Volume 6, Issue 12, December 2016 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

A Survey on Social Visual Image Ranking Techniques for Web Search

Nayan Sawant*, Prof. N. R. Wankhade
Department of Computer Engineering, KCT's Late G.N. Sapkal College of Engineering,
Maharashtra, India

Abstract*— In this paper we present deep literature survey on the Social Visual Image Ranking techniques for web search. These days researchers are gaining interest in matching textual query with the visual image their surrounding texts or tags for Web image search. The returned results are often unsatisfactory due to their deviation from user intentions. Many proposed system for ranking the images are studied in this paper and their limitations are suggested with our proposed approach.

Keywords*— specific Social-Visual Ranking(SVR), Social image search, Image re-ranking, Social relevance.

I. INTRODUCTION

The advent of social multimedia tagging – assigning tags or keywords to images, music, or video clips by common users – is significantly reshaping the way people generate , manage, and search multimedia resources. Good examples are Flickr, which hosts more than 2 billion images with around 3 million new uploaded photos per day [1], and YouTube, which serves 100 million videos and 65,000 uploads daily [2]. Apart from their usage for general-purpose search, these rich multimedia databases are triggering many innovative research scenarios in areas as diverse as personalized information delivery [3], concept similarity measurement [4], tag recommendation [5]. One would expect user-contributed tags to be a good starting point for all these applications. Despite the success of social tagging, however, tags contributed by common users are known to be ambiguous, limited in terms of completeness, and overly personalized [6], [7]. This is not surprising because of the uncontrolled nature of social tagging and the diversity of knowledge and cultural background of its users. Although the relevance of a tag given the visual content can be subjective for a specific user, an objective criterion is desirable for general-purpose search and visual content understanding. We consider a tag relevant to an image if the tag accurately describes objective aspects of the visual content, or in other words, users with common knowledge relate the tag to the visual content easily and consistently. Other tags are subjective or overly personalized and thus we consider those irrelevant, as illustrated in Figure 1.

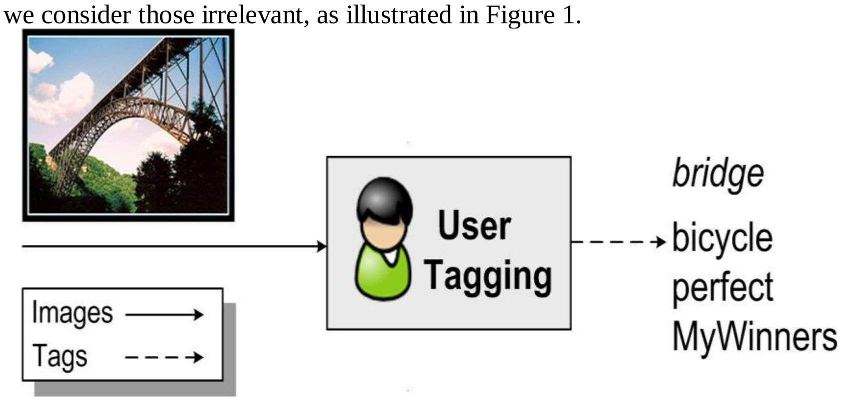
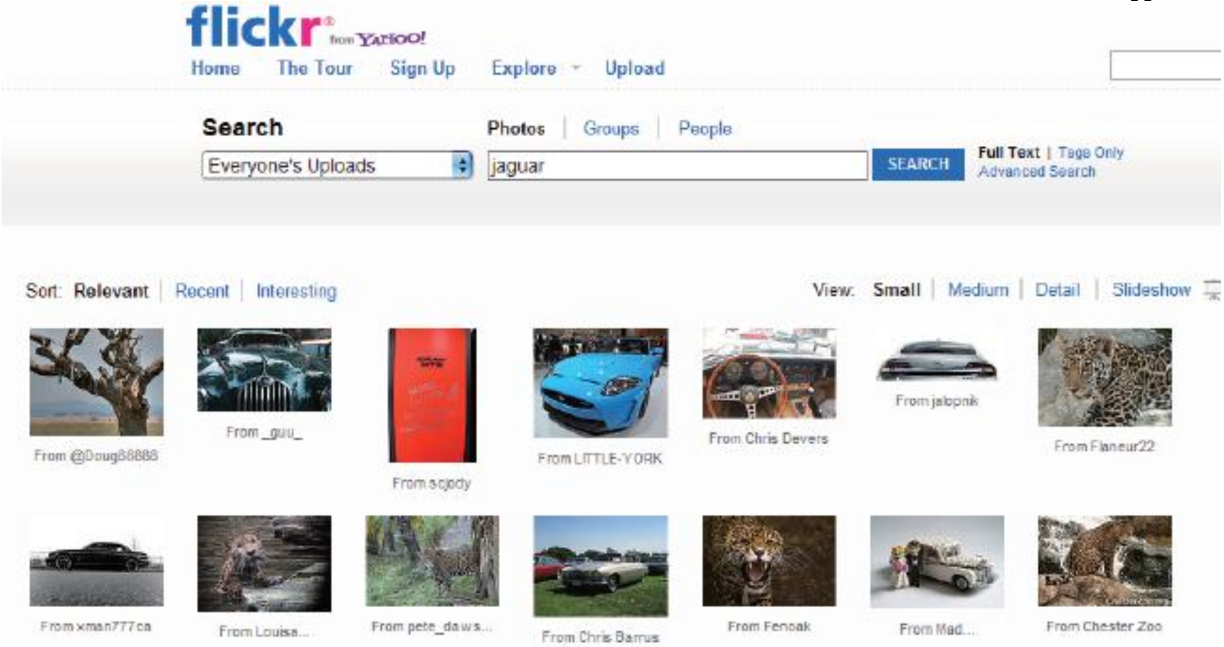


Fig 1. Dataflow of user tagging

A lot of existing research work focuses on improving the relevance between the textual query and visual images. However, there exists semantic gap between user intention and textual query. Let's take the query "jaguar" as an example, as shown in Fig.2. Different users have different intentions when inputting the query "jaguar". Some are expecting leopard images, while others are expecting automobile images. This scenario is quite common, particularly for queries with heterogeneous concepts or general (non-specific) concepts. This raises a fundamental but yet unsolved problem in Web image search: how to understand user intentions when users conducting image search? In the past years, interest analysis is very difficult due to the lack of personal data. With the development of social media platforms, such as Flickr and Facebook, the way people can get social data has been changed: users' profiles, interests and their favorite images are exposed online and open to public, which are crucial information sources to implicitly understand user interests.



II. LITERATURE SURVEY

Existing methods to automatically predict tag relevance with respect to the visual content often heavily rely on supervised machine learning methods [9]–[11]. In general, the methods boil down to learning a mapping between lowlevel visual features, e.g., color and local descriptors, and high-level semantic concepts, e.g., airplane and classroom. Since the number of training examples are limited for the supervised methods, the methods are not scalable to cover the potentially unlimited array of concepts existing in social tagging. Moreover, uncontrolled visual content contributed by users creates a broad domain environment having significant diversity in visual appearance, even for the same concept [12]. The scarcity of training examples and the significant diversity in visual appearance might make the learned models unreliable and difficult to generalize. Therefore, in a social tagging environment with large and diverse visual content, a lightweight or unsupervised learning method which effectively and efficiently estimates tag relevance is required.

To improve the visual relevance, many methods are proposed which are based on incorporating visual factors into image ranking [13,14]. In [13] the author is concerned with the matter of multimodal fusion in video search. First, an object-sensitive approach is utilized to query analysis to enhance the baseline results of text-based video search. Then, authors propose a PageRanklike graph-based approach to text-based search result reranking. to raised exploit the underlying relationship between video shots, the planned re- ranking scheme at the same time leverages textual connection, semantic construct connection, and low-level-feature-based visual similarity. During this PageRank-like scheme, a set of graphs is built with the video shots as vertexes, and therefore the abstract and visual similarity between video shots as “hyperlinks.” A changed topic-sensitive PageRank algorithmic rule is then applied on these graphs to propagate the connection scores through all connected video shots.

An essential problem in these methods is to measure the visual similarity[15]. As an effective approach, VisualRank[16] determines the visual similarity by the number of shared SIFT features[17]. After a similarity based image link graph was generated, an iterative computation similar to PageRank[18] is utilized to rerank the images. Visual Rank obtains a better performance than text-based image search in the measurement of relevance for queries with homogeneous visual concepts. However, for queries with heterogeneous visual concepts, VisualRank does not work well[19]. With the development of social media platform, the concept of social image retrieval was proposed, which brings more information and challenges to us[20]. Most of works in social image search focus on tags [21]. However, the quality of recommendation is based on the technique of tag annotation[22], which is not mature enough. Overall, understanding user intention is significant but challengeable in social media platform. Many social media sites such as Flickr offer millions of groups for users to share images with others. There are tons IV of works based on improving the user experience [23]. Group information is an efficient way to estimate user interests.

Authors in [24]Propose a generic approach that contributes to up the informativeness of image tags by combining generalizations concerning the spatial arrangement tendencies of physical objects within the universe and statistics of linguistic communication use patterns that have been mined from the online. The approach, that we have a tendency to see as Reading between the Tags, provides for every tag related to a picture, first, a prediction regarding quality , i.e., whether or not or not the tag denotes a physical entity, and, then, regarding the real-world size of that entity, i.e., large, medium or tiny. Mining takes place employing a set of Language Use Frames (LUFs) that consisted of linguistic communication neighborhoods characteristic of tag categories.

III. DISCUSSION

Studying the existing system we find that combining social relevance and visual relevance faces the following challenges:

1. Social data sparseness. In social media platform, most users only possess a small number of favored images, from which it is difficult to discover user intentions. With the hypothesis that users in the same community share similar interests, a community-specific method is more practical and effective than a user-specific method.
2. The tradeoff between social relevance and visual relevance. Although social relevance may guarantee the interest of returned images for the user, the quality and representativeness of images, cannot be ignored. Both of which are essential for good search results. Thus, some social relevance and visual relevance are needed to be addressed and subtly balanced.
3. Complex factors. To generate the final image ranking, one needs to consider the user query, returned images from current search engines, and many complex social factors derived from social media platforms. How to integrate these heterogeneous factors in an effective and efficient way is quite challenging.

This disadvantages can be overcome by social re-ranking algorithm can be implemented where user information is firstly introduced the conventional ranking method considering the semantics, social clues & visual information of images. Also a tag-based image search approach with social re-ranking can be used. The visual information can be systematically fuse, social users information and image view times to boost the diversity performance of the search result also 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 and 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 users 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. The proposed architecture is show in fig 3.

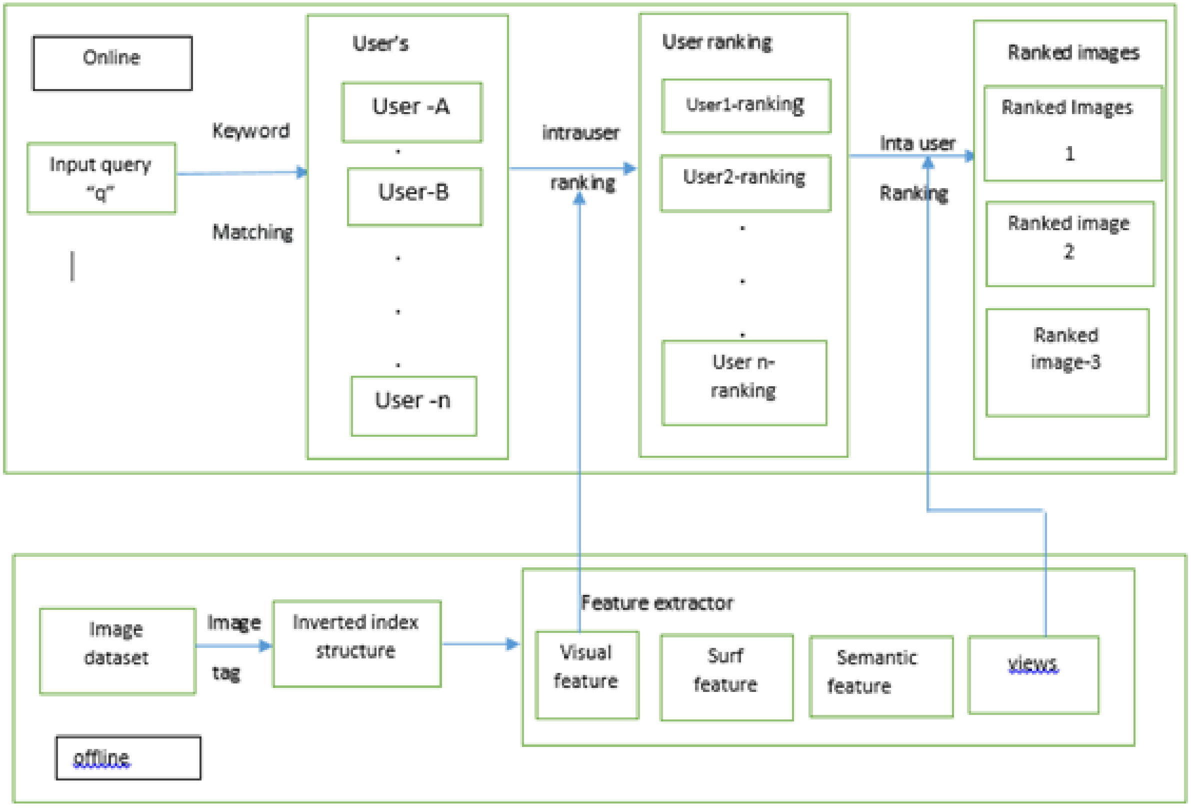


Fig 3. Proposed system Architecture

IV. CONCLUSIONS

In this paper we studied the existing mechanisms which are used for Social Visual Image Ranking techniques for web search. 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.

ACKNOWLEDGMENT

My sincere thanks go to KCTs Late G.N. Sapkal College of Engineering for providing a strong platform to develop my skill and capabilities. We would like to thanks all those who directly or indirectly help us in presenting the paper. I hereby take this opportunity to express our heartfelt gratitude towards the people whose help is very useful to complete our project. I would like to express our heartfelt thanks to my guide Prof. N. R. Wankhade whose experienced guidance became very valuable for me.

REFERENCES

[1] E. Auchard, "Flickr to map the world's latest photo hotspots," Reuters [Online], 2007, Nov., available: http://www.reuters.com/article/technologyNews/idUSHO94233920071119?sp=true.

[2] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, and S. Moon, "I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system," in *Proc. ACM SIGCOMM on Internet measurement*, 2007, pp. 1–14.

[3] D. A. Shamma, R. Shaw, P. L. Shafton, and Y. Liu, "Watch what I watch: using community activity to understand content," in *Proc. ACM MIR*, 2007, pp. 275–284.

[4] L. Wu, X.-S. Hua, N. Yu, W.-Y. Ma, and S. Li, "Flickr distance," in *Proc. ACM Multimedia*, 2008, pp.31–40.

[5] B. Sigurbj ¨ornsson and R. van Zwol, "Flickr tag recommendation based on collective knowledge," in *Proc. WWW*, 2008, pp. 327–336.

[6] S. A. Golder and B. A. Huberman, "Usage patterns of collaborative tagging systems," *Information Science*, vol. 32, no. 2, pp. 198–208, 2006.

[7] K. K. Matusiak, "Towards user-centered indexing in digital image collections," *OCLC Systems and Services*, vol. 22, no. 4, pp. 283–298, 2006.

[8] X. Li, C. G. M. Snoek, and M. Worring, "Learning tag relevance by neighbor voting for social image retrieval," in *Proc. ACM MIR*, 2008, pp. 180–187.

[9] K. Barnard, P. Duygulu, D. Forsyth, N. de Freitas, D. M. Blei, and M. I. Jordan, "Matching words and pictures," *Jour. Machine Learning Research*, vol. 3, no. 6, pp. 1107–1135, 2003.

[10] E. Chang, G. Kingshy, G. Sychay, and G. Wu, "CBSA: contentbased soft annotation for multimodal image retrieval using Bayes point machines," *IEEE Trans. CSVT*, vol. 13, no. 2, pp. 26–38, 2003.

[11] J. Li and J. Z. Wang, "Real-time computerized annotation of pictures," *EEE Trans. PAMI*, vol. 30, no. 6, pp. 985–1002, 2008.

[12] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. PAMI*, vol. 22, no. 12, pp. 1349–1380, 2000.

[13] W. H. Hsu, L. S. Kennedy, and S.-F. Chang. Video search reranking via information bottleneck principle. In Proceedings of the 14th annual ACM international conference on Multimedia, MULTIMEDIA '06, pages 35–44, New York, NY, USA, 2006. ACM.

[14] J. Liu, W. Lai, X.-S. Hua, Y. Huang, and S. Li. Video search re-ranking via multi-graph propagation. In Proceedings of the 15th international conference on Multimedia, MULTIMEDIA '07, pages 208–217, New York, NY, USA, 2007. ACM.

[15] R. I. Kondor and J. Lafferty. Diffusion kernels on graphs and other discrete structures. In In Proceedings of the ICML, pages 315–322, 2002.

[16] Y. Jing and S. Baluja. Visualrank: Applying pagerank to large-scale image search. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 30(11):1877 –1890, nov. 2008.

[17] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener. Graph structure in the web. Computer Networks, 33(1C6):309 – 320, 2000.

[18] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. Technical Report 1999-66, Stanford InfoLab, November 1999. Previous.

[19] Y. Liu, T. Mei, and X.-S. Hua. Crowdreranking: exploring multiple search engines for visual search reranking. In Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, SIGIR '09, pages 500–507, New York, NY, USA, 2009. ACM.

[20] M. Wang, B. Ni, X.-S. Hua, and T.-S. Chua. Assistive tagging: A survey of multimedia tagging with human-computer joint exploration. ACM Comput. Surv., 44(4):25:1–25:24, Sept. 2012.

[21] M. Wang, K. Yang, X.-S. Hua, and H.-J. Zhang. Towards a relevant and diverse search of social images. Multimedia, IEEE Transactions on, 12(8):829 –842, dec. 2010.

[22] J. Sang, J. Liu, and C. Xu. Exploiting user information for image tag refinement. In Proceedings of the 19th ACM international conference on Multimedia, MM '11, pages 1129–1132, New York, NY, USA, 2011. ACM.

[23] L. A. F. Park and K. Ramamohanarao. Mining web multi-resolution community based popularity for information retrieval. In Proceedings of the sixteenth ACM conference on Conference on information and knowledge management, CIKM '07, pages 545–554, New York, NY, USA, 2007. ACM.

[24] Larson, Martha, Christoph Kofler, and Alan Hanjalic. "Reading between the tags to predict real-world size-class for visually depicted objects in images." Proceedings of the 19th ACM international conference on Multimedia. ACM, 2011.

