ATTRIBUTE BASED WEB IMAGE SEARCH RERANKING MODEL

[1]P. Manikanta, [2]A. Divya

[1]M.Tech Student, [2]Assistant Professor,
Department of Computer Science and Engineering,
P.V.P.Siddhartha institute of technology, Kanuru, Vijayawada-520007, Andhra Pradesh, India.

Abstact: **Image searching on web is very popular now days for getting intended images. People generally use available and popular search engines like Google search engines, Bing search, and Yahoo search engine. This popular search engines have common method i.e. Text based Retrieval. The noisy or irrelevant images may be present in the retrieved results. Image search re ranking is an effective approach to refine the text-based image search result. Exploit semantic attributes for image search Re ranking. 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 is then used to model the relationship between images by integrating low-level visual features and attribute features. Hypergraph ranking is then performed to order the images. . Its basic principle is that visually similar images should have similar ranking scores. This modeling connection among more close samples will be able to Domain the robust semantic similarity and thus accelerate the great ranking performance.**

Index Terms: **Search, hypergraph, attribute-assisted**

I. INTRODUCTION

INTRODUCTION OF THE PROJECT

With the dramatic increase of online images, image retrievals attracted significant attention in both academia and industry many image search engines such as Google and Bing have relied on matching textual information of the images against queries given by users. However, text based image retrieval suffers from essential difficulties that are caused mainly by the incapability of the associated text to appropriately describe the image content.

Recently, visual reranking has been proposed to refine text-based search results by exploiting the visual information contained in the images the existing visual reranking methods can be typically categorized into three categories as the clustering based, classification based and graph based methods. The clustering based reranking methods stem from the key observation that a wealth of visual characteristics can be shared by relevant images. With intelligent clustering algorithms (e.g., mean-shift, *K*-means,and *K*-medoids), initial search results from text-based retrieval can be grouped by visual closeness. However, for queries that return highly diverse results or without clear visual patterns, the performance of the clustering-based methods is not guaranteed. In the classification based methods, visual reranking is formulated as binary classification problem aiming to identify whether each search result is relevant or not. Pseudo Relevance Feedback (PRF) is applied to select training images to learn a classifier or a ranking model. However, in many real scenarios, representative examples obtained via PRF for the training dataset are very noisy and might not be adequate for constructing effective classifiers. Graph based methods have been proposed recently and received increasing attention as demonstrated to be effective. The multimedia entities in top ranks and their visual relationship can be represented as a collection of nodes and edges. propose to exploit stronger semantic relationship in the graph for image search reranking.

On the other hand, semantic attributes have received tremendous attention recently, where their effectiveness was demonstrated in broad applications, including face verification object recognition fine-grained visual categorization classification with humans-in-the-loop and image search. Semantic attributes could be shape, color, texture, material, or part of objects, such as "round," "red,""mental," "wheel" and "leg" etc. As a kind of intermediate level descriptor, an attribute has semantic meaning as opposed to low-level visual features, but it is easy to model compared to a full object, *e.g.*, "car". Thus, attributes are expected to narrow down the semantic gap between low-level visual features and high-level semantic meanings. Furthermore, attribute based image representation has also shown great promises for discriminative and descriptive ability due to intuitive interpretation and cross-category generalization property. Hence, attribute-based visual descriptor has achieved good performance in assisting the task of image classification. Besides that, an attribute is potentially any visual property that humans can precisely communicate or understand, even if it does not correspond to a traditionally defined object part. For instance, "red-dot in center of wings" is a valid attribute, even though there is not a single butterfly part that corresponds to it. Furthermore, the type of the most effective features should vary across queries.

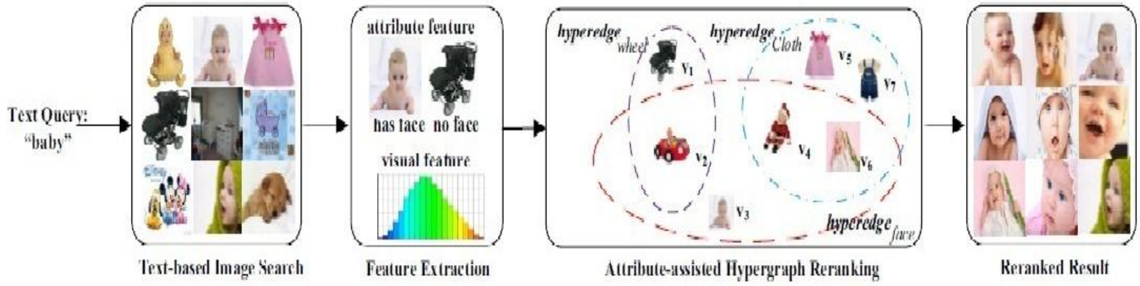

Figure 1: Structure of attribute assisted reranking model

For example, for queries that are related to color distribution, such as sunset, sunrise and beach, color features will be useful. For some queries like building and street, edge and texture features will be more effective. It can be understood that semantic attribute could also be viewed a description or modality of image data. Using multimodal features can guarantee that the useful features for different queries are contained. Therefore, all these superiorities drive us to exploit semantic attributes for image representation in the task of web image search reranking. Motivated by the above observations, we move one step ahead of visual reranking and propose an attribute-assisted reranking approach. Figure 1 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.

However, even these merits, we have to face the challenges it brings to this unified formulation at the same time: 1)simply learning classifiers by fitting them to all visual features often fails to generalize the semantics of the attributes correctly because low-level features are extracted by region or interest point detector instead of aiming to depict specific attribute. Some need to select representative features which are in favor to describe current semantic attributes. Hence we propose a regularize on the hyperedge weights which performs a weighting or selection on the hyperedges. In this way, for attributes or hyperedges that are informative, higher weights will be assigned. In contrast, noisy hyperedges will be implicit removed when the weights converges to zeros after hypergraph learning. Finally, we can obtain the reranked list of the images with respect to relevance scores in descending order.

PROBLEM STATEMENT

Image search Reranking is a real approach to recover the text-based image search result. Exploiting low-level visual features for image retrieval. The graph based image classification technique are more effective, hence a Hypergraph is used to model the association among images by incorporating low-level features and attribute features.

II.LITERATURE SURVEY

Describing objects by their attributes

Doing so allows us not only to name cognizant objects, but also: to report unfamiliar features of a familiar entity ("spotty dog", not just "dog") to say something about unknown entities and to learn how to identify new objects with little or no graphical samples. Rather than focusing on identity assignment, this mark inferring features the core problem of recognition. These features can be meaningful ("spotty") or discriminative ("dogs have it but sheep do not"). Learning features shows a main new challenge: generalization across object groupings, not just across instances within a category. This paper present innovative feature selection method for learning attributes that simplify well across categories. Paper objectives are supported by thorough assessment that offers insights into the boundaries of the standard recognition model of naming and shows the new abilities delivered by our attribute based framework.

Image ranking and retrieval based on multi-attribute queries

Proposed a novel method for ranking and retrieval of images based on multi-attribute queries. The image retrieval methods that are already in place train separate classifiers for each word and heuristically uses their outputs for accessing multiword queries. Moreover, these methods also ignore the interdependencies among the query terms. In contrast, a proposed principle approach for multi-attribute retrieval which explicitly frame the connections that are present between the features. Given a multi attribute query, this also utilize other attributes in the jargon which are not exist in the query, for ranking/retrieval. Furthermore, the integration of ranking and retrieval within the same formulation, by presenting them as structured prediction problems. Extensive experimental evaluation presents that our method expressively overtakes several state of the-art ranking and retrieval methods.

Image retrieval via probabilistic hypergraph ranking

We propose a new transductive learning framework for image retrieval, in which images are taken as vertices in a weighted hypergraph and the task of image search is formulated as the problem of hypergraph ranking. Based on the similarity matrix computed from various feature descriptors, we take each image as a 'centroid' vertex and form a hyperedge by a centroid and its k-nearest neighbors. To further exploit the correlation information among images, we propose a probabilistic hypergraph, which assigns each vertex vi to a hyperedgeej in a probabilistic way. In the incidence structure of a probabilistic hypergraph, we describe both the higher order grouping information and the affinity relationship between vertices within each hyperedge. After feedback images are provided, our retrieval system ranks image labels by a transductive inference approach, which tends to assign the same label to vertices that share many incidental hyperedges, with the constraints that predicted labels of feedback images should be similar to their initial labels. We compare the proposed method to several other methods and its effectiveness is demonstrated by extensive experiments on Corel5K, the Scene dataset and Caltech 101.

Harvesting image databases from the web

The objective of this work is to automatically generate a large number of images for a indicated object class (for example, bear). A multi-modal approach employing both text, meta data and visual features is used to access many high-quality images from the web. Applicant images are found by a text based web search inquiring on the object identifier (the word bear). The task is then to remove irrelevant images and re-rank the residue. No visual information is used at this phase. Second, as the (noisy) training data utilizes the top - ranked images and a SVM visual classifier is learnt to advance the ranking more. The major uniqueness is in combining text/meta-data and visual features in order to achieve a fully automatic ranking of the images. Examples like selection of animals (e.g. camels, sharks, and penguins), vehicles (cars, airplanes, cycle) and other classes (guitar, wristwatch), totaling 18 classes. The results are assessed by precision/recall curves on perfectly annotated data and by comparison to previous approaches including those of Bergetal.

III.METHODOLOGY

We propose an attribute-assisted hypergraph learning method 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 v belongs to a hyperedge$e$, but also the prediction score that v is affiliated to a specific e. The weight is incorporated into graph construction as tradeoff parameters among various features. Our modified hypergraph is thus able to improve reranking performance by mining visual feature as well as attribute information.

The hypergraph model has been widely used to exploit the correlation information among images. 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, andthus, the generated hypergraph contains n vertices. Let $V= \{v_1, v_2, . . . ,v_n\}$ denote$n$vertices and$E = \{e_1,e_2, . . . ,e_m\}$ represent $m$hyperedges where the images sharing the same attribute are connected by one hyperedge. For various hyper-edges, we set the weight vector to be $w= [w_1, w_2, ...,w_m]$ in the hypergraph, where $\sum_{i=1}^{m}(w_i) = 1$. 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 hyperedge in our framework is K . The incidence matrix H of a hypergraph is defined as follows:

$$h(v_i, e_j) = \begin{pmatrix} A(j,i) & if\ v_i \epsilon e_j , \\ 0 & otherwise \end{pmatrix}$$

According to this formulation, A(I,j) describes the similarity between v_i and v_j and determines that a vertex v_i is assigned to e_j. We define the similarity matrix A between two images as follows:

$$A(I,j)=\gamma_1 A_{attribute}+ \gamma_2 A_{local}+ \gamma_3 A_{global}= \gamma_1 \exp\left(-\frac{Dis_{attribut}(i,j)}{D_{attribute}}\right) + \gamma_2 \exp\left(-\frac{Dis_{local}(i,j)}{D_{local}}\right)+ \gamma_3 \exp\left(-(\Sigma_{k=1}^{3} \frac{Dis_{gk}(i,j)}{D_{localgk}}\right)$$

where $Dis_g(i,j)$, $Dis_{local}(i,j)$ and $Dis_{attribute}(i,j)$ are pair-wise Euclidean distance between $v_i$andv_j of global feature, local feature and attribute feature, respectively. And D is the mean value of elements in the distance matrix. The initial weight $w(e_i)$ is set to $w(e_i) =\Sigma_{vj\in ei}A(i, j)$.

For hypergraph learning, it is formulated as a regularization framework: $\arg_{f,w} \min\{\lambda R_{emp}(f)+\Omega(f)\}$

where f is the relevance score to be learned, $\Omega(f)$ is the normalized cost function, $R_{emp}(f)$ is empirical loss and λ is a regularizer on the weights. Instead of fixing hyperedge weights, we assume that they have Gaussian distribution, such that the weights w can be learned together with the relevance score f . Defining $\Theta=D_v^{-1/2}HWD_e^{-1}H^TD_v^{-1/2}$

where I is the identity matrix, Δ=I- Θ is a positive semidefinite matrix called hypergraphLaplacian.To force the assigned relevance score to approach initial relevance score of y,a regularization term is defined as follows: $\| f -y\|^2=\Sigma_{u\epsilon V}(f(u)-y(u))^2$

The normalized rank is widely used to estimate the samples relevance probability. For the i-th image of total N sample images, the initial score will be y_i=1-i/N. Then the learning task is to minimize the sum of two cost terms with respect to f, which is

$$\Phi(f)= f^T\Delta f +\lambda\| f -y\|^2$$

where λ is the regularization parameter. Differentiating $\Phi(f)$ with respect to f, we have $f=(1-\alpha)(1-\alpha^{\theta})^{-1}y$

Algorithm

Attribute assisted hypergraph learning

Step 1: Initialization.

1.1 Set W as a diagonal matrix with initial values.

1.2 Construct the hypergraphLaplacian Δ and compare the matrices D_v,D_e and H accordingly.

Step 2: Label Update.

Compute the optimal f based on the equation, 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**(-)** accordingly.

Step 5: Let $t = t + 1$. If t > T, quit iteration and output the results. Otherwise go to step 2.

Four types of features, 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. We used a bag-of-words style feature for each of these four feature types. Color descriptors were densely extracted for each pixel as the 3-channel LAB values. 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. 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 an 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 one thousand-dimensional bag-of-words feature.

We learn a Support Vector Machine (SVM) classifier for each attribute. However, simply learning classifiers by fitting them to all visual features often fails to generalize the semantics of the attributes correctly. For each attribute, we need to select the features that are most effective in modeling this attribute. It is necessary to conduct this selection based on the following two observations: 1) such a wealth of low level features are extracted by region or interest point detector, which means these extraction may not aim to depict the specific attribute and include redundant information. Hence we need select representative and discriminative features which are in favor to describe current semantic attributes. 2) the process of selecting a subset of relevant features has been playing an important role in speeding up the learning process and alleviating the effect of the *curse of dimensionality*. We here apply the feature selection method as described. In particular, 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, we help the classifier avoid being confused about "metallic", as both types of example for this "wheel" classifier have "metallic" surfaces.

IV.CONCLUSION

Image search re ranking 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. 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, we propose a novel attribute assisted retrieval model for re ranking images. Based on the classifiers for all the predefined attributes, each image is represented by an attribute feature consisting of the responses from these classifiers. We perform hyper graph ranking to re-order the images, which is also constructed to model the relationship of all images. Its basic principle is that visually similar images should have similar ranking scores and a visual-attribute joint hyper graph learning approach has been proposed to simultaneously explore two information sources. The experimental results demonstrate the effectiveness of our proposed attribute assisted Web image search re ranking method. This project has a wide scope for development by including some extra features. We can add high level visual features means big size images upload to the system and view by the admin and user pages or forms. We can add download options for images to users.

REFERENCES

[1] L. Yang and A. Hanjalic. "Supervised reranking for web image search". In Proceedings of ACM Conference on Multimedia, 2010.
[2] X. Tian, L. Yang, J. Wang, Y. Yang, X. Wu and X.-S. Hua. "Bayesian video search reranking", Transaction on Multimedia, vol. 14, no. 7, pp. 131-140, 2012.
[3] F. Shroff, A. Criminisi and A. Zisserman. "Harvesting imag e databases from the web". In Proceedings of the IEEE International Conference on Computer Vision, 2007.
[4] B. Siddiquie, R.S.Feris and L. Davis. "Image ranking and retrieval based on multi-attribute queries". In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2011.
[5] A. Farhadi, I. Endres, D. Hoiem and D. Forsyth. "Describing objects by their attributes". In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2009.
[6] R. Yan, A. Hauptmann, and R. Jin, "Multimedia search with pseudo relevance feedback," in Proc. ACM Int. Conf. Image Video Retr., 2003, pp. 238–247.
[7] M. Wang, L. Yang and X.-S. Hua. "MSRA-MM: Bridging research and industrial societies for multimedia information retrieval", 2009.
[8] K. Jarvelin and J. Kelkalainen. "IR evaluation methods for retrieving highly relevant documents". In Proceedings of ACM SIGIR conference on Research and Development in Information Retrieval, 2000.
[9] 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.
[10] Y. Huang, Q. Liu, S. Zhang and D. N. Meta as. "Image retrieval via probabilistic hypergraph ranking". In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2010.
[11] Junjie Cai, Zheng-Jun Zha, Meng Wang, Shiliang Zhang, and Qi Tian, "An Attribute Assisted Reranking Model for Web Image Search", IEEE Transactions on Image Processing, Vol 24, No 1, January 2015.
[12] M. Wang, H. Li, D. Tao, K. Lu, and X. Wu, "Multimodal graph-based reranking for web image search," *IEEE Trans. Image Process.*, vol. 21, no. 11, pp. 4649–4661, Nov. 2012.
[13] Y. Su and F. Jurie, "Improving image classification using semantic attributes," *Int. J. Comput. Vis.*, vol. 100, no. 1, pp. 59–77, 2012.

[14] H. Zhang, Z.-J. Zha, Y. Yang, S. Yan, Y. Gao, and T.-S. Chua, “Attribute augmented semantic hierarchy: Towards bridging semantic gap and intention gap in image retrieval,” in Proc. ACM Conf. Multimedia, 2013, pp. 33–42.
[15] application in image classification,” *IEEE Trans. Image Process.*, vol. 21, no. 7, pp. 3262–3272, Jul. 2012.
[16] F. X. Yu, R. Ji, M.-H. Tsai, G. Ye, and S.-F. Chang, “Weak attributes for large-scale image retrieval,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2012, pp. 2949–2956
[17] L. Zheng and S. Wang, “Visual phraselet: Refining spatial constraints for large scale image search,” IEEE Signal Process. *Lett.*, vol. 20, no. 4, pp. 391–394, Apr. 2013.
[18] J. Cai, Z-J. Zha, W.Zhou, and Q. Tian, “Attribute –assisted reranking for web image retrival,” in proc. ACM Int. Conf. Multimedia, 2012,pp. 873-876.
[19] http://www.w3schools.com
[20] The Unified Modeling Language User Guide” by Wesley, Grady Booch, James Rumbaugh, Ivar jacobson, 1999.