Intent Search: Survey on Various Methods of Image Re-ranking

Payal P. Thakur, Prof. M. A. Kalyankar

Abstract **––** Image re-ranking is an effective way for improving the result of web-based image search. To refine the text-based image search result image search re-ranking is the best approach. In this Paper, we give the different methods of image re-ranking such as Query Specific Semantic Signature, Graph Based Learning, One Click Internet Image Search and Multimodal Sparse Coding For Click Prediction. And then we propose the One Click Internet Image Search, it requires the user to give only one click on a query image and images from a pool retrieved by text based search are re-ranked based on their visual and textual similarities to the query image. Adaptive similarity**,** Keyword Expansion, Image pool Image, Visual query expansion by using these four steps we have srearch and re-ranked the images.

Index Terms**-—Hypergraph, Image Search, Keyword Expansion, Re-ranking**

I. INTRODUCTION

Due to the large number of images on the web, image re-ranking has become an active challenging topic. Image re-ranking, as an effective way to improve the results of web-based image search. Many commercial Internet scale image search engines use only keywords as queries. Users type query keywords in the hope of finding a certain type of images. The search engine returns thousands of images ranked by the keywords extracted from the surrounding text. It is well known that text-based image search suffers from the ambiguity of query keywords. Fig. 1 shows the top ranked images from Bing image search using "apple" as query. They belong to different categories, such as "green apple," "red apple," "apple logo," and "iPhone" because of the ambiguity of the word "apple." The ambiguity issue occurs for several reasons. First, the query keywords' meanings may be richer than users' expectations. For example, the meanings of the word "apple" include apple fruit, apple computer, and apple iPod[1]. Second, the user may not have enough knowledge on the textual description of target images.

Manuscript received March 24, 2015

Payal P. Thakur, Department of Comp Sci & Engg, Dr. Sau K.G.I.E.T. Amravati University, Maharashtra, India,

Prof. M. A. Kalyankar, Department of Comp Sci & Engg, Dr. Sau K.G.I.E.T. Amravati University, Maharashtra, India,

Fig 1. Top-ranked images returned from Bing image search using "apple" as query.

II. METHODS USED FOR IMAGE RE-RANKING

A. Using Query-Specific Semantic Signature:

In this method[2] user asked to select a query image from the pool, the remaining images are re-ranked based on their visual similarities with the query image. A major challenge is that the similarities of visual features do not well correlate with images' semantic meanings which interpret users' search intention. Recently people 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. In this paper, we propose a novel image re-ranking framework, 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 obtained from the semantic space specified by the query keyword. The proposed query-specific semantic signatures significantly improve both the accuracy and efficiency of image re-ranking. The original visual features of thousands of dimensions can be projected to the semantic signatures as short as 25 dimensions.

B. Graph Based Learning Methods:

Graph-based learning methods have been widely used in the fields of image classification [21], ranking [22] and clustering. In these methods, a graph is built according to the given data, where vertices represent data samples and edges describe their similarities. The Laplacian matrix [23] is constructed from the graph and used in a regularization scheme. The local geometry of the graph is preserved during the optimization, and the function is forcefully smoothed on the graph. However, a simple graph-based method cannot capture higher order information. Unlike a simple graph, a hyper edge in a hypergraph links several (two or more) vertices, and thereby captures this higher-order information. Hyper graph learning has achieved excellent performance in many applications. A hypergraph-based image retrieval approach has been proposed.

C. One-Click Internet Image Search:

one-click query image contains four steps[1]:

1) The query image is categorized into one of the predefined adaptive weight categories which reflect users' search intention at a coarse level. Inside each category, a specific weight schema is used to combine visual features adaptive to this kind of image to better re-rank the text-based search result.

2) Based on the visual content of the query image selected by the user and through image clustering, query keywords are expanded to capture user intention.

3) Expanded keywords are used to enlarge the image pool to contain more relevant images.

4) Expanded keywords are also used to expand the query image to multiple positive visual examples from which new query specific visual and textual similarity metrics are learned to further improve content-based image re-ranking.

D. Multimodal Sparse Coding For Click Prediction:

Both strategies of early and late fusion of multiple features are used in this method through three main steps[8].

First, we construct a web image base with associated click annotation, collected from a commercial search engine. Second, we consider both early and late fusion in the proposed objective function. The early fusion is realized by directly concatenating multiple visual features, and is applied in the sparse coding term. Late fusion is accomplished in the manifold learning term. Finally, an alternating optimization procedure is conducted to explore the complementary nature of different modalities. The weights of different modalities and the sparse codes are simultaneously obtained using this optimization strategy.

In this paper we discuss the One click Internet Image Search

III. LITERATURE REVIEW

The key component of image re-ranking is to compute visual similarities reflecting semantic relevance of images. Many visual feature were developed for image search in recent years. Chum et al. [3] used RANSAC to verify the spatial configurations of local visual features and to purify the expanded image examples. However, it was only applicable to object retrieval. It required users to draw the image region of the object to be retrieved and assumed that relevant images contained the same object. Under the frame work of pseudo relevance feedback. Zhang et al. [4] proposed geometry preserving visual phases which captured the local and long range spatial layouts of visual words. One of the major challenges of content-based image retrieval is to learn the visual similarities which reflect the semantic relevance of images well. Image similarities can be learned from a large training set where the relevance of pairs of images is known. Krapac et al.[5] proposed the query-relative classifiers, which combined visual and textual information, to re-rank images retrieved by an initial text-only search. However, since users were not required to

select query images, the users' intention could not be accurately captured when the semantic meanings of the query keywords had large diversity. Deng et al. [6] learned visual similarities from a hierarchical structure defined on

semantic attributes of training images. Since web images are highly diversified, defining a set of attributes with hierarchical relationships for them is challenging. In general, learning a universal visual similarity metric for generic images is still an open problem to be solved. We conducted the first study that combines text and image content for image search directly on the Internet in [33], where simple visual features and clustering algorithms were used to demonstrate the great potential of such an approach. Following our intent image search work in and, a visual query suggestion method is developed in [34]. Its difference from and is that instead of asking the user to click on a query image for re-ranking, the system asks users to click on a list of keyword-image pairs generated offline using a data set from Flickr and search images on the web based on the selected keyword.

IV. METHODOLOGY

In this paper, we propose a one click Internet image search approach. we propose a novel Internet image search approach. It requires the user to give only one click on a query image and images from a pool retrieved by text based search are re-ranked based on their visual and textual similarities to the query image. We believe that users will tolerate one-click interaction, which has been used by many popular text-based search engines. In this paper we discuss that how to capture user intention from one-click query image. For that Following four steps are followed.

The flow chart of our approach is shown in Fig.2. The user first submits query keywords q. A pool of images is retrieved by text-based search (Fig. 2a). Then the user is asked to select a query image from the image pool. The query image is classified as one of the predefined adaptive weight categories. Images in the pool are re-ranked (Fig. 2b) based on their visual similarities (Fig. 2c) specified by the category to combine visual features. In the keyword expansion step (Fig. 2d), words are extracted from the textual descriptions (Fig.2d), images with the same candidate word may have a large diversity in visual content. Images assigned to the same cluster have higher semantic consistency since they have high visual similarity to one another and contain the same candidate word. Among all the clusters of different candidate words. The largest visual similarity to the query image is selected as visual query expansion. The image pool is enlarged through combining the original image pool retrieved by the query keywords q provided by the user and an additional image pool retrieved by the expanded keywords q (Fig. 2f), Images in the enlarged pool are re-ranked using the learned query-specific visual and textual similarity metrics (Fig.2g)

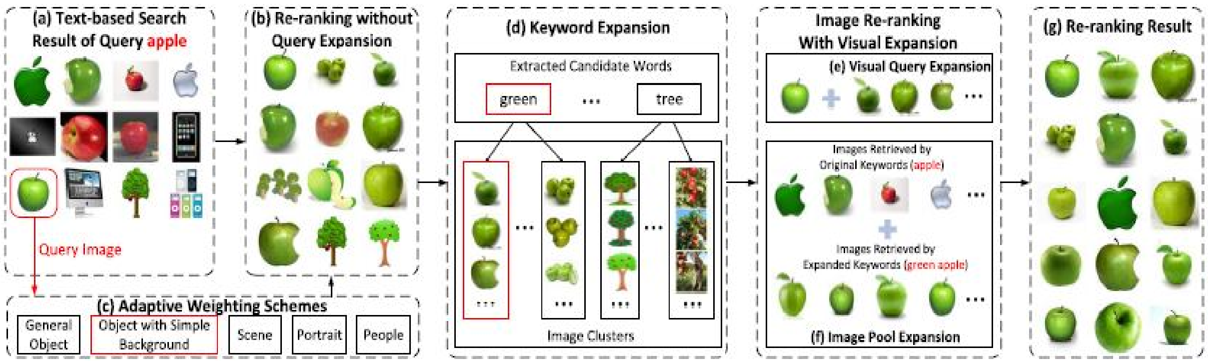

Fig 2. Steps For One Cick Internet Image Searching

A. Adaptive similarity:

For describing a different aspects of images we design a set of visual features. The similarity between two images is usually expressed by aggregating the similarities between corresponding local features. How to integrate various visual features to compute the similarities between the query image and other images is an important problem. Adaptive Similarity is proposed, motivated by the idea that a user always has specific intention when submitting a query image. For example, if the user submits a picture with a big face in the middle, most probably he/she wants images with similar faces and using face-related features is more appropriate. In our approach, the query image is first categorized into one of the predefined adaptive weight categories, such as "portrait" and "scenery." Inside each category, a specific pretrained weight schema is used to combine visual features adapting to this kind of images to better re-rank the text-based search result. This correspondence between the query image and its proper similarity measurement reflects the user intention. This initial re-ranking result is not good enough and will be improved by the remaining three steps.

B. Keyword expansion:

Automatic query expansion (AQE) is a tried and tested method in web search for tackling the vocabulary problem by adding related words to the search query and thus increase the likelihood that appropriate documents are contained in the result. Query keywords input by users tend to be short and some important keywords may be missed because of users' lack of knowledge on the textual description of target images. In our approach, query keywords are expanded to capture users' search intention, inferred from the visual content of query images, which are not considered in traditional keyword expansion approaches. A word w is suggested as an expansion of the query if a cluster of images are visually similar to the query image and all contain the same word w.1 The expanded keywords better capture users' search intention since the consistency of both visual content and textual description is ensured.

C. Image pool Image:

Normal way of image retrieval is the text based image retrieval. The image pool retrieved by text-based search accommodates images with a large variety of semantic meanings and the number of images related to the query image is small. Generally the process of image search includes searching of image based on keyword typed. When query is entered in the search box for searching the image, it is forwarded to the server that is connected to the internet. The server gets the URL's of the images based on the tagging of the textual word from the internet and sends them back to the client. In this case, re-ranking images in the pool is not very effective. Thus, more accurate query by keywords is needed to narrow the intention and retrieve more relevant images. A naive way is to ask the user to click on one of the suggested keywords given by traditional approaches only using text information and to expand query results like in Google "related searches." This increases users' burden. Moreover, the suggested keywords based on text information only are not accurate to describe users' intention. Keyword expansions suggested by our approach
using both visual and textual information better capture users' intention. They are automatically added into the text query and enlarge the image pool to include more relevant images.

D. Visual query expansion:

One query image is not diverse enough to capture the user's intention. In Step 2, a cluster of images all containing the same expanded keywords and visually similar to the query image are found. They are selected as expanded positive examples to learn visual and textual similarity metrics, which are more robust and more specific to the query, for image re-ranking. Compared with the weight schema in Step 1, these similarity metrics reflect users' intention at a finer level since every query image has different metrics.

Different from relevance feedback, this visual expansion does not require users' feedback. All four of these steps are automatic with only one click in the first step without increasing users' burden. This makes it possible for Internet scale image search by both textual and visual content with a very simple user interface.

V. CONCLUSION

In this paper, we propose a novel Internet image search approach which only requires one-click user feedback Without additional human feedback, textual and visual expansions are integrated to capture user intention. Expanded keywords are used to extend positive example images and also enlarge the image pool to include more relevant images.

VI. REFERENCE

[1] Xiaoou Tang, Ke Liu, Jingyu Cui, Fang Wen, and Xiaogang Wang," Intent Search: Capturing User Intention for One-Click Internet Image Search", VOL. 34, NO. 7, JULY 2012

[2] Xiaogang Wang, Shi Qiu, Ke Liu, and Xiaoou Tang, "Web Image Re-Ranking Using Query-Specific Semantic Signatures", VOL. 36, NO. 4, APRIL 2014

[3] O. Chum, J. Philbin, J. Sivic, M. Isard, and A. Zisserman, "Total Recall: Automatic Query Expansion with a Generative Feature Model for Object Retrieval," Proc. IEEE Int'l Conf. Computer Vision, 2007.

[4] Y. Zhang, Z. Jia, and T. Chen, "Image Retrieval with Geometry-Preserving Visual Phrases," Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition, 2011.

[5] J. Krapac, M. Allan, J. Verbeek, and F. Jurie, "Improving Web Image Search Results Using Query-Relative Classifiers," Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition, 2010.

[6] J. Deng, A.C. Berg, and L. Fei-Fei, "Hierarchical Semantic Indexing for Large Scale Image Retrieval," Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition, 2011.

[7] B. Luo, X. Wang, and X. Tang, "A World Wide Web Based Image Search Engine Using Text and Image Content Features," Proc. IS&T/SPIE Electronic Imaging, Internet Imaging IV, 2003.

[8] Jun Yu, Yong Rui, and Dacheng Tao, "Click Prediction for Web Image Re-ranking Using Multimodal Sparse Coding", VOL. 23, NO. 5, MAY 2014