



Survey on Semantic Based Image Retrieval Using Image Annotation

Seena Elizabeth Philip*

PG Scholar

Dept. of Computer Science and Engg
Karunya University, India

Kumudha Raimond

Professor

Dept. of Computer Science and Engg
Karunya University, India

Abstract—Data mining is one of the areas in computer science which is being used for image mining. According to the rapid growth of multimedia image storing, there is a need for image retrieval. Image retrievals are of three types: text based image retrieval, content based image retrieval and semantic based image retrieval. Semantic based image retrieval is a method which uses the semantic meaning of the images in response to the user's query. Many image annotation techniques and retrieval methods have emerged recently. The main idea of Automatic Image Annotation (AIA) technique is to automatically learn semantic concept models from large number of image samples and use the concept models to label new images. The aim of this paper is to study the various methods used in semantic based image retrieval.

Keywords— Multi-labelling, Image annotation, Semantic Template, SBIR

I. INTRODUCTION

Many research works are going on in the area of data mining. Data mining is the process of extracting features from a collection of large database. It can be either automatic or semiautomatic. Image mining also plays a vital role in every aspect because there is a rapid growth in storing multimedia images.

Image retrieval is one of the emerging areas of research because of the rapid growth of image storing. Many methods have been implemented to retrieve images as shown in Fig. 1; first approach is the text based image retrieval system (TBIR) in which retrieval is based on the keywords provided in the description. The drawback is that it is impractical to annotate huge amounts of data and also it consumes more time. In order to overcome this difficulty content based image retrieval (CBIR) or visual based image retrieval system (VBIR) came into existence, where images are automatically indexed and retrieved with low level content features like colour, shape and texture. The drawback is that it did not seek for the concepts, i.e., there is a significant gap between the low level content features and semantic concepts used by humans to interpret the images. Now in recent days research is going on in the area of semantic based image retrieval (SBIR), where retrieval is fully based on the semantic meaning of the images, i.e., it looks for the high level semantics within the image regardless of their low level features. AIA is used in SBIR, the main idea of AIA technique is to automatically learn semantic concept models from large samples and use the concept models to label new images. Once the images are annotated with semantic labels, images can be retrieved by keywords. It employs the advantages of both the text based annotation and content based annotation.

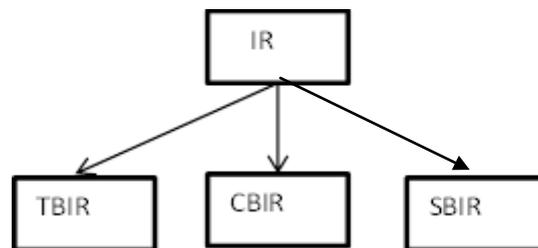


Fig. 1 Classifications of image retrieval

This paper makes an attempt to introduce this rising space to the researchers for the quick summary of current analysis, its progress and framework.

II. DIFFERENT METHODS USED FOR SEMANTIC BASED IMAGE RETRIEVAL

In [1], a region based image retrieval system with high level semantics was proposed. In CBIR, images are indexed by their own visual contents such as colour, texture and shape, therefore it lacks retrieval accuracy. In this semantic template is used for avoiding the problem of feature discretization. The difference from other methods is that it introduces a hybrid tree simplification method to handle the noise and tree fragmentation problems and thus improves the classification performance of the tree. If we query an image with specified region, first it extracts the low level features and hybrid level concept of the specified region and then the images are retrieved based on the concept similar as that of

the query region from the image database. Low level feature is normalized and similarity is measured between retrieved images and query image. Finally, retrieved and ranked images are given back to the user.

In [2], a clustering based multi-labelling classification for image annotation and retrieval was proposed. In this clustering based classification, the original training set is divided into several disjoint clusters of data, then for each cluster a multi-label classifier is modelled. First it groups the training data into user-specified number of clusters k , using clustering algorithm, labels are not considered in this step. Then at the second step clustering based multi-label classifier (CBMLC) uses the multi-label algorithm on the data of each cluster and produces k multi-label classification models. For the classification of a new instance, it first finds the closest cluster of instance and uses the corresponding multi-label model to classify. The drawback is that, it is fully based on the region, not on the concept or semantics of the image. In [3], a region based supervised annotation for semantic image retrieval is proposed. A key to retrieve semantic based image is AIA. For any new image to be annotated, the association is applied on its low level features to automatically obtain the semantic labels. In this a supervised image annotation method based on regional features is proposed, for this supervised multi-class labelling (SML) and multi-instance learning techniques (MIL) are used. The relationship between low level features and semantic concepts is found from the training set and then it is modified using relevance with other regions in the image. This is tested on COREL & TRECVID datasets and achieves good annotation. But it is not dealing with any retrieval process of an image.

In [4], an approach based on probabilistic latent semantic analysis (PLSA) to achieve the task of AIA and retrieval is adopted. For training, the images are represented as a bag of visual words, and then a probabilistic framework is designed for capturing semantic aspects from visual and textual data. An adaptive asymmetric learning algorithm is used to fuse both the data. The process of adaptive learning is shown in Fig 2.

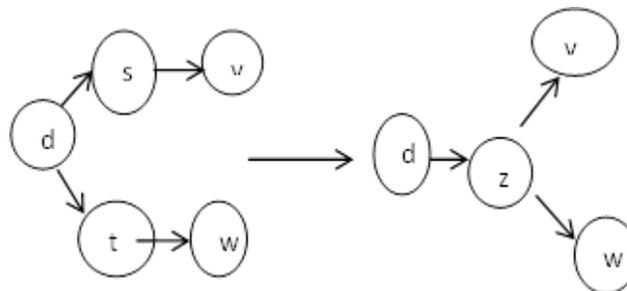


Fig-2: Adaptive learning process

Where v is the visual word, d is the caption of an image, w is the count of the word, z is the aspect distribution, s and t are aspects of two models

In [5], a method for SBIR using Decision Tree (DT) in which the problem with feature discretization is solved by using semantic template is proposed. For this a DT is proposed, which can be obtained by splitting the training data into different subsets based on all the possible values of an attribute. The root node contains all the instances in the training set. Then, an input attribute is selected to split the training set into different subsets corresponding to different possible values of the attribute. This process is continued to each derived subset up to the leaves of the tree. There are different ways to split the data resulting in different trees. First, we have to select the most significant attribute for decision making, for the DT algorithms like $c4.5$ and $ID3$ selects the attribute with highest information gain by first calculating the entropy value. Given a set S , containing 'm' possible outcomes, the entropy of set S is defined as

$$H(S) = \sum_{i=1}^m - P_i * \log_2 P_i \quad (1)$$

Where P_i is the proportion of instances in S that takes the i^{th} value of outcome. Information gain measures the reduction in entropy (gain in information) induced by splitting the dataset on a certain attribute E .

$$\text{Gain}(S,E) = H(S) - \sum_{v \in E} \frac{|S_v|}{|S|} H(S_v) \quad (2)$$

Where v is a value of E , S_v is the subset of instances of S with E having value v , S is the number of instances in S . After this rules are derived for this and images are retrieved by entering appropriate keyword and images are retrieved based on concepts. This is a strong tool for learning image semantics. The dataset used is COREL image set. But it is yet an open research question in determining the best learning tool for image semantic learning. Also it is difficult to learn image concepts in other domains by using different types of image features and defining new semantic template. Liag Chen Y et.al., in [6], reported that most DT classifiers are designed to classify the data with categorical or Boolean class labels. But many practical problems' data and its class labels are naturally organized as a hierarchical structure. In the hierarchy, the upper levels are less specific but easier to predict, while the lower levels are more specific but harder to predict. To build a DT, it is important to determine the way to classify the data so that the class labels can be as specific as possible while also ensuring the highest possible prediction accuracy. This paper has proposed a novel DT based classification algorithm with hierarchical class labels. HLC (hierarchical labelling classifier) algorithm is used for this purpose. It follows the standard framework of classical DT induction methods, such as $ID3$. U.C. Irvine repository is used as a database.

In [7], a method for combining classification and regression techniques in DT for multi labelling image annotation is proposed. Two main steps are followed: one for constructing DT and the other for image retrieval (IR). It avoids the problem with semantic template. A new algorithm called Hierarchical Decision Tree (HDT) is used. In the DT construction process, first the data are normalized; it is done for reducing redundancy and dependency. Next the continuous values are discretised, then the rules are extracted and finally ontology is prepared. After the DT construction gets over all the concepts and their ranks, the approach will turn from single labelling to multi labelling. From the ranks, it is easy to find how much a tuple belongs to a class. To retrieve the images, three steps are followed first tuples are matched with rules, then indices are created for concepts and finally the system is queried to extract the images. A standard dataset SAIAPR-TC12 is used and the performance is measured by using precision, recall and F-measure. The result shows that it is very effective in multi-labelling approach for semantic based image retrieval system.

III. SOME OF THE IMAGE ANNOTATION TECHNIQUES

The majority of machine learning systems for object recognition are limited by their requirement of single labelled images for training, which are difficult to create or obtain in quantity. In [8], a new object recognition system named MLBoost is proposed which learns from multi-label data through boosting and improves on state-of-the-art multi-label annotation and labelling systems. The system is trained using images with accompanying text and the process is unsupervised. Having once been trained it is able to give segment labels and a list of descriptive words (an annotation) for any novel image. MLBoost algorithm is used to learn different weak learners through training data. COREL dataset is used and no quantitative evaluation measures are used, since hand labelling each segment of all images is necessary. An advantage is that it exploits the availability of the captions of images. Disadvantage is that the time complexity and space complexity is more.

In [9], a novel Multi-label Correlated Green's Function (MCGF) approach is used to annotate images over graphs by employing the labels correlations. A new adaptive decision boundary method for multi-label assignment is also proposed and integrated into MCGF algorithm to solve the label assignment ambiguity problem in semi supervised learning. The new semi-supervised method can easily combine other image feature extraction and distance metric learning methods to improve the image annotation performance. The new MCGF method is employed to annotate image/video on three commonly used data sets: TRECVID, Microsoft Research Cambridge data set and Barcelona data set. Average precision is used for evaluation. By incorporating the correlations among labels, the performance has improved significantly. Over all three data sets, the new method outperforms four other state-of-the-art methods.

In [10], multiple targets classification system for visual concepts detection and image annotation is presented. Multiple targets classification (MTC) is a variant of classification where an instance may belong to multiple classes at the same time. The system is composed of two parts: feature extraction and classification/annotation. The feature extraction part provides global and local descriptions of the images. These descriptions are then used to learn a classifier and to annotate an image with the corresponding concepts. For this a Predictive Clustering Trees (PCTs) are used, which are capable to classify an instance to multiple classes at once, thus exploiting the interactions that may occur among the different visual concepts (classes). Moreover, ensembles (random forests) of PCTs are constructed, to improve the predictive performance. Testing is done on the image database from the visual concept detection and annotation task part of ImageCLEF 2010. The extensive experiments conducted on the benchmark database show that the system has very high predictive performance and can be easily scaled to large number of images and visual concepts. The hierarchical multi-label classification is a variant of traditional classification in which the instances can belong to several labels that are in turn organized in a hierarchy. Existing hierarchical multi-label classification algorithms ignore possible correlations between the labels. Moreover, most of the current methods predict instance labels in a "flat" fashion without employing the ontological structures among the classes. In [11], a HiBLADE (Hierarchical multi-label Boosting with LAbel DEpendency), a novel algorithm that takes advantage of not only the pre-established hierarchical taxonomy of the classes, but also effectively exploits the hidden correlation among the classes that is not shown through the class hierarchy, thereby improving the quality of the predictions. According to this approach, first, the pre-defined hierarchical taxonomy of the labels is used to decide upon the training set for each classifier. Second, the dependencies of the children for each label in the hierarchy are captured and analysed using Bayes method and instance-based similarities. F-measure is used for evaluation. The experimental results showed that the algorithm HiBLADE outperforms the flat classification method and the local classifiers method that builds independent classifier for each class.

Semantic hierarchies have been introduced recently to improve image annotation. This is used as a framework for hierarchical image classification, and thus to improve classifiers accuracy and reduce the complexity of managing large scale data. In [12], a method for semantic hierarchies for hierarchical image classification is proposed. First a method based on the hierarchical structure to train hierarchical classifiers called One-Versus-Opposite-Nodes is used, which allows decomposing the problem into several independent tasks and therefore scales well with large databases. Also two methods are used for computing a hierarchical decision function that serves to annotate new image samples. The former is performed by a top-down classifiers voting, while the second is based on a bottom-up score fusion. The experiments on Pascal VOC'2010 dataset showed that the methods improved the image annotation results well. Precision and recall are used for evaluation.

V. CONCLUSION

This paper tries to cover different methods used for SBIR. All the methods tried to retrieve images based on their semantic meaning. But there is a problem with semantic discretization. But we found that multi-labelling image annotation was achieved by combination of classification and regression in HDT. Instead of selecting just one class with

highest rank, all concepts and their ranks were considered in final leaf nodes. Hierarchical network was employed to represent the relationships of semantic classes. The system covered the weakness of single-labelling approach in establishing trade-off between accuracy and details. Multi-labelling approach considered all concepts within a hierarchical structure and the images were retrieving according to their ranks. Semantic template approach was employed to handle data discretization problem. It is found that retrieval using HDT is the best solution for semantic based image retrieval.

ACKNOWLEDGMENT

At the outset I express my gratitude to the almighty god who has been with me during each and every step that I have taken towards the completion of this survey paper. I feel it pleasure to place my heart-felt thanks and gratitude to my guide to Dr. KUMUDHA RAIMOND M.E, Ph.D., for her great supervision and encouragement during this work

REFERENCES

- [1] Liu Ying, Dengsheng Zhang, Guojun Lu, *Region-based image retrieval with high-level semantics using decision tree* Learning, Pattern Recognition 41 (2008) 2554 – 2570
- [2] Gulisong Nasierding, Grigorios Tsoumakas, Abbas Kouzani: *Clustering Based Multi-Label Classification for Image Annotation and Retrieval*. In: Proc. 2009 IEEE International Conference on Systems, Man, and Cybernetics (SMC 2009), 2009.
- [3] Shi Fei, Jiajun Wang, Zhiyong Wang, *Region-based supervised annotation for semantic image retrieval*, Int. J. Electron. Commun. (AEÜ) 65 (2011) 929– 936
- [4] Li Zhixin, Zhiping Shi, Xi Liu , Zhiqing Li , Zhongzhi Shi, *Fusing semantic aspects for image annotation and retrieval*, J. Vis. Commun. Image R. 21 (2010) 798–805
- [5] Liu Ying, Dengsheng Zhang, Guojun Lu, and Ah-Hwee Tan, *Integrating Semantic Templates with Decision Tree for Image Semantic Learning*, T.-J. Cham et al. (Eds.): MMM 2007, LNCS 4352, Part II, pp. 185 – 195, 2007.© Springer-Verlag Berlin Heidelberg 2007
- [6] Liag Chen Y, Hsiao-Wei Hu, Kwei Tang, *Constructing a decision tree from data with hierarchical class labels*, Expert Systems with Applications 36 (2009) 4838–4847
- [7] Ali Fakhari, Amir Masoud Eftekhari Moghadam, *Combination of classification and regression in decision tree for multi-labeling image annotation and retrieval*, Applied Soft Computing 13 (2013) 1292–1302.
- [8] Johnson M, R Cipolla - BMVC, 2005 - userweb.cs.utexas.edu, *Improved Image Annotation and Labeling through Multi-label Boosting*.
- [9] H Wang, H Huang, C Ding - Computer Vision, 2009 IEEE 12th international conference, 2009 ieeexplore.ieee.org, *Image annotation using multi-label correlated Green's function*.
- [10] Dimitrovski Ivica, Dragi Kocev, Suzana Loskovska, and Sašo Džeroski, *Detection of Visual Concepts and Annotation of Images using Predictive Clustering Trees*, ICPR 2010, LCNS 6388, pp 152-161, springer 2010.
- [11] Alaydie Noor, Chandan K. Reddy, and Farshad Fotouhi, *Exploiting Label Dependency for Hierarchical Multi-label Classification*.
- [12] Bannour Hichem, Céline Hudelot, *Hierarchical Image Annotation Using Semantic Hierarchies*, 21st ACM international conference on information and knowledge management, pages 2431-2434.