International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 - 5161

Available at http://inpressco.com/category/ijcet

Review Article

Web-based Image Search using Image Re-Ranking Technique: A Review Paper

Harshil Jain[Ȧ*], Riya Lodha[Ȧ] and Khushali Deulkar[Ȧ]

[Ȧ]Computer Engineering Department, D.J.Sanghvi College of Engineering, Mumbai University, Mumbai, India

Accepted 10 Sept 2014, Available online 01 Oct 2014, **Vol.4, No.5 (Oct 2014)**

Abstract

Image re-ranking is a useful method for web-based image search. The search based on only keywords queried by the users is not efficient and results in imprecise output. The web-based image search used by Bing and Google uses image re-ranking. In image re-ranking, users' intention is captured by one-click on the query image. This helps in providing better search results to the users. In this paper, we review the method in which a query keyword is first used to retrieve a plethora of images based on the keyword. Image re-ranking framework automatically learns different semantic spaces offline for different query keywords. To get semantic signatures for images, their visual features are projected into their related semantic spaces. Images are re-ranked by comparing their semantic signatures and the query keyword during the online stage. The query-specific semantic signatures, in the reviewed paper, significantly improve both the accuracy and efficiency of the re-ranking process. Hence, it is proved to be a better method than the conventional web-based image search techniques.

Keywords: *image re-ranking, framework, query image, query keyword, image search, semantic signature.*

1. Introduction

Web image search engines use keywords as queries and search images based on the text associated with them. It is difficult for users to accurately describe the visual content of target images only using keywords and hence text-based image search suffers from the ambiguity of query keywords. For example, using apple as a query keyword, the retrieved images belong to different categories, such as apple laptop, apple logo, apple fruit. To capture users' search intention, additional information has to be used in order to solve the ambiguity. Text-based keyword expansion is one way to make the textual description of the query more detailed. Existing methods find either synonyms or other linguistic-related words from thesaurus. However, the intention of users can be highly diverse and cannot be accurately captured by these expansions, even with the same query keywords. Content-based image retrieval with relevance feedback is widely used in order to solve this ambiguity. Users are required to select multiple relevant and irrelevant image examples and the visual similarity metrics are learned through online training from them. Images are re-ranked based on the learned visual similarities. However, for web-scale commercial systems, users' feedback has to be limited to the minimum without online training.

In the method reviewed in this paper, a query keyword is first used to retrieve a set of images based on the keyword. Then the user is asked to pick an image from these images. Also, the rest of the images are ranked based on their visual similarities. The major challenge is the correlation of similarities of visual features and images' semantic meaning, which are needed to interpret users' intention to search. Recently, it has been proposed to match images in a semantic space that used attributes or reference classes closely related to the semantic meanings of images as basis. However, characterizing the highly diverse images from the web is difficult because it is impossible to learn a universal visual semantic space.

2. Conventional Techniques for Image Re-Ranking

Computing the visual similarities that reflect the semantic relevance of images is the key component of image re-ranking. Many visual features have been developed in recent years. However, the effective low-level visual features are different for different query images. Therefore, Cui *et al*. classified query images into eight predefined intention categories and gave different feature weighting schemes to different types of query images. But it was difficult for the eight weighting schemes to cover the large diversity of all the web images. It was also likely for a query image to be classified to a wrong category. Query-specific semantic signature was first proposed in in order to reduce the semantic gap.

There is a lot of work on using visual features to re-rank images retrieved by initial text-only search, however, without requiring users to select query images. Jing and Baluja proposed VisualRank to analyze the visual link structures of images and to find the visual themes for re-ranking. Cai *et al*. re-ranked images with attributes which were manually defined and learned from manually labeled

*Corrresponding author: **Harshil Jain**

training samples. These approaches assumed that there was one major semantic category under a query keyword. Images were re-ranked by modeling this dominant category with visual and textual features.

2.1 Re-Ranking without Query Images

Query-specific semantic signature can be applied to image re-ranking without selecting query images. This application also requires the user to input a query keyword. But it assumes that images returned by initial text-only search have a dominant topic and images belonging to that topic should have higher ranks. Existing approaches typically address two issues: (1) how to compute the similarities between images and reduce the semantic gap; and (2) how to find the dominant topic with ranking algorithms based on the similarities. The query-specific semantic signature is effective in this application since it can improve the similarity measurement of images.

The query-specific semantic signature is also effective in this application, where it is crucial to reduce the semantic gap when computing the similarities of images. Due to the ambiguity of query keywords, there may be multiple semantic categories under one keyword query. These approaches cannot accurately capture users' search intention without query images selected by users. Recently, for general image recognition and matching, there have been a number of works on using projections over predefined concepts, attributes or reference classes as image signatures. The classifiers of concepts, attributes, and reference classes are trained from known classes with labeled examples. But the knowledge learned from the known classes can be transferred to recognize samples of novel classes which have few or even no training samples. Since these concepts, attributes, and reference classes are defined with semantic meanings, the projections over them can well capture the semantic meanings of new images even without further training.

Rasiwasia *et al*.mapped visual features to a universal concept dictionary for image retrieval. Attributes with semantic meanings were used for object detection and recognition, face recognition, action recognition, image search and 3D object retrieval.

Lampert *et al*. predefined a set of attributes on an animal database and detected target objects based on a combination of human-specified attributes instead of training images. Parikh and Grauman proposed relative attributes to indicate the strength of an attribute in an image with respect to other images. Some approaches transferred knowledge between object classes by measuring the similarities between novel object classes and known object classes (called reference classes). For example, Torresani *et al*. proposed an image descriptor which was the output of a number of classifiers on a set of known image classes, and used it to match images of other unrelated visual classes.

3. A Conventional Image Re-Ranking Framework

Online image re-ranking limits users' effort to just one-click feedback is an effective way to improve search results and its interaction is simple enough. Major web image search engines have adopted this strategy. Its diagram is shown in Fig. 1. Given a query keyword input by a user, a pool of images relevant to the query keyword is retrieved by the search engine according to a stored word-image index file. Usually the size of the returned image pool is fixed, e.g., containing 1000 images.

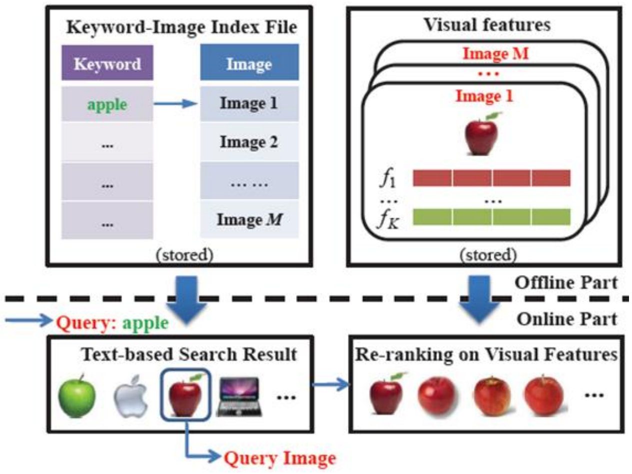

Fig. 1 The conventional image re-ranking framework

The user is asked to select a query image from the pool. This image reflects the user's search intention and the remaining images in the pool are re-ranked based on their visual similarities with the query image. The word-image index file and visual features of images are pre-computed offline and stored. The main online computational cost is on comparing visual features. To achieve high efficiency, the visual feature vectors need to be short and their matching needs to be fast. Some popular visual features are in high dimensions and efficiency is not satisfactory if they are directly matched.

In the current approaches, all the concepts/ attributes/ reference-classes are universally applied to all the images and they are manually defined. They are more suitable for offline databases with lower diversity (such as animal data-bases and face databases), since image classes in these databases can share similarities in a better way. A huge set of concepts or reference classes are required to model all the web images, which is impractical and ineffective for online image re-ranking. Intuitively, only a small subset of the concepts is relevant to a specific query. Many concepts irrelevant to the query not only increase the computational cost but also deteriorate the accuracy of re-ranking. However, how to find such relevant concepts automatically and use them for online web image re-ranking was not well explored in the conventional methods.

4. New Image Re-Ranking Framework

The new image re-ranking framework focusses on the semantic signatures associated with the images. These semantic signatures are derived from the visual features associated with the images but are much shorter than the visual features.

The diagram of the approach is shown in Fig. 2. It has offline and online parts. At the offline stage, the reference classes (which represent different concepts) related to

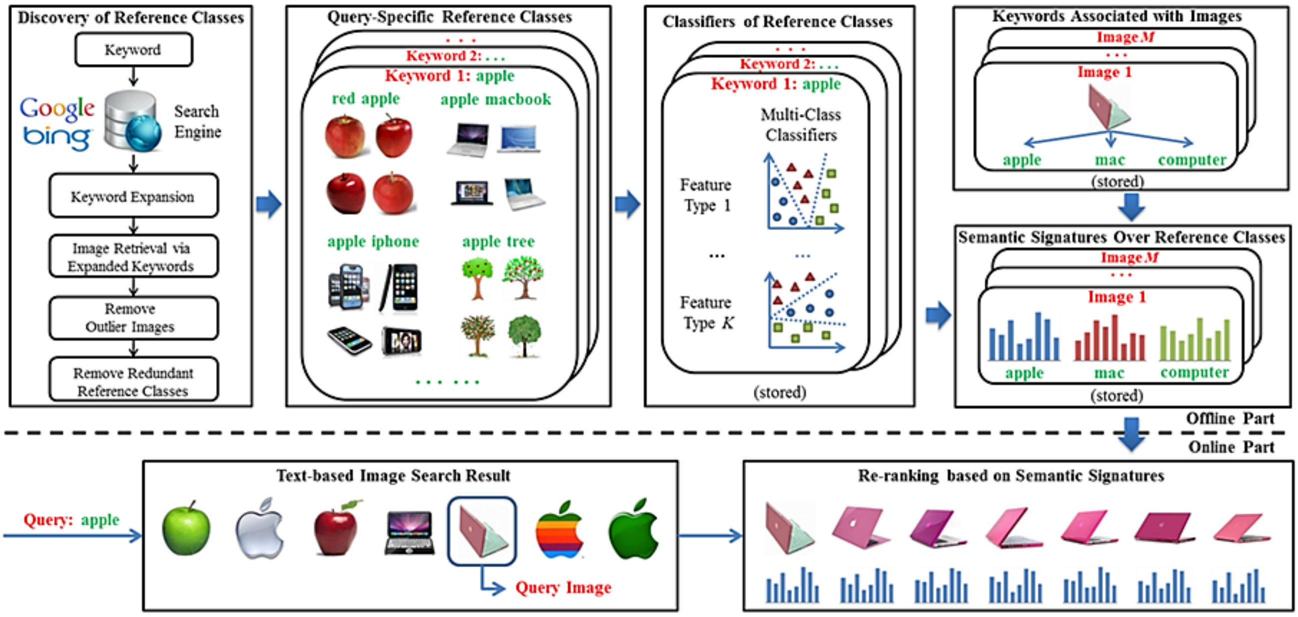

Fig. 2 Diagram of new image re-ranking framework

query keywords are automatically discovered and their training images are automatically collected in several steps. For a query keyword (for example apple), automatic selection of a set of most relevant keyword expansions (such as red apple and apple macbook) is performed utilizing both textual as well as visual information. This set of keyword expansions defines the reference classes for the query keyword.

In order to automatically obtain the training examples of a reference class, the keyword expansion (e.g., red apple) is used to retrieve images by the search engine based on textual information again. Images retrieved by the keyword expansion (red apple) are much less diverse than those retrieved by the original keyword (apple). The retrieved top images are used as the training examples of the reference class after automatically removing outliers. Some reference classes (such as apple laptop and apple macbook) have similar semantic meanings and their training sets are visually similar. The redundant reference classes are removed in order to improve the efficiency of online image re-ranking.

To better measure the similarity of semantic signatures, the semantic correlation between reference classes is estimated with a web-based kernel function. For each query keyword, its reference classes forms the basis of its semantic space. A multi-class classifier on visual and textual features is trained from the training sets of its reference classes and stored offline. Under a query keyword, the semantic signature of an image is extracted by computing the similarities between the image and the reference classes of the query keyword using the trained multiclass classifier. If there are K types of visual/textual features, such as color, texture, and shape, one could combine them together to train a single classifier, which extracts one semantic signature for an image. A separate classifier for each type of feature can also be trained.

Then, the K classifiers based on different types of features extract K semantic signatures, which are combined at the later stage of image matching.

An image may be associated with multiple query key-words, which have different semantic spaces according to the word-image index file. Therefore, it may have different semantic signatures. The query keyword input by the user decides which semantic signature to choose. As an example shown in Fig. 2, an image is associated with three keywords apple, mac and computer. When using any of the three keywords as query, this image will be retrieved and re-ranked. However, under different query keywords, different semantic spaces are used. Therefore an image could have several semantic signatures obtained in different semantic spaces. They all need to be computed and stored offline.

At the online stage, the search engine, according to the query keyword, retrieves a pool of images. Since all the images in the pool are associated with the query keyword according to the word-image index file, they all have pre-computed semantic signatures in the same semantic space specified by the query keyword. Once the user chooses a query image, these semantic signatures are used to compute image similarities for re-ranking. The semantic correlation of reference classes is incorporated when computing the similarities.

The conventional framework compares images based upon their visual features. The length of these visual features is much longer than that of the semantic signatures used in the new framework. Hence, the computational cost is higher. Compared with the conventional image re-ranking diagram in Fig. 1, the new approach is much more efficient at the online stage, because the main online computational cost is on comparing semantic signatures and the lengths of semantic signatures are much shorter than those of low-level visual features.

Conclusions

In this paper, we have reviewed an Internet based image search approach. We have also discussed the conventional web-based image search techniques and pointed out their shortcomings. The reviewed image re-ranking framework overcomes the shortcomings of the previous methods and also significantly improves both the accuracy and efficiency of the re-ranking process. It captures users' intention using a query image. It learns query-specific semantic spaces to significantly improve the effectiveness and efficiency of online image re-ranking. The visual features of images are projected into their related semantic spaces automatically learned through keyword expansions offline. The extracted semantic signatures are shorter than the original visual features.

In future work, image re-ranking can be further improved by incorporating other metadata and log data along with the textual and visual features for finding the keyword expansions used for defining the reference classes. The log data of user queries provides useful co-occurrence information of keywords for keyword expansion. Finally, in order to further improve the quality of re-ranked images, they should be re-ranked not only by content similarity but also by the visual quality of the images.

References

X. Tang, K. Liu, J. Cui, F. Wen, and X. Wang, (2012), Intent Search: Capturing User Intention for One-Click Internet Image Search, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 34, no. 7, pp. 1342-1353.

J. Cui, F. Wen, and X. Tang, (2008), Real Time Google and Live Image Search Re-Ranking, Proc. 16th ACM Int'l Conf. Multimedia.

Xiaogang Wang; Shi Qiu; Ke Liu; Xiaoou Tang, (2014), Web Image Re-Ranking UsingQuery-Specific Semantic Signatures, Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.36, no.4, pp.810,823,doi: 10.1109/TPAMI.2013.214.

X. Wang, K. Liu, and X. Tang, (2010), Query-Specific Visual Semantic Spaces for Web Image Re-Ranking, Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR).

Y. Jing and S. Baluja, (2008), Visual Rank: Applying Page Rank to Large-Scale Image Search, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 30, no. 11, pp. 1877-1890.

J. Cai, Z. Zha, W. Zhou, and Q. Tian, (2012), Attribute-Assisted Reranking for Web Image Retrieval, Proc. 20th ACM Int'l Conf. Multimedia.

N. Rasiwasia, P.J. Moreno, and N. Vasconcelos, (2007), Bridging the Gap: Query by Semantic Example, IEEE Trans. Multimedia, vol. 9, no. 5, pp. 923-938.

C. Lampert, H. Nickisch, and S. Harmeling, (2009), Learning to Detect Unseen Object Classes by Between-Class Attribute Transfer, Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR).

D. Parikh and K. Grauman, (2011), Relative Attributes, Proc. Int'l Conf. Computer Vision (ICCV).

L. Torresani, M. Szummer, and A. Fitzgibbon, (2010), Efficient Object Category Recognition using Classemes, Proc. European Conf. Computer Vision (ECCV).