TAG BASED IMAGE SEARCH BY SOCIAL RE-RANKING

CHINNAM SATYASREE, B.SRINIVASA RAO
Mtech Scholar, Assistant Professor
Department of of CSE
ISTS Women's Engineering College, Rajanagaram, Rajamahendravaram, A.P, India.

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 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. 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.

1. INTRODUCTION

With the development of social media based on Web 2.0, amounts of images and videos spring up everywhere on the Internet. This phenomenon has brought great challenges to multimedia storage, indexing and retrieval. Generally speaking, tag-based image search is more commonly used in social media than content based image retrieval [47] and context-and-content based image retrieval. In recent years, the re-ranking problem in the tag-based image retrieval has gained researchers' wide attention. Nonetheless, the following challenges block the path for the development of re-ranking technologies in the tag-based image retrieval.

1) 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 ontology 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 [26, 48]. Thus, the same image can be interpreted in several ways with several different tags according to the background behind the image. Thus, many seemingly irrelevant tags are introduced.

2) Query ambiguity. Users cannot precisely describe their request with single words and tag suggestion system always recommend words that are highly correlated to the existing tag set, thus add little information to a users' contribution. Besides, polysemy and synonyms are the other causes of the query ambiguity. Thus, a fundamental problem in the re-ranking of the tag-based social image retrieval is how to reliably solve these problems. As far as the "tag mismatch" problem is concerned, tag refinement [2, 3, 21, 23, 25, 27], tag relevance ranking [18, 34, 36, 46] and image relevance ranking approach [4, 8, 16, 22, 28, 34, 35] have been dedicated to overcome this problems. As for the "query ambiguity" problem, an effective approach is to provide diverse retrieval results that cover multiple topics underlying a query. Currently, image clustering [9, 11] and duplicate removal [5-7, 10, 29, 30, 32] are the major approaches in settling the diversity problem. However, the essence of social images is ignored. The social images uploaded and tagged by users are user-oriented. These user-oriented images which share the same user and tagged with same query are always taken in a fixed time interval at a specific spot. It is well-known that, images taken in the same time interval and fixed spot are fairly similar. To diversify the top ranked search results, it's better to re-rank the results by removing the duplicate images from the same user.

Starting from this intuition and above analysis, we propose a social re-ranking algorithm which user information is firstly introduced into the traditional ranking method considering the semantics, social clues and visual information of images. The contributions of this paper can be described as follows: 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 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.

Comparing with the preliminary work [44], we have made some improvements as follows: 1) In order to improve the robustness of the algorithm to obtain the co-occurrence word set with respect to the given query in [44], a new self-adaptive algorithm is introduced in this paper, in which relative frequency of each tag about the given query is required and a self-adaptive parameter is decided by this relative frequency. 2) In the intra-user re-ranking process, we take the views into consideration to learn the relevance score of each image on the basis of [44]. In order to achieve this, a new iterative algorithm to obtain the relevance score is proposed.

3) Comparing with the algorithm proposed in [44], this paper is more considerate.
Discussions about weight selection and image features in the regularization framework are complemented. Through this discussion, we find that our performance doesn't rely on the adjustment of parameters and feature selection. It's robust and relatively stable. Besides, in order to find an optimal number of representative images which are selected from each user's image set, many new comparison experiments and comprehensive discussions are added.

2. SYSTEM ANALASYS

2.1 EXISTING SYSTEM:

- Lee and Neve proposed to learn the relevance of tags by visually weighted neighbor voting, a variant of the popular baseline neighbor voting algorithm.
- Agrawal and Chaudhary proposed a relevance tag ranking algorithm, which can automatically rank tags according to their relevance with the image content. A modified probabilistic relevance estimation method is proposed by taking the size factor of objects into account and random walk based refinement is utilized.
- Li et al. presented a tag fusion method for tag relevance estimation to solve the limitations of a single measurement on tag relevance. Besides, early and late fusion schemes for a neighbor voting based tag relevance estimator are conducted.
- Zhu et al. proposed an adaptive teleportation random walk model on the voting graph which is constructed based on the images relationship to estimate the tag relevance.
- Sun et al. proposed a tag clarity score measurement approach to evaluate the correctness of a tag in describing the visual content of its annotated images.

2.2 DISADVANTAGES OF EXISTING SYSTEM:

- 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 ontology or taxonomy in social image tagging. Every user has own habit to tag images. Even for the same image, tags contributed by different users will be of great difference. Thus, the same image can be interpreted in several ways with several different tags according to the background behind the image. Thus, many seemingly irrelevant tags are introduced.
- Query ambiguity. Users cannot precisely describe their request with single words and tag suggestion system always recommend words that are highly correlated to the existing tag set, thus add little information to a users' contribution. Besides, polysemy and synonyms are the other causes of the query ambiguity.

2.3 PROPOSED SYSTEM:

- We propose a social re-ranking algorithm which user information is firstly introduced into the traditional ranking method considering the semantics, social clues and visual information of images. The contributions of this paper can be described as follows:
- 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.
- 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.
- 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.

2.4 ADVANTAGES OF PROPOSED SYSTEM:

- In order to improve the robustness of the algorithm to obtain the co-occurrence word set with respect to the given query, a new self-adaptive algorithm is introduced in this paper, in which relative frequency of each tag about the given query is required and a self-adaptive parameter is decided by this relative frequency.
- In the intra-user re-ranking process, we take the views into consideration to learn the relevance score of each image. In order to achieve this, a new iterative algorithm to obtain the relevance score is proposed.
- This system is more considerate when compared to existing systems.
- Discussions about weight selection and image features in the regularization framework are complemented. Through this discussion, we find that our performance doesn't rely on the adjustment of parameters and feature selection. It's robust and relatively stable. Besides, in order to find an optimal number of representative images which are selected from each user's image set, many new comparison experiments and comprehensive discussions are added.

3. LITERATURE SURVEY

4.

1) **Interoperability of personal health records**

AUTHORS: J. L ahteenm€ aki, J. Lepp anen, and H. Kaijanranta,

The establishment of the Meaningful Use criteria has created a critical need for robust interoperability of health records. A universal definition of a personal health record (PHR) has not been agreed upon. Standardized code sets have been built for specific entities, but integration between them has not been supported. The purpose of this research study was to explore the hindrance and promotion of interoperability standards in relationship to PHRs to describe interoperability progress in this area. The study was conducted following the basic principles of a systematic review, with 61 articles used in the study. Lagging interoperability has stemmed from slow adoption by patients, creation of disparate systems due to rapid development to meet requirements for the Meaningful Use stages, and rapid early development of PHRs prior to the mandate for integration among multiple systems. Findings of this study suggest that deadlines for implementation to capture Meaningful Use incentive payments are supporting the creation of PHR data silos, thereby hindering the goal of high-level interoperability.

2) Applying cloud computing model in PHR architecture. AUTHORS: S. Kikuchi, S. Sachdeva, and S. Bhalla,

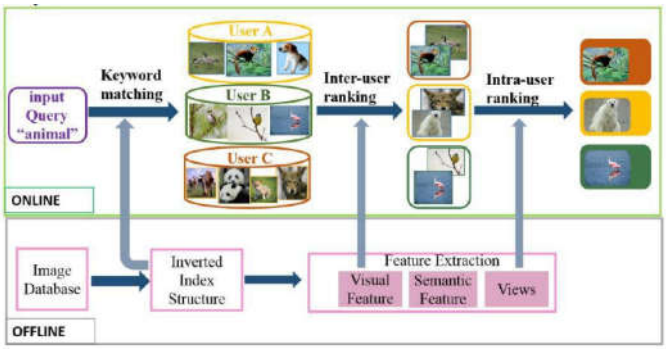
In recent years, some practical and commercial Personal Health Records and some related services such as Google Health [1] and Microsoft HealthVault [2] have been launched. On the other hand, Cloud Computing has matured more and become the major streams to realize a more effective operational environment. However so far, there have been few studies in regards to applying Cloud architecture in the PHR explicitly despite generating volume data. In this paper, we review our trial on the general architecture design by applying the Cloud components for supporting healthcare record areas and clarify the required conditions to realize it.

3. Health Information Privacy, Security, and Your EHR. AUTHORS: M. Bellare
If your patients lack trust in Electronic Health Records (EHRs) and Health Information Exchanges (HIEs), feeling that the confidentiality and accuracy of their electronic health information is at risk, they may not want to disclose health information to you. Withholding their health information could have life-threatening consequences. To reap the promise of digital health information to achieve better health outcomes, smarter spending, and healthier people, providers and individuals alike must trust that an individual's health information is private and secure. Your practice, not your EHR developer, is responsible for taking the steps needed to protect the confidentiality, integrity, and availability of health information in your EHR system.

4. A Secure Anti-Collusion Data Sharing Scheme for Dynamic Groups in the Cloud AUTHORS: C. Ng and P. Lee. Revdedup
Benefited from cloud computing, users can achieve an effective and economical approach for data sharing among group members in the cloud with the characters of low maintenance and little management cost. Meanwhile, we must provide security guarantees for the sharing data files since they are outsourced. Unfortunately, because of the frequent change of the membership, sharing data while providing privacy-preserving is still a challenging issue, especially for an untrusted cloud due to the collusion attack. Moreover, for existing schemes, the security of key distribution is based on the secure communication channel, however, to have such channel is a strong assumption and is difficult for practice. In this paper, we propose a secure data sharing scheme for dynamic members. Firstly, we propose a secure way for key distribution without any secure communication channels, and the users can securely obtain their private keys from group manager. Secondly, our scheme can achieve fine-grained access control, any user in the group can use the source in the cloud and revoked users cannot access the cloud again after they are revoked. Thirdly, we can protect the scheme from collusion attack, which means that revoked users cannot get the original data file even if they conspire with the untrusted cloud. In our approach, by leveraging polynomial function, we can achieve a secure user revocation scheme. Finally, our scheme can achieve fine efficiency, which means previous users need not to update their private keys for the situation either a new user joins in the group or a user is revoked from the group.

5. **ADVANCE SECURITY TO CLOUD DATA STORAGE AUTHORS:** P. Lee, and W. Lou
The proposed system is an effective and flexible distributed Scheme with explicit dynamic data support to ensure the correctness of user's data in the cloud. To fully ensure the data integrity and save the cloud users computation it is of critical importance to enable public auditing service for cloud data storage, so that users may depend on independent third party auditor to audit the outsourced data. The Third party auditor can periodically check the integrity of all the data stored in the cloud .which provides easier way for the users to ensure their storage correctness in the cloud.

SYSTEM DESIGN



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 are also firstly fused into a traditional regularization framework to enhance the relevance performance of retrieved results. Discussions and experiments have demonstrated that our 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. In addition to this, many information in Flickr dataset are still ignored, such as title information, time stamp and so on. For future work, we will investigate the similarity among user groups in Flickr dataset. Therefore, we can fuse these relationships to enhance the diversity performance of image ranking system.

REFERENCES

1. Flickr. http://www.Flickr.com/.
2. 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.
3. 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.,
4. 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.
5. 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.
6. 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.
7. Ksibi, AB. Ammar, CB. Amar. Adaptive diversification for tag-based social image retrieval. International Journal of Multimedia Information Retrieval, 2014, 3.1: 29-39.

8. 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.
a. 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. IEEETransactions 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.
13. X. Qian, G. Liu, D. Guo. Object categorization using hierarchical wavelet packet texture descriptors. in Proc. ISM 2009, pp.44-51.
a. Xueming Qian, Yisi Zhao, Junwei Han: Image Location Estimation by Salient Region Matching. IEEE Transactions on Image Processing 24(11): 4348-4358 (2015)
b. D. Wu, J. Wu, M. Lu. A Two-Step Similarity Ranking Scheme for Image Retrieval. In Parallel Architectures, Algorithms and Programming, pp. 191-196, IEEE, 2014.
c. D. Zhou, O. Bousquet, T. Lal, J. Weston. Learning with local and global consistency. Advances in neural information processing systems, 16(16), 321-328, 2004.
d. G. Agrawal, R. Chaudhary. Relevancy tag ranking. In Computer and Communication Technology, pp. 169-173, IEEE, 2011.
e. L. Chen, S. Zhu, Z. Li. Image retrieval via improved relevance ranking. In Control Conference, pp. 4620-4625, IEEE, 2014.
14. L. Wu, R. Jin. Tag completion for image retrieval. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 35(3), 716-727, 2013.
15. Y. Yang, Y. Gao, H. Zhang, J. Shao. Image Tagging with Social Assistance. In Proceedings of International Conference on Multimedia Retrieval (p. 81), ACM, 2014.
a. L. Chen, D. Xu, I. Tsang. Tag-based image retrieval improved by augmented features and group-based refinement. Multimedia, IEEE Transactions on, 14(4), 1057-1067, 2012.
b. S. Lee, W. D. Neve. Visually weighted neighbor voting for image tag relevance learning. Multimedia Tools and Applications, 1-24, 2013.
c. Z. Xu, X. Luo, Y. Liu, L. Mei. Measuring Semantic Relatedness between Flickr Images: From a Social Tag Based View. The Scientific World Journal, 2014.
d. X. Li. Tag relevance fusion for social image retrieval. Multimedia Systems, 1-12, 2014.
16. X. Qian, X. Liu, C. Zheng. Tagging photos using users' vocabularies. Neurocomputing, 111, 144-153, 2013. [27] D. Mishra. Tag Relevance for Social Image Retrieval in Accordance with Neighbor Voting Algorithm. IJCSNS, 14(7), 50, 2014.
a. Y. Hu, M. Li. Multiple-instance ranking: Learning to rank images for image retrieval. In Computer Vision and Pattern Recognition, CVPR 2008. IEEE Conference on (pp. 1-8).
17. F. Sun, M. Wang, D. Wang. Optimizing social image search with multiple criteria: Relevance, diversity, and typicality. Neurocomputing, 95, 40-47, 2012.
18. B. Wang, Z. Li, M. Li. Large-scale duplicate detection for web image search. In Multimedia and Expo, 2006 IEEE International Conference on (pp. 353-356).
19. K. Weinberger, M. Slaney. Resolving tag ambiguity. In Proceedings of the 16th ACM international conference on Multimedia (pp. 111-120), ACM, 2008.

A. Ksibi, G. Feki, A. Ammar. Effective Diversification for Ambiguous Queries in Social Image Retrieval. In Computer Analysis of Images and Patterns (pp. 571-578), 2013.
A. Khosla, A. Sarma. What makes an image popular?. In Proceedings of the 23rd international conference on World wide web (pp. 867-876), 2014.
B. Haruechaiyasak. Improving social tag-based image retrieval with CBIR technique (pp. 212-215), Springer Berlin Heidelberg, 2010.
20. X. Zhu, W. Nejdl. An adaptive teleportation random walk model for learning social tag relevance. In Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval (pp. 223-232), ACM, 2014.
21. J. Yu, D. Tao, M. Wang. Learning to Rank Using User Clicks and Visual Features for Image Retrieval. IEEETrans.Cybern.(2014).
a. S. Ji, K. Zhou, C. Liao, Z. Zheng, GR. Xue. Global ranking by exploiting user clicks. In Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval (pp. 35-42), ACM, 2009.
22. G. Dupret. A model to estimate intrinsic document relevance from the clickthrough logs of a web search engine. In Proceedings of the third ACM international conference on Web search and data mining (pp. 181-190). ACM, 2010.
a. S. Cen, L. Wang, Y. Feng, H. Bai. Efficient image reranking by leveraging click data. In Multimedia and Expo Workshops, 2014 IEEE, International Conference on (pp. 1-4).
b. X. Hua, M. Ye. Mining knowledge from clicks: MSR-Bing image retrieval challenge. In Multimedia and Expo Workshops (ICMEW), 2014 IEEE International Conference on (pp. 1-4).
c. L. Chen, D. Xu, I. Tsang. Tag-based web photo retrieval improved by batch mode re-
23. tagging. In Computer Vision and Pattern Recognition, 2010 IEEE Conference on(pp. 3440-3446). IEEE.
A. Sun, S. Bhowmick, **"**Image tag clarity: In search of visual representative tags for social images,**"** in 1st SIGMM Workshop on Social Media, New York, 2009, pp. 19–26.
24. X. Qian, X. Hua, Y. Tang, and T. Mei, **"**social image tagging with diverse semantics**"**, IEEE Trans. Cybernetics, vol.44, no.12,2014, pp. 2493-2508.
a. X. Qian, D. Lu, X. Liu,**"** Tag based image retrieval by user-oriented ranking**"**. Proceedings of International Conference on Multimedia Retrieval. ACM, 2015.
25. K.Jarvelin and J.Kekalainen. Cumulated Gain-Based Evaluation of IR Techniques. In ACM Transactions on Information System, 2002. Y. Gu, X. Qian, Q. Li. Image Annotation by Latent Community Detection and Multikernel Learning. IEEE Transactions on Image Processing 24(11): 3450-3463 (2015).