Image Classification using Tag and Segmentation based Retrieval

Shrikant Badghaiya
Department of Computer Science
OIST, College
Bhopal, India

Atul Barve
Department of Computer Science
OIST, College
Bhopal, India

ABSTRACT
In today's scenario when social media sites are widely used and high resolution images are shared. Tagging is an important approach for the retrieval of images in various applications. These are also used to manage personal media data. There are various techniques implemented for the image retrieval such as feature based, histogram based and transformation based but these techniques are not efficient in terms of classification ratio and accuracy. Here in this paper various tag based image technique are discussed so by analyzing the tag based image retrieval techniques a new proposed methodology can be implemented in the future.

Keywords
TBIR, CBIR, tag annotation, AFSVM, NDCG.

1. INTRODUCTION
Due to past years we have been examined that an internet traffic enduring increase in image data, most important to huge repositories. Due to progressively image have been produced in digital form around the globe image retrieval attracts among studies in the field of digital libraries, image processing application, multimedia data, remote sensing information and various related image retrieval application. Content-based image retrieval methods have made an effort to ease the access to image data. To meeting, several feature extraction techniques have been proposed with the intention of get better the quality of content-based image retrieval and image classification systems. There have been enormous responsibilities to generate text-based image search engines by quite a lot of companies e.g. Google, Yahoo or Microsoft. Even though the pure text-based search performs before now to a certain extent well, the image search results are not adequate at all outcomes, since text cannot sufficiently describe an image's content, e.g. imagine an expressionistic painting and its features. Hence, image retrieval systems exclusively based on textual search often perform poorly. A variety of kinds of CBIR features can be grouped into the following categories: color with/without spatial constraints [1], texture [2], shape and structure [3].

The content based image retrieval (CBIR) plans to get back visually comparable of given images from image database for a given query image (or sketch) using image retrieval system. Retrieval of mandatory-query-similar images from in large quantities obtainable and easy to get to use digital images is a demanding need of today image retrieval application. The image retrieval techniques based on visual image content has been in-focus for more than a decade. There is an increasing interest in developing new methods to help users retrieve their personal photos including the raw photos that are not associated with any textual descriptions. Most image retrieval methods can be classified into two categories: content-based image retrieval [5], [6] (CBIR) and keyword/tag-based image retrieval [7], [8] (TBIR). Many search-engines, including modern web-search-engines retrieve similar images by searching and matching textual metadata associated with digital images. For quicker response time, association of metadata is carried out as an off-line process known as image-annotation.

The need for image retrieval evolved two solutions – image annotation or image tagging and content based image retrieval. The annotation of images is in the form of textual explanation, generally these images are carried out either automatically or manually e.g. Google image labeler. The automatic image marginal note tagging analyses the image content for producing and associating textual description with images under contemplations. The textual query is then matched with annotations for image retrieval in the image retrieval system. The content based image retrieval techniques aspire to get answer to a given query image or sketch with query-similar resultant images obtained from the image database, as shown in Fig.1.

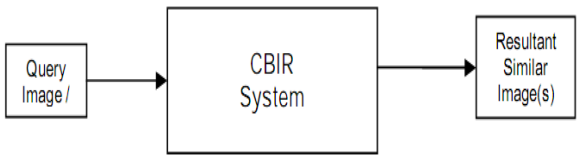

Fig.1: Block diagram of a CBIR system.

The detailed block diagram of a CBIR system available in the paper is shown in the Fig.2. The system consists of two databases – image database and feature database. The image feature database is a group of the image features on which image search is to be conceded out. Whereas adding a new image in the image database, image features are taken out from the image database and stored in the feature database. The features are represented in a convenient format before storing them appropriately for faster search. Features of the query image are extracted and compared with the features of the images, available in feature database. The similarity comparison and searched image are carried out with the image feature database for finding similar-featured-images.

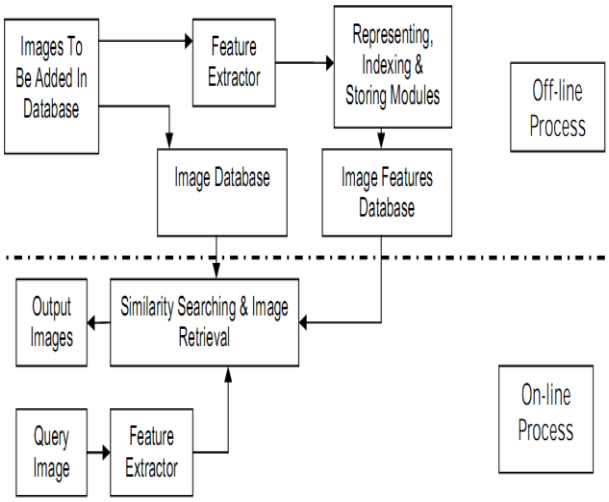

Fig.2: Detailed block diagram of a CBIR system.

Computer vision methods and image processing research area is the image retrieval problem of CBIR technique that is the difficulty of looking for digital images in large databases set. Content based represents that the search will investigate the actual contents of the image dataset to some extent than the metadata i.e. keywords, tags, and/or descriptions join together with the given image. In CBIR, images are indexed by their visual content, such as color, texture, shape. Also having humans by hand go through keywords for images in a large database can be incompetent, and may not imprison every keyword that illustrates the digital image. Thus a retrieval system that can filter images based on their content would provide better indexing and return more correct results [4].

2. LITERATURE SURVEY

Xian Zhou and Thomas Huang [9] propose the method of similarity matrix, which has keywords as columns and rows. Similarity matrix is used to find out the relation among the keywords. Assumption is that the keywords occurring frequently in the same documents are relevant. Authors have assumed that some of the images in the database have been already annotated through the procedure of document analysis i.e. text data in HTML pages or manual annotation. Then the weights are improved over time depending on relevance feedback from the user. They used color moments and color histograms to represent color content and Tamura's features, wavelet moments and co-occurrence matrix to represent texture. Structural feature is represented by using water filling feature and salient points.

In this paper [10], they proposed a new method on Social image analysis and retrieval is significant for selecting people organize and right of entry the growing quantity of user-tagged multimedia data. In view of the fact that user tagging is acknowledged to be abandoned, indefinite, and excessively adapted an elementary difficulty of user tagged data is how to interpret the relevance of a user-contributed tag with respect to the visual content the tag is describing. Intuitively, if different persons label visually similar images using the same tags, these tags are likely to reflect objective aspects of the visual content. Starting from this intuition, here they propose in this paper a neighbour voting algorithm which accurately and efficiently learns tag relevance by accumulating votes from visual neighbours. Under a set of well defined and realistic assumptions, we prove that our algorithm is a good tag relevance measurement for both image ranking and tag ranking. Three experiments on 3.5 million Flickr photos demonstrate the general applicability of our algorithm in both social image retrieval and image tag suggestion. Our tag relevance learning algorithm substantially improves upon baselines for all the experiments [10]. Since user tagging is known to be subjective and overly personalized, a fundamental problem in social image analysis and retrieval is how to accurately interpret the relevance of a tag with respect to the visual content the tag is describing. In this paper, we propose a neighbour voting algorithm as an initial step towards conquering the problem. Our key idea is to learn the relevance of a tag with respect to an image from tagging behaviours of visual neighbours of that image. In particular, our algorithm estimates tag relevance by counting neighbour votes on tags. We show that when (1) the probability of correct user tagging is larger than the probability of incorrect user tagging and (2) content-based visual search is better than random sampling, our algorithm produces a good tag relevance measurement for both image ranking and tag ranking. Moreover, since the proposed algorithm does not require any model training for any visual concept, it is efficient in handling large-scale image data sets. Our study demonstrates that the proposed algorithm predicts more relevant tags even when the visual search is unsatisfactory. In summary, all the three experiments show the general applicability of tag relevance learning for both image ranking and tag ranking.

Liu et al. [11] proposed a constrained nonnegative matrix factorization method for multi-label learning. Unlike the existing approaches for multi-label learning that assume complete and perfect class assignments, the offered approach is able to deal with noisy and incorrect tags assigned to the images. Even though a matrix completion scheme was proposed for transductive categorization, it is different from the offered work in that it applies Euclidean distance to measure the difference between two training illustrations, despite the fact that the proposed approach commences a distance metric to better capture the similarity between two instances.

Lin Chen et al [12] proposed a new method Tag-Based Image Retrieval Improved by Augmented Features and Group-Based Refinement. Here author try to build up a new tag-based image retrieval (TBIR) structure to get better the image retrieval routine of personal images captured by the equivalent user and well categorized by user group or organization. Mainly these methods can be groups of unprocessed end user photos not joined together with any textual explanations or groups of *Flickr* photos that are associated with noisy attach a label to. For any specified image retrieval query tag (for e.g., "car"), the reversed file technique is make use of without human intervention agreed on the appropriate training data set on web images that are associated with the query tag and the inappropriate training data set on web images that are also not correlated with the query tag. By means of these significant and unrelated web images as positive and negative training data set in that order, they recommend a new arrangement technique called support vector machine with enhanced features (AFSVM) technique to become skilled at an turn out to be accustomed classifier by leveraging the pre-learned SVM classifiers method of an appropriate accepted tags that are related with a large number of significant training data set on web images [12]. It is concentrating that the appropriate final answer of AFSVM method is to retrain SVM classifiers method again it is based on expanded characteristics which unite with the original qualities and the pronouncement importance get hold of from the pre-learned SVM classifiers method of relevant well-liked tags or keyword.

Based on the some hypothesis that the end user photos confined by the equivalent consumer within a small stage of impressive occasion and related action can be think about to form a semantically related group of raw consumer photos and a new dataset gathered after that they present to use the Laplacian make conformed least squares (LapRLS) technique in the subsequent group-based get rid of impurities to different phase. And then it may take care of the decision values of one collection of examined photos from AFSVM classifiers as their preliminary significance keep counts and then apply LapRLS to refine the initial relevance scores of test consumer photos by make the most of the illustration resemblance of the images dataset within the particular group. Mainly these scores are based on the distilled very well at particular significance scores; they can straightforwardly rank the raw consumer photos or *Flickr* photos with noisy tags images. To get enhanced re-rank of these consumer photos of the *Flickr* photos and finally they additional recommend a new technique to get superior calculation with the relevance score by furthermore considering the total number of tags created by the *Flickr* users [12]. Experimental result using the Kodak dataset shows that contains a group of raw consumer photos and a new dataset collected from Flickr.com. The dataset are effectiveness of this structure including these classification methods, the group-based sophistication method using LapRLS, and the new relevance score for re-ranking Flickr photos [12].

Aixin Sun et al [13] proposed Tag-Based Social Image Retrieval: An Empirical Evaluation. Take the first step to propose a generic, flexible and extensible framework for TagIR and undertake a systematic and comprehensive empirical evaluation of 288 search methods in total and systematically compared their performances using Precision at top-K, Mean Average Precision, Recall and Normalized Discounted Cumulative Gain for ranking images using this framework. This framework consists of five orthogonal dimensions that play pivotal roles in social image tagging, namely, tag relatedness for measuring the degree of effectiveness of a tag describing the tagged image, tag favoritism for enumerating the grade of discrimination of a tag with respect to the entire tagged image collection, tag length normalization corresponding to document length normalization in web search, tag-query matching model for computing a matching score between an image tag and a query tag, and query model for rewriting tag queries [13].

Their experimental results suggest that for single-tag queries, tag relatedness, tag-query matching model, and query model are the most crucial dimensions for superior TagIR experiences. However, for multi-tag queries, where the information need is much more specifically defined, all these dimensions become significantly less important. The presence of all tags matching a multi-tag query largely guarantees a very good ranking of search outcomes. Further complex formulations in excess of these dimensions typically lead to degradation of search results ranking [13].

3. PROPOSED METHODOLOGY

1. Query for the input image database.
2. Apply classification algorithm for the classification of images into groups.
3. Generate Tag matrix for each of the classified image.
4. Apply Tag matrix Completion algorithm.
5. Generate and retrieve final images according to tags.

4. EXPECTED OTUCOME

The Proposed technique implemented here can be concluded on the basis of following parameters:

1. Precision: it is defined as the total image correctly retrieved using tag matrix based completion to the total images retrieved.
2. Recall: It is defined as the total images correctly retrieved to the total images in the dataset.
3. Final Score : $\frac{2*precision\ *recall}{precision\ +recall}$
4. Accuracy: it is defined as the total number of image retrieved correctly classified from the dataset.
5. Error rate: it is defined as the change or difference in original image to the final or resultant image.
6. Retrieval rate: it is defined as the total images correctly retrieved from the dataset.

5. CONCLUSION

With the attractiveness of digital cameras and other social media, recent years have witnessed a rapid growth of personal photo albums. People capture photos to record their lives and share them on the web. The photo sharing and tagging are widely used in current scenario. In this paper we are presenting a review on a tag completion and retrieval of image. Here in this paper various classification techniques are discussed with their issues and advantages.

The classification of images using tag matrix completion is an efficient technique and further improvement can be done in the methodology for the better accuracy and classification of the images. In the future an efficient technique is implemented for the classification of images so that the retrieval rate increases.

6. REFERENCES

[1] R. Veltkamp, H. Burkhardt, and H.-P. Kriegel. State-of-the-Art in Content-Based Image and Video Retrieval. Kluwer Academic Publishers, 2001.

[2] P. Howarth, A. Yavlinsky, D. Heesch, and S. M. Rger. Medical image retrieval using texture, locality and colour. In C. Peters, P Clough, J Gonzalo, G.J.F. Jones, M. Kluck, and B. Magnini, editors, Multilingual Information Access for Text, Speech and Images, 5th Workshop of the Cross-Language Evaluation Forum, pages 740–749, 2004.

[3] A. K. Jain and A. Vailaya. Image retrieval using color and shape. Pattern Recognition, 29(8):1233–1244, August 1996.

[4] J. Eakins and Margaret Graham, "Content based image retrieval," university of Northumbria at Newcastle.

[5] M.S. Lew, "Content-Based Multimedia Information Retrieval: State of the Art and Challenges," ACM Trans. Multimedia Computing, Comm. and Applications, vol. 2, pp. 1-19, 2006.

[6] L. Yang, R. Jin, L. Mummert, R. Sukthankar, A. Goode, B. Zheng, S.C.H. Hoi, and M. Satyanarayanan, "A Boosting Framework for Visuality-Preserving Distance Metric Learning and Its Application to Medical Image Retrieval," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 32, no. 1, pp. 30-44, Jan. 2010.

[7] L. Chen, D. Xu, I.W. Tsang, and J. Luo, "Tag-Based Web Photo Retrieval Improved by Batch Mode Re-Tagging," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 3440-3446, 2010.

[8] Y. Liu, D. Xu, I.W. Tsang, and J. Luo, "Textual Query of Personal Photos Facilitated by Large-Scale Web Data," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 33, no. 5 pp. 1022-1036, May 2011.

[9] Zhou X.S., Thomas S. Huang "Unifying Keywords and Visual Contents in Image Retrieval", IEEE Multimedia, Vol. 9, Issue 2, April 2002 pages 23-33.

[10] Xirong Li, Cees G. M. Snoek, Marcel Worring, "Learning Social Tag Relevance by Neighbor Voting". IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 11, NO. 7, NOVEMBER 2009.

[11] Y. Liu, R. Jin, and L. Yang, "Semi-Supervised Multi-Label Learning by Constrained Non-Negative Matrix Factorization," Proc. 21st Nat'l Conf. Artificial Intelligence, pp. 421-426, 2006.

[12] Lin Chen, Dong Xu, Ivor W. Tsang, and Jiebo Luo "Tag-Based Image Retrieval Improved by Augmented Features and Group-Based Refinement", IEEE Transactions On Multimedia, Vol. 14, No. 4, pp. 1057 – 1067, August 2012.

[13] Aixin Sun, Sourav S. Bhowmick, and Khanh Tran Nam Nguyen and Ge Bai "Tag-Based Social Image Retrieval: An Empirical Evaluation", Journal Of The American Society For Information Science And Technology, vol. 62, issue 12, pp. 2364–2381, 2011.