

A Hybrid Approach for Content Based Image Retrieval System

Mrs. Madhavi Kshatri¹, Prof. Yogesh Rathore²

¹ (Mtech 4 sem student, Dept. of CSE, RIT Engineering College, Raipur, C.G, INDIA)

² (Sr. Lecturer Dept. of CSE, RIT Engineering College, Raipur, C.G, INDIA)

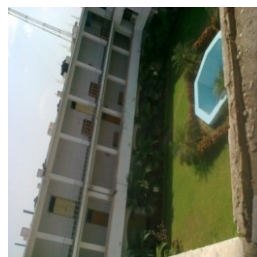
Abstract: Processing of a two-dimensional picture are generally refers to term digital image processing through a digital computer. Also say it implies digital processing of some two-dimensional data. Include methodology for query image in my project work are based on matching and retrieval & Euclidean distance are used for distance calculation. Initially for feature extraction different feature like row sum, column sum, back ward diagonal sum, diagonal sum, for ward diagonal sum are used. In next Only histogram method are used where Threshold value are used for comparison. In final step row sum, column sum, back ward diagonal sum, diagonal sum, far ward diagonal sum and histogram method are combining applied to images stored in the database. The feature vector of database images is stored as .mat file. Distance calculation between the feature vector of the images stored in the database and the feature vector of the query image is accepted.

Keywords: Row sum, Column Sum, Forward & backward diagonal sum, Histogram & Euclidean Distance.

I Introduction

1.1 Image retrieval

Vision is the most advanced of our senses, so it is not surprising that images play the single most important role in human perception [1]. Various approaches to image retrieval are someway derived from usual information retrieval and are planned to manage the new flexible and huge amount of image data which exist. Image retrieval has been one of the most interesting and vivid research areas in the field of computer vision [2]. For Example:-



Query Image



Retrieve Image

Figure 1.1 Image Retrieval

1.2 Content-based image retrieval (CBIR)

Content used to refer to something that is contained: the contents of a box or the topics or subjects covered in a documents or books. Something that is to be expressed through some medium, as writing, speech, or any of various arts, a poetic form adequate to a poetic content. Here, in the case of Content of Image "Content-based" means that the search will analyze the actual contents of the image rather than the metadata such as keywords, tags, and/or descriptions associated with the image Content-based visual information retrieval (CBVIR) also known as query by image content (QBIC) or Content-based image retrieval (CBIR). Content-based image retrieval (CBIR) is the application for problem of image retrieval. In large databases of digital images, searching problem is also corresponding problem of image retrieval.

Content-based image retrieval (CBIR) is an active yet challenging research area. The performance of a CBIR system is inherently constrained by the features adopted to represent the images in the database [3]. The similarity of image depends on the feature representation and feature dissimilarity function. However, users have difficulties in representing their information needs in queries to content-based image retrieval systems [4]. CBIR is an important alternative and complement to traditional text-based image searching and can greatly enhance the accuracy the information being returned [5].

1.4 Literature Review – Various research are done in this area. Some few literature are (1) **“Application of the Image analysis technique for textile identification** “by Robert Drobina, Mieczysław S. Machnio Technical University of Bielsko-Biala Institute of Textile Engineering and Polymer Materials ul Willowa 2,43-309 Bielsko-Biala, Poland .Published on AUTEX Research Journal, Vol. 6, No 1, March 2006 . In this paper applying image correction techniques enables a detailed identification of the structure and geometry of linear textile fabrics. Elaborating the digitisation algorithm, combined with numerical methods, allows the numerical characteristics of a textile product’s structure to be obtained. (2) **“Integrated color, texture and shape information for content-based image retrieval** “ by Ryszard S. Choras Tomasz Andrysiak,Michal Choras Published on Pattern Analysis and Application October 2007, Volume 10,Issue 4 . In this paper the main focus in this paper is on integrated color, texture and shape extraction methods for CBIR & developed original CBIR methodology that uses Gabor filtration for determining the number of regions of interest (ROIs), in which fast and effective feature extraction is perform.(3)**“A paper on automatic fabrics fault processing using image processing technique in matlab** “by R.Thilepa Department of EEE Adhiyamaan Educational & Research Institute, Hosur-tamil nadu & m.thanikachalam department of civil engineering,velammal engineering College, Chennai .Published on An International Journal (SIPIJ) Vol.1, No.2, and December 2010. The main objective of this paper is to elaborate how defective fabric parts can be processed using Matlab with image processing techniques.etc.

1.3 Proposed Method -

This Paper proposes an efficient and accurate CBIR system which is based on the combination of spatial feature like Row Sum, Column Sum, diagonal Sum and histogram. The combination of this entire feature is combined to form a feature vector of the image. Euclidean distance measure is used to compare the query image with the feature vector of the images stored in database image. The proposed system results the images from the database which are most similar to input query image. Query Image which is not present in the database is added or appended to the database for future queries.

II Overview of Feature Extraction

Content based image retrieval (CBIR) system is a software tool to search and access digital image from a large database.

The content of the image is very important and used to describe and search the image. The content of any image can be described by a low dimension data also known as feature vector. Since feature vector is a approximate representation of whole image but with reduced dimension and hence it is unique identity of the image therefore it can be used effectively for searching and indexing the image in a large database with less computational as well as less storage requirement.Color, shape, texture, histogram are the most common feature vector of the image and CBIR system based on these feature or combination of these feature are proposed in the literature.

III Equation Used For Proposed Method

3.1 Row Sum - If we take the image and calculate the sum of each row then they must be nearly same for two similar images therefore Row sum can be used as a feature vector for the image & any image can be computed using following formula-

$$\text{Row Sum (i)} = \sum_{j=1}^N I(i, j) \quad \text{-----} \quad (1)$$

Where , M = No. of Rows in an image
N=Number of Column in an image
i=1, 2, 3.....M

3.2 Column Sum- Column Sum is another way to represent the distribution of pixel values in the image therefore by computing column sum; we can determine the flow of pixel value distribution in an image.

$$\text{Column Sum (j)} = \sum_{i=1}^M I(i, j) \quad \text{----} \quad (2)$$

Where, M=No. of Rows in an Image
N=Number of Column in an image

$j=1, 2, 3, \dots, N$

3.3 Diagonal sum - For two similar images, each diagonal sum of first image is nearly equal to corresponding diagonal sum of second image. Forward diagonal sum vector represent the sum of all the forward diagonal element.

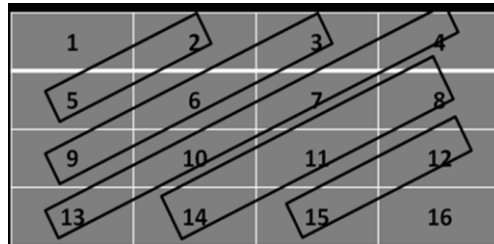


Figure 3.1 Forward Diagonal Backward diagonal sum- Backward diagonal sum vector represent the sum of all the forward diagonal element.

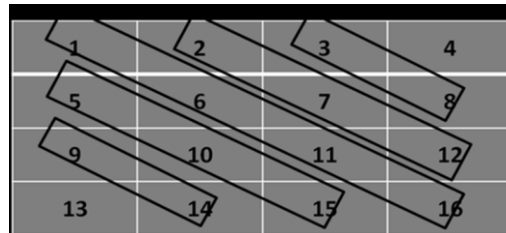


Figure 3.2 Backward Diagonal

$$\text{Forward/Backward Diagonal (k)} = \sum D_{m \times m} \quad \text{---- (3)}$$

Where, D = Diagonal element of Square image $I_{m \times m}$
 m = size of square matrix
 $k = 1, 2, \dots, N-1$

3.4 Histogram Histograms of an image represent the pixel distribution in an image and hence it is widely used as a feature vector in image retrieval system. Normalized histogram of an image is given by the following formula

$$P(r_k) = \frac{n_k}{n} \dots \dots \dots \quad \text{----- (4)}$$

Where, $r_k = k^{\text{th}}$ gray level
 n_k = number of pixel in k^{th} gray level
 n = Total number of pixel in an image

Euclidean Distance (ED) In this system Euclidean distance is used for obtaining the similarity between query images and database image the formula for computing the Euclidean distance is given by below equation -

$$ED = \sqrt{\sum_{i=1}^k (fqI_i - fdbI_i)^2} \dots \dots \dots \quad \text{----- (5)}$$

Where, fqI_i = feature vector of Query Image
 $fdbI_i$ = feature vector of Database Image
 k = No. of features

IV. Figure & Table

4.1 Precision & Recall

Precision -Recall table is used for the performance estimation of retrieval result .In the system Precision is defined as the sum of applicable image retrieved to the sum of images retrieved .If retrieved images belongs to the similar class then retrieved images are considered related image. We have created our own databases of the images taken from standard database of Caltech and found out the efficiency of different methods for the comparative analysis of all the methods employed. All image are divided in Group 1,2 etc in according to categories Table **for Precision-Recall Method** is as shown:

Table for Precision-Recall Method Precision-Recall for Row Sum, Column Sum and Diagonal Sum Method

Group ID	Precision	Recall
Group 1	0.81	0.1100
Group 2	0.51	0.1275
Group 3	0.37	0.0925
Group 4	0.75	0.1875
Group 5	0.69	0.1725

Table 4.1 Precision-Recall for Row Sum, Column Sum and Diagonal Sum Method

Precision-Recall for Histogram Method

ID	Group	Precision	Recall
Group 1		0.86	0.1015
Group 2		0.63	0.1575
Group 3		0.58	0.1450
Group 4		0.69	0.1725
Group 5		0.56	0.1400

Table 4.2 Precision-Recall for Histogram Method

Precision-Recall for Hybrid Method

Group ID	Precision	Recall
Group 1	0.88	0.1100
Group 2	0.66	0.1700
Group 3	0.52	0.1300
Group 4	0.76	0.1900
Group 5	0.64	0.1600

Table 4.3 Precision-Recall for Hybrid Method

Where, Group 1 =Groups of Animals –Example horse, elephant, dinosaur
 Group 2= Group of Flower
 Group 3= Group of Building
 Group 4 =Group of Bus
 Group 5= Group of Mountain

Graph for average Precision Curve for all the CBIR method

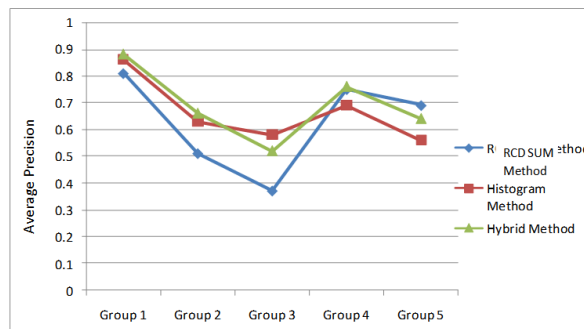


Figure 5.1 Average Precision graph for all the CBIR method

Graph for average Precision Curve for all the CBIR method

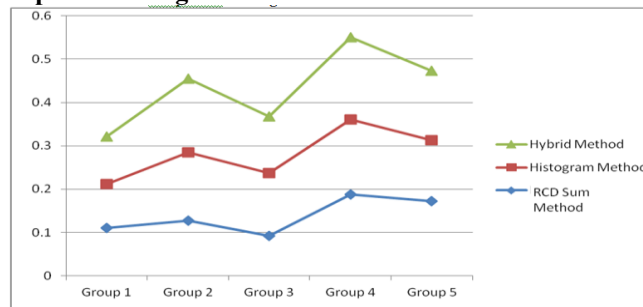


Figure 4.2 Average Recall Graph for all the CBIR method

Result -Result obtained from CBIR system using Row Sum, Column Sum, Backward and forward Diagonal Sum –

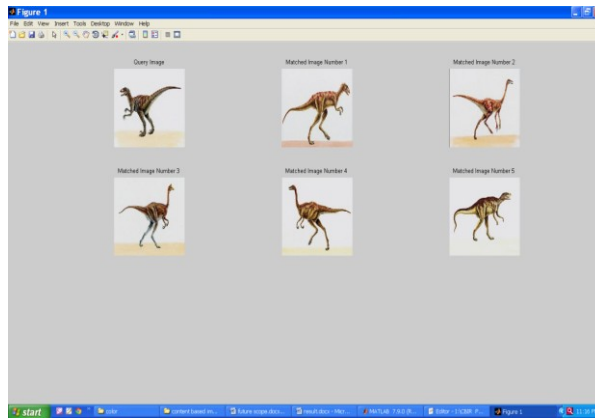


Figure 4.3 Query Image and image retrieved using Row sum, Column Sum, Backward and forward diagonal Sum

Result obtained from CBIR system using Histogram feature-

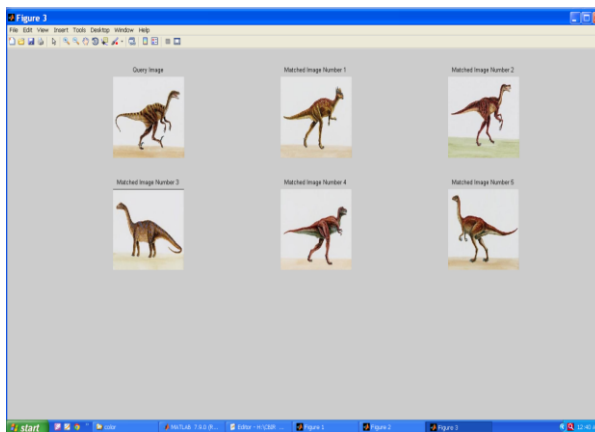


Figure 4.4 Query Image and image retrieved using Histogram feature

Result obtained from CBIR system using Hybrid Method-

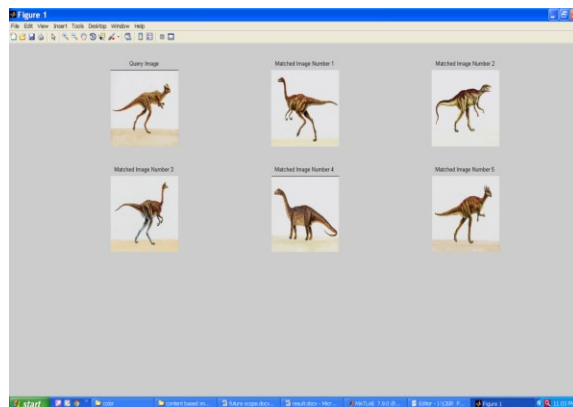


Figure 4.5 Query Image and image retrieved using Hybrid Method

V. Conclusion

Content Based image retrieval (CBIR) is the area where searching is done using image content. Image retrieval is one of important third area for researchers from image processing and information technology domain. The desire of better and faster image retrieval is still not satisfied. Show the table

Precision & Recall value is best in Hybrid method & compared to individual two methods. Hybrid method is providing best result.

The precision-Recall obtained by the proposed CBIR is 80%.

References

Journal Papers:

- [1]. Vijay Bagdi, Sulabha Patil Professor, R.V. Dharaskar “Analysis Of Various Cbir Techniques For Retriving Forensically Viable Information From The Images”, International Journal of Engineering Research & Technology (IJERT), Vol. 2 Issue 1, January-2013 ISSN: 2278-0181.
- [2]. Xiang Sean Zhou,Thomas S. Huang, “Edge-Based Structural Features for Content-Based Image Retrieval”, Appeared in Pattern Recognition Letters, April 2001
- [3]. Christopher C. Yang , “Content-based image retrieval: a comparison between query by example and image browsing map approaches”,Journal of Information Science, page no .254-267 ,June 2004,30 .
- [4]. Ch.Kavitha ,Dr.B.Prabhakara Rao & Dr.A.Govardhan ,“Image Retrieval Based On Color and Texture Features of the Image Sub-blocks”,International Journal of Computer Applications -(0975 - 8887) Volume 15- No.7, February 2011
- [5]. Dr.H.B.Kekre ,Sudeep Thepade, Priyadarshini Mukherjee, Miti Kakaiya, Shobhit Wadhwa, Satyajit Singh ,“Edge Texture Based CBIR using Row Mean of Transformed Column Gradient Image” ,Journal of Computer Applications (0975 - 8887) Volume 7- No.10, October 2010

Book:

- [6]. Rafael C.Gonzalez & Richard E.Woods, Digital Image Processing - A Remote Sensing Perspective, Second Edition, Pearson Prentice-Hall Education, 2006.