ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

Enhancing the Performance of Image Retrieval through Similarity Measure

Ritu A. Mundada[1], Akash D. Waghmare[2]

M.E Student, Department of Computer Science & Engineering, SSBT'S COET, Bambhori, Jalgaon, India[1]

Assistant Professor, Department of Computer Science & Engineering, SSBT'S COET, Bambhori, Jalgaon, India[2]

ABSTRACT: Now a days, digital world increase with hand held devices such as mobile , storage technologies, many social networking sites and large volume of images stored on web. With significantly large number image database, it is difficult to mine data and retrieve relevant images from database. Feature Based Image Retrieval is a very important and dominant topic in the image retrieval mechanism. It retrieve the image using low level feature, such as color, texture, shape and similarity measures for the comparison of images. The problem in usual approach is to extract certain number of relevant features from an image that reduces dimensionality, yet preserving useful information. In the proposed system, Hue, Saturation, Value space histogram are used for color information extraction. Gabor filter is used for texture feature extraction. Gabor texture descriptors are adopted for texture feature in image, Mean, median and standard deviation are calculated using the laplacian filter. Color moments and texture features based on the global features . The performance of the method improved using the mean, median, standard deviation. Experimental results show that the proposed method has a higher retrieval accuracy than other conventional methods.

KEYWORDS: Content Based Image Retrieval, Feature Extraction, Laplacian Filter etc…

I. INTRODUCTION

Content Based Image Retrieval (CBIR) is the retrieval of images based on low level feature called visual feature such as color, texture and shape. Development is required because in many large image databases, traditional methods of image indexing have proven to be insufficient, laborious, and extremely time consuming. These old methods of image indexing, ranging from storing an image in the database and associating it with a keyword or number, to associating it with a categorized description, have become obsolete. In content based image retrieval, each is stored in the database and its features is extracted. Now compared this features with the query image. CBIR is method that search, browse and retrieve the image using the actual contents of the image like visual features of an image such as color, shape, texture and spatial layout instead of that using the keyword or any description related related to the image. Query By Image Content (QBIC) and Content Based Visual Information Retrieval (CBVIR) is another name of content based image retrieval. Content based image retrieval is efficient and effective because most webs based image search engines rely totally on metadata and this produces a lot of garbage and noise in the results while retrieving the image. Also having human manually enters keywords for images in a large database can be inefficient, expensive and may not capture every keyword that describes the image. Thus a system that can filter images on their content would provide better indexing and return more accurate results[7].For each image present in the image database, its features are extracted and the obtained feature vector is stored in the feature database. When a query image is enter, its feature space is compared with those in the feature database one by one and the similar images with the smallest feature distance is retrieved. The image retrieval method is important because german say that, "picture say much more than the thousand and millions of word. Query to the database is of many types such as Query-by-text, Query-by-Sketch, Query-by-Example. It involves two steps:1.Feature Extraction, 2.Matching.

Section II represent the work related to the image retrieval. Further Section III describes the proposed system that is used to retrieve the image from the database. Section IV presents the experimental result that obtain after the testing. Discussion presented in section V. After which reference is followed.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

A. **Feature Extraction**

When the input is given to the algorithm is very large for processing and it is redundant the input data. A reduced representation set of feature is obtained after transformation. Transforming the input data into the set of features is called feature extraction. Proposed CBIR system is based on direct wavelet decomposition of image. RGB color space are used in it and utilizes the query by example method which is query to the database. database images are decomposed into multi-level coefficients ranges from -1 to -J levels, and develop color feature database and perform similarity match between images using the similarity measure. By using the special processing of decomposition, each resulting sub image is a coefficient matrix, large coefficients with more energy can be distributed in the up-left area, therefore using f-norm theory the dimension of the image is decreases and highly efficient image matching is done. Feature extraction is most important method or level in image retrieval.

B**. Color Feature**

Color similarities among images i.e. it supports querying image with respect to color is presented in content based image retrieval system. Every image has the important feature which is color. Color information is represented as each pixel in an image has a three-dimensional color vector and different color space approaches. The RGB image is converted into HSV plane. Hue is used to distinguish color, Saturation gives a measure of the percentage of white light added to a pure color and Value is represent the perceived light intensity. The advantage of HSV is that each of its attributes corresponds directly to the basic color concept, make it conceptually simple. The image is divided into a block of any size. Each image in the database is computed to obtain the color histogram, which shows the proportion of the pixels of each color within the image. In database that color histogram of each image is then stored. When the user does the search in database by specifying the query image, the system registers the proportion of each color of the query image and goes through all images in the database to find those whose color histograms match those of the query most closely.

C**. Texture Feature**

Texture is a second low level feature which is difficult to describe as well as subjected to the difference of human perception. Texture extraction by using the segmentation process is difficult because it is enable to extract the whole or entire texture feature of the image except than element. Using histogram texture measure is computed as it does not carry information regarding to the relative position of pixel with respect to each other. Texture has significantly efficient information than color histograms and corresponds to human perception. The problem is that it is very sensitive to transforms such as scaling illumination and view angle. The number of texture related features is only 32 whereas for color and luminance total 1200 features this is second problem with texture. So for requiring extract texture from image which gives the information of relative position of pixel Gray Level Co-occurrence matrix(GLCM)is used.

D. **Shape Feature**

The third and last important low level feature of image is shape in image retrieval system.So an object, can form by a set of shape which is most similar to objects have a high correlation in the set of shapes. Using segmentation process image is retrieve by using the feature shape and classify the shape, where each shape should have their own representation and should variant to scaling, rotation, and transition. In shape-based image retrieval the user need to choose a reference image or sketch a desired shape. Using this method match the exact shape and similar shape with respect to query image is identify.

Two types of feature descriptor are available such as Scale Invariant Feature Descriptor and Speeded-Up Robust Feature Descriptor. This descriptors are used to detect the neighborhood of pixels near to the key points .key point is an one type of image feature which describe a lot of information about the point it need to describe. The problem is that it is fail to describe the high level semantic feature .indoor retrieval SIFT is used. Many thousand point

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

extracted from images and distance computation is done which is quite costly. It is also called the point detector method.

II. RELATED WORK

Image shape matching is prime concern in object recognition and identification methods. An image matching is a means of determining the resemblance of one image with the other image. Images are matched based on their shape and texture and it finds variety of applications ranging from image retrieval, object recognition, remote sensing, image classification, image analysis and so on. In general, image matching techniques are classified into structure- based [1] [2] and feature-based [3][4] methods. Structure-based methods compare the shape/ structure and the size of the images, whereas the feature- based methods examine the image features like color and texture in addition to size and shape. Therefore, the image shape and size are the most essential component in automatic image matching systems. Moreover, the image comparison can not be done directly on image shapes, it require an image registration process to align the image within the same coordinate space.

Kalavathi, in [12], proposed image comparison method based on Fourier Mellin transformation which accurately compare two images and computes the overlapping similarity and hausdorff distance. For some images, the proposed method has failed to produce accurate result when the spatial difference between the images is high. This may be avoided by modifying the Fourier Mellin (FM) transformation registration. The proposed method has efficiently compared the given images and has produced accurate comparison result. The drawback of this method is that for some images it fails to register the images correctly when the difference in spatial coordinates is high.

Huang, in [13], proposed water filling algorithm applied on the edge map of the original image. The purpose of this algorithm is to efficiently extract information embedded in the edges. The new features are more generally applicable than texture or shape features. Algorithm is to extract features from the edge map directly without edge linking or shape representation. The idea is to look for measures for the edge length and edge structure and complexity by a very efficient graph traversal algorithm. Water filling algorithm is represent edge/structural information embedded in the edge maps, such as edge length, edge connectivity and complexity, loop structure, etc., which are proven to be effective in CBIR, even with simple Sobel flters as the edge detector. The proposed edge/structural features are more general than texture or shape features in that they require neither uniform texture region nor a closed shape contour. They can be effective on non-uniform real-world images and natural scenes. Since water filling features are extracted from the binary edge map, edge detection algorithm plays an important role in these features. Also since edge extraction process is a lossy transformation in terms of information contents, false positive during retrieval is sometimes rather high, which indicates the integration with other features is necessary for large databases.

Chary, in [14], proposed a Feature extraction method for color image similarity. Retrieval of images within a large image collection based on color projections and different mathematical approaches are introduced and applied for retrieval of images. Images are sub grouping using threshold values and RGB color combinations considered for retrieval of images. Proposed system uses combinations of color feature to overcome the problem description.

Elbakry, in [17], proposed Bag-of-Visual Word (BoVW) a visual feature descriptor that can be used successfully in Content-based Image Retrieval (CBIR) applications. In this, an image retrieval system that uses local feature descriptors and BoVW model to retrieve efficiently and accurately similar images from standard databases. The proposed system uses SIFT and SURF techniques as local descriptors to produce image signatures that are invariant to rotation and scale. The most important step in the proposed system is to extract the local descriptors from the processed image and key points are extracted from the image.The system combines the robust techniques, such as SIFT, SURF, and BoVW, to enhance the retrieval process. In the system, we used a k-means algorithm to cluster the feature descriptors in order build a visual vocabulary. But they have many limitation when dealing with
the broad content of image.

Shanmugapriya, in[15], proposed method uses three approaches to retrieve the relevant images from the database. Images can be retrieved based on Color, Texture, both Color and texture respectively. The proposed method uses

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

algorithms such as auto color correlogram to retrieve color based images, Gaussian mixture models to retrieve texture based images and Query point movement for relevance feedback. The experimental results conforms that the proposed method gives maximum accuracy when compared to existing work. This method lacks when the structure of object is similar between each other. system can be redesigned to accept semantic in addition to content based queries. To obtain integrated system texture features must be derived from other algorithms.

Shirazi, in[16], proposed feature based algorithm with combination of color, shape, texture and region. The systems based on color feature usually retrieve images having similar color. The problem in texture based retrieval is that they produce in-accurate retrieval results. Approaches which are relying on a single specific algorithm like color, texture or shape, such type of approaches can work successfully on specific images but when varied types of images are taken into account their performance is degraded.

Babber, in [11], proposed the matching algorithm which is based on the image feature point. By searching correct feature point and setting bidirectional threshold value, the matching process can be quickly and precisely implemented with optimistic result. The resemblance of two images is defined as the overall similarity between two families of image features[1]. Same proportion image matching algorithm using bi-directional threshold image matching technique is used. Small window of pixels in a reference image (template) is compared with equally sized windows of pixels in other (target) images. In FBM, instead of matching all pixels in an image, only selected points with certain features are to be matched. Area based matching provide low speed. feature based matching algorithm is faster in comparison to the area based matching technique. feature based matching time complexity depend on number of feature to be selected as well as right or wrong threshold. If the number of feature are high then sometimes it takes more computational time in comparison to area based feature. The number of features extracted from an image depends largely on the contents of an image. If there are high variations then features computed are high. This reduces time efficiency to match.

Although these systems have been successful in the past, their widespread use has exposed some of their limitations such as the problems of low level image feature, high level image feature, semantic gap, accuracy in the form of precision, recall and so on. Content-Based Image retrieval(CBIR), low level features of images are extracted such as color, texture and shape. These features are used in similarity measurement to retrieve relevant images from an image database. Improvement in accuracy of image retrieval system has been addressed in. By adding much more image feature like mean, median, standard deviation and so on. Improve the accuracy that has been discussed.

III. PROPOSED SOLUSION

Mean and standard deviation are statistical measurements computed in histogram. Laplacian filter is used for the noise removal. Initially image come with the much more noise. With this noise image cannot be retrieve properly. So Laplacian filter are applied first to the image when enter.
Algorithm consist if the following steps:

1. Input query image.
2. The input RGB image is acquired and converted into grayscale image.
3. Apply Laplacian filter to grayscale image.
4. Quantize the filtered image into 32 bins.
5. Calculate the mean and standard deviation of pixels in each bin.
6. mean[i][j]=(Math.abs(W.mat[r][c]))/((img1.getRows())*(img1.getCols())).
7. Standard Deviation σ= std[i][j]=Math.sqrt((Math.pow(((Math.abs(W.mat[r][c]))-(mean[i][j])),2).
8. σ / ((img1.getRows())*(img1.getCols()).
9. median= m[middle].
10. if not median:(m[middle-1]+m[middle])/2.0.
11. Calculate similarity measure using laplacian formula.
12. tans=(Math.sqrt(dismean))+(Math.sqrt(disstd)).

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

13. tans=(Math.sqrt(dismean))+(Math.sqrt(dismedia)).
14. tans=(Math.sqrt(dismedian))+(Math.sqrt(disstd)).
13 A query image will be acquired from the user as an example to retrieve similar images from the database by using the extracted features.
14 Sort the array in ascending order. The most similar images are then displayed.

After applying these 14 steps we extract more relevant image with respect to query image.

IV. EXPERIMENTAL RESULT

The database used in the evaluation is MARKOV dataset . It consists of 500 images, a subset of the Corel database, which have been manually selected to be a database of 5 classes of 100 images each. The images are of size 120 - 120 pixels. The database was extensively used to test many CBIR systems because the size of the database and the availability of class information allows for performance evaluation. A database containing no of images with any one of the formats .bmp, .jpg ,.jpeg is required. The simulation environment used in proposed system is Java NetBeans IDE 7.4. The input for proposed system is 500 images in Dataset.

The performance of system was measured using precision and recall measures. Recall measures the ability of the system to retrieve all the images that are relevant while precision measures the ability of the system to retrieve only the images that are relevant as shown in Figure 1. Precision range is between 0 to 1. It is observed that the precision value calculated for proposed system is higher than the existing system.

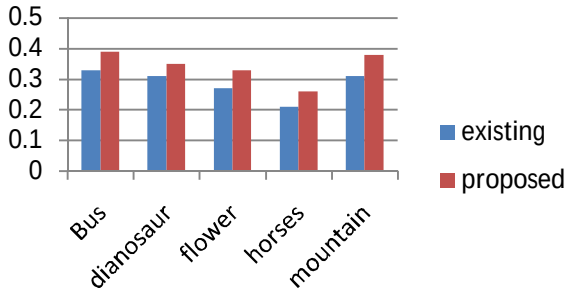

Figure 1: Performance comparison of precision.

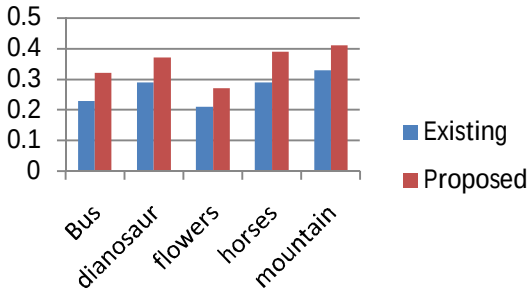

Figure 2: Performance comparison of recall.

The Figure 3, Figure 4, Figure 5, Figure 6 and Figure 7 represents the comparison of precision for sample images of Bus, Dinosaurs, Flowers, Horses and Mountain respectively. By taking the average precision of ten images of existing and proposed system against the number of images increases in all categories shows in The Figure 3, Figure 4, Figure 5, Figure 6 and Figure 7. As the number of image increases proposed system retrieve more relevant images than the existing system shows the Figure 3, Figure 4, Figure 5, Figure 6 and Figure 7.

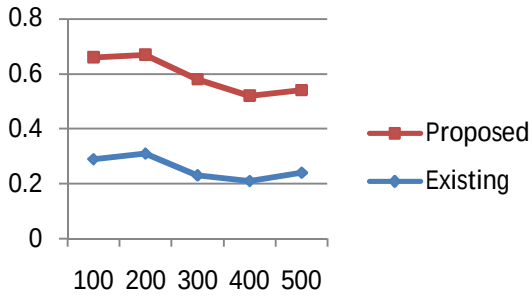

Figure 3: Performance comparison of precision for bus.

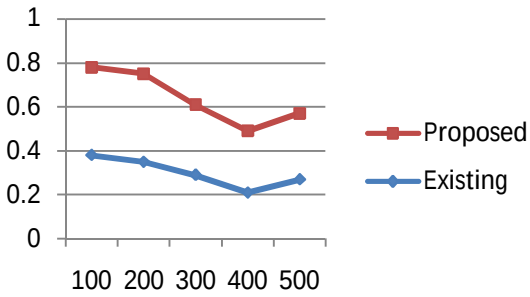

Figure 4: Performance comparison of precision for dianosaurs.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

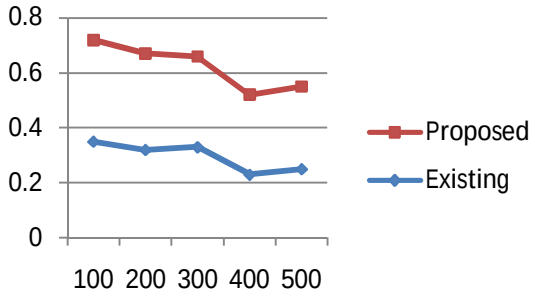

Figure 5: Performance comparison of precision for flowers.

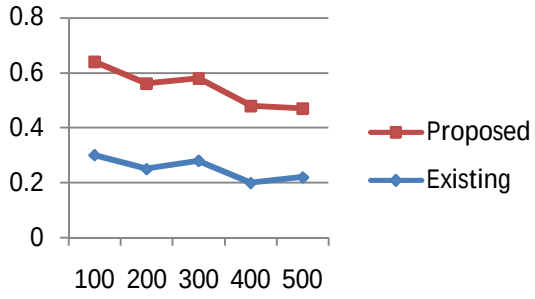

Figure 6: Performance comparison of precision for horses.

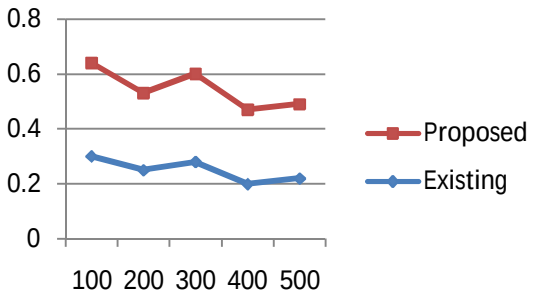

Figure 7: Performance comparison of precision for mountains.

The Figure 8, Figure 9, Figure 10, Figure 11 and Figure 12 represents the comparison of recall for sample images of Bus, Dinosaurs, Flowers, Horses and Mountain respectively. Each figure shows the categories of images against average recall. By taking the average recall of ten images of existing and proposed system against the number of images increases in all categories shows in Figure 8, Figure 9, Figure 10, Figure 11 and Figure 12 respectively. As the number of image increases proposed system retrieve more relevant images than the existing system shows in Figure 8, Figure 9, Figure 10, Figure 11andFigure12respectively.

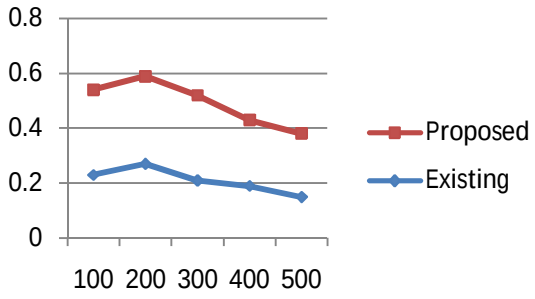

Figure 8: Performance comparison of Recall for bus.

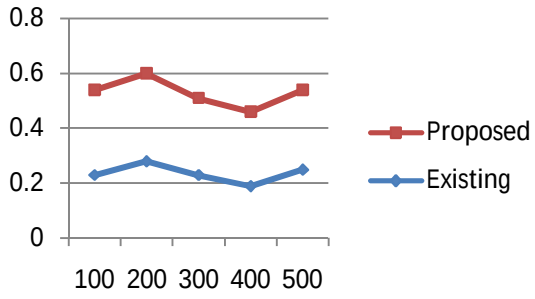

Figure 9: Performance comparison of Recall for dinosaurs.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

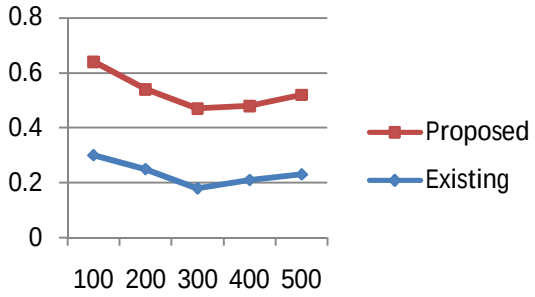

Figure 10: Performance comparison of Recall for flowers.

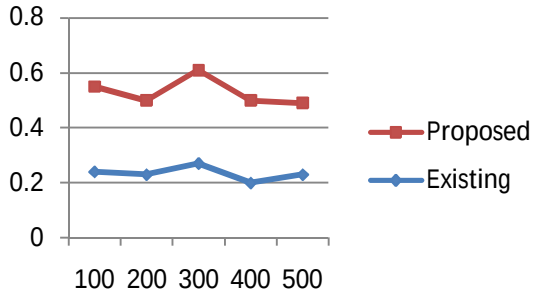

Figure 11: Performance comparison of Recall for horses.

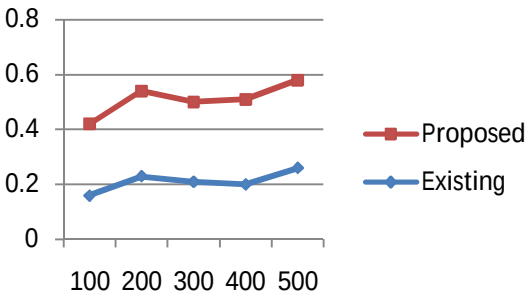

Figure 12: Performances comparisons of Recall for mountain.

V. DISCUSSION

The existing method compares only the structural similarities of the images but not compare the image features. Text based image retrieval method provide lesser and unsatisfactory result. In some cases, clustering techniques are used which are enable to give proper accuracy as number of images increases. As number of images increases, it need to be re-luster again. Hence system fails to gives the accuracy. So the existing system is not accurate. The proposed image retrieval system, compare the images using mathematical approaches like mean, median and standard deviation. After taking the combination of three mean, median, standard deviation gives the more better accuracy. Mean represents brightness and standard deviation represents contrast of image. By taking the average of ten images of existing system and proposed system for comparing the precision for sample images of bus, dinosaurs, flowers, horses and mountain as shown in Figure 3, Figure 4, Figure 5, Figure 6 and Figure 7. The large number of relevant image retrieved from large dataset which indicated using the red upper line as shown in Figure 3 to Figure 12. By taking the average of ten images of existing system and proposed system for comparing the recall for sample images of bus, dinosaurs, flowers, horses and mountain as shown in Figure 8, Figure 9, Figure 10, Figure 11 and Figure 12 respectively. The Figure 1 shows the average number of precision against categories of all images. The recall of proposed system as well as existing system of image retrieval by taking the average of the ten images is as shown in Figure 2. The Figure 2 shows the average number of recall against categories of all images. Proposed system improves the recall and precision of images better than the existing system.

VI. CONCLUSION AND FUTURE WORK

The proposed system is based on texture feature and Laplacian filter to a get more accurate image. The histogram is applied to filtered image to get featured in the image. The mean, median and standard deviation of the pixels in bins of the histogram image are calculated to get feature vector. Algorithm tested with various queries and good results are obtained. The combination of mean, median and standard deviation features are giving accurate result. The result

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

contains precision and recall more appropriate results similar to the query image. The existing system provides 30% - 40% accuracy while proposed system provides 40%-70% accuracy.
In future, the work can be extended further to improve the image saliency in the context of the other image called the co-saliency that explore the local structure changes e.g. human pose, appearance change.

REFERENCES

1. T. Kato, "database architecture for content-based image retrieval", In Proceedings of the International Society for Optical Engineering, vol.1662, pp.112-113, 1992.
2. Arnold W.M. Smeulders., Marcel Worring, and Ramesh Jain , "content based image retrieval at the end of early years,'' IEEE Transaction on pattern analysis and machine Intelligence , VOL. 22, NO. 12. year 2000.
3. Ying Liua, Dengsheng Zhanga, Guojun Lua,Wei, Ying Mab, "A survey of content-based image retrieval with high-level semantics", Vol.40, pp.262 282, 2007.
4. Brendan Klare, "Sketch to Photo Matching: A Feature-based Approach",IEEE International Conference on Computer Vision,Vol.3, pp.44-49,February 2009.
5. Suresh Pabboju, "A Novel Approach for Content-Based Image Indexing and retrieval System using Global and Region Features '',International Journal of Computer ScienceandNetwork security,Vol.9,No.3,pp.687-694,2009.
6. Parulpreet singh, "An Approach of Image Retrieval Using Content Based Retrieval System '',International Journal of Advanced Research in Computer Science and Soft-ware Engineering,Vol.3,February 2013.
7. M.Malathi, "Image Retrieval and Image Categorization by Content based Information" ,International Conference on Electronics, Communication and Information Systems,Vol.12,2012.
8. A. M. Patil, "Content Based Image Retrieval Using Color and Shape Features" ,International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 1, Issue 5, November 2012.
9. Shankar M. Patil,"Content Based Image Retrieval Using Color, Texture Shape", International Journal of Computer Science Engineering Technology,Vol.3, Issue 9, Sept.2012.
10. Prashan Premaratne, "New structural similarity measure for image comparison", Emerging Intelligent Computing Technology and Applications, pp.292-297, 2012.
11. Geetanjali Babbar, "Development of an Improved Feature based Algorithm for Image Matching",International Journal of Computer Applications, Vol.14, No.8, February 2011.
12. kalavathi, "Image Shape Matching Based on FM Transformation Technique", An Indian Journal of applied research, Vol.4, Issue 10, pp.2249-555, October 2014.
13 Thomas S. Huang, "edge-Based Structural Features for Content-Based Image Retrieval", IEEE Tran on Circuits and Systems for Video Technology, Vol.8, April 2001.
14 R.Venkata Ramana Chary, "Feature Extraction Methods For Color Image Similarity", Advanced Computing an International Journal, Vol.3, No.2, March 2012.
15 Shanmugapriya, "A New Content Based Image Retrieval System Using Gmm And Relevance Feedback", Journal of Computer Science, Vol.2, pp.330-340, Nov. 2014.
16 Syed Hamad Shirazi, "Content-Based Image Retrieval Using Texture Color Shape and Region", International Journal of Advanced Computer Science and Applications, Vol.7, pp.418-426, 2016.
17 Hazem Elbakry, "Content-Based Image Retrieval using Local Features Descriptors and Bag-of-Visual Words",International Journal of Advanced Computer Science and Applications, Vol.6, pp.212-219, No.9,2015.