OAIJSE

OPEN ACCESS INTERNATIONAL JOURNAL OF SCIENCE & ENGINEERING

TAG-BASED IMAGE SEARCH BY SOCIAL RE-RANKING

Miss. Swathi Kalaskar[1], Prof. Vandana Navale[2]

PG Students[1], Assistant Professor[2], Department of Computer Engineering, Dhole Patil College of Engineering, Wagholi, Near Eon IT park., Pune-412207, Maharashtra, India [1][2]

Abstract: **Social media sharing websites like Flickr allow users to annotate images with free tags, which significantly contribute to the development of the web image retrieval and organization. Tag-based image search is an important method to find images contributed by social users in such social websites. However, how to make the top ranked result relevant and with diversity is challenging. In this paper, we propose a social re-ranking system for tag-based image retrieval with the consideration of image's relevance and diversity. We aim at re-ranking images according to their visual information, semantic information and social clues. The initial results include images contributed by different social users. Usually each user contributes several images. First, we sort these images by inter-user re-ranking. Users that have higher contribution to the given query rank higher. Then we sequentially implement intra-user re-ranking in the ranked user's image set, and only the most relevant image from each user's image set is selected. These selected images compose the final retrieved results. We build an inverted index structure for the social image data set to accelerate the searching process.**

Keywords: Re-ranking, Social Clues, Tag-based Image Retrieval, Social Media, Image search.

---∴∵∴---

I INTRODUCTION

Many commercial image search engines in the internet use only keywords as queries. Users type query keywords as input in the hope of finding a certain type of images they search for. The search engine returns images in thousands that are ranked by the keywords extracted from the surrounding text. It is well known that text-based image search process suffers a lot from the ambiguity of query keywords. The keywords provided by users tend to be short and mostly not commonly known. For example, the average query length of the top 2, 000 queries of Picsearch is 1.369 words, and 95% of them contain only one or three words. They cannot describe the content of images accurately and perfectly. The search results are noisy and ambiguous consist of images with quite different semantic meanings. Fig1 shows the top ranked images that are ranked from Bing image search using "Jaguar" as a query. They belong to different categories, such as "Blue Jaguar car", "Black Jaguar car", "Jaguar logo", and "Jaguar animal", due to the ambiguity of the word " Jaguar". The ambiguity issue occurs for so many reasons. First, the query keywords that the user searching for, meanings may be richer than users' expectations. Consider this, the meanings of the word "Jaguar" includes Jaguar animal and Jaguar car and Jaguar logo. Second, the user may not have enough knowledge about the textual description of target images he/she searching for.

The most importantly, in many scenarios, it is difficult for users to explain the visual content of queried images using keywords accurately. In order to solve the ambiguity issues, additional information has to be use.

Figure 1: Top ranked images returned from Bing image search using "Jaguar" as query

Figure 2: An example of a social image with its associated tags

There is an explosion of social media content available online, such as Flickr, Youtube and Zooomr. Such media repositories promote users to collaboratively create, evaluate and distribute media information. They also allow

users to annotate their uploaded media data with descriptive keywords called tags. As an example, Fig. 2 illustrates a social image and its associated user-provided tags. These valuable metadata can greatly facilitate the organization and search of the social media. By indexing the images with associated tags, images can be easily retrieved for a given query [1]. However, since user-provided tags are usually noisy and incomplete, simply applying text-based retrieval approach may lead to unsatisfactory results. Therefore, a ranking approach that is able to explore both the tags and images' content is desired to provide users better social image search results. Currently, Flickr provides two ranking options for tag-based image search. One is "most recent", which orders images based on their uploading time, and the other is "most interesting", which ranks the images by "interestingness", a measure that integrates the information of click-through, comments, etc. In the following discussion, we name these two methods time-based ranking and interestingness-based ranking, respectively. They both rank images according to measures (interestingness or time) that are not related to relevance and it results in many irrelevant images in the top search results.

II RELATED WORK

1. Boost Search Relevance For Tag-Based Social Image Retrieval.

In this paper [2], Author proposes a relevance-based ranking scheme for social image search, aiming to automatically rank images according to their relevance to the query tag. It integrates both the visual consistency between images and the semantic correlation between tags in a unified optimization framework. Author proposes an iterative method to solve the optimization problem, and the relevance based ranking can thus be accomplished.

2. Social Image Search with Diverse Relevance Ranking.

In this paper [3], Author proposes a social re-ranking system for tag based image retrieval with the consideration of image's relevance and diversity. We aim at re- ranking images according to their visual information, semantic information and social clues. The initial results include images contributed by different social users. Usually each user contributes several images. First, we sort has images of inter-user re-ranking. Users that have higher contribution to the given query rank higher. Then we sequentially implement intra-user re-ranking in the ranked user's image set, and only the most relevant image from each user's image set is selected. These selected images compose the final retrieved results. Author builds an inverted index structure for the social image data set to accelerate the searching process.

3. Towards relevant and diverse search of social images.

In This Paper [4], Author presents a diverse relevance ranking scheme which simultaneously takes relevance and diversity into account by exploring the content of images and their associated tags. First, it estimates the relevance scores of images with respect to the query term based on both visual information of images and semantic information of associated tags. Then semantic similarities of social images are estimated based on their tags. Based on the relevance scores and the similarities, the ranking list is

generated by a greedy ordering algorithm which optimizes Average Diverse Precision (ADP), a novel measure that is extended from the conventional Average Precision (AP).

4. Hierarchical clustering of WWW image search results using visual, textual and link information.

In this paper [5], Author proposes a hierarchical clustering method using visual, textual and link analysis. By using a vision-based page segmentation algorithm, a web page is partitioned into blocks, and the textual and link information of an image can be accurately extracted from the block containing that image. By using block-level link analysis techniques, an image graph can be constructed. We then apply spectral techniques to find a Euclidean embedding of the images which respects the graph structure. Thus for each image, we have three kinds of representations, i.e. visual feature based representation, textual feature based representation and graph based representation.

5. The Google Similarity Distance.

In this paper [6], author presents a new theory of similarity between words and phrases based on information distance and Kolmogorov complexity. To fix thoughts we use the world-wide-web as a database, and Google as search engine. The method is also applicable to other search engines and databases. This theory is then applied to construct a method to automatically extract similarity, the Google similarity distance, of words and phrases from the world-wide-web using Google page counts. The world-wide-web is the largest database on earth, and the context information entered by millions of independent users averages out to provide automatic semantics of useful quality. We give applications in hierarchical clustering, classification, and language translation.

III SYSTEM ARCHITECTURE

Our social re-ranking system includes two main sections: online and offline as shown in following figure The offline section contains two parts: a) Inverted index structure construction for image dataset. An inverted index structure is built to accelerate the retrieval speed. b) Feature extraction. In this project, we extract the visual feature, semantic feature and views for the images dataset. Semantic feature refers to the co-occurrence word set of query tags and the tags of the images.

Our online parts consist of the following three steps:

A) Keyword matching. For an input query, our system will return the initial retrieval results by keyword matching. And the following two online steps are all conducted to re-rank the initial results. B) Inter-user re-ranking. The inter-user re-ranking is applied to rank the corresponding users with the consideration of their contributions to the given query. C) Intra-user re-ranking. A regularization framework is proposed to determine the relevance level of each image by fusing the visual, semantic and views information into a unified system. Then we sequentially select the most relevant image in each ranked user's image set. These selected images constitute our re-ranking results. Here in after the details are displayed.

1. Proposed System work:

a) We propose a tag-based image search approach with social re-ranking. We systematically fuse the visual information,

social user's information and image view times to boost the diversity performance of the search result.

b) We propose the inter-user re-ranking method and intra-user re-ranking method to achieve a good trade-off between the diversity and relevance performance. These methods not only reserve the relevant images, but also effectively eliminate the similar images from the same user in the ranked results.

c) In the intra-user re-ranking process, we fuse the visual, semantic and views information into a regularization framework to learn the relevance score of every image in each user's image set. To speed up the learning speed, we use the co-occurrence word set of the given query to estimate the semantic relevance matrix.
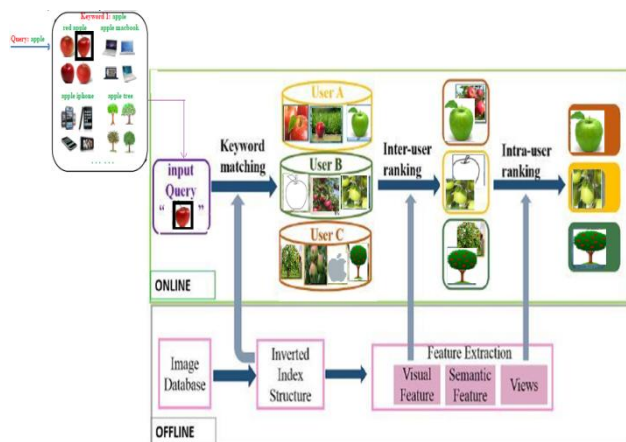

Figure 3: System Architecture

In this project, we propose a social re-ranking method which fuses the user information into the traditional tag-based image retrieval framework. We first get the initial results by keyword matching process. Then the inter-user proposed method is effective and time-saving. However, in the inter-user ranking process only user's contribution is considered and the similarity among users is ignored. Intra-user re-ranking are introduced to re-rank the initial results. Inter-user re-ranking algorithm is applied to rank users according to their contribution to the given query. After the inter-user re-ranking, we further introduce intra-user re-ranking to sequentially select the most relevant image from each image dataset of the ranked users. That's to say, the final retrieved images all have different user. The most relevant image uploaded by the highest contribution user is the first in the retrieved results.

IV CONCLUSION

In this paper, we propose a social re-ranking method for tag-based image retrieval. In this social re-ranking method, inter-user re-ranking and intra-user re-ranking are carried out to obtain the retrieved results. In order to enhance the diversity performance, user information is firstly introduced into our proposed approach and obtains satisfactory results. Besides views of social image is also firstly fused into a traditional regularization framework to enhance the relevance performance of retrieved results.

REFERENCES

[1]Xueming Qian, Member, IEEE, Dan Lu, And Xiaoxiao Liu, "Tag Based Imagesearch By Social Re-Ranking", 12 May 2016.

[2] Vilas Dilip Mane, Prof. Nilesh P. Sable, "Tag Based Image Search By Social Reranking", 12 December 2016.

[3] Shweta Gonde, Uday Chouarisia, Raju Barskar, "A Survey On Web Image Search Using Reranking", 5 May 2014.

[4] A. Ksibi, AB. Ammar, CB. Amar. "Adaptive Diversification For Tag-Based Socialimage Retrieval". International Journal Of Multimedia Information Retrieval, 2014, 3.1: 29-39.

[5] V. Rajakumar, Vipeen V. Bopche, "Image Search Reranking", 5 Dec. 2013.

[6] X. Qian, H. Wang, G. Liu, X. Hou, "Hwvp: Hierarchical Wavelet Packet Texture".

[7] Y. Gao, M. Wang, H. Luan, J. Shen, S. Yan, And D. Tao. "Tag-Based Socialimage Search With Visual-Text Joint Hypergraph Learning", Proceedings Of The Acm International Conference On Multimedia Information Retrieval, 2011:1517-1520.

[8] K. Yang, M. Wang, X. Hua, And H. Zhang. "Social Image Search With Diverse Relevance Ranking", Proceedings Of The Ieee International Conference On Magnetism And Magnetic Materials", 2010:174-184.

[9] M. Wang, K. Yang, X. Hua, And H. Zhang. "Towards Relevant And Diverse Search Of Social Images", IEEE Transactions On Multimedia, 12(8):829-842,2010.

[10] R. Leuken, L. Garcia, X. Olivares, And R. Zwol. "Visual Diversification Of Imagesearch Results", In Proc. Conf., 2009, Pp341350.