RESEARCH ARTICLE   **OPEN ACCESS**

Content Based Satellite Cloud Image Retrieval Using Texture Features

D.Chandraprakas., M.Narayana,

[1]*ECE Research Scholar, Jjtuniversity. Rajastan.*
[2]*Professor JPNEC, Mahaboonnar Telangana.*

ABSTRACT
This paper proposes, the retrieving satellite cloud images from satellite cloud database images using the texture features. The traditional image retrieval technique is inefficient in retrieving these images. Content-based image retrieval is an approach from data mining community which provides the solution of managing this huge quantity of data. In this research, a Content-Based Image Retrieval (CBIR) system has been developed using texture features from the satellite image repository. The system allows the user to search for an image on the basis of any texture features. The texture features are extracted using gray level co-occurrence matrix method and Euclidean distance metric is used to compute the similarity between the images. The system is robust as it provides search based on the multiple features and it can also retrieve the images in which are rotated with an angle 'Ɵ'. The performance of the system was evaluated by analyzing the retrieval results using precision
Keywords: Image retrieval, texture feature , similarity, Euclidean distance, Content-based image retrieval.

I. INTRODUCTION
1.1 motivation and overview
With the rapidly growth of the Internet, and the easily available image capturing devices in market such as digital cameras & scanners, there is increasing more numbers of availability of digital images on Internet. So the efficient image searching, browsing and retrieval tools are required by users from various domains, including remote sensing, fashion, crime prevention, publishing, medicine, architecture, etc. For this purpose, many image retrieval systems have been developed. That's why, now a days researchers paying their attention on a very interesting system which is nothing but the Content-based image retrieval technique. In CBIR technique, image retrieval is based on similarities in their lower level features, which are textures, colours, shapes, etc. CBIR is the computer vision application to the difficulty in finding of images from large databases. Content Based Image Retrieval (CBIR) is techniques that allow us to access the digital images from a large collection of image databases by using the image features. The increasing number of digital images requires new methods to access. The images can be retrieved using the image features color, texture and shape. Image retrieval approaches are based on the computation of the similarity between the users query and images. Figure 1 shows the Architecture of CBIR system, each image is stored in the large image database and its features are extracted, compared to the features of the query image. In on-line image retrieval, the user gives a query image to the CBIR system to search the required images. The system exhibit a feature vector for the query image. Then similarities between the feature vectors of the query image and the images in the feature database are compared. In last, the system returns the images that are most similar to the query images. The challenge in CBIR is to develop the methods that will reduce the retrieval time and increase the retrieval accuracy. Evaluation of retrieval performance is a crucial problem in content-based image retrieval (CBIR) [1]. The most common evaluation measures used in CBIR are precision and recall.
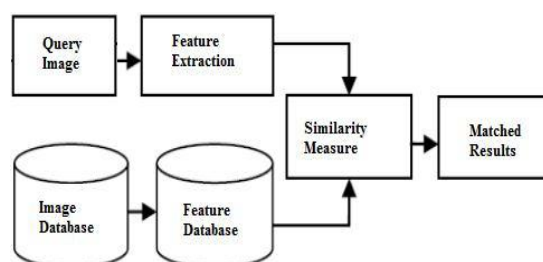

Figure1. Architecture of CBIR system

II. REVIEW OF LITERATURE
2.1 Image retrieval
Satellite images have become a common component of our daily life either on the internet, in car driving and even in our hand-held mobile handsets. There is new image and video content appearing every second through multiple competing television and internet channels, Manual interaction with this large volume of data is becoming more and more inappropriate, which creates an urgent need for automatic treatment to store, organize and retrieve this content. [2] The traditional methods for

retrieving images from geo-databases are: geographic location, date of acquisition and spectral /spatial properties of acquisition devices [3].Our needs from the satellite scenes are specific contents. Therefore we need to retrieve images that contain our intended contents. There are other challenges in the field of satellite images itself [4].These images are geo-referenced images; this means that all images form in reality a huge continuous image covering the entire earth surface. It is not always proper to deal with such content as isolated images. Retrieval of data means to get desired data from the database and to reduce the semantic gap. It may image, text , audio or video as per requirement of user. The basic types of retrievals are mentioned in Figure 2.
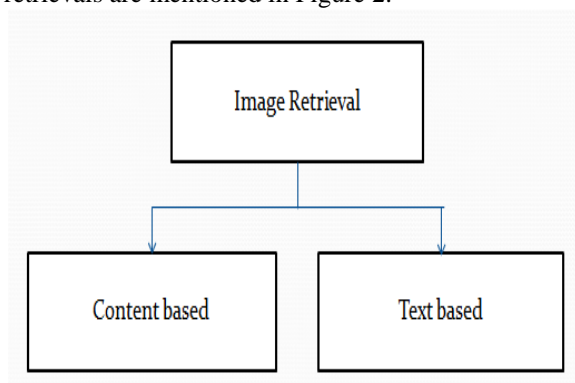

Figure 2: types of image retrievals

2.2 Text based image retrieval
Image retrieval is a solution to search for an image in a large collection of images.the traditional approach of image retrieval is the text based approach .this approach relies on the textual descriptions of the images. Each and every image in the database is to be annotated with the textual keywords and the image search is based on these keywords. The limitations of the text based approach of image retrieval are lack of consistency of the text to describe the image contents and manual annotations of the images [5].

2.3 content based image retrieval
The term Content-based image retrieval was coined in 1992 by T. Kato to describe experiments into automatic retrieval of images from a database, based on the colors, texture and shapes. Since then, this term has been used to describe the process of retrieving desired images from a large collection on the basis of syntactical image features. The technique used three main technologies of: pattern recognition, signal processing, and computer vision, [6].

In content-based image retrieval (CBIR), the image databases are indexed with descriptors derived from the visual content of the images. Most of the CBIR systems are concerned with approximate queries where the aim is to find images visually similar to a specified target image. In most cases the aim of CBIR systems is to replicate human perception of image similarity as well as possible, [7].

2.4 Feature Extraction
It is the process where features such as shape, texture, color, etc. are used to describe the content of the image. The features further can be classified as low-level and high-level features. In this stage visual information is extracts from the image and saves them as features vectors in a features database .For each pixel, the image description is found in the form of feature value (or a set of value called a feature vector) by using the feature extraction .These feature vectors are used to compare the query with the other images and retrieval, [8]. The color histogram several other color feature representation like color moments and color sets have been applied [10]. Ruietal proposed a modified Fourier descriptor which is robust to noise and invariant to geometric transformation [11].

2.5 Similarity Matching
It is a process that entails the information about each image is stored in its feature vectors for computation process and these feature vectors are matched with the feature vectors of query image (the image to be search in the image database whether the same image is present or not or how many are similar kind images are exist or not) which helps in measuring the similarity. This step involves the matching of the above stated features to yield a result that is visually similar with the use of similarity measure method called as Distance method. There are various distances methods available such as Euclidean distance, City Block Distance, and Canberra Distance, [9].

2.6 Resultant Retrieved images
It is the process that searches the previously maintained information to find the matched images from database. The output will be the similar images having same or very closest features as that of the query image, [8].

III. TEXTURE FEATURE EXTRACTION
We all know about the term texture but for defining it is a hard time. One can differentiate the two different textures by recognizing the similarities and differences. Commonly there are three ways for the usage of the textures: based on the textures the images can be segmented to differentiate between already segmented regions or to classify them. We can reproduce textures by producing the descriptions. The texture can be analyzed in three different ways. They are spectral, structural and statistical:

3.1 STATISTICAL:

Normally the textures may be random but with the consistent properties. such textures can be described by the statistical properties. Moment of intensity places a major role in describing texture in a region. Suppose in a region we construct the histogram of the intensities then the moment of the 1-D (one dimension) histogram can be computed. The mean intensity which we have discussed is the first moment. The variance describes how similar the intensities are within the region then this variance is second central moment Skew describes the symmetry of the intensity distribution about the mean then this skew is the third central moment [13]. Kirtosis describes the flatness of the distribution then this kirtosis is the fourth central moment.

IV. METHODOLOGY

4.1 Texture Feature Extraction Based On GLCM:

The following figure shows how gray co matrix calculates the first three values in a GLCM. In the output GLCM, element (1,1) contains the value 1 because there is only one instance in the input image where two horizontally adjacent pixels has a values 1 and 1, respectively. GLCM (1,2) contains the value 2 because there are two instance where two horizontally adjacent pixels have the values 1 and 2. Element (1,3) in the GLCM has the value 0 because there are no instances of two horizontally adjacent pixels with the values 1 and 3. Gray co matrix continuous processing input image, scanning the image for other pixel pairs (i j) and recording the sums in the corresponding elements of the GLCM.
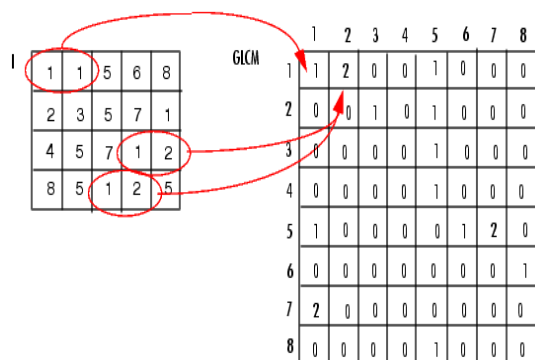

**Figure 3 :** gray level co-occurrence matrix

4.2 Process used to create the GLCM:

Specifying the offsets: By default, a single GLCM with a spatial relationship, or offset, define as two horizontally adjacent pixels created by the gray co matrix function .Gray co matrix can create a multiple GLCMS per single image because a single horizontal offset might not be sensitive to texture with a vertical orientation to create multiple GLCMS

an array of offsets to the gray co matrix function is specified. These offsets define pixel relationships of varying distance and direction [15]. For example, you can define an array of offsets that specify four directions (horizontal, vertical, and two diagonals) and four distances in this case, the input image is represented by 16 GLCMS. When you calculate statistics from these GLCMS, you can take the average.

These offsets can be specified as a p-by-2 array of integers. Each row in the array is a two element vector. And it is represented in the form of the [Row_offset,column_offset]. Row_offset specifies the number of rows between the pixel of interest and its neighbor. colum_offset specifies the number of columns between the pixel of interest and its neighbour.

Offsets = [01;02;03;04;……
-11;-22;-33;-44;…..
-10;-20;-30;-40;……
-1-1;-2-2;-3-3;-4-4];

Spatial relationships of pixels
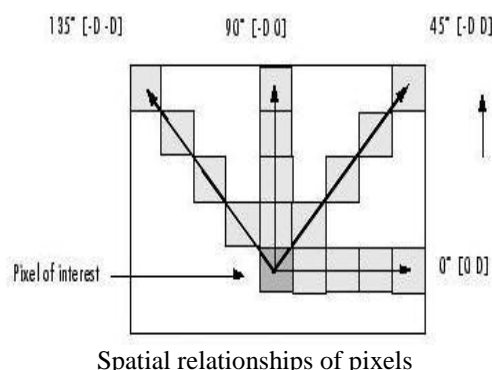
4.3 Deriving statistics from a GLCM:

After treating the GLCMS gray co props function may be used to derive several statistics from them. These derived statistics gives information about the texture of an image. By calling the gray props function you can specify the statistics you want. The following table illustrates the statistics you have been derived. Then the Table 1 is as follows:

Statistics	Description
Homogeneity	It measures the closeness of the distribution of elements between the GLCM and GLCM diagnol.
Energy	It is known as the uniformity or the angular second moment .it provides the sum of the squares of the elements of the GLCM.

Contrast	In the gray level co-occurrence matrix it measures the local variations.

4.4 Determining the pixel values in the range filtered output image:

The stdfilt and entropyfilt functions operate similarly, defining a neighborhood around the pixel of interest and calculating the statistic for the neighborhood to determine the pixel value in the output image. The stdfilt function calculates the standard deviation of all the values in the neighborhood.

The entropyfilt function calculates the entropy of the neighborhood and assigns that value to the output pixel. Note that, by default, the entropyfilt function defines a 9-by-9 neighborhood around the pixel of interest. To calculate the entropy of an entire image, use the entropy function.
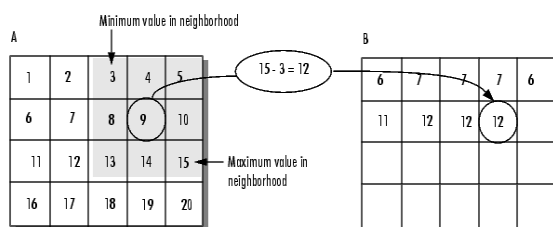

Figure 4: Texture filter function

The following Figure 4 illustrates how the texture filter functions can detect regions of texture in an image. In the figure, the background is smooth; there is very little variation in the gray-level values. In the foreground, the surface contours of the coins exhibit more texture. In this image, foreground pixels have more variability and thus higher range values. Range filtering makes the edges and contours of the coins more visible.

GLCM creates a matrix with the directions and the distance between the pixels andextracts meaningful statistics from the matrix as texture features. GLCM texture features commonly shown in the following.

GLCM is composed of the probability value, it is defined as $p(i,j|d,\theta)$ which expresses the probability of the couple pixels at ѳ direction and the d interval. When θ and d are determined $p(i,j|d,\theta)$ is shown by $p_{i,j}$. Distinctly GLCM is a symmetry matrix its level is determined by the image gray level. Elements in the matrix are computed by the equation shown as the following.

$$P(i,j|d,\theta) = \frac{P(i,j|d,\theta)}{\sum_i \sum_j P(i,j|d,\theta)}$$

(1)

GLCM expresses the texture feature according the correlation of the couple pixels gray level at different positions. It quantificationally describes the texture features. But here mainly four things are considered they are energy, contrast, entropy and the inverse difference

Energy: $E=\sum\sum p(i,j)^2$ (2)

It is a gray scale image texture measure of the homogeneity changing reflecting the distribution of the image gray-scale uniformity of the image and the texture

Contrast: $I=\sum\sum(i-j)^2 p(i,j)$ (3)

Contrast is the main diagonal near the moment of inertia, which measures the value of the matrix is distributed and images of local changes in the number, reflecting the image clarity and the texture of the shadow depth if the contrast is large then the texture is deeper.

Entropy: $S=\sum\sum p(i,j)\log p(i,j)$ (4)

Entropy measures image texture randomness, when the space co-occurrence matrix for all values are equal, it achieved the minimum value; on the other hand, if the value of co-occurrence matrix is very uneven, its value is greater. Therefore, the maximum entropy implied by the image gray distribution is random

Inverse difference: $H=\sum\sum 1/1 + (i-j)2 \ p(i,j)$ (5)

It measures local changes in image texture number. Its value in large is illustrated that image texture between the different regions of the lack of change and partial very evenly.

Here p(x, y) is the gray level value at the co-ordinate (x, y).

V. CO-OCCURRENCE MATRIX

The matrix that is defined over an image to be the distribution of the co-occurring values at a given offset. A co-occurrence matrix can also be defined mathematically over an image from (m*n). and in this now the offset is parameterized as $(\Delta x, \Delta y)$,

The co-occurrence matrix will provide you the second order method for generating the texture features. This probability is also termed as the special gray level dependency represents the conditional joint probabilities of all gray levels in the pair wise combination in the spatial window. There are the two parameters illustrated below they are the inter pixel distance and the orientation angles and they are represented by the terms 'd' for the inter pixel distance and the orientation or the angles is represented by ' '

Therefore the probability measure can be taken into the consideration as follows:

$Pr(x)=\{C_{ij}| (d,\theta)\}$ (6)

Where Cij (the co-occurrence probability between gray levels i and j) is defined as

$C_{ij}=(P_{ij})/(\sum_{I,j=1\ldots to\ G}P_{ij})$ (7)

Where Pij represents the number of co-occurrence of gray level i and j within the given window, given a certain (d,ө) pair; and G is the quantized number of gray level (In this work, G=16)

The orientation o the angle is defined as the set of the nearest neighbors as same as the most researches for example degree of 0,45,90,135.then now coming to the inter pixel distance 'd' is defined as it is quite broad range of 1 to 3 or 1 to 5 or 1 to 9 therefore this project chooses the middle one.

5.1 Feature extraction based on the CCM:

A co-occurrence matrix or co-occurrence distribution is a matrix or distribution that is defined over an image to be the distribution of co-occurring values at a given offset. Mathematically a co-occurrence matrix x is defined over an n*m image I, parameterized by an offset. The value of the image originally referred to the gray scale value of the specified pixel. The value could be anything, from a binary on/off value to 32-bit color and beyond. Note that 32-bit color will yield $2^{32}x2^{32}$ co-occurrence matrix. Really any matrix or pair of matrices can be used to generate a co-occurrence matrix, though their main applicability has been in the measuring of texture in images, so the typical definition, as above, assumes that the matrix is in fact an image[17]. It is also possible to define the matrix across two different images. such a matrix can then be used for color mapping. Note that the (Δx, Δ y) parameterization makes the co-occurrence matrix sensitive to rotation.

We choose one offset vector, so a rotation of the image not equal to 180 degrees will result in a different co-occurrence distribution for the same (rotated) image. This is rarely desirable in the applications co-occurrence matrices are used in, so the co-occurrence matrix is often formed using a set of offsets sweeping through 180 degrees (i.e. 0,45,90, and 35 degrees) at the same distance to achieve a degree of rotational invariance

Assuming color image is divided into N*N image sub-block, for anyone image sub-block.

$$T_{(i-j)}(1\le i \le N, 1 \le j \le N) \tag{8}$$

Using the main color image extraction algorithm to calculate the main color $C_{(i,j)}$. for any two four connected image sub-block `$T_{(i \cdot j)}$ and $T_{(k, l)}(|i − k| = 1$ and j=1 or |j-1|=1 and i=k).If it corresponds to the main color and in the HSV space to meet the following condition.

Cj and the Ci belong to the same color of the magnitude that is its HSV components Hi=Hj, Si = Sj,Vi = Vj

Cj and Ci don't belong to the same color the same color of the magnitude but satisfy condition Si*3+Vi=Vj and |Hi - Hj|=1; or satisfy Hi = Hj ,Si = Sj and Vi, Vj {0,1}

We can say the image sub block Ti,j and Tk,l are color connected. According to the concept of the color connected regions, we can make each sub block of the entire image in to a unique color of connected set S= {Ri}(1≤i≤M)

The set S corresponds to the color connected region. For each color connected region {Ri}(1≤i≤M) the color components R g in the RGB color space and H in the HSV color space are respectively extracted the CCM at the direction the same[20].

Operation is done with one.

$$\delta = 1; \theta = 0°, 45°, 90°, 135°$$

The statistic features extracted from the CCM are as follows:-

Energy E=$\sum\sum[m(i,j)]^2$ (9)
Contrast I=$\sum\sum(i-j)^2 m(i,j)$ (10)
Entropy S=$\sum\sum m(i,j) \log[m(i,j)]$ (11)
Where ,if m(i,j)=0,log[m(i,j)]=0   (12)
Inverse difference H=$\sum\sum m(i,j)/1+(i-j)^2$ (13)

VI. SIMILARITY CALCULATION

6.1. DISTANCE MATRIC

Along with efficient features the proper choice of distance matric plays a key role in CBIR. This has necessistated us to explore the different similarity measures to find out best distance matric for image retrieval .in conventional image retrieval technique, Euclidean distance is used to find the similarity between the query image and database image.similarity score is used to find the best match of query image from the database image. The distance metric gives the minimum distance the query image and its nearest shape in the database is the best metric.for better classification, the maximum intra-class distance should be less than the minimum of the intra-class distance.we assume p and q represents the feature vectors for database image and query imagerespectively in each distance metric.the present work evaluates and compares the CBIR performance for computing distances d(p,q) using the following distance metrics :

6.2 EUCLIDEAN DISTANCE

Euclid stated that the shortest distance between two points on a plane is a straight line and is known as Euclidean distance. Euclidean distance metric as in equation (5.1) was often called Pythagorean metric since it is derived from Pythagorean theorem. Euclidean distance metric defined for p=2.in euclidean distance metric difference of each feature of query and database image is squared which increases the divergence between the query and database image.

$$D_{euc}=\sqrt{\sum (p_j-q_j)^2} \tag{14}$$

The similarity between the query image feature vectors and the feature vectors of the images in the database is computed using the Euclidean distance formula

6.3 Texture similarity calculation

In , the similarity between the images are calculated using the Euclidean distance The similarity between the texture feature vector of the query image and the texture feature vectors of the images in the database is calculated by the formula in the equation[18]

$$Sim_{tex}=\sqrt{\sum(QI_n-DBI_n)^2} \qquad (15)$$

where n=1 to 4 ,QI is the query image and the DBI is the database image.

Implementation
1. Algorithm

In this project, experimental data set contains 1000 images from MOSDAC database of images. experimental images of satellite cloud images for testing. In section 3, we study two kinds of feature extraction techniques: feature extraction techniques based on the HSV color space and texture feature extraction technology. At texture feature extraction techniques, we introduces two different extraction methods the gray co-occurrence matrix and CCM. texture are just in part describing the characteristics of images. image database varies some images dramatic ups and downs in gray level, showing a very strong texture characteristics, and some images from a number of smooth but the colors are different regional composition.
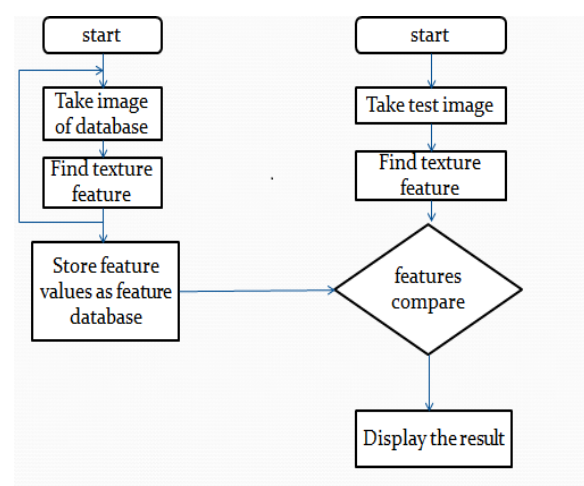
FLOWCHART

Figure 5: Flow chart of texture based satellite cloud image retrieval

VII. RESULTS & DISCUSSIONS
1. System performance and evaluation

The standard measures of evaluating the performance of a CBIR system are precision and recall. This method is used to evaluate the performance of the system in for the typhoon image retrieval. In this research, the performance of the system is evaluated on the basis precision in the retrieving the images.

Precision: Precision is defined as the ratio of the number of the relevant images retrieved to the total number of the images retrieved. It is a measure of the accuracy of the retrieval and can be defined as :

Precision=no. of relevant images retrieved/total no.of relevant images in the database.
Where ,the precision gives information about the relative effectiveness of the system.

Recall, Recall=no.of relevant images retrieved/total no.of similar images in database.
Where,the recall gives information about the absolute accuracy of the system.

The performance of satellite cloud image retrieval described in terms of precision and recall values. For the results of satellite cloud image which is shown in figure 6 ,the performance rate is high.in whole database we have 20 rotated versions of satellite cloud images. the retrieval rate is high by using the proposed method. the satellite cloud image is texture dominant image.as we are using the GLCM method for image retrieval, its retrieves the texture dominant images effectively. So, for the satellite cloud test image more no.of similar images are retrieved.

2. Texture feature extraction

The texture feature vector is computed by taking average of all the values of texture descriptors contrast, correlation, homogeneity and energy, computed from the 4 GLCMs. The texture feature vector of the infrared satellite image is presented in the Table 2.

Table2: Texture feature vector

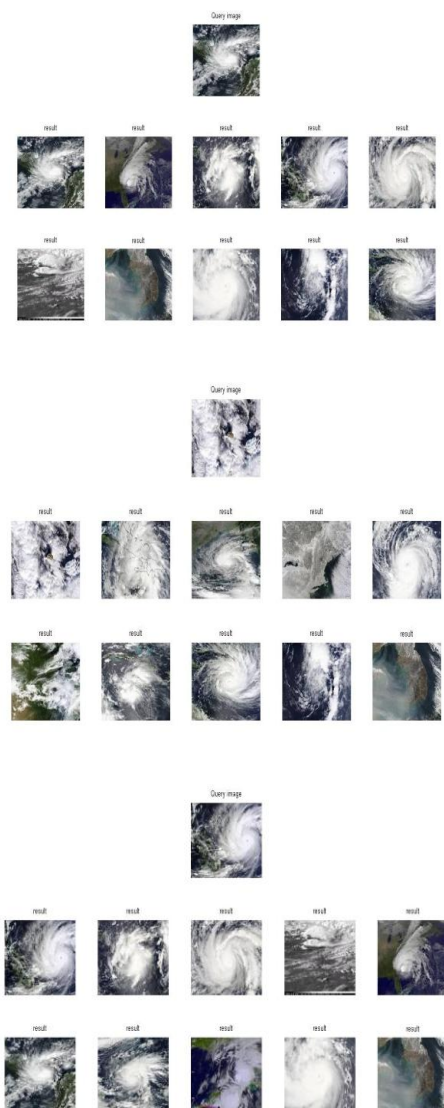|	contrast	correlation	Homogeneity	Energy
Avarege value	14.1639	3.7985	3.1109	414.9777
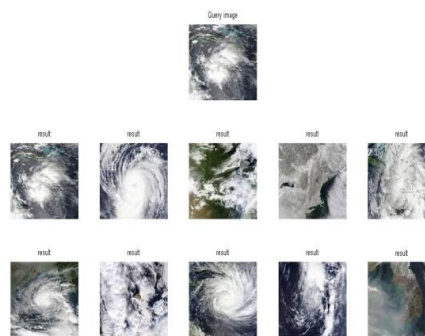
Figure 6: Retrieval of Top 10 similar satellite images for different query images.

VIII. CONCLUSION & FUTURE SCOPE

In this paper, presented the retrieving satellite cloud images from the database using the texture based image retrieval.similar images can be retrieved quickly and accurately by inputting an image.GLCM has been used in various areas to improve the performance of the system and to achieve better results in different applications.

In future work, we hope to build a generalized query method which increases the system searching ability and provide more accurate content discriptions of places of interest.This work can be further extended to some domain-based applications such as finger print recognition, retina identification, and object detection etc for large image database.

REFERENCES

[1]. YogitaMistry, Dr.D.T. Ingole, "Survey on Content BasedImage Retrieval Systems", IJIRCCE Vol. 1, Issue 8, October 2013.

[2]. M. Martins, L. Guimaraes, and L. Fonseca, "Texture feature neural classifier for remote sensing image retrieval systems," XV Brazilian Symposium on Computer Graphics and Image Processing, pp. 90 – 96,, 7-10 Oct. 2002.

[3]. Y. Li and T. Bretschneider, "Semanticsensitive satellite image retrieval," Geoscience and Remote Sensing, IEEE Transactions on, vol. 45, pp. 853 – 860, April 2007.

[4]. Y. Li and T. Bretschneider, "Semantics-based satellite image retrieval using low-level features," IEEE International Geoscience and Remote Sensing Symposium, vol. 7, pp. 4406 – 4409, 2004.

[5]. Y. Alemu, Jong-bin Koh, M. Ikram, and Dong-Kyoo Kim, "Image Retrieval in Multimedia Databases: A Survey," Intelligent Information Hiding and Multimedia Signal Processing, 2009. IIH-MSP '09. Fifth International Conference on, 2009, pp. 681-689.

[6]. Fierro-Radilla, A., Perez-Daniel, K., NakanoMiyatakea, M., Perez-Meana, H., &Benois-Pineau, J. (2014). An Effective Visual Descriptor Based on Color and Shape Features for Image Retrieval. In A. Gelbukh, F. C. Espinoza, & S. N. Galicia-Haro (Eds.), Human-Inspired Computing and Its Applications: 13th Mexican International Conference on Artificial Intelligence, MICAI 2014, Tuxtla Gutiérrez, Mexico, November 16-22, 2014. Proceedings, Part I (pp. 336–348). Cham: Springer International Publishing. Retrieved from http://dx.doi.org/10.1007/978-3-319-13647-9_31.

[7]. Viitaniemi, V. (2002). Image Segmentation in Content-Based Image Retrieval (Master Thesis).Helsini University OF Technology, Department of Electrical and Communications Engineering, Finland.

[8]. Kekre, H. .(2011). Survey of CBIR Techniques and Semantics.International Journal of Engineering Science and Technology (IJEST, 3(5).

[9]. Danish, M., Rawat, R., & Sharma, R. (2013). A Survey: Content Based Image Retrieval Based On Color, Texture, Shape &Neuro Fuzzy. Int. Journal Of Engineering Research And Application, 3(5), 839–844.

[10]. Gonzalez, R. C., & Woods, R. .(2010). Digital Image Processing (3rd Ed). India: Prentice-Hall.

[11]. Jaswal, G., &Kaul, A. (2009). Content Based Image Retrieval – A Literature Review.In Communication and Control. India: National Institute of Technology.

[12]. Gwo-jong yu,yuh-shyanchen,kuei-ping shih 2004. A content based image retrieval System for outdoor ecology learning afirefly watching system.proceeding of the 18th international conference on advanced information networking and application.

[13]. Yusra .T. Mshari, Hameed A. Younis, "Content Based Image Retrieval using Haar Wavelet to Extracted Color Histogram and Texture Features", IJCSMC, Vol. 4, Issue. 8, August 2015.

[14]. Ashwani Kr. Yadav, R. Roy, "Survey on Content-based Image Retrieval and Texture Analysis with Applications",IJSIP Vol. 7, No. 6 (2014), pp. 41-50.

[15]. Syed HamadShirazi, "Content-Based Image Retrieval Using Texture Color Shape and Region", IJACSA Vol. 7, No. 1, 2016.

[16]. Vincent Arvis, Christophe Debain, "Generalization of The Co-occurrence Matrix For Colour Images: Application To Colour Texture Classification", Image Anal Stereol 2004; 23:63-72

[17]. Mrs. Smita Jawale, "Gabor Wavelet Transform And Cooccurrence Matrix Based Texture Features For CBIR", IJERA (VNCET-30 Mar'12).

[18]. WangXing-yuan, Chen Zhi-feng, "An effective method for color image retrieval based on texture", Computer Standards& Interfaces 34 (2012) 31–35.

[19]. Bino Sebastian V, A. Unnikrishnan, "Grey level cooccurrence matrices: generalisation and some new features", IJCSEIT Vol.2, No.2, April 2012.

[20]. M. Martins, L. Guimaraes, and L. Fonseca, "Texture feature neural classifier for remote sensing image retrieval systems," XV Brazilian Symposium on Computer Graphics and Image Processing, pp. 90 – 96,, 7-10 Oct. 2002.