International Conference on Recent Innovations in Engineering and Management
Dhananjay Mahadik Group of Institutions (BIMAT) Kolhapur, Maharashtra (ICRIEM-16)
23rd March 2016, www.conferenceworld.in ISBN: 978-81-932074-5-1

COMPARATIVE ANALYSIS OF RERANKING TECHNIQUES FOR WEB IMAGE SEARCH

Suvarna V. Jadhav[1], A.M.Bagade[2]

[1,2]*Department of Information Technology, Pune Institute of Computer Technology, Pune,(India)*

ABSTRACT

The Image reranking is effective for improving the web image search. The search engines are mostly based on text and constrained due to user search by keyword which results into ambiguity among images. The noisy or irrelevant images may be present in the retrieved results. The purpose of web image search re-ranking is to reorder retrieved elements to get optimal rank list. The existing visual reranking schemes improve text-based search results by making the use of visual information. These methods are based on low-level visual features, and do not take into account the semantic relationship among images. Semantic attribute assisted re-ranking is proposed for web image search. Using the classifiers for predefined attributes, each image is represented by attribute features. The hypergraph is used to model the relationship between images. Hypergraph ranking is carried out to order the images. The basic principle is that similar images should have similar ranking. This paper presents a detail review of different image retrieval and reranking approaches. The purpose of the survey is to provide an overview and analysis of the functionality, merits, and demerits of the existing image reranking systems, which can be useful for researchers for developing effective system with more accuracy.

Keywords*: *Attribute assisted, hypergraph learning, image search, reranking, semantic attributes.

I. INTRODUCTION

World Wide Web is a vital part of our daily life. Image search has become a keystone of many commercial search engines. Flooding of wide range of images makes it necessary to develop some strategic solution so that exact images can be extracted and easily accessible.

Image search reranking is defined as the enhancement of search results by using image visual information to reorder the initial text-based search results. Most of the frequently-used commercial web image search engines are based on indexing and searching of textual information associated with images, like image file names, surrounding texts, universal resource locator, etc. Although text-based image search is useful for large-scale image collections, it suffers from the weakness that textual information cannot comprehensively and significantly illustrate the content of images. Some irrelevant images may get retrieved in the search results. To deal with the difficulties in text-based image search, visual reranking can be used. It integrates visual information of images to process the text-based search results. Generally, the top returned images are reordered via various reranking approaches by mining their visual patterns. Also semantic attributes, as a intermediate level descriptor narrow down the semantic gap low level visual features and high level semantic meanings. These describe image regions that are common within object category. Hence semantic attributes have achieved good performance in assisting image classification.

II. IMAGE RETRIEVAL

The ever-growing number of digital images on the Internet, retrieving relevant images from a large collection of database images has become an important research topic. Over the past decades, many image retrieval systems have been developed, such as text-based image retrieval (TBIR), content-based image retrieval (CBIR) and hybrid approach.

2.1. Text-Based Approach

The TBIR has been widely used in popular image search engines e.g. Google, Bing and Yahoo! .Specifically, a user is required to input a keyword as a textual query to the retrieval system. Then the system returns the ranked relevant images whose surrounding texts contain the query keyword, and the ranking score is obtained according to some similarity measurements between the query keyword and the textual features of relevant images. Text-based search techniques have been verified to perform well in textual documents; they often result in mismatch when applied to the image search. The reason is that metadata cannot represent the semantic content of images

2.2 Content-Based Approach

Most search engine works on Text Based Approaches but there exist alternative approach, content based image retrieval that requires a user to submit a query image, and return images that are similar in content .Google is one of the search engine that works on content based Image re-ranking. The extracted visual information is natural and objective, but completely ignores the role of human knowledge in the interpretation process. As the result, a red flower may be regarded as the same as a rising sun, and a fish the same as an airplane etc.

2.3 Hybrid approach

Recent research combines both the visual content of images and the textual information obtained from the Web for the WWW image retrieval. Such methods exploit the usage of the visual information for refining the initial text-based search result. Especially, through user's relevance feedback, i.e., the submission of desired images or visual content-based queries, the re-ranking for image search results can achieve significant performance improvement [15].

III. RERANKING STRATEGIES

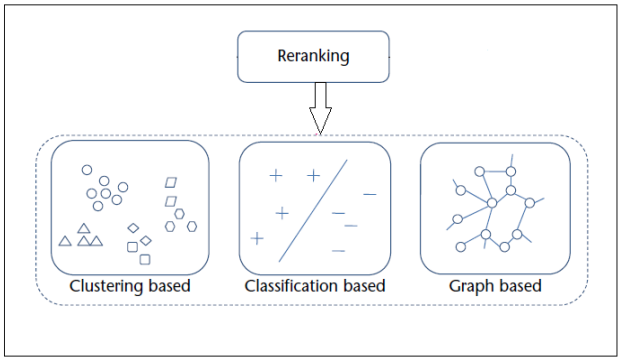

Fig.1: Classification of Reranking Strategies [12]

3.1 Clustering Based Reranking

Clustering methods are based on the observation that query-relevant images often share high visual similarity. By using various clustering algorithms, this kind of methods reorganize the initial text search result by grouping visually similar samples together.

3.1.1. Bag-Based Reranking

Given a textual query in traditional text-based image retrieval relevant images are to be reranked using visual features after the initial text-based search. In this paper [9], a new bag-based reranking framework is proposed for large-scale TBIR. The relevant images are clustered using both textual and visual features. By treating each cluster as a bag and the images in the bag as instances, problem is formulated as multi-instance (MI) learning problem. The top ranked bags are used as pseudopositive training bags, while pseudonegative training bags can be obtained by randomly sampling a few irrelevant images that are not associated with the textual query. Advantage is that the automatic bag annotation method can achieve the best performance as compared with other traditional image reranking methods for large scale TBIR. Disadvantage is that it ambiguity may occur in the labels of instances. Visual features and user search intention can be used to improve the performance.

3.1.2. Intent Search

To use visual information in order to solve the ambiguity in text-based image retrieval, novel Internet image search approach i.e. Intent search [8] is proposed that only requires the user to click on one query image with minimum effort and images from a pool retrieved by text-based search are reranked based on both visual and textual content. Key contribution is to capture the users' search intention from this one-click query image. All these things are automatic, without extra effort from the user. Besides this key contribution, a set of visual features which are both effective and efficient in Internet image search are designed. Disadvantage is that more user burden is added for labeling the regions that the user things more important.

Clustering methods are suitable for queries that have obvious near-duplicate images in the initial text-based results. Limitation is that, for these queries that return visually diverse images without salient patterns, this kind of methods cannot achieve good performance.

3.2 Classification Based Reranking

Instead of using multiple clusters, reranking can be simplified using binary classification. There are normally three steps: select the pseudo-positive and pseudo-negative samples from the initial text-based search results; train a classifier using the selected samples; and reorder the samples according to the relevance scores predicted by the trained classifier. For the first step, pseudo relevance feedback (PRF) is typically used to select training samples. PRF is a concept introduced from text retrieval. It assumes that a fraction of the top ranked documents in the initial search results are pseudo-positive. The pseudo-negative samples are selected from either the least-relevant samples in the initial ranking list or the database, with the assumption that few samples in the database are relevant to the query. In the second step, various classifiers, such as support vector machine can be used.

3.2.1 Active Reranking

The visual information is insufficient to infer the user's intention, especially when the query term is ambiguous. The reranking with user interaction [10] is proposed in which some images are selected according to an active sample selection strategy, and then user is required to label them. With the knowledge of user's intention

including both the labeling information and the learned discriminative submanifold, the reranking process is conducted. Advantage is that, compare to Intent search [8] it can learn the user's intention more extensively and completely. Limitation is that more user burden for labeling the samples.

3.2.2 Prototype Based

The previous methods for image search reranking suffer from the unreliability of the assumptions under which the initial text-based image search result is employed in the reranking process such as labeling of samples by user. A prototype-based reranking method [7] is proposed to address this problem in a supervised, but scalable fashion. By applying different meta rerankers to an image from the initial result, reranking scores are generated, which are then aggregated using a linear model to produce the final relevance score and the new rank position for an image in the reranked search result. Advantage is that the prototype-based reranking approach is hardly affected by outliers, since the weights for different images are learned from human-labeled data. But it has limitations on user as it needs labeling the samples by users.

3.2.3 Multimodal Sparse Coding

Instead of labeling, user click information can been used in image reranking, because clicks have been shown to more accurately describe the relevance of retrieved images to search queries. However, a critical problem for click-based methods is the lack of click data, since only a small number of web images have actually been clicked on by users. Therefore, for predicting image clicks multimodal hypergraph learning-based sparse coding method [6] is proposed, and applied the obtained click data to the reranking of images. In this way, click predictions are used to improve the performance. Limitation is that semantic spaces are not taken into account.

3.2.4 Query-Specific Semantic Signatures:

A major challenge in user click based approach is that the similarities of visual features do not well correlate with images semantic meanings which interpret users search intention. Recently researchers proposed to match images in a semantic space which used attributes or reference classes closely related to the semantic meanings of images as basis. However, learning a universal visual semantic space to characterize highly diverse images from the web is difficult and inefficient. A novel image re-ranking framework [2] is proposed, which automatically offline learns different semantic spaces for different query keywords. The visual features of images are projected into their related semantic spaces to get semantic signatures. At the online stage, images are re-ranked by comparing their semantic signatures. Advantage is semantic spaces show better performance along visual features.

Although in classification based, classifiers are effective, sufficient training data is required to achieve satisfactory performance because a lot of parameters must be estimated. However, in visual reranking, the training data usually obtained via PRF is noisy due to the imperfect text-based search result and insufficient, restricting the performance of this kind of method for real-world applications.

3.3 Graph Based Reranking

In the graph methods, graph is constructed to mine the relations between the images. The graph is constructed with the samples as the nodes and the edges between them being weighted by visual similarity. Then, reranking is performed on the graph by propagating the ranking scores through the edges. In graph methods, the relationships of all samples are represented by the graph. Therefore, the graph construction plays the key role in

this kind of method.

3.3.1 Visual Rank

It is a framework [13] to efficiently model similarity of Google image search results with graph. The framework casts the reranking problem as random walk on an affinity graph and reorders images according to the visual similarities. The final result list is generated via sorting the images based on graph nodes weights. The advantage of this is for quantifying the effectiveness of visual features by using bias vector visual rank is computed. It is not showing the relationship between the image similarity and likelihood for transaction more extensively is the disadvantages.

3.3.2 Bayesian Reranking

It is a framework [14] formulating the reranking formulating the reranking process as an energy minimization problem. The objective is to optimize the consistency of ranking scores over visually similar samples and minimize the inconsistency between the optimal list and the initial list. Thus, performance is significantly dependent on the statistical properties of top ranked search results which considers only feature vectors.

3.3.3 Multimodal Graph-Based

It is a web image search reranking approach [11] that explores multiple modalities in a graph based learning scheme. Different from the conventional methods that usually adopt a single modality or integrate multiple modalities into a long feature vector, this approach can effectively integrate the learning of relevance scores, weights of modalities, and the distance metric and its scaling for each modality into a unified scheme for reranking. Limitation is that semantic spaces are not taken into account,which narrow down the semantic gap between low level visual features and high level semantic meanings.

IV. SEMANTIC ATTRIBUTES

Semantic attributes can be used as a set of mid level semantic preserving concepts. Different from low level visual features every attribute has explicit meaning. Due to advantages of being semantic aware and easy to model, attributes have been studied recently and these are revealing their importance in various applications. Thus, attributes are expected to narrow down the semantic gap between low level visual features and high level semantic meanings.

Su et al. [16] propose to alleviate the semantic gap between visual words and high level concepts, focusing on polysemy phenomenon of particular visual words. Kumar et al. [17] define a set of binary attributes called similes for face recognition. Each attribute in similes are exclusively trained for one specific category. Zhang et al. [18] proposed an attribute augmented semantic hierarchy for content based image retrieval. It employs attributes to describe the multiple facts of concepts and hybrid feedbacks of attributes and images are collected.

V. HYPERGRAPH LEARNING

In a simple graph, samples are represented by vertices and an edge links the two related vertices. Learning tasks can be performed on a simple graph. Assuming that samples are represented by feature vectors in a feature space, an undirected graph can be constructed by using their pair wise distances, and graph-based semi-supervised learning approaches can be performed on this graph to categorize objects. It is noted that this simple graph cannot reflect higher-order information. Compared with the edge of a simple graph, a hyperedge in a

hypergraph is able to link more than two vertices. The advantage of hypergraph can be summarized that not only does it take into account pair wise relationship between two vertices, but also higher order relationship among three or more vertices containing grouping information.

VI. CONCLUSION

In this paper, various image retrieval methods are discussed and the reranking methods which are proposed by earlier researchers for the better development in the web image search are discussed. These methods are categorized in different reranking strategies depending on the used approaches such as clustering based reranking, classification based reranking and graph based reranking. The clustering based reranking consist of bag based approach in which visual features and user search intention is not considerd, while in Intent search with user search intention has the limitation of user burden. The classification based approaches such as active reranking, prototype based reranking and multimodal sparse coding are making use of visual features effectively, but the semantic meanings are not used, which are used in query specific semantic spaces for more effective reranking. Also the graph based methods like VisualRank, Bayesian reranking are analyzed, these methods make use of simple graphs for relationship among the images. Compared with the simple graph, hyperedges in the hypergraph are able to link more than two vertice i.e. comparing the relation between more than two entities. By making the use of semantic attributes and hypergraph learning, the new search engine can be developed which will improve accuracy as well as effectiveness of the re-ranking process by using attribute based features of images. From the analysis, it's come to know that attribute-assisted re-ranking is one of the best re-ranking methods.

REFERENCES

[1] Junjie Cai, Zheng-Jun Zha, Meng Wang, Shiliang Zhang, and Qi Tian; "An Attribute Assisted Reranking Model for Web Image Search" *IEEE Transactions on Image Procesing, Vol.24, No.1, pp. 261-272, January 2015.*

[2] Xiaogang Wang, Shi Qui, Ke Liu, and Xiaoou Tang, "Web Image Reranking Using Query Specific Semantic Signatures" *IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.36, No. 4, pp. 810-823, April 2014.*

[3] Xinmei Tian, Linjun Yang, Yijuan Lu, Qi Tian and Dacheng Tao, "Image Search Reranking With Hierarchical Topic Awareness", *IEEE Transactions On Cybernetics, Vol. 45, No. 10, pp.2177-2189,October 2015.*

[4] Zhong Ji, Yanwei Pang and Xuelong Li, "Relevance Preserving Projection and Ranking for Web Image Search Reranking*", IEEE Transactions On Image Processing, Vol. 24, No. 11, pp.4137-4147,November 2015.*

[5] Lingxi Xie, Qi Tian, Wengang Zhou, and Bo Zhang, "Heterogeneous Graph Propagation for Large-Scale Web Image Search", *IEEE Transactions On Image Processing, Vol. 24, No. 11, pp. 4287-4298,November 2015.*

[6] Jun Yu,Yong Rui and Dacheng Tao, "Click Prediction for Web Image Reranking Using Multimodal Sparse Coding" *IEEE Transactions On Image Processing, Vol. 23, No. 5, pp.2019-2037,May 2014.*

[7] Linjun Yang and Alan Hanjalic, "Prototype-Based Image Search Reranking", *IEEE Transactions On Multimedia, Vol. 14, No. 3, pp.871-882, June 2012.*

[8] Xiaoou Tang, , Ke Liu, Jingyu Cui, Fang Wen and Xiaogang Wang , "Intent Search: Capturing User Intention for One-Click Internet Image Search", *IEEE Transactions On Pattern Analysis and Machine intelligence, vol. 34, No. 7, pp.1342-1353,July 2012.*

[9] Lixin Duan, Wen Li, Ivor Wai-Hung Tsang, and Dong Xu, "Improving Web Image Search by Bag-Based Reranking", *IEEE Transactions On Image Processing, Vol. 20, No. 11,pp.3280-3290, November 2011.*

[10] Xinmei Tian, Dacheng Tao, Xian-Sheng Hua and Xiuqing Wu, "Active Reranking for Web Image Search", *IEEE Transactions On Image Processing,Vol.19, No. 3,pp.805-820, March 2010.*

[11] Meng Wang, Hao Li, Dacheng Tao, Ke Lu, and Xindong Wu ,"Multimodal Graph-Based Reranking for Web Image Search", *IEEE transactions On Image Processing, Vol. 21, No. 11, pp. 4649-4661,November 2012.*

[12] Xinmei Tian and Dacheng Tao, "Visual Reranking: From Objectives to Strategies", *IEEE Transaction on Multimedia, July, 2011.*

[13] F. Jing and S. Baluja, "VisualRank: Applying pagerank to large-scale image search," *IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 30, no. 11, pp.1877–1890, Nov. 2008.*

[14] X. Tian, L. Yang, J. Wang, X. Wu, and X. Hua, "Bayesian Visual Reranking*," IEEE Transaction on Multimedia, vol. 13, no. 4, pp. 639-652, Aug. 2012.*

[15] Kirti Yadav, Sudhir Singh,"Improving Web Image Search Re-Ranking Using Hybrid Approach," *International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 6, June 2014.*

[16] Y.Su and F. Jurie, "Improving image classification using semantic attributes," *Int. J. Comput. Vis, vol. 100, no.1, pp. 59–77, 2012.*

[17] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, "Attribute and simile classifiers for face verification," in Proc. IEEE Int. Conf. Comput. Vis., pp. 365–372, Sep./Oct. 2009

[18] H. Zhang, Z.-J. Zha, Y. Yang, S. Yan, Y. Gao, and T.-S. Chua, "Attribute augmented semantic hierarchy: Towards bridging semantic gap and intention gap in image retrieval," in Proc. ACM Conf. Multimedia, pp. 33– 42,2013.