STUDY OF SOCIAL IMAGE RE-RANKING ACCORDING INTER AND INTRA USER IMPACT

Miss .Gaikwad Karuna B.
Department of Computer Engineering,
SCSMCOE, Nepti
Ahmednagar,Maharashtra
karunagaikwad777@gmail.com

Miss. Dagadkhair Priyanka V.
Department of Computer Engineering,
SCSMCOE, Nepti
Ahmednagar , Maharashtra
pari.priyanka7777@gmail.com

Miss .Patahare Varsha R.
Department of Computer Engineering,
SCSMCOE, Nepti
Ahmednagar,Maharashtra
varshapathare22@gmail.com

Miss. Aware Ashvini R.
Department of Computer Engineering,
SCSMCOE, Nepti
Ahmednagar , Maharashtra
awareashwini6@gmail.com

Abstract - Social media provides most of the websites for example Facebook that allows users to describe their images with special tag. That helps in the web image retrival system tag based image search is very important method to find out images uploaded by the social users in most of social websites that helps to make the result related and improve the variety of the images. we considered the visual information and click count of the images in this system first we sort the images by inter-user ranking the user who upload the maximum image will rank higher. After this we implement intra-user ranking on the ranked user image set and only the most related image from each user image database is selected . These selected images combine the final retrived results. We build an inverted index for social image databases to improve the speed of searching of an images.

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

I. INTRODUCTION

Many commercial image search engines in internet uses the keywords as queries. Users enter keywords for finding the several images whatever they wants. The search engine returns losts of images in thousands that are ranked by the keywords extracted from the nearer text. But the keywords based search having the problem of the query ambiguity because the keywords entered by the the users are to short and mostly not commonly known. The search results are noisy and ambiguous consist of images with quite different semantic meanings[1]. Figure1 shows the top ranked images that are ranked from Bing image search using “Jaguar” as 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 enoughknowledge about the textual description of target images he/she searching for.

Figure 1: Top ranked images returned from Being image search using “Jaguar” as query

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 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 describe their uploaded media data with descriptive keywords called tags . As an example, Figure2 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[2]. 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[3]. 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

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

In this paper, Author introduces a relevance-based ranking system for social image search, to automatically rank images according to their relevance to the keyword. It combine both the visual consistency between images and the semantic correlation between tags in a join optimization framework[3].

B.Social Image Search with Diverse Relevance Ranking.

In this paper, Author introduces a social re-ranking system for tag based image retrieval that considers two key features of the images i.e. image relevance and its diversity. We ranked images by considering 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 by inter-user re-ranking. Users that have higher contribution to the given query rank higher. Then we sequentially implement intra-user re-ranking on 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 dataset to accelerate the searching process[4]..

C. Towards relevant and diverse search of social images.

In This Paper, Author introduces 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 calculates 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)[5].

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

In this paper, 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[8].

E. The Google Similarity Distance.

In this paper. 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 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[11].

III. PROPOSED ALGORITHM

A. K-means Clustering Algorithm:

K-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed apriority. The main idea is to define k centres, one for each cluster. These centres should be placed in a cunning way because of different location causes different result. So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest center. When no point is pending, the first step is completed and an early group age is done. At this point we need to re-calculate k new centroids as bary canter of the clusters resulting from the previous step. After we have these k new centroids, a new binding has to be done between the same data set points and the nearest new center. A loop has been generated. As a result of this loop we may notice that the k-centers change their location step by step until no more changes are done or in other words centers do not move any more. Finally, this algorithm aims at minimizing an objective function knows as squared error function given by:

Algorithmic steps for k-means clustering
Let X = {x1,x2,x3,……..,xn} be the set of data points and V = {v1,v2,…….,vc} be the set of centers.

1) Initialize both the clusters with first dataset. so C1={8} and C2={6}
2) Now the mean values for C1 and C2 will be M1=8 and m2=6 Now from the consecutive datasets i.e. 2
3) We compare the value of the data with the respective mean values,
4) So the new clusters formed will be C1={8} and c2={6,2} Now recompute the mean value for every cluster m1=8 and m2=(6+2)/2=4
5) Now for the next data set, compare the values with means,so the cluster formed are, c1={8} c2={6,2,1}
6) Now again repeat the steps for recomputing mean for every cluster, so finally the clusters formed will be C1={8,7,6,6,9} and C2={6,2,1,3,4.5,5}

B.Re-ranking Algorithm:
The ranked image for the query tag q:
1. Keyword matching
2. Inter-user ranking
3. Intra-user ranking

The details of these three main part of online system will be described as follows.

Keyword matching for the query, from the inverted file index{}, we can obtain the corresponding images that all tagged with query q, which is denoted bye X. It can be further described by taking the social user's information into account as follows.
X={x(u1)….x(uz),…Xz}

IV. 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:
1) Inverted index structure construction for image dataset. An inverted index structure is built to accelerate the retrieval speed. 2) 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:
1) 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. 2) 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. 3) 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.

A. Proposed System work:
1) 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.
2) 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.
3) In the intra-user re-ranking process, we join 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

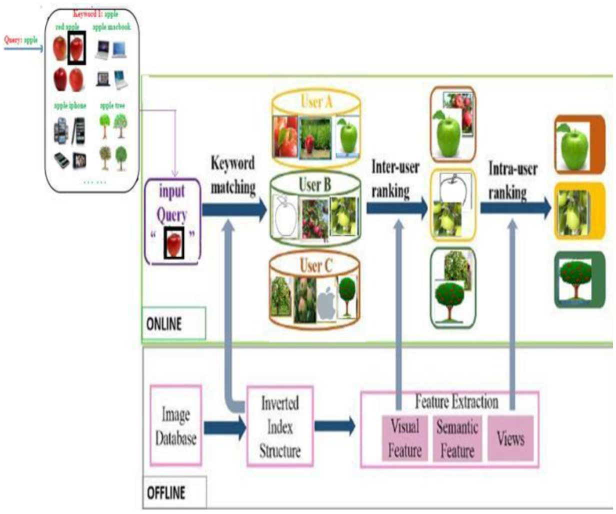

Figure 3: System Architecture

given query to estimate the semantic relevance matrix. 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 and 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.

V. CONCLUSION

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 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. However, in the inter-user ranking process only user's contribution is considered and the similarity among users is ignored.

References

[1] D. Liu, X. Hua, L. Yang, M. Wang, and H. Zhang,"Tag ranking Proceedings of the IEEE International Conference on World Wide Web", 2009: 351-360.

[2] X. Li, C. Snoek, and M. Worring,"Learning tag relevance by neighbor voting for social image retrieval. Proceedings of the ACM International Conference on Multimedia information retrieval", 2008: 180-187.,

[3] D. Liu, X. Hua, M. Wang, and H. Zhang,"Boost Search Relevance For Tag-Based Social Image Retrieval. Proceedings of the IEEE International Conference on Multimedia and Expo", 2009:1636-1639.

[4] 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.

[5] 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.

[6] A. Ksibi, AB. Ammar, CB. Amar,"Adaptive diversification for tag-based social image retrieval. International Journal of Multimedia Information Retrieval", 2014, 3.1: 29-39.

[7] Y. Gao, M. Wang, H. Luan, J. Shen, S. Yan, and D. Tao,"Tag-based social image search with visual-text joint hypergraph learning. Proceedings of the ACM International Conference on Multimedia information retrieval", 2011:1517-1520.

[8] D. Cai, X. He, Z. Li, W. Ma, and J. Wen.,"Hierarchical clustering of WWW image search results using visual, textual and link information", In Proc. ACM Multimedia Conf., 2004, pp.

[9] K. Song, Y. Tian, T. Huang, and W. Gao,"Diversifying the image retrieval results", In Proc. ACM Multimedia Conf., 2006, pp. 707–710.

[10] R. Leuken, L. Garcia, X. Olivares, and R. Zwol,"Visual diversification of image search results",In Proc. WWW Conf., 2009, pp.341–350.

[11] R. Cilibrasi and P. Vitanyi,"The Google Similarity Distance", IEEE Transactions on Knowledge and Data Engineering, 19(3):1065-1076, 2007.

[12] X. Qian, H. Wang, G. Liu, X. Hou, "HWVP: Hierarchical Wavelet Packet Texture Descriptors and Their Applications in Scene Categorization and Semantic Concept Retrieval", Multimedia Tools and Applications, May 2012.