Journal of Engineering and Applied Sciences 12 (Special Issue 7): 8021-8024, 2017
ISSN: 1816-949X

Image Retrieval using K-Means Clustering and Image Annotation

Rucha S. Patil and Avinash J. Agrawal
Department of Computer Science and Engineering,
Shri Ramdeobaba College of Engineering and Management, Nagpur, India

Abstract: An image retrieval system is one of the important computer systems for browsing and retrieving images from a large database. There are two approaches for image retrieval, Text Based Image Retrieval (TBIR) and Content Based Image Retrieval (CBIR). The drawback of TBIR is manual annotation which is impossible and expensive task for large database. The problems in TBIR have raised the interest of researchers to come up with techniques for retrieving images. Image annotation task assign a set of semantic tags or keywords to an image based on some models learned from certain training data. Automatically assigning keywords to images is of great interest as it is an active topic of research in computer vision and pattern recognition. In this study, we proposed a framework for content based image retrieval using k-means clustering and also performed image annotation using adaptive threshold.

Key words: Image annotation, content based image retrieval, k-means clustering, adaptive threshold, text based image retrieval, performed

INTRODUCTION

Digital image processing is the technology which is used for applying number of computer algorithms for processing a digital image. Digital image processing in a straight line deals with an image which is poised with many image points. These image points are also known as pixels are of spatial coordinate that specify the position of the point in the image and intensity values. An image retrieval system is a computer system for browsing, examining and retrieving images from a huge database of digital images.

In today's modern age, image retrieval is used for efficient service in various areas like commerce, architecture, journalism, crime prevention, fashion and historical research. A large collection of these images is referred to as image database. In image retrieval the aim is to find the images from the database that are relevant to user query. With the devastating growth of image databases, effectual image coding, manipulation, indexing and retrieval are required for efficiently dealing with large image databases and make images easily manageable.

There are two main scenarios of image retrieval where in the former user query is image and goal is to final the similar images from database relevant to query image. For the latter, user query is the keyword and goal is to find the images from database relevant to query keyword. In this case, the images in the database should be indexed with the keywords and these keywords of the images are created by human operator. But, manual annotation of images in large database is an inefficient and expensive task, hence, the interest in automatic image annotation increase.

For years Automatic Image Annotation (AIA) has become an active and challenging research topic in computer vision and pattern recognition areas due to its potential impact on both image understanding and semantic based image retrieval. The aim of AIA is to find suitable annotation words to represent the visual content of an untagged tagged image. But, it is difficult task due to lack of correspondence between keywords and image regions. Hence, many machine learning algorithms have been proposed to solve this problem by making use of ground truth (Datta *et al*., 2008).

LITERATURE REVIEW

This study discusses related research in the specific area of content-based image retrieval and image annotation.

Ma *et al*. (2010), they extract four kinds of effective global features Grid color moment, Local Binary Pattern (LBP), Gabor wavelet texture, edge using canny edge detector from every image, they create image similarity graph and then form a hybrid graph with the image-tag bipartite graph. After building the hybrid graph, they propose a novel and effective random walk model that

Corresponding Author: Rucha S. Patil, Department of Computer Science and Engineering,
Shri Ramdeobaba College of Engineering and Management, Nagpur, India

employs a fusion parameter to balance the importance between the image contents and the tags and also provide a natural solution for including the pseudo relevant feedback into image retrieval and annotation tasks. The experimental results on a large flickr dataset show the advantage of their proposed framework.

Rahman *et al.* (2012) texture features is extracted using invariant Gabor descriptor from the image and similarity measure Canberra distance was applied for retrieval of images. The Brodatz texture database and food database were used for experimentation.

Tian (2015) presented a number of promising SVM techniques for Automatic Image Annotation (AIA) have been presented, so, as to complement the prevailing surveys in literature. They specifically spotlight on SVM for automatic image annotation from three traits including SVM ensemble for AIA, SVM with concoction of kernels for AIA and hybrid SVM for AIA, respectively. The main purpose of their study was to exemplify the pros and cons of SVM combined with a great deal of current works.

Jeon *et al.* (2003) have illustrated that cross-media relevance models are a decent choice for annotating and retrieving images. Three different models were proposed and tested. They found that Fixed Annotation-based Cross-Media Relevance Model (FACMRM) model precision is twice as good as a state of the art translation model in annotating images. They also presented how to achieve ranked retrieval using some of their models.

Zhang *et al.* (2010) introduced a regularization based feature selection algorithm to influence both the sparsity and clustering properties of features and include it into the image annotation task. They also proposed a novel approach iteratively gain similar and dissimilar pairs from both the keyword similarity and the relevance feedback. Several experiments are designed to compare the performance between features, feature combinations and regularization based feature selection approaches applied on the image annotation task which gives vision into the properties of features in the image annotation task. The experimental results determine that the group sparsity based technique is more precise and stable than others.

PROPOSED WORK

Figure 1 shows the propose system for image retrieval using k-means clustering. In this proposed system, we will extract the features from database images and store it in the feature vector. In next step, we perform k-means clustering on feature vector of database which forms 'k' clusters. When, the query image is given to the system, the features are extracted from the query image and store

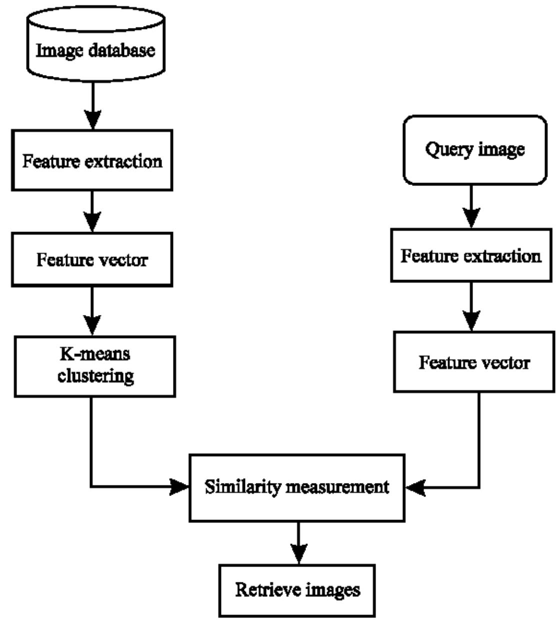

Fig. 1: Proposed system for image retrieval using k-means clustering

it in the vector. The similarity measurement is calculated between query feature vector and the centroid of 'k' clusters. For similarity measurement, we use Euclidean distance, the cluster which has smallest distance to query feature vector is selected and the images are retrieve from the selected cluster relevant to the query image.

We will also perform image retrieval without clustering in which Euclidean distance is calculated between query image and each image in the database. We will compare the retrieval time require by system with clustering and without clustering. Following sub-section describe the methods used for feature extraction, k-means clustering, similarity measurement and image annotation.

Feature extraction

Color moments: In this method, color features are extracted using two moments of each color channel of an image which are mean and standard deviation (Lee *et al.*, 1998; Munje and Kapgate, 2014). The first moment mean (E_i) is given by Eq. 1. The second moment standard deviation (σ_i) is given by Eq. 2:

$$E_i = \frac{1}{m.n}\sum_{j=1}^{m.n} P_{ij} \tag{1}$$

$$\sigma_i = \left[\frac{1}{m.n}\sum_{j=1}^{m.n}(P_{ij}-E_i)^2\right]^{1/2} \tag{2}$$

Where:
P_{ij} = A value of each color channel at jth image pixel
m.n = The total number of pixels per image

Gray-Level Co-occurrence Matrix (GLCM): To create the GLCM we have convert RGB query image into gray-scale image. GLCM depends upon the orientation and distance between the pixels. The GLCM matrix finds how often a pixel with the intensity value i occurs in a specific spatial relationship to a pixel with the value j. Haralick *et al.* (1973) proposed 28 kinds of textural features each extracted from the gray level co-occurrence matrix. The most common texture features extracted from GLCM are contrast, entropy, energy and homogeneity.

Contrast: Contrast measures the local variations in the gray-level co-occurrence matrix:

$$\text{Contrast} = \sum_{i,j} (i-j)^2 p(i, j) \tag{3}$$

Entropy: Entropy gives measures of complexity of the image and this complex texture tends to higher entropy:

$$\text{Entropy} = \sum_{i} \sum_{j} p(i, j) \tag{4}$$

Energy: Energy is the sum of squared elements in the GLCM and it is by default one for constant image:

$$\text{Energy} = \sum_{i,j} (i, j)^2 \tag{5}$$

Homogeneity: Homogeneity measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal:

$$\text{Homogeneity} = \sum_{i,j} \frac{p(i, j)}{1+(i-j)^2} \tag{6}$$

Histogram of edge directions: The general shape information in the image is capture by the edge histogram. To obtain edge information, we will use prewitt edge detection algorithm. The edge directions are quantized into a number of bins (Zhang *et al*., 2008; Rahman *et al*., 2007) and histogram is normalized with respect to the number of pixels in the image for achieving scale invariance.

K-means clustering: Clustering is a technique of grouping together data samples that are similar in some way according to some criteria that we pick its form of unsupervised learning (Liu and Yu, 2009).

K-means clustering is a method commonly used to automatically partition a data set into k groups. It proceeds by selecting k initial cluster centres and then iteratively refining the results. The algorithm stops when there is no further change in assignment of instances to clusters.

Algorithm for K-means:

a. Decide on a value for k
b. Initialize the k cluster centers (randomly if necessary)
c. Decide the class memberships of the N objects by assigning them to the nearest cluster center
d. Re-estimate the k cluster centres by assuming the memberships found above are correct
e. If none of the N objects changed membership in the last iteration, exit. Otherwise go to step c

In this way, the clusters of similar kind are formed and it helps to reduce the elapsed time of the system.

Similarity measurement: We will use Euclidean distance for similarity measurement. The Euclidean distance between the query image and the data base image is given by Eq. 7:

$$D(q, d) = \sum_{i=1}^{n} (q-d_i)^2 \tag{7}$$

With clustering: In this first distance of query image with 'K' cluster centroids will be calculated. The cluster which has smallest distance with query image will be selected. Later the distance between query image and database image belong to selected cluster is calculated. This process is repeated until all the images in the selected cluster have been compared with the query image. After completion of this algorithm, we have array of Euclidean distance which is then sorted.

Without clustering: Where q is the query image and d be the database image. The distance between query and data base image is calculated and least value of distance is taken. This process is repeated until all the images in the database have been compared with the query image. After completion of this algorithm, we have array of Euclidean distance which is then sorted.

Image annotation: We will also perform image annotation for query image. An image annotation will perform by using adaptive thresholding method. For experiment results of the proposed system we will use standard corel5k database. In corel5k database images are labeled with 2-5 tags or keyword. So, the images retrieved by our system will contain tags, these tags will be input for image annotation method. In adaptive thresholding method, the unique tags will be selected and also find the instances of these unique tags in images which are retrieved by the system. Then, we will calculate the sum of these unique tag instances and find the mean. The calculated mean is used as threshold, the unique tags which have instances greater this threshold will annotate the query image.

CONCLUSION

The proposed system will able to retrieve the images similar to the query image as well as annotating the query image. It will also compare the retrieval time required by our system with clustering and without clustering. The proposed system use the combination of color moments, GLCM, histogram of edge direction and K-means clustering algorithm that considers the similarities among images in database. K-means clustering technique is helpful to reduce the elapsed time of the system. The system will also perform image annotation using adaptive thresholding.

SUGGESTION

In the future, the proposed system can be combined with concept of relevance feedback to improve the performance.

REFERENCES

Datta, R., D. Joshi, J. Li and J.Z. Wang, 2008. Image retrieval: Ideas, influences and trends of the new age. ACM. Comput. Surv., 40: 34-94.

Haralick, R.M., K. Shanmugam and I.H. Dinstein, 1973. Textural features for image classification. IEEE Trans. Syst. Man Cybern., SMC-3: 610-621.

Jeon, J., V. Lavrenko and R. Manmatha, 2003. Automatic image annotation and retrieval using cross-media relevance models. Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval, July 28-August 01, 2003, ACM, Toronto, Canada, ISBN:1-58113-646-3, pp: 119-126.

Lee, S.M., H.J. Bae and S.H. Jung, 1998. Efficient content-based image retrieval method using colour and texture. ETRI. J., 20: 272-283.

Liu, H. and X. Yu, 2009. Application research of k-means clustering algorithm in image retrieval system. Proceedings of the 2nd Symposium on International Computer Science and Computational Technology (ISCSCT'09), December 26-28, 2009, Huangshan Publishing House, China, ISBN:978-952-5726-07-7, pp: 274-277.

Ma, H., J. Zhu, M.R.T. Lyu and I. King, 2010. Bridging the semantic gap between image contents and tags. IEEE. Trans. Multimedia, 12: 462-473.

Munje, P.N. and D. Kapgate, 2014. Novel content-based image retrieval techniques for medical database. Intl. J. Adv. Comput. Sci. Technol., 3: 376-379.

Rahman, M.H., M.R. Pickering, M.R. Frater and D. Kerr, 2012. Texture feature extraction method for scale and rotation invariant image retrieval. Electron. Lett., 48: 626-627.

Rahman, M.M., P. Bhattacharya and B.C. Desai, 2007. A framework for medical image retrieval using machine learning and statistical similarity matching techniques with relevance feedback. IEEE. Trans. Inf. Technol. Biomed., 11: 58-69.

Tian, D., 2015. Support vector machine for automatic image annotation. Intl. J. Hybrid Inf. Technol., 8: 435-446.

Zhang, J., G.L. Li and S.W. He, 2008. Texture-based image retrieval by edge detection matching GLCM. Proceedings of the 10th IEEE International Conference on High Performance Computing and Communications (HPCC'08), September 25-27, 2008, IEEE, New York, USA., ISBN: 978-0-7695-3352-0, pp: 782-786.

Zhang, S., J. Huang, Y. Huang, Y. Yu and H. Li *et al*., 2010. Automatic image annotation using group sparsity. Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'10), June 13-18, 2010, IEEE, San Francisco, California, ISBN: 978-1-4244-6984-0, pp: 3312-3319.