International Journal of Engineering Research and General Science Volume 2, Issue 6, October-November, 2014
ISSN 2091-2730

Surveyof Knowledge Discovery for Object Representation with Spatio-Semantic Feature Integration

Madhuri B. Dhas[1] Prof.S.A.Shinde[2]
[1]PG Student [2]Asistant professor
[1] Department Of Computer Engineering [2]Department Of Computer Engineering
VPCOE, Baramati,Savitribi Phule University VPCOE,Baramati,Savitribi Phule University
Baramati, India Baramati, India

madhuri.dhas13@gmail.com, shinde.meetsan@gmail.com

ABSTRACT:A social media network is becoming popular these days, where user interacts with each other to form social networks. The photo sharing websites include Flicker, Picasa, YouTube support users to create, annotate, share, detect and comment on Media data. The terms data mining, knowledge discovery are related to this web multimedia objects. Multimedia object classification is a necessary step of multimedia information retrieval. There is urgent need to efficiently index and organize these web objects, to facilitate convenient browsing and search of the objects, and to effectively reveal interesting patterns from the objects. For all these tasks, classifying the web objects into manipulable semantic categories is an essential preprocessing procedure. One important issue for classification of objects is the representation of images. There exist an excessive number of visual features proposed for the representation of images. To perform supervised classification tasks, the knowledge is extracted from unlabeled objects through unsupervised learning. Instead of using traditional Bag-of-words (BoW), a higher level image representation model called Bag-of-visual-phrases(BoP) is used to incorporate the spatial and semantic information.

KEYWORDS
Knowledge Discovery, Correlation Knowledge, Spatio-Semantic Feature Integration.

I.INTRODUCTION

Recently the task of image retrieval has received more attention from the web community and web multimedia retrieval since there are so many useful images on web pages. Image retrieval is a branch of information retrieval whose task is to retrieve some pieces of information to fulfill a user's information needs according to semantic relevance measurements.In multimedia retrieval process Knowledge discovery concerns the entire knowledge extraction phase, including storage of data and accessing that data, what are the efficient and scalable algorithms used to analyze massive datasets, interpretation and visualization of the results, interaction between human and machine. It also deals with the support for learning and analyzing the application domain.

Multimedia mining and retrieval deals with the extraction of prior knowledge, multimedia data relationships, or other patterns that are not explicitly stored in multimediafiles. In retrieval process, indexing and classification of multimedia data with efficient information fusion of the different modalities is essential for improvement of system's overall performance. An automated method for generating images annotations, taking into accounttheir visual features. The semantic rules map the combinations of visual characteristics (colour, texture, shape, position, etc.) with semantic concepts; capture the meaning and understanding of a domain.

This topic contributes:

• An unsupervised learning method for discoveringcross-domain correlation knowledge in which anovel cross-domain method is possible to discover the correlationknowledge between different domains via unsupervised learning on unlabeled data from multiple feature spaces,and then apply it to perform supervised classificationtasks.

• A novel two-level image representation model in which Unlike the basic bag-of-visual-words (BoW) model, the higher level visual components (words and phrases) to incorporate the spatial and semantic information into image representation model (*i.e.*,bag-of-visualphrases (BoP)). By combining visual words with phrases, the distinguishing power of image features is enhanced.

• Two different strategies (Enlarging Strategy and Enriching Strategy) to utilize the correlationknowledge to enrich the feature space for classification. Bytransferring such knowledge, both the strategies can handlethe situation when one information source is missing, especiallythe most common situation that the textual descriptionsof a small portion of web images are missing.

By effectively transferring the cross-domain correlation knowledge to new learning tasks, this method cannot only be applied in some specific domains (*e.g.*,, disaster emergency management), but also be used in the general domain (*e.g.*,social-media images organization, etc.).

II. RELATED WORK

The Fayyad et al.[3] KDP(Knowledge Discovery Process) model consists of nine steps, which are given bellow:

1. Developing and understanding the application domain. It includes learning the relevant prior knowledge and discovers the knowledge.
2. Creating a target data set. The data miner selects a subset of variables (attributes) anddata points (examples) that will be used to perform discovery tasks. This usually includesquerying the existing data to select the desired subset.
3. This step consists of Data cleaning and preprocessing which includes removing outliers, deals with noiseand missing terms and values in the data, and maintaing account for time sequence information and knownchanges.
4. It includes the Data reduction and projection in which finding useful attributes by applyingdimension reduction and different transformation methods.
5. Choosing the data mining task where data miner matches the goals defined in step 1 by applyingregression, clustering, classification, etc.
6. Choosing the data mining algorithm in this the data miner selects methods to search for patterns inthe data and decides which models and parameters of the methods used may be appropriate.
7. Data mining step generates patterns in a particular representational form, such asregression models classificationrules, decision trees, trends, etc.
8. Interpreting mined patterns includes the process where the analyst performs visualization of the extracted patternsand models, and visualization of the data based on the extracted models.
9. Consolidating discovered knowledge consists of incorporating the discoveredknowledge into the performance system, documenting and reporting it to the interestedparties. It also includes checking and resolving potential conflicts with previouslybelieved knowledge.

Pravin M. Kamde, Dr. Siddu. P. Algur[4] introduced different models for Multimedia classification and clustering based on the supervised and unsupervised learning.

Classification models

Machine learning (ML) and meaningful information extraction can only be realized. Decision trees can be translated into a set of rules by creating a separate rule for each path from the root to a leaf in the tree. The rules can also be directly induced from training data using a variety of rule-based algorithms. Artificial Neural Networks (ANNs) are another method of inductive feasible learning, based on computational models of biological neurons and networks. The Support VectorMachines (SVMs) [5] is the newest technique that considers the notion of a "margin".

Clustering Models

In unsupervised classification, the issue is to group a given collection of unlabeled multimedia files into meaningful clusters according to the multimedia content without a priori knowledge.Clustering algorithms can be divided into different methods such as partitioning methods, hierarchical methods, density-based, grid based, and model-based methods. A survey of clustering techniques can be found in[4].Density-based clustering algorithms try to find clusters based on density of data points in a region. The main idea of density-based clustering is that, for each instance of a cluster, the close neighbourhood of a given radius has to contain at least a minimum number of instances.Grid-based clustering algorithms first quantize the clustering space into a finite number of cells (hyper-rectangles) and then perform the required operations on the quantized space.

Association Rules

Rule based classification is based on the occurrences similarity of pattern or entity in the specific domain data set. Support, Confidence and Interest are three mainmeasures of the association.The support factor indicates the relative occurrences of both X and Y within the overall data set of transactions. It is the ratio of the number of instances that are satisfying both X and Y over the total number of instances. The confidence factor is defined as it is the probability of Y given X and itis the ratio of the number of instances that are satisfying both X and Y over the number of instances satisfying X. The support factor simply indicates the frequencies of the occurring patterns in the rule, and the confidence factor describes the strength of implication of the rule. Ameasure of human interest in the rule is called interest factor. The little research has been conducted on mining multimedia data [4][6] with different types of associations: association between image content and non image content features. Association mining in multimedia data can be transformed into problems of association mining in traditional transactional databases. That's why mining the frequently occurring patterns among different images becomes mining the frequent patterns in a set of transactions.

Hsin-Chang Yang, Chung-Hong Lee [8] proposed the self organizing map that is usful for image retrieval. Traditional content-based image retrieval (CBIR)[11]systems most of the times fails to meet a user's need due to the 'semantic gap' between the extracted features of the systems and the user's query. It is a difficult task to extract semantics of image. Most existing techniques apply some predefined semantic categories and assign the images to appropriate categories through some learning processes.These techniques always need human intervention and rely on content-based features. A novel approach to bridge the semantic gap which is the major deficiency of CBIR systems [11].By applying a text mining process, which adopts the self-organizing map (SOM) learning algorithm as a kernel process on the environmental texts of an image to extract the semantic information from this image.Some implicit semantic information of the images can be discovered after the text mining process. A semantic relevance measure has been given to achieve the semantic-based image retrieval task.

Yimin Wu and AidongZhang [6] introduced an adaptive classification method for multimedia retrieval using relevance feedback.Relevance feedback method caneffectively improve the performance of CBIR [11].A relevance feedback approach can be able to efficiently capture the user's query concept from a very limited number of training samples. To address this issue, a novel adaptive classification method using random forests, which is a machine learning algorithm with proven good performance on many traditional classification problems with random forests, this feedback method reduces the relevance feedback to a two-class classification problem and classifies database objects as relevant or irrelevant. Fromthe relevant object set, it returns the top k nearest neighbors of the query to the user.

$$P(1|\vec{o}) = \frac{\sum_{j=1}^{J} h_j(\vec{o})}{J}$$

where P(1|o) is the number of classifiers that classify o as relevant over the total number of classifiers. The larger the P(1|o), the more confident it is to output o as relevant. So,this method returns some classified-irrelevant objects with the largest P(1|o) values, in case less than k objects were classified as relevant.

Features:

- This system will able to address the multimodal distribution of relevant points, because it trains a nonparametric and nonlinear classifier, i.e., random forests, for relevance feedback.
- It does not overfit training data because it uses an ensemble of tree classifiers to classify multimedia objects.

Christian Ries[7] proposed a classification framework for digital images which is capable of identifying images which belong to a particular class. Simply to design filters which find images in a given database which feature content (e.g brand logos).This framework learnsdifferent class models in an unsupervised classification manner. The user is only required to provide images which contain some common object or concept as positive training examples without further annotation or knowledge. Then it finds common properties of the positive training images based on color and visual words. It consists of two main stages: A color-based pre-filter (or region of interest detector) and a classifier trained on histograms of visual words ("bags-of-words").As the learning process of the color model is to be unsupervised, different problems has been considered like identification of colors of the object without manual annotation and dealing with color deviations due to different lighting conditions. It is not so easy to classify images or localize objects based on color models.A classifier trained on histograms of visual words ("bags-of-words") this stage uses bag-of-words models to classify images. By compute spatial histograms of visual words for positive and negative training images and then train a binary classifier using these histograms. To find positive images among large scale databases for a very low false positive rate the classifier AdaBoost is used.There exist different local feature descriptors which can be used for the bag-of-words model. Also, the clustering process which yields visual vocabulary and the AdaBoost classifier depend on many parameters.

Dong ping Tian[13] has been studied different Image Feature Representation techniques. It includes how to partition an image and how to organize the image features with challenging questions. In general, there are mainly three methods to transform an image into a set of regions: regular grid approach, unsupervised image segmentation and interest point detectors. Images can be segmented by a regular grid or by the JSEG , and to detect the salient regions detected by the Difference of Gaussian (DoG) detector.Bag of visual words representation has been widely used in image annotation and retrieval. This visual-word image representation is analogous to the bag-of-words representation of text documents in terms of form and semantics. The procedure of generating bag-of-visual-words can be as follows: First, region features are extracted by partitioning an image into blocks or segmenting an image into regions. Second, clustering and discretizing these features into visual word that represents a specific local pattern shared by the patches in that cluster. Third, mapping the patches to visual words and then we can represent each image as a bag-of-visual-words.Compared to previous work, J. Yang, Y.Jiang, [15] have thoroughly studied the bag-of-visual-words from the choice of dimension, selection, and weighting of visual words in this representation. For more detailed information, please refer to the corresponding literature. Figure 1 shows the basic procedure of generating visual-word image representation based on vector-quantized region features.

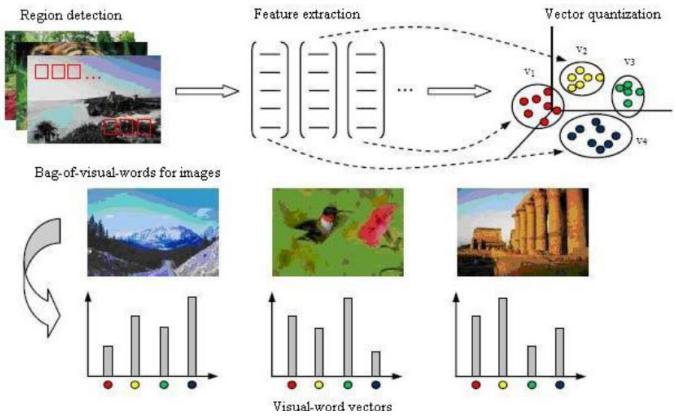

Figure1:Procedure of generating visual-word image representation based on vector-quantized region features

III. SOURCES OF DATA

Now a days there are n number of data sources are available for the classification purpose. The different data sources are social media, news articles, review sites, blogs, datasets, etc.

1. Social Media

Social media become a huge platform to share the images among the people. It is a large network where at a time millions of people can share their photos and views about the particular photo. There are different type of social media sites are available like www.facebook.com, www.tweeter.com, www.hi5.com, www.linkedin.com etc. which contains images with their descriptions.

Figure2: Facebook image with its textual description

2. News Articles

The websites like www.abpmajha.com, www.aajtak.com and www.lokmat.com, www.bhaskar.com has a news article that contains images of different news with its detail information.

The Reserve Bank of India didn't give a reason for changing the gold-import rule.

Figure3: Times of India news

3. Review Sites

Before purchasing any product it is very important to get details of that product. There are various ecommerce sites like

www.flipkart.com, www.cnet.com, www.snapdeal.com where customers want to know about the product and it's details while purchasing the product.

Figure4: Example of the product view and it's details

4. Blogs

A web log is called as blog it is a personal webpage on which particulars can write their likes, dislikes, opinions, hyperlinks to various sites etc. daily. Tweeter is one of the popular micro blogging service in which user creates status messages in a limited word count

which called as tweets. The tweeter will get flooded while the elections were going on. Tweets can also use as data source for multimedia object classification.

CONCLUSION

It explored related research efforts that generally focused on information retrieval tasks. Our intention is to recognize the trends in the surveyed area and categorize them in a new way that would integrate and add understanding to the work in the field with respect to the Flickr social media network. Spatio-Semantic Feature Integration can be used for multimedia object classification

IV. ACKNOWLEDGMENT

I express great many thanks to Prof. S. A. Shinde for his great effort of supervising and leading me, to accomplish this fine work. Also to college and department staff, they were a great source of support and encouragement. To my friends and family, for their warm, kind encourages and loves. To every person gave us something too light my pathway, I thanks for believing in me.

REFERENCES:

[1] Wenting Lu, Jingxuan Li, Tao Li, WeidongGuo, Honggang Zhang, and Jun Guo "Web Multimedia Object Classifcation Using Cross-Domain Correlation Knowledge," *IEEE transactions on multimedia*, vol.15, no. 8,December 2013.

[2] AncaLoredana Ion,"Methods for Knowledge Discovery in Images," *ISSN 1392-124X Information Technology and Control*, Vol.38, No.1, 2009.

[3]Fayyad et al., "Knowledge discovery and data mining: towards a unifying framework," *Proceedings of the 2nd International Conference on Knowledge Discovery andData Mining*, pp. 82–88, 2000.

[4]Pravin M. Kamde, Dr. Siddu. P. Algur, "A survey on Web Multimedia Mining," *The International Journal of Multimedia & Its Applications (IJMA)* ,Vol.3, No.3, August 2011

[5] Cristianini, N. &Shawe-Taylor, J., "An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods," *Cambridge University Press, Cambridge*, 2000.

[6]Yimin Wu and Aidong Zhang,"An Adaptive Classification method for Multimedia Retrieval," *The Adaptive Web*,pp. 291-324, Springer, 2007.

[7]Christian Ries, "Unsupervised One-class Image Classification,"*ACM International Conference on Multimedia Retrieval* (ICMR 2012), 2012.

[8]Hsin-Chang Yang,Chung-Hong Lee, "Image semantics discovery from web pages for semantic-based image retrieval using self-organizing maps," *Expert Systems with Applications*, vol 22, pp.266–279, 2008.

[9]Khatib, W., Day, F., Ghafoor, A., & Berra, P. B., "Semantic modeling and knowledge representation in multimedia databases," *IEEE Transactions on Knowledge and Data Engineering*, vol no.11,pp64–80. 2000.

[10] Chang, S. F., Chen, W., &Sundaram, H., "Semantic visual templates: linking visual features to semantic," In Proc. *IEEEinternational conference on image processing (ICIP'98)*, Chicago, IL(pp. 531–535),2011.

[11] T. K. Shih, J. Y. Huang, C. S. Wang, *et al.*, "An intelligent content-based image retrieval system based on color, shape and spatial relations," In Proc. *National Science Council, R. O.C., Part A: Physical Science and Engineering*, vol. 25, no. 4, pp. 232-243,2001.

[12] T. W. S. Chow and M. K. M. Rahman, "A new image classification technique using tree-structured regional features,"*Neurocomputing*, vol. 70, no. 4-6, pp. 1040-1050,2007.
[13]Dong ping Tian, " A Review on Image Feature Extraction and Representation Techniques*," International Journal of Multimedia and Ubiquitous Engineering*, Vol. 8, No. 4, July, 2013.
[14]Y. N. Deng and B. S. Manjunath, "Unsupervised segmentation of color-texture regions in images and video," *IEEE PAMI*, vol. 23, no. 8, pp. 800-810,2008.
[15] J. Yang, Y.Jiang, A. G. Hauptmann and C. Ngo, "Evaluating bag-of-visual-words representations in scene classification,", In Proc. *Workshop on MIR*, pp. 197-206,2007