Annotating Images from Large Scale Web Community

C. Ranjeeth Kumar, Prisgal Saha G, Sameema N, Shruthy Raj

Abstract- Automatic image annotation is a process by which metadata is assigned in form of captioning or keywords to a digital image. Large annotation databases are difficult to build because some of the images have partial annotations and noise tags problem. In order to solve the problems with the annotation of large databases, in our approach we remove noise and invalid images from the dataset and extract the visual features from the images. Annotation results are improved using WordNet based annotation refinement method.

Index Terms- Automatic Annotation, Classification, Feature Extraction, Tagging

I. INTRODUCTION

As the Internet technologies are growing. Images and videos are booming on the Internet. Organizing and managing the huge information on internet have become an issue. Through multimedia information retrieval we can solve this problem to a certain degree and automatic image annotation task is the key step of multimedia information retrieval. Automatic image annotation refers to automatically generate captions and tags according to images' visual contents. The traditiona l methods of
image retrieval such those used by libraries have relied on manually annotated images. The application of automatic image annotation in image processing is used in image retrieval systems to organize and locate images of interest from database. There are two major approaches of automatic image annotation: one approach uses the commercial image search engines (such as Google), which mainly uses the context information on the Internet to annotate images such as image filenames, URL, ALT tags as well as the surrounding text around images etc. However, the annotation results of this approach could not be satisfied because it does not consider the internal visual feature of the images. Another approach assigns tags to images by considering images' visual contents, which can be called as content-based automatic image annotation. This kind of methods can build good links between visual contents and annotation.

II. RELATED WORK

Automatic image annotation is similar to object recognition task, but there still have some differences between these two tasks.

Manuscript Received on March 2015.

Prof. Mr. C. Ranjeeth Kumar, Department of Information Technology, Sri Ramakrishna Engineering College, Coimbatore, India.
Ms. G. Prisgal Saha, Department of Information Technology, Sri Ramakrishna Engineering College, Coimbatore, India.
Ms. N. Sameema, Department of Information Technology, Sri Ramakrishna Engineering College, Coimbatore, India.
Ms. Shruthy Raj, Department of Information Technology, Sri Ramakrishna Engineering College, Coimbatore, India.

Automatic image annotation does not particularly care about the exact position of each object in the images. Object recognition systems usually find the specific foreground object and then construct different classifiers for different object. Image annotation and object recognition are both meaningful and challenging tasks at present. Xiao Ke,Shaozi Lic,and Guolong Chen proposed a web community based image annotation model to obtain good performance on large-scale web community images[1]. To solve the problems with the annotation of large databases ,Yong W, Tao M, Shaogang G, and Xian-Sheng H combining the global, regional, and contextual features to improve annotation performance by an extended Cross Media Relevance Model(CMRM)[3]. Two approaches for browsing and retrieving images are:
1. Content based image retrieval.
2. Retrieving images by their associated textual keywords. Ameesh M, Vladimir P,and Sanjiv K proposed a technique that utilizes global low-level image features and a simple combination of basic distance measures to find nearest neighbors of a given image[2]. The keywords are then assigned using a greedy label transfer mechanism. Jeon J, Lavrenko V, and Manmatha R proposed Cross Media Relevance Model (CMRM) which used joint probabilities of semantic labels and visual words to annotate images[6]. Automatic approaches are proposed to retrieve images based on a training set of images. The training set contains regions in a image described using a small vocabulary of blobs. Annotation words are predefined in the dataset, which makes it difficult to extend keywords to a large scale semantic set. Therefore, it is necessary to build and annotate the large scale web image database. Wang XJ, Zhang L, Jing F and Ma WY proposed an Internet image annotation system AnnoSearch [5]. AnnoSearch need an initial keyword provided by the user, and then find unlabeled image's most similar images. Similar images and their annotations are used to annotate this unlabeled image. The drawback of AnnoSearch is that it need assign at least one initial keyword as seed, which is quite inconvenient. Our model does not need any initial keywords, which is suitable for use in a large web image database. Li XR, Chen L, Zhang L, Lin FZ, Ma WY proposed a Web image annotation method based on searching and mining technologies[4]. In this method, content-based image retrieval (CBIR) is used to obtain similar collections of each unlabeled image, and then find several keywords with the biggest annotation probabilities from relevant texts in similar collections by using clustering method. The disadvantage of this model is that it is quite difficult to determine the clustering number in a large web image database. We do not use any clustering

algorithms, so our model does not have the above problems.

III. ELIMINATION OF NOISY IMAGES

Initially a WANG dataset which consist of both noise and noiseless images are taken. Elimination of noisy images are done with the help of Peak Signal To Noise Ratio (PSNR) value. PSNR value is found via Mean square error (MSE) value. Both PSNR and MSE are used to compare the squared error between the original and the reconstructed image to find the mean square error ratio. PSNR value can only be calculated via MSE,

MSE = $(1/(r*c*d))*SUM\ ((Y1-Y2)^2)$

Where *r* is the row, *c* is the column, *d* is the dimension, *Y1* is the original image and *Y2* is the filter applied image.

PSNR = $10*log10\ (MAX^2/MSE)$

Where *MAX* is the maximum pixel value possible for an image (always 255 for 8 bit images).

Typical values for the PSNR is between 30dB and 50dB in images, provided the bit depth is 8 bit where higher is better.

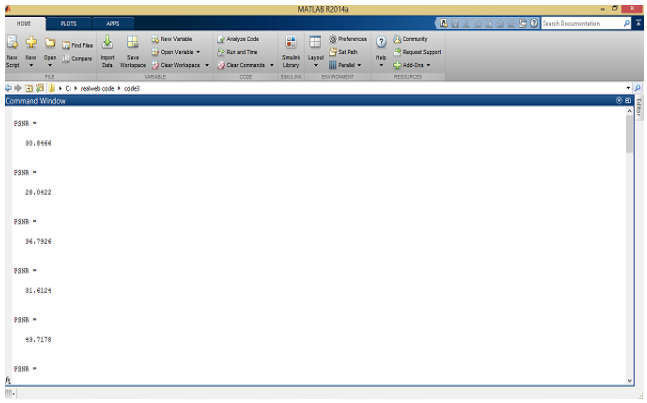


Fig. 1 Reading PSNR value

PSNR values for each image is read as in Fig.1.The images which do not have PSNR value between 30 and 50 eliminated from the dataset and the remaining images are taken as the training dataset as shown in the below fig.2.

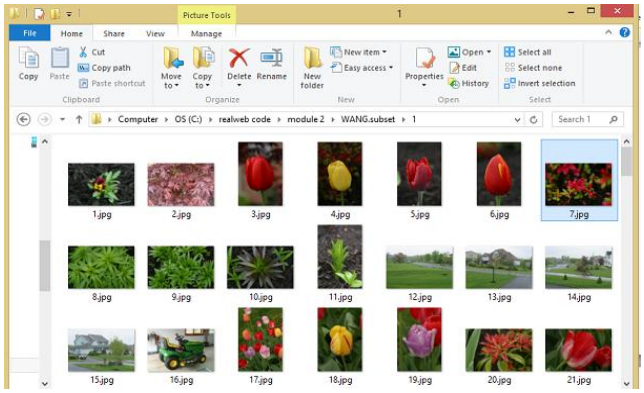


Fig. 2.Noise eliminated dataset

IV. FEATURE EXTRACTION

Feature extraction a type of dimensionality reduction that efficiently represents interesting parts of an image as a bounded feature vector. It is useful when image sizes are large and a reduced feature representation is required to quickly complete tasks such as image matching and retrieval. There are two types of feature extraction namely:

1. Gradient feature extraction.
2. Texture feature extraction.

A. Gradient Feature Extraction

Gradient feature represents the rate of change in color or intensity of an image. In our approach, Multi Coordinate Histogram of Oriented Gradients (MCHOG) is used to extract the gradient feature of an image. Gradient represents the rate of change of x axis with respect to y axis. Image histogram shows the frequency of pixel's intensity values. Multi Coordinate Histogram of Oriented Gradients requires filtering the color or intensity of the image by using the function

hx=[-1 0 1]
hy=-hx'

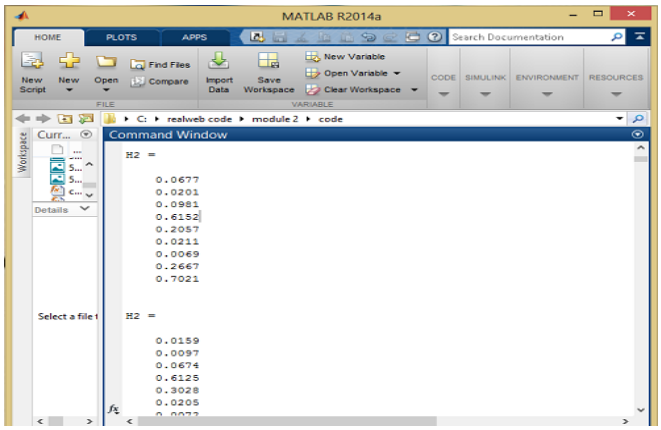


Fig. 3.Histogram values generation

After that cell histogram is created where each pixel has weight in histogram based on the values found in the gradient computation as shown in fig.3. Each pixel's value contributes to the magnitude computation.

B. Texture Feature Extraction

Gabor filters with different frequencies and orientations are used to extract the features(spatial arrangements between the color and the intensities) from an image. Initialize the Gabor filter sizes to zero and create an array as long as the image. Read an image and apply the Gabor filters to that image to extract the features. Convolute the extracted features and the images in array and store in the database as shown in fig.4.

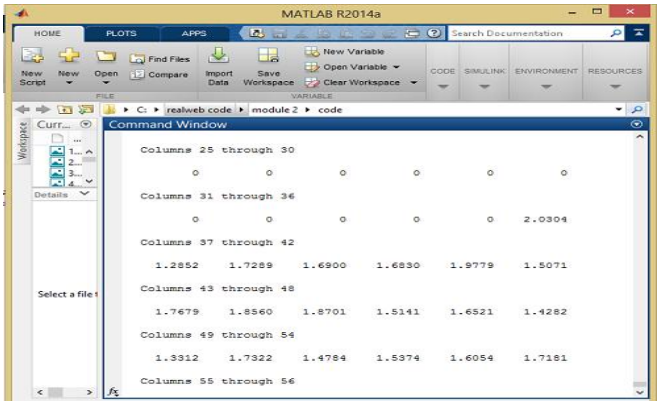
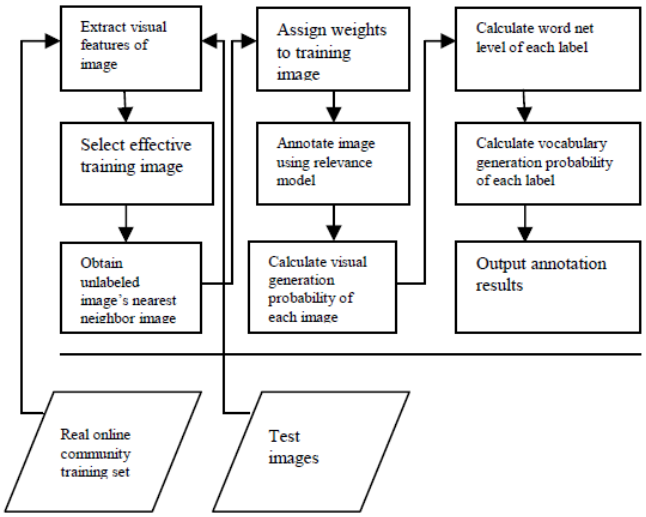


Fig. 4.Texture feature values

V. CLASSIFICATION AND ANNOTATION



In this subsection, Real Community Annotation Model(RCAM) is used to automatically annotate images in a database, the architecture of Real Community Annotation Model(RCAM) model is shown in Fig.5.In the training set and an unlabeled image from a web community, visual features of each image should be first extracted. Secondly, effective training images should be selected and then nearest neighbor image collection of unlabeled image could be obtained. Thirdly, different weights are assigned to each image in the nearest neighbor image collection by using weighted KNN method. Fourthly, Real Community Annotation Model(RCAM) is used to annotate the unlabeled image. In Real Community Annotation Model (RCAM), similarities between each unlabeled image and selected effective training images are needed to be calculated, so visual generation probability of the unlabeled image is calculated. Each label's WordNet level should be calculated, and then each label would be given different weights according to their WordNet levels by using image annotation refinement based on word net level. After that, relations of different labels in each unlabeled image and selected effective training images are considered, so vocabulary generation probability of each label is calculated so that annotations on large scale web community is done**.**

IV. CONCLUSION

We have studied the problem of annotating images from large scale web community. We observed that Automatic image annotation is a challenging problem in pattern recognition and also improper annotation results in large database. The proposed methodology consists noise elimination from dataset for more efficiency. Our proposed work also uses feature extraction using MCHOG and Gabor texture method to extract important features from images for efficient annotation. Experimental result shows that annotating image is achieved better in our system.

REFERENCES

[1] Xiao Ke,Shaozi Lic,and Guolong Chen,"Real web community based automatic Image annotation,"vol.39,no.6, October 2012,pp.945-956,.
[2] M. Ameesh , P. Vladimir ,and K.. Sanjiv,"Baselines for image annotation,"Int J Comput Vision., vol.90,no.1,May 2010, pp.88–105.
[3] W. Yong , M. Tao, G. Shaogang, and H. Xian-Sheng ,"Combining global, regional and contextual features for automatic image annotation,"Pattern Recogn.,vol-42,no.2,May2008,pp.259–66.
[4] Li XR, Chen L, Zhang L, Lin FZ, Ma WY,"Image annotation by large-scale content-based image retrieva," In: Proceedings of the l4th ACM international conference on multimedia, Santa Barbara; 2006, p607–10.
[5] XJ. Wang,L. Zhang,F. Jing,and WY. Ma,"AnnoSearch:image auto-annotation by search," Proceedings of the IEEE computer society conference on computer vision and pattern recognition ,2006, pp.1483-1490.
[6] J. Jeon, V. Lavrenko, and R. Manmatha," Automatic image annotation and retrieval using cross-media relevance models,"Proceedings of the 26th annual international ACM SIGIR,August 2003, pp.119–26.

AUTHOR PROFILE

Mr. C. Ranjeeth Kumar, received his Bachelor degree in Information Technology from Anna University, Chennai and his Master degree in Computer Science and Engineering from Anna University, Coimbatore. He has published many papers in journals and conferences. He is currently working as Assistant Professor, Department of Information Technology, Sri Ramakrishna Engineering College, Coimbatore. His research areas are image processing.

Ms. G. Prisgal Saha, is currently doing her under graduation in information technology from Anna University Chennai. Her research area is image processing**.**

Ms. N Sameema, is currently doing her under graduation in information technology from Anna University Chennai. Her research area is image processing.

Ms. Shruthy Raj, is currently doing her under graduation in information technology from Anna University Chennai. Her research area is image processing.