

Survey Paper On Different Techniques Of Social Tag Relevance

Deepshikha Mishra¹, Prof. Uday Prtap Singh²

¹ M.Tech Scholar, Dept. of Computer Science and Engineering, LNCT, Bhopal, India

² Professor, Department of Computer Science and Engineering, LNCT, Bhopal, India

Abstract- Social image retrieval is important for exploiting the increasing amounts of amateur-tagged multimedia such as Flickr images. Intuitively, if different persons label similar images using the same tags, these tags are likely to reflect objective aspects of the visual content. Interpreting the relevance of a user-contributed tag with respect to the visual content of an image is an emerging problem in social image retrieval. An algorithm is proposed that scalably and reliably learns tag relevance by accumulating votes from visually similar neighbours. Treated as tag frequency, learned tag relevance is seamlessly embedded into current tag-based social image retrieval paradigms.

Preliminary experiments on two thousand Flickr images demonstrate the potential of the proposed algorithm. The tag relevance learning algorithm substantially improves upon baselines for all the experiments. The results suggest that the proposed algorithm is promising for real-world applications.

Keywords- neighbour voting, tag relevance, user contributed tag, social image tagging.

1. INTRODUCTION

Image sharing websites such as Flickr and Facebook are hosting billions of personal photos. Tagging is a significant feature of social bookmarking systems which enables users to add, annotate, edit and share bookmarks of a web documents. Social image tagging, assigning tags to images by common users, is reshaping the way people manage and access such large-scale visual content. Image tagging basically refers to a process of categorizing or mapping of images on the basis of their contents either visual or context. Along with the rapid growth of personal albums in social networking sites, it has been seen that tagging is the most promising and practical way to facilitate the huge photos database semantically searchable. To tag an image firstly the training set is manually tagged and then the tags of the testing set are automatically predicted.

In a social tagging environment with large and diverse visual content, a light weight or unsupervised learning method which effectively and efficiently estimates tag relevance is required. Two simplest and easiest ways for multi feature tag relevance learning are-the classical borda count and uniform tagger. Image tagging can be done in two ways:-

1. Manual Image Tagging

2. Automatic Image Tagging.

An image retrieval system is a computer system for browsing, searching and retrieving images from a large database of digital images.

Improving Image Tagging- Image tagging can be improved by tagging the images on the basis of their features and tags should be relevant to the image and with the help of which image can be retrieved from pool of the databases.

Improving Image Retrieval- Image retrieval can be improved on the basis of the content as well as the features, characteristics, color etc of the image.

The techniques which are used for image retrieval includes:-

1.1 Content Based Image Retrieval- In content based image retrieval, images are searched and retrieved on the basis of similarity of their visual contents to a query image using features of the image. A feature extraction module is used to extract low-level image features from the images in the collection. Commonly extracted image features include color, texture and shape

1.2 Text Based Image Retrieval- Text-based image retrieval is also called description-based image retrieval. Text-based image retrieval is used to retrieve the XML documents containing the images based on the textual information for a specific multimedia query. To overcome the limitations of CBIR, TBIR represents the visual content of images

by manually assigned keywords/tags. It allows a user to present his/her information need as a textual query, and find the relevant images based on the match between the textual query and the manual annotations of images.

1.3 Multimodal Fusion Image Retrieval-

Multimodal fusion image retrieval involves data fusion and machine learning algorithms. Data fusion, also known as combination of evidence, is a technique of merging multiple sources of evidence. By using multiple modalities, we can learn the skimming effect, chorus effect and dark horse effect.

1.4 Semantic Based Image Retrieval-Image retrieval based on the semantic meaning of the images is currently being explored by many researchers. This is one of the efforts to close the semantic gap problem. In this context, there are two main approaches: Annotating images or image segments with keywords through automatic image annotation or adopting the semantic web initiatives.

2. Related Work

To overcome the problem of image tagging in 2008 Xirong Li, Cees G.M. Snoek, and Marcel Worring proposed a method proposed a novel algorithm that scalably and reliably learns tag relevance by accumulating votes from visually similar neighbors. The advantages of the proposed algorithm are three-fold: 1) reliable, since only common tags are propagated between neighbors without introducing new tags to an image. Such self-validation scheme reduces the risk of incorrectly propagating irrelevant tags; 2) scalable, since the proposed method does not require any model training for any visual concepts; 3) exible, since learned tag relevance information, treated as tag frequency, can be seamlessly embedded into current tag-based retrieval framework.

In 2008 David Grangier and Samy Bengio proposed a algorithm in which a discriminative model for the retrieval of images from text queries. The approach formalizes the retrieval task as a ranking problem and introduces a learning procedure optimizing a criterion related to the ranking performance. The proposed model hence addresses the retrieval problem directly and does not rely on an intermediate image annotation task, which contrasts with previous research. Moreover, the learning procedure builds upon recent work on the online learning of kernel-based classifiers. This yields an efficient scalable algorithm, which can benefit from recent kernels developed for image comparison.

In Eva Horster, Rainer Lienhart, Malcolm Slaney proposed an algorithm in this they have employed the

image content as a source of information to retrieve images. They study the representation of images by Latent Dirichlet Allocation (LDA) models for content-based image retrieval. Image representations are learned in an unsupervised fashion, and each image is modeled as the mixture of topics/object parts depicted in the image. This allows us to put images into subspaces for higher-level reasoning which in turn can be used to find similar images. Different similarity measures based on the described image representation are studied.

In Gustavo Carneiro, Antoni B. Chan, Pedro J. Moreno, and Nuno Vasconcelos proposed a probabilistic formulation for semantic image annotation and retrieval. The supervised formulation is shown to achieve higher accuracy than various previously published methods at a fraction of their computational cost. Finally, the proposed method is shown to be fairly robust to parameter tuning SML has the advantage of combining classification and retrieval optimality with 1) scalability in database and vocabulary sizes, 2) ability to produce a natural ordering for semantic labels at annotation time, and 3) implementation with algorithms that are conceptually simple and do not require prior semantic image segmentation. They have also presented the results of an extensive experimental evaluation, under various previously proposed experimental protocols, which demonstrated superior performance with respect to a sizable number of state-of-the-art methods, for both semantic labeling and retrieval.

In Jing Huan, S Ravi Kumar, Mandar Mitra, Wei-Jing Zhu, Ramin Zabih They define a new image feature called the color correlogram and use it for image indexing and comparison. This feature distills the spatial correlation of colors, and is both effective and inexpensive for content-based image retrieval. The correlogram robustly tolerates large changes in appearance and shape caused by changes in viewing positions, camera zooms, etc. Experimental evidence suggests that this new feature outperforms not only the traditional color histogram method but also the recently proposed histogram refinement methods for image indexing/retrieval. They have described a new image feature that can be used to index and compare images. Since this feature captures the spatial correlation of colors in an image, it is effective in discriminating images. It thus rectifies the major drawbacks of the classical histogram method. The correlogram can also be computed efficiently. The experiments on a large image database evaluated using fair performance measures show that the correlogram performs very well.

In Xirong Li, Le Chen, Lei Zhang, Fuzong Lin, Wei-Ying Ma proposed an algorithm in which they target at solving the automatic image annotation problem in a novel *search* and *mining* framework. Given an uncaptioned image, first in the *search* stage, we perform content-based image retrieval (CBIR) facilitated by high-dimensional indexing to find a set of visually similar images from a large-scale image database. The database consists of images crawled from the World Wide Web with rich annotations, e.g. titles and surrounding text. Then in the *mining* stage, a search result clustering technique is utilized to find most representative keywords from the annotations of the retrieved image subset. These keywords, after salience ranking, are finally used to annotate the uncaptioned image. Based on search technologies, this framework does not impose an explicit training stage, but efficiently leverages large-scale and well-annotated images, and is potentially capable of dealing with unlimited vocabulary. Based on 2.4 million real Web images, comprehensive evaluation of image annotation on Corel and U. Washington image databases show the effectiveness and efficiency of the proposed approach. They have presented a practical and effective image annotation system. They formulated the image annotation as searching for similar images and mining key phrases from the descriptions of the resultant images, based on two key techniques: Multi-Index – a practical solution to index 2.4M Web image database and SRC – the search result clustering technique.

In Theo Gevers and Arnold W. M. Smeulders proposed an algorithm in which the aim is at combining color and shape invariants for indexing and retrieving images. To this end, color models are proposed independent of the object geometry, object pose, and illumination. From these color models, color invariant edges are derived from which shape invariant features are computed. Computational methods are described to combine the color and shape invariants into a unified high-dimensional invariant feature set for discriminatory object retrieval. Experiments have been conducted on a database consisting of 500 images taken from multicolored man-made objects in real world scenes. From the theoretical and experimental results it is concluded that object retrieval based on composite color and shape invariant features provides excellent retrieval accuracy. Object retrieval based on color invariants provides very high retrieval accuracy whereas object retrieval based entirely on shape invariants yields poor discriminative power. Furthermore, the image retrieval scheme is highly robust to partial occlusion, object clutter and a change in the object's pose. From the theoretical and experimental results, it is concluded that object

search based on composite color and shape invariant features provides excellent recognition accuracy. Object search based on color invariants provides very high retrieval accuracy whereas object search based entirely on shape invariants yields poor discriminative power.

In Yong Rui and Thomas S. Huang proposed an algorithm in which a comprehensive survey of the technical achievements in the research area of image retrieval, especially content-based image retrieval, an area that has been so active and prosperous in the past few years. The survey includes 100+ papers covering the research aspects of image feature representation and extraction, multidimensional indexing, and system design, three of the fundamental bases of content-based image retrieval. Furthermore, based on the state-of-the-art technology available now and the demand from real-world applications, open research issues are identified and future promising research directions are suggested. There are three databases in this system architecture. The image collection database contains the raw images for visual display purpose. During different stages of image retrieval, different image resolutions may be needed. In that case, a wavelet-compressed image is a good choice. Image processing and compression research communities contribute to this database. The visual feature database stores the visual features extracted from the images using technique feature extraction. This is the information needed to support content-based image retrieval. Computer vision and image understanding are the research communities contributing to this database. The text annotation database contains the key words and free-text descriptions of the images. It is becoming clear in the image retrieval community that content-based image retrieval is not a replacement of, but rather a complementary component to, the text-based image retrieval. Only the integration of the two can result in satisfactory retrieval performance. The research progress in IR and DBMS is the main thrust to this database.

In Lin Chen, Dong Xu, Ivor W. Tsang, and Jiebo Luo proposed a new tag-based image retrieval framework to improve the retrieval performance of a group of related personal images captured by the same user within a short period of an event by leveraging millions of training web images and their associated rich textual descriptions. For any given query tag (e.g., "car"), the inverted file method is employed to automatically determine the relevant training web images that are associated with the query tag and the irrelevant training web images that are not associated with the query tag. Using these relevant and irrelevant web images as positive and negative training data respectively, we propose a new

classification method called support vector machine (SVM) with augmented features (AFSVM) to learn an adapted classifier by leveraging the prelearned SVM classifiers of popular tags that are associated with a large number of relevant training web images. Treating the decision values of one group of test photos from AFSVM classifiers as the initial relevance scores, in the subsequent group-based refinement process, we propose to use the Laplacian regularized least squares method to further refine the relevance scores of test photos by utilizing the visual similarity of the images within the group. Based on the refined relevance scores, our proposed framework can be readily applied to tag-based image retrieval for a group of raw consumer photos without any textual descriptions or a group of Flickr photos with noisy tags. Moreover, they propose a new method to better calculate the relevance scores for Flickr photos.

3. Comparison of Different Approaches

1) Text Based Image Retrieval-

Text-based image retrieval is used to retrieve the XML documents containing the images based on the textual information for a specific multimedia query.

Advantages-

- a) TBIR represents the visual content of images by manually assigned keywords/tags.
- b) It allows a user to present his/her information need as a textual query, and find the relevant images based on the match between the textual query and the manual annotations of images.

2) Content Based Image Retrieval-

In content based image retrieval images are searched on the basis of the content or we can say on the basis of the features of the image.

Advantages-

- a) A feature extraction module is used to extract images.
- b) Extracted image features include color, texture and shape.
- c) The extracted image features are represented in a multidimensional feature vector, referred to as the image signature.

3) Multimodal Fusion Image Retrieval-

Multimodal fusion image retrieval involves data fusion and machine learning algorithms.

Advantages-

- a) Machine learning algorithms are used to study and classify the combination of modalities that represent images or regions.
- b) The skimming effect is used when the top-ranked documents are fused to increase the

recall and precision of the retrieved documents.

4) Semantic Based Image Retrieval-

A semantic multimedia retrieval system consists of two components-the first component links low level physical attribute to multimedia data to high level semantics class label. The second component is represented by domain knowledge.

Advantages-

- a) Methods for learning and propagating labels are assigned by human users.
- b) Retrieval methods with relevance feedback during a retrieval session.

4. Conclusion

In this paper I presented the different methods used for image tagging and retrieval. Each method applies its own methods and approaches to improve image tagging and retrieval. Each paper I discussed in this paper has its own advantages and disadvantages. Image retrieval and image tagging can be improved on the basis of the content as well as the features, characteristics, color etc of the image. Tagging of images is done on the basis of ranking and ranking is done on the basis of priority. Priority of an image is labeled by the features matching the content of the searched image.

References

1. X. Li, C. G. M. Snoek, and M. Worring, "Learning tag relevance by neighbor voting for social image retrieval," in *Proc. ACM MIR*, 2008, pp. 180–187.
2. D. Grangier and S. Bengio, "A discriminative kernel-based approach to rank images from text queries," *IEEE Trans. PAMI*, vol. 30, no. 8, pp. 1371–1384, 2008.
3. E. H'orster, R. Lienhart, and M. Slaney, "Image retrieval on large-scale image databases," in *Proc. CIVR*, 2007, pp. 17–24.
4. G. Carneiro, A. B. Chan, P. J. Moreno, and N. Vasconcelos, "Supervised learning of semantic classes for image annotation and retrieval," *IEEE Trans. PAMI*, vol. 29, no. 3, pp. 394–410, 2007.
5. J. Huang, S. Kumar, M. Mitra, W. Zhu, and R. Zabih, "Image indexing using color correlograms," in *Proc. CVPR*, 1997, pp. 762–768.
6. X. Li, L. Chen, L. Zhang, F. Lin, and W.-Y. Ma, "Image annotation by large-scale content-based image retrieval," in *Proc. ACM Multimedia*, 2006, pp. 607–610.
7. Theo Gevers and Arnold W. M. Smeulders, "PicToSeek: Combining Color and Shape

- Invariant Features for Image Retrieval”, *IEEE Trans.* vol.9, no.1, Jan 2000.
8. Yong Rui and Thomas S. Huang, “Image Retrieval: Current Techniques, Promising Directions, and Open Issues”, *Journal VCIR* 10, pp. 39–62, 1999.
 9. Lin Chen, Dong Xu, Ivor W. Tsang, and Jiebo Luo, “Tag-Based Image Retrieval Improved by Augmented Features and Group-Based Refinement” *IEEE Trans. Multimedia*, vol. 14, no. 4, Aug 2012.

IJERT