ISSN 23499842(Online), Volume 1, issue 1 March 2015
International Journal of Innovative Trends and Emerging Technologies

IMAGE OBJECT CLUSTERING BY USING SEMANTIC FEATURE EXTRACTION AND USER TAG REFINEMENT

G.ARUN PRASANTH[1], Dr.P.MARIKKANNU[2] M.Tech.Phd
[1]PG Scholar, [2]Assistant Professor
[1,2]Department Of Information Technology, [1]Anna University
[2]Anna University Regional Center, Coimbatore
[1]g.arunprasanth88@gmail.com

ABSTRACT

Nowadays, the growth of an image and videos in social network are quite enormous. To retrieve the required image in social network are quite difficult. User may not get required image as low level of retrieval mechanism. To get the required image, user produces a query either by keyword or image for image retrieval search engine. The existing image retrieval algorithm is depends up on content of image and description for that image. Retrieval efficiency becomes low due to small changes in viewpoints, lighting condition and irrelevant description of an image. In this work, we propose to increase the accuracy of image object retrieval by combining content of image, associated user defined tag and feature extraction of an image. The proposed frame work discovers the relevant feature by means of visual and image graph. The proposed method can applied to keyword based search and image based search retrieval. Experimental results confirm that the proposed method improves the performance of image retrieval application.

Keywords: Semantic feature extraction, clustering an image object, user tag refinement.

INTRODUCTION

People are sharing their personal photos through social network like facebook and flicker. Users are willing to provide tags and comment on the photos while sharing it. In order to retrieve the required photo for user, the search option is limited and unable to retrieve the required photo which was shared in the social network. To retrieve the image, technologies used nowadays are either by content based and keyword based method. By this method, the retrieval of image is quite difficult as user will not get required image

For Content based image retrieval system, Bow model is popular and represents quantizes high dimensional local feature into discrete visual words but this model fails to address the issue related to noisily quantize visual feature and various viewpoint, occulusion, lighting condition, etc., Due to varying condition, the feature of the target images might have different visual words and it also difficult to obtain the desired images through query process.

For keyword-based image retrieval system, textual features such as tags are more semantically relevant than visual features. However, it is still difficult to retrieve all the target images by keywords only because users might know non-specific keywords such as “Trip” . Meanwhile, in most photo sharing websites, tags and other forms of text are freely entered and therefore often inaccurate, wrong or ambiguous.

In response to the above challenges for content-based and keyword-based image retrieval in social media, we propose a general framework, which integrates both visual feature and textual information. In particular, we augment each image in the image collections with semantic additional features that are semantically relevant to the search targets such as specific VWs for certain landmarks or refined tags for certain scenes and events. Aiming at large-scale image collections for serving different queries, we mine the semantic features in an unsupervised manner by incorporating both visual and (noisy) textual information. We construct graphs of images by visual and textual information. We then automatically propagate and select the informative semantic features across the visual and textual graphs The two processes are formulated as optimization formulations iteratively through the subtopics in the image collections. Meanwhile, we also consider the scalability issues by leveraging distributed computation frameworks.

We demonstrate the effectiveness of the proposed framework by applying it to two specific tasks, i.e., image object retrieval and tag refinement. The first task- image object retrieval is a challenging problem because the target object may cover only a small region in the database images as shown in Fig. 1. We apply the semantic feature discovery framework to augment each image with visual word and Second task is that user tag refinement which augments each image with semantically related texts. Similarly,

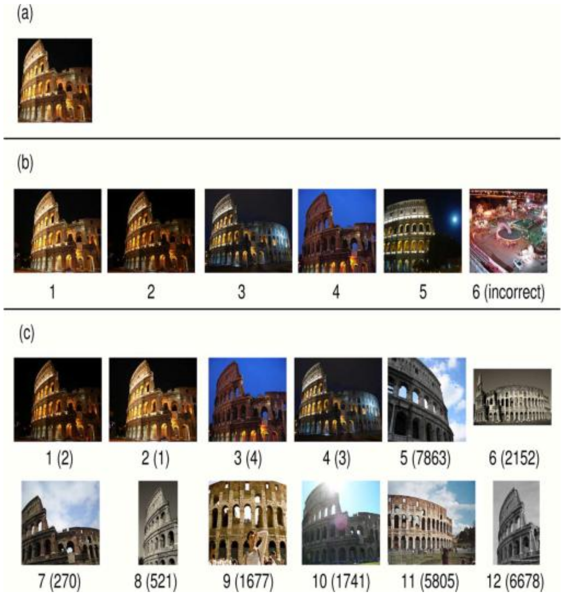

Fig .1. Image object retrieval performance of the BoW model and the proposed approach. (a) An example of object-level query image. (b) The retrieval results of a BoW model, which generally suffers from the low recall rate. (c) The results of the proposed system, which obtains more accurate and diverse images, with the help of automatically feature discovered

we apply the framework on the textual domain by exchanging the role of visual and textual graphs so that we can propagate (in visual graph) and select (in textual graph) relative and representative tags for each image. In particular, the unsupervised auxiliary visual words discovery greatly out performs BoW models and is complementary to conventional pseudo-relevance feedback.

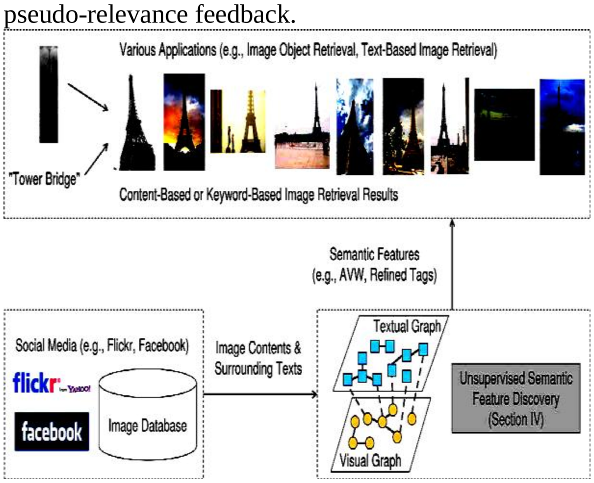

Fig. 2. A system diagram of the proposed method. Based on multiple modalities such as image contents and tags from social media, we propose an supervised semantic feature discovery which exploits both textual and visual feature information. The general framework can discover semantic features in large-scale community-contributed photos. Therefore, we can apply semantic features to various applications.

RELATED WORK

A.SEMANTIC FEATURE DISCOVERY FOR IMAGE OBJECT RETRIEVAL AND TAG REFINEMENT

The image search engines mostly use text information to retrieve images. It is difficult to capture user's intention from the text information. From the images retrieved by text based search, user click on one query image. This query image captures user's intention by using adaptive similarity measure, keyword expansion, image pool expansion and visual query expansion. Re-ranking images based on both visual and textual content improves the performance of the image retrieval system. A set of visual features which are effective and efficient in image search are designed. This approach consists of multiple steps, which can be improved by other techniques. The problem with this system is that sometimes duplicate images show up as similar images to the query. This can be improved by including duplicate detection in the future. To further improve the quality of re-ranked images, this system can be combined with photo quality assessment work. Using visual information to re-rank and improve text based image search results is the best way to improve the performance of the image retrieval system. After query by keyword, user can click on one image, indicating this is the query image. Then all the returned images are re-ranked according to their similarities with the query.

B.IMAGE AUTO ANNOTATION BY SEARCH

Most image search engines use only text information. Users type keywords to find certain type of images. The search engine returns thousands of images ranked by the text keywords extracted from the surrounding text. However, many of returned images are noisy, disorganized, or irrelevant. This is because; the search engine did not use any visual information. The Adaptive Similarity method is proposed in this thesis, which is motivated by the idea that a user always has a specific intention when submitting a query image. The query image is first categorized into one of several predefined categories. Inside each category, a specific weight schema is designed to combine the feature. The correspondence between query image and its proper similarity measurement reflects user intention when using image to query. The specific weighting schema inside each intention category is obtained by minimizing the rank loss for all query images. Using adaptive similarity measurement according to the query image, improves the overall retrieval performance.

C.A TEXT RETRIEVAL APPROACH TO OBJECT MATCHING

In general, images have the related text annotations which could be obtained from where images are stored. So, conventional image retrieval systems utilize the text information of the images, and work as text (keyword) retrieval systems. Some systems use the text and simple image information (eg., image size, image format etc.) and other systems provide the user input interface for relevance feedback. Existing image search systems allow users to search for images via keywords and/or via query by image example. Generally, the system presents pages of representative thumbnail images to the user. The user then marks one or more images as relevant to the query. The visual image features for these images are then used in defining a visual query. However, it is often observed that there are many wrong results from the keyword-based image retrieval. To solve such a problem, integration of results of text and image contents is used. This approach retrieves images using keyword first, and then automatically re-ranks images using visual features of retrieved results. Such a strategy is called Content-based image retrieval (CBIR). In Content-based image retrieval, the image search engine find a set of images from a given image collection that is similar to the given query image. Traditional methods for CBIR are based on a vector space model. These methods represent an image as a set of features and the difference between two images is measured through a similarity function between their feature vectors. Most image retrieval systems are based on features representing color, texture, and shape that are extracted from the image pixels. Color is an important attribute for describing the contents of image. Clustering methods are used to group the images and to select the representative image features. The important hypothesis used in this approach is that the more popular images have the higher probability to be desirable images. Based on this hypothesis, the images are ranked. The advantage of this approach is that, it can use the contents of image in determining the rank of web images. This approach performs better than keyword-based image retrieval.

D.IMPROVING PARTICULAR OBJECT RETRIEVAL IN DATABASES

Nearest neighbor search is the most straightforward approach to find matching image in CBIR. But, nearest neighbor search suffers from a lack of adaptability. The two major drawbacks in nearest neighbor search. First, the nearest neighbor assigns equal weight to both the relevant features and irrelevant features. It is therefore reasonable to select a subset of features or re-weight the features before the nearest neighbor search. This problem is solved by Relevance feedback in this thesis. Through either feature re-weighting or query refinement, relevance feedback provides more accurate retrieval results.

The second drawback to nearest neighbor search is the fixed similarity metric. To solve this problem, an adaptive nearest neighbor algorithm is proposed in which the similarity metric can be locally adapted to the features relevant for each query point and globally optimized using dimensionality reduction. Relevance feedback is an important component when designing image databases. Relevance feedback interactively determines a user's desired output or query concept by asking the user whether certain proposed images are relevant or not. For a relevance feedback algorithm to be effective, it must grasp a user's query concept accurately and quickly, while asking the user to label a small number of images. A support vector machine active learning algorithm for conducting effective relevance feedback for image retrieval is proposed in this paper. This algorithm selects the most informative images to query a user and quickly learns a boundary that separates the images that satisfy the user's query concept from the rest of the dataset. The SVM active learning for image retrieval is particularly well suited to the query refinement task in image retrieval. The use of multi-resolution image feature organization is much helpful for image retrieval.

E.AUTOMATIC QUERY EXPANSION WITH A GENERATIVE FEATURE MODEL FOR OBJECT RETRIEVAL

Shape is the most important requirement at the primitive level. Two main types of shape features are global features such as aspect ratio, moment invariants and local features such as sets of consecutive boundary segments. Shape is a well-defined concept and there is considerable evidence that natural objects are primarily recognized by their shape. In this report, the algorithm chooses a set of patches in an image, and for each patch computes a fixed-length feature vector. This gives a set of vectors per image, where the size of the set can vary from image to image.

PROPOSED WORK

A. Creation of image graphs and image clusters

The proposed framework starts by constructing a graph which embeds image similarities from the image collection. We calculate the image

ISSN 23499842(Online), Volume 1, issue 1 March 2015
International Journal of Innovative Trends and Emerging Technologies

similarity since we observe that most of the textual and visual features are sparse for each image and the correlation between images are sparse as well. To cluster images on the image graph, we apply affinity propagation for graph-based clustering. The images are represented by Visual Words (VW) and text tokens from their associated (noisy) tags. if an image is close to the canonical image (center image), it has a higher AP score, indicating that it is more strongly associated with the cluster.

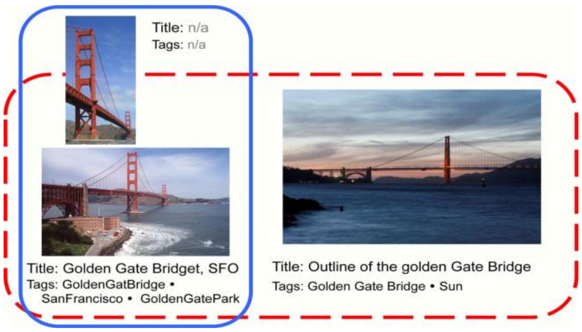

Fig. 3.The roles of semantic related features in image object retrieval. Images in the blue rectangle are visually similar, whereas those images in the red dotted rectangle are textually similar.

B. Auxiliary Visual Word Propagation& Selection

Then based on the image clusters, we propagate auxiliary VWs in and select representative VWs. we propose to augment each image with additional VWs propagated from the visual and textual clusters. Propagating the VWs from both visual and textual domains can enrich the visual descriptions of the images and be beneficial for further image object queries. Though the propagation operation is important to obtain different VWs, it may include too many VWs and thus decrease the precision. To mitigate this effect and remove those irrelevant or noisy VWs, we propose to select those representative VWs in each visual cluster. We observe that images in the same visual cluster are visually similar to each other, therefore, the selection operation is to retain those representative VWs in each visual cluster. Finally, we combine both selection and propagation methods. The propagation operation obtains semantically relevant VWs to improve the recall rate, whereas the selection operation removes visually irrelevant VWs and improves memory usage and efficiency.

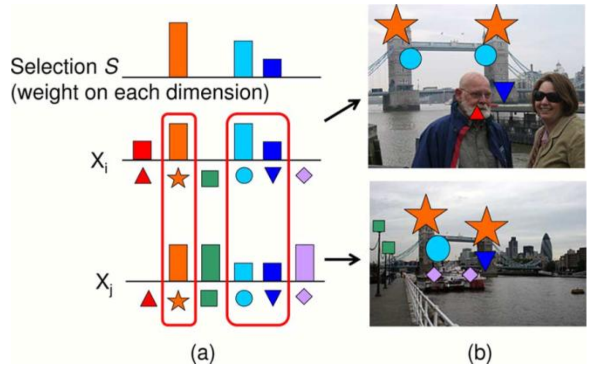

Fig. 4. Illustration of the selection operation for auxiliary visual words. The VWs should be similar in the same visual cluster; therefore, we select those representative visual features (red rectangle). (b) illustrates the importance (or representativeness) for different VWs. And we can further remove some noisy features (fewer representatives) which appeared on the people or boat.

C. Tag refinement

Textual features are generally semantically richer than visual features. However, tags (or photo descriptions) are often missing, inaccurate, or ambiguous as annotated by the amateurs. Traditional keyword-based image retrieval systems are thus limited in retrieving these photos with noisy or missing textual descriptions. Hence, there arise strong needs for effective image annotation and tag refinement. To tackle this problem, we propose to annotate and refine tags by jointly leveraging the visual and textual information. The proposed method concentrates on obtaining more (new) semantically related tags from semantically related images. We further select representative tags to suppress noisy or incorrect tags.

D. Tag Propagation& Selection

In order to obtain more semantically relevant tags for each image, we propose to propagate tags through its visually similar images. We will then remove noisy tags and preserve representative ones. Following the auxiliary feature propagation, we construct the extended textual cluster to propagate relevant tags. After the previous tag propagation step, each image can obtain more different tags. However, it is possible to obtain some incorrect ones. Similar to visual feature selection, we propose to retain important (representative)tags and suppress the incorrect ones. To select important tags for each image, we can directly adopt the same selection formulation as mentioned for images. We select representative tags in each textual cluster since images in the same textual cluster are semantically similar to each other. Through tag selection, we can

ISSN 23499842(Online), Volume 1, issue 1 March 2015
International Journal of Innovative Trends and Emerging Technologies

highlight the representative tags and reject the noisy ones.

E. Image augmentation

We augment each image with AWV additional and important features relevant to the target image by considering semantically related VWs in its textual cluster and representative VWs in its visual cluster.

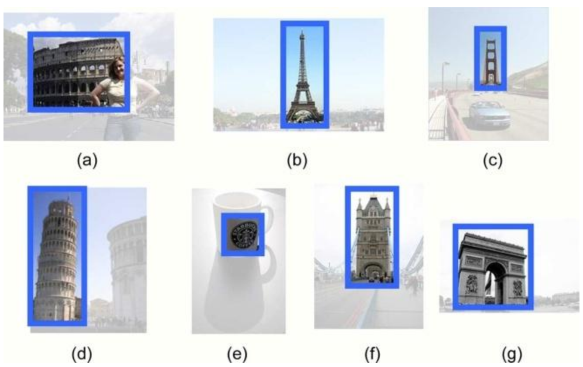

Fig. 5. The query objects are enclosed by the blue rectangles and the corresponding query keywords are listed below each object image. (a) Colosseum, (b) Eiffel tower, (c) Golden gate bridge, (d) Leaning tower of Pisa, (e) Starbucks, (f) Tower bridge, (g) Arc de Triomphe.

EXPERIMENTAL SETUP

A. Dataset

We use Flickr550 as our main dataset in the experiments. To evaluate the proposed approach, we select 56 query images (1282 ground truth images) which belong to the following 7 query categories: Colosseum, Eiffel Tower (Eiffel), Golden Gate Bridge (Golden), Leaning tower of Pisa (Pisa),
Starbucks logo (Starbucks), Tower Bridge (Tower), and Arc de Triomphe (Triomphe). Also, we randomly pick up 10000 images from Flickr550 to form a smaller subset called Flickr11K.

B. Performance Metrics

In the experiments, we use the average precision, a performance metric commonly used in the previous work to evaluate the retrieval accuracy. It approximates the area under a non-interpolated precision-recall curve for a query. A higher average precision indicates better retrieval accuracy. Since average precision only shows the performance for a single image query, we also compute the mean average precision (MAP) over all the queries to evaluate the overall system performance.

C. Evaluation Protocols

As suggested by the previous work, our image object retrieval system adopts 1 million visual words as the basic vocabulary. The retrieval is then conducted by comparing (indexing) the AVW features for each database image. To further improve the recall rate of retrieval results, we apply the query expansion technique of pseudo-relevance feedback, which expands the image query set by taking the top-ranked results as the new query images. This step also helps us understand the impacts of the discovered AVWs because in our system the ranking of retrieved images is related to the associated auxiliary visual words. They are the key for our system to retrieve more diverse and accurate images

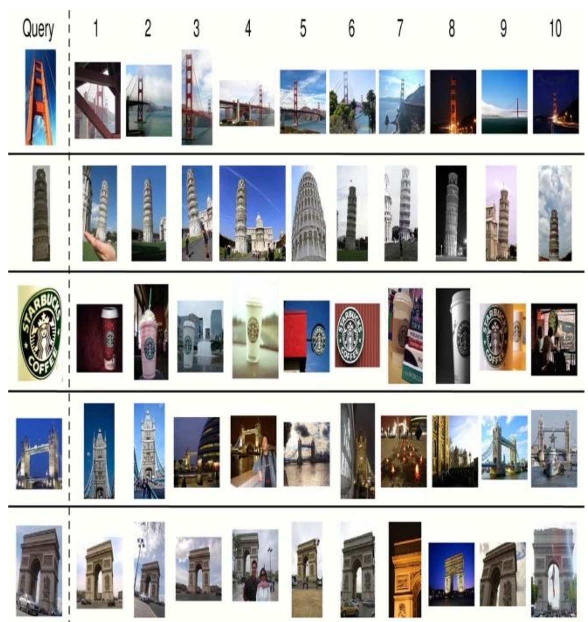

Fig. 6. More search results by auxiliary VWs. The number represents its retrieval ranking. The results show that the proposed AVW method, though conducted in an unsupervised manner in the image collections, can retrieve more diverse and semantic related results.

CONCLUSION

Business organizations are opting for computer in solving business process problems to attain efficiency in operation. A computerized system will produce error free information with less man power and effort. The system has been tested using all possible test data that can work in any conditions.

No programming skill is required to handle the system.

In this work, we present a general framework for semantic feature discovery which utilizes both the visual and textual graphs to propagate and select important (visual or textual) features. First, we show the problems of current BoW model and the needs for semantic visual words to improve the recall rate for image object retrieval. We propose to augment each database image with semantically related auxiliary visual words by propagating and selecting those informative and representative VWs in visual and textual clusters (graphs).Note that we formulate the processes as unsupervised optimization problems. Experimental results show that we can greatly improve the retrieval accuracy compared to the BoW model for image object retrieval. Besides, we extend the proposed method to textual domain. It can not only help to retain representative tags for each image but also automatically derive meaningful tags to annotate unlabeled images. Experiments in text-based image retrieval show that tag refinement can improve the retrieval accuracy effectively (10.7% relatively).

We are to investigate more advanced contextual features, such as geo-tags, time, user attributes, along with the proposed framework to leverage the rich contexts from the emerging social media.

REFERENCES

1. Ames M.and Naaman M.,(2007), ‘Why we tag: Motivations for annotationin mobile and online media,’ in Proc. SIGCHI Conf. Human FactorsComput. Syst., pp. 971–980.
2. Chum O., Philbin J. andZisserman A.,(2007), ‘Total recall:Automatic query expansion with a generative feature model for objectretrieval,’ in Proc. IEEE Int. Conf. Comput. Vis., pp. 1–8.
3. Hays J.and Efros A.,(2008), ‘IM2GPS: Estimating geographic informationfrom a single image,’ in Proc. IEEE Conf. Comput. Vis. PatternRecognit., pp. 1–8.
4. Ma H., Zhu J., King I.,(2010), ‘Bridging the semantic gapbetween image contents and tags,’ IEEE Trans. Multimedia, Vol. 12,no. 5, pp. 462–473.
5. PhilbinJ.,Chum O.,Isard M. and Zisserman A.,(2008),‘Lost inquantization: Improving particular object retrieval in large scale imagedatabases,’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit.,pp. 1–8.
6. Sivic J. and Zisserma A.,(2003), ‘Video Google: A text retrieval approachto object matching in videos,’ in Proc. IEEE Int. Conf. Comput. Vis., Vol. 2, pp. 1470–1477.
7. Wang J., Zhang L. and Jing F.,(2006), ‘Annosearch: Imageautoannotation by search,’ in Proc. IEEE Conf. Comput. Vis. PatternRecognit., pp. 1483–1490.
8. Wang J., Liu M. and Li Y., (2010),‘ARISTA-Imagesearch to annotation on billions of web photos,’ in Proc. IEEE Conf.Comput. Vis. Pattern Recognit., pp. 2987–2994.