IJCTA, 9(26) 2016, pp. 41-46

Content Based Image Retrieval System Data Mining Combined K-means, Relevance Feedback Techniques

N. Pavithra[1] and K.Sivaranjani[2]

ABSTRACT

Content Based image retrieval system (CBIR) low- level visual image features that is color, texture, and shape are automatically extract for image descriptions and indexing purposes. In this paper popularity of network and development of multimedia technology, the established information retrieval techniques are not working efficiently according to users command in search and retrieving images from record. In today's world there is increased need of content based image retrieval system in number of different domains such as education, medical imaging, crime anticipation, whether forecasting, secluded sensing and association of earth property. It works on the features of images like color and texture. In this paper an enhancement to basic content based image retrieval procedure with indexing maintain by using K-means clustering data mining and relevance feedback technique. The enhanced feature helps in retrieving images from large folder initially. In this process an index is practical on database of images based on clustering technique. for the duration of this process clustering paper uses features like texture, color, shape, relevance pointer and wavelet based histogram system to find similarity among the images. Based on association value the images are separated into clusters, then the new image which is to be confirmed with database is compared with these clusters and based on its likeness corresponding images in cluster are retrieved.

Keywords: content based image retrieval(CBIR), query refinement, relevance feedback ,k-mean clustering; shape; color; texture; color structure descriptor(CSD);textbased image retrieval(TBIR).

I. INTRODUCTION

Image retrieval process is searching and retrieving images from large data sets .As the images grow diverse and complex, retrieving the right images becomes a complex challenge. With the development of internet and the inventions of image capture devices like digital cameras, image scanners , the size of the digital image collection has increased rapidly it is very complicated for storing and retrieving digital images which have different application such as fashion architecture, design, medicine, crime prevention, virtual museums, military and security features and personal photo albums. Therefore an efficient image recovery system is required. The image retrieval system is classified into two types such as Text based image retrieval and Content based image retrieval. The text based image reclamation was introduced in the year 1970s in this technique the images are manually tagged by text descriptors which are used by database management system(DBMS) for the stage image retrieval. But the text based retrieval method has a lot of demerits such as loss of information, time overriding, more expensive task, inaccuracy and decreased efficiency.

The aim of this development is to review the current state of the art in content-based image retrieval (CBIR), a technique for retrieve images on the origin of automatically-derived covering quality such as color, texture and shape. Our findings are based both on a review of the important literature and on discussions with researchers in the field.

[1] Research Scholar, Department of Computer Science, Bishop Heber College, Tiruchirappalli, Tamilnadu, India, *E-mail: nvpavibb@gmail.com*.

[2] Asst. Professor, Department of Information Technology, Bishop Heber College, Tiruchirappalli, Tamilnadu, India.

II. METHODOLOGY

2.1. Content Based Image Retrieval using Image Mining Techniques

In this section presenting the notion of use of image mining techniques for effective content based image retrieval. Most of the CBIR techniques employ the color and texture features only to retrieve required images from the image database. Image mining process helps to determine different image patterns for each of the images in image database and query image as well. Similarity among target images with the query image is decided on the basis of the pattern which is similar to them. We can Euclidian distance measure or MLE i.e. Maximum likelihood Estimation. Following steps are involved in content based image retrieval using image mining techniques.

Step 1: Retrieval Based on Color Feature

The color is the majority important feature used in retrieving the images mostly based on color histogram. Color are define on the selected color space. Universal color features in CBIR system include color histogram, color moment and color correlogram.

Color Histogram Color histogram describe the proportion of specific colors in an image and has been widely used procedure used for content based image retrieval because of its efficiency and usefulness. Color histograms have many advantages such as quickness and low demand of memory space, Color histogram provides HSV color space and RGB color space

RGB color space: The important color space is RGB color space which stands for Red-Green-Blue. This gap consists of main colors such as light Red, Green, and Blue.

HSV color space: HSV color gap is distinct based on the terms of three part workings: Hue, dispersion, Value.

Color Correlogram Color correlogram are the particular feature about the color in a row of an image. Color correlogram and the color rationality vector has the advantage of merge the spatial link of color regions to describe the universal distribution of local spatial correlation of colors.

Color Moments: The color moment are used to make different the image based on the color feature of the image. It provides the quantify of color similarities free among the images.

Step 2: Retrieval Based on Shape Feature

It the main requirement at the primitive level. It is identified that all the normal objects are accepted by their shapes. In image retrieval methods, depending on the applications some needs the shape image such as alteration, rotation and scaling. 2 important types of shape features are commonly used –universal features such as circularity, feature ratio and instant invariants and the limited features such as successive limit segments. The shape representation can be secret into two category :

- Boundary based which uses the boundary or the draw round of the shape.
- Area based which use the unbroken expanse of the shape.

2.2. Algorithms of K-MEANS CLUSTERING

Clustering is a way of group together data sample that are similar in some way according to some criteria that we pick its form of unsupervised learning So, it's a method of data investigation – a way of looking for patterns or structure in the data that are of attentiveness Clustering algorithms are generally used in an unsubstantiated fashion. They are on hand with a set of data instances that must be group according to some notion of comparison. The algorithm has way in only to the set of features describing each object; it is not given any information as to where each of the instances should be placed within the partition.

K-means clustering is a method normally used to automatically partition a information set keen on k groups. It proceeds by selecting k original cluster centers and then iteratively refinement the results. The algorithm converges when there is no additional change in assignment of instances to clusters.

THE K-MEANS ALGORITHM

Algorithm: k-means. The k-means algorithm for partition based on the mean value of the objects in the cluster.

Input: The number of clusters k and a record containing n objects.

Output: A set of k clusters that minimizes the squared-error decisive factor.

Method:

(1) Randomly decide k objects as the primary cluster centers:

(2) Repeat

(3) (Re) disperse each entity to the cluster to which the item is the most similar, based on the mean value of the objects in the huddle;

(4) modernize the cluster means, i.e., estimate the mean value of the objects for all cluster;

(5) Until no alteration.

2.3. Relevance feedback

A relevance feedback technique allows the user to interact with the image retrieval algorithm by representing the information of an image which the user thinks that it is relevant to the query. Keyword based image retrieval in performed by similar the keyword based on the user input with the image in the database. However some images may not have correct keywords to explain the image and therefore the image retrieval becomes difficult. Relevance feedback provides the solution to the more than declare difficulty and uses user pointer to reduce the error and redundancy. It also uses Bayesian classifier to create positive and negative feedback. To produce more accuracy log based clustering methods are used.

Relevance Feedback (User Purpose) Algorithm

Step 1. Start

Step 2. Firstly feature drawing out of all the images of the record is performing.

Step 3. User input a query image.

Step 4. Feature extraction of query image is performed.

Step 5. Relevance feedback procedure is implemented using query reweighting, query expansion, query point movement done.

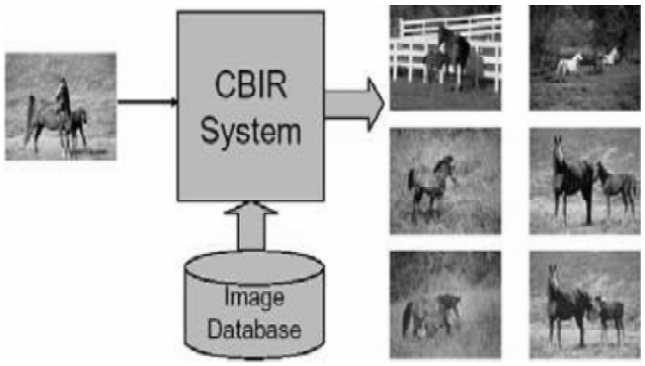
Step 6. Result is displayed.

Step 7. If user satisfied then searching get complete.

Step 8. Else repeat process of query reweighting, query development, query point interest group technique.

Step 9. User puts his pointer after exploratory the searched result.

2.4. Content-Based Approach:- Index images using images



CBIR Search pictures as Pictures

Advantages

- Image features, such as color, texture, and shape information, of images are extracted automatically.
- similarity of images are based on the distance between features

III. EXPERIMENT AND RESULTS

The proposed method has been implemented using .NET Technology and tested on a general-purpose database containing 500 images in JPEG format of size 384x256 and 256x386. The search is usually based on similarity rather than the exact match. The excellence of the image retrieval, with different quantization schemes has been evaluated by randomly selecting sample the image database of 600 hundred images is used to evaluate the performance of CBIR system. We have selected a database of almost thousand images and then divide the database into number of concepts of different complexities to verify the results. The comparison among the all fast query point movement techniques and by applying K-means clustering the elapsed time for all these methods is reduced effectively The comparison of time is enlisted in the table for twenty images in the database . The time is measured in seconds for each method without K-means.It's clear that by applying K-means technique the elapsed time for target search methods reduced effectively. There is a big difference in elapsed time for all these methods. it's clear that the elapsed time for the CBIR system is reduced effectively.

RESULT

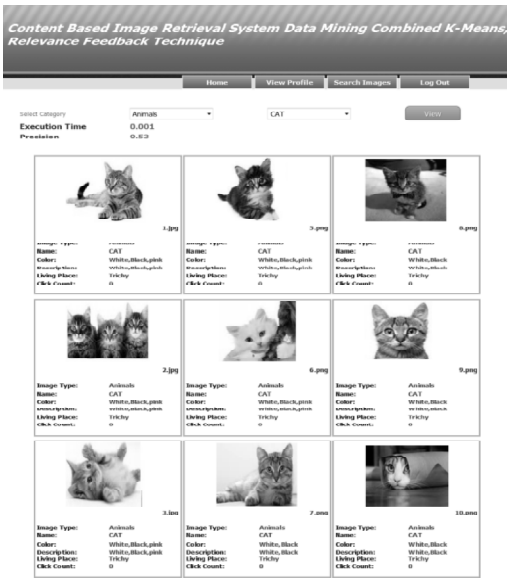


Table 1
General analysis Image retrieval elapsed time seconds

Algorithms	*Total No of Images*	*Elapsed Time*
K-Means & RF	Images 600	0.001

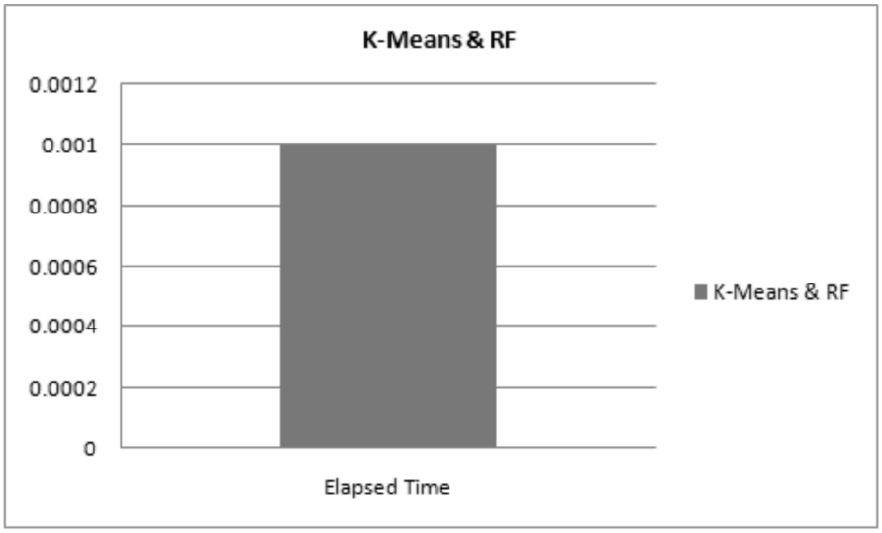


Chart 1 : General analysis Image retrieval elapsed time seconds

In a similar way precision value can be kept in a value by retrieving only few images or precision should either be used mutually or the number of images retrieved must be specified.

Table 2
General analysis Image retrieval precision

Algorithms	*Total No of Images*	*Precision*
K-Means &RF	Images 600	0.05

$$\text{Precision} = \frac{\text{No of relevant images retrieved}}{\text{Total no of images retrieved}}$$

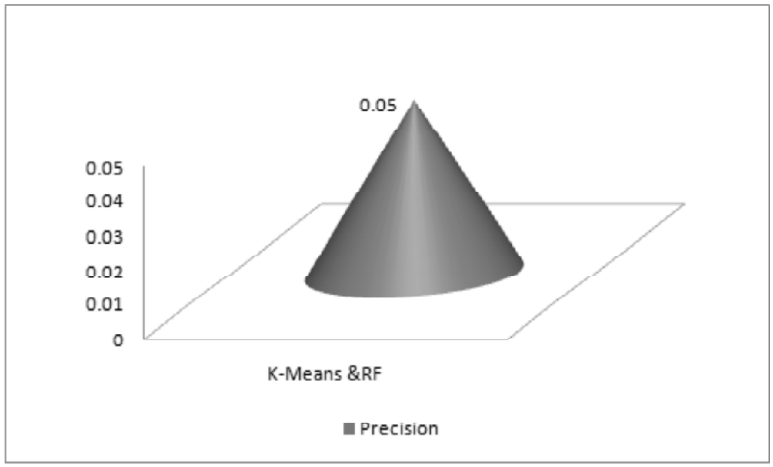


Chart 2 : General analysis Image retrieval precision

V. CONCLUSION

This paper summarizes the specific contribution from the work and outlines some direction of future work. Review of previous study reveals that as manual observations are required in the text-based retrieval using content-based approach is more desirable than text-based approach. Color is low level feature. In addition to color, texture feature is used. For CBIR this algorithm is based on color, texture and shape features of static image. The classification technique nearby very little amount of memory for features storage and a prominent rate of computation and give good results in terms of accuracy. k-means clustering is fairly useful for appropriate image retrieval queries (0.001) seconds, precision (0.05). The K-means clustering algorithms to group

The images content into different clusters based on the color feature and k-means clustering have been often used in the pattern recognition. The results are proved by using the K- Means clustering concept in terms of the elapsed time, Precision.

REFERENCES

[1] Ahmad Alzu'bi "Semantic content-based image retrieval: A comprehensive study" Journal of Visual Communication and Image Representation. Volume 32, October 2015, Pages 20–54

[2] Y. Chen and J. Z. Wang, "A Region-Based Fuzzy Feature Matching Approach to Content Based Image Retrieval", IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 24, No.9, pp. 1252-1267, 2012.

[3] R. C. Gonzalez and E.W. Richard, Digital Image Processing, Prentice Hall. 2011.

[4] N. Jhanwar, S. Chaudhurib, G. Seetharamanc and B. Zavidovique, "Content based image retrieval using motif co-occurrence matrix", Image and Vision Computing, Vol.22, pp-1211–1220, 2004.

[5] P.W. Huang and S.K. Dai, "Image retrieval by texture similarity", Pattern Recognition, Vol. 36, pp- 665–679, 2013.

[6] S. Harpreet, W. Equits, M. Flickner and W. Niblack, "Efficient Color Histogram Indexing for Quadratic Form Distance Functions", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No. 7, 2010.

[7] P. S. Hiremath and J. Pujari, "Content Based Image Retrieval based on Color, Texture and Shape features using Image and its complement", 15th International Conference on Advance Computingand Communications. IEEE. 201

[8] "Introduction to Data Mining and Knowledge Discovery", Third Edition, Two Crows Corporation.

[9] S. Livens, P. Scheunders, G. V. D. Wouwer and D. V. Dyck, "Wavelets for texture analysis, an overview", Proceedings of Sixth International Conference on Image Processing and Its Applications, Vol. 2, pp-581–585, 2013.

[10] Nidhi Singh "A Novel Approach for Content Based Image Retrieval" 2nd International Conference on Computer, Communication, Control and Information Technology.Department of Computer Science & Technology, ABES Engineering College, U.P., India. Procedia Technology Volume 4, 2012, Pages 245–250. June 2012.

[11] Paresh Marwaha et al et al -"Content Based Image Retrieval in Multimedia Databases"-Jaypee Institute of Information Technology University, Noida, India Vol. 1, No. 2- 2012.

[12] G. Raghupathi, R.S. Anand, and M.L Dewal, "Color and Texture Features for content Based image retrieval", Second International conference on multimedia and content based image retrieval, July-21-23, 2010.

[13] Samar Zutshi et al "Proto-Reduct Fusion BASED Relevance Feedback IN Cbir",(Monash University)–2010.

[14] Tang li et al -"Developing a Shape-and-Composition CBIR Thesaurus for the Traditional Chinese Landscape"-University of Maryland College of Information Science College Park, MD, United States Library Student Journal, July 2012.

[15] G. V. D. Wouwer, P. Scheunders and D. V. Dyck, "Statistical texture characterization from discrete wavelet representation", IEEE Transactions on Image Processing, Vol.8, pp-592–598, 2012.