

Automatic Image Annotation Based on Dominant Color and GLCM using Fuzzy c means Clustering

Abstract— With the detonative growth of the digital technologies in the web large amount of visual data are created and stored. The majority of image available on the web have little or no metadata associated with it describing the semantic concept associated with the images. There is a need of efficient and effective technique to find visual information on demand. One of the promising approach to enhance the image retrieval is automatic image annotation which refers to process of assigning relevant keywords to the image to bridge the semantic gap between low level content features of image such as color, texture and shape and semantic concepts understood by the humans such as keywords, description or image classification. The paper discusses implementation of the automatic image annotation using fuzzy c means clustering to annotate the image based on Dominant color and Gray level cooccurrence matrix texture feature. The experiments are conducted on 50 beach images and 50 images of the corel dataset to identify the best cluster number and similarity measure such as Cityblock, Sqeuclidean, Canberra used in the fuzzy c means clustering for annotating the image. The experiments results and comparison with other clustering method and feature descriptor showed that the proposed automatic image annotation scheme can effectively improve the annotation performance. The analysis of annotation results using various similarity measure and feature descriptor are also presented in the paper.

Keywords—image retrieval; annotation; dominant color; fuzzy c means clustering; similarity measure

Dhatri Pandya

Computer Engineering Department
Sarvajnik College of Engg and Technology
Surat, India

Prof. Bhumika Shah

Computer Engineering Department
Sarvajnik College of Engg and Technology
Surat, India

I. INTRODUCTION

Handling large volumes of digital information become vital as online resources and their usage continuously grows at high speed online image sharing applications are getting extremely popular. This shows the potential of these online image collections and the need to search them on the basis of words. So to enhance the image retrieval results annotation is required as label associated with image represents the

semantic information associated with it. In general, research efforts on image retrieval can be divided into three types of approaches [1]. (1) Text based annotation (2) Content Based Image Retrieval (3) Automatic Image Annotation. The first approach is the traditional text based annotation. In this approach, images are annotated manually by humans and images are then retrieved in the same way as text documents. It is time consuming and expensive to manually annotate the large amounts of images. Also human annotations are too subjective and vague. The second type of approach centers on content based image retrieval (CBIR), which is a task of searching images from a database and retrieval of an image, which are seemed to be visually similar to a given example or query image. It is impractical for general users to use a CBIR system because it furnishes users to provide query images [2]. The “semantic gap” is a well-known problem in multimedia. The challenge is to accurately classify and effectively search multimedia content from automatically extracted low-level visual features. Semantic gap is defined as “the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation”. To overcome the semantic gap that exist between low level features present in the image and semantic concepts understood by humans and to enhance the image retrieval various annotation techniques have been defined according to the literature [1]. The main idea of Automatic Image Annotation techniques is to automatically learn semantic concept models from large number of image samples, and use the concept models to label new images. Annotation based image retrieval is perceived as better as users are allowed to query the image in natural language. The key aspect of automatic image annotation is it offers the advantages of both text based annotation and content based image retrieval. The paper is organized as follows. Section II discusses the clustering approach used for annotation. Section III covers the proposed scheme and discusses the feature representation, clustering and similarity measure used in proposed scheme.

A. RELATED WORK

Several algorithms have been proposed for automatic image annotation according to the literature [1]. Probabilistic model based algorithm has been proposed by [3][4][5] which annotates the image by estimating the joint probability of words to images regions. Classification based approach is based on the number of training samples fed to the classifier using artificial neural network, support vector machine and

based on the positive and negative instances of trained data[11][12]. Parametric approach is based on multivariate Gaussian distribution but the estimation of the parameters is complex[7]. Graph based approach considers the relation between visual features of image but does not consider the relation between words assigned to the image as proposed in the literature[6].

II. CLUSTERING APPROACH

Clustering deals with finding a structure in a collection of unlabeled data and can be considered as the most significant unsupervised learning problem. In hard or crisp clustering, data is divided into unique clusters, where each data element belongs to exactly one cluster. Clustering algorithms are usually used in classifying low-level features of the images prior to annotation. Clustering is the process of organizing objects into groups by maximizing the intraclass similarity and minimizing the interclass similarity. A cluster can be defined as a collection of objects that are similar between them and are dissimilar to objects belonging to other clusters. The three main tasks carried out in the annotation process using clustering approach are as follows (a) Feature representation which extracts the low level features of image such as color and texture features which are the more expressive visual features of the images (b) Clustering partitions the feature space into number of distinct clusters based on similarity between feature vector of image. (c) Annotation is carried out by matching feature vector from number of training samples.

The distance measure measures the discrepancy between two objects. Various similarity measure is being used in the clustering of feature vector according to the literature defined in [13]. The similarity between two images can be computed by measuring the distance between feature vector of the images. The distance measure widely used in the clustering process for annotation are city block, euclidean, Canberra using color k means clustering proposed by the Jamil et al[8]. It annotates the image based on RGB feature vector extracted from the image and applying k means clustering. In general the annotation process using clustering approach consists of training the images using label me annotation tool [15] for generating ground truth labels for the training image. The testing phase consist of feature extraction using RGB vector and generating clusters based on RGB (Red Green Blue) feature vector and transferring the labels from training phase to testing phase based on the closest matching class.

The annotation results of the image are dependent on the feature representation and clustering algorithm used. The annotation algorithm based on color k means clustering extracts only RGB feature vector which may generate the ambiguity in the annotation results for the object such as sky, sea, cloud as only few color are extracted. As sky and sea have similar color but their texture are different so annotation results may be different for each dataset and also dependent on the predefined class. This demands for the proposing automatic image annotation scheme based on the fuzzy c means clustering and using dominant color and GLCM texture feature which leads to better feature representation of image and soft clustering that will improve the annotation process.

Due to the limitations of k means clustering of assigning feature vector to only one unique class the annotation results may be affected giving ambiguous results as for some categories of image two different objects are similar in color and texture so annotation results are affected. So by utilizing better feature representation and clustering algorithm the annotation results will be improved.

III PROPOSED SCHEME

The proposed approach discusses the implementation of the automatic image annotation based on dominant color and GLCM (Gray Level Cooccurrence Matrix) texture feature using fuzzy c means clustering. As only RGB color feature which will not lead to better feature representation of image as it is possible that image may contain multiple objects with various colors such as yellow, green, blue, magenta those features are not extracted by RGB feature representation. Due to the heterogeneity of data and loss of sensitivity for large number of data the distance measure such as city block, euclidean, Canberra sometimes gives unpredictable results. So a new similarity measure based on HSV color wheel is used in the proposed scheme for the annotation of the image.

A. FEATURE REPRESENTATION

The color and texture feature of image are most expressive visual features of image. Various color descriptors has been defined by Manjunath et al [9]. The textual feature representation of image carries useful information for discrimination purposes. It is essential to develop significant features for texture such as surface depth and surface orientation, contrast and homogeneity feature that describes statistics about the texture of the image. Textures in images quantify: Grey level differences (contrast), size of area where change occurs, (window) Directionality or lack of it [14].

The dominant color descriptor gives the distribution of the salient colors in the image. The purpose of dominant color is to provide an effective, compact, and perceptive representation of colors present in a region of interest. The dominant color descriptor allows color specification in any of the color spaces supported by MPEG-7[9]. A set of 12 dominant colors in a region of interest or in an image provide a compact description that is easy to index. Colors in a given region are clustered into a small number of representative colors. The feature descriptor based on dominant color used in the proposed scheme is obtained by extracting RGB feature representation from the image and a histogram is plotted for each of the planes. Histogram is transformed into a probability mass function (pmf) for each of the planes. After obtaining the pmf of each plane, probability of obtaining a colored pixel is obtained by multiplying the probability of obtaining the value in red plane by probability of obtaining the value in green plane by probability of obtaining the value in blue plane. A weighted average of all the features/colors present is then taken using probabilities obtained as weights to obtain the dominant color. The difference between the dominant color

descriptor and the color histogram descriptor is that the representative colors are computed from each image instead of being fixed in the color space, thus allowing the feature representation to be accurate as well as compact[9].

The average color for each class is obtained by comparing each pixel in multiple images present for a given class and checks which dominant color the pixels belong too then the probability of each dominant color is calculated and a weighted average of all the pixel values

In the proposed scheme the texture analysis based on signal processing method and statistical method is used. Texture feature extraction based on signal processing method is done using 1D DFT .In statistical texture analysis, texture features are computed from the statistical distribution of observed combinations of intensities at specified positions relative to each other in the image. The GLCM tabulates how frequently different combinations of pixel (grey levels) occur in an image.

Gray Level Co-Occurrence Matrix (GLCM) has proved to be a popular statistical method of extracting textural feature from images. A number of texture features may be extracted from the GLCM.Two texture features are computed from GLCM contrast and homogeneity feature as defined in the literature [14].

B. FUZZY C MEANS CLUSTERING

The goal of the fuzzy c-means algorithm is the assignment of data points into clusters with varying degrees of membership. This membership reflects the degree to which the point is more representative of one cluster than another. FCM clustering techniques are based on fuzzy behaviour and they provide a technique which is natural for producing a clustering where membership weights have a natural interpretation but not probabilistic at all. It is an approach of clustering which permits one piece of data to fit into two or more clustering based on degree of membership[10]. It partitions set of n objects $X=\{x_1,x_2,\dots,x_n\}$ in R^n dimensional space into c ($1 < c < n$) fuzzy clusters with $Z = \{ Z_1 , Z_2 , \dots , Z_c \}$ cluster centers. The fuzzy clustering of objects is described by a fuzzy matrix u with n rows and c columns in which n is the number of data objects and c is the number of clusters. U_{ij} , the elements in the i-th row and j-th column in u, indicates the degree of membership function of the i-th object with the j-th cluster[18].

$$u_{ij} \in [0,1] \forall i = 1,2, \dots, n; \forall j = 1,2, \dots, c$$

Fuzzy c-means algorithm is based on the minimization of the following object function:[10]

$$J_m = \sum_{j=1}^c \sum_{i=1}^n u_{ij}^m d_{ij}^2$$

Where $d_{ij} = \|x_i - z_j\|; m \in [1, \infty)$ is a parameter (weighting exponent) which determines the fuzziness of the resulting clusters and d_{ij} is the Euclidean distance from object x_i to the cluster center z_j . The z_j is obtained using the equation stated in the following[10]

$$z_j = \frac{\sum_{i=1}^n u_{ij}^m x_i}{\sum_{i=1}^n u_{ij}^m}; 1 \leq j \leq c$$

A new similarity measure based on HSV Color wheel is used in the proposed scheme for matching features of training to testing image. It first find the probability for a set of dominant colors and texture feature belongs to which predefined class if it does not belong to any class then the hsv color wheel is rotated in anticlockwise direction. Usually the Hsv color wheel is divided into 12 colors and h value ranges from 0 to 360 degrees. The probability array consists of getting probability of one of the 12 colors present in the image.The process is repeated for the specified number of clusters and distance is measured by computing the difference between color extracted from HSV color wheel and probability of the colors belonging to particular class.

C. ALGORITHM FOR THE PROPOSED SCHEME

The annotation process consists of two phases testing phase and training phase.The detailed algorithm of the proposed scheme is stated in the below section.

Training Phase

1. Generate the masks for each predefined object using label me annotation tool
2. Generate the ground truth labels for each mask generated for the training image
3. Extract the RGB color feature of each predefined class and texture feature using GLCM (or 1D DFT) for each mask from the training image.
4. Define a set of 15 colors (a set of 12 colors and 3 for saturation) and find out the number of pixels belonging to each color from the RGB color feature vector Store the color and texture features in feature vector file.
5. Find the probabilities of each color and store it.
6. Store the color and texture feature in feature vector file.

Testing Phase

Initialization

1. Define the classes
2. Specify distance measure and initialization mode, cluster number used for clustering.
3. Predefined the cluster color for each class.
4. Load the feature vector file for training images.

2. Feature testing and clustering

Feature testing

1. Specify the set of colors
2. Extract RGB feature vector and texture feature based on GLCM of test image
3. Apply fuzzy c means clustering on rgb feature vector of image and texture feature vector of testing image.

Return [C,FcPM,FcPM_hard]

4. Find out the no of pixels present in color range as present in the color map
5. Find the probability of each color present in the testing image from the set of 12 colors and based on the maximum probability find the dominant color.
6. Extract the texture feature based on GLCM.
7. Store the probability of color and texture features in matrix

Return [prob,prop,FCPM,FCPM_hard]

3.Annotation

1. Find distance of the centroids from the features, then find the closest class.
2. Find the similarity between training feature vector and feature vector of the testing images based on using HSV color wheel based similarity measure
3. Annotate the image based on the closed class using step2
4. Calculate the accuracy of the annotation results using precision and recall measure.





D. EXPERIMENTS

The experiments are conducted on 50 beach images and 50 images of the corel dataset using proposed scheme .consisting of 8 predefined classes for beach images and 15 predefined classes for the corel dataset using various distance measure and fuzzy c means clustering using Matlab 2012a.

To generate the ground truth labels for the training image label me annotation tool is used as defined in [15].The annotation results using various distance measure based on beach images and corel dataset is depicted in the table 1 using k means and proposed scheme

The results shown using k means approach are based on results obtained using distance measure such as city block that show the high precision rate for the annotation of the image and results obtained using proposed scheme are obtained using similarity measure based on hsv (hue,saturation,value) color wheel.

Table 1.Annotation results

| | | | |
|-----------------|--|---|---|
| |  |  |  |
| k means | Beach,grass ,tree,rock | Sky,beach ,grass,tree | Sea,cloud ,beach,hill ,tree |
| Proposed Scheme | Sky,sea cloud,beach ,rock | Sky,sea,clou d,beach,rock | Sky,sea, cloud,beach, tree |
| |  |  |  |
| k means | Wolf,snow ,road | Car,grass, tree,road, snow | Ship, cloud elephant, ,sky,snow |
| Proposed Scheme | Wolf,snow | Car,tree ,road | Ship,cloud, sky, |

The performance results of the automatic image annotation using proposed scheme based on HSV color wheel is stated in table 2

Table 2 Performance results using Proposed scheme

| Cluster number | BEACH | | COREL | |
|----------------|-----------|--------|-----------|--------|
| | Precision | Recall | Precision | Recall |
| 8 | 0.80 | 0.76 | 0.48 | 0.17 |
| 30 | 0.76 | 0.78 | 0.45 | 0.23 |
| 50 | 0.65 | 0.81 | 0.46 | 0.23 |
| 100 | 0.56 | 0.82 | 0.44 | 0.22 |

The performance of the annotation results are measured using precision and recall. The precision predicts the number of keywords correctly predicted while the recall measure predicts the number of keywords transferred to the image.

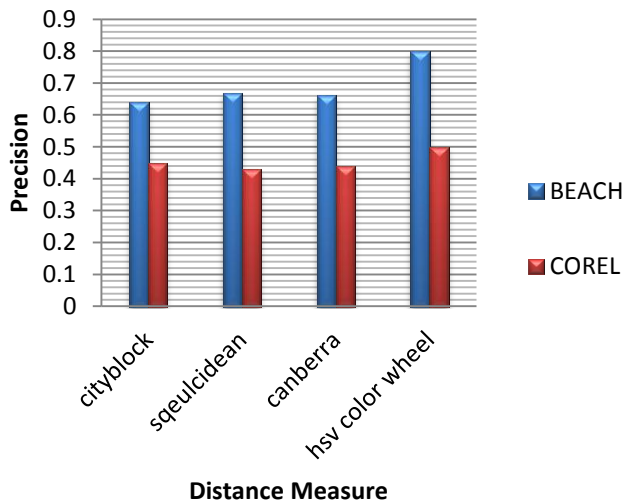


Fig 1 Performance of the Annotation Results

The comparison of the annotation results using k means ,dominant color+1D Discrete Fourier Transform using fuzzy c means clustering, dominant color and GLCM using hsv color wheel is shown in the figure 2.

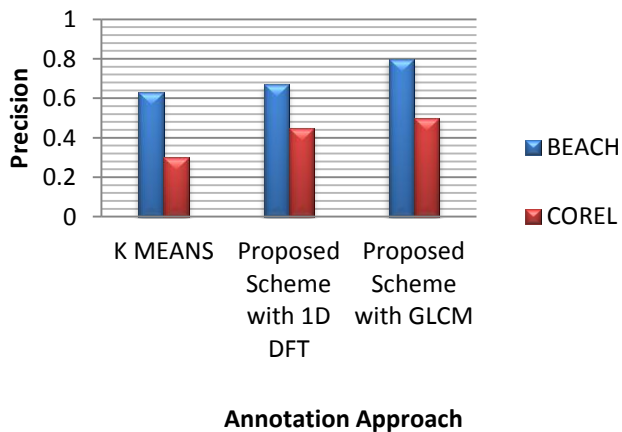


Fig 2 Performance of the Annotation Results

The annotation results for fuzzy c means clustering using dominant color and GLCM gives high precision on beach images and corel dataset and annotation results using hsv color wheel based similarity measure gives high precision as compared to other distance measure such as cityblock,sqeuclidean,Canberra.

CONCLUSION

Automatic Image Annotation is promising approach to enhance the Image Retrieval. In this paper we have proposed the annotation process based on dominant color and GLCM texture and 1D DFT using fuzzy c means clustering. By conducting experiments on various distance measure and using proposed scheme it can be concluded that annotation results for beach image and corel dataset gives promising results as compared to k means clustering. Future work will lead to applying other clustering algorithms and better feature representation method and developing annotation based image retrieval.

REFERENCES

- [1] D. Zhang, Md. M. Islam, G. Lu, "A review on automatic image annotation techniques", Pattern Recognition, vol. 45, no. 1, pp. 346–362, Jan. 2012.
- [2] Manaf, S.A, Nordin, Md Jan, "Review on statistical approaches for automatic image annotation," Electrical Engineering and Informatics, 2009. ICEEI '09. International Conference on , vol.01, no., pp.56,61, 5-7 Aug. 2009.
- [3] J. Jeon, V. Lavrenko, R. Manmatha, "Automatic Image Annotation and Retrieval using Cross-Media Relevance Models", 26th Annual Int. ACM SIGIR Conference, Toronto, Canada, 2003.
- [4] S.L. Feng, R. Manmatha, V. Lavrenko," Multiple Bernoulli relevance models for image and video annotation", in: Proceedings of the CVPR04, 2004, pp. 1002–1009.
- [5] P. Duygulu, K. Barnard, N. de Freitas, D. Forsyth, "Object recognition as machine translation: learning a lexicon for a fixed image vocabulary", In Seventh European Conference on Computer Vision (ECCV), Vol. 4, pp. 97-112, 2002.
- [6] J. Liu, M. Li, Q. Liu, H. Lu, S. Ma, "Image annotation via graph learning," Pattern Recogn., 42(2), 2009,pp. 218-228.
- [7] G. Carneiro, A.B. Chan, P.J. Moreno, N. Vasconcelos, "Supervised learning of semantic classes for image annotation and retrieval", IEEE PAMI 29 (3) (2007) 394–410.
- [8] Nurusuriati Jamil, Siti, Aisyah Sadam,"Automatic image annotation using color k means clustering.", SPRINGER, pp.645-652,2009
- [9] Manjunath, B.S.,Ohm, J.-R., Vasudevan, V.V.; Yamada, A., "Color and texture descriptors," Circuits and Systems for Video Technology, IEEE Transactions on , vol.11, no.6, pp.703,715, Jun 2001.
- [10] Jafar, O.A.M., Sivakumar, R., "A study on fuzzy and particle swarm optimization algorithms and their applications to clustering problems," Advanced Communication Control and Computing Technologies (ICACCCT), 2012 IEEE International Conference on , vol., no., pp.462,466, 23-25 Aug. 2012.
- [11] Yufeng Zhao,Yao Zhao, Zhenfeng Zhu; Jeng-Shyang Pan, "A Novel Image Annotation Scheme Based on Neural Network," Intelligent Systems Design and Applications, 2008. ISDA '08. Eighth International Conference on, vol.3, no., pp.644, 647, 26-28 Nov. 2008.
- [12] Shaohua Wan, "Image Annotation Using the Simple Decision Tree," Management of e-Commerce and e-Government (ICMeCG), 2011 Fifth International Conference on , vol., no., pp.141,146, 5-6 Nov. 2011.
- [13] Sung-HyukCha,"Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions, International Journal Of Mathematical Models And Methods In Applied Sciences, 2007,pp.300-307.
- [14] P.Mohanaiah, P.Sathyanarayana, L.Guru Kumar,"ImageTexture Feature Extraction using GLCM approach", International Journal of Scientific and Research Publications",2013,pp.1-5.
- [15] <http://people.csail.mit.edu/brussell/research/AIM-2005-025-new.pdf>
- [16] <http://lear.inrialpes.fr/data>
- [17] <http://people.revoledu.com/kardi/tutorial/Similarity/index.html>

[18] http://www.softcomputing.net/nabic09_submission_24.pdf