ISSN 2250-3153

Automatic Image Annotation by Classification Using Mpeg-7 Features

Manjary P.Gangan[*], Dr. R. Karthi [**]

[*] Department of Computer Science and Engineering, Vidya Academy of Science and Technology, Thrissur, 680501, India
[**] Department of Computer Science, Amrita Vishwa Vidyapeetham, Amrita University, Coimbatore, 641105, India

Abstract- Automatic image annotation (AIA) is a technique to provide semantic image retrieval. In AIA, the image contents are automatically labelled with a pre-defined set of keywords which are exploited to represent the image semantics. This paper proposes an Automatic image annotation method using MPEG-7 features. The feature vectors are provided for training the KNN-classifier. When a query image is provided by the user, the extracted features are provided to the trained KNN -classifier for annotation. The system also compares the results of different combinations of mpeg 7 descriptors.

The Automatic image annotation proposed in this paper is on fruit images. This system has several applications like automatic labeling and price computing of fruits and vegetables in a grocery store, morphological analysis of fruits, for scientific studies, enhanced learning for kids and Down syndrome patients using fruits pattern recognition.

Index Terms- Automatic Image Annotation, MPEG-7 features, KNN classifier

I. INTRODUCTION

The process of image retrieval has started from 1970's. Image retrieval methods can be broadly categorized into text based image retrieval and content based image retrieval. Text based methods are very simple and effective. Content Based Image Retrieval (CBIR) emerged in 1990's. Using these methods; images are retrieved by analyzing and comparing the low-level image features such as color, shape, texture etc. CBIR system searches an image database for images with similar content, as that of the given target image. These systems usually have a feature extraction phase and an image retrieval phase.

The feature extraction phase identifies the relevant regions in images and the features describing colors, textures and/or shape information of these regions or whole image. In image retrieval phase, images are retrieved by selecting properties such as colors, shapes and/or texture of image regions or a combination of these. One of the important problems in the CBIR systems is the 'semantic gap'. It refers to the gap between low-level content of information that can be extracted from images and the interpretation of higher level concept of images by humans. Automatic image annotation can bridge this semantic gap. In this method, users can specify their query concepts easily by using the relevant keywords. Image classification is one promising approach to enable automatic image annotation. The performance of image classifiers largely depends on image content representation, automatic feature extraction, an effective algorithm for image classifier training and feature subset selection. More the visual features used for training the classifier, more effectively and efficiently it characterizes different visual properties of images. This may further enhance the classifier's ability on recognizing different image concepts or object classes and result in higher classification accuracy.

The main idea of Automatic image annotation (AIA) techniques is to identify the concepts from large number of image samples, and use the concept models to label new images. Once images are annotated with semantic labels, they can be retrieved by keywords, which is similar to text document retrieval. The key characteristic of AIA is that it offers keyword searching based on image content and it employs the advantages of both the text based annotation and CBIR. AIA usually consist of two main steps i.e.; feature extraction and annotation. An image is an unstructured array of pixels represented using some features like color, texture, shape etc. For annotation, there is a need for appropriate feature selection from these pixels. Based on the different features selected and different algorithm techniques used, the performance of the semantic learning techniques vary.

II. OVERVIEW OF PROPOSED SYSTEM

In the proposed method of automatic image annotation, the fruit images are annotated using MPEG 7 features. The major steps in this method include preprocessing, feature extraction and algorithm for annotation. The MPEG 7 features are extracted from the images in the dataset. These feature vectors are then provided as a training set to the KNN-classifier. The same steps are performed for the query image provided by the user and the extracted features are provided to the trained KNN -classifier for classification and annotation. A comparison on different combinations of mpeg 7 descriptors for AIA is also performed.

This paper is organized as follows. Section 3 contains a review of existing methods, for the system as a whole or for a part of the system. Section 4 describes the working of proposed system. Section 5 contains the experimental results. It is followed by conclusion and references.

III. LITERATURE SURVEY

In all the related works, the main modules that have been identified are segmentation of image objects or regions, feature extraction from the images and the algorithms used for annotation. A Statistics based annotation called Continuous Relevance Model (CRM) was proposed by V. Lavrenko et al. in

[3]. In this model a joint probability is computed for image features and words, using a training set of annotated images. Another method for automatic image annotation was proposed by Wong et al. in [5], based on the image metadata information like aperture, exposure time, subject distance, focal length, etc. The metadata extraction is followed by automatic annotation using decision trees and rule induction. Antonio Torralba et al. proposed in [7], a web-based annotation tool that allows online users to label objects, for sharing and labeling of images for computer vision research purposes.

In [6] Ja-Hwung Su et al. proposed a model called Annotation by Image-to-Concept Distribution Model (AICDM) for image annotation which finds the associations between visual features and human concepts from image-to-concept distribution by integrating the methods of entropy, term frequency – inverse document frequency and association rules. A region-based approach using high-level semantics and decision tree learning was proposed for image retrieval in [4] by Ying Liu et al. In [8] Balasubramani et al. discusses in detail about the MPEG 7 features, CLD and EHD. In [9] Manjunath et al. presents an overview of MPEG-7 color and texture descriptors and effectiveness of the mpeg 7 descriptors in similarity retrieval, as well as extraction, storage, and representation complexities. Dong Kwon Park et al. shows in [10], how the edge histogram descriptor for MPEG-7 can be efficiently utilized for image matching.

IV. PROPOSED SYSTEM

The proposed system has the four modules namely, Preprocessing, Feature Extraction and Algorithm module.

Preprocessing

In the preprocessing stage, the input image was segmented and the region of interest (i.e.; the fruit region) is cropped out. Segmentation was done by Otsu's method.

Feature Extraction

The MPEG-7 descriptors are effective in similarity retrieval, storage, and representation complexities because they are semantic rich features. Moreover, these descriptions do not depend on the way the content is coded or stored. The MPEG-7 features used were the Color Layout Descriptor, Dominant Color Descriptor, Homogenous Texture Descriptor and Edge Orientation Histogram

Color Layout Descriptor (CLD)

It is a very compact descriptor and a resolution-invariant representation of color. It efficiently represents the spatial distribution of colors and has no dependency on image format. CLD uses a frequency domain feature which introduces perceptual sensitivity of human vision system for similarity calculation.

Dominant Color Descriptor (DCD)

DCD provides a compact description of the representative colors in an image or image region. It is accurate than the conventional histogram, sufficient to represent the color information of a region. The feature dimension is low and computation is inexpensive.

Homogenous Texture Descriptor (HTD)

HTD provides a precise quantitative description of a texture. It is composed of 62 fields, coming from the Gabor filter response.

Edge Orientation Histogram (EOH)

EOH describes spatial distribution of four edges and one non-directional edge in the image. It is scale invariant.

Classifier

The extracted features are given to a KNN classifier for training. The features extracted from the query image are given to the trained KNN classifier to classify the query image to given set of fruit classes.

V. EXPERIMENTAL RESULTS

The Data Set used consists of the fruit images that are collected from Google and the Tropical fruit database. The dataset consists of 141 images, in which 70 images were assigned for training and 71 images for testing, with 7 classes of fruits. Among 71 test images, 35 images were of single fruits, 18 were groups of fruits of same class and the rest were images of slices of fruits.

A. Preprocessing (Segmentation)

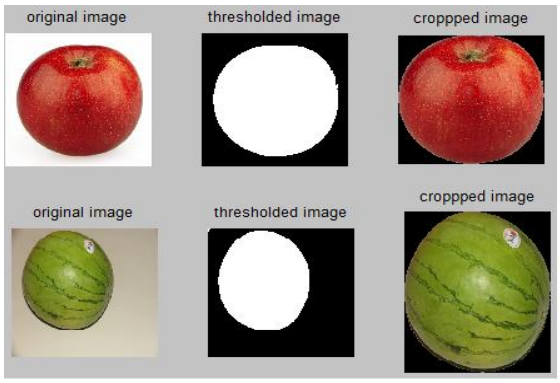

Fig. 1: Segmented and cropped fruit image

B. Observations

Table 1: Observations

Sl. No:	Case	Example images	Features used	No: of test images used	No: of correct outputs
1	• Single fruit images • Three features extracted for an image	 Fig.2: Watermelon	• CLD • HTD • EHD	35	26
2	• Single fruit images • four features extracted for an image	 Fig.3: Apple	• CLD • DCD • HTD • EHD	35	29

3	• Group of fruits of same type • four features extracted for an image	 Fig.4: Oranges  Fig.5: Banana	• CLD • DCD • HTD • EHD	18	6
4	• Slice of fruits and four features extracted for an image	 Fig.6: Banana  Fig.7: Apple	• CLD • DCD • HTD • EHD	18	3

VI. SCOPE OF FUTURE WORK

- improve segmentation by using some advanced segmentation techniques like contour based segmentation
- decrease the complexity of the system
- Include more number of MPEG 7 descriptors like scalable color descriptor, Region Based Shape Descriptor, etc.
-

VII. CONCLUSION

The current system is just a prototype for implementation of automatic image annotation, using a fruit image database. A comparison of annotation results is made by taking different number of mpeg 7 feature descriptors. The next steps will be to overcome the pitfalls of previous stages, and design a tool for a larger dataset with higher accuracy and implement and attach a CBIR tool for retrieval of similar images for a query image.

REFERENCES

[1] Dengsheng Zhang, Md. MonirulIslam, GuojunLu, "A review on automatic image annotation techniques", *Pattern Recognition*, May 2011

[2] Syaifulnizam Abd Manaf, Md. Jan Nordin "Review on Statistical Approaches for Automatic Image Annotation", *International Conference on Electrical Engineering and Informatics*, August 2009.

[3] V. Lavrenko, S. L. Feng, R. Manmatha, STATISTICAL MODELS FOR AUTOMATIC VIDEO ANNOTATION AND RETRIEVAL", *The International Conference on Acoustics, Speech and Signal Processing*, May 2004.

[4] Ying Liu, Dengsheng Zhang, Guojun Lu, "Region-based image retrieval with high-level semantics using decision tree learning", *Pattern Recognition* , Vol. 41, December 2007.

[5] R. C. F. Wong, C. H. C Leung, "Automatic Semantic Annotation of Real-World Web Images" , *IEEE Transactions on Pattern analysis and Machine intelligence*, Vol. 30, November 2008.

[6] Ja-Hwung Su, Chien-Li Chou, Ching-Yung Lin, Vincent S. Tseng, "Effective Semantic Annotation by Image-to-Concept Distribution Model", *IEEE Transactions on multimedia*, Vol. 13, June 2011

[7] Antonio Torralba, Bryan C. Russell, and Jenny Yuen "LabelMe: Online Image Annotation and Applications", *Proceedings of the IEEE*, August 2010

[8] Balasubramani R., V.Kannan,"Efficient use of MPEG-7 Color Layout and Edge Histogram Descriptors in CBIR Systems", *Global Journal of Computer Science and Technology* , 2009

[9] B. S. Manjunath,, Jens-Rainer Ohm, Vinod V. Vasudevan, and Akio Yamada "Color and Texture Descriptors", *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY*, VOL. 11, NO. 6, JUNE 2001

[10] Dong Kwon Park , Yoon Seok Jeon, Chee Sun Won "Efficient Use of Local Edge Histogram Descriptor", *Proceeding MULTIMEDIA ACM workshops on Multimedia*, 2000

[11] Antti Ilvesmäki, Jukka Iivarinen, "DEFECT IMAGE CLASSIFICATION WITH MPEG-7 DESCRIPTORS", *Proceedings of the 13th Scandinavian Conference on Image Analysis*, 2003.

[12] Horst Eidenberger, "How good are the visual MPEG-7 features?", *SPIE & IEEE Visual Communications and Image Processing Conference*, 2003.

[13] Ying Liua, Dengsheng Zhanga, Guojun Lua,Wei-Ying Mab ,"A survey of content-based image retrieval with high-level semantics", *Pattern Recognition* 40, 28 April 2006.

[14] RITENDRA DATTA, DHIRAJ JOSHI, JIA LI, and JAMES Z. WANG "Image Retrieval: Ideas, Influences, and Trends of the New Age", *ACM Computing Surveys*, April 2008.

[15] Thomas Sikora, "The MPEG-7 Visual Standard for Content Description—An Overview", *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY*, VOL. 11, JUNE 2001

[16] Yong Man Ro, Munchurl Kim, Ho Kyung Kang, B.S. Manjunath, and Jinwoong Kim, "MPEG-7 Homogeneous Texture Descriptor", *ETRI Journal*, Volume 23, June 2001

AUTHORS

First Author – MANJARY P. GANGAN, Department of Computer Science and Engineering, Vidya Academy of Science and Technology, Thrissur, 680501, India, manjaryp@gmail.com
Second Author – Dr. R. KARTHI, Department of Computer Science, Amrita Vishwa Vidyapeetham, Amrita University, Coimbatore, 641105, India, r_karthi@cb.amrita.edu