A SPATIAL FEATURE BASED EFFICIENT CBIR SYSTEM FOR COLOR IMAGES

Miss. Laxmi Goswami
Asst. Professor
Dr.C.V. Raman University
Bilaspur (C.G.)

Miss. Amrita Singh
Asst. Professor
Dr. C.V. Raman University
Bilaspur (C.G.)

Aijaz U.R. Khan
Asst. Professor
M.M. College of Technology
Raipur (C.G.)

Abstract—Row sum, column sum and histogram are some of the important spatial feature of any image and represent the pixel distribution of any image. Row sum, column sum and histogram feature of two similar images are nearly same and hence if these feature are combined then a very efficient CBIR system can be designed. This paper present a row sum, column sum and histogram based hybrid approach for designing a CBIR system. Different statistical parameter were computed to show the accuracy of this system.

Keywords—Row sum, Column sum, Histogram, feature vector, Euclidean Distance.

I. INTRODUCTION

Content based image retrieval system can be defined as the software tool which is used to search and access digital images in large database [1]. Feature vector of the image are basically low dimensional data which represent the content of the image[2]. Some of the most widely used features of the image are color, texture, shape, mean, entropy, standard deviation etc. Since these feature has low dimension and contain all the necessary information about the image therefore these feature can be used for image searching and indexing in a large database of image. Due to low dimensionality, these feature require less storage and less computation power. In the past various CBIR system were proposed which are based on color feature [3][4]. Some author also suggested an CBIR system based on computing color histogram[5][6]. J. Berens, G. D. Finlayson and G. Qiu presented a CBIR system based on color histogram in compressed transform domain[7]. Some authors [8][9] used color mean or color average for designing CBIR system. In some images shape of the object[10-12] is also one of the important feature and has been used for CBIR system. Texture is also important feature and hence have been used for designing CBIR system[13]. Some frequency domain approaches have also been proposed like discrete cosine transform based [14] and wavelet transform based approach[15]. This paper describe a very simple and accurate image retrieval system which is based on computing the row sum and column sum of R, G, B channel of color image separately along with the histogram of Red channel, Blue Channel and Green Channel. Since Row sum, column sum and histogram of each channel in color image contain very crucial information about the image therefore an efficient CBIR system can be designed using these features.
II. PROPOSED CBIR ALGORITHM

A. Image database Creation

In any CBIR system design, the first part is extracting the features of the image. In this algorithm, row sum, column sum and histogram are used as the spatial features. Since in this approach, coral database has been taken whose image dimension is 384x256 so in order to reduce the computational complexity, all the images are first resized to 256x256.

The steps for spatial feature extraction are as follows:
1. Input the image.
2. Resize the image to dimension 256x256.
3. Extract the Red channel, Green channel and blue channel.
4. Extract the row and column for each channel and store it in a matrix.
5. Calculate the Row sum and column sum for each channel.
6. Prepare the feature database for each image.

Block diagram of database creation is shown in Figure 1.

Since in this algorithm, an image of size 256x256 has been taken therefore there are 256 row sum and 256 column sum is obtained for each channel and represented RSR(row sum for red channel), CSR(column sum for red channel), RSG(row sum for green channel), CSG(column sum for green channel), RSB(row sum for blue channel) and CSB(column sum for blue channel) by and the dimension of histogram for each channel is also 256 and is represented by HR, HG and HB for red, green and blue channel respectively. Feature vector for each image has a dimension 9x256 while the dimension of the image is 256x256. Image feature database for this algorithm is shown in Figure 2.
A. Image Retrieval System

Once the creation of feature database for all the images are carried out then the next step is to design a CBIR system. The block diagram of proposed method is shown in Figure 2. Illustration of Image Feature database

B. Metric used for Similarity Measurement

Similarity measurement has a very important role in any image retrieval system. Euclidian Distance is most widely used parameter for measuring the similarity between query image and database images. Euclidean Distance between features of query image and database image is computed using equation given below

\[ E_d(q,d_b) = \sqrt{\sum_{i=1}^{k} (F_{q_i} - F_{d_b_i})^2} \]

- \( F_{q_i} \) = Feature vector of Query Image
- \( F_{d_b_i} \) = Feature vector of Database Image
- \( k \) = No. of features
IV. PROPER THRESHOLD SELECTION

It is very important for any image retrieval system that a proper threshold value $T$ of Euclidean distance must be set. In order to set the proper threshold value $T$, maximum ED distance between database and query image has been taken as a reference point to calculate the value of threshold $T$. It was found through extensive experimentation that taking 40% of maximum ED distance produce better result than other value and hence this value is chosen for threshold $T$.

V. EXPERIMENTAL RESULT

The proposed algorithm was implemented in a system having configuration core 2 duo processor, 2GB RAM. MATLAB (Ver 7.0) platform is used for programming. A database having as many as 800 images of different categories has been prepared from coral database. Images are divided into 8 different categories as shown in table I.

<table>
<thead>
<tr>
<th>Group ID</th>
<th>Image category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bus</td>
</tr>
<tr>
<td>2</td>
<td>Dinosaur</td>
</tr>
<tr>
<td>3</td>
<td>elephant</td>
</tr>
<tr>
<td>4</td>
<td>Tribal</td>
</tr>
<tr>
<td>5</td>
<td>Beaches</td>
</tr>
<tr>
<td>6</td>
<td>Flower</td>
</tr>
<tr>
<td>7</td>
<td>Horse</td>
</tr>
<tr>
<td>8</td>
<td>Food</td>
</tr>
</tbody>
</table>

All the images are first converted into dimension of 256x256 then feature vectors of all the images have been computed and stored in feature vector matrix as described in section II. In the database, 100 images for each category have been taken. A GUI has been designed for this method.

Figure 3 Image Retrieval Process
which take the query image as input and gives five closest
match images as output.

![Precision Vs Recall Curve for proposed CBIR](image1)

![Average Precision Curve for proposed CBIR](image2)

Precision and recall are the two parameters which have been
used for testing the performance of the proposed system.
Average precision has been computed for different image
categories and a graph is plotted between average precision
and image categories and shown in Figure 5. Moreover
precision and recall have also been computed and a plot
between precision and recall is drawn and shown in Figure
4.

\[
\text{Precision} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}}
\]

\[
\text{Recall} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of relevant images in database}}
\]

The proposed CBIR system starts by inputting the query
image. Then Euclidean distance between feature vectors of
database image is computed. An image is considered as the right match for the
query image or belong to the same category of the query image if the ED distance between query image and
considered image is less than the preset threshold T.

VI. CONCLUSION

Basically, the main criterion for judging the performance of
any CBIR system is its efficiency and speed. From the
result obtained by extensive experimentation and the
computed statistical parameter, it is clear that the CBIR
system proposed in this paper is able to retrieve the right
image efficiently and accurately from the database of large
image. It has also been observed that this system retrieve
the right image in less than a 3 second from the database of
800 images which clearly shows that it is very fast in
speed.
Figure 6 Some output of the proposed CBIR system

References


