|| Volume 3 || Issue 4 || April 2018 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH
AND ENGINEERING TRENDS

Novel Image Re-ranking Using Data Mining and Image Processing Techniques

Shinde Sarika Maruti[1], Prof. Manjushri Mahajan[2]
G. H. Raisoni College Of Engineering & Management, Wagholi Pune, Maharashtra, India[1,2]
ssshindesarika30@gmail.com[1] , manjushri.mahajan@raisoni.net[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 images relevance and diversity[1]. 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 on the ranked users image set, and only the most relevant image from each users image set is selected. These selected images compose the final retrieved results. We build an inverted index structure for the social image dataset to accelerate the searching process.**

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

I INTRODUCTION

The following parts present the existing works related to the above three aspects respectively.

Tag Processing Strategy

It has been long acknowledged that tag ranking and refinement play an important role in the re-ranking of tag-based image retrieval, for they lay a firm foundation on the development of re-ranking in tag based image retrieval (TBIR)[3].

Relevance Ranking Approach

To directly rank the raw photos without undergoing any intermediate tag processing, utilized an optimization framework to automatically rank images based on their relevance to a given tag. Visual consistency between images and semantic information of tags are both considered. Proposed a hyper graph learning approach, which aims to estimate the relevance of images. They investigate the bag-of-words and bag-of-visual words of images, which are extracted from both the visual and textual information of image. Proposed a Support Vector Machine classifier per query to learn relevance scores of its associated photos. proposed a two-step similarity ranking scheme that aims to preserve both visual and semantic resemblance in the similarity ranking. In order to achieve this, a self-tune manifold ranking solution that focuses on the visual-based similarity ranking and a semantic-oriented similarity re-ranking method are included[3].

- **Diversity Enhancement**

The relevance based image retrieval approaches can boost the relevance performance; however the diversity performance of searching are often ignored. Many researchers dedicated their extensive efforts to solve this problem. Proposed a hierarchical clustering method to cluster the search results into different semantic clusters by using visual, textual and link analysis. Similarly, in, studied three visually diverse ranking methods to re-rank the image search results based on the visual characteristics of these images. Different from clustering, proposed a re-ranking method to meet users ambiguous needs by analysing the topic richness[3].

II LITERATURE REVIEW

1. **D. Liu, X. Hua, M. Wang, and H. Zhang, "Boost Search Relevance For Tag-Based Social Image Retrieval", 2009.**

In this paper, 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 propose an iterative method to solve the optimization problem, and the relevance based ranking can thus be accomplished[12].

2. **K. Yang, M. Wang, X. Hua, and H. Zhang, "Social Image Search with Diverse Relevance Ranking", 2010.**

In this paper, Author 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 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[10].

3. **M. Wang, K. Yang, X. Hua, and H. Zhang, "Towards relevant and diverse search of social images", 2010.**

In This Paper, 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)[11].

4. **D. Cai, X. He, Z. Li, W. Ma, and J. Wen, "Hierarchical clustering of WWW image search results using visual, textual and link information", 2004.**

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[15].

5. **R. Cilibrasi and P. Vitanyi, "The Google Similarity Distance", 2007.**

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[14].

6. **T. Chaitanya Reddy, K. Chaitanya, "Ranking of Images Based on Tags", Jul – Aug 2017.**

The Previous system problem is user tagging is known to be uncontrolled, ambiguous, and overly personalized, a fundamental problem is how to interpret the relevance of a user-contributed tag with respect to the visual content the tag is describing. We propose solution to the system is a social re-ranking method for tag based image retrieval. It is a new approach of tag image re-ranking for social dataset. It can be used for retrieving images on the basis of tagging. This approach for Social image analysis and retrieval is important for helping people organize and access the increasing amount of user-tagged multimedia. Tag-based image search is an important method to find images contributed by social users in social websites[3].

7. **Apeksha P. Bochare, Prof. B. D. Phulpagar, "A Survey on Tag Based Image Search by Social Re-ranking", February – 2017.**

Day to day the amount of tablet or mobile devices and web services are increasing speedily and the victory of many online social media websites, which allow users to create and share media information such as images and videos as well as describe the created content with tags. Social media sharing websites like Flickr or "Picasa" allow users to annotate or tag images with free tags, which notably support to the development of the web image retrieval and organization. Tag-based image search is an important and effective method than Content based image search to find images contributed by social users in such social websites like Flickr and Picasa[1].

8. **B. Usharani , "Relevance and Diversity of Images by Using Tags", 2017.**

Now a day, there is a growth of digital images and video archives. Some of these are very important from user point of view. The collection of these digital images may be for personal use and may be for public use. For searching images in the database, there is the need of secure, fast and efficient techniques. Tag-based image search is an effective technique to find images in the databases and in the social websites[6].

9. **Neethu Mathew, "Review on Tag based Image Search by Social Re-Ranking", August 2017.**

There are many photos sharing websites like Flickr which allow users to annotate pictures with descriptive keywords called tags. Tag based image search helps to find images contributed by users in such social media sharing websites, which significantly support to the development of the web image retrieval and organization. How to make top positioned ranked result relevant is the challenging problem[4].

II LITERATURE ANALYSIS

Sr. No	Paper Name	Author Name	Year	Description
1	Boost Search Relevance For Tag-Based Social Image Retrieval	D. Liu, X. Hua, M. Wang, and H. Zhang	2009	Author proposed 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 image and the semantic correlation between tags in a unified optimization framework.
2	Social Image Search with Diverse Relevance Ranking	K. Yang, M. Wang, X. Hua, and H. Zhang.	2010	Author proposed a social re-ranking system for tag based image retrieval with the consideration of images relevance and diversity. Author aimed at re-ranking images according to their visual information, semantic information and social clues. The initial results include images contributed by different social users.
3	Towards relevant and diverse search of social images	M. Wang, K. Yang, X. Hua, and H. Zhang.	2010	Author presented 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.
4	Hierarchical clustering of WWW image search results using visual, textual and link information	D. Cai, X. He, Z. Li, W. Ma, and J. We.	2004	Author proposed 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.
5	The Google Similarity Distance.	R. Cilibrasi and P. Vitanyi.	2007	Author presented 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.

III SYSTEM ARCHITECTURE

The proposed system will avoid the unwanted result to user. We propose a social re-ranking system for tag-based image retrieval with the consideration of image's relevance and diversity. 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 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.

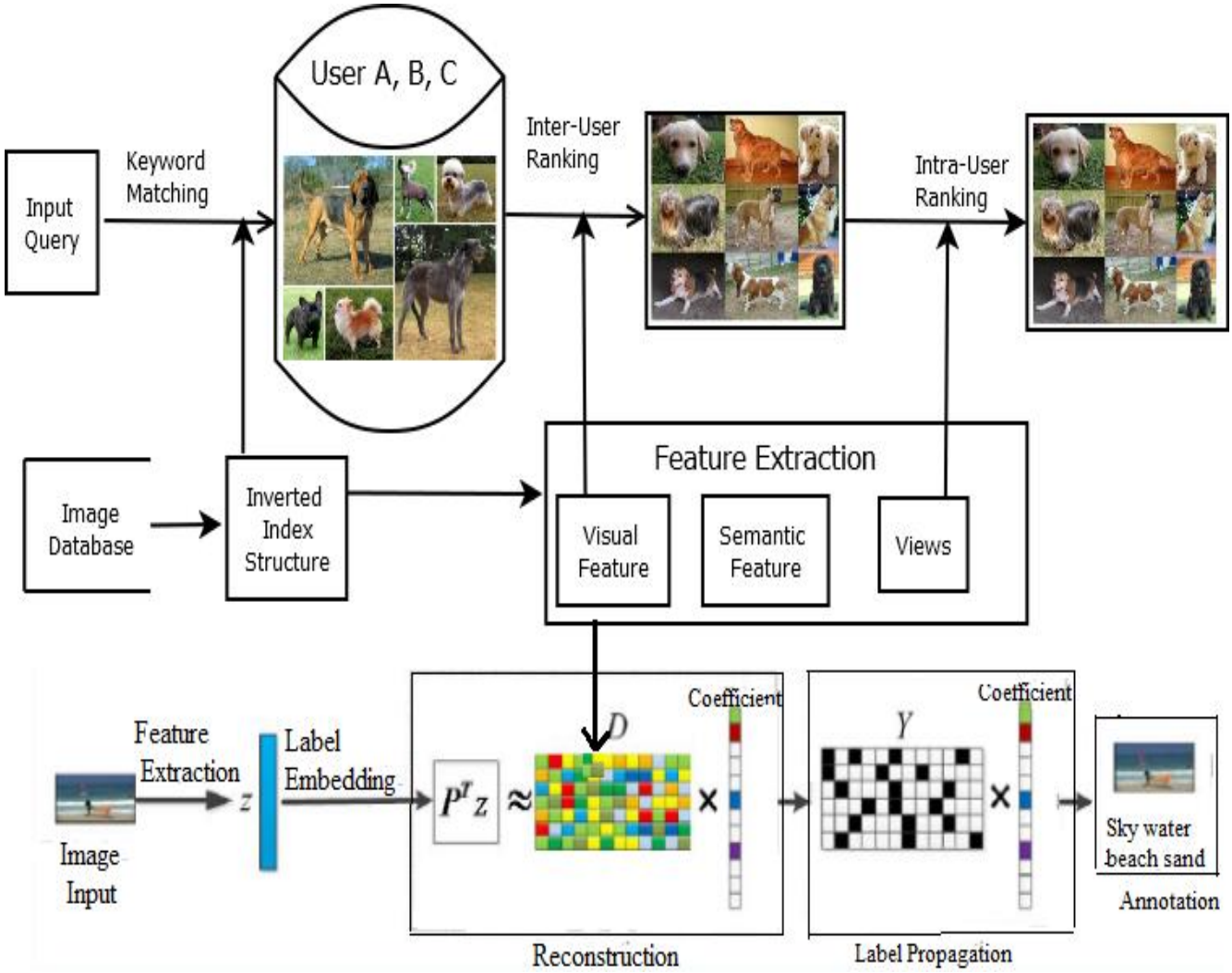


Figure1: System Architecture

IV METHODOLOGY

1 Tag Based Image Retrieval:

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. Tag-based image search is more com-monly used in social media than content based image retrieval and context-and-content based image retrieval.

2 Social Tags:

Tag mismatch. Social tagging requires all the users in the social network to label their uploaded images with their own keywords and share with others. Different from ontology based. image annotation, there is no predefined on-tology or taxonomy in social image tagging. Every user has his own habit to tag images. Even for the same image, tags contributed by different users will be of great difference.

3 Image Search:

The tag-based image search can be easily accomplished by using the tags as query terms. However, the weakly relevant tags, noisy tags and duplicated information make the search result unsatisfactory. Most of the literatures regarding the re-ranking of the tag-based image retrieval focus on tag processing, image relevance ranking and diversity enhancement of the retrieval results.

4 Social Re-ranking:

We build an inverted index structure for the social image dataset to accelerate the searching process. Experimental results on Flickr dataset show that our social re-ranking method is effective and efficient. We first get the initial results by keyword matching process.

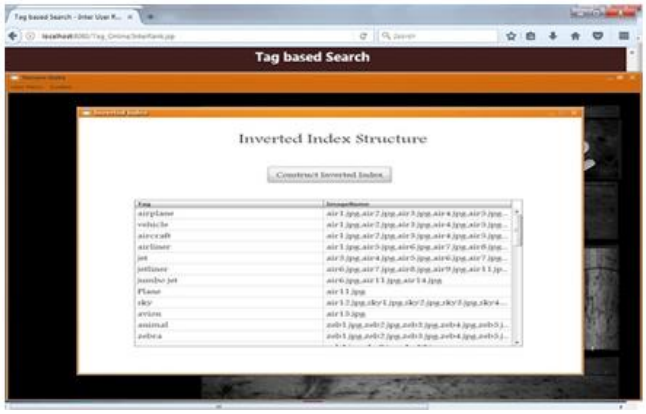
V ADVANTAGES

- Tag-based image search is an important method to find images contributed by social users in such social websites.

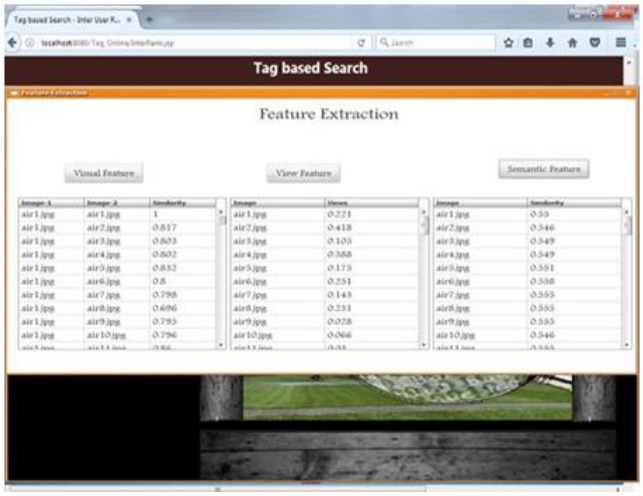
- This project proposes a social re-ranking system for tag-based image retrieval with the consideration of image's relevance and diversity.
- We aim at re-ranking images according to their visual information, semantic information and social clues.
- The initial results include images contributed by different social users. Usually each user contributes several images.

VI RESULTS

1. Inverted Index Search



2. Feature Extraction



3 Image User Re-ranking Zebra

VII CONCLUSION AND FUTURE WORK

In this Project, we propose a social re-positioning strategy for tag-based picture retrieval. In this social re-positioning technique, between client re-positioning and intra-client re-positioning are done to get the recovered outcomes. To improve the jumper sity execution, client data is right off the bat brought into our proposed approach and acquires agreeable outcomes. Additionally, perspectives of social picture are likewise right off the bat melded into a customary regularization structure to upgrade the importance execution of recovered outcomes. Notwithstanding, in the between client positioning procedure just client's commitment is considered and the closeness among clients is disregarded. Furthermore, numerous data in Flickr dataset are as yet overlooked, for example, title data, time stamp et cetera. For future work, we will research the comparability among client bunches in Flickr dataset. In this way, we can meld these connections to upgrade the assorted variety execution of picture positioning framework.

REFERENCES

[1] Apeksha P. Bochare, Prof. B. D. Phulpagar, "A Survey on Tag Based Image Search by Social Re-ranking", International Journal of Recent Trends in Engineering & Research (IJRTER) Volume 03, Issue 02; February – 2017.

[2] A. P. Bochare, Dr. B. D. Phulpagar,"Social Re-ranking using Tag Based Image Search", International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 04 Issue: 0 | July -2017.

[3] T.Chaitanya Reddy, K.Chaitanya, "Ranking of Images Based on Tags", International Journal of Computer Science Trends and Technology (IJCST) – Volume 5 Issue 4, Jul – Aug 2017.

[4] Neethu Mathew , "Review on Tag based Image Search by Social Re-Ranking", International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887 Volume 5 Issue VIII, August 2017.

[5] Snehal N. Paraj, Suresh K. Shirgave, "Survey on Tag Based Image Search", International Journal of Computer Engineering In Research Trends Volume 4, Issue 11, November - 2017.

[6] B.Usharani , "Relevance and Diversity of Images by Using Tags", Asian Journal of Science and Applied Technology ISSN: 224-0698 Vol. 6 No. 2, 2017.

[7] Gunjan, Er. Madan Lal, "Investigation of Various Image Stenography Techniques in Spatial Domain", International Journal of Computer Engineering In Research Trends, Volume 3, Issue 6, June-2016.

[8] "A Survey on Security in Tag based Image search with De-duplication in Social Re-ranking", September-2016.

[9] Weizhi Ma, Min Zhang, Yiqun Liu, Shaoping Ma, and Lingfeng Chen, "Beyond Your Interests: Exploring the

Information Behind User Tags", Springer International Publishing Switzerland 2015.

[10] 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.

[11] 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.

[12] 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.

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

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

[15] 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.