International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2018): 7.426

Retrieval Techniques for Feature Extraction in CBIR System

Vanita M. Yadao[1], Ashish B. Kharate[2]

[1]PG Student, Department of Electronic & Telecommunication Engineering, Sant Gadge Baba Amravati University, Amravati (M.S.), India
[2]Associate Professor, Dept.of Electronics and Telecommunication Engineering, Sant Gadge Baba Amravati University, Amravati (M.S.), India

Abstract: ***CBIR (Content-Based Image Retrieval) uses the visual contents of a picture like global features-color feature, shape feature, texture feature, and local features-spatial domain present to signify and index the image. CBIR method combines global and local features. In this paper worked on Hear Discrete Wavelet Transform (HDWT) for decaying an image into horizontal, vertical and diagonal region and Gray Level Coo-ccurrence Matrix (GLCM) for feature extraction. In this paper for classification process, Support Vector Machine (SVM) used. The experimental results show improved results in comparison to previous methods. In this paper, proposed a calculation which consolidates the advantages of a few different calculations to improve the exactness and execution of recovery.***

Keywords: CBIR; Similarity Matrix; DWT; SVM; GLCM; Global feature; Local feature; Color Correlogram; Color Histogram

1. Introduction

The development in digital photography, storage limit and speed of network made possible in storing a high quality large amount of images. Applications of digital images include military, medical, virtual museums and individual photograph collections. However, users have few troubles in establishing and searching huge numbers of images in the databases, as the present commercial database systems are designed for text document and not well suitable for the digital images. Therefore, an effective way for image retrieval is desired. In image retrieval system, computer system for surfing, finding and retrieving images from a vast record of the digital images. Image retrieval is categorized into two kinds of retrieval are Content Based Image Retrieval and Text-Based Image Retrieval. Text-Based Image Retrieval is having disadvantages of effectiveness, loss of knowledge, more expensive job and also time consuming. Overcome these difficulties by applying CBIR (Content Based Image Retrieval) systems for image retrieval. The image retrieval system acts as a classifier to distribute the images in the image database into two sessions, either relevant or irrelevant

2. Proposed Methodology

The novel approach of CBIR system is based on color correlogram, color moment, color histogram, Gabor filter, SVM and GLCM.

A. Global Features

Color Histogram-Color histogram gives HSV color space and RGB color space. The matching method, then retrieves the images whose color histograms equivalent those of the query most narrowly.

Color Moment-Color moments are used to distinguish images on the basis of their color features. These moments give a dimension for color similarity between the images. These similarity values can be matched to the values of images indexed in a catalog for content based image retrieval. Color Correlogram-Color correlogram are the color feature information. The advantages of the color correlogram that contains the spatial correlation of colors can be used to describe the global sharing of local spatial correlation of colors and is simple calculating Texture feature that is the quite hard to explain, and subjected to dissimilarity of human perception. Extraction of Texture feature is computationally exhaustive, and the working speed is very important in CBIR method, as time of the response wants to be small sufficient for superior interactivity. The main aim is to present a quick and proficient extraction of texture feature technique for CBIR systems in proposed systems Fourier Descriptor (FD) and space of curvature scale descriptors are contour-based, since they are extracted from the contour, while image moments are region-based extracted from the whole shape region

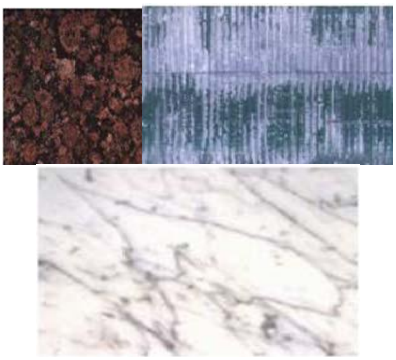


Figure 1: Texture Analysis

The process of the color quantization is that optimizes utilize different colors in an image not including affecting the image visual properties. For a true color image, the distinct amount of colors is up to 224 = 16777216 and the direct removal feature of color from the true color will direct to a huge calculation. In order to decrease the computation, the quantization of color can be used to correspond to the image, without an important image feature reduction.

Support Vector Machine (SVM)

The main element of Support Vector Machine is to create hyper planes or a collection of hyper planes with the help of support vectors in a higher dimension space. SVM used for classification. It divides the space into two half spaces.A 'good separation' is reached through hyper planes that have the major closest data distance to the points. Here decent separation means superior the division between two hyper planes gives lesser generalization error. That's by it is known as a maximum margin classifier. If geometric gap between the hyper planes more elevated than classification error is low.

Gray-level Co-occurrence Matrix (GLCM)

GLCM creates a matrix with distances and directions among the pixels, and then removes significant figures from the matrix as texture features. In this paper, worked on four features including energy, contrast, correlation homogeneity. Homogeneity- It is a grayscale image texture calculates of homogeneity varying, shiny the distribution of images grayscale regularity of weight and texture. Contrast-Contrast is the main diagonal near the instant of inertia.

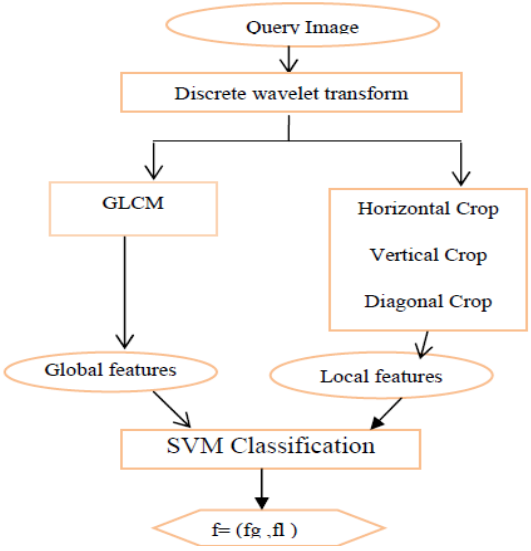


Figure 1: Block Diagram of Proposed System

3. Proposed Algorithm

1) Take RGB image.
2) Convert RGB image into HSV image.
3) Extract the Blue, Red, and Green Components from an image
4) Apply Haar Wavelet transformation at 1st level to get estimated vertical and coefficient, diagonal and detail coefficients of horizontal.
5) Apply Color histogram using assigning the 8 level ever to type, saturation and value provide a quantized the HSV space with the histogram of 8x8x8=512 bins.
6) Gray Level Co-occurrence Matrix (GLCM) computed over horizontal, vertical and diagonal submatrices.
7) Divide the image into sub-regions: horizontal (H), vertical (V) and diagonal (D) region:

$fG=(f_{GH}, f_{GV}, f_{GD})$
$f_{GH}=(f_{GHene}, f_{GHcor}, f_{GHcon}, f_{GHidm})$
$f_{GV}=(f_{GVene}, f_{GVcor}, f_{GVcon}, f_{GVidm})$
$f_{GD}=(f_{GDene}, f_{GDcor}, f_{GDcon}, f_{GDidm})$
$fL=(f_{LH}, f_{LV}, f_{LD})$
$f_{LH}=(f_{LHh\_crop}, f_{LHv\_crop}, f_{LHc\_crop})$
$f_{LV}=(f_{LVh\_crop}, f_{LVv\_crop}, f_{LVc\_crop})$
$f_{LD}=(f_{LDh\_crop}, f_{LDv\_crop}, f_{LDc\_crop})$

Where G is global feature and L is local feature, ENE is energy, cor is correlation, con is contrast and idm is homogeneity

8) Combine estimated coefficient of Red, Green, and Blue element.
9) In the same way, merge the vertical and horizontal coefficients of the Blue, Red, and Green element.
10) Allocate the weights 0.003 to estimate coefficients, 0.001 to vertical and 0.001 to horizontal coefficients.
11) Change the estimated, vertical and horizontal coefficients into HSV plane.
12) Repeat step1 to step8 on an image in database.
13) Determine the similarity matrix of the query image and image database using relative deviation, standard deviation.
14) Repeat the steps from 9 to 10 for totally images in the database.
15) Classify the images using SVM classifier and combine global and local features.

$f_{QUERY}=(fG,fL)$

1) Image Database

Figure 2: Image Database

2) Read Query Image

Figure 3: Query Image

3) Retrieved Images

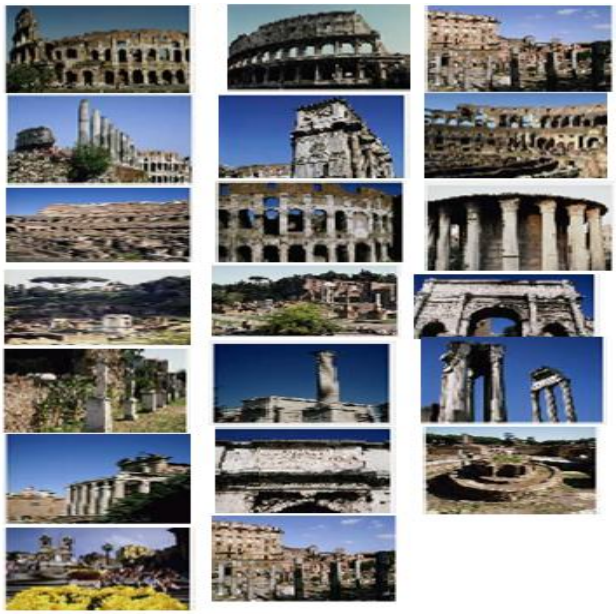


Figure 4: Retrieved Images using Relative Deviation

4. Result Analysis

Performance Evaluation

To compute local features, in our paper first segment the processed image into blocks and reach a descriptor for every block. After dividing the image into sub-regions, two statistical measures are computed for each region. These measures are mean (μ) and standard deviation (σ)

$$\mu = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} f(i,j)}{M \times N}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{M}(|f(i,j)| - \mu)^2}{M \times N}}$$

Similarity Measurement

For similarity comparison, these proposed works have to use *RSD*, Relative standard Derivation using below equation. Methods used to arrange images, moreover, compute the distinction or comparison between two vectors. In this paper used Euclidean distance which is the most predictable metric for calculating the lack of involvement between two vectors. Given two vectors *Q* and *D*, where

$$Q = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{bmatrix} and D = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}$$

Then Relative Standard Deviation flanked by them is given by

$$RSD = \sqrt{\sum_{i=1}^{n}(Q_i - D_i)^2}$$

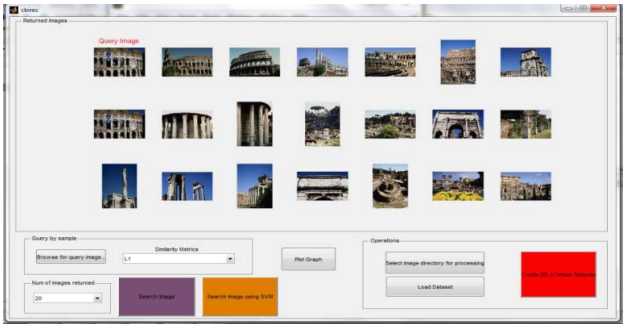
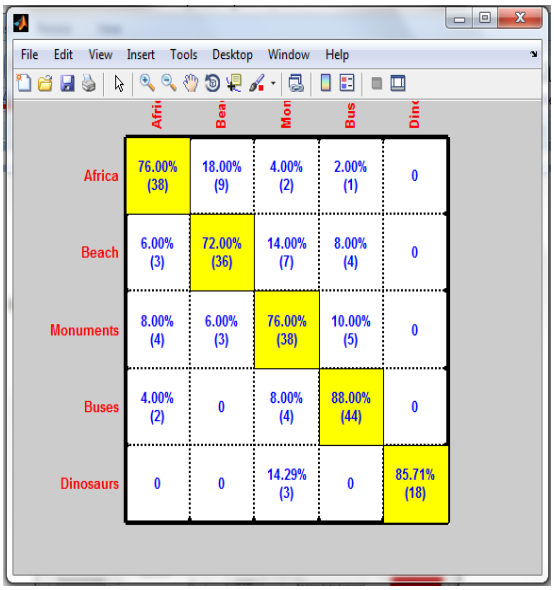
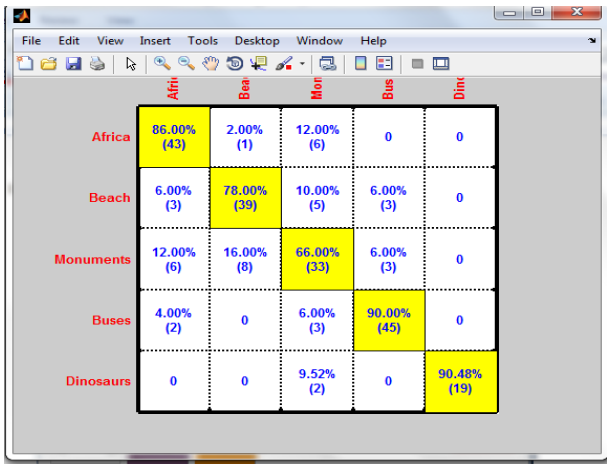


Figure 5: L1 Similarity Metrics Results on monuments images



Graph 1: Confusion Matrix Overall accuracy = 76.47%



Graph 2: Confusion Matrix Overall accuracy = 86.88%

5. Conclusion

Content based Image Retrieval System is a method to locate the related image in the image collection when the given query image. Our proposed paper used texture, color and shape feature extraction, color features are extracted using three methods such as the color Correlogram, color moment, Color histogram extracted Texture features are using gray level cooccurrence matrix. This process calculated energy, correlation, contrast and homogeneity for texture analysis. Shape features are extracted using noise removal. In this

paper, the relative standard derivation and SVM algorithm, normalized deviation, standard deviation.

References

[1] Dr. R. Usha Rani[#1],"Image Quantification Learning Technique through Content based Image Retrieval", International Journal on Future Revolution in Computer Science & Communication Engineering ISSN: 2454-4248Volume: 4 Issue: 1 203 – 206 IJFRCSCE | January 2018,

[2] V. Ramya, "Content Based Image Retrieval System using Clustering with Combined Patterns", International Journal of Scientific Research in Computer Science, Engineering and Information Technology © 2018 IJSRCSEIT | Volume 3 | Issue 1 | ISSN : 2456-3307

[3] Abdolraheem Khader Alhassan, Ali Ahmed Alfaki, "Color and Texture Fusion-Based Method for Content-Based Image Retrieval", IEEE 2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE), Khartoum, Sudan

[4] Abdolreza Rashno,Saeed Sadri, "Content-based Image Retrieval with Color and Texture Features in Neutrosophic Domain", IEEE 2017 3rd International Conference on Pattern Recognition and Image Analysis (IPRIA 2017) April 19-20, 2017

[5] Ms. Foram S. Patel , Prof. Dipali Kasat,"Hashing Based Indexing Techniques for Content Based Image Retrieval A Survey", International Conference on Innovative Mechanisms for Industry Applications (ICIMIA 2017)

[6] S. Vanitha Sivagami and K. Muneeswaran "Image Object Retrieval Using Distribution of Mixed Shape Descriptors", Middle-East Journal of Scientific Research 24 (4): 1057-1062, 2016

Author Profile

Vanita M. Yadao, is currently appeared M.E. from H.V.P.M's College Engineering & Technology, Amravati, in Sant Gadge Baba Amravati University, Amravati (M.S.) Department of Electronic & Telecommunication Engineering , Her Research interest includes Imager Rerieval, Pattern Rcognition

Ashish B. Kharate, He is currently an assistant professor in the H.V.P.M's College of Engineering and Technology, Amravati. He has working experience of more than 15 years in H.V.P.M. COET. His Research interest includes Image Retrieval, Patteern Rcognition and data Mining, Digital electronics in VLSI Techniques.