International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 04 Issue: 06 | June -2017 www.irjet.net p-ISSN: 2395-0072

CBIR Processing Approach on Colored and Texture Images using KNN Classifier and Log-Gabor Respectively

Arti, Astt.Prof. Nancy

Department of CSE, International Institute of Engineering and Technology (Samani,KKR) India

---

Abstract: Content Based Image Retrieval (CBIR), also called as Query By Image Content (QBIC). Content Based Image Retrieval is the method to retrieve stored image from database by supplying query image instead of text. This is achieved using proper feature extraction and matching process. Here we have implemented two methods of content based image retrieval using color and texture. In feature extraction of color is done using classifiers and similarity measure,color moment. While feature extraction of texture is done using wavelet texture features and Log-Gabor features. Finally we have retrieved top images using euclidean distance and chisquare distance and we have made comparative analysis. Content Based Image Retrieval has endless discussion to do. Here we can say that results or retrieval ratio depends upon image class for some images. we can have better precision and time complexity while some images give average result. Finally comparative analysis given in table 1 says that overall precision and time complexity given by combined approach using classifier, similarity measurev and log-gabor respectively color and texture gives better result as compared to wavelets and gabor filter. Different types of classification we can use Neural network, Support Vector Machine (SVM), KNN, Bayesian etc. In this paper, we are using K Nearest Neighbor (KNN) classifier to find out the relevant images and after that we use Spearman's Rank Correlation Function to reduce the time complexity and improve F-measure. Hence if we want to improve retrieval efficiency we have to use some other approach. Here for effective retrieval we can use other features like shape.

Keywords: CBIRS, Image databases, Color string comparison, Feature extraction, Query image, Target Image.

I. INTRODUCTION

With the headway in internet and multimedia technologies, a immense amount of multimedia data in the form of audio, video and images has been utilized as a part of numerous fields like medical treatment, satellite data, video and still images repositories and surveillance system. This has made a progressing interest of frameworks that can store and retrieve mixed media information in a powerful way.

Numerous multimedia information storage and retrieval systems have been developed till now to cater these requests.

The most common retrieval systems are Text Based Image Retrieval (TBIR) systems, where search is based on automatic or manual explanation of images. A conventional TBIR searches database for the similar text surrounding the image as given in the query string. The TBIR systems are fast as the string matching is computationally less time consuming process. In addition annotation of images is not always correct and takes a lot of time. For finding alternative way of looking and overcoming the limitations forced by TBIR systems more natural and easy to understand content based image retrieval systems (CBIR) were developed. A CBIR system uses visual contents of images described in the form of low level features like color, texture, shape and spatial locations to represent images in databases. The system retrieves comparable images when an query image or sketch is presented as input to the system. Querying in this way removes need for describing the visual content of images in words and is near to human perception of visual data. A portion of representative CBIR systems is Query by Image Content (QBIC).

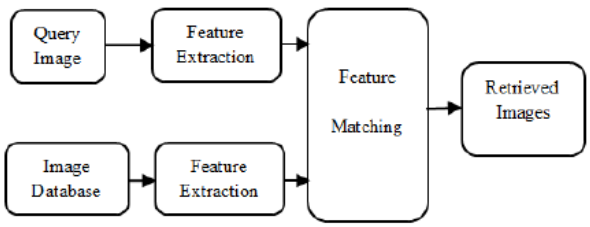


Fig.(1) Block Diagram of CBIR

In a typical CBIR system (Figure 1.2), image low level features like color, texture, shape and spatial locations are represented in the form of a multidimensional feature vector. The query image is converted into the internal representation of feature vector using the same feature extraction routine that was used for building the feature

database. The similarity measure is employed to calculate the distance between the feature vectors of query image and those of the target images in the feature database. Finally, the retrieval is performed using an indexing scheme which facilitates the efficient searching of the image database. Recently, user's relevance feedback is also incorporated to further improve the retrieval process in order to produce perceptually and semantically more meaningful retrieval results. In this, we discuss these fundamental techniques for content- based image retrieval.

II. RELATED WORK

M. Kaipravan and Rejiram R, et al [28] In this paper, we present a CBIR system based on integration of both color and texture feature. Due to the poor discriminating power of color histogram, color moments that encode some spatial information are used to extract the color feature from the image. Gabor filter is used to represent the texture feature. Then we assign weights to each feature and calculate the similarity of combined features using Manhattan distance measure. we have proposed an efficient image retrieval system based on combination of color moments and Gabor texture feature. Gabor filter are adopted to extract texture feature and feature vector of length 48 is obtained. We calculate the similarity with combined features of color and texture using Manhattan distance as the similarity measure. The proposed method has higher retrieval accuracy than other conventional methods.

P. B. Patil and M. B. Kokare, et al [12] Due to the semantic gap between low-level image features and high level concepts, we have presented a framework for effective image retrieval by proposing a novel idea of cumulative learning using Support Vector Machines (SVM). It creates a knowledge base model to increase the training samples by simply accumulating the samples based on user interactions. As we know relevance feedback (RF) is online process, so we have optimized the learning process by considering the most positive image selection on each feedback iteration. To learn the system we have used SVM. The main significances of our system are to address the small training sample and to reduce retrieval time. In this paper, an active relevance feedback framework has been proposed to handle the small training data problem in RF and optimizing the testing set in order to reduce the retrieval time. The proposed relevance feedback framework, we found that RF using SVM with combined texture features RCWF and DT-CWT gives better retrieval performance than contourlet and curvelet texture features.

R. Sudhakar, K. R. Krishnan and S. Muthukrishnan, et al[3] with an increase in the awareness of internet usage, there has been an explosion of data on the web. The problem of retrieving near approximate images using textual queries has always been an area of research. This focuses on bridging the gap between textual search input given by the user and the images retrieved from the database, by making use of visual features instead of the file name.

The work concentrates on employing a simple keyword extraction technique rather than using complex NLP techniques. Also, an automatic segmentation method is proposed to avoid human intervention. It refining the segmentation algorithm and improving the template matching techniques would further improve the retrieval efficiency. the spot and eye detection methods can be applied on medical image datasets to identify circular portions such as cysts, tumors or organs. This paper considers a single isolated object.

Roshi Choudhary et al. [7] proposed an approach to perform content based image retrieval. It is an integrated approach used to extract color and texture feature from images. By using single feature, correct results can never produced. So multi feature extraction is more beneficial to perform image retrieval. To extract the color feature, higher order of color moment is used which is the descriptor of color. To extract texture, LBP is used which is the descriptor of texture. Local binary pattern is mainly used to face recognition.

Vinee. V. Kawade et al. [10] announced a user based system for CBIR in which genetic algorithm is applied. The different features of color image like mean, standard deviation and the image bitmap are used for retrieval. The texture features like edge histogram of an image and the entropy of gray level co-occurrence matrix are used. Moreover, the genetic algorithm is applied to help user in identifying the images which satisfy his needs for reducing gap between the users' expectation and the retrieval results. Experimental results show remarkable improvement in the performance after applying IGA.

L. Chai, H. Zhang, Z. Qin, J. Yu and Y. Qi, et al[4] Content-based image retrieval (CBIR) has got an intense interest and seen considerable progress over the last decade. But most of the time it is only applied in laboratory. One important reason for this is the diversity of images.. At present, and even in the foreseeable future, a general purpose CBIR system is not really possible. In this paper, we propose a region-based method fit for the content-based retrieval of product images. The method focuses on two key issues: fast extraction of the main region, in which the product locates, as well as efficient shape and color features extraction. To show the validity of the proposed region-based method, compared experiments are carried out and illustrated on the PI 100 dataset.

S. Selvarajah and S. R. Kodithuwakku, et al [20] Representation of visual features and similarity match are

important issues in CBIR. Colour and texture features are important properties in CBIR systems. In this paper, a combined feature descriptor for CBIR is proposed to enhance the retrieval performance for CBIR. This method is developed by exploiting the wavelets and colour histogram moments. First, Haar wavelet is used to decompose colour images into wavelet coefficients. Then Second image feature extraction and similarity matching are performed by means of histogram moments.

Kommineni Jenni et al. [8] presented a Content Based Image Retrieval approach based on the database classification using Support Vector Machine (SVM) and color string coding feature selection. In SVM method, the feature extraction was done based on the basis of color string coding and string comparison. Here, they succeed in transferring the images retrieval problem to strings comparison. Thus the computational complexity is decreases obviously and increased the accuracy in obtaining results for image retrieval. Using database classification we can improve the performance of the content based image retrieval.

Ammar Huneiti et al. [9] proposed a CBIR method by extracting both color and texture feature vectors using the Discrete Wavelet Transform (DWT) and the Self Organizing Map (SOM) artificial neural networks. At query time texture vectors are compared using a similarity measure which is the Euclidean distance and the most similar image is retrieved. In addition, other relevant images are also retrieved using the neighborhood of the most similar image from the clustered data set via SOM. Results showed that the proposed method is able to retrieve images with higher average precision values than other methods proposed in literature by just comparing the texture similarity and without any need to compare color similarities.

S. Selvarajah and S. R. Kodithuwakku, et al [13] Representation of visual features and similarity match are important issues in CBIR. Color and texture features are important properties in CBIR systems. In this paper, a combined feature descriptor for CBIR is proposed to enhance the retrieval performance for CBIR. This method is developed by exploiting the wavelets and colour histogram moments. First, Haar wavelet is used to decompose colour images into wavelet coefficients. Then image feature extraction and similarity matching are performed by means of histogram moments.

Siddarth Ladhake et al. [6] provides a system for large scale database is designed and implemented. Here, proposed system exploits semantic binary code generation techniques, fine and coarse similarity measure technique, which improves accuracy, image retrieval speed. Finally,the performance of image retrieval is improved in terms of accuracy, retrieval time and efficiency.

Devyani Soni et al. [14] proposed an efficient color space Based Approach for Image Retrieval Using fusion of Color Histogram and color correlogram. During experimentation, both HSV color model as well as RGB color model was used for the same process of retrieval and it was observed that HSV color space gives more accurate result as compared to RGB color space.

Priyadarshini Patil et al. [2] proposed and implemented an efficient image retrieval technique using both color and texture features of an image. Here they compare and analyze performance of an image retrieval using both these features. And we see that CBIR using color features gives high precision where as CBIR using texture feature features give high recall.

III. PROPOSED OF CBIR

Content Based Image Retrieval is a technique that enables a user to extract an image based on a query from the database containing huge amount of images. Here, we have to test a query image from our own built Dataset and provides the accurate result to the user.

IV. PROPOSED WORK

In this paper, we are going to propose Content based Image Retrieval System on the colored images. Here, we built our own dataset. The Dataset is divided into two categories: Train Data and Test Data. The steps involved are:

- The Query image is given by the user corresponding to which user want the results.
- Read that particular Query image.
- Extract features from the Query image on the basis of color and relatively find the Prediction class using **K Nearest Neighbor (KNN)** classifier. Prediction class is used to find out the relevant images.
- Now to order these relevant images we use **Spearman's Rank Correlation Function** to calculate the distances of each relevant image with Query image**.** It will sort all these relevant images and fetch top-n images (top-n < size of Dataset) and print these images.

K Nearest Neighbor: K nearest neighbors is a simple algorithm that stores all available cases and classifies new cases based on a similarity measure (e.g., distance functions). KNN has been used in statistical estimation and pattern recognition.

Spearman's rank correlation coefficient or **Spearman's rho**: is a nonparametric measure of rank correlation

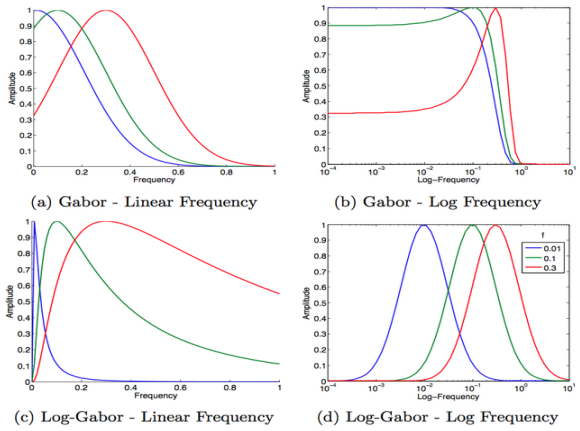
(statistical dependence between the ranking of two variables). It assesses how well the relationship between two variables can be described using a monotonic function. Mathematical notation of Spearman Rank formula is

$$(R) = 1 - \frac{6\sum d^2}{n^3 - n}$$

The result will always be between 1 and minus 1.

Here,
d= Distances
n= No. of cases

Log Gabor: Field introduced the Log-Gabor filter



(a) Gabor - Linear Frequency (b) Gabor - Log Frequency

(c) Log-Gabor - Linear Frequency (d) Log-Gabor - Log Frequency

and showed that it is able to better encode natural images compared with the original Gabor filter. Additionally, the Log-Gabor filter does not have the same DC problem as the original Gabor filter. A one dimensional Log-Gabor function has the frequency response:

$$G(f) = \exp\left(\frac{-(\log(f/f_0))^2}{2(\log(\sigma/f_0))^2}\right)$$

V. EXPERIMENT AND RESULTS SYSTEM DETAILS

A. Hardware details

We have validated our results on machine with the configuration of installed memory (RAM 3GB), 64-bit Operating System, having processor Intel(R) Core(TM) i3-2310M CPU @ 2.10GHz. Here, we have our own created Dataset e.g., images of Flags of different countries.

B. Software details

MATLAB 7.0
WINDOW 7

The Experimental work is done on the MATLAB. MATLAB is a software package for high performance numerical computation and visualization. It provides an interactive environment with hundreds of built-in functions for technical computation, graphics and animation. The name MATLAB stands for **MATrix LABoratory**. MATLAB is an efficient program for vector and matrix data processing. It contains ready functions for matrix manipulations and image visualization.

MATLAB provides a suitable environment for image processing. Although MATLAB is slower than some languages (such as C), its built in functions and syntax makes it a more versatile and faster programming environment for image processing. In this paper, proposed work is done on MATLAB as it contains ready-made functions so this tool is easy to use. We are going to compare existing and proposed approach.

A **precision rate** can be defined as the number of relevant images retrieved by a search divided by the total number of images retrieved by that search. The equation is as follows:

$$\text{Precision} = \frac{\text{Relevant Correctly Retrieved}}{\text{All Retrieved}}$$

Where A is relevant correctly retrieved and B is falsely retrieved.

A **recall rate** is defined as the number of relevant images retrieved by a search divided by the total number of existing relevant images (which should have been retrieved). The equation is as follows have been retrieved). The equation is as follows.

Table1 : Comparison of Existing and Proposed approach

Approach	Precision	Recall	F-Measure	Time complexity
Existing	0.625	0.625	0.625	1.62
Proposed	1	1	1	1.34

The validated proposed results are shown below from figure 1(a,b)

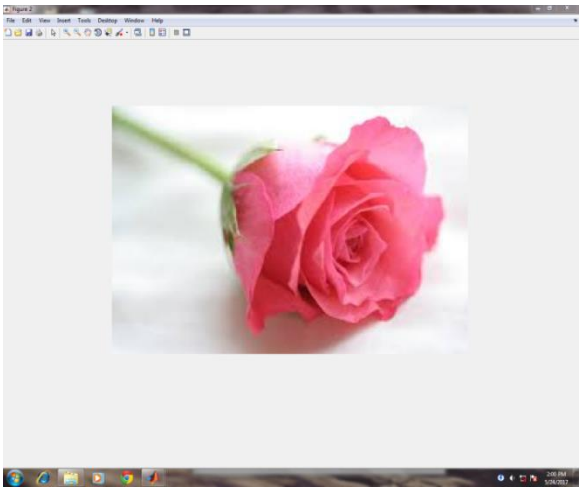


Figure 1(a): Query image of Rose to be searched

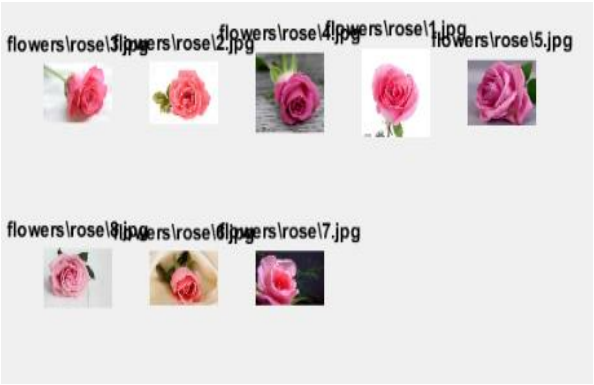


Figure 1(b): Results corresponding to the Query Image of Rose

The results of Texture Base Proposed approach is shown below in figure 2(a,b)

Figure 2(a): Query Image for Texture base feature

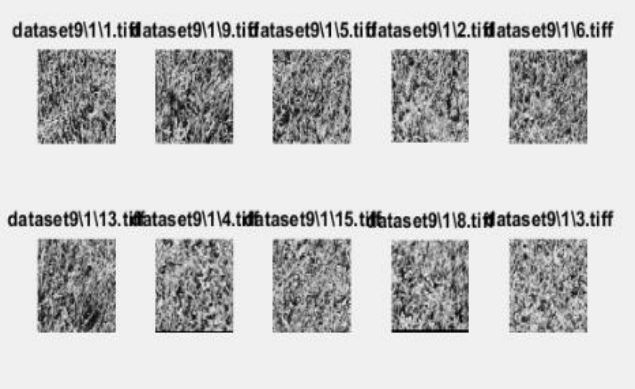
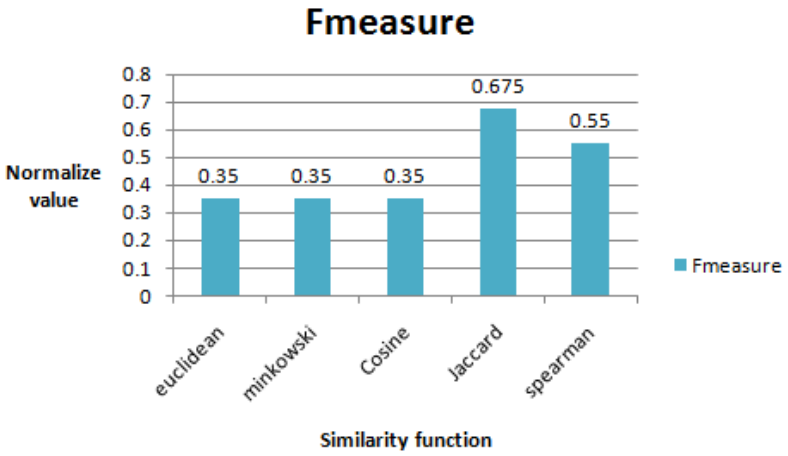


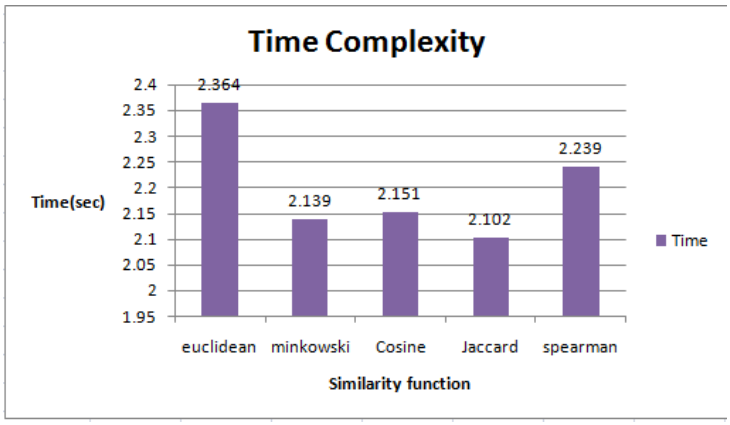
Figure 2(a): Query Image for Texture base feature

These results can also be explained with the help of graphs:

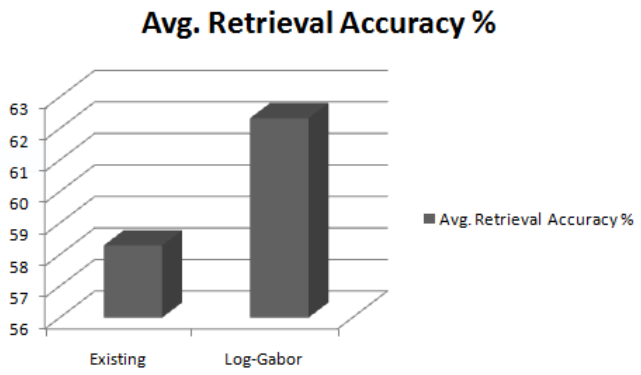
Here Graph1,2 and 3 shows the F-Measure and Time Complexity and Accuracy respectively



Graph 1: Comparison of Existing and Proposed Approaches on the basis of F-measure.



Graph 2: Comparison of Existing and Proposed Approaches in terms of Time Complexity.



Graph 3: Comparison of Existing and Proposed Approaches in terms of Accuracy

VI. CONCLUSION AND FUTURE SCOPE

From above discussion we conclude that the proposed approach perform better than existing approach with reduced time complexity and improves F-measure value. Also we use Nearest Neighbor (KNN) Classification to calculate the relevant images from Dataset and Spearman's Rank Correlation Function for calculating the distances and Log-Gabor feature of texture. Gabor filters have not zero mean, which produces a non-uniform coverage of the Fourier domain. This distortion causes fairly poor pattern retrieval accuracy. To address this issue, we propose a simple yet efficient image retrieval approach based on a novel log-Gabor filter scheme. We make emphasis on the filter design to preserve the relationship with receptive fields and take advantage of their strong orientation selectivity. As here we have own Dataset which contain limited number of images. In future, we can use this concept for the huge database as well and use some other classifiers to enhance the results.

REFRENCES

[1] Priyadarshini Patil and Bhagya Sunag, Analysis of Image Retrieval Techniques Based On Content, IEEE International Advance Computing Conference (IACC).

[2]. R. Sudhakar, K. R. Krishnan and S. Muthukrishnan, "A hybrid approach to content based image retrieval using visual features and textual queries," 2011 Third International Conference on Advanced Computing, Chennai, 2011, pp. 241-247.

[3]. L. Chai, H. Zhang, Z. Qin, J. Yu and Y. Qi, "Multi-feature content-based product image retrieval based on region of main object," 2011 8th International Conference on Information, Communications & Signal Processing, Singapore, 2011,pp.1-5.

[4] Anuja khodaskar, Siddarth Ladhake, Promising Large Scale Image Retrieval by using Intelligent Semantic Binary Code Generation Technique, International Conference on Intelligent Computing Communication & Covergence (ICCC-2014) Conference Organized by Interscience Institute of Management and Technology, Bhubaneswar, Odisha, India.

[5] Roshi Choudhary, Nikita Raina, Neeshu Chaudhary, Rashmi Chauhan, Dr. R H Goudar, An Integrated Approach to Content Based Image Retrieval, International Conference on Advances in
Computing, Communications and informatics (lCACC1) 2014

[6] Kommineni Jenni, Satria Mandala, Mohd Shahrizal Sunar, Content Based Image Retrieval Using Colour Strings Comparison, 2nd International Symposium on Big Data and Cloud Computing (ISBCC'15)

[7] Ammar Huneiti, Maisa Daoud, Content-Based Image Retrieval Using SOM and DWT, Journal of Software Engineering and Applications, 2015, 8, 51-61 Published Online February 2015

[8] Vinee. V. Kawade and Arti. V. Bang, Content Based Image Retrieval Using Interactive Genetic Algorithm, Annual IEEE India Conference (INDICON) 2014

[9]. P. B. Patil and M. B. Kokare, "Interactive content-based texture image retrieval," 2011 2nd International Conference on Computer and Communication Technology (ICCCT-2011), Allahabad,2011,pp.71-76.

[10]. S. Selvarajah and S. R. Kodithuwakku, "Combined feature descriptor for Content based Image Retrieval," 2011 6th International Conference on Industrial and Information Systems, Kandy, 2011, pp.164-168.

[11] Devyani Soni , K. J. Mathai, An Efficient Content Based Image Retrieval System based on Color Space Approach Using Color Histogram and Color Correlogram, Fifth International Conference on Communication Systems and Network Technologies 2015.

[12]. S. Selvarajah and S. R. Kodithuwakku, "Combined feature descriptor for Content based Image Retrieval," 2011 6th International Conference on Industrial and Information Systems, Kandy, 2011, pp.164-168.

[13]. M. Kaipravan and Rejiram R, "A novel CBIR system based on combination of color moment and Gabor filter," 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE), Ernakulam,2016,pp.170-174