Volume 4, Issue 11, November 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

A Review on Content Based Image Re-Ranking using Semantic Analysis

Rachana C. Patil*, Asst. Prof. K. P. Gaikwad
Computer Engg. Dept., S.N.D.C.O.E.R.C, Yeola, Dist-Nasik Savitribai Phule University,
Pune, Maharashtra, India

Abstract—A CBIR is an effective way to improve the relevancy rate for image search. Image search engines such as Google and Bing mostly use keyword queried by the user and they rely on surrounding text for searching images. But sometime it is not efficient and results in imprecise output therefor ambiguity of query images is difficult to describe accurately by using keyword for e.g. Sony is query keyword then categories can be "Sony TV", "Sony laptop", etc. Given a query keyword, lots of images are first retrieved by the search engine with the help of text information. When user select a query image from the lots of images then the remaining images are re-ranked based on their visual features, A major challenge is that, semantic meanings of images do not well correlated with the similarities of visual features, which interpret user's search intention. Another challenge is, to characterize highly diverse images by learning a visual semantic space is difficult and insufficient.

In this paper, we review the framework in which, a query keyword is first used to retrieve a pool of images based on the given keyword. Framework of image re-ranking automatically learns different visual semantic space for different query keywords by using keyword expansions, at offline stage. For getting the semantic signatures of images, the visual as well as textual features of images are then projected into their related semantic spaces. At online stage, by comparing the semantic signatures of images which are obtained from the visual semantic space with query keywords, images are re-ranked. This concept is significantly improved accuracy as well as efficiency of the re-ranking process. Hence, it is called to be a better approach than the conventional image search techniques.

Keywords— Image search, semantic space, semantic signature, image re-ranking, keyword expansion

I. INTRODUCTION

Query keywords are mostly used by web scale image search engine which search the images based on the surrounding text. It is most commonly known that they have pain from the doubtful results of given query keyword. To accurately describe the visual content or query images by using the key ward then it is difficult for users. Hence, in this case, searching the image based on text suffers from the ambiguity of query keyword. For example, using a query as "Sony", then the different categories of images are retrieved, such as "Sony TV", "Sony laptop", "Sony as a name of person", "Sony handset" and "Sony digital camera". To search the result as per the user's intension, additional information has to be used in order to solve the ambiguity. There is one way to make the textual description of the query more detailed, known as text-based keyword expansion. Nowadays, text-based image retrieval systems are present, in which by using text-based keywords, images are manually annotated. Instead of looking into the contents of the image, if we query by a keyword then they matches the query to the keywords present in the database by using this system.

This technique has its some disadvantages:

a) It is not feasible to manually annotate the huge collection of images.
b) By using keywords, the important feature that is present in an image cannot be described completely.

These are the disadvantages of text-based image retrieval techniques which are overcome by new technique known as Content-Based Image Retrieval (CBIR). CBIR is a technology that helps to classify digital image archives according to their visual content. To distinguish the different regions present in an image based on their similarity in colour, pattern, texture, shape, etc. in that case this system is used and decides the similarity between two images by estimate the closeness of these different regions. By using CBIR low-level image features can be automatically extracted from the images and to some extent they describe the image in a more detail compared to the text-based approach.

The existing method is used to find either synonyms or other linguistic-related words from thesaurus. But the user's intension can be highly diverse and cannot be accurately captured by these expansions, even with the same query keywords. In order to solve this ambiguity, the concept of contents-based image retrieval with relevance feedback is widely used. The CBIR is a technique which uses visual contents to search images from large scale database according to user's interests. The images contents are available in various forms such as shape, colour, texture etc. .CBIR is more efficient technique for image retrieval.

In the method reviewed in this paper, based on given query keyword firstly the pool of image are retrieved. Then the user is asked to pick an image from these images. Also, the remaining images are ranked based on their visual similarities. The correlation between the similarities of visual features and semantic meaning of images, which are required for searching the result as per user's intention .nowadays, it has been proposed to match images in a semantic space which used reference classes that are closely related to the semantic meanings of images.

II. OUR APPROACH

In this paper, we proposed the novel framework instead of constructing a universal concept dictionary for web image re-ranking. For different query keywords it learns different visual semantic spaces as individually and automatically. We know that, the query keyword which is provided by the user can be significantly narrowed down when the semantic space related to the images to be re-ranked. We can more accurately model the images to be re-ranked by using the query specific visual semantic spaces; hence they have to eliminate other unlimited number of non-relevant concepts.

To get the semantic signatures the visual features of images are projected into their related visual semantic spaces. Images are re-ranked by comparing the semantic signatures which is obtained from the visual semantic space of the query keyword, at the online stage. A large scale benchmark database is also introduced in this paper for the performance evaluation of image re-ranking. By using 120 query keywords, the Bing image search retrieved 120,000 labelled images of around 1500 categories (which are defined by semantic concepts). Since, by our approach 20% - 35% relative improvement has been achieved on re-ranking precision with the help of benchmark database with much faster speed.

III. RELATED WORK

To calculate the image similarity, the content based image retrieval uses visual features. Hence, to learn visual similarity metrics by using relevance feedback to capture the user's search intention. Therefore, to select multiple relevant as well as irrelevant images, it required more users' effort and also needs online training. Cui et al [5, 4] proposed an image re-ranking approach that limited user's effort to just one-click feedback. Recently, the popular web scale image search engines such as Google and Bing have been introduced for image re-ranking approach, as the "find similar images" function.

If we want to capture the visual similarities between images then the key component of image re-ranking has been used. In recent years, many image features [6, 2, 8] have been developed. However, for different query images, suppose one category of image having low-level visual features that are effective but it may not work well for another. To address this, Cui et al [5, 4] approach is used to classify the query images into 8 predefined intension categories and gave different feature weighting schemes top different types of query image.

Recently, there are number of works done by using predefined concepts or attributes as image signature for general image recognition and matching. Rasiwasia et al [9] approach mapped visual features to a universal concept dictionary. Lampert et al [7] approach used predefined attributes with semantic meanings to detect novel object classes. By measuring the similarities between novels object classes (called reference classes), some approaches [1, 11, 10] are used to transferred knowledge between object classes. All these concepts were applied to all the images and their training data was manually selected. For offline databases, they are more suitable. A huge set of concepts or reference classes are required if we want to model the web images.

IV. CONVENTIONAL IMAGE RE-RANKING FRAMEWORK

To improve the image search results, the effective way is generated which are known as online image re-ranking concepts [5, 3]. The re-ranking strategy has been adopted by the most of the image search engine [5].

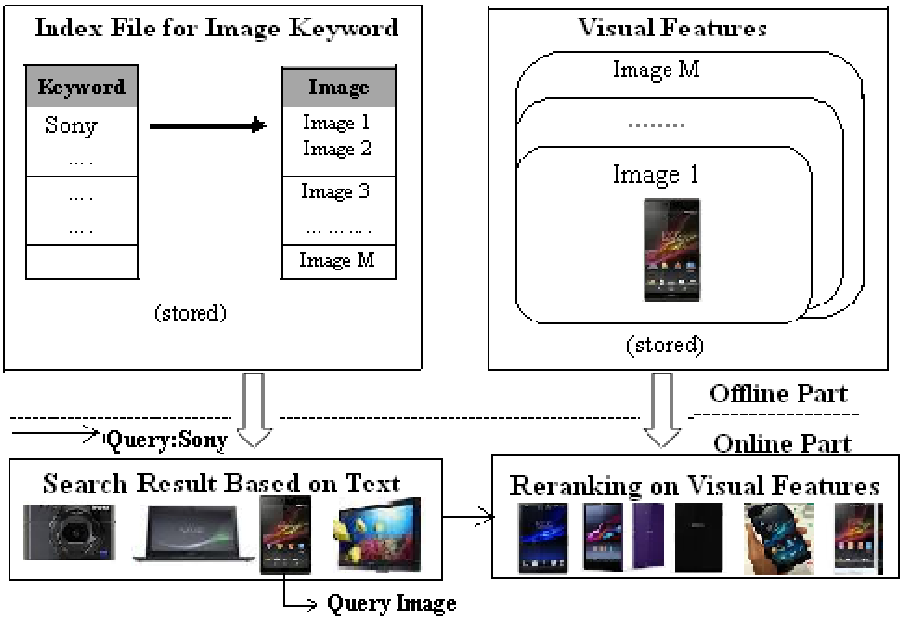

Fig 1 Conventional Image Re-ranking Framework

As shown in figure 1, the user gives the input as a query keyword, consequently the index file which contains the word-image, then by using the search engine they retrieved the images which are relevant to the query keyword from given lots of images. As per the user's search intension, select the query image from the lots of images by asking a user then re-ranked the remaining images based on their visual similarities with the query image. At the offline stage, the visual features of images are computed and it is stored by search engine. It is necessary; the visual feature vectors need to be short and their matching needs to be fast, in order to achieve high efficiency.

In the current approaches, all the reference classes are universally applied to all the images and these are manually defined. For lower diversity such as face databases these concepts are more suitable, since image classes in these databases can share similarities in a better way. For images which are impractical and ineffective for online image re-ranking, a huge set of concepts or reference classes are required to model all the web images. Naturally, small subset of the concepts is relevant to a specific query. Most of the concepts are irrelevant to the query which is not only increase the computational cost but also decline 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.

V. NEW IMAGE RE-RANKING FRAMEWORK

The diagram of new image re-ranking framework is shown in figure 2. The reference classes which are used to represent different semantic concepts of query keywords are automatically discovered at the offline stage. Therefore, by considering both textual and visual information, a set of most appropriate or relevant keyword expansions are automatically selected for a given query keyword. For example, "Sony" as a keyword then a set of most relevant keyword expansions such as "Sony laptop", "Sony handsets", "Sony TV" and "Sony digital camera" are automatically selected. The reference classes of a query keyword are defined by this set of keyword expansions. If we want to automatically obtain the training example of reference class then the keyword expansion (e.g. "Sony laptop") is used to retrieve images by the search engine. Instead of images are retrieved by the original keyword, if those are retrieved by the keyword expansion then these are much less different. The top retrieved images are used as the training examples of the reference class after removing the outliers automatically. There are some reference classes which have similar semantic meanings as well as their training sets are visually similar. For example, "Sony handset" and "Sony mobile". Therefore, redundant reference classes are removed, in order to improve the efficiency of online image re-ranking.

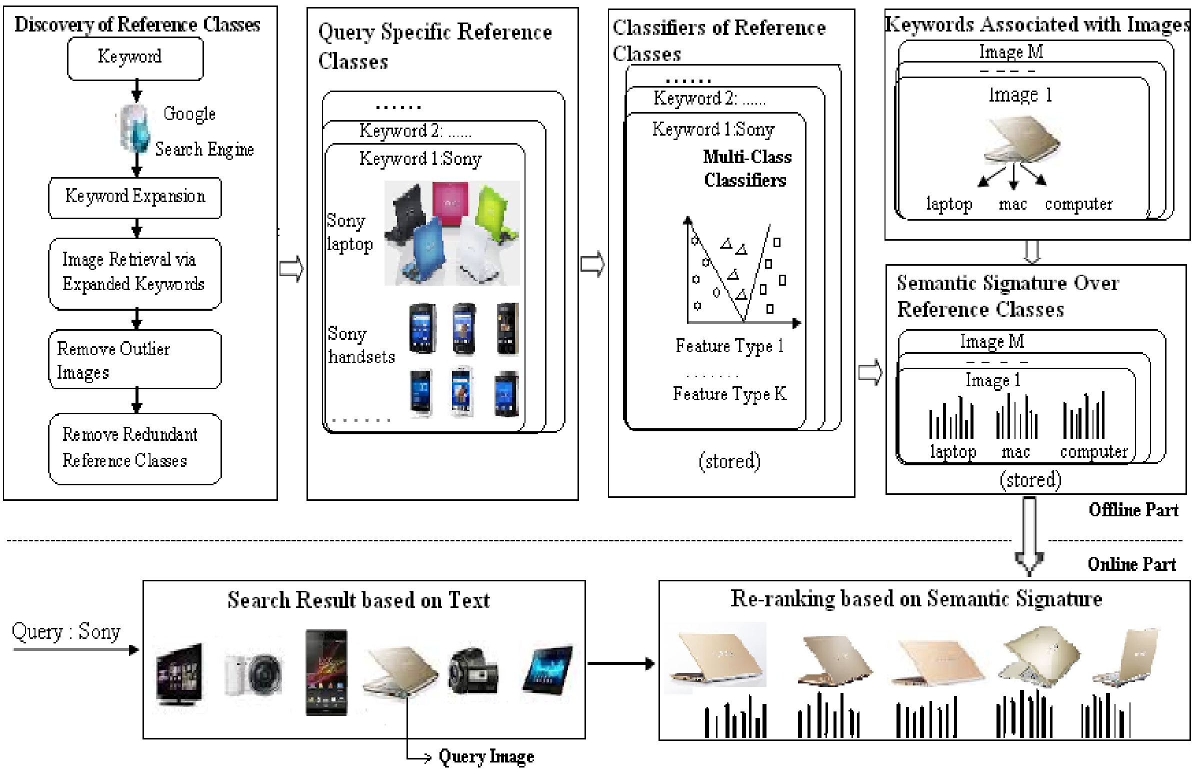

Fig 2 New image re-ranking framework

A multi-class classifier on low level visual features for each query keyword is trained from the training sets of its reference classes which is stored offline. If there are multiple types of visual features then one could combine them to train the single classifier. Due to this, it can increase the re-ranking accuracy but will also increase storage as well as reduce the online matching efficiency because of the increased size of semantic signatures. Most of the time, an images are relevant to the multiple query keywords. Therefore it could have several semantic signatures which are obtained in different semantic spaces. Each image which is stored in the database is associated with a few relevant keywords, according to the word image index file. By computing the visual similarities between the image and the reference classes of the keyword, a semantic signature of the image is extracted for each relevant keyword. There are N semantic signatures to be computed if an image has N relevant keywords, and stored offline.

According to the query keyword, the search engine retrieves a set of images, at the online stage. Hence all the set of images are associated with the given query keyword according to the word- image index file. As specified by the query keyword, all images have pre-computed semantic signatures in the same semantic space. When the user chooses a query image then to compute image similarities for re-ranking, the semantic signature concept is used. In conventional framework the images are compared based upon their visual features. The length of visual features available in conventional framework is longer than that of the semantic signatures which are used in new framework. Differentiate the conventional image re-ranking diagram in figure 1 with the new approach as shown in figure 2 is very much efficient at the online stage. Because the lengths of semantic signatures as well as online computational cost by comparing semantic signatures are much shorter than other low-level visual features.

VI. METHODOLOGY

Algorithm:-

1. In this paper, there are 2 parts online and offline parts.
2. At offline stage, by using a query keyword their number of reference classes are defined which representing different concepts related to that query keywords which are automatically discovered. There are a set of most relevant keyword expansions are available (such as “Sony laptop” and “Sony MacBook”) for a query keyword (e.g. “Sony”), are automatically selected using both textual and visual information.
3. For different keywords, the number of keyword expansions defines reference classes.
4. On the basis of training set of reference classes, a multi class classifier is accomplished.
5. If there are k types of visual and textual features are available such as color, shape, texture then they can combine them to accomplish a single classifier.
6. At online stage, according to query keyword pools of images are retrieved. When user selects query image then concept of semantic signatures are used to calculate similarities of image with pre-computed semantic signatures.

VII. CONCLUSION

In this paper, we have reviewed a novel image re-ranking framework by learning the query-specific semantic spaces it helps to significantly improve both the effectiveness and efficiency of online image re-ranking. At offline stage, through keyword expansions the visual features of images are projected into their related visual semantic spaces automatically. We have also discussed the conventional image search techniques and find out their drawbacks. The reviewed image re-ranking framework overcomes the drawbacks of existing method and improves search result as per user’s intention.

REFERENCES

[1] E. Bart and S. Ullman. Single-example learning of novel classes using representation by similarity. In Proc. BMVC, 2005.
[2] Y. Cao, C. Wang, Z. Li, L. Zhang, and L. Zhang. Spatial-bag-of features. In Proc. CVPR, 2010.
[3] B. Luo, X. Wang, and X. Tang. A world wide web based image search engine using text and image content features. In Proceedings of the SPIE Electronic Imaging, 2003.
[4] J. Cui, F. Wen, and X. Tang, “Intent Search: Interactive on-Line Image Search Re-Ranking,” Proc. 16th ACM Int’l Conf. Multimedia, 2008.
[5] J. Cui, F. Wen, and X. Tang, “Real Time Google and Live Image Search Re-Ranking,” Proc. 16th ACM Int’l Conf. Multimedia, 2008.
[6] D. Lowe. Distinctive image features from scale-invariant keypoints. Int’l Journal of Computer Vision, 2004.
[7] C. Lampert, H. Nickisch, and S. Harmeling. Learning to detect unseen object classes by between-class attribute transfer. In Proc. CVPR, 2005.
[8] J. Philbin, M. Isard, J. Sivic, and A. Zisserman. Descriptor Learning for Efficient Retrieval. In Proc. ECCV, 2010.
[9] N. Rasiwasia, P. J. Moreno, and N. Vasconcelos. Bridging the gap: Query by semantic example. IEEE Trans. on Multimedia, 2007.
[10] M. Rohrbach, M. Stark, G. Szarvas, I. Gurevych, and B. Schiele. What helps wherevand why? Semantic relatedness for knowledge transfer. In Proc. CVPR, 2010.
[11] Q. Yin, X. Tang, and J. Sun. An associate-predict model for face recognition. In Proc. CVPR, 2011.

