International Journal of Contemporary Research in Computer Science and Technology (IJCRCST) e-ISSN: 2395-5325
Volume 2, Issue 3 (March'2016)

AN INTELLECTUAL BASED RE-RANKING FOR IMAGE SEARCH IN WEB

S.Aiswarya,
Assistant Professor,
Department of Computer science and engineering,
Velammal Institute of Technology,Chennai,India.

B.Gayatri,
Student,
Department of Computer science and engineering,
Velammal Institute of Technology,Chennai,India.

S.Gayathri,
Student,
Department of Computer Science and Engineering,
Velammal Institute of Technology,Chennai,India.

M.Archana Devi,
Student,
Department of Computer science and engineering,
Velammal Institute of Technology,Chennai,India.

Abstract: Image search is a specialized data search used to find images. Due to the lack of visual information and context of information, the searching results are unsatisfactory at most of the times. Therefore, image re-ranking which incorporates visual features of images to improve text based image searching is introduced. In this paper, we propose to exploit semantic attributes for image search re-ranking. i.e. The is image represented by an attribute feature. Then we propose a visual-attribute joint hyper graph learning approach to simultaneously explore two information sources. A hypergraph is constructed to model the relationship of all images. We conduct experiments on more than 1,000 queries in MSRA-MMV2.0 data set. The experimental results demonstrate the effectiveness of our approach.

Keywords: Image Search, Feature Extraction, ranking factors

I.INTRODUCTION

Search engines can provide many search services to users including text, image and video search services. Today, searches for textual content generally encompass the majority of the searches performed by the users. However, searches for image content are increasingly becoming popular with the user as image search services are becoming more readily available to the users. Searching for images in a large database is still highly challenging. Image search technology at the major search engines does mostly rely upon searches where images are associated with keywords rather than for specific details about the images themselves. Manual image annotation is time-consuming, laborious and expensive. Hence we introduce a approach called re-ranking. It is defined as re-ordering the visual documents based on the initial text based search and their visual patterns.

Re-ranking is the process performed for image optimization. Traditional image search are based on matching the text queries. Sometimes the text on those pages aren't very relevant to the image presented on those pages. A way of ranking images using a mix of ranking factors might make it more likely that the image presented to people searching for them are related to the query they used during their search. Currently, image search technology is based on keyword search and is not based on searching for the media itself. The existing approach was based on low level features extraction which leads to too many deficiencies. Different from the existing methods, a hypergraph is then used to model the relationship between images by integrating low-level features and attribute features.

II.EXISTING SYSTEM

There currently does not exist sufficient methods for determine the relevance of each image for a corresponding search query. Conventional methods may not take into considerations the relevancy of text with corresponding images and also the quality factor is not taken into consideration. Accordingly, conventional methods for providing images in response to image search queries may not provide the best possible image search results to requesting user. According to the statistical analysis model used, the existing re-ranking approaches can roughly be categorized into three categories including the clustering based, classification based and graph based methods. The algorithm used in the existing system such as k-means, mean shift, k-mediods suffer from a serious of disadvantages. Some of them are searching result is highly diverse, Visual pattern not clear, Classification problem when identify whether each relevant or not.

III.PROPOSED SYSTEM

Graph based methods have been proposed recently and received increasing attention as demonstrated to be effective. The proposed system is developed to perform a function that includes receiving a search query having text and identifying image related to search query. Furthermore, the function can include re-ranking the images using one or more ranking factor and providing the Re-ranked images to the requester. The multimedia entities in top ranks and their visual relationship can be represented as a collection of nodes and edges. After a query "apple" is submitted, an initial result is obtained via a text-based search engine. It is observed that text-based search often returns "inconsistent" results. Then the re-ranked result list is created first by ordering the clusters according to the

cluster conditional probability and next by ordering the samples within a cluster based on their cluster membership value. In a fast and accurate scheme is proposed for grouping Web image search results into semantic clusters.

A) Image Features Extraction

In image processing, feature extraction starts from an initial set of measured data and builds derived values (features) intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations. When the input data to an algorithm is too large to be processed and it is suspected to be redundant ,then it can be transformed into a reduced set of features. This process is called *feature extraction*. The extracted features are expected to contain the relevant information from the input data, so that the desired task can be performed by using this reduced representation instead of the complete initial data. We used four types of features, including color and texture, which are good for material attributes; edge, which is useful for shape attributes. We performed K-means clustering with 128 clusters. The color descriptors of each image were then quantized into a 128-bin histogram. Texture descriptors were computed for each pixel as the 48-dimensional responses of text on filter banks.

B) Semantic Attributes

Semantic attributes can be regarded as a set of mid-level semantic preserving concepts. Different from low-level visual features, each attribute has an explicit semantic meaning, *e.g.*, "animals". Attribute concepts also differ from specific semantics since they are relatively more general and easier to model, *e.g.*, attributes "animal" and "car" are easier to model and distinguish than the concrete semantic concepts "Husky" and "Gray Wolves". Due to the advantages of being semantic-aware and easier to model, attributes have been studied recently and are revealing their power in various applications such as object recognition and image/video search. Thus, attributes are expected to narrow down the semantic gap between low-level visual features and high-level semantic meanings.

Figure 1: Some examples images for semantic attributes

C) Attribute-Assisted Hyper Graph Construction:

We propose an attribute-assisted hypergraph learning method to reorder the ranked images which returned from search engine based on textual query.In this paper, we regard each image in the data set as a vertex on hypergraph G =(V, E,w). Assume there are n images in the data set, and thus, the generated hypergraph contains n vertices. Let V = {v1,v2,...,vn} denote n vertices and E ={ e1,e2,...,em}represent m hyperedges where the images sharing the same attribute are connected by one hyperedge. For various hyperedges, we set the weight vector to be w =[w1,w2,...,wm] in the hypergraph, where m i =1 wi = 1. In each hyperedge, we select K images which offer more preference to corresponding attribute based on the descending order of classifier scores. So, the size of a hyperedgein our framework is K.

Algorithm 1 Attribute-Assisted Hypergraph Learning

Step 1: Initialization.
1.1 Set W as a diagonal matrix with initial values.
1.2 Construct the hypergraph Laplacian Δ and compute the matrices D_v,D_e and H accordingly.
Step 2: Label Update.
Compute the optimal f based on the equation 17, which is:

$$f = (1-\alpha)(I-\alpha\Theta)^{-1}y$$

Step 3: Weight Update.
Update the weights w_i with the iterative gradient descent method introduced.
Step 4:
After obtaining W, update the matrix Θ accordingly.
Step 5:
Let $t = t+1$. If t > T, quit iteration and output the results, otherwise go to step 2.

Figure 2: Algorithm for Attribute-Assisted Hypergraph

D) Web Image Search Re-Ranking:

Web image search re-ranking is emerging as one of the promising techniques for automotive boosting of retrieval precision. The basic functionality is to reorder the retrieved multimedia entities to achieve the optimal rank list by exploiting visual content in a second step. In particular, given textual query, an initial list of multimedia entities is returned using the text-based retrieval scheme. Subsequently, the most Relevant results are moved to the top of the result list while the less relevant ones are reordered to the lower ranks. As such, the overall search precision at the top ranks can be enhanced dramatically. Deciding what images are associated with which queries first depends upon a search engine associating images with keywords that might be used as search queries. It might take all of the text from those pages and store it in a database, or text that is only a certain distance away from the pictures. Ranking factors are then used to determine the relevance of a picture to the query, and the order that these associated images are presented to a searcher.

Figure 3: Re-ranked image by attributes

E) Re-Ranking Factors

Images can be identified through any conventional means such as identifying image tag within a web page's source code. The text can be identified that is related to an identified image but is found on a different web page as the identified image. The images stored in the database are ranked based on one or more ranking factor. An image's ranking can be upgraded or downgraded based on the ranking factors. Below are some ranking factors that can be employed when ranking images:

- Collecting and using the number of websites that contain an identical image as a ranking factor
- Collecting and using the number of websites that contain a similar image as a ranking factor
- Collecting and using the size of images as a ranking factor
- Detecting image feature levels as a ranking factor

IV.EXPERIMENTAL RESULTS AND ANALYSIS

We use the MSRA-MM V2.0 dataset as our experimental data. This dataset consists of about 1 million images from 1,097 diverse yet representative queries collected from the query log of Bing [7]. Our proposed hypergraph re-ranking approach performs the best in comparison with the conventional re-ranking strategies.We can see that the presented hypergraph learning approach assisted with 2 regularizer performs better than Hypergraph Re-ranking. It achieves around 3.8% improvements at MNDCG@20. From the experimental comparison, we could see that our approach is more favorable in the task of Web image search re-ranking, which improves the baseline steadily and outperforms the other re-ranking strategies. The true positive (tp) and true negative (tn) are divided by the total number of samples (n) and classification accuracy is denoted as: Accuracy =(tp+tn)/ n

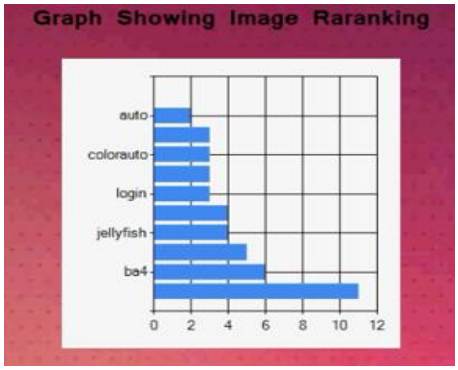

Figure 4: Performance Measures

V.CONCLUSION

With the help of image re-ranking technique, the quality of the image search can be improved.The method of providing images in a ranked order comprises of: receiving a search query including at leastone item of text;identifying at least one image related to the search query;ranking the at least one image using one or more ranking factor,wherein at least one of the one or more ranking factors considers a distance the at least one item of text is to the at least one image on a web page , and providing the ranked one or more image.

VI.REFERENCES

[1] Junjie Cai, Zheng-Jun Zha, Member, IEEE, Meng Wang, Shiliang Zhang, and Qi Tian, Senior Member, IEEE, "An Attribute-Assisted Reranking Model for Web Image Search", IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 24, NO. 1, JANUARY 2015.

[2] J. Cai, Z.-J. Zha, W. Zhou, and Q. Tian, "Attribute-assisted reranking for web image retrieval," in Proc. ACM Int. Conf. Multimedia, 2012, pp. 873–876.

[3] L. Yang and A. Hanjalic, "Supervised reranking for web image search," in Proc. Int. ACM Conf. Multimedia, 2010, pp. 183–192.

[4] Y. Huang, Q. Liu, S. Zhang, and D. N. Metaxas, "Image retrieval via probabilistic hypergraph ranking," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2010, pp. 3376–3383.

[5] K. Barnard, P. Duygulu, N. de Freitas, D. Forsyth, D. Blei, and M. I. Jordan. Matching words and pictures. Journal of Machine Learning Research, 3, pages 1107–1135, 2003.

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

[7] M. Wang, L. Yang, and X.-S. Hua, "MSRA-MM: Bridging research and industrial societies for multimedia," Tech. Rep. MSR-TR-2009-30, 2009.

[8] K. Järvelin and J. Kekäläinen, "IR evaluation methods for retrieving highly relevant documents," in Proc. ACM SIGIR Conf. Res. Develop. Inf. Retr., 2000,. 41–48

[9] W. H. Hsu, L. S. Kennedy, and S.-F. Chang, "Video search reranking via information bottleneck principle," in Proc. ACM Conf. Multimedia, 2006, pp. 35–44.

[10] J. Yu, D. Tao, and M. Wang, "Adaptive hypergraph learning and its application in image classification," IEEE Trans. Image Process., vol. 21, no. 7, pp. 3262–3272, Jul. 2012.