Scientific Journal of Impact Factor (SJIF): 4.14

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 3, Issue 5, May -2016

A MULTIMODALIMAGE DESCRIPTORS SYSTEMFOR IMAGERETRIEVALUSING 3D HISTOGRAM AND ORB

Andanagouda Patil[1], Dr Suresh H N[2]

[1]*M.Tech, Signal Processing, Dept. of EIE, BIT, Bengaluru-560004,*
2 Professor & Coordinator for PG Studies and Research,Dept. Of Electronics & Instrumentation, BIT, Bengaluru-560004

Abstract:- *A prominent research area of retrieving images from a huge database with the help of query image without using any tag words is a Content Based Image Retrieval (CBIR). Various approaches have been proposed in this area. Since the use of global features many systems will fail those cannot extract the important features in the image. On having a think process on this, we came up with a new approach using different kinds of descriptors which can extract as many as features as possible. Which intern gives us the better results. Hence the multimodal feature descriptors are used here in developing a CBIR system. We made use of 3D Histogram Technique along with the corner detection with ORB. The results shows that the accuracy is more compared to the other papers.*

I. Introduction

Managing large image data base is a tedious job. Retrieving images from that huge data base is more tedious one. The cutting edge technology that solves this problem is the "content based image retrieval".Content based image retrieval is based on concept.Images are retrieved by indexing the images in many image search enginesi.e. based on the metadata such as keywords, tags, description, and captions. In the data base along with the image, human must generate and store metadata. Manually, labelling each image for large data base is a tedious job. The images retrieved on using metadata will not be accurate because the manual annotation may vary for different images. This type of search requires meaningful description, text labels, and keywords.

To overcome the above issue of retrieving images from the large database the image retrieving techniqueis used that is based on content of the image. The CBIR is an application of computer vision techniques to the image retrieval problem. This is also known as query by image content (QBIC). Content based means, the content of the image is analysed during the search. The content here refers to the

Colours, shapes, textures, or any other information that can be derived from the image.
The content based image search uses visual contents of an image such as, colour shape, texture to represent and index the image. User will give an input example image or a sample sketch to the system as a query. The system extracts the features of the example image or sketch. The similarities between the feature vectors of the example image and thefeature vectors of the images in the data base are then calculated and retrieval is performed with indexing scheme. The relevance feedback is taken from the user to modify the retrieval process to provide meaningful retrieval results. Image content may include visual content and the semantic content. The visual content may be general or domain specific. The general visual content include, shape, colour, texture etc. The domain specific visual content includes human faces, domain knowledge is also required in domain specific. The semantic content is obtained by textual annotation based on visual content.
The visual content is obtained by dividing an image into parts. The simple way to divide an image is by using partition that cuts image into tiles of equal size and shape. A more complex way of dividing an image is to undertake object segmentation completely to obtain semantically meaningful objects. The general CBIR system is as shown in figure 1.

A basic CBIR system follows the following steps in order to retrieve images from the data base [2].

1. Input query image or a sample sketch.
2. Detect features of query image and the images in the database.
3. Matching the features.
4. Result presentation
5. Feedback

II. Related Work

In order to extract the images from the large database, extracting the features of the images play a vital role.

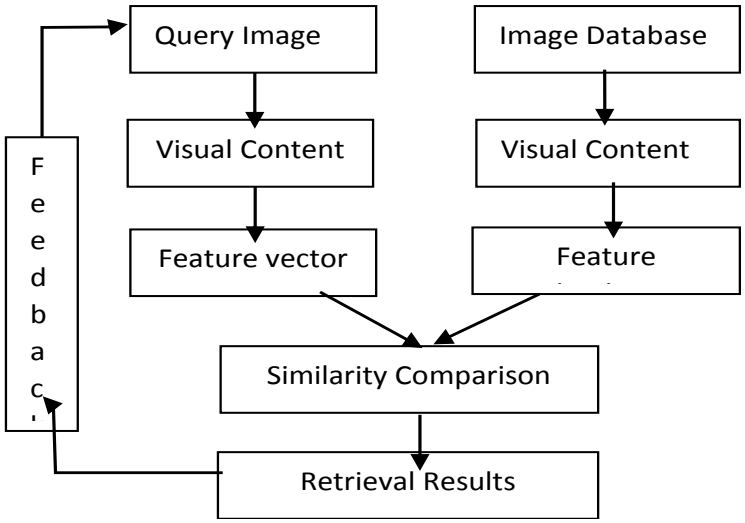

Fig 1. : General CBIR System

According to the survey carried on the brief description of the variant feature extraction techniques is discussed here which are efficient and helpful in real-time retrieval of the images. Features in an image are Colour features, Texture Features, Shape Features, and Spatial Features[1][2].

Features are nothing but the output of image descriptor. Feature vectors are the list of numbers which are used to represent the image and quantify. Colour components in the image are the main features in detecting the image. The chromatic part of an image is considered and usually the image is described in RGB colour model (Red, Green, and Blue). Using the histogram technique with the fixed bin size the colour feature vectors are stored[3]. Histogram is a colour feature vector that describes the total number of times each colour occurred in the colour image. The other colour models used in this technique are YCbCr, HSV, Etc...[4]. In this paper I have used 3D HSV colour model because it is well known to be unchanged in uniformity and produces far better results[5].

Detecting the corners is also a very good feature which holds more information in an image. Among many other descriptorsthe most reliable, efficient and simple descriptor is Scale invariant feature transform (SIFT) which has received most attention. The two main components of the SIFT are the key point detector and key pointdescriptor. The former one will detects the key points which are invariant to transformation in an image and the later one will describe where the region of that key point is appeared[6][7].

When it comes to the real time implantation and while detecting the key points the SIFT is not fast enough to compute the key points. The other main disadvantage of SIFT is, it is patented. The alternative of this is Oriented FAST and rotated BRIEF (ORB) which is fusion of FAST algorithm for corner detection and BRIEF(Binary Robust Independent Elementary Features)[8]

In [9][10] for extracting the colour feature the colour moments is used and the local binary pattern is used to extract the texture features using grayscale images. On combining both the features a single feature vector is formed. Further in [11] colour and texture features are used on irregular region of interest.

Once the image feature descriptors are found, the next step in CBIR is to match the descriptors between query image and the images in the data base and retrieve the similar images from the database. Various distance functions like, Normalized Euclidean,Euclidean, Minkowski,Mahalanobis, Chebyshev,Canberra,functions etc. are used to calculate the distance between two features and to find the similarities. In [4][10] it is said that the results from Manhattan distance are close to human perception and is best compared to earlier ones. The most similar images are ranked and accordingly the images are displayed.

Along with the above descriptors there are many other descriptors through which the image features can be extracted[5][12]–[14]. Retrieving machine with only one variant of features are not efficient. Hence in this paper we make use of a multimodal

feature set which extracts the feature descriptors using multiple feature sets. These are stored as the vectors and will be used while matching the feature vectors. So far we understood briefly what is CBIR and the various feature sets and the main motto of this paper. We further discuss about our multimodal feature sets in III. Method proposed. In IV.Experimental Results, we show the experimental results by comparing the results with other research works. In V. Future work and Conclusion, about the future research work on multimodal feature extraction with conclusion is explained.

III. Proposed Method

As said earlier, the two most important descriptors are extracting colour histogram and corner detection using ORB. In our proposed multimodal method, we make use of these two features of an image to retrieve images from database effectively. The use of colour histogram is a simple method but an effective descriptor. In most of the cases images with similar colour description are considered to be relevant to each other. The ORB came from OpenCV labs which is a best alternative for SIFT inmatching performance and computation cost.

Our proposed modal will have four steps to extract the images from the database.

1. Define multimodal image descriptor.
2. Extract features from the database.
3. Define type of searcher.
4. Extract the features from the query image.
5. Perform the search.

*1. Defining Multimodal image descriptor:*Among the multimodal descriptors the 1st modal used is a 3D colour histogram based on HSV colour space (Hue, Saturation and Value). HSV over RGB is because the HSV model is good at mimicking as human perceive colour. Since we have selected the colour space, to find the histogram of the image the main request is the number of bins to be used. The use of bins mainly depends upon the number size of the database. More the images in the database the more number of bins gives the better results. It also depends on the similarity between colour distributions in the database images. We made use of 8 bins of Hue, 12 bins for Saturation and 3 for Value. With knowing all the prerequisites, practically the image is always in a RGB model so our first step is to convert the image into HSV model. Our next step is to calculate the histogram for the image. But in our proposed model the 3D histogram is not calculated for whole image. Instead, the image is divided in to five regions and the histogram is calculated for each region separately. The regions are divided as top right corner, top left corner, bottom right corner, bottom left corner, and centre of the image. Which is called as the region based feature descriptor.

The other part of the multimodal descriptor is detecting the corners with the use of ORB model. This model finds the key points using the FAST and then the Harris corner algorithm is applied which detects the top Npoints among them. It also computes the orientation. Intensity weighted centroid of patch is computed at the centre with located corner. The orientation is given from the direction of the vector from corner point to centroid. It uses the BRIEF and rBRIEF for computing the descriptors. Here also the corners are detected for each region separately.

2. Extract features from the database: Extracting the features from the database and storing them is basically called as the "Indexing". On using the first step the feature vectors are stored in the .CSV file. It will have the file name followed by the numbers which are feature vectors. The CSV file will have the multimodal feature vectors which includes the key points and also the histogram values for future comparison.

3. Define type of searcher: Since we have got the feature vectors, our next step is to compare the features to find the similarities. The similarities are found between the query image and the images in the database image. In order to do that openCV provides the Chi-squared distance function. This will divide the squared difference of each individual bin by first histograms bin count. But a slight verification for this, instead of using the first histogram bin count, the sum of bin count values is used to divide the squared difference of every bint count.

4. Extract the features from the query image: Using the second step the multimodal features descriptors are calculated for the query image as well.

5. Perform the search: The last part of the CBIR is performing the search from the feature vectors. The feature vectors of the database in the index file and the feature vectors of the query image are matched and the best top ten results are displayed.

IV. Experimental Results

In this paper the experiment is carried on to examine the performance and efficiency of the system that has been proposed using the WANG database. This database with 1000 images with 10 clusters is one of the benchmark database for the content based image retrieval systems[11]. According to the literature survey the main parameter that are calculated in the CBIR is average precision. The various parameters like average precision, recall, MTBF, Sensitivity, specificity etc. shows the performance of the proposed method. Those are given by,

1. Average Precision:

$$Precision = \frac{No.of\ relevent\ images\ retrieved}{Total\ number\ of\ images\ retrieved}$$

2. Sensitivity/Recall:

$$\text{Sensitivity} = \frac{No.of\ relevent\ images\ retrieved}{Total\ no.of\ relevant\ images\ in\ DB}$$

The values obtained during the evaluation of the performance of the proposed work for multimodal feature set is shown in tables 1 and 2 respectively.

3. MTFB(Mean time between Failure)

$$\text{MTFB} = \frac{No.of\ times\ the\ system\ checked}{Number\ of\ failures}$$

The mean time between the failures should be more in the system to justify the good performance of the system. The failure of the system has never occurred which can be said that the mean time between the failure is more.

4. Accuracy: Accuracy is proportional to the correct results retrieved among the total number of images retrieved.

$$\text{Accuracy} = \frac{No.of\ relevant\ images\ retrieved}{Total\ no.of\ images\ retrieved} * 100$$

Parameter	CEDD	ROI	CCV	**Proposed Multimodal Method**
Precision %	74.50%	74%	75%	**89%**

Table 1: Average Precision value comparison between proposed method and the various other methods

As shown above in the precision value comparison table proposed model holds the more effective performance compared to the other methods. Our method gives almost accurate results.

Parameter	**Proposed Multimodal Method**	CEDD	CCV	ROI
Accuracy	**83.5%**	75%	78%	83%

Table 2: Accuracy comparison between proposed method and the various other methods

The above figure shows the accuracy between the proposed model and the other model which clearly shows that our model is more accurate compared to the other systems.

Parameter	CEDD	ROI	CCV	**Proposed Multimodal Method**
Recall	0.15	0.63	0.44	**0.80**

Table 3: Recall comparison between proposed method and the various other methods

The Recall varies a lot because of the limit that is assigned at retrieving only the top images compared rom the database. Even though the database has similar 100 images but only top 10 or 20 images are retrieved.

IV. Conclusion and Future Work

Extracting features from image requires an efficient algorithm which should be more a generic one. Application specific search engines are easy to develop based on the content available with respect to the application. Thus to make a system a more generic we developed a multimodal feature set using 3D Colour Histogram and Corner based ORB detector which gave a better results compared to the other models. Adding more features to the multimodal is our future work. The other descriptors are shape, texture etc..

REFERENCES

[1] P. D. Pardhi, P. L. Paikrao, and D. S. Chaudhari, "Introduction to Query Techniques for Large CBIR Systems," no. 1, pp. 181–184, 2012.
[2] R. Chaudhari and A. M. Patil, "Content Based Image Retrieval Using Color and Shape Features," vol. 1, no. 5, pp. 386–392, 2012.
[3] "A COLOR IMAGE RETRIEVAL SYSTEM BASED ON WEIGHTED AVERAGE University of Electronic Science and Technology of China."
[4] N. Mariam, "A Modified Approach in CBIR Based on Combined Edge Detection , Color and Discrete Wavelet Transform," pp. 2201–2205, 2015.
[5] P. D. Scholar, "A Novel approach to Self Order Feature Reweighting in CBIR to reduce Semantic Gap using Relevance Feedback," pp. 1437–1442, 2014.
[6] R. Bandaru, "Retrieve the Similar Matching Images Using Reduced SIFT with CED Algorithm," 2014.
[7] N. Puviarasan, R. Bhavani, and A. Vasanthi, "Image Retrieval Using Combination of Texture and Shape Features," vol. 3, no. 3, pp. 5873–5877, 2014.
[8] A. Mordvintsev, "OpenCV-Python Tutorials Documentation," 2014.
[9] R. Choudhary, N. Raina, N. Chaudhary, R. Chauhan, and R. H. Goudar, "An Integrated Approach to Content Based Image Retrieval," pp. 2404–2410, 2014.
[10] F. Malik and B. Baharudin, "Analysis of distance metrics in content-based image retrieval using statistical quantized histogram texture features in the DCT domain," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 25, no. 2, pp. 207–218, 2013.
[11] Y. Velazco-paredes, R. Flores-quispe, and R. E. Pati, "Region-Based Image Retrieval Using Color and Texture Features on Irregular Regions of Interest," 2015.
[12] P. Rachana, S. Ranjitha, and H. N. Suresh, "A Stride Towards Developing an CBIR System Based on Image Annotations and Extensive Multimodal Feature Set," no. 1, 2014.
[13] G. F. Ahmed and R. Barskar, "A Study on Different Image Retrieval Techniques in Image Processing," no. 4, pp. 247–251, 2011.
[14] R. C. Veltkamp, M. Tanase, and D. Sent, "Chapter 5 FEATURES IN CONTENT-BASED IMAGE RETRIEVAL SYSTEMS : A SURVEY," pp. 97–124.

