International Journal of Scientific Research in
Computer Science and Engineering
Vol.4, Issue.3, pp.20-25, June (2016)

Review Paper

E-ISSN:2320-7639

An Attribute-Assisted Reranking Model for Web Image Search

S.T. Tangudubilli[1*], A.S. Kumar[2]

[1]Dept. of CSE, Sanketika Vidya Parishad Engineering College, Visakhapatnam, India
[2]Dept. of CSE, Sanketika Vidya Parishad Engineering College, Visakhapatnam, India

Corresponding Author: santhoshi1@gmail.com

Available online at www.isroset.org

Received: 14 May 2016, Revised: 25 May 2016, Accepted: 18 Jun 2016, Published: 30 Jun 2016

Abstract— Image search reranking is a successful approach to refine the text-based image search result. Most existing reranking ways are based on low level visual features. This paper proposes to make use of semantic attributes for image search reranking. Depend on the classifiers for all the predefined attributes, each image is represented by an attribute feature containing the responses from these classifiers. A hypergraph is then used to model the relationship between images by combining low level visual features and attribute features. Hypergraph ranking is then performed to order the images. The basic principle is that visually similar images should have similar ranking scores. In this paper, we propose a visual attribute joint hypergraph learning approach at the same time to explore two information sources. A hypergraph is created 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 indicate the productiveness of our approach.

Keywords- Text base query; Attribute-assisted; Image retrieval; Query image; hyper graph learning; Image reranking

I. INTRODUCTION

In everyday life, the searching of an image have become a part of our working. Which will gives the very adequate understand ability of our working. Based on the basis of this approach we are using the search engine basis searching. This will gives the high consequence set of images. But this gives result is not the adequate from the user requirement. As per user they said that, it will not give direct output of the images which they wants. Therefore we use the concept of applicable searching as per the user need which will gives the user to choice which type of image he/she searching. Hence the searching mechanism should be very adequate as per the existing system. In such a system that will make easy searching of images that is advantageous for the users based on the re-ranking strategy [1]. This strategy helps user can getting top nine images based on the hyper graph instead of the number of images. In such a system the user can click on the intents to search images to show the related result.

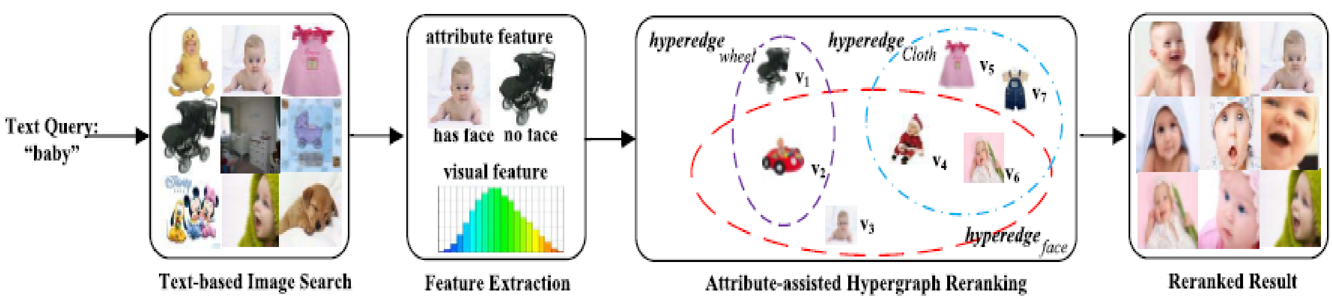

Figure 1. An attribute assisted web image search

The concept of the filtering which is used to give the choices to the user. The filtering is nothing but a pool of image when user select its interest, then it will be filtering the result set into user interested images. This will give the applicable searching of the images. Which create the more interaction with the user while searching, because of this way, if user do not have any knowledge about the text based query searching then this will give the additional knowledge in the user knowledge [2].

The searching of image is searched on the visual semantic signature which is the resemblance of the form low level feature extraction of size, shape, color etc., this will differ the images from similar characteristics. That is nothing but our query specified or user requires image searching. Our work emphasis on two parts:

1. Offline 2. Online

The offline process is used for the text based query input. Which retrieve images from the search engine? This procedure also done for extracting the semantic signature. To make efficient data set. The online procedure is done for retrieval of images from the search engine. It will also do the filtering of the images using the query image. And helps to remove the unnecessary search on the image.

The initial search results from text-based retrieval can be grouped by visual closeness. In the classification based methods, visual reranking is organized as binary classification problem which aims to check whether search result is applicable or not [3]. Image search reranking use the firm relationship in the graph. All these aspects make us to utilize semantic attributes for image representation. Fig. 1 illustrates the flowchart. First a query "baby" is submitted to the text based search engine and then an initial result is obtained based on the attributes [4]. Web image search reranking is emerged as one of the promising techniques for boosting of retrieval precision.

A hyper graph is then used to shows the relationship between images by including low-level visual features and attributes features [5]. Visual-attribute joint hyper graph learning approaches the simultaneous exploration of two information sources. Visual representation and semantic description are synchronously uses in a model is called hyper graph. The selection of attribute features can be, at the same time performed through the process of hyper graph learning so that, the effects of semantic attributes can be used further more. By comparing with the previous method, a hyper graph is reassembling to model the relationship between all the images, in which each vertex represent an image and a hyper edge represents [6].

II. Related Work

To develop the act of searching images visual search reranking is very good choice. In this section, existing visual search reranking ways are explained along with sematic attributes and hypergraph learning. To advance the accuracy of the text-based image search ranking, visual reranking has been suggested to refine the search result from the text-based image search engine by consolidating the information transferred by the visual method.

A. Test Based Search

When the user enters the query in the search engine it gets the related images with respect to that query in conclusion image set. The search engines never day's uses different image search algorithm. Mostly they are text based. That mean the conclusion image set contain only the images which have name identical to that query. The image set contains all the images that are retrieved from the image database which have the name similar to input query. All this happened because of using text based algorithm in which ASCII values choose the ranking of characters. In database there are many images relevant to our query so their ranking is essential to get ideal result. In order to rank the text based search, algorithm uses the ASCII values [7].

As per the ranking of ASCII value image names of concluded images are ranked. The main advantage of text based searching is that, it helps to get all that images from database having the name identical to our query. But disadvantage is that, it unable to concentrate on image contain. The resultant image set contain the images which not related to our search of interest, only the image name is same to the query that's why they are in concluded image set. In short, text based search can't check relevance of images [8]. Some algorithms are their which check image relevance but they have some drawback.

B. Content Based Image Search

It is designed to work more with the actual pieces of the image. Some types use images as samples, some take various pieces of color info, etc. Different types are there which includes, Region-based, Object-based, Example-based, and Feedback based.

Region-based Image Retrieval:

It is a low-level content-based searching. It can be interpret portions of images. This works with low-level images. This can be partition image and search only one portion. But this can't work with objects. For high detailed images it is impossible [9].

Object-based Image Retrieval:

It can be worked with the pieces of an image, like Region-based Image. Image Retrieval can interpret images including high-detail. High-detailed images are easy to search. It can use pre-defined shapes, to get images for the query. Implementation is very intense. User-interface also does not fit typical search ideas of simplicity [10].

Example-based Image IR:

In this users would give a sample image, or a portion of an image, that the system uses as a base for the search. The system then finds the images that are similar to the base image. Easy for the user until the user realizes that the picture they want looks nothing like the one they have. It can be simple input for the user[11].

Feedback-based Image IR:

This takes slightly, time consuming for the user. System shows user a sample of pictures and asks for rating from the user. Using these ratings, system re-queries and rerun until the right image is found. Any image can be found with enough feedback. It may take a lengthy time to find the image that the user wants.

C. Visual Reranking

It is the re-arrangement of images on the basis of visual similarities. Visual reranking has been suggested to purify the search result from the text-based image search engine by consolidating the information transferred by the visual method. According to the statistical analysis model used, the existing reranking ways can roughly be classified into three categories including the clustering based, classification based and graph based methods.

Clustering-Based Methods:
Clustering analysis is very useful to evaluate the inter-entity similarity. The images in the initial results are primarily organized automatically into several many near duplicate media documents. However, for queries that return extremely diverse results or without clear visual patterns, the performance is not guaranteed.

Classification-Based Methods
In the classification based methods, the visual reranking is been developed as binary classification problem aiming to identify whether each search result is applicable or not. For example, a classifier or a ranking model is learned with the pseudo purposed feedback. Yet, in many real scenarios, training examples obtained via PRF are very noisy and might not be enough for training efficient classifier. To address this problem, learned a query independent text based re-ranker. The top ranked results that obtained from the text based reranking are selected as positive training examples. Negative training examples are picked up randomly from other queries. A binary SVM classifier is used to re-rank the results on the basis of visual features.

Graph-Based Methods
Graph based methods have been suggested recently and received increased thought as demonstrated to be effective. Visual rank framework casts the reranking problem as the random walk on a similarity graph and reorders images according to the visual similarities. The final result list is generated along sorting the images based on graph nodes weights. The objective is to improve the consistency of ranking scores over visually similar samples and minimize the deviation between the optimal list and the initial list. Hence, the performance is significantly dependent on the statistical properties of the top ranked search results. Motivated by this observation, a semi-supervised framework to refine the text based image retrieval results along leveraging the data distribution and the partial supervision information obtained from the top ranked images is suggested.

D. Semantic Attributes

Attributes are expected to narrow down the semantic gap between low-level visual features and high-level semantic meanings. Moreover, the type of the most efficient features should vary across the queries. For example, for queries that are related to the color distribution, such as the sunset, sunrise and the beach, color features will be useful. For some queries like the building and the street, edge and texture features will be more effective. It can be understood that semantic attribute could also be viewed a description or a method of image data. Using multimodal features, one can assure that the useful features for different queries are contained. Therefore, all these perfection drive us to accomplish semantic attributes for image representation in the task of web image search reranking. Based on the classifiers for all the predefined attributes, each image is represented by an attribute feature that consists of the responses from these classifiers.

E. Hypergraph Learning

Visual representation and semantic description are simultaneously accomplished in a unified model called hypergraph. A hyperedge in a hypergraph is able to link more than two vertices. Different from the existing methods, a hypergraph is then used to model the relationship between the images by integrating low-level features and the attribute features. The selection of attribute features can be conducted simultaneously through the process of hypergraph learning in a way that the effects of semantic attributes could be more tapped and integrated in the reranking framework.

Graph based methods have been suggested recently and received increasing attention as indicated to be impressive. The multimedia entities in top ranks and their visual relationship can be described as a collection of nodes and edges. 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. Regularized logistic regression trained for each attribute within each class. As attribute features are formed by prediction of several classifiers, semantic description of each image might be inaccurate and noisy. Compared with the previous method, a hypergraph is reestablished to model the relationship of all the images, in which each vertex denotes an image and a hyperedge represents an attribute and a hyperedge connects to multiple vertices. 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 ways can be performed on this graph to categorize objects. It is noted that this simple graph cannot reflect the higher-order information. Compared with the edge of a simple graph, a hyperedge in a hypergraph is able to link more than two vertices [12].

III. Attribute Based Image Search Re-Ranking

A. Learning Scalable Discriminative Dictionary with Sample Relatedness

This method proposes a new dictionary learning method which encodes the image visual features into the binary ones, and more particular it effectively alleviates the above limitations. Our approach is motivated by the fact that humans flexibly adapt the number and nature of the attributes they use to the accordance and variety of the observed objects, and to the complexity of the task. For example, from the great number of possible attributes to describe a set of animals, such as furry, four-legged and can swim, humans effectively only use a limited number. The principle to select attributes is simple: the chosen attributes should provide sufficient information to reflect shared and discriminative properties. This method follows this principle and combines three main things. First, this model discovers binary features by factorizing low-level features of training images into a dictionary of arbitrary (infinite) size – the actual visual patterns present in the data from the dictionary, which adapts to the complexity of the data. The resulting Adaptive Dictionary algorithm is practical even for large data sets. Second, this model uses the Adaptive Dictionary algorithm in a discriminative framework that not only strives for good representations, but also biases towards learning dictionary which provides discriminative binary features. In the model, the dictionary, binary representations of training samples and classifiers are learned jointly in a max-margin framework.

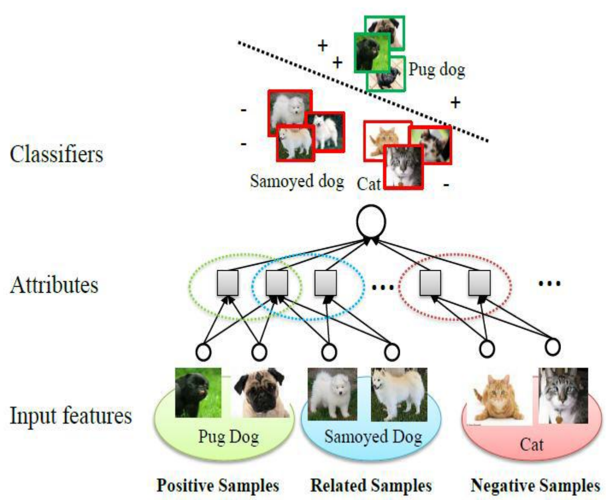

Figure 2. Illustration of the proposed dictionary learning method

Third, to build up the generalization ability of dictionary, this method exploits the knowledge about sample correspondence to guide the learned binary features to take the relational structure between samples. Specifically, this method encourages closely related samples to have more similar binary features than less related ones. Hence, the dictionary generalizes by applying related examples while still being discriminative. Figure 2 shows a graphical illustration of this method. The comprehensive experiments in suggest that the resulting learned dictionary is indeed discriminative and generalizes well.

It uses three types of samples for training: positive samples, samples related to the positive class and negative samples. “Attributes” of related samples (pug dog and samoyed dog) are encouraged to be shared, but the “attributes” of unrelated samples (pug dog and cat) may be different. In short, this way has the following benefits: (1) the size of learned dictionary automatically adapts to the complexity of the training data. Thus there is no need of bother to determine an appropriate number of bases in the dictionary as regularization parameter in this method works across a variety of data sets. (2) No need to pre-define an attribute vocabulary and tediously annotate the attributes for the training samples. (3) The model can consolidate arbitrary levels of sample relatedness from a variety of sources. In this way, the structure captured by the learned dictionary and features can be tailored to specific needs and data.

B. Attribute-augmented Semantic Hierarchy for Image Retrieval

When a semantic hierarchy is available to structure the concepts of images, we can more boost image retrieval by applying the hierarchical relations between the concepts. This method presents a novel. Each semantic concept is linked to a set of related attributes. These attributes are specifications of the multiple facets of the corresponding concept. Unlike the traditional flat attribute structure, the concept-related attributes span a local and hierarchical semantic space in the context of the concept. For instance, the attribute “wing” of concept “bird” refers to appendages that are feathered; while the same attribute refers to metallic appendages in the context of “jet”.

We develop a hierarchical semantic similarity function to precisely characterize the semantic similarities between images. The function is computed as a hierarchical aggregation of their similarities in the local semantic spaces of their common semantic concepts at multiple levels. In order to better capture users’ search intent, a hybrid feedback mechanism is also developed, which collects hybrid feedbacks on attributes and images.

These feedbacks are then used to refine the search results based on A2SH. Compared to the attribute-based image retrieval system based on flat structure, A2SH organizes images as well as concepts and attributes from general to specific and is moreover expected to achieve a further efficient and effective retrieval.

C. Attribute-Assisted Hypergraph Based Image Search Rearanking

Image Feature

Four types of features are useful, including color and texture, which are good for material attributes; edge, which is useful for shape attributes; and scale-invariant feature transform (SIFT) descriptor, which is useful for part attributes. A bag-of-words style feature is used for each of these four feature types. Color descriptors were densely extracted for each pixel as the 3-channel LAB values. K-means clustering represented 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. The texture descriptors of each image were then quantized into a 256-bin histogram. Edges were found using a standard canny edge detector and their orientations were quantized into 8 unsigned bins. This gives rise to a 8-bin edge histogram for each image. SIFT descriptors were densely extracted from the 8 × 8 neighboring block of each pixel with 4 pixel step size. The descriptors were quantized into a 1,000-dimensional bag-of-words feature. Since semantic attributes usually appear in one or more certain regions in an image, split each image into 2 × 3 grids and extracted the above four kinds of features from each grid respectively. Finally, obtain a 9,744-dimensional feature for each image, consisting of a 1, 392 × 6-dimensional feature from the grids and a 1,392-dimensional feature from the image. This feature was then used for learning attribute classifiers.

D. Attribute Learning

Support Vector Machine (SVM) classifier use for each attribute. Yet, simply learning classifiers by fitting them to all visual features often fails to generalize the semantics of the attributes correctly. For each attribute, need to select the features that are most adequate in modelling this attribute. Feature selection method is apply in this case. Specifically, if we want to learn a "wheel" classifier, we select features that perform well at distinguishing examples of cars with "wheels" and cars without "wheels". By doing so, it is help the classifier avoid being confused about "metallic", as both types of example for this "wheel" classifier have "metallic" surfaces. Features are selected using consistent logistic regression trained for each attribute within each class, then pool examples over all classes and train using the selected features. Such regression model is utilized as the preliminary classifiers to learn sparse parameters. The features are then selected by pooling the union of indices of the sparse nonzero entries in those parameters. For example, first select features that are good at differentiating cars with and without "wheel", then use the same procedure to select features that are good at separating motorbikes with and without wheels, buses with and without wheels, and trains with and without wheels. Then pool all those selected features and learn the "wheel" classifier over all classes using those selected features. In this way, effective features are selected for each attribute and the selected features are then used for learning the SVM classifier.

Attribute-Assisted Hypergraph

Attribute-assisted hyper graph learning method is used to reorder the ranked images which returned from search engine based on textual query. Different from the typical hypergraph, it presents not only whether a vertex belongs to a hyperedge, but also the prediction score that is affiliated to a specific. The weight is incorporated into graph construction as tradeoff parameters among various features. This modified hypergraph is thus able to improve reranking performance by mining visual feature as well as attribute information. Fig. 2 illustrates the flowchart of our proposed method. After a query "baby" is submitted, an initial result is obtained via a text-based search engine. It is observed that text-based search often returns inconsistent results. Some visually similar images are scattered in the result while other irrelevant results are filled between them, such as "dog" and "disney baby". Based on the returned images, both visual features and attribute features are extracted. In particular, the attribute feature of each image consists of the responses from the binary classifiers for all the attributes. These classifiers are learned offline. Visual representation and semantic description are simultaneously exploited in a unified model called hypergraph. Hypergraph is rearranged to model the relationship of all the images, in which each vertex denotes an image and a hyperedge represents an attribute and a hyperedge connects to multiple vertices. The weight of each edge based on the visual and attribute similarities of images which belongs to the edge. The relevance scores of images are learned based on the hypergraph. 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. Essentially, modeling relationship among more close samples will be able to preserve the stronger semantic similarity and thus facilitate ranking performance. Finally, the reranked list of the images set with respect to relevance scores in descending order.

IV. CONCLUSION

Image search reranking has been studied for several years and various approaches have been developed recently to boost the performance of text-based image search engine for general queries. This paper serves as a attempt to include the attributes in reranking framework. It is observe that semantic attributes are expected to narrow down the semantic gap between low-level visual features and high level semantic meanings. Motivated by that, a novel attribute-assisted retrieval model for reranking images is proposed. Based on the classifiers for all the predefined attributes, each image is represented by an attribute feature consisting of the responses from these classifiers. A hypergraph can be the effective approach to model the relationship between images by integrating low-level visual features and semantic attribute features. Hypergraph ranking performed to re-order the images, which is also constructed to model the relationship of all images.

REFERENCES

[1]. J. Cai, ZJ Zha, M. Wang, S. Zhang, Q. Tian, *"An Attribute-Assisted Reranking Model for Web Image Search*", IEEE Transactions On Image Processing, Vol. 24, No. 1, pp.261-272, 2015.

[2]. L. Yang, A. Hanjalic, *"Prototype-Based Image Search Reranking",* IEEE Transactions On Multimed, Vol.14, No.3,pp. 871-88, 2012.

[3]. S. Changpinyo, E. Sudderth, "Learning image attributes using the Indian Buffet Process", Technical report ofBrown University, USA, pp.1-146, 2012.

[4]. B. Siddiquie, R.S. Feris, L.S. Davis, "*Image ranking and retrieval based on multi-attribute queries*", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, pp.801-808, 2011 .

[5]. J. Feng, S. Jegelka, S. Yan and T. Darrell, "*Learning Scalable Discriminative Dictionary with Sample Relatedness*", 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, pp.1645-1652, 2014.

[6]. N. Kumar, A. C. Berg, P. N. Belhumeur and S. K. Nayar, "Attribute and simile classifiers for face verification," 2009 IEEE 12th International Conference on Computer Vision, Kyoto, pp.365-372, 2009.

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

[8]. K. Järvelin, J. Kekäläinen, "*IR evaluation methods for retrieving highly relevant documents*", 23rd annual international ACM SIGIR conference on Research and development in information retrieval, Greece , pp.41–48, 2000.

[9]. W. H. Hsu, L. S. Kennedy, S.-F.Chang, "*Video search reranking via information bottleneck principle",* Proceedings of the 14th ACM international conference on Multimedia, Santa Barbara, pp.35–44, 2006.

[10]. Y. Huang, Q. Liu, S. Zhang, D.N. Metaxas, *"Image retrieval via probabilistic hypergraph ranking",* IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, pp.3376–3383, 2010.

[11]. C. H. Lampert, H. Nickisch, S. Harmeling, "*Learning to detect unseen object classes by betweenclass attribute transfer*", IEEE Conference on Computer Vision and Pattern Recognition, Miami, pp. 951–958, 2009.

[12]. R. Yan, A. Hauptmann, and R. Jin, "*Multimedia search with pseudo- relevance feedback*", Proceedings of the 2nd international conference on Image and video retrieval, Urbana-Champaign, pp. 238–247, 2003.

