are of 25 dimensions on average.

The total cost linearly increases with the number of query keywords, which can be processed in parallel. Given 1000 CPUs, we will be able to process 100,000 query key- words in one day. With the fast expansion of GPUs, which achieve hundreds of times of speedup than CPU, it is practicable to process the industrial scale queries. If separate classifier are able for different types of visual features, the semantic signatures are of 100−200 dimensions. However, our approach needs extra offline working out and storage. According to our experimental study, it takes 20 hours to learn the semantic places of 120 keywords using a machine with Intel Xeon W5580 3.2G CPU.

V. EXPERIMENTS

For testing we have collected 500 images associated with 10 keywords which have collected from Google Image Search and Microsoft image search and other. In this database we have many concept such as mobile, laptop, flower, wallpaper, devices and people etc. and cover a large number of keyword.

The Good images are sub-divided into sub-classes (Images which are visually similar) and main classes (Images which are semantically similar). It is important to treat images that are from same sub-class as the query image to be relevant, while other are "not so relevant".

TABLE 1 DESCRIPTIONS OF DATA SET

	Image for re-ranking			
Data set	Keyword	Images	collection date	Search Engine
1	10	200	Dec-15	Google Image
2	5	100	Jan-16	Google Image

VI. CONCLUSION

We propose a given small sketch about re-ranking manner for searching the multimedia type of data on web. We propose a fresh image re-ranking framework, which learns query-specific semantic spaces to significantly progress the usefulness and efficiency of online image re- ranking. In this re-ranking system, we can achieve the results relatively faster than the traditional search method. Due to this, we need less time as compared to traditional one. This image re-ranking system, which learns query-specific semantic spaces to significantly improve the usefulness and competence of online image re-ranking.

REFERENCES

[1] Xiaogang Wang, Member, IEEE , Shi Qiu, Ke Liu, and Xiaoou Tang, Fellow, IEEE Web Image Re-Ranking Using Query-Specific Semantic Signatures

[2] Meng wang, Hao Li Dacheng Tao, Ke Lu and Xindong Wu, fellow IEEE"multimodal graph based rankings for web image search "November 2012

[3] Nishana S S et al./ International Journal of Computer Science & Engineering Technology (IJCSET) A Survey On Visual Search Reranking 2013

[4] Daniela Oelke_ Peter Bak_ Daniel A. Keim_ Mark Last† Guy Danon‡ University of Konstanz, Ben-Gurion U University of the Negev"Visual Evaluation of Text Feature for Document Summarization and Analysis "2012

[5] V rajkumar, Vipeen V Bopche"image search re-ranking" December 2013

[6] Winston H Hsu Taiwan University, Lyndon S. Kennedy Columbia University, Shih-Fu Chang-Columbia University" video search re ranking through random walk over document level context graph" 2007

[7] IEEE International Conference on Computer Vision and Pattern Recognition. 761–768. CAO, Y., WANG, H., WANG, C., LI, Z., ZHANG, L., AND ZHANG, L. 2010. Mindfinder: interactive sketch-based image search on millions of images. In Proceedings of ACM Multimedia.

[8] CARBONELL, J. AND GOLDSTEIN, J. 1998. The use of mmr, diversity-based re-ranking for reordering documents and producing summaries. In Proceedings of ACM SIGIR conference on Research and Development in Information Retrieval.

[9] Vidit Jain, Manik Varma." Learning to Re-Rank: Query-Dependent Image Re-Ranking Using Click Data". ACM 978-1-4503-0632-4, April 2011.

[10] Lixin Duan, Wen Li, Ivor Wai-Hung Tsang, and Dong Xu, Member, IEEE. "Improving Web Image Search by Bag-Based Re-ranking".IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 20, NO. 11, NOVEMBER 2011.

[11] Linjun Yang, Member, IEEE, and Alan Hanjalic, Senior Member, IEEE." Prototype-Based Image Search Re-ranking".IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 14, NO. 3, JUNE 2012.

[12] Prof. A.G. Dongre1, Anand Kumar Dubey2, Rohin Bhat3, Vijay Thombare4, Maruti Bote5 Web Image Re-ranking using Query Specific Semantic Signatures Vol. 4, Issue 4, April 2015