Volume 5, Issue 9, September 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Re-Ranking of Web Image Prediction using Multimodal Sparse Code and Voting Strategy

[1]Supriya S. Jadhav, [2]Pooja H. Lokhande, [3]Yogita S. Shedge, [4]Amol G. Baviskar
[1, 2, 3] B.E. COMP & SPPU, [4]Asst. Prof. in CSE
[1, 2, 3, 4] JSPM'S BSIOTR, Wagholi, Pune, Maharashtra, India

Abstract*— This paper give you the information about the web image search using the Multimodal Sparse code and voting strategy .The performance of Text-based image search which has been improved by using Image re-ranking. Existing re-ranking algorithms have two main drawbacks, first the textual meta-data associated with images is often mismatched with their actual visual content and second the extracted visual features do not accurately describe the semantic similarities between images. Pseudo-relevance feedback (PRF) tool is used in many existing re-ranking system. A critical problem for click-based methods is the lack of click data, so only a small number of web images have actually been clicked on by users. So our aim is to solve this problem by predicting image clicks. Click prediction is effective to improving the performance of prestigious graph-based image re-ranking algorithms is useful in real world data.

Keywords:- Image re-ranking, click, manifolds, sparse code, hypergraph.

I. INTRODUCTION

Image retrieval is a key issue of user point of view. Normal way of image retrieval is the text based image retrieval technique (TBIR). TBIR-needs more semantic textual description of web images .This technique is popular but needs very specific description of the query which is difficult and not always possible. Image retrieval has been adopted in most of the major search engines like Google, Yahoo!, Bing, etc. A large number of image search engines mainly employ the surrounding texts located at the images and the image names to index the images.

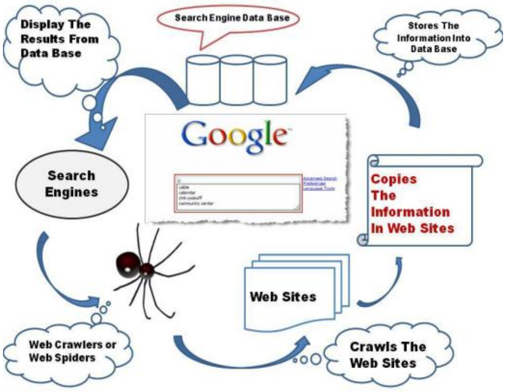

Fig 1 Working of Google search engine.

The search engine navigates through the pages and collects the images. It gives the client the top ranked image which is the one with maximum number of clicks from the user and a set of images. This is the technique of text based image retrieval system.

However, this limits the capability of the search engines in retrieving the semantically related images using a given query. Well-recognized image search engines, such as Bing, Yahoo and, usually Google use textual meta-data included in the surrounding text, captions titles, and URLs, to index web images. Although the performance of text-based image retrieval for many searches is valid, the accuracy and efficiency of the retrieved results could still be improved significantly. One major problem impacting performance is the unequal between the actual content of image and the textual data on the web page. One method used to solve this problem is image re-ranking, in which both textual and visual information are merge to return improved results to the user. The ranking of images based on a text-based search is considered a reasonable baseline, although with noise. Extracted visual information is then used to re-rank related images to the top of the list. All existing re-ranking methods use a tool known as pseudo-relevance feedback (PRF), where a proportion of the top-ranked images are assumed to be related, and subsequently used to build a model for re-ranking. This is in contrast to relevance feedback, where users explicitly give feedback by labelling the top results as positive or negative. In the classification-based PRF technique, the top-ranked images are regarded as pseudo-positive and low-ranked images regarded as pseudo-negative.

In summary, we present the important contributions of this paper:

- An effectively utilize search engine derived images annotated with clicks, and then predict the clicks for new input images without clicks. Based on the obtained clicks, we rerank the images, a strategy which could be beneficial for improving commercial image searching.
- System proposed a novel method called multimodal hypergraph learning-based sparse coding. This method uses both early and late fusion in multimodal learning. At the same time learning the sparse codes and the weights of different hypergraphs, the performance of sparse coding is performed.
- We conduct comprehensive experiments to empirically analyze the proposed method on real world web image datasets, collected together from a commercial search engine. Their corresponding clicks are collected from internet users. The experimental results demonstrate the effectiveness of the proposed method.

II. EXISTING SYSTEM

Most existing reranking technique use a tool known as pseudo-relevance feedback (PRF), where a proportion of the top ranked images are assumed to be relevant, and used to build a model for re-ranking. This is in contrast to relevance feedback, where user provides feedback by labeling the top results as positive or negative. In the classification based PRF method, the top ranked images are considered as pseudo-positive and low ranked images considered as pseudo-negative examples to train a classifier, and then ranked. Hsu et al. also adopt this pseudo-positive and pseudo negative image method to develop a clustering based re ranking algorithm.

Disadvantages of Existing System:

- One major problem is the mismatches between the actual content of image and the textual data on the web page.
- The problem with these techniques is the reliability of the obtained pseudo-positive and pseudo-negative images is not assured.

III. PROPOSED SYSTEM

In this system we proposed a method named multimodal hyper graph learning-based sparse code for click prediction, and we apply the predicted clicks to re-rank web images. Both strategies of early and late fusion of multiple features are used in this method through three main steps.

- First Construct a web image base with associated click annotation, collected from a search engine. The search engine has recorded clicks for each image which is present on database. Indicate that the images with high clicks are relevant to the queries, while remaining images are non-relevant with zero clicks that are not relevant to the query.
- Now consider both early and late fusion in the proposed objective function. The early fusion is based on directly calculating multiple visual features, and is applied in the sparse coding term. In Late fusion manifold learning term is used. For web images without clicks or no clicks, it constructs a hyper graph learning to create a group of manifold; this preserves the local smoothness by using the hyper edges. Otherwise a graph that has an edge between two or more vertices, a set of vertices is connected by the hyper edge in a hyper graph. Common graph-based learning technique all time considers the pair wise relationship between two vertices, it ignoring the higher-order relationship among three or more vertices. Using this term it can preserve the local smoothness of the constructed sparse codes.
- Finally, an alternating optimization procedure is carried out to find the complementary nature of different modalities. The weights of different modalities and the sparse codes are simultaneously abstracted using this optimization strategy. A voting strategy is then used to predict if an input image will be clicked or not, based on its sparse code.

IV. SYSTEM ARCHITECTURE

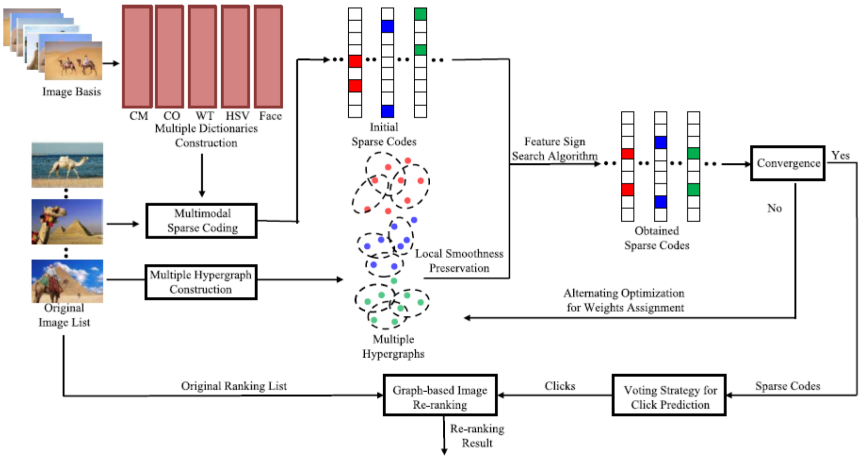

Fig 2 The framework of multimodal hypergraph learning based sparse coding for click prediction.

First, multiple features are extracted from both the input images and image bases. Second, multiple hypergraph Laplacians are constructed, and the sparse codes are built. Meanwhile, the locality of the obtained sparse codes is preserved by using manifold learning on hypergraphs. Then, the sparse codes of the images and the weights for different hypergraphs are obtained by simultaneously optimization through an iterative two-stage optimization procedure. A voting strategy is used to achieve click data propagation. Finally, the obtained sparse codes are integrated with the graph-based schema for image re-ranking.

V. CONCLUSIONS

A new multimodal hyper graph learning based sparse coding method for the click prediction of images is used. The obtained sparse codes can be used for image re-ranking by collecting them with a graph-based schema .System adopt a hyper graph to build a group of manifolds, which explore the complementary characteristics of different features through a group of weights of images. Unlike a graph that has an edge between two vertices, a set of vertices are connected by a hyper edge in a hyper graph. This helps preserve the local smoothness of the constructed sparse codes. Then, an alternating optimization procedure is performed and the weights of different modalities and sparse codes are simultaneously retrieved using the optimization strategy. Finally, a voting strategy is used to predict the click from the corresponding sparse code. Experimental results on real-world data sets have indicated that the proposed method is effective in determining click prediction. Additional experimental results on image re-ranking recommend that this method can improve the results returned by search engines.

REFERENCES

[1] L. Duan, W. Li, I. W. Tsang, and D. Xu, "Improving web image search by bag-based re-ranking," *IEEE Trans. Image Process*., vol. 20, no. 11, pp. 3280–3290, Nov. 2011.

[2] X. Tian, L. Yang, J. Wang, X. Wu, and X. Hua, "Bayesian visual re-ranking," *IEEE Trans. Multimedia*, vol. 13, no. 4, pp. 639–652,Aug. 2010.

[3] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in Proc. *IEEE Int. Conf. Comput.Vis.,* Oct. 2009, pp. 2272–2279.

[4] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach.Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[5] C. Wang, S. Yan, L. Zhang, and H. Zhang, "Multi-label sparse coding for automatic image annotation," in Proc. *IEEE Conf. Comput. Vis. PatternRecognit.,* Jun. 2009, pp. 1643–1650.

[6] M. Wang, X. S. Hua, R. Hong, J. Tang, G. Qi, and Y. Song, "Unified video annotation via multigraph learning," *IEEE Trans. Circuits Syst.Video Technol.*, vol. 19, no. 5, pp. 733–746, May 2009.

[7] R. Zass and A. Shashua, "Probabilistic graph and hypergraph matching," *in Proc. Int. Conf. Comput. Vis. Pattern Recognit.,* 2008, pp. 1–8.

[8] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.

[9] H. Lee, A. Battle, R. Raina, and A. Y. Ng, "Efficient sparse coding algorithms," *in Proc. Adv. Neural Inf. Process. Syst.,* 2006, pp. 801–808

[10] C. Snoek, M. Worring, and A. Smeulders, "Early versus late fusion in semantic video analysis," *in Proc. ACM Int. Conf. Multimedia*, 2005,pp. 399–402.

[11] D. Donoho, "For most large underdetermined systems of linear equations, the minimal l1-norm solution is also the sparsest solution," Dept. Statist., Stanford Univ., Stanford, CA, USA, Tech. Rep., 2004.

[12] J. Sivic and A. Zisserman. Video Google: A text retrieval approach to object matching in videos. In Proc. ICCV, 2003.